Transcriptomic and genomic analysis highlights the basis of ZYMV tolerance in zucchini

Ciro Amoroso
University of Naples Federico II
Giuseppe Andolfo
University of Naples Federico II
Claudio Capuozzo
University of Naples Federico II
Antimo Di Donato
University of Naples Federico II
Cecilia Martinez
University of Almería
Laura Tomassoli
Agricultural Research Council
Maria Ercolano (ercolano@unina.it)
University of Naples Federico II

Research Article

Keywords: Cucurbita pepo, RNA-seq, Transcriptome, Molecular markers, Plant recovery, Plant Virus

Posted Date: January 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-134049/v1

Transcriptomic and genomic analysis highlights the basis of ZYMV tolerance in zucchini

Amoroso CG[1], Andolfo G[1], Capuozzo C[1], Di Donato A[1], Martinez C[2], Tomassoli L[3], Ercolano MR*

[1] Department of Agricultural Science – University of Naples "Federico II" – Portici (NA)

[2] Department of Biology and Geology, Research Centers CIAIMBITAL and CeiA3, University of Almería, 04120 Almería, Spain

[3] Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria, Research Centre of Plant Control and Certification, Rome, Italy

* Corresponding author: Maria Raffaella Ercolano, Department of Agricultural Sciences, University of Naples Federico II. Via Università 100, 80055 Portici (Naples), Italy. ercolano@unina.it

Abstract

Background: *Cucurbita pepo* is high susceptible to Zucchini yellow mosaic virus (ZYMV) and the resistance found in several wild species does not provide complete or broad-spectrum resistance. In this study, a source of tolerance introgressed in *C. pepo* (381e) from *C. moschata,* in True French (TF) background, was investigated 12 days post inoculation (DPI) at transcriptomic and genomic levels.

Results: A comparative RNA-seq experiment on TF (susceptible to ZYMV) and 381e (tolerant to ZYMV), allowed to evaluate 33,000 expressed transcripts and to identify 146 differentially expressed genes (DEGs) in 381e, mainly involved in photosynthesis, transcription, cytoskeleton organization and callose synthesis. By contrast, the susceptible line True French triggered oxidative processes related to response to biotic stimulus, and key regulators of plant virus intercellular movement. Moreover, the genome mapping of transcripts allowed the identification of two chromosome regions rich in SNPs (Single Nucleotide Polymorphisms), putatively introgressed from *C. moschata,* containing genes exclusively expressed in 381e.

Conclusion: The transcriptome reprogramming of 381e ZYMV challenged revealed a globally restoring of cellular activities and a reduced virus movement and replication. Furthermore, genes putatively involved in ZYMV tolerance were detected in *C. moschata* introgressed genomic regions. Our work provides new insight into the plant virus recovery process and a better understanding of the molecular basis of 381e tolerance.

Keywords: Cucurbita pepo, RNA-seq, Transcriptome, Molecular markers, Plant recovery, Plant Virus

1. Background

Cucurbita pepo is a high polymorphic species including economically important crops grown worldwide. This species is high susceptible to Zucchini yellow mosaic virus (ZYMV), a quickly spread potyvirus aphid transmitted, leading to leaf yellowing and deformations, stunting and fruit defects [1-3]. Till now, no sources of resistance to ZYMV have been found in this species though in many areas it results one of the most serious threat limiting crop production.

A wide range of ZYMV resistant genetic resources was found in other *Cucurbita* species [4,5]. Several *C. moschata* accessions carrying resistance genes to ZYMV were identified: Nigerian local carrying *"zym0"* and *"zym4"* resistance gene [6]; Menina and Bolina possessing a resistance gene named "*Zym1"* [7]; Soler accession containing the recessive gene *"zym6"* [8] and four genes that confer resistance to ZYMV were also identified in *"OttoF1"* variety [9].

The genetic inheritance of resistance/tolerance derived from Menina in the *C. pepo* background showed that *Zym-1* gene could interact with minor genes able to modulate the defense response [6,10]. Recently, Capuozzo et al., [11] investigated *Zym-1* segregation in different *C. pepo* populations highlighting segregation anomaly left over from the original interspecific cross between *C. pepo* and *C. moschata*. Genetic analysis for tolerance to ZYMV, at both the phenotypic and genotypic level, indicated clearly that a major gene, *Zym-1* is essential for a tolerance expression in 381e genotype background but putative additional genes with additive and/or epistatic effects may mitigate or delay the symptoms. Pachner et al., reported that the combined deployment of 7 genes in *C. pepo*, would be required for maximal expression of ZYMV tolerance [12]. The involvement of putative modifying minor genes made difficult the understanding of genetic control and the subsequently the transfer of ZYMV tolerance in breeding lines. In addition, the tolerant cultivar 381e during ZYMV challenging showed a systemic symptomatic infection 4 days after inoculation (DPI) and a plant recovery, characterized by the emergence of newly developing leaves about 12 days after inoculation that can hamper genetic analysis [10].

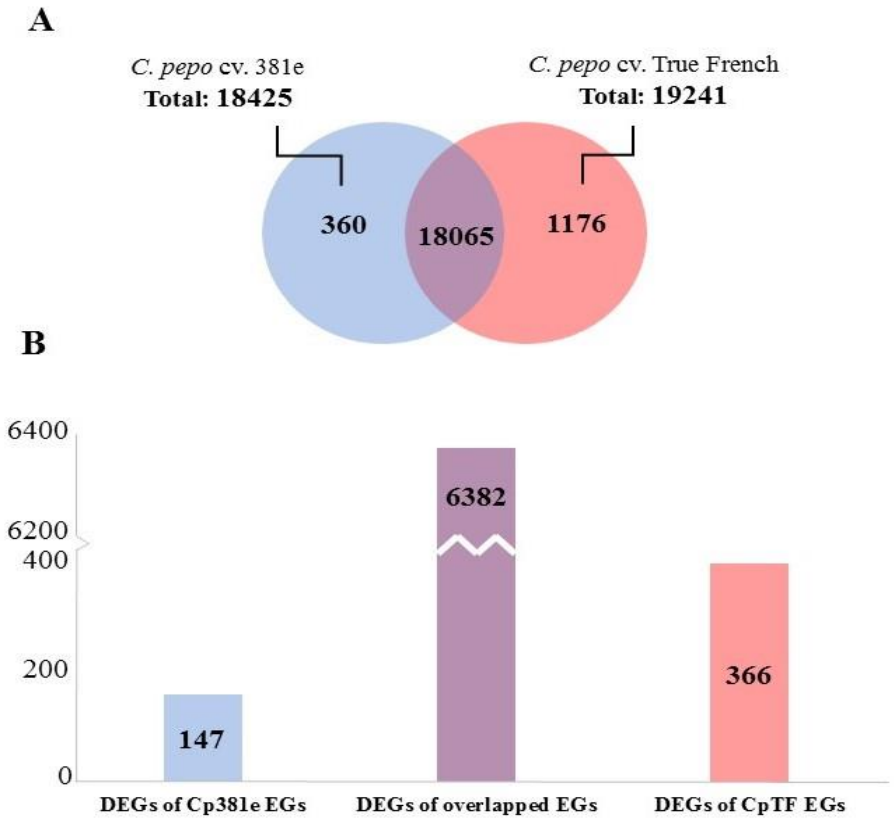
Cucurbit genomic analysis and comparative gene expression analysis resulted very useful for overcoming the difficulties related to introgression of desirable traits and for dissecting the molecular basis of host tolerance [13-15]. A reference zucchini genome version 4.1 (BGV004370) was recently provided and several *C. pepo* next-generation sequencing technologies as RNA-sequencing (RNA-seq) have been explored [16-19]. To date, RNA-seq technology provides new opportunities for mapping and quantifying transcripts, creating new opportunities for comprehension of pathways activated during plant-pathogen interaction and of genomic regions underlying the genes involved.

The main goal of this work was to investigate the mechanism of ZYMV tolerance by the means of genome-wide transcriptional analysis of the susceptible TF cultivar and of a tolerant derived line 381e. The remodulation of genome expression in virus-infected plants were assessed at 12 DPI to better highlight key genes activated in the plant recovering process in the tolerant accession. In addition, to facilitate the identification the candidate genes involved in ZYMV tolerance in *C. moschata* introgressed regions a genomic scanning was performed.

