

ORIGINAL ARTICLE

Plant Genetic Resources

Exploring the genomic diversity and breeding applications of the *Solanum sessiliflorum* transcriptome via phylogenetic analysis

Priscila O. Silva^{1,2} | Lucas Eduardo R. Silva² | Débora G. Gouveia² |
 Priscila M. Rodrigues¹ | José Francisco C. Gonçalves³ | Adriano Nunes-Nesi² |
 Wagner L. Araújo² | João Henrique F. Cavalcanti¹ 

¹Instituto de Educação, Agricultura e Ambiente, Universidade Federal do Amazonas, Humaitá, Amazonas, Brazil

²National Institute of Science and Technology on Plant Physiology under Stress Conditions, Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil

³Laboratory of Plant Physiology and Biochemistry, National Institute for Amazon Research (INPA), Manaus, Amazonas, Brazil

Correspondence

João Henrique F. Cavalcanti, Instituto de Educação, Agricultura e Ambiente, Universidade Federal do Amazonas, 69800-000, Humaitá, Amazonas, Brazil.
 Email: jcavalcanti@ufam.edu.br

Assigned to Associate Editor Laura Shannon.

Funding information

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; National Institute of Science and Technology in Plant Physiology under Stress Conditions, Grant/Award Number: 406455/2022-8; Fundação de Amparo à Pesquisa do Estado do Amazonas, Grant/Award Numbers: 062.00899/2019, 01.02.016301.04229/2023-51; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Numbers: 175472/2023-8, 304360/2022-8, 313534/2020-9

Abstract

Compelling evidence supports the potential application of wild crop relatives in *Solanum* breeding. Efforts have been made to generate genomic data from wild *Solanum* plants to assess the insertion of advantageous traits into crop species. South America hosts a broad range of plants that have not been widely evaluated at the molecular level. *Solanum sessiliflorum* is a wild species that shows tolerance to *Ralstonia solanacearum* and nematode infections. The Illumina platform was used to construct the transcriptome of *S. sessiliflorum*. The data were analyzed using bioinformatics tools for both phylogenetic comparison and bioprospecting of disease-related genes. Our leaf transcriptome assembly of *S. sessiliflorum* enables phylogenetic comparisons and stress-tolerant gene bioprospecting. De novo assembly generated 114,184 unigenes. A comparison of the *S. sessiliflorum* unigene dataset with other *Solanum* nucleotide genomic resources revealed greater similarity with *Solanum tuberosum* than with *Solanum melongena*. Additionally, *S. sessiliflorum* within the *Leptostemonum* group, along with *S. melongena*, possesses features related to *Solanum* clade evolution. Bioprospection of disease response targets identified 122 potential candidate unigenes retrieved from the *S. sessiliflorum* dataset. The abundant expression of fragments of disease-related and hormonal defense genes appears to constitute a housekeeping mechanism to avoid pathogen attacks on leaves. In total, 1091 unigenes were classified as transcription factors

Abbreviations: BLAST, Basic Local Alignment Search Tool; BUSCO, benchmarking universal single-copy ortholog; CWR, crop wild relatives; IL, introgression line; NCBI, National Center for Biotechnology Information; ORF, open reading frame; QTL, quantitative trait loci; TF, transcription factor; UTR, untranslated region.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Crop Science* published by Wiley Periodicals LLC on behalf of *Crop Science* Society of America.

(TFs), with a large number of TFs associated with biotic resilience. These results highlight the potential for exploring the genomic diversity of *S. sessiliflorum*, which will be useful for applications in breeding programs.

Plain Language Summary

The *Solanum* genus encompasses a diverse array of species, including several economically important crops like potato, tomato, and eggplant. Wild *Solanum* species offer a valuable genetic reservoir for improving cultivated crops. *Solanum sessiliflorum* is one such wild relative that exhibits promising traits, particularly in disease resistance. To explore the genetic makeup of *S. sessiliflorum* and identify potential genes for crop improvement, we conducted a comprehensive transcriptome analysis. Our findings revealed a close genetic relationship between *S. sessiliflorum* and *Solanum tuberosum* (potato). Moreover, we identified 122 candidate genes associated with disease resistance. These results underscore the potential of *S. sessiliflorum* as a valuable genetic resource for breeding programs aimed at developing more resilient and productive crop varieties. By leveraging the genetic diversity of this wild species, we can address critical challenges in agriculture, such as pests and diseases.

1 | INTRODUCTION

The *Solanum* genus is one of the largest crop groups used to feed human populations worldwide. Therefore, it has received considerable attention because of its importance in food security. Eggplant (*Solanum melongena*), potato (*Solanum tuberosum*), and commercial tomato (*Solanum lycopersicum*) are used in plant breeding programs to improve agronomic traits and increase plant productivity (Tirnaz et al., 2022). However, artificial selection through crop breeding reduces the genetic diversity of modern plant populations (Allaby et al., 2018; Meyer & Purugganan, 2013). Modern monocropping contrasts with the idea that productivity or environmental resilience depends on the genetic diversity most likely found in wild species or, to a lesser extent, landraces (Khoury et al., 2022; Panis et al., 2020). Use of genetic pool variations found in crop wild relatives (CWRs) may assist in breeding programs (Gebhardt, 2016; Henderson & Salt, 2017). For example, closely related tomato varieties, such as *Solanum chilense*, *S. pimpinellifolium*, and *S. pennellii*, are drought-tolerant (Zsögön et al., 2017). Thus, both morphological traits (such as a robust root system) and physiological responses (such as closing stomata) are key for drought tolerance (Prince et al., 2015).

However, incorporating these traits into crops remains challenging (Renzi et al., 2022). Although a range of methodologies is commonly used for breeding, the major strategy relies on cross-hybridization with CRW by controlling pollination (Dempewolf et al., 2017). Interspecific hybrids, such as those of eggplant and its wild ancestor, *Solanum insanum*,

are of particular interest among members of the *Solanum* genus and are both more vigorous and productive than are their parental genotypes (Kaushik et al., 2016; Plazas et al., 2016; Syfert et al., 2016). Moreover, one of the most successful breeding hybridizations from CWR to *Solanum* was observed between tomatoes and wild *S. pennellii*. Generation of introgression lines (ILs) has resulted in the acquisition of quantitative loci in tomato (de Oliveira Silva et al., 2017; Lana-Costa et al., 2020). For instance, among the 76 ILs generated, IL 3–5 and IL 10–1 were drought-resistant, as under water deprivation, they showed higher fruit yields compared with tomato M82 (Dariva et al., 2020). The authors detected a strong correlation between fruit yield, photosynthetic parameters (e.g., An and gs), and leaf water potential at drought stress, suggesting that these traits are important selection criteria that can be further applied in breeding programs. Notably, the hybridization approach requires long-term effort and has limitations, such as a range of sexual incompatibilities (Goulet et al., 2017; Guri & Sink, 1988; Knapp & Peralta, 2016).

Therefore, omics tools, primarily sequencing-based technologies, can be used to identify compatible CRWs and/or traits for breeding (Dempewolf et al., 2017; Y. Yang et al., 2021). Over the last decade, the genomes and transcriptomes of several wild *Solanum* species were assembled, revealing potential application features for stress responses and productivity (Palumbo et al., 2021; Vendramin et al., 2019). As described above, although phenotyping and physiological observations revealed that *S. pimpinellifolium*, *S. chilense*, and *S. pennellii* are drought-resistant, genomic investigations enabled molecular exploitation of this feature

(Szymański et al., 2020). Molitor et al. (2021) reported features in the genome assembly of *Solanum siliense* with key roles in the drought response. Gene duplication events and genome fragment inversion on scaffold eight changed the gene expression profiles of key genes involved in drought stress, including *Yucca* and *Bam1*. Furthermore, the metabolome approach applied to tomato ILs resulted in the identification of metabolic quantitative trait loci (QTL) associated with photosynthetic rate, plant architecture, and yield (de Oliveira Silva et al., 2017; Lana-Costa et al., 2020). Taken together, these studies highlight the value of omics, mainly genome sequencing tools, in assessing genomic biodiversity among wild species for bioprospecting and identifying key factors associated with stressful conditions or plant productivity, which can be applied in genetic engineering and assisted crop breeding programs (Abdullah-Zawawi et al., 2022).

