Receptor-Like Kinase (RLK) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee

Dênia Pires de Almeida, Isabel Samila Lima Castro, Tiago Antônio de Oliveira Mendes, Danúbia Rodrigues Alves, Geleta Dugassa Barka, Pedro Ricardo Rossi Barreiros, Laércio Zambolim, Ney Sussumu Sakiyama, Eveline Teixeira Caixeta

**ABSTRACT:** The biotrophic fungus *Hemileia vastatrix* causes coffee leaf rust (CLR), one of the most devastating diseases in *Coffea arabica*. Coffee, like other plants, has developed effective mechanisms to recognize and respond to infections caused by pathogens. Plant resistance gene analogs (RGAs) have been identified in certain plants as candidates for resistance (R) genes or membrane receptors that activate the R genes. The RGAs identified in different plants possess conserved domains that play specific roles in the fight against pathogens. Despite the importance of RGAs, in coffee plants these genes and other molecular mechanisms of disease resistance are still unknown. This study aimed to sequence and characterize candidate genes from coffee plants with the potential for involvement in resistance to *H. vastatrix*. Sequencing was performed based on a library of bacterial artificial chromosomes (BAC) of the coffee clone 'Híbrido de Timor' (HdT) CIFC 832/2 and screened using a functional marker. Two RGAs, HdT_LRR_RLK1 and HdT_LRR_RLK2, containing the motif of leucine-rich repeat-like kinase (LRR-RLK) were identified. Based on the presence or absence of the HdT_LRR_RLK2 RGA in a number of differential coffee clones containing different combinations of the rust resistance gene, these RGAs did not correspond to any resistance gene already characterized (S1-9). These genes were also analyzed using qPCR and demonstrated a major expression peak at 24 h after inoculation in both the compatible and incompatible interactions between coffee and *H. vastatrix*. These results are valuable information for breeding programs aimed at developing CLR-resistant cultivars, in addition to enabling a better understanding of the interactions between coffee and *H. vastatrix*.

**Keywords:** *Coffea arabica*, coffee leaf rust, resistance gene analogs, molecular markers, plant breeding

**Introduction**

For many centuries, rust diseases have been a major threat to many crops, causing severe damage to farming activities. In coffee (*Coffea arabica*), leaf rust, caused by the biotrophic fungus *Hemileia vastatrix*, is the most devastating disease worldwide (Zambolim, 2016). Resistant cultivars have been developed and successfully sustained in different coffee-producing regions of the world (Pestana et al., 2015; Alkimim et al., 2017; Valença et al., 2017). However, owing to the high adaptive potential of the pathogen, the emergence of new physiological races and a corresponding ‘breakdown’ of resistance has been observed in many coffee cultivars (Várzea and Marques, 2005; Capucho et al., 2012; Cresssey, 2013).

The pathogen, in general, employs several strategies to infect its host. One strategy is the secretion of effector proteins capable of suppressing plant defense responses and allowing the colonization of host tissues. In response to the infection, plants have evolved an innate immune system consisting of two lines of defense which limit the proliferation of pathogens in their tissues. Pattern-triggered immunity (PTI) is triggered by pathogen-associated molecular patterns (PAMP) and the second line of defense, effector-triggered immunity (ETI) fit the gene-for-gene theory proposed by Flor (1971), and Dodds and Rathjen, 2010. This theory explains the interaction between coffee and *H. vastatrix*, whereby the resistance of coffee plants is conditioned by at least nine dominant genes with main effects (S₁-S₉). The S₁, S₂, S₄, and S₅ genes are found in *C. arabica*, whereas S₆, S₇, S₈, and S₉ are present in *C. canephora* and S₃ is found in *C. liberica* (Noronha-Wagner and Bettencourt, 1967; Bettencourt and Noronha-Wagner, 1971; Bettencourt et al., 1980; Bettencourt and Rodrigues, 1988).

In our study, a novel and important candidate gene corresponding to a leucine-rich repeat (LRR) receptor-like serine/threonine-protein kinase was identified, and cloning was performed by screening a library of bacterial artificial chromosome (BAC) clones corresponding to ‘Híbrido de Timor’ (HdT) CIFC 832/2 (Cação et al., 2013). In addition, we developed functional markers linked to this receptor-like kinase (RLK) gene that can be used in marker-assisted selection in coffee breeding programs, allowing for early selection of individual and pyramidal genes for more durable resistance.
Materials and Methods

Screening of BACs

A library of 56,832 BAC clones derived from one of the main resistance sources, ‘HdT’ CIFC 832/2 (Cação et al., 2013), was used for screening the molecular marker HT24F133. This marker corresponds to the LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP interacting kinase 1-like, which was found to be upregulated in resistant coffee lines (Barka et al., 2017).

The BAC library was replicated in 384-well titration plates using a plate replicator sterilized in a laminar airflow hood. Old cultures were copied onto a new 384-well titration plate with 70 μL of fresh LB medium (containing 12.5 μg mL−1 chloramphenicol) in each well. Multiplication of the cultures was done by incubating them at 37 °C for 18 h on a shaker at 180 rpm. The plasmid DNA of the selected BAC library was screened using PCR with the molecular marker HT24F133. This marker corresponds to the LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP interacting kinase 1-like, which was found to be upregulated in resistant coffee lines (Barka et al., 2017).