2. Results

2.1 *C. pepo* transcriptional profile upon ZYMV inoculation

The transcriptomic reprogramming of two zucchini isogenic cultivars (TF and derived ZYMV tolerant line 381e) inoculated with (ZYMV), were evaluated by analysing 33,388 genes expressed at 12 DPI (days post inoculation). In TF-ZYMV interaction, transcriptome variation resulted in 19,241 total expressed genes. As for the cultivar 381e, 18,425 total transcripts were observed upon inoculation with ZYMV (Fig 1A). Private expressed genes (360) in 381e were more than three times less than TF (1,176). Out of a total of 18,065 overlapping transcripts 6,382 were differentially expressed in the both cultivars; 147 were differentially expressed in 381e and 366 in TF (Fig. 1B). Among the 3,681 upregulated genes (FDR-adjusted p-value, 0.01), 147 resulted exclusively up regulated in 381e, suggesting that they play an important role in recovery during virus infection. The two transcriptomes virus challenged were further investigated to understand in which biological processes were involved the differentially expressed genes (DEGs) identified.



93

94 **Figure 1.** Transcriptional Profile. A) Total Expressed Genes in *C. pepo* 381e (blue), TF (orange), and overlapped
95 genes (violet). B) Differentially Expressed Genes for each area of Venn diagram.

96 2.2 Overview of cellular processes activated in the tolerant and susceptible line

97

98 According to phenotypic recovery observed at 12DPI, the tolerant cultivar 381e showed the restoring of
99 main cellular activities (transcription, protein translation, photosynthesis, metabolites biosynthesis and
100 general cellular organization) belonging to 139 GO term enriched categories. By the contrast, the susceptible
101 cultivar TF showed 32 overrepresented GO terms, principally receptors involved in response to biotic stress,
102 enzymes for ROS production, peptide degradation and lipid oxidation (Figure 2).

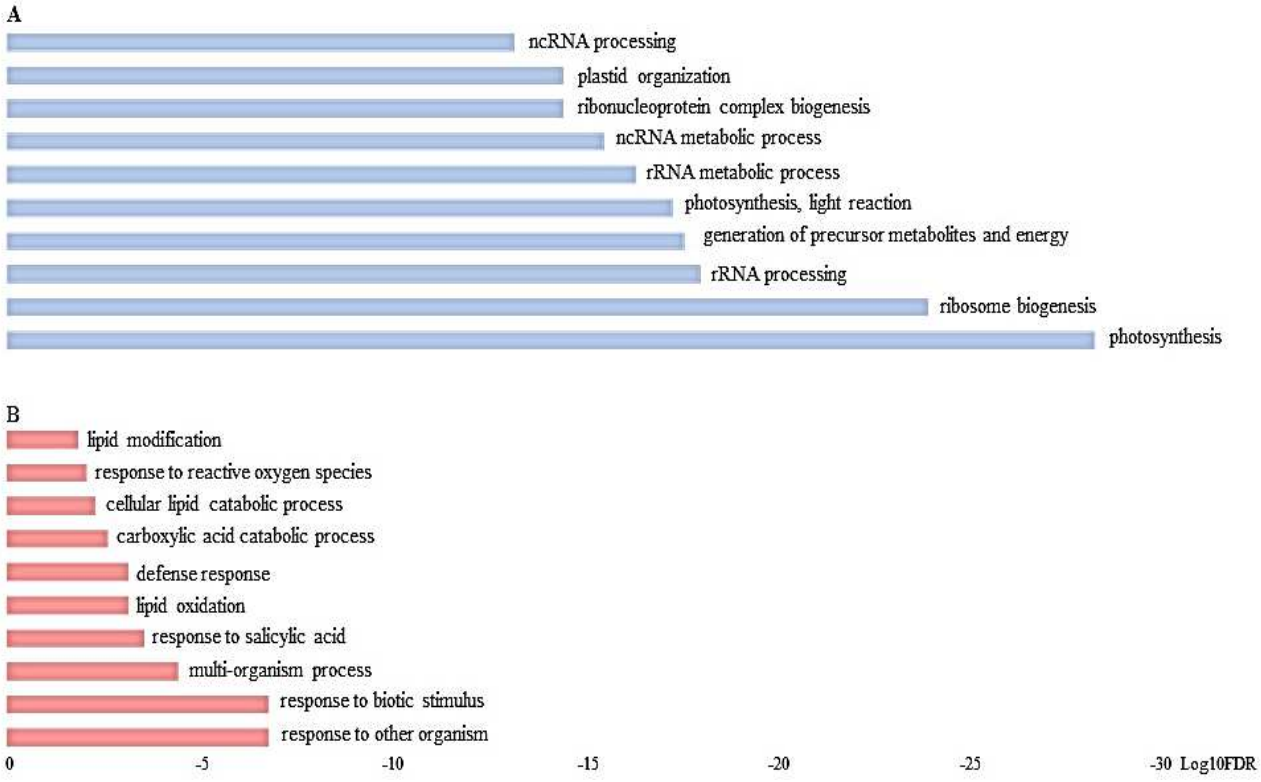


Figure 2. Top ten of non-redundant significantly enriched GO terms in biological process category. A) Upregulated genes in 381e and B) Upregulated genes in True French. The categories size is releted to LOG_{10} FDR of enriched GOs.

The plant vegetative restoring in 381e is promoted by photosynthetic activity through the up regulation of genes involved in the formation of the photosystems (i.e. Cup000001g001195.1), in the electron transport chain and in membrane transporters. In addition, genes directly involved in mitochondrial activity, in the glycolysis process (PEP, Acetyl-CoA, GAPDH, Hexokinaes) in phospholipids synthesis (CDP-alcohol phosphatidyltransferase) and in shikimic acid network synthesis, resulted strongly up regulated in 381e, suggesting an overall enhanced photo-respiratory activity. A strong up regulation of genes involved in helicase activity in 381e, including DEAD-box ATP-dependent RNA helicases, DNA replication licensing factor such as the exclusively expressed Cup000013g012986.1 and two "DNA-binding protein" (Cup000085g037850.1 Cup000022g020052.1) was observed. Moreover, several HSPs proteins were found differentially regulated in both lines. In particular, 381e activated Cup001195g045603.1 an Heat Shock protein-

(HSP 90-5) that could be involved in newly synthesized proteins transporting, from ribosomes to chloroplasts and an HSP-70 (Cup000010g010288.1), involved in proteins trafficking through the chloroplast membrane to stroma in cooperation with 14-3-3 protein (Cup000060g034360.1), that resulted overexpressed [20]. Furthermore, 7 aquaporin membrane transporters (i.e. Cup000679g045055.1) resulted strongly upregulated in 381e (LogFC >1).

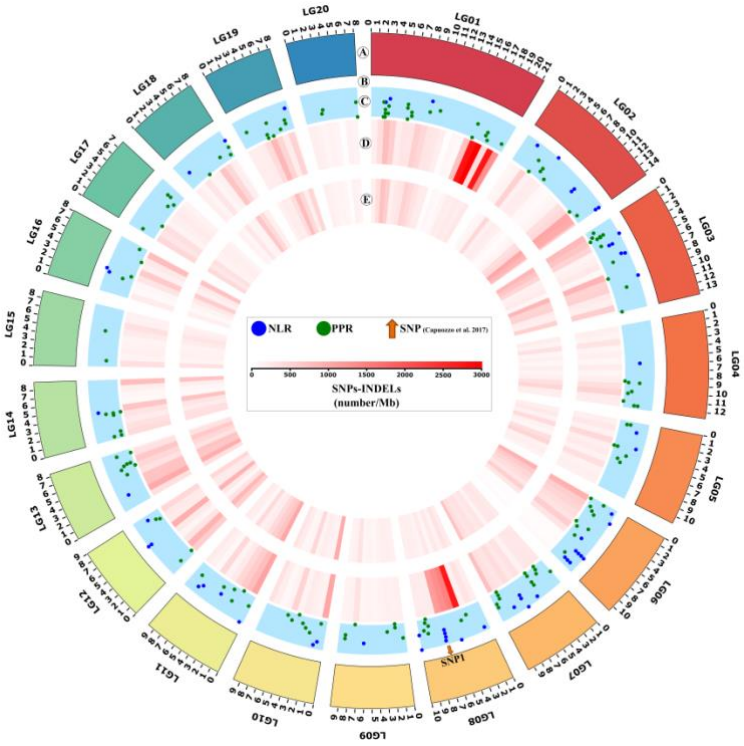
By contrast, TF showed an high activation of genes involved in the chloroplast architecture (i.e. Cup000003g002638.1), oxygen scavenging such as 2 L-ascorbate peroxidase for H_2O_2 formation (Cup000004g003828.1; Cup000052g032798.1), catalase and superoxide dismutase (i.e. Cup000032g025333.1) and a protein ridA (Cup000182g042991.1) to prevent damage from BCAT3 (Branched-chain-amino-acid aminotransferase 3) that is produced in response to stress. It is interesting to note the up regulation of several cytochromes type p450, associated with the endoplasmic reticulum, involved in metabolism of chemical compounds extraneous to the organism. Among them, a cytochrome CYP82C2 (Cup000001g001147.1), directly involved in the jasmonic acid network synthesis was high expressed, suggesting that the plant defence system is still alerted [21]. In addition, two synaptotagmin and SNAP receptors (Cup000004g004336.1; Cup000029g023659.1; Cup000011g011417.1) resulted up regulated and exclusively expressed in TF.