In addition to available genetic sources, an interspecific gene pool for crossing is required for *Solanum* breeders. South America, particularly the Amazon rainforest, has broad plant diversity, and many species have not been evaluated in ecological, physiological, and/or molecular studies, including *Solanum sessiliflorum*. *Solanum sessiliflorum*, popularly referred to as cocona or indigenous tomato, has high nutritional value and is useful for extracting bio-compounds for applications in the pharmaceutical industry (Moreno et al., 2015; Silva Filho et al., 2005). *Solanum sessiliflorum* is found throughout the Amazon region, showing a large population with broad polymorphisms in fruit color, form, and size (Silva Filho et al., 2005). In addition, a few studies reported the use of *S. sessiliflorum* as a rootstock for tomato growth in soils with *Ralstonia solanacearum*, bacterial wilt, and nematode infection (Carvalho et al., 2020; Ramírez et al., 2018), suggesting that this wild *Solanum* member is tolerant to biotic stress. However, considering the high level of genetic diversity within species, not all populations may exhibit the desired stress tolerance.

Studies of genomic information can improve the understanding of the natural genetic diversity in wild *Solanum* species, which has potential applications in crop breeding programs. We hypothesized that transcriptome analysis of *S. sessiliflorum* would provide insight into its genomic diversity and evolution, revealing valuable information for phylogenetic resolution within the *Solanum* genus. Additionally, identification of putative disease resistance genes and transcription factors (TFs) in *S. sessiliflorum* offers potential targets for breeding programs aimed at enhancing biotic stress tolerance in *S. sessiliflorum* crops. Thus, the objective of the present study was to conduct a comprehensive de novo transcriptome analysis of this wild perennial herbaceous plant. This study contributes to the understanding of the genetic basis of stress tolerance in wild *Solanum* species and provides valuable information for the develop-

Core Ideas

- *Solanum sessiliflorum*, a wild species native to South America, is resistant to bacterial wilt and nematodes.
- Transcriptome assembly reveals clues about *Solanum* genus evolution resolution and cross-hybridization.
- Genomic information may allow trait/gene prospecting for application in crop breeding.

ment of improved crop varieties through breeding and genetic engineering approaches.

2 | MATERIALS AND METHODS

2.1 | Plant material and RNA isolation

Seeds of *S. sessiliflorum*, kindly provided by Dr. Danilo Fernandes Silva-Filho at the Instituto Nacional de Pesquisas da Amazônia (Manaus, Brazil), were surface sterilized in 0.5% (v/v) sodium hypochlorite for 15 min and washed five times with distilled water. The seeds were sown in a commercial substrate in a greenhouse until they reached 60 days of age. The fifth leaf was harvested from three biological replicates and immediately frozen in liquid nitrogen. Leaf samples were powdered in liquid nitrogen, and total RNA was extracted from 100 mg of each sample using an RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. Total RNA quality was assessed using a Bioanalyzer 2100 (Agilent Technologies) to assess total RNA integrity, quantity, and ribosomal ratio.

2.2 | Illumina sequencing and raw data processing

Sequencing was conducted by Macrogen, Inc. Total RNA was used to amplify cDNA using a TruSeq Stranded mRNA kit (Illumina). The quality of the cDNA library was determined by assessing the PCR-enriched fragment size. The cDNA library was sequenced on a NovaSeq 6000 system with 150 bp paired ends, generating 100 million reads per sample (10 Gb/sample). The bioinformatics analysis was conducted by using Galaxy Training hands-on tutorial (de novo transcriptome reconstruction with RNA-Seq) depicted at GitHub (<https://shiltemann.github.io/training-material/topics/transcriptomics/>), a friendly open source, web-based platform for analytical research and omics data (Afgan et al., 2018; Hiltmann et al., 2023). The Galaxy Europe (<https://usegalaxy.eu/>) provides tools for RNA-seq

manipulation (i.e., align reads and estimate abundance, trinity, and trimmomatics) described according to Table S1 following default parameters (Raghavan et al., 2022). Raw read processing was performed using Trimmomatic to remove adapters and low-quality sequences using the program Trimmomatic 0.38 (Bolger et al., 2014). Filtering provided high-quality reads based on Phred quality scoring (Q30%). The de novo transcriptome was assembled in Trinity v.2.4.0 (Grabherr et al., 2011) and submitted to the CD-HIT-est tool (W. Li & Godzik, 2006) to remove redundant transcript sequences. The completeness of the *S. sessiliflorum* transcriptome assembly was evaluated using benchmarking universal single-copy orthologs (BUSCOs) v.5.4.6 against the Eudicot dataset via a web-based platform (Simão et al., 2015). Expression data were generated using the default parameters of the RSEM v.1.2.0 (B. Li & Dewey, 2011) package. Raw RNA-sequencing data were deposited in the Sequence Read Archive database at the National Center for Biotechnology Information (NCBI) with BioProject ID #PRJNA1167730 (*S. sessiliflorum* leaf transcriptome).

2.3 | Transcriptome annotation and phylogenetic analysis

Solanum sessiliflorum transcriptome annotation was performed using the default parameters in BLASTN of NCBI BLAST (Basic Local Alignment Search Tool) and BLASTX of DIAMOND software, with *e*-value cutoffs of $<10^{-5}$ (refs). Unigenes were submitted to the NCBI Nucleotide (<https://www.ncbi.nlm.nih.gov/nucleotide/>), NCBI non-redundant protein (<https://www.ncbi.nlm.nih.gov/protein/>), UniProt Knowledgebase (<http://www.uniprot.org/>), Kyoto Encyclopedia of Genes and Genomes (<http://www.genome.jp/kegg/ko.html>), and gene ontology (GO) (<http://www.geneontology.org/>). Transcriptome comparison analysis was performed using BLAST2GO software, with the *S. sessiliflorum* unigene dataset compared with the genomic data of the Solanaceae dataset. The Solanaceae genomic dataset was downloaded from https://solgenomics.genomes/sequenced_accessions. Concatenated phylogenetic trees were constructed using six nuclear (*its* and *wax*) and six plastid genes (*matk*, *ndhf*, *psbA-trnH*, and *trnS-G*), as described by Särkinen et al. (2013). A phylogenetic tree was constructed using the maximum likelihood approach in MEGA 11 software.

2.4 | In silico search for biotic stress genes and expression data

To identify the genes involved in the disease response of *S. sessiliflorum*, 112 manually curated pathogens deposited

in the Plant Resistance Genes Database (http://prgdb.crg.eu.gov/Main_Page) were used. These genes were subjected to BLAST+ blastn, cutoff, and 10^{-20} *e*-values, and the results were downloaded from <http://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>. For absolute gene expression analysis, the abundance of unigenes across samples was estimated as an expression measure using the RSEM algorithm (<http://deweylab.github.io/RSEM/>; RSEM version v1.3.1), according to Li and Dewey (2011). Expression profiles were extracted from the RSEM output per unit of unigene for each sample. The data are presented as transcripts per million.

2.5 | Prediction of TF families

The prediction of TF families of *S. sessiliflorum* was predicted using the Plant Transcription Factor Database (v4.0; <http://planttfdb.gao-lab.org/prediction.php>) (Jin et al., 2017).

3 | RESULTS AND DISCUSSION

3.1 | Transcriptome sequencing and assembly of *S. sessiliflorum*

The estimated exponential growth of the global population in the next few decades has resulted in greater pressure on agriculture to increase food security worldwide (Saath & Fachinello, 2018). Therefore, research aimed at increasing crop production and resilience traits is critical in modern agriculture. The development of next-generation sequencing led to a revolution in plant breeding through a range of approaches, from the generation of genetic information on plant diversity to quantitative gene expression through the identification of key candidates for increasing yield or stress tolerance (Zhang et al., 2023). Indeed, these approaches have been successful, particularly for members of the *Solanum* genus, whose many wild relatives are important agronomic crops (Poiczai et al., 2022). Nonetheless, further research is required to investigate their diversity and evolutionary relationships to fully utilize wild *Solanum* species in breeding programs. We performed transcriptome assembly of *S. sessiliflorum*, a perennial herbaceous plant native to the Amazon rainforest (Figure S1).