The BAC library was replicated in 384-well titration plates using a plate replicator sterilized in a laminar airflow hood. Old cultures were copied onto a new 384-well titration plate with 70 μL of fresh LB medium (containing 12.5 μg mL−1 chloramphenicol) in each well. Multiplication of the cultures was done by incubating them at 37 °C for 18 h on a shaker at 180 rpm. The plasmid DNA of the selected BAC library was screened using PCR with the molecular marker HT24F133. Clones were identified by grouping and subsequent group decomposition of the 384 clones until a single clone was identified as proposed by Cação et al. (2013). The plasmid DNA of the selected BAC clone 104-O-23 was extracted using a Wizard® Plus SV Miniprep DNA purification kit [Promega, Madison, USA], following the manufacturer’s recommendations.

New generation sequencing of the BAC clones and contig assembly

The DNA fragment isolated from the BAC library was sequenced in an Illumina HiSeq2000/2500 100PE platform (paired-end run with 101 bases per read), and analyzed using multiplexed shotgun sequencing. DNA samples were prepared for sequencing using Illumina’s Nextera® XT DNA. Sequences were processed and contig-assembled using the SPAdes algorithm [version 2.5.1] [Bankevich et al., 2012].

The quality of each base call was determined by estimating a quality score similar to the phred score based on the image output without considering the reference sequence. More precisely, Bustard [version 1.8.28], estimates the $P$ probability of a base call of being wrong and reports the corresponding quality score $Q = -10 \log_{10} [P(1-P)]$. A quality assessment of scaffolding assembly was performed using a QUAST Genome Assembly Assessment Tool [version 5.0.2] [Gurevich et al., 2013].

Prediction and annotation of genes

Gene prediction was formulated using the generalized hidden Markov model in AUGUSTUS [version 3.22], which provides a general feature format (GFF3) file containing information such as the position of genes, exons, introns, and transcripts for each gene [Stanke et al., 2004]. Contigs longer than 500 bp were selected for gene prediction in AUGUSTUS [Stanke et al., 2004]. Predicted open reading frames (ORFs) were annotated based on the detection of conserved domains in PFAM [version 32.0] protein families [Finn et al., 2014]. The transmembrane domains of proteins were analyzed using the TMHMM Server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM/). To check for coding sequences, the predicted ORFs were subjected to a BLASTn [Basic Local Alignment Search Tool] [Johnson et al., 2008] search in the C. canephora (http://coffee-genome.org/blast) and C. arabica (https://blast.ncbi.nlm.nih.gov/blast) databases.

Gene expression analysis

Sequences of the identified genes were used to design primers using GenScript Real-time PCR (TaqMan) Primer Design (https://www.genscript.com/tools/real-time-pcr-tagman-primer-design-tool). A pair of primers, named P1, (F: TGCATCGGAGTGGGAGGATTT, R: TTTGGCCTCCTCAAGGCACA, 113 bp, and $T_m$: 60 °C) was used for expression analysis of the candidate gene HdT_LRR_RLK1 and P2 (F: GCTCACAGGTCCGATTCCTCTG, R: TTTGGGAATAGGCCCGGAAAGA, 94 bp, and $T_m$: 60 °C) for the candidate gene HdT_LRR_RLK2.

‘Caturra Vermelho’ CIFC 19/1 was used as the susceptible coffee line and ‘HdT’ CIFC 832/1 as the resistant line. The experiment was conducted in a growth chamber using a completely randomized design, with three biological replicates. One-year-old greenhouse-grown young plants of ‘Caturra’ and ‘HdT’ were challenged with the pathogen race XXXIII of H. vastatrix as proposed by Capucho et al. [2009]. The samples were collected at 0, 12, 24, and 96 h after inoculation (h.a.i).

For RNA extraction, the inoculated leaves were collected and macerated in liquid nitrogen. Total RNA was extracted using 100 mg of the macerated tissue with RNeasy Plant Mini Kit [Qiagen], following the manufacturer’s recommendations. The RNA was quantified using Qubit RNA BR and a NanoDrop spectrophotometer. RNA integrity was evaluated using agarose gel electrophoresis (1.5 %), followed by staining with ethidium bromide. The samples were stored in an ultra-freezer at −80 °C until use.

The cDNA was synthesized with 3 μg of total RNA, which was pretreated with 1 μL of DNase for 15 min (50 U μL−1, amplification grade DNase I) to remove possible genomic DNA contamination. The first cDNA strand was synthesized using the ImProm-IT™ Reverse Transcription system protocol RT-PCR kit, according to the manufacturer’s guidelines and was subsequently stored at −20 °C until use.

Real-time quantitative PCR was performed in a 7500 Real-Time PCR System. Reactions were induced in a final volume of 10 μL with 50 ng μL−1 of cDNA and 2 μM of forward and reverse primers in 1 × GoTaq qPCR Master Mix. The reaction conditions were: 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. The melting curve stage was set to default conditions.

The expression level of the genes was quantified by relating the threshold cycle value to a standard
curve generated by three biological replicates and three technical replicates. For data normalization, we used two selected constitutive genes (UBQ10 and GAPDH) that had stable expression levels.

Statistical analyses were performed using Prism 6 (GraphPad Software, La Jolla, USA). All data are presented as the mean ± SEM, and, unless otherwise stated, p values (obtained using Student’s t-test) < 0.05 were considered significant.