In our study 29 serine/threonine protein kinases, of which 22 were localized on the plasma membrane, resulted up regulated in 381e while 42 in TF. In addition, TF differentially expressed several genes involved in the cell wall degradation (ie. pectin-esterases, beta-glucosidases and Xyloglucan endo-transglycosylases), indicating the collapse of the primaries structure of the plants. On the other hand, it is interesting to note that genes involved in cytoskeleton organisation (Cup000006g006297.1; Cup000010g010383.1; Cup000021g019231.1) resulted exclusively up regulated in 381e as well as a gene involved in glycosaminoglycan biosynthesis (Cup000019g017783.1), two Rab GTP-ases (Cup000067g035924.1; Cup000038g027690.1) and a gene having an important role in the callose synthesis (Cup000021g019191.1).

2.3 Genomic localization of transcript variants

144 In order to identify the putative genomic regions introgressed into C. *pepo* from C. *moschata* cv. Menina,
145 we performed a variant calling with respect of small variants (SNVs: SNPs single nucleotide polimorfisms;
146 INDELs insertions and deletions) for the two near isogenic lines (TF and 381e).

147 A total of 150,605 high-quality small variants were mapped to the *C. pepo* BGV004370 reference genome
148 (Figure 3) [16]. The number of small variants in 381e and TF common expressed genes were normalised respect
149 to the LG (Linkage Group) length (Figure 3). In 381e, on average 8 short variants were mapped for expressed
150 locus on 1 and 8 LGs, while in TF 2 variants per transcript were found on such LGs (Additional Table 2).
151 About 30% of small variants identified in '381e' were mapped on two linkage groups, whilst in TF the
152 percentage of variants on the same LGs was close to 13% (Additional Tables 1 and 2 and Figure 3). Close-up
153 view on 1 and 8 LGs of two *C. pepo* cultivars underlined two genomic regions potentially introgressed from
154 resistant *C. moschata* cv. Menina (Figure 4). Indeed, the 20% (16,891) of small variants identified in 381e were
155 concentrated in two short regions of 1 and 8 LGs, while in TF less than 3% (1,410) were mapped in the same
156 areas. Over 65% (10,321) of small variants identified on LG01 were located on a genomic region of about 4 Mb
157 in 381e. The second putative introgression region of 3,5 Mb showed 82% (6,570) of total variants mapped on
158 LG08 in 381e. Interestingly, Capuozzo et al., have mapped on the same region of LG08 the marker (SNP1)
159 associated to ZYMV tolerance in '381e' (Figure 4 and Additional Table 2) [11].



160

161 **Figure 3**. Circos plot integrating the genomic positions of ZYMV-related markers, pathogen recognition genes
162 and small variants (SNP: single nucleotide polymorphisms and InDel: insertion and deletion). Inset legend
163 providing information represented by each data ring. Track A denotes the 20 pseudochromosomes (LG:
164 linkage group) of *Cucurbita pepo*. The length of each circle segment represents the size of pseudochromosomes
165 expressed in megabases (Mb). B) The marker associated to ZYMV tolerance identified by Capuozzo et al. [11]
166 is indicated by orange arrow. C) Genomic positions of nucleotide binding like proteins (NLR; blue spots) and
167 receptor like proteins and receptor like kinase (PPR; green spots), annotated by Andolfo et al. [19] , are shown
168 as scatterplots. Heatmaps for the density levels (number/Mb) of small variants identified in (D) *C. pepo* 381e
169 and (E) *C. pepo* TF cultivars.

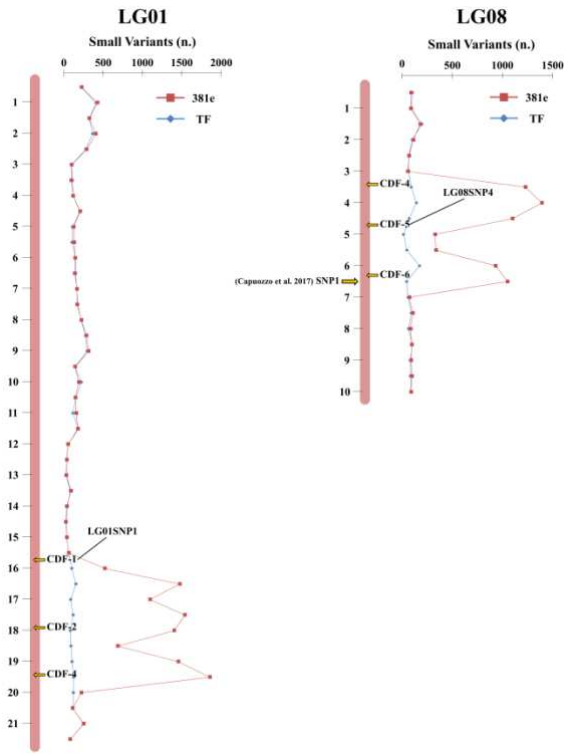


Figure 4. High resolution map integrating genomic and genetic marker information. The distribution of small variants along 1 and 8 pseudochromosomes was reported (381e in red and TF in blue). The genomic position of six coding DNA fragments (CDFs) and two CAPS markers (LG01SNP1 and LG08SNP4) was indicated.

2.4 Variants validation and marker-phenotype correlation

To confirm the DNA variants identified on putative introgression regions located on LG01 and LG08 a molecular validation was performed. A total of six coding DNA fragments (CDFs) located on two genomic regions inherited from *C. moschata* were sequenced and analysed (Figure 4). The about 70 SNP calling-related variants annotated into the 6 CDFs were confirmed from Sanger sequencing (Additional Table 3). Two high confidence variants (LG08SNP4 and LG01SNP1) were converted in CAPS markers (Additional Table 3; Additional Figure 1) to conduct a marker-phenotype correlation analysis on a subset of the F2 segregating population (381e x TF) phenotyped by Capuozzo et al. [11]. LG08SNP4 was physically close (~ 2 Mb) to the previously mapped "SNP1" and exhibited a very high correlation coefficient (Pearson test p-value << 0,01; r=0,991) with phenotypic data [11]. In particular, marker correlation with ZYMV-resistance/susceptibility traits

was in accordance for 73 genotypes out of 82 genotypes analysed. A good correlation coefficient with phenotypic results was displayed also by LG01SNP1 (r=0,827; Pearson test p-value < 0,05). In addition, an interesting correlation trend (r=0,895; Pearson test p-value < 0,05) was identified between the LG01SNP1 marker located on chromosome 1 and LG08SNP4 marker located on chromosome 8.

3. Discussion

From its first report in the late 1970s, ZYMV continued its worldwide expansion, causing cucurbits important yield losses. In this work, we explored the molecular basis of ZYMV tolerance identified in *Cucurbita moschata* cv Menina analysing the *C. pepo* derived tolerant cv 381e and its susceptible counterpart TF, through an integrated transcriptomic and genomic experiment. Though *C. moschata* is sparingly cross-fertile with *C. pepo*, the gene expression may be hampered in new genomic background [6]. In addition, approximately at 12 days after ZYMV inoculation, 381e plants showed a vegetative recovery [10], while damaged TF plants displayed yellowing of foliage, internodes shortening and leaf deformations. In our experiment, an opposite response to infection at 12 DPI in terms of expressed and differentially regulated genes between the two analyzed contrasting genotypes was clearly observed. ZYMV virus usually takes about 9 days to infect the plant systemically through the phloem [22]; and the early systemic infection of 381e, about 4 DPI, was able to prompt a plant recovery. At initial, stage of infection, the plant moderates its immune response and the general fitness, triggering to a decreasing in symptoms of viral infection [23]. The 381e physiological perturbations imposed various constraints to face virus and limiting the synthesis of host proteins essential for replication and movement. Tolerant plants react to virus infection through changes in physiological and developmental processes to limit the synthesis of factors required for virus multiplication [24-26]. We found several DEAD-box RNA helicases and DNA helicase such as Cup000013g012986.1 up regulated in 381e. DEAD-box RNA helicases can function as viral RNA sensors or as effectors by blocking RNA virus replication. RH30 (Dead-box helicase) in tomato was required to restrict bushy stunt virus (TBSV) replication [27]. Moreover, RNA and DNA helicases can also exert a positive role in stress tolerance [28]. A general gene down regulation could be useful to impede virus movement and replication, leading to a less virus titer and promoting a consequent

plant recovery [23]. By contrast, the up regulation of factors correlated with transcription process observed in TF plants may facilitate the use the host nuclear factors for ZYMV replication [29].

