SCIENTIFIC REPORTS

OPEN

Transcriptome Analysis Reveals Candidate Genes involved in Blister Blight defense in Tea (*Camellia sinensis* (L) Kuntze)

Kuldip Jayaswall[†], Pallavi Mahajan, Gagandeep Singh, Rajni Parmar, Romit Seth, Aparnashree Raina, Mohit Kumar Swarnkar, Anil Kumar Singh[‡], Ravi Shankar & Ram Kumar Sharma

To unravel the molecular mechanism of defense against blister blight (BB) disease caused by an obligate biotrophic fungus, *Exobasidium vexans*, transcriptome of BB interaction with resistance and susceptible tea genotypes was analysed through RNA-seq using Illumina GAIIx at four different stages during ~20-day disease cycle. Approximately 69 million high quality reads were assembled *de novo*, yielding 37,790 unique transcripts with more than 55% being functionally annotated. Differentially expressed, 149 defense related transcripts/genes, namely defense related enzymes, resistance genes, multidrug resistant transporters, transcription factors, retrotransposons, metacaspases and chaperons were observed in RG, suggesting their role in defending against BB. Being present in the major hub, putative master regulators among these candidates were identified from predetermined protein-protein interaction network of *Arabidopsis thaliana*. Further, confirmation of abundant expression of well-known RPM1, RPS2 and RPP13 in quantitative Real Time PCR indicates salicylic acid and jasmonic acid, possibly induce synthesis of antimicrobial compounds, required to overcome the virulence of *E. vexans*. Compendiously, the current study provides a comprehensive gene expression and insights into the molecular mechanism of tea defense against BB to serve as a resource for unravelling the possible regulatory mechanism of immunity against various biotic stresses in tea and other crops.

Tea (*Camellia sinensis* (L) Kuntze) is one of the most popular non-alcoholic beverage crops, globally. It is consumed everyday by millions of people worldwide for its biologically active polyphenols, vitamins, flavanones, catechins and medicinal properties[1–3]. While genomic resources are essential to gain insights of various pathways for dissection of complex traits, it is only very recently that few transcriptomic studies on tea have been performed[4–6], with no such study in response to biotic stresses. Being a perennial plant, tea comes across a wide range of abiotic and biotic stresses during its life span. Among the biotic stresses, fungal pathogens are the most prevalent, causing severe crop loss annually[7,8]. Furthermore, leaf diseases are among the major bottlenecks as commercial tea production is mainly dependent on young succulent leaves[9].

Blister Blight (BB) disease caused by a basidiomycete obligate biotrophic pathogen *Exobasidium vexans* Massee, is amongst the most serious leaf diseases, significantly affecting commercial production in major tea producing countries, including India, Indonesia, Sri Lanka and Japan. This pathogen mainly attacks young succulent, harvestable tender leaves that not only cause more than 40% total yield loss[10], but also affects the tea quality significantly by reducing total phenols and catechin content[11,12]. Management of the disease faces serious challenges of short but multiple disease cycles with several generations within a single crop season, therefore, requires repeated applications of fungicides[13,14]. Although, the application of protectant and eradicant fungicides have shown encouraging results for controlling BB, however, plants face a serious problem of phytotoxicity and fungicide residues. Available bio-controls derived from antagonists namely *Trichoderma harzianum, Gliocladium*

Biotechnology Department, CSIR-Institute of Himalayan Bioresource Technology, Palampur, Himachal Pradesh, India, 176061. †Present address: ICAR-Directorate of Onion and Garlic Research, Rajgurunagar, Pune, Maharashtra, India, 410505. ‡Present address: ICAR-Indian Institute of Agricultural Biotechnology, PDU Campus, IINRG, Namkum, Ranchi-834010, (JH), India. Correspondence and requests for materials should be addressed to R.K.S. (email: rksharma.ihbt@gmail.com)
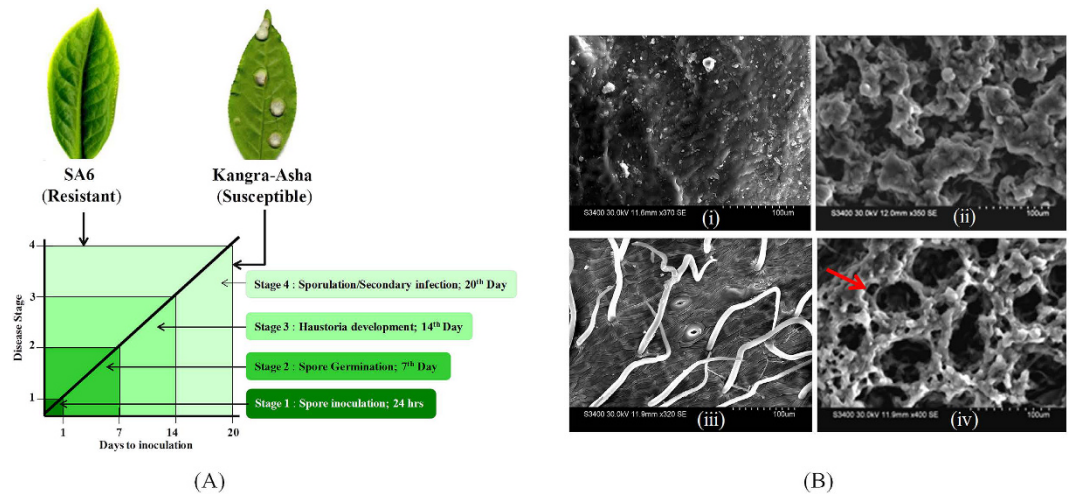
Figure 1. Schematic representation of transcriptome analysis during Blister Blight (BB) transition infected tea leaves. (**A**) Four successive stages of mRNA extraction each from FTL of susceptible (SG; Kangra Asha) and resistant (RG; SA6) tea genotype interaction during 20 days life span of *E. vexan*. Designation of the eight samples was depicted in RG and SG genotypes. (**B**) Scanning electron microscopy of Stage 4 of RG (i & iii) and SG (ii & iv), arrow indicate the hypertrophy/bursting of blister lesion consisting of dense hyphae beneath the lower leaf epidermis in SG at S4.

virens, Serratia marcescens, Pseudomonas fluorescens and *Bacillus subtilis* were not found to be very effective[15,16]. Furthermore, genetic improvement of resistance against blister blight disease through conventional approaches has suffered due to rare availability of resistant tea accessions, highly heterozygous nature, self-incompatibility and long gestation period of tea[17–19].

Well understood immune system of *Arabidopsis* and other crop plants suggests that biotrophic pathogens after entering through stomata proliferate in intercellular spaces and obtain nourishment through specialized haustoria. These pathogens, reduce plant immunity by delivering effectors into plant cells. Unlike mobile defender cells and somatic adaptive immune system reported in animals, plants depend on innate/acquired immunity of each cell and systemic signalling from infection sites *via* indirect activation of many resistances (R) genes through well-developed guard hypothesis[20]. In the plant plasma membrane, Pattern Recognition Receptors (PRR) recognize the Microbial/Pathogen Associated Molecular Pattern (MAMP/PAMP) and also activate other Nucleotide Binding (NB) and Leucine Rich Repeats (LRRs). Disease resistance in plant occurs by induction of several resistance proteins and defense enzymes, which act as major immune regulators through physiological and biochemical alterations[21–23]. Additionally, WRKY, NAM transcription factors, LTR retrotransposons, chaperons and metacaspases are among the major defense regulators[20,24], wherein, WRKY and NAM transcription factors play significant role in large scale transcriptional reprogramming by binding to promoter elements of defense related genes and regulating their expression during plant immunity[25–27]. Transcriptome analyses have revealed that putative sites and sequence motifs, ubiquitously conserved in upstream regions of genes are up-regulated during SAR or R-mediated basal defense[28,29]. Moreover, activation of several retrotransposons during disease transition elicits defense responses and defense gene activation[30,31]. Although, few genes regulating blister blight resistance have been identified[32], but genome wide transcriptome study to understand the global molecular basis of the immune system against BB has not been elucidated in tea, so far.

In this study, the global gene expression pattern was analysed for the first time using high-throughput Illumina sequencing of young leaf tissues during different stages of BB transition in resistant (RG) and susceptible (SG) genotypes. Comparative transcriptome analysis of RG and SG leads to the identification of putative pathways, genes and their interactions, and suggests good candidates involved in BB defense in tea. Comprehensive efforts of the current study jointly with conventional breeding techniques would thus accelerate genetic improvement of tea and other perennial crops.

