Exploiting wild tomato genetic resources for resistance to Tomato Yellow Leaf Curl Virus

Carolina Myluska Caro Rios

Thesis committee

Promotor
Prof. Dr Richard G.F. Visser
Professor of Plant Breeding
Wageningen University

Co-promotors
Dr Yuling Bai
Associate professor, Laboratory of Plant Breeding
Wageningen University

Dr Richard J.M. Kormelink
Senior scientist, Laboratory of Virology
Wageningen University

Other members
Prof. Dr Jaap Bakker, Wageningen University
Dr Wim H. Vriezen, Bayer CropScience Vegetable Seeds, Roermond
Dr Joan E. Wellink, Wagenigen University
Prof. Dr Bas J. Zwaan, Wageningen University

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences

Exploiting wild tomato genetic resources for resistance to Tomato Yellow Leaf Curl Virus

Carolina Myluska Caro Rios

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A. P. J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 7 December 2015
at 1:30 p.m. in the Aula.

Carolina Myluska Caro Rios
Exploiting wild tomato genetic resources for resistance to Tomato Yellow Leaf Curl Virus,
187 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)

With references, with summary in English

ISBN 978-94-6257-593-6

Table of contents

Chapter 1

General Introduction

Tomato yellow leaf curl disease (TYLCD) is a major constraint in tomato (*Solanum lycopersicum*) production worldwide production since the 1980's (Moriones and Navas-Castillo 2000). It is considered a devastating disease in many other economically important horticultural crops as well, including common bean (*Phaseolus vulgaris*), sweet pepper (*Capsicum annuum*), chilli pepper (*Capsicum chinense*), tobacco (*Nicotiana tabacum*); ornamental crops such as lisianthus (*Eustoma grandiflora*) and *Petunia* species. In addition, it and can also infect various common weeds (Diaz-Pendon et al. 2010).

TYLCD is caused by a complex of several virus species belonging to the *Begomovirus* genus, such as *Tomato Yellow Leaf Curl Virus* (TYLCV). *Begomovirus* is the largest genus of the *Geminiviridae* family, which comprises more than 200 species. In addition to TYLCV, the most invasive and best studied species, TYLCV-like viruses and strains have been identified and classified according to the description in the International Committee on Taxonomy of Viruses (ICTV) (King et al. 2011).

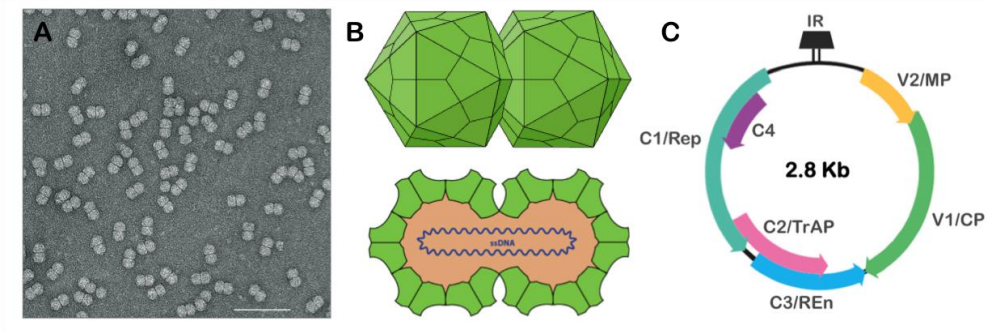
The virus: TYLCV genome organization

Unlike most plant infecting viruses which contain single-stranded RNA genomes, all viruses in the *Geminiviridae* family possess a single-stranded DNA (ssDNA) genome (Briddon and Stanley 2009). Most viruses in the *Begomovirus* genus have a bipartite genome with two circular ssDNA molecules, DNA-A and DNA-B (Gronenborn 2007, Zhou 2013). Viruses within the TYLCV-like virus complex have a single genome component of approximately 2.7-2.8kb similar to the DNA-A of the bipartite begomoviruses (Figure 1), except for *Tomato yellow leaf curl Thailand virus* (TYLCTHV) and *Tomato yellow leaf curl Kanchanaburi virus* (TYLCKaV), which have two components (Gronenborn 2007, Hosseinzadeh et al. 2014).

The TYLCV genome contains six open reading frames (ORFs) (Figure 1). These ORFs encode proteins for encapsidation (V1/CP), virus movement (V2/MP), replication initiator (C1/Rep), transcriptional activator (C2/TrAP), replication enhancer (C3/REn) and a determinant for symptom expression and virus spreading (C4). They are partially overlapped and organized in two transcriptional directions, separated by a conserved inverted repeat termed intergenic region (IR) (Gronenborn 2007). The IR of about 200-nucleotides functions as the origin of replication and contains the (bidirectional) promoters of V1/CP, V2/MP, C1/Rep, and C2/TrAP (Gronenborn 2007). All six ORFs are essential for the establishment

of a successful infection, from efficient replication to long distance movement within the host plant (Castillo et al. 2007).

Figure 1. *Tomato Yellow Leaf Curl Virus*, a begomovirus of the Geminiviridae family. A. Electron microscopy image of purified begomovirus particles (bar=100nm, Glick et al. 2009). B. Drawing of a begomovirus virion capsid; each geminate particle contains one ssDNA (ViralZone:www.expasy.org/viralzone, SIB Swiss Institute of Bioinformatics). C. Schematic representation of the TYLCV genome organization. Open reading frames (ORFs) are represented by arrows, coding in two directions: virion sense (V) and complementary sense (C). V1/CP, coat protein; V2/MP, virus movement; C1/Rep, replication initiator; C2/TrAP, transcriptional activator; C3/REn, replication enhancer; C4, symptom determinant and spreading; IR, intergenic region.



For many begomoviruses, a successful induction of viral symptoms has frequently been found in association with two classes of DNA satellite molecules, termed alpha- and beta-satellites. Beta-satellites are circular ssDNA molecules of approximately 1,350-nucleotides long and reported as pathogenicity determinants of some TYLCV species. They encode a single gene, known as βC1, which plays an important role in disease symptom induction (Zhou 2013). Monopartite begomoviruses are capable of establishing an infection in the absence of beta-satellites, but symptom expression may be affected; beta-satellites seem to be essential for reaching high viral titres and required for symptom development (Cui et al. 2004). Moreover, the beta-satellite βC1 gene is involved in suppression of transcriptional- (TGS) and post-transcriptional gene silencing (PTGS), counteracting plant host defence mechanisms and resulting in severe infections (Yang et al. 2011, Zhou 2013). Alpha-satellites are self-replicating circular ssDNA molecules containing one single gene, which codes for a replication-associated protein. Alphasatellites have mainly been found in monopartite begomoviruses

associated with betasatellites (Zhou 2013), but still remains unknown their exact contribution in the begomovirus/betasatellite infections (Shahid et al. 2014).

The vector: *Bemisia tabaci,* the sweetpotato whitefly

In nature, TYLCV is exclusively transmitted by the sweetpotato whitefly (*Bemisia tabaci* Genn.) in a persistent and circulative manner (Gronenborn 2007). Whitefly populations can be classified into at least 24 genetic groups (biotypes) that differ in various biological characteristics, host range, reproductive rate, host adaptability, plant-virus transmission capabilities, degrees of phytotoxic symptom induction, insecticide resistance, morphology and/or behaviour (Brown et al. 1995, Gottlieb et al. 2010, Firdaus et al. 2013, Ning et al. 2015). Biotype B has high fecundity and a very wide host range, while biotype Q is able to acquire higher resistance to insecticides; both biotypes are considered the most important and damaging vector types on tomato and are able to efficiently transmit TYLCV (Brown 2007, Gottlieb et al. 2010, Pan et al. 2012).

Although it was reported that a single whitefly is able to acquire TYLCV and infect tomato plants, approximately 30–50 whiteflies per plant are needed to transfer TYLCV and reach up to 100% efficiency (Czosnek 2007a, Lapidot 2007). Whiteflies are not able to transmit the virus immediately after uptake; the virus first needs to cross layers of barriers from the mid-gut to the salivary glands. TYLCV genomic DNA can be detected in whitefly individuals 30 minutes after acquisition feeding and has a 12-24 hours latency period in *B. tabaci* before the vector becomes infective (Rubinstein and Czosnek 1997). The virus does not replicate within the vector (Hogenhout et al. 2008); the latency period is therefore essential for virions to reach and accumulate at sufficient titres in the salivary glands to infect host plants via injection of saliva upon vector feeding (Czosnek 2007a).

Virus-vector-plant interactions: mediators of plant disease development

During feeding on an infected plant, whiteflies uptake TYLCV virions and become viruliferous vectors. GroEL, a protein produced by an endo-symbiotic bacteria in *B. tabaci,* permits the virus survival in the haemolymph of the vector by its interaction with TYLCV-V1/CP (Morin et al. 2000, Gottlieb et al. 2010).

Viruliferous whiteflies deliver the viral particles into the host phloem sieve tubes. In the host cell, single-stranded (ssDNA) is released from the virions and double-

stranded DNA (dsDNA) is produced, process mediated by host DNA polymerases. The dsDNA is transcribed by the host RNA polymerase II, producing C1/Rep (Hanley-Bowdoin et al. 2013). C1/Rep protein then binds to the conserved nine nucleotides in the intergenic region (IR) to initiate DNA replication via a rolling circle mechanism (RCR), producing multiple viral ssDNA copies from dsDNA circular intermediates (Gutierrez 2000, Moriones and Navas-Castillo 2000). During this process, C3/REn protein plays an important role in viral DNA replication, directly or indirectly enhancing both viral DNA accumulation and symptom expression through its ability to interact with C1/Rep (Castillo et al. 2003). The transcriptional activator (C2/TrAP) protein is essential for infectivity (Wartig et al. 1997). Newly produced ssDNA products can then enter the replication pool or can be encapsidated to produce progeny virions (Figure 2) (Gutierrez 2000, Gutierrez 2002, Khan 2005).

For successful invasion of its host, TYLCV systemic infection occurs when the virus is transported outside the nuclei to the cytoplasm, entering into the phloem sieve tube for long distance trafficking to new infection sites (Moriones and Navas-Castillo 2000, Scholthof 2005, Lucas 2006). Shuffling of TYLCV ssDNA and dsDNA replicons from the nucleus into the cytoplasm is mediated by V1/CP. V2/MP codes for the movement protein, which traffics the viral DNA across a plasmodesma; when the function of V2/MP is impaired, systemic infection of tomato plants is not achieved. C4 may also mediate the movement of viral DNA into adjacent cells (Figure 2) (Wartig et al. 1997, Rojas et al. 2001, Hanley-Bowdoin et al. 2013).

Once a susceptible tomato plant is successfully infected, TYLCV symptoms start to appear 2-3 weeks after exposure to *B. tabaci*. At an early stage of infection (Figure 3), plants show slight yellowing of the leaflet margins in the apical leaves. In a later stage of the disease (Figure 3), plant leaflet ends display curling upwards and cupping, and reduction of leaves and plant size. Finally, severe stunting and yellowing of the plant together with abortion of flowers and fruits will be followed by a stop in plant growth (Lapidot et al. 2007).

Figure 2. Life cycle of the begomovirus TYLCV. After successful acquisition of TYLCV virions, the whitefly vector is able to transmit the virus to a healthy plant during feeding. Viral V1/CP interacts with GroEL, ensuring survival of the virus in the haemolymph of the vector. In the plant cell, infection begins when single-stranded DNA (ssDNA) is released from virions, and copied to generate double-stranded DNA (dsDNA). After transcription by host RNA polymerase II, C1/Rep initiates virus replication via a rolling circle mechanism (RCR), while C3/REn greatly enhances TYLCV DNA accumulation through its interaction with C1/Rep. Viral DNA is transported outside the nuclei to the cytoplasm by V1/CP while V2/MP traffics it across a plasmodesma. C4 may also mediate the cell-to-cell movement. Transcription of the viral genome is activated by C2/TrAP; translation of viral proteins is mediated by host factors. ssDNA is encapsidated into virions by V1/CP, becoming available for subsequent vector acquisition (modified from Hanley-Bowdoin et al. 2013).

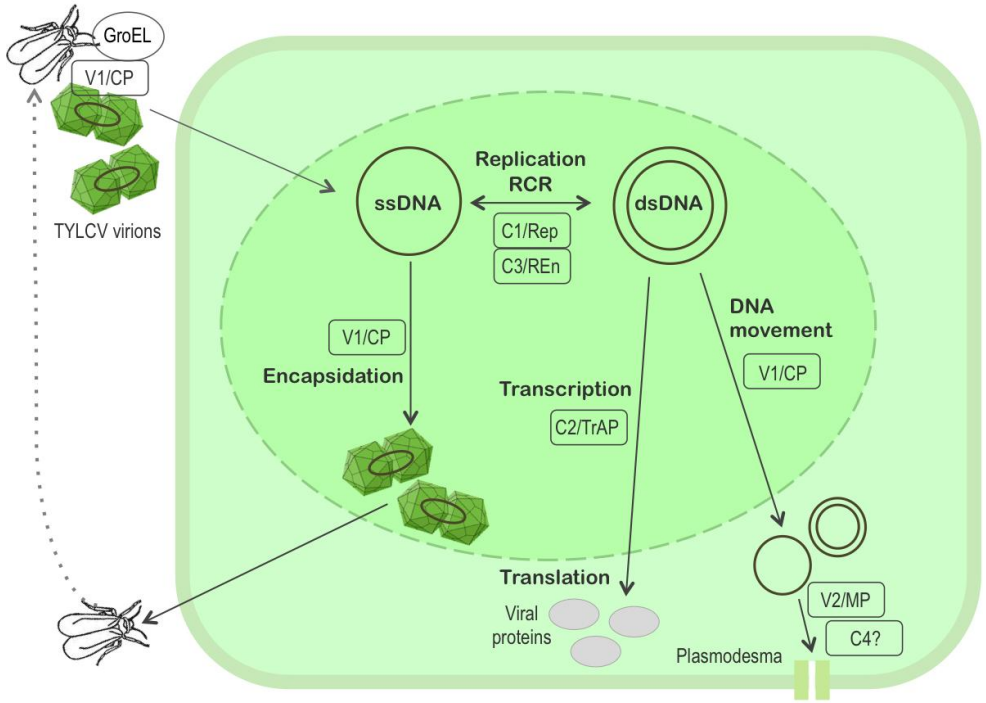


Figure 3. Disease symptoms caused by a successful infection of *Tomato yellow leaf curl virus*. Shown are leaves of a tomato plant: (a) healthy, non-TYLCV infected, (b) at an early stage of TYLCV infection and (c) at a later stage of the disease.

Emergence and spreading of TYLCD

Considering that begomoviruses are exclusively transmitted by *B. tabaci*, the global distribution of TYLCV is closely related to that of whiteflies (Seal et al. 2006). TYLCV originated in the Middle-East Mediterranean region (Pan et al. 2012), and the first official report of its presence dated to 1964 in Israel (Gronenborn 2007). Later on, due to a worldwide increase of the insect vector population, global exchange of plant materials and the rapid evolution of virus variants, TYLCD became one of the most destructive diseases in tomato production areas (Varma and Malathi 2003, Seal et al. 2006). It rapidly spread to Africa, Europe, Caribbean Islands, America, Japan, and Mexico (Moriones and Navas-Castillo 2000, Lefeuvre et al. 2010). Recently, more tomato production regions have reported TYLCV epidemics (Wu et al. 2006, Botermans et al. 2009, Melzer et al. 2010, Van Brunschot et al. 2010, Barboza et al. 2014, CABI 2014).

Genetic variation of begomoviruses arises through mutations, recombination and pseudo-recombination, while recombination is a frequent phenomenon driving the evolution of the TYLCV-like virus complex. TYLCD causing viruses display high degrees of inter- and intra-species recombination within TYLCV-like viruses and between different *Begomovirus* species during mixed infections. Rapid accumulation of emerging recombinants in the progeny contributes to the

generation of genetic diversity, which enables better adaptation of the viruses to ever-changing environmental conditions (Moriones et al. 2007, Diaz-Pendon et al. 2010, Lefeuvre et al. 2010). In addition, most whitefly biotypes are able to transmit a range of begomoviruses (Gottlieb et al. 2010), increasing the possibilities of their rapid recombination and evolution.

Control measurements for TYLCV are usually focused on reducing or avoiding the whitefly vector population; heavy spray of insecticides and the use of nets in isolated greenhouses are common practices (Lapidot 2007). Breeding TYLCV-resistant tomato cultivars, as part of an integrated management approach, is an economically and environmentally sustainable alternative that would reduce the chemical protection dependence and result in safer products (Chomdej et al. 2007).

Virus resistance mechanisms: a multi-layered process

The predominant form of resistance is known as **non-host resistance (NHR)** or native resistance. A first layer of NHR is an inherent and durable protection system present in most plant species (e.g. cell-wall thickening, phytoalexin accumulation, secondary metabolites) acting against a wide variety of pathogens. Non-adapted pathogens are not able to overcome these basic defence mechanisms; thus, only a few pathogens or parasites are able to infect plants (Uma et al. 2011, de Ronde et al. 2014). This layer of resistance is still poorly understood in terms of the genetic and molecular basis, mainly because most non-host species are not crossable with host species (den Boer 2014).

Whenever the pathogen is able to overcome this first basic, passive defence, a second layer of NHR is activated by the host recognition of specific pathogen proteins or structures termed MAMPs or PAMPs (microbe or pathogen associated molecular patterns). Such recognition is mediated by transmembrane pattern recognition receptors (PRRs) and results in a signalling cascade leading to **PAMP-triggered immunity (PTI)** (Jones and Dangl, 2006) (Figure 4). In the case of viruses -being intracellular obligate parasites- they already overcome the plant cell wall barrier when entering their host via mechanical inoculation or through their insect vectors; thus, viral recognition does not take place in the apoplast contrary to fungi, oomycetes and bacterial pathogens (de Ronde et al. 2014). PTI for plant viruses is not characterized, but typical PTI responses, similar to animal antiviral pathways have been recorded (Nicaise et al. 2014). Recently though, BRI1-associated kinase 1 (BAK1), regarded as a general regulator of plant immunity, has been proposed to be

involved in PTI antiviral defence in Arabidopsis; *bak1* mutants display increased susceptibility to three RNA viruses. However, the MAMPs involved in such interaction have not been identified yet (Kørner et al. 2013).

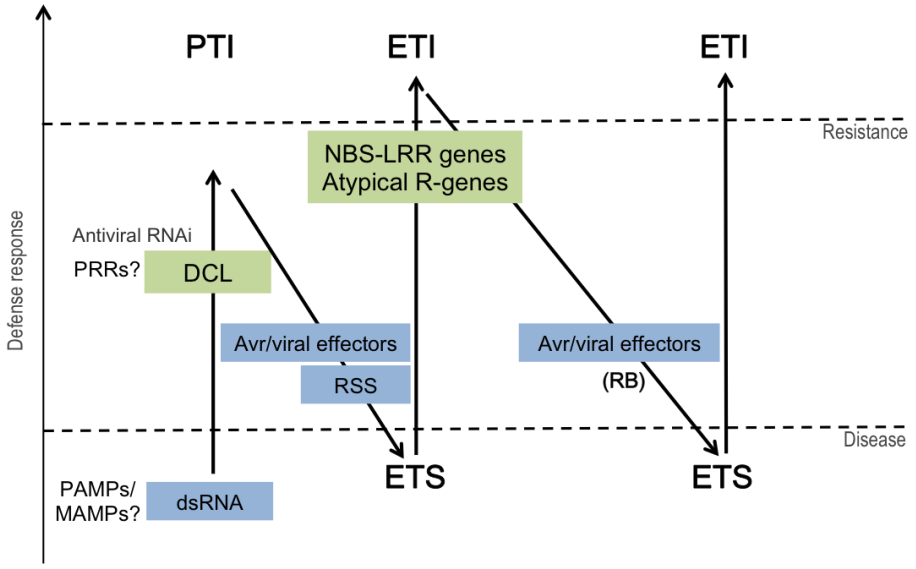
On the other hand, the RNA silencing pathway is generally regarded as the first line of antiviral defence and can also be regarded as PTI. **RNA interference (RNAi)** is triggered by viral double stranded RNA (dsRNA) molecules, which arise from intermediates or secondary RNA folding structures. Such dsRNA molecules act then as PAMPs, and their recognition takes place intracellularly. dsRNA is cleaved by a host RNase type III-like enzyme called Dicer (DCL) to produce 21 to 24 nucleotide-sized small-interfering RNA (siRNA); DCLs are then regarded as PRRs. One strand of these RNA duplex molecules is consecutively incorporated into an RNA-induced silencing complex (RISC), which will degrade viral RNA target molecules with sequence complementarity. This degradation is mediated by slicer, a member of the Argonaut (AGO) protein family and the active core component of RISC (de Ronde et al. 2014, Sharma et al. 2013). For RNA viruses, 21- or 22-nt siRNA are generated that direct the degradation and thereby silencing of viral (m)RNA transcript sequences; a mechanism generally termed post-transcriptional gene silencing (PTGS). As for DNA viruses, additional 24-nt siRNA are produced that -once uploaded into RISC- direct the methylation of cytosines in corresponding (promotor) regions within viral DNA genomes and leads to interference with gene transcription, generally called transcriptional gene silencing (TGS) (Hanley-Bowdoin et al. 2013). A strong plant antiviral RNAi response relies on the amplification of primary siRNAs into secondary siRNAs; this amplification is mediated by host RNA-dependent RNA polymerase proteins (RDRs). Although RNAi might not completely prevent virus infection (de Ronde et al. 2014), it is effective against both RNA and DNA viruses.

In order to suppress PTI and counteract the plant RNAi defence mechanism, viruses have evolved RNA silencing suppressor (RSS) proteins, regarded as pathogenicity factors or effectors. RSS act in diverse manners; they can affect the processing of dsRNAs, suppress the silencing signal amplification or disturb stabilization of siRNAs, or can suppress the RISC complex activity (Alvarado and Scholthof, 2009). As a result of this interference, plant viruses are able to achieve a succesful infection known as Effector triggered susceptibility (ETS). For instance, the viral transcriptional activator protein (C2/TrAP) of geminiviruses inactivates the host adenosine kinase (ADK), interfering in the TGS process (Lozano-Durán et al. 2011). C1/Rep also acts as a silencing suppressor, disrupting TGS by lowering

the expression of plant DNA methyl transferases, thus reducing DNA methylation levels (Rodríguez-Negrete et al. 2013, Wang et al. 2014).

As plant viruses do not encode effector proteins per se, viral effectors were proposed to be defined as viral proteins interfering with host defence signalling to promote virulence (Mandadi and Scholthof 2013). For example, the TYLCV V2/MP targets the host papain-like cysteine protease CYP1, thereby interfering with CYP1-mediated defence activity and facilitating viral infection and/or spread. In such compatible interaction, V2/MP would hence be regarded as a TYLCV effector protein (Bar-Ziv et al. 2012). Likewise, the replication initiator protein (C1/Rep) of begomoviruses binds and inhibits the host retinoblastoma-related protein (RBR), activating the expression of plant genes required for viral replication (Hanley-Bowdoin et al. 2013). Interaction of TYLCV-CP and the plant HSP70 protein is necessary for TYLCV nuclear translocation, viral replication, intracellular movement, and long-distance transport (Gorovits et al. 2013).

Figure 4. Zig-zag model for the plant-virus arms race. The model shown and described in this chapter is modified from Jones and Dangl (2006). Plant pattern recognition receptors (PRR) recognize microbe- or pathogen-associated molecular patters (MAMPs or PAMPs) to trigger **PAMP-triggered immunity** (**PTI**). The antiviral PTI mechanism, **RNA interference** (**RNAi**), is triggered after recognition of the viral double-stranded RNA (dsRNA) by the host RNAse Dicer (DCL), the plant PRR. Viruses encoding RNA silencing suppressors (RSS) would interfere with viral PTI (RNAi), successfully infecting the plant and resulting in effector-triggered susceptibility (**ETS**). Likewise, viral effector proteins will also enable ETS. In a second layer of defence, *R*-genes recognize viral effectors, triggering a strong resistance response known as **effector-triggered immunity** (**ETI**). Pathogens able to evolve new effectors (resistance-breaking strains, RB) would suppress ETI resulting in ETS. *R*-genes able to recognize the newly evolved effectors will trigger again an ETI response. Green/blue boxes correspond to plant/virus factors involved in the RNAi (PTI), ETS and ETI responses.



In the next, second layer of defence, plant resistance genes (*R*-genes) encode for resistance proteins that will sense pathogen effectors or avirulence (Avr) factor proteins and activate the so-called **Effector triggered immunity (ETI)**. The major group of *R*-genes against viral pathogens belong to the nucleotide binding site-leucine-rich repeat (NBS-LRR) family (De Ronde et al., 2014), which after recognizing Avr proteins trigger a defence response generally involving a hypersensitive response (HR). This is a local programmed cell death response, but NBS-LRR genes mediated resistance also induces systemic acquired resistance (SAR), a defence signalling directed to distally located tissues (Nicaise et al. 2014). Examples of this type of antiviral *R*-genes are *Tm-2,* which confers resistance to the tobamoviruses *Tobacco mosaic virus* (TMV) and *Tomato mosaic virus* (ToMV), and *Sw-5b* against the tospovirus *Tomato spotted wilt virus* (TSWV) and several related tospovirus species (Hallwass et al. 2014). In addition to a resistance response mediated by the NBS-LRR class of genes, other distinct resistance mechanisms (generally HR-independent) have been found to activate ETI against viruses and are governed by atypical *R*-genes. This is the case of the RTM genes, which impair long distance spread of several potyviruses in Arabidopsis (Cosson et al. 2012). A comprehensive compilation of all reported and cloned *R*-genes against viruses so far can be found in De Ronde et al. (2014). As for the RNAi pathway, an ETI response has been observed in some *Nicotiana* species inducing immunity against *Tomato*

bushy stunt virus by counteracting the effect of P19, the virus encoded RSS. Such 'extreme resistance' involves salicylic acid and ethylene and no HR is triggered (Sansregret et al. 2013).

To avoid ETI, pathogens evolve to diversify effector molecules, or acquire additional ones that can directly suppress ETI (Alvarado and Scholthof, 2009). Viruses are dynamically evolving in order to antagonize host defence systems (Ishibashi et al. 2014). ToMV-GeRo is a ToMV resistance-breaking strain able to overcome the *Tm-2* resistance in some Arabidopsis ecotypes (Rasul 2012). Amino acid changes in ToMV-GeRo necessary to overcome the resistance have been identified in the N-terminal half of the movement protein (MP), the ToMV avirulence factor/effector protein (Lanfermeijer et al. 2005). A third layer of ETI would then be necessary to successfully protect the plant against such resistance-breaking virus strains. Considering the high plasticity of viral pathogens, a robust plant immune system is necessary to effectively control viral diseases.

In addition to the RNAi and *R*-gene mediated resistance pathways, susceptibility factors (*S*-genes) have become an alternative and interesting source for generating durable resistance. For successfully invading a host, viruses require host factors (susceptibility factors) to complete their infection cycle. The inability of such virus-plant molecular interaction leads to resistance, and is generally termed recessive resistance. This is the case of the potyvirus resistance conferred by (mutants of) host translation initiation factors of the 4E or 4G family (eIF4E/eIF4G), which do not interact with the "cap-like" structure of potyviral transcripts, inhibiting potyviral translation (Sanfaçon, 2015).

Downstream genes involved in the TYLCV-resistance pathway

Resistance mechanisms trigger a cascade of defence responses as part of a complex network. Several studies investigating host plant transcriptional responses triggered by Geminivirus infections have been performed. Defence-related genes involved in resistance pathways such as ubiquitination, proteinase inhibition and ethylene response factors have been identified from comparative transcriptome profiling between TYLCV resistant and susceptible cultivars (Chen et al. 2013). Integrated metabolomics and transcriptomics data pointed out profound activation of phenylpropanoid pathway, salicylic acid (SA) biosynthesis pathway as well as polyamine metabolic pathway in a resistant genotype upon TYLCV infection (Sade et al. 2015).

The involvement of host defence-related genes in the TYLCV resistance network has been functionally confirmed. Permease I-like protein encodes a transmembrane transporter preferentially expressed in *S. habrochaites*-derived resistant plants upon TYLCV inoculation; silencing of its gene in resistant tomato plants compromised the resistance (Eybishtz et al. 2009). Upon silencing the hexose transporter *LeHT1* in TYLCV resistant plants, enhanced virus movement and accumulation has been recorded; this gene is also described to be involved in the signal transduction pathway leading to TYLCV resistance (Eybishtz et al. 2010, Sade et al. 2013). Silencing the lipocalin-like *SlVRSLip* protein gene in resistant plants leads to the loss of resistance (Sade et al. 2012). *SlVRSLip* acts downstream of *LeHT1,* while *Permease I*-like protein acts independent of the *LeHT1- SlVRSLip* associated resistance pathway (Sade et al. 2012).

Mapped TYLCV resistance genes

The term resistance is defined as the ability of plants to prevent or reduce growth and/or development of the pathogen after establishment of the infection process. In other words, resistant plants challenged by a pathogen respond by hindering or preventing the establishment or colonization by the pathogen (Parlevliet and Zadoks 1977). *R*-genes of the type NBS-LRR generally lead to an absolute resistance that often is associated with a visual HR-response.

Resistance against TYLCV exists in nature and has been found in several accessions of wild tomato species. Major resistance genes have been identified and mapped from *S. habrochaites*, *S. chilense* and *S. peruvianum.* In total, six genes conferring TYLCV resistance have been mapped and termed *Ty-1, Ty-2, Ty-3, Ty-4, ty-5* and *Ty-6* (Table 1). In plants containing one of these genes, viral replication is still detected but at much lower levels compared to susceptible genotypes (Pico et al. 1999, Maruthi et al. 2003, Pérez de Castro et al. 2005, Glick et al. 2009). Thus, these genes are considered to confer partial or complete resistance against TYLCV.

Table 1. Mapped tomato genes for TYLCV resistance

	Genetic source		Chromosome location	Genetic mechanism[b]	Inheritance pattern	Reference
	Accession/ Line[a]	Species				
	LA1969					
Ty-1/	LA2779	*S. chilense*	6	Major gene	Incomplete dominance	Zamir et. al. 1994
Ty-3	LA1932					Verlaan et al. 2013
	LA1938					Caro et al. 2015
Ty-2	B6013	*S. habrochaites*	11	Major gene	Dominance	Hanson et al. 2006
Ty-4	LA1932	*S. chilense*	3	Minor gene	Incomplete dominance	Ji et al. 2008
ty-5	TY172	*S. peruvianum?*	4	Major gene	Recessive	Anbinder et al. 2009
Ty-6	LA2779	*S. chilense*	10	Major gene	Recessive	Hutton, 2013

[a] Tomato breeding line TY172 is derived from 4 different accessions of *S. peruvianum*.
[b] Major genes were described as the gene/locus that confers the highest effect on the resistance genotype.

Ty-1 originates from *S. chilense* accession LA1969. It has first been reported and mapped to the short arm of tomato chromosome 6 by Zamir et al. (1994), followed by a study that mapped *Ty-1* to the long arm of chromosome 6 (de Castro et al. 2007). Recently *Ty-1* has been successfully fine-mapped and cloned, and shown to present an allele of a gene coding for a RNA–dependent RNA polymerase (RDR) (Verlaan et al. 2011, Verlaan et al. 2013).

Ty-3 has been identified in *S. chilense* accessions LA1932, LA1938, and LA2779 and was first mapped to the long arm of chromosome 6. It confers resistance against TYLCV and to the bipartite begomovirus *Tomato mottle virus* (ToMoV) (Ji et al. 2007). *Ty-3* mapped to an overlapping interval with the *Ty-1* region, indicating the possibility of allelism. Recently, the fine mapping of *Ty-1* and *Ty-3* and their cloning has confirmed this idea (Verlaan et al. 2011, Verlaan et al. 2013). Both *Ty-1* and *Ty-3* have been the major focus of breeding programs so far and are incorporated into commercial hybrids worldwide.

The *Ty-1/Ty-3* gene from *S. chilense* LA1969 is the first and only TYLCV dominant resistance gene cloned so far, and encodes a RNA-dependent RNA polymerase involved in antiviral RNA silencing. *Ty-1* and *Ty-3* belong to a RDRγ type with homology to RDR3, RDR4 and RDR5 of *A. thaliana,* genes that were not assigned any function until then (Verlaan et al. 2013). Plants containing *Ty-1/Ty-3* show enhanced levels of TYLCV-specific siRNA targeting the promoter region of V1, and concomitant cytosine methylation of its promoter region, suggesting an enhanced TGS resistance mechanism (Butterbach et al. 2014).

Ty-2-mediated resistance has first been identified by Hanson et al. (2000), who reported the presence of a resistance introgression derived from *S. habrochaites* accession B6013 in the tomato resistant line H24. The gene has later been named *Ty-2* (Hanson et. al, 2006) and initially mapped to the bottom of the long arm of chromosome 11 (Barbieri et al. 2010) in a region of 14.6 cM. More recently, *Ty-2* has been further fine-mapped to a 4.5-cM interval (Ji et al. 2009). A complete or partial dominant inheritance pattern has been suggested for *Ty-2* (Hanson et. al. 2000).

Ty-4 originates from *S. chilense* accession LA1932 and has been mapped to the long arm of chromosome 3. *Ty-4* has a minor effect on TYLCV resistance, accounting only for 15.7% of the total variance while *Ty-3,* deriving from the same accession accounts for 59.6% of the variance (Ji et al. 2009).

ty-5 has been identified in the tomato breeding line TY172, which was derived from four different accessions of *S. peruvianum* (PI126926, PI126930, PI390681, and LA0441). Genetic analysis has shown that *ty-5*-mediated resistance is controlled by a major QTL on chromosome 4 (Anbinder et al. 2009). *ty-5* is inherited in a recessive manner and co-segregates with the marker *SlNAC1* (Hutton et al. 2012). A recent cloning study has demonstrated that *ty-5* gene codes for a *pelota* homolog, involved in host protein translation. A T-to-G transversion in the coding region of the *pelota* allele is linked with TYLCV resistance (Levin et al. 2013).

Ty-6, a newly identified TYLCV resistant locus, is derived from *S. chilense* accession LA2779. Preliminary mapping results indicate its location on the long arm of chromosome 10 (Hutton 2013), in a region of approximately 3 Mb (Hutton, personal communication).

Introgression breeding for TYLCV resistance

Ty-1, *Ty-2* and *Ty-3* are dominant TYLCV resistance genes, and are the main sources of resistance for tomato breeding programs in different parts in the world. The *Ty-4* resistance gene contributes to a lesser extent while *ty-5* and *Ty-6* (likely) are genes of recessive nature; these genes are not yet exploited in tomato breeding programs.

Breeding for TYLCV resistance has mainly been based on transferring the resistance originating from wild tomato accessions, using different techniques to facilitate inter-specific hybridization (e.g. embryo rescue). In some cases, undesired traits coupled with the trait of interest are introgressed as well, a phenomenon known as linkage drag. In order to break this linkage drag, repeated backcrossing

is needed. However, structural rearrangements frequently occur when alien chromosome regions are introgressed from wild relatives into cultivated tomato (Szinay et al. 2010). These non-homologous rearrangements can suppress recombination rates and hamper the introgression of the resistance gene without linkage drag.

This is the case for the *Ty-1* resistance introgressed from *S. chilense* into cultivated tomato, which had an adverse effect on several production and quality traits (Rubio et al. 2012). The suppression of recombination reported in the *Ty-1* region was later uncovered using BAC-FISH to be caused by two inversions between the *S. chilense* LA1969 resistance donor and *S. lycopersicum*, disturbing chromosomal pairing during meiosis (Verlaan et al. 2011).

Aims of the thesis

So far, only three out of the six aforementioned TYLCV resistance genes have been cloned, i.e. *Ty-1, Ty-3* and *ty-5*. Studies have indicated that the *Ty-1/Ty-3* RDR gene resistance mechanism is based on enhanced transcriptional gene silencing (TGS; Butterbach et al. 2014). Although reports on the cloning of *ty-5* have been publicly presented, very limited information has been released regarding its TYLCV resistance mechanism.

Although the *Ty-1* gene is thought to likely confer a broad spectrum Geminivirus resistance based on its involvement in the innate immune RNAi response (Butterbach et al. 2014), some tomato production regions have reported TYLCV strains overcoming this resistance (García-Cano et al. 2008). For this reason, there is a necessity to uncover the resistance mechanisms associated with the other *Ty*-genes and their allelic variants in order to design an efficient strategy to pyramid resistance genes with different (levels of) resistance or tolerance mechanisms.

The aim of this thesis was to characterize the known TYLCV resistance loci, to assess candidate genes possibly involved in their defence responses and to analyse the feasibility of their successful introgression into the cultivated tomato. In addition, we aimed to map and identify new tomato genes for resistance to TYLCV from wild tomato relatives, which would potentially also complement the proposed diverse resistance profile for breeding. Altogether, the objectives of this thesis aimed to provide further insights into various TYLCV resistance genes/mechanisms

to contribute to a breeding strategy for cultivars that are able to mount an effective, broad-spectrum, durable resistance response.

To this end, we first examined the genetic variation of the *Ty-1/Ty-3* gene in a large panel of cultivars, landraces and accessions of *Solanum* related species. The results described in **Chapter 2** show that many *S. chilense* accessions probably carry a resistant allele of the *Ty-1* gene and additionally, each accession can harbour more than one TYLCV resistance locus. We also show that the catalytic domain of the RDR protein is conserved among the tested tomato *Solanum* species, and identified three *Ty-1/Ty-3*–specific amino acids shared by seven TYLCV resistant lines and accessions. Additionally, we assessed the expression behaviour of resistant and susceptible *Ty-1* alleles.

Chapter 3 describes the fine-mapping of the resistance locus *Ty-2* introgressed from *S. habrochaites* B6013 to an interval of approximately 300kb. Expression and functional analysis were carried out in order to assess predicted genes in the region and propose candidates for the *Ty-2* gene. A severe suppression of recombination region of approximately 200kb was detected within the *Ty-2* introgression, but the causes of this phenomenon could not be uncovered by a BAC-FISH painting approach. In order to better visualize the genome structure of the *Ty-2* region, the draft *de novo* sequence of an accession of *S. habrochaites*, LYC4, was compared with the *S. lycopersicum* sequence in **Chapter 4**. Together with BAC sequence data of the region from a *Ty-2*-containing line and a genetic inheritance analysis, these studies revealed an inversion to be the cause for the recombination suppression. These findings highlight the importance of understanding genome structures when interspecific crosses are performed, as a premise for introgression breeding.

Aimed at the identification of novel genes and/or mechanisms of resistance to TYLCV, we investigated the newly identified resistance from *S. pimpinellifolium* accession G1.1554. In **Chapter 5** a QTL mapping strategy is described for identification of the genetic factors conferring resistance in this accession using a RIL mapping population. Using a combination of different ~omics platforms, two QTLs conferring resistance were identified, and also indications of the secondary metabolites and volatiles to be differentially and constitutively present in resistant lines compared to susceptible ones are presented.

Finally, **Chapter 6** summarizes and discusses the implications of the findings presented in this thesis for a successful introgression-breeding program. Likewise,

the importance of pyramiding different layers of innate immunity to TYLCV and future prospects on resistance breeding are discussed.

Chapter 2

Assessing the genetic variation of *Ty-1* and *Ty-3* alleles conferring resistance to Tomato Yellow Leaf Curl Virus in a broad tomato germplasm

Myluska Caro[1,2], Maarten G. Verlaan[1,a], Richard Finkers[1], Anne-Marie A. Wolters[1], Samuel F. Hutton[3], John W. Scott[3], Richard Kormelink[4], Richard G.F. Visser[1] and Yuling Bai[1,c]

[1]Wageningen UR Plant Breeding, Wageningen University & Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
[2]Graduate School Experimental Plant Sciences, Wageningen University & Research Centre, Droevendaalsesteeg1, 6708 PB Wageningen, The Netherlands
[3]Gulf Coast Research and Education Center, University of Florida, 14625 CR 672, Wimauma, FL 33598, USA
[4]Laboratory of Virology, Wageningen University & Research Centre, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
[a]Present address: Rijk Zwaan Breeding B. V., Burgemeester Crezéélaan 40, 2678 KX De Lier, The Netherlands

This chapter has been published in a modified version in Molecular Breeding 2015; 35(6): 132

Abstract

Tomato Yellow Leaf Curl Virus (TYLCV) hampers tomato production worldwide. Our previous studies have focussed on mapping and ultimately cloning of the TYLCV resistance genes *Ty-1* and *Ty-3*. Both genes are derived from *Solanum chilense* and were shown to be allelic. They code for an RNA-dependent RNA polymerase (*RDR*) belonging to the *RDR*γ type defined by a DFDGD catalytic domain. In this study, virus-induced gene silencing (VIGS) was used to silence *Ty-1*/*Ty-3* in tomato lines carrying TYLCV resistance introgressed from *S. chilense* LA1932, LA1938 and LA1971. Results showed that silencing *Ty-1/Ty-3* compromised the resistance in lines derived from *S. chilense* LA1932 and LA1938. The LA1971-derived material remained resistant upon silencing *Ty-1/Ty-3*. Further, we studied the allelic variation of the *Ty-1/Ty-3* gene by examining cDNA sequences from nine *S. chilense* derived lines/accessions and more than 80 tomato cultivars, landraces and accessions of related wild species. The DFDGD catalytic domain of the *Ty-1/Ty-3* gene is conserved among all tomato lines and species analysed. In addition, the 12 base pair insertion at the 5-prime part of the *Ty-1/Ty-3* gene was found not to be specific for the TYLCV resistance allele. However, compared to the susceptible *ty-1* allele, the *Ty-1/Ty-3* allele is characterized by three specific amino acids shared by seven TYLCV-resistant *S. chilense* accessions or derived lines. Thus, *Ty-1/Ty-3* specific markers can be developed based on these polymorphisms. Elevated transcript levels were observed for all tested *S. chilense* *RDR* alleles (both *Ty-1* and *ty-1* alleles), demonstrating that elevated expression level is not a good selection criterion for a functional *Ty-1/Ty-3* allele.

Keywords: Breeding, Resistance, RNA-dependent RNA polymerase (*RDR*), Tomato, Tomato Yellow Leaf Curl Virus (TYLCV), Virus-induced gene silencing (VIGS)

Introduction

Tomato Yellow Leaf Curl Virus (TYLCV), a *begomovirus* of the *geminiviridae* family, is a phloem-limited single-stranded DNA virus that is vectored by the whitefly (*Bemisia tabaci*). TYLCV is one of the causal viruses of Tomato Yellow Leaf Curl Disease (TYLCD). In the last two decades TYLCD has been a major constraint on tomato (*Solanum lycopersicum*) production in many warm and (sub) tropical regions worldwide and nowadays it is still a huge problem for many farmers. Tomato plants affected by TYLCD show yellowing and curling of apical leaves, and when plants are severely affected, flowers are abscised and plants stop growing completely (Cohen and Lapidot, 2007). Controlling vector whitefly populations is expensive, labour intensive, and often ineffective; thus, using resistant tomato cultivars is a good solution to control TYLCV. No resistance has yet been described in cultivated tomato, and breeders have screened wild tomato relatives to identify resistance sources from which resistance loci have been introgressed (Ji. et al. 2007; Vidavski, 2007). "

To date, six TYLCV resistance/tolerance genes have been described, *Ty-1* to *Ty-6* (Figure S1) (Zamir et al. 1994; Hanson et al. 2006; Ji et al. 2007; Anbinder et al. 2009; Ji et al. 2009; Hutton et al. 2012; Hutton and Scott 2013). Most of these loci originated from accessions of *S. chilense*. The *Ty-1* gene is derived from LA1969 and the *Ty-3* gene from LA2779. Both *Ty-1* and *Ty-3* are located on the long arm of tomato chromosome 6 and have been shown to be allelic (Verlaan et al. 2011; Verlaan et al. 2013). LA1932 is reported to carry an allele at this locus, *Ty-3a*, (Scott et al. 1996; Ji et al. 2007), and is also the donor of *Ty-4*, which maps to chromosome 3 (Ji et al. 2009). *Ty-6* is derived from LA2779 (also the donor of the *Ty-3* allele) and recently mapped to chromosome 10 (Hutton, 2013). The other two known TYLCV resistance genes do not originate from *S. chilense*. *Ty-2* was introgressed from *S. habrochaites f. glabratum* accession "B6013" and is located on chromosome 11 (Yang et al. 2014). *Ty-5* was first described in TY172, a breeding line said to be derived from crosses of four *S. peruvianum* accessions. However, whether the *Ty-5* originated from *S. peruvianum* is still in debate; there is recent evidence that this gene is recessively inherited and resulted from a loss-of-function mutation that likely occurred in cultivated tomato (Hutton et al. 2012; Levin *et al.* 2013). *Ty-5* maps on chromosome 4, and because of its recessive nature, the symbol *ty-5* was proposed to refer to this gene (Hutton et al. 2012).

Recently we cloned the *Ty-1* and *Ty-3* genes (Verlaan et al. 2013), which code for RNA-dependent RNA polymerases (*RDR*) belonging to the *RDR*ɣ type. *RDR*s are defined by a conserved motif in the catalytic domain, DFDGD for the *RDR*ɣ, and

DLDGD for the *RDR*α type. In *Arabidopsis thaliana*, the *RDR*α type has been well-studied and shown to be involved in stress response, pathogen resistance, female gamete formation, and transgene silencing amongst many other functions (excellently reviewed in Willmann et al. 2011). In contrast to the *RDR*α type, no functions for *RDR*γ have been described in literature. Because the *RDR*α type is known to be involved in the amplification of the siRNA signal, it is possible that the *RDR*γ type has a similar function in siRNA amplification. Our results suggested that *Ty-1*, representative for the *RDR*γ type and a novel class of *R*-genes, confers resistance through enhanced transcriptional gene silencing (Butterbach et al. 2014).

In *S. chilense*, multiple accessions have been described as symptomless after TYLCV inoculation. For some accessions, including LA1960, LA1971 and LA1938, the causal genes for resistance were mapped to chromosome 6 in the chromosome region where *Ty-1* is located, suggesting allelism to *Ty-1*/*Ty-3* (Pérez de Castro *et al.*, 2013; Agrama and Scott, 2006, Hutton and Scott, unpublished data). For other accessions, the causal genes have not been mapped (Pico et al. 1999). Thus, it is intriguing whether the TYLCV resistance in various *S. chilense* accessions is governed by allelic variants of the *Ty-1/Ty-3* gene.

The aim of this study is to explore the allelic variation of *Ty-1*/*Ty-3* in wild tomato relatives, with the focus on *S. chilense* accessions. In a previous study we showed that the susceptible *ty-1* allele differs from the resistant *Ty-1*/*Ty-3* allele at multiple amino acid positions (Verlaan et al. 2013). The most striking difference was a 4 amino acid insertion near the start of the protein in the *Ty-1*/*Ty-3* allele, while in the catalytic domain there were no differences. In this study we compare the full-length cDNA sequence of seven different tomato introgression lines that have *S. chilense*-derived TYLCV resistance and two *S. chilense* accessions to identify *Ty-1/Ty-3* specific polymorphisms. The insertion and catalytic domain of the protein are also explored in 87 lines/accessions of tomato and its wild relatives to see if these *S. chilense* features are unique. Further, we silenced the alleles of the *Ty-1*/*Ty-3* gene in several TYLCV resistant tomato lines carrying introgressions from different *S. chilense* accessions to check whether the silencing compromises the TYLCV resistance in these lines.

Results

Ty-1/Ty-3 alleles in multiple *S. chilense*-derived introgression lines

In a previous study, TYLCV resistance in *S. chilense* accessions LA1932, LA1960 and LA1971 was studied and shown to be controlled by a major dominant gene located on chromosome 6 (Pérez de Castro *et al.* 2013), indicating that the causal

genes in these accessions are likely allelic to *Ty-1*. For the resistance in LA1938, breeding practice showed that there is a linkage of the resistance from LA1938 with the self-pruning (*sp*) locus, which is located on the long arm of chromosome 6 (Agrama & Scott, 2006). Suppression of recombination made breakage of this linkage difficult. Using an F_2 of the cross between LA1938-derived line F11E976-BK (also known as Fla.976) and a susceptible tomato cultivar, the resistance was shown to be linked with the *Ty-3*-associated markers used in Ji et al. (2007).

To assess whether TYLCV resistance derived from the aforementioned *S. chilense* accessions was based on *Ty-1*/*Ty-3* alleles, a VIGS approach was applied to silence the *Ty-1* gene with the TRV2-180 and/or TRV2-190 silencing constructs as described in Verlaan et al. (2013). Tomato introgression lines derived from these accessions were used for VIGS (Table 1). Tomato Moneymaker (MM) plants were used as a susceptible control. Two weeks after TYLCV inoculation, all MM plants showed typical TYLCV symptoms, while plants of tomato introgression lines infiltrated with the empty vector (EV) remained symptom free. In the lines derived from *S. chilense* LA1932 and LA1938, all but two out of 31 plants infiltrated with TRV2-180/190 silencing constructs showed typical symptoms (Table 1). The two symptom-free plants may have been escapes from the TYLCV infection or due to a low silencing level. Together with the mapping data from previous studies (Pérez de Castro et al. 2013), the collapse of TYLCV resistance by VIGS clearly indicates that resistance in the tested lines derived from *S. chilense* LA1932 and LA1938 is based on *Ty-1/Ty-3*.

Table 1. Silencing *Ty-1* compromises TYLCV resistance in multiple *Solanum chilense*-derived lines. VIGS constructs TRV2-180 and TRV2-190 targeting different parts of the gene were used to silence the *Ty-1* allele (Verlaan et al. 2013). Empty TRV vector was used as control.

			Silencing construct			Control		
Resistance source	Tomato line	Reported gene	Plants tested	**S**[a]	R[a]	Plants tested	**S**	R
S. chilense LA1932	1538	*Ty-3*	16	**14**	2	4	**0**	4
	B26	*Ty-3*	4	**4**	0	2	**0**	2
S. chilense LA1938	Fla.976	*Ty-3*	11	**11**[b]	0	5	**0**	5
S. chilense LA1971	1594	unknown	15	**0**	15	5	**0**	5
S. chilense LA2779	Fla.8680	*Ty-3*	14	**13**[b]	1	3	**0**	3
	Fla.8383	*Ty-6*	5	**0**	5	2	**0**	2

[a] Susceptible (S): showing TYLCV symptoms; disease score 2-4
Resistant (R): symptom-free; disease score 0-1

[b] All susceptible plants had a disease score of 2.

In contrast, plants of line 1594 with TYLCV resistance derived from LA1971 remained symptomless after infiltration with both TRV2-180 and TRV2-190 silencing constructs (Table 1). Since all PDS control plants of this line were showing photo bleaching, we assumed that VIGS was working in this line as well. Thus, we expected the majority of the TRV silencing construct-infiltrated plants of the line 1594 to become susceptible if the resistance in this line is conferred by a *Ty-1* allelic variant. Two lines derived from *S. chilense* LA2779 were included in the VIGS experiment. In the line Fla. 8680, which carries the *Ty-3* allele (Verlaan et al. 2013), all TRV silencing construct-infiltrated plants except for one showed TYLCV symptoms. But resistance was uncompromised in Fla. 8383 which carries the *Ty-6* allele (Hutton and Scott 2013), indicating that *Ty-6* is different from the *Ty-1/Ty-3* gene. For each line, at least two VIGS experiments were performed with comparable results.

Genetic variation of the *RDR* in tomato and its wild relatives

There is a four amino acid insertion, from positions 12 to 16, present in the *Ty-1/Ty-3* alleles (Verlaan et al. 2013) compared with the MM allele. To determine whether this insertion may be present in a variety of *S. chilense*-derived TYLCV resistant lines, cDNA was made from six *S. chilense*-derived lines containing *Ty-1/Ty-3* and two wild *S. chilense* accessions (Figure S2A). Primers were designed to amplify the region of interest and sequence analysis showed that these four amino acids (Proline, Serine, Cysteine, Isoleucine) are present in all lines. However, there is one synonymous SNP (T-G) among the *S. chilense* derived lines (Figure S2A).

To check for the presence/absence of the four amino acid insertion amongst cultivated tomato and wild tomato species, the re-sequenced genome reads of 84 accessions of different species were mapped to the reference genome of *S. lycopersicum* cv. Heinz 1706 and compared for the insertion. In addition, draft *de novo* assemblies of three tomato wild relatives (*S. arcanum* LA2157, *S. habrochaites* LYC4 and *S. pennellii* LA716) were included in the analysis (Figure S2B). All cultivated tomato lines in the test panel (including *S. lycopersicum* var. *lycopersicum* and *S. lycopersicum* var. *cerasiforme*) and the majority of the wild species do not have the insertion (Figure S2C). Several related wild species in the test panel do however have the insertion, i. e. *S. arcanum* LA2157, S. *corneliomulleri* LA118, *S. peruvianum* LA1954, two accessions of *S. huaylasense* (LA1983 & LA1365), *S. habrochaites* LYC4, and *S. pennellii* LA716 (Figure S2B & S2D). Within the 12bp insertion, one non-synonymous SNP was detected in *S. habrochaites* LYC4, leading to an amino acid change (P→S). In many disease tests

(data not shown) *S. arcanum* LA2157 and *S. habrochaites* LYC4 exhibited clear virus symptoms after TYLCV infection.

To further explore crucial *Ty-1/Ty-3* allele-specific polymorphisms, the sequences of the coding regions of the *RDR* from tomato, nine *S. chilense*-derived lines and five wild tomato accessions were obtained and analysed (Figure S4). Within *S. chilense* all genotypes have different *RDR* alleles, as shown by the presence of accession-specific SNPs or combination of SNPs. Importantly, five SNPs (two in exon 12, one each in exon 13, 14 and 18, yellow marked in Figure S4) specific to the *Ty-1/Ty-3* alleles were identified. Additionally, four SNPs (two in exon 2, two in exon 6, green marked in Figure S4) were found to be unique for the *Ty-3* allele. The *RDR* cDNA sequence of *S. chilense* G1.1556 contained intron 17, which would result in a premature stop codon. No full-length cDNA sequence was obtained from *S. chilense* LA2779-derived line Fla.8383, but the sequence of exons 12-14 was identical to the MM sequence, indicating the presence of a susceptible *RDR* allele.

RDR protein sequences were derived from the cDNA sequences and aligned (Figure S5). A small number of *Ty-1/Ty-3* specific amino acids were observed, which were shared by TYLCV-resistant *S. chilense* accessions LA1969 (*Ty-1*), LA2779 (*Ty-3*), LA1932 (*Ty-3A*), LA1938, LA1971 and introgression lines BTI-87 and Gh13 reported to contain *Ty-3* alleles (Menda et al. 2014; Mejía et al. 2005). These amino acids are L563, V616, and Q919 (numbering based on the *Ty-1* allele, SEQ2 in Patent No. WO2012125025). They are absent in *S. chilense* accessions G1.1556 and G1.1558 that do not contain *Ty-1* or *Ty-3* (Figure 1). A phylogenetic analysis using an unrooted tree grouped together the proteins of seven *S. chilense Ty-1/Ty-3* alleles responsible for TYLCV resistance (Figure 2).

Figure 1. *Ty-1/Ty-*3 allele-specific polymorphisms. Partial alignment of protein sequences of the *Ty-1 RDR* alleles; red arrows indicate three *Ty-1/Ty-3* specific amino acids.