The response to viral infection in 381e required a global host transcriptome reprogramming to avoid the development of most severe symptoms. Indeed, an up-regulation of genes involved in energy production and cell repairing at 12 dpi was observed. In our experiment, the tolerant cultivar 381e showed an increased photosynthetic and photo-respiratory activity to promote the recovery and symptom alleviation. The partially resistant cultivar Jaguar, 15 day after ZYMV inoculation, showed a strong up regulation several proteins involved in photosystems activities [30]. Interestingly, we found a high up regulated gene coding for an HSP90.5 (Cup001195g045603.1) with a LogFC of 4.6 privately expressed in 381e. A properly control of HSP90.5 expression is required for plant growth and development and is especially essential for the chloroplast thylakoid formation and the regulation of the receptor proteins involved in plant immunity [31-33]. In addition, the overexpressed HSP70 (Cup000010g010288.1) cooperate with the overexpressed 14-3-3 chaperone (Cup000060g034360.1) to keep the chloroplast or mitochondrial precursor proteins in an unfolded state, acting as a motor for transporting the precursor proteins interacting with the plasmodesmata in translocation pathway [34,35]. By contrast, the TF photosynthetic system was affected by the down regulation of the genes involved in chloroplasts and thylakoids formation.

Symptom recovery is generally accompanied by activation of pathogen receptor genes and other defence-related genes during the plant-pathogen interaction [23,36,37]. In particular, we found a resistance protein (CNL) annotated by Andolfo et al., [19] and located on linkage group 8 scaffold 3 (Cup000003g003423.1) up regulated and privately expressed (LogFC 1.51) in 381e [19,38,39]. It is interesting to note that in such transcript we found a variant that could lead to functional modification. In our study, numerous genes codifying for serine/threonine kinases proteins differentially expressed in both lines were also found. These proteins play a central role in signalling during the pathogen's recognition and the subsequent activation of plant defence mechanisms. A serine/threonine kinases belonging to RIO protein family, down regulated and privately expressed in TF (Cup000024g021133.1), were located on linkage group one (scaffold 24) corresponding to one of regions putative introgressed from Menina in 381e. These proteins are generally involved in biogenesis of

small ribosomal subunits and were found to interact with the tomato mosaic virus (ToMV) movement protein (MP) to promote virus movement [40,41].

The ZYMV virus manages to enter the cells through wounds or transported by aphids and moves from cell to cell through the plasmodesms to reach the phloem where induce systemic infections. A delicate equilibrium between RNA silencing and virus counter-defense responses in recovered leaves may help maintain virus at levels below the threshold required for symptom induction. In TF, we found the up regulation of a CYP82C2 (Cup000001g001147.1) that in Arabidopsis have a role in JA-induced defense-related genes [42]. We also found exclusively expressed in TF two synaptotagmins (SYTA) that are key regulators of plant virus intercellular movement, necessary for the ability of diverse cell-to-cell movement proteins and able to regulate endocytosis and protein-mediated trafficking through plasmodesmata [43,44].

Genes involved in cytoskeleton organization, vesicular traffic and in callose deposition such as Cup000021g019191.1 (LogFC 1.59) could affect the containment of the virus spread from cell to cell regulating plasmodesmal permeability, resulted high activated in 381e [45,46]. By contrast, an upregulation of genes involved in cell structure degradations (such as the pectin esterases) was observed in TF.

Comparing the two lines at genomic level, having nearby the same makeup, since 381e is a backcross line derived from TF, we identify two regions in the tolerant line 381e enriched in SNPs. These regions could be considered as introgressed from Menina and so putatively involved in the tolerance. The level of resistance of 381e to ZYMV is not as high as that of *C. moschata cv.* Menina, and some other gene could be involved in the resistance regulation [6,47]. A marker located on chromosome 8, showing 90% co-segregation with resistance in a F2 population derived from the cross between 381e and TF, was previously identified [11]. The correlation of LG08SNP4 with resistance trait in F2 population (381e x TF) was also very high, supporting the finding of Capuozzo et al., [11]. Whilst the lower correlation of LG01SNP1with resistant phenotypes suggests that genes with minor effects on the virus tolerance expression could be located there.

4. Conclusion

In conclusion, the network of expressed genes at 12 days after virus infection in 381e revealed recovery of physiological processes, defense response activation, virus replication and movement limiting. Two introgressed regions from Menina in 381e, contain NBS genes, serine/threonine and DEAD-box RNA helicases, could be involved in tolerance to ZYMV. In particular, on linkage group eight might be located a major gene responsible for tolerance and on linkage group one a minor gene with additive effect. Further efforts should be performed to assess the contribution of different loci to ZYMV tolerance.

5. Materials and Methods

5.1 Plant material and ZYMV-inoculation

Two zucchini near-isogenic lines: TF (susceptible to ZYMV) and 381e (resistant to ZYMV) reported in additional Figure 2, were used for RNA-seq experiment and DNA sequencing. The latter variety had been derived from six generations of backcrossing to TF, selecting for resistance, followed by four successive generations of self-pollination and selection for resistance [10]. Seeds of parental were sown in multi-cellular trays consisting of 4 cm diameter pots that had been filled with peat, one per pot. After the seedlings were transferred into pots 15 cm diameter and were grown under glasshouse condition at 22-24°C using supplement lighting to maintain 12 hours photoperiod at the Research Centre for Plant Pathology (CRA-PAV) in Rome. Each individual plant was assigned an identification number. An isolate of ZYMV from a naturally infected plant of field-grown summer squash was used for experiments. The isolate causes the typical symptoms of ZYMV disease including yellow mosaic, vein banding, blistering and malformation of leaves. Symptomatic leaves of artificially infected zucchini plants were crushed, and the raw juice was extracted at a ratio 1:10 w/v in 0.1 M phosphate buffer pH 7.2. Both cotyledons of test plants, sprinkled with the abrasive powder “celite”, were inoculated with approximately 20 µl of the diluted extract, and subsequently washed with distilled water. Finally, the DNA of a F2 segregating population (381e x TF) obtained by Capuozzo et al., was used for molecular analysis [11].

284 5.2 Sample Collection and nucleic acid isolation

285

286 Twelve days after treatment, infected zucchini leaf samples of three independent replicates were
287 collected. Leaves were removed from the plants, weighed and immediately frozen in liquid nitrogen and
288 stored at −80°C. RNA was isolated using the RNAeasy Plant Kit according to the manual instructions (Quiagen
289 Valencia, USA). RNA sample concentration was determined using a NanoDrop ND-1000 Spectrophotometer
290 (Nano-Drop Technologies, Wilmington, DE, USA) and by Bioanalyzer (Agilent Technologies). RNA integrity
291 was checked by horizontal electrophoresis on a 1.2% (w/v) agarose gel. Two μg of each sample were prepared
292 with 20 μl of 10 X RNA Loading Buffer composed by 400 μl Formamide, 10 μl 37% formaldeyde, 2 μl loading
293 buffer 10X (50% glycerol, 0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol; Sigma) and 1 μl of 10 mg/μl
294 SYBR Safe DNA Gel Stain (Invitrogen). Gel visualization was performed using UV light (UV Gel Doc
295 BIORAD). DNA was extracted from leaves of the 381e and TF not infected controls using DNeasy Plant Mini
296 Kit (Qiagen). DNA sample concentration was determined using a NanoDrop ND-1000 Spectrophotometer
297 (Nano-Drop Technologies, Wilmington, DE, USA).

298 In order to obtain high quantity of nucleid acid, total RNA from the six replicates was pooled in one
299 sample, obtaining three independent replicas for sequencing. Total purified RNA was converted to cDNA
300 libraries (QuantiTect Reverse Transcription Kit, Quiagen) and sequenced on Illumina HiSeq1500 platform at
301 the LabMedMolGe (Laboratory of Molecular Medicine and Genomics Department of Medicine and Surgery,
302 University of Salerno). The process generates millions (36 M) of short (100 bp) reads sequenced from both ends
303 of each cDNA fragment (paired-end sequencing).