Results

Transcriptome sequencing. To dissect the molecular mechanism of defense against BB, eight young leaf samples (first two leaves; FTL) of four successive stages from susceptible (SG) and resistant (RG) genotypes was collected during disease progression. Considering the 20-day lifespan of *E. vexans*[8], four successive stages were categorized as exobasidiospore inoculation or landing of exobasidiospores on the upper surface of leaves [24 hrs post inoculation (PI), Stage1 (S1)]; penetration/germination of spores inside the host tissue [7th day PI; Stage2 (S2)], haustoria and mycelial development [14th day PI; Stage3 (S3)] and sporulation of exobasidiospore [20th day PI; Stage 4 (S4)] (Fig. 1A). Progression of BB disease confirmed through SEM analysis, clearly revealed hypertrophy/bursting of blister lesion consisting of dense hyphae beneath the lower leaf epidermis at SG_S4 (Fig. 1B). To cater sample bias, five subsamples were collected at each stage of RG & SG; independently processed for RNA isolation and pooled at equimolar concentration for making RNA-seq libraries. Transcriptome sequencing of
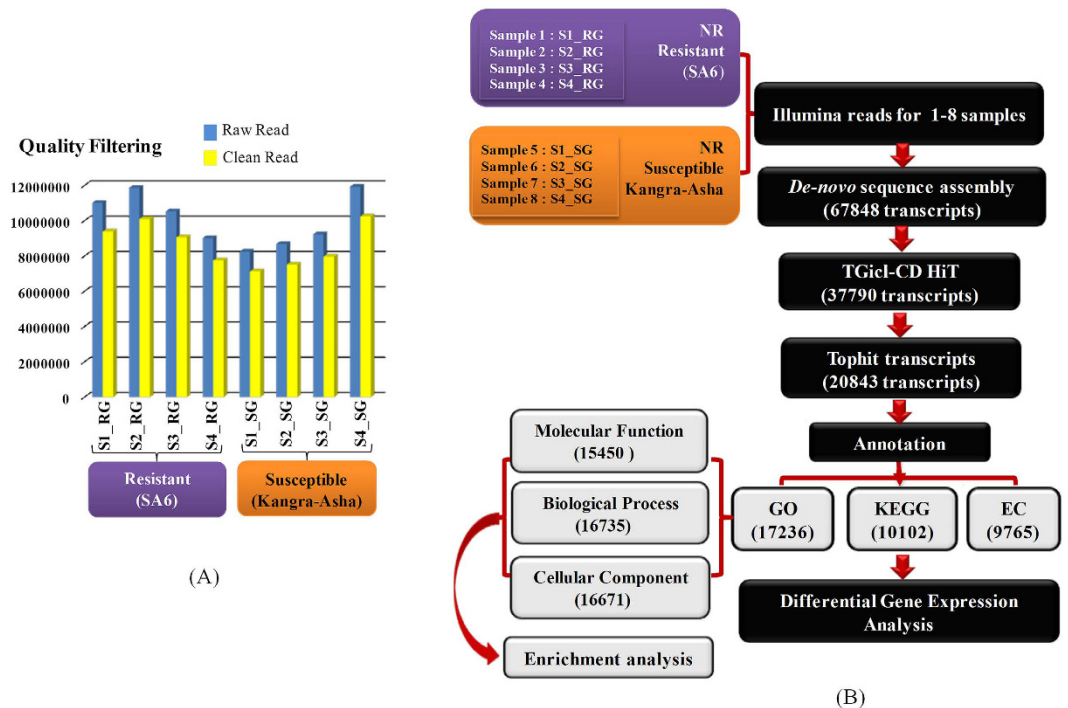
Figure 2. Summary of transcriptome data and analysis approach followed to dissect the defense mechanism against BB in tea. (**A**) Sample wise raw/clean reads obtained in SG and RG. (**B**) Workflow of *de novo* assembly and annotation details of transcriptome data generated in illumina sequencing.

Assembly Statistics	Details
K-mer	21
N 75 length	675
N 50 length	1,731
Minimum length (bp)	325
Maximum length (bp)	11,318
Average length (bp)	1087.58
Total Contigs	37790

Table 1. Overview of the *De novo* assembly.

8 different FTL libraries generated 42,474,318 (RG) and 38,132,457 (SG) paired end (PE) raw reads. Quality filtering and removal of fungal specific reads by mapping to phylogenetically related fungal genome database of *Exobasidium vaccinia* (http://genome.jgi.doe.gov), 36,339,509 and 32,887,574 high quality reads of RG and SG, respectively were achieved (Fig. 2A). To increase the total coverage, frequency and average transcript length, overall 69227083 PE reads obtained from eight libraries were pooled before assembly. Best primary assembly of short reads obtained at a k-mer size of 21 nucleotides[6], which yielded 67848 primary assembled transcripts with an average length of 1087.58 bp and N50 about 1731. In all the assembly steps, the minimum length cutoff for assembling transcripts was followed as described by Gahlan *et al.*[33]. The raw reads derived from Illumina GAIIx of all the analysed samples have been deposited in National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) with accession number SRP067826 under BioProject, PRJNA306068.

Homology search and functional annotation. A total of 37790 unique transcripts was generated using a hierarchical clustering approach involving TGICL-CAP3 and CD-HIT[33] from the primary assembled transcripts, which were further utilized for downstream analysis to dissect a BB defense (Fig. 2B, Table 1). BLASTX[34] search of 37790 transcripts against NCBI non-redundant (NR) protein sequence database, annotated 20843 transcripts with significant hits. Annotation with Gene Ontology (GO)[35], Kyoto Encyclopedia of Genes and Genomes (KEGG)[36,37], Enzyme Commission Codes (EC)[36] and Plant Transcription Factor Database (PlantTFDB)[38] for each query sequence identified 17237, 10102, 9765, 11515 transcripts, respectively (Fig. 2B; Supplementary Table S1).

Global gene expression dynamics of defense against BB. Overall transcriptional activity was estimated using edgeR[39] based on the number of genes up-regulated with fold change ≥2 in RG and SG at different stages during BB progression. Defining the global gene expression dynamics of different sets of genes during disease transition, we found 2690 (S1_RG), 1906 (S2_RG), 4848 (S3_RG) and 7398 (S4_RG) transcripts/
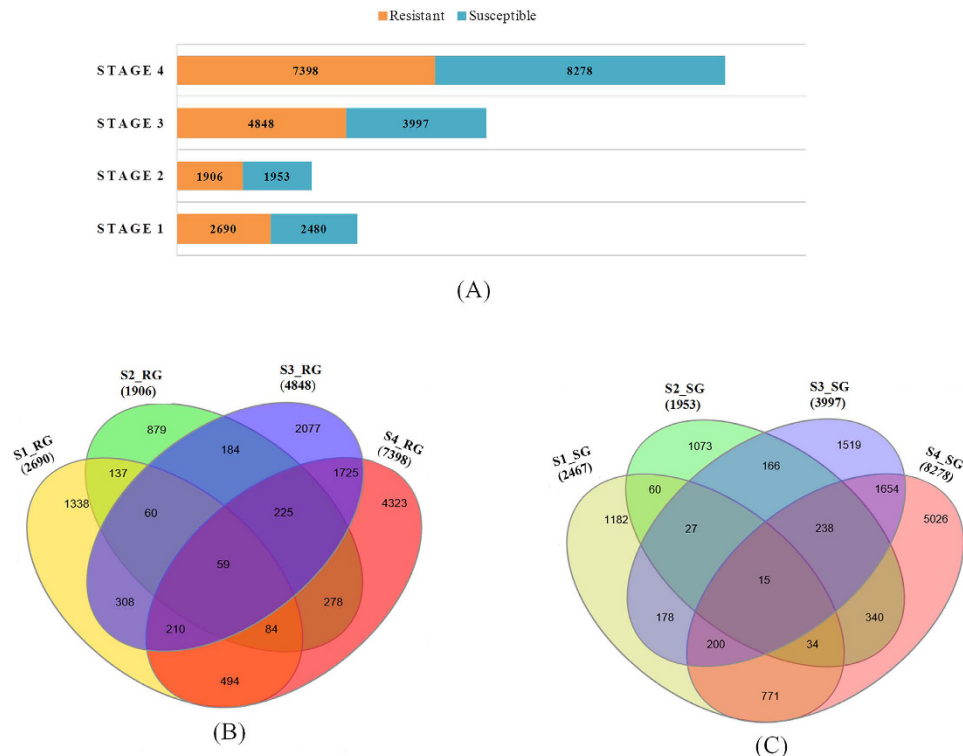
Figure 3. Illustration of overall upregulated transcripts during Blister blight (BB) disease transition.
(**A**) Graphical representation of overall upregulated transcripts in both RG and SG; (**B**) number of upregulated transcripts in Resistant genotype (RG) illustrated in the form of Venn Diagram; (**C**) number of upregulated transcripts in Susceptible genotype (SG) illustrated in the form of the Venn Diagram.

genes up-regulated in RG, while, 2467 (S1_SG), 1953 (S2_SG), 3997 (S3_SG) and 8278 (S4_SG) in SG (Fig. 3A). Correlation of RG and SG during the disease transition revealed that greater abundance of transcripts was recorded at S4. In the case of RG during BB transition, 1338, 879, 2077 and 4323 transcripts were uniquely up-regulated in stages S1, S2, S3 and S4 respectively. Interestingly, 59 transcripts were up-regulated commonly in all the stages as illustrated in the venn diagram (Fig. 3B). On the contrary, in SG, 1182, 1073, 1519 and 5026 transcripts were uniquely upregulated in stages S1, S2, S3 and S4 respectively, wherein 15 common transcripts were found to be upregulated in all the stages (Fig. 3C). Despite the higher number of transcripts in SG (8278 transcripts, S4_SG), most of the defense related genes were expressed in S4_RG.