**PCR amplification in coffee clones**

For characterization of the identified gene HdT_LRR_RLK2, a set of coffee clones bearing different resistance gene combinations was analyzed. These differential coffee clones proposed by the Coffee Rust Research Center (Centro de Investigação das Ferragens do Café, Portugal) have been used to characterize the physiological races of *H. vastatrix* (Várzea and Marques, 2005). We also included CIFIC 832/2, the genotype that originated from the BAC library, as a positive control. ‘Caturra’ CIFC 19/1 and ‘Catuaí Amarelo’ IAC 64 (accession UVF 2148-57) were used as negative controls, as they are susceptible to all known Brazilian races of *H. vastatrix*. All clones were vegetatively propagated in the greenhouse in Viçosa, MG, Brazil (20°45'28.7" S, 42°52'11.7" W, altitude of 648 m above sea level).

Genomic DNA was extracted from the young second pair of leaves, following the protocol described by Diniz et al. (2012). DNA integrity was checked by subjecting the samples to electrophoresis in a 1 % gel, followed by staining with ethidium bromide (0.5 μg mL⁻¹); DNA was quantified using a NanoDrop spectrophotometer and stored at –20 °C until use.

PCR was optimized such that the reaction mixture contained 50 ng of genomic DNA, 0.1 μM of each primer P2 to candidate gene HdT_LRR_RLK2, 0.15 mM of each dNTP (Promega), 1 mM MgCl₂, 1 U of Taq DNA polymerase, and 1 × PCR reaction buffer, to yield a final volume of 20 μL. DNA was amplified in a thermocycler, programmed with initial denaturation at 94 °C for 5 min, followed by 34 amplification cycles of denaturation at 94 °C for 30 s, annealing at 66 °C for 30 s, an extension at 72 °C for 1 min and a final extension at 72 °C for 10 min. The amplification products were visualized after electrophoresis on a 6 % polyacrylamide denaturing gel, followed by the staining of the gel with silver nitrate.

**Results**

**BAC clone identification, sequence assembly, and gene prediction**

The clone 104-O-23 [with a ~120 kb insert] was selected from the BAC library owing to the presence of the HT24F133 marker. This marker was developed to amplify a candidate gene identified by Barka et al. (2017), which corresponds to “LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP interacting kinase I-like” and has the potential to be involved in the resistance of coffee to *H. vastatrix*. The BAC clone 104-O-23 was sequenced, and the assembled genomic region had an N50 value of 78.320 and L50 value of 20. The reads were assembled as 145 contigs, with 44.1 % GC content and 55.9 % AT content. In the 145 contigs, 1484 ORFs were found. The comparison with the *C. canephora* and *C. arabica* genomes allowed for the annotation of 37 ORFs in 10 contigs (Table 1).

Based on the annotation, contig 37 (16,570 bp), contig 51 (8,285 bp), and contig 65 (17,594 bp) revealed candidate genes related to host defense to pathogens. Aiming to identify the candidate gene involved with the resistance of coffee to rust as proposed by Barka et al. (2017), we selected contig 37 owing to the presence of genes coding the LRR motif.

Contig 37 had four genes (Table 2). The gene g339.t1 denoted in this study as HdT_LRR_RLK1 is homologous to sequences in *C. canephora* and *C. arabica* (coverage > 50 %) on chromosome 1. Genes g340.t1 and g341.t1 have no homologs to any other genome; however, g342.t1 denoted as HdT_LRR_RLK2 is homologous to sequences in *C. arabica* on chromosome 1e and 1c (coverage > 50 %) (Figure 1A, Table 1). The structures of genes HdT_LRR_RLK1 and HdT_LRR_RLK2 are represented in Figure 1B and C, respectively, showing promoter regions, presence of exons and introns, and anti-sense (3'-5') orientation.

The protein domains in the genes were identified using the PFAM database. HdT_LRR_RLK1 had four LRR domains, one leucine-rich repeat N terminal [LRRNT] domain and protein tyrosine kinases-Phosphokinase (TYR) domain (Figure 1D). HdT_LRR_RLK2 had two LRR domains, and one cytochrome P450 and protein kinase-Pkinase domain each (Figure 1E). These two genes have typical RLK domains, and in general, contain an extracellular domain, a single transmembrane domain, and a cytoplasmic kinase domain. LRR kinases represent a large and functionally diverse family of transmembrane proteins critical for signal recognition and transduction at the plant cell plasma membrane (Huang et al., 2018). The HdT_LRR_RLK1 and HdT_LRR_RLK2 identified in the BAC clone had protein transmembrane helices, evident from the TMHMM analysis (Figure 1D and E).

**Gene expression analysis**

The genes HdT_LRR_RLK1 and HdT_LRR_RLK2 had differential expression in both the compatible and incompatible coffee-*H. vastatrix* interactions (Figure 2). The expression peak of the gene HdT_LRR_RLK1 was observed at 24 h.a.i in the compatible interaction. The expression of HdT_LRR_RLK2 was significantly (p < 0.05) higher at 24 h.a.i in the incompatible interaction, as the appressorium and penetration hypha were differentiated. Based on these results, the HdT_LRR_RLK2 gene was considered to be a candidate resistance gene.
Table 1 – Annotation of genes found in the bacterial artificial chromosomes (BAC) based on *Coffea canephora* and *C. arabica* genomes.