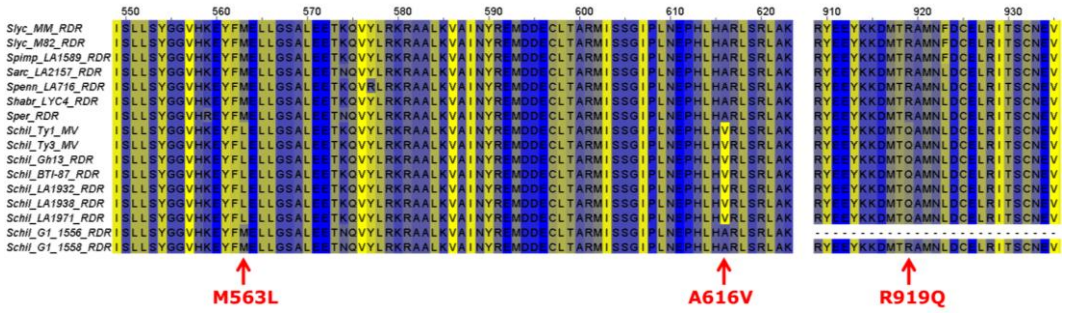
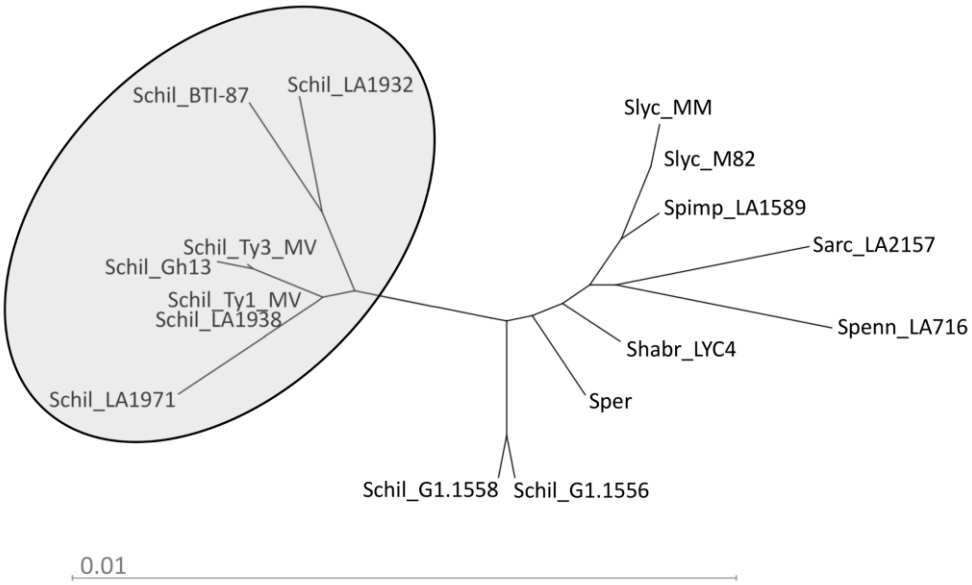


Figure 2. Unrooted phylogenetic tree of protein sequences of *Ty-1/Ty-3 RDR* of accessions of *S. chilens*e and other (wild) tomato species. The *RDR* proteins of the TYLCV-resistant *S. chilense* accessions cluster in one clade, as indicated.



The catalytic domain of the *Ty-1* gene is conserved in tomato and its wild relatives.
The catalytic domain of the *Ty-1*/*Ty-3* allele is characterized by a five amino acid motif, DFDGD (position 723-727) (Verlaan et al. 2013). SNPs in this domain could potentially have an effect on the functioning of this protein. Sequence analysis of an amplified cDNA fragment among all tested *S. chilense* derived lines showed that there were no SNPs present in the catalytic domain and furthermore no differences were found in four amino acids up- or downstream of the catalytic domain (Figure S3). The sequence coding for the catalytic DFDGD motif was also compared among all available sequences used for Figure S2. This region was found to be highly conserved and no polymorphisms were detected among the susceptible and resistant lines analysed.

Elevated expression level of the *RDR* alleles in *S. chilense* accessions.
Analysis of expression of the *RDR* in *S. chilense*-derived resistant lines revealed significant differences compared to the tomato susceptible allele. The expression

level of the alleles was measured at different time points in the presence or absence of the virus (Figure S6). Six resistant lines derived from *S. chilense* LA1969, LA2779, LA1932, LA1938 and LA1971 showed significantly higher relative expression of *RDR* compared to the susceptible allele, despite the presence of the virus. *RDR* transcript levels of line Fla.8383, derived from *S. chilense* LA2779 remained very low, resembling the expression levels of the *ty-1* allele from cultivated tomato. This is in agreement with the result that Fla.8383 carries a susceptible *RDR* allele and the TYLCV resistance in this line is conferred by another gene located on chromosome 10 (*Ty-6*). Similarly, transcript levels of the accessions *S. arcanum* LA2157, *S. habrochaites* LYC4 and *S. pennellii* LA716 were comparable to those of the susceptible allele.

However, two accessions of *S. chilense* (G1.1556 and G1.1558) with resistance governed by genes of recessive nature (data not shown) also showed a significantly higher level of expression of the *RDR* compared to MM. These results suggest that even though a high expression of the *RDR* is necessary for the *Ty-1/Ty-3* mediated resistance, it is not exclusively responsible of the resistant response.

Discussion

Recently we cloned the *S. chilense* derived TYLCV resistance genes *Ty-1* and *Ty-3* and found that they are allelic and code for an RNA-dependent RNA polymerase (*RDR*) of the DFDGD class (Verlaan et al. 2013). In this study we show, based on fine-mapping and/or VIGS, that functional *Ty-1*/*Ty-3*-like alleles are present in *S. chilense* accessions LA1932, LA1938 and likely also in LA1960 and LA1971. We also show that the DFDGD motif in the catalytic domain of the *Ty-1* gene is conserved among cultivated tomato and several wild species in the tomato clade. Three *Ty-1/Ty-3* specific amino acids were identified among TYLCV-resistant *S. chilense* accessions, each genotype representing different *RDR* alleles. These specific amino acids in concomitance with high gene expression level are indicative of *Ty-1/Ty-3* mediated resistance. An insertion of 12 base pairs at the 5-prime part of the coding sequence is however found in *S. chilense*-derived alleles but also in several other wild *Solanum* species of which some are known to be susceptible to TYLCV.

One *S. chilense* accesssion can harbour more than one TYLCV resistance gene

Many *S. chilense* accessions, including LA1969, LA1932, LA1938, LA2779, LA1960 and LA1971 are resistant to TYLCV. The resistant *Ty-1* and *Ty-3* alleles were originally identified in LA1969 and LA2779, respectively. Previous studies have

mapped the resistance from LA1932, LA1960 and LA1971 accessions to chromosome 6, a region overlapping with the *Ty-1/Ty-3* interval (Pérez de Castro et al. 2013); therefore, it was expected that these accessions harbour a *Ty-1/Ty-3* allele. The data from our VIGS experiments showed the existence of *Ty-1/Ty-3* allelic variants which control TYLCV resistance in LA1932 and LA1938. However, this hypothesis was not confirmed by the VIGS experiments for accession LA1971, since silencing the *Ty-1/Ty-3* gene did not compromise the resistance in line 1594 derived from it. Expression analysis of the *RDR* in this line showed the highest transcript levels among all the resistant lines tested, about 80 times higher compared to the susceptible *ty-1* allele levels (Figure S6). Complete suppression of VIGS-targeted genes in tomato is rarely observed (Sahu et al. 2012), thus the inability of this silencing approach to repress such high expression levels might have caused the unexpected resistant phenotype. An alternative possibility is that line 1594 may carry, in addition to a *Ty-1* allele, another TYLCV resistance gene derived from LA1971. It is worthwhile to note that the LA1971-derived line 1594 used in the VIGS experiments in this study is "sister" of the lines described in the previous paper of Pérez de Castro et al. (2013). Checking the LA1971 introgressions in this line, it appeared that line 1594 has multiple introgressions located on chromosomes 6, 7, 10 and 11 (see Fig. 2 in Pérez de Castro et al. 2013). Interestingly, the introgressions on chromosome 10 and 11 overlap with the intervals where *Ty-6* derived from *S. chilense* LA2779 (Hutton and Scott 2013) and *Ty-2* from *S. habrochaites* B6013 (Yang et al. 2014) are mapped, respectively. Therefore, the presence of a second resistance gene could explain why line 1594, with *S. chilense* LA1971 derived resistance, remained symptomless after silencing of *Ty-1*/*Ty-3* followed by TYLCV inoculation. Similarly, a resistant response after silencing the *RDR* allele in the *S. chilense* LA2779-derived line Fla.8383 was observed. Sequence analysis revealed that this line does not contain the *Ty-1/Ty-3* resistant allele-specific polymorphisms, and transcript levels of the *RDR* in this line resemble those of the susceptible *ty-1* allele (Figure S6). As Fla.8383 is devoid of a functional *Ty-1/Ty-3* allele the TYLCV resistance in this line is probably conferred by *Ty-6*. We are further genotyping these lines to verify our hypothesis. Alternatively, a mapping approach on populations segregating for only one introgression would be helpful in solving the puzzle. Future cloning of *Ty-2* and *Ty-6* would allow silencing of these two genes in lines 1594 and Fla.8383 to confirm this hypothesis.

The wild tomato species *S. chilense* is self-incompatible and thus heterogeneous, leading to multiple alleles of the same gene present in one accession (Bai et al. 2004). As shown in LA1932, resistant alleles of both *Ty-1/Ty-3* and *Ty-4* are present

(Ji et al. 2009). Also in LA2779, both *Ty-3* and *Ty-6* have been identified. Similarly, LA1971 may carry alleles of *Ty-1/Ty-3* and other *Ty*-genes, e.g. *Ty-6*. Thus, depending on selection procedures and heterogeneity present in *S. chilense*, it is possible that advanced *S. chilense*-derived lines carry different resistance genes for TYLCV resistance. Pyramiding of different *Ty*-genes could possibly provide higher resistance levels and/or broaden the resistance to a wider range of begomoviruses. Therefore, when a species is shown to be resistant to multiple viruses, it is possible—even probable—that more than one gene is contributing to the broad-spectrum resistance. For example, *S. chilense* accession LA1932 was found to be resistant to *Tomato mottle virus* (ToMoV) and TYLCV (Ji et al. 2009; Scott et al. 1996). It is worthwhile to test whether *Ty-1* or *Ty-4* confers resistance to both viruses. These genes should be studied more deeply in order to understand their specificity and effectiveness.

It is unfortunate that we did not have enough seeds of an advanced introgression line derived from the accession LA1960 for VIGS. Previous mapping data showed that a *Ty-1* allele is likely present in this accession (Pérez de Castro et al. 2013). However, we cannot rule out the possibility of the presence of other *Ty* genes in accession LA1960. As shown in Pérez de Castro et al. (2013), the introgression line generated by selecting for TYLCV resistance carries multiple LA1960 fragments, including the *Ty-1* region on chromosome 6 and the *Ty-6* region on chromosome 10.

Ty-1/Ty-3 mediated resistance is determined by allele-specific polymorphisms in concomitance with high expression levels of the *RDR*

In our previous study (Verlaan et al. 2013) we detected a 12 base pairs insertion in the 5 prime part of the coding sequence in the resistant *Ty-1*/*Ty-3* allele and proposed this polymorphism as the most striking difference between the *Ty-1*/*Ty-3* and *ty-1* alleles. Here we show that this 12 base pairs insertion is present in a set of 8 lines/accessions containing different *S. chilense* alleles as well as in the related wild species *S. arcanum*, *S. corneliomulleri*, *S. peruvianum, S. huaylasense*, *S. habrochaites* and *S. pennellii*, evidencing that this feature is not *S. chilense*-specific. Since some of these species, e.g. *S. arcanum* LA2157 and *S. habrochaites* LYC4 have exhibited virus symptoms after TYLCV infection (data not shown), we conclude that this insertion cannot be used as a *Ty-1/Ty-3* specific marker. By further analysing the *RDR* coding regions, we succeeded in finding five SNPs present in different exons that are specific to the *Ty-1/Ty-3* allele. These SNPs can be exploited to generate in-gene markers. Further, four SNPs were shown to be unique to *Ty-3*, useful for allele-specific marker development. In addition, the origin of the *Ty-1/Ty-3* alleles can be traced by accession-specific SNPs (Figure S4).

In a previous study we found that the resistant *Ty-1* allele was more highly expressed than the susceptible *ty-1* allele (Verlaan et al. 2013). In this study we observed comparable results where the expression of all *Ty-1/Ty-3* resistant alleles are significantly higher than the susceptible allele. However, the expression level varied among different *RDR* alleles (Figure S6). Surprisingly, elevated expression levels of the *RDR* were also detected *in S. chilense* accessions G1.1556 and G1.1558, which carry a susceptible *ty-1* allele. Therefore, we conclude that the expression level of the *RDR* is not solely responsible for the resistance but this feature together with the *Ty-1/Ty-3* allele specific amino acid sequence determine the resistance response.

The same set of tomato accessions/lines was also used to compare the typical DFDGD catalytic domain of the *RDR*γ type to which the *Ty-1* gene belongs. No SNPs were found in the domain nor in 12 base pairs up- or downstream of this domain. Further, no differences were found among a *Ty-2* carrying line, a wild *S. pimpinellifolium* and the same nine *S. chilense*-derived lines described before. The region was also compared amongst the same set of cultivated lines and wild tomato accessions. It was found that the catalytic domain was conserved and no SNPs were found in any of the accessions/lines tested. This could indicate this gene is important for the plant and that SNPs in the catalytic domain have a negative effect on plant fitness.

In conclusion, this study shows that probably many *S. chilense* accessions carry a TYLCV resistance locus on chromosome 6, allelic to *Ty-1*/*Ty-3*. Fine mapping and/or more VIGS experiments could prove whether this is really true. The catalytic domain of the *Ty-1*/*Ty-3* gene is conserved among *Solanum* species. The 12 base pair insertion in *Ty-1*/*Ty-3* is present in *S. chilense* and in six other wild *Solanum* species, and not exclusively linked to TYLCV resistance. To develop allele-specific markers, SNPs unique to the resistant *Ty-1/Ty-3* alleles can be used. Moreover, our study shows that (1) VIGS can be applied as a tool for testing allelism, and (2) more than one TYLCV resistance gene can be present in one *S. chilense* accession.

Materials and methods

Plant materials

Two breeding lines (developed in Spain) were used for the VIGS experiment: 1538, derived from *S. chilense* LA1932 and corresponding to line 2 in Pérez de Castro et

al. (2013) and 1594, derived from *S. chilense* LA1971 and corresponding to line 5 in Pérez de Castro et al. (2013). The line B26, progeny of one homozygous resistant F_2 plant derived from LA1932, was also included (Table 1).

In Florida, four breeding lines with begomovirus resistance derived from different *S. chilense* sources were developed through the University of Florida tomato breeding program. Resistance to either TYLCV and/or tomato mottle virus was selected phenotypically over multiple seasons. Fla. 8680 (Verlaan et al. 2013) and Fla. 8383 both have resistance derived from accession LA2779. Fla. 8783 is a small-fruited line with resistance from accession LA1932, and Fla. 976 has resistance derived from LA1938. Resistance in each line, with exception of Fla. 8383, was determined previously to co-segregate with a *S. chilense* introgression on chromosome 6 and spanning the *Ty-1/Ty-3* locus.

TYLCV inoculation

For TYLCV tests, an infectious TYLCV-IL clone (pTYCz40a) originating from Israel was used for agroinoculation using the method as described in Verlaan et al. (2011). Briefly, *A. tumefaciens* LBA4404 was transformed, cultured in LB, pelleted and resuspended in infiltration medium at an OD_{600} of 0.5. Seeds were sown and plants were kept under greenhouse conditions at a temperature of 23 °C and relative humidity of 60% during a 16-hour day / 8-hour night regime.Three week old seedlings were infiltrated by pressure inoculation in the leaves with a needle-less syringe. For the VIGS experiments the agro infiltration was done two weeks after TRV inoculation.

Briefly, plants were inoculated at 3-4 true-leaf stage during seven days in a climatic chamber inside muslin-covered cages. After this period, plants were transplanted in a greenhouse with controlled temperature until the end of the assay. Symptom severity was scored at 15, 25, 35, 45 and 55 days post inoculation using a scale (Friedmann et al. 1998) from 0 (no visible symptoms) to 4 (very severe symptoms; plants cease to grow). The limit to classify individual plants as resistant or susceptible was established at symptom score 2, based on previous studies (Pérez de Castro et al. 2007). Plants scored under 2 were considered resistant, given that no significant yield losses were expected as a consequence of infection, while plants scored 2 or higher were considered susceptible.

TRV based VIGS

For the silencing experiments TRV constructs and procedures as described in (Verlaan et al. 2013) were used. Briefly, *A. tumefaciens* strain GV3101 containing

the TRV replicons were cultured, pelleted and resuspended in infiltration medium. Agro infiltration was performed on cotyledons of 10-day old seedlings using pressure inoculation.

RNA isolation and cDNA synthesis

For sequence analysis 3 week old seedlings were agro inoculated as described above. Three weeks after agroinoculation top leaves of plants were harvested and grinded in liquid nitrogen using mortar and pestle. Total RNA was extracted by using the RNeasy Plant Mini Kit (Qiagen) as described by the manufacturer. One µg RNA was digested using DNase I (Amp. Grade) following the manufacturers protocol (Invitrogen) and cDNA was synthesized using the iScript cDNA Synthesis Kit following the protocol (Bio-Rad).

Sequence analysis of the *S. chilense* derived lines and accessions

For amplifying the region containing the 5 prime deletion primers Del-F1 (5'-TTCAAGTATATACAGGAAAAATGGGTGATCCG-3') and Del-R1 (5'-CTGAGGGCTTGCACAGGCCAAT-3') were used. For amplifying the region containing the catalytic domain, primers DFDGD-F4 (5'-GGGCGTGTTTTGGTCTACAG-3') and DFDG-R4 (5'-GCTATCAGCTGCCAGAGACAT-3') were used. PCR amplification was performed according to standard protocols in an Eppendorf Mastercycler Pro. Amplified fragments were sequenced and analyzed using SeqMan Pro 9 (DNA Star). Alignments were made with MEGA version 5 (Tamura et al. 2011).

The *RDR* cDNA sequence from *S. lycopersicum* 'Moneymaker'(MM), the *Ty-1* allele from *S. chilense* LA1969 and the *Ty-3* allele from *S. chilense* LA2779 were described by Verlaan et al. (2013) and published in Patent No. WO2012125025 & US2014208459 (SEQ 1 = *Ty-1*; SEQ 3 = MM). The *RDR* cDNA sequence from *S. peruvianum* was obtained from the SGN *S. peruvianum* de-novo transcriptome (a19742). Genomic sequences/contigs were available for *Solanum lycopersicum* M82, *S. pimpinellifolium* LA1589, *S. arcanum* LA2157, *S. pennellii* LA716, *S. habrochaites* LYC4 (NCBI WGS whole genome shotgun contigs data), *S. chilense* introgression line Gh13 and *S. chilense* introgression line BTI-87 (SGN database Tomato Inbred Lines). *RDR* exons were extracted from the genomic sequence based on homology with the *S. lycopersicum* 'Moneymaker ' *RDR* allele. In the *S. habrochaites* LYC4 genomic sequence part of the *RDR* gene was missing, i.e. a large part of intron 8 and exon 9. This was confirmed by PCR analysis and sequencing. RNA was isolated from TYLCV-infected *S. chilense* LA1932, LA1938, LA1971, G1.1556 and G1.1558, 19 days post

infection. cDNA was prepared and full length *RDR* cDNA sequences were obtained by PCR with primers Ty-F7 and Ty-R5 (Verlaan et al. 2013). Sequences were determined from nested PCR products with primers RDR-F3+R10, RDR-F7+R7, RDR-F6+R4, RDR-F4+R5 (Figure S4 in Verlaan et al. 2013).

The *RDR* cDNA sequence from *S. chilense* G1.1556 was smaller than the expected size of approximately 3 kb. Nested PCR with primers RDR-F7 and RDR-R7 was not successful. Therefore, exons 9 -14 were amplified from genomic DNA, although we could not verify whether they are included in the transcript. The PCR product with primers RDR-F3 and RDR-R10 was larger than the expected 1068 bp and proved to contain intron 17.

De novo assembled wild species genomes and re-sequencing collection and analysis of the deletion and the catalytic domain

Data of the 84 accessions of the 100 tomato genome re-sequencing consortium (The 100 Tomato Genome Sequencing Consortium 2014) were obtained from the European Nucleotide Archive (http://www.ebi.ac.uk/ena/; project PRJEB5235). The *de novo* assemblies, of *S. arcanum* LA2157, *S. habrochaites* LYC4, and *S. pennellii* LA716 were obtained from the same resource and are available under the project numbers PRJEB5226, PRJEB52267 and PRJEB52268 respectively. In short, 84 tomato and related wild species were re-sequenced with a read depth of approximately 42x (The 100 Tomato Genome Sequencing Consortium 2014). For a list of sequenced species, and their variants, we refer to http://www.tomatogenome.net. Sequence reads were mapped to the reference genome of *S. lycopersicon* cv. Heinz version SL2.40 (The Tomato Genome Consortium 2012) using BWA (Li and Durbin, 2009), SNP and INDELS were called using samtools (Li et al. 2009) and saved in the variant call format (VCF) (Danecek et al. 2011). Variants were visually inspected using the Integrative Genomics Viewer (IGV) (Robinson et al. 2011).

Phylogenetic analysis

A unrooted neighbour joining tree was constructed from multiple sequence alignment using MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/).

Quantitative RT-PCR

For the gene expression experiment, leaf samples of the top part of each plant were taken 0 and 19 days after TYLCV inoculation; the mock treatment consisted of infiltration media without bacteria. Total RNA extraction, cDNA synthesis and Quantitative Real Time PCR were performed as described in Verlaan et al. (2013).

For RT-PCR of *Ty-1/ty-1*, primers 180-F1 and 180-R2 were used. The actin (ACT) gene was used as reference, using primers ACT-F and ACT-R; gene expression levels were calculated using the ΔΔCt method (Verlaan et al. 2013).

Acknowledgements

The infectious TYLCV clone was kindly provided by Professor Eduardo Rodríguez Bejarano (Universidad de Málaga). We thank Dick Lohuis for his help with agro-infiltrations, Marc Hendriks and Marjon Arens for RNA isolation and sequencing. This project was financed by the Centre for BioSystems Genomics (CBSG) which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (http://www.cbsg.nl). Olga Julián was granted a scholarship by Generalitat Valenciana.

Supplementary files

Figure S1. Physical chromosome locations of mapped tomato genes conferring resistance to TYLCV. Schematic representation of chromosome location of *Ty-1, Ty-3, Ty3-A* (incompletely dominant, Verlaan et al. 2013, Ji et al. 2007, Scott et al. 1996), *Ty-2* (dominant, Yang et al. 2014), *Ty-4* (incompletely dominant, Ji *et al.* 2009), *ty-5* (recessive, Anbinder et al. 2009) and *Ty-6* (Hutton and Scott 2013). Source of *ty-5* is tomato breeding line TY172, derived from 4 different accessions of *Solanum peruvianum*. Grey shaded regions represent pericentromeric heterochromatin; approximate physical positions are shown on the left side of chromosomes and represent millions of basepairs.

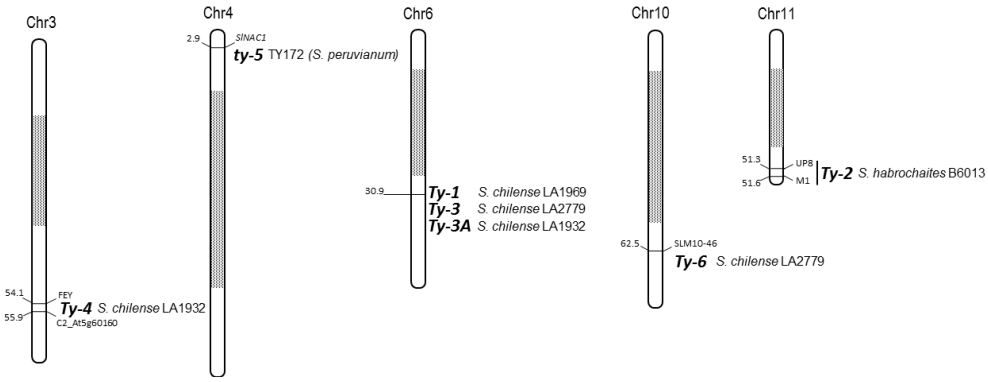


Figure S2. Alignment of sequences of the region containing the 5-prime insertion in the *Ty-1* allele. All *Solanum chilense* derived lines have the 12 base pair insertion (A). There is one non-synonymous SNP in *S. arcanum* LA2157 and *S. habrochaites* LYC4 (B). Of the multiple species tested, six had the insertion, e.g. *S. arcanum, S. corneliomulleri*, *S. peruvianum*, *S. huaylasense*, *S. habrochaites* and *S. pennellii* (B, C and D). Sequences from (A) have been obtained from cDNA, sequences from (B) have been obtained from a *de novo* assembly of these three accessions (C) and (D) have been obtained from whole genome re-sequencing. Note: Read-mapping information of *S. habrochaites* and *S. pennellii* against Heinz was ambiguous and thus cautions need to be taken for using data of these two species.

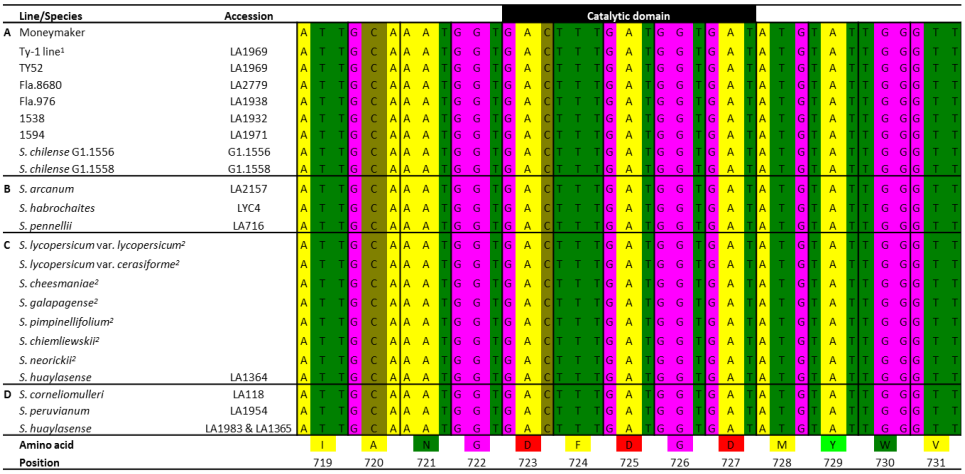
	Line/Species	Accession																								
A	Moneymaker		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	Ty-1 line[1]	LA1969	G	T	T	C	C	T	T	C	T	T	G	T	A	T	A	C	T	G
	TY52	LA1969	G	T	T	C	C	T	T	C	T	T	G	T	A	T	A	C	T	G
	Fla.8680	LA2779	G	T	T	C	C	T	T	C	T	T	G	T	A	T	A	C	T	G
	Fla.976	LA1938	G	T	T	C	C	T	T	C	T	T	G	T	A	T	A	C	T	G
	1538	LA1932	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	1594	LA1971	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	S. chilense G1.1556	G1.1556	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	S. chilense G1.1558	G1.1558	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
B	*S. arcanum*	LA2157[2]	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	G	T	G
	S. habrochaites	LYC4[2]	G	T	T	T	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	S. pennellii	LA716	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
C	*S. lycopersicum* var. *lycopersicum*[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. lycopersicum var. *cerasiforme*[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. cheesmaniae[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. galapagense[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. pimpinellifolium[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. chiemliewskii[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. neorickii[3]		G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
	S. huaylasense	LA1364	G	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	C	T	G
D	*S. corneliomulleri*	LA118	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	S. peruvianum	LA1954	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	S. huaylasense	LA1983 & LA1365	G	T	T	C	C	T	T	C	G	T	G	T	A	T	A	C	T	G
	Amino acid			V			P			S			C			I			L	
	Position			11			12			13			14			15			16	

[1]Ty-1 line derived from Tygress (see: Verlaan et al., 2013)

[2]Leads to amino acid change P->S L->V

[3]Several Accessions (see: http://www.tomatogenome.net)

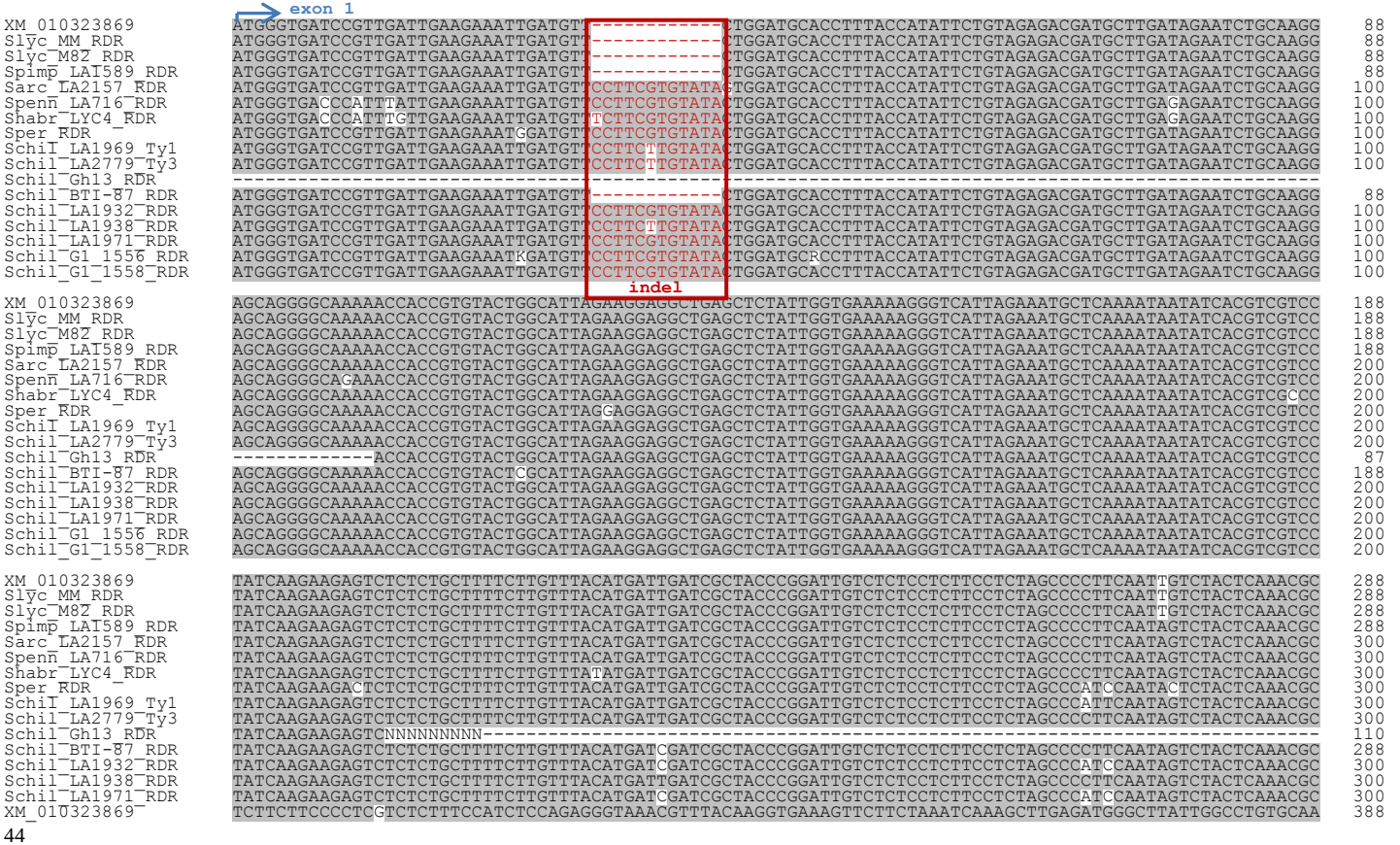
Figure S3. Alignment of cDNA sequences of the region containing the catalytic domain. All *S. chilense* derived lines have an identical sequence in this region. Accession from 14 *Solanum* species also have the same sequence. All species in the full genome data set were also analysed but no SNPs were observed.

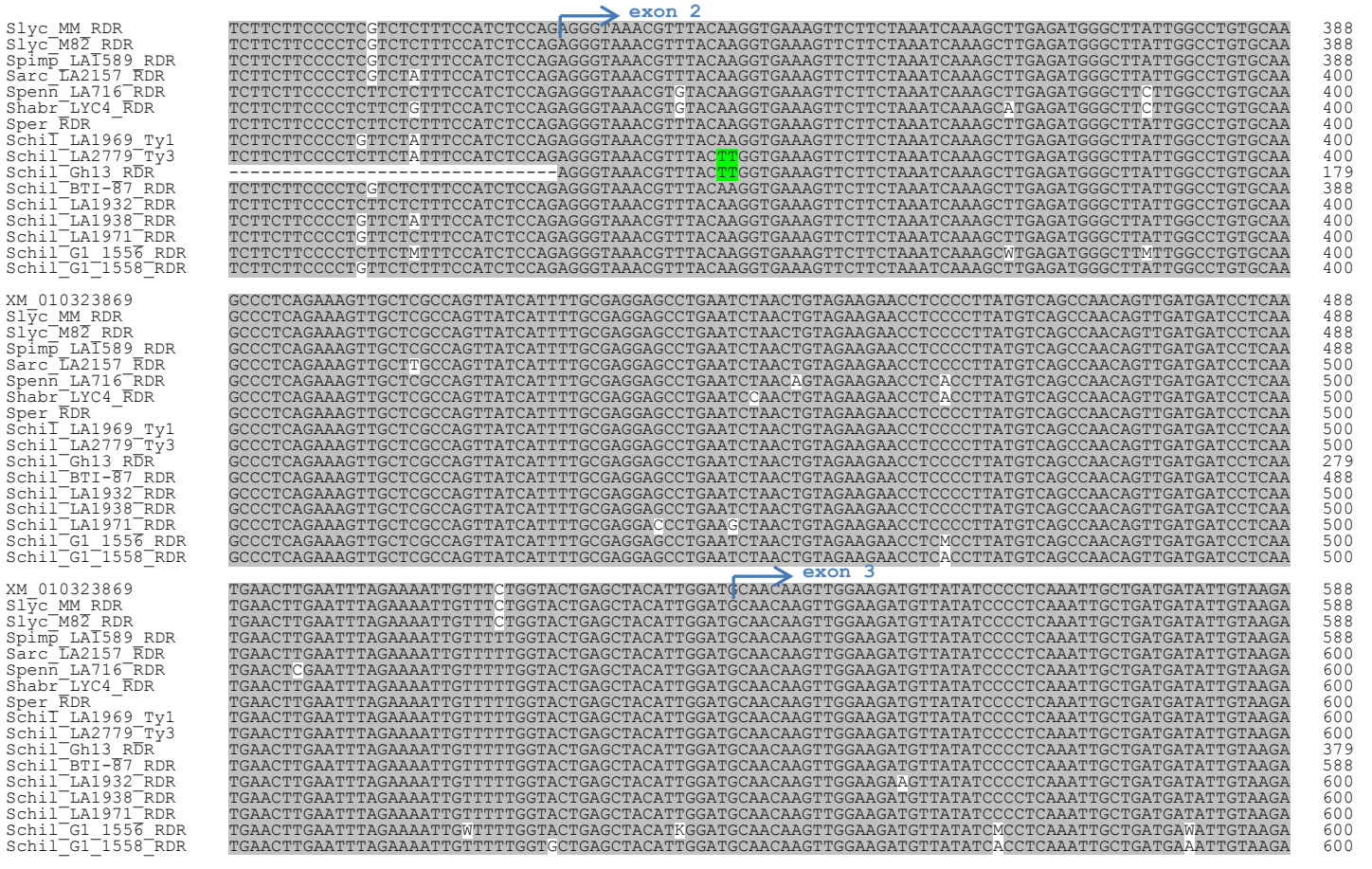


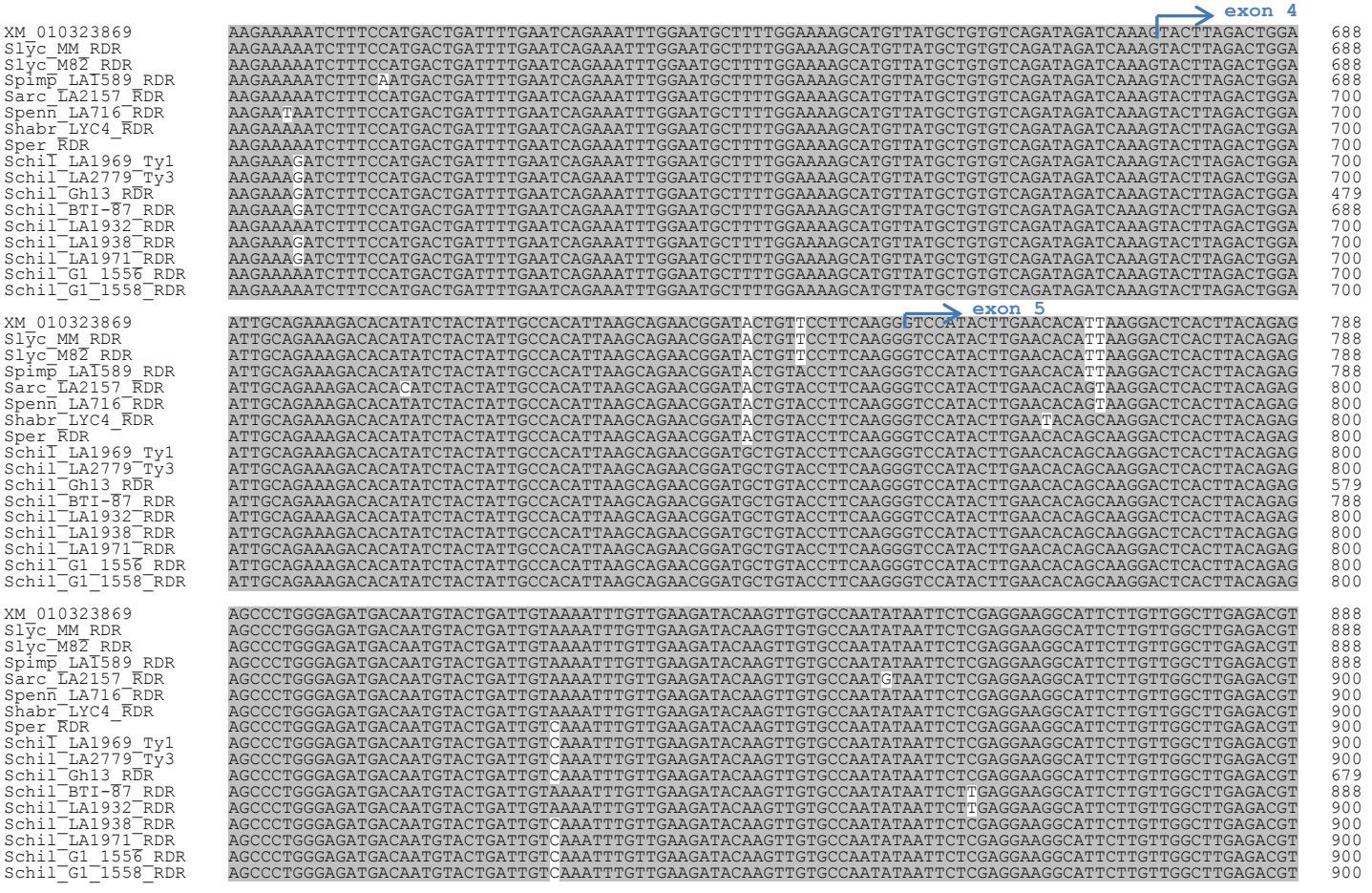
¹Ty-1 line derived from Tygress (see: Verlaan et al., 2013)

²Several Accessions (see: http://www.tomatogenome.net)

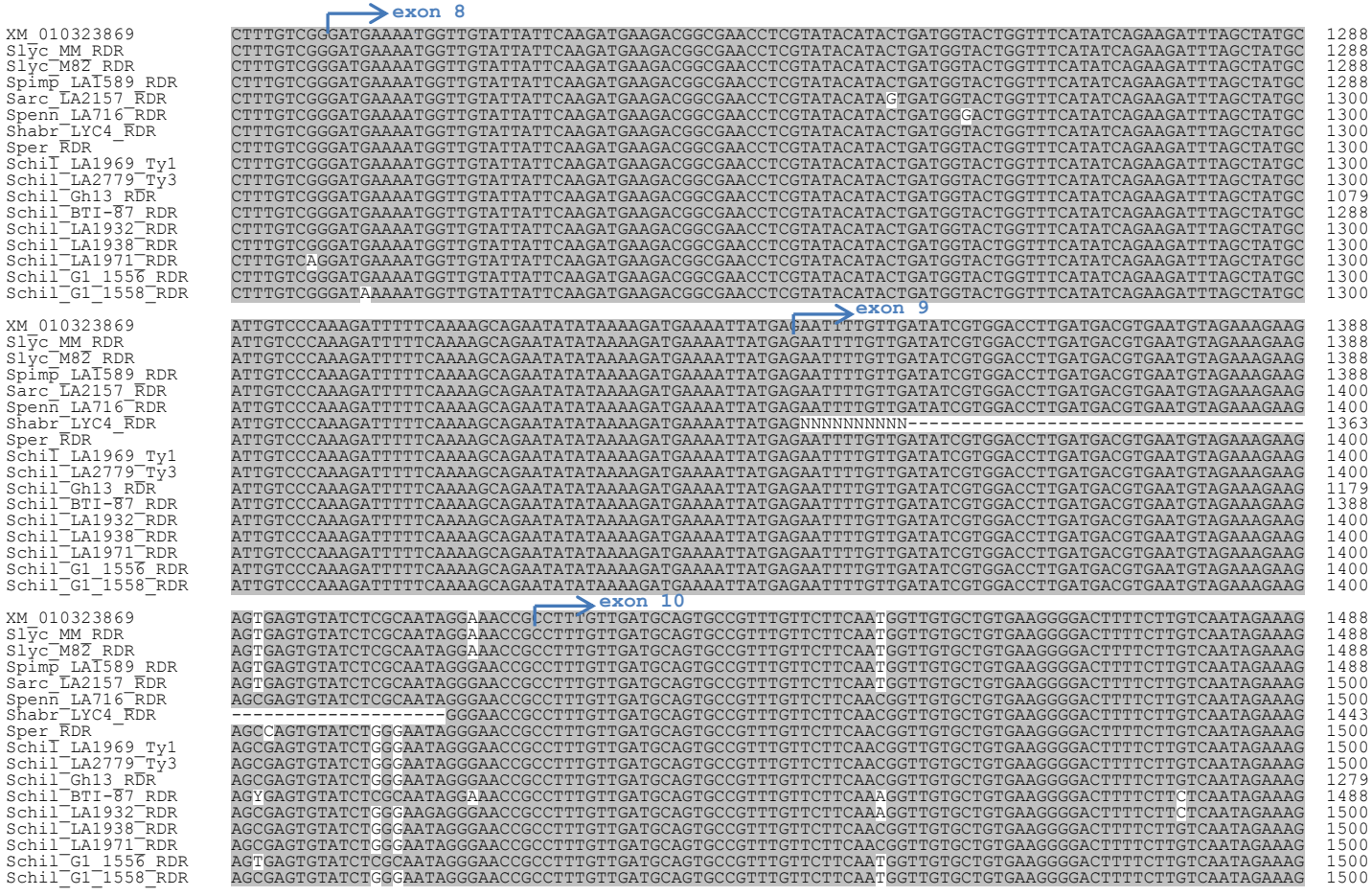
Figure S4. Figure S4. Alignment of full-length cDNA sequences of the *Ty-1/Ty-3* RDR. Sequences of two *S. lycopersicum* lines, nine *S. chilense*-derived lines/accessions and five related *Solanum* accessions were obtained and compared to explore for allele-specific polymorphisms. Start positions of the exons are indicated. The 5' indel and the catalytic domain are highlighted in red. The premature stop codon in *S. pennellii* LA716 is highlighted in red. Five *Ty-1/Ty-3* specific SNPs are highlighted in yellow; four *Ty-3* specific SNPs are highlighted in green.



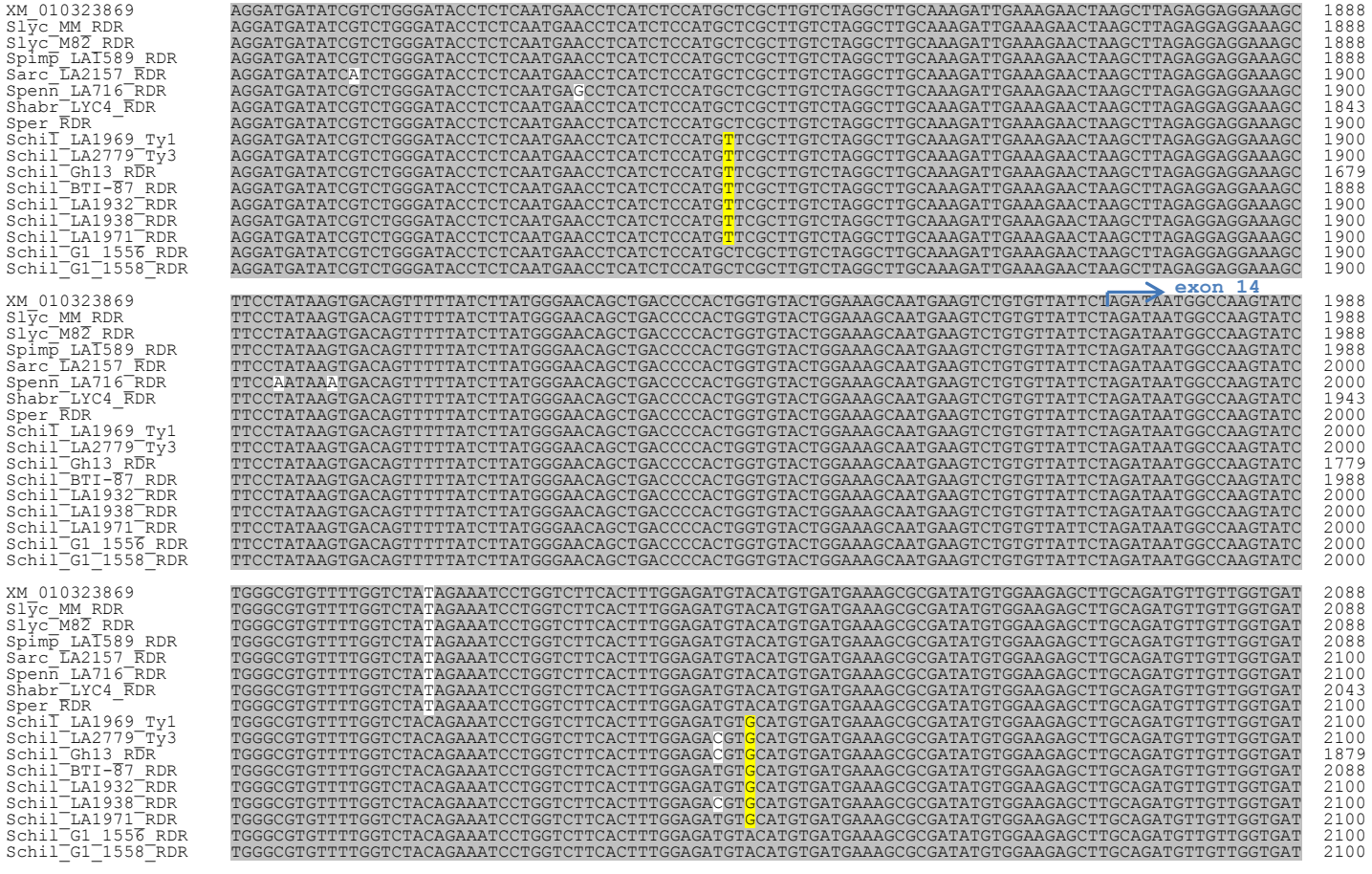

exon 2
exon 3
Slyc_MM_RDR
Slyc_M82_RDR
Spimp_LA1589_RDR
Sarc_LA2157_RDR
Spenn_LA716_RDR
Shabr_LYC4_RDR
Sper_RDR
Schil_LA1969_Ty1
Schil_LA2779_Ty3
Schil_Gh13_RDR
Schil_BTI-87_RDR
Schil_LA1932_RDR
Schil_LA1938_RDR
Schil_LA1971_RDR
Schil_G1_1556_RDR
Schil_G1_1558_RDR
XM_010323869

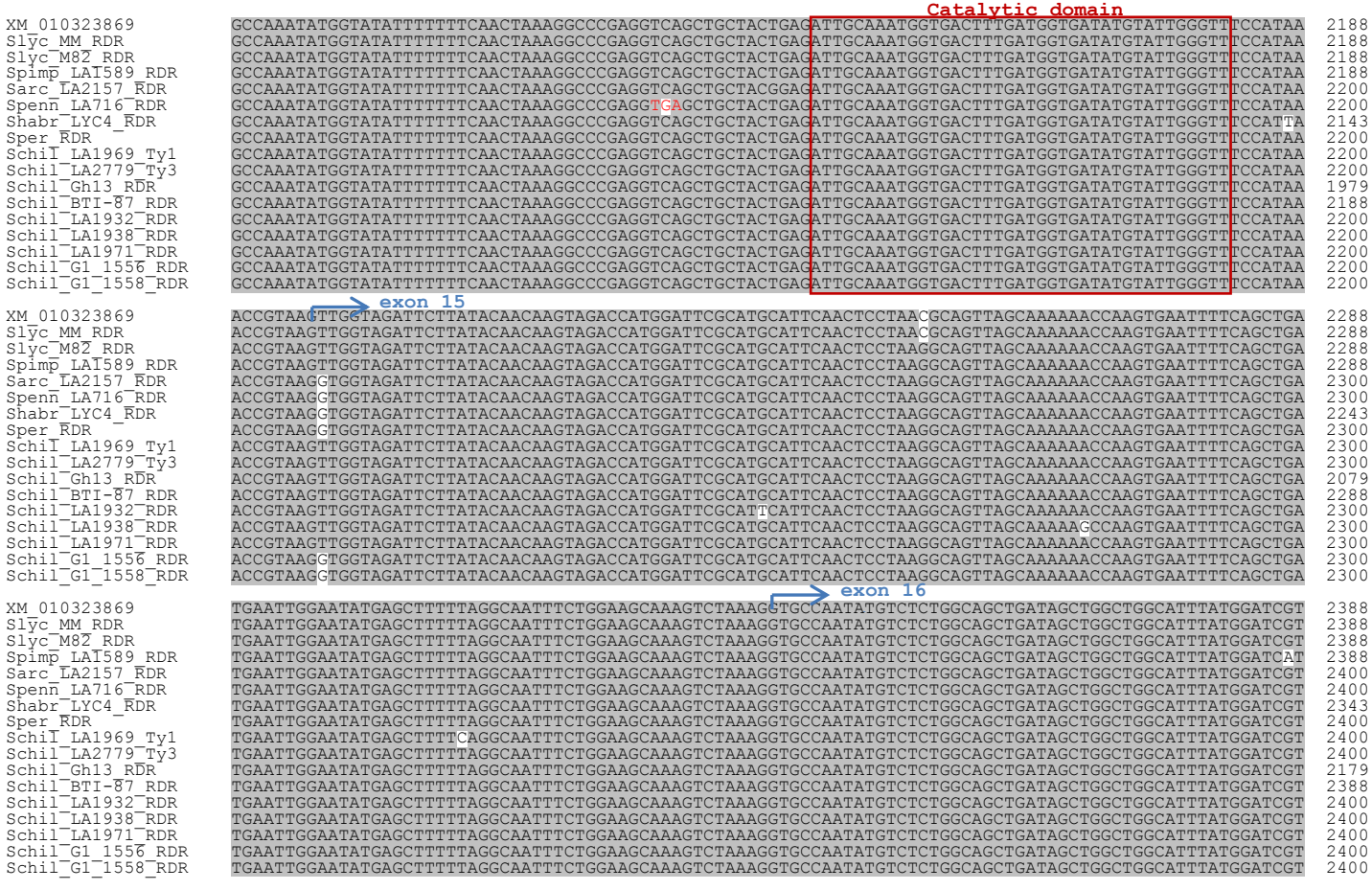

exon 4
exon 5
XM_010323869 AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 688
Slyc_MM_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 688
Slyc_M82_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 688
Spimp_LA1589_RDR AAGAAAAATCTTTCAATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 688
Sarc_LA2157_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Spenn_LA716_RDR AAGAATAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Shabr_LYC4_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Sper_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_LA1969_Ty1 AAGAAAGATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_LA2779_Ty3 AAGAAAGATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_Gh13_RDR AAGAAAGATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 479
Schil_BTI-87_RDR AAGAAAGATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 688
Schil_LA1932_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_LA1938_RDR AAGAAAGATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_LA1971_RDR AAGAAAGATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_G1_1556_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
Schil_G1_1558_RDR AAGAAAAATCTTTCCATGACTGATTTTGAATCAGAAATTTGGAATGCTTTTGGAAAAGCATGTTATGCTGTGTCAGATAGATCAAAGTACTTAGACTGGA 700
XM_010323869 ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTTCCTTCAAGGGTCCATACTTGAACACATTAAGGACTCACTTACAGAG 788
Slyc_MM_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTTCCTTCAAGGGTCCATACTTGAACACATTAAGGACTCACTTACAGAG 788
Slyc_M82_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTTCCTTCAAGGGTCCATACTTGAACACATTAAGGACTCACTTACAGAG 788
Spimp_LA1589_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTACCTTCAAGGGTCCATACTTGAACACATTAAGGACTCACTTACAGAG 788
Sarc_LA2157_RDR ATTGCAGAAAGACACACATCTACTATTGCCACATTAAGCAGAACGGATACTGTACCTTCAAGGGTCCATACTTGAACACAGTAAGGACTCACTTACAGAG 800
Spenn_LA716_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTACCTTCAAGGGTCCATACTTGAACACAGTAAGGACTCACTTACAGAG 800
Shabr_LYC4_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTACCTTCAAGGGTCCATACTTGAATACAGCAAGGACTCACTTACAGAG 800
Sper_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATACTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_LA1969_Ty1 ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_LA2779_Ty3 ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_Gh13_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 579
Schil_BTI-87_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 788
Schil_LA1932_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_LA1938_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_LA1971_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_G1_1556_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
Schil_G1_1558_RDR ATTGCAGAAAGACACATATCTACTATTGCCACATTAAGCAGAACGGATGCTGTACCTTCAAGGGTCCATACTTGAACACAGCAAGGACTCACTTACAGAG 800
XM_010323869 AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 888
Slyc_MM_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 888
Slyc_M82_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 888
Spimp_LA1589_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 888
Sarc_LA2157_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATGTAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Spenn_LA716_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Shabr_LYC4_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Sper_RDR AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_LA1969_Ty1 AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_LA2779_Ty3 AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_Gh13_RDR AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 679
Schil_BTI-87_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTTGAGGAAGGCATTCTTGTTGGCTTGAGACGT 888
Schil_LA1932_RDR AGCCCTGGGAGATGACAATGTACTGATTGTAAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTTGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_LA1938_RDR AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_LA1971_RDR AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_G1_1556_RDR AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900
Schil_G1_1558_RDR AGCCCTGGGAGATGACAATGTACTGATTGTCAAATTTGTTGAAGATACAAGTTGTGCCAATATAATTCTCGAGGAAGGCATTCTTGTTGGCTTGAGACGT 900