To provide genomic information about this species, we performed deep RNA sequencing of three biological replicates of foliar samples using the Illumina NovaSeq 6000 system and paired-end reads. A total of 11,965,669,104 read bases covering at least 101,913,872 reads with quality scores greater than the Q30 cutoff were sequenced (Table S2). These high-quality reads were used for de novo *S. sessiliflorum* transcriptome assembly of 114,184 unigenes with a GC

TABLE 1 RNA-sequencing paired-end sequencing and de novo assembly of *Solanum sessiliflorum*.

Total unigenes	114,184
GC unigene content	38.03%
Total length (bases)	77,613,237
Transcript mean length (bases)	679.72
Median contig length (bases)	428
Max length (bases)	17,113
Min length (bases)	201
N50 (bases)	1083
Total of ORFs	26,210 (100%)
Complete ORFs	13,810 (52.69%)
Truncated in 5'UTR ORFs	7234 (27.6%)
Truncated in 3'UTR ORFs	1491 (5.69%)
Truncated in both sides ORFs	3675 (14.02%)

Abbreviations: ORF, open reading frame; UTR, untranslated region.

content of 38.03%. In total, 77,613,237 high-quality bases were sequenced; the minimum contig lengths were 17,113 and 201, with a mean transcript length of 679.72 bases and an N50 of 1083 bases (Table 1). The length distribution of all unigenes is shown in Figure S2. A previous study indicated that *S. sessiliflorum* is diploid, with $2n = 24$ chromosomes, which is consistent with other important *Solanum* crop species (Purcell, 1966). Furthermore, although its exact genome size is unknown, estimates suggest that it is approximately 900 Mb long, similar to other members of this genus (Wei et al., 2020; Zhang et al., 2023). Thus, the transcriptome used in this study was approximately 10-fold longer than the genome.

Coding sequence annotation of the dataset revealed 26,210 open reading frames (ORFs). Of these, 13,810 (52.69%) were complete, 7234 (27.6%) were truncated at the 5' untranslated region (UTR), 1491 (5.69%) were truncated at the 3' UTR, and 3675 (14.02%) were truncated on both sides (Table 1). The total number of ORFs in *S. sessiliflorum* appears to be lower than those in other wild *Solanum* transcriptomes, such as *S. aethiopicum*, *S. dulcamara*, and *S. incanum* (which have 40,353, 32,157, and 35,943 ORFs, respectively) (D'Agostino et al., 2013; Gramazio et al., 2016). However, we focused solely on leaf tissues, whereas previous studies used a mix of tissue samples, which may explain the differences in ORF counts. The quality of the assembly was evaluated using BUSCO against the Eudicots dataset using 2326 BUSCO groups. The transcriptome contained over 88.9% (2067) complete BUSCOs, with only 3.6% (84) and 7.5% (175) fragmented and missing BUSCOs, respectively. Of the 2067 BUSCOs, 2004 were "complete and single-copy BUSCOs" and 63 were "complete and duplicated BUSCOs."

TABLE 2 Annotation result of various databases.

Overall	78,477 (68.73%)
NCBI nucleotide sequence	69,712 (61.05%)
NCBI non-redundant protein	50,874 (44.55%)
UniProt knowledgebase	31,372 (27.47%)
Gene ontology	37,453 (32.8%)
EggNOG	49,772 (43.59%)
Kyoto Encyclopedia of Genes and Genomes	50,105 (43.88%)

Abbreviations: EggNOG, Evolutionary Genealogy of Genes; NCBI, National Center for Biotechnology Information.

3.2 | Functional annotation and phylogenetic inference

To provide biological information on the generated genomic data, the unigene contigs of *S. sessiliflorum* were compared with a range of datasets in BLASTX with a 10^{-5} cutoff (Table 2).

Although the gene structure is not fully correlated with gene function, sequence comparison remains a reliable approach for inferring the correspondence between sequence similarity and functionality using pairwise alignment tools (Yunes & Babbitt, 2019). In total, 78,477 (68.73%) of 114,184 unigenes were annotated in at least one public database. Moreover, 69,712 (61.05%), 50,874 (44.55%), and 31,372 (27.47%) unigenes were annotated against the NCBI nucleotide sequences, NCBI nonredundant proteins, and UniProt Knowledgebase databases, respectively. However, BLASTX covered only 37,453 (32.8%) of the 114,184 unigenes in the GO database, 67.2% of which were unannotated (Figure S3). This unknown identification percentage is concerning for species lacking robust genomic information and is well-documented in other wild *Solanum* species (Lateef et al., 2018).

GO classification revealed 37,453 unigenes in the three major enrichment categories (Figure S4). In terms of biological function, the most prominent enriched GO terms were metabolic processes (32.35%) with respect to anabolic and catabolic biochemical reactions and pathways, followed by cellular processes (14.58%) and biological regulation (14%) associated with cellular local and biological process homeostasis, respectively. Collectively, these terms constituted >50% of biological function annotations in the UniGene dataset (Figure S4A). Catalytic and binding activities represented >75% of molecular function annotations in the transcriptome of *S. sessiliflorum* (Figure S4B). Both of these activities are associated with biological reactions, most of which involve enzymes and interactions between cellular components. Finally, the cellular component term was associated with the majority of unigenes related to cell parts (44.6%) and organelles (22.4%) (Figure S4C). The cellular component

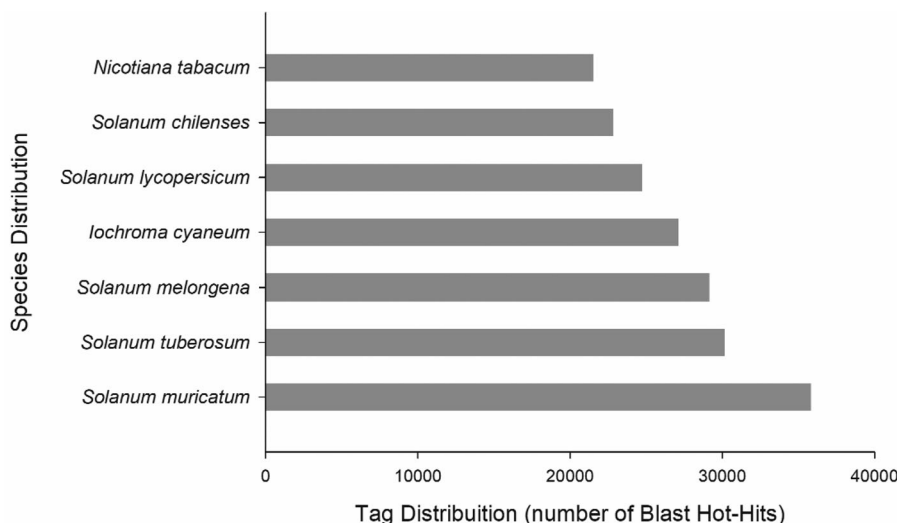


FIGURE 1 Megablast nucleotide top hit species distribution of the *Solanum sessiliflorum* unigene dataset against the Solanaceae genomic resource.