304

305 5.3 RNA-Sequencing data analysis

306

307 The reads obtained from illumina sequencing, were analysed using the online platform A.I.R.
308 (https://transcriptomics.sequentiabiotech.com). FastQ files uploaded on A.I.R. platform were assessed by PCA
309 analysis both to recognize the file pairs and evaluate the quality of the samples. After the validation an

experimental label was provided, and the option pair-end sequences analysis was chosen. FastQ files of inoculated and not inoculated condition for each sample were moved in the corresponding sections and mapped against the *C. pepo* reference genome. Trimming of the reads, trimmed quality analysis, mapping, BAM quality check and statistical analysis was provided by the platform. Clean FastQ files were analyzed by DESeq2 algorithm to obtain DEGs. The DEGs list obtained from DESeq2 analysis, the PCA chart and other general statistics data and lists, .bam and .bai files were downloaded. Genes having an FDR < 0.05 in DEGs list was filtered to conduct further analyses. The *C.pepo* reference genome and relative gene annotation were reported by Montero-Pau et al., (2017) and Andolfo et al., (2017) respectively [16,19]. Gene Ontology Enrichment Analysis was performed using AgriGO version 2.0, REViGO was used to remove redundant GO terms [48,49]. Genes were then considered significantly differentially expressed if the false discovery rate (FDR) of the statistical test was less than 0.05.

5.4 Variant calling and genome plot construction

In order to detect small variants: SNPS and INDEL less than 50 base pairs on the transcripts of the susceptible 'TF' and resistant 381e accessions, a SNP calling analysis against the *Cucurbita pepo* reference genome was performed. SNP calling, based on the two filtered .bam files, produced two VCF (Variant Call Format) files (http://www.broadinstitute.org/igv/viewing_vcf_files) with the SNPs of each genotype in comparison to the genome. To obtain the SNPs between both genotypes we found the common and unique SNPS between both VCF files, removing at the same time SNPs at low quality/coverage and inconsistent SNPs. Finally, the circular multi-track plot was carried out using the 'RCircos' R software package [50]. Vcf files were used to obtain .bed files using bedtools, and *C. pepo* genome released by Montero-Pau et al. was used as reference [16].

5.5 Validation of predicted variants

A pool of 6 coding DNA regions was selected to perform molecular validation in 381e and TF. PCR was executed with 25 ng of genomic or complementary DNA, 10 pmol primers, 1 U of Taq DNA polymerase Kit (Invitrogen, Carlsbad, CA, USA), 10 pmol dNTPs, and 2 mM MgCl2 in 25 µl reaction volumes. Amplification was performed using the following cycling conditions: 1 min at 94 °C, followed by 30 cycles of 1 min at 94 °C, 1 min 30 s at 60 °C and 2 min at 72 °C, with a final extension for 7 min at 72 °C. Amplicons were separated by electrophoresis on agarose gel (1.5 %), and photographed by a GelDoc apparatus. Primers were designed with Primer3 (http://frodo.wi.mit.edu), with a length between 18 and 27 bp. The length of the amplified fragments ranged from 300 to 1,000 bp, and the Tm of the specific primers was 59 °C for all pairs of primers (Online Resource S4). Amplicons were sequenced using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and run on automated DNA sequencers (ABI PRISM 3100 DNA Sequencer, Applied Biosystems). Sequence data deriving from *C. pepo* v4.1 reference genome were aligned with corresponding sequences originated from amplicons using MUSCLE 3.6 [51].

Additional Materials: Additional Table 1: Number of short variants (SNV and InDel) identified in TF and 381e. Additional Table 2: Number of short variants (SNV and InDel) identified in TF and 381e on LG01 and LG08. Additional Table 3: Validation of 6 polymorphic loci located on putative genomic regions introgressed from Menina. (1): List of expressed genes; (2): Differentially expressed genes (DEGs); (3): Gene ontology annotation (GO); (4): Up regulated genes, (5); Down regulated genes. Additional Figure 1: Gel electrophoretic separation of digested PCR products. Additional Figure 2: Comparison of the ZYMV tolerant line 381e (left) and the susceptible True French (right) 12 days after ZYMV inoculation

Declarations

Data Availability: The data set supporting the results of this article is available in the Sequence Read Archive (SRA) repository of NCBI under the GEO accession number: GSE159530.

Author Contributions: Conceptualization, M.R.E.; methodology, G.A, C.C and A.D.D; validation, C.G.A; formal analysis, C.G.A, G.A, C.C, A.D.D, C.M; investigation, L.T; resources Supervision GA and AD.D; M.R.E.; writing—original draft preparation C.G.A, G.A and M.R.E; writing—review and editing, C.M and L.T; project funding acquisition, M.R.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of University and Research (GenHORT project).

Acknowledgments: We thank La Semiorto Sementi S.r.l. for plant material and Romanos Zois, Erasmus student from Wageningen University (NL), for support in molecular analysis.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ZYMV	Zucchini Yellow Mosaic Virus
DPI	Days Post Inoculation
DEGs	Differentially Expressed Genes
SYTA	Synaptotagmin
CNL	Coiled-Coil Nucleotide binding domain Leucine rich repeat
TF	True French
FDR	False Discovery Rate
GO	Gene Ontology
ROS	Reactive Oxygen Species
PEP	Phosphoenolpyruvate carboxylase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HSP	Heat Shock Protein
SNV	Single Nucleotide Variants
INDEL	Insertion and Deletion
LG	Linkage Group
NLR	Nucleotide binding Like Protein
PPR	Proteins and Receptor like Kinase
CDF	Coding DNA Fragments
NBS	Nucleotide Binding Site

References

1. Castle SJ, Perring TM, Farrar CA, Kishaba A. Field and laboratory transmission of watermelon mosaic virus 2 and zucchini yellow mosaic virus by various aphid species. Phytopatol. 1992; *82*: 235-240
2. Desbiez C, Lecoq H. Zucchini yellow mosaic virus. Plant Path. 1997**;** *46*: 809-829.doi:10.1046/j.1365-3059.1997.d01-87.x
3. Katis NI, Tsitsipis JA, Lykouressis DP, Papapanayotou A, Margaritopoulos JT, Kokinis GM, Perdikis DC, Manoussopoulos IN. Transmission of zucchini yellow mosaic virus by colonizing and non-colonizing aphids in Greece and new aphid species vectors of the virus. Journal of Phytopathology. 2006; *154*: 293–302.
4. Robinson RW, Weeden NF, Provvidenti R. Inheritance of resistance to zucchini yellow mosaic virus in the interspecific cross Cucurbita maxima x C. ecuadorensis. Cucurbit Genetics Cooperative Report**.** 1988; *11*: 74-75
5. Leibman D, Wolf D, Saharan V, Zelcer A, Arazi T, Yoel S, Gaba V, Gal-On A. A high level of transgenic viral small RNA is associated with broad potyvirus resistance in cucurbits. Mol. Plant–Microbe Interact. 2011; *24*: 1220–1238.
6. Pachner M, Paris H.S, Lelley T. Genes for resistance to zucchini yellow mosaic in tropical pumpkin. J Hered**.** 2011; *102*: 330-335. doi:10.1093/jhered/esr006
7. Paris HS, Cohen S, Burger Y, Yoseph R. Singlegene resistance to zucchini yellow mosaic virus in Cucurbita moschata. Euphytica. 1988; *37*: 27-29.
8. Pachner M, Lelley T. Different genes for resistance to zucchini yellow mosaic virus (ZYMV) in Cucurbita moschata. In: Lebeda A, Paris H.S, (eds) Progress in cucurbit genetics and breeding research. Proc Cucurbitaceae. 2004; 8th EUCARPIA meeting on cucurbit genetics and breeding, Olomouc, 237–243
9. Nacar Ç, Fidan H, Ekbiç E, Aras V, Denli, N, Keles D. Development of suitable sources of resistance to ZYMV in Cucurbita pepo. Cucurbitaceae. 2012; Proceedings of the 10th EUCARPIA Meeting on Genetics and Breeding of Cucurbitaceae, Antalya, Turkey, 15-18 October, 633-637
10. Paris HS, Cohen S. Oligogenic inheritance for resistance to Zucchini yellow mosaic virus in Cucurbita pepo. Annals of Applied Biology. 2000; *136*: 209-214. doi:10.1111/j.1744-7348.2000.tb00027.x
11. Capuozzo C, Formisano G, Iovieno P, Andolfo G, Tomassoli L, Barbella MM, Pico B, Paris HS, Ercolano MR. Inheritance analysis and identification of SNP markers associated with ZYMV resistance in Cucurbita pepo. Mol Breeding. 2017; *37*: 99. DOI 10.1007/s11032-017-0698-5
12. Pachner M, Paris HS, Winkler J, Lelley T. Phenotypic and marker-assisted pyramiding of genes for resistance to zucchini yellow mosaic virus in oilseed pumpkin (*Cucurbita pepo*). Plant Breed. 2015; *134*: 121-128. doi:10.1111/pbr.12219