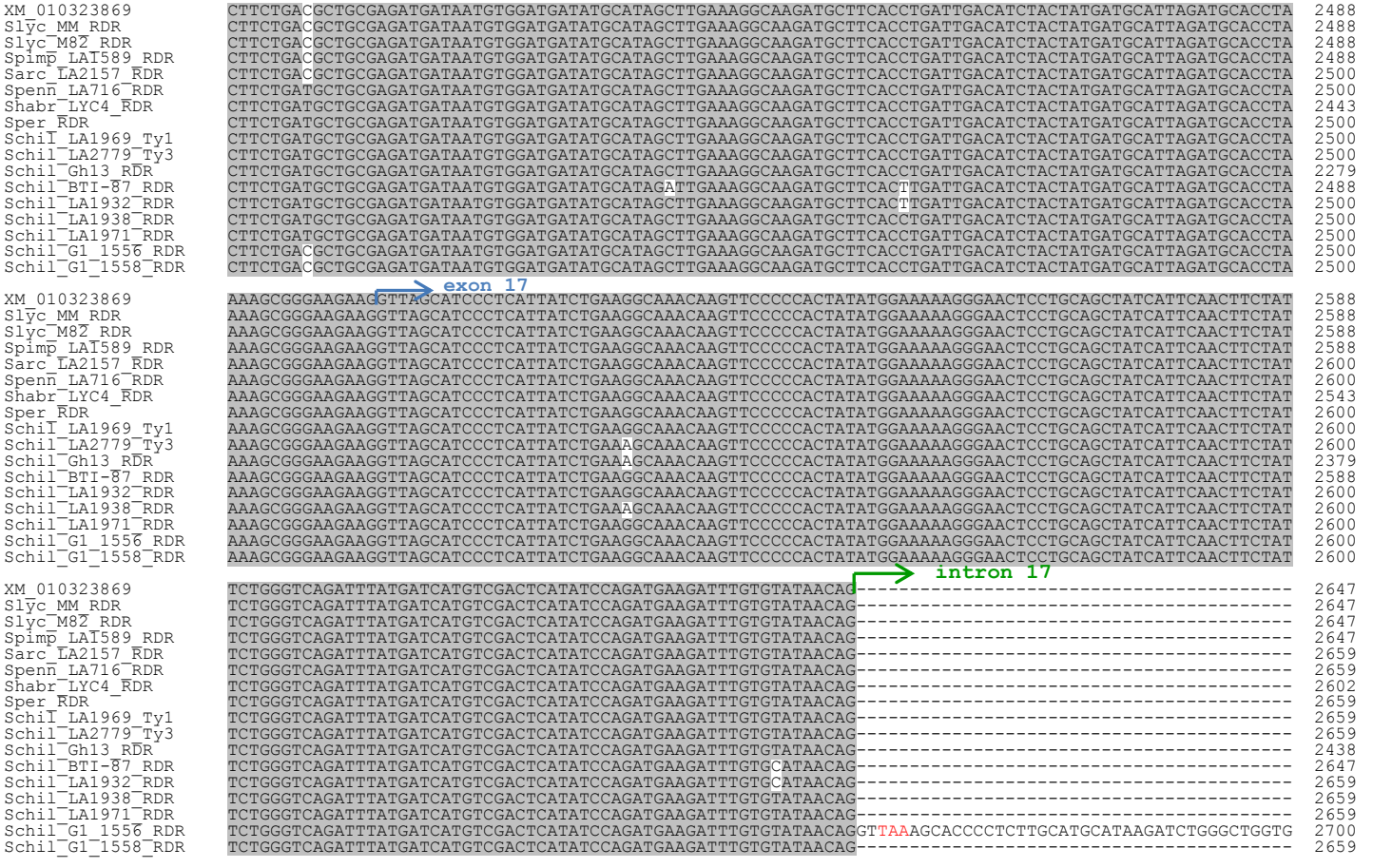

exon 6
exon 7
XM_010323869
Slyc_MM_RDR
Slyc_M82_RDR
Spimp_LA1589_RDR
Sarc_LA2157_RDR
Spenn_LA716_RDR
Shabr_LYC4_RDR
Sper_RDR
Schil_LA1969_Ty1
Schil_LA2779_Ty3
Schil_Gh13_RDR
Schil_BTI-87_RDR
Schil_LA1932_RDR
Schil_LA1938_RDR
Schil_LA1971_RDR
Schil_G1_1556_RDR
Schil_G1_1558_RDR

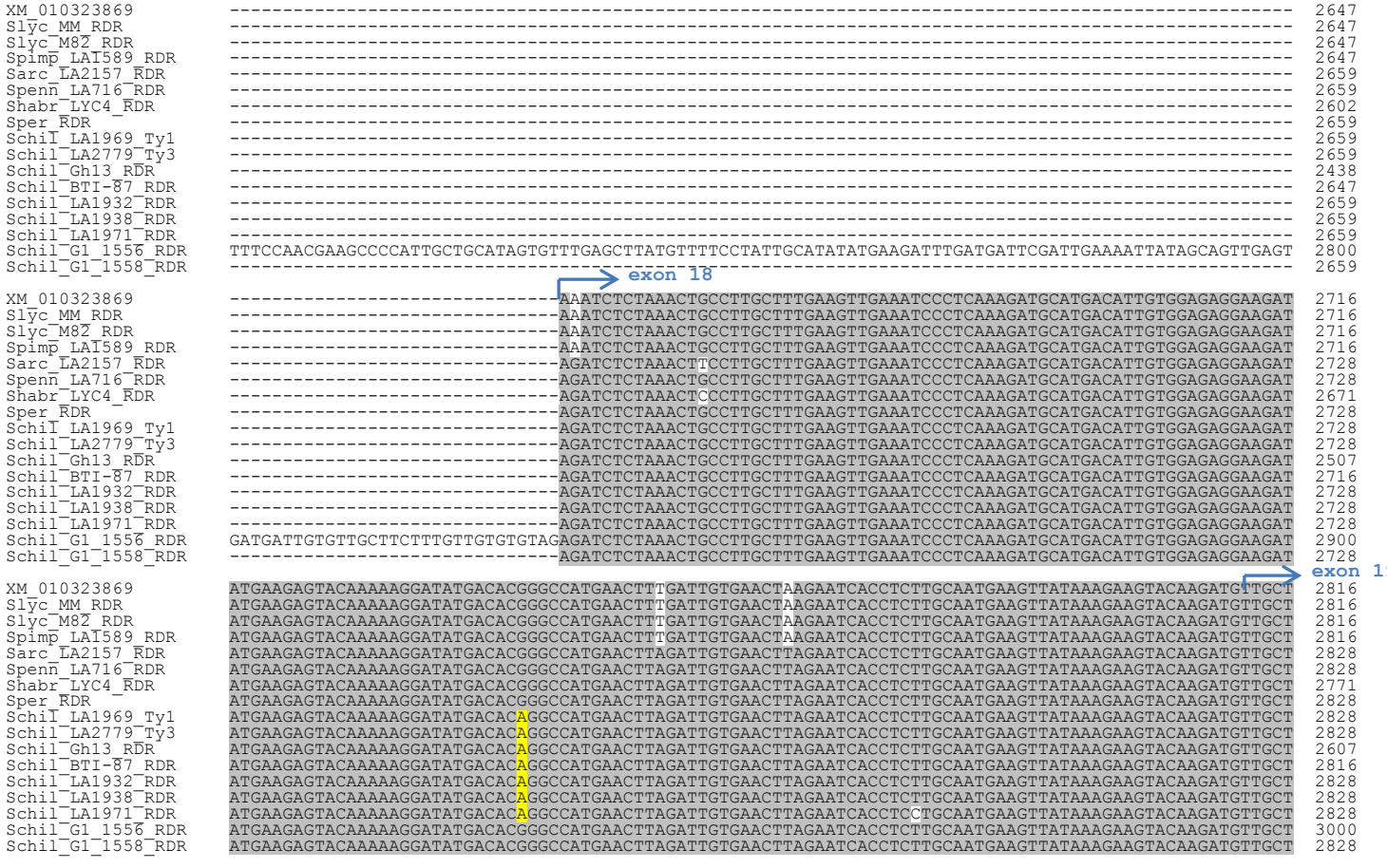

exon 8
exon 9
exon 10
XM_010323869 CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1288
Slyc_MM_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1288
Slyc_M82_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1288
Spimp_LA1589_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1288
Sarc_LA2157_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATAGTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Spenn_LA716_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGGACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Shabr_LYC4_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Sper_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_LA1969_Ty1 CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_LA2779_Ty3 CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_Gh13_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1079
Schil_BTI-87_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1288
Schil_LA1932_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_LA1938_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_LA1971_RDR CTTTGTCAGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_G1_1556_RDR CTTTGTCGGGATGAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
Schil_G1_1558_RDR CTTTGTCGGGATAAAAATGGTTGTATTATTCAAGATGAAGACGGCGAACCTCGTATACATACTGATGGTACTGGTTTCATATCAGAAGATTTAGCTATGC 1300
XM_010323869 ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1388
Slyc_MM_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1388
Slyc_M82_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1388
Spimp_LA1589_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1388
Sarc_LA2157_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Spenn_LA716_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Shabr_LYC4_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGNNNNNNNNNN---------- 1363
Sper_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_LA1969_Ty1 ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_LA2779_Ty3 ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_Gh13_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1179
Schil_BTI-87_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1388
Schil_LA1932_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_LA1938_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_LA1971_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_G1_1556_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
Schil_G1_1558_RDR ATTGTCCCAAAGATTTTTCAAAAGCAGAATATATAAAAGATGAAAATTATGAGAATTTTGTTGATATCGTGGACCTTGATGACGTGAATGTAGAAAGAAG 1400
XM_010323869 AGTGAGTGTATCTCGCAATAGGAAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1488
Slyc_MM_RDR AGTGAGTGTATCTCGCAATAGGAAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1488
Slyc_M82_RDR AGTGAGTGTATCTCGCAATAGGAAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1488
Spimp_LA1589_RDR AGTGAGTGTATCTCGCAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1488
Sarc_LA2157_RDR AGTGAGTGTATCTCGCAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Spenn_LA716_RDR AGCGAGTGTATCTCGCAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Shabr_LYC4_RDR ----------GGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1443
Sper_RDR AGCCAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Schil_LA1969_Ty1 AGCGAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Schil_LA2779_Ty3 AGCGAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Schil_Gh13_RDR AGCGAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1279
Schil_BTI-87_RDR AGYGAGTGTATCTCGCAATAGGAAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAAAGGTTGTGCTGTGAAGGGGACTTTTCTTCTCAATAGAAAG 1488
Schil_LA1932_RDR AGCGAGTGTATCTGGGAAGAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAAAGGTTGTGCTGTGAAGGGGACTTTTCTTCTCAATAGAAAG 1500
Schil_LA1938_RDR AGCGAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Schil_LA1971_RDR AGCGAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAACGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Schil_G1_1556_RDR AGTGAGTGTATCTCGCAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500
Schil_G1_1558_RDR AGCGAGTGTATCTGGGAATAGGGAACCGCCTTTGTTGATGCAGTGCCGTTTGTTCTTCAATGGTTGTGCTGTGAAGGGGACTTTTCTTGTCAATAGAAAG 1500


exon 11
exon 12
exon 13
XM_010323869 ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1588
Slyc_MM_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1588
Slyc_M82_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1588
Spimp_LA1589_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1588
Sarc_LA2157_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Spenn_LA716_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGATCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Shabr_LYC4_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1543
Sper_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_LA1969_Ty1 ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_LA2779_Ty3 ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_Gh13_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1379
Schil_BTI-87_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1588
Schil_LA1932_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_LA1938_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_LA1971_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATACCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_G1_1556_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATATCAACTTTTGACTCATTGGAGATAGTTG 1600
Schil_G1_1558_RDR ATTGGATCACGAAAAATTCATATTAGACCCTCAATGGTGAAGGTTGAGATAGACCCAACAATTTCAAGTATATCAACTTTTGACTCATTGGAGATAGTTG 1600
XM_010323869 CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAACTTTTGGG 1688
Slyc_MM_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAACTTTTGGG 1688
Slyc_M82_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAACTTTTGGG 1688
Spimp_LA1589_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAACTTTTGGG 1688
Sarc_LA2157_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAGCTTTTGGG 1700
Spenn_LA716_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAGCTTTTGGG 1700
Shabr_LYC4_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAGCTTTTGGG 1643
Sper_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAGAGAATACTTTATGGAGCTTTTGGG 1700
Schil_LA1969_Ty1 CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1700
Schil_LA2779_Ty3 CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1700
Schil_Gh13_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1479
Schil_BTI-87_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1688
Schil_LA1932_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1700
Schil_LA1938_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1700
Schil_LA1971_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTCTGGAGCTTTTGGG 1700
Schil_G1_1556_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAGCTTTTGGG 1700
Schil_G1_1558_RDR CAATCAGTCATAGACCAAATAAGGCATATCTGTCCAAGAATTTAATCTCTCTGCTGAGCTACGGAGGAGTCCATAAAGAATACTTTATGGAGCTTTTGGG 1700
XM_010323869 AAGTGCGCTGGAAGAGACGAAACAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1788
Slyc_MM_RDR AAGTGCGCTGGAAGAGACGAAACAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1788
Slyc_M82_RDR AAGTGCGCTGGAAGAGACGAAACAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1788
Spimp_LA1589_RDR AAGTGCGCTGGAAGAGACGAAACAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1788
Sarc_LA2157_RDR AAGTGCACTGGAAGAGACGAATCAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Spenn_LA716_RDR AAGTGCACTGGAAGAGACGAAACAAGTACGTTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Shabr_LYC4_RDR AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1743
Sper_RDR AAGTGCACTGGAAGAGACGAATCAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_LA1969_Ty1 AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_LA2779_Ty3 AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_Gh13_RDR AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1579
Schil_BTI-87_RDR AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1788
Schil_LA1932_RDR AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_LA1938_RDR AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_LA1971_RDR AAGTGCACTGGAAGAGACGAAACAAGTATATTTGAGGAAACGGGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_G1_1556_RDR AAGTGCACTGGAAGAGACGAATCAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800
Schil_G1_1558_RDR AAGTGCACTGGAAGAGACGAATCAAGTATATTTGAGGAAACGTGCAGCTCTAAAAGTTGCTATCAACTATAGAGAAATGGATGATGAATGTCTAACAGCA 1800


exon 14

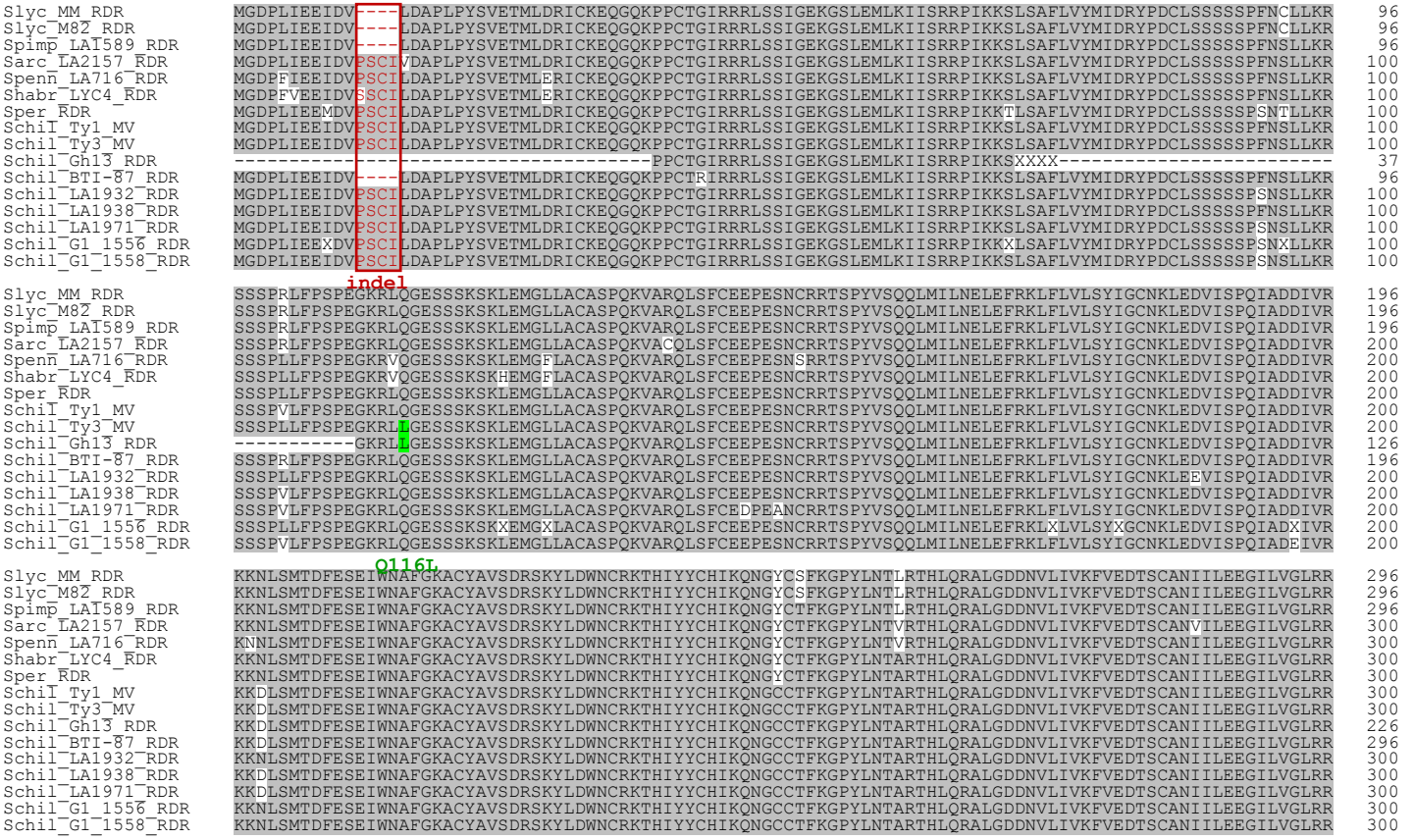

Catalytic domain
exon 15
exon 16


exon 17
intron 17

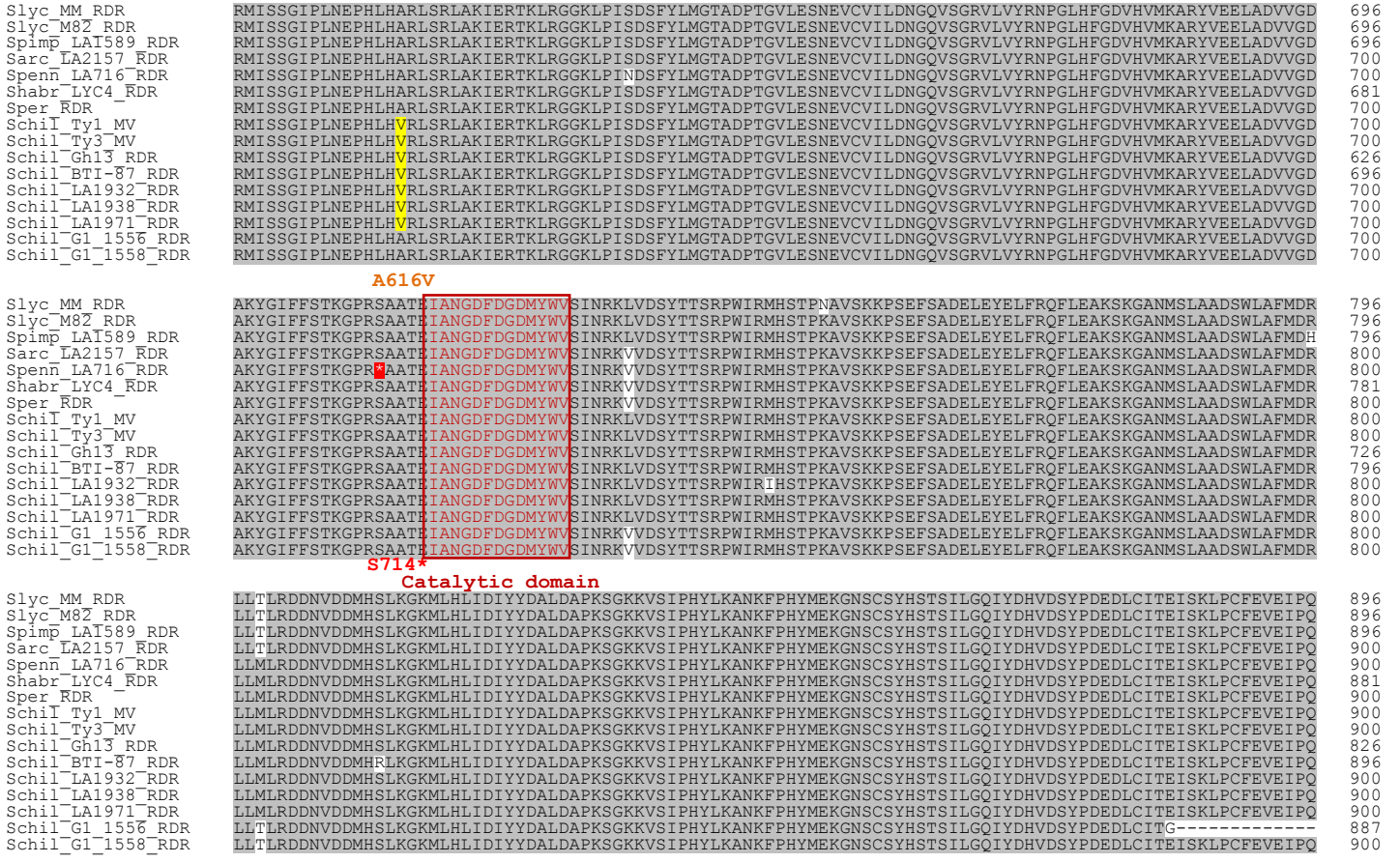

exon 18
exon 19

```
XM_010323869          ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGACGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2916
Slyc_MM_RDR           ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGACGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2916
Slyc_M82_RDR          ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGACGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2916
Spimp_LA1589_RDR      ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGACGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2916
Sarc_LA2157_RDR       ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGACGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Spenn_LA716_RDR       ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Shabr_LYC4_RDR        ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGTGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2871
Sper_RDR              ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Schil_LA1969_Ty1      ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAAGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Schil_LA2779_Ty3      ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAAGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Schil_Gh13_RDR        ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAAGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2707
Schil_BTI-87_RDR      ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2916
Schil_LA1932_RDR      ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Schil_LA1938_RDR      ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Schil_LA1971_RDR      ATATGGTGCTGTAGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTTGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928
Schil_G1_1556_RDR     ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  3100
Schil_G1_1558_RDR     ATATGGTGCTGTGGAGTTTGAACAAACAGTAAGAAAGACTGAAGACATTTTCGATGAGGCCCTTGCAATATATCATGTAACATATGATAATGCAAGGATC  2928

XM_010323869          ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3016
Slyc_MM_RDR           ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3016
Slyc_M82_RDR          ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGCTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3016
Spimp_LA1589_RDR      ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGCTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3016
Sarc_LA2157_RDR       ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Spenn_LA716_RDR       GCGTACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Shabr_LYC4_RDR        GCATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  2971
Sper_RDR              TCATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Schil_LA1969_Ty1      ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Schil_LA2779_Ty3      ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Schil_Gh13_RDR        ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  2807
Schil_BTI-87_RDR      ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCACAAGGAAAAAGACTTGCCCATTT  3016
Schil_LA1932_RDR      ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCACAAGGAAAAAGACTTGCCCATTT  3028
Schil_LA1938_RDR      ACATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Schil_LA1971_RDR      GCATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028
Schil_G1_1556_RDR     TCATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3200
Schil_G1_1558_RDR     TCATACAGCATAGAGAAATGTGGTTTTGCTTGGAAAGTAGCTGGTTCTGCGCTTTGCAGGATCCACGCCATGTATCGCAAGGAAAAAGACTTGCCCATTT  3028

XM_010323869          TGCCATCGGTTTTGCAGGAAATACTCTAG  3045
Slyc_MM_RDR           TGCCATCGGTTTTGCAGGAAATACTCTAG  3045
Slyc_M82_RDR          TGCCATCGGTTTTGCAGGAAATACTCTAG  3045
Spimp_LA1589_RDR      TGCCATCGGTTTTGCAGGAAATACTCTAG  3045
Sarc_LA2157_RDR       TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Spenn_LA716_RDR       TGCCATCGCTTTTGCAGGAAATACTCTAG  3057
Shabr_LYC4_RDR        TGCCATCGGTTTTGCAGGAAATACTCTAG  3000
Sper_RDR              TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Schil_LA1969_Ty1      TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Schil_LA2779_Ty3      TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Schil_Gh13_RDR        TACCATCGGTTTTGCAGGAAATACTCTAG  2836
Schil_BTI-87_RDR      TGCCATCGGTTTTGCAGGAAATACTCTAG  3045
Schil_LA1932_RDR      TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Schil_LA1938_RDR      TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Schil_LA1971_RDR      TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
Schil_G1_1556_RDR     TGCCATCGGTTTTGCAGGAAATACTCTAG  3229
Schil_G1_1558_RDR     TGCCATCGGTTTTGCAGGAAATACTCTAG  3057
```

Figure S5. Alignment of protein sequences of the *Ty-1/Ty-3 RDR*. Protein sequences are derived from cDNA sequences of accessions and derived lines as in Figure S4. The 5' indel and the catalytic domain are highlighted in red. A premature stop codon in *S. pennellii* LA716 is highlighted in red. Three *Ty-1/Ty-3* specific amino acids are highlighted in yellow; two *Ty-3* specific amino acids are highlighted in green.

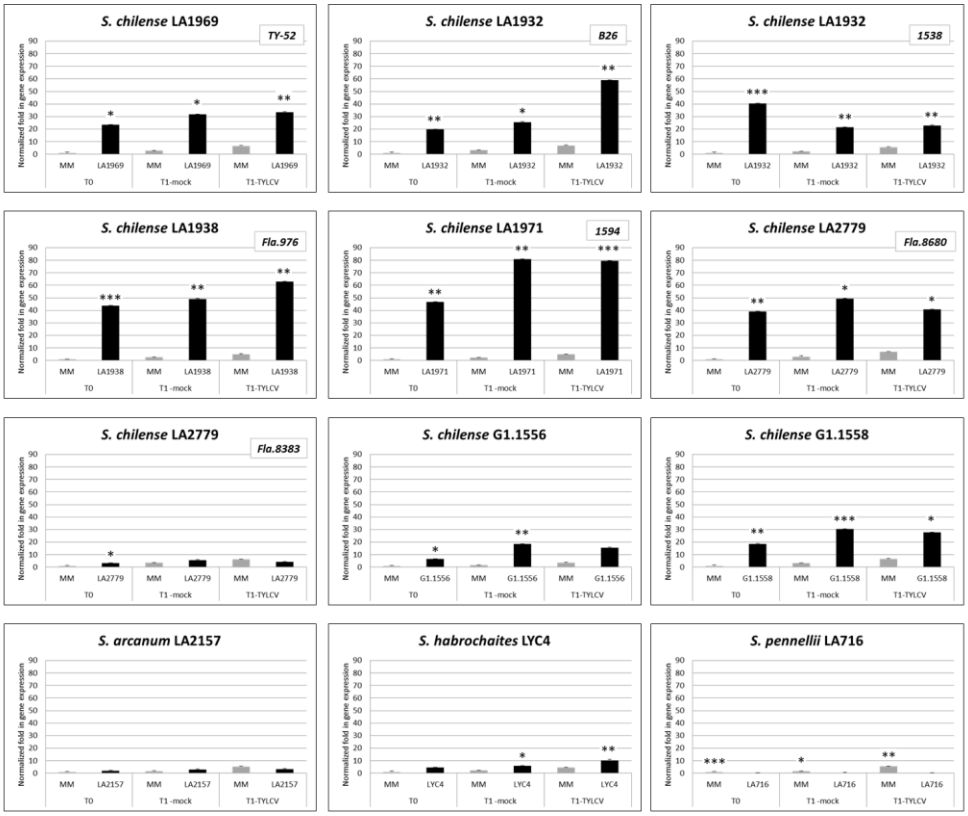



Slyc_MM_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCNDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQTDLDSVTIERIEDI 396
Slyc_M82_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCNDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQTDLDSVTIERIEDI 396
Spimp_LA1589_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCNDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQTDLDSVTIERIEDI 396
Sarc_LA2157_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCNDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQTDLDSVTIERIEDI 400
Spenn_LA716_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCNDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQTDLDSVTIERIEDI 400
Shabr_LYC4_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCNDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQADLDSVTIERIEDI 400
Sper_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQADLDSVTIERIEDI 400
Schil_Ty1_MV YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQVDLDSVTIERIEDI 400
Schil_Ty3_MV YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTTTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQVDLDSVTIERIEDI 400
Schil_Gh13_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTTTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQTDLDSVTIERIEDI 326
Schil_BTI-87_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQADLDSVTIERIEDI 396
Schil_LA1932_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQADLDSVTIERIEDI 400
Schil_LA1938_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQVDLDSVTIERIEDI 400
Schil_LA1971_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTVSQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQVDLDSVTIERIEDI 400
Schil_G1_1556_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQVDLDSVTIERIEDI 400
Schil_G1_1558_RDR YRFFVYKDDKERKKSPAMMKTKTASLKCYFVRFESIGTCDDGESYVFSTKTISQARCKFMHVHMVSNMAKYAARLSLILSKTIKLQVDLDSVTIERIEDI 400
K350I
Slyc_MM_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRVSVSRNRKPPLLMQCRLFFNGCAVKGTFLVNRK 496
Slyc_M82_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRVSVSRNRKPPLLMQCRLFFNGCAVKGTFLVNRK 496
Spimp_LA1589_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRVSVSRNREPPLLMQCRLFFNGCAVKGTFLVNRK 496
Sarc_LA2157_RDR LCRDENGCIIQDEDGEPRIHSDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRVSVSRNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Spenn_LA716_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSRNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Shabr_LYC4_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYEXXX----------EPPLLMQCRLFFNGCAVKGTFLVNRK 481
Sper_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Schil_Ty1_MV LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Schil_Ty3_MV LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Schil_Gh13_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 426
Schil_BTI-87_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRXSVSRNRKPPLLMQCRLFFKGCAVKGTFLLNRK 496
Schil_LA1932_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGKREPPLLMQCRLFFKGCAVKGTFLLNRK 500
Schil_LA1938_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Schil_LA1971_RDR LCQDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Schil_G1_1556_RDR LCRDENGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRVSVSRNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Schil_G1_1558_RDR LCRDKNGCIIQDEDGEPRIHTDGTGFISEDLAMHCPKDFSKAEYIKDENYENFVDIVDLDDVNVERRASVSGNREPPLLMQCRLFFNGCAVKGTFLVNRK 500
Slyc_MM_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 596
Slyc_M82_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 596
Spimp_LA1589_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 596
Sarc_LA2157_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETNQVYLRKRAALKVAINYREMDDECLTA 600
Spenn_LA716_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETKQVRLRKRAALKVAINYREMDDECLTA 600
Shabr_LYC4_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 581
Sper_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHREYFMELLGSALEETNQVYLRKRAALKVAINYREMDDECLTA 600
Schil_Ty1_MV IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 600
Schil_Ty3_MV IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 600
Schil_Gh13_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 526
Schil_BTI-87_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 596
Schil_LA1932_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 600
Schil_LA1938_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 600
Schil_LA1971_RDR IGSRKIHIRPSMVKVEIDPTISSIPTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFLELLGSALEETKQVYLRKRAALKVAINYREMDDECLTA 600
Schil_G1_1556_RDR IGSRKIHIRPSMVKVEIDPTISSISTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETNQVYLRKRAALKVAINYREMDDECLTA 600
Schil_G1_1558_RDR IGSRKIHIRPSMVKVEIDPTISSISTFDSLEIVAISHRPNKAYLSKNLISLLSYGGVHKEYFMELLGSALEETNQVYLRKRAALKVAINYREMDDECLTA 600
M563L


A616V
S714*
Catalytic domain

```
Slyc_MM_RDR          RCMTLWRGRYEEYKKDMTRAMNFDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM   996
Slyc_M82_RDR         RCMTLWRGRYEEYKKDMTRAMNFDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM   996
Spimp_LA1589_RDR     RCMTLWRGRYEEYKKDMTRAMNFDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM   996
Sarc_LA2157_RDR      RCMTLWRGRYEEYKKDMTRAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM  1000
Spenn_LA716_RDR      RCMTLWRGRYEEYKKDMTRAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARIAYSIEKCGFAWKVAGSALCRIHAM  1000
Shabr_LYC4_RDR       RCMTLWRGRYEEYKKDMTRAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDVALAIYHVTYDNARIAYSIEKCGFAWKVAGSALCRIHAM   981
Sper_RDR             RCMTLWRGRYEEYKKDMTRAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARISYSIEKCGFAWKVAGSALCRIHAM  1000
Schil_Ty1_MV         RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM  1000
Schil_Ty3_MV         RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM  1000
Schil_Gh13_RDR       RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM   926
Schil_BTI-87_RDR     RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM   996
Schil_LA1932_RDR     RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM  1000
Schil_LA1938_RDR     RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARITYSIEKCGFAWKVAGSALCRIHAM  1000
Schil_LA1971_RDR     RCMTLWRGRYEEYKKDMTQAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARIAYSIEKCGFAWKVAGSALCRIHAM  1000
Schil_G1_1556_RDR    ------------------------------------------------------------------------------------------------------   887
Schil_G1_1558_RDR    RCMTLWRGRYEEYKKDMTRAMNLDCELRITSCNEVIKKYKMLLYGAVEFEQTVRKTEDIFDEALAIYHVTYDNARISYSIEKCGFAWKVAGSALCRIHAM  1000
                                                                                                    R919Q

Slyc_MM_RDR          YRKEKDLPILPSVLQEIL   1014
Slyc_M82_RDR         YRKEKDLPILPSVLQEIL   1014
Spimp_LA1589_RDR     YRKEKDLPILPSVLQEIL   1014
Sarc_LA2157_RDR      YRKEKDLPILPSVLQEIL   1018
Spenn_LA716_RDR      YRKEKDLPILPSLLQEIL   1018
Shabr_LYC4_RDR       YRKEKDLPILPSVLQEIL    999
Sper_RDR             YRKEKDLPILPSVLQEIL   1018
Schil_Ty1_MV         YRKEKDLPILPSVLQEIL   1018
Schil_Ty3_MV         YRKEKDLPILPSVLQEIL   1018
Schil_Gh13_RDR       YRKEKDLPILPSVLQEIL    944
Schil_BTI-87_RDR     YHKEKDLPILPSVLQEIL   1014
Schil_LA1932_RDR     YHKEKDLPILPSVLQEIL   1018
Schil_LA1938_RDR     YRKEKDLPILPSVLQEIL   1018
Schil_LA1971_RDR     YRKEKDLPILPSVLQEIL   1018
Schil_G1_1556_RDR    ------------------    887
Schil_G1_1558_RDR    YRKEKDLPILPSVLQEIL   1018
```

Figure S6. Relative expression of the *Ty-1/Ty-3 RDR* in different accessions of *S. chilense*. Normalized fold gene expression of the target gene in derived introgression lines or *Solanum* accessions as determined by qRT-PCR; *S. chilense* LA1969, LA1932, LA1938, LA1971, LA2779, G1.1556, G1.1558 and related species *S. arcanum* LA2157, *S. habrochaites* LYC4 and *S. pennellii* LA716 are also included in the analysis. Time points T0 and T1 (0 and 19 days after TYLCV or *mock* inoculation respectively) and genotypes (Moneymaker (MM) vs. each *RDR* allele source) are shown on the *x*-axis. Values are normalized against the Moneymaker T0 sample; *bars* represent means and standard error of four biological replicas. *Asterisks* above the *bars* represent significant differences between genotypes per time point and *mock* or *TYLCV* treatment (*$P < 0.05$, **$P < 0.01$, *** $P < 0.001$).



Chapter 3

Fine Mapping of the Tomato Yellow Leaf Curl Virus Resistance Gene *Ty-2* on Chromosome 11 of Tomato

Myluska Caro[1,§], Xiaohui Yang[2,3,4,§], Samuel F. Hutton[4], John W. Scott[4], Yanmei Guo[2], Xiaoxuan Wang[2], Harunur Rashid[1], Dora Szinay[5], Hans de Jong[5], Richard G.F. Visser[1], Yuling Bai[1] and Yongchen Du[2]

[1]Wageningen UR Plant Breeding, Wageningen University & Research Center, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
[2]Institute of Vegetable and Flowers, the Chinese Academy of Agricultural Sciences, Beijing, China
[3]Institute of Vegetables and Flowers, Shandong Academy of Agricultural Sciences, Shandong Key Laboratory for Biology of Greenhouse Vegetables, Jinan, China
[4]Gulf Coast Research and Education Center, University of Florida, 14625 CR 672, Wimauma, FL 33598
[5]Laboratory of Genetics, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

§ Myluska Caro and Xiaohui Yang contributed equally to this work.

This chapter has been published in Molecular Breeding 2014, 34:749–760

Abstract

Resistances to begomoviruses, including bipartite tomato mottle virus (ToMoV) and monopartite tomato yellow leaf curl virus (TYLCV), have been introgressed to cultivated tomato (*Solanum lycopersicum*) from wild tomato accessions. A major gene, *Ty-2* from *S. habrochaites* f. *glabratum* accession 'B6013', that confers resistance to TYLCV, was previously mapped to a 19 cM region on the long arm of chromosome 11. In the present study, approximately 11,000 plants were screened and nearly 157 recombination events were identified between the flanking markers C2_At1g07960 (82.5 cM, physical distance 51.387 Mb) and T0302 (89 cM, 51.878 Mb). Molecular marker analysis of recombinants and TYLCV evaluation of progeny from these recombinants, localized *Ty-2* to an approximately 300,000 bp interval between markers UP8 (51.344 Mb) and M1 (51.645 Mb). No recombinants were identified between TG36 and C2_At3g52090, a region of at least 115 kb, indicating severe recombination suppression in this region. Due to the small interval, Fluorescence in situ hybridization (FISH) analysis failed to clarify whether recombination suppression is caused by chromosomal rearrangements. Candidate genes predicted based on tomato genome annotation were analyzed by RT-PCR and virus induced gene silencing. Results indicate that the NBS gene family present in the *Ty-2* region is likely not responsible for the *Ty-2*-conferred resistance and that two candidate genes might play a role in the *Ty-2*-conferred resistance. Several markers very tightly linked to the *Ty-2* locus are presented and useful for marker assisted selection in breeding programs to introgress *Ty-2* for begomovirus resistance.

Keywords: Breeding, Resistance, Tomato, Tomato yellow leaf curling virus (TYLCV), Virus-induced gene silencing

Introduction

Tomato-infecting begomoviruses, including the monopartite tomato yellow leaf curl virus (TYLCV) and numerous bipartite viruses, are transmitted by the adult sweet potato whitefly [*Bemisia tabaci* (Gennadius) biotype B], which is also known as the silverleaf whitefly (*B. argentifolii* Bellows & Perring). These viruses cause serious losses to tomato (*Solanum lycopersicum* L.) production in many tropical and subtropical regions in the world (Ji et al. 2007a; Cohen and Lapidot, 2007). Whitefly control measures such as the use of insecticides and/or fine-mesh screens or UV-absorbing plastic films/screens can limit disease damage, but epidemics can still occur. Also whitefly resistance to the used chemicals has been reported (Antignus et al. 2001; Horowitz et al. 2007). Thus, deployment of resistant cultivars offers an attractive method to control these diseases. Cultivated tomato is susceptible to TYLCV, so breeding efforts rely on the transfer of resistance genes from wild tomato relatives. Species that have demonstrated resistance include *S. pimpinellifolium, S. peruvianum, S. chilense, S. habrochaites*, and *S. cheesmaniae* (Ji et al. 2007b; Pico et al. 1996; Scott, 2007; Vidavski, 2007). So far, as many as five resistance loci have been mapped, i.e. the dominant genes including *Ty-1*, *Ty-2*, *Ty-3*, *Ty-4*, and recessive gene *ty-5* (Zamir et al. 1994; Hanson et al. 2000; Ji et al. 2007a; Ji et al. 2009b; Anbinder et al. 2009). *Ty-1* and *Ty-3* were each derived from *S. chilense* and mapped to nearby positions on chromosome 6 (Ji et al. 2007c); however, Verlaan et al. (2013) demonstrated that *Ty-1* and *Ty-3* are alleles of the same gene. *Ty-4*, also derived from *S. chilense*, was mapped to chromosome 3 (Ji et al. 2008). The recessively inherited *ty-5* gene, first identified in the breeding line TY172 and later found in material derived from 'Tyking,' was mapped to chromosome 4 (Anbinder et al. 2009; Hutton et al. 2012). The *ty-5* gene is likely derived from a complex of *S. peruvianum* accessions (Anbinder et al., 2009). However, there is also evidence showing that *ty-5* is a loss-of-function mutation that likely occurred in cultivated tomato (Levin et al. 2013). *Ty-2* was derived from *S. habrochaites* f. *glabratum* accession 'B6013' (Kalloo and Banerjee, 1990; Ji et al. 2009a) and was previously mapped to the long arm of chromosome 11 near markers TG36 (84 cM) and TG393 (103 cM) (Hanson et al., 2000). Further research indicated that *Ty-2* was localized to an introgression spanning markers TG36 (84 cM) to TG26 (92 cM) (Hanson et al. 2006). Later, *Ty-2* was delimited to a shorter introgression spanned by markers C2_At1g07960 (82.5 cM) and T0302 (89 cM) (Ji et al. 2009a), a distance of at least 500,000 bp on the tomato genome assembly. The fusarium wilt race 2 resistance gene (*I-2*) is close to the *Ty-2* region (Simons et al. 1998) and there may be difficulty in combining these important resistances in *cis*.

Reducing the *Ty-2* introgression would be helpful in combining these two important disease resistances in a single line (Ji et al. 2009a).

Because *Ty-1* and *Ty-2* are both dominant and provide high levels of resistance to many strains of TYLCV, they are widely utilized by breeders. Yet neither gene is effective against bipartite begomoviruses, and the resistance of both has been overcome by some strains of TYLCV (Ji et al. 2007a). There is evidence, however, that *Ty-2* can provide an enhanced level of resistance to bipartite begomoviruses when pyramided with *Ty-3* (Mejía et al. 2005), potentially making it a more attractive tool to breeders. Very recently, the cloning of *Ty-1* and *Ty-3* showed that they code for a DFDGD-class RNA-dependent RNA polymerase (RDR) for which no clear function has yet been described. Also, in the same study it was shown that *Ty-2* does not encode for a RDR (Verlaan et al. 2013). Thus, cloning additional genes for TYLCV resistance offers a unique opportunity to advance the insight into novel types of resistance genes. The objective of this research is to fine map *Ty-2* towards the cloning of the gene.

Results

Ty-2 maps between UP8 and M1, a region of about 300kb

Phase I screening of approximately 4,000 seedlings in Fall 2009 resulted in the identification of 30 plants having a recombination event between C2_At1g07960 and T0302 (Table 1). Progeny lines of Group A recombinants were phenotypically evaluated. Eight recombinants (A1 and A2) that were segregating for the upper portion (C2_At1g07960 to M1) of the *Ty-2* introgression also segregated for resistance, while the five recombinants (A3 and A4) that were lacking this upper portion of the introgression were all susceptible. Likewise, evaluation of Group B RILs in Fall 2010 confirmed that those containing the upper portion of the introgression were resistant (B1 to B3, Table 1), while those lacking this region were susceptible (B4 to B7, Table 1). Thus, the genotype and phenotype results of important recombinants clearly delimited *Ty-2* to the region below UP8 (data of A1 to A2, and B1 to B3) and above M1 (data of B3 and B4).

Table 1. Genotype for the UP8 to T0302 marker interval of *Ty-2* recombinants identified from Phase I screening and their phenotype as determined by testing their progenies.

Group	No. of recom_ binants	Genotype[a]							Spring 2010 Progeny phenotype		Fall 2010 RIL Phenotype	
		UP8	C2_At1 g07960	C2_At3 g52090	M1	M2	M3	T0302	Total plants	Resist. plants	Tested plants	Phenotype[b]
A-1	5	ll	hl	hl	hl	ll	ll	ll	75	55		
A-2	3	ll	hl	hl	hl	hl	hl	ll	94	71		
A-3	2	ll	ll	ll	ll	hl	hl	hl	66	0		
A-4	3	ll	ll	ll	ll	ll	ll	hl	92	0		
B-1	1	ll	hh	hh	hh	hh	hh	ll			7	R
B-2	4	ll	hh	hh	hh	ll	ll	ll			41	R
B-3	1	ll	hh	hh	ll	ll	ll	ll			23	R
B-4	2	ll	ll	ll	hh	hh	hh	hh			28	S
B-5	5	ll	ll	ll	ll	hh	hh	hh			31	S
B-6	2	ll	ll	ll	ll	ll	hh	hh			27	S
B-7	2	ll	ll	ll	ll	ll	ll	hh			30	S

[a]hh= homozygous for the *S. habrochaites* allele; ll= homozygous for *S. lycopersicum* allele; hl= heterozygous
[b] R= resistant; S= susceptible

In Spring 2011, Phase II screening of approximately 7,000 additional plants identified 127 recombinants for the C2_At1g07960 and T0302 marker interval, but only 26 of these had recombinations above marker M1 (Table 2). Of those 26 recombinations, occurred between the markers C2_At3g52090 and M1, and the other two cross-overs occurred between P1-16 and TG36. Selection for homozygosity of the recombined introgression of these 26 individuals resulted in three genotypic categories of RILs that were evaluated in Fall 2011. All Category I and II RILs containing the upper portion of the introgression were resistant, while all Category III RILs lacking this region were susceptible (Table 2), confirming the location of *Ty-2* above M1. Subsequent development and testing of additional markers between C2_At2g28250 and C2_At1g07960 determined that the upper end of the *S. habrochaites* introgression is likely between UP8 and C2_At1g07960 (Figure 1). Thus, our results map *Ty-2* to the approximately 300,000 bp region between UP8 and M1.

Table 2. Genotype and phenotype of 26 recombinant inbred lines (RILs) identified from Phase II recombinant screening.

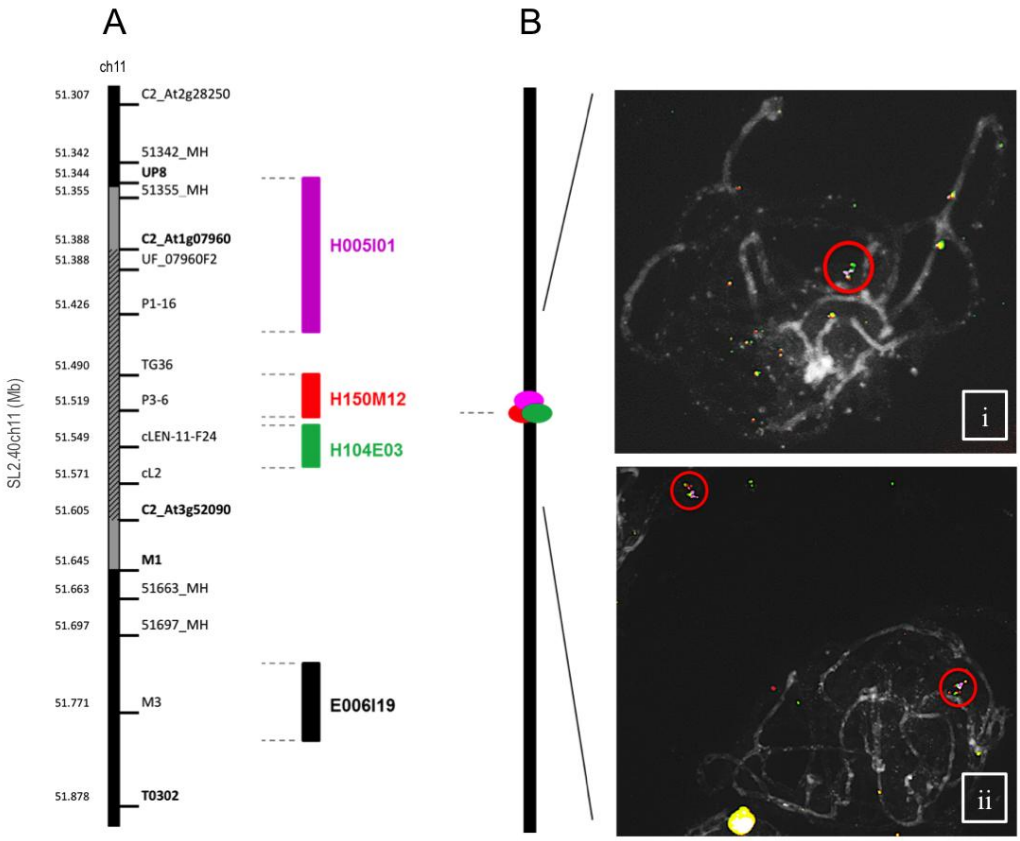
Category	No. RILs	Molecular marker[a]										
		UP8	C2_At1 g07 960	P1-16	TG36	cLEN-11-F24	cL1	cL2	C2_At3 g52090	M1	T0302	Pheno type
I	13	ll	hh	hh	hh	hh	hh	hh	hh	ll	ll	R
II	2	ll	hh	hh	hl	hl	hl	hl	hl	ll	ll	R
III	11	ll	ll	ll	ll	ll	ll	ll	ll	hh	hh	S

hh=homozygous for *S. habrochaites* alleles; ll= homozygous for *S. lycopersicum* alleles; hl= heterozygous alleles
R, resistant; S, susceptible

Skewed allele frequency in the *Ty-2* introgression

During the Phase II development of RILs, an interesting segregation pattern was obtained from progeny of the two Category II recombinants in the marker interval between P1-16 to TG36 (Table 2). Each of these recombinants was homozygous for *S. habrochaites* alleles (*hh*) between markers C2_At1g07960 and P1-16 and heterozyzgous (*hl*) at markers TG36 to M1. To generate RILs, progeny of these two individuals were screened with marker M1 to select plants homozygous for the *S. lycopersicum* allele (*ll*). For each recombinant, 48 seedlings were screened, and several individuals were selected. Segregation at M1 had an acceptable fit to a 1:2:1 ratio (1*hh*:2*hl*:1*ll*). However, subsequent screening of these individuals with additional markers between P1-16 and M1 showed that all selected plants, although homozygous for the *S. lycopersicum* allele at M1, remained heterozygous for all markers tested in the TG36 to C2_At3g52090 interval (Table 2). Further screening of nearly 100 progeny from each of the two Category II recombinants confirmed this result, and no progeny were identified that were homozygous for the *S. lycopersicum* alleles in the TG36 to C2_At3g52090 interval. Within this interval, the allele frequency of *hh*:*hl* segregated in a 1:3 ratio. Although all progeny of these two recombinants showed clear TYLCV resistance, the failure to recover homozygous *S. lycopersicum* alleles between the TG36 to C2_At3g52090 interval from the genotyping of 200 plants prevented the further narrowing of the *Ty-2* region.

Figure 1. Genetic map of chromosome 11 (part). **A** Map position of the *Ty-2* gene is shown (grey box between markers UP8 and M1) and the region where suppression of recombination was identified (shaded region between markers C2_At1g07960 and C2_At3g52090). Bacterial artificial chromosome (BAC)s selected for fluorescence in situ hybridization (FISH) are shown in color blocks. **B** Schematic drawing of arrangements of BACs observed in FISH experiments. FISH images showing BAC signals in F_2 plants i) homozygous for the susceptible *S. lycopersicum* 'Moneymaker' alleles and ii) for the *S. habrochaites* alleles of loci in the *Ty-2* introgression. Overlapping BACs are observed for both genotypes.



Suppression of recombination in the *Ty-2* introgression

In summary, Phase I and II screening of approximately 11,000 progeny identified 157 recombinants for the approximately 500,000 bp region between C2_At1g07960 and T0302. Only 29 of these cross-overs occurred above marker M1; of these, 27

occurred in the approximately 35,000 bp region between C2_At3g52090 and M1, and only two were in the approximately 60,000 bp region between P1-16 and TG36. No recombinants were found between TG36 and C2_At3g52090, indicating suppression of recombination.

To clarify whether the suppression was population dependent, a recombinant screening was carried out in Wageningen (the Netherlands) in another F_2 population derived from a round tomato F_1 hybrid. In this F_2 population, markers from 51355_MH through T0302 segregated (Figure 1), showing that a large introgression of *S. habrochaites* is present in the commercial hybrid carrying the *Ty-2* gene. The presence of the *Ty-2*-conferred resistance was confirmed by challenging 110 F_2 plants with TYLCV and genotyping them with markers between 51355_MH and T0302. Among the 110 plants tested, 25 showed TYLCV symptoms similar to the susceptible control, MM and were homozygous for the susceptible allele at all tested markers. The other 85 plants showing slight or no symptoms were scored as resistant. Resistant plants were either homozygous or heterozygous for *S. habrochaites* alleles at all tested markers. Thus, there was no skewing of allele frequency in the region between 51355_MH and T0302 in this F_2 population. By screening an additional 1900 plants of this F_2 population with markers UF_07960F2 and T0302, 18 recombinants were identified (data not shown) and all recombination events occurred downstream of the marker C2_At3g52090 (Figure 1), confirming a severe suppression of recombination in the region between markers C2_At1g07960 and C2_At3g52090.

As with the *Ty-1* introgression (Verlaan et al. 2011), we hypothesized that differences in chromosome structure between the two parental lines might be the cause for the suppression of recombination. We previously showed that Fluorescent *in situ* Hybridization (FISH) can be used as a molecular tool to reveal inversions or chromosomal rearrangements among several *Solanum* species (Szinay *et. al.* 2012). Therefore, we applied FISH analysis in order to visualize the chromosome structure of the *Ty-2* introgression. Four BACs located within the 300 kb *S. habrochaites* introgression were selected (Figure 1) and labeled for FISH as described in Verlaan et al. (2011). Unfortunately, FISH images showed overlapping

fluorescing foci from their corresponding BACs, indicating that this 300 kb region is too small for resolution using the FISH technique.

Differential expression of the candidate genes

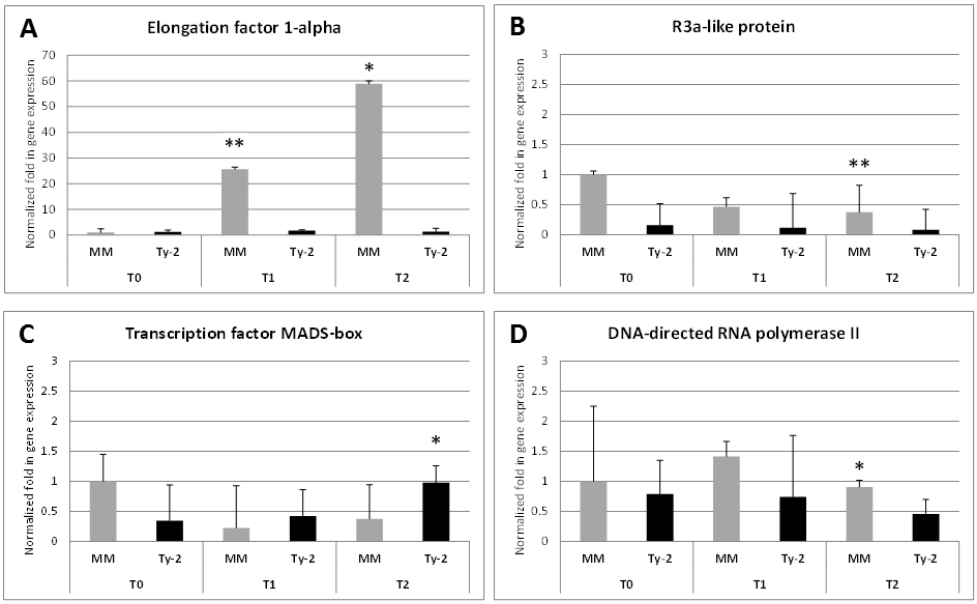
Within the 300-kb *Ty-2* region, 35 genes were annotated in the tomato sequence version SL2.40, Sol Genomics Network (Table S2). In order to examine the effects of TYLCV infection on the expression of the predicted genes in the *Ty-2* region, relative expression levels were quantified at 0, 9 and 20 days after TYLCV inoculation. Resistant plants carrying the *Ty-2* introgression and susceptible MM plants were sampled. Priority was given to genes expressed according to the RNAseq coverage information (Sol Genomics Network) and transcript levels of 25 out of the 35 predicted genes were quantified by RT-PCR (Table S3). Using ubiquitin as the housekeeping gene, three predicted genes were shown to be differentially expressed in the *Ty-2* plants compared to the susceptible genotype and upon TYLCV inoculation (Figure 2). These are Solyc11g069700.1, encoding an elongation factor 1-alpha; Solyc11g069770.1, encoding a transcription factor of MADS-box family; and Solyc11g069930.1, encoding an R3a-like resistance protein. Among these genes, the elongation factor 1-alpha and the disease resistance protein R3a-like showed lower expression in the resistant genotype. Twenty days after virus infection, relative amount of transcripts of the elongation factor in the susceptible genotype was over 40 times higher than in the resistant plants (Figure 2A), and across all time points, the expression level of this gene in the resistant plants remained very low. Expression of the predicted R3a-like homolog in the resistant plants remained almost 4 times lower than in the susceptible genotype (Figure 2B) across all time points. In contrast, expression of the transcription factor MADS-box in the resistant plants was 2.5 times higher than in the susceptible phenotype 20 days after TYLCV infection (Figure 2C). In addition to these three genes, Solyc11g069910.1, the gene encoding a DNA-directed RNA polymerase II showed a down-regulation in the *Ty-2* line upon TYLCV inoculation, almost 2 times lower than in the susceptible MM plants (Figure 2D).

