GELETA DUGASSA BARKA

IDENTIFICATION, MOLECULAR CHARACTERIZATION AND DIFFERENTIAL EXPRESSION STUDIES OF GENES ACTIVATED DURING *Coffea arabica* L. - *Hemileia vastatrix* Berk. & Broome INTERACTIONS

Thesis submitted to Universidade Federal de Viçosa, as part of the requirements of the Postgraduate Program in Genetics and Breeding, to obtain the title of *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS - BRAZIL
2017

Ficha catalográfica preparada pela Biblioteca Central da
Universidade Federal de Viçosa - Campus Viçosa

T

B254i
2017

Barka, Geleta Dugassa, 1982-
Identification, molecular characterization and differential expression studies of genes activated during *Coffea arabica* L. - *Hemileia vastatrix* Berk. & Broome interactions / Geleta Dugassa Barka. - Vicosa, MG, 2017.
xvii, 149f. : il. (algumas color.) ; 29 cm.

Inclui anexos.
Inclui apêndices.
Orientador: Eveline Teixeira Caixeta.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Café - Melhoramento genético. 2. Café - Doenças e pragas. 3. *Coffea arábica L.* - *Hemileia vastatrix*. I. Universidade Federal de Viçosa. Outros Órgãos. Programa de Pós-graduação em Genética e Melhoramento. II. Título.

CDD 22 ed. 633.732

GELETA DUGASSA BARKA

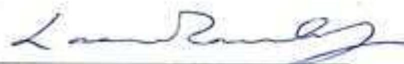
IDENTIFICATION, MOLECULAR CHARACTERIZATION AND DIFFERENTIAL EXPRESSION STUDIES OF GENES ACTIVATED DURING *Coffea arabica* L. - *Hemileia vastatrix* Berk. & Broome INTERACTIONS

Thesis submitted to Universidade Federal de Viçosa, as part of the requirements of the Postgraduate Program in Genetics and Breeding, to obtain the title of *Doctor Scientiae*.

APPROVED ON: 20, February 2017

Antônio Carlos Baião de Oliveira

Tiago Antônio de Oliveira Mendes


Laércio Zambolim
(Co-advisor)

Jorge Luis Badel Pacheco
(Co-advisor)

Eveline Teixeira Caixeta
(Advisor)

Dedicated to:

My father Dugassa Barka Debelo & my mother Degitu Inkosa Gurmu

“Never be afraid to trust an unknown future to a known God.”

-Corrie ten Boom

ACKNOWLEDGEMENTS

Thank you Jesus for everything!

I would like to express my heartfelt gratitude and appreciation to my advisor Prof. Eveline Teixeira Caixeta for her tremendous mentor and guidance all the way through. Your encouragement during hardships and consistent advice on my research and career were priceless. I also thank my co-advisor Prof. Laércio Zambolim for his support and constructive criticisms in conducting the project and in wrapping up the manuscript.

I thank Post Graduate Program in Genetics and Breeding of Universidade Federal de Viçosa for granting my enrollment as a PhD student.

My special thanks go to TWAS-CNPq (Third World Academy of Science - Conselho Nacional de Desenvolvimento Cíentifico e Tecnológico CNPq) for the scholarship and financial support.

The staff of BioCafé laboratory deserves my utmost admiration for the unique friendship and cooperative spirit I enjoyed. I thank Daniela Lelis for her tireless assistance in availing resources and support during laboratory work.

My heartfelt gratitude to Fekadu Gebretensay, Wogayehu Tilahun, Kátia Pestana, Tiago Vieira and others who eased language difficulties during my arrival and for the enjoyable moments we shared. To Juan Carlos and Júlia Rosa, thank you for the great friendship and funs we had. My sincere thanks goes to all my lab mates at BioCafé and others not listed by name but had direct or indirect contribution for my achievement.

I would like to extend my sincere appreciation and thank my brother Ibsa Dugassa, my mother, my father and the whole family for their unconditional love, follow up and advice during my studies of the last four years.

BIOGRAPHY

Geleta Dugassa Barka was born on fourth of October 1982 in Jardega Jarte district, Oromia, Ethiopia. He attended his elementary school at Sute Elementary School from 1988 to 1994. He studied at Alibo Junior Secondary School from 1994 to 1996 and Shambu Secondary School from 1996 to 2000. He joined Haramaya University (the then Alemaya University) in September 2000 and graduated with B.Ed. in Biology in July 2004. Soon after graduation, he was employed at Wolisso preparatory School, Oromia State, where he served for three years. In October 2007, he joined Addis Ababa University for his postgraduate studies and graduated with M.Sc. in Biotechnology in September 2009. He was then recruited as plant biotechnology researcher at Jimma Research Center of Ethiopian Institute of Agricultural Research (EIAR) where he served for one year (September 2009 - October 2010). From October 2010 to February 2013, he was working at Wollega University as a lecturer at the Department of Plant Sciences. In March 2013, he was awarded TWAS-CNPq postgraduate fellowship and admitted for Postgraduate Program (D.Sc.) in Genetics and Breeding at Universidade Federal de Viçosa, MG, Brazil. He submitted his thesis for defense in February 2017.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS .... **ix**

LIST OF ILLUSTRATIONS .... **xi**

LIST OF TABLES .... **xiii**

ABSTRACT .... **xiv**

RESUMO .... **xvi**

1. GENERAL INTRODUCTION **1**

2. LITERATURE REVIEW **4**

2.1. Economic importance of coffee leaf rust [*Hemileia vastatrix* Berk. & Broome] .. 4

2.2. Plant defense against pathogens 4

2.2.1. Plant immunity and mechanisms of resistance development 4

2.2.2. Genes involved in defense signaling 7

2.3. Techniques to study differential gene expression 8

2.4. BAC end sequencing and its applications 11

3. REFERENCES **13**

CHAPTER 1 .... **18**

[i]MOLECULAR RUST RESISTANCE COMPONENTS HAVE DISTINCTIVE EXPRESSION PROFILES IN *Coffea arabica* - *Hemileia vastatrix* INTERACTIONS **18**

COMPONENTES MOLECULARES DA RESISTÊNCIA À FERRUGEM TÊM PERFIS DE EXPRESSÕES DISTINTOS EM INTERAÇÕES ENTRE *Coffea arabica* E *Hemileia vastatrix* **19**

1. INTRODUCTION **20**

2. MATERIALS AND METHODS **23**

2.1 Plant materials and pathogen inoculation 23

2.2 RNA extraction and cDNA synthesis 24

2.3 Suppression subtractive hybridization (SSH) and cloning of ESTs 24

2.4 Sequencing of ESTs 25

2.5 Annotation and homology search 25

2.6 Subcellular localization of ESTs 26

2.7 Quantitative analysis by RT-qPCR 26

2.8 Statistical analysis 30

3. RESULTS **30**
3.1 Isolation of differentially expressed genes 30
3.2 Cloning and sequencing of ESTs 30
3.3 Annotation and metabolic categorization 31
3.4 Upregulated and downregulated genes 34
3.5 Subcellular localization prediction 34
3.6 RT-qPCR quantification 34
4. DISCUSSION **37**
5. REFERENCES **45**
CHAPTER 2 .... **69**
STRUCTURAL AND FUNCTIONAL ANALYSIS OF GENES WITH POTENTIAL INVOLVEMENT IN RESISTANCE TO COFFEE LEAF RUST: A FUNCTIONAL MARKER BASED APPROACH **69**
ANÁLISE ESTRUTURAL E FUNCIONAL DE GENES COM POTENCIAL EM ENVOLVEMENTO NA RESISTANCIA AO FERRUGEM DO CAFEEIRO: UMA ABORDAGEM BASEADA EM MARCADOR FUNCIONAL **70**
1.INTRODUCTION **71**
2. MATERIALS AND METHODS **73**
2.1 Plant genetic materials 73
2.3 PCR conditions 74
2.4 BAC clone screening 74
2.5 Sequencing and contig assembly 75
2.6 Gene prediction and annotation 75
2.7 Motif search 75
2.8 Sequence alignment and comparative analysis 76
2.9 Point mutation analysis 76
2.10 Functional and phylogenetic analysis 76
3. RESULTS **76**
3.1 Resistance gene screening among differential coffee clones 76
3.2 BAC clone identification, sequence assembly and gene prediction 80
3.4 CARF005 amplicon verification 84
3.5 Motif search 84
3.6 Annotation 84
3.7 Gene characterization 86

3.8 Structural and functional site comparative analysis 91
3.9 Interlocus comparison of S_H genes 93
3.10 Phylogenetic analysis 93
4. DISCUSSION 97
5. REFERENCES 102
6. GENERAL CONCLUSIONS AND RECOMMENDATIONS 111
7. APPENDIX I 113
Appendix I Figure 1 113
Appendix I Table 1 114
8. APPENDIX II 145
Appendix II Figure 1 145
Appendix II Figure 2 146
Appendix II Table 1 147
Appendix II Table 2 148
Appendix II Table 3 149

LIST OF ABBREVIATIONS AND SYMBOLS

ASM	Acybenzolar-S-Methyl
avr	Avirulence
BAC	Bacterial Artificial Chromosome
CaNDR1	*Coffea arabica* Non-race specific disease resistance 1
cDNA-AFLP	Complementary DNA –Amplified Fragment Length Polymorphism
CDPKs	Calcium-dependent protein kinases
CIFC	Centro de Investigação das Ferrugens do Cafeeiro
CNL	Coiled-coil Nucleotide binding site leucine rich repeat
C_q	Quantification cycle
dNTPs	Dinucleotide triphosphates
E	Efficiency
EST-SAGE	Expressed sequence tag-sequencing and serial analysis of gene expression
ETI	Effector Triggered Immunity
GA	Gibberellic Acid
GAL	Galactosidase
GO	Gene Ontology
HDT	Hibrido de Timor
HR	Hypersensitive Reaction
IPTG	Isopropyl-β-D-thiogalactoside
KAPP	Kinase associated protein phosphatase
LRR	Leucine Rich Repeat
MAMPs	Microbial Associated Molecular Patterns
MAPKs	Mitogen-Activated Protein Kinases
MIQE	Minimum Information for Publication of RT-qPCR Experiments
NBS-LRR	Nucleotide Binding Site - LRR

NGS	Next Generation Sequencing
NPR1	Non-Expresser pathogenesis related protein 1
PAMPs	Pathogen Associated Molecular Patterns
PE	Paired end
PK	Protein kinase
PPO	Polyphenol oxidase
PR	Pathogenesis-Related protein
PRRs	Pattern recognition receptors
PTI	Pathogen Triggered Immunity
R	Resistance
RAP-PCR	Arbitrarily primed – PCR
RDA	Representational difference analysis
RGAs	Resistance Gene Analogs
RLK	Receptor-Like Kinase
ROS	Reactive Oxygen Species
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SAR	Systemic Acquired Resistance
S_H	Coffee R genes against specific avirulence factors of *H. vastatrix*
SSH	Suppression subtractive hybridization
TM	Trans-membrane
TNL	Toll nucleotide binding site leucine rich repeat
UPGM	Unweighted Pair Group Method with Arithmetic Mean
V	Virulence
WAKs	Wound Associated Kinases
WIPKs	Wound Induced Protein Kinases
YAC	Yeast Artificial Chromosome

LIST OF ILLUSTRATIONS

GENERAL INTRODUCTION

Figure 1. The different types of major resistance proteins in plants……………………….6

Figure 2. Downstream mRNA processing scheme for differential gene expression studies……………………………………………………………………………........9

Figure 3. PCR-Select cDNA subtraction protocol sketch………………………………10

CHAPTER 1

Figure 1. Number of genes upregulated and downregulated during compatible and incompatible interaction at 12 and 24 h.a.i. before redundant ESTs were excluded…….31

Figure 2. Genes specifically and/or commonly expressed at different hours after *H. vastatrix* inoculation in both interactions (after redundant ESTs were removed)……….32

Figure 3. Proportion of different functional categories in each library as described in any of the four databases with $\leq 10^{-5}$ E-value…………………………………………...33

Figure 4. Localization of differentially expressed genes as found in any of the databases mined. Subcellular localization was based on the presence of N-terminal sequences for all the categories…………………………………………………………………...34

Figure 5. RT-qPCR quantification of seven resistance-signaling genes (a-g) at 12, 24, 48 and 72 h.a.i. in resistant (HDT, CIFC-832/1) and susceptible (Catuaí Vermelho IAC 44) genotypes inoculated with *H. vastatrix* race IIurediniospore……………………………………………………………………...36

CHAPTER 2

Figure 1. Mapping of five resistance gene analogs clustered on chromosome 0 of *C. canephora*…………………………………………………………………………83

Figure 2. Mapping of contig 9 to transcriptome of differentially expressed genes during *C. arabica-H. vastatrix* (race XXXIII) incompatible interaction to show the region of active gene (gene 9, 10, 11 and 12) expression……………………………………………87

Figure 3. Conserved domain and motif architecture comparative illustration in gene 5 and 11…………………………………………………………………………………………88

Figure 4. Alignment of gene 5 and 11 encoded proteins and their detected protein binding regions……………………………………………………………………………........90

Figure 5. *In silico* secondary structure, protein and nucleotide binding site prediction for gene 5 and 11……………………………………………………………………………....92

Figure 6. Gene conversion/recombination analysis among homologous genes…………94

Figure 7. Phylogenetic history of gene 5 and 11 in three related genomes………………96

LIST OF TABLES

CHAPTER 1

Table 1. Primer sequences of genes validated by RT-qPCR……………………………27

CHAPTER 2

Table 1. Screening for CARF005 marker using 22 differential coffee clones currently used for S_H gene characterization ………………………………..…………………78

Table 2. Size and structure of five resistance gene analogs and their mapping to chromosome 0 of *C. Canephora* genome………………………………………………81

Table 3. Annotation and functional comparison of gene 5 and 11………………………85

Table 4. Pair-wise synonymous and non-synonymous nucleotide substitution analysis among the six resistance gene analogs (gene 5 and 11 and their respective two top hits as mined by BLASTn in NCBI) ……………………………………………………...95

ABSTRACT

BARKA, Geleta Dugassa, D.Sc., Universidade Federal de Viçosa, February, 2017. **Identification, molecular characterization and differential expression studies of genes activated during *Coffea arabica* L. - *Hemileia vastatrix* interactions Berk. & Broome.** Advisor: Eveline Teixeira Caixeta. Co-advisors: Laércio Zambolim and Jorge Luis Badel Pacheco.

Coffee is one of the most valued cash crops making the economies of many developing countries and sustaining the livelihoods of millions around the world. Despite many decades of breeding efforts with great achievements in incorporating resistance components into elite cultivars, coffee leaf rust (caused by *Hemileia vastatrix*) is increasingly damaging coffee production. Understanding the molecular mechanisms of rust resistance is believed to play a vital part in enhancing resistant cultivar development. The objective of the present work is to understand the expression patterns of resistance genes activated following pathogen inoculation and characterize some major resistance genes. Suppression subtractive hybridization (SSH) was used to identify genes differentially over expressed and repressed at 12 and 24 hours after pathogen inoculation during incompatible and compatible interactions between *C. arabica* and *H. vastatrix*. From 433 clones of expressed sequence tags (ESTs) sequenced, 352 were annotated and categorized of which the proportion of genes expressed during compatible interaction were relatively smaller. RT-qPCR analysis of seven resistance-signaling genes showed similar expression patterns for most of the genes in both interactions, indicating these genes are involved in basal (non-specific) defense during which immune reactions are similar. In another experiment, resistance gene analogs (RGAs) conferring coffee rust resistance were identified from a BAC library, sequenced and characterized. Five RGAs were annotated and mapped to chromosome 0 of *C. canephora*. Four of the RGAs are actively expressed during *C. arabica-H. vastatrix* incompatible interaction. The result obtained in this work suggests that one of the RGAs sequenced (gene 11) is a new S_H gene (S_H10) not yet identified biologically. We also report an S_H gene (S_H10) in differential host clone 644/18 H. Kawisari for the first time. Moreover, comparative analysis of two RGAs belonging to the CC-NBS-LRR gene family showed intense diversifying selection due to nonsynonymous substitution and genetic recombination. Phylogenetic analysis of orthologous genes showed high interaspecies variability among the two genes in related species than in coffee. Overall, differential gene expression

analysis provided a compiled expression profile of genes upregulated and downregulated at 12 and 24 h. a. i. during incompatible and compatible interactions. Likewise, the NBS-LRR genes sequenced in this work are the largest and most complete gene reported in Arabica coffee to date, which makes the work extremely important for molecular breeding of coffee rust resistance.