13. Anagnostou K, Jahn M, Perl-Treves R. Inheritance and linkage analysis of resistance to zucchini yellow mosaic virus, watermelon mosaic virus, papaya ringspot virus and powdery mildew in melon. Euphytica. 2000; *116:* 265–270. doi.org/10.1023/A:1004005716806

14. Park Y, Katzir N, Brotman Y, King J, Bertrand F, Havey M. Comparative mapping of ZYMV resistances in cucumber (*Cucumis sativus* L.) and melon (*Cucumis melo* L.). Theor Appl Genet. 2004; *109*: 707–712. doi.org/10.1007/s00122-004-1684-y

15. Iovieno P, Andolfo G, Schiavulli A, Catalano D, Ricciardi L, Frusciante L, Ercolano MR, Pavan S. Structure, evolution and functional inference on the *Mildew Locus O* (*MLO*) gene family in three cultivated Cucurbitaceae *spp*. BMC Genomics. 2015; *16:* 1112. https://doi.org/10.1186/s12864-015-2325-3

16. Montero-Pau J, Blanca J, Bombarely A, Ziarsolo P, Esteras C, Martí-Gómez C, Ferriol M, Gómez P, Jamilena M, Mueller L, Picó B, Cañizares J. De novo assembly of the zucchini genome reveals a whole-genome duplication associated with the origin of the *Cucurbita* genus. Plant Biotechnol J. 2018; *16*: 1161-1171. doi:10.1111/pbi.12860

17. Wyatt L, Strickler S, Mueller L, Mazourek M. An acorn squash (*Cucurbita pepo* ssp. *ovifera*) fruit and seed transcriptome as a resource for the study of fruit traits in *Cucurbita*. Hortic Res. 2015; *2*:14070. https://doi.org/10.1038/hortres.2014.70

18. Xanthopoulou A, Psomopoulos F, Ganopoulos I, Manioudaki M, Tsaftaris A, Nianiou-Obeidat I, Panagiotis M. De novo transcriptome assembly of two contrasting pumpkin cultivars. Genomics Data. 2016; *7*: 200–201. https://doi.org/10.1016/j.gdata.2016.01.006.

19. Andolfo G, Di Donato A, Darrudi R, Errico A, Aiese Cigliano R, Ercolano MR. Draft of Zucchini (Cucurbita pepo L.) Proteome: A Resource for Genetic and Genomic Studies. Front Genet. 2017; *8:* 181

20. Inoue H, Li M, Schnell DJ. An essential role for chloroplast heat shock protein 90 (Hsp90C) in protein import into chloroplasts. Proc Natl Acad Sci U S A. 2013; *110*: 3173-3178. doi:10.1073/pnas.1219229110

21. Rajniak J, Barco B, Clay N, Sattely E. A new cyanogenic metabolite in *Arabidopsis* required for inducible pathogen defence. Nat. 2015; *525*: 376–379. doi.org/10.1038/nature14907

22. Zellnig G, Pöckl MH, Möstl S, Zechmann B. Two- and three-dimensional characterization of Zucchini Yellow Mosaic Virus induced structural alterations in *Cucurbita pepo* L. plants. J Struct Biol. 2014; *186*: 245–252.

23. Bengyella L, Waikhom SD, Allie F, Rey C. Virus tolerance and recovery from viral induced symptoms in plants are associated with transcriptome reprograming. Plant Mol Biol. 2015; *89*: 243-252. doi:10.1007/s11103-015-0362-6

24. Hanley-Bowdoin L, Bejarano E, Robertson D, Mansoor S. Geminiviruses: masters at redirecting and reprogramming plant processes. Nat Rev Microbiol. 2013; *11*: 777–788. doi.org/10.1038/nrmicro3117

25. Ghoshal B, Sanfaçon H. Symptom recovery in virus-infected plants: Revisiting the role of RNA silencing mechanisms. Virology. 2015; 479-480: 167-179. doi.org/10.1016/j.virol.2015.01.008.

26. Nigam D, LaTourrette K, Souza Pedro FN, Garcia-Ruiz H. Genome-Wide Variation in Potyviruses. Front in Plant Sci. 2019; *10*: 1439. doi=10.3389/fpls.2019.01439

27. Wu CY, Nagy PD. Blocking tombusvirus replication through the antiviral functions of DDX17-like RH30 DEAD- box helicase. PLoS Pathog. 2019; 15:128. doi.org/10.1371/journal.ppat.1007771

28. Pandey S, Prasad A, Sharma N, Prasad M. Linking the plant stress responses with RNA helicases. Plant Sci. 2020; *299*. doi.org/10.1016/j.plantsci.2020.110607.

29. Kumar S, Karmakar R, Gupta I, Patel AK. Interaction of potyvirus helper component-proteinase (HcPro) with RuBisCO and nucleosome in viral infections of plants. Plant Physiol Biochem. 2020; *151*: 313–322

30. Nováková S, Flores-Ramírez G, Glasa M, Danchenko M, Fiala R, Skultety L. Partially resistant Cucurbita pepo showed late onset of the Zucchini yellow mosaic virus infection due to rapid activation of defense mechanisms as compared to susceptible cultivar. Front Plant Sci. 2015; *6*: 263. doi:10.3389/fpls.2015.00263

31. Oh SE, Yeung C, Babaei-Rad R, Zhao R. Cosuppression of the chloroplast localized molecular chaperone HSP90.5 impairs plant development and chloroplast biogenesis in Arabidopsis. BMC Res Notes. 2014; *7*:643 doi.org/10.1186/1756-0500-7-643

32. Hubert DA, Tornero P, Belkhadir Y, Krishna P, Takahashi A, Shirasu K, Dangl JL. Cytosolic HSP90 associates with and modulates the *Arabidopsis* RPM1 disease resistance protein. The EMBO Journal. 2003; *22*: 5679-5689. doi:10.1093/emboj/cdg547

33. Liu Y, Burch-Smith T, Schiff M, Feng S, Dinesh-Kumar SP. Molecular chaperone Hsp90 associates with resistance protein N and its signalling proteins SGT1 and Rar1 to modulate an innate immune response in plants. J Biol Chem. 2004; *279*: 2101-2108. doi:10.1074/jbc.M310029200

34. Aoki K, Kragler F, Xoconostle-Cazares B, Lucas WJ. A subclass of plant heat shock cognate 70 chaperones carries a motif that facilitates trafficking through plasmodesmata. Proc Natl Acad Sci U S A. 2002; *99*: 16342-16347. doi: 10.1073/pnas.252427999.

35. Zhang XP, Glaser E. Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone, Trends Plant Sci. 2002; *7*: 14–21.

36. Osuna-Cruz CM, Paytuvi-Gallart A, Di Donato A, Sundesha V, Andolfo G, Aiese Cigliano R, Sanseverino W, Ercolano MR. PRGdb 3.0: a comprehensive platform for prediction and analysis of plant disease resistance genes, Nucleic Acids Research. 2018; *46*: 1197–1201. doi.org/10.1093/nar/gkx1119

37. Niehl A, Wyrsch I, Boller T, Heinlein M. Double-stranded RNAs induce a pattern-triggered immune signaling pathway in plants. New Phytol. 2016; *211*: 1008-1019. doi:10.1111/nph.13944

38. Andolfo G, Jupe F, Witek K, Etherington G, Ercolano MR, Jones JDG. Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. BMC Plant Biol. 2014; *14***:** 120. doi.org/10.1186/1471-2229-14-120

39. Andolfo G, Di Donato A, Chiaiese P, De Natale A, Pollio A, Jones JDG, Frusciante L, Ercolano MR. Alien Domains Shaped the Modular Structure of Plant NLR Proteins. Genome Biology and Evolution. 2019; *11*: 3466–3477. doi.org/10.1093/gbe/evz248

40. Gao Q, Xu S, Zhu X, Wang L, Yang Z, Zhao X. Genome-wide identification and characterization of the RIO atypical kinase family in plants. Genes Genom. 2018; *40*: 669–683. doi.org/10.1007/s13258-018-0658-4

41. Yoshioka K, Matsushita Y, Kasahara M, Konagaya K, Nyunoya H. Interaction of tomato mosaic virus movement protein with tobacco RIO kinase. Mol Cells. 2004; *17*:223–229

42. Uchiyama A, Shimada-Beltran H, Levy A, Zheng JY, Javia PA, Lazarowitz SG. The Arabidopsis synaptotagmin SYTA regulates the cell-to-cell movement of diverse plant viruses, Front in Plant Sci. 2014; *5*: 584