Silencing of the differentially expressed candidate genes

To determine the implication of the candidate genes on TYLCV resistance, specific VIGS constructs (Table S4) were designed to silence these four candidate genes in MM and the line carrying the *Ty-2* gene. Two weeks after infiltration with TRV vector for gene silencing, plants were challenged with TYLCV. Plants infiltrated with TRV vectors but non-infected with TYLCV and plants infiltrated with an empty (EV) TRV vector were used as controls. Except for *R3a*-like genes, plants infiltrated with TRV vectors targeting these genes all showed an abnormal

phenotype when compared to the control plants (Figure 3). Silencing the elongation factor 1-alpha had a lethal effect (Figure 3A); silencing the transcription factor MADS-box resulted in plants with yellowish leaves (Figure 3B) and silencing the DNA-directed RNA polymerase II protein led to stunted plants with smaller and curled leaves (Figure 3C). These phenotypes were observed in plants of the *Ty-2* line and MM before TYLCV inoculation, thus determined by the silencing of the target gene itself rather than induced by TYLCV infection.

Figure 2. Relative expression of candidate genes. Normalized fold in gene expression of differentially expressed candidate genes as determined by RT-PCR; **A** Elongation factor 1-alpha, **B** R3a-like protein, **C** Transcription factor MADS-box, **D** DNA-directed RNA polymerase. Time points are shown in the x-axis; T0, T1, and T2 (0, 9, and 20 days after TYLCV inoculation). Values are normalized against the Moneymaker T0 sample; bars represent means and standard deviation of three biological replicas. Asterisks above the bars represent significant differences between genotypes per time point (*P<0.05, **P<0.01).



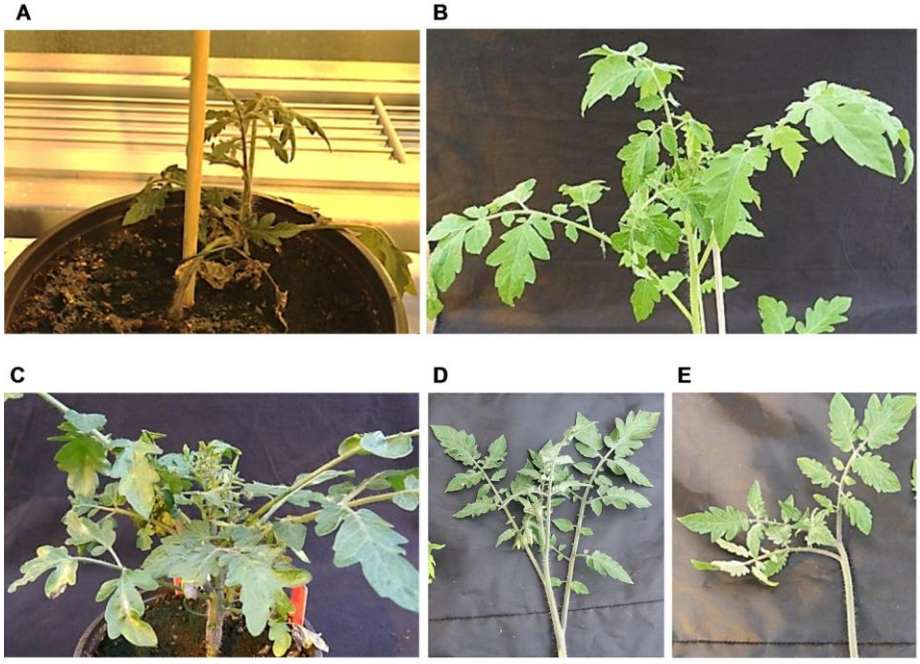
Silencing of the differentially expressed candidate genes

To determine the implication of the candidate genes on TYLCV resistance, specific VIGS constructs (Table S4) were designed to silence these four candidate genes in

MM and the line carrying the *Ty-2* gene. Two weeks after infiltration with TRV vector for gene silencing, plants were challenged with TYLCV. Plants infiltrated

with TRV vectors but non-infected with TYLCV and plants infiltrated with an empty (EV) TRV vector were used as controls. Except for *R3a*-like genes, plants infiltrated with TRV vectors targeting these genes all showed an abnormal phenotype when compared to the control plants (Figure 3). Silencing the elongation factor 1-alpha had a lethal effect (Figure 3A); silencing the transcription factor MADS-box resulted in plants with yellowish leaves (Figure 3B) and silencing the DNA-directed RNA polymerase II protein led to stunted plants with smaller and curled leaves (Figure 3C). These phenotypes were observed in plants of the *Ty-2* line and MM before TYLCV inoculation, thus determined by the silencing of the target gene itself rather than induced by TYLCV infection.

Figure 3. Candidate genes silencing effects. Pictures were taken from resistant plants carrying *Ty-2*. Targeting three candidate genes showed abnormal phenotypes: **A** Elongation factor 1-alpha. Targeting this gene had a lethal effect. **B** Transcription factor MADS-box. Yellowish leaves and smaller and weaker plants were observed upon TYLCV infection. **C** DNA-directed RNA polymerase II. A stunted plant with shorter internodes and curled small leaves was observed. **D** R3a homologs. Resistance was not compromised; no phenotype was observed. **E** TRV-empty vector control plant.



Candidate genes in the *Ty-2* region mapped in this study include 5 genes encoding CC-NBS-LRR proteins, of which two are *R3a*-like. *R3a* is a member of the *R3* complex locus on chromosome 11 of potato, which confers race-specific resistance to the oomycete *Phytophtora infestans* (Huang et al. 2005). To determine whether these *R3a* homologues are required for *Ty-2* mediated resistance to TYLCV, a VIGS construct (VG930) targeting both Solyc11g069670.1 and Solyc11g069930.1 was generated (Figure S1). After silencing, plants were TYLCV inoculated and subsequently monitored for development of symptoms. Until 45 days after virus infection, no viral symptoms were recorded (Figure 3D). The phenotype observed was similar to that displayed by the control plants (inoculated with an empty TRV vector and infiltrated only with TRV silencing vector). Our data indicate that silencing of the *R3a* homologues does not compromise the resistance conferred by *Ty-2*.

Discussion

Recombination suppression is a common phenomenon in genomic regions introgressed from wild tomato species (Ji et al. 2007b and 2009a). therefore, it was not surprising to observe this in the introgressed segment containing the *Ty-2* gene. In a previous study, the *Ty-2* gene was delimited to a region between the markers C2_At2g28250 and T0302, a distance of at least 500 kb (Ji et al. 2009a). Although approximately 11,000 plants were genotyped in the present study, only 157 recombinants within this region were obtained. These recombinants did allow the further delimiting of *Ty-2* to a shorter region spanned by markers UP8 and M1 of ≈300,000 bp. The reason for this suppression is unknown, but perhaps there is an inversion as there was on chromosome 6 in a region where *Ty-1* and *Ty-3* have been mapped (Verlaan et al. 2011). Unfortunately, FISH was not powerful in this case to visualize any potential chromosomal rearrangements. Alternatively, a region of duplication or a cold spot for recombination could also explain the suppression of recombination. Previous studies have shown that recombination frequency is positively related to the length of alien segments and that, in some cases, cross direction also has significant impact on the frequency of recombination (Canady et al., 2006; Li et al. 2010). In order to increase the frequency of recombination in the *Ty-2* region, it would be helpful to use populations derived from lines with a larger *Ty-2* introgressed segment. In case that chromosomal rearrangement is present in *S. habrochaites*, the best option would be to use an intraspecific crosses with susceptible *S. habrochaites* accession to facilitate the cloning of *Ty-2*.

Whatever the reason for the suppression of recombination, the inability to further reduce the size of the introgressed chromosome segment has an important impact on practical breeding for two reasons. First, there is the possibility of linkage drag. No reports of linkage drag associated with *Ty-2* have been published to date, but in large fruited tomato germplasm, a rough blossom scar where teratomas emerge has been associated with resistance from *Ty-2* (Ryohei Arimoto, personal communication). Our lines with the shortest introgressions need to be tested to determine if this problem has been eliminated. Secondly, the large chromosome segment introgressed from wild species can hamper combining important genes in *cis*. The fusarium wilt race 2 resistance gene, *I-2*, has been cloned (Simmons et al. 1998) and is located on chromosome 11 at approximately 52 Mb according to version SL 2.40 of the tomato genome assembly. Considering that this locus is more than 400,000 bp below *Ty-2*, it should not be a significant problem to combine the two genes in *cis*, although a directed effort will be needed.

Due to the suppression of recombination and skewing of allele frequencies, it is difficult to further delimit *Ty-2* into a smaller region in order to pinpoint the candidate. Therefore, we have performed gene expression and VIGS experiments in order to predict potential candidates for *Ty-2*. There are 35 genes predicted in the target region; among these are genes involved in plant-defence mechanisms or signaling pathways against viruses or other pathogens, such as ABC transporters, kinases, receptor-like proteins or cytochrome P450 (Krattinger et. al., 2009; Tena et al. 2011; Larkan et al. 2013; Howe et al. 2000). In order to more accurately determine potential candidates for *Ty-2,* we have performed gene expression and VIGS experiments. Of the 35 genes predicted in the target region, 25 were checked for expression and 4 out of these showed to be differentially expressed in the *Ty-2* line upon TYLCV infection. These genes encode for an elongation factor 1-alfa, a R3a-like protein, a DNA-directed RNA polymerase II and a transcription factor of the MADS-box family.

Host translation elongation factors are involved in the multiplication of viruses in multiple organisms (Lai, 1998). Elongation factor 1-alpha has been found to interact with several viral proteins (Buck, 1999; Thivierge et al. 2008) and recently recorded in a metabolite profile of a TYLCV resistant line upon TYLCV infection (Moshe et al. 2012). Matsuda and Dreher (2004) suggested EF1-alpha to enhance the translation of *Turnip yellow mosaic virus* RNA, therefore decreased amounts of gene products might prevent or interfere with viral replication, thus leading to resistance. We observed a reduced expression of the EF-1 alpha on *Ty-2* resistant plants, before and after TYLCV infection. However, silencing this gene led to the

collapse of the plants, preventing us to elucidate its implication on the *Ty-2* mediated resistance.

The most interesting altered phenotype was shown by silencing Solyc11g069910.1, the gene encoding a DNA-directed RNA polymerase II (Pol II). DNA-dependent RNA polymerases mediate epigenetic silencing as a resistance mechanism against geminiviruses. DNA-dependent RNA polymerases IV and V (and indirectly Pol II) are involved in the RNA-directed DNA methylation (RdDM) process, which can lead to transcriptional silencing, not only of viral invading DNA but also of host nuclear genes, transposons and repetitive elements (Carr et al. 2010; Haag and Pikaard, 2011). It might be possible that this gene is targeted by the virus, interfering with the RdDM process and causing epigenetic changes in the host and/or viral DNA, consequently producing TYLCV-like symptoms, e.g small and curling leaves of stunting plants.

The MADS-box family is described to mainly play fundamental roles in plant development (Kaufmann et al. 2009), but it is also involved in various stress-related processes (Lee et al. 2008). Silencing Solyc11g069770.1, a transcription factor MADS-box, led to yellowish leaves. Although it is speculative, our results may suggest that TYLCV suppresses the expression of the transcription factor MADS-box leading to yellowish leaves.

In the *Ty-2* region, three genes are predicted to encode CC-NBS-LRR, NBS-LRR and NBS resistance proteins. Additionally, two genes coding for a disease resistance R3a-like protein (fragment) and disease resistance R3a-like protein are predicted and each contains the NB-ARC dominion. *Ty-2* is a dominant resistance gene, and to date, most of the cloned dominant resistance genes encode proteins containing the conserved NB-ARC domain, making these genes likely candidates. However, silencing both R3a-like homologues did not compromise the resistance conferred by *Ty-2*, suggesting that this gene may not belong to a NBS gene family.

Ty-2 has shown complete dominance for TYLCV resistance (Ji et al. 2009a), but has been ineffective against some TYLCV strains and against bipartite begomoviruses (Mejía et al. 2005). The *Ty-3* locus has generally shown less dominance, but a wider range of resistance against TYLCV strains and bipartite begomoviruses (Ji et al. 2007a). Hybrids with the heterozygous combination of both genes may prove to be effective and durable against a wide array of begomoviruses. Although *Ty-2* alone provided no resistance to bipartite begomoviruses in Guatemala, pyramiding *Ty-2*

and *Ty-3* together provided a higher level of resistance than *Ty-3* alone (Mejia et al. 2010). Vidavski (2007 and 2008) also showed that combining different begomovirus resistance genes can have unanticipated synergistic effects, and the combination of *Ty-2* with other genes should be tested further in this regard. Tightly-linked PCR markers can be used to effectively tag these TYLCV resistance genes, and expedite the process of pyramiding these resistance genes of various origins into a single elite genotype, thus improving the resistance to TYLCV as well as broadening the resistance against a wider range of begomoviruses.

Materials and Methods

Plant Materials used in Florida, US

H9205 is an H.J. Heinz Inc. processing tomato hybrid with TYLCV resistance conferred by *Ty-2* in its heterozygous status. In a previous study by Ji et al. (2009a), F_2 progeny from H9205 was screened for recombination, and three F_2 recombinants (i.e. 82, 108, 134) were identified. One of these plants (no. 134) was heterozygous for *Ty-2* contained within an introgression from C2_At1g07960 (82.5 cM) to T0302 (89 cM). This plant was advanced to the F_3 generation, and progeny heterozygous for the same region were self-pollinated to produce an F_4 population used in the present study. In total, 11,000 individual F_4 plants were screened in two phases for recombination within the *Ty-2* region.

For Phase I of the recombinant screening, 4,000 F_4 progeny were screened in Fall 2009, and 30 plants were identified that contained cross-over events between the markers C2_At1g07960 and T0302. Recombinants were categorized into two groups; Group A was composed of individuals carrying one chromosome with a recombined introgression and one chromosome with no introgression; Group B was composed of individuals carrying one recombined and one non-recombined introgression. These plants were transplanted to the field in mid–October and allowed to self-pollinate, and seeds were harvested. In Spring 2010, progeny lines of Group A recombinants, along with resistant and susceptible controls, were inoculated and transplanted to small pots in the greenhouse. Plants were evaluated for disease severity, and from each resistant line, F_5 seed was harvested from one or two plants that were homozygous for the recombined introgression. In Fall 2010, 24 seedlings from each of the Group B recombinants were grown in a greenhouse. Recombinant inbred lines (RILs) were developed by genotyping each plant and selecting individuals which were homozygous for the recombined introgression. Plants of each RIL, along with controls, were inoculated and transplanted to 3.8-L pots in the greenhouse for evaluation of disease severity.

For Phase II testing, approximately 7,000 additional F_4 plants were screened in Spring 2011 for recombination between markers C2_At1g07960 and T0302. Plants with recombination events between C2_At1g07960 and M1 were selected and transferred to the field in April, 2011 and selfed seed was harvested from each plant. In Fall 2011, 48 seedlings from each recombinant were genotyped, and individuals homozygous for the recombined introgression were selected representing RILs for TYLCV inoculation and field evaluation of disease severity. RILs were evaluated in a randomized complete block design with two blocks and 4- to 5-plant plots.

In all inoculated experiments, 'Horizon' was used as the susceptible control, and an F_5 breeding line homozygous for *Ty-2* was used as the resistant control. The *Ty-2* breeding line was developed from H9205 by self-pollinating to the F_5 generation while selecting for homozygosity of the entire introgression originally present in the hybrid.

Plant materials used in Wageningen, The Netherlands

One advanced breeding line and one F_2 population, both derived from commercial hybrids harboring the *Ty-2* gene in the genetic background of *S. lycopersicum*, were provided by breeding companies within the cooperative framework of the Centre for BioSystems Genomics (CBSG). The F_2 population was used for recombinant screening. F_2 plants selected from recombinant screenings were selfed and their F_3 progenies were used for further testing with TYLCV. The advanced breeding line was used for gene expression and virus-induced gene silencing experiments. For all the experiments, plants were grown under greenhouse conditions (23 °C, 60% humidity and 16-h/8-h day/night cycle).

Inoculation and disease evaluation

In Florida, US, all plants tested were inoculated with whiteflies viruliferous for the Israeli strain of TYLCV and subsequently assessed for disease severity according to the method described by Griffiths and Scott (2001) with some modifications. Briefly, four- to six-week old seedlings were exposed to viruliferous whiteflies for one to two weeks in a growth chamber. Following inoculation, the whiteflies were killed and the plants were transplanted to 3.8-L pots in the greenhouse or to the field. Plants were rated for TYLCV disease severity approximately 40 days after exposure to whiteflies. Plants without symptoms similar to the resistant control were rated R and plants with severe symptoms similar to the susceptible control were rated S. There were no intermediate reactions.

In Wageningen, The Netherlands, TYLCV infection was done using *Agrobacterium*-mediated inoculation using the infectious TYLCV-IL clone as previously described by Verlaan et al. (2011). In all disease tests, *S. lycopersicum* cv. Moneymaker (MM) was used as the susceptible control.

Molecular markers

All markers used in this study were PCR-based, including sequence-characterized amplified region (SCAR) markers and cleaved amplified polymorphic sequence (CAPS) markers (Table S1). These were either publicly available or were designed from version SL2.40 of the tomato genome assembly by BatchPrimer3 online (http://probes.pw.usda.gov/batchprimer3/index.html).

Quantitative RT-PCR

For gene expression analysis, leaf samples from the top part of 3 plants per genotype were taken at 0, 9 and 20 days after TYLCV inoculation at Wageningen (The Netherlands). Two genotypes were used, the Wageningen *Ty-2* line (see description above) and tomato cultivar *S. lycopersicum* cv. Moneymaker (MM). Total RNA was isolated from leaf tissue using Qiagen RNA easy Plant Mini Kit, according to manufacturer's protocol. cDNA was synthesized using the iScript cDNA Synthesis kit (Bio-Rad). Quantitative real-time RT-PCR was performed using a Bio-Rad iCycler iQ5, in a 10µl reaction (employing SYBR Green Supermix) and according to the Bio-Rad protocol. Primers were designed to amplify a 100-200 bp region of each candidate gene from tomato *Ty-2* cDNA. Primer3 online software was used for primer selection and conditions were settled following recommendations of Thornton & Basu (2011). As a reference the ubiquitin gene was used with primers UBI-F (5'-GGACGGACGTACTCTAGCTGAT-3') and UBI-R (5'-AGCTTTCGACCTCAAGGGTA-3').

Virus Induced Gene Silencing (VIGS)

cDNA sequences of candidate genes predicted in the *Ty-2* region were obtained from the Sol Genomics Network database. Primers were designed to amplify a 150-450 bp region from cDNA of the Wageningen *Ty-2* line using Phusion DNA Polymerase. Fragments targeting the candidate genes for silencing were amplified and cloned into pENTR-TOPO (Invitrogen), sequenced for confirmation and subsequently cloned into TRV2 vector (Liu et al. 2002) using the Gateway system. Plasmids were transformed into *Agrobacterium tumefasciens* strain GV3101. For sequence alignments, MEGA version 5 software was used.

VIGS experiments were performed as described in Verlaan et al. (2013). Briefly, TRV infection was done through *Agrobacterium*-mediated infiltration on cotyledons

of 10-day-old seedlings using syringes without needle. Two weeks after TRV inoculation, agro infiltration with TYLCV was performed.

Fluorescent *in situ* Hybridization (FISH) analysis

Slide preparations, BAC isolation and FISH were carried out as described in Verlaan et al. (2011).

Acknowledgements

We thank Dr. Yuanfu Ji for his help in developing the H9205-derived populations and Prof. Eduardo Rodrígue Bejarano (Universidad de Málaga, Spain) for providing the infectious TYLCV clone used at Wageningen in this study. This work was supported by the National Natural Science Foundation of China (31272160), the National Basic Research Program of China (2012CB113906), and the National High Technology Research and Development Program of China (863 Program, No. 2012AA100101). The works done at Wageningen in the Netherlands was supported by an internal grant from Wageningen UR Plant Breeding.

Supplementary files

Table S1. Molecular markers used for mapping *Ty*-2 on tomato chromosome 11

Marker	Physical position (Mb)[a]	Primer[b]	Restriction enzyme	Source of primers[c]
C2_At2g28250	SL2.40ch11: 51.307	F-AGACTTCATCATCGTCATGTGGTTCCG R-TTTGGAGGTGCTTTGCCATACCAAG	*DdeI*	SGN
51342_MH	SL2.40ch11: 51.342	F-ACCCCCACTCCATGATATT R-GCTGGAGAAGCTGGACCATA	None	This study
UP8	SL2.40ch11: 51.344	F-GCGCTGCTAGACATTTCGAT R-CTGAAGTTGCTTGAATGCTCA	None	This study
51355_MH	SL2.40ch11: 51.355	F-GCTAGAGCTTTCAAATCACTCTCAA R-GCTCATTGGCATTCACCTTCT	*FspBI*	This study
51372_MH	SL2.40ch11: 51.372	F-GTTGGGAGCAACTCAGGTGA R-CCAGCACTAGGACAGCTTCC	*Eco32I*	This study
UP15	SL2.40ch11: 51.381	F-TCTCAAAGCGTTGATCGTTG R- GCTTGCTCTTGTTGGTCTCC	*EcoRV*	This study
UF_07960F2	SL2.40ch11: 51.388	F-CGTGCCACCCCTTCATAATA R-CCCTTGCGAGGAAAATACAG	*BanI*	This study
C2_At1g07960	SL2.40ch11: 51.387	F-AAAGCCATTGTTACCGTCTCCGTG R-AGCCATAAGTGGTGTGGAGGACTT	*RsaI*	Ji et al. (2009)
P1-16	SL2.40ch11: 51.426	F-CACACATATCCTCTATCCTATTAGCTG R-CGGAGCTGAATTGTATAAACACG	None	This study
P1-19	SL2.40ch11: 51.432	F-TAACACCAAATCGCGTCTGA R-TTGGGAAAACTATAGCATCG	*AseI*	This study
TG36	SL2.40ch11: 51.490	F-AACCACCACAAGAAAGATCCC R-TCCTGAAATGGAAGATTGCC	*RsaI*	Schmitz et al. (2002)
T0386_MH	SL2.40ch11: 51.499	F-CATTTGCTTTACTGCTAGTGTGC R-GGTTGACCATCTCGAACTCC	None	This study
T0386A	SL2.40ch11: 51.503	F-ATGCTGATGAAAGATTGGGCGCTG R-TTAGGCTTTGGCTTCTCGACCACT	*HinfI*	Ji et al. (2009a)
P3-6	SL2.40ch11: 51.519	F-TGGTGTTTTGTGCGGTAAGA R-TGAAATCGCATGTCCAAAGA	*HaeIII*	This study
P8687	SL2.40ch11: 51.534	F-TACCGTTGCGTAATCTAA R-TTCCACTCAGCATCCCTA	*HinfI*	This study
P4-2	SL2.40ch11: 51.543	F-TCATTCACGGGGAAATTAGG R-CAAGGGACCCAACTTTTTGA	*HinfI*	This study
cLEN-11-F24	SL2.40ch11: 51.549	F-TTATGGACAGCATGGTCCTCGGAA R-GAAGTCTGGGAGCGATAGTAGTCT	*MnlI*	Ji et al. (2009)
cL1	SL2.40ch11: 51.559	F-ATTGCCTACATCTGGTTC R-AAGATACCCACAAGACAA	*HaeIII*	This study
cL2	SL2.40ch11: 51.571	F-GGTAGGGATAAGGTCTGT R-CCTTAGCCGTTACACTCT	*BclI*	This study
P7-8	SL2.40ch11: 51.600	F-TTGACCACGTTTTGGAAATG R-GCAAGAAGACGCTTTTCGAT	*HindIII*	This study
C2_At3g52090	SL2.40ch11: 51.605	F-AGGGATACGAAGATCATGAATGCAGC R-ACTCTTCAGATGATCAAGTTCCTTGTC	*None*	SGN
P8-8	SL2.40ch11: 51.628	F-AGTGGAACTTAATGGCTTTCC R-CGCAATTGACGCATACATTC	*TaqI*	This study
51632_MH	SL2.40ch11: 51.632	F-GGCACTGATGGAGGAGAGTT R-AGCTCACCTGTTGACCTTCA	*DraI*	This study
P8-11	SL2.40ch11: 51.635	F-CGACAGTGTTTTCACCAGCTC R-ACCGAGTATGCACCACCAAT	*RsaI*	This study

M1	SL2.40ch11: 51.645	F-CGCTCGGGCAAATAGTTCGTAATGG R-TTCATGGTCTAGAAATGTCCCCTGT	*BstUI*	This study
M2	SL2.40ch11: 51.661	F-TCAGGGAAGTCTATGTAAACGC R-ATGTGGTAGATAGAAGGGAAGC	*HindIII*	This study
51663_MH	SL2.40ch11: 51.663	F-CCCTCTTGCTTAGTGGGTGA R-ACGCTCCAAATCAGAGGTTG	*Hin6I*	This study
C2_At4g32930	SL2.40ch11: 51.688	F-TCCTCTTCCTATTGGCAAGGGC R-TGGACACTCCCCCTTTTCATCATAC	*Cfr13I*	SGN
51697_MH	SL2.40ch11: 51.697	F-CCCTCAAACCCAAGTGCTTAC R-CTCCAACTTTGCGACTGTTCT	*RsaI*	This study
51752_MH	SL2.40ch11: 51.752	F-ACTCTTGCTTCACTCCTTGGA R-ACCATACCTCAACTTGGAAACA	*SspI*	This study
M3	SL2.40ch11: 51.771	F-TGAATGGAACAGGGCAGAGTAAG R-CTAGTGTCCTTGGTGGTAGTCAT	*TaqI*	This study
BAC_119J05	SL2.40ch11: 51.830	F-AACTTACGGCACCTCAATTTTTC R-GTGCCCCCTATGCAAGTAATTC	*None*	Ji et al. (2009)
T0302	SL2.40ch11: 51.878	F-TGGCTCATCCTGAAGCTGATAGCGC R-TGATKTGATGTTCTCWTCTCTMGCCTG	*None*	Ji et al. (2009)
T0302-1	SL2.40ch11: 51.878	F-TGGCTCATCCTGAAGCTGAT R-TGGAAAGGGATCGAAGAGAA	*None*	This study
TG105A	SL2.40ch11: 52.07	F-ACATTTGGACAAATAGCAGAAGTC R-TGAGAGCAGACAGCAGGCATCATC	*HpyCH4IV*	Ji et al. (2007b)
TG26	SL2.40ch11: 52.53	F-GTCGGTAACAGTTCTATGTTGCGG R-TATTTGGTTCAGTCGTGGAGCC	*HinfI*	Ji et al. (2009a)
TG393	SL2.40ch11: 53.25	F-TGGATTTGATTAGCCGAAGG R-CCAAGAATCCCAGAAGGAGA	*DpnII*	SGN

From Tomato WGS Chromosomes SL2.40 database at http://solgenomics.net/.

[b] K= T/G; W = T/A; M = C/A.

[c] SGN = Sol Genomics Network (http://www.sgn.cornell.edu).

Table S2. Candidate genes in the *Ty-2* region. List shows predicted genes in the 300kb *Ty-2* region (based on the tomato genome sequence, Sol Genomics Network, SGN, www.solgenomics.net); 14 genes silenced and assessed for functionality so far are shown in bold.

Gene number	SGN gene name	Annotation	Expression[a]	VIGS[b]
1	**Solyc11g069620.1**	**CC-NBS-LRR, resistance protein**	nd	**no phenotype**
2	Solyc11g069630.1	Receptor-like protein kinase At5g59670	nd	
3	Solyc11g069640.1	Carbonic anhydrase family protein	nd	
4	Solyc11g069650.1	Unknown protein	nd	
5	Solyc11g069660.1	CC-NBS-LRR, resistance protein	down	
6	**Solyc11g069670.1**	**Disease resistance protein R3a-like protein (fragment)**	nd	**no phenotype**
7	Solyc11g069680.1	Acyltransferase-like protein	nd	
8	Solyc11g069690.1	Protein disulfideisomerase	nd	
9	**Solyc11g069700.1**	**Elongation factor 1-alpha**	down**	**death**
10	Solyc11g069710.1	ABC transporter G family member 3	similar	
11	**Solyc11g069720.1**	**26S protease regulatory subunit 6B homolog**	similar	**no phenotype**
12	**Solyc11g069730.1**	**Unknown Protein**	similar	**yellow leaves**
13	Solyc11g069740.1	Nitrate transporter	ns	
14	Solyc11g069750.1	Nitrate transporter	ns	
15	Solyc11g069760.1	High affinity nitrate transporter protein	ns	
16	**Solyc11g069770.1**	**Transcription factor MADS-box**	up*	**smaller yellowish plants**
17	Solyc11g069780.1	2-phosphoglycerate kinase	up	
18	**Solyc11g069790.1**	**Chaperonin**	similar	**no phenotype**
19	**Solyc11g069800.1**	**Cytochrome P450**	ns	**no phenotype**
20	**Solyc11g069810.1**	**OTU domain containing protein**	similar	**no phenotype**
21	Solyc11g069820.1	ABC transporter G family member 28	down	
22	**Solyc11g069830.1**	**Arsenite ATPase transporter (Eurofung)**	down	**no phenotype**
23	**Solyc11g069840.1**	**Os03g0859900 protein**	similar	**no phenotype**
24	**Solyc11g069850.1**	**Telomere repeat-binding protein 4**	similar	**no phenotype**
25	Solyc11g069860.1	Glutaredoxin	down	
26	Solyc11g069870.1	Ripening-related protein 3	ns	
27	Solyc11g069880.1	Ripening-related protein 3	down	
28	Solyc11g069890.1	BEL1-like homeodomain protein 8	similar	
29	Solyc11g069900.1	Unknown Protein	similar	
30	**Solyc11g069910.1**	**DNA-directed RNA polymerase II subunit J**	down*	**stunting, curling**
31	Solyc11g069920.1	Nbs, resistance protein fragment	similar	
32	**Solyc11g069930.1**	**Disease resistance protein R3a-like protein**	down*	**no phenotype**
33	Solyc11g069940.1	Glutaredoxin	nd	
34	Solyc11g069950.1	Cell division protease ftsH homolog	nd	
35	Solyc11g069960.1	Receptor like kinase, RLK	nd	

[a] Relative expression of genes on the resistant *Ty-2* line vs. susceptible (cv. Moneymaker) genotype. *indicates statistically differential up- or down-regulation (*=P<0.05, **=P<0.01); ns, not signal detected via RT-PCR; nd, not determined.

[b]VIGS: altered phenotype observed upon VIGS (virus-induced gene silencing).

Table S3. Primers used for expression level analysis of candidate genes (RT-PCR)

Gene number[a]	Primer name	Sequence (5'-3')	Target gene (SGN gene name)
1	rt620-F rt620-R	TGATAGAAGGGAAGCCGTGAA GTTTGAGAATCGTAAGCAGAACTACC	Solyc11g069620.1
5	rt660-F rt660-R	TGTGGAACAAGGTGGGCTTC TGAGGTGGGAAAGATGTAGTGAATG	Solyc11g069660.1
9	rt700-F rt700-R	AGTATGCCTGGGTTCTTGAC TGATGAAATCCCTATGACCA	Solyc11g069700.1
10	rt710-F rt710-R	GAAGTCAGAGTGGATGAAGGAGGTG TAAAGGTGGAGATGGCAACGAAC	Solyc11g069710.1
11	rt720-F rt720-R	GTTAGAGGAGGGATGAGAAGTGTGG AAGATTTCAGTGTCTGGAGCAACAA	Solyc11g069720.1
13	rt740-F rt740-R	AAATGGATGTTGGTAATGCTGGAGT GTCACAAATCGCGCCCATAG	Solyc11g069740.1
14	rt750-F rt750-R	GGATGAGAGGAAGGTTATGGGTCTT ATTGAAACTATGGCGATGGGTAATG	Solyc11g069750.1
15	rt760-F rt760-R	GAGAGGCAGACTATGGACATTATGGA ATAGCCAGAGGAAGTGTGGTAGCC	Solyc11g069760.1
16	rt770-F rt770-R	GATGACACTGGCTCCCTCAG TCTTCAAACTCTTCTTCAATCTCCA	Solyc11g069770.1
17	rt780-F rt780-R	TCTGAAGCGAAAGCGAAGAAA TCTGGAACTAATCAAATCAACAGCA	Solyc11g069780.1
18	rt790-F rt790-R	CTCTGCTGGAAATGATGAAAGC AGTTGTCTCTAAGGAGGAGGATGACT	Solyc11g069790.1
19	rt800-F rt800-R	GCTCGGTCGTTATTTGGAGTT AGGTAATCGGAAAGTATGGAGGAG	Solyc11g069800.1
20	rt810-F rt810-R	CCTAGAATCAATGGAGAAATACCATCAA ATCTGACAAAGCACGAAACT	Solyc11g069810.1
21	rt820-F rt820-R	ATGTTGGGCTGGAAATGGTTATG GCTTCACGACGAGTTGCTCTG	Solyc11g069820.1
22	rt830-F rt830-R	AGGAGGGAAGGGAGGTGTGG TGAAACAACAAGAGTCGGATGACC	Solyc11g069830.1
23	rt840-F rt840-R	CAGAGAACCAAGAAACCCGACAC GTTGGAATTGGACTTTGCATGATTT	Solyc11g069840.1
24	rt850-F rt850-R	GAGACCCGAGGAACTAACCAAGACA GCTCCATCTGCCAACACCATACTT	Solyc11g069850.1
25	rt860-F rt860-R	TTCTACATCGTTATCCTCCATACCC ATGAATCTCCTCCTCCAGCAAC	Solyc11g069860.1
26	rt870-F rt870-R	ACAACTTTCATCTCTTCATCTTACTCCA ACTTCCTCCTCCTCCACAAATACTG	Solyc11g069870.1

27	rt880-F rt880-R	CCCTCCTTCTGTCCCTTCTCC GGTTAGTTGAGCGGGCGTTG	Solyc11g069880.1
28	rt890-F rt890-R	GAAACCGTGGCTGGTCTTAGTG ATCTGTGATGGCGTTCCTCAGT	Solyc11g069890.1
29	rt900-F rt900-R	GACAACGACTTCTAGCTTTGCTACG GCCCTCGTTCCAATAGGGTTT	Solyc11g069900.1
30	rt910-F rt910-R	AAGAGAGGACCATACAATCGGGAAC CATAGGTGAGGACTGGCTTGTTGTT	Solyc11g069910.1
31	rt920-F rt920-R	GGAGAGTGTTGCTCAGACGATG TTCAAGTGCTGGATGATTCCTATTT	Solyc11g069920.1
32	rt930-F rt930-R	CGGAAGCGTAAGAATGATGTTGAG TGATTTGATGCCTGCTTATTCTCTG	Solyc11g069930.1

[a] The same as the gene number in Table S2.

Table S4. Sequences of primers used for generation of VIGS constructs.

Gene number[a]	Primer name	Sequence (5'-3')	Target gene (SGN gene name/Annotation)
1	VG620-F	CACCTCGGAAGCGTAAGAATGATGT	Solyc11g069620.1
	VG620-R	TCTACAGCATGTCGAAGCTCA	CC-NBS-LRR, resistance protein
6	VG670-F	CACCTCGGAAGCGTAAGAATGATGT	Solyc11g069670.1
	VG670-R	TCTACAGCATGTCGAAGCTCA	Disease resistance protein R3a-like protein (fragment)
9	VG700-F	CACCGGGTCCAACCCTTCTTGAGG	Solyc11g069700.1
	VG700-R	AGGCTCCTTCTCGAGTTCCT	Elongation factor 1-alpha
11	VG720-F	CACCGCACCAGTGTCCAGATGTCA	Solyc11g069720.1
	VG720-R	CTTTCCAGAACCAGGGGGTC	26S protease regulatory subunit 6B homolog
12	VG730-F	CACCACGAGTATTCGCGAGTTATCCA	Solyc11g069730.1
	VG730-R	TGTGAACATCCAATTGAGCCC	Unknown Protein
16	VG770-F	CACCCGCAGAAGAATCCGCATTGA	Solyc11g069770.1
	VG770-R	CCATTAGAGCCAACACCCCC	Transcription factor MADS-box
18	VG790-F	CACCACAAGGGTGCAAAGAACCGA	Solyc11g069790.1
	VG790-R	TGTCGTCACGGCCTTTAACA	Chaperonin
19	VG800-F	CACCGCAGTGATCAAGCTGCGTTC	Solyc11g069800.1
	VG800-R	GGATCTCCACCCCTGCTTTC	Cytochrome P450
20	VG810-F	CACCGCTTTCAGTCGCTGAACCAC	Solyc11g069810.1
	VG810-R	GGCCATGGGAACATATCCGT	OTU domain containing protein
22	VG830-F	CACCTGGTTAGTGGGACACAACGG	Solyc11g069830.1
	VG830-R	CTGCGTGGAACTCTTCCCTT	Arsenite ATPase transporter (Eurofung)
23	VG840-F	CACCGCATCAACATCTATGGAGCCC	Solyc11g069840.1
	VG840-R	TGCTGGAGAAGAACGTGTGC	Os03g0859900 protein
24	VG850-F	CACCTTCGCGAGGAGGATTTGGTC	Solyc11g069850.1
	VG850-R	CAGCTGGACTACGATGCACA	Telomere repeat-binding protein 4
30	VG910-F	CACCTTCGTTGTTCCAGAAGGCGT	Solyc11g069910.1
	VG910-R	AGGTGAGGACTGGCTTGTTG	DNA-directed RNA polymerase II subunit J
32	VG930-F	CACCTCGGAAGCGTAAGAATGATGT	Solyc11g069930.1
	VG930-R	TCTACAGCATGTCGAAGCTCA	Disease resistance protein R3a-like protein

[a] The same as the gene number in Table S2.

Figure S1. Target regions for silencing the *R3a* homologs in tomato chromosome 11. Nucleotide sequence alignments of predicted R3a homologs in the *Ty-2* region are shown: Disease resistance protein R3a-like fragment (Solyc11g069670.1), Disease resistance protein *R3a*-like protein (Solyc11g069930.1), and TRV-based VIGS construct VG930. Regions highlighted in black represent sequences targeted for VIGS for each predicted gene. cDNA sequences were obtained from SGN public database.

```
Solyc11g069670.1  :     TCGGACGCGTAAGAATGGTGTTGAGCTCTTAGAGAAGCTGAAATCGACTT
Solyc11g069930.1  :     TCGGAAGCGTAAGAATGATGTTGAGCTCTTAAAGAAGCTGAAATTGACTT
VG930             :     TCGGAAGCGTAAGAATGATGTTGAGCTCTTAAAGAAGCTGAAATTGACTT

Solyc11g069670.1  :     TGCTTGGACTTCAAGCTGTGCTAAGTGATGCAGAGAATAAGCAGGCATCA
Solyc11g069930.1  :     TGCTTGGTCTTCAAGCTGTGCTAACTGAGCAGAGAATAAGCAGGCATCA
VG930             :     TGCTTGGTCTTCAAGCTGTGCTAACTGATGCAGAGAATAAGCAGGCATCA

Solyc11g069670.1  :     AATCAATTTGTGAGAGAATGGCTTAATGAGCTTCGACATGCTGTAGACTC
Solyc11g069930.1  :     AATCAATTTGTGAGAGAGTGGCTTAATGAGCTTCGACATGCTGTAGACTC
VG930             :     AATCAATTTGTGAGAGAGTGGCTTAATGAGCTTCGACATGCTGTAGAA--


Solyc11g069670.1  :     TGCTGAAAACATGATTG
Solyc11g069930.1  :     TGCTGAAAACTTGATTG
VG930             :     -----------------
```

Chapter 4

Detection of an inversion in the *Ty-2* region between *S. lycopersicum* and *S. habrochaites* by a combination of *de novo* genome assembly and BAC cloning

Myluska Caro[1,§], Anne-Marie A. Wolters[1,§], Shufang Dong[2,§], Richard Finkers[1], Jianchang Gao[2], Richard G.F. Visser[1], Xiaoxuan Wang[2], Yongchen Du[2], Yuling Bai[1]

[1]Wageningen UR-Plant Breeding, Wageningen University & Research Centre, P.O. Box 386, 6700AJ Wageningen, The Netherlands
[2]The Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences Zhongguancunnandajie 12, Beijing 100081, People's Republic of China

§ Myluska Caro, Anne-Marie A. Wolters and Shufang Dong equally contributed to this work.

This chapter has been published in Theoretical and Applied Genetics, 2015 :2561

Abstract

Among tomato and its wild relatives inversions are often observed, which result in suppression of recombination. Such inversions hamper the transfer of important traits from a related species to the crop by introgression breeding. Suppression of recombination was reported for the TYLCV resistance gene, *Ty-2,* which has been introgressed in cultivated tomato (*Solanum lycopersicum*) from the wild relative *S. habrochaites* accession B6013. *Ty-2* was mapped to a 300-kb region on the long arm of chromosome 11. The suppression of recombination in the *Ty-2* region could be caused by chromosomal rearrangements in *S. habrochaites* compared with *S. lycopersicum.* With the aim of visualizing the genome structure of the *Ty-2* region, we compared the draft *de novo* assembly of *S. habrochaites* accession LYC4 with the sequence of cultivated tomato ('Heinz'). Furthermore, using populations derived from intraspecific crosses of *S. habrochaites* accessions, the order of markers in the *Ty-2* region was studied. Results showed the presence of an inversion of approximately 200 kb in the *Ty-2* region when comparing *S. lycopersicum* and *S. habrochaites*. By sequencing a BAC clone from the *Ty-2* introgression line, one inversion breakpoint was identified. Finally the obtained results are discussed with respect to introgression breeding and the importance of *a priori de novo* sequencing of the species involved.

Keywords: *Ty-2* region, TYLCV, *Solanum habrochaites*, chromosomal rearrangement, comparative genomics

Introduction

In genetics, introgression (also known as introgressive hybridization) is the transfer of a gene from one species into the gene pool of another species. Such a transfer starts with an interspecific hybridisation and is followed by backcrossings with one of the parental species. In breeding, introgression is an important strategy to broaden the genetic base of highly inbred crops such as tomato by transferring economically important traits from a related species to the crop. One of the major problems in introgression breeding is caused by chromosomal rearrangements, such as inversions and translocations, between the donor species and the crop (Szinay et al. 2010). Genetic maps created from different intraspecific or interspecific crosses using the same markers can indicate co-linearity or a change in order of markers in distinct chromosomal regions. In the Solanaceae family, genetic maps have been used to detect chromosomal rearrangements (e.g. Wu and Tanksley 2010; Doğanlar et al. 2014). Also, the study of pachytene synaptonemal complexes in interspecific F1 hybrids can indicate the presence of chromosomal rearrangements (Anderson et al. 2010). Meanwhile, cross-species BAC fluorescence *in situ* hybridisation (FISH) analysis has been shown to be a powerful instrument to identify chromosomal rearrangements in the Solanaceae family (van der Knaap et al. 2004; Iovene et al. 2008; Tang et al. 2008) and more specifically, among related species of *Solanum* (Lou et al. 2010; Verlaan et al. 2011; Peters et al. 2012; Szinay et al. 2012; Shearer et al. 2014). In introgression breeding this technology can be used as a diagnostic tool to monitor meiotic disturbances in the pairing of homoeologous chromosomes from crops and their related species. Nowadays, the released full genome sequences of closely related species have facilitated comparative genome analysis. Occurrences of both large scale and small scale rearrangements have been reported between tomato and potato genomes (The Potato Genome Sequencing Consortium 2011; The Tomato Genome Consortium 2012). For a co-linearity study of two genomes, a reference genome sequence should not only be available for the cultivated species, but also for the wild donor species. To this end, *de novo* assembly of genome sequences of three wild relatives of tomato has been undertaken recently (The 100 Tomato Genome Sequencing Consortium 2014).

Among tomato and its wild relatives, inversions are often observed (Szinay et al. 2012) which can cause meiotic pairing disturbances between homologues. Crossovers are unlikely to occur in the inverted region, which results in suppression of recombination (Szinay et al. 2010). Thus, the inverted region will be genetically inherited as one locus during the introgression and many unwanted

sequences in the inverted region will be transferred together with the gene of interest from the wild donor to the crop species, a phenomenon known as linkage drag. A good example is the *Ty-1* gene which originated from *S. chilense* and confers resistance to tomato yellow leaf curl virus (TYLCV) (Verlaan et al. 2013). Chromosomal rearrangements between *S. chilense* and the cultivated tomato were detected by BAC-FISH (Verlaan et al. 2011). These rearrangements caused severe suppression of recombination in the *Ty-1* region and thus hampered the *Ty-1* introgression (Verlaan et al. 2011). Suppression of recombination was reported for another TYLCV resistance gene, *Ty-2,* which has been introgressed in cultivated tomato (*Solanum lycopersicum*) from the wild relative *S. habrochaites* accession B6013 (Kalloo and Banerjee 1990). The gene has been mapped on the long arm of chromosome 11 (Hanson et al. 2000; Ji et al. 2009), and fine mapped to a 300-kb region (Yang et al. 2014). Attempts to clone the gene have been hampered by the occurrence of severe suppression of recombination in a large part of this region. Such suppression of recombination could be caused by chromosomal rearrangements in *S. habrochaites* compared with *S. lycopersicum,* as shown previously for the *Ty-1* gene (Verlaan et al. 2011). However, for the *Ty-2* region, FISH on pachytene chromosomes using three BACs spanning the introgression region and one BAC outside the region resulted in overlapping fluorescing signals (Yang et al. 2014). Because of this, the order of the BACs could not be determined, and therefore no conclusion could be drawn on the cause of the suppression of recombination.

In order to visualize the genome structure of the *Ty-2* region, in this study we combined *de novo* genome assembly and BAC cloning. First, the draft *de novo* assembly of *S. habrochaites* accession LYC4 was compared with the genome sequence of cultivated tomato ('Heinz') to determine whether a chromosomal rearrangement has occurred in the *Ty-2* region. Secondly, BAC cloning of the *Ty-2* introgression line was performed. Furthermore, recombinant screening of F_2 populations derived from intraspecific crosses of *S. habrochaites* accessions was carried out. Taken together, the results showed the presence of an inversion of approximately 200 kb in the *Ty-2* region when comparing *S. lycopersicum* and *S. habrochaites*.

Results

Bioinformatic comparison of the *Ty-2* region from *S. lycopersicum* and *S. habrochaites*.

The region containing the *Ty-2* resistance gene was determined to span a 300-kb sequence at the distal end of the long arm of chromosome 11, flanked by markers UP8 and M1 (Figure 1A, Yang et al. 2014). This corresponds to the region between nucleotides 51,344,943 and 51,646,517 on chromosome 11 of tomato genome version SL2.40, or between nucleotides 54,261,443 and 54,563,017 on chromosome 11 of tomato genome version SL2.50. We prefer to use the coordinates of the SL2.40 version in this paper for easy reference to previous articles.

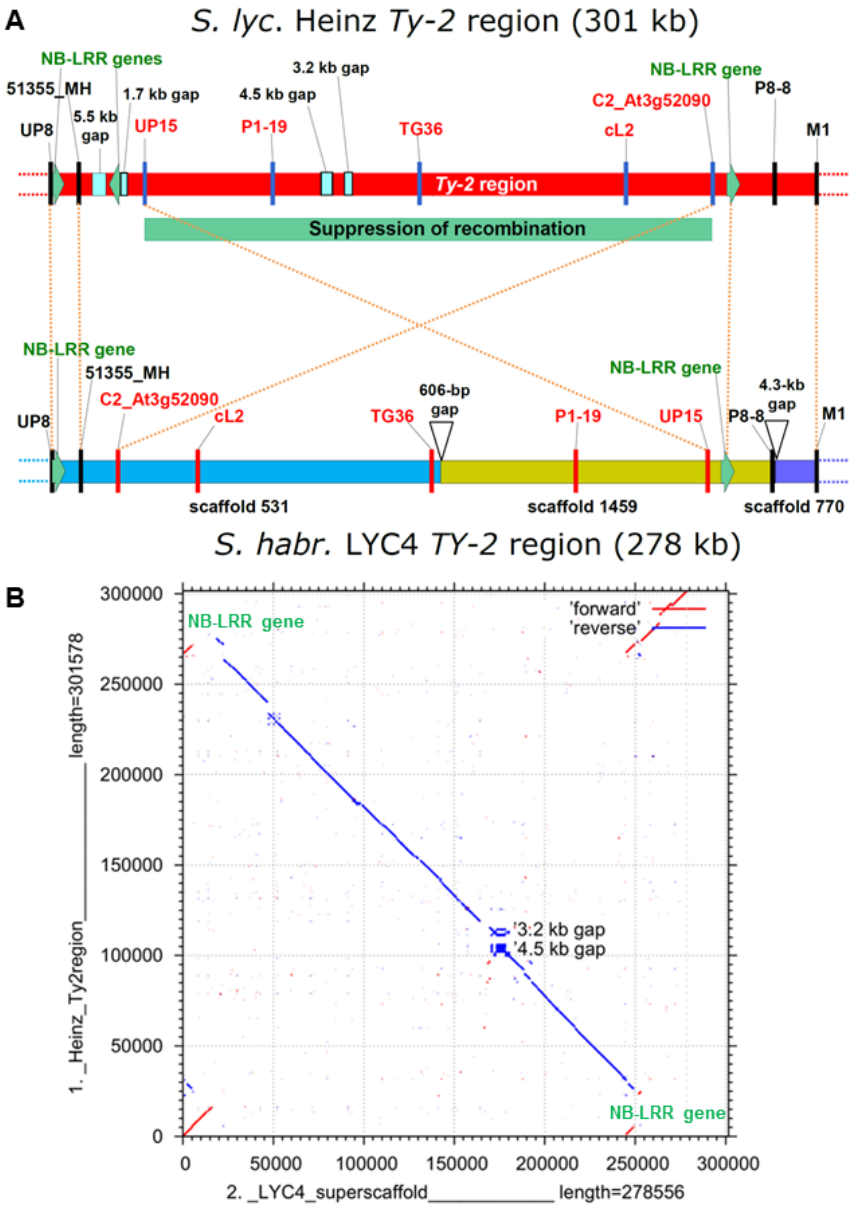
A BLAST analysis of this region was performed against the draft *de novo* assembly of the *S. habrochaites* LYC4 genome. Three large scaffolds (531, 1459 and 770) spanned most of the *Ty-2* region (Figure 1A). Interestingly, *S. habrochaites* LYC4 scaffold 531 contained both the flanking marker UP8 and marker C2_At3g52090 at a distance of only 26 kb, whereas in *S. lycopersicum* 'Heinz' the distance between these two markers is 262 kb. Additionally, *S. habrochaites* LYC4 scaffold 1459 contained both markers P8-8 and UP15 at a distance of 24 kb, whereas in *S. lycopersicum* 'Heinz' the distance between these two markers is 247 kb.

The three scaffolds 531, 1459 and 770 of *S. habrochaites* LYC4 were connected to form a superscaffold. To confirm the linkage between the scaffolds PCRs were performed. PCR products spanning the gaps between the scaffolds were obtained and sequenced. The gap between scaffolds 531 and 1459 proved to be small (606 bp, Figure 1A). The gap between scaffolds 1459 and 770 was larger, approximately 4.3 kb. Thus, by closing the gaps we confirmed the orientation of the three scaffolds.

By aligning the *Ty-2* regions of *S. lycopersicum* 'Heinz' and *S. habrochaites* LYC4 we observed an inversion of ±200 kb in the central part (Figure 1B). Within this inversion there is good co-linearity between 'Heinz' and LYC4, except for some gaps (unknown sequences) in the assemblies, of which the largest ones are indicated in Figure 1A.This inversion coincides with the 'suppression of recombination' block in progeny of the interspecific cross between *S. lycopersicum* and the *Ty-2* donor *S. habrochaites* B6013 (Yang et al. 2014).

Figure 1. Comparison between the *Ty-2* genomic region of *S. lycopersicum* 'Heinz' and the superscaffold spanning the *Ty-2* region in *S. habrochaites* LYC4. **A** Visual representation of the *Ty-2* regions in 'Heinz' and LYC4. Markers in the 'suppression of recombination' block are indicated in red (UP15, P1-19, TG36, cL2 and C2_At3g52090), and the other markers are in black (UP8, 51355_MH, P8-8 and M1). Gaps in the 'Heinz' sequence are shown in light blue bars and the sizes of these gaps are estimated by the

number of “N” in the tomato genome. NBS-LRR genes are indicated as green arrows. Orange dotted lines connect homologous sequences in LYC4 compared with ‘Heinz’.
B Dot plot of the alignment of the *Ty-2* regions of ‘Heinz’ and LYC4. Red lines indicate co-linearity; blue lines indicate inversion. The gaps in the ‘Heinz’ sequence disrupt the co-linearity of the two sequences.

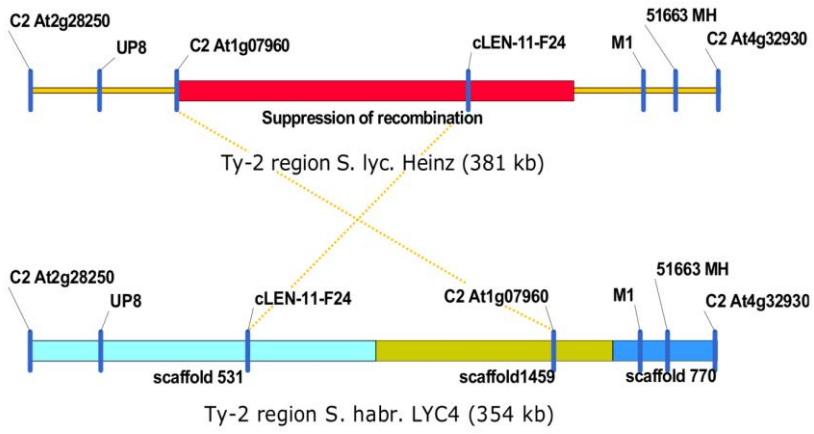


Recombinant screening within *S. habrochaites* species

Previously, Yang et al. (2014) reported a severe suppression of recombination in the *Ty-2* region in an interspecific cross between *S. lycopersicum* and a BC_4S_2 introgression line derived from *S. habrochaites* B6013 (donor of the *Ty-2* gene). Among 11,000 F_4 plants no recombinants were observed between markers TG36 and C2_At3g52090 (Figure 1).

We investigated whether suppression of recombination in the *Ty-2* region is also occurring in intraspecific *S. habrochaites* crosses. For this, we analysed F_2 populations of three crosses between four different *S. habrochaites* accessions. All crosses had one parent in common, which is accession G1.1560. This accession was chosen as the common parent because it shows a relatively high level of polymorphisms compared with the other three accessions that are more similar to each other. Recombinant screening was performed on 91 to 287 F_2 progeny per cross using selected CAPS markers in the *Ty-2* region that had been shown to be polymorphic between G1.1560 and the other *S. habrochaites* accessions (Figure 2). These include one marker above the 'suppression of recombination' block (C2_At2g28250), two markers within the block (C2_At1g07960/UF_07960 and cLEN-11-F24), and three markers below the block (M1, 51663_MH and C2_At4g329530). Markers C2_At1g07960 and UF_07960 are derived from the same gene, but amplify different fragments. Polymorphisms could be detected in one or the other marker, depending on the crossing population.

Figure 2. Confirmation of inversion in the *Ty-2* region by genetic analysis. CAPS markers used for recombinant screening of intraspecific *S. habrochaites* F_2 populations are indicated. UP8 is included as a reference to delineate the *Ty-2* region, but was not used as marker.



First, we analysed occurrence of recombination between markers C2_At1g07960/ UF_07960 and cLEN-11-F24. The physical distance between these markers in the *S. lycopersicum* 'Heinz' genome is 162 kb, while the genetic distance between these markers is 4.5 cM in the Tomato-EXPEN 2000 genetic map. In total, 21 recombinants were found between these two markers (Table 1), seven in population 1 (PV960357, 91 plants), six in population 2 (PV970303, 287 plants), and eight in population 3 (PV960350, 96 plants). This indicated that there is no suppression of recombination in this region in intraspecific *S. habrochaites* crosses, although the genetic distance between these two markers varies among the crosses (2 to 8 cM).