RESUMO

BARKA, Geleta Dugassa, D.Sc., Universidade Federal de Viçosa, fevereiro de 2017. **Identificação, caracterização molecular e estudos de expressão diferencial de genes ativados durante as interações *Coffea arabica* L. - *Hemileia vastatrix* Berk. & Broome**. Orientadora: Eveline Teixeira Caixeta. Coorientadores: Laércio Zambolim e Jorge Luis Badel Pacheco.

O café é uma das culturas de maior valor econômico mundial, além de proporcionar e qualidade de vida para milhões de pessoas em países em desenvolvimento. Embora existam vários programas de melhoramento para essa cultura e cultivares comerciais disponíveis que apresentam fatores de resistência a estresse biótico, verifica-se ainda prejuízos significativos devidos à ferrugem do cafeeiro causada por *Hemileia vastatrix*. A compreensão dos mecanismos moleculares da resistência à ferrugem desempenha papel importante na eficiência do desenvolvimento de novas cultivares resistentes. O objetivo do presente trabalho é identificar, caracterizar e compreender os padrões de expressão de genes de resistência do cafeeiro que são ativados após inoculação com *H. vastatrix*. Foi utilizada a metodologia de Hibridação Subtrativa de Supressão (HSS) para identificar os genes diferencialmente expressos (*upregulated* e *dowregulated*) às 12 e 24 horas após inoculação (h.a.i.) de *Coffea arabica* com o patógeno, em interações compatíveis e incompatíveis. A partir de 433 clones obtidos e sequenciados, 352 foram anotados e categorizados. Observou-se proporção relativamente menor de genes expressos em interação compatível. A análise RT-qPCR de sete genes de sinalização de resistência mostrou padrões de expressão semelhantes para a maioria dos genes em ambas as interações, indicando que esses genes estão envolvidos na resistência (não específica) durante a qual as reações imunes são semelhantes. Na segunda etapa do trabalho, *resistance gene analogs* (RGAs), que conferem resistência à ferrugem do café, foram identificados, sequenciados e caracterizados a partir de uma biblioteca BAC do cafeeiro. Cinco RGAs foram anotados e mapeados no cromossomo 0 (zero) de *C. canephora*. Destes, quatro RGAs são ativados na interação incompatível entre *C. arabica* e *H. vastatrix*. Os resultados obtidos no trabalho sugerem que um desses genes RGA sequenciado (gene 11) é um novo gene S_H (S_H10) ainda não identificado biologicamente. Com base nesses dados, foi verificado pela primeira vez o novo gene S_H (S_H10) no clone diferenciador 644/18 H. Kawisari. Foi realizado a análise comparativa entre os cincos RGAs e verificado alta similaridade entre dois destes, os quais são pertencentes à família

de genes CC-NBS-LRR. Foi verificado intensa seleção diversificada promovida pela substituição não sinônima e pela recombinação genética. Foi realizada a análise filogenética de genes ortólogos para as espécies de café, tomate e uva e observou-se alta variabilidade intraespecífica destes dois genes CC-NBS-LRR para as espécies, exceto para o café. Estes genes sequenciados são as maiores e mais completas sequências disponíveis para o *C. arabica*. Estes resultados são de extrema importância para o melhoramento genético molecular visando a resistência à ferrugem do cafeeiro. De modo geral, a compreensão dos padrões de expressão de genes de resistência e a caracterização molecular de novos RGAs são resultados valiosos e estabelece nova base para estudos futuros.

1. GENERAL INTRODUCTION

Coffea arabica (L.), a native cash crop to *Kaffa* of south west Ethiopia, is an allotetraploid (2n = 4x = 44) dicot which belongs to family *Rubiaceae* (Davis et al. 2006). Coffee is the second most traded commodity, of which 51.37 million of sacks were produced in 2016 corresponding 30% of the world´s coffee production (CONAB 2017). Coffee is not only a beverage to enjoy, as perceived by most consumers, but also has cultural and social values. Though its consumption is on rise from time to time, its production is under threat due to climate change (Iscaro 2014) and diseases (www.agriskmanagementforum.org/content/major-pests-and-diseases-coffee). Among the main diseases, coffee leaf rust or simply coffee rust caused by biotrophic fungus *H. vastatrix* is challenging the sustainability of coffee production in almost all coffee producing regions. This fungal disease has been seriously affecting the incomes of coffee farmers with a percussive effect on the economies of coffee producing countries and consumers as well. As climate change is reaching every parts of the world, coffee leaf rust is adversely affecting coffee producing regions around the globe (Stone 2014).

To deal with biological, cultural or traditional means of control, it is an essential step to understand factors playing in coffee rust epidemiology and strategies for intervention (Rutherford and Phiri 2006). The use of chemical pesticides (copper-based fungicides) is considered the most reliable and popular due to their effective protection. However, it is now understood that with non-target effects and resistance of pathogens, the risks to human lives and to the environment are so evident that there is a need for a shift to crop protection techniques which are less dependent on chemicals (Adejumo 2005).

Since 1920s, when the disease was affecting coffee production in Africa and Asia, different approaches of countering the disease have been in place. With its promising potential, developing resistant varieties by classical hybridization is of paramount importance and indisputably came with the most breakthrough being technically efficient and environmentally friendly. One of such prominent result was from random natural interspecific hybridization of *C. arabica* and *C. canephora* resulting in Híbrido de Timor (HDT), a genotype originally from Timor Island discovered by Coffee Rust Research Center (CIFC) in 1950's. As molecular genetics is unprecedentedly evolving, hence its immense applications in developing resistant cultivars, moving resistance genes within

and between species is likely to ease the hurdles in classical breeding. In line with the classical breeding through hybridization between and within coffee species and cultivars, identification, characterization and expression pattern profiling of resistance genes could ease the development of resistant cultivar, which would in turn pay-off coffee producers.

The fact that coffee is one of the economically valuable commodities, its increased production at minimum cost has a fundamental importance for both producers and consumers. As coffee rust was recognized as one of the challenges facing coffee production, works on resistance breeding assisted by molecular genetics has been appealing ever since. To that end; isolation, characterization and expression pattern studies of resistance and related genes advance further understanding in molecular breeding for the development of resistant varieties in the most sustainable and efficient way. So far, isolation of differentially expressed genes due to *H. vastatrix* infection has been reported in different studies (Fernandez et al. 2004; Guzzo et al. 2009; Diola et al. 2011). However, comparative studies of differential gene expression in contrasting conditions following pathogen inoculation, database mining and homology search, identification, characterization, expression and regulation patterns of coffee rust resistance related genes are still ways forward. Furthermore; identification, characterization and understanding resistance genes in resistant and susceptible cultivars is an essential component of resistance breeding as it enables develop resistant coffee varieties with the most resource effective methods.

On the other hand, the discovery and characterization of disease resistance genes has been on the rise in resistance-breeding researches. Despite the complete sequencing of several plant genomes in the last 15 years or so, there are few reports in characterizing RGA (resistance gene analogs) to understand their structure, function and evolution as part of an effort to develop novel disease resistant cultivars (Ribas et al. 2011; Marone et al. 2013; Rajesh et al. 2015; Sekhwal et al. 2015). As NBS-LRR (nucleotide binding site-leucine rich repeat) RGAs are one of main resistance genes in plant defense and widely distributed across the entire genome, they are primary targets in plant genetic engineering for resistance development. Since the complete sequence of *C. canephora* genome (Denoeud et al. 2014) and assembly of *C. arabica-H. vastatrix* interaction transcriptome (Freitas 2015) has facilitated the mapping, discovery, characterization and phylogeny

studies of resistance genes to coffee rust. Such efforts have an invaluable importance in controlling coffee rust disease, which is currently affecting coffee production worldwide. Moreover, as the sequencing platforms and data bases are ever advancing, data storage, exchange and processing are indispensable in determining the complexity of molecular pathology in broader dimensions. As such understanding the nature, diversity and expression patterns of genes activated or suppressed following *H. vastatrix* inoculation makes essential part of resistance breeding programs. To that end, the general objective of this work was to investigate differential expression pattern and characterize genes involved in the defense system of *C. arabica* against *H. vastatrix*.

The specific objectives were:

- To identify differentially expressed genes during the compatible and incompatible interactions between coffee and *H. vastatrix*
- To quantitatively evaluate the expression of some genes involved in resistance perception and transduction during incompatible and compatible interactions at different hours after *H. vastatrix* inoculation;
- To sequence and characterize a RGA (CARF005) functional marker and associated regulatory elements
- To determine whether CARF005 belongs to genes expressed in defense against *H. vastatrix*
- To map rust resistance gene loci to *C. canephora* genome and analyze the distribution and physical distance between RGAs
- To screen CARF005 gene in differential host coffee clones conferring S_H1-9 resistance genes
- To investigate interspecies and intraspecies diversity/polymorphism of sequenced RGAs in coffee and related species.

2. LITERATURE REVIEW

2.1. Economic importance of coffee leaf rust [*Hemileia vastatrix* Berk. & Broome]

Coffee leaf rust is severely damaging coffee production throughout the entire coffee producing countries. It is a rust affecting leaves by defoliation, which weakens the plant resulting in reduced yield in the following year. During the last two years, it has threatened the economies and employment of coffee producing countries in Central America (http://www.ico.org/new_historical.asp). In some regions in Brazil with high prevalence, the loss is estimated to be as high as 50% (Zambolim 2016). As a result, the demand and use of fungicide has been increased during the last couple of years partly due to financial institutions providing credits for the vulnerable farmers in Central America (Avelino et al. 2015). In general, the efficiency of rust control using fungicide depends on funds, technology and operations, in addition to climate conditions. The cost of coffee leaf rust control could be as high as US$ 500 /ha / year (Sera et al. 2007), making it difficult for the low income farmers. This makes the use of resistant cultivars to be economically and ecologically the most preferable choice in controlling the disease.

2.2. Plant defense against pathogens

2.2.1. Plant immunity and mechanisms of resistance development

Plants have multiple layers of sophisticated defense mechanisms that recognize potentially dangerous pathogens and rapidly respond before those organisms have a chance to cause serious damage (N. M. Sanabria et al. 2010). Once the pathogen attacks plant tissue, the host plant challenges the advancement of the infection in a series of defense reactions. Basal resistance is the first line of pre-formed and inducible defense that protects plants against various groups of pathogens (Thomma et al. 2011). It can be triggered when plant cells recognize pathogen-associated molecular patterns (PAMPs), including specific proteins, lipopolysaccharides, and cell wall components commonly found in microbes (Freeman 2008). Recognition of PAMPs is executed by the transmembrane pattern recognition receptors (PRRs) resulting in PAMP-Triggered Immunity (PTI) that stops further colonization of the pathogen. Successful pathogens deploy effectors that would deceive basal defense for further infection and colonization. The interference of effectors in such counter attack against PTI may result in effector

triggered susceptibility (ETS) followed by the plant response in a more rigorous and amplified form of PTI known as ETI to continuously check the ever evolving pathogen effectors (Jones and Dangl 2006). This chain of effector-resistance gene co-evolution is at least attributed to mutation and horizontal transfer of genes of the pathogen and selection pressure on plant for resistance (Anderson et al. 2010).

The response of ETI, which instantly follows suppression of PTI (or as also called basal defense), is in the form of hypersensitive response (HR). HR is characterized by deliberate plant cell suicide at the site of infection. Once the HR has been triggered, plant tissues become highly resistant to a broad range of pathogens for extended period. Ultimately, HR leads to the establishment of systemic acquired resistance (SAR) which represents a heightened state of resistance in which distal part of the plant may develop resistance to further infection (Freeman 2008). Moreover, in addition to providing life-long protection for the plant, this immune memory is transmitted to subsequent generations (Jones and Dangl 2006).

In coffee-rust interaction, it has been reported that there are at least nine different types of virulence genes (V_{1-9}) for which there are believed to exist nine corresponding resistance genes (S_H1-9) as proposed by gene-for-gene theory (Mayne 1932). It implies a one-for-one relationship between a pathogen avirulence (avr) gene and a plant resistance gene. Recognition of the avr gene by the resistance gene (R-gene) triggers a classical signal-transduction cascade to induce plant susceptibility (Mayne 1932). Such interactions in which susceptibility is developed by the recognition of the pathogen effector factor is known as compatible interaction unlike the incompatible interaction during which no effector recognition leads to resistance development (Flor 1971; Keen 1990). Dominantly inherited, in plants generally, the largest class of R-genes encode nucleotide binding site leucine-rich-repeat (NBS-LRR) protein that recognizes the corresponding pathogen avr factor (McHale et al. 2006; Jones and Dangl 2006). Intracellular signaling similar to drosophila toll/mammalian interleukin-1 receptor (TNL, Toll-NBS-LRR) and the coiled-coil (CNL, CC-NBS-LRR) are the two major amino acid sequences preceding NBS domain on the N-terminal committed to specific signal transduction (DeYoung and Innes 2006; Jones and Dangl 2006; Tan and Wu 2012). The other domains linked to LRR include (among the others) leucine-zipper (a transmembrane protein, TM), protein kinase (PK) and WRKY TIR on the N-terminal (Liu et al. 2007).

These domains are predominantly involved in resistance signal transduction by conformational change (Leipe et al. 2004). On the carboxyl-terminal region is the LRR mediating specific protein-protein interaction to recognize pathogen effectors (Van der Hoorn 2001; Kushalappa et al. 2016). Nucleotide polymorphism and high variability of LRR region of the R-gene is responsible for pathogen specificity (Jeff Ellis et al. 2000; McHale et al. 2006). Inter and intraspecific extreme variability of NBS-LRR has been attributed to gene duplication, unequal crossing over, recombination, deletion, sequence exchange, point mutation and selection pressure due to continuous response to diverse pathogen population (Yang et al. 2008; Ribas et al. 2011). From over 70 resistance genes categorized into 14 groups with different domains, most of the cloned and characterized have LRR domain providing diverse resistance gene variants for the engineering of disease resistance cultivars and evolution studies (Liu et al. 2007) (Figure 1).

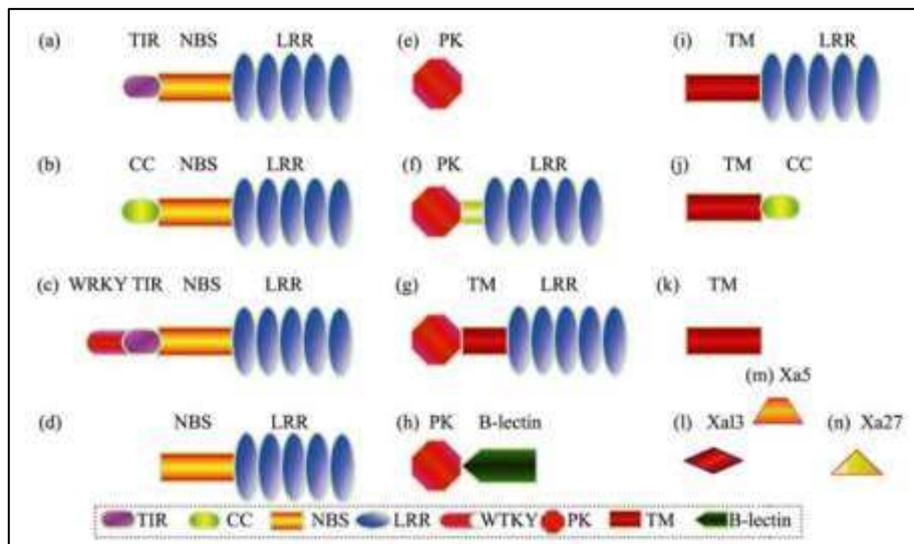


Figure 1. The different types of major resistance proteins in plants. 14 (a-n) different types of R genes based on conserved domains. Three novel proteins with no conserved domains are shown as Xa13, Xa5 and Xa27. TIR: Toll-interleukin-1 receptor; CC: coiled-coil; TM: transmembrane domain; PK: Protein Kinase; WRKY: WRKY domain; B-lectin: bulb-type mannose specific binding lectin domain (Liu et al. 2007).

