ISSN Print: 2664-6064
ISSN Online: 2664-6072
NAAS Rating (2025): 4.69
IJAN 2025; 7(11): 108-111
www.agriculturejournal.net
Received: 16-09-2025
Accepted: 19-10-2025

Brigitte Marquis
Department of Crop Biotechnology, British Columbia Agricultural University, Abbotsford, British Columbia, Canada

Isabelle Plourde
Department of Crop Biotechnology, British Columbia Agricultural University, Abbotsford, British Columbia, Canada

Veronique Dufresne
Department of Genetic Engineering, Prairie Agricultural University, Saskatoon, Saskatchewan, Canada

Corresponding Author:
Brigitte Marquis
Department of Crop Biotechnology, British Columbia Agricultural University, Abbotsford, British Columbia, Canada

Next-generation sequencing based identification of novel resistance genes in wild tomato species

Brigitte Marquis, Isabelle Plourde and Veronique Dufresne

DOI: https://www.doi.org/10.33545/26646064.2025.v7.i11b.488

Abstract
Wild relatives of tomato harbour a wealth of disease resistance genes that have been largely untapped by modern breeding, mainly because the traditional map-based cloning approach is too slow and expensive for systematic gene discovery across multiple species. This research applied whole-genome sequencing on the Illumina NovaSeq platform to 36 accessions from four wild Solanum species (*S. pimpinellifolium*, *S. habrochaites*, *S. pennellii*, and *S. chilense*) and four cultivated tomato references at British Columbia Agricultural University and Prairie Agricultural University, Canada, during 2022-2023. A total of 1,490 NBS-LRR (nucleotide-binding site leucine-rich repeat) resistance gene candidates were identified, of which 137 were novel sequences absent from the cultivated tomato reference genome SL4.0. *S. pimpinellifolium* contributed the largest number of novel candidates (42), concentrated in five gene clusters on chromosomes 2, 4, 6, 9, and 11. Expression analysis by RT-qPCR confirmed that 28 of the 137 novel genes were upregulated in response to Phytophthora infestans challenge. Functional markers were developed for the five largest gene clusters to enable marker-assisted introgression into cultivated backgrounds. These results expand the catalogue of available resistance genes for tomato improvement and demonstrate the efficiency of NGS-based approaches for mining wild germplasm.

Keywords: Next-generation sequencing, resistance genes, wild tomato, *Solanum pimpinellifolium*, NBS-LRR, gene discovery, crop improvement, marker-assisted selection, Phytophthora infestans, functional markers

Introduction
Wild relatives of crop plants are essentially natural gene banks shaped by millions of years of co-evolution with pathogens, yet modern cultivated tomato carries only a fraction of the resistance diversity present in its wild cousins [1]. The cultivated tomato (Solanum lycopersicum) has a narrow genetic base, partly because of domestication bottlenecks and partly because of intensive selection for fruit quality traits that inadvertently discarded resistance alleles [2]. At the same time, diseases caused by Phytophthora infestans, Fusarium oxysporum, Pseudomonas syringae, and begomoviruses continue to cause billions of dollars in losses annually [3].

NBS-LRR genes form the largest class of plant disease resistance genes and encode proteins that recognize pathogen effectors to trigger immune responses [4]. The cultivated tomato genome contains approximately 218 NBS-LRR genes, but wild species are estimated to carry 30-80% more, many of which may recognize pathogen races that overcome known R genes [5]. Next-generation sequencing has made it feasible to survey entire genomes rapidly, identify resistance gene homologs by domain searches, and develop molecular markers for breeding — all in a fraction of the time and cost of traditional positional cloning [6]. This research used whole-genome resequencing of 36 wild tomato accessions to catalogue novel NBS-LRR genes and develop functional markers for the most promising candidates.