Table 1. Recombinants in the 'suppression of recombination block' in three populations derived from crosses between different *S. habrochaites* accessions. Marker order in the **A** panels is based on the *S. habrochaites* LYC4 *de novo* sequence; markers in the **B** panels are ordered according to the *S. lycopersicum* 'Heinz' genome sequence. Markers highlighted in dark grey are located in the 'suppression of recombination block'. Numbers under marker names correspond to their positions on chromosome 11 of tomato 'Heinz' genome sequence SL2.40 (in Mbp). Markers were scored in the following way: A, homozygous for *S. habrochaites* G1.1560 allele; B, homozygous for other parent *S. habrochaites* allele; H, heterozygous; AH, no distinction possible between homozygosity for G1.1560 allele or heterozygosity.

Population PV960357 (G1.1560 x G1.1606)

A

	Markers with physical position (Mbp) in chromosome 11 of Heinz SL2.40					
Rec	C2_At2g28250	cLEN-11-F24	UF_07960F2	M1	51663_MH	C2_At4g32930
#	51.307	51.549	51.388	51.645	51.663	51.688
1	A	A	H	H	AH	H
2	A	A	H	H	AH	H
3	A	A	H	H	AH	H
4	H	H	B	B	B	B
5	H	H	B	B	B	B
6	H	H	B	B	B	B
7	H	H	A	A	AH	A

B

	Markers with physical position (Mbp) in chromosome 11 of Heinz SL2.40					
Rec	C2_At2g28250	UF_07960F2	cLEN-11-F24	M1	51663_MH	C2_At4g32930
#	51.307	51.388	51.549	51.645	51.663	51.688
1	A	H	A	H	AH	H
2	A	H	A	H	AH	H
3	A	H	A	H	AH	H
4	H	B	H	B	B	B
5	H	B	H	B	B	B
6	H	B	H	B	B	B
7	H	A	H	A	AH	A

Population PV970303 (G1.1290 x G1.1560)

A

	Markers with physical position (Mbp) in chromosome 11 of Heinz SL2.40					
Rec	C2_At2g28250	cLEN-11-F24	UF_07960F2	M1	51663_MH	C2_At4g32930
#	51.307	51.549	51.388	51.645	51.663	51.688
1	H	H	B	B	B	B
2	H	H	A	A	AH	A
3	H	H	A	A	AH	A
4	H	H	A	A	AH	A
5	A	A	H	H	AH	H
6	A	A	H	H	AH	H

B

	Markers with physical position (Mbp) in chromosome 11 of Heinz SL2.40					
Rec	C2_At2g28250	UF_07960F2	cLEN-11-F24	M1	51663_MH	C2_At4g32930
#	51.307	51.388	51.549	51.645	51.663	51.688
1	H	B	H	B	B	B
2	H	A	H	A	AH	A
3	H	A	H	A	AH	A
4	H	A	H	A	AH	A
5	A	H	A	H	AH	H
6	A	H	A	H	AH	H

Population PV960350 (G1.1257 x G1.1560)

A

	Markers with physical position (Mbp) in chromosome 11 of Heinz SL2.40					
	C2_At2g28250	cLEN-11-F24	C2_At1g07960	M1	51663_MH	C2_At4g32930
Rec				51.64		
#	51.307	51.549	51.387	5	51.663	51.688
1	H	H	B	B	B	B
2	B	H	A	A	AH	A
3	H	H	A	A	AH	A
4	H	H	A	A	AH	A
5	B	B	H	H	AH	H
6	H	H	A	A	AH	A
7	H	H	B	B	B	B
8	A	A	H	H	AH	H

B

	Markers with physical position (Mbp) in chromosome 11 of Heinz SL2.40					
	C2_At2g28250	C2_At1g07960	cLEN-11-F24	M1	51663_MH	C2_At4g32930
Rec						
#	51.307	51.387	51.549	51.645	51.663	51.688
1	H	B	H	B	B	B
2	B	A	H	A	AH	A
3	H	A	H	A	AH	A
4	H	A	H	A	AH	A
5	B	H	B	H	AH	H
6	H	A	H	A	AH	A
7	H	B	H	B	B	B
8	A	H	A	H	AH	H

To determine marker order in *S. habrochaites*, markers flanking the 'suppression of recombination' block (UP8, C2_At2g28250, M1, 51663_MH and C2_At4g32930) were included in the analysis (Table 1). When the markers are ordered according to the *S. lycopersicum* 'Heinz' genome three crossovers in a relatively small region of 338 kb are required to explain the obtained recombinant genotypes. However, a single recombination is sufficient to explain these genotypes when the order of markers C2_At1g07960 and cLEN-11-F24 is reversed. This strongly suggests that an inversion of the region containing these two markers is present in multiple *S. habrochaites* accessions compared with *S. lycopersicum*.

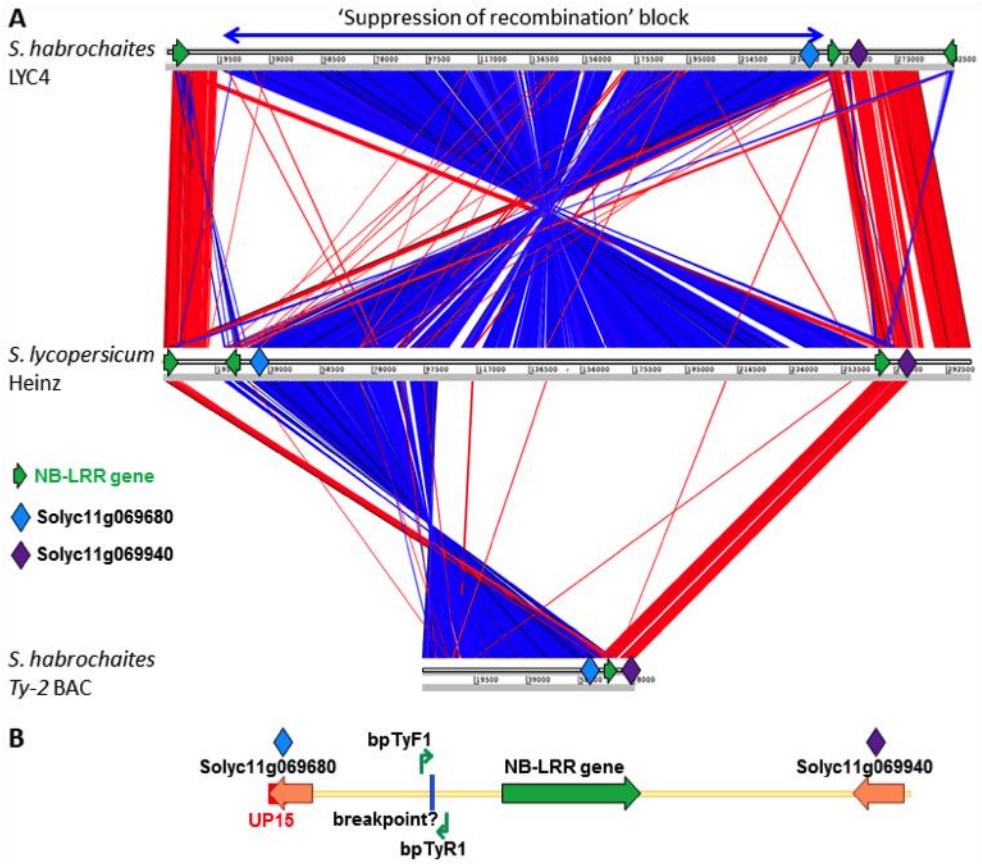
To investigate whether suppression of recombination in the *Ty-2* region is unique to the cross described by Yang et al. (2014) we analysed 88 F_2 plants from a different interspecific cross, between *S. lycopersicum* 'Moneymaker' (MM) and TYLCV-susceptible *S. habrochaites* accession G1.1257 (parent of population 3, PV960350). No recombination events were found between markers C2_At1g07960/UF_07960 and cLEN-11-F24, suggesting a suppression of recombination in this population.

Analysis of inversion breakpoints

So far, only a draft version of the *de novo* assembly of the *S. habrochaites* LYC4 genome is available. Alignment of the LYC4 superscaffold to the 'Heinz' genome sequence showed that the inversion was flanked by NBS-LRR-like genes in inverted orientation in the 'Heinz' genome (Figure 3A). One could argue that the inversion in the LYC4 superscaffold is due to misassembled sequences. To obtain evidence for the presence of an inversion in the *Ty-2* region in *S. habrochaites* compared with *S. lycopsercium* a BAC library was made of a *Ty-2* introgression line. This line contains a small introgression of the *Ty-2* region from *S. habrochaites* 'B6013' (donor of the *Ty-2* gene) in an otherwise *S. lycopersicum* background. A BAC containing the UP15 marker (Figure 1A) was obtained and sequenced. Alignment of this sequence to the *S. lycopersicum* 'Heinz' sequence (Figure 3A) showed that a large part was homologous to the upper end of the inversion in *S. lycopersicum*, which is as expected based on the location of the UP15 marker which is derived from gene Solyc11g069680 (Figure 3B). However, additionally it contained a sequence homologous to gene Solyc11g069940, which is close to M1 (Figure 1), a marker at the other side of the inversion. Thus, the BAC contained predicted genes homologous to Solyc11g069680 and Solyc11g069940 in close proximity (18 kb) (Figure S1). In between these genes an NBS-LRR type of gene is predicted that shows homology to both Solyc11g069660 and Solyc11g069930.

A detailed analysis of the BAC sequence was performed to determine the location of the inversion breakpoint (Figure 4A). Primers flanking the potential breakpoint were designed on the *Ty-2* BAC sequence (Figure 3B and 4A). They amplified a 732-bp fragment in the *Ty-2* introgression line (Figure 4B). As expected, no PCR product was obtained with *S. lycopersicum* MM DNA. In the 'Heinz' sequence the reverse primer bpTyR1 is present in the same orientation as in the *Ty-2* BAC sequence, while the forward primer bpTyF1 is present in the inverse orientation. However, reverse primer bpTyR1 is also present on the other side of the inversion in 'Heinz', between markers 51355_MH and UP15. A PCR product of approximatley 8.9 kb might be obtained if the adequate PCR conditions for long PCR products would be applied. Remarkably, also no PCR product was obtained for *S. habrochaites* LYC4 (Figure 4B). When comparing the *Ty-2* BAC sequence with the LYC4 superscaffold sequence we found that the forward primer bpTyF1 was in the expected position above an NBS-LRR gene (Figure 4A). However, the reverse primer bpTyR1 was present below the NBS-LRR gene, in the same orientation as the forward primer bpTyF1. This explains why no amplification product was obtained for LYC4.

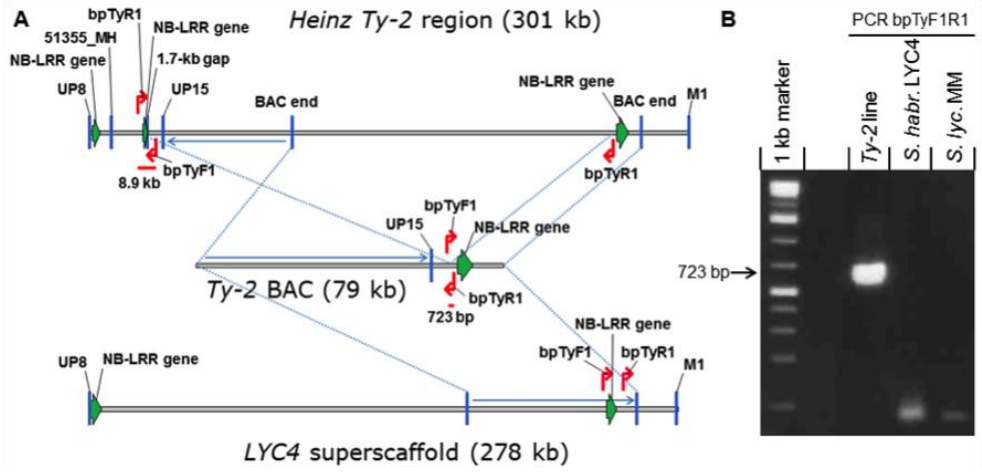
Figure 3. Comparative sequence analyses. **A** Visualization of the alignment of the *S. habrochaites* LYC4 *Ty-2* region and the BAC sequence from the *Ty-2* introgression line with the *S. lycopersicum* Heinz *Ty-2* region. Red lines indicate homology in the direct orientation (+ strand). Blue lines indicate homology in the inverse orientation (– strand). The blue lines indicate the presence of an inversion when comparing *S. habrochaites* with *S. lycopersicum*. NBS-LRR gene sequences (indicated as green arrows) align to different positions in both direct and inverse orientation. Positions of genes homologous to Solyc11g06980 and Solyc11g069940 are indicated. **B** Part (18 kb) of the *Ty-2* BAC containing the lower inversion breakpoint. Primers (bptyF1 and bpTyR1) spanning the putative inversion breakpoint are indicated.



The 732-bp sequence of the PCR product obtained from the *Ty-2* introgression line (Figure 4B) was aligned with the sequence upstream of the unique primer bpTyF1 in the 'Heinz' genome (Figure S1A). These sequences showed a poor alignment, except for the first 154 bp starting from primer bpTyF1. In order to verify this breakpoint region in cultivated tomato, primer bpTyR2 was developed based on the 'Heinz' genome sequence (Figure S1B). A PCR with primers bpTyF1 and bpTyR2 resulted in the expected 687-bp product in *S. lycopersicum* MM but not in the *Ty-2* introgression line (Figure S1C). The sequence of the 687-bp PCR product of MM was identical to the sequence in 'Heinz'. Although the alignment of the sequences of the two PCR products show an abrupt end of co-linearity it is preliminary to conclude that this is the exact breakpoint of the inversion. To verify this conclusion, we need to know the sequence of the upper breakpoint region in the *Ty-2* introgression line. Figure S1B shows the regions habouring the upper and lower

breakpoints in *S. lycopersicum* 'Heinz', and the lower breakpoint in the BAC clone from the *Ty-2* introgression line. Gene Solyc11g069680 is present in the upper breakpoint region, while gene Solyc11g069940 is present in the lower breakpoint region in the 'Heinz' genome. Both genes are adjacent to NBS-LRR gene fragments. The *Ty-2* BAC sequence contains orthologs of both Solyc11g069680 and Solyc11g069940, separated by a NBS-LRR gene.

Figure 4. Analysis of inversion breakpoint. **A** Location of breakpoint primers bpTyF1 and bpTyR1 (red arrows) flanking the inversion breakpoint in the BAC sequence obtained from the *Ty-2* introgression line (middle sequence). Location of these primers is also shown in the *S. lycopersicum* 'Heinz' region (upper sequence) and *S. habrochaites* LYC4 superscaffold (lower sequence). Green arrows indicate sequences homologous to NBS-LRR genes. Blue dotted lines indicate co-linear regions. **B** PCR with primers bpTyF1 and bpTyR1 flanking the inversion breakpoint. Only in the *Ty-2* introgression line a PCR product is obtained.



Discussion

Chromosomal rearrangements are frequently associated with resistance gene clusters

We report the presence of an inversion of a ±200 kb region on the long arm of chromosome 11 in *S. habrochaites* compared with *S. lycopersicum*. This inversion is different from the 294-kb inversion underlying the *fasciated* locus on the long arm

of chromosome 11, which is polymorphic within the cultivated *S. lycopersicum* germplasm (Huang and Van der Knaap 2011).

There are numerous examples of chromosomal rearrangements/inversions associated with (introgression of) R-genes or R-gene clusters. Two R-gene clusters in the *Mi* locus on the short arm of chromosome 6 in *Solanum peruvianum* are separated by approximately 300 kb region, which is inverted compared to *S. lycopersicum* (Seah et al. 2007). The introgression of the *Ty-1* locus on the long arm of chromosome 6 from *S. chilense* in *S. lycopersicum* background shows an inversion and suppression of recombination (Verlaan et al. 2011). The *H1* locus on the distal end of chromosome 5 of potato (Finkers-Tomczak et al. 2011) shows repression of recombination in a region of at least 170 kb. The *R1* locus in the same region was shown to be present in a region that was inverted in tomato compared with potato (Achenbach et al. 2010). The donor of the *R1* gene, *S. demissum*, contained haplotypes that were highly diverged in the R-gene cluster region, while the flanking non-resistance gene regions were conserved (Kuang et al. 2005). A 70-kb inversion between the resistant R1 and the susceptible r1 haplotypes was reported by Ballvora et al. (2007). The clubroot resistance region in *Brassica rapa* has an internal inversion compared with *Arabidopsis* of about 310 kb (Suwabe et al. 2012). Suppression of recombination in these R-gene regions may be a consequence of the chromosomal rearrangement. On the other hand, suppressed recombination may also be caused by the pericentromeric position of the introgression rather than the inversion, as is the case for the *Mi-1* locus (Seah et al. 2007).

For other resistance gene loci suppressed recombination has been reported, but it is unknown whether this is a consequence of chromosomal rearrangements, and/or of (peri)centromeric locations. These include the *Tm-2a* gene from *S. peruvianum* introgressed in *S. lycopersicum* (Pillen et al. 1996), the *MXC3* gene in poplar, the *Lr20-Sr15-Pm1* resistance locus and *Sr22*, *Lr9*, *Lr24* and *Lr35* resistance genes in wheat, the *Mla* and *Mlg* powdery mildew resistance gene clusters and the *Rrs2* resistance gene in barley (reviewed in Hanemann et al. 2009), and the *Rhg1*/*Rfs2* locus in soybean (Afzal et al. 2012).

Chromosome rearrangements complicate the fine-mapping and cloning of resistance genes, especially when they involve large regions containing many genes. In the case of the *Ty-2* resistance gene it was shown previously that it is unlikely to be a typical NBS-LRR gene, because silencing of the NBS-LRR

candidates in the *Ty-2* region did not result in compromised TYLCV resistance (Yang et al. 2014).

Advantage of *de novo* genome assemblies of wild relatives of crop species

After the assembly of the genome of cultivated crop species the focus has shifted to sequencing related wild species at low read depth to obtain information on sequence variation by mapping reads to the reference genome. The assumption is that there is a high degree of co-linearity within a species and between closely related species, and that a large set of SNP markers developed after re-sequencing can be used to fine map traits of interest. However, as shown by Huang and van der Knaap (2011) chromosomal rearrangements may occur even within a cultivated species. Re-sequencing data consisting of small reads do not provide positional information of SNP markers, or SNP marker order. Therefore, such data do not uncover the presence of chromosomal rearrangements in wild species, especially those that are not closely related to the cultivated species as shown in tomato (Szinay et al. 2012). FISH using BAC clones has been demonstrated to be a powerful tool in the study of chromosomal rearrangements (Lou et al. 2010; Verlaan et al. 2011; Peters et al. 2012; Szinay et al. 2012; Shearer et al. 2014). However, in the *Ty-2* region, FISH was not successful due to the small size of the inversion (Yang et al. 2014). In this study, we show that a *de novo* genome assembly has been very helpful to analyse the chromosomal structure of a wild species, which can be exploited to explain unexpected recombination phenomena in crosses with the cultivated species.

Also within a wild species there may be accessions that show small-scale rearrangements, as we observed when comparing the inversion breakpoint between *S. habrochaites* LYC4 and the *Ty-2* BAC sequence derived from *S. habrochaites* B6013. Therefore, BAC libraries may still be required to zoom in on the gene of interest in specific accessions.

Perspectives for resistance gene cloning

Introgression of the smallest possible DNA fragment containing the gene of interest from a donor species into the crop species is often a time-consuming process, and the success can be limited when chromosomal rearrangements exist in related species used for interspecific crosses. Since genome structure and genomic co-linearity of the introgressed region between donor species and recipient crops are often unknown, breeders are 'blind' and cannot foresee complications in their introgression breeding programs. With the example of the *Ty-1* gene (Verlaan et al.

2011) and the *Ty-2* gene in this study, we demonstrated that FISH and genomic approaches can be applied to investigate chromosomal rearrangements in genetic mapping and introgression breeding. Furthermore, the occurrence of chromosomal rearrangements stresses the importance of a *de novo* genome assembly when wild *Solanum* species are sequenced.

The fact that *Ty-2* is located in a chromosomal region which is inverted in *S. habrochaites* compared with *S. lycopersicum* has consequences for the strategy of cloning this gene. For marker-assisted breeding it is not necessary to clone the gene conferring resistance to TYLCV, because linked markers in the 'suppression of recombination' block do not show segregation in the progeny. However, this large block introgressed from *S. habrochaites* contains at least 35 genes (Yang et al. 2014), of which it is unknown whether they have an adverse effect on plant growth, performance and yield in diverse growing conditions. Negative effects on agronomic and quality traits have been observed to be associated with introgression from the *Tm-2a*, *Sw-5* and *Ty-1* virus resistance genes (Rubio et al. 2012), probably due to linkage drag.

Here we show that recombination in the *Ty-2* region is occurring in intraspecific crosses between different *S. habrochaites* accessions. Therefore, in order to further fine-map the TYLCV resistance gene we are generating F2 progenies from a cross between resistant *S. habrochaites* accession B6013 and susceptible *S. habrochaites* accessions that show enough polymorphisms for efficient and detailed recombinant screening. In the near future the fine-mapped position of the *Ty-2* gene will show whether it is located in the inversion. If the gene is outside the inversion, it should be possible to eliminate the inversion in an introgression line carrying the *Ty-2* gene.

Materials and methods

Plant materials and DNA isolation

S. lycopersicum 'Moneymaker' (MM), *S. habrochaites* accessions LYC4, G1.1560 (=CGN15790), G1.1257 (=CGN15370), G1.1606 (=CGN24036) and G1.1290 (=CGN15391) were obtained from the Centre for Genetic Resources (CGN), Wageningen, Netherlands.

F_2 family PV95279 was obtained from an interspecific cross between *S. lycopersicum* MM and *S. habrochaites* accession G1.1257. F_2 families of intraspecific crosses between *S. habrochaites* accessions were PV960350 (G1.1257 x

G1.1560), PV960357 (G1.1560 x G1.1606), and PV970303 (G1.1290 x G1.1560). Seeds from the F_2 populations and parental accessions were sown in plastic cell trays and kept in a germination chamber for germination. The temperature of this chamber was between 25°C to 27°C with 90% relative humidity. Genomic DNA was extracted from 2-3 weeks old seedlings using a cetyltrimethyl ammonium bromide (CTAB) protocol (Fulton et al. 1995).

Sequence alignment

The *S. habrochaites* LYC4 sequence from the *Ty-2* region was extracted from the draft *de novo* assembly (The 100 Tomato Genome Sequencing Consortium 2014). Pairwise comparisons between sequences were made using the WebAct tool (Abbott et al. 2005) using default settings. Resulting alignments were visualized using the Artemis comparison tool ACT (Carver et al. 2005) with the footprint slider set at 101 (filter the regions of similarity based on the length of sequence over which the similarity occurs). The dot plot was obtained by aligning the Ty-2 regions from *S. lycopersicum* 'Heinz' and *S. habrochaites* LYC4 using MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/, Katoh and Standley 2103). Sequence analyses were performed with DNASTAR Lasergene 8 and Vector NTIAdvance 11 (Invitrogen).

Recombinant screening

CAPS markers used for fine-mapping of the *Ty-2* region (Ji et al. 2009; Yang et al. 2014) were tested for polymorphisms between different *S. habrochaites* accessions. PCR products obtained from the different accessions were sequenced, and the sequences were aligned to discover SNPs. When possible, co-dominant CAPS markers were developed to distinguish the different parental alleles. Primer sequences are presented in Table S1. PCRs were performed in 96-wells plates. PCR products were digested with restriction enzymes from Thermo Scientific and New England Biolabs.

Construction and screening of BAC library

The cultivated tomato line 12g-60, homozygous for the smallest introgression containing the *Ty-2* resistance gene from *Solanum habrochaites* B6013, was selected for the construction of a bacterial artificial chromosome (BAC) library. *Hin*dIII fragments were cloned into vector CopyControl™ pCC1BAC™ (*Hin*dIII Cloning-Ready) (Epicentre), and transformed to *E. coli* strain TransforMax™ EPI300™ (Epicentre), according to a previously described protocol (Rouppe van der Voort et al. 1999). The BAC library consisted of 99,840 clones with an average

insert size of 100 kb, corresponding to 10 times coverage of the tomato genome. The library was stored in 260 384-well microtiter plates, and all 384 clones in one plate were mixed to form a BAC pool. The BAC pool DNA was isolated by alkaline lysis method and screened by PCR using 17 primer pairs within and flanking the *Ty-2* region (Table S2). Afterwards, individual colonies from the 384-well plates corresponding to the positive BAC pools were identified using the same markers, and DNA was isolated from the positive colonies.

DNA sequencing and analysis

BAC ends were sequenced to confirm that they originated from the *Ty-2* region. Complete sequences of the selected BAC clones (16-100 kb) were obtained by constructing a library of subclones (1-3 kb). Both ends of the subclones were sequenced using the ABI 3730xl platform and then assembled (BGI, Beijing, China). Putative genes in the BAC sequence were predicted with the online Softberry program FGENESH (Solovyev et al. 2006). Results were compared with the 'Heinz' 1706 genome annotations derived from the International Tomato Annotation Group (ITAG2.3 version). Primers used to analyse the putative inversion breakpoints were bpTyF1 (5'-AAACTCACACCGCTCCGTTGTC-3'), bpTyR1 (5'- CCTCTTCCGATCTTTGGGTACA-3') and bpTyR2 (5'TGTTGGCATGTGACTTATAGGTA-3').

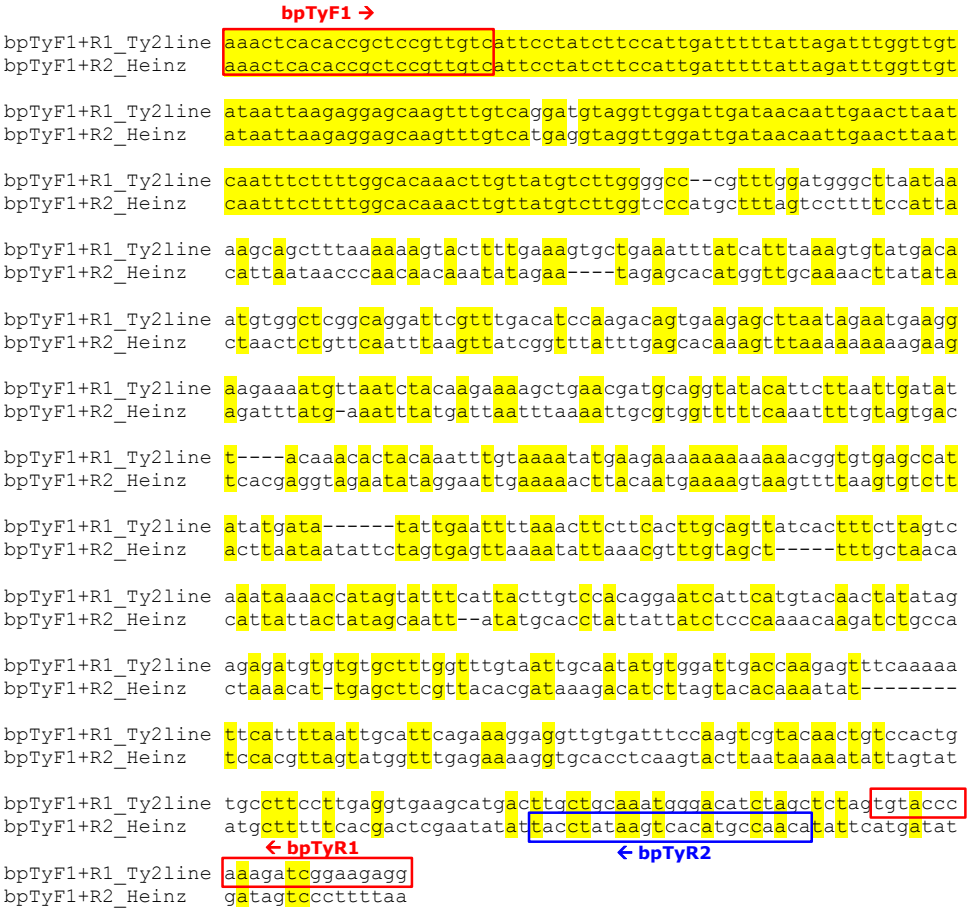
Acknowledgements

We thank Marc Hendriks, Anindita Chakraborty, Sayantani Chatterjee and Yucheng Huang for their assistance in marker analysis. This project was partially funded by the Netherlands Top Consortium for Knowledge and Innovation (TKI project TU-370), the National Natural Science Foundation of China (31372070), The Major Project of Chinese National Programs for Fundamental Research and Development (2011CB100600), and an internal grant of Wageningen UR Plant Breeding.

Supplementary files

Figure S1. Comparison of sequences in the breakpoint regions of the *Ty-2* inversion (a) Sequence alignment of the PCR products obtained with primers bpTyF1 and bpTyR1 in the *Ty-2* line, and with primers bpTyF1 and bpTyR2 in *S. lycopersicum* 'Heinz'. Identical nucleotides are highlighted in yellow. The sequence in cultivar Moneymaker (MM) is identical to the 'Heinz' sequence. (b) Graphical representations of the upper and lower breakpoint regions in the *S. lycopersicum* 'Heinz' genome, and of the lower breakpoint region in the BAC sequence from the *Ty-2* line. (c) PCR products obtained with primers bpTyF1 and bpTyR1 in the *Ty-2* line and with primers bpTyF1 and bpTyR2 in *S. lycopersicum* Moneymaker (MM).

A



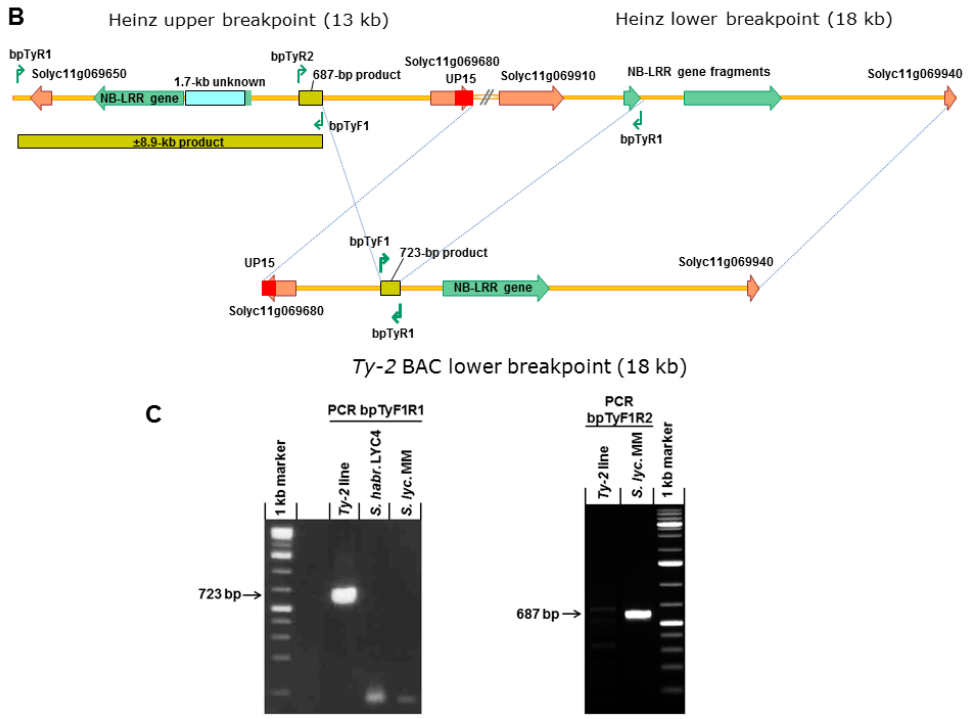


Table S1. Primers used for recombinant screening.

Marker	Physical position on SL2.40ch11 (Mbp)	Restriction enzyme	Primers
C2_At2g28250*	51.307	–	F-AGACTTCATCATCGTCATGTGGTTCCG R-TTTGGAGGTGCTTTGCCATACCAAG
C2_At1g07960	51.388	*RsaI*	F-AAAGCCATTGTTACCGTCTCCGTG R-AGCCATAAGTGGTGTGGAGGACTT
UF_07960F2	51.388	*BanI*	F-CGTGCCACCCCTTCATAATA R-CCCTTGCGAGGAAAATACAG
cLEN-11-F24	51.549	*RsaI*	F-TTATGGACAGCATGGTCCTCGGAA R-GAAGTCTGGGAGCGATAGTAGTCT
M1	51.645	*SsiI*	F-CGCTCGGGCAAATAGTTCGTAATGG R-TTCATGGTCTAGAAATGTCCCCTGT
51663_MH	51.663	*Hin6I*	F-CCCTCTTGCTTAGTGGGTGA R-ACGCTCCAAATCAGAGGTTG
C2_At4g32930	51.688	*SspI*	F-TCCTCTTCCTATTGGCAAGGGC R-TGGACACTCCCCCTTTTCATCATAC

* SNP marker; PCR product sequenced

Table S2. Primers used for BAC clone screening.

Marker	Physical position on SL2.40ch11 (Mbp)	Primers
UP8	51.345	F-GCGCTGCTAGACATTTCGAT R-CTGAAGTTGCTTGAATGCTCA
UP15	51.382	F-TCTCAAAGCGTTGATCGTTG R-GCTTGCTCTTGTTGGTCTCC
C2_At1g07960	51.388	F-AAAGCCATTGTTACCGTCTCCGTG R-AGCCATAAGTGGTGTGGAGGACTT
D1-3	51.399	F-GGGGTGTGGTTCTCTTTGCGTT R-GCCGGTACTTGCGAGCTTCTTC
D1-2	51.414	F-TGTCTGTTGTCTCTGACCCGTA R-ATTCCACACTTCTCATCCCTCC
P1-16	51.427	F-CACACATATCCTCTATCCTATTAGCTG R-CGGAGCTGAATTGTATAAACACG
D8-g40145	51.441	F-CCCCTATTGTTTTCTCTGTT R-CCCATGTCCTATAATTTGTC
D7-g60068	51.461	F-TTGGTGGTGTGTTCTGTTTA R-CGTTAGGTGGAGTAGGTGCT
D2-3	51.490	F-ATAACTGCATGGGAAGACCG R-CTCCGTAAGCAACCGAAGAC
D3-2	51.503	F-TCAAAGGACGAGATACAATC R-AATCAACAAAGGCTTAACAG
D4-3	51.549	F-TACTTGACCCTGCTGTTATT R-GTCTGGGAGCGATAGTAGTC
D5-2	51.586	F-ATCACTTCCTTCACCCGTAA R-CTCCGATTTCAACTCCATTT
C2_At3g52090	51.606	F-AGGGATACGAAGATCATGAATGCAGC R-ACTCTTCAGATGATCAAGTTCCTTGTC
S1-2	51.615	F-AAATGTAGTTGATAGAAGGG R-ATGCTAAAAGGTAAGGAGGT
S1-3	51.633	F-AAGGAGAAAAACGGAAGAGC R-CTACAATAGCCACAGGGTCA
S1-4	51.645	F-GAAAAGGAAGTGTTGGAACA R-CAATCTCATATAAATGGGGA
M1	51.645	F-CGCTCGGGCAAATAGTTCGTAATGG R-TTCATGGTCTAGAAATGTCCCCTGT

Chapter 5

Mapping in the era of sequencing: high density genotyping and its application for mapping TYLCV resistance in *Solanum pimpinellifolium*

Myluska Caro[1,3,§], Marcela Víquez-Zamora[1,2,3,§], Richard Finkers[1,2], Yury Tikunov[1,2], Arnaud Bovy[1,2], Richard GF Visser[1,2], Yuling Bai[1], Sjaak van Heusden[1,2,*]

[1] Wageningen UR Plant Breeding, Wageningen University & Research Centre, P.O. Box 386, Wageningen 6700 AJ, the Netherlands
[2] Centre for Biosystems Genomics, P.O. Box 98, Wageningen 6700 AB, the Netherlands
[3] Graduate School Experimental Plant Sciences, Wageningen 6708 PB, the Netherlands

[§] Myluska Caro and Marcela Víquez-Zamora contributed equally to this work.

This chapter has been published in BMC Genomics 2014, 15:1152

Abstract

A RIL population between *Solanum lycopersicum* cv. Moneymaker and *S. pimpinellifolium* G1.1554 was genotyped with a custom made SNP array. Additionally, a subset of the lines was genotyped by sequencing (GBS). A total of 1974 polymorphic SNPs were selected to develop a linkage map of 715 unique genetic loci. We generated plots for visualizing the recombination patterns of the population relating physical and genetic positions along the genome.
This linkage map was used to identify two QTLs for TYLCV resistance which contained favourable alleles derived from *S. pimpinellifolium*. Further GBS was used to saturate regions of interest, and the mapping resolution of the two QTLs was improved. The analysis showed highest significance on Chromosome 11 close to the region of 51.3 Mb (*qTy-p11*) and another on Chromosome 3 near 46.5 Mb (*qTy-p3*). Furthermore, we explored the population using untargeted metabolic profiling, and the most significant differences between susceptible and resistant plants were mainly associated with sucrose and flavonoid glycosides. The SNP information obtained from an array allowed a first QTL screening of our RIL population. With additional SNP data of a RILs subset, obtained through GBS, we were able to perform an *in silico* mapping improvement to further confirm regions associated with our trait of interest. With the combination of different ~omics platforms we provide valuable insight into the genetics of *S. pimpinellifolium*-derived TYLCV resistance.

Keywords: flavonoids, genotype by sequencing (GBS), hexose, *in silico*, TYLCV, *S. pimpinellifolium,* SNPs.

Introduction

Solanum pimpinellifolium is a source for introgression breeding in tomato *(S. lycopersicum)*. This species is one of the closest wild relatives of *S. lycopersicum*, and it is present in the pedigree lineage of some commercial cultivars such as the sequenced 'Heinz 1706' (The tomato genome consortium 2012). Linkage maps from crosses between *S. lycopersicum* and *S. pimpinellifolium* were generated by various researchers (Grandillo and Tanksley 1996, Chen and Foolad 1999, Lippman and Tanksley 2001, Doganlar et al. 2002, Sharma et al. 2008, Ashrafi et al. 2009, Sim et al. 2012). Their work represents a small piece of the successful use of genome-wide linkage analyses to map underlying genetic factors of traits between the two species.

Recombinant inbred lines (RILs) derived from inter-specific crosses consist of individuals with parental mosaics and are an efficient resource for mapping quantitative trait loci (QTL) (Broman 2005). Genotyping with molecular markers allows the visualization of recombination patterns which is crucial for the elucidation of loci associated with segregating traits (Paran et al. 1995, Mézard 2006). This has become more efficient due to the availability of vast numbers of markers such as single nucleotide polymorphisms (SNPs). In tomato, the availability of high throughput SNP arrays allows massive parallel whole-genome screening of genotypes (Sim et al. 2012, Víquez-Zamora et al. 2013).

Nowadays, next generation sequencing technologies are offering new ways to increase genotyping throughput by several orders of magnitude (Huang et al. 2009). Even more, it is possible to combine different genotyping platforms to increase the power of the analyses. Furthermore, due to published complete tomato genomes (The tomato genome consortium 2012), next generation re-sequencing approaches can be applied in related germplasm (Causse et al. 2013). Studies on evolutionary and domestication, as well as the genetic basis underlying important traits can be benefited from these genomic tools (Aflitos et al. 2014).

TYLCV is the causal agent of an aggressive tomato disease that can result in production losses up to one hundred percent, and its rapid spread worldwide is threatening the production of tomatoes. Development of TYLCV resistant tomato cultivars is an important strategy to avoid the damage caused by TYLCV. However, no TYLCV resistance has been identified in the cultivated tomato germplasm, except for the resistance allele of *ty-5* which is possibly originated from a mutation in the cultivated tomato (Anbinder et al. 2009). Breeding for resistance to TYLCV has been focused on the introgression of tolerance or resistance genes from tomato wild relatives such as *S. pimpinellifolium, S. chilense*, *S. habrochaites* and *S. peruvianum* (Pico et al. 2001, Verlaan et al. 2013). Several *S. pimpinellifolium* accessions are known to confer resistance to the virus (Banerjee

and Kalloo 1987, Kasrawi et al. 1988, Chagué et al. 1997, Pico et al. 2000, Pilowski and Cohen 2000, Pérez de Castro et al. 2007), but attempts to map the causal factor in this species were not very successful. Thus, *S. pimpinellifolium*-derived TYLCV resistance is currently not well-exploited in tomato breeding programs (Ji et al. 2007). In our study we genotyped a RIL population between *S. lycopersicum* cv. Moneymaker and *S. pimpinellifolium* G1.1554 with a custom made SNP array (Víquez-Zamora et al. 2013), and a subset of 60 lines was also genotyped by sequencing using Illumina HiSeq 2000 (150 Tomato Genome ReSequencing project; www.tomatogenome.net). Furthermore, we explored the population with an untargeted metabolic profiling and compared resistant vs. susceptible lines in order to get more insights on compounds that might play a role in the resistance. Our study shows how we can combine different ~omics approaches to identify genetic loci underlying resistance to Tomato Yellow Leaf Curl Virus (TYLCV) in *S. pimpinellifolium* using a RIL population. .

Results

Linkage map and genome-wide visualizations

A custom made SNP Array was assembled from polymorphisms mainly found between two cherry and two round tomatoes (Víquez-Zamora et al. 2013). This array was used to genotype a RIL population between *S. lycopersicum* cv. Moneymaker and *S. pimpinellifolium* G1.1554. A total of 1974 polymorphic SNPs were identified between the parents. These SNPs were used to develop a linkage map based on their segregation patterns among the 100 RILs. The resulting map included 715 loci with an average distance of 1.85 cM between loci (Figure 1). The greatest gap was approximately 40 cM on Chromosome 1 and covered the region between 76 and 83 Mb.

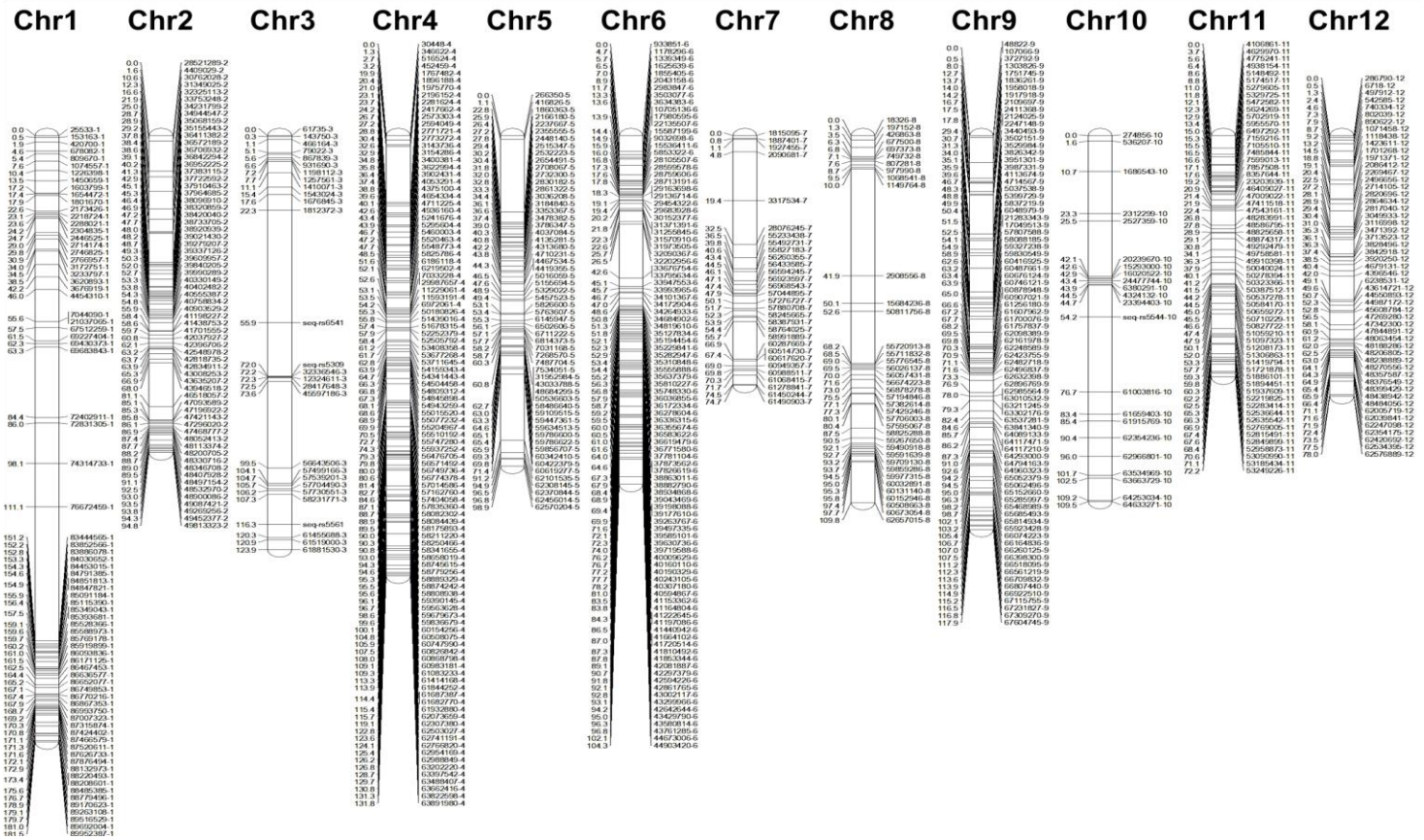
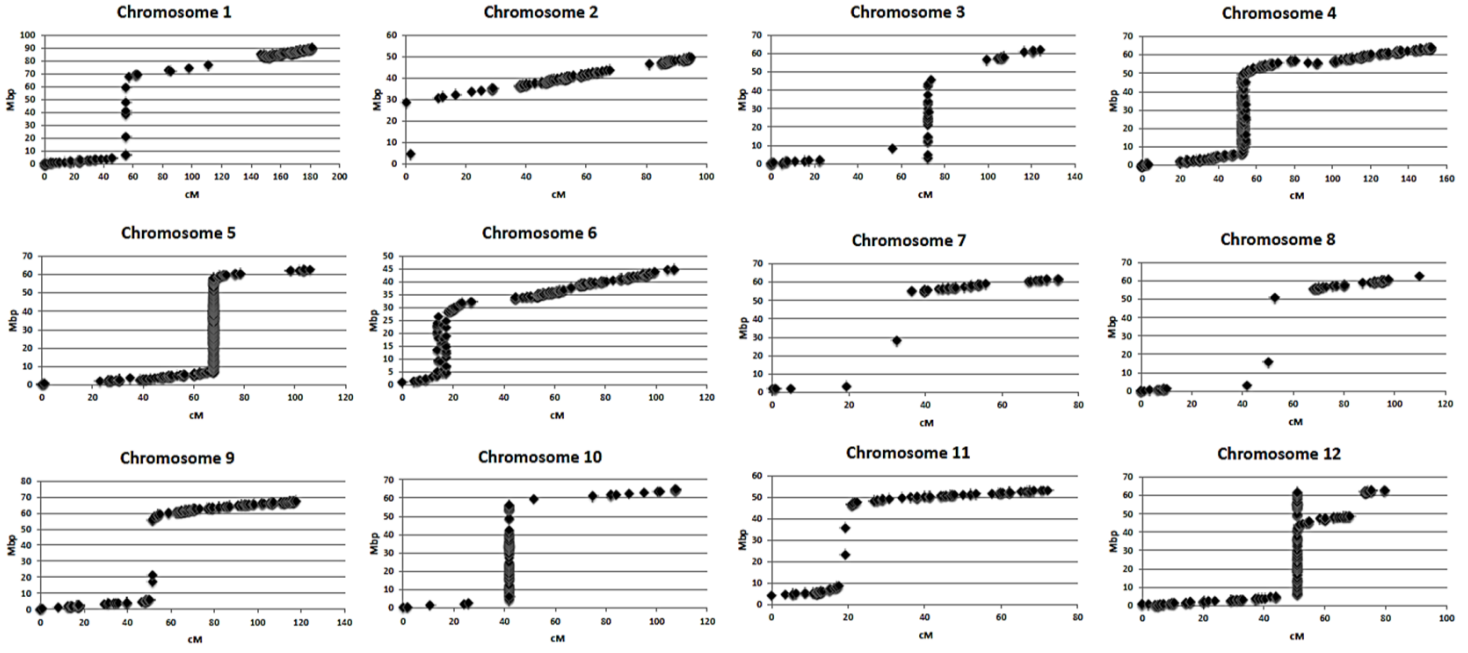


Figure 1. Linkage map of a RIL population originating from a cross between *Solanum lycopersicum* cv. Moneymaker and *Solanum pimpinellifolium* G1.1554. The map shows 715 SNPs representing single recombination positions. Markers are named according to their physical positions.

Figure 2. Scatter plots combining linkage maps (genetical positions in cM) and physical positions (Mb) from the RIL population created from a cross between *Solanum lycopersicum* cv. Moneymaker and *Solanum pimpinellifolium* G1.1554.



In order to visualize the recombination patterns along each chromosome, the physical positions of the SNP markers were determined using the published tomato genome (The tomato genome consortium 2012). For each SNP and its flanking sequence, a BLAST was performed to the genome sequence version SL2.40. Except for markers on chromosome 12, colinear orders were observed between the genetic and physical maps, as shown in scatter plots per chromosome between the linkage (cM) and physical map (Mb) (Figure 2). These scatter plots further allowed the visualization of cold- and hot-spots of recombination. When a large physical distance corresponds to only a small difference in cM, we can assume cold-spots of recombination. These cold-spots were always the heterochromatin pericentromeric regions and could be as long as 50 to 80 Mb. In contrast, hot-spots of recombination could be present if there is a large cM difference corresponding to small physical distance between markers.

The mosaic pattern of each RIL was calculated and composition of lines varied between 20% and 80% of alleles coming from each parent. In addition, we calculated the SNP allele frequency within the RIL population per marker location along each chromosome. The frequency distribution was mostly 50-50% as expected. However, we found skewness in the distribution of two regions. A preference for *S. pimpinellifolium* alleles was seen near the centromere of Chromosome 2, and a preference for *S. lycopersicum* alleles on Chromosome 9 (Figure S1).

QTLs and in silico mapping

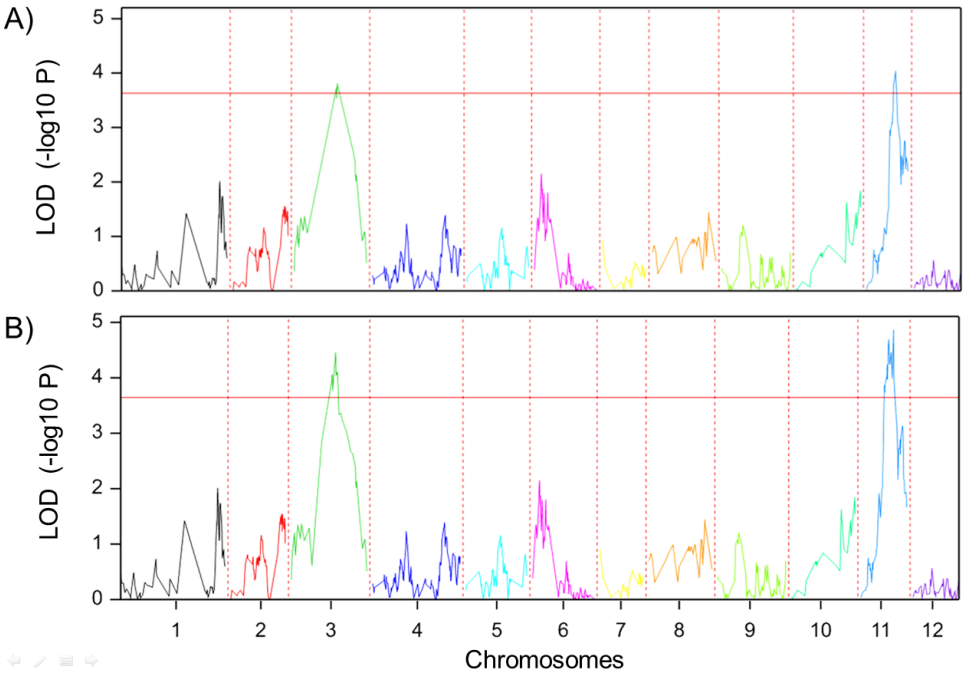
The genotypic file and the linkage map obtained above were then used to map multiple traits. One of the traits screened using our RIL population was TYLCV resistance. Eighty-one RILs were infected with TYLCV. Typical virus symptoms appeared from 30 days after inoculation (dpi); plants were scored according to their symptom development up to 45 dpi and classified as Resistant (R) or Susceptible (S). The susceptible parent 'Moneymaker', as expected, displayed severe TYLCV symptoms such as plant stunting and reduced leaf size with upwards curling and yellowing. The resistant parent, *S. pimpinellifolium* G1.1554, remained without symptoms until the end of the experiment. Five out of 81 tested RILs showed no symptoms after virus inoculation (disease score = 0), and four RILs showing very mild symptoms (disease score $\leq$ 1) were considered resistant. The remaining 72 RILs were classified as susceptible, showing clear TYLCV symptoms including the characteristic leaf curling and yellowing with disease scores ranging from 2 to 4 (Figure S2).

In order to identify the genomic regions involved in the resistance, single trait QTL analysis was performed. Two putative QTLs associated with the resistance were identified, one on Chromosome 3, hereafter referred to as *qTy-p3*, and one on Chromosome 11, hereafter referred to as *qTy-p11* (Figure 3A). For *qTy-p3,* 20 markers showed significant association with a LOD value ranging from 3.68 to 3.81, locating the QTL between 4.74 and 45.59 Mb of chromosome 3; the most significant marker for *qTy-p3* was L_45597186-3. For *qTy-p11*, 6 significantly associated markers were identified with a LOD value from 3.79 to 4.04, in a region between 50.82 and 51.20 Mb of chromosome 11. The most significant marker for this QTL was L_51208173-11 (Figure 3; Figure S3).

Sixty lines from the RIL population were re-sequenced, and the resulting genome sequences were aligned to the published tomato genome, version SL2.40 (The tomato genome consortium 2012). The fully resistant lines were included among the 60 sequenced RILs. JBrowse (Skinner et al. 2009) was used to visualize SNP variants within the RILs and allowed us to retrieve the corresponding SNP information of all aligned reads in regions of interest.

We selected 43 additional SNPs to saturate Chromosome 3 resulting in approximately one marker per 0.6 Mb. For Chromosome 11, we included two markers in the region of 7.5-8.3 Mb and 27 in the region between 49-53Mbp. As a result, the Chromosomes 3 and 11 linkage groups were improved, as was the *in silico* mapping for the subset of 60 lines.

Figure 3. QTL mapping of *qTy-p3* and *qTy-p11* (Chromosome 3 and Chromosome 11) conferring resistance to TYLCV from *S. pimpinellifolium* G1.1554. Y-axis represents values according to the interval mapping, horizontal red line delimits threshold of 3.6. A) QTL mapping in GenStat only with the SNPs obtained from the SNP array. B) QTL mapping after the inclusion of more SNP information obtained from sequences in chromosomes 3 and 11.



The outcome of the QTL analysis with the enriched genotypic data and improved genetic map is depicted in Figure 3B. Using this extended dataset, the analysis confirmed the QTLs *qTy-p3* and *qTy-p11*, The calculated threshold was very similar to the previous calculated threshold (3.64). For *qTy-p3* the LOD values ranged from 3.7 to 4.5, comprising a region with 53 significantly linked markers. The most significantly linked marker position for *qTy-p3* was then refined from 45597186 bp in the first QTL mapping to 46454095 bp and 46520535bp (both LOD of 4.46) in the improved version. For *qTy-p11* the LOD values for the 26 significantly linked markers (in the improved map) ranged from 3.86 to 4.86, and the most significant marker position was refined from 51208173 bp to 51347236 bp and 51373277 bp (both LOD of 4.86). Together, both QTLs explained almost 28% of the phenotypic effect (13.46 for *qTy-p3* and 14.18 for *qTy-p11*).

A QTL analysis using cofactors (MQM) was performed. When the most significant markers of Chromosome 3 were used as cofactors, the LOD values of *qTy-p11* decreased but were still significant. However when the most significant markers of Chromosome 11 were used as cofactors, the values of *qTy-p3* decreased to non-significant levels. Therefore, the greater impact of *qTy-p11*for the resistance was confirmed. Although all resistant RILs were homozygous for the *S.*

pimpinellifolium allele at both QTLs, 14 RILs had disease scores of 2-4 (susceptible). Thus both QTLs with the favourable alleles are necessary for resistance, but their presence did not necessarily result in resistant plants.