The direct interaction of plants against fungal pathogens is mainly manifested by the expression of antifungal cell wall components (Freeman 2008). Hence, fungal cell degradation is the major step in defending target plant against further attacks by the

fungus. A group of hydrolytic proteins called pathogenesis related proteins (PR) mainly carries out these activities. PRs are all induced in both susceptible and resistant cultivars. Even though high accumulations are evident in susceptible cultivars, it is a delayed response and has no resistance function (Eugenia Rivera et al. 2002). It is known that distinct sets of PRs are induced in response to different pathogens. In *Arabidopsis thaliana,* PR-1, PR-2 (a β-1,3-glucanase) and PR-5 (thaumatin) are induced by salicylic acid in response to biotrophic pathogens, whereas PR-3 (chitinase), PR- 4 (chitinase) and PR-12 (defensin) are induced by jasmonic acid in defense against necrotrophic pathogens (Spoel and Dong 2012). Yet in other study, the synthesis of PRs is also induced by immune signals in the absence of pathogen challenge (Eugenia Rivera et al. 2002). Plants generally have multiple isoforms of PRs encoded by multiple gene families (Eugenia Rivera et al. 2002). β-1, 3-glucanase proteins are among the PRs which are broadly classified as class I and II across different types of plants. These features are common among the other PRs in which most class I are basic proteins while others are acidic proteins with distinctive sites of action and localization (Leubner-Metzger 2003). The other enzyme active after pathogenic or physical wounding is polyphenol oxidase (PPO) which is mainly involved in the oxidation of induced phenolic compounds (Melo et al. 2006).

2.2.2. Genes involved in defense signaling

One of the major components of plant molecular defense is the perception and detection of race-specific effectors and non-race-specific elicitors (PAMP) of the invading pathogen and the changes that follow inside the plant tissue under attack. Elicitors are diverse classes of molecules whose detection and downstream signal perception rely on the specific receptors on the surface of plant cells and in the cytoplasm (Peck 2003). These signal receptors, broadly known as receptor-like protein kinase (RLK) family, are transmembrane proteins with cytoplasmic serine/threonine kinase and diverse extracellular domains (Morris and Walker 2003). Some of the rapidly responsive kinases are mitogen activated protein kinases (MAPKs), calcium dependent protein kinases (CDPKs), wall-associated kinases (WAKs) and wound induced protein kinases (WIPKs) (Wurzinger et al. 2011; Matschi et al. 2015). Reactive nitrogen oxide and different types of peroxidases and oxidases are also involved in direct and indirect defenses (Lamb and Dixon 1997). The latter two groups of oxidizing enzymes have a direct antimicrobial function and an indirect role being involved in resistance signaling pathways (Levine et

al. 1994; Lamb and Dixon 1997). Though their involvement in signal transduction has been reported in different model plants, how they transmit information is little understood (Romeis 2001; Jonak et al. 2002; Peck 2003). The other defense signaling pathways are those involving secondary messengers yet whose functions are regulated by antagonistic and synergetic interactions of jasmonic acid, salicylic acid and ethylene in defense of biotrophic pathogens (Yang et al. 2015). Some of the overall outcomes of these signaling pathways are the expression of transcription factors, cell wall fortification, HR and induction of different types of resistance genes (Delledonne 2005).

CaWRKYs are transcription factors reported to induce the expression of NPR1 (non-expressor of pathogenesis-related gene1) that enhance resistance against bacterial speck in *Arabidopsis thaliana* (D. Yu et al. 2001). They function as signaling gene to trigger expression of defense genes in both compatible and incompatible interactions in *C. arabica* against *H. vastatrix*. CaWRKY 1 - 20 are expressed in coffee and other plants, time required for their activation after pathogen inoculation differing in resistant and susceptible cultivars (Petitot et al. 2008).

Though limited understanding of the exact role they play in resistance signaling, the expression of CaR111 and *C. arabica* non race specific disease resistant protein (CaNDR1) has been found to significantly enhance the response to virulent rust strains indicating their importance in resistance signaling pathways (Ganesh et al. 2006). Specifically, an orthologous protein CaNDR1 cloned from *C. arabica* was reported to induce nonspecific resistance genes in *N. benthamiana* and enhanced its resistance against coffee leaf rust (Cacas et al. 2011).

2.3. Techniques to study differential gene expression

Differential gene expression has been used to study the relative expression of genes in experimental samples as compared to the control. In most of the strategies, cDNA is used as a starting material to be used in several methods of downstream differential expression studies (Figure 2). cDNA libraries are created for experimental and control samples so that differential expression is detected either by direct sequencing or involving different techniques of hybridization (Casassola et al. 2013). In RNA fingerprinting by arbitrarily primed – PCR (RAP-PCR) and EST sequencing and serial analysis of gene expression

(EST-SAGE) cDNA, segments are electrophoresed and compared with control groups to identify differentially expressed genes (Moody 2001). cDNA-AFLP, nucleic acid microarray, SSH and representational difference analysis (RDA) are based on the principle of hybridization where the target sequences are separated by a complementary sequence (Diatchenko et al. 1999; Guindalini and Tufik 2007). With advancement of sequencing platforms, direct RNA sequencing (RNA-Seq) as the case in transcriptomic studies, coupled with RT-qPCR is used to discover and identify differentially expressed genes making the method more practical (Casassola et al. 2013). This combination of techniques is also used to sequence RNA starting from small amount of transcripts without prior knowledge of the genes of an organism and at the same time quantifies differentially expressed genes (Howald et al. 2012). With the suppression effect of the PCR, SSH provides a powerful technical capability in selectively eliminating undesirable DNA fragments while differentially expressed gene fragments are exponentially amplified and enriched with high efficiency (Byers et al. 2000).

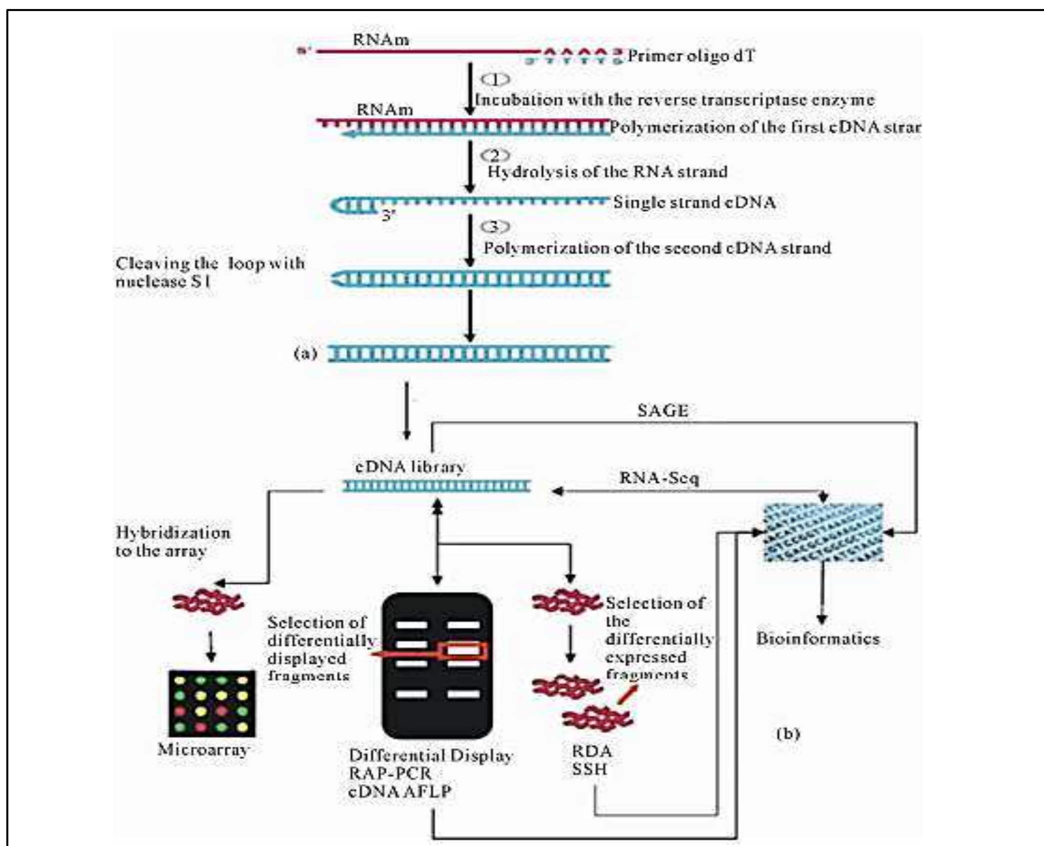


Figure 2. Downstream mRNA processing scheme for differential gene expression studies. cDNA synthesis by reverse transcription (a) and the different types of differential gene expression studies (b). Adapted from Addison Wesley Longman, Inc.

Subtractive hybridization and its derivatives in differential expression studies

Identification of genes differentially expressed at a given developmental stage of pathological process is obviously a critical step in gene expression profiling. Different variants of subtractive hybridization have been adapted as part of an effort towards resistance gene characterization, is always a priority in plant pathology. One of the commonly used is suppression subtractive hybridization (SSH) to separate upregulated genes from those suppressed (Figure 3). As all subtractive hybridization derivatives are not based on prior knowledge of the genes to be screened and hence leading to the discovery of novel genes upregulated at a given biological stage, such techniques are opted to be preferential choices in differential expression studies (Byers et al. 2000).

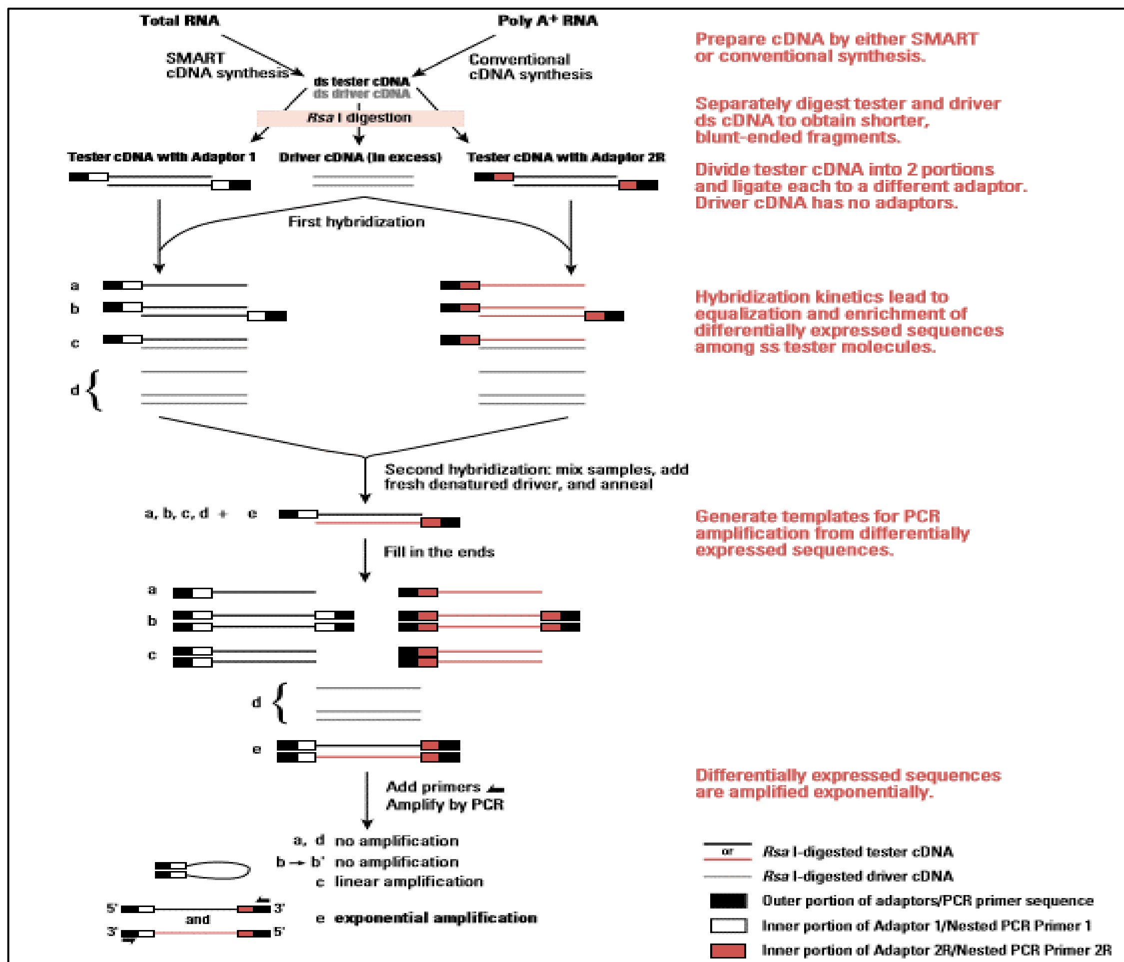


Figure 3. PCR-Select cDNA subtraction protocol sketch. Hybridization kinetics leading to differentially expressed sequences exponentially amplified. Clontech Laboratories, Inc.

2.4. BAC end sequencing and its applications

BAC (bacteria artificial chromosome), cloning tool based on *Escherichia coli* fertility factor (F-factor), is an extremely important molecular tool for physical mapping, map based cloning, analysis of gene structure and function and high throughput genome sequencing (Kelley et al. 1999; Dereeper et al. 2013). Due to evolving sequencing platforms and next generation sequencing (NGS), construction has been routinely applied in genome wide surveying studies in Arabica and Robusta genotypes (Combes et al. 2004; Cação et al. 2013; Dereeper et al. 2013) and recently in complete genome sequencing of *C. canephora* (Denoeud et al. 2014). Insert stability, accommodation of large DNA fragment and easy manipulation makes BAC library a primary choice in functional and structural genomics studies (Kelley et al. 1999; Cação et al. 2013).

Among the other applications of BAC library, BAC end sequencing for the assembly of whole genome, gene mapping, gene characterization and sequencing of coding and regulatory sequences have indispensable contribution in understanding the genome of complex plants (Shizuya et al. 1992; Zhang and Wu 2001; Farrar and Donnison 2007; Figueiredo et al. 2011). Universal primers that amplify the flanking vector sequences and portion of the DNA fragment incorporated for cloning may be used in BAC end sequencing. Or else, sequencing of the harboring vector and the inserted fragment of DNA altogether could be sought as a single shoot where post sequencing intensive trimming is performed. As BAC end sequencing is the preferred method to obtain minimal overlapping clones, it is considered as an option in sequencing large genome size. Though challenging in obtaining sufficient quantity of high quality template DNA by mini-prep, *E. coli* F-factor replicon limits the number of BAC copy number to one to two per cell, which in turn minimizes DNA rearrangements resulting in high quality sequencing (Kelley et al. 1999). The capacity of BAC vector to accommodate large segments of genome (30-300kb, on average) has also facilitated the speed and efficiency of cloning and sequencing of genomes with high polyploidy levels (Villalobos et al. 2004; Denoeud et al. 2014). The construction and readily availability of Arabica coffee BAC libraries in a number of laboratories has accelerated many studies involving resistance gene cloning (Combes et al. 2004; Cação et al. 2013). Furthermore, construction of arabica coffee BAC libraries has been improving in minimizing the composition of organelle DNAs (a factor affecting nuclear genome targeting works) and genome coverage since the work of

Combes et al. (2004) to the recent report of Cação et al. (2013). These BAC libraries of *C. arabica* cultivars resistant to most diseases (Leroy et al. 2005; Dereeper et al. 2013) and additional sources of genetic information from sequenced genome of *C. canephora* (Denoeud et al. 2014) are extremely important resources in coffee breeding programs.

3. REFERENCES

Alvarenga, M.S., Caixeta, T.E., Hufnagel, B., Thiebaut, F., Maciel-Zambolim, E., Zambolim, L., & Sakiyama, S.N. (2010). In silico identification of coffee genome expressed sequences potentially associated with resistance to diseases. *Genet. Mol. Biol.* **33**, 795–806.