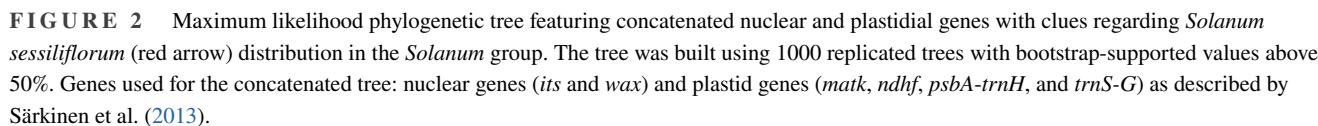
enrichment terms in the GO database were most likely based on signal peptide similarity. Thus, the 5' regions were likely unable to provide a statistically accurate ability to specify the cellular component. A large number of orthologs in the Evolutionary Genealogy of Genes database were searched to identify proteins distributed in eukaryotic orthologous groups, clusters of orthologous groups, and non-supervised orthologous groups. The Evolutionary Genealogy of Genes database was used to annotate 49,772 (43.59%) of the unigenes (Table 2). Using this approach, the unigenes were mapped to the annotations of the corresponding orthologous group clusters in three main categories: information storage and processing, cellular processes, signaling, and metabolism (Figure S5). Similarly, considering that annotation based on the Kyoto Encyclopedia of Genes and Genomes database can also be used to infer orthology to obtain the eukaryotic orthologous (KO; reference pathway) number of the Kyoto Encyclopedia of Genes and Genomes annotation, 50,105 (43.88%) unigenes from the *S. sessiliflorum* transcriptome were annotated (Table 2). To address the low ontology shown in Figure S4, the unigene dataset was further analyzed using BLAST2GO software to compare the sequence similarity of the *S. sessiliflorum* transcriptome with those of the other *Solanum* datasets (Figure 1).

This approach identified *Solanum muricatum* as the top BLAST hit, followed by *S. tuberosum* and *S. melongena*. Interestingly, the *S. sessiliflorum* transcriptome was more similar to that of *Iochroma cyanenum* than to those of *S. lycopersicum* or *Solanum chilenses*. This finding reinforces the idea that the *Solanum* phylogenetic tree is complex in terms of classification and suggests a polytomy origin (Gagnon et al., 2022; Särkinen et al., 2013; Weese & Bohs, 2007). Solanaceae trees exhibit several characteristics related to backbone reso-

lution, as highlighted recently by Powell et al. (2022). Thus, introduction of new genomic information, such as the *S. sessiliflorum* transcriptome, may be valuable for clarifying the *Solanum* genus clade.

3.3 | Phylogenetic tree and features of *S. sessiliflorum* in the *Solanum* clade

Phylogenetic trees provide graphical outputs for visualizing evolutionary relationships. We constructed a concatenated tree using both nuclear and plastid genes to analyze the evolutionary position of *S. sessiliflorum* (Figures 2 and Figure S6). Six nuclear and plastid markers were selected for phylogenetic analyses by Särkinen et al. (2013). *Solanum sessiliflorum* was closely related to eggplant when a generalist angiosperm tree was constructed, suggesting that both species are sister groups (Figure S6). Hence, to integrate *S. sessiliflorum* features into the *Solanum* phylogeny, a more specific tree was constructed (Figure 2). *Solanum sessiliflorum* clustered with other native wild species from South America (specifically Bolivia, Brazil, and Peru), including *S. sisymbriifolium*, *S. talarense*, and *S. wright*. Thus, systematically, *S. sessiliflorum* belongs to CLADE II in the *Leptostemonum* group but is outside the eastern hemisphere spiny cluster (formerly known as the Old World sessiliflorum clade), where eggplant and its closely related wild species (e.g., *S. triblobatum* and *S. incanum*) are clustered. These data aligned with the orthodox taxonomic classification (The Brazilian Flora Group, 2018). The use of a concatenated tree with both plastid and nuclear gene markers may infer the relationship distance among *Solanum* members. Therefore, our data suggest that, although *S. sessiliflorum* belongs to the *Leptostemonum*



3.4 | Identification and differential expression profile of disease resistance genes

A range of studies has focused on genetic resources found in wild species for the bioprospection of disease response genes (Ghatak et al., 2017; Herraiz et al., 2016; Poczai et al., 2022; Sato et al., 2021). *Solanum sessiliflorum* was previously reported to be tolerant to bacterial wilt and nematode infections in the soil (Carvalho et al., 2020; Fernandes, 2016; Ramírez et al., 2018). Although the data provided evidence for the disease mechanism observed in *S. sessiliflorum*, the specific molecules involved in this resistance remain unclear. To explore the putative disease genes in the *S. sessiliflorum* transcriptome, we used the plant resistance gene database, a curated server, to identify disease-related genes (Osuna-Cruz et al., 2018). Thus, 122 biotic-stress-related unigenes displaying high similarity to seven protein classes associated with resistance genes were retrieved (Table S3 and Tables 3 and 4) (van Ooijen et al., 2007). Our analysis revealed fewer unigenes associated with these protein classes in *S. sessiliflorum* compared to those identified previously by X. Yang et al. (2014), who cataloged these protein classes in several *Solanum* species (Table 3).

TABLE 3 Summary of plant resistance genes comparison among *Solanum sessiliflorum* and other *Solanum* members.

	This work	Yang et al. (2014)			
	<i>S. sessiliflorum</i>	<i>S. melongena</i>	<i>S. torvum</i>	<i>S. lycopersicum</i>	<i>S. tuberosum</i>
CNL	39	110	194	99	219
TNL	15	46	66	29	93
RLK	7	221	255	134	156
RLP	10	84	128	77	132
TN	4	1	–	–	–
NL	13	16	21	41	46
Kinase	19	31	29	16	23
Other	15	112	122	109	105
Total	122	621	815	505	774

TABLE 4 Top 20 most abundant disease-related genes response from leaf tissue of *Solanum sessiliflorum*.

Contig ID	Gene name	Transcripts per million
c133549_g1_i1	Hop-interacting protein THI032	671.29
c107119_g1_i1	Cinnamoyl-CoA reductase 1	164.98
c125035_g1_i1	Anthocyanidin reductase	153.71
c84093_g2_i1	Somatic embryogenesis receptor kinase 2	112.46
c133599_g1_i1	MLO-like protein 1	100.62
c1535_g1_i2	Indole-3-pyruvate monooxygenase YUCCA10	87.01
c107805_g1_i1	Cinnamoyl-CoA reductase 1	71.99
c124872_g1_i1	Cinnamoyl-CoA reductase 1	68.62
c82838_g1_i2	MLO-like protein 8	47.84
c106994_g1_i1	Cinnamoyl-CoA reductase 1	34.34
c47075_g1_i1	Phenylacetaldehyde reductase	25.75
c88934_g1_i3	Elongator complex protein 6	25.09
c89070_g2_i1	LRR receptor-like serine/threonine-protein kinase	22.94
c58492_g1_i1	Somatic embryogenesis receptor kinase 3A	22.81
c89070_g4_i1	Receptor-like protein Cf-9 homolog	22.55
c88757_g1_i6	RESTRICTED TEV MOVEMENT 1-like	20.22
c89301_g4_i3	TMV resistance protein N-like	19.32
c89380_g2_i1	Brassinosteroid insensitive 1	15.43
c89245_g1_i2	Anthocyanidin reductase	14.71

However, compared with X. Yang et al. (2014), who performed searches using BLAST with an *e*-value cutoff of 10^{-5} , we used more rigorous parameters (i.e., an *e*-value cutoff of 10^{-20}) to avoid misleading disease gene annotations. The transcriptome was generated from the leaves of unstressed plants. Therefore, resistance genes that appeared to be tissue-specific and/or induced only by stress were not considered in the experimental design. The expression profiles of these 122 unigenes were determined, and the 20 most highly expressed genes are listed in Table 4. Numerous cinnamoyl-CoA reductase one unigenes appear to play a housekeeping role for lignin biosynthesis in *S. sessiliflorum*. Furthermore, the presence of the YUCCA unigene is consistent with the

findings of Molitor et al. (2021), who suggested that the YUCCA gene is involved in the stress response of *S. sitiens*.

In contrast, strategies used to identify a single gene for classical breeding or genetic engineering studies are weakly supported. The architecture and organization of chromosomes, which are heritable for new generations, are likely associated with agronomic traits related to crop improvement (Meyer et al., 2021). The introduction of QTL from wild species into *Solanum* crop species to introduce advantageous features is a suitable strategy for plant breeding (X. Li et al., 2024; Lana-Costa et al., 2020). In this context, the introduction of novel genes and/or agronomic traits into a target crop, assisted by cross-hybridization from a CRW gene pool or by

TABLE 5 Top 20 most abundant putative transcription factors from leaf tissue of *Solanum sessiliflorum*.