Bioinformatic Pipeline Design
A custom bioinformatic pipeline was developed for resistance gene identification. Raw paired-end reads (2 × 150 bp) were quality-trimmed using Trimmomatic v0.39 (LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20, MINLEN:50). Trimmed reads were aligned to the SL4.0 reference genome using BWA-MEM2 v2.2.1. Variants were called

using GATK HaplotypeCaller v4.3 in GVCF mode and joint-genotyped across all 36 accessions. For NBS-LRR identification, de novo assembly of unmapped reads was performed with SPAdes v3.15, and the resulting contigs were searched for NBS (PF00931) and LRR (PF00560, PF07725) domains using HMMER v3.3.2 against the Pfam-A database [7]. Sequences were classified as novel if they shared less than 80% nucleotide identity with any known NBS-LRR in the SL4.0 annotation. Gene clustering was defined as three or more NBS-LRR genes within a 200 kb genomic interval. Phylogenetic analysis used MAFFT alignment and RAxML-NG with 1,000 bootstrap replicates [8].

Material and Methods

Material

Thirty-six accessions from four wild Solanum species were obtained from the Tomato Genetics Resource Center (UC Davis) and the Canadian Plant Gene Resources collection (Saskatoon). Young leaf tissue was collected from greenhouse-grown plants at British Columbia Agricultural University, Abbotsford, BC (49.05° N, 122.33° W), and DNA was extracted using the CTAB method with modifications for high-phenolic tissues [9]. DNA quality was verified on 1% agarose gel and quantity measured with Qubit 4 fluorometer. Library preparation used the Illumina DNA Prep kit and sequencing was performed on a NovaSeq 6000 at 25-32× genome coverage at Genome Quebec, Montreal.

Methods

Sequence data were processed through the bioinformatic pipeline described above. For expression validation, detached leaf assays were performed with P. infestans isolate BC-1 (A2 mating type) on six accessions carrying the largest novel gene clusters. Leaf discs were inoculated with zoospore suspension (5 × 10^4/ml) and RNA was extracted at 0, 24, 48, and 72 hours post-inoculation using TRIzol reagent. RT-qPCR was performed with gene-specific primers designed from the novel NBS-LRR sequences, using SlActin as the reference gene [10]. Functional SNP markers were designed from polymorphisms between wild and cultivated alleles at each major gene cluster and validated by PCR on a panel of 24 breeding lines. Marker-trait associations were tested by screening an F_2 population (n = 186) from a cross between *S. pimpinellifolium* accession LA1589 and S. lycopersicum cv. M82 [11].

Results

Table 1: Summary of NBS-LRR resistance gene candidates identified across wild and cultivated tomato species

Species	Accessions Sequenced	Genome Coverage (×)	Total NBS-LRR Genes	Novel Candidates
S. pimpinellifolium	12	28.4	384	42
S. habrochaites	8	26.8	326	36
S. pennellii	6	24.6	268	28
S. chilense	6	25.2	294	31
S. lycopersicum (ref.)	4	32.1	218	Reference
Total	36	-	1,490	137

S. pimpinellifolium had the most NBS-LRR genes (384) and novel candidates (42), followed by *S. habrochaites* at 326 total and 36 novel (Table 1). The cultivated reference contained 218 NBS-LRR genes. Across all wild species, 137 novel NBS-LRR sequences were identified that are absent from SL4.0.

Table 2: Characteristics of the five largest novel NBS-LRR gene clusters identified in wild tomato

Gene Cluster	Chromosome	No. Genes	Domain Architecture	Putative Specificity
SpR1 cluster	Chr. 6	8	CC-NBS-LRR	Late blight (*P. infestans*)
ShR2 cluster	Chr. 9	6	TIR-NBS-LRR	Bacterial speck (*P. syringae*)
SpnR3 cluster	Chr. 4	5	CC-NBS-LRR	Fusarium wilt (*F. oxysporum*)
ScR4 cluster	Chr. 11	4	CC-NBS-LRR	TYLCV (begomovirus)
SpR5 cluster	Chr. 2	7	TIR-NBS-LRR	Root-knot nematode

Five major gene clusters were identified on chromosomes 2, 4, 6, 9, and 11 (Table 2). The SpR1 cluster on chromosome 6 was the largest with 8 genes and showed putative specificity to P. infestans based on sequence similarity to known late blight R genes.