| Contig | ORF | Locus ID | Annotation in *C. canephora* | E-value | Query Coverage (%) | Annotation in *C. arabica* | Locus ID | E-value | Query Coverage (%) |
|--------|-----|---------|------------------------------|---------|-------------------|-----------------------------|---------|---------|-------------------|
| Contig_32_g287.t1 | Cc01_g18850 | Putative uridine kinase C227.14 | 2e-46 | 29.38 | Putative uridine kinase C227.14 isof orm X4 | Ca01e | 113715709 | 8e-12 | 100 |
| Contig_32_g287.t1 | - | - | - | - | - | - | - | - | - |
| Contig_32_g289.t1 | Cc01_g18870 | NADH ubiquinone oxidoreductase 24-kDa subunit, Putative | 8e-60 | 15.82 | NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial | Ca01e | 113715770 | 4e-55 | 100 |
| Contig_32_g289.t1 | Cc01_g18860 | Auxin-induced protein 22D | 2e-74 | 11.93 | Auxin-induced protein 22D like | Ca01e | 113715747 | 2e-98 | 99 |
| Contig_32_g290.t1 | Cc01_g18880 | NADH-ubiquinone oxidoreductase 24-kDa subunit, Putative | 0.0 | 40.55 | NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial | Ca01e | 113715770 | 4e-55 | 100 |
| Contig_32_g291.t1 | Cc01_g18890 | Histone acetyltransferase GCN5 | 0.0 | 27.09 | Histone acetyltransferase GCN5, transcript variant X2 | Ca01e | 113715778 | 4e-89 | 100 |
| Contig_32_g292.t1 | Cc06_g17020 | Putative Receptor-like protein 12 | 3e-28 | 35.55 | Histone acetyltransferase GCN5, transcript variant X2 | Ca01e | 113715778 | 0.0 | 100 |
| Contig_32_g293.t1 | Cc01_g18920 | Protein of unknown function (DUF506) | 0.0 | 32.53 | Uncharacterized protein LOC113715638 | Ca01e | 113715638 | 0.0 | 100 |
| Contig_36_g335.t1 | Cc01_g18900 | Protein of unknown function (DUF506) | 0.0 | 32.53 | Uncharacterized protein LOC113715638 | Ca01e | 113715638 | 0.0 | 100 |
| Contig_36_g338.t1 | Cc01_g18940 | Auxin-induced protein 22D | 0.0 | 46.08 | Protein LON GFOU a-like isoform X1 | Ca01c | 113729570 | 0.0 | 75 |
| Contig_36_g338.t1 | Cc01_g18940 | Auxin-induced protein 22D | 0.0 | 18.35 | Protein LON GFOU a-like isoform X1 | Ca01c | 113729570 | 0.0 | 75 |
| Contig_37_g339.t1 | Cc01_g03450 | Leucine-rich repeat receptor-like protein kinase family protein | 0.0 | 96.28 | Probable leucine-rich repeat receptor-like protein kinase At1g35710 | Ca01e | 113700612 | 0.0 | 100 |
| Contig_37_g340.t1 | Cc03_g08290 | Late blight resistance protein homolog R1B-14 | 4e-111 | 5.00 | Glycine-rich cell wall structural protein 1.8-like | Ca08c | 113701758 | 5e-64 | 21 |
| Contig_37_g341.t1 | Cc03_g08290 | Late blight resistance protein homolog R1B-14 | 1e-77 | 2.00 | Agamous-like MADS-box protein AGL23 | Ca09c | 113701009 | 1e-113 | 11 |
| Contig_37_g342.t1 | Cc01_g03490 | Leucine-rich repeat receptor-like protein kinase family protein | 0.0 | 17.52 | Probable leucine-rich repeat receptor-like protein kinase At1g35710 | Ca01e | 113700612 | 0.0 | 100 |
| Contig_48_g45.t1 | - | - | - | - | - | - | - | - | - |
| Contig_48_g46.t1 | Cc10_g10570 | - | - | - | - | - | - | - | - |
| Contig_50_g52.t1 | Cc11_g11240 | - | - | - | - | - | - | - | - |
| Contig_51_g62.t1 | Cc10_g06400 | Probable WRKY transcription factor 51 | 7e-97 | 28.94 | Probable WRKY transcription factor 51 | Ca01c | 113715139 | 9e-166 | 100 |
| Contig_51_g63.t1 | Cc10_g09400 | P-loop containing nucleotide triphosphate hydrolases superfamily protein | 2e-16 | 1.88 | Probable WRKY transcription factor 51 | Ca01c | 113715139 | 0.0 | 100 |
| Contig_60_g133.t1 | Cc07_g08780 | Protein of unknown function (DUF6) | 4e-108 | 21.09 | Uncharacterized protein LOC113699174 | Ca07c | 113699304 | 1e-99 | 100 |
| Contig_60_g134.t1 | Cc07_g08790 | alpha/beta-Hydrolases superfamily protein | 0.0 | 32.58 | Uncharacterized protein LOC113699174 | Ca07c | 113699304 | 0.0 | 100 |
| Contig_60_g135.t1 | Cc07_g08780 | Mannose-6-phosphate isomerase, type I | 0.0 | 27.17 | Mannose-6-phosphate isomerase 1-like | Ca07c | 113699275 | 0.0 | 100 |
| Contig_60_g136.t1 | Cc07_g08770 | Histone acetylation protein 2 | 4e-97 | 40.59 | Elongator complex protein 5-like | Ca07c | 113700071 | 5e-89 | 99 |
| Contig_60_g137.t1 | Cc07_g08770 | Mannose-6-phosphate isomerase, type I | 0.0 | 27.17 | Mannose-6-phosphate isomerase 1-like | Ca07c | 113699275 | 0.0 | 100 |
| Contig_64_g152.t1 | Cc11_g00930 | - | - | - | - | - | - | - | - |
| Contig_64_g152.t1 | Cc10_g10570 | Calcium-dependent protein kinase 24-like | 5e-143 | 100 |
| Contig_65_g153.t1 | Cc07_g04920 | Trichome birefringence-like 19 | 0.0 | 100 | Trichome birefringence-like 19 | Ca07c | 113737487 | 5e-143 | 100 |
| Contig_65_g154.t1 | Cc07_g04930 | Trichome birefringence-like 19 | 0.0 | 12.91 | Pentatricopeptide repeat-containing protein At1g62670, mitochondrial-like | Ca07e | 113700679 | 0.0 | 99 |
| Contig_65_g155.t1 | Cc07_g04930 | Trichome birefringence-like 19 | 0.0 | 12.91 | Pentatricopeptide repeat-containing protein At1g62670, mitochondrial-like | Ca07e | 113700679 | 0.0 | 99 |
| Contig_65_g156.t1 | Cc07_g04930 | Trichome birefringence-like 19 | 0.0 | 12.91 | Pentatricopeptide repeat-containing protein At1g62670, mitochondrial-like | Ca07e | 113700679 | 0.0 | 99 |
Figure 1 – Structure and position of genes in reference to coffee genomes. A) Chromosome of Coffea arabica and C. canephora with regions matching the coding sequence (CDS) of the gene; B) Structure of partial HdT_LRR_RLK1 showing the positions of CAT BOX = 141 nucleotides, TATA BOX = 27 nucleotides downstream of the 5' UTR region = 2.939-2.941 nucleotides, transcription start site (TSS) = 2.997 nucleotides, Exon = 122-2.997 nucleotides, △ = Start codon 2.939-2.941, CDS = 122-2.941 nucleotides; C) Structure of HdT_LRR_RLK2 showing the positions of CAT BOX = 193 nucleotides and TATA BOX = 29 nucleotides downstream of the 5' UTR region = 35.189-35.191, TSS = 35.247, E1 = 34.897-35.247, △ = Start codon 35.189-35.191, E2 = 33.812-33.913, E3 = 33.203-33.565, E4 = 32.123-33.082, E5 = 31.419-31.986, E6 = 30.782-31.353, E7 = 30.250-30.464, Stop codon 30.462-30.464, CDS = 30.462-30.586, CDS = 30.782-31.353, CDS = 31.419-31.986, CDS = 32.123-33.082, CDS = 33.203-33.565, CDS = 33.812-33.913, CDS = 34.897-35.191, 3' UTR = 30.462-30.464, TTS = 30.250; D) HdT_LRR_RLK1 contains an N-terminal signal peptide domain (SP-LRRNT2, residues 43-83), four leucine-rich repeat domains (LRR8, 111-170, 209-267, 279-339, 616-675), a transmembrane domain (TM, 750-770), and an intracellular kinase domain (Pkinase Tyr, 844-931); E) HdT_LRR_RLK2 contains a N-terminal signal peptide domain (P450, residues 88-151), two leucine-rich repeat domains (LRR8, 144-204 and 392-452), a transmembrane domain (TM, 738-760), and an intracellular kinase domain (Pkinase, 780-896).