Gene ontology enrichment and pathway analysis. Gene Ontology (GO)[35] categories of up-regulated transcripts within each stage of disease progression in RG and SG were identified using annot8r[36]. Biological processes and metabolic pathways GO enrichment analysis revealed that GO terms, namely response to stimulus were found to be highly enriched in S1_RG, where in response to chemical stimulus in the form of organic substance were found to be moderately enriched; response to endogenous stimulus and response to defense of stress category remained enriched. On the other hand, cellular aromatic compound and metabolic process, including salicylic acid metabolism and biosynthetic processes were also found to be enriched in S1_RG (Supplementary Figure S1); while in SG, secondary metabolic process and the flavonoid biosynthesis process were enriched (Supplementary Figure S2). At S2_RG, response to stimuli, especially chemical stimulus of biological process was found to be moderately enriched, whereas oxidation reduction (part of metabolic process), cellular respiration (part of cellular metabolic process) and ion transport (cation transport, and transition metal ion transport) were enriched. Additionally, hydrogen peroxide regulation, a part of regulation of oxygen and reactive oxygen species and metabolic process which is positively regulated by a hydrogen peroxide metabolic process were enriched only in RG (Supplementary Figure S3). At S3_RG, major categories including signalling and response to stimulus were found to be enriched, with response to stimulus, including response to organic substance under the category of response to chemical stimulus being highly enriched. Also, response to fungus (response to biotic stimulus) and cellular response to endogenous stimulus were found to be enriched. Furthermore, phenylpropanoid biosynthetic and metabolic process nested in cellular aromatic compounds and metabolic process category was also enriched in S3_RG (Supplementary Figures4), whereas in S3_SG, only chloroplast RNA processing GO accession was enriched (Supplementary Figure S5). Interestingly, at S4_RG, response to stimulus (highly enriched) remained as a major category with maximally enriched biological GO processes. Defense response and cellular response to chemical stimulus (enriched) under the major subcategories namely response to stress and response to chemical stimulus remained moderately enriched, respectively. Jasmonic acid mediated signalling pathway nested in GO sub-category response to endogenous stimulus, including cellular response to jasmonic acid stimulus and
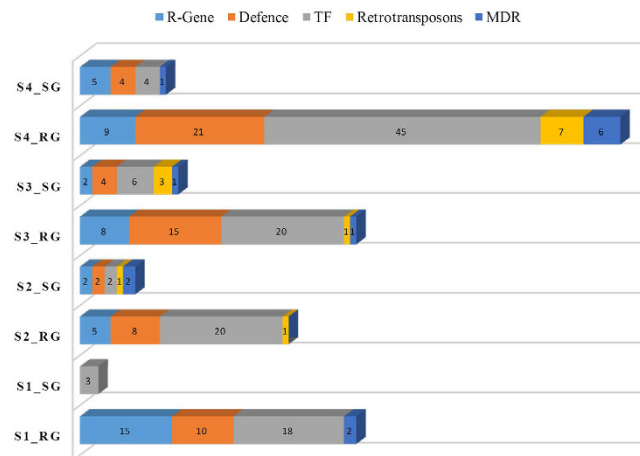
Figure 4. Graphical Representation of immune responsible genes identified in transcriptome data classified into five categories: R-gene, Defense related. TF, Retrotransposons, MDR and Other Defense related genes.

endogenous stimulus were also enriched (Supplementary Figure S6). Nevertheless, during S4_SG histone lysine methylation and DNA replication were enriched in metabolic process; cell cycle, cell wall organization or biogenesis, negatively regulating cell cycle process and cell wall metabolic process were also enriched. Similarly, enrichment of multicellular organismal development, negatively regulating system development and organ development were observed and anatomical structure development under developmental process remained moderately enriched (Supplementary Figure S7).

In molecular function GO enrichment analysis, we found the genes such as Oxidoreductase, Catalytic monoxygenase, Aminopropane-1-carboxylate oxidase, Flavanol synthase, NADH dehydrogenase and Xyloglucosyl transferase and Nuclease and Endonuclease involved in various aspects of defense responses were enriched in RG. Of these, catalytic activity was found to be the most significantly enriched at S4_RG while, Oxidoreductase was significantly enriched irrespective of stages. Flavanol synthase at S1_RG, NADH dehydrogenase and Xyloglucosyl transferase at S2_ SG and Nuclease and Endonuclease were enriched at S3_SG.

Overall GO enrichment analysis revealed salicylic acid biosynthesis and metabolism with a secondary metabolic pathway, possibly activate and trigger the primary immune response in S1_RG, while flavonoid synthesis and other defense related genes trigger immunity in S1_SG. As the disease progresses towards S2 & S3, major secondary metabolite pathways, biotic stress controlling enzymes and free radical scavenging processes provides immunity in RG. In S2_SG, oxidoreductase and cell wall strengthening enzyme play role in defense, but later at S3_SG nuclease activates and probably leads to immune system failure. At S4_RG, enriched expression of jasmonic acid, secondary metabolite, biotic stress controlling enzymes and other strong immune responsible genes provide defense in RG, while, expression of cell division and developmental stage regulating genes clearly indicates cell(s) death, initiating new cell division and development in SG.

Differential gene expression of defense related genes.   We were specifically interested in identifying the gene expression pattern during BB defense, 149 immune responsible genes classified in KEGG[37], GO[35], EC and PlantTFDB[38] annotations, were studied to elucidate key pathway genes involved in BB defense. These include genes encoding defense related enzymes (30), resistance genes (25), multidrug resistant transporters (9), transcription factors (65), retrotransposons (9), and other defense genes (15) comprising metacaspases (2) and chaperons (2) (Fig. 4). To assess the overall similarity between the immune responsible dataset, the Euclidian distance was calculated between each stage of RG and SG based on transcript abundance values, which further confirms the significance of immune responsive genes in RG at S4 (Fig. 5; Supplementary Table S2).

Differential expression of resistance (R) genes.   The Plant genome encodes many R genes that recognize divergent pathogen effectors and triggers the hypersensitive cell death resistance response. In total, 25 R genes exhibiting significant differential expression during disease transition were identified, among which 14 (S1_RG), 5 (S2_RG) and 9 each at S3_RG and S4_RG were up-regulated. Although, many of these R genes were common irrespective to stage during disease transition, four of the R genes, namely, RIN4, RPM1, RPS2 and RPP13 were most represented (Fig. 4A). A Heat map of R gene clustering indicates that RIN4 possibly regulates other RPM1, RPS2 and RPP13 resistant proteins for early immune development as previously reported in *A. thaliana*[21]. Expression data of these R genes indicate transcriptional dynamics during disease transition.

Differential expression of defense related enzymes.   In total, 1337 enzymes, encoded by tea transcripts during different stages of BB transition were annotated using Enzyme Commission (EC) database. While analysing 30 defense related enzymes, 10 (S1_RG), 8 (S2_RG), 15 (S3_RG) and 21 (S4_RG) were found to be highly up-regulated (Fig. 5B). Of these, 21 up-regulated defense related enzymes at S4_RG with fewer up-regulated irrespective of the BB transition stage, which might account for the long lasting immunity in
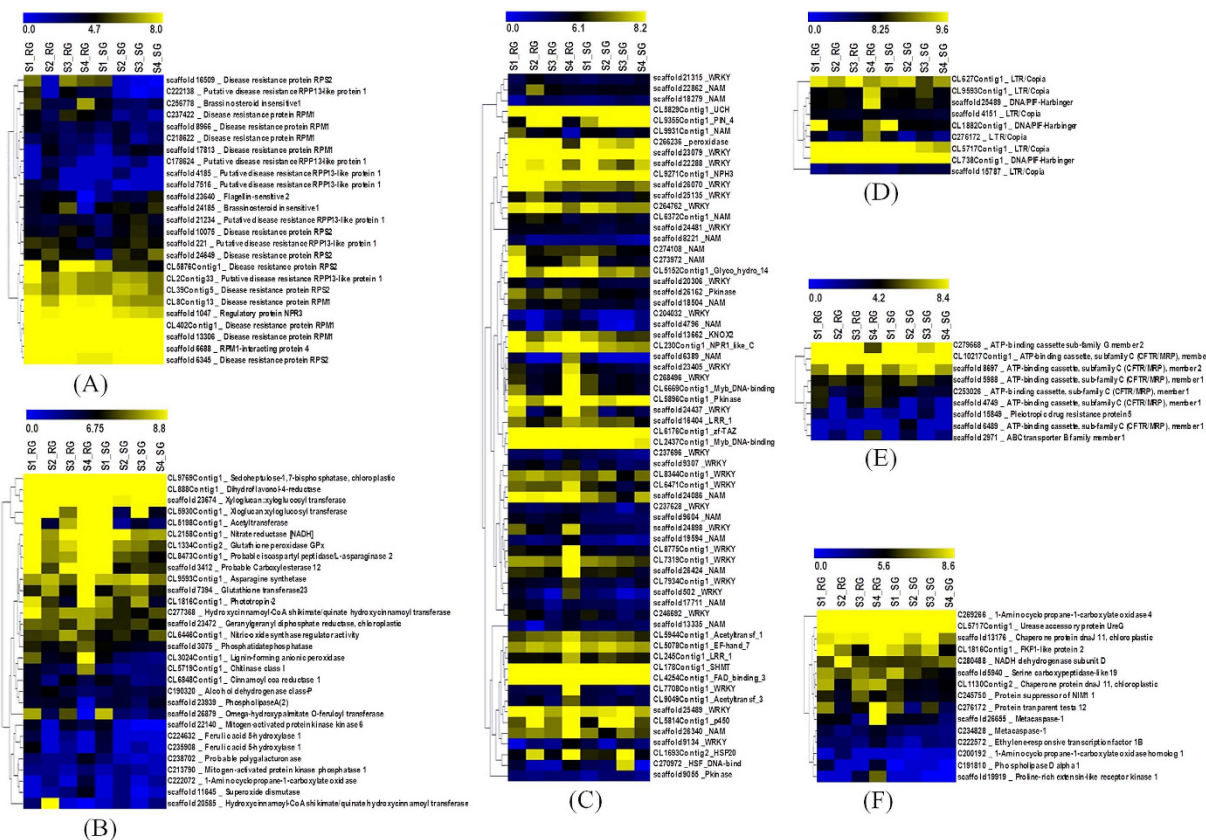
Figure 5. Heat map depicting differential expression profiles of 149 immune related genes at different BB disease transition stages in the RG and SG. (**A**) Resistance genes; (**B**) Defense related enzymes; (**C**) Transcription factors; (**D**) Retrotransposons; (**E**) Multidrug Resistance Defense Transporters. (**F**) Other important defense regulatory genes.