Contig ID	Gene name	Transcripts per million
c99320_g1_i1	Zinc finger CCCH domain-containing protein 48	320.94
c99499_g1_i1	Barley B recombinant-like protein	305.45
c99295_g1_i1	Probable transcription factor KAN4	279.53
c99110_g1_i1	Zinc finger protein ZAT6	259.1
c99016_g1_i1	Oxidative stress 2	209.8
c98805_g1_i1	Ethylene-responsive transcription factor ERF010	199.33
c99010_g1_i1	BEL1-like homeodomain protein	189.51
c98456_g1_i1	Heat Shock Protein 24	186.56
c98690_g1_i1	STOREKEEPER protein	182.86
c98511_g1_i1	Myb-related protein 38	181.25
c98736_g1_i1	Transcription factor TCP19	169.83
c98392_g1_i1	Homeobox-leucine zipper ANTHOCYANINLESS 2	155.5
c98338_g1_i1	Heat shock transcription factor 34	149.646
c98258_g1_i1	Pathogenesis-related transcriptional activator PTI6	132.84
c98220_g1_i1	Zinc finger CCCH domain-containing protein 48	130.22
c98225_g1_i1	Probable transcription factor	128.44
c98277_g1_i1	Transcription factor PRE1	127.193
c98318_g1_i1	WRKY75	122.32
c98219_g1_i1	Protein indeterminate-domain 7	121.533
c98130_g1_i1	Scarecrow-like protein 6	1.189.933

controlling pollination, involves intraspecific sexual barriers (Zhang et al., 2017). Thus, although we focused on the transcriptome and did not evaluate the chromosome structure and fragments linked to specific traits, the transcriptome data provided information on the relationship distances that drive new crosses between *S. sessiliflorum* and *Solanum* crop species. The use of *S. sessiliflorum* and other *Solanum* species has also been reported (Ramírez, 2020). A cross between *S. sessiliflorum* and *S. candidum* resulted in fertile hybrids with low fruit production (Heiser, 1987). In turn, crossing *S. sessiliflorum*, which acts as a female, and *S. quitoense*, which acts as a male, results in hybrids with both greater plant height and higher levels of nematode tolerance (Rabelo, 2015; Ramírez et al., 2018).

3.5 | TF members of the *S. sessiliflorum* transcriptome

TFs play crucial roles in gene expression and thus are the regulatory hubs of plant metabolism. Using an online TF prediction tool (<http://planttfdb.gao-lab.org/prediction.php>), 1060 putative single unigenes covering major families of plant TFs were identified (Figure S7). Among the most abundant TF families, BLBH (96), C2H2 (91), ERF (70), MYB-related (57), bZIP (55), WRKY (52), NAC (46), GRAS (40), and MYB (39) accounted for >50% of the TFs identified. The

expression profiles of these are shown in Table S4, and the top 20 most highly expressed TF unigenes are listed in Table 5.

In this context, most of the expressed TF unigenes were related to leaf ontogenesis, development, or hormone metabolism, although other TFs associated with biotic resilience were highly abundant. Thus, to explore the framework associations between unigene disease genes and the TFs identified in Tables 4 and 5, Pearson correlation analysis was performed (Table S5; Figure 3). Pathogenesis-related TFs, such as WRKY75, PT6I, and ZAT6, appear to be suitable candidates for coordinating disease gene responses. Consistent with these findings, WRKY75 was positively correlated with all disease-related genes, whereas PT6I was associated with CCR1, SERK, and LLR-RLK. In fact, PT6I and ZAT6 displayed opposite profiles. In addition, several GRAS TFs (e.g., SCR and PRE1) are involved in the stress response. PRE1 is correlated with EPL6, a stress response player with an epigenetic role, and ANR, a drought stress response, whereas SCR is more generally correlated with CCR1, an MLO protein, and well-known resistance players such as TMVE (Shi et al., 2021). Recent studies reported the transcriptional dynamics of tomato resistance to bacterial wilt, wherein CC1 plays a key role in lignin biosynthesis, mediated by L-arabidnose exogenous application (Fu et al., 2024). The authors discussed the role of lignin as a physical barrier against pathogenic infections. Moreover, lignin acts as a physical barrier against

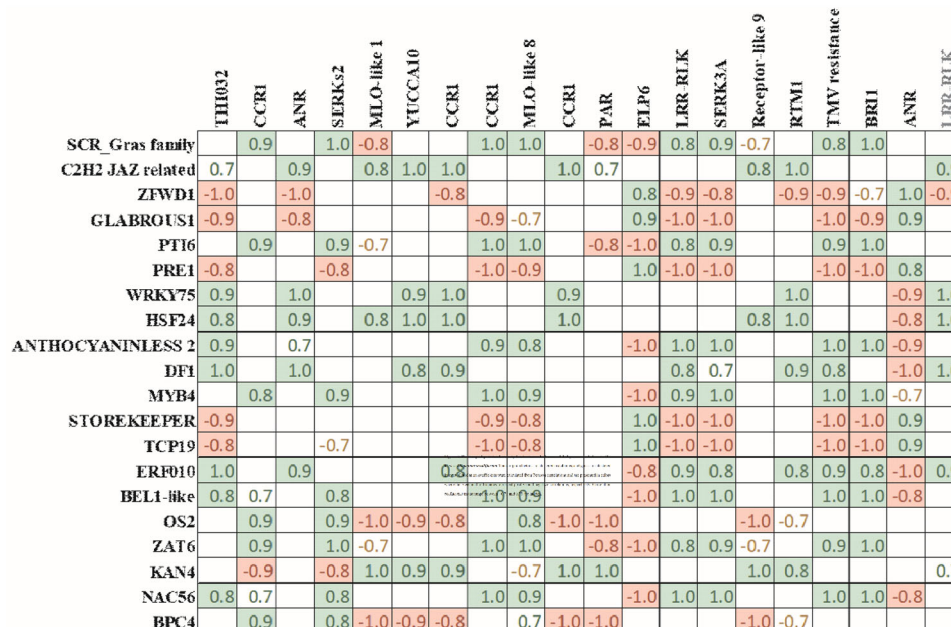


FIGURE 3 Heatmap of putative transcription factor and disease-related gene correlations in the leaves of *Solanum sessiliflorum*. Transcription factors are clustered in columns, and genes are clustered in rows. Correlation coefficients were calculated from Pearson correlations and are presented in colors where the green and red squares represent positive and negative correlations, respectively. Correlation coefficients in the range between -0.75 and 0.75 are hidden.

nematode attack (Veronico et al., 2018). Taken together, the high expression of CC1 and the TF regulating its expression in *S. sessiliflorum* may be associated with biotic responses. This association should be further investigated to determine the mechanisms or connections with the stress response traits of *S. sessiliflorum*. In turn, other TF unigenes associated with ontogenesis, such as KAN4 and TCP19, were correlated with disease genes despite the lack of known stress-related functions associated with these genes. Taken together, these findings provide an initial gene regulatory network for *S. sessiliflorum*, offering insights into the potential regulatory mechanisms for further investigation of cell metabolism.

Collectively, our study of the *S. sessiliflorum* transcriptome provides insights into its genetic diversity and potential applications for crop improvement. Through transcriptome analysis, we identified evolutionary relationships within the *Solanum* genus and candidate genes associated with biotic stress tolerance. These findings offer promising opportunities for enhancing crop resilience and productivity by harnessing the genetic resources of wild *Solanum* species.