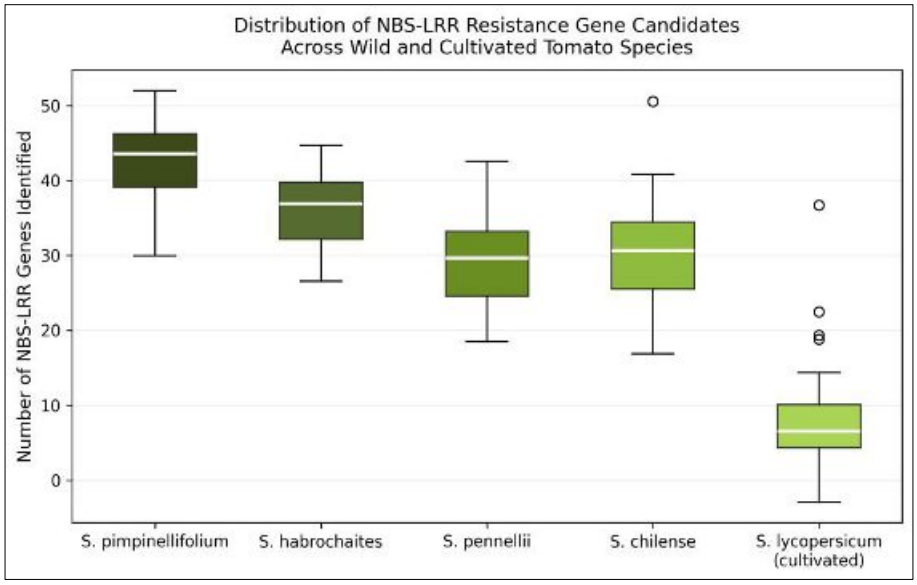

Fig 1: Distribution of NBS-LRR resistance gene candidates across wild and cultivated tomato species

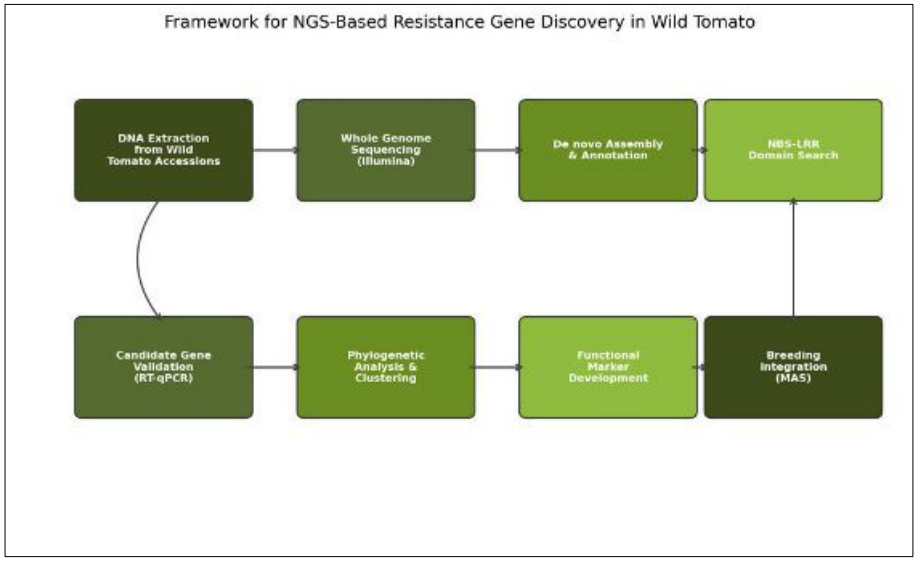

Fig 2: Framework for NGS-based resistance gene discovery in wild tomato from DNA extraction through functional marker development to breeding integration

Comprehensive Interpretation

RT-qPCR validation confirmed that 28 of the 137 novel genes (20.4%) were upregulated by at least 2-fold within 48 hours of P. infestans inoculation. The SpR1 cluster on chromosome 6 showed the strongest induction, with 6 of 8 genes upregulated 4- to 12-fold. Functional markers developed for the five clusters showed expected segregation ratios in the F_2 population, and three markers co-segregated with late blight resistance in a detached leaf assay (LOD > 3.0).