Anderson JP, Gleason CA, Foley RC, et al (2010). Plants versus pathogens: an evolutionary arms race. Funct Plant Biol 37:499. doi: 10.1071/FP09304

Adejumo, T. (2005). Crop protection strategies for major diseases of cocoa, coffee and cashew in Nigeria. A review. *African Journal of Biotechnology* **4** (2):143-150.

Bettencourt, A. J. and Rodrigues Jr, C.J. (1988). Principles and practice of coffee breeding for resistance to rust and other disease. *In*: Resistance to Leaf Rust in Coffee Carrying SH_3 Gene and other SH Genes. *Brazilian Archives of Biology and Technology* **50** (5): 753-757.

Byers, J.R., Hoyland, A.J., Dixon, J. and Freemont, J.A. (2000). Subtractive hybridization – genetic takaways and the search for meaning. Current status review. *Int. J. Exp. Path*. **81**:391-404.

Cabral, P.G.C., Zambolim, E.M., Zambolim, L., Lelis. T.P., Capucho, A.S. and Caixeta, E.T. (2009). Identification of a new race of *Hemileia vastatrix* in Brazil. *Australasian Plant Dis Notes* **4**:129-130.

CONAB. (2017). *Acompamento da safra brasileira : café*. Brasília. http://www.conab.gov.br

Casassola, A., Brammer, P. S., Chaves, S. M., Martinelli, A. J., Grando, F.M. and Denardin, D.N. (2013). Gene expression: A review on methods for the study of Defense-related gene differential expression in plants. *American Journal of Plant Sciences* **4**:64-73.

Chen, Z., Chen, C. and Yu, D. (2001). Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression. The American Society of Plant Biologists, *The Plant Cell* **13**:1527-1539.

Cruz, F., Kalaoun, S., Nobile, P., Colombo, C., Almeida, J., Barros, L., Romano, E., Grossi-de-as, M., Vaslin, M. and Alves-Ferreira, M. (2009). Evaluation of coffee reference genes for relative expression. Studies by quantitative Real-Time RT-PCR. *Mol Breeding* **23**:607-616.

Davis, P.A., Govaerts, R., Bridson, M.D. and Stoffelen, P. (2006). An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Botanical Journal of the Linnean Society* **152**:465-512.

Delledonne, M. (2005). NO news is good news for plants. *Curr. Opin. Plant Biol.* **8**:390-396.

Diatchenko, L., Lukyanov, S., Lau, F. Y. and Siebert, D. P. (1999). Suppressive Subtraction Hybridization: A versatile method for identifying differentially expressed genes. *Methods in Enzymology* **303**:349-380.

Diola, V., Brito, G. G., Caixeta, T. E., Maciel-Zambolim, E., Sakiyama, S. N. and Loureiro, E. M. (2011). High density genetic mapping for coffee leaf rust resistance. *Tree Genetics and Genomics.* DOI 10.1007/s11295-011-0406-2.

Dong, X. and Spoell H. S. (2012). How do plants achieve immunity? Defense without specialized immune cells. Macmillan Publishers Limited *Nature Reviews, Immunology* **12**:89.

Ebel, J. and Scheel, D. (1997). Signals in host-parasite interactions. *In*: Coffee resistance to the main diseases: leaf rust and coffee berry disease. *Braz. J. Plant Physiol.* **18**(1):119-147.

Fernandez, D., Cacas, L. J., Petitot, S. A., Bernier, L., Estevan, J., Conejeros, G. and Mongrand, S. (2011). Identification and characterization of the nonrace specific disease resistance 1 (NDR1) orthologous protein in coffee. *BMC Plant Biology* **11**:144.

Fernandez, D., Ganesh, D., Petitot, A. S., Silva, M. C., Alary, R. and Lecouls, A. C. (2006). Monitoring of the early molecular resistance responses of coffee (*Coffea arabica* L.) to the rust fungus (*Hemileia vastatrix*) using real-time quantitative RT-PCR. *Plant Science* **170**:1045-1051.

Fernandez, D., Lecouls, A.C. and Petitot, A. S. (2008). Sub-genomic origin and regulation patterns of a duplicated *WRKY* gene in the allotetraploid species *Coffea arabica. Tree Genetics & Genomics* **4**:379-390.

Fernandez, D., Santos, P., Agostini, C., Bom, C.M., Petitot, S. A., Silva, C.M., Guerra-Guimaraes, L., Ribeiro, A., Argout, X. and Nicole, M. (2004). Coffee (*Coffea arábica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). *Molecular Plant Pathology* **5**(6):527-536.

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **9**:275-296.

Freeman, B. C. and G. A. Beattie (2008). An Overview of Plant Defenses against Pathogens and Herbivores. *The Plant Health Instructor*. DOI: 10.1094/PHI-I-2008-0226-01.

Gichuru, K.E., Ithiru, M.J., Silva, C.M., Pereira, P.A. and Varzea, P.M.V. (2012). Additional physiological races of coffee leaf rust (*Hemileia vastatrix)* identified in Kenya. Brazilian Phytopathology Society, *Tropical Plant Pathology*, **37**(6):424-427.

Guindalini, C. and Tufik, S. (2007). Use of Microarrays in the Search of Gene Expression Patterns-Application to the Study of Complex Phenotypes. *Revista Brasileira de Psiquiatria* **29** (4):370-374.

Guzzo, D. S., Harakava, R., and Tsai, M. S. (2009). Identification of coffee genes expressed during systemic acquired resistance and incompatible interaction with *Hemileia vastatrix*. *J Phytppathol* **157**:625-638.

Howlad, C., Tanzer, A., Chrast, J., Kokocinski, F., Der- rien, T., Walters, N., M. Gonzales, M. J., Frankish, A., Aken, L. B., Hourlier, T., Vogel, H-J., White, S., Searle, S., Harrow, J., Hubbard, J. T., Guigo, R. and Reymond, A. (2012). Combining RT-PCR-seq and RNA-Seq to Catalog All Genic Elements Encoded in the Human Genome. *Genome Research* **22**:1698-1710.

Iscaro, J. (2014). The impact of climate change on coffee production in Colombia and Ethiopia. *Global Majority E-Journal* **5** (1): 33-43.

Jonak, C., Okresz, L., Bogre, L., Hirt, H. (2002). Complexity, crosstalk and integration of plant MAP kinase signaling. *Curr. Opin. Plant Biol.* **5**:415-424.

Jones, G. D. and Dangl, L. J. (2006). The plant immune system: A review, *Nature* **446**:323-329. doi:10.1038/nature05286

Juliana, M. (2012). Brazilian Coffee Industry. The Brazil Business, URL: http://thebrazilbusiness.com/article/brazilian-coffee-industry

Kumar, D. and Sreenath, H. L. (2012). Isolation of Nucleotide Binding Site (NBS)-Leucine Rich Repeat (LRR) Resistant Gene Analogs (RGAs). *In*: Arabica Coffee (*Coffea Arabica* L. Cv S.288). *J Biotechnol Biomater* **2**:146. doi:10.4172/2155-952X.1000146

Leipe, D. D., Koonin, E. V., Aravind, L. (2004). STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *J Mol Biol*. **343**:1-28.

Lamb, C. and Dixon, R.A. (1997). The oxidative burst in plant disease resistance. Annu Rev Plant Physiol. *Plant Mol Biol* **48**: 251-275.

Leubner-Metzger, G. and Meins, F. J. (1999). Function and regulation of plant β-1, 3-glucanases. *In*: Wu, C and Bradford, J. K., Class I chitinase and β-1, 3-glucanase are differentially regulated by wounding, methyl jasmonate, ethylene, and gibberellin in tomato seeds and leaves. *Plant Physiology* **133**:263-273.

Levine, A., Tenhaken, R., Dixon, R. and Lamb, C.J. (1994). H_2O_2 from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell* **79**: 583-593.

Malkin, E. (2014). A coffee crop withers. Fungus cripples coffee production in Central America. *The New York Times*. Accessed on 13/12/2015. Available online at: http://www.nytimes.com/2014/05/06/business/international/fungus-cripples-coffee-production-across-central-america.html?_r=0

Martin, G. B., Bogdanove, A. J. and Sessa, G. (2003). Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology* **54**(1):23-61.

Mazzafera, P., Shimizu, M. M. and Melo A. G. (2005). Polyphenoloxidase activity in coffee leaves and its role in resistance against the coffee leaf miner and coffee leaf rust. Elsevier. *Phytochemistry* **67**:277-285.

Moody, E. D. (2001). Genome Techniques: An Overview of Methods for the Study of Gene Expression. *Journal of Animal Science* **79**:128-135.

Morris, E. and Walker, J. (2003). Receptor-like protein kinases, thekeys to response. *Curr. Opin. Plant Biol.* **4**:339-342.

Ori, N., Sessa, G., Lotan, T., Himmelhoch, S. and Fluhr, R. (1990). A major stylar matrix polypeptide (sp41) is a member of the pathogenesis-related protein superclass. *EMBO J.* **9**: 3429-3436

Peck, S. C. (2003). Early phosphorylation events in biotic stress. *Curr. Opin. Plant Biol.* **6**:334-338.

Perez-Garcia, A., Vicente, A., Olea, F., Codina, C. J. and Rivera, E. M. (2002). Differential expression of β-1, 3-glucanase in susceptible and resistant melon cultivars in response to infection by *Sphaerotheca fusca*. Elsevier Science Service Ltd, *Physiological and Molecular Plant Pathology* **61**:257-265.

Pfaffl, W.M. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29** (9):2002-2007.

Ribas, F. A., Cenci, A., Combes, M., Etienne, H. and Lashermes, P., (2011). Organization and molecular evolution of a diseases resistance gene cluster in coffee trees. *BMC Genomics* **12**:240.

Romeis, T., Ludwig, A. A., Martin. R., and Jones, J. D. G. (2001). Calcium-dependent protein kinase plays an essential role in the plant defense response. *EMBO J* **20**:5556-5567.

Rutherford, A. M. and Phiri, N. (2006). Pests and diseases of coffee in Eastern Africa: A technical and Advisory manual. CAB International. p.38.

Sanabria, N. M., Huang, J.-C., & Dubery, I. A. (2010). Self/non-self perception in plants in innate immunity and defense. *Self/Nonself*, *1*(1), 40–54. doi:10.4161/self.1.1.10442

Schieber, E., (1972). Economic impact of coffee rust in Latin America. *Annual Review of Phytopathology* **10**: 491-510. DOI:10.1146/annurev.py.10.090172.002423

Sera, H. G., Sera, T., Ito, S. D., Azevedo, A. J., Mata, S. J., Dói, S. D., Filho, R. C. and Kanayama, S. F.(2007). Resistance to leaf rust in coffee carrying SH_3 Gene and others SH Genes. *Brazilian Archives of Biology and Technology* **50**: (5) 753-757.

Silva, C.M., Várzea, V., Guerra-Guimarães, L., Azinheira, G.H., Fernandez, D., Petitot, A., Bertrand, B., Lashermes, P. and Nicole, M. (2006). Coffee resistance to the main disease: Leaf rust coffee berry disease. *Braz. J. Plant Physiol*. **18** (1):119-141.

Stone, D. (2014). Fungus, climate change threatening big part of global coffee supply. *National Geographic*, URL:www.news.nationalgeographic.com/news/2014/05/140531-coffee-rust-columbia-brazil-cost-problems/

Thomma, P.H.J.B., Nurnberger, T. and Joosten, H.A.J.M. (2011). Of PAMPs and effectors: The Blurred PTI-ETI dichotomy. American Society of plant biologists. *The Plant Cell* **23**:4-15.

CHAPTER 1

[i]MOLECULAR RUST RESISTANCE COMPONENTS HAVE DISTINCTIVE EXPRESSION PROFILES IN *Coffea arabica* - *Hemileia vastatrix* INTERACTIONS

ABSTRACT

Countering the economic hurdle caused by coffee leaf rust disease is most appealing at this time as it has posed a major threat to coffee production around the world. Establishing differential expression profiles at different times following pathogen invasion in both innate and acquired immunities unlocks the molecular components of resistance and susceptibility. Suppression subtractive hybridization (SSH) was used to identify genes differentially over-expressed and repressed during incompatible and compatible interactions between *Coffea arabica* and *Hemileia vastatrix*. From 433 clones of expressed sequence tags (ESTs) sequenced, 352 were annotated and categorized of which the proportion of genes expressed during compatible interaction were relatively smaller. The result showed upregulation and downregulation of various genes at 12 and 24 hours after pathogen inoculation in both interactions. The use of four different databases in searching for gene homology resulted in different number of similar sequences. BLASTx against EMBL-EBI (European Molecular Biology Laboratory-European Bioinformatics Institute) database being with the maximum (100%) hits for all the annotated sequences. RT-qPCR analysis of seven resistance-signaling genes showed similar expression patterns for most of the genes in both interactions, indicating these genes are involved in basal (non-specific) defense during which immune reactions are similar. Using SSH, we identified different types of resistance related genes that could be used for further studies towards resistant cultivar development. The potential role of some of the resistance related proteins found were also discussed.

Keywords: Coffee, gene expression, incompatible, plant-pathogen interaction, RT-qPCR

---

[i]Accepted (with correction) for publication in the European Journal of Plant Pathology (EJPP)

COMPONENTES MOLECULARES DA RESISTÊNCIA À FERRUGEM TÊM PERFIS DE EXPRESSÕES DISTINTOS EM INTERAÇÕES ENTRE *Coffea arabica* E *Hemileia vastatrix*

RESUMO

A ferrugem do cafeeiro é considerada a principal doença dessa cultura e promove prejuízos significativos. Dessa forma, reduzir as perdas econômicos causados por essa doença é de extrema importância. O estabelecimento de perfis de expressão diferencial em diferentes tempos após a inoculação de patógenos, tanto em imunidade natural quanto em imunidade adquirida, permite o entendimento dos componentes moleculares de resistência e de susceptibilidade. A hibridização subtrativa de supressão (SSH) foi utilizada para identificar genes diferencialmente expressos (*upregulated* e *downregulated*) em interações compatíveis e incompatíveis entre *C. arabica* e *H. vastatrix*. A partir de 433 clones de sequências *expressed sequence tags* (ESTs), 352 foram anotadas e categorizadas com proporção relativamente menor de genes expressos em interação compatível. Observou-se nas bibliotecas SSH, expressão *upregulation* e *downregulation* para vários genes nos tempos de 12 e 24 horas após a inoculação do patógeno (h.a.i.) em ambas as interações. Foram utilizados quatro bancos de dados para verificar a homologia entre os genes e observou-se resultados diferentes entre os bancos de dados. A maior similaridade (100%) dos hits das sequências anotadas foi observada pelo BLASTx no banco de dados EMBL-EBI (Instituto Europeu de Biologia Molecular-Instituto Europeu de Bioinformática). A análise RT-qPCR de sete genes de sinalização de resistência mostrou padrões de expressão semelhantes para a maioria dos genes em ambas as interações. Isso indica que esses genes estão envolvidos na defesa basal (não específica) durante a qual as reações imunes são semelhantes. Usando o SSH, foi identificado diferentes tipos de genes relacionados à resistência os quais poderiam ser utilizados em estudos futuros para o desenvolvimento de cultivares resistentes. Os papéis potenciais de algumas proteínas relacionadas à resistência encontradas também foram discutidos em detalhes.

Palavras-chaves: Café, expressão gênica, incompatível, interação planta-patógeno, RT-qPCR

1. INTRODUCTION

Coffee leaf rust (*H. vastatrix*) is one of the main disease of coffee worldwide as its races are rapidly emerging and its aggressive isolates are presenting new challenges. Therefore, the current level of epidemiology is being the most severe than any time before in some regions (Avelino et al. 2015; Zambolim 2016). The triggering force behind the surge of new races is not well understood as urediniospores are asexual means of reproduction (Zambolim 2016). However, some reports have started to emerge attributing the hidden sexual spore with in the asexual urediniospore to be among the factors behind the evolution of different races (Carvalho et al. 2011). What is well known in any pathosystem is that the emergence of new races and resistance breakdown is related to co-evolution during plant-pathogen interaction. Plants develop the ability to recognize pathogens and develop elaborate defense mechanisms to avoid pathogen attack. Likewise, the pathogens evolve to avoid such recognitions leading to evolution race between the plant and the pathogen (Burdon and Thrall 2009).