43. Lewis DJ, Lazarowitz SG. Arabidopsis synaptotagmin SYTA regulates endocytosis and virus movement protein cell-to-cell transport. Proceedings of the National Academy of Sciences. 2010; *107*: 2491-2496. doi:10.1073/pnas.0909080107

44. Li W, Zhao Y, Liu C, Yao G, Wu S, Hou C, Zhang M, Wang D. Callose deposition at plasmodesmata is a critical factor in restricting the cell-to-cell movement of *Soybean mosaic virus*. Plant Cell Rep. 2012**;** *31*: 905–916. doi.org/10.1007/s00299-011-1211-y

45. Bo X, Yunfei D, Masahiro M, Kanaoka KO, Zonglie H. Expression of Arabidopsis callose synthase 5 results in callose accumulation and cell wall permeability alteration. Plant Sci. 2012; *183*: 1-8. doi.org/10.1016/j.plantsci.2011.10.015

46. Liu F, Jiang H, Ye S, Chen W, Liang W, Xu Y, Sun B, Sun J, Wang Q, Cohen JD, Li C. The *Arabidopsis* P450 protein CYP82C2 modulates jasmonate-induced root growth inhibition, defense gene expression and indole glucosinolate biosynthesis. Cell Res. 2010**;** *20*: 539–552. doi.org/10.1038/cr.2010.36

47. Paris HS, Brown R. The Genes of Pumpkin and Squash. HortSci. 2005; *40*: 1620-1630.

48. Tian T, Yue L, Hengyu Y, Qi Y, Xin Y, Zhou D, Wenying X, Zhen S. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. 2017; gkx382. doi: 10.1093/nar/gkx382

49. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLoS ONE. 2011; *6*. doi.org/10.1371/journal.pone.0021800

50. Zhang H, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. BMC Bioinformatics. 2013; *14*: 244. doi.org/10.1186/1471-2105-14-244

51. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acid Research. 2004; *32*: 1792-97.

Figures

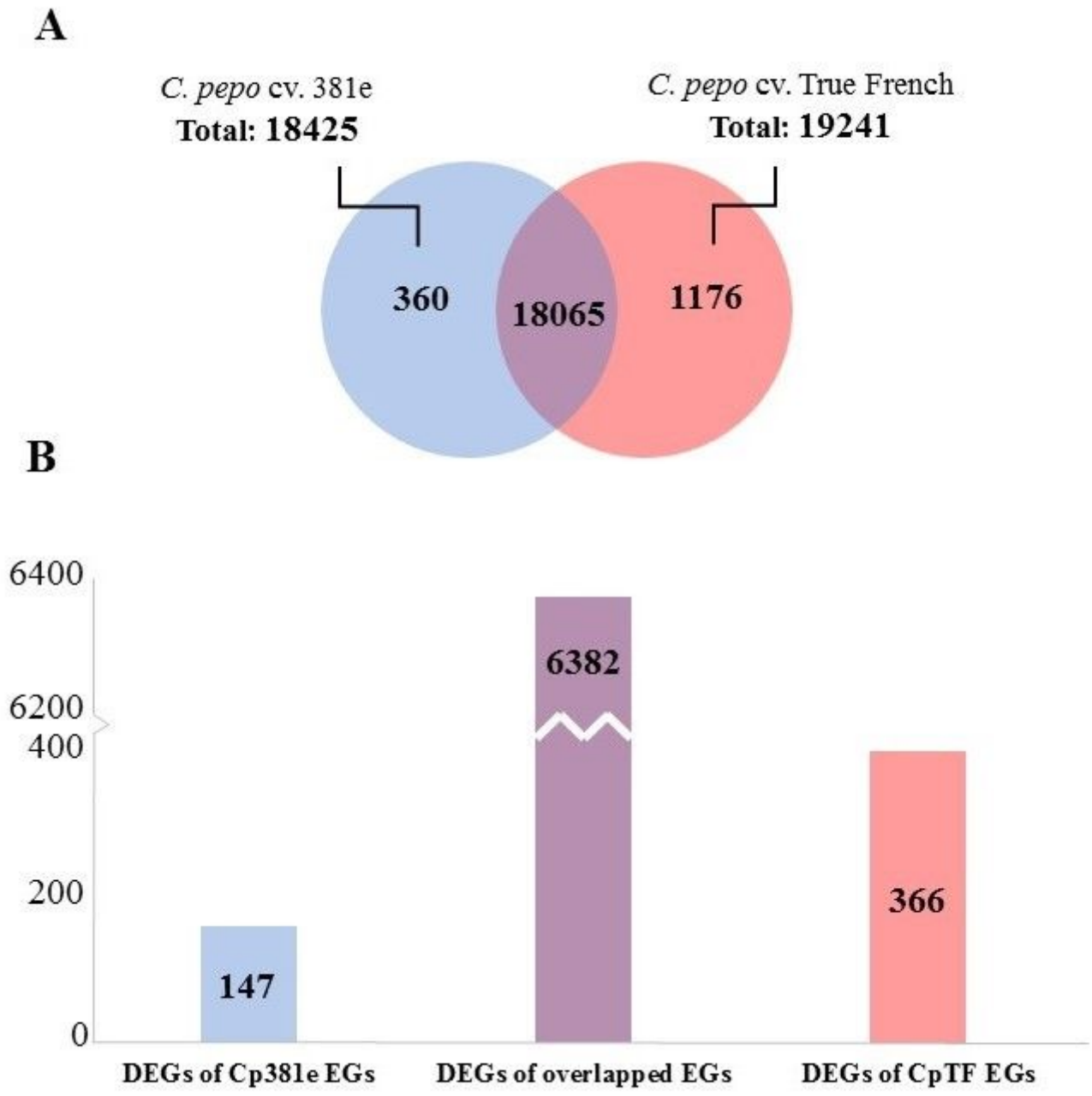


Figure 1

Transcriptional Profile. A) Total Expressed Genes in C. pepo 381e (blue), TF (orange), and overlapped genes (violet). B) Differentially Expressed Genes for each area of Venn diagram. 2.2 Overview of cellular processes activated in the tolerant and susceptible line

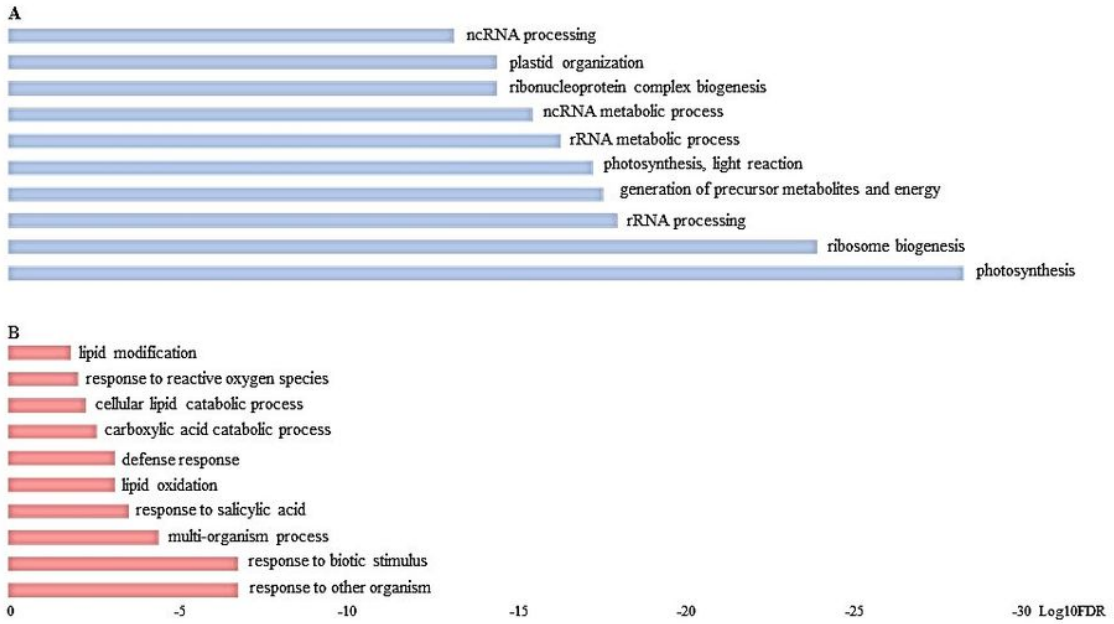


Figure 2

Top ten of non-redundant significantly enriched GO terms in biological process category. A) Upregulated genes in 381e and B) Upregulated genes in True French. The categories size is releted to LOG10 FDR of enriched GOs.