RG[20] (Fig. 5B). Up-regulation of acetyl-transferase, xyloglucosyl-transferase, peroxidase and carboxyl-esterase suggests that S4_RG is among the most active stage of BB transition (Supplementary Table S2). Nevertheless, Sedo heptulose 1–7 bisphosphate, a carbon fixing enzyme expressed in both RG and SG possibly provides photosynthate for immune development. Defense enzymes, namely, Xyloglucan:Xyloglucosyl transferase, Dihydroflavonal-4-reductase, Glutathione peroxidase, Glutathione transferase, Asparagine synthase, Quinate hydroxycinnamoyl transferase, Alcohol dehydrogenase, Nitrate reductase, Nitric oxide synthase regulatory protein up-regulated in RG also provide immunity during the early stages of disease transition in SG as revealed by heat map clustering. However, Acetyltransferase, Carboxylesterase, Chitinase, Lignin forming anionic peroxidase, Omega-Hydroxypalmitate-o-feruloyl transferase, ACC oxidase, Polygalacturonase, Superoxide dismutase, Inositol 1,4,5 triphosphate, Mitogen activated protein kinase kinase 6, Cinnamoyl-CoA reductase showed elevated expression in RG.

Differential expression of defense responsive transcription factor. Transcription factors are key regulatory proteins of biotic and abiotic stress, which mediate the transcriptional regulation[25]. Expression dynamics of transcription factor genes involved in the BB defense identified 66 defense responsive transcription factors exhibiting significant differential expression in RG and SG (Fig. 5C). Among these, WRKY & NAM (NAC domain) families were most represented, while stage wise comparison during BB transition recorded 15, 20, 20 and 44 TFs up regulated in S1_RG, S2_RG, S3_RG and S4_RG, respectively. The highest expression of TFs at S4_RG further confirms the relevance of S4 in BB defense to cope with disease severity due to secondary infection.

Differential expression of retrotransposons. Retrotransposons remain quiescent during normal condition, but become more active during the onset of various stresses[30]. The new insertion within or flanking to coding regions generates mutation that can lead to changes in the dynamics of gene expression and reshape the genome. It has been reported that LTR retrotransposons are activated during pathogenic infection[40]. In total, 9 transcripts of retrotransposons family exhibited significant differential expression during the BB transition in RG, wherein LTR/Copia were found to be highly upregulated (Fig. 5D). Furthermore, stage wise comparison identified 1 (S2_RG), 2 (S3_RG) and 7 (S4_RG) up-regulated transcripts. Current data indicate that during the higher biotic pressure (S4_RG), more number of up-regulated retrotransposons might play an important role in reshaping the genome leading to immunity in RG.

Differential expression of multidrug resistance transporters. Plant genome contains a large number of genes encoding putative multidrug resistance transporters, reported to be involved in the transport of a wide range of compounds including auxins, flavonoids, glutathione conjugates, metal chelators, herbicide and antibiotics[41]. We found 9 MDR transporters exhibiting significant differential expression in disease transition (Fig. 5E). Out of these MDR transcripts, 2 in each S1_RG & S3_RG, and 6 in S4_ RG were found to be up-regulated, which further indicates its importance during BB transition, wherein up-regulated MDRs at stage S4 possibly helps in acquiring long lasting immunity of tea in RG.

Other important defense regulatory genes. Transcripts encoding chaperone, 1-Aminocyclopropane-1-carboxylate oxidase, cytochrome p450, metacaspase and circadian rhythm exhibited significant differential expression during the BB transition with most of them revealing a higher number of up-regulated transcripts (12) in S4_RG (Fig. 5F). Overall differential gene expression of other important defense regulatory genes indicated S4 was the most transcriptionally active stages irrespective of RG and SG.

Protein-Protein Interactions against BB defense. For identification of master regulators among the defense related genes, predetermined protein-protein interaction (PPI) network of Arabidopsis was used to predict the key genes/proteins involved in BB defense in tea. Out of 149 defense responsive genes, 126 unique TAIR-IDs were successfully assigned, among which 88 were mapped and found to be interacting with 746 nodes and 3626 edges. Average number of undirected neighbours in the network for each gene was approximately 9.721. On closer analysis, it was observed that 62 of 88 mapped genes were found in the major group containing 661 nodes with average number of neighbours 10.6, while 26 remained in solitary clusters. Based on the interactome, R genes (4), defense enzyme (7), transcription factor (11), reterotransposon (2), multidrug resistance (9) and other defense genes (2) were present in the major hub, hence can be considered as putative key genes for BB defense (Table 2). Furthermore, PPI network analysis revealed that R genes interacting with 35 other genes, defense related enzymes with 125, transcription factors with 149, reterotransposon with 175, multidrug resistant category with 149 genes and other defense related genes with 41 genes; therefore, can be potential candidates for combining of resistance in promising tea genotypes (Fig. 6).

Experimental validation of differential expression data by qRT-PCR. To validate a differential gene expression pattern obtained through RNA-seq, 13 immune related genes were selected for qRT-PCR analysis (Supplementary Table S3). Among these, expression of transcripts encoding defense related enzymes, multidrug resistant transporters, transcription factors, metacaspases and chaperons obtained in qRT-PCR was found to be largely corresponding with RNA-seq data (Fig. 7). Hence, indicating that RNA-seq approach provided reliable differential gene expression information to understand molecular mechanism of immune response against the BB defense in tea.

Discussion

It is imperative to use NGS based global transcriptome sequencing approach to identify key pathway genes involved in BB defense in tea as follows in earlier studies for dissecting various biotic stresses[42,43]. *E. vexans* Massee is an obligate biotroph thus difficult to study its virulence, and therefore, elucidation of molecular mechanism and identification of key immune gene(s) involved in BB defense in RG could be an appropriate approach to control severe crop loss. Furthermore, identification of diverse R genes will also provide opportunity to study interactions with a wide range of *E. vexans* effectors for ascertaining possible immune signalling during BB defense in tea.

In this study, GO enrichment of the defense related aromatic compound and salicylic acid biosynthesis processes in RG, and secondary metabolite biosynthesis (lignin or phytoalexins like antimicrobial compounds) in S1_SG suggests the activation of the immune system during the initial stages of disease transition irrespective of the RG & SG. The Presence of hydrogen peroxidase activity suggests efficient scavenging of free radicals enabling restricted penetration/germination of spores inside the host tissue at S2_RG. Phenylpropanoid and aromatic compound synthesis perhaps lead to synthesis of antimicrobial compounds that might have provided protection during haustoria and mycelial growth at S3_RG. Concurrently, chloroplast RNA processing, known to reduce photosynthesis, could be responsible for failure of the immune system during S3_SG[44]. Additionally, enrichment of defense response, jasmonic acid stimulus, secondary metabolic process and jasmonic acid mediated signalling during sporulation of exobasidiospore at later stages of BB transition (S4_RG) confirmed that RG developed immunity by jasmonic acid signalling pathway. Although the presence of developmental processes, cell cycle, histone lysine methylation is known to provide small immunity achieved by histone lysine methylation, jasmonic and salicylic acid signalling of immune response, it was totally absent during sporulation of exobasidiospore (S4_SG). Furthermore, flavonol synthesis and secondary metabolite production during S1_SG and presence of xyloglucosyl transferase at S2_SG as evident in the molecular function GO term, might provide a preliminary defense by lignin deposition restricting the entry of pathogen and subsequent spread of disease as reported earlier[45,46], whereas, immune system appeared to have failed in S3_SG due to expression of nucleases and endonucleases activity. In general, molecular function GO analysis revealed a declining trend of immunity during progression of BB from S1_SG to S4_SG. However, up-regulation of monooxygenase and ACC oxidase activity in S4_RG possibly provides defense by removing free radicals and production of ethylene[47,48]. Enrichment of defense related genes possibly suggests stronger immunity due to activation of *E. vexans* secondary infection, immune prime cell. Conclusively, the presence of strong immune system machinery might be responsible for long lasting immunity at S4_RG.