AUTHOR CONTRIBUTIONS

Priscila O. Silva: Conceptualization; formal analysis; investigation; writing—original draft; writing—review and editing. **Lucas Eduardo R. Silva:** Formal analysis; methodology; writing—original draft. **Débora G. Gouveia:** Data curation; investigation; writing—original draft. **Priscila M. Rodrigues:** Data curation; writing—original draft. **José**

Francisco C. Gonçalves: Data curation; formal analysis; methodology; writing—original draft; writing—review and editing. **Adriano Nunes-Nesi:** Data curation; investigation; methodology; writing—review and editing. **Wagner L. Araújo:** Formal analysis; methodology; writing—original draft; writing—review and editing. **João Henrique F. Cavalcanti:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; supervision; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

The authors thank Prof. Diego Batista Silva of the Universidade Federal da Paraíba for valuable critical reading and suggestions for improving this manuscript. This study was financially supported by the Foundation for Research Assistance of Amazonas State (FAPEAM; Grant number 062.00899/2019 and 01.02.016301.04229/2023-51 to JHFC). This study was partially supported by the National Institute of Science and Technology in Plant Physiology under Stress Conditions (INCT Plant Stress Physiology; Grant 406455/2022-8) of the National Council for Scientific and Technological Development. Scholarship granted by CNPQ (grant number 175472/2023-8) to POS and DGG Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to LERS and FAPEAM to PMR. CNPQ contributed also with additional financial support (Grant number: 407276/201-1 and number 406455/2022-8 to ANN and WLA, respectively). The authors are also grateful for the

research fellowships granted by CNPQ-Brazil to AN-N (Grant 313534/2020-9), WLA, and JHFC (Grant # 304360/2022-8).

The Article Processing Charge for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ORCID

João Henrique F. Cavalcanti  <https://orcid.org/0000-0002-1136-6564>

REFERENCES

- Abdullah-Zawawi, M.-R., Tan, L.-W., Ab Rahman, Z., Ismail, I., & Zainal, Z. (2022). An integration of transcriptomic data and modular gene co-expression network analysis uncovers drought stress-related hub genes in transgenic rice overexpressing OsAbp57. *Agronomy*, 12, 1959. <https://doi.org/10.3390/agronomy12081959>
- Afgan, E., Baker, D., & Batut, B. (2018). The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, 46, W537–W544.
- Allaby, R. G., Ware, R. L., & Kistler, L. (2018). A re-evaluation of the domestication bottleneck from archaeogenomic evidence. *Evolutionary Applications*, 12, 29–37. <https://doi.org/10.1111/eva.12680>
- Bolger, A., Scossa, F., Bolger, M. E., Lanz, C., Maumus, F., Tohge, T., Quesneville, H., Alseekh, S., Sorensen, I., Lichtenstein, G., Fich, E. A., Conte, M., Keller, H., Schneeberger, K., Schwacke, R., Ofner, I., Vrebalov, J., Xu, Y., Osorio, S., ... Fernie, A. R. (2014). The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature Genetics*, 46, 1034–1038. <https://doi.org/10.1038/ng.3046>
- Carvalho, L. T. D. S., Melo, D. M., Vargas, P. F., Santos, H. C. A., & Ferreira, J. V. (2020). Tomato grafting onto Solanaceae genotypes to control bacterial wilt (*Ralstonia solanacearum* Smith 1896). *Pesquisa Agropecuária Tropical*, 50, e63476. <https://doi.org/10.1590/1983-40632020v5063476>
- D'Agostino, N., Golas, T., Van de Geest, H., Bombarely, A., Dawood, T., Zethof, J., Driedonks, N., Bargsten, J., Nap, J., Mariani, C., & Rieu, I. (2013). Genomic analysis of the native European *Solanum* species, *S. dulcamara*. *BMC Genomics*, 14, 1–14. <https://doi.org/10.1186/1471-2164-14-356>
- Dariva, F. D., Copati, M. G. F., Pessoa, H. P., Alves, F. M., Dias, F. O., Picoli, E. A. T., Cunha, F. F., & Nick, C. (2020). Evaluation of anatomical and physiological traits of *Solanum pennellii* Cor. associated with plant yield in tomato plants under water-limited conditions. *Scientific Reports*, 10, 16052. <https://doi.org/10.1038/s41598-020-73004-4>
- Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., & Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Science*, 57, 1070–1082. <https://doi.org/10.2135/cropsci2016.10.0885>
- de Oliveira Silva, F. M., Lichtenstein, G., Alseekh, S., Rosado-Souza, L., Conte, M., Suguiyama, V. F., & Nunes-Nesi, A. (2017). The genetic architecture of photosynthesis and plant growth-related traits in tomato. *Plant, Cell & Environment*, 41, 327–341. <https://doi.org/10.1111/pce.13084>
- Fernandes, B. S. (2016). Uso da enxertia para o controle da murcha bacteriana [*Ralstonia solanacearum* Smith (1896) (Yabuuchi) e t al. 1996] no tomateiro [Master's dissertation, Universidade Federal do Amazonas].
- Fu, H., Gao, Y., Shu, H., Mushtaq, N., Altaf, M. A., Lu, X., Cheng, S., Wang, Z., Zhu, G., & Yu, W. (2024). Dynamic transcription analysis reveals the underlying key mechanism of L-arabinose-enhanced tomato resistance to bacterial wilt. *Scientia Horticulturae*, 337, 113583. <https://doi.org/10.1016/j.scienta.2024.113583>
- Gagnon, E., Hilgenhof, R., Orejuela, A., McDonnell, A., Sablok, G., Aubriot, X., Giacomini, L., Gouvêa, Y., Bragionis, T., Stehmann, J. R., Bohs, L., Dodsworth, S., Martine, C., Pocza, P., Kanpp, S., & Särkinen, T. (2022). Phylogenomic discordance suggests polytomies along the backbone of the large genus *Solanum*. *American Journal of Botany*, 109, 580–601. <https://doi.org/10.1002/ajb2.1827>
- Gebhardt, C. (2016). The historical role of species from the Solanaceae plant family in genetic research. *Theoretical and Applied Genetics*, 129, 2281–2294. <https://doi.org/10.1007/s00122-016-2804-1>
- Ghatak, A., Chaturvedi, P., Paul, P., Agrawal, G. K., Rakwal, R., Kim, S. T., Weckwerth, W., & Gupta, R. (2017). Proteomics survey of Solanaceae family: Current status and challenges ahead. *Journal of Proteomics*, 169(3), 41–57. <https://doi.org/10.1016/j.jprot.2017.05.016>
- Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: Old ideas, new techniques. *Plant Physiology*, 173, 65–78. <https://doi.org/10.1104/pp.16.01340>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., ... Regev, A. (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644.
- Gramazio, P., Blanca, J., Ziarolo, P., Herraiz, F. J., Plazas, M., Prohens, J., & Vilanova, S. (2016). Transcriptome analysis and molecular marker discovery in *Solanum incanum* and *S. aethiopicum*, two close relatives of the common eggplant (*Solanum melongena*) with interest for breeding. *BMC Genomics*, 17, Article 300. <https://doi.org/10.1186/s12864-016-2631-4>
- Guri, A., & Sink, K. C. (1988). Interspecific somatic hybrid plants between eggplant (*Solanum melongena*) and *Solanum torvum*. *Theoretical and Applied Genetics*, 76, 490–496. <https://doi.org/10.1007/BF00260897>
- Heiser, C. B. (1987). Origins of *Solanum lasiocarpum* and *S. repandum*. *American Journal of Botany*, 74, 1045–1048. <https://doi.org/10.1002/j.1537-2197.1987.tb08713.x>
- Henderson, I. R., & Salt, D. E. (2017). Natural genetic variation and hybridization in plants. *Journal of Experimental Botany*, 68, 5415–5417. <https://doi.org/10.1093/jxb/erx377>
- Herraiz, F. J., Blanca, J., Ziarolo, P., Gramazio, P., Plazas, M., Anderson, G. J., Prohens, J., & Vilanova, S. (2016). The first de novo transcriptome of pepino (*Solanum muricatum*): Assembly, comprehensive analysis and comparison with the closely related species *S. caripense*, potato and tomato. *BMC Genomics*, 17, 1–17. <https://doi.org/10.1186/s12864-016-2656-8>
- Hilteimann, S., Rasche, H., Gladman, S., Hotz, H.-R., Larivière, D., Blankenberg, D., Jagtap, P. D., Wollmann, T., Bretaudeau, A., Goué, N., Griffin, T. J., Royaux, C., Bras, Y. L., Mehta, S., Syme, A., Coppens, F., Driesbeke, B., Soranzo, N., Bacon, W., ... Batut, B. (2023). Galaxy training: A powerful framework for teaching!