In resistant cultivars, resistance proteins induce an immunity by which the network of defense systems recognize molecular signatures of the pathogen as a sign of invasion leading to cease of pathogen colonization (Saunders et al. 2012). The immune reaction could be manifested ranging from physical stress to death of the infected tissue (Nimchuk et al. 2003). The first line of defense is induced by pattern recognition receptors (PRRs), which recognizes microbe-associated molecular patterns (MAMPs) (Coll et al. 2011). MAMP- triggered immunity confers basal defense, the only defense in compatible interaction unlike in incompatible interaction in which cascades of pathogen specific immune reactions are eventually triggered (Jones and Dangl 2006; Muthamilarasan and Prasad 2013). The recognition of MAMPs followed by reaction of the defense system at the site of infection, induces rapid resistance signal transduction and death of local cells (hypersensitive response, HR) prompting strong protection against the invading pathogens in resistant cultivars (Rojas et al. 2014; Kushalappa et al. 2016). HR is often associated with the direct or indirect recognition of pathogen avirulence (avr) proteins by the corresponding resistance (R) protein in incompatible interaction (Jones and Dangl 2006; Bozkurt et al. 2010). Ultimately, the activation of defense responses in the surrounding tissue and throughout the whole plant results in the development of systemic

acquired resistance (SAR) (Hunt et al. 1996). In incompatible interaction, HR is manifested as an effective defense response in stopping biotrophic pathogen invasion and spread by programmed cell death (Niks and Marcel 2009; Gill et al. 2015). Even though a characteristic HR cell death is generally observed during compatible interaction, HR does not result in effective defense response and lacks pathogen specificity (Gaudet et al. 2007). Once pathogen components are recognized by the cognate receptors, cascades of signal transduction responses are followed which involve changes in calcium level, extracellular alkalization, production of reactive oxygen species (ROS), activation of kinases, transcriptional reprogramming and changes in hormone concentration (Seybold et al. 2014; Wu et al. 2014). During incompatible interaction, apoplastic secretome protein extraction in different organisms (Sheen 1998), differential expression screening studies in coffee-rust interaction (Diola et al. 2013; Guerra-Guimarães et al. 2015) and *A. thaliana* treated with fungus elicitors (Ndimba et al. 2003) indicated early expression of signaling genes.

Kinase associated protein phosphatase (KAPP) 2C family and the LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP-interacting kinase 1 are involved in early signal perception of pathogen effectors (Sheen 1998; Afzal et al. 2008). The exact physiological role of kinase associated protein phosphatase 2C isoenzymes has not yet been fully understood though these kinase domains are featuring in defense systems of different organisms ranging from unicellular yeast (Maeda et al. 1993; Bögre 2003) to complex mammals (Flajolet et al. 2003). In *A. thaliana*, the kinase associated protein phosphatase (KAPP) is an enzyme that dephosphorylates the Ser/Thr receptor-like kinase (RLK) (Umbrasaite et al. 2011). Pathogen infection is also marked by increased level of endogenous auxin/IAA triggering the expression of auxin binding and auxin-responsive genes to escalate the expression of resistance signaling genes through the cascading amplification of phosphorylation mediated by different kinases (Carna et al. 2014). On the other hand, auxin/IAA proteins favor pathogenesis by repressing the response of auxin resistance signaling as seen in *A. thaliana* (Padmanabhan et al. 2008). Similar studies show that the role of auxin-repressed protein in *A. thaliana* is linked to the concentration of auxin/IAA proteins (Ulmasov 1997; Korasick et al. 2014). In its entirety, the function of gibberellic acid (GA) signaling F-box protein or GA-insensitive dwarf2 (gid2), as known in rice (Sasaki et al. 2003), is less understood in establishing signal perception during compatible interaction (Bari and Jones 2009). However, the role of GA in plant

development and resistance signaling has started to emerge as it has a cross- talk with other hormones in inducing basal resistance and susceptibility. As reported in rice, GA has a negative role in basal resistance (Yang et al. 2008).

Histone proteins, on the other hand, are other essential proteins involved in regulating gene expression and transposon silencing in plants and animals (Law and Jacobsen 2010). They are molecular components of nucleosomes where methylation and acetylation plays regulatory role by determining the accessibility of chromatin to regulatory proteins (Ding and Wang 2015). In this context, therefore, the abundance of histone proteins (Ac-like transposase) is associated to the state of the plant immunity that signals the need for the expression of required genes against the pathogen effectors or limit the amplification of defense signals depending on the required response (Junqi Song and Bent 2014).

Coffee rust resistance is a condition when the defense system recognizes any of the 50 or so races of *H. vastatrix* nine virulence factors (v1-9) in the presence of the corresponding nine major R genes (S_H1-9) (Rodrigues et al. 1975; Gichuru et al. 2012; Alwora and Gichuru 2014; Zambolim 2016). Crucial to the progresses achieved so far with regards to resistance coffee breeding is the discovery of Híbrido de Timor (HDT), a rust resistant natural interspecific hybrid between *C. arabica* and *C. canephora* (Bettencourt 1973). As HDT is resistant to coffee leaf rust and other major diseases (Pereira et al. 2005), it has been used as an important source of R genes in rust resistance coffee breeding programs (Bettencourt 1973; Rodrigues et al. 1975). Among the derivatives of HT, CIFC 832/1 and CIFC 832/2 are of considerable importance in crossing with various rust susceptible coffee cultivars around the world as they are resistant to all races of *H. vastatrix* (Rodrigues et al. 1975; Diniz et al. 2012). Despite its susceptibility to coffee leaf rust (Guzzo et al. 2009; www.ico.org/leafrust), Catuaí Vermelho IAC 44 is another important cultivar due to its high vigor, cup quality, wide adaptation capacity and high productivity (http://www.consorciopesquisacafe.com.br/). The transfer of rust and other disease R genes from HDT derived genotypes to susceptible but with high economic value coffee cultivars like Catuaí Vermelho IAC 44 has an indispensable input in mitigating coffee rust damages around the world. Indeed, the ultimate objective of breeding programs in this regard is to develop resistant varieties without compromising other agronomic qualities. To that end, expression profiling and molecular characterization of R genes

could help open another level of understanding of phytopathosystem and in turn leads to durable rust resistance development against *H. vastatrix*.

Differential expression of resistance and resistance-signaling genes has been reported in compatible and incompatible interactions between coffee cultivars and *H. vastatrix* races (Nimchuk et al. 2003; Fernandez et al. 2004; Glazebrook 2005; Ganesh et al. 2006; Guzzo et al. 2009; Diniz et al. 2012; Diola et al. 2013). However, identification of resistance related genes differentially over expressed and repressed at a given time altogether during incompatible and compatible interaction is less exploited, hence of great importance to execute. Such efforts are essential inputs to have an insight into the understanding of how R genes function. It also paves a way to identify candidate R genes in developing resistant cultivars. Therefore, the objective of the present work was to identify coffee genes differentially upregulated and downregulated at 12 and 24 h.a.i. and quantify some resistance-signaling genes at 0 (control), 12, 24, 48 and 72 hours after inoculation (h.a.i.) during compatible and incompatible interactions between *C. arabica* and *H. vastatrix*.

2. MATERIALS AND METHODS

2.1 Plant materials and pathogen inoculation

Coffee rust resistant cultivar HDT (CIFC-832/2) and susceptible Catuaí Vermelho IAC 44 were used in all the experiments. One-year-old greenhouse grown young Catuaí Vermelho IAC 44 seedlings and clone derived HDT (CIFC-832/2) plants were used for pathogen inoculation. Race II of *H. vastatrix* fungus urediniospore was rubbed off against the intact abaxial leaf surface to induce immune challenge. Pathogen inoculated plants were immediately transferred to moist dark chamber at 22°C (±2) with 85% relative humidity. For expression studies of selected genes, inoculated plants were kept in dark for 48hours during which 12 and 24 h.a.i. samples were collected and then taken to light condition until 72 h.a.i. samples were collected. Initially, eight expression studies were carried out on two interactions in such a way that forward and reverse expression studies were setup at 12 and 24 h.a.i. In the same way, five expression studies (0, 12, 24, 48 and 72 h.a.i.) were set up for RT-qPCR validation of seven resistance-signaling genes (HT12F50, HT12F100, HT12R109, HT24F85, HT24F123, HT24F133, HT24R75)

during incompatible and compatible interactions. Uninoculated plants (controls) were randomly picked and used along with the treated plants.

2.2 RNA extraction and cDNA synthesis

Collected leaves were immediately stored at -80°C until total RNA was extracted. Total RNA was extracted using Concert™ Plant RNA Reagent (Invitrogen) following the manufacturer's recommendations. Nanodrop (NanoDrop Technologies, Wilmingon, DE, USA) scanning at 260/280 nm and 1.5% UltraPure™ agarose (Invitrogen) gel electrophoresis stained with ethidium bromide (0.5 µg/ml) was routinely run to check the quantity and integrity of RNA before any downstream use. Subsequently, total mRNA was separated using Dynabeads® mRNA Purification Kit (Dynal Biotech-Life Technologies) and subjected to cDNA synthesis by SMART-PCR (polymerase chain reaction) cDNA Synthesis Kit (Clontech).

2.3 Suppression subtractive hybridization (SSH) and cloning of ESTs

cDNA subtraction and enrichment for selective amplification of differentially expressed genes during compatible and incompatible interactions was done using *Select cDNA Subtraction Kit* (Clontech). To separate and investigate upregulated and downregulated genes due to pathogen infection, cDNAs from the two contrasting expressions were labeled as tester and then as driver subsequently. First, cDNAs from incompatible interaction were labeled as tester and subtracted from mock inoculated (control) samples, which were labelled as driver, resulting in resistance related differentially expressed ESTs. This procedure was repeated for both interactions and at each point of time (12 and 24 h) after pathogen inoculation by shuffling the labelling of tester and driver otherwise (Diatchenko et al. 1996). Amplified ESTs were inserted into pGEM-T easy vector (Promega) and then incorporated into *E. coli* DH5α by heat shock transformation. Plate LB (Luria-Bertani) medium with ampicillin (200 mg/ml), X-GAL (20 mg/ml) and 2% (w/v) IPTG (Isopropyl-β-D-thiogalactoside) was used to select transformed white colonies. Selected white colonies were picked using toothpick, transferred to 3ml liquid LB medium with ampicillin, and shaken at a speed of 180 rpm and temperature of 37°C for 12-16 h. Plasmid DNA was extracted using centrifugation protocol of Wizard® SV Plus Minipreps DNA Purification System (Promega). The quality and quantity of

extracted plasmid DNA was measured by 1% gel electrophoresis and Nanodrop. Insertion of the DNA segment (clone) of the eight libraries was detected by PCR using SP6 and T7 primers of pGEM-T easy vector. PCR reaction was for 35 cycles in 94ºC for 30 s, 45 ºC for 1 min and 72 ºC for 2 min in a total reaction volume of 25 µl using 1x PCR buffer, 200 µM each dNTPs, 0.4 µM of each primer, 1.4 mM $MgCl_2$, 1 unit Taq polymerase (Invitrogen), and 200 ng plasmid DNA.

2.4 Sequencing of ESTs

433 clones were sequenced using 16-capillary 3130xl Genetic Analyzer (Applied Biosystems, Foster City, USA) fluorescence-based capillary electrophoresis system. Sp6 and T_7 primers were used in sequencing PCR reactions. PCR reactions were based on BigDye® Terminator v3.1 cycle sequencing kit at a reaction condition of 96ºC for one min followed by 15 cycles of 96ºC and 50ºC for 15 s and extension reaction of 60ºC for 4 min in a reaction volume of 20 µl. Sequence quality of > 20 QC was considered for downstream processing. Sequences were trimmed-off low quality, adapter, vector and primer sequences using Vecscreen server of NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/tools/vecscreen/).

2.5 Annotation and homology search

Eight expression libraries were categorized into two groups as ESTs of incompatible and compatible interactions. The clones subtracted were subjected to BLASTx (Nucleotide Basic Local Alignment Search Tool) against ESTs available in NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi), NCBI BLAST2GOx (BLAST for Gene Ontology), EMBL BLASTx, Brazilian Coffee Genome Project BLASTx (LGE, http://bioinfo03.ibi.unicamp.br/coffea/) and *C. canephora* EST data repositories. To avoid the potential contamination by fungal genes, BLASTx was run against ESTs of three related species of fungi; *Melampsora- laricis-populina* (54445 ESTs), *Puccinia graminis* (269 ESTs) and *Hemileia vastatrix* (726 ESTs). All the three ESTs were accessed at http://www.ncbi.nlm.nih.gov/nucest (accessed on 11 August 2016) by searching for each species separately. Genes with homology to any of these three species were excluded from the libraries. Subsequently, redundant genes (7.81%) were eliminated to avoid unnecessary duplications of identical genes with in a library.

Associated GOs (gene ontologies) were pooled together from all the accessed databases and grouped based on their function and cellular localization. A cutoff e-value of 10^{-5} or less was considered significantly similar in annotating ESTs. Upregulated and downregulated genes were analyzed in both categories across all the databases. Putative and hypothetical functions of ESTs were searched in all the databases to map differentially expressed genes in all the libraries during the two contrasting interactions.

2.6 Subcellular localization of ESTs

Protein subcellular localization of all 352 ESTs was done using TargetP 1.1 online localization prediction tool (http://www.cbs.dtu.dk/services/TargetP/) according to Emanuelsson et al. (2000). Significance cut-off for four different categories of subcellular localities were set to standard for plant network as 0.73 for cTP (chloroplast transit peptide), 0.86 for mTP (mitochondrial targeting peptide), 0.43 for SP (signal peptide, involved in secretary pathway) and 0.84 for others (other subcellular compartment).

2.7 Quantitative analysis by RT-qPCR

Seven genes involved in resistance-signaling against *H. vastatrix* were selected from libraries constructed at 12 and 24 h.a.i. for quantitative validation (Table 1). Three of these genes were downregulated at either 12 or 24 h.a.i. while the others were upregulated at either of these time points during incompatible interaction as seen during identification of differentially expressed genes. Three reference genes (S24, UBQ10 and GAPDH), whose expressions were found to be stable (Cruz et al. 2009) were used. Primer design was done using NCBI primer designing tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) with preferential selection of primer pairs having minimum self-complementarity. The primers were *Sigma-Aldrich* made having 22-23 bp size (Table 1).

Table 1. Primer sequences of genes validated by RT-qPCR[a].