Table 2 – Position of the genes found in the contig 37 of the BAC clone in Coffea canephora and Coffea arabica genome.

| Contig 37 | Chr     | Start   | End     | I (%) | E-value | Coverage |
|-----------|---------|---------|---------|-------|---------|----------|
| g339.t1   | 1       | 7.939.294 | 7.936.475 | 98.00 | 0.0     | 97.94    |
| g340.t1   | 3       | 163.880.291 | 163.880.595 | 90.00 | 4e-111  | 10.53    |
| g341.t1   | 3       | 7.943.782  | 7.944.010  | 93.00 | 5e-93   | 20.44    |
| g342.t1   | 1       | 7.971.448  | 7.972.037  | 90.00 | 0.0     | 17.86    |

| Contig 37 | Chr     | Start   | End     | I (%) | E-value | Coverage |
|-----------|---------|---------|---------|-------|---------|----------|
| g339.t1   | Chr1e   | 6.699.099 | 6.701.921 | 98    | 0.0     | 100      |
|           | Chrc    | 7.258.267 | 7.260.275 | 91    | 0.0     | 100      |
| g340.t1   | Chrc8c  | 2.631.263 | 2.631.487 | 87    | 5e-64   | 21       |
| g341.t1   | Chrc9e  | 29.984.070 | 29.984.372 | 91    | 1e-113  | 11       |
| g342.t1   | Chrc1e  | 6.869.945 | 6.870.871 | 87    | 0.0     | 99       |
|           | Chrc1c  | 7.107.697 | 7.107.226 | 92    | 0.0     | 98       |

Gene HdT_LRR_RLK2 in differential coffee clones

The presence of the gene HdT_LRR_RLK2 in the differential coffee clones was analyzed using the P2 primers. The marker was detected in four differential coffee clones, CIFIC 832/2 (S19, S17, S18, S19, S21), CIFIC 832/1 (S19, S17, S18, S19, S21), CIFIC 1343/269 (S16), and H419/20 (S19, S16, S19, S21) [Figure 3, Table 3]. The S16 gene was present in these four clones, suggesting that S16 could be our candidate resistance gene. However, S16 had also been identified in the clones H420/10 and CIFIC.
Table 3 – Screening for gene HdT_LRR_RLK2 marker in a set of coffee clones bearing different resistance gene combinations (differential clones). CIFC 832/2 is the positive control; ‘Caturra’ CIFC 19/1 and UFV 214857 are the negative controls.