- PLoS Computational Biology, 19, e1010752. <https://doi.org/10.1371/journal.pcbi.1010752>
- Jin, J., Tian, F., Yang, D.-C., Meng, Y. Q., Kong, L., Luo, J., & Gao, G. (2017). PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, 45, D1040–D1045.
- Kaushik, P., Prohens, J., Vilanova, S., Gramazio, P., & Plazas, M. (2016). Phenotyping of eggplant wild relatives and interspecific hybrids with conventional and phenomics descriptors provides insight for their potential utilization in breeding. *Frontiers in Plant Science*, 7, 677. <https://doi.org/10.3389/fpls.2016.00677>
- Khoury, C. K., Brush, S., Costich, D. E., Curry, H. A., de Haan, S., Engels, J. M., Guarino, L., Hoban, S., Mercer, K. L., Miller, A. J., Nabhan, G. P., Perales, H. R., Richards, C., Riggins, C., & Thormann, I. (2022). Crop genetic erosion: Understanding and responding to loss of crop diversity. *New Phytologist*, 233, 84–118. <https://doi.org/10.1111/nph.17733>
- Knapp, S., & Peralta, I. E. (2016). The tomato (*Solanum lycopersicum* L., Solanaceae) and its botanical relatives. *The Tomato Genome* (pp. 7–21). https://doi.org/10.1007/978-3-662-53389-5_2
- Lana-Costa, J., de Oliveira Silva, F. M., Batista-Silva, W., Carolino, D. C., Senra, R. L., Medeiros, D. B., Martins, S. C. V., Gago, J., Araújo, W. L., & Nunes-Nesi, A. (2020). High photosynthetic rates in a *Solanum pennellii* chromosome 2 QTL is explained by biochemical and photochemical changes. *Frontiers in Plant Science*, 11, 794. <https://doi.org/10.3389/fpls.2020.00794>
- Lateef, A., Prabhudas, S. K., & Natarajan, P. (2018). RNA sequencing and de novo assembly of *Solanum trilobatum* leaf transcriptome to identify putative transcripts for major metabolic pathways. *Scientific Reports*, 8, 15375. <https://doi.org/10.1038/s41598-018-33693-4>
- Li, B., & Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12, 1–16. <https://doi.org/10.1186/1471-2105-12-323>
- Li, W., & Godzik, Adam. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659.
- Li, X., Liu, X., Pan, F., Hu, J., Han, Y., Bi, R., Zhang, C., Liu, Y., Wang, Y., Liang, Z., Zhu, C., Guo, Y., Huang, Z., Wang, X., Du, Y., Liu, L., & Li, J. (2024). Dissection of major QTLs and candidate genes for seedling stage salt/drought tolerance in tomato. *BMC Genomics*, 25(1). <https://doi.org/10.1186/s12864-024-11101-8>
- Meyer, R. C., Weigelt-Fischer, K., Knoch, D., Heuermann, M., Zhao, Y., & Altmann, T. (2021). Temporal dynamics of QTL effects on vegetative growth in *Arabidopsis thaliana*. *Journal of Experimental Botany*, 72, 476–490.
- Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. *Nature Reviews Genetics*, 14, 840–852. <https://doi.org/10.1038/nrg3605>
- Molitor, C., Kurowski, T. J., Fidalgo de Almeida, P. M., Eerolla, P., Spindlow, D. J., Kashyap, S. P., Singh, B., Prasanna, H., Thompson, A. J., & Mohareb, F. R. (2021). De novo genome assembly of *Solanum* reveals structural variation associated with drought and salinity tolerance. *Bioinformatics*, 37, 1941–1945. <https://doi.org/10.1093/bioinformatics/btab048>
- Moreno, C., Quiñones, J. R., & Jiménez, P. (2015). Phenological growth stages of *Solanum sessiliflorum* according to BBCH scale. *Annals of Applied Biology*, 168, 151–157. <https://doi.org/10.1111/aab.12251>
- Osuna-Cruz, C. M., Paytuvi-Gallart, A., Di Donato, A., Sundesha, V., Andolfo, G., Aiese Cigliano, R., Sanseverino, W., & Ercolano, M. R. (2018). PRGdb 3.0: A comprehensive platform for prediction and analysis of plant disease resistance genes. *Nucleic Acids Research*, 46, 461. <https://doi.org/10.1093/nar/gkx1119>
- Palumbo, F., Vannozzi, A., & Barcaccia, G. (2021). Impact of genomic and transcriptomic resources on Apiaceae crop breeding strategies. *International Journal of Molecular Sciences*, 22(18), 9713. <https://doi.org/10.3390/ijms22189713>
- Panis, B., Nagel, M., & Van den Houwe, I. (2020). Challenges and prospects for the conservation of crop genetic resources in field Genebanks, in vitro collections and/or in liquid nitrogen. *Plants*, 9, 1634. <https://doi.org/10.3390/plants9121634>
- Plazas, M., Vilanova, S., Gramazio, P., Rodríguez-Burruezo, A., Fita, A., Herraiz, F. J., Ranil, R., Fonseka, R., Nirán, L., Fonseka, H., Kouassi, B., Kouassi, A., Kouassi, A., & Prohens, J. (2016). Interspecific hybridization between eggplant and wild relatives from different gene pools. *Journal of the American Society for Horticultural Science*, 141, 34–44. <https://doi.org/10.21273/JASHS.141.1.34>
- Poccai, P., D'Agostino, N., Deanna, R., & Portis, E. (2022). Solanaceae VII: Biology, genetics, and evolution. *Frontiers in Genetics*, 13, 932421. <https://doi.org/10.3389/fgene.2022.932421>
- Poccai, P., D'Agostino, N., Deanna, R., & Portis, E. (2023). Editorial: Solanaceae VIII: biodiversity, climate change and breeding. *Frontiers in Genetics*, 14, 1348372. <https://doi.org/10.3389/fgene.2023.1348372>
- Powell, A. F., Zhang, J., Hauser, D., Vilela, J. A., Hu, A., Gates, D. J., Mueller, L. A., Li, F.-W., Strickler, S. R., & Smith, S. D. (2022). Genome sequence for the blue-flowered Andean shrub *Ichroma cyaneum* reveals extensive discordance across the berry clade of Solanaceae. *The Plant Genome*, 15, e20223. <https://doi.org/10.1002/tpg2.20223>
- Prince, S. J., Song, L., Qiu, D., Maldonado dos Santos, J. V., Chai, C., Joshi, T., Patil, G., Valliyodan, B., Vuong, T. D., Murphy, M., Krampis, K., Tucker, D. M., Biyashev, R., Dorrance, A. E., Maroof, M. S., Xu, D., Shannon, J. G., & Nguyen, H. T. (2015). Genetic variants in root architecture-related genes in a Glycine soja accession, a potential resource to improve cultivated soybean. *BMC Genomics*, 16(1). <https://doi.org/10.1186/s12864-015-1334-6>
- Purcell, C. M. (1966). Chromosome Number of Cubiu, *Solanum sessiliflorum* Dunal, a Solanaceous Fruit of the Brazilian Amazonas. *The Journal of Agriculture of the University of Puerto Rico*, 50, 247–251. <https://doi.org/10.46429/jaupr.v50i3.3463>
- Rabelo, J. S. (2015). *Prospecção de espécies de solanáceas para resistência ao nematoide das galhas*. Monografia (Graduação em Agronomia)-Universidade Federal do Ceará.
- Raghavan, V., Kraft, L., Mesny, F., & Rigerte, L. (2022). A simple guide to de novo transcriptome assembly and annotation. *Briefings in Bioinformatics*, 23, bbab563. <https://doi.org/10.1093/bib/bbab563>
- Ramírez, F. (2020). Cocona (*Solanum sessiliflorum* Dunal) reproductive physiology: A review. *Genetic Resources and Crop Evolution*, 67, 293–311. <https://doi.org/10.1007/s10722-019-00870-x>
- Ramírez, F., Kallarackal, J., & Davenport, T. L. (2018). Lulo (*Solanum quitoense* Lam.) reproductive physiology: A review. *Scientia Horticulturae*, 238, 163–176. <https://doi.org/10.1016/j.scienta.2018.04.046>
- Renzi, J. P., Coyne, C. J., Berger, J., von Wettberg, E., Nelson, M., Ureta, S., Hernández, F., Smýkal, P., & Brus, J. (2022). How could the use of crop wild relatives in breeding increase the adaptation of crops to marginal environments? *Frontiers in Plant Science*, 13, 886162. <https://doi.org/10.3389/fpls.2022.886162>