Identification of candidate genes

In order to identify candidate genes for TYLCV resistance, we re-explored the QTL regions using the physical positions of the SNP markers flanking the QTLs. For Chromosome 11, we targeted the region between 50.2 and 51.4 Mb. For *qTy-p11*, a total of 124 predicted genes were identified using Marker2sequence (Chibon et al. 2012) based on the tomato genome sequence (Sol Genomics Network, SGN). Four putative disease-resistance proteins were predicted in the *qTy-p11* region, three of them clustering in the region from position 51347236 to 51373277. Furthermore, approximately 74.9 kb of *qTy-p11* overlaps with the region reported to contain the *Ty-2* resistance allele from *S. habrochaites* accession B6013 (Yan et al. 2014).

The *qTy-p3* QTL region is physically large, from 2.48 to 47.44 Mb (45 Mb), including the centromeric region. This QTL region harbours more than six hundred annotated genes. In the vicinity of position 46454095 bp (the marker with the highest LOD score) there are genes related to sugars (e.g. high-affinity sugar transporters) and flavonoids (e.g. flavanone 3-hydroxylase-like protein).

RIL population metabolic profiling

Using the RIL population (not TYLCV infected), we performed untargeted metabolic profiling on leaf material. Primary metabolites were evaluated using GC-TOF-MS. Few differences were observed between parents and individuals of the population showing a similarity in the primary metabolism. However, the LC-TOF-MS and the SPME-GC-MS platforms uncovered more differences and revealed several QTLs for secondary metabolites and volatiles. More than 200 QTLs were found with putatively identified compounds; an mQTL for sucrose was mapped near *qTy-p11,* and several mQTLs for flavonoid glycosides were present near the region of *qTy-p3* (Table S1).

Furthermore, since there were TYLCV-susceptible and resistant lines with both QTLs having the homozygous *S. pimpinellifolium* alleles, we performed a T-Test with all metabolic data in order to find metabolites that were significantly different between the two groups of RILs. Five compounds showed significant differences (p-value lower than 0.05) and had higher accumulations in the resistant plants. Three of them were putatively identified as glycosylated forms of kaempferol (LCS146),

laricitrin (LCS149) and quercetin (LCS151) having a 4.3, 3.8 and 2.8-fold change, respectively. The other two compounds were acetoxytomatine (C724) and sucrose (C121) with 1.6 and 1.5-fold difference, respectively.

Discussion

High-throughput genetic mapping

The custom made SNP array was designed to distinguish different *S. lycopersicum* cultivars, nevertheless a vast amount of polymorphisms were detected between *S. pimpinellifolium* and *S. lycopersicum* cv. Moneymaker making it possible to construct a high density genetic linkage map. In general, positions on the genetic linkage map were consistent with the physical positions on the tomato genome showing the accuracy and robustness of the map and the quality of the tomato sequence.

High and low recombination rates were consistent with the known distribution of euchromatic and heterochromatic regions, as shown by Sim *et al* (2012). Chromosomes 1, 3, 4, 5 and 10 had large regions without recombination including the centromeres. Centromeric patterns were also observed for chromosomes 6, 7, 8, 9 and 11, but there were some possible distortions that could profit from more markers in the region. Still, the distortions of Chromosome 6 might be influenced by the distinct heterochromatin distribution that follows an alternating pattern (Iovene et al. 2008). Chromosome 12 also showed a non-recombining centromeric pattern, but this is a clear representation of the likely scaffold misalignment reported previously (Víquez-Zamora et al. 2013). Strong clustering of markers on the genetic map but with a clear physical distance between these markers shows a suppression of recombination in these areas (Figure 2).

The allele frequencies showed a preference for the *S. pimpinellifolium* alleles near the centromere on Chromosome 2. This part of the chromosome is linked to rDNA genes. Therefore, there could be a preference for *S. pimpinellifolium* rDNA. A preference was also found for the 'Moneymaker' alleles on Chromosome 9 which might be related to deleterious effects of carrying the *S. pimpinellifolium* alleles in this region or to structural DNA differences. Species in the same genus can have DNA configuration differences generating structural changes in the rearrangement of chromosomes after a cross (Mézard, 2006). Differences in local recombination frequencies could be related to the pairing of homologous chromosomes, DNA sequence similarity or divergence, including the presence or absence of genes

involved in the recombination process, chromatin conformation or to differences in timing during meiosis (Tam et al. 2011).

Actual research is enriched by the combination of different software packages. The combination of JBrowse (Skinner et al. 2009), loaded with gene models from the Sol Genomics Network (http://solgenomics.net/), with previous information of possible genes of interest obtained from Marker2sequence (Chibon et al. 2012) allowed an efficient targeted *in silico* mapping.

TYLCV resistance mapping and ~omics platforms combination

The sequenced subset of 60 lines created suitable tools for mapping regions of interest. We enriched regions on Chromosome 3 and Chromosome 11 that were associated with TYLCV resistance, and the *in silico* approach proved to be successful in increasing the power of QTL detection. After the addition of more SNPs coming from the known sequences, we confirmed that *qTy-p3* and *qTy-p11* were not artefacts but had real effects. This allowed us to target the location of the QTL region for *qTy-p11* and it showed the most significant region for *qTy-p3* (Figure 3), even though a large region of Chromosome 3, including the centromere, looks to have an essential impact on the expression of the resistance.

The effect of both QTLs together explained only 28% of the phenotypic effect on the resistance of our RIL population, suggesting additional genetic factors playing a role on the resistance which might have been undetected in our analysis. The accuracy of QTL localization using RILs depends on population size, where a genome-wide coverage of the parents should be present in the mapping population (Keurentjes et al. 2007). The fact that both *qTy-p3* and *qTy-p11* were needed for resistance but their presence does not necessarily lead to resistant plants also suggests the possible interaction of extra factors. TYLCV resistance derived from a number of *S. pimpinellifolium* accessions (e.g. LA121, LA373, UPV16991) has been previously suggested to be quantitatively inherited and to show variable gene penetrance (Pérez de Castro et al. 2007). Further genotyping, targeting the regions of low marker coverage, is being assessed in order to detect the presence of one or more additional QTLs, or potential modifier genes. These interactions might be associated with the secondary metabolism of the plants.

A number of TYLCV resistance loci have been reported from different wild *Solanum* species, including *S. chilense*, *S. habrochaites* and *S. peruvianum* (Ji et al. 2007). Recently, the *Ty-1* gene from *S. chilense* LA1969 has been cloned and is a

representative for a novel class of resistance genes, an RNA-dependent RNA polymerase of the RDRγ class (Verlaan et al. 2013, Butterbach et al. 2014). TYLCV resistance in *S. chilense* accessions LA1932 and LA2779, *S. habrochaites* accession B6013 and TY172, a tomato line derived from different accessions of *S. peruvianum* have been mapped to Chromosomes 3 and 10 [*Ty-4* (Ji et al. 2007) and *Ty-6* (Hutton, 2013)], Chromosome 11 [*Ty-2* (Yang et al. 2014)] and Chromosome 4 [*ty-5* (Anbinder et al. 2009)], respectively.

Several accessions from *S. pimpinellifolium* have been screened and identified to confer resistance to TYLCV (Banerjee and Kalloo 1987, Kasrawi et al. 1988, Pico et al. 2000, Pilowsky and Cohen 2000, Ji et al. 2007, Pérez de Castro et al. 2007). However, the genetics of the trait are complex and only one report on mapping resistance originating from *S. pimpinellifolium* (accession 'Hirsute INRA') has been reported using RAPD markers (Chagué et al. 1997). This resistance was mapped to Chromosome 6, close to the *Ty-1* gene. The QTLs identified in the present study represent newly mapped loci conferring resistance derived from *S. pimpinellifolium* G1.1554 and provide a starting point for assessing putative candidate genes in the identified regions. A cluster of disease resistance-like proteins is present near *qTy-p11* (based on the cultivated tomato genome sequence). Furthermore, this region on Chromosome 11 overlaps with 75 kb of the upper part of the mapped region of *Ty-2*, a TYLCV resistance allele derived from *S. habrochaites* accession B6013 (Yang et al. 2014). Although *Ty-2* has not yet been cloned, annotated genes in this common region (e.g. elongation factor 1-alpha) might provide further insights for assessing candidate genes for TYLCV resistance derived from these wild tomato species, and/or additional genes involved in the resistance pathway. Plant defense mechanisms are the result of complex gene networks which trigger or mediate the signaling pathways leading to resistance. Besides the reported *Ty*-loci, genes playing a role in these networks have been identified from their differential expression in resistant vs. susceptible genotypes and induced by TYLCV infection, e.g. *Permease I*-like protein and the hexose transporter *LeHT1* (Eybishtz et al. 2009, Eybishtz et al. 2010). Silencing these genes through Virus-induced gene silencing (VIGS) in a resistant genotype led to the collapse of the resistance, demonstrating the role and importance of these genes in the defense network of the plant.

In general, the presence of compounds such as amino acids and organic acids was very similar between the two species. Differences are more pronounced in the secondary metabolism. Our metabolic data show that the compounds present at higher amounts in the resistant plants are mainly flavonoid glycosides (Table 1).

Flavonoids are phenolic compounds known to be involved in resistance to diverse stress conditions, including plant viruses (Bol et al. 1990). For instance quercetin, one of the metabolites detected at higher levels in the resistant lines is a flavonoid known to inhibit HSP70 (Heat-shock protein 70) transcription in animal and plant cells. In *N. benthamiana,* Tomato yellow leaf curl Sardinia virus (TYLCSV) had a delayed infection speed after silencing a member of the HSP70 family, showing that high levels of this protein are required for infection of the virus (Czosnek et al. 2013). Inhibition of HSP70 expression by quercetin resulted in decreased amounts of nuclear TYLCV coat protein in tomato, demonstrating the potential involvement of this flavonoid in the virus resistance pathway (Gorovits et al. 2013). Furthermore, an additional QTL analysis suggests that glycosides of the flavonoid kaempferol co-localise with the TYLCV resistance QTL on Chromosome 3 and that sucrose could be related to the QTL on Chromosome 11 (Table S1). Kaempferol is known for its antibacterial properties. Besides, we observed the presence of this compound and other flavonoids attached to hexoses in the resistant RILs; transporters of hexoses have been reported to play crucial roles in disease resistance (Eybishtz et al. 2010, Sade et al. 2013). Some of these compounds likely linked to the resistance also showed an mQTL on chromosome 1 besides the one on chromosome 3, and the mQTL of sucrose also showed significance on chromosome 7. These regions will be further targeted in a fine mapping effort following up this research.

It should be noted that the different concentrations of the compounds observed in resistant vs. susceptible lines were measured prior to TYLCV infection. Sade et al. (2014) showed that the expression of genes controlling the synthesis of these phenolic compounds is associated with TYLCV resistance. Genes in the flavonoid biosynthesis pathway of a resistant line derived from *S. habrochaites* increased their expression after TYLCV infection leading to the accumulation of flavonoids and contributing to the resistance.

In conclusion, a RIL population obtained from a cross between *S. lycopersicum* cv. Moneymaker and *S. pimpinellifolium* G1.1554 was successfully genotyped with a custom made SNP array. Furthermore, the re-sequencing of a subset of the RILs allowed the possibility of *in silico* mapping of TYLCV resistance. Two QTLs were related to the resistance, one showing the highest significance on Chromosome 11 close to the region of 51.3 Mb and the other close to 46.5Mbp on Chromosome 3. However, there might be extra loci or genetic factors playing a role that could be unravelled if the population size is increased or when advanced populations are

further explored. The resistance towards TYLCV suggests an interaction between flavonoids and hexoses favouring the trait.
We concluded that investments in sequencing can redeem the value of screenings of germplasm due to the fact that both SNPs and sequences can be targeted at the same time. Therefore, screenings can start with a defined number of retrieved SNPs per chromosome, and thereafter, regions of interest can be further targeted. However, data storage, software acquisition and qualified human resources for data analysis and interpretation of combined ~omics platforms are going to make the difference to get robust analyses.

Materials and Methods

Recombinant Inbred Lines (RILs)

From a cross between *S. lycopersicum* cv. Moneymaker and *S. pimpinellifolium* G1.1554 (CGN reference CGN 15528) a set of 100 RILs was generated through single seed descent (SSD) until the sixth generation (Voorrips et al. 2000). These RILs, which have been used for many different experiments e.g. Khan *et al.* (2012), were used in this study.

DNA extraction

Genomic DNA from young leaflets was extracted using a CTAB based protocol (Stewart and Via 1993, Kabelka et al. 2002) adjusted for high throughput isolation. Two young leaflets were ground with a Retsch 300 mm shaker (Retsch BV, Ochten, The Netherlands) using 1 ml micronic tubes (Micronic BV, Lelystad, The Netherlands). DNA pellets were washed in 76% EtOH with 10mM NH_4Ac before re-suspending the DNA in TE buffer.

Genome wide genotyping

Genome wide genotyping was done as described by Víquez-Zamora *et al.* (2013). In short, DNA samples were sent to ServiceXS (http://www.servicexs.com/), Leiden, The Netherlands. A custom made Infinium HD Ultra Assay protocol (Infinium® HD Assay: Ultra Protocol Guide, 2009) was used for hybridization onto a BeadChip. The Genotyping Module 1.9.4 of Illumina's GenomeStudio® V2011.1 software package was used to analyse the genotyping results under default settings. All samples corresponding to the RIL population and the parents were selected for a separate analysis in which manual inspection and adjustment were performed in order to discard questionable SNPs for the population and to optimize call rates. All polymorphic SNPs for the RIL population were named after their position on

the SL2.40 version (http://solgenomics.net/) of the tomato genome sequence published online (The tomato genome consortium 2012).

Genotype by sequencing (GBS)

A subset of 60 lines was selected for resequencing (lines with extreme values for TYLCV resistance were included). Whole genomic DNA was isolated from each line (see above). Shallow sequencing of 500 bp inserts was carried out using Illumina HiSeq 2000 (100 bp paired end reads) at an average coverage of 3x. Bases with Q < 20 were trimmed before read mapping with BWA (Li et al. 2009, Li and Durbin 2012) against the SL2.40 genome sequence of *S. lycopersicum* cv. Heinz with a maximum insert size of 750 bp (50% deviation), reporting at most 30 hits and removing PCR duplicates. SAMTOOLS (Li et al. 2009) was used for variant calling without skipping InDels and a minimum gap distance of 5bp. In addition, GATK (McKenna et al. 2010), was used to call variants for all 60 genotypes in one single analysis.

The JBrowse by Skinner *et al.* (2009) was used for the embedding and visualization of the SNP variants. The available gene models (ITAG 2.3) were obtained from the Sol Genomics Network (http://solgenomics.net/). Subsequently, a script was generated in order to combine the information of SNPs within the RILs. Access to the JBrowse with the information of the sequences can be obtained through: http://www.tomatogenome.net/ril_variants. Furthermore, the program Marker2sequence (Chibon et al. 2012) was used to look for genes between specific genome coordinates based on their annotation.

TYLCV screening

Virus inoculation. *Agrobacterium*-mediated inoculation was performed to infect plants with TYLCV. Plantlets at the 3-4 leaf stage (approximately 21 days after sowing) were inoculated with *A. tumefaciens* LBA4404 bearing a tandem repeat of an infectious TYLCV-IL (Israel isolate) clone. Bacterial growth was performed as previously described by Verlaan *et al.* (2011) and bacteria were injected into true leaves using syringes without needle. Plants were grown under greenhouse conditions at 23 ºC, 60% humidity and 16-h/8-h day/night cycle.

Disease test. Disease symptoms were recorded 20, 25, 35 and 45 days post inoculation. Plants were scored for symptom severity according to the scale described by Friedmann et al. (1998). A first screening of the RILs was conducted using one plant per line. Thereafter, a second screening followed for the RILs classified as resistant to confirm the phenotype where four plants per resistant line

were assessed. TYLCV disease symptoms rating was: 0 = no visible symptoms, inoculated plants show same growth and development as non-inoculated plants; 1 = very slight yellowing and minor curling of leaflet margins on apical leaf; 2 = some yellowing and minor curling of leaflet ends; 3 = a wide range of leaf yellowing, curling and cupping, with some reduction in size, yet plants continue to develop; 4 = very severe plant stunting and yellowing, and pronounced cupping and curling; plants cease to grow (Figure S2).

Metabolic profiling

The RIL population was grown in triplicate under the same greenhouse conditions. Seven weeks after sowing, fully developed leaves were detached and main veins were removed. Samples were frozen in liquid nitrogen and thereafter ground into fine powder.

Untargeted metabolic profiling of leaves was performed with three platforms: 1) Liquid chromatography (LC), using a C18-reversed phase column, coupled to a Quadrupole-time-of-flight (TOF) mass spectrometer (MS) and a photodiode array detector (PDA) to detect semi-polar compounds such as flavonoids, alkaloids, phenylpropanoids, saponins, phenolic acids and polyamines according to De Vos et al. (2007). 2) Gas chromatography (GC) coupled to electron impact time of flight (TOF)-MS for detection of primary metabolites according to Lisec et al. (2006). 3) Solid phase microextraction (SPME)-GC-MS for the analysis of volatiles according to Tikunov et al. (2005).

Metabolomics data processing

Metabolites were quantified and identified according to Tikunov et al. (2010). Each dataset was processed using MetAlign (www.metalign.nl) for baseline correction, noise estimation, and ion-wise mass spectral alignment of the corresponding chromatograms. MSClust software was used to extract compounds mass spectra and for data reduction (Tikunov et al. 2012).

The putative identification of metabolites was based upon their spectra, retention time, molecular weight and fragmentation patterns. For LC-MS data, compound characteristics were analysed and compared using the Dictionary of Natural Products (http://dnp.chemnetbase.com) and in-house tomato metabolite databases. GC-MS data were annotated using the NIST Mass Spectral Search Program v2.0 (http://chemdata.nist.gov/mass-spc/ms-search/) by matching mass spectra extracted to the NIST mass spectra collection and the Golm Metabolome Database

(http://gmd.mpimp-golm.mpg.de/) for mass spectra matching followed by retention index comparison.

Linkage Analysis

Linkage maps were constructed using JoinMap® 4.1 (Kyazma©: http://www.kyazma.nl/, Van Ooijen 2011) with the specifications by Víquez-Zamora et al. (2013) using the Haldane's mapping function. Genetic linkage groups were compared to the physical maps based on the tomato genome version SL2.40 using MapChart 2.2 (Voorrips 2002). The software GenStat 16th edition was used to perform mapping of QTLs for TYLCV resistance and the MapQTL software was used to map metabolite QTLs (mQTLs). The genotypic and phenotypic information is available at:
http://www.plantbreeding.wur.nl/Publications/SNP/RILs_genotype-TYLCVphenotype.xlsx.

Identified QTLs for TYLCV resistance were named according to their chromosomal position as in Kadirvel et al. (201); *qTy-p3* and *qTy-p11* (*p* as from *S. pimpinellifolium*) for QTL on Chromosomes 3 and 11, respectively. The Marker2sequence application was used to mine regions for candidate genes (Chibon et al. 2012). Furthermore, the information of the sequences was embedded into JBrowse 1.11.1 (Skinner et al. 2009) to visualize the detected structural variants. The SL2.40 tomato genome assembly and ITAG 2.31 tomato genome annotation was loaded together with the BAM and VCF files of the 60 genotypes.

Acknowledgements

This project was carried out within the research programme of the Centre of BioSystems Genomics (CBSG) which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research. And the metabolic profiling of the leaves was carried out within a project developed in collaboration with the Netherlands Metabolomics Centre (NMC). Furthermore, we received support from the COST action 'Quality Fruit' (COST FA1106).

We would like to thank Fien Meijer-Dekens for collecting and maintaining all the tomato samples at Wageningen UR-Plant Breeding, Ben Vosman for his support with the NMC project and Maarten Verlaan for his support and guidance in the elaboration of the TYLCV resistance experiments. We also would like to thank various people within the department of Bioinformatics of Wageningen UR for their

support in the development of the SNP markers and in allocating them into the different versions of the Tomato Genome Sequence. We would like to make a special acknowledgement to the department of Bioscience at Wageningen UR, specially to Ric de Vos, Henriette van Eekelen, Harry Jonker, Roland Mumm and Robert Hall for all their help and support in running the samples. Additionally, MVZ acknowledges the 'Instituto Nacional de Innovación y Transferencia en Tecnología Agropecuaria' in Costa Rica, for their support through part of the doctoral project.

Supplementary files

Figure S1. Probabilities of marker frequencies calculated in GenStat. A skewness in the direction of the chromosome region from *S. pimpinellifolium* G1.1554 is observed for Chromosome 2. A skewness in the direction of the chromosome region from *S. lycopersicum* cv. Moneymaker is observed for Chromosome 9.

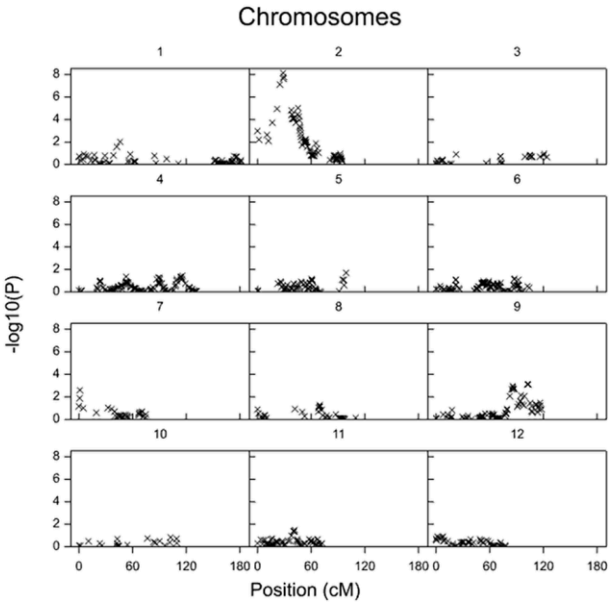


Figure S2. Disease scores of TYLCV symptom development. Plants were scored according to symptom severity: 0, no visible symptoms; 1, very slight yellowing and minor curling of leaflet margins; 2, yellowing and minor curling of leaflet ends; 3, leaf yellowing, curling and cupping; 4, severe leaf yellowing, curling and cupping, plant stunting (Friedmann et al. 1998).

Figure S3. Physical map of Chromosome 3 and Chromosome 11 between 49-53 Mb. A) Chromosomes with initial SNPs. B) Chromosomes with incorporated SNPs in green. Black arrows indicate the most significant marker related to TYLCV resistance on each case. Black frames indicate the length covering significant markers for *qTy-p3* and *qTy-p11*.

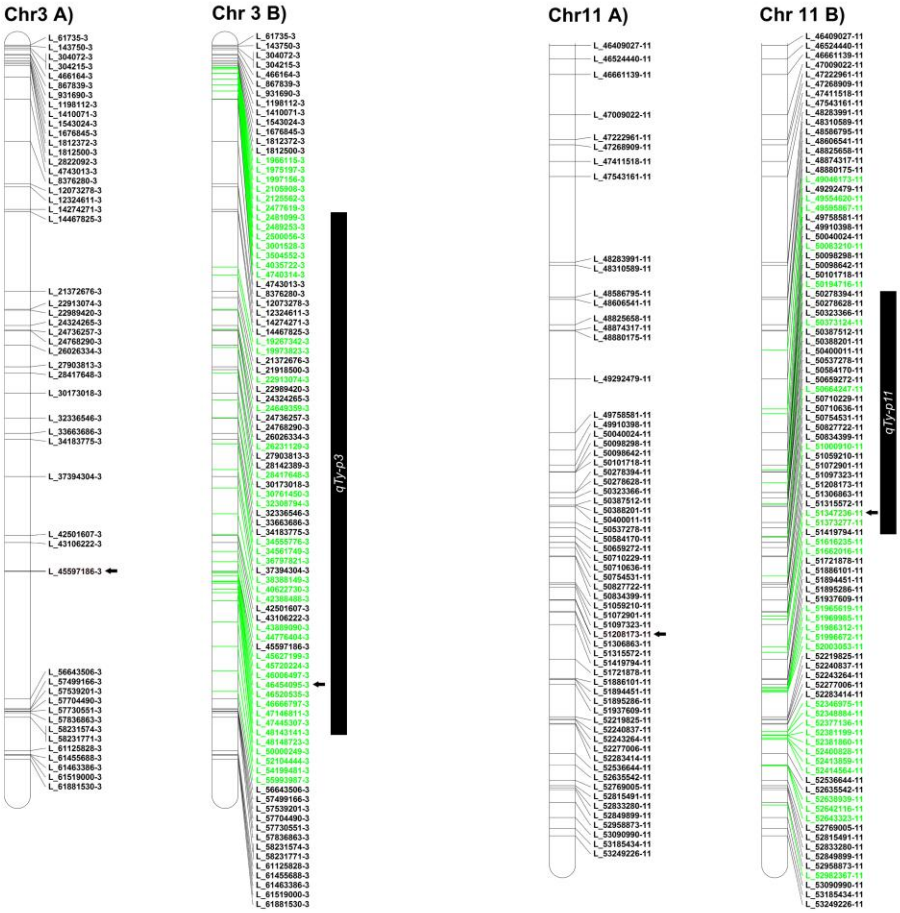


Table S1. QTLs found in non-infected leaves among the population between *S. lycopersicum* var. Moneymaker and *S. pimpinellifolium* G1.1554

Putative Compound Identification	Id. lev*	Specific negative ion, m/z**	Compound class	Platform	Trait Name	Chr	Position (cM)	Marker (closest)	LOD	% Expl.
Delphinidin deoxyhexose-feruloyl-hexose	5	947.2435	Flavonoid	LC-QTOF-MS	LCS131	Chr1	7.44	1074557-1	3.01	13.30
Cinnamicacid	5	147.0427	Phenolic Acid	LC-QTOF-MS	LCS2	Chr1	26.69	2446525-1	3.05	13.50
Caryophyllene oxyde	3	161	Sesquiterpene	GC-SPME-MS	SPME12691	Chr1	36.48	3233797-1	3.24	14.3
Humulene	2	146	Sesquiterpene	GC-SPME-MS	SPME11646	Chr1	37.48	3620893-1	4.32	18.6
Caryophyllene	2	124	Sesquiterpene	GC-SPME-MS	SPME11232	Chr1	38.48	3620893-1	4.57	19.5
O-Cymene	2	91	Terpenoid/Alkylbenzene	GC-SPME-MS	SPME3329	Chr1	38.51	3620893-1	3.25	14.3
Myrcene	2	119	Terpenoid	GC-SPME-MS	SPME3543	Chr1	40.51	3676919-1	3.61	15.7
Glycoalkaloid	5	1344.6145	Alkaloid	LC-QTOF-MS	C816	Chr1	51.96	7044090-1	3.59	15.70
Quercetin3-O-rutinoside-7-O-glucoside	2	771.1974	Flavonoid	LC-QTOF-MS	C365	Chr1	55.58	7044090-1	3.38	14.80
Kaempferol-hexose-hexose-hexose	5	771.1974	Flavonoid	LC-QTOF-MS	LCS52	Chr1	55.58	7044090-1	3.49	15.30
N296	4	693.3505	n.a.	LC-QTOF-MS	C435	Chr1	107.12	76672459-1	4.15	17.90
N740	4	493.2294	n.a.	LC-QTOF-MS	LCS171	Chr1	134.15	83444565-1	9.05	34.90
Geraniol	4	51	Terpenoid	GC-SPME-MS	SPME9044	Chr1	139.15	83444565-1	5.47	22.9
3-Methyl-2-butenal	2	56	Leucine/Isoleucine derivative	GC-SPME-MS	SPME774	Chr1	151.15	83444565-1	4.42	18.9
(E)-4-Oxo-2-hexenal	3	57	Lipid derivative alcohol	GC-SPME-MS	SPME2724	Chr1	152.75	83886078-1	6.7	27.2
(Z)-2-Hexenol	2	100	Lipid derivative alcohol	GC-SPME-MS	SPME1864	Chr1	152.75	83886078-1	18.36	58.2
1-Penten-3-one	1	51	Lipid derivative alcohol	GC-SPME-MS	SPME536	Chr1	152.75	83886078-1	7.27	29.2
2,2,6-Trimethylcyclohexanone	2	69	Cyclic molecule	GC-SPME-MS	SPME4755	Chr1	152.75	83886078-1	6.69	27.2
T-2-hexenal	2	84	Lipid derivative alcohol	GC-SPME-MS	SPME1506	Chr1	152.75	83886078-1	23.68	67.5
(4Z)-Heptenal	2	83	Lipid derivative alcohol	GC-SPME-MS	SPME2271	Chr1	154.28	84453015-1	13.11	46.3
(E)Hex-3-enol	1	70	Lipid derivative alcohol	GC-SPME-MS	SPME1557	Chr1	154.28	84453015-1	14.49	49.7
2-ethylthiophene	2	52	Heterocyclic compound	GC-SPME-MS	SPME2001	Chr1	154.28	84453015-1	14.58	50
E-2-pentenal	1	85	Lipid derivative alcohol	GC-SPME-MS	SPME826	Chr1	154.28	84453015-1	17.24	55.9

Pentanal	1	50	Lipid derivative alcohol	GC-SPME-MS	SPME602	Chr1	154.28	84453015-1	4.07	17.6
(Z)-2-pentenol	2	63	Lipid derivative alcohol	GC-SPME-MS	SPME944	Chr1	154.34	84453015-1	7.86	31.2
Heptanal	1	71	Lipid derivative alcohol	GC-SPME-MS	SPME2290	Chr1	154.34	84453015-1	7.91	31.3
Hexa-2,4-dienal	1	61	Lipid derivative alcohol	GC-SPME-MS	SPME2379	Chr1	154.34	84453015-1	11.93	43.2
Penten-3-ol	2	37	Lipid derivative alcohol	GC-SPME-MS	SPME486	Chr1	154.34	84453015-1	9.41	36
3-penten-2-one	2	69	Lipid derivative alcohol	GC-SPME-MS	SPME744	Chr1	154.60	84791385-1	3.53	15.4
Hexanoic acid, 2-oxo-, methyl ester	4	97	Carboxilic fatty acid	GC-SPME-MS	SPME4900	Chr1	154.60	84791385-1	3.87	16.8
β-Acoradien-15-ol	4	63	Sesquiterpene	GC-SPME-MS	SPME10498	Chr1	154.60	84791385-1	6.56	26.8
2-ethylfuran	2	49	Lipid derivative alcohol	GC-SPME-MS	SPME621	Chr1	154.86	84851813-1	12.59	45
3-Hexenoic acid, (E)	3	99	Fatty acid	GC-SPME-MS	SPME3127	Chr1	154.86	84851813-1	13.22	46.6
4-methylpentanol	2	41	Lipid derivative alcohol	GC-SPME-MS	SPME1419	Chr1	154.86	84851813-1	6.29	25.8
Hexanal	1	61	Lipid derivative alcohol	GC-SPME-MS	SPME1223	Chr1	154.86	84851813-1	4.16	17.9
Pentanol	1	53	Lipid derivative alcohol	GC-SPME-MS	SPME912	Chr1	154.86	84851813-1	5.03	21.3
Phenylethanal	1	90	Phenolic	GC-SPME-MS	SPME4932	Chr1	154.86	84851813-1	5.63	23.5
(Z)-2-pentenol	4	67	Lipid derivative alcohol	GC-SPME-MS	SPME853	Chr1	155.86	85091184-1	7.96	31.5
2-Heptanol	2	70	Alcohol	GC-SPME-MS	SPME2235	Chr1	156.44	85115390-1	3.52	15.4
Cis-3-nonen-1-ol	4	45	Lipid derivative alcohol	GC-SPME-MS	SPME5848	Chr1	157.51	85349043-1	5.68	23.6
Ethyl Acetate	2	62	Ester	GC-SPME-MS	SPME281	Chr1	157.51	85349043-1	4.82	20.4
Phenol	1	37	Phenolic	GC-SPME-MS	SPME3091	Chr1	157.51	85349043-1	5.32	22.3
Linalyl oxide	1	72	Terpenoid	GC-SPME-MS	SPME5527	Chr1	159.12	85528366-1	5.05	21.3
P-mentha-1,5-dien-8-ol	4	68	Terpenoid	GC-SPME-MS	SPME7758	Chr1	159.12	85528366-1	4.35	18.7
Trans linalool furanoxide	4	111	Terpenoid	GC-SPME-MS	SPME5785	Chr1	159.12	85528366-1	4.15	17.9
Linalool	2	86	Terpenoid	GC-SPME-MS	SPME5953	Chr1	160.17	85919899-1	15.39	51.9
α-terpinol	2	62	Terpenoid	GC-SPME-MS	SPME7947	Chr1	160.17	85919899-1	8.96	34.7
1-p-Menthen-9-al	2	84	Terpenoid	GC-SPME-MS	SPME8586	Chr1	160.96	86093836-1	5.48	22.9
Quercetin3-O-glucoside	1	463.0887	Flavonoid	LC-QTOF-MS	C554	Chr1	161.49	86171125-1	4.04	17.40
Kaempferol-hexose-deoxyhexose,-hexose-coumaroyl	5	901.2403	Flavonoid	LC-QTOF-MS	LCS146	Chr1	167.14	86749853-1	4.92	20.80

Laricitrin-deoxyhexose-coumaroyl	5	785.1927	Flavonoid	LC-QTOF-MS	LCS149	Chr1	168.71	86993750-1	3.64	15.90
Quercetin-hexose-deoxyhexose,-hexose, -C10H8O3 (176)	2	947.2434	Flavonoid	LC-QTOF-MS	C625	Chr1	169.23	87007323-1	5.13	21.60
Quercetin-hexose-deoxyhexose,-hexose,-coumaroyl	2	917.2349	Flavonoid	LC-QTOF-MS	C643	Chr1	169.23	87007323-1	6.93	28.00
Delphinidin-deoxyhexose-coumaroyl-hexose	5	917.2350	Flavonoid	LC-QTOF-MS	LCS136	Chr1	169.23	87007323-1	4.86	20.60
Camphene	4	92	Terpenoid	GC-SPME-MS	SPME8873	Chr1	171.60	87626733-1	7.95	31.4
Isocitricacid	1	191.0191	organic acid	LC-QTOF-MS	C240	Chr1	178.87	89170623-1	3.47	15.20
N458	4	623.1622	n.a.	LC-QTOF-MS	C596	Chr2	20.60	33753248-2	3.04	13.40
Glucose	3	157	Sugar	GC-TOF-MS	GCTOF6232	Chr2	21.92	33753248-2	3.05	13.4
Glucopyranose	4	204	Sugar	GC-TOF-MS	GCTOF9109	Chr2	21.92	33753248-2	3.65	15.7
(E)-Geranylacetone	4	109	Acyclic carotenoids	GC-SPME-MS	SPME11158	Chr2	30.20	35155443-2	3.5	15.3
Eugenol	1	117	Phenylpropanoid	GC-SPME-MS	SPME10297	Chr2	45.34	37964685-2	27.22	72.5
Eugenol-hexose-pentose	1	457.1724	Phenylpropanoid glycosilated volatile	LC-QTOF-MS	LCS132	Chr2	46.34	38096910-2	12.91	45.80
Pentadecanal	2	124	Lipid derivative	GC-SPME-MS	SPME13089	Chr2	48.23	39021430-2	3.06	13.5
Tridecanal	4	79	Lipid derivative	GC-SPME-MS	SPME13102	Chr2	48.23	39021430-2	3.03	13.4
Quercetin-dihexose-deoxyhexose	5	771.1979	Flavonoid	LC-QTOF-MS	LCS79	Chr2	77.96	46518057-2	3.68	16.00
Kaempferol-hexose-deoxyhexose,-hexose-coumaroyl	2	901.2407	Flavonoid	LC-QTOF-MS	C728	Chr2	89.51	48407928-2	3.47	15.20
Methylbutenol	2	68	Leucine/Isoleucine derivative	GC-SPME-MS	SPME254	Chr2	91.12	48497154-2	3.56	15.5
Acetoxy-tomatine+FA	1	1136.5490	Alkaloid	LC-QTOF-MS	C724	Chr2	94.81	49813323-2	3.63	15.80
4-Oxoisophorone	2	152	Cyclic ketone	GC-SPME-MS	SPME6959	Chr3	77.05	46454095-3	4.79	20.3
Methylbutenol	2	68	Leucine/Isoleucine derivative	GC-SPME-MS	SPME254	Chr3	77.05	46454095-3	3.16	13.9
N238	4	431.1921	n.a.	LC-QTOF-MS	C416	Chr3	80.31	47146811-3	4.15	17.90
1-Nonanol	4	98	Lipid derivative alcohol	GC-SPME-MS	SPME7260	Chr3	93.81	54199481-3	3.06	13.5
Geraniol	4	51	Terpenoid	GC-SPME-MS	SPME9044	Chr3	97.33	54199481-3	3.51	15.3
Isopentanol	1	54	Leucine/Isoleucine derivative	GC-SPME-MS	SPME702	Chr3	101.33	55993987-3	5.09	21.5
Laricitrin-hexose,hexose	3	665.1724	Flavonoid	LC-QTOF-MS	LCS88	Chr3	111.25	57499166-3	3.16	13.90
Quercetin-deoxyhexose-feruloyl	5	785.1929	Flavonoid	LC-QTOF-MS	LCS151	Chr3	111.25	57499166-3	4.07	17.60
N338	4	793.1805	n.a.	LC-QTOF-MS	LCS102	Chr3	113.37	57730551-3	4.78	20.30
Quercetin-3-O-glucoside	1	463.0887	Flavonoid	LC-QTOF-MS	C554	Chr3	114.43	58231574-3	3.02	13.40

Kaempferol-3-O-rutinoside	1	593.1501	Flavonoid	LC-QTOF-MS	C585	Chr3	114.43	58231574-3	3.61	15.80
Quercetin-hexose-deoxyhexose,-hexose, -C10H8O3 (176)	2	947.2434	Flavonoid	LC-QTOF-MS	C625	Chr3	114.43	58231574-3	3.02	13.40
Quercetin-hexose-deoxyhexose, -C12H12O5(236)	2	845.2148	Flavonoid	LC-QTOF-MS	C773	Chr3	114.43	58231574-3	3.51	15.40
N458	4	623.1622	n.a.	LC-QTOF-MS	C596	Chr3	114.43	58231574-3	3.79	16.50
Quercetin-hexose,-hexose (3,7-O)	5	625.1405	Flavonoid	LC-QTOF-MS	C362	Chr3	114.43	58231574-3	4.06	17.50
N429	5	773.1933	n.a.	LC-QTOF-MS	C466	Chr3	114.43	58231574-3	3.60	15.70
Kaempferol-3-O-glucoside	1	447.0937	Flavonoid	LC-QTOF-MS	C601	Chr3	115.43	58231574-3	4.86	20.60
Quercetin-hexose-deoxyhexose,-pentose	2	741.1871	Flavonoid	LC-QTOF-MS	C473	Chr3	115.43	58231574-3	6.83	27.70
Kaempferol-hexose	5	447.0937	Flavonoid	LC-QTOF-MS	LCS125	Chr3	115.43	58231574-3	4.67	19.90
Kaempferol-hexose-deoxyhexose,-pentose	2	725.1921	Flavonoid	LC-QTOF-MS	LCS101	Chr3	116.43	58231574-3	6.64	27.10
Laricitrin-deoxyhexose-coumaroyl	5	785.1927	Flavonoid	LC-QTOF-MS	LCS149	Chr3	116.43	58231574-3	4.89	20.70
Kaempferol3-O-rutinoside-7-O-glucoside	1	755.2031	Flavonoid	LC-QTOF-MS	LCS71	Chr3	117.43	58231574-3	6.01	24.80
Kaempferol3-O-rutinoside	5	593.1516	Flavonoid	LC-QTOF-MS	C406	Chr3	117.43	58231574-3	6.00	24.80
Kaempferol -hexose-deoxyhexose,-hexose-coumaroyl	5	901.2403	Flavonoid	LC-QTOF-MS	LCS146	Chr3	117.43	58231574-3	8.92	34.50
Heptahydroxyflavone,-trimethylether	3	375.0712	Flavonoid	LC-QTOF-MS	LCS53	Chr4	92.78	58658019-4	3.36	14.80
P-mentha-1,5-dien-8-ol	4	94	Terpenoid	GC-SPME-MS	SPME7486	Chr4	99.56	59836679-4	3.23	14.2
1,3,8-p-Menthatriene	4	78	Terpenoid	GC-SPME-MS	SPME5859	Chr4	99.62	59836679-4	3.94	17
Quercetin-hexose-deoxyhexose,-pentose	2	741.1871	Flavonoid	LC-QTOF-MS	C473	Chr4	125.10	62954169-4	3.14	13.80
Methyl salicylate	1	104	Phenylpropanoid	GC-SPME-MS	SPME8127	Chr4	125.44	62954169-4	3.75	16.3
Hexanoic acid, 2-oxo-, methyl ester	4	97	Carboxilic fatty acid	GC-SPME-MS	SPME4900	Chr5	38.98	3786347-5	3.27	14.4
Threitol	5	103	Sugar alcohol	GC-TOF-MS	GCTOF2469	Chr5	38.98	3786347-5	3.18	13.9
N-Acetylglutamic acid	1	174	Amino acid	GC-TOF-MS	GCTOF2777	Chr5	38.98	3786347-5	3.39	14.7
Pentanal	1	50	Lipid derivative alcohol	GC-SPME-MS	SPME602	Chr5	90.42	62101535-5	3.34	14.7
L-Glutamic acid	1	246	Amino acid	GC-TOF-MS	GCTOF3032	Chr5	91.18	62101535-5	3.26	14.2
Methylheptenone	4	77	Organic compound	GC-SPME-MS	SPME3411	Chr5	96.75	62456014-5	3.86	16.8
Benzophenone	2	181	Phenolic ketone	GC-SPME-MS	SPME12882	Chr6	16.83	28105507-6	3.08	13.6
Quercetin-hexose-deoxyhexose,-hexose,-coumaroyl	2	917.2349	Flavonoid	LC-QTOF-MS	C643	Chr6	46.97	34172904-6	3.07	13.60
Isopentanol	1	54	Leucine/Isoleucine derivative	GC-SPME-MS	SPME702	Chr6	52.86	35282947-6	3.41	14.9
Pentanol	1	53	Lipid derivative alcohol	GC-SPME-MS	SPME912	Chr6	52.86	35282947-6	3.39	14.9
4-methylpentanol	2	41	Lipid derivative alcohol	GC-SPME-MS	SPME1419	Chr6	69.43	39198088-6	3.94	17

Laricitrin-hexose,hexose	3	665.1724	Flavonoid	LC-QTOF-MS	LCS88	Chr6	89.11	42081887-6	3.59	15.70
(E)Hex-3-enol	1	70	Lipid derivative alcohol	GC-SPME-MS	SPME1557	Chr7	21.43	3317534-7	4.4	18.9
Sucrose	1	341.1074	Sugar	LC-QTOF-MS	C121	Chr7	70.35	61068415-7	5.82	24.20
Fenchene	4	52	Terpenoid	GC-SPME-MS	SPME6919	Chr8	2.34	197152-8	3.89	16.9
1,3,8-p-Menthatriene	4	78	Terpenoid	GC-SPME-MS	SPME5859	Chr8	3.34	426863-8	8.93	34.5
(E)-Ocimene	2	66	Terpenoid	GC-SPME-MS	SPME4844	Chr8	3.53	426863-8	15.37	51.8
2-Carene	2	66	Terpenoid	GC-SPME-MS	SPME3962	Chr8	3.53	426863-8	22.67	65.9
Camphene	2	107	Terpenoid	GC-SPME-MS	SPME2895	Chr8	3.53	426863-8	6.58	26.8
Limonene	1	65	Terpenoid	GC-SPME-MS	SPME4484	Chr8	3.53	426863-8	11.01	40.7
P-Cymen-9-ol	2	132	Terpenoid	GC-SPME-MS	SPME7806	Chr8	3.53	426863-8	6.7	27.2
P-Cymol	2	66	Terpenoid	GC-SPME-MS	SPME4397	Chr8	3.53	426863-8	19.42	60.2
Pinene	2	74	Terpenoid	GC-SPME-MS	SPME2621	Chr8	3.53	426863-8	5.15	21.7
Pinene	4	136	Terpenoid	GC-SPME-MS	SPME6601	Chr8	3.53	426863-8	13.24	46.7
P-mentha-1,5-dien-8-ol	4	94	Terpenoid	GC-SPME-MS	SPME7486	Chr8	3.53	426863-8	9.72	37
Verbenone	4	108	Terpenoid	GC-SPME-MS	SPME9327	Chr8	3.53	426863-8	6.78	27.5
α-Phellandrene	2	107	Terpenoid	GC-SPME-MS	SPME4012	Chr8	3.53	426863-8	25.74	70.5
α-Terpinene	2	119	Terpenoid	GC-SPME-MS	SPME4204	Chr8	3.53	426863-8	20.65	62.5
β-Phellandrene	2	123	Terpenoid	GC-SPME-MS	SPME4643	Chr8	3.53	426863-8	27.11	72.4
O-Cymene	2	50	Terpenoid/Alkylbenzene	GC-SPME-MS	SPME3280	Chr8	4.53	426863-8	5.98	24.7
O-Cymene	2	91	Terpenoid/Alkylbenzene	GC-SPME-MS	SPME3329	Chr8	4.53	426863-8	3.77	16.4
P-mentha-1,5-dien-8-ol	4	51	Terpenoid	GC-SPME-MS	SPME7353	Chr8	4.53	426863-8	6.88	27.9
L-Glutamine	5	145.0609	Amino acid	LC-QTOF-MS	C88	Chr8	70.02	56057431-8	3.69	16.10
LycoperosideHorHydroxytomatineIV +FA	5	1094.5402	Alkaloid	LC-QTOF-MS	C567	Chr8	72.64	56878278-8	3.23	14.20
Citricacid	1	191.0200	organic acid	LC-QTOF-MS	C291	Chr8	75.46	57194846-8	3.16	13.90
3-Caffeoylquinicacid (Chlorogenicacid)	5	353.0876	phenolic acid	LC-QTOF-MS	C395	Chr8	75.46	57194846-8	3.04	13.40
LycoperosideHorHydroxytomatineI	5	1048.5354	Alkaloid	LC-QTOF-MS	C765	Chr8	83.37	57595067-8	3.58	15.60
Dehydrotomatine (S)I	5	1076.5258	Alkaloid	LC-QTOF-MS	C652	Chr8	88.46	58825288-8	3.01	13.30
Benzylalcohol-hexose-pentose	2	401.1456	Phenolic glycosilated volatile	LC-QTOF-MS	C380	Chr8	94.51	59977315-8	3.22	14.20
Protocatechuicacid	5	153.0204	phenolic acid	LC-QTOF-MS	C311	Chr8	94.51	59977315-8	3.65	15.90
α-tomatin	1	1078.5415	Alkaloid	LC-QTOF-MS	C734	Chr8	97.69	60673054-8	3.32	14.60
Tomatidinedihexosedipentose +FA	1	1048.5322	Alkaloid	LC-QTOF-MS	C749	Chr8	97.69	60673054-8	3.28	14.40
Glycoalkaloid	5	1344.6145	Alkaloid	LC-QTOF-MS	C816	Chr8	97.69	60673054-8	3.32	14.60

β-Damascenone	2	190	Cyclic carotenoids	GC-SPME-MS	SPME10714	Chr9	6.52	1303826-9	4.2	18.1
(Z)-2-Hexenol	2	100	Lipid derivative alcohol	GC-SPME-MS	SPME1864	Chr9	8.02	1303826-9	4.34	18.6
2-ethylthiophene	2	52	Heterocyclic compound	GC-SPME-MS	SPME2001	Chr9	8.02	1303826-9	3.08	13.6
E-2-pentenal	1	85	Lipid derivative alcohol	GC-SPME-MS	SPME826	Chr9	8.02	1303826-9	3.69	16.1
L-Aspartic acid	1	100	Amino acid	GC-TOF-MS	GCTOF2612	Chr9	38.88	4113674-9	4.21	17.9
Phenylethanol	1	37	Aromatic alcohol	GC-SPME-MS	SPME6326	Chr9	52.47	57807588-9	6.15	25.3
Quercetin-hexose,-hexose (3,7-O)	5	625.1405	Flavonoid	LC-QTOF-MS	C362	Chr9	63.94	60746121-9	4.29	18.40
Kaempferol-hexose-deoxyhexose,-hexose-coumaroyl	2	901.2407	Flavonoid	LC-QTOF-MS	C728	Chr9	66.64	61256180-9	3.16	13.90
N152	4	443.1924	n.a.	LC-QTOF-MS	LCS41	Chr9	67.16	61607962-9	6.97	28.20
(E)-Geranylacetone	4	109	Acyclic carotenoids	GC-SPME-MS	SPME11158	Chr9	69.21	62098389-9	3.05	13.5
Laricitrin-deoxyhexose-coumaroyl	5	785.1927	Flavonoid	LC-QTOF-MS	LCS149	Chr9	70.33	62248589-9	3.16	13.90
Kaempferol-3-O-rutinoside-7-O-glucoside	1	755.2031	Flavonoid	LC-QTOF-MS	LCS71	Chr9	70.85	62423755-9	9.64	36.70
Quercetin3-O-rutinoside-7-O-glucoside	2	771.1974	Flavonoid	LC-QTOF-MS	C365	Chr9	70.85	62423755-9	12.99	46.00
Quercetin-dihexose-deoxyhexose-pentose	3	903.2413	Flavonoid	LC-QTOF-MS	LCS54	Chr9	70.85	62423755-9	12.84	45.60
Quercetin-dihexose-deoxyhexose-pentose	5	903.2408	Flavonoid	LC-QTOF-MS	C353	Chr9	70.85	62423755-9	9.99	37.80
Kaempferol3-O-rutinoside	5	593.1516	Flavonoid	LC-QTOF-MS	C406	Chr9	70.85	62423755-9	10.15	38.20
Kaempferol-hexose-hexose-hexose	5	771.1974	Flavonoid	LC-QTOF-MS	LCS52	Chr9	70.85	62423755-9	12.04	43.50
Isorhamnetin-hexose-hexose (3-O)	3	639.1605	Flavonoid	LC-QTOF-MS	LCS111	Chr9	76.87	62896769-9	3.06	13.50
Isorhamnetin-hexose,-hexose (3,7-O)	5	639.1574	Flavonoid	LC-QTOF-MS	LCS107	Chr9	76.87	62896769-9	3.53	15.40
Hexanol-pentose-hexose	3	395.1934	Lipid glycosilated volatile	LC-QTOF-MS	C572	Chr9	104.21	65923428-9	3.86	16.80
Glycoalkaloid	5	1344.6145	Alkaloid	LC-QTOF-MS	C816	Chr9	104.21	65923428-9	3.16	13.90
Hexanoic acid, 2-oxo-, methyl ester	4	97	Carboxilic fatty acid	GC-SPME-MS	SPME4900	Chr9	105.40	66074223-9	3.45	15.1
β-Acoradien-15-ol	4	63	Sesquiterpenes	GC-SPME-MS	SPME10498	Chr9	105.40	66074223-9	3.13	13.8
N427	5	773.1922	n.a.	LC-QTOF-MS	C480	Chr9	106.40	66164836-9	4.63	19.80
5-Caffeoylquinicacid	1	353.0875	phenolic acid	LC-QTOF-MS	C337	Chr10	1.00	536207-10	3.03	13.40
4-Caffeoylquinicacid	5	353.0883	Acid	LC-QTOF-MS	C361	Chr10	1.00	536207-10	8.44	33.00
Citricacid	1	191.0200	organic acid	LC-QTOF-MS	C291	Chr10	3.62	536207-10	3.72	16.20
N143	5	402.9155	n.a.	LC-QTOF-MS	C29	Chr10	18.65	2312299-10	3.09	13.70
2-methylbutanol	1	45	Leucine/Isoleucine derivative	GC-SPME-MS	SPME726	Chr10	23.33	2312299-10	3.71	16.2
Butanol	2	55	Alcohol	GC-SPME-MS	SPME429	Chr10	23.33	2312299-10	3.47	15.2

Methylbutenol	2	68	Leucine/Isoleucine derivative	GC-SPME-MS	SPME254	Chr10	23.33	2312299-10	3.66	16
Caffeicacid	5	179.0362	organic acid	LC-QTOF-MS	C295	Chr10	25.33	2527359-10	6.42	26.30
Kaempferol-hexose-deoxyhexose,-hexose-coumaroyl	2	901.2407	Flavonoid	LC-QTOF-MS	C728	Chr10	25.52	2527359-10	4.22	18.20
N71	4	337.0771	n.a.	LC-QTOF-MS	LCS20	Chr10	42.90	16020522-10	3.46	15.20
β-Damascenone	2	190	Cyclic carotenoids	GC-SPME-MS	SPME10714	Chr10	44.48	4324132-10	8.36	32.8
3-methyl-2-butenol	2	68	Leucine/Isoleucine derivative	GC-SPME-MS	SPME1019	Chr10	47.74	23394403-10	3.52	15.4
Limonene	1	65	Terpenoid	GC-SPME-MS	SPME4484	Chr10	49.74	seq-rs5544-10	3.84	16.7
Coumaroylquinicacid	1	337.0940	phenolic acid	LC-QTOF-MS	C414	Chr10	62.21	59477572-10	4.51	19.30
Coumaroylquinicacid	5	337.0936	phenolic acid	LC-QTOF-MS	C462	Chr10	63.21	59477572-10	4.56	19.40
4-Oxoisophorone	2	152	Cyclic ketone	GC-SPME-MS	SPME6959	Chr10	95.36	62966801-10	3.18	14
2-ethylhexanol	2	81	Lipid derivative alcohol	GC-SPME-MS	SPME4243	Chr11	0.00	4106861-11	9.81	37.2
2-methylbutanol	1	45	Leucine/Isoleucine derivative	GC-SPME-MS	SPME726	Chr11	0.00	4106861-11	10.82	40.2
3-methyl-2-butenol	2	68	Leucine/Isoleucine derivative	GC-SPME-MS	SPME1019	Chr11	0.00	4106861-11	3.2	14.1
Butanol	2	55	Alcohol	GC-SPME-MS	SPME429	Chr11	0.00	4106861-11	11.66	42.5
Geranial	2	137	Terpenoid	GC-SPME-MS	SPME9389	Chr11	0.00	4106861-11	3.3	14.5
Heptanol	2	68	Lipid derivative alcohol	GC-SPME-MS	SPME2968	Chr11	0.00	4106861-11	4.78	20.3
Hexanol	1	70	Lipid derivative alcohol	GC-SPME-MS	SPME1892	Chr11	0.00	4106861-11	6.12	25.2
Isopentanol	1	54	Leucine/Isoleucine derivative	GC-SPME-MS	SPME702	Chr11	0.00	4106861-11	7.54	30.1
Methylbutenol	2	68	Leucine/Isoleucine derivative	GC-SPME-MS	SPME254	Chr11	0.00	4106861-11	3.48	15.2
Pentanol	1	53	Lipid derivative alcohol	GC-SPME-MS	SPME912	Chr11	0.00	4106861-11	7.3	29.3
Benzylalcohol-hexose-pentose	2	401.1456	Phenolic glycosilated volatile	LC-QTOF-MS	C380	Chr11	0.00	4106861-11	5.60	23.30
Xylose	1	103	Sugar	GC-TOF-MS	GCTOF3183	Chr11	3.71	4629970-11	3.16	13.8
2-Heptanol	2	70	Alcohol	GC-SPME-MS	SPME2235	Chr11	8.86	5174517-11	4.23	18.2
Caryophyllene	2	124	Sesquiterpenes	GC-SPME-MS	SPME11232	Chr11	11.07	5279605-11	3.58	15.6
Caryophyllene oxyde	3	161	Sesquiterpenes	GC-SPME-MS	SPME12691	Chr11	11.07	5279605-11	3.78	16.4
Humulene	2	146	Sesquiterpenes	GC-SPME-MS	SPME11646	Chr11	11.07	5279605-11	3.11	13.7
Benzaldehyde	1	63	Phenolic/Aromatic aldehyde	GC-SPME-MS	SPME2995	Chr11	11.86	5329725-11	4.02	17.4
β-Ionone	2	145	Cyclic carotenoids	GC-SPME-MS	SPME11902	Chr11	18.82	23203939-11	3.13	13.8

_O-Feruloylquinicacid	3	367.1035	phenolic acid	LC-QTOF-MS	C449	Chr11	21.02	47009022-11	3.90	16.90
3-O-Feruloylquinicacid	5	367.1040	Acid	LC-QTOF-MS	C491	Chr11	21.02	47009022-11	3.42	15.00
Methyl salicylate	1	104	Phenylpropanoid	GC-SPME-MS	SPME8127	Chr11	46.45	50710636-11	3.74	16.3
N50	4	609.1888	n.a.	LC-QTOF-MS	LCS17	Chr11	56.66	51347236-11	3.75	16.30
Sucrose	1	341.1074	Sugar	LC-QTOF-MS	C121	Chr11	73.08	52635542-11	4.26	18.30
Protocatechuicacid	5	153.0204	phenolic acid	LC-QTOF-MS	C311	Chr11	73.08	52635542-11	3.68	16.00
1-Nonanol	4	98	Lipid derivative alcohol	GC-SPME-MS	SPME7260	Chr12	48.99	6238531-12	3.31	14.5
Octanol	2	71	Lipid-derived	GC-SPME-MS	SPME5262	Chr12	49.12	6238531-12	4.2	18.1
Quercetin3-O-rutinoside	1	609.1450	Flavonoid	LC-QTOF-MS	C512	Chr12	49.12	6238531-12	4.04	17.50
Kaempferol3-O-rutinoside	1	593.1501	Flavonoid	LC-QTOF-MS	C585	Chr12	51.70	44987172-12	3.42	15.00
Kaempferol-3-O-glucoside	1	447.0937	Flavonoid	LC-QTOF-MS	C601	Chr12	51.70	44987172-12	4.28	18.40
N429	5	773.1933	n.a.	LC-QTOF-MS	C466	Chr12	51.70	44987172-12	3.37	14.80
Kaempferol-hexose	5	447.0937	Flavonoid	LC-QTOF-MS	LCS125	Chr12	51.70	44987172-12	4.29	18.40
Petunidin-deoxyHexose-coumaroyl-hexose +H2O	5	949.2614	Flavonoid	LC-QTOF-MS	LCS89	Chr12	52.32	44987172-12	3.21	14.10
Guaiacol	1	37	Phenylpropanoid	GC-SPME-MS	SPME5805	Chr12	74.45	62420692-12	5.65	23.5

*Annotation level: 1=Identified compounds. 2=Putatively annotated compounds (e.g. without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries). 3=Putatively characterized compound classes (e.g. based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class). 4. Unknown compounds—although unidentified or unclassified these metabolites can still be differentiated and quantified based upon spectral data. 5. Unknown compounds—similar mass to a putatively characterized compound.