Up-regulation of R genes, defense related enzymes, transcription factors, retrotransposons and chaperones during BB transition in RG as revealed in the current study suggests the putative defense mechanism against BB in tea was comparable with model plant *Arabidopsis*. In general, R genes encode five classes of proteins, among

Defense category	Codes	Transcript ID	Arabidopsis ID	No. of Nodes	Description
R-Genes	1R	scaffold6688	AT3G25070	5	RPM1-interacting protein 4
	2R	C237422	AT3G07040	5	Disease resistance protein RPM1
	3R	scaffold10075	AT4G26090	12	Disease resistance protein RPS2
	4R	scaffold6345	AT4G19050	13	Disease resistance protein RPS2
Enzymes	1E	scaffold11645	AT1G08830	36	Superoxide dismutase
	2E	scaffold22140	AT5G56580	35	Mitogen-activated protein kinase kinase 6
	3E	CL2158Contig1	AT1G37130	12	Nitrate reductase [NADH]
	4E	CL1816Contig1	AT1G68050	18	Phototropin-2
	5E	CL6848Contig1	AT1G68540	8	Cinnamoyl CoA reductase 1
	6E	CL1334Contig2	AT2G31570	5	Glutathione peroxidase GPx
	7E	C190320	AT1G77120	11	Alcohol dehydrogenase class-P
Transcription Factors	1TF	scaffold16404_2_ORF1	AT4G33430	21	LRR_1
	2TF	CL1693Contig2_4_ORF2	AT5G12020	16	HSP20
	3TF	scaffold9307_1_ORF2	AT1G80840	5	WRKY
	4TF	CL8344Contig1_6_ORF4	AT4G23810	11	WRKY
	5TF	CL5829Contig1_6_ORF1	AT3G20630	23	UCH
	6TF	C270972_2_ORF1	AT3G22830	13	HSF_DNA-bind
	7TF	CL178Contig1_3_ORF1	AT4G37930	23	SHMT
	8TF	scaffold13662	AT1G62990	15	KNOX2
	9TF	CL5814Contig1	AT1G74110	5	p450
	10TF	scaffold26162	AT2G48010	9	Pkinase
	11TF	scaffold24086	AT5G13180	8	NAM
Retrotransposons	1Re	CL738Contig1	AT1G62750	19	DNA/PIF-Harbinger
	2Re	C276172	AT3G26590	156	LTR/Copia
Multi Drug resistant	1MDR	C253026	AT1G04120	12	ATP-binding cassette, sub-family C (CFTR/MRP), member 1
	2MDR	CL10217Contig1	AT3G59140	14	ATP-binding cassette, subfamily C (CFTR/MRP), member 2
	3MDR	scaffold8697	AT2G47800	12	ATP-binding cassette, subfamily C (CFTR/MRP), member 2
	4MDR	scaffold4749	AT3G13090	15	ATP-binding cassette, subfamily C (CFTR/MRP), member 1
	5MDR	scaffold2971	AT3G28390	20	ABC transporter B family member 1
	6MDR	scaffold6489	AT3G13080	51	ATP-binding cassette, sub-family C (CFTR/MRP), member 1
	7MDR	scaffold15849	AT1G15520	6	Pleiotropic drug resistance protein 5
	8MDR	scaffold5988	AT3G60160	14	ATP-binding cassette, sub-family C (CFTR/MRP), member 1
	9MDR	C279668	AT2G13610	5	ATP-binding cassette sub-family G member 2
Other defence related genes	1OD	C234828	AT1G02170	6	Metacaspase-1
	2OD	scaffold5940	AT4G12910	35	Serine carboxypeptidase-like 19

Table 2. Summarization of protein-protein interaction (PPI) network of key genes of BB defense identified in transcriptome analysis in tea with *Arabidopsis* proteome.

which, the NB-LRR (Nucleotide Binding Leucine Rich Repeat) composed of a conserved Nucleotide Binding domain and a variable Leucine Rich Repeat, makes up the largest class[49]. Up-regulation of various R genes during the BB transition in RG possibly increased its recognition capacity to various effectors that limit the pathogen growth by cell wall modification, secondary metabolite production and finally, programmed cell death (hypersensitive response) by activation of salicylic acid signalling pathway, probably involved in tea immunity against BB. In general, disease resistance is governed by specific plant and biotrophic pathogen interaction between pathogen avr gene and resistance R locus alleles[50]. Although, little is known on signalling events required to activate NB-LRR mediated ETI (Effective Triggered Immunity) in tea, presence of few important R genes up-regulated in RG possibly binding to unknown effector(s) of *E.vexans* that might be playing a role in conferring immunity. Among the various R genes reported in *Arabidopsis*, RIN4, a plasma membrane associated protein is guarded by NB-LRR protein and manipulated by AvrRpm1 and AvrB effectors; eventual modification of RIN4 activates RPM1 NB-LRR protein. AvrRpt2 activates a cysteine protease inside the host cell which cleaves RIN4, consequently activating the RPS 2 NB-LRR protein[20]. Up-regulation of RPM1, RPS2-like disease resistance genes in RG suggests guard model might possibly serve as one of the important mechanisms in tea defense against BB. However, separate reverse genetics studies will be required to prove guard hypothesis in tea. Up-regulation of RPP13 (another R gene) in RG insinuates its important role in defense against BB as reported in *A thaliana*[51].
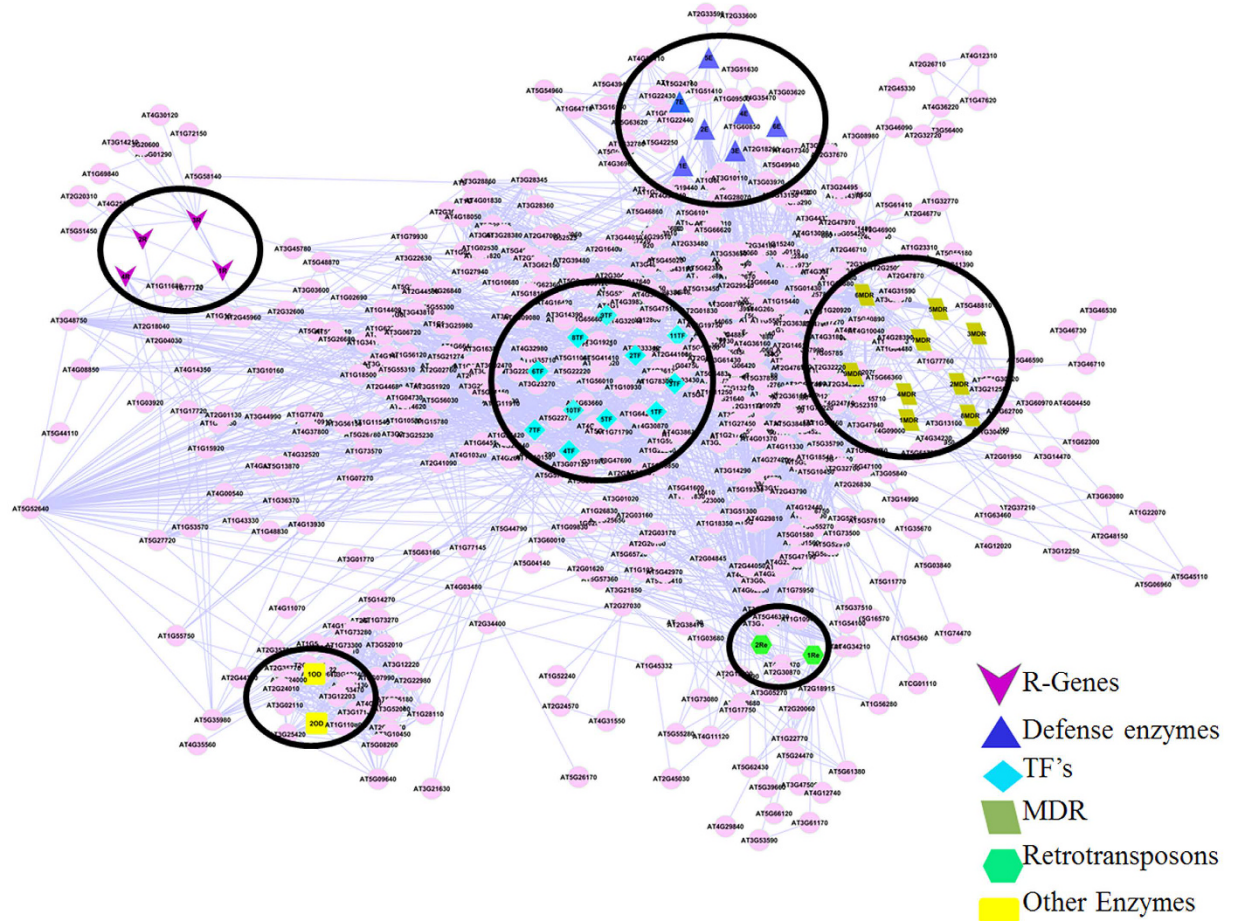
Figure 6. Interactome analysis of putative immune related genes involved in BB defense in tea with a PPI network of *Arabidopsis thaliana*. Categories of Defense related genes (R genes; Defense responsible enzymes; Transcription factors; Retrotransposons; Multidrug Resistance Defense Transporters; Other important defense regulatory genes) depicted with different signs and also indicated in the major hubs PPI network analysis.