Primer name (code)	Putative sequence identification	Annotation (LGE: Brazilian Coffee Genome Project, http://bioinfo03.ibi.unicamp.br/coffea/ and EMBL-EBI: http://www.ebi.ac.uk/Tools/sss/ncbiblast/nucleotide.html)	Sequence	Tm (ºC)	Amplicon size (bp)
S24[b]	GT021438	CarCTFrHh24FL2_17-C09-M13F similar to 40S ribosomal protein s24-1-*Arabidopsis thaliana*	F: 5'- GCCCAAATATCGGCTTATCA-3' R: 5'- TCTTCTTGGCCCTGTTCTTC-3'	63.5 63.3	92
UBQ10[b]	KF56925	*Phaseolus vulgaris* clone BE1176 polyubiquitin mRNA, complete cDNA	F: 5´- CAGACCAGCAGAGGCTGATT-3´ R: 5´- AGAACCAAGTGAAGGGTGGA-3´	64.6 63.5	100
GAPDH[b]	FN431983	*Tectona grandis* partial g3pdhcy2 gene for glyceraldehyde-3-phosphate dehydrogenase	F: 5'-AGGCTGTTGGGAAAGTTCTTC-3' R: 5'- ACTGTTGGAACTCGGAATGC-3'	63.4 64.0	70
HT12F50	EM_EST:GT003697	TransId-204764 CACATN1 *Coffea arabica* cDNA clone CACATN1-1B26TVB similar to Auxin-binding protein ABP20 precursor - *Prunus persica* (Peach), mRNA sequence	F: 5'- ACGCGGGGCTCCACCACTTAATC-3' R: 5'- CACAGCGAAATCCAGACGGGCAA-3'	67.0 65.8	341
HT12F100	gb|ABR18801	Kinase-associated protein phosphatase [*Solanum peruvianum*]	F: 5'-TGCTCCGCCATGTCCATCACACA-3' R: 5'-GTCGCGGCCGAGGTGCAAAGAAG-3'	66.7 68.6	116
HT12R109	KF588660	*Actinidia deliciosa* GA signaling F-Box (SLY1_12) gene, complete cds	F: 5'- CGGCGGCTTTACTCGCTCTACCT-3' R: 5'- TGCCGGAAGACTTGGCCGGAAC-3'	66.5 67.5	130
HT24F85	EM_EST:GR990444	TransId-100452 CACAT45FR *Coffea arabica* cDNA clone CACAT45FR_32_B02_018.F similar to auxin-responsive protein / indoleacetic acid-induced protein 9 (IAA9), identical to	F: 5'-GGAAAAGGAACCGACAGCACGCA-3' R: 5'-TGCACAAGGAAAGGGAACAGGGG-3'	66.0 65.0	238

		SP:Q38827 Auxin-responsive protein IAA9 (Indoleacetic acid-induced protein 9)			
HT24F123	ref|XM_004232861	PREDICTED: *Solanum lycopersicum* auxin-repressed 12.5 kDa protein-like, transcript variant 2 (LOC101258429), mRNA	F: 5'-CTCACCCTGAACGTGGCCTGGGA-3' R: 5'- GGTTGCAGGACTGGGTGGCATAG-3'	68.3 65.4	136
HT24F133	ref|XM_006338244	PREDICTED: LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP-INTERACTING KINASE 1-like (LOC102591832), mRNA	F: 5'-CGGGCAGGTACACAAGAGAGCCA-3' R: 5'-AGCAGCGAACCAGAAAGGGGCAA-3'	66.4 67.2	154
HT24R75	gb|AAD12209	Act-like transposase [*Arabidopsis thaliana*],similar to Histone H1 - Plantago major (Common plantain), mRNA sequence	F: 5'-ATGCGGAGAAGAAGAGGGCTGGT-3' R: 5'- TTGGCTGCTTTGCCTTCGCTGGA-3'	65.7 67.7	174

[a]The efficiency and R^2 of all primers were 0.81-0.99 and 0.915-0.996, respectively.
[b]Reference genes (Cruz et al., 2009).

Two step RT-qPCR was carried out following the MIQE minimum standard guidelines for fluorescence based quantitative real-time PCR experiments (Bustin et al. 2009). cDNA was synthesized following Im-Prom-II^{TM} Reverse Transcription System cDNA synthesis Kit (Promega) using 1µg total RNA. Amplification of target fragments were optimized by testing various annealing temperatures around Tm ranging from 54 °C to 66 °C using Applied Biosystems (Foster City, California 94404, USA) thermocycler temperature gradient program. Amplification of expected targets were verified by 1% agarose gel electrophoresis. Primer concentration of 1µM was selected for both reference and target genes in RT-qPCR quantification. Standard curve was developed using pooled cDNA from both genotypes serially diluted by a factor of 5 at five dilution points (1:1, 1:5, 1:25, 1:125 and 1:625) starting from 600 ng/µl. Real-time PCR reaction volume was 10 µl containing 2 µl H_2O, 1 µl (1 µM) primer, 5 µl (50% v/v) SYBR green master mix and 1 µl (120 ng/µl) cDNA. Reaction parameters were: 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 60 °C for 1 min. Finally melting curve stage was set to default conditions of Applied Biosystems 7500 Real Time PCR System (Foster City, California 94404, USA). Primer specificity was confirmed by examining melting curve of reference genes in control samples and target genes at the four h.a.i. during efficiency test and final plate run, respectively. Amplification efficiency of ≥81% was used for both reference and target genes. Three biological and three technical replicates run for all genes quantified. For both interactions, three different plants were randomly selected and inoculated by freshly collected urediniospores for the four time points under investigation. No template controls were included for all target genes in all run plates. All the parameters of standard curve were used in relative quantification reactions. Technical errors within each biological replicate was tested using the quantification cycle (C_q) values by using mean test while cDNA collected at 0 hours from susceptible and resistant samples were used for comparison against the remaining four time points (Nicot et al. 2005). For all reference and target genes, technical replicate C_q values were averaged. Using the amplification efficiency (E) and threshold cycle of reference and target genes, relative expression was calculated in control and unknown sample in comparison to reference genes according to Pfaffl (2001) method.

2.8 Statistical analysis

For the three biological replicates, three technical replicate Cq values were averaged and normalized to reference gene Cq values by qBase relative quantification tool (Hellemans et al. 2007). Before any downstream analysis, primer dimer correction was performed. Cq values of target genes from control samples (0 h.a.i.) in both interactions were normalized to reference genes and subjected to one-way analysis of variance (ANOVA). Within interaction normalized mean comparison was made using Dunnett test while Tukey's multiple mean comparison was used do compare corresponding normalized means between interactions using GraphPad Prism version 7.00 for windows (GraphPad software, La Jolla California, www.graphpad.com). Differential expression was shown as relative expression of a gene at a given time after pathogen inoculation in comparison to control sample as upregulated, downregulated or not changed significantly.

3. RESULTS

3.1 Isolation of differentially expressed genes

Four expression libraries (HT12F, HT12R, HT24F and HT24R) during coffee rust resistant cultivar HDT CIFC832/2 and another four libraries (Ca12F, Ca12R, Ca24F and Ca24R) during susceptible cultivar Catuaí Vermelo IAC 44 interaction with *H. vastatrix* were constructed. These acronyms were used to represent the genotypes used; how long (in hours) the plant was treated with the pathogen (*H. vastatrix* race II) before the samples were collected followed by F (upregulated genes) or R (downregulated genes). The cDNAs of differentially expressed genes from both genotypes were separated at 12 and 24 h.a.i. The number of upregulated genes in each library from incompatible interaction were comparable while the number of downregulated genes at 24 h.a.i. were much lowered in susceptible cultivar (Figure 1). Nevertheless, for the other three libraries of compatible interaction, it followed the same trend as in incompatible interaction.

3.2 Cloning and sequencing of ESTs

The number of isolated genes in each library was substantially greater than what they were after subsequent downstream processing. After repeated multiplication of white colonies on selective LB media, 433 ESTs (Figure 1) were identified and sequenced.

Redundant EST sequences were found in all subtraction libraries and excluded during annotation. After screening and eliminating redundant sequences, the number of non-redundant genes were 80% in upregulated libraries and 69% in downregulated libraries. The fragments sequenced were ranging from 77 bp to 1190 bp. However, only insert fragments with 154 bp or more were considered for homology search and further analyses.

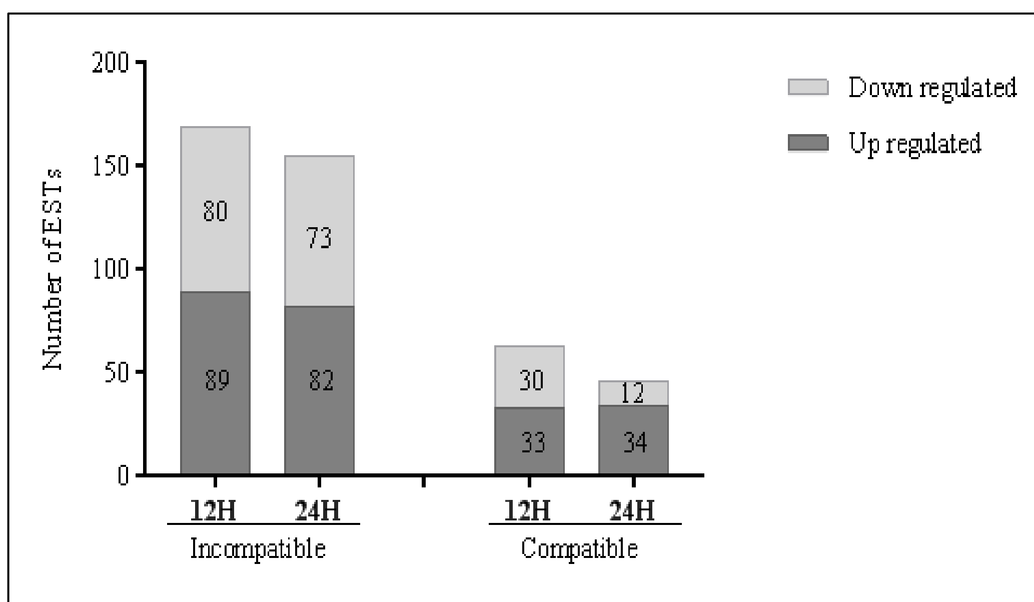


Figure 1. Number of genes upregulated and downregulated during compatible and incompatible interaction at 12 and 24 h.a.i. before redundant ESTs were excluded.

3.3 Annotation and metabolic categorization

Non-redundant 352 ESTs differentially expressed due to pathogen inoculation in resistant and susceptible genotypes were annotated and categorized based on their metabolic roles (Appendix I Table I). BLASTx search of these clones resulted in different number of matches in different databases (Appendix I Figure I). Matches were found for all the sequenced gene fragments in EMBL database though some hits (18.24%) were either not significant (e-value $\geq 10^{-5}$) or matched with unrelated species and hence excluded. Contrarily, only 32.63% of the ESTs were with significant similarities using BLASTx against NCBI database. Yet, using BLAST2GO as homology search tool, the number of hits with significant similarities were different from simple BLASTx search (57.75% larger than the output by BLASTx). As these two databases were too robust, homology search was carried out in two more databases (Brazilian Coffee Genome Project, LGE

EST and *C. canephora* EST) exclusively devoted to coffee and related genes. In this latter search, significant matches were found for 57.63% and 68% of the ESTs using BLASTx against LGE EST and *C. canephora* EST, respectively. Out of 352 ESTs with significant hits in any of the four databases searched (100% in EMBL, data not shown), 140 (39.55%) ESTs were shared between LGE, NCBI and *C. canephora* EST databases (Appendix I Figure I). We are interested to identify genes exclusively upregulated during each interaction after redundant ESTs were excluded. We found almost all ESTs expressed at 12 h.a.i. and at 24 h.a.i. were interaction specific (Fig. 2).

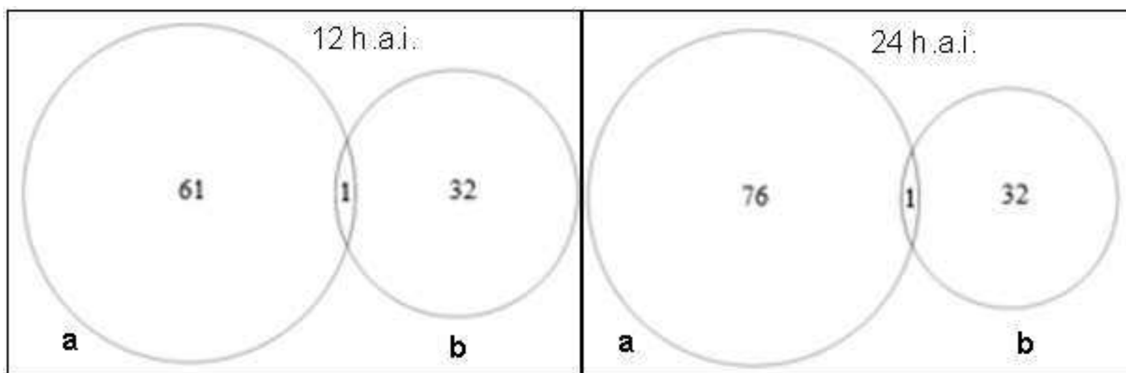


Figure 2. Genes specifically and/or commonly expressed at different hours after *H. vastatrix* inoculation in both interactions (after redundant ESTs were removed). The number of differentially upregulated genes at 12 h.a.i. during incompatible (a) and compatible interaction (b) and 24 h.a.i. during incompatible (a) and compatible interaction (b). Overlapping regions show number of genes expressed in both libraries and h.a.i.: hours after pathogen inoculation.

Furthermore, to annotate and map GO terms associated with all ESTs (352) with significant similarities, a category of four broad biological function was considered for expression profiling. This classification was as follows; resistance and antimicrobial functions (A), resistance signal induction and transduction (B), cell maintenance and homeostasis (C) and no gene ontologies associated (D) (Figure 3). In both interactions, no transcripts with resistance and antimicrobial functions were found in SSH reverse libraries unlike gene transcripts involved in resistance-signaling, which were found in all but Ca24R library. Homology search in four databases resulted in no associated gene ontology (GO) terms for 33.21% of ESTs while the majority of the annotated ESTs (49.83%) were found to have cell maintenance and homeostasis role.

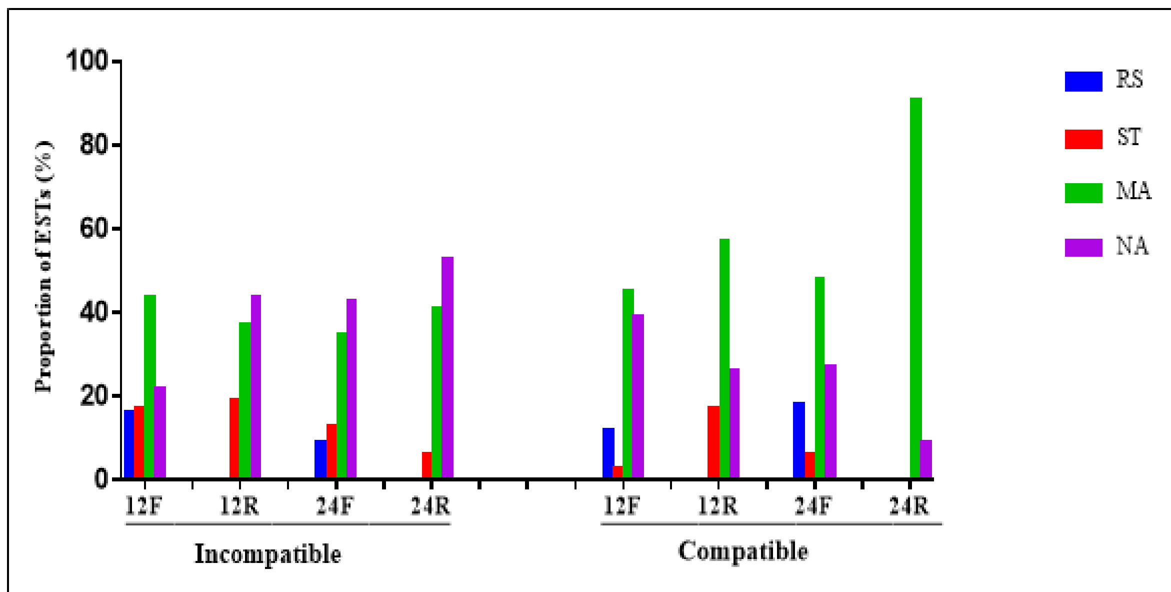


Figure 3. Proportion of different functional categories in each library as described in any of the four databases with ≤10^{-5} E-value. 12F: upregulated genes at 12 h.a.i., 24F: upregulated genes at 24 h.a.i, 12R: downregulated genes at 12 h.a.i., 24R: downregulated genes at 24 h.a.i., RS: Resistance and antimicrobial function, ST: Signal induction and transduction, MA: Maintenance and homeostasis, NA: No GO associated.

In a separate data mining strategy; 619 LRR ESTs (http://www.ncbi.nlm.nih.gov/nucest/?term=LRR), 587 LRR GSSs (leucine rich repeat genome survey sequences, http://www.ncbi.nlm.nih.gov/nucgss/?term=LRR), 105152 LRR proteins (http://www.ncbi.nlm.nih.gov/protein/?term=LRR), 231 NBS-LRR ESTs (http://www.ncbi.nlm.nih.gov/protein/?term=NBS-LRR), 222 NBS-LRR GSSs (http://www.ncbi.nlm.nih.gov/nucgss/?term=NBS-LRR) and 7011 NBS-LRR proteins of plants (all accessed on 10 August 2016) downloaded and BLASTed against the 352 sequenced ESTs. There was a significant (E- value 10^{-5}) BLAST hit for HT24F120 and HT24F133 against NBS-LRR EST and LRR protein, respectively. Similarly, cell component associated GO term search resulted in 42.6% of annotated ESTs with no functional site and localization in any of the databases mined (Figure 4).