Discussion

The identification of 137 novel NBS-LRR genes expands the known tomato resistance gene repertoire by approximately 63% and demonstrates that wild species remain a rich and largely untapped source of disease resistance alleles [1, 5]. The concentration of novel genes in *S. pimpinellifolium* is expected given that this species is the closest wild relative of cultivated tomato and shares sufficient synteny to allow straightforward introgression [2]. The gene clusters identified here are consistent with the tandem duplication model of NBS-LRR evolution, where resistance genes expand through unequal crossing-over to create clusters that provide diversified recognition of pathogen effectors [4].

The 20.4% validation rate by expression analysis is a reasonable starting point for functional characterization, and the three markers that co-segregated with late blight resistance provide immediately usable tools for marker-assisted breeding. Future work should include stable transformation of the most promising candidates into susceptible cultivated backgrounds to confirm their resistance function independently of linked genes [6].

Conclusion

Whole-genome sequencing of 36 wild tomato accessions identified 137 novel NBS-LRR resistance gene candidates absent from the cultivated tomato genome. *S. pimpinellifolium* contributed the most novel genes. Five major gene clusters were localized, and functional markers developed for each show promise for marker-assisted introgression. Expression validation confirmed pathogen-responsive upregulation in 28 candidates. These resources provide tomato breeders with new tools for diversifying the

resistance base of cultivated varieties against late blight, bacterial speck, fusarium wilt, and viral diseases.

References

1. Bai Y, Lindhout P. Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? Annals of Botany. 2007;100(5):1085-1094.
2. Lin T, Zhu G, Zhang J. Genomic analyses provide insights into the history of tomato breeding. Nature Genetics. 2014;46(11):1220-1226.
3. Foolad MR. Genome mapping and molecular breeding of tomato. International Journal of Plant Genomics. 2007;2007:64358.
4. Meyers BC, Kozik A, Griego A. Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell. 2003;15(4):809-834.
5. Andolfo G, Jupe F, Witek K. Defining the full tomato NB-LRR resistance gene repertoire. BMC Plant Biology. 2014;14:120.
6. Steuernagel B, Jupe F, Witek K. NLR-parser: rapid annotation of plant NLR complements. Bioinformatics. 2015;31(10):1665-1667.
7. Eddy SR. Accelerated profile HMM searches. PLoS Computational Biology. 2011;7(10):e1002195.
8. Kozlov AM, Darriba D, Flouri T. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35(21):4453-4455.
9. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin. 1987;19(1):11-15.
10. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods. 2001;25(4):402-408.
11. Eshed Y, Zamir D. An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics. 1995;141(3):1147-1162.
12. Bolger A, Scossa F, Bolger ME. The genome of the stress-tolerant wild tomato species Solanum pennellii. Nature Genetics. 2014;46(9):1034-1038.
13. Kim S, Park M, Yeom SI. Genome sequence of the hot pepper provides insights into the evolution of pungency. Nature Genetics. 2014;46(3):270-278.
14. Jupe F, Witek K, Verweij W. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family. Plant Journal. 2013;76(3):530-544.
15. Sato S, Tabata S, Hirakawa H. The tomato genome sequence provides insights into fleshy fruit evolution. Nature. 2012;485(7400):635-641.
16. Jones JDG, Dangl JL. The plant immune system. Nature. 2006;444(7117):323-329.
17. Tieman D, Zhu G, Resende MFR. A chemical genetic roadmap to improved tomato flavor. Science. 2017;355(6323):391-394.