Additionally, higher expression of most of the R genes in RG indicates strongly activated immune machinery, maintained till S4_RG where more biotic pressure exists due to secondary infection of BB.

Penetration of pathogen (hyphae) generally, is prevented due to the presence of a tightly sealed cell wall in RG. This results in depletion of nutrient available to pathogen leading to inhibition of hyphal growth and pathogen toxin diffusion. During the polymerization reaction of cell wall, low molecular weight phenolic compounds which are precursor of lignin and free radicals (ROS) affect the pathogenicity. In response to *E. vexans* infection, lignin metabolism, cellulose synthesis, secondary metabolite synthesis, detoxification of toxic compound, ethylene bio-synthesis, free radical scavenger, long lasting immune memory priming responsible enzymes were found to be up regulated at various stages of disease transition and possibly provide immunity as reported previously[20,21]. In reference to *A. thaliana,* MPK6 and MPK1, m-RNA and proteins were reported to be accumulated in inactive form in an immune primed cell for long lasting immunity[52]. During secondary cycle/infection of the pathogen, MPK1 and MPK6 were activated in primed cell, thus enabling cell to take strong and rapid action at a very low level of stimulus compared to un-primed cell. Higher MPKK 6 transcript basal level in primed cell irrespective of the disease transition stages might be the reason for strong immune response in RG during secondary infection.

During pathogenic pressure, various transcription factors have been reported to reprogramme transcriptional dynamics of the plant. Among others, WRKY TFs form a large family of regulatory proteins which are known to play an important role in regulating plant immunity[25]. Up-regulated members of WRKY, NAM and other tran-scription factor families in RG suggests their importance in BB defense[27]. Additionally, R gene activation most likely increases salicylic acid level subsequently leading to NPR1 activation *via* WRKY TF binding upstream of NPR1 gene boxes. The high salicylic acid level might disturb the redox potential of NPR1 protein reducing TGA disulfide bridge, and results a conformational change which allows the TGA family transcription factor to interact with NPR1. TGA-NPR1 complex binds to the promoter of PR gene and possibly provides immunity in RG[25].

In general, the biotic pressure increases the genome instability[53]. Increased pathogen pressure promotes the formation of R gene by locally changing the epigenetic chromatin landscape to stabilize R gene clusters and allow gene rearrangement. Additionally, transposons activity may contribute to the evolution of R genes and could influence their expression[54]. In our study, up-regulation of LTR and DNA/PIF-Harbinger class of
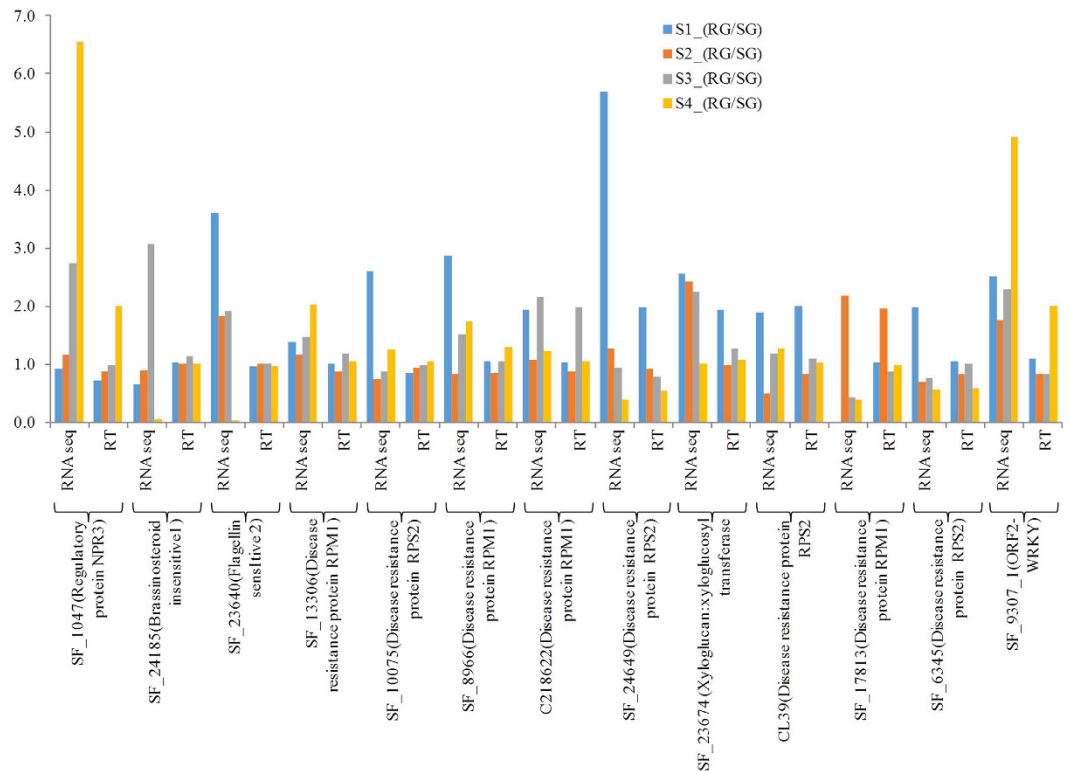
Figure 7. Comparisons between RNA-seq and qRT–PCR expression profiles. Log 2 transformed relative mRNA levels of DEGs by RNA-Seq and analysed by qRT-PCR of defense related genes at different stages during BB progressions in RG and SG.

retrotransposons might lead to genome reshuffling and formation of new polymorphic R genes; hence, could be responsible for providing strong immunity. Though the basal level of expression of most retrotransposons was found to be higher in RG irrespective to BB transition, almost all the retrotransposons were activated as a result of elevated *E.vexans* disease pressure due to the secondary infection at S4_RG.

ATP-binding cassette (ABC) transporters mediate the translocation of a wide range of molecules across the biological membrane. ABC transporters are reported to be responsible for the vacuolar import of chlorophyll catabolites and xenobiotics, playing a key role in cellular detoxification by vacuolar sequestration of endogenous or exogenous toxic compounds[55]. Higher basal level of most MDR transporter's transcripts in RG suggests these perhaps detoxifies the cellular toxin produced by *E. vexans*.

Identification of other important defense related genes such as chaperones and metacaspases might facilitate their conformational changes to induce downstream immune signalling as reported earlier[20]. Chaperone regulates the R protein, eventually leading to programmed cell death (apoptosis) at the site of infection. In plants, tissue or organ apoptosis occurs during normal senescence or response to pathogens. During incompatible interaction, PCD (Programmed Cell Death) occurs and prevents the spreading of pathogen. All these important and defense related regulatory genes were up-regulated at various stages of disease transition with higher basal level of transcripts in RG that may be responsible for regulating tea immunity.

PPI network analysis of 149 defense related proteins with model plant *Arabidopsis* was conducted to reaffirm key proteins modulating BB defense in tea as follows in earlier studies[21]. Being in the PPI network hub and regulating more than 5 other proteins, 35 proteins could thus be designated as key putative proteins involved in BB defense in tea. A Direct association of RIN4 with RPM1 and RPS2 and indirect interaction with NB-ARC domain related to disease resistance indicates that guard hypothesis reported in *Arabidopsis* might be acting in tea defense against BB[20]. Defense related enzymes, namely, superoxide dismutase, mitogen-activated protein kinase kinase-6, nitrate reductase [NADH], phototropin-2, Alcohol dehydrogenase, glutathione peroxidase GPx and cinnamoyl CoA reductase-1, being present in hub possibly regulate other enzymes involved in BB defense. Transcription factors disclose various genes like LRR1, HSP20, WRKY, UCH, HSF_DNA-bind, SHMT, KNOX2, p450, Pkinase and NAM present in network regulating other genes, thus might be playing vital role in BB defense. Among these, NAM (NAC Domain) and WRKY were also found in regulating secondary metabolites: flavonoid, theanine and caffeine biosynthesis in tea[56]. This suggests that these TFs may regulate the BB defense either directly or indirectly via secondary metabolite biosynthesis. In case of retrotransposons it has been found that 2 genes, LTR/Copia and DNA/PIF-Harbinger associated with 175 other Arabidopsis genes in the network and may have an important role in BB defense. Multi Drug resistant genes, namely ATP-binding cassette, sub-family C (CFTR/MRP), 1 pleiotropic drug resistance protein, 1ABC transporter B family member and 1ABC transporter G family member have shown direct interactions with resistance in predicting network of *Arabidopsis,* hence affirming

Figure 8. Diagrammatic representation of putative defense mechanism in tea against BB based on different categories of immune related candidates identified in RNA-seq data.

to be valuable candidates in BB defense in tea. Other genes mapped from identified genes related to defense are metacaspase-1and serine carboxypeptidase-like-19 present in centre regulating various genes in the network. The overall interactome analysis predicted 35 key putative proteins, thus can be categorized as good candidates for regulating BB defense in tea.