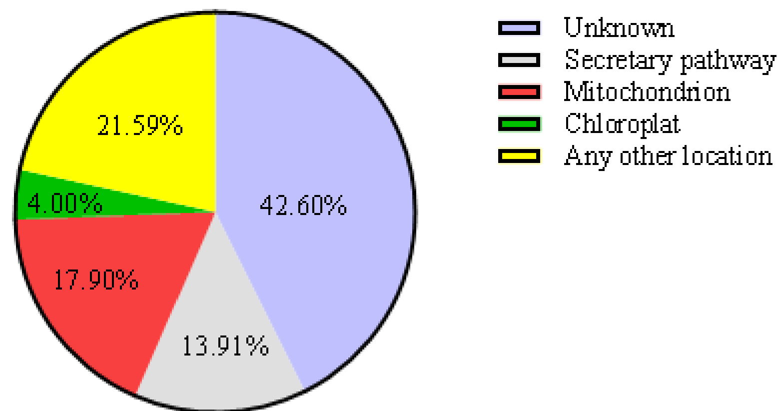


Figure 4. Localization of differentially expressed genes as found in any of the databases mined. Subcellular localization was based on the presence of N-terminal sequences for all the categories. Localization was done by using TargetP 1.1 server (http://www.cbs.dtu.dk/services/TargetP/).

3.4 Upregulated and downregulated genes

The number of differentially expressed genes during incompatible interaction was greater than genes expressed during compatible interaction. Numerically, it was more than two times larger than what was expressed otherwise. Likewise, differentially downregulated genes were by far more abundant during incompatible interaction at both time points following pathogen inoculation (Figure 1). Identification and annotation of all 352 ESTs showed most of these genes have cell maintenance and homeostasis role (Figure 3). In both interactions, homology search resulted in higher number of significant matches for upregulated ESTs in all of the databases accessed (Appendix I Figure I).

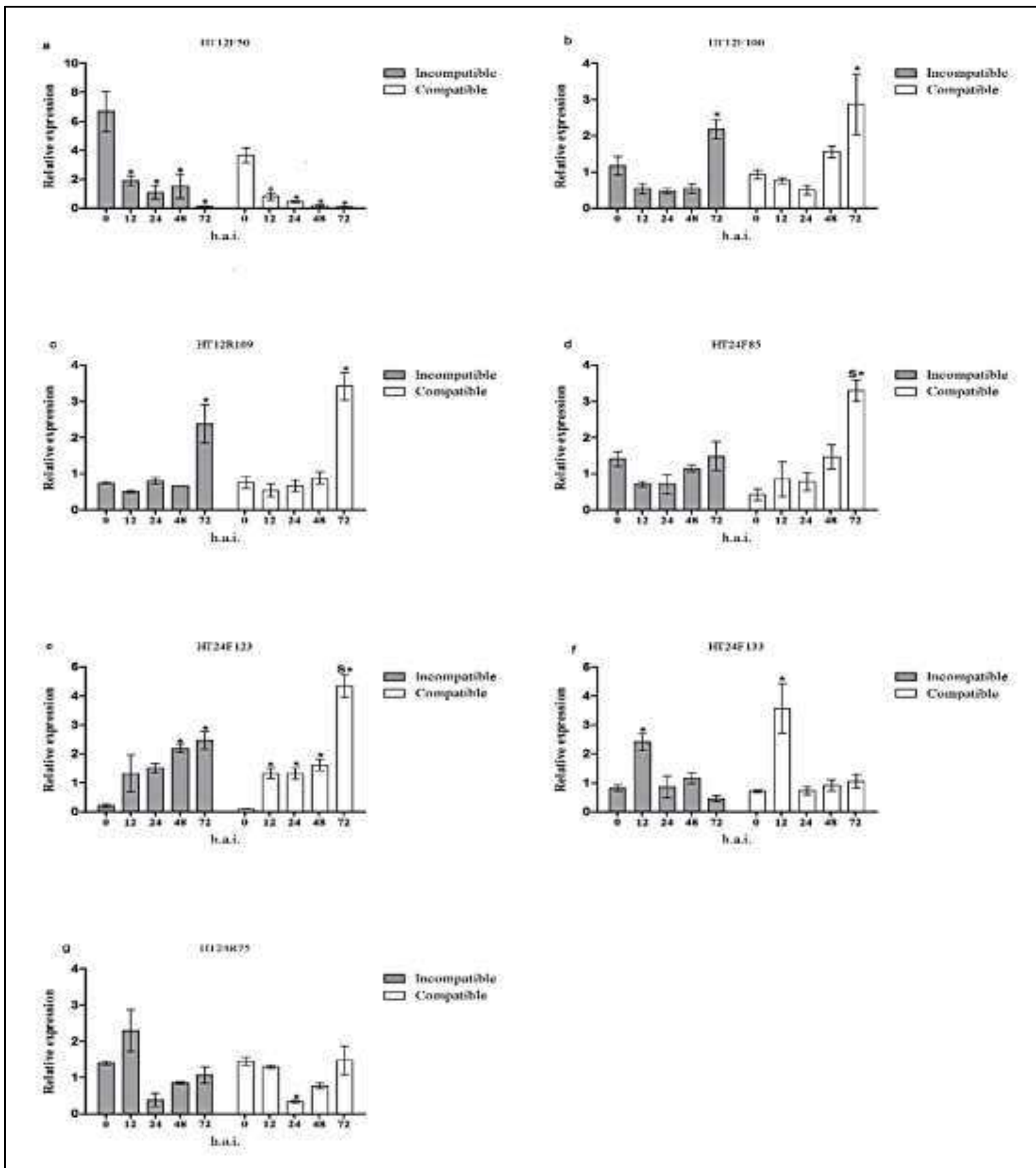
3.5 Subcellular localization prediction

Subcellular localization prediction of EST with significant hit ($\leq 1 \times 10^{-5}$ e-value cut-off) resulted in 42.60% unknown location using Targetp 1.1 plant network cutoffs. Proteins predicted as 'any other location' were the second largest proportion (21.59%) of proteins as shown by TargetP 1.1 subcellular prediction (Emanuelsson et al. 2000).

3.6 RT-qPCR quantification

The seven target genes selected from subtractive hybridization libraries of incompatible interaction showed different expression levels along the pathogen treatment times, but the patterns, in general, were similar during both interactions (Figure 5a-g). HT12F50

showed consistent downregulation along the time course, 72 h.a.i. being its lowest point in both interactions. The levels of HT12F100 and HT12R109 transcript abundance was not significantly changed until 72 h.a.i., at which an elevated expression level was observed for both genes in the two interactions. Looking into inter-interaction comparison, the only significant difference in expression level was seen at 72 h.a.i. for HT24F85 and HT24F123 during compatible interaction. The level of HT24F123 expression was increasing with time after inoculation, while HT24F133 expression peaked at 12 h.a.i, when appressorium is supposed to form, but remaining constant in other times during both interactions.



*Significantly upregulated or downregulated relative to uninoculated samples with in interaction ($p<0.05$).
s Significant difference in expression level between the same h.a.i. across interactions.

Figure 5. RT-qPCR quantification of seven resistance-signaling genes (a-g) at 12, 24, 48 and 72 h.a.i. in resistant (HDT**,** CIFC-832/1) and susceptible (Catuaí Vermelho IAC 44) genotypes inoculated with *H. vastatrix* race II urediniospore. Quantities of transcripts were shown in relative expression compared with control (0-hour) samples after Cq normalization against reference genes (S24, UBQ10 and GAPDH) by qBase relative quantification and Dunnett and Tukey's mean comparisons of GraphPad Prism version 7.00. Results were normalized means ± SEM of three replicates taken from three independent biological replicates.

4. DISCUSSION

Identification of differentially expressed genes in *C. arabica* due to pathogen inoculation showed different expression profiles in terms of the number and types of genes identified in both interactions similar to other work by Guerra-Guimarães et al. (2015). This specificity was observed at different hours for upregulated genes due to *H. vastatrix* inoculation in samples from the same cultivar and yet the specificity was much more when comparing samples at different hours after pathogen inoculation in different genotypes with contrasting interactions. Similar results were reported using different genotypes for both interactions at different hours after pathogen inoculation (Fernandez et al. 2004; Guzzo et al. 2009).

Genes encoding enzymes that degrade fungal cell wall components were upregulated in both interactions. These pathogen related proteins (PR) include class III chitinase and acidic endochitinase, which are considered to be the front lines in defending against fungal pathogens, were also reported in other coffee cultivars and plant species (Guzzo et al. 2009; Legay et al. 2011; Martínez et al. 2012; Dolatabadi et al. 2014). Anti-fungal activity of chitinase enzymes against *H. vastatrix* and other species was reported in different plants (Jach et al. 1995; Martínez et al. 2012; Dolatabadi et al. 2014). Unlike in other rust resistant coffee cultivars treated with the same race of *H. vastatrix* (Fernandez et al. 2004; Guzzo et al. 2009) and resistance inducer ASM (Guzzo et al. 2009), the transcripts of β-1,3-glucanase gene was not found in both interactions. In tobacco, it has been reported that, chitinase and β-1,3-glucanase synergistically provide the maximum defense against fungal pathogen as chitinase is less effective in degrading the harder structure of chitin alone (Jach et al. 1995). The differential expression of other PRs like protease inhibitors and different types of antimicrobial genes in both interactions indicate that these genes are part of basal immunity which is characteristic of most plants (Jones and Dangl 2006; Guerra-Guimarães et al. 2015). Protease enzyme transcripts were found in upregulated libraries in both interactions at both points of time after inoculation, mainly to neutralize foreign proteins from the pathogen as reported in grapevine (Legay et al. 2011). Chalcone synthase and polyubiquitin were the other genes upregulated during incompatible interaction with defense or cell maintenance functions (Fernandez et al. 2004). No ABC types of resistance protein transporter genes were induced in either

interactions despite these proteins were reported in different coffee cultivars (Guzzo et al. 2009) and *Vitis quinquangularis* against *Erysiphe necator* (Gao et al. 2012).

As an anti-fungal agent owing to their oxidative action of phenolic compounds leading to cell wall lignification during HR, the activity of peroxidase enzyme has been reported to be elevated after 20 hours of *H. vastatrix* inoculation in coffee during incompatible interaction (Silva et al. 2008). Peroxidase genes were activated at 12 and 24 h.a.i during incompatible interaction, yet the pick of their biological activity are to be refined. Catalase was another anti-fungal agent whose expression and activity was detected in this work and others (Koç and Üstün 2012; Helepciuc et al. 2014). However, it was detected during compatible interaction only as one of the upregulated genes. On the other hand, NB-LRR are considered to be the main classes of major resistance proteins encoded by R genes in different coffee cultivars against bacteria (Kumar 2012) and *H. vastatrix* (Kobe and Deisenhofer 1994; Guzzo et al. 2009; Ribas et al. 2011) with direct and indirect resistance functions. Though no genes with the association of NB-LRR identity tags in their names were found in our work, like in other report (Fernandez et al. 2004), these diversified R genes with NB-LRR feature are believed to characterize many of the genes in upregulated libraries. Evidently, differential expression of LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP-interacting kinase 1-like (HT24F133), as validated by RT-qPCR quantification (discussed below) is an intuitive assertion for induction of signaling in response to NBS-LRR R gene(s) among the upregulated genes in this particular expression library (24 h. a.i., incompatible interaction). At least one EST (HT24F120) has shown to have similarity (96%) with NBS-LRR like EST motif (gi = 24977765), though this EST is not yet fully characterized and annotated to fully describe its exact role during such interaction (Vidal et al. 2010). Our result could also be explained by the fact that NBS-LRR genes are among the R genes whose expression levels are tissue specific (Carazzolle et al. 2011).

The expression of genes involved in signal perception and transduction as a vital component of SAR was evident as they were expressed during both interactions. The proportions of these genes were comparable to genes with direct anti-fungal role in upregulated libraries of incompatible interaction (6.53% to 8.57%) while there were much less number of such genes in upregulated libraries of compatible interaction altogether (22 to 3 ESTs). The major functions of resistance-signaling genes is the linkage of effector

recognition and defense responses through signal transduction involving secondary messengers (Nürnberger and Scheel 2000; Petre et al. 2014). Defense response mediated by ethylene and jasmonic acid are often considered to be the effective resistance response in developing basal or SAR immunity against necrotrophic pathogens like what salicylic acid does against biotrophs (Glazebrook 2005; Seifi et al. 2014). Ethylene and jasmonic associated signaling gene transcripts were found in both upregulated libraries during incompatible interaction, ascertaining their involvement in defending biotrophic pathogens as well. The result also indicates that their defense signaling may not necessarily be independent of one another; but involved in different signaling pathways as synergistic and antagonistic regulatory interactions (Thaler et al. 2012; Mur et al. 2013; Zhu and Lee 2015). Auxin responsive genes were expressed at 12 and 24 h.a.i. during incompatible interaction in which case the repressive effect of auxin expression was limited to 24 h.a.i. which could route to activation of other signaling pathways as seen in *A. thaliana* (Kovtun et al. 2000). Different types of kinases and GTP-binding proteins known to characterize upregulated libraries in both interactions as seen in different cultivars and other plants (Guzzo et al. 2009; Medeiros et al. 2009; Gao et al. 2012). A large number of genes involved in signal transduction were downregulated at 12 and 24 h.a.i. during incompatible interaction and at 12 h.a.i. during compatible interaction. This selective repression of some signaling genes indicates that their expression is less important (or effective at low level) when compared to other signaling genes that are favored for up-regulation to counter the advancing pathogen in both basal and SAR defenses (Kovtun et al. 2000; Denancé et al. 2013).

Exclusive consideration of genes with annotations shows that the majority of the genes downregulated and upregulated were with cell maintenance and homeostasis functions in both interactions (Figure 3). These few but very important gene products may control the activity of several cell maintenance metabolisms during biotic and abiotic stresses (Chauhan et al. 2013). The number of both upregulated and downregulated genes involved in photosynthesis pathways are merely comparable (seven up- and eight downregulated) during incompatible interaction at the two time points studied. All of these genes are associated with starch biosynthesis metabolisms (Saithong et al. 2013). This result is similar to the work of Bilgin et al. (2010) in which pigment and light-reaction genes were downregulated while genes involved in redox reactions were upregulated following biotic stress. Moreover, most of induced genes were not R genes

in both interactions. This could be explained from unique defense pathways pertaining to specific species/cultivar in responding to the pathogen invasion by few but effective R genes.

Most of the resistance related genes with putative function in all of the databases searched are localized in cytosol where chloroplast and its genes may take part in light-requiring signaling for HR development (Guttman et al. 2002; Jelenska et al. 2007). Moreover; catalase and peroxidase enzymes are localized in apoplast, cell wall and cytoplasm which make their anti-fungal role active at every site of attack (Silva et al. 2008). Different studies show that most signaling proteins such as kinases and resistance proteins with direct anti-microbial function have apoplastic, cytoplasmic and transmembrane localizations in most plants (Piedras et al. 2000; Zipfel 2009; Schneider and Collmer 2010). Class III chitinase was the most abundant anti-fungal gene product whose localization is in apoplast to directly block the establishment of fungal hyphae and induce fungal elicitors for additional expression of different types of chitinases (Neuhaus et al. 1996; Stotz et al. 2014; Jashni et al. 2015). Plant specific subcellular localization prediction indicated that most of the genes whose expression levels were influenced by pathogen inoculation were mitochondrial (17.90%). The result showed the expression of gene products linked to redox pathways are highly affected due to the metabolic cost of defense response (Nie et al. 2015). Gene products localized in mitochondrion and intimately linked chloroplast, including the different types of peroxidase species, are involved in HR resulting in the apoptosis due to ROS. Their temporal expression and regulation are globally linked to the nucleus by mitochondrial proteome for organelle communication (retrograde signaling) (Schwarzlander et al. 2012). The fact that chloroplasts contain light-dependent reaction centers, the overall response of an infected plant is highly influenced by the chloroplast proteins in one or the other way. This is clearly evident as chloroplasts are the sources of stress induced hormones and different types of secondary metabolites induced in response to pathogen attack (Abramovitch et al. 2006; Delprato et al. 2015). Still for a reasonable number of functionally annotated ESTs, subcellular localization is not yet known (42.60%) indicating the importance further studies in this front.