| Differential clone (CIFC) | Susceptible to Hemileia vastatrix physiological race | $S_r$ gene | HdT_LRR_RLK2 |
|--------------------------|-----------------------------------------------------|------------|-------------|
| 1 832/2                  | None                                                | 6,7,8,9,?  | +           |
| 2 832/1                  | None                                                | 6,7,8,9,?  | +           |
| 3 Caturra CIFC 19/1      | All                                                 | 5          | -           |
| 4 UFV 214857             | All                                                 | 5          | -           |
| 5 63/1                   | I, II, III, VIII, X, XII, XIII, XIV, XVI, XVII, XXII, XXIII, XXIV, XXVI, XXVII, XXVIII, XXIX, XXX, XXXI, XXXIII, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL, XLI, XLII | 5          | -           |
| 6 HW17/12                | XVI, XXIII, XXXVIII                                 | 1,2,4,5    | -           |
| 7 1343/269               | XXII, XXV, XXVI, XXVII, XXIX, XXI, XXII, XXVII, XXX, XL | 6          | +           |
| 8 H153/2                 | XII, XVI                                            | 1,3,5      | -           |
| 9 H419/20                | XXX, XXXI, XXXVII, XXXIX                           | 5,6,9      | +           |
| 10 H420/10               | XXX, XXXVII, XXXIX                                  | 5,6,7,9    | -           |
| 11 H420/2                | XXX, XXXI, XXXVII, XXXVIII, XLII                    | 5,8        | -           |
| 12 110/5                 | X, XIV, XV, XXII, XXIII, XXV, XXVI, XXVII, XXXV, XXXV, XXXVIII, XXXIX | 4,5        | -           |
| 13 128/2                 | III, X, XII, XVI, XVII, XXI, XXII, XXIV, XXVI, XXVII | 1          | -           |
| 14 134/4                 | X, XVI, XXII, XXIII, XXIV, XXVI                      | 1,4        | -           |
| 15 635/2                 | X, XIV, XV, XVI, XXII, XXIII, XXIV, XXVI, XXVII, XXVIII, XXIX | 4,5        | -           |
| 16 635/3                 | X, XVI, XXII, XXIII                                 | 1,4,5      | -           |
| 17 87/1                  | III, X, XII, XVI, XXII, XXXVIII, XXIX, XL           | 1,5        | -           |
| 18 1006/10               | XII, XVI, XXII, XXIII, XXXVII, XL                   | 1,2,5      | -           |
| 19 7963/117              | XXXIII                                              | 5,7 or 5,7,9 | -  |
| 20 4106                  | -                                                   | 6,7,8,9,?  | -           |
| 21 644/18                | XIII                                                | ?          | -           |
| 22 H147/1                | XIV, XVI                                            | 2,3,4,5    | -           |
| 23 32/1                  | I, VIII, XII, XVI, XXI, XXII, XXIV, XXVIII, XXXI, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL, XLI, XLII | 2,5        | -           |
| 24 33/1                  | VII, VIII, XII, XIV, XVI                           | 3,5        | -           |
| 25 H152/3                | XIV, XVI, XXII, XXV, XXVI, XXVII, XXVIII, XXXV      | 2,4,5      | -           |
| 26 849/1                 | I, II, III, IV, V, X, XII, XIII, XIV, XV, XVI, XVII, XVIII, XXX, XXI, XXII, XXIII, XXIV, XXV, XXVI, XXVII, XXVIII, XXX, XXXI, XXXII, XXXIII, XXXIV, XXXV, XXXVI, XXXVII, XXXVIII, XXXIX, XL, XLI, XLII | ?          | -           |

No = resistance to all known Brazilian races; All = susceptible to all known Brazilian races; + = presence of the HdT_LRR_RLK2 marker; - = absence of the HdT_LRR_RLK2 marker.

Figure 2 – RT-qPCR quantification of two candidate genes (HdT_LRR_RLK1 and HdT_LRR_RLK2) at 0, 12, 24, and 96 h after inoculation (h.a.i) in resistant (‘HDT’ CIFC 832/1) and susceptible (‘Caturra’ CIFC 19/1) genotypes inoculated with Hemileia vastatrix race XXXIII urediniospores. Error bars = three independent biological replicates. * Shows significant difference in the expression levels at the same h.a.i between the interactions, * shows significantly upregulated or downregulated expression levels relative to the uninoculated samples (0 h.a.i).

Discussion

Several genetic studies have identified coffee genes associated with rust resistance. At least nine genes ($S_1$ to $S_9$) have been inferred according to the Flor (1971)
The gene HdT_LRR_RLK2 has domains typical of plant RLKs. RLKs are the major components of pattern-recognition receptor (PRR) complexes that comprise a superfamily of transmembrane proteins, many of which function in pathogen detection as PRRs [Tang et al., 2015]. Plants have evolved various defense mechanisms to combat diseases. The plant’s innate immune system mainly consists mainly of two interconnected branches, termed pattern-triggered immunity and effector-triggered immunity [Jones and Dangl, 2006; Cui et al., 2015; Boutrot and Zipfel, 2017].

Leucine-rich repeat-receptor-like kinases (LRR-RLKs), which contain up to 30 LRRs in their extracellular domain, constitute the largest RLK family [Shiu and Bleecker, 2001]. The LRR_RLK gene family contributes to basal immunity to adapted pathogens and to non-host resistance to non-adapted pathogens through the induction of both local and systemic immune responses with broad-spectrum and potentially durable disease resistance [Boutrot and Zipfel, 2017; Hu et al., 2018; Ranf, 2018].