Conclusions

This study, for the first time provides comprehensive defense responsive expressions against the BB defense in tea. An overall analysis initiated after inoculating *E. vexans'* spores revealed activation of various R genes along with defense related enzymes to prevent spread of disease after infection, followed by transporters, involved in overcoming the virulence caused by the pathogen in RG. To cope up with disease severity, retrotransposons are activated and by genome reshuffling possibly produce new type of R gene(s) so that plant is able to recognize a new kind of virulence produced by pathogen. Transcription factors are activated in the plant to further enhance the activity of R and other defense associated genes. Overall, R genes, defense related enzymes, retrotransposons, transcription factors and other defense associated molecules provide immunity in RG (Fig. 8). Furthermore, 149 putative key defense related transcripts/genes identified in expression and PPI network analysis can serve as good candidates for future research and also to be used for combining and characterizing defense system against BB in quality tea clones through molecular breeding and genetic engineering technologies. This information is valuable considering the biotrophic nature of fungi and will provide a powerful approach to identifying potential RG from the random gene pool and to unravel immune regulation of various biotic stresses in tea and other crops.

Methods

Plant Material. Tea [*Camellia sinensis* (L.) O. Kuntze] accessions namely SA6 and Kangra-Asha, known source of resistance (RG) and susceptible (SG) response against biotroph fungus *E. vexans* causing blister blight (BB) disease) was used to understand the possible molecular mechanism of defense against BB. SG and RG tea accessions were maintained at the CSIR-Institute of Himalayan Bioresource Technology, Palampur, India (1,300 m altitude; 32°06′ N, 76°33′ E). Vegetatively propagated 3 years old plants of RG and SG were shifted to polyhouse at 22 °C with high humidity (>75%). Plants were sufficiently watered and inoculated with *E. vexans* exobasidiospore suspension at a density of 10^6 cells/ml at maximum[11]. Leaf tissues (first two leaves; FTL) of RG and SG were collected during the BB transition at different time intervals (24 hour, 7th day, 14th day and 20th day after inoculation; AI), snap-frozen in liquid N_2 and stored at −80 °C until used for RNA isolation. To avoid the bias in each stage of BB transition, five subsamples were collected for each sample from RG and SG.

RNA extraction, cDNA library preparation and transcriptome sequencing. Total high quality RNA was extracted using IRIS method[57]. About 100–200 mg of tissue was used to extract the total RNA. The quality and quantity of RNA were assessed with Agilent Bioanalyser using Chip RNA 7500 series II (Agilent Technologies, USA).

cDNA libraries were constructed with approximately 4 μg of RNA (pooled in equimolar concentration from five subsamples)for each sample using the Illumina TruSeq RNA Sample prep Kit v2 LS (Illumina Inc., San Diego, CA) according to the manufacturer's protocol. The libraries were quantified and checked for quality using Bioanalyser chip DNA 1000 series II. Each library was diluted to 8 picomoles in Elution Buffer (Qiagen) and used for Paired End sequencing using the Illumina Genome Analyser GX II platform (San Diego, CA). Fastq format of PE reads were generated utilizing Illumina's CASAVA package GERALD tool removing 3 bp from the

3′ ends to minimize sequencing errors. Poor quality reads and reads with lower K-mer frequency was filtered using a FilteR tool (http://scbb.ihbt.res.in/SCBB_dept/filter.php). Transcriptome assembly was carried out using SOAP*denovo,* wherein, high quality reads were cleaved into smaller fragments (21 K-mers) and put together into final contigs using de Bruijn graphs[58]. Further, clustering of assembled contigs/scaffolds was also done to cut down redundancy by merging them using TGICL- CAP3 and CD-HIT-EST at 90% similarity cut-offs[59], and consensus seqnces were subjected to BLASTX against NR database to get best hits. On the basis of blast tophit dissimilar group clustering were performed using a shell script to identify unigenes, the tophits were searched for a common ID[33] and grouped together.

Functional annotation and classification.

Assembled transcripts were searched against associated GO[35] (http://www.geneontology.org), KEGG[37] (http://www.kegg.jp/kegg/download/) and EC entries. The best hit among the multiple hits for the query transcript sequence was selected on the basis of highest bitscore and E-value (1e–5). Annotation tool, annot8r[36] was used for GO, EC and KEGG based annotations and statistics. Classification and identification of various transcription factors were performed against the Plant transcription factor database (http://plntfdb.bio.uni-potsdam.de)[38] using BLASTX with an E-value threshold of 1e–5. GO term enrichment analysis was performed using agriGO[60].

Read mapping and expression analysis.

Reads Per Kilo base per Million of mapped reads (RPKM), a normalized quantitative procedure for gene expression of RNA Seq data[61] was used to measure the expression level of each assembled transcript. The filtered reads were uniquely mapped back to various assembled transcripts to estimate the transcript abundance using BOWTIE2[62]. Mapped read counts on transcripts were analysed for differential gene expression based on log2 fold count after normalization using EdgeR[39]. Expression data were collected for each of the transcript for all the eight stages of BB transition in resistant (RG) and susceptible (SG) genotypes. The transcripts with fold change ($\geq$2) irrespective of stages during disease transition were selected for enrichment analysis.

Validation and expression analysis by RT-PCR.

Gene expression results were validated by Real time PCR (RT-PCR). RNA was pre-treated with RNase-free DNase I (Invitrogen, USA) to remove contaminating DNA followed by the cDNA synthesis by using 2.0 μg of total RNA with the help of superscript III (Invitrogen, USA) as per manufacturer's instructions. PCR conditions, including primer details for RT-PCR are mentioned in Supplementary Table S3. Cycling conditions were optimized to obtain amplification within the exponential phase. Each reaction was carried out triplicates, and 18 s rRNA and RUBISCO were used as an internal control for normalization.

Interactome analysis of key transcripts associated with BB defense.

To further confirm putative master regulators of defense against BB, putative defense responsible differentially expressed transcripts in RG and SG that include R-genes, defense related enzymes, transcription factors, retrotransposons and multidrug resistant transporters were utilized for intractome analysis. Due to limited genomic information (whole genome/transcriptome) in the tea plant, predetermined protein-protein interactions (PPI) network of *Arabidopsis thaliana* (www.arabidopsis.org/index.jsp) was used for mapping of recognized transcripts[51]. Functional and regulatory defense related transcripts were searched with TAIR to find putative targets using BLASTX (1e–5) and mapped to PPI network. A correlation edge was considered as a conserved correlation edge when the correlation between the defense related gene pair in tea was supported by a significant correlation edge between its respective orthologs in the *Arabidopsis thaliana* PPI network (AtPIN) using Cytoscape software (version 2.8)[63]. First neighbour of mapped IDs was selected for predicting their interaction in network to create the regulatory network. Additionally, based on the interactome statistics, putative master regulators were considered to be interacting with at least five other proteins from each defense category.

References

1. Dou, Q. P. Molecular mechanisms of green tea polyphenols. *Nutrition and cancer* **61,** 827–835 (2009).
2. Alipoor, B. & Rad, A. H. A Review on the Therapeutic Effects of Tea. *Asian Journal of Clinical Nutrition* **4,** 1–15 (2012).
3. Thirugnanasambantham, K., Prabu, G., Palanisamy, S., Chandrabose, S. R. S. & Mandal, A. K. A. Analysis of dormant bud (Banjhi) specific transcriptome of tea (*Camellia sinensis* (L.) O. Kuntze) from cDNA library revealed dormancy-related genes. *Applied Biochemistry and Biotechnology* **169,** 1405–1417 (2013).
4. Shi, C.-Y. *et al.* Deep sequencing of the *Camellia sinensis* transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds. *BMC genomics* **12,** 131 (2011).
5. Wang, X.-C. *et al.* Global transcriptome profiles of *Camellia sinensis* during cold acclimation. *BMC Genomics* **14,** 415 (2013).
6. Paul, A. *et al.* RNA-seq-mediated transcriptome analysis of actively growing and winter dormant shoots identifies non-deciduous habit of evergreen tree tea during winters. *Scientific Reports* **4,** 5932 (2014).
7. Mouli, B. C. & Kumar, P. R. Fungi occurring on tea and Grevillea robusta in tea plantations of Southern India. *Indian Phytopathology* **41,** 503 (1988).
8. Keith, L., Ko, W.-H. & Sato, D. M. Identification guide for diseases of tea (*Camellia sinensis*). *Plant Disease* **33,** 4 (2006).
9. Muraleedharan, N. & Chen, Z. M. Pests and diseases of tea and their management. *Journal of Plantation Crops* **25,** 15–43 (1997).
10. Saravanakumar, D., Vijayakumar, C., Kumar, N. & Samiyappan, R. PGPR-induced defense responses in the tea plant against blister blight disease. *Crop Protection* **26,** 556–565 (2007).
11. Baby, U. I., Balasubramanian, S., Ajay, D. & Premkumar, R. Effect of ergosterol biosynthesis inhibitors on blister blight disease, the tea plant and quality of made tea. *Crop Protection* **23,** 795–800 (2004).
12. Gulati, A., Gulati, A., Ravindranath, S. D. & Chakrabarty, D. N. Economic yield losses caused by Exobasidium vexans in tea plantations. *Indian Phytopathol* **46,** 155–159 (1993).
13. Gadd, C. H. & Loos, C. A. The basidiospores of Exobasidium vexans. *Transactions of the British Mycological Society* **31,** 229–233 (1948).
14. Agnihothrudu, V. & Moulli, B. C. Blister blight of tea, its control and future lines of research. in *Proceedings of International Symposium on Tea Science* **26,** 655–659 (1991).