- Saath, K. C. de O., & Fachinello, A. L. (2018). Crescimento da demanda mundial de alimentos e restrições do fator terra no Brasil. *Revista de Economia e Sociologia Rural*, 56(2), 195–212. <https://doi.org/10.1590/1234-56781806-94790560201>
- Särkinen, T., Bohs, L., Olmstead, R. G., & Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): A dated 1000-tip tree. *BMC Evolutionary Biology*, 13, Article 214. <https://doi.org/10.1186/1471-2148-13-214>
- Sato, K., Uehara, T., Holbein, J., Sasaki-Sekimoto, Y., Gan, P., Bino, T., Yamaguchi, K., Ichihashi, Y., Maki, N., Shigenobu, S., Ohta, H., Franke, R. B., Siddique, S., Grundler, F. M. W., Suzuki, T., Kadota, Y., & Shirasu, K. (2021). Transcriptomic analysis of resistant and susceptible responses in a new model root-knot nematode infection system using *Solanum torvum* and *Meloidogyne arenaria*. *Frontiers in Plant Science*, 12, 680151. <https://doi.org/10.3389/fpls.2021.680151>
- Shi, A., Gepts, P., Song, Q., Xiong, H., Michaels, T. E., & Chen, S. (2021). Genome-wide association study and genomic prediction for soybean cyst nematode resistance in USDA common bean (*Phaseolus vulgaris*) Core Collection. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.624156>
- Silva Filho, D. D., Yuyama, L. K. O., Aguiar, J. P. L., Oliveira, M. C., & Martins, L. H. P. (2005). Caracterização e avaliação do potencial agrônomo e nutricional de etnovarietades de cubiu (*Solanum sessiliflorum* Dunal) da Amazônia. *Acta Amazonica*, 35. <https://doi.org/10.1590/S0044-59672005000400003>
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31, 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Syfert, M. M., Castañeda-Álvarez, N. P., Khoury, C. K., Särkinen, T., Sosa, C. C., Achicanoy, H. A., Bernau, V., Prohens, J., Daunay, M. C., & Knapp, S. (2016). Crop wild relatives of the brinjal eggplant (*Solanum melongena*): Poorly represented in genebanks and many species at risk of extinction. *American Journal of Botany*, 103, 635–651. <https://doi.org/10.3732/ajb.1500539>
- Szymański, J., Bocobza, S., Panda, S., Sonawane, P., Cárdenas, P. D., Lashbrooke, J., Kamble, A., Shahaf, N., Meir, S., Bovy, A., Beekwilder, J., Tikunov, Y., Fuente, I. R., Zamir, D., Rogachev, I., & Aharoni, A. (2020). Analysis of wild tomato introgression lines elucidates the genetic basis of transcriptome and metabolome variation underlying fruit traits and pathogen response. *Nature Genetics*, 52, 1111–1121. <https://doi.org/10.1038/s41588-020-0690-6>
- The Brazilian Flora Group. (2018). Brazilian Flora 2020: Innovation and collaboration to meet Target 1 of the Global Strategy for Plant Conservation (GSPC). *Rodriguesia*, 69, 1513–1527. <https://doi.org/10.1590/2175-7860201869402>
- Tirnaz, S., Zandberg, J., Thomas, W. J. W., Marsh, J., Edwards, D., & Batley, J. (2022). Application of crop wild relatives in modern breeding: An overview of resources, experimental and computational methodologies. *Frontiers in Plant Science*, 13, 1008904. <https://doi.org/10.3389/fpls.2022.1008904>
- van Ooijen, G., van den Burg, H. A., Cornelissen, B. J. C., & Takken, F. L. W. (2007). Structure and function of resistance proteins in solanaceous plants. *Annual Review of Phytopathology*, 45, 43–72. <https://doi.org/10.1146/annurev.phyto.45.062806.094430>
- Vendramin, V., Ormanbekova, D., Scalabrin, S., Scaglione, D., Maccaferri, M., Martelli, P., Salvi, S., Jurman, I., Casadio, R., Cattonaro, F., Tuberosa, R., Massi, A., & Morgante, M. (2019). Genomic tools for durum wheat breeding: De novo assembly of Svevo transcriptome and SNP discovery in elite germplasm. *BMC Genomics*, 20, 278. <https://doi.org/10.1186/s12864-019-5645-x>
- Veronico, P., Paciolla, C., Pomar, F., Leonardis, S., García-Ulloa, A., & Melillo, M. T. (2018). Changes in lignin biosynthesis and monomer composition in response to benzothiadiazole and root-knot nematode *Meloidogyne incognita* infection in tomato. *Journal of Plant Physiology*, 230, 40–50. <https://doi.org/10.1016/j.jplph.2018.07.013>
- Weese, T. L., & Bohs, L. (2007). A three-gene phylogeny of the genus *Solanum* (Solanaceae). *Systematic Botany*, 32, 445–463. <https://doi.org/10.1600/036364407781179671>
- Wei, Q., Wang, J., Wang, W., Hu, T. H., Hu, H. J., & Bao, C. L. (2020). A high-quality chromosome-level genome assembly reveals genetics for important traits in eggplant. *Horticulture Research*, 7, 153. <https://doi.org/10.1038/s41438-020-00391-0>
- Yang, X., Cheng, Y. F., Deng, C., Ma, Y., Wang, Z. W., Chen, X. H., & Xue, L. B. (2014). Comparative transcriptome analysis of eggplant (*Solanum melongena* L.) and turkey berry (*Solanum torvum* Sw.): Phylogenomics and disease resistance analysis. *BMC Genomics*, 15, 412. <https://doi.org/10.1186/1471-2164-15-412>
- Yang, Y., Saand, M. A., Huang, L., Abdelaal, W. B., Zhang, J., Wu, Y., Li, J., Sirohi, M. H., & Wang, F. (2021). Applications of multi-omics technologies for crop improvement. *Frontiers in Plant Science*, 12, 563953. <https://doi.org/10.3389/fpls.2021.563953>
- Yunes, J. M., & Babbitt, P. C. (2019). Effusion: Prediction of protein function from sequence similarity networks. *Bioinformatics*, 35, 442–451. <https://doi.org/10.1093/bioinformatics/bty672>
- Zhang, H., Chen, H., Tan, J., Huang, S., Chen, X., Dong, H., Zhang, R., Wang, Y., Wang, B., Xiao, X., Hong, Z., Zhang, J., Hu, J., & Zhang, M. (2023). The chromosome-scale reference genome and transcriptome analysis of *Solanum torvum* provides insights into resistance to root-knot nematodes. *Frontiers in Plant Science*, 14, 1210513. <https://doi.org/10.3389/fpls.2023.1210513>
- Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., & Song, B.-H. (2017). Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, 10, 5–24. <https://doi.org/10.1111/eva.12434>
- Zsögön, A., Cermak, T., Voytas, D., & Peres, L. E. (2017). Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: Case study in tomato. *Frontiers in Plant Science*, 256, 120–130. <https://doi.org/10.1016/j.plantsci.2016.12.012>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Silva, P. O., Silva, L. E. R., Gouveia, D. G., Rodrigues, P. M., Gonçalves, J. F. C., Nunes-Nesi, A., Araújo, W. L., & Cavalcanti, J. H. F. (2025). Exploring the genomic diversity and breeding applications of the *Solanum sessiliflorum* transcriptome via phylogenetic analysis. *Crop Science*, 65, e70036. <https://doi.org/10.1002/csc2.70036>