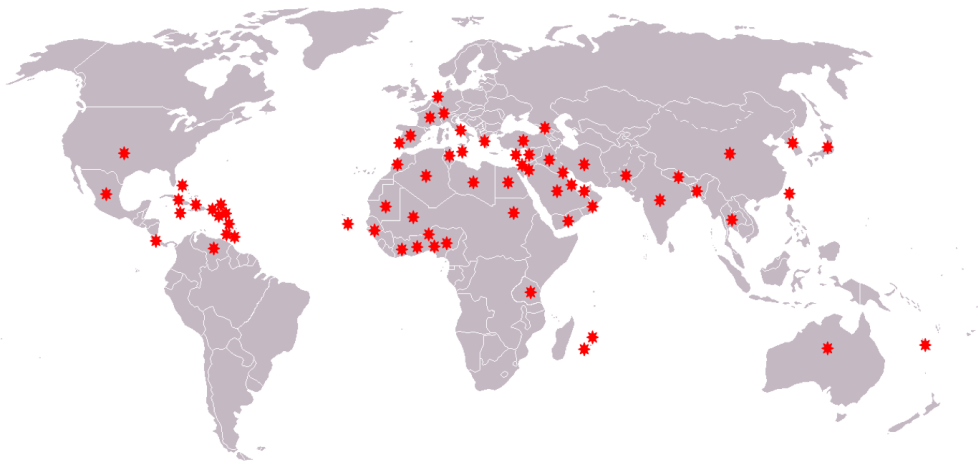
**Compounds analysed using LC-QTOF-MS platform represented by measured accurate masses of corresponding negatively charged parent molecule ions or their formic acid adducts (denoted by +FA). Volatile compounds and primary metabolites measured by SPME-GC-MS and GC-TOF-MS, respectively, represented by selected nominal negative mass ion fragments picked automatically by MSClust software.

Chapter 6

General discussion

The *Geminiviridae* family contains some of the most important and devastating viruses for horticultural crops worldwide. Over the past decade, the severity of the begomovirus TYLCV has increased to epidemic proportions and acquired global importance (Czosnek, 2007b). Up to now, more than 65 countries report (or have reported) the presence of TYLCV (CABI, 2015) (Figure 1), responsible for economical and yield losses ranging from 5 up to 100%. For tomato farmers TYLCV has become one of the major limitations for tomato production in tropical and sub-tropical regions.

Figure 1. Geographical worldwide distribution of TYLCV. Red stars indicate countries in which presence of TYLCV has been officially reported (CABI, TYLCV datasheet 2015).



As part of a sustainable strategy to cope with TYLCV, breeding efforts are focused on developing elite breeding lines containing resistance genes to protect the crop against this aggressive virus. Within this framework, this project aimed to identify, assess and characterize tomato genes and their associated mechanisms contributing to different layers of innate immunity to TYLCV.

Plant immunity against viruses

In order to successfully invade a plant, viruses (as any plant pathogens) need to overcome the plant immune system (**Chapter 1**). Host plants under viral attack, trigger a defence response that involves several layers, complementary in terms of timing (early or late infection), location (first infected leaf followed by systemic

tissue) and targeting virus-derived molecules (viral genome and/or viral protein (Nicaise, 2014) (Figure 2). A defence response is thus the final outcome of complex regulatory gene networks and a cassette of genes is required for the (pathogen-specific) resistance pathway.

Dominant resistance genes: the main virus-resistance driving force for breeding

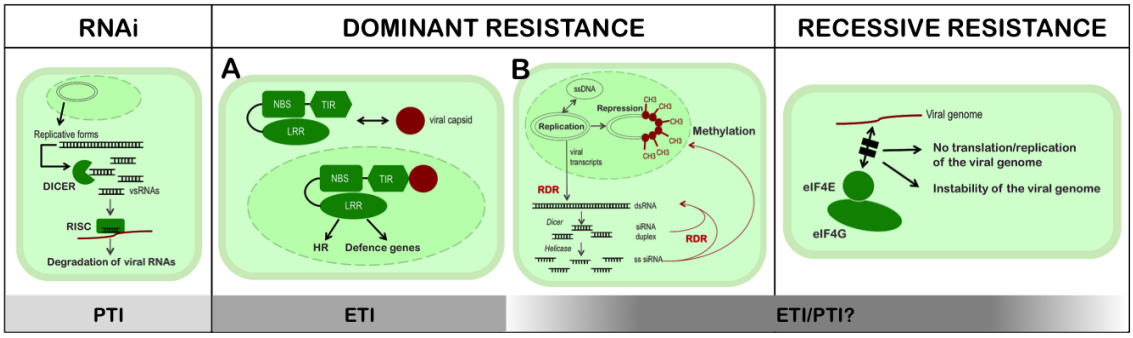
So far more than 200 dominant *R*-genes against viruses have been reported, and only 24 genes have been cloned and characterized so far from different hosts i.e. Arabidopsis, potato, tomato, pepper, field mustard, soybean, muskmelon, tobacco, kidney bean, trifoliate orange and black gram (Whitham et al. 2000, Cosson et al. 2010, de Ronde et al. 2014). From these 24 genes, 18 belong to the nucleotide binding site-leucine-rich repeat (NBS-LRR) class. Within this class, *R*-genes are subdivided based on the presence of either a N-terminal Toll-interleukin-1 receptor (TIR) domain or a N-terminal coiled-coil (CC) domain (Maule et al. 2007). *R* gene-mediated resistance relies on the specific recognition of a viral avirulence (*avr*) gene product, and in most cases is thought to involve an indirect interaction. One of the most common models to explain for *R*-gene mediated resistance is the so-called 'guard hypothesis'. In this model, interaction of the *avr* with a host target protein (guard) leads to a conformational change that is being perceived by the *R*-gene product (guardee) and triggers a downstream defence response (Maule et al. 2007). Activation of single dominant *R*-genes of the NBS-LRR type is generally associated with a hypersensitive response (HR) that involves a programmed cell death (PCD) of tissue cells at and surrounding the site of pathogen entry, in order to contain and prevent the pathogen from further spreading.

From the 18 NBS-LRR-type virus resistance genes cloned so far, only 2 confer resistance to DNA viruses, i.e. *PvVTT1* (resistance to the begomovirus *Bean dwarf mosaic virus*) and *CYR1* (resistance to the begomovirus *Mungbean yellow mosaic virus*). The other 16 NBS-LRR cytoplasmic proteins confer resistance to a diverse range of RNA viruses (de Ronde et al. 2014).

Replication of RNA viruses generally takes place in the cytoplasm, while that of DNA viruses takes place mostly in the nucleus. Molecular interactions between viral and cellular factors occur in cellular compartments (termed e.g. viroplasms or virus replication centres) in which viral replication, genome expression and particle assembly take place. Viroplasms are also intended to recruit factors required for viral replication and to control (evade from) host antiviral defence mechanisms (Schmid et al 2014). The latter often involves the association/surrounding of viroplasm with host intracellular membranes (Nicaise et al. 2014). The fact that

most of the cloned *R*-genes are effective against (cytoplasmic replicating) RNA viruses and to a lesser extent to (nuclear replicating) DNA viruses points to the question whether there is an association between the viral replication compartment and the location of candidate host protein interactors. The *R*-gene products (NBS-LRR) of *Rx1* and *N* against the RNA viruses PVX and TMV, respectively, are activated in the cytoplasmic compartment. However, a nucleocytoplasmatic distribution is required for their complete functionality (Nicaise 2014). Whether the cellular compartment of the virus-host interaction plays a role on the evolution of NBS-LRR resistance genes against RNA-viruses distinct to that of DNA viruses is unknown.

Figure 2. Main antiviral plant immunity mechanisms. The PAMP-triggered immunity (**PTI**) against viruses, RNA interference (RNAi), mediates viral RNA degradation. **Dominant resistance** is mostly conferred by typical *R*-genes of the NBS-LRR class (A. Molecular interaction between avirulence factor and specific *R*-gene product, e.g. *N* gene against *Tobacco mosaic virus*, TMV. Such resistance mechanism is regarded as **ETI**), and for atypical, non-canonical *R*-genes (B. *Ty-1* RDR, involved in enhanced transcriptional gene silencing of TYLCV). **Recessive resistance** relies on the inability of the virus to interact with host factors (*S*-genes) required for the virus cycle (e.g. eIF4E and eIF4G). As for the dominant resistance conferred by *Ty-1* and the *S*-gene mediated recessive resistance, it remains unclear whether they correspond to **ETI** or **PTI** resistance mechanisms (modified from Nicaise et al. 2014).



Atypical *R*-genes: a different resource for virus resistance

Only six non-canonical plant *R*-genes (non-NBS-LRR) conferring resistance to viruses have been reported and cloned so far (de Ronde et al. 2014). Their activated responses do not involve an HR or the defence signalling of systemic-acquired

resistance (SAR) characteristic of a typical NBS-LRR gene resistance response. The first reported non-NBS-LRR virus resistance genes were the RTM (Restricted *Tobacco etch virus* Movement) genes RTM1, RTM2 and RTM3 from *Arabidopsis thaliana*. RTM1 encodes a protein with similarity to the lectin jacalin (Chisholm et al. 2000), the RTM2 gene codes for a protein with similarity to small heat-shock proteins (HSP) (Whitham et al. 2000) and RTM3 was later found to belong to a new protein family containing a MATH domain (Cosson et al. 2010). The RTM-mediated resistance response does allow an on-going viral replication and cell-to-cell movement, but restricts the long distance movement of several potyviruses (Cosson et al. 2012). The tomato gene *Tm-1* from *S. habrochaites* inhibits the replication of the tobamovirus ToMV (*Tomato mosaic virus*) by binding virus replication proteins and inactivating the viral RNA-dependent RNA polymerase (Ishibashi et al. 2007). The JAX1 (*Jacalin-type lectin required for potexvirus resistance 1*) gene from Arabidopsis confers resistance to a broad range of potexviruses by impairing viral accumulation at the cellular level, thus acting during the initial stages on infection. It has been proposed that JAX-1 inhibits viral infection by the recognition of viral proteins, or by the inactivation of virus replication bodies. Similar to RTM1, JAX1 also encodes a lectin-like gene, but both genes confer different levels of resistance and seem to inhibit different phases of viral infection (Yamaji et al. 2012). In all the aforementioned cases, neither a programmed-cell-death response nor systemic-acquired resistance are triggered, supporting the idea that the resistance involves mechanisms distinct from those triggered by NBS-LRR genes.

Similarly, an HR response has never been reported for any TYLCV resistance. We have evaluated the resistance responses from *S. chilense*, *S. habrochaites* and *S. pimpinellifolium* (*Ty-1, Ty-2, qTy-p3, qTyp11* in **Chapters 2, 3** and **5**) and none of these cases were shown to reveal a classical *R*-gene mediated triggering of PCD. Instead, plants carrying these loci remained asymptomatic and still supported virus replication, but in much lower levels compared to susceptible plants. Low TYLCV titres have also been recorded in resistant plants containing *Ty-4, ty-5* and *Ty-6* (this thesis, unpublished data).

In addition to the presence of low virus titres, *Ty-1/Ty-3*- and *Ty-2*-containing lines revealed, although in different amounts, the presence of TYLCV-specific siRNAs, indicative for an ongoing viral infection. The siRNA profile from TYLCV infected *Ty-2* plants was comparable with that of the susceptible control but different from the *Ty-1* line, suggesting that *Ty-2* resistance likely involves a mechanism distinct from the *Ty-1*/*Ty-3* enhanced transcriptional gene silencing (TGS) (Butterbach et al. 2014). Moreover, silencing three disease resistance genes in *Ty-2* containing

tomato plants that mapped to the *Ty-2* region, one of the CC-NBS-LRR type and two homologs with similarity to the R3a resistance gene, did not compromise the resistance (**Chapter 3**), suggesting that *Ty-2* likely does not belong to the NBS gene family. Altogether these data strongly suggest that, besides the resistance conferred by the recessive gene *ty-5*, the TYLCV resistance conferred by *Ty-2, qTyp3, qTyp1*, and likely *Ty-4* and *Ty-6* is possibly determined by a novel class of resistance genes distinct from the classical *R-Avr* interaction, as already nicely exemplified by the mechanism of *Ty-1*-mediated enhanced transcriptional gene silencing.

The first atypical geminivirus resistance gene cloned, *Ty-1,* is a RNA-dependent RNA polymerase of the γ type and involves antiviral RNA silencing. It has been proposed that *Ty-1* enhances transcriptional gene silencing (TGS) by amplifying the siRNA silencing signal. We have confirmed that at least 4 *S. chilense* accessions (LA1969, LA2779, LA1932, LA1938) contain different functional alleles of the same RDR gene and that one accession can even harbour more than 1 resistance gene. In order to investigate the natural genetic variation of the *Ty-1* gene, its features have been analysed from *S. chilense* accessions and other related *Solanum* species within the tomato clade (**Chapter 2**).

Verlaan et al. (2013) previously postulated that an insertion of 4 amino acid residues in the *Ty-1* and *Ty-3* alleles was the most striking difference from their alleles of the susceptible 'Moneymaker' genotype. However, in our studies, this insertion was also found in several related *Solanum* accessions that are susceptible to the virus, ruling out the association of this polymorphism with *Ty-1*-mediated resistance. Instead, five specific SNPs for the *Ty-1/Ty-3* alleles and four SNPs unique for the *Ty-3* allele were identified. These data will be very useful for development of new *Ty-1* and/or *Ty-3* specific markers for breeding programs. Furthermore, three *Ty-1/Ty-3* specific amino acids were identified that linked to *Ty-1*-mediated resistance and were only shared by several other resistant *S. chilense*-derived lines and accessions. This indicates that the non-functionality of the *ty-1* (susceptible) allele is likely determined by the absence of these 3 amino acid that seem required for a functional RDR "resistance" protein. No clear association between resistance/susceptibility and the levels of RDR expression was found in our studies (**Chapter 2**). A similar scenario has earlier been observed for the RTM-conferred resistance to potyviruses; the functionality of the RTM alleles is not correlated with their expression levels. Instead, RTM1 alleles of susceptible hosts contain a 6- amino acid deletion at its C-terminal end and a 4-amino acid

change in the jacalin domain compared to its functional, resistance allele. Similarly, non-functional alleles of the RTM2 gene contain amino acid changes in the alpha helix, the HSP domain and the C-terminal end of the gene. Amino acid changes in the MATH and the coiled-coil (CC) domain of non-functional RTM3 alleles indicate the importance of these domains for resistance. All domains are known to be involved in protein-protein interactions, suggesting that these mutations possibly disrupt interactions needed for protein functionality and/or stability (Cosson et al. 2012).

Within the *Ty-1* gene, the DFDGD motif from the catalytic domain showed to be conserved in tomato and among its wild relatives. Phylogenetic analysis of five RDRs in *Salvia miltiorrhiza* clusters SmRDR5 together with Arabidopsis AtRDR3, AtRDR4 and AtRDR5, all belonging to the RDRγ clade to which the *Ty-1* RDR gene also belongs. Sequence analysis further revealed that four motifs (1, 7, 10 and 12) are conserved in all five SmRDRs (Shao and Lu, 2014). The remaining less-conserved motifs could be further explored to determine if the 3 *Ty-1/Ty-3* specific amino acids are located within similar motifs, and whether these are associated with possible gene-specific functions as for the TYLCV resistance. Analysis of *Ty-1* mutants changed for these 3 specific amino acids (individually or combined) will provide more insights on the possible role of each amino acid in the resistance mechanism. The molecular processes in the plant-pathogen interaction in which these polymorphisms might be involved are still to be uncovered.

S-genes: a rising and promising source of durable resistance

Single dominant *R* genes have predominantly been used in breeding for disease resistance in all crops. The so called '*S*-genes' concept, which makes use of host susceptibility factors, has lately been proposed to be a promising complementary approach to breed for resistance to different pathogens.

Being obligate intracellular parasites, TYLCV as most viruses depend on host (susceptibility) factors to complete their infection cycle. In order to successfully invade the host, a series of complex molecular interactions occur. Geminiviruses need to (i) sequester the host cellular machinery and reprogramme the cell cycle to complete their viral replication cycle, and (ii) to counteract (virus-induced) host defence mechanisms. To this end the virus hijacks, modulates and inhibits many host cellular pathways to achieve a successful infection and generate virus progeny for further dissemination (Gutierrez 2002, Castillo et al. 2003, Gorovits et al. 2013a, Miozzi et al. 2014).

When a susceptibility factor in the host plant is lacking or has been altered, viral protein recognition or binding fails; this inability of the virus to interact with such host factors results in resistant plants. Considering the fact that susceptibility factors are dominant, a recessive state of the gene is required to prevent infection by a successful virus-host interaction, thus leading to a recessive resistance. One of the best examples is the eukaryotic translation initiation factor 4E (eIF4E), which interacts with the viral protein VPg of the *Potyvirus* genus and is required for a successful viral infection. Amino acid substitutions in the protein sequence leads to an isoform with an altered cap recognition pocket, i.e. loss of 4E-VPg molecular interaction, leading to a resistance phenotype (Wang and Krishnaswamy, 2012). Other potential susceptibility factors have been identified. The host eukaryotic elongation factor 1B (eEF1B) interacts with the RNA-dependent RNA polymerase (RdRp) protein of *Tobacco Mosaic Virus* (TMV) and has been proposed to be necessary for virus replication and/or cell-to cell movement in *N. benthamiana* and *C. annuum* (Hwang et al. 2013). *Pelo* is a highly conserved gene among yeast, plants and animals, and has been shown to be required for high efficiency of protein synthesis of *Drosophila C virus* (DCV) in Drosophila. Furthermore, *pelo* deficiency also restricts replication of the RNA viruses *Cricket Paralysis Virus* (CrPV) and *Drosophila X virus*, and of the DNA virus *invertebrate iridescent virus 6* (IIV6) (Wu et al. 2014). The *Pelota* tomato homolog is involved in protein synthesis and it has been proposed that a loss of function allele of this gene (*ty-5* gene) would impede TYLCV multiplication, leading to resistance (Levin et al. 2013). These findings strongly suggest a conserved, broad-spectrum/non-race specific nature of susceptibility factors, giving rise to a recessively inherited resistance mechanism.

Nevertheless, virus isolates overcoming recessive resistance have also been reported. This is the case of a strain of *Lettuce mosaic virus* (LMV), which has overcome the resistance conferred by the lettuce *mo1* gene, a eukaryotic translation initiation factor 4E (eIF4E). The gain of virulence of the LMV resistance-breaking strain is correlated with a mutation in the cylindrical inclusion (CI) viral protein, promoting virus replication and/or movement (Sorel et al. 2014, Sanfaçon 2015). The ability of viruses to adapt to adverse conditions should not be underestimated, and the understanding of such mechanisms could provide insights for an adequate selection of resistance sources.

Functional genomics tools including genome-wide cDNA microarrays and RNA-sequencing will contribute to the understanding of plant-pathogen interactions,

and reveal virus-responsive candidate genes potentially involved in the infection process of TYLCV-like viruses. Host genes functioning in different cellular processes for a successful TYLCV-like virus infection have already been identified (Gorovits et al. 2007, Chen et al. 2013, Miozzi et al. 2014, Sade et al. 2015). Plant genes involved in susceptibility may also be considered as promising candidates in breeding programs in order to exploit durable resistance against TYLCV-like viruses. More than 50% of the known genes conferring resistance to viruses are recessive, suggesting that such source of resistance is more common to viruses than to any other plant pathogen (Nicaise et al. 2014). Nonetheless, since susceptibility factors can mediate resistance/susceptibility to other pathogens or have other non-resistance related functions in the plant, pleiotropic effects are often observed, such as dwarfism or spontaneous necrotic lesions. This negative aspect on the use of *S*-genes in plant breeding programs needs thus to be considered and addressed on a case-by-case basis (Gawehns et al. 2012).

Towards durable and broad-spectrum geminivirus resistance

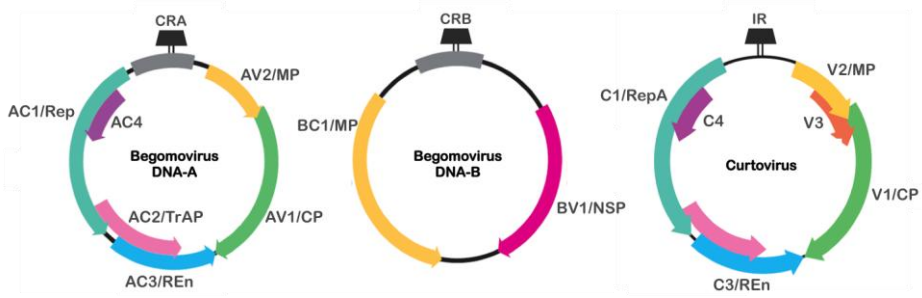
The best-characterized and exploited disease resistance is that conferred by dominant *R*-genes, mostly accompanied by a visual HR response. The resistance response is strain-specific and the (in)direct interaction between virus and *R*-gene product is imperative for pathogen recognition and triggering of a defence response. The pathogen-recognition nature of this resistance mechanism makes *R*-genes very specific and of narrow spectrum.

The resistance mechanism of the *Ty-1* gene was proposed to be based on enhanced transcriptional silencing of the invading virus. Butterbach et al. (2014) proved this hypothesis by showing an increased cytosine methylation of the TYLCV DNA genome collected from *Ty-1* resistant plants relative to susceptible tomato (MM) plants. Considering the generic role of RDRs in the amplification of RNAi, the *Ty-1* resistance was thought to be effective against a wide range of mono- and bipartite begomoviruses which was confirmed and demonstrated by a similar increased cytosine methylation of the bipartite *Tomato severe rugose virus* (ToSRV) genome. Nonetheless, *Ty-1*-mediated resistance has been overcome in some tomato production regions (García-Cano et al. 2008). Whether this involves a genuine resistance-breaking (mutant virus) strain, similar as the evolvement of resistance-breaking virus strains with single dominant *R* genes, or is a combination of various (a)biotic factors remains to be investigated. In light of the latter, a co-infection with the *Cucumber mosaic virus* (CMV) RNA virus has been shown to compromise *Ty-1*-mediated TYLCV resistance, and explained by the interference of the CMV 2b

RNAi suppressor in the TGS mechanism (Butterbach et al. 2014; Hamera et al. 2012). Cultivated tomato is a host for several RNA and DNA viruses and under field conditions mixed infections are very common. This stresses the importance of building a resistance strategy that not only protects the crop against TYLCV, but preferably to a variety of different viruses present in natural field conditions.

The resistance mechanism of the *Ty-2* has not yet been uncovered, but hypotheses on the involvement of enhanced RNAi or an NBS-LRR-mediated resistance have been demonstrated unlikely **(Chapters 2 and 3)**. The *Ty-2*-mediated resistance has been reported to be ineffective against some TYLCV strains and bipartite begomoviruses (Mejía et al. 2005). A *Ty-2*-containing line has also been observed to be susceptible to the leafhopper-transmitted monopartite *Beet curly top virus* (BCTV) from the *Curtovirus* genus (this thesis, unpublished data). The fact that the *Ty*-2 gene confers resistance to TYLCV but not to BCTV compelled us to look further into the molecular differences between these two monopartite geminiviruses. Compared to TYLCV, the BCTV genome contains 1 additional ORF (total of 7 ORFs) overlapping with both V2 and V1 proteins, named V3 (Figure 3). This protein together with V1 (coat protein) has been proposed to be implicated in virus movement through the plant. Bipartite begomoviruses contain a protein with a similar function but encoded by the DNA-B component and essential for their cell-to-cell movement. Inoculation of the DNA-A component from two such bipartite begomoviruses without their DNA-B component but mixed with a curtovirus, to *trans*-complement for this movement protein, has permitted an efficient infection in *N. benthamiana*. This demonstrated the importance and conserved nature of this protein to support the cell-to-cell movement of both groups of geminiviruses and enable their systemic infection (Briddon and Markham, 2001). Although speculative, the fact that *Ty-2* is resistant to TYLCV but susceptible to both BCTV and to bipartite begomoviruses supports the idea that this viral movement protein might play a role in the infectivity of *Ty-2*-containing lines or could potentially interact with host proteins and thereby interfere in the resistance cascade, overall counteracting the resistance response. If so, a multiple sequence comparison of viral protein sequences to identify other begomoviruses coding for a similar protein could potentially serve as a tool to predict the *Ty-2* resistance response and its resistance spectrum.

Figure 3. Schematic representation of the genomic organization of geminiviruses from the *Begomovirus* and *Curtovirus* genus. Open reading frames (ORFs) are represented by arrows. Bipartite begomoviruses contain a DNA-A and DNA-B component while monopartite begomoviruses contain one component that resembles that of the bipartite viral DNA-A component (adapted from King et al. 2011).



The durability of resistance genes is not only determined by the ability of the resistance mechanism to successfully prevent infection by a large number of strains or isolates, but also by the evolutionary potential of the given pathogen population. Recombination and mutation are major processes driving the evolution of TYLCV-like viruses, and has led to an enormous genetic diversity within this virus family (Ge et al. 2007, Moriones et al. 2007). The rapid evolution and adaptation of virus populations will not only give rise to the appearance of resistance-breaking strains and/or species but could easily broaden their host ranges and result in new viral diseases and epidemics in crops that were not earlier infected by these viruses.

A strategy to generate durable resistance could be to pyramid different genes conferring resistance to the pathogen and/or even to different strains. Each stacked gene would then contribute with a different mode of resistance, covering different layers of immunity to TYLCV. In view of the appearance of resistance-breaking isolates, there is a necessity to find new alternative resistance sources and genes. To this end, a search for new sources of TYLCV resistance was performed in wild tomato species and the resistance in *S. pimpinellifolium* accession G1.1554 was investigated (**Chapter 5**). The spectrum of this resistance still remains to be investigated, but introgression of the resistance QTLs into elite lines could potentially increase the resistance response. This already has been shown to be the case when pyramiding the *Ty-2* and *Ty-3* genes, that generated lines reported to confer an enhanced level of resistance to bipartite begomoviruses (Mejía et al. 2005).

The mechanisms underlying recessive resistance against viruses have been regarded as durable resistance (Diaz-Pendon et al. 2004, Maule et al. 2007). This makes the usage of *S*-genes an attractive alternative to be introduced in a TYLCV-resistance breeding program. An important feature of these recessive genes is that breaking down of the resistance by new virus variants may be more difficult to achieve. After all, the virus will have to be able to recognize the altered plant receptor site to trigger a compatible interaction (susceptibility), i.e. to invade and initiate a viral infection. Resistance associated with the non-functionality or mutation of a susceptibility factor can therefore be generally regarded as durable resistance. However, some potyvirus resistance-breaking isolates have already been reported to overcome recessive resistance genes that code for eukaryotic translation initiation factors (Sanfaçon 2015).

Within a single plant species, multiple accessions can display a resistance response against the same virus or to multiple viruses. This has been observed for different accessions of *S. chilense* for which TYLCV resistance loci have been mapped to tomato chromosome 6. We have shown that resistance in lines derived from different accessions is conferred by alleles of the same (RDR) gene (**Chapter 2**). Whether each allele displays the same or different virus specificity remains to be investigated. Our results have also shown the presence of more than one TYLCV resistance gene in some of the accessions. This observation/fact needs to be taken into consideration when further exploring the resistance spectrum of each source, to prevent that this resistance, which sometimes is a joint response of several distinct genes, is assigned to various viruses.

TYLCV resistance breeding: a challenging task

The tomato genome has gone through a genetic bottleneck during its domestication process. The genetic diversity of cultivated tomato is estimated to be less than 5% of the diversity in its closest wild relatives (Menda et al. 2014). Considering the narrow genetic diversity of cultivated tomato, the use of wild and genetically distant species for introgressing desirable traits is necessary. Breeding for TYLCV resistance has been based on the introgression of resistance loci from wild tomato species, but in both *Ty-1* and *Ty-2* cases, undesired traits have been introduced with the resistance trait due to linkage drag.

The high variability of the viruses causing TYLCD has led to the appearance of new TYLCV-like virus strains and species. This has resulted in diverse responses of the same resistance source in different geographic regions, which further

hampers the breeding for stable resistance. Tomato breeding lines derived from *S. chilense* have been obtained that respond differently to TYLCV infections (Pico et al. 1999). H24, a breeding line containing the resistance gene *Ty-2*, displays a clear resistance response against TYLCV strains/species from Taiwan, northern Vietnam, southern India and Israel, while it is susceptible to the northern India, Thailand and the Philippines strains (Ji et al. 2007). This response could be due to the fact that the resistance is highly specific to certain TYLCV strains, but can also be a consequence of additional introgressions present in these lines responsive to different isolates. We have shown that one *S. chilense* accession can harbour more than one resistance gene (**Chapter 2**). The presence of these additional genetic factors may be important for building a broad spectrum and more durable resistance, however can also encumber breeding efforts since the effect of the individual resistance genes may become masked.

The mechanisms underlying TYLCV resistance uncovered up to now are complex and poorly understood. Considering the dynamic nature of this virus it still is and remains a tremendous challenge for plant breeders to achieve effective and durable genetic control of TYLCV, to reduce its disease epidemics.

Troubleshooting of introgression breeding

When interspecific crosses are performed, some drawbacks are commonly observed considering the phylogenetic distances of the crossed species. Segregation distortion is a common phenomenon, which might occur due to selection for structural or genetic factors influencing the viability of spores, gametophytes and/or young sporophytes. Segregation distortion has been reported in Solanum species as eggplant, e.g. in a cross between *S. melongena* and *S. linnaeanum* (Dongalar et al. 2014). Skewness in favour of *S. melongena* alleles was observed on linkage groups 3 and 7, and a preference for *S. linnaeanum* alleles on linkage group 2. Such a phenomenon was also observed in a cross between *S. lycopersicum* x *S. pimpinellifolium* G1.1554, even though these species are very closely related (**Chapter 5**). Preference for the *S. pimpinellifolium* alleles was detected near the centromere of chromosome 2, while preference for the *S. lycopersicum* alleles was prevalent on chromosome 9. Likewise, the interval between markers P1-16 and TG36 (approx. 63 kb) within the *Ty-2* region from the cross between *S. lycopersicum* and *S. habrochaites* (**Chapter 3**) showed skewness towards the *S. habrochaites* allele. This phenomenon, together with a severe suppression of recombination in the region so far has impaired the further fine mapping and cloning of the *Ty-2* gene (**Chapter 4**).

Clusters of *R*-genes have been associated with regions of suppression of recombination, in many cases as a consequence of chromosomal rearrangements. Such an association has mainly been reported for NBS-LRR gene clusters, as in the case of e.g. the *Tm-2* (*Tomato mosaic virus* resistance) cluster on the short arm of chromosome 9 of tomato Heinz 1706 (Andolfo et al. 2014). This recombination suppression phenomenon has now not only been observed on NBS-LRR genes against plant viruses, but seems to occur for non-canonical virus resistance genes as well. Attempts to identify the *Ty-1* and *Ty-2* resistance genes against TYLCV and the *Tm-1* gene against ToMV, all examples of non-canonical resistance genes, have been difficult due to severe suppression of recombination (Verlaan et al. 2011, Yang et al. 2014, Ishibashi et al. 2007). This strengthens the idea that atypical, non-NBS-LRR classes of *R*-genes against viruses are likely associated with large genetic regions of recombination suppression as well, with a possible linkage drag, carrying detrimental genetic factors within this region. Linkage drag has been reported for the *Ty-1* (Verlaan et al. 2011) and *Ty-2* introgressions (Yang et al. 2014).

The cloning of the *Ty-1* gene in 2013 (Verlaan et al. 2013) involved a long and time-consuming process after it was first reported and mapped by Zamir et al. in 1994. It required large mapping populations over many generations of selection. The main reason why accurately mapping the *Ty-1* gene failed during so many years was the suppression of recombination in its region, caused by chromosomal rearrangements between the resistant donor (*S. chilense* LA1969) and the cultivated tomato (Verlaan et al. 2011). The occurrence of this phenomenon was only uncovered after using BAC-FISH, a technique that earlier was successfully used to study the genome organization of the *Solanaceae* family to uncover chromosome evolution (Szinay et al. 2012). The use of BAC-FISH approach thus represents a useful diagnostic tool for introgression breeding. In analogy, the further fine-mapping of the *Ty-2* gene has so far also been impaired by severe recombination suppression, due to an inversion in the introgressed region from the *S. habrochaites* resistant donor compared to the tomato genome. Contrary to the *Ty-1* story, uncovering this inversion using FISH failed, due to the small region size, but was observed by comparative analysis of the sequences from a related *S. habrochaites* accession and a BAC clone of a *Ty-2* introgression line (**Chapter 4**).
The release of the first version of the complete tomato genome sequence (variety Heinz 1706) in 2009 enabled a comparison of the position of markers in genetic maps to those on the physical map. This tool is now commonly used for a large number of genetic and genomic studies, and has facilitated and accelerated

breeding studies and applications in the tomato crop. Mapping of reads from re-sequenced lines using the tomato genome as reference can be successful in regions with limited SNPs, and insertions and deletions can be successfully detected using paired-end sequencing. However, regions containing large insertions and genome regions that are highly divergent are more difficult to align (Menda et al. 2014). Considering the wide use of wild tomato accessions for introgression breeding, the (public) release of *de novo* sequences of three wild tomato relatives of each phylogenetic group from the tomato clade (Aflitos et al. 2014) will be very useful. However, across accessions within the same species some small re-arrangements are still observed, in which case the use of BAC libraries provides an additional powerful tool for introgression breeding **(Chapter 4)**. Sequences from additional wild tomato species and more high-quality *de novo* assemblies will further ease and support introgression breeding, while pre-breeding programs developing intraspecific crosses to obtain smaller introgressions for breeding programs will limit and/or prevent the additional problem of linkage drag.

Concluding remarks

Selection of resistance sources is mainly based on the visual observation of a successful suppression or reduction of viral symptoms, involving mechanisms that enable tomato plants to deal with or overcome the negative effects of virus replication while plant development and yields are not compromised. TYLCV resistance genes currently available in breeding lines display an efficient and proper resistant response, and in some cases a broad spectrum and durable resistance has been achieved (as for the case of the *Ty-1* gene and alleles). However, all reported *Ty*-genes have been shown to allow virus replication to some extent, defining them as symptomless carriers. The large monoculturing of commercial tomato containing these resistance genes is potentially very harmful considering that these plants will still serve as reservoirs of virus populations and mixed infections could rapidly end up in new potential resistance-breaking strains. Thus, it is imperative that sustainable tomato breeding programs aimed at introgression of TYLCD resistance consider the use of various sources or genes to cover a diverse range of resistance mechanisms to a broad spectrum of geminiviruses, while continuing the search for new alternative resistance sources. The large genetic diversity of the tomato gene pool enables this and strengthens the idea that new sources will be discovered; promising accessions have already been identified for *S. arcanum, S. cheesmaniae, S. chilense, S. chmielewskii, S. corneliomulleri, S. galapagense, S. habrochaites, S. lycopersicoides, S. neorickii, S. pennellii, S. peruvianum* and *S. pimpinellifolium*. Although in some of these cases

the occurrence of allelism has to be studied, these resistant accessions might represent a golden opportunity to achieve durable, broad-spectrum resistance.

The TYLCD pathosystem is a complex ecosystem that goes far beyond the plant-virus interaction. Crop management practices to combat this disease thus cannot only rely on the use of *R* genes but additionally should involve strategies to reduce vector populations and/or prevent the evolution of new resistance-breaking begomovirus isolates for a sustainable control of TYLCD.

The most exploited resistance system is based on the use of single dominant *R* genes, but the use of *S*-genes has slowly received a growing interest. However, other factors involved in the plant-virus interactions are still somewhat underestimated. A successful virus infection not only depends on the use of host factors, but also on the virus ability to avoid inhibitory host responses. Even in compatible interactions, inhibition of virus multiplication seems to take place, although this inhibition is weak and thus still allows virus replication and disease development. For instance, the proteins Pseudouridine Synthase 4 (Pus4) and Actin Patch Protein 1 (App1), when overexpressed in *N. benthamiana* plants have been shown to inhibit the RNA-virus *Brome Mosaic Virus* (BMV) systemic spread and to reduce virus accumulation (Zhu et al. 2007). Further studies on such negative regulators of virus replication in compatible interactions would then uncover yet unexplored potential sources of resistance to be implemented in breeding programs. Likewise, an understanding of the interactions between viruses and their satellite molecules, as well as of natural mixed infections, still need to be further explored, and to assess their impact on the resistant lines bred up to now. A robust and wide-ranging antiviral host defence would thus cover different layers of plant immunity, including *R*-genes, SAR, *S*-genes, PTI and ETI, targeting all stages of a virus infection cycle.

The rapid advances and availability of next-generation sequence (NGS) technologies and bioinformatic tools allow breeders not only to use this information for MAS or *in silico* mapping (**Chapter 5**), but also to predict regions with potential chromosomal rearrangements (**Chapter 4**). NGS for *de novo* whole genome shotgun (WGS) sequencing and re-sequencing (WGRS) tools should be more widely implemented and used by 'next-generation breeders' to develop crossing plans, allowing them much better to consider whether inter- or intra-specific crosses have to be performed for the introgression of shorter alien regions.

Reference list

Abbott JC, Aanensen DM, Rutherford K, Butcher S, Spratt BG (2005) WebACT-an online companion for the Artemis Comparison Tool. Bioinformatics 21: 3665-3666

Achenbach UC, Tang X, Ballvora A, de Jong H, Gebhardt C (2010) Comparison of the chromosome maps around a resistance hot spot on chromosome 5 of potato and tomato using BAC-FISH painting. Genome 53: 103-110

Aflitos S, Schijlen E, de Jong H, de Ridder D, Smit S, Finkers R, Wang J, Zhang G, Li N, Mao L et al (2014) Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing. The Plant Journal 80: 136-148

Afzal AJ, Srour A, Saini N, Hemmati N, El Shemy HA, Lightfoot DA (2012) Recombination suppression at the dominant Rhg1/Rfs2 locus underlying soybean resistance to the cyst nematode. Theoretical and Applied Genetics 124: 1027-1039

Agrama H, Scott J (2006) Quantitative trait loci for Tomato yellow leaf curl virus and Tomato mottle virus resistance in tomato. Journal of the American Society for Horticultural Science 131: 267-272

Alvarado V, Scholthof HB (2009) Plant responses against invasive nucleic acids: RNA silencing and its suppression by plant viral pathogens. Seminars in cell & developmental biology. Elsevier, pp 1032-1040

Anbinder I, Reuveni M, Azari R, Paran I, Nahon S, Shlomo H, Chen L, Lapidot M, Levin I (2009) Molecular dissection of Tomato leaf curl virus resistance in tomato line TY172 derived from Solanum peruvianum. Theoretical and Applied Genetics 119: 519-530

Anderson LK, Covey PA, Larsen LR, Bedinger P, Stack SM (2010) Structural differences in chromosomes distinguish species in the tomato clade. Cytogenetic and Genome Research 129: 24-34

Andolfo G, Jupe F, Witek K, Etherington GJ, Ercolano MR, Jones JD (2014) Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. BMC plant biology 14: 120

Antignus Y, Nestel D, Cohen S, Lapidot M (2001) Ultraviolet-deficient greenhouse environment affects whitefly attraction and flight-behavior. Environmental Entomology 30: 394-399

Ashrafi H, Kinkade M, Foolad MR (2009) A new genetic linkage map of tomato based on a Solanum lycopersicum × S. pimpinellifolium RIL population displaying locations of candidate pathogen response genes. Genome 52: 935-956

Bai Y, Van der Hulst R, Huang C, Wei L, Stam P, Lindhout P (2004) Mapping Ol-4, a gene conferring resistance to Oidium neolycopersici and originating from

Lycopersicon peruvianum LA2172, requires multi-allelic, single-locus markers. Theoretical and Applied Genetics 109: 1215-1223

Ballvora A, Jöcker A, Viehöver P, Ishihara H, Paal J, Meksem K, Bruggmann R, Schoof H, Weisshaar B, Gebhardt C (2007) Comparative sequence analysis of Solanum and Arabidopsis in a hot spot for pathogen resistance on potato chromosome V reveals a patchwork of conserved and rapidly evolving genome segments. BMC Genomics 8: 112

Banerjee MK, Kalloo MK (1987) Sources and inheritance of resistance to leaf curl virus in Lycopersicon. Theoretical and Applied Genetics 73: 707-710

Barbieri M, Acciarri N, Sabatini E, Sardo L, Accotto G, Pecchioni N (2010) Introgression of resistance to two Mediterranean virus species causing tomato yellow leaf curl into a valuable traditional tomato variety. Journal of Plant Pathology: 485-493

Barboza N, Blanco-Meneses M, Hallwass M, Moriones E, Inoue-Nagata A (2014) First Report of Tomato yellow leaf curl virus in Tomato in Costa Rica. Plant Disease 98: 699-699

Bar-Ziv A, Levy Y, Hak H, Mett A, Belausov E, Citovsky V, Gafni Y (2012) The Tomato yellow leaf curl virus (TYLCV) V2 protein interacts with the host papain-like cysteine protease CYP1. Plant signaling & behavior 7: 983-989

Den Boer E (2014) Genetic investigation of the nonhost resistance of wild lettuce, Lactuca saligna, to lettuce downey mildew, Bremia lactucae (PhD thesis)

Bol J, Linthorst H, Cornelissen B (1990) Plant pathogenesis-related proteins induced by virus infection. Annual Review of Phytopathology 28: 113-138

Botermans M, Verhoeven JTJ, Jansen C, Roenhorst J, Stijger C, Pham K (2009) First report of Tomato yellow leaf curl virus in tomato in the Netherlands. Plant Disease 93: 1073-1073

Briddon R, Markham P (2001) Complementation of bipartite begomovirus movement functions by topocuviruses and curtoviruses. Archives of virology 146: 1811-1819

Briddon RW, Stanley J (2009) Geminiviridae. eLS

Broman KW (2005) The genomes of recombinant inbred lines. Genetics 169: 1133-1146

Brough CL, Gardiner WE, Inamdar NM, Zhang X-Y, Ehrlich M, Bisaro DM (1992) DNA methylation inhibits propagation of tomato golden mosaic virus DNA in transfected protoplasts. Plant Molecular Biology 18: 703-712

Brown J, Frohlich D, Rosell R (1995) The sweetpotato or silverleaf whiteflies: biotypes of Bemisia tabaci or a species complex? Annual Review of Entomology 40: 511-534

Brown JK (2007) The Bemisia tabaci complex: genetic and phenotypic variation and relevance to TYLCV-vector interactions. Tomato Yellow Leaf Curl Virus Disease. Springer, pp 25-56

Buchmann RC, Asad S, Wolf JN, Mohannath G, Bisaro DM (2009) Geminivirus AL2 and L2 Proteins Suppress Transcriptional Gene Silencing and Cause Genome-Wide Reductions in Cytosine Methylation. Journal of Virology 83: 5005-5013

Buck KW (1999) Replication of tobacco mosaic virus RNA. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 354: 613-627

Butterbach P, Verlaan MG, Dullemans A, Lohuis D, Visser RGF, Bai Y, Kormelink R (2014) The TYLCV Resistance Gene Ty-1 confers resistance in tomato through enhanced transcriptional gene silencing. Proceedings of the Natural Academy of Sciences 111: 12942-12947

CABI, 2015. Invasive Species Compendium. Wallingford , UK: CAB International. www.cabi.org/isc

Canady MA, ,1 Ji Y, Chetelat RT (2006) Homeologous Recombination in Solanum lycopersicoides Introgression Lines of Cultivated Tomato. Genetics 174: 1775-1788

Carr JP, Lewsey MG, Palukaitis P (2010) Signaling in induced resistance. Advances in virus research 76: 57-121

Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J (2005) ACT: the Artemis comparison tool. Bioinformatics 21: 3422-3423

Castillo AG, Collinet D, Deret S, Kashoggi A, Bejarano ER (2003) Dual interaction of plant PCNA with geminivirus replication accessory protein (Ren) and viral replication protein (Rep). Virology 312: 381-394

Castillo AG, Morilla G, Lozano R, Collinet D, Perez-Luna A, Kashoggi A, Bejarano E (2007) Identification of Plant Genes Involved in TYLCV Replication. Tomato Yellow Leaf Curl Virus Disease. Springer, pp 207-221

Causse M, Desplat N, Pascual L, Le Paslier MC, Sauvage C, Bauchet G, Bérard A, Bounon R, Tchoumakov M, Brunel D et al (2013) Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. BMC Genomics 14: 791

Chagué V, Mercier J, Guenard M, De Courcel A, Vedel F (1997) Identification of RAPD markers linked to a locus involved in quantitative resistance to TYLCV in tomato by bulked segregant analysis. Theoretical and Applied Genetics 95: 671-677

Chellappan P, Masona M, Vanitharani R, Taylor N, Fauquet C (2004) Broad Spectrum Resistance to ssDNA Viruses Associated with Transgene-Induced Gene Silencing in Cassava. Plant Molecular Biology 56: 601-611

Chen FQ, Foolad MR (1999) A molecular linkage map of tomato based on a cross between Lycopersicon esculentum and L. pimpinellifolium and its comparison with other molecular maps of tomato. Genome 42: 94-103

Chen T, Lv Y, Zhao T, Li N, Yang Y, Yu W, He X, Liu T, Zhang B (2013) Comparative transcriptome profiling of a resistant vs. susceptible tomato (Solanum lycopersicum) cultivar in response to infection by tomato yellow leaf curl virus. PloS one 8: e80816

Chibon P, Schoof H, Visser RG, Finkers R (2012) Marker2sequence, mine your QTL regions for candidate genes. Bioinformatics 28: 1921-1922

Chisholm ST, Mahajan SK, Whitham SA, Yamamoto ML, Carrington JC (2000) Cloning of the Arabidopsis RTM1 gene, which controls restriction of long-distance movement of tobacco etch virus. Proceedings of the National Academy of Sciences 97: 489-494

Chomdej O, Chatchawankanpanich O, Kositratana W, Chunwongse J (2007) Response of resistant breeding lines of tomato germplasm and their progenies with Seedathip3 to Tomato Yellow Leaf Curl Virus, Thailand isolate (TYLCTHV-[2]). Songklanakarin Journal Science and Technology 29: 1469-1477

Cohen S, Lapidot M (2007) Appearance and expansion of TYLCV: A historical point of view. Czosnek H (ed) Tomato yellow leaf curl virus disease. Springer, The Netherlands, pp 3–12

Cosson P, Schurdi-Levraud V, Le QH, Sicard O, Caballero M, Roux F, Le Gall O, Candresse T, Revers F (2012) The RTM resistance to potyviruses in Arabidopsis thaliana: natural variation of the RTM genes and evidence for the implication of additional genes. PLoS One 7: e39169

Cosson P, Sofer L, Le QH, Léger V, Schurdi-Levraud V, Whitham SA, Yamamoto ML, Gopalan S, Le Gall O, Candresse T (2010) RTM3, which controls long-distance movement of potyviruses, is a member of a new plant gene family encoding a meprin and TRAF homology domain-containing protein. Plant physiology 154: 222-232

Cui X, Zhou X (2004) AC2 and AC4 proteins of Tomato yellow leaf curl China virus and Tobacco curly shoot virus mediate suppression of RNA silencing. Chinese Science Bulletin 49: 2607-2612

Czosnek H (2007a) Interactions of Tomato yellow leaf curl virus with its whitefly vector. Tomato Yellow Leaf Curl Virus Disease: 157-170

Czosnek H (2007b) Tomato yellow leaf curl virus disease: management, molecular biology, breeding for resistance. Recherche 67: 02

Czosnek H, Eybishtz A, Sade D, Gorovits R, Sobol I, Bejarano E, Rosas-Díaz T, Lozano-Durán R (2013) Discovering host genes involved in the infection by the tomato yellow leaf curl virus complex and in the establishment of

resistance to the virus using Tobacco Rattle Virus-based post transcriptional gene silencing. Viruses 5: 998-1022

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, Group GPA (2011) The variant call format and VCFtools. Bioinformatics 27: 2156-2158

de Castro AP, Blanca JM, Díez MJ, Vinals FN (2007) Identification of a CAPS marker tightly linked to the Tomato yellow leaf curl disease resistance gene Ty-1 in tomato. European Journal of Plant Pathology 117: 347-356

De Ronde D, Butterbach P, Kormelink R (2014) Dominant resistance against plant viruses. Frontiers in plant science 5

De Vos RCH, Moco S, Lommen A, Keurentjes JJB, Bino RJ, Hall RD (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nature Protocols 2: 778-791

Diaz-Pendon JA, Cañizares MC, Moriones E, Bejarano ER, Czosnek H, Navas-Castillo J (2010) Tomato yellow leaf curl viruses: ménage à trois between the virus complex, the plant and the whitefly vector. Molecular plant pathology 11: 441-450.