15. Premkumar, R. Annual Report of UPASI Tea Research Foundation, pp. 38–39 (2003).
16. Balasubramanian, S., Parathiraj, S. & Haridas, P. Effect of vermicompost based Trichoderam (Vermiderma) on the recovery of pruned bushes and on the control of certain disease in tea (*Camellia sinensis* (L) O. Kuntze.). *Journal of Plantation Crops* **34,** 512 (2006).
17. Sharma, R., Bhardwaj, P., Negi, R., Mohapatra, T. & Ahuja, P. Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). *BMC Plant Biology* **9,** 1 (2009).
18. Sharma, R. K. *et al.* AFLP-based genetic diversity assessment of commercially important tea germplasm in India. *Biochemical Genetics* **48,** 549–564 (2010).
19. Kamunya, S. M. *et al.* Genomic mapping and testing for quantitative trait loci in tea (Camellia sinensis (L.) O. Kuntze). *Tree Genetics & Genomes* **6,** 915–929 (2010).
20. Spoel, S. H. & Dong, X. How do plants achieve immunity? Defence without specialized immune cells. *Nature Reviews Immunology* **12,** 89–100 (2012).
21. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444,** 323–329 (2006).
22. Liu, Y.-B., Lu, S.-M., Zhang, J.-F., Liu, S. & Lu, Y.-T. A xyloglucan endotransglucosylase/hydrolase involves in growth of primary root and alters the deposition of cellulose in *Arabidopsis*. *Planta* **226,** 1547–1560 (2007).
23. Kasprzewska, A. Plant chitinases-regulation and function. *Cellular and Molecular Biology Letters* **8,** 809–824 (2003).
24. Nuruzzaman, M., Sharoni, A. M. & Kikuchi, S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Frontiers in Microbiology* **4,** 248, doi: 10.3389/fmicb.2013.00248 (2013).
25. Pandey, S. P. & Somssich, I. E. The role of WRKY transcription factors in plant immunity. *Plant Physiology* **150,** 1648–1655 (2009).
26. Eulgem, T. Regulation of the *Arabidopsis* defense transcriptome. *Trends in Plant Science* **10,** 71–78 (2005).
27. Jensen, M. K. *et al.* Transcriptional regulation by an NAC (NAM–ATAF1, 2–CUC2) transcription factor attenuates ABA signalling for efficient basal defence towards *Blumeria graminis* f. sp. *hordei* in *Arabidopsis*. *The Plant Journal* **56,** 867–880 (2008).
28. Maleck, K. *et al.* The transcriptome of Arabidopsis thaliana during systemic acquired resistance. *Nature Genetics* **26,** 403–410 (2000).
29. Zipfel, C. *et al.* Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* **428,** 764–767 (2004).
30. Grandbastien, M.-A. *et al.* The expression of the tobacco Tnt1 retrotransposon is linked to plant defense responses. *Genetica* **100,** 241–252 (1997).
31. Grandbastien, M.-A. Activation of plant retrotransposons under stress conditions. *Trends in Plant Science* **3,** 181–187 (1998).
32. Bhorali, P. *et al.* Molecular analysis and expression profiling of blister blight defenserelated genes in tea. *Indian Journal of Genetics and Plant Breeding* (*The*) **72,** 226–233 (2012).
33. Gahlan, P. *et al.* De novo sequencing and characterization of *Picrorhiza kurrooa* transcriptome at two temperatures showed major transcriptome adjustments. *BMC Genomics* **13,** 126 (2012).
34. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **215,** 403–410 (1990).
35. Consortium, G. O. Gene Ontology database http://www.geneontology.org. *Accessed March* **22** (2006).
36. Schmid, R. & Blaxter, M. L. annot8r: GO, EC and KEGG annotation of EST datasets. *BMC Bioinformatics* **9,** 180 (2008).
37. Kanehisa, M. & Susumu, G. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **28,** 27–30 (2000).
38. Riaño-Pachón, D. M., Ruzicic, S., Dreyer, I. & Mueller-Roeber, B. PlnTFDB: an integrative plant transcription factor database. *BMC Bioinformatics* **8,** 42 (2007).
39. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).
40. Wessler, S. R., Bureau, T. E. & White, S. E. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Current Opinion in Genetics & Development* **5,** 814–821 (1995).
41. Conte, S. S. & Lloyd, A. M. Exploring multiple drug and herbicide resistance in plants—Spotlight on transporter proteins. *Plant science* **180,** 196–203 (2011).
42. Unamba, C. I. N., Nag, A. & Sharma, R. K. Next Generation Sequencing technologies: The doorway to the unexplored genomics of non-model plants. *Frontiers in Plant Science* **6,** 1074, doi: 10.3389/fpls.2015.01074 (2015).
43. Thakur, K. *et al.* De novo transcriptome sequencing and analysis for *Venturia inaequalis*, the devastating apple scab pathogen. *PloS one* **8,** e53937 (2013).
44. Kishine, M. *et al.* Ribosomal RNA processing and an RNase R family member in chloroplasts of *Arabidopsis*. *Plant Molecular Biology* **55,** 595–606 (2004).
45. Bennett, R. N. & Wallsgrove, R. M. Secondary metabolites in plant defence mechanisms. *New Phytologist* **127,** 617–633 (1994).
46. Liu, Y.-B., Lu, S.-M., Zhang, J.-F., Liu, S. & Lu, Y.-T. A xyloglucan endotransglucosylase/hydrolase involves in growth of primary root and alters the deposition of cellulose in *Arabidopsis*. *Planta* **226,** 1547–1560 (2007).
47. Mishina, T. E. & Zeier, J. The Arabidopsis flavin-dependent monooxygenase FMO1 is an essential component of biologically induced systemic acquired resistance. *Plant Physiology* **141,** 1666–1675 (2006).
48. Wang, K. L.-C., Li, H. & Ecker, J. R. Ethylene biosynthesis and signaling networks. *The plant cell* **14,** S131–S151 (2002).
49. Meyers, B. C., Kozik, A., Griego, A., Kuang, H. & Michelmore, R. W. Genome-wide analysis of NBS-LRR–encoding genes in Arabidopsis. *The Plant Cell* **15,** 809–834 (2003).
50. Dodds, P. & Thrall, P. Recognition events and host–pathogen co-evolution in gene-for-gene resistance to flax rust. *Functional Plant Biology* **36,** 395–408 (2009).
51. Bittner-Eddy, P. D., Crute, I. R., Holub, E. B. & Beynon, J. L. RPP13 is a simple locus in Arabidopsis thaliana for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *The Plant Journal* **21,** 177–188 (2000).
52. Beckers, G. J. M. *et al.* Mitogen-activated protein kinases 3 and 6 are required for full priming of stress responses in *Arabidopsis thaliana*. *The Plant Cell* **21,** 944–953 (2009).
53. Boyko, A. *et al.* Transgenerational changes in the genome stability and methylation in pathogen-infected plants (virus-induced plant genome instability). *Nucleic Acids Research* **35,** 1714–1725 (2007).
54. Whisson, S. C., Vetukuri, R. R., Avrova, A. O. & Dixelius, C. Can silencing of transposons contribute to variation in effector gene expression in *Phytophthora infestans*? *Mobile Genetic Elements* **2,** 110–114 (2012).
55. Stukkens, Y. *et al.* NpPDR1, a pleiotropic drug resistance-type ATP-binding cassette transporter from Nicotiana plumbaginifolia, plays a major role in plant pathogen defense. *Plant Physiology* **139,** 341–352 (2005).
56. Li, C.-F. *et al.* Global Transcriptome and Gene Regulation Network for Secondary Metabolite Biosynthesis of Tea Plant (*Camellia sinensis*). *BMC Genomics* **16,** 1 (2015).
57. Ghawana, S. *et al.* An RNA isolation system for plant tissues rich in secondary metabolites. *BMC Research Notes* **4,** 85 (2011).
58. Li, R. *et al.* De novo assembly of human genomes with massively parallel short read sequencing. *Genome Research* **20,** 265–272 (2010).
59. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22,** 1658–1659 (2006).
60. Du, Z., Zhou, X., Ling, Y., Zhang, Z. & Su, Z. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research* gkq310, doi: 10.1093/nar/gkq310 (2010).
61. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5,** 621–628 (2008).

62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9,** 357–359 (2012).
63. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L.& Ideker, T. Cytoscape; 2.8: new features for data integration and network visualization. *Bioinformatics*. **27,** 431–432 (2011).

Acknowledgements

Author Contributions

Conceived and designed the experiments: R.K.S. Performed the experiments: K.J., R.P., G.S., R.Se., M.K.S., A.K.S. and R.K.S. Analysed the data: K.J., P.M. and R.K.S. Contributed reagents/materials/Bioinformatics analysis tools: K.J., R.S., P.M., G.S., A.S.R. and R.K.S. Wrote the paper: K.J. and R.K.S. Editing and approval of final version of manuscript: R.K.S.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Jayaswall, K. *et al.* Transcriptome Analysis Reveals Candidate Genes involved in Blister Blight defense in Tea (*Camellia sinensis* (L) Kuntze). *Sci. Rep.* **6**, 30412; doi: 10.1038/srep30412 (2016).