Based on the candidate resistant gene identified, a functional molecular marker was developed. After validation of this marker in a segregated coffee population for resistance to *H. vastatrix*, the functional marker can be efficiently used for marker-assisted selection of resistant seedlings in coffee breeding programs.

**Acknowledgments**

We thank the Coordination for the Improvement of Higher Level Personnel (CAPES), the Brazilian Coffee Research and Development Consortium (CBP&D/Café), the Minas Gerais State Agency for Research and Development (FAPEMIG), the Brazilian National Council for Scientific and Technological Development (CNPq), and the National Institutes of Science and Technology of Coffee (INCT/Café) for their financial support. We are also grateful to Luiz Filipe Protassio Pereira (Brazilian Agricultural Research Corporation – Embrapa, and Agronomic Institute of Paraná – IAPAR) for providing the Bacterial Artificial Chromosomes Library.
Authors’ Contributions

Conceptualization: Almeida, D.P.; Caixeta, E.T.
Data acquisition: Almeida, D.P.; Castro, I.S.L.; Alves, D.R.; Barka, G.D.; Sakiyama, N.S.
Data analysis: Almeida, D.P.; Zambolim, L.; Mendes, A.O.; Barreiros, P.R.R.M.
Design of methodology: Almeida, D.P.; Caixeta, E.T.; Mendes, A.O.
Writing and editing: Almeida, D.P.; Caixeta, E.T.

References

Alkimim, E.R.; Caixeta, E.T.; Sousa, T.V.; Pereira, A.A.; Oliveira, A.C.B.; Zambolim, L.; Sakiyama, N.S. 2017. Marker-assisted selection provides arabica coffee with genes from other Coffea species targeting on multiple resistance to rust and coffee berry disease. Molecular Breeding 37: article 6.

Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pyshkin, A.D.; Syomin, A.V.; Vyahhi, N.; Tesler, G.; Alekseyev, M.A.; Pevzner, PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19: 455-477.

Barka, G.D.; Caixeta, E.T.; Almeida, R.F.; Alvarenga, S.M.; Zambolim, L. 2017. Differential expression of molecular rust resistance components have distinctive profiles in Coffea arabica - Hemileia vastatrix interactions. European Journal of Plant Pathology 149: 543-561.

Bettencourt, A.J.; Noronha-Wagner, M.; Lopes, J. 1980. Genetic factor that conditions the resistance of clone 1343/269 (Híbrido de Timor) to Hemileia vastatrix et Br. – Factor genético que condiciona a resistência do clone 1343/269 (Híbrido de Timor) à Hemileia vastatrix et Br. Broteria Genética XIII 1: 53-58 [in Portuguese].

Bettencourt, A.J.; Noronha-Wagner, M. 1971. Genetic factors conditioning resistance of Coffea arabica L. to Hemileia vastatrix Berk & Br. Agronomia Lusitana 31: 285-292.

Bettencourt, A.J.; Rodrigues, C.J. 1988. Principles and practice of coffee breeding for resistance to rust and other diseases. Coffee Agronomy 4: 199-234.

Bouhot, F.; Ziepke, C. 2017. Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. Annual Review of Plant Pathology 55: 257-286.

Caçao, S.M.B.; Silva, N.V.; Domingues, D.S.; Vieira, L.G.E.; Diniz, L.E.C.; Vinecky, F.; Alves, G.S.C.; Andrade, A.C.; Carpentieri-Pipolo, V.; Pereira, L.F.P. 2013. Construction and characterization of a BAC library from the Coffea arabica genotype Timor Hybrid CIFC 832/2. Genetica 141: 217-226.

Capucho, A.S.; Caixeta, E.T.; Zambolim, E.M.; Zambolim, L. 2009. Inheritance of coffee leaf rust resistance in Timor Hybrid UFV 443-03. Pesquisa Agropecuária Brasileira 44: 276-282 [in Portuguese, with abstract in English].

Capucho, A.S.; Zambolim, E.M.; Freitas, R.L.; Haddad, F.; Caixeta, E.T.; Zambolim, L. 2012. Identification of race XXXIII of Hemileia vastatrix on Coffea arabica Catimor derivatives in Brazil. Australasian Plant Disease Notes 7: 189-191.

Cressey, D. 2013. Coffee rust regains foothold: researchers marshal technology in bid to thwart fungal outbreak in Central America. Nature 493: 587.

Cui, H.; Tsuda, K.; Parker, J.E. 2015. Effector-triggered immunity: from pathogen perception to robust defense. Annual Review of Plant Biology 66: 487-511.

Diniz, I.; Talhínhas, P.; Azinheira, H.G.; Várzea, V.; Medeira, C.; Maia, I.; Petitot, A.S.; Nicole, M.; Fernandez, D.; Céu, S.M. 2012. Cellular and molecular analyses of coffee resistance to Hemileia vastatrix and nonhost resistance to Uromyces vignae in the resistance-donor genotype HDT832/2. European Journal of Plant Pathology 133: 141-157.

Dodds, P.N.; Rathjen, J.P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. Nature Reviews Genetics 11: 539-548.

Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherton, K.; Holm, L.; Mistry, J.; Sonnhammer, E.L.L.; Tate, J.; Punta, M. 2014. Pfam: the protein families database. Nucleic Acids Research 42: D222-D230.

Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29: 1072-1075.

Hu, P.; Liu, J.; Xu, J.; Zhou, C.; Cao, S.; Zhou, W.; Huang, Z.; Yuan, S.; Wang, X.; Xiao, J.; Zhang, R.; Wang, H.; Zhang, S.; Xing, L.; Cao, A. 2018. A melanin-like/leucine-rich repeat receptor protein kinase gene, RLK-V, regulates powdery mildew resistance in wheat. Molecular Plant Pathology 19: 2561-2574.

Huang, Y.; Jamieson, P.; Shan, L. 2018. The APEX approaches: a unified LRR/RLK network revealed. Trends in Plant Science 23: 372-374.

Hista, J.D.; Dangl, J.L. 2006. The plant immune system. Nature 444: 323-329.

Mahé, L.; Combes, M-C.; Várzea, V.M.P.; Guilhaumon, C.; Lashermes, P. 2007. Development of sequence characterized DNA markers linked to leaf rust (Hemileia vastatrix) resistance in coffee (Coffea arabica L.). Molecular Breeding 21: 105-113.

Noronha-Wagner, M.; Bettencourt, A.J. 1967. Genetics study of resistance of coffea spp. to leaf rust. Identification and behaviour of four factors conditioning disease reaction in Coffea arabica to twelve physiologic races of Hemileia vastatrix. Canadian Journal of Botany 45: 2021-2031.

Pestana, K.N.; Capucho, A.S.; Caixeta, E.T.; Almeida, D.P.; Zambolim, E.M.; Cruz, C.D.; Zambolim, L.; Pereira, A.A.; Oliveira, A.C.B.; Sakiyama, N.S. 2015. Inheritance study and linkage mapping of resistance loci to Hemileia vastatrix in Híbrido de Timor UFV 443-03. Tree Genetics and Genomes 11: 72.

Prakash, N.S.; Marques, D.V.; Várzea, V.M.P.; Silva, M.C.; Combes, M.C.; Lashermes, P. 2004. Introngression molecular analysis of a leaf rust resistance gene from Coffea liberica into C. arabica L. Theoretical and Applied Genetics 109: 1311-1317.
Ranf, S. 2018. Pattern recognition receptors: versatile genetic tools for engineering broad-spectrum disease resistance in crops. Agronomy 8: 134. https://doi.org/10.3390/agronomy8080134.

Ribas, A.F.; Cenci, A.; Combes, M.C.; Etienne, H.; Lashermes, P. 2011. Organization and molecular evolution of a disease-resistance gene cluster in coffee trees. BMC Genomics 12: 240.

Shiu, S.H.; Bleecker, A.B. 2001. Plant receptor-like kinase gene family: diversity, function, and signaling. Science’s STKE 113: re22

Stanke, M.; Steinkamp, R.; Waack, S.; Morgenstern, B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Research 32: W309-W312.

Talhinhas, P.; Batista, D.; Diniz, I.; Vieira, A.; Silva, D.N.; Loureiro, A.; Tavares, S.; Pereira, A.P.; Azinheira, H.G.; Guerra-Guimarães, L.; Várzea, V.M.P.; Silva, M.C. 2017. The coffee leaf rust pathogen Hemileia vastatrix: one and a half centuries around the tropics. Molecular Plant Pathology 18: 1039-1051.

Tang, S.; Teng, Z.; Zhai, T.; Fang, X.; Liu, F.; Liu, D.; Zhang, J.; Liu, D.; Wang, S.; Zhang, K.; Shao, Q.; Tan, Z.; Paterson, A.H.; Zhang, Z. 2015. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (Gossypium hirsutum L.). Euphytica 201: 195-213.

Valencia, A.; Morales, A.Y.; Moncada, M.P.; Cortina, H.A.; Herrera, J.C. 2017. Introgression of the SH3 gene resistant to rust (Hemileia vastatrix) in improved lines of CASTILLO® variety (Coffea arabica L.). Journal of Plant Breeding and Crop Science 9: 130-138.

Várzea, V.M.P.; Marques, D.V. 2005. Population variability of Hemileia vastatrix vs. coffee durable resistance. p. 53-74. In: Zambolim, L.; Zambolim, E.M.; Várzea, V.M.P., eds. Durable resistance to coffee leaf rust. UFV, Viçosa, MG, Brazil.

Zambolim, L. 2016. Current status and management of coffee leaf rust in Brazil. Tropical Plant Pathology 41: 1-8. Silva, M.J.; Pastina, M.M.; Souza, V.F.; Schaffert, R.E.; Carneiro, P.C.S.; Noda, R.W.; Damasceno, C.M.B.; Parrella, R.A.C. 2017. Phenotypic and molecular characterization of sweet sorghum accessions for bioenergy production. PLoS ONE 12: e0183504.

Takano-kai, N.; Jiang, H.; Kubo, T.; Sweeney, M.; Matsumoto, T.; Kanamori, H.; Padhukasahasram, B.; Bustamante, C.; Yoshimura, A.; Doi, K. 2009. Evolutionary history of GS3, a gene conferring grain length in rice. Genetics 182: 1323-1334.

Varshney, R.K.; Nayak, S.N.; May, G.D.; Jackson, S.A. 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends in Biotechnology 27: 522-530.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer, Berlin, Germany.

Yu, G.; Smith, D.K.; Zhu, H.; Guan, Y.; Lam, T.T.Y. 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8: 28-36.