Diaz-Pendon JA, Truniger V, Nieto C, Garcia-Mas J, Bendahmane A, Aranda MA (2004) Advances in understanding recessive resistance to plant viruses. Molecular plant pathology 5: 223-233

Doğanlar S, Frary A, Daunay MC, Huvenaars K, Mank R, Frary A (2014) High resolution map of eggplant (Solanum melongena) reveals extensive chromosome rearrangement in domesticated members of the Solanaceae. Euphytica 198: 231-241

Doğanlar S, Frary A, Ku HM, Tanksley SD (2002) Mapping quantitative trait loci in inbred backcross lines of Lycopersicon pimpinellifolium (LA1589). Genome 45: 1189-1202

Eybishtz A, Peretz Y, Sade D, Akad F, Czosnek H (2009) Silencing of a single gene in tomato plants resistant to Tomato yellow leaf curl virus renders them susceptible to the virus. Plant Molecular Biology 71: 157-171

Eybishtz A, Peretz Y, Sade D, Gorovits R, Czosnek H (2010) Tomato yellow leaf curl virus infection of a resistant tomato line with a silenced sucrose transporter gene LeHT1 results in inhibition of growth, enhanced virus spread, and necrosis. Planta 231: 537-548

Fauquet CM, Briddon RW, Brown JK, Moriones E, Stanley J, Zerbini M, Zhou X (2008) Geminivirus strain demarcation and nomenclature. Archives of Virology 153: 783-821

Finkers-Tomczak A, Bakker E, de Boer J, van der Vossen E, Achenbach U, Golas T, Suryaningrat S, Smant G, Bakker J, Goverse A (2011) Comparative sequence analysis of the potato cyst nematode resistance locus H1 reveals

a major lack of co-linearity between three haplotypes in potato (Solanum tuberosum ssp.). Theoretical and Applied Genetics 122: 595-608

Firdaus S, van Heusden AW, Hidayati N, Supena EDJ, Mumm R, de Vos RC, Visser RG, Vosman B (2013) Identification and QTL mapping of whitefly resistance components in Solanum galapagense. Theoretical and Applied Genetics 126: 1487-1501

Friedmann M, Lapidot M, Cohen S, Pilowsky M (1998) A Novel Source of Resistance to Tomato Yellow Leaf Curl Virus Exhibiting a Symptomless Reaction to Viral Infection. Journal of the American Society for Horticultural Science 123: 1004-1007

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Molecular Biology Reporter 13: 207–209

García-Cano E, Resende R, Boiteux L, Giordano L, Fernández-Muñoz R, Moriones E (2008) Phenotypic expression, stability, and inheritance of a recessive resistance to monopartite begomoviruses associated with tomato yellow leaf curl disease in tomato. Phytopathology 98: 618-627

Gawehns F, Cornelissen BJ, Takken FL (2013) The potential of effector-target genes in breeding for plant innate immunity. Microbial biotechnology 6: 223-229

Ge L, Zhang J, Zhou X, Li H (2007) Genetic structure and population variability of tomato yellow leaf curl China virus. Journal of virology 81: 5902-5907

Glick E, Levy Y, Gafni Y (2009) The viral etiology of tomato yellow leaf curl disease-a review. Plant Protection Science 45: 81-97

Gorovits R, Akad F, Beery H, Vidavsky F, Mahadav A, Czosnek H (2007) Expression of stress-response proteins upon whitefly-mediated inoculation of Tomato yellow leaf curl virus in susceptible and resistant tomato plants. Molecular Plant-Microbe Interactions 20: 1376-1383

Gorovits R, Moshe A, Ghanim M, Czosnek H (2013a) Recruitment of the host plant heat shock protein 70 by tomato yellow leaf curl virus coat protein is required for virus infection. PloS one 8: e70280

Gorovits R, Moshe A, Kolot M, Sobol I, Czosnek H (2013b) Progressive aggregation of Tomato yellow leaf curl virus coat protein in systemically infected tomato plants, susceptible and resistant to the virus. Virus Research 171: 33-43

Gottlieb Y, Zchori-Fein E, Mozes-Daube N, Kontsedalov S, Skaljac M, Brumin M, Sobol I, Czosnek H, Vavre F, Fleury F (2010) The transmission efficiency of tomato yellow leaf curl virus by the whitefly Bemisia tabaci is correlated with the presence of a specific symbiotic bacterium species. Journal of Virology 84: 9310-9317

Grandillo S, Tanksley SD (1996) Genetic analysis of RFLPs, GATA microsatellites and RAPDs in a cross between L. esculentum and L. pimpinellifolium. Theoretical and Applied Genetics 92: 957-965

Griffiths PD, Scott JW (2001) Inheritance and linkage of tomato mottle virus resistance genes derived from Lycopersicon chilense accession LA 1932. Journal of the American Society of Horticultural Sciences 126: 462–467

Gronenborn B (2007) The tomato yellow leaf curl virus genome and function of its proteins. Tomato Yellow Leaf Curl Virus Disease: 67-84

Gutierrez C (2000) DNA replication and cell cycle in plants: learning from geminiviruses. The EMBO journal 19: 792-799

Gutierrez C (2002) Strategies for geminivirus DNA replication and cell cycle interference. Physiological and molecular plant pathology 60: 219-230

Haag JR, Pikaard CS (2011) Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nature reviews Molecular cell biology 12: 483-492

Hallwass M, Oliveira AS, Campos Dianese E, Lohuis D, Boiteux LS, Inoue-Nagata AK, Resende RO, Kormelink R (2014) The Tomato spotted wilt virus cell-to-cell movement protein (NSM) triggers a hypersensitive response in Sw-5-containing resistant tomato lines and in Nicotiana benthamiana transformed with the functional Sw-5b resistance gene copy. Molecular plant pathology 15: 871-880

Hamera S, Song X, Su L, Chen X, Fang R (2012) Cucumber mosaic virus suppressor 2b binds to AGO4 related small RNAs and impairs AGO4 activities. The Plant Journal 69: 104-115

Hanemann A, Schweizer GF, Cossu R, Wicker T, Röder MS (2009) Fine mapping, physical mapping and development of diagnostic markers for the Rrs2 scald resistance gene in barley. Theoretical and Applied Genetics 119: 1507-1522

Hanley-Bowdoin L, Settlage SB, Robertson D (2004) Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. Molecular plant pathology 5: 149-156

Hanson PM, Bernacchi D, Green S, Tanksley SD, Muniyappa V, Padmaja AS, Chen H, Kuo G, Fang D, Chen J (2000) Mapping a wild tomato introgression associated with tomato yellow leaf curl virus resistance in a cultivated tomato line. Journal of the American Society for Horticultural Science 15: 15–20

Hanson PM, Green SK, Kuo G (2006) Ty-2, a gene on chromosome 11 conditioning geminivirus resistance in tomato. Tomato Genetics Cooperative Report 56: 17-18

Hogenhout SA, Ammar ED, Whitfield AE, Redinbaugh MG (2008) Insect Vector Interactions with Persistently Transmitted Viruses. Annual Review of Phytopathology 46: 327-359

Horowitz AR, Denholm I, Morin S (2007) Resistance of the TYLCV whitefly vector Bemisia tabaci to insecticides. In: Czosnek H (ed) Tomato yellow leafcurl virus disease. Springer, New York, pp 309–329

Hosseinzadeh MR, Shams-Bakhsh M, Osaloo SK, Brown JK (2014) Phylogenetic relationships, recombination analysis, and genetic variability among diverse variants of tomato yellow leaf curl virus in Iran and the Arabian Peninsula: further support for a TYLCV center of diversity. Archives of virology 159: 485-497

Howe GA, Lee GI, Itoh A, Li L, DeRocher AE (2000) Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning and expression of allene oxide synthase and fatty acid hydroperoxide lyase. Plant Physiology 123: 711-724

Huang S, Van Der Vossen EA, Kuang H, Vleeshouwers VG, Zhang N, Borm TJ, Van Eck HJ, Baker B, Jacobsen E, Visser RG (2005) Comparative genomics enabled the isolation of the R3a late blight resistance gene in potato. The Plant Journal 42: 251-261

Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang T et al (2009) High-throughput genotyping by whole-genome resequencing. Genome Research 19: 1068-107

Huang Z, van der Knaap E (2011) Tomato fruit weight 11.3 maps close to fasciated on the bottom of chromosome 11. Theoretical and Applied Genetics 123: 465-474

Hutton SF, Scott JW (2013) Fine-Mapping and Cloning of Ty-1 and Ty-3; and Mapping of a New TYLCV Resistance Locus, "Ty-6" In: Tomato Breeders Round Table Proceedings 2013 (Chiang Mai, Thailand)

Hutton SF, Scott JW, Schuster DJ (2012) Recessive resistance to tomato yellow leaf curl virus from the tomato cultivar Tyking is located in the same region as Ty-5 on chromosome 4. Journal of the American Society for Horticultural Science 47: 324-327

Hwang J, Oh C-S, Kang B-C (2013) Translation elongation factor 1B (eEF1B) is an essential host factor for Tobacco mosaic virus infection in plants. Virology 439: 105-114

Illumina: Infinium® HD Assay: Ultra Protocol Guide. California, USA: ©Illumina, Inc; 2009:1–224. Catalog #WG-901-4007

Iovene M, Wielgus SM, Simon PW, Buell CR, Jiang J (2008) Chromatin structure and physical mapping of chromosome 6 of potato and comparative analysis with tomato. Genetics 180: 1307-1317

Ishibashi K, Masuda K, Naito S, Meshi T, Ishikawa M (2007) An inhibitor of viral RNA replication is encoded by a plant resistance gene. Proceedings of the National Academy of Sciences 104: 13833-13838

Ishibashi K, Kezuka Y, Kobayashi C, Kato M, Inoue T, Nonaka T, Ishikawa M, Matsumura H, Katoh E (2014) Structural basis for the recognition–evasion arms race between Tomato mosaic virus and the resistance gene Tm-1. Proceedings of the National Academy of Sciences 111: E3486-E3495

Ji Y, Schuster DJ, Scott JW (2007) Ty-3, a begomovirus resistance locus near the Tomato yellow leaf curl virus resistance locus Ty-1 on chromosome 6 of tomato. Molecular Breeding 20: 271-284

Ji Y, Scott JW, Hanson P, Graham E, Maxwell DP (2007b) Sources of resistance, inheritance, and location of genetic loci conferring resistance to members of the tomato-infecting begomoviruses. In: Czosnek H (ed) Tomato yellow leaf curl virus disease: Management, molecular biology, breeding for resistance. Kluwer, Dordrecht, The Netherlands, pp 343–362

Ji Y, Salus MS, van Betteray B, Smeets J, Jensen KS, Martin CT, Mejía L, Scott JW, Havey MJ, Maxwell DP (2007c) Co-dominant SCAR markers for detection of the Ty-3 and Ty-3a loci from Solanum chilense at 25 cM of chromosome 6 of tomato. Report Tomato Genetics Cooperative 57: 25–28

Ji Y, Scott JW, Maxwell DP, Schuster DJ (2008) Ty-4, a tomato yellow leaf curl virus resistance gene on chromosome 3 of tomato. Reports Tomato Genetics Cooperative 58: 29–31

Ji Y, Scott JW, Schuster DJ, Maxwell DP (2009) Molecular mapping of Ty-4, a new Tomato yellow leaf curl virus resistance locus on chromosome 3 of Tomato. Journal of the American Society for Horticultural Science 134: 281-288

Ji Y, Scott JW, Schuster DJ (2009a) Toward fine mapping of the Tomato yellow leaf curl virus resistance gene Ty-2 on chromosome 11 of tomato. Journal of the American Society for Horticultural Science 44: 614–618

Jones JD, Dangl JL (2006) The plant immune system. Nature 444: 323-329

Kabelka E, Franchino B, Francis DM (2002) Two loci from Lycopersicon hirsutum LA407 confer resistance to strains of Clavibacter michiganensis subsp. michiganensis. Phytopathology 92: 504-510

Kadirvel P, de la Peña R, Schafleitner R, Huang S, Geethanjali S, Kenyon L, Tsai W, Hanson P (2013) Mapping of QTLs in tomato line FLA456 associated with resistance to a virus causing tomato yellow leaf curl disease. Euphytica 190: 297-308

Kalloo G, Banerjee MK (1990) Transfer of tomato leaf curl virus resistance from Lycopersicon hirsutum f. glabratum to L. esculentum. Plant Breeding 105: 156–159

Kasrawi MA, Suwwan MA, Mansour A (1988) Sources of resistance to tomato-yellow-leaf-curl-virus (TYLCV) in Lycopersicon species. Euphytica 37: 61-64

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30: 772-780

Kaufmann K, Muino JM, Jauregui R, Airoldi CA, Smaczniak C, Krajewski P, Angenent GC (2009) Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the Arabidopsis flower. PLoS biology 7: e1000090

Keurentjes JJ, Bentsink L, Alonso-Blanco C, Hanhart CJ, Blankestijn-De Vries H, Effgen S, Vreugdenhil D, Koornneef M (2007) Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. Genetics 175: 891-905

Khan SA (2005) Plasmid rolling-circle replication: highlights of two decades of research. Plasmid 53: 126-136

Khan N, Kazmi RH, Willems LAJ, van Heusden AW, Ligterink W, Hilhorst HWM (2012) Exploring the Natural Variation for Seedling Traits and Their Link with Seed Dimensions in Tomato. PLoS ONE 7: e43991

King AM, Adams MJ, Lefkowitz EJ (2011) Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier

Kørner CJ, Klauser D, Niehl A, Domínguez-Ferreras A, Chinchilla D, Boller T, Heinlein M, Hann DR (2013) The immunity regulator BAK1 contributes to resistance against diverse RNA viruses. Molecular plant-microbe interactions 26: 1271-1280

Krattinger SG, Lagudah ES, Spielmeyer W, Singh RP, Huerta-Espino J, McFadden H, Bossolini E, Selter LL, Keller B (2009) A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. Science 323: 1360-1363

Kuang H, Wei F, Marano MR, Wirtz U, Wang X, Liu J, Shum WP, Zaborsky J, Tallon LJ, Rensink W, Lobst S, Zhang P, Tornqvist CE, Tek A, Bamberg J, Helgeson J, Fry W, You F, Luo MC, Jiang J, Buell RC, Baker B (2005) The R1 resistance gene cluster contains three groups of independently evolving, type I R1 homologues and shows substantial variation among haplotypes of Solanum demissum. The Plant Journal 44: 37-51

Lai M (1998) Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology 244: 1-12

Lanfermeijer FC, Warmink J, Hille J (2005) The products of the broken Tm-2 and the durable Tm-22 resistance genes from tomato differ in four amino acids. Journal of experimental botany, 56: 2925-2933

Lapidot M (2007) Screening for TYLCV-resistance plants using whitefly-mediated inoculation. Tomato Yellow Leaf Curl Virus Disease: 329-342

Larkan N, Lydiate D, Parkin I, Nelson M, Epp D, Cowling W, Rimmer S, Borhan M (2013) The Brassica napus blackleg resistance gene LepR3 encodes a receptor-like protein triggered by the Leptosphaeria maculans effector AVRLM1. New Phytologist 197: 595-605

Lee S, Woo Y-M, Ryu S-I, Shin Y-D, Kim WT, Park KY, Lee I-J, An G (2008) Further characterization of a rice AGL12 group MADS-box gene, OsMADS26. Plant Physiology 147: 156-168

Lefeuvre P, Martin DP, Harkins G, Lemey P, Gray AJA, Meredith S, Lakay F, Monjane A, Lett JM, Varsani A (2010) The spread of Tomato yellow leaf curl virus from the Middle East to the world. PLoS Pathogens 6: e1001164

Levin I, Karniel U, Fogel D, Reuveni M, Gelbart D, Evenor D, Chen L, Nahon S, Shlomo H, Machbosh Z, and Lapidot M (2013) Cloning and Analysis of the Tomato yellow leaf curl virus Resistance Gene Ty-5. Proc. Tomato Breeders Roundtable, Chaing-Mai, Thailand http://tgc.ifas.ufl.edu/2013/abstracts/LevinAbstract%20TBRT%202013.pdf

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25: 1754-1760

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26: 589-595

Li W, Royer S, and Chelelat RT (2010) Fine Mapping of *ui6.1*, a Gametophytic Factor Controlling Pollen-Side Unilateral Incompatibility in Interspecific Solanum Hybrids. Genetics 185: 1069-1080

Lippman Z, Tanksley SD (2001) Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species Lycopersicon pimpinellifolium and L. esculentum var. Giant Heirloom. Genetics 158: 413-422

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR (2006) Gas chromatography mass spectrometry-based metabolite profiling in plants. Nature Protocols 1: 387-396

Liu Y, Schiff M, Dinesh-Kumar S (2002) Virus-induced gene silencing in tomato. The Plant Journal 31: 777-786

Lou Q, Iovene M, Spooner DM, Buell CR, Jiang J (2010) Evolution of chromosome 6 of Solanum species revealed by comparative fluorescence in situ hybridization mapping. Chromosoma 119: 435-442

Lozano-Durán R, Rosas-Díaz T, Gusmaroli G, Luna AP, Taconnat L, Deng XW, Bejarano ER (2011a) Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in Arabidopsis thaliana. The Plant Cell Online 23: 1014-1032

Lozano-Durán R, Rosas-Díaz T, Luna AP, Bejarano ER (2011b) Identification of host genes involved in geminivirus infection using a reverse genetics approach. PloS one 6: e22383

Lucas WJ (2006) Plant viral movement proteins: agents for cell-to-cell trafficking of viral genomes. Virology 344: 169-184

Luna AP, Morilla G, Voinnet O, Bejarano ER (2012) Functional analysis of gene silencing suppressors from Tomato yellow leaf curl disease viruses. Molecular Plant-Microbe Interactions 25: 1294-1306

Mandadi KK, Scholthof K-BG (2013) Plant immune responses against viruses: how does a virus cause disease? The Plant Cell 25: 1489-1505

Maruthi MN, Czosnek H, Vidavski F, Tarba S-Y, Milo J, Leviatov S, Venkatesh HM, Padmaja AS, Kulkarni RS, Muniyappa V (2003) Comparison of resistance to Tomato leaf curl virus (India) and Tomato yellow leaf curl virus (Israel) among Lycopersicon wild species, breeding lines and hybrids. European Journal of Plant Pathology 109: 1-11

Matsuda D, Dreher TW (2004) The tRNA-like structure of Turnip yellow mosaic virus RNA is a 3′-translational enhancer. Virology 321: 36-46

Maule AJ, Caranta C, Boulton MI (2007) Sources of natural resistance to plant viruses: status and prospects. Molecular Plant Pathology 8: 223-231

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M et al: The genome analysis toolkit (2010) A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20: 1297-1303

Mejía L, Teni RE, Vidavski F, Czosnek H, Lapidot M, Nakhla MK, Maxwell DP (2005) Evaluation of tomato germplasm and selection of breeding lines for resistance to begomoviruses in Guatemala. In Proceedings 1st International Symposium on Tomato Diseases. Momol MT, Ji P, Jones JB (eds) Acta Horticulturae 695: 251-256

Mejía L, Teni RE, García BE, Fulladolsa AC, Méndez L, Melgar S, Maxwell DP (2010) Preliminary observations on the effectiveness of five introgressions for resistance to begomoviruses in tomatoes. TGC Report, 60: 41–53

Melzer M, Ogata D, Fukuda S, Shimabuku R, Borth W, Sether D, Hu J (2010) First report of Tomato yellow leaf curl virus in Hawaii. Plant Disease 94: 641-641

Menda N, Strickler SR, Edwards JD, Bombarely A, Dunham DM, Martin GB, Mejia L, Hutton SF, Havey MJ, Maxwell DP, Mueller LA (2014) Analysis of wild-species introgressions in tomato inbreds uncovers ancestral origins. BMC Plant Biology 14: 287

Mézard C (2006) Meiotic recombination hotspots in plants. Biochemical Society Transactions 34: 531-534

Miozzi L, Napoli C, Sardo L, Accotto GP (2014) Transcriptomics of the Interaction between the Monopartite Phloem-Limited Geminivirus Tomato Yellow Leaf Curl Sardinia Virus and Solanum lycopersicum Highlights a Role for Plant Hormones, Autophagy and Plant Immune System Fine Tuning during Infection. PloS one 9: e89951

Morin S, Ghanim M, Sobol I, Czosnek H (2000) The GroEL protein of the whitefly Bemisia tabaci interacts with the coat protein of transmissible and nontransmissible begomoviruses in the yeast two-hybrid system. Virology 276: 404-416

Moriones E, Navas-Castillo J (2000) Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. Virus Research 71: 123-134

Moriones E, García-Andrés S and Navas-Castillo J (2007) Recombination in the TYLCV complex: a mechanism to increase genetic diversity. Implications for plant resistance development. Tomato Yellow Leaf Curl Virus Disease: 119-138

Moshe A, Pfannstiel J, Brotman Y, Kolot M, Sobol I, Czosnek H, Gorovits R (2012) Stress responses to Tomato Yellow (Leaf Curl Virus TYLCV) infection of Resistant and Susceptible tomato plants are different. Metabolomics: 2153-0769

Moyle LC (2008) Ecological and evolutionary genomics in the wild tomatoes (Solanum sect. Lycopersicon). Evolution 62: 2995-3013

Nicaise V (2014) Crop immunity against viruses: outcomes and future challenges. Frontiers in plant science 5

Ning W, Shi X, Liu B, Pan H, Wei W, Zeng Y, Sun X, Xie W, Wang S, Wu Q (2015) Transmission of Tomato Yellow Leaf Curl Virus by Bemisia tabaci as Affected by Whitefly Sex and Biotype. Scientific Reports 5

Pan H, Chu D, Yan W, Su Q, Liu B, Wang S, Wu Q, Xie W, Jiao X, Li R (2012) Rapid spread of Tomato yellow leaf curl virus in China is aided differentially by two invasive whiteflies. PLoS One 7: e34817

Paran I, Goldman I, Tanksley SD, Zamir D (1995) Recombinant inbred lines for genetic mapping in tomato. Theoretical and Applied Genetics 90: 542-548

Parlevliet J, Zadoks J (1977) The integrated concept of disease resistance: a new view including horizontal and vertical resistance in plants. Euphytica 26: 5-21

Pérez de Castro A, Díez MJ, Nuez F (2005a) Caracterización de entradas de Lycopersicon peruvianum y L. chilense por su resistencia al Tomato yellow leaf curl virus (TYLCV). Actas Portuguesas de Horticultura 8: 48-54

Pérez de Castro A, Díez MJ, Nuez F (2005b) Evaluation of breeding tomato lines partially resistant to Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus derived from Lycopersicon chilense. Canadian journal of plant pathology 27: 268-275

Pérez de Castro A, Díez MJ, Nuez F (2007) Inheritance of Tomato yellow leaf curl virus resistance derived from Solanum pimpinellifolium UPV16991. Plant Disease 91: 879-885

Pérez de Castro A, Julián O, Díez M (2013) Genetic control and mapping of Solanum chilense LA1932, LA1960 and LA1971-derived resistance to Tomato yellow leaf curl disease. Euphytica 190: 203-214

Peters SA, Bargsten JW, Szinay D, van de Belt J, Visser RGF, Bai Y, de Jong H (2012) Structural homology in the Solanaceae: analysis of genomic regions in support of synteny studies in tomato, potato and pepper. The Plant Journal 71: 602-614

Pico B, Diez MJ, Nuez F (1996) Viral diseases causing the greatest economic losses to the tomato crop. 2. The tomato yellow leaf curl virus-A review. Scientia Horticulturae 67: 151–196

Pico B, Ferriol M, Diez MJ, Nuez F (1999) Developing tomato breeding lines resistant to Tomato yellow leaf curl virus. Plant Breeding: 118, 537-542

Pico B, Sifres A, Elia M, Díez MJ, Nuez F (2000) Searching for new resistance sources to tomato yellow leaf curl virus within a highly variable wild Lycopersicon genetic pool. Acta Physiologiae Plantarum 22: 344-350

Pico B, Ferriol M, Diez MJ, Vinals FN (2001) Agroinoculation methods to screen wild Lycopersicon for resistance to Tomato yellow leaf curl virus. Journal Plant Pathology 83: 215-220

Pillen K, Ganal MW, Tanksley SD (1996) Construction of a high-resolution genetic map and YAC-contigs in the tomato Tm-2a region. Theoretical and Applied Genetics 93: 228-233

Pilowsky M, Cohen S (2000) Screening additional wild tomatoes for resistance to the whitefly-borne tomato yellow leaf curl virus. Acta Physiologiae Plantarum 22: 351-353

Raja P, Sanville BC, Buchmann RC, Bisaro, DM (2008) Viral genome methylation as an epigenetic defense against geminiviruses. Journal of Virology 82: 8997-9007

Rasul I (2012) Characterization of the Tm-2² locus of tomato and its durability (PhD thesis)

Ribeiro, SG, Lohuis H, Goldbach R, Prins M (2007) Tomato chlorotic mottle virus is a target of RNA silencing but the presence of specific short interfering RNAs does not guarantee resistance in transgenic plants. Journal of Virology 81: 1563-1573

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. Nature Biotechnology 29: 24-26

Rodríguez-Negrete E, Lozano-Durán R, Piedra-Aguilera A, Cruzado L, Bejarano ER, Castillo AG (2013) Geminivirus Rep protein interferes with the plant DNA methylation machinery and suppresses transcriptional gene silencing. New Phytologist 199: 464-475

Rojas MR, Jiang H, Salati R, Xoconostle-Cázares B, Sudarshana M, Lucas WJ, Gilbertson RL (2001) Functional analysis of proteins involved in movement of the monopartite begomovirus, Tomato yellow leaf curl virus. Virology 291: 110-125

Rouppe van der Voort J, Kanyuka K, van der Vossen E, Bendahmane A, Mooijman P, Klein-Lankhorst R, Stiekema W, Baulcombe D, Bakker J (1999) Tight physical linkage of the nematode resistance gene Gpa2 and the virus resistance gene Rx on a single segment introgressed from the wild species Solanum tuberosum subsp. andigena CPC 1673 into cultivated potato. Molecular Plant-Microbe Interactions 12: 197-206

Rubinstein G, Czosnek H (1997) Long-term association of tomato yellow leaf curl virus with its whitefly vector Bemisia tabaci: effect on the insect transmission capacity, longevity and fecundity. Journal of General Virology 78: 2683-2689

Rubio F, García-Martínez S, Alonso A, Grau A, Valero M, Ruiz JJ (2012) Introgressing resistance genes into traditional tomato cultivars: effects on yield and qualtity. Acta Horticulturae 935: 29-33

Sade D, Eybishtz A, Gorovits R, Sobol I, Czosnek H (2012) A developmentally regulated lipocalin-like gene is overexpressed in Tomato yellow leaf curl virus-resistant tomato plants upon virus inoculation, and its silencing abolishes resistance. Plant molecular biology 80: 273-287

Sade D, Brotman Y, Eybishtz A, Cuadros-Inostroza Á, Fernie AR, Willmitzer L, Czosnek H (2013) Involvement of the hexose transporter gene LeHT1 and of sugars in resistance of tomato to tomato yellow leaf curl virus. Molecular Plant 6: 1707-1710

Sade D, Shriki O, Cuadros-Inostroza A, Tohge T, Semel Y, Haviv Y, Willmitzer L, Fernie AR, Czosnek H, Brotman Y (2014) Comparative metabolomics and

transcriptomics of plant response to Tomato yellow leaf curl virus infection in resistant and susceptible tomato cultivars. Metabolomics 1-17

Sade D, Shriki O, Cuadros-Inostroza A, Tohge T, Semel Y, Haviv Y, Willmitzer L, Fernie AR, Czosnek H, Brotman Y (2015) Comparative metabolomics and transcriptomics of plant response to Tomato yellow leaf curl virus infection in resistant and susceptible tomato cultivars. Metabolomics: 1-17

Sahu PP, Puranik S, Khan M, Prasad M (2012) Recent advances in tomato functional genomics: utilization of VIGS. Protoplasma 249: 1017-1027

Sanfaçon H (2015) Plant Translation Factors and Virus Resistance. Viruses 7: 3392-3419

Sansregret R, Dufour V, Langlois M, Daayf F, Dunoyer P, Voinnet O, Bouarab K (2013) Extreme resistance as a host counter-counter defense against viral suppression of RNA silencing

Schmid M, Speiseder T, Dobner T, Gonzalez RA (2014) DNA virus replication compartments. Journal of virology 88: 1404-1420

Schmitz G, Tillmann E, Carriero F, Fiore C, Cellini F, Theres K (2002) The tomato blind gene encodes a MYB transcription factor that controls the formation of lateral meristems. Proceedings of th eNational Academy of Sciences 99: 1064–1069

Scholthof HB (2005) Plant virus transport: motions of functional equivalence. Trends in plant science 10: 376-382

Scott JW (2001) Geminivirus resistance derived from Lycopersicon chilense accessions LA 1932, LA 1938 and LA 2779. Tomato Breeders' Round Table Proceedings 2001. Antigua, Guatemala

Scott JW, Agrama HA, Jones JP (2004) RFLP-based analysis of recombination among resistance genes to fusarium wilt races 1, 2 and 3 in tomato. Journal of the American Society of Horticultural Science 129: 394-400

Scott JW (2007) Breeding for resistance to viral pathogens. In: Razdan MK, Mattoo AK (eds) Genetic improvement of solanaceous crops. Science Publisher, Inc., Enfield, NH, pp 447-474

Scott JW, Stevens MR, Barten JHM, Thome CR, Polston JE, Schuster DJ, Serra CA (1996) Introgression of resistance to whitefly-transmitted geminiviruses from Lycopersicon chilense to tomato. In: Bemisia 1995; taxonomy, biology, damage control, and management. D. Gerling and R.T. Mayer, eds. Intercept Press, Andover, UK pp. 357-367

Seah S, Telleen AC, Williamson VM (2007) Introgressed and endogenous Mi-1 gene clusters in tomato differ by complex rearrangements in flanking sequences and show sequence exchange and diversifying selection among homologues. Theoretical and Applied Genetics 114: 1289-1302

Seal S, vandenBosch F, Jeger M (2006) Factors influencing begomovirus evolution and their increasing global significance: implications for sustainable control. Critical reviews in plant sciences 25: 23-46

Shahid MS, Ikegami M, Waheed A, Briddon RW, Natsuaki KT (2014) Association of an alphasatellite with Tomato yellow leaf curl virus and Ageratum yellow vein virus in Japan is suggestive of a recent introduction. Viruses 6: 189-200

Shao F, Lu S (2014) Identification, molecular cloning and expression analysis of five RNA-dependent RNA polymerase genes in Salvia miltiorrhiza. PloS one 9: e95117

Sharma A, Zhang L, Nio-Liu D, Ashrafi H, Foolad MR (2008) A solanum lycopersicum solanum pimpinellifolium linkage map of tomato displaying genomic locations of R-genes, RGAs, and candidate resistance/defense-response ESTs. International Journal of Plant Genomics 2008

Sharma VK, Basu S, Chakraborty S (2015) RNAi mediated broad-spectrum transgenic resistance in Nicotiana benthamiana to chilli-infecting begomoviruses. Plant cell reports: 1-11

Shearer LA, Anderson LK, de Jong H, Smit S, Goicoechea JL, Roe BA, Hua A, Giovannoni JJ, Stack SM (2014) Fluorescence in situ hybridization and optical mapping to correct scaffold arrangement in the tomato genome. G3: Genes | Genomes | Genetics 4: 1395-1405

Sim SC, Durstewitz G, Plieske J, Wieseke R, Ganal MW, van Deynze A, Hamilton JP, Buell CR, Causse M, Wijeratne S et al (2012) Development of a large snp genotyping array and generation of high-density genetic maps in tomato. PLoS ONE 7: e40563

Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, Van der Lee T, Bleeker M, Onstenk J, de Both M, Haring Jurriaan Mes M, Cornelissen B, Zabeau M, Vos P (1998) Dissection of the fusarium I2 gene cluster in tomato reveals six homologs and one active gene copy. The Plant Cell 10: 1055–1068

Skinner ME, Uzilov AV, Stein LD, Mungall CJ, Holmes IH (2009) JBrowse: A next-generation genome browser. Genome Research 19: 1630-1638

Solovyev V, Kosarev P, Seledsov I, Vorobyev D (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biology 7 Suppl 1: 10.1-10.12

Sorel M, Svanella-Dumas L, Candresse T, Acelin G, Pitarch A, Houvenaghel M, German-Retana S (2014) Key mutations in the cylindrical inclusion involved in Lettuce mosaic virus adaptation to eIF4E-mediated resistance in lettuce. Molecular Plant-Microbe Interactions 27: 1014-1024

Stewart Jr CN, Via LE (1993) A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. BioTechniques 14: 748-750

Suwabe K, Suzuki G, Nunome T, Hatakeyama K, Mukai Y, Fukuoka H, Matsumoto S (2012) Microstructure of a Brassica rapa genome segment homoeologous to the resistance gene cluster on Arabidopsis chromosome 4. Breeding Science 62: 170-177

Szinay D, Bai Y, Visser R, de Jong H (2010) FISH applications for genomics and plant breeding strategies in tomato and other Solanaceous crops. Cytogenetic and Genome Research 129: 199-210

Szinay D, Wijnker E, van den Berg R, Visser RG, de Jong H, Bai Y (2012) Chromosome evolution in Solanum traced by cross-species BAC-FISH. New Phytologist 195: 688-698

Tam SM, Hays JB, Chetelat RT (2011) Effects of suppressing the DNA mismatch repair system on homeologous recombination in tomato. Theoretical and Applied Genetics 123: 1445-1458

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28: 2731-2739

Tang X, Szinay D, Lang C, Ramanna MS, van der Vossen EAG, Datema E, Klein Lankhorst R, de Boer J, Peters SA, Bachem C, Stiekema W, Visser RGF, de Jong H, Bai Y (2008) Cross-species bacterial artificial chromosome-fluorescence in situ hybridization painting of the tomato and potato chromosome 6 reveals undescribed chromosomal rearrangements. Genetics 180: 1319-1328

Tena G, Boudsocq M, Sheen J (2011) Protein kinase signaling networks in plant innate immunity. Current opinion in plant biology 14: 519-529

The 100 Tomato Genome Sequencing Consortium (2014) Exploring genetic variation in the tomato (Solanum section Lycopersicon) clade by whole-genome sequencing. The Plant Journal 80: 136-148

The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475: 189-195

The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635-641

Thivierge K, Cotton S, Dufresne PJ, Mathieu I, Beauchemin C, Ide C, Fortin MG, Laliberté J-F (2008) Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. Virology 377: 216-225

Thornton B, Basu C (2011) Real-time PCR (qPCR) primer design using free online software. Biochemistry and Molecular Biology Education 39: 145-154

Tikunov Y, Lommen A, De Vos CHR, Verhoeven HA, Bino RJ, Hall RD, Bovy AG (2005) A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. Plant Physiology 139: 1125-1137

Tikunov YM, de Vos RC, Paramás AMG, Hall RD, Bovy AG (2010) A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. Plant Physiology 152: 55-70

Tikunov YM, Laptenok S, Hall RD, Bovy A, de Vos RCH (2012) MSClust: A tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. Metabolomics 8: 714-718

Uma B, Rani TS, Podile AR (2011) Warriors at the gate that never sleep: non-host resistance in plants. Journal of plant physiology 168: 2141-2152

Van Brunschot S, Persley D, Geering A, Campbell P, Thomas J (2010) Tomato yellow leaf curl virus in Australia: distribution, detection and discovery of naturally occurring defective DNA molecules. Australasian Plant Pathology 39: 412-423

van der Knaap E, Sanyal A, Jackson SA, Tanksley SD (2004) High-resolution fine mapping and fluorescence in situ hybridization analysis of sun, a locus controlling tomato fruit shape, reveals a region of the tomato genome prone to DNA rearrangements. Genetics 168: 2127-2140

Van Ooijen JW (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genetics research 93: 343-349

Vanitharani R, Chellappan P, Pita JS, Fauquet CM (2004) Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. Journal of Virology 78: 9487-9498

Varma A, Malathi V (2003) Emerging geminivirus problems: A serious threat to crop production. Annals of Applied Biology 142: 145-164

Verlaan MG, Hutton SF, Ibrahem RM, Kormelink R, Visser RGF, Scott JW, Edwards JD, Bai Y (2013) The tomato yellow leaf curl virus resistance genes Ty-1 and Ty-3 are allelic and code for DFDGD-class RNA-dependent RNA polymerases. PLoS Genetics 9: e1003399

Verlaan MG, Szinay D, Hutton SF, de Jong H, Kormelink R, Visser RGF, Scott JW, Bai Y (2011) Chromosomal rearrangements between tomato and Solanum chilense hamper mapping and breeding of the TYLCV resistance gene Ty-1. The Plant Journal 68: 1093-1103

Vidavski F (2007) Exploitation of resistance genes found in wild tomato species to produce resistant cultivars; Pile up of Resistant Genes. In Tomato Yellow

Leaf Curl Virus Disease, H. Czosnek, ed (Springer Netherlands), pp. 363-372

Vidavski F, Czosnek H, Gazit S, Levy D, Lapidot M (2008) Pyramiding of genes conferring resistance to Tomato yellow leaf curl virus from different wild tomato species. Plant Breeding 127: 625-631

Víquez-Zamora M, Vosman B, van de Geest H, Bovy A, Visser RGF, Finkers R, van Heusden AW (2013) Tomato breeding in the genomics era: Insights from a SNP array. BMC Genomics 14: 354

Voinnet O, Pinto YM, Baulcombe DC (1999) Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses of plants. Proceedings of the National Academy of Sciences 96: 14147-14152

Voorrips RE (2002) Mapchart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93: 77-78

Voorrips RE, Verkerke W, Finkers R, Jongerius R, Kanne J (2000) Inheritance of taste components in tomato. Acta Physiologiae Plantarum 22: 259-261

Wang H, Hao L, Shung CY, Sunter G, Bisaro DM (2003) Adenosine Kinase Is Inactivated by Geminivirus AL2 and L2 Proteins. The Plant Cell Online 15: 3020-3032

Wang H, Buckley KJ, Yang X, Buchmann RC, Bisaro DM (2005) Adenosine Kinase Inhibition and Suppression of RNA Silencing by Geminivirus AL2 and L2 Proteins. Journal of Virology 79: 7410-7418

Wang A, Krishnaswamy S (2012) Eukaryotic translation initiation factor 4E mediated recessive resistance to plant viruses and its utility in crop improvement. Molecular plant pathology 13: 795-803

Wang B, Li F, Huang C, Yang X, Qian Y, Xie Y, Zhou X (2014) V2 of tomato yellow leaf curl virus can suppress methylation-mediated transcriptional gene silencing in plants. Journal of General Virology 95: 225-230

Wartig L, Kheyr-Pour A, Noris E, De Kouchkovsky F, Jouanneau F, Gronenborn B, Jupin I (1997) Genetic analysis of the monopartite tomato yellow leaf curl geminivirus: roles of V1, V2, and C2 ORFs in viral pathogenesis. Virology 228: 132-140

Whitham SA, Anderberg RJ, Chisholm ST, Carrington JC (2000) Arabidopsis RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock–like protein. The Plant Cell 12: 569-582

Willmann MR, Endres MW, Cook RT, Gregory BD (2011) The Functions of RNA-Dependent RNA Polymerases in Arabidopsis. The Arabidopsis Book, e0146

Wu F, Tanksley SD (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genomics 11: 182

Wu J, Dai F, Zhou X (2006) First report of Tomato yellow leaf curl virus in China. Plant Disease 90: 1359-1359

Yadav RK, Chattopadhyay D (2011) Enhanced viral intergenic region-specific short interfering RNA accumulation and DNA methylation correlates with resistance against a geminivirus. Molecular Plant-Microbe Interactions 24: 1189-1197

Yamaji Y, Maejima K, Komatsu K, Shiraishi T, Okano Y, Himeno M, Sugawara K, Neriya Y, Minato N, Miura C (2012) Lectin-mediated resistance impairs plant virus infection at the cellular level. The Plant Cell 24: 778-793

Yang X, Caro M, Hutton SF, Scott JW, Guo Y, Wang X, Rashid MH, Szinay D, de Jong H, Visser RGF, Bai Y, Du Y (2014) Fine mapping of the tomato yellow leaf curl virus resistance gene Ty-2 on chromosome 11 of tomato. Molecular Breeding 34: 749-760

Yang X, Xie Y, Raja P, Li S, Wolf JN, Shen Q, Bisaro DM, Zhou X (2011) Suppression of methylation-mediated transcriptional gene silencing by βC1-SAHH protein interaction during geminivirus-betasatellite infection. PLoS Pathogens 7: e1002329

Zamir D, Ekstein Michelson I, Zakay Y, Navot N, Zeidan M, Sarfatti M, Eshed Y, Harel E, Pleban T, van Oss H (1994) Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, Ty-1. Theoretical and Applied Genetics 88: 141-146

Zhou X (2013) Advances in understanding begomovirus satellites. Annual review of phytopathology 51: 357-381

Zhu J, Gopinath K, Murali A, Yi G, Hayward SD, Zhu H, Kao C (2007) RNA-binding proteins that inhibit RNA virus infection. Proceedings of the National Academy of Sciences 104: 3129-3134

Summary

Tomato Yellow Leaf Curl Virus (TYLCV) is one of the most severe tomato diseases, ranked as the third most important viral disease based on its scientific and economic importance (Scholthof et al. 2011). Resistance to TYLCV is currently available in commercial tomato breeding lines, mainly based on the introgression of the resistance genes *Ty-1/Ty-3* and *Ty-2*. This thesis aimed to characterize these resistance genes/loci and to identify and map new alternative TYLCV resistance genes, in order to assist introgression breeding for more durable forms of TYLCV resistance.

Ty-1 and *Ty-3*, derived from *Solanum chilense* accessions LA1969 and LA2779 respectively, are alleles of one gene coding for an RNA-dependent RNA polymerase (*RDR*$_{\gamma}$) conferring resistance to TYLCV. **Chapter 2** aimed to identify and examine allelic forms of this gene in a panel of wild tomato species, landraces and cultivars. Using a VIGS approach, we show that the TYLCV resistance in *S. chilense* accessions LA1932 and LA1938 is conferred by an allele of *Ty-1/Ty-3*; introgression lines derived from these accessions rendered to susceptibility after silencing the gene. Our results also showed that one accession might carry more than one TYLCV resistance locus, as for the case of *S. chilense* LA1971. In order to analyse the genetic variation of the gene, we examined the cDNA sequences of the *RDR* in our complete panel. Three *Ty-1/Ty-3* specific amino acids shared by seven TYLCV-resistant *S. chilense* accessions (or derived lines) were identified when compared to the susceptible allele; such polymorphisms can be used to develop *Ty-1/Ty-3* specific markers. Instead, the characteristic DFDGD motif in the catalytic domain of the *RDR* is conserved among all tested tomato lines and wild relatives, indicating that SNPs in this motif might have a negative effect on plant fitness. To further examine differences between susceptible and resistant *Ty-1/Ty-3* alleles, we analysed the expression of the *RDR* from different *S. chilense*-derived lines, accessions and related species. High expression of the *RDR* is necessary for a resistance response, but is not exclusively responsible for *Ty-1/Ty-3*-mediated resistance.

The TYLCV resistance gene *Ty-2* derives from *S. habrochaites* accession B6013, and has been mapped to the long arm of chromosome 11. In order to fine-map the *Ty-2* gene, we screened nearly 11,000 plants for recombinants in our *Ty-2* target region as described in **Chapter 3.** Our molecular marker analysis together with TYLCV disease tests allowed us to locate the *Ty-2* gene in a region of approximately 300kb, between markers UP8 (51.344 Mbp) and M1 (51.645 Mbp). These analyses also evidenced a region with severe suppression of recombination

within the *Ty-2* introgression (of around 200kb) between markers C2_At1g07960 (51.387 Mbp) and C2_At3g52090 (51.605 Mbp). Fluorescence in situ hybridization (FISH) failed to elucidate our hypothesis that such phenomenon was caused by chromosomal rearrangements. Since we were therefore unable to shorten the region using our mapping strategy, expression and functional analysis of a set of genes predicted in the region were performed to assist us in the identification of candidate genes for *Ty-2*. Our results suggested the possible implication of two candidate genes (annotated as an Elongation factor 1-alpha and a DNA-directed RNA polymerase II) in the *Ty-2*-mediated resistance. Furthermore, the involvement of NBS-LRR genes present in the region was likely out ruled, since VIGS-silencing the genes in our *Ty-2* line did not compromise the resistance.

Further analyses of the *Ty-2* introgression region were performed in order to visualize its genome structure. In **Chapter 4** we compared the *Ty-2* region of *S. habrochaites* accession LYC4 (using the draft *de novo* sequence) with the *S. lycopersicum* sequence. Markers UP8 and C2_At3g52090 were contained in one *S. habrochaites* LYC4 scaffold at a distance of 26kb, markers otherwise located at a 247kb distance according to the *S. lycopersicum* Heinz sequence. Alignment of the *Ty-2* region from both sequences revealed an inversion of around 200kb in the central part. Furthermore, we analysed F_2 populations of different *S. habrochaites* intraspecific crosses in order to check if the suppression of recombination was also occurring for these crosses. Recombinant screening of F_2 individuals suggested a marker order similar to that of *S. habrochaites* LYC4. To confirm such an inversion in the target region, a BAC library of a *Ty-2* introgression line was made and a BAC clone containing the marker UP15 (51.381) was sequenced. When aligned to the *S. lycopersicum* Heinz sequence, the BAC-clone-sequence homologous regions were split partly to the upper side of the inversion, and another to its lower end; no sequences aligned to the inversion region. Altogether, our results confirm that the cause of the suppression of recombination previously described in our *Ty-2* introgression line is due to an inversion between *S. habrochaites* and *S. lycopersicum*.

This thesis also aimed at the identification of new TYLCV resistance sources and the underlying genes conferring resistance. In order to map the identified resistance in *S. pimpinellifolium* accession G1.1554, a RIL population consisting of 100 lines (derived from the cross of aforementioned accession and *S. lycopersicum* 'Moneymaker') was genotyped using a SNP array in **Chapter 5**. A total number of 1974 polymorphic SNPs were identified and used to develop a linkage map. A TYLCV disease test was performed on 81 RILs, which were scored according to

their symptom development up to 45 days after virus inoculation. Five RILs were classified as resistant (symptom-free) while 72 RILs were susceptible to the virus. QTL mapping analysis revealed two putative QTLs associated with the resistance, located in chromosomes 3 and 11 (named *qTy-p3* and *qTy-q11*, respectively). Further re-sequencing of a subset of lines (60 RILs) allowed us to saturate the QTL regions to improve our linkage groups and to *in silico* genotype the RILs subset. Such enriched dataset allowed us to confirm the identified QTLs and to refine the mapped regions. The effect of both QTLs together accounted for 28% of the phenotypic variation; however, 14 RILs homozygous for the *S. pimpinellifolium* allele in the QTL regions were found to be susceptible, evidencing the presence of additional genetic factor(s) undetected in our study. In addition, untargeted metabolic profiling of the RILs revealed differences between resistant and susceptible lines mainly associated with sucrose and flavonoid glycosides. Our study shows that a combination of different ~omics approaches can provide useful information for mapping and characterizing TYLCV resistance genes.

In **Chapter 6** the results provided in this thesis are summarized and brought into a broader context, and discussed as part of a long-term strategy to contribute with different layers of immunity to TYLCV as well as their implications for introgression breeding.

Acknowledgments

It is hard to believe that the moment of writing the acknowledgements for my PhD thesis has come. Seems to be the easiest part of a thesis, but is certainly very difficult to write the final words of one of the most important chapters of my life. This long walk has been full of many, many people that have enriched this thesis and my life in all senses, professionally and personally: professors, teachers, colleagues from here and there, friends, family.

I would first like to express my deepest gratitude to my supervisor, Dr. Yuling Bai. The journey together began when I started my MSc thesis under your supervision; your trust in my potential as a researcher allowed me to continue my PhD studies in the same topic, TYLCV resistance. Our discussions were always challenging, allowing me to develop a critical sense of analysis. Thanks Yuling for your support at the beginning, during and especially at the end of my PhD, for offering me your support and for encouraging me to pursue my professional ambitions.

To my promotor, Dr. Richard Visser, thanks for your support to start my PhD studies at the Plant Breeding department of Wageningen University. Your support and guidance during the process and especially at the writing stage of my thesis allowed me to conclude my PhD in good timing, many thanks!

I would also like to thank Dr. Richard Kormelink for the valuable suggestions and always passionate and energetic discussions. Richard and his Virology team broadened my scientific view into an in-depth analysis, especially from the (crucial) pathogen side. Many thanks Richard!

I need to deeply acknowledge and thank one of the scientific guides and key persons in our research group. Her guidance, input, suggestions and support were of great importance for the development of this thesis. She is a 1st-class scientist and the support of the group in many other senses too. I think I do not speak only for myself when saying how important your presence is for the group; she is also my good friend, thanks Anne-marie Wolters!

Many colleagues were also a crucial element for the completion of my thesis. Sjaak van Heusden, thanks for your willingness to work together and your 24/7 positive attitude that is admirable in you. Thanks to Richard Finkers, Dora Szinay and Hans de Jong from Wageningen University; Samuel Hutton, John Scott and Xiaohui Yang, from the University of Florida; and Yanmei Guo, Xiaoxuan Wang, Yongchen Du, Shufang Dong and Jianchang Gao, from the IVF-CAAS Institute in

China. Our collaboration together yielded successfully published, nice pieces of work. Thank you for both your input, and trust.

To our day-by-day supporters, our important 'tool-box' colleagues: Dorette and Danny; Annelies, Gert and Johan, thank you for the smiles and your good vibes, and for the always-available help. Wendy thanks for the long recombinant-screening journeys. Marc, a big thanks to you, for investing your time in helping us to make sense of some of the challenging stories presented in this book. Marjon, not only your technical expertise was very important for the project, but also your input in our discussions, your willingness to help and go deeper in our approach. Your strength and energy during this period were an inspiration. Thanks a lot. Fien, also thank you for your patience and help when trying to figure out the enormous pedigrees, when trying to find ancient seed bags and for the numerous selfing, crosses and breeding plans.

To my colleagues and friends from Unifarm: Bertus and Andre, your help in my experiments was essential for each chapter of this thesis. Maarten and Teus, thanks a lot for being there to help, always with a big, contagious smile. Casper, thanks for the care and love to my tomato plants, and for the laughs and good moments! My dear friend Henk, you were my greenhouse angel! Thanks for your help, for the nice conversations, sharing of stories and your advices in tough, hard moments. Going to the greenhouse only meant pleasant moments, thanks to you all!

To my colleagues from the Virology department: Dick Lohuis, thanks a lot for your help and patience when asking for advice! Corien, Athos, Dryas, Paulus, thanks a lot for the rich discussions during the lunch meetings and other occasions.

My students were not only part of this research project but also part of my development as a professional. Thanks to Harun and Anita; Yohanis and his whole-day smile; Ron and his contagious enthusiasm. Thanks to Anindita for her careful approach in the lab, and loving attitude towards her colleagues. Thanks to Yucheng and Sayantani, and the great sense of humor! And to Andrea, my last student, who is starting a new journey in the Plant Breeding world, all the success!

To my Plant Breeding colleagues: first of all, the secretaries, Nicole, Janneke and Letty. I hope you realise the importance of your presence in a PhD-student journey. Thank you for all your help and for the numerous laughs together!

Then to those people in the same ship towards a common goal: Yan Zhe, thank you for the nice moments in China and here, back in Wageningen. Thanks a lot for your support and help in the experiments during this short yet pleasant overlapping time of our PhD's. I am confident that you will be successful in your professional life, and in your personal future!

Marcela, thanks a lot for the time we spent together during the development of one of the chapters in this thesis. Thanks for the discussions, ideas and especially for being the driving force to make it happen. But most importantly, for your friendship and for your encouraging words to be a woman full of strength. Mil gracias Marce!

Thanks to my lunch-break buddies for the great company, for the moments full of joy and laughs and for our usual 'bizarre' talks: Anne, Xuan, Tim, Peter Dihn, Dália, Gilda; to Valentina, full of unusual stories that will make every conversation an amusing experience. Thanks for all the love dear Vale! Atiyeh, thanks for your friendship my dear, you have an intense inner strength which I am sure will take you very far. Christos, you probably were my first friend in the department. Your particular, unique sense of humour makes everyone around you have such a great time! Thanks for your objective comments when I asked you for scientific feedback, and for the great moments and support during these years. I hope we keep our friendship for much, much longer, wherever we are. Jordi, Cynara, Marine, Xiao, Manos, Behzad, thanks for the daily smiles! Kaile, Ageeth and Ernest, thanks also for the nice moments during my last days in PBR in our trip to Bordeaux. Great thanks my friends!

My current office mates: Peter, Charlie and Ehsan, thanks for the great environment; this was not an office but 'the place to be'. Thanks Mas for your company and friendship; since the very first day you came into our office you lightened up our lives, not only with your sweet presents but also with your contagious laughter, your positivism and open friendship. You have a special light, always remember it! My dear Michi, it will be hard to mention so many things we lived together during these years. We have been office mates for long, but that is the least important part of our journey together. We were colleagues, discussion partners, we shared so many ideas for our own research projects, but also our happy and tough moments during this exciting yet challenging research path. But most importantly, you are one of my best friends and your presence during these years is invaluable. We have shared so many beautiful moments, moments I

cherish deeply in my heart. Thanks, many thanks my friend, I could not have done it without you.

Thanks to my former colleagues, now in their respective new successful lives. Dong Li, thanks for your technical advices, for the Chinese lessons and for all the funny, great moments. Nuria, thanks for the time we shared in and outside Plant Breeding, I know you are doing great now and I would not expect less.

Very special thanks to Maarten Verlaan. I started my journey in PBR with Maarten as my MSc Thesis supervisor; later on we became PhD colleagues and now we are colleagues, once again. Many of the things I know now came from you. Thanks for the time you spent in introducing me into the TYLCV-resistance world, for the nice discussions, our work collaboration and for the stressful moments buying concert tickets, thanks a lot Maarten!

Many thanks to my friend Olga Julián, with whom I spent uncountable hours together grinding samples, doing dozens of PCRs and many other lab duties, and shared many interesting discussions during your stay in PBR. Besides being a good professional, you are a good friend, a person with an admirable perseverance who always knows what you want and how to get it. I know many more years of our special friendship will come, thank you very much Olguita.

To my family in Wageningen, my friends: Alejandra and Claudia, my former housemates. The great chemistry we had living together indirectly helped me to successfully finish my PhD, since going to the 'Peruvian Embassy' meant to go home, to feel at home. I love you! To my friends Lorena, Natalia, Alvaro, Cristian and Jennie, Elizabeth and Timo, thanks a lot for offering me your friendship, and for sharing so many great moments. And to the ones that are not in Wageningen anymore but were an important piece of my life during my PhD journey: Martincito and Golda, my great, good friends, you know your support and our time together in Wageningen made all this as special as it can be. Dunia, you were the first Peruvian I met in PBR and the chemistry was immediate. Thanks for your support, for your wise advices and for your friendship. Verónica, Claudia V., Lore G., Yván, Francisco, Gabriela, Ricardo, Víctor Hugo, Luchito, you also helped to make this PhD journey a very pleasant one. And special thanks to my friend Carlos, who very kindly helped me to build up the cover of this thesis. Many thanks Carlitos!

And to my friends in Peru, that always receive me with an intact, true love. To my life friends/sisters: Ana María, Jassmin, Vanessa, Fiorela, Claudia U. and Maju,

this thesis also belongs to you. And also to Carla, Viqui, Roxana, Silvia, Patricia, Fabiola, Ernesto, Manuel, Carlos, Iván, Jossepi… and so many more. Your love, every time I saw you during holidays gave me the necessary strength to continue until the end of this PhD road.

My central column during these years and all the years of my life: my family. Mi mamá Berta: Tobicita, siempre has sido mi ángel y soporte en todo lo que he decidido emprender, me has dado apoyo, fuerza y coraje en todos mis proyectos de vida. Sólo tus palabras me dan el aliento necesario para sentir que soy capaz de hacer imposibles. Mi papá Osito, desde siempre has estado a mi lado, entregando todo en beneficio nuestro, apoyándonos en lo posible e imposible. Como dice nuestra canción, 'Yo te seguiré' es una promesa que estará vigente por siempre. Mi hermano Mauricio, tu ímpetu, energía y fuerza ahora se han convertido en tu éxito, y yo no podría estar más orgullosa de tí. Esta tesis es el fruto de mi trabajo pero está íntegramente dedicada a ustedes, es el final de un largo trabajo… mío, y suyo. Mil gracias, los adoro… hasta el infinito y más.

And to my Fernando, thanks a lot for your help and support, your encouraging words when the deadlines for this thesis were approaching and I was about to fall, and a word, a joke (or a dinner) would cheer me up… and because the load of this whole journey was lightened down with you by my side, muchas gracias esinito.

Education certificate

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: Carolina Myluska Caro Rios
Date: 7 December 2015
Group: Laboratory of Plant Breeding
University: Wageningen University & Research Centre

1) Start-up phase	date
► **First presentation of your project**	
Improved and durable resistance: tomato genes and their associated mechanisms underlying the TYLCV resistance	Apr 08, 2013
► **Writing or rewriting a project proposal**	
Writing draft of proposal 'Improved and durable resistance: tomato genes and their associated mechanisms undelying the TYLCV resistance'	2012
► **Writing a review or book chapter**	
"Resistance to Tomato Yellow Leaf Curl Virus in tomato", submitted to Molecular Breeding	Dec 2015
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*8.5 credits**

2) Scientific Exposure	date
► **EPS PhD student days**	
EPS PhD student day (Get2Gether), Soest (NL)	Jan 29-30, 2015
► **EPS theme symposia**	
EPS theme 2 Symposium 'Interactions between Plants and Biotic Agents' & Willie Commelin Scholten Day, Wageningen University	Feb 10, 2012
EPS theme 2 Symposium 'Interactions between Plants and Biotic Agents' & Willie Commelin Scholten Day, Utrecht University	Feb 20, 2015
► **NWO Lunteren days and other National Platforms**	
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 02-03, 2012
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 22-23, 2013
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 14-15, 2014
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 13-14, 2015
► **Seminars (series), workshops and symposia**	
Plant Breeding Research Day "Next generation sequencing – What's in it for me"	Feb 28, 2012
Open day '100 Years Plant Breeding', Wageningen	Aug 31, 2012
Seminar Prof. Sir D.C.Baulcombe 'Plant versus virus: defense, counter defense and counter counter defense'	Oct 10, 2012
Plant Breeding Research Day	Sep 24, 2014
Symposium 'All-inclusive breeding: Integrating high throughput science'	Oct 16, 2014
Symposium 'Crop Pathology and Plant-Microbe interactions'	May 08, 2015
Plant Breeding Research Day	Sep 29, 2015
Mini-symposium farewell address of Prof. J. H. S. G. M de Jong	Oct 01, 2015
► **Seminar plus**	
► **International symposia and congresses**	
Next Generation Plant Breeding conference (Ede, The Netherlands)	Nov 11-14, 2011
SOL 2013 (Beijing, China)	Oct 13-17, 2013
SOL 2015 (Bordeaux, France)	Oct 25-29, 2015
► **Presentations**	
Summer school 'Natural Variation of plants' (Poster)	Aug 21-24, 2012
SOL 2013 (Beijing, China) (Poster)	Oct 13-17, 2013
Training course (Peru) "Mejoramiento Genético en Cultivos Alimenticios a través del Uso de Herramientas Moleculares". Oral Presentation	Dec 02-13, 2013
ALW meeting 'Experimental Plant Sciences', Lunteren (Oral)	Apr 14-15, 2014
ALW meeting 'Experimental Plant Sciences', Lunteren (Oral)	Apr 13-14, 2015
SOL 2015 (Bordeaux, France) (Oral)	Oct 25-29, 2015
► **IAB interview**	
► **Excursions**	
EPS council excursion to vegetable breeding company Rijk Zwaan	Sep 27, 2013
Subtotal Scientific Exposure	*16.2 credits**

3) In-Depth Studies	date
► **EPS courses or other PhD courses**	
EPS course: Mixed model based QTL mapping in GenStat	14-16 May 2012
Summer school 'Natural Variation of plants'	21-24 August 2012
Transplant workshop: Exploiting and understanding Solanaceous genomes	13-14 October 2014
EPS course: Genome assembly	28-29 April 2015
► **Journal club**	
► **Individual research training**	
Subtotal In-Depth Studies	*3.3 credits**

4) Personal development		*date*
►	**Skill training courses**	
	WGS course Techniques for Writing and Presenting a Scientific Paper	Apr 01-04, 2014
	WGS course Career Assessment	Apr 23, 2015
	WGS course Entrepreneurship in and outside science	Sep 15, 22, Oct 13, 2015
	WGS course Last stretch of the PhD programme	May 22, 2015
►	**Organisation of PhD students day, course or conference**	
	Training course (Peru) "Mejoramiento Genético en Cultivos Alimenticios a través del Uso de Herramientas Moleculares"	Dec 02-13, 2013
►	**Membership of Board, Committee or PhD council**	
	Subtotal Personal Development	*4.0 credits**

TOTAL NUMBER OF CREDIT POINTS*	**32.0**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

** A credit represents a normative study load of 28 hours of study.*