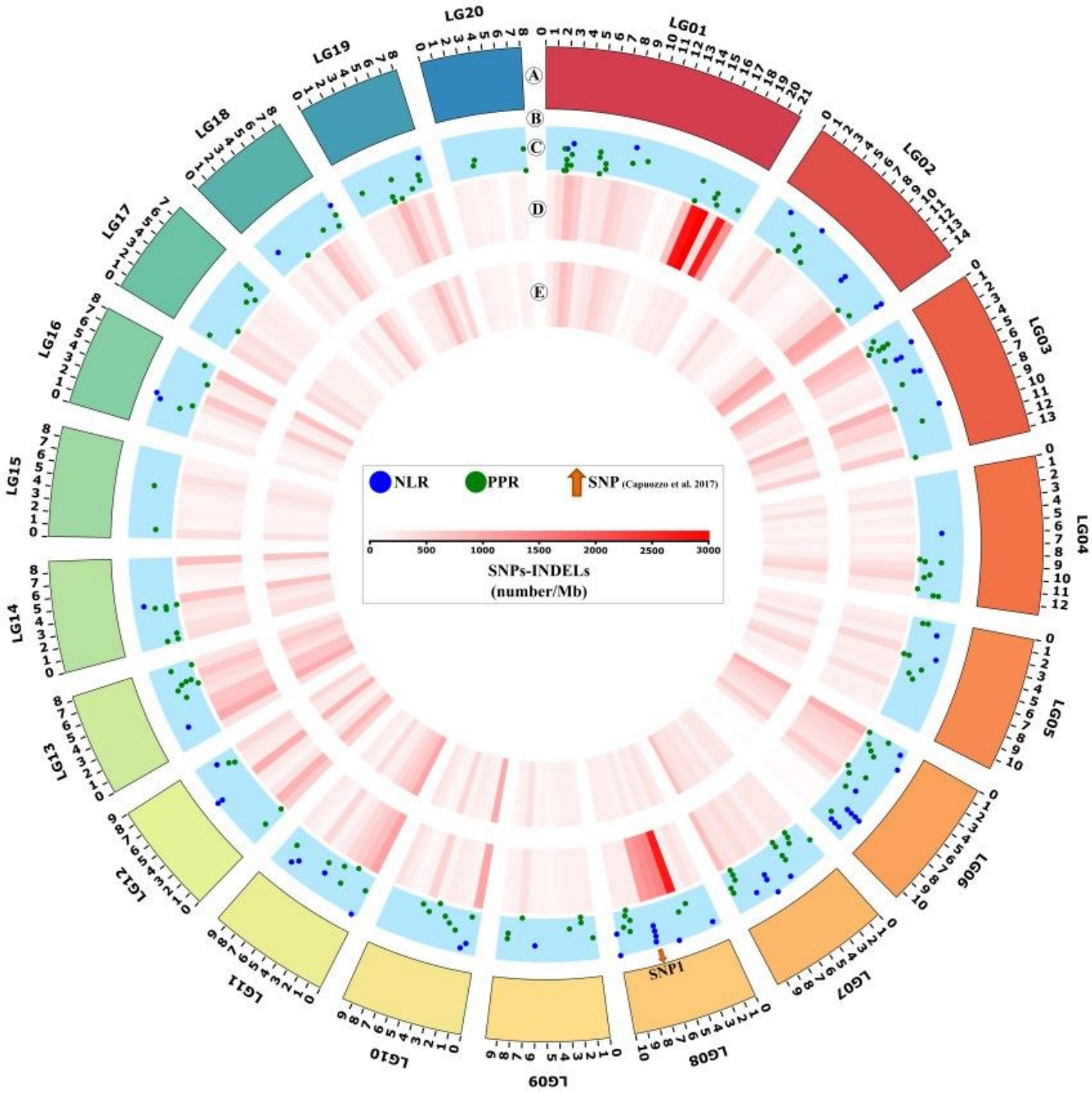


Figure 3

Circos plot integrating the genomic positions of ZYMV-related markers, pathogen recognition genes and small variants (SNP: single nucleotide polymorphisms and InDel: insertion and deletion). Inset legend providing information represented by each data ring. Track A denotes the 20 pseudochromosomes (LG: linkage group) of Cucurbita pepo. The length of each circle segment represents the size of pseudochromosomes expressed in megabases (Mb). B) The marker associated to ZYMV tolerance identified by Capuozzo et al. [11] is indicated by orange arrow. C) Genomic positions of nucleotide

binding like proteins (NLR; blue spots) and receptor like proteins and receptor like kinase (PPR; green spots), annotated by Andolfo et al. [19] , are shown as scatterplots. Heatmaps for the density levels (number/Mb) of small variants identified in (D) C. pepo 381e and (E) C. pepo TF cultivars.

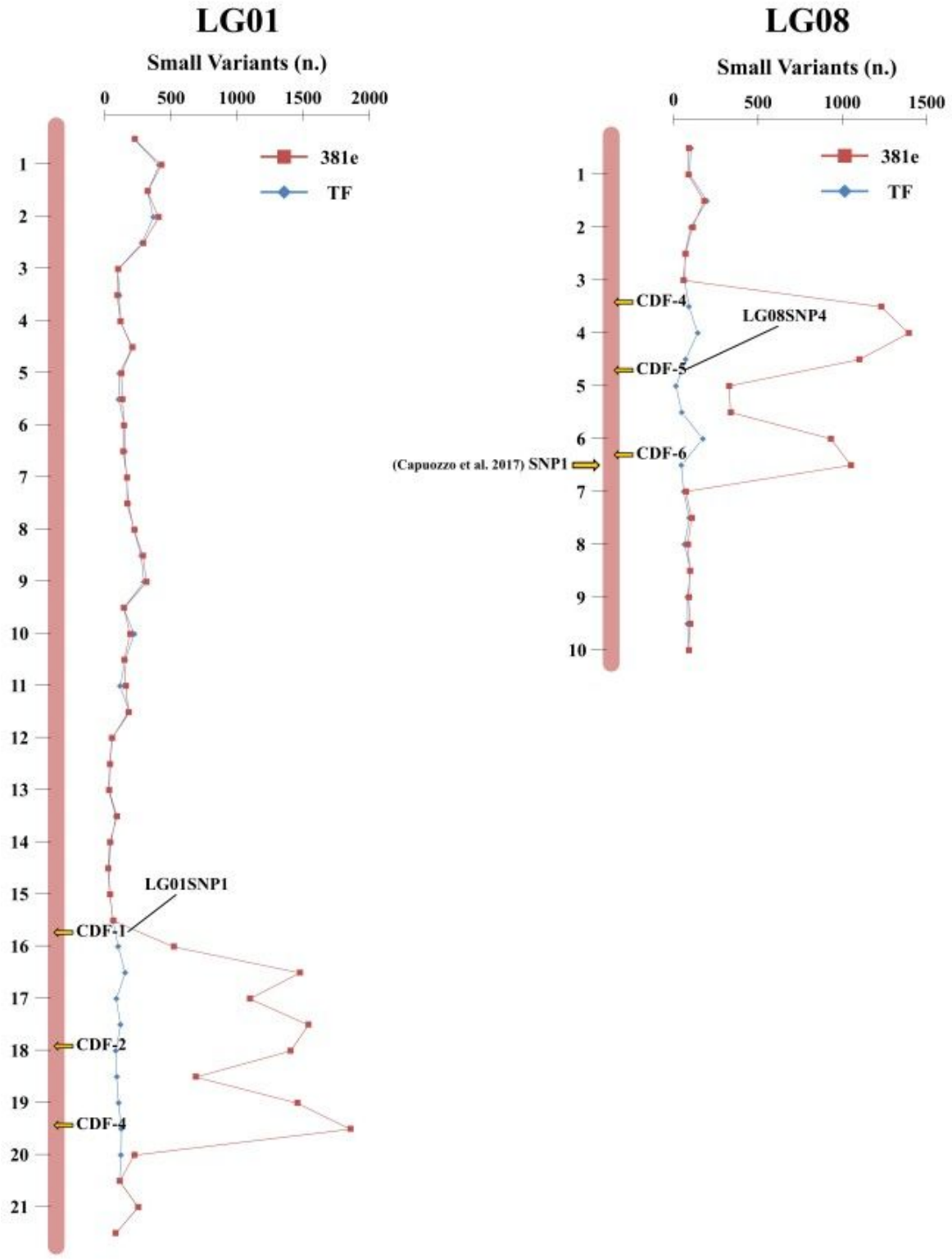


Figure 4

High resolution map integrating genomic and genetic marker information. The distribution of small variants along 1 and 8 pseudochromosomes was reported (381e in red and TF in blue). The genomic

position of six coding DNA fragments (CDFs) and two CAPS markers (LG01SNP1 and LG08SNP4) was indicated.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Suplementarymaterials.xlsx