Continuous generation of nucleotide and protein sequences has enriched databases and provided a great research potential for gene function prediction and annotation. There

were different number of significant matches during homology search for most sequences in all libraries using EMBL and NCBI databases. On top of that, the discrepancy of GO terms was found in all databases; including the two coffee genome devoted databases, LGE and *C. canephora*, as well as EMBL and NCBI for a given EST. This result was in accordance to previous work on biological database integration (Gomez-Cabrero et al. 2014). A simple BLASTx search engine at LGE database resulted not only descriptions associated to each significant match, but also with associated GO terms unlike the same task at NCBI, EMBL and *C. canephora* which could only fetch short descriptions or simply an EST identification number. The development of standardized UniProt and structured GO annotation vocabulary incorporated with BLAST2GO, as per the objective of its inception (Camon 2004), provides an interface to deal with the biochemistry of annotated proteins. However, its restricted access to advanced level of annotation limits its fair availability to all users at different levels. The absence of a single run and unified access route to different databases forces the switch between different interactive interfaces, and to manually seek and combine results from different sources. Therefore, the development and availability of homology search tools like BLAST2GO and their integration to databases would help bypass the tedious and time consuming annotation works.

From the putative resistance related genes annotated, seven genes were selected and analyzed with RT-qPCR. For most of the RT-qPCR quantified genes, the overall expression trend showed increased level of transcript abundance in later time points during both interactions. Kinase associated protein phosphatase (HT12F100) and LRR-receptor like protein kinase (HT24F133) are among the main signaling genes in response to different types of biotic stresses in different plant species (Sheen 1998; Durian et al. 2016). The expression of kinase associated protein phosphatase showed no significant change during the early hours of infection, with expression peak at 72 h.a.i., in both interactions. Studies in *A. thaliana* and other plants show that phosphatase proteins involved in defense signaling have negative regulation in plant innate immunity (Shi et al. 2013; Segonzac et al. 2014). Receptor-like kinases (RLK) are one of the major defense proteins, that are structurally diversified super families, evolved into LRR proteins with intracellular kinase domain (Goff and Ramonell 2007). The extracellular domain of these proteins are in continuous evolution to recognize the ever changing pathogen effectors (Kaku et al. 2006; Zhang et al. 2006). The concentration peak of HT24F133 at 12 h.a.i.

during both interactions indicates that pathogen recognition and defense signaling occurred at the time of appressorial differentiation as also reported by Diniz et al. (2012). The coordinated activity of kinase associated phosphatases and RLK plays a decisive role in triggering resistance signaling (Alves et al. 2014). Delayed activation (after haustorial formation) of kinase associated phosphatase could be attributed to their negative modulator of stress-responsive signaling kinases at a time when elevated expression is no more required (Rodrigues et al. 2013). GA (gibberellic acid) signaling F-box gene (HT12R109), a hormone responsive gene was the other gene with expression pattern similar to kinase associated phosphatase. Post-haustorial activation of this gene may be associated to its involvement in host resistance development where HR is the major defense during the incompatible interaction (Ellis 2006). Based on our result, similar expression pattern during both interactions, GA signaling is one of the innate immunity component shared between the two genotypes. The involvement of F-box gene as regulators of defense responses has been reported in grapevine where it showed upregulated expression following *Botrytis cinerea* infection (Paquis et al. 2011). The role of GA in plant defense against pathogen attack either individually or in conjunction with other hormones has recently emerged (De Bruyne et al. 2014; Kazan and Lyons 2014). GA and its signaling f-box proteins also have a role in cross communication between signals to control development and disease defense (De Bruyne et al. 2014).

The expression of auxin-binding protein abp20 precursor (HT12F50) gene was consistently lowered at all time points, 72 h.a.i. being the lowest level during both interactions. In part, this result was similar to the work of Xue et al. (2015) on fusarium wilt in common bean in which auxin regulated protein was kept low until 72 h.a.i. during incompatible interaction. This protein precursor has been supposed to be transmembrane localized and controls the flow of auxin from cytosol to endoplasmic reticulum (Feng and Kim 2015). This gene is exceptional in that its deactivation was probably an important step for the other defense signaling genes to be upregulated. However, auxin-binding proteins were recently reported to have no role in either auxin signaling or in plant development stages in *A. thaliana* (Gao et al. 2015), hence its role in plant defense signaling is largely obscure (Feng and Kim 2015). Of all the expression patterns analyzed, a remarkably interesting defense signaling cross-talk was observed between auxin-responsive (HT24F85) and auxin-repressed protein-like (HT24F123) genes. Auxin responsive gene expression level was maintained at basal level throughout all the time

points studied during incompatible and compatible (an exception is at 72 h.a.i.) interactions. The only change in expression level was at 72 h.a.i. during compatible interaction. However, the exact role of auxin responsive protein still remains less conclusive as it has complex hormonal cross-talk signaling role in plant defense against different pathogens (Carna et al. 2014; Verma et al. 2016). On the contrary, auxin-repressed protein-like gene expression was significantly increased during haustorial differentiation (48 h.a.i. and afterwards) while it was significantly elevated at all time points during compatible interaction. Auxin-repressed protein gene is a repressor of plant growth by inhibiting the expression of auxin responsive factor gene and R gene activator in tobacco (Zhao et al. 2014) and wheat (Song et al. 2015) against fungal pathogens. However, our result showed activation of this gene in resistant (at 48 and 72 h.a.i.) and susceptible (at all time points) plants. Therefore, as it was upregulated during both interactions, its role as one of the activator of R gene expression against *H. vastatrix* in coffee requires further studies. Steady upregulation of auxin-repressed protein during both interactions, along the time course in similar fashion, was an indication that this signaling gene is indiscriminately involved in basal defense (Groszmann et al. 2015). When auxin expression is inhibited, the expression of auxin-responsive genes are expected to be low, leading to increased auxin-repressed gene expression (Tiwari 2004; Song et al. 2015). In host defense to biotrophic pathogen, downregulation of auxin responsive genes was reported to be part of salicylic acid (SA) defense signaling (Wang et al. 2007; Zhao et al. 2014). The exclusively coinciding upregulation of the two genes (auxin-responsive and auxin-repressed genes) at 72 h.a.i. during compatible interaction, which are supposed to be antagonistic otherwise, requires further studies.

A relatively different expression pattern was followed by Ac-like transposase (similar to histone H1) (HT24R75). The expression of this gene was not significantly affected at all time points studied during incompatible interaction unlike during compatible interaction in which it was significantly lowered at 24 h.a.i. The change in expression level of this gene may not be important to induce resistance during incompatible interaction while its downregulation at 24 h.a.i. during compatible interaction is also hard to neglect as it may contribute to some unsuccessful defense responses. According to Dereeper et al. (2013), substantial portion (11.9 %) of *C. canephora* (diploid parent of *C. arabica*) genome is occupied by transposable elements. The role of AC-like transposase in plant defense has recently come to light as stress adaptive capacitor in *M. oryzae* (Chadha and Sharma

2014) and determinant of susceptibility under different phytopathosystems by enhancing gene expression or chromosome rearrangement (Hua-Van et al. 2011; Yu et al. 2012). In general, similar expression patterns (up or downregulation) of most genes (HT12F50, HT12F100, HT12R109, HT24F123 and HT24F133) validated by RT-qPCR showed that these genes are not essential to prevent the establishment of *H. vastatrix* and, therefore, they are involved in basal defense response. Such shared expression patterns of resistance related genes in host and non-host plants has been reported in barley against *P. graminis* (McGrann et al. 2009).

So far identification of differentially expressed resistance related genes has been reported by SSH during compatible and incompatible interactions between *C. arabica* and *H. vastatrix* (Fernandez et al. 2004; Guzzo et al. 2009). This work extended the scope by emphasizing on explicit identification of upregulated and downregulated genes during compatible and incompatible interactions at 12 and 24 h.a.i. Moreover, the result paved a way forward in comprehensive understanding of some genes commonly over expressed and suppressed at different times in both interactions. Also, most of the genes upregulated and downregulated showed to be specific to a particular interaction. In general, the proportion of genes upregulated and downregulated in resistant cultivar showed that there was strong resistance metabolic dynamism in SAR for complete and long lasting resistance development during incompatible interaction. RT-qPCR analysis of seven resistance-signaling genes showed similar expression patterns for most of the genes in both interactions, indicating these genes are involved in basal (non-specific) defense during which immune reactions are similar.

5. REFERENCES

Abramovitch, R. B., Anderson, J. C., & Martin, G. B. (2006). Bacterial elicitation and evasion of plant innate immunity. *Nature Reviews Molecular Cell Biology*, *7*(8), 601–611. doi:10.1038/nrm1984

Afzal, A. J., Wood, A. J., & Lightfoot, D. A. (2008). Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Molecular plant-microbe interactions : MPMI*, *21*(5), 507–17. doi:10.1094/MPMI-21-5-0507

Alvarenga, M.S., Caixeta, T.E., Hufnagel, B., Thiebaut, F., Maciel-Zambolim, E., Zambolim, L., & Sakiyama, S.N. (2010). In silico identification of coffee genome expressed sequences potentially associated with resistance to diseases. *Genet. Mol. Biol.* **33**, 795–806.

Alves, M., Dadalto, S., Gonçalves, A., de Souza, G., Barros, V., & Fietto, L. (2014). Transcription Factor Functional Protein-Protein Interactions in Plant Defense Responses. *Proteomes*, *2*(1), 85–106. doi:10.3390/proteomes2010085

Alwora, G. O., & Gichuru, E. K. (2014). Advances in the Management of Coffee Berry Disease and Coffee Leaf Rust in Kenya. *Journal of Renewable Agriculture*, *2*(1), 5. doi:10.12966/jra.03.02.2014

Anderson, J. P., Gleason, C. A., Foley, R. C., Thrall, P. H., Burdon, J. B., & Singh, K. B. (2010). Plants versus pathogens: an evolutionary arms race. *Functional Plant Biology*, *37*(6), 499. doi:10.1071/FP09304

Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., et al. (2015). The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Security*, *7*(2), 303–321. doi:10.1007/s12571-015-0446-9

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, *19*(5), 455–477. doi:10.1089/cmb.2012.0021

Bari, R., & Jones, J. D. G. (2009). Role of plant hormones in plant defence responses. *Plant molecular biology*, *69*(4), 473–88. doi:10.1007/s11103-008-9435-0

Bettencourt, A. J., & Coronha-Wagner. (1971). Genetic factors conditioning resistance of Coffea arabica L. to Hemileia vastatrix Berk et Br. *Agronomia Lusitana*, *31*, 285–

292.

Bettencourt, J. . (1973). *Considerações gerais sobre o "Híbrido de Timor." Journal of Chemical Information and Modeling* (1st ed., Vol. 53). Sao Paulo: Instituto Agronômico. doi:10.1017/CBO9781107415324.004

Bettencourt AJ, R. C. (1981). *Melhoramento genético do cafeeiro - Transferência de fatores de resistência à Hemileia vastatrix Berk et Br. para as principais cultivares de Coffea arabica L* (1st ed.). Lisboa.

Bettencourt, & Rodrigues. (1988). Principles and practice of coffee breeding for resistance to rust and other diseases. *Elsevier Applied Science*, *3*, 199–234.

Bilgin, D. D., Zavala, J. A., Zhu, J., Clough, S. J., Ort, D. R., & Delucia, E. H. (2010). Biotic stress globally downregulates photosynthesis genes. *Plant, Cell & Environment*, *33*(10), 1597–1613. doi:10.1111/j.1365-3040.2010.02167.x

Bögre, L. (2003). Growth signalling pathways in Arabidopsis and the AGC protein kinases. *Trends in Plant Science*, *8*(9), 424–431. doi:10.1016/S1360-1385(03)00188-2

Bozkurt, T. O., Mcgrann, G. R. D., Maccormack, R., Boyd, L. A., & Akkaya, M. S. (2010). Cellular and transcriptional responses of wheat during compatible and incompatible race-specific interactions with Puccinia striiformis f. sp. tritici. *Molecular Plant Pathology*, no-no. doi:10.1111/j.1364-3703.2010.00633.x

Burdon, J. J., & Thrall, P. H. (2009). Coevolution of Plants and Their Pathogens in Natural Habitats. *Science*, *324*(5928), 755–756. doi:10.1126/science.1171663

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry*, *55*(4), 611–22. doi:10.1373/clinchem.2008.112797

Byers, R. J., Hoyland, J. A., Dixon, J., & Freemont, A. J. (2000). Subtractive hybridization--genetic takeaways and the search for meaning. *International journal of experimental pathology*, *81*(6), 391–404. http://www.ncbi.nlm.nih.gov/pubmed/11298187

Cação, S. M. B., Silva, N. V., Domingues, D. S., Vieira, L. G. E., Diniz, L. E. C., Vinecky, F., et al. (2013). Construction and characterization of a BAC library from the Coffea arabica genotype Timor Hybrid CIFC 832/2. *Genetica*, *141*(4–6), 217–226. doi:10.1007/s10709-013-9720-y

Cacas, J.-L., Petitot, A.-S., Bernier, L., Estevan, J., Conejero, G., Mongrand, S., &

Fernandez, D. (2011). Identification and characterization of the Non-race specific Disease Resistance 1 (NDR1) orthologous protein in coffee. *BMC Plant Biology*, *11*(1), 144. doi:10.1186/1471-2229-11-144

Camon, E. (2004). The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Research*, *32*(90001), 262–266. doi:10.1093/nar/gkh021

Carazzolle, M. F., Rabello, F. R., Martins, N. F., de Souza, A. A., do Amaral, A. M., Freitas-Astua, J., et al. (2011). Identification of defence-related genes expressed in coffee and citrus during infection by Xylella fastidiosa. *European Journal of Plant Pathology*, *130*(4), 529–540. doi:10.1007/s10658-011-9775-5

Carna, M., Repka, V., Skupa, P., & Sturdik, E. (2014). Auxins in defense strategies. *Biologia*, *69*(10), 1255–1263. doi:10.2478/s11756-014-0431

Carvalho, C. R., Fernandes, R. C., Carvalho, G. M. A., Barreto, R. W., & Evans, H. C. (2011). Cryptosexuality and the Genetic Diversity Paradox in Coffee Rust, Hemileia vastatrix. *PLoS ONE*, *6*(11), e26387. doi:10.1371/journal.pone.0026387

Casassola, A., Brammer, S. P., Chaves, M. S., Martinelli, J. A., Grando, M. F., D ', N., & Denardin, Á. (2013). Gene Expression: A Review on Methods for the Study of Defense-Related Gene Differential Expression in Plants. *American Journal of Plant Sciences*, *4*(December), 64–73. doi:10.4236/ajps.2013.4.12A3008

Cenci, A., Combes, M.-C., & Lashermes, P. (2010). Comparative sequence analyses indicate that Coffea (Asterids) and Vitis (Rosids) derive from the same paleo-hexaploid ancestral genome. *Molecular Genetics and Genomics*, *283*(5), 493–501. doi:10.1007/s00438-010-0534-7

Chadha, S., & Sharma, M. (2014). Transposable Elements as Stress Adaptive Capacitors Induce Genomic Instability in Fungal Pathogen Magnaporthe oryzae. *PLoS ONE*, *9*(4), e94415. doi:10.1371/journal.pone.0094415

Chauhan, H., Khurana, N., Agarwal, P., Khurana, J. P., & Khurana, P. (2013). A Seed Preferential Heat Shock Transcription Factor from Wheat Provides Abiotic Stress Tolerance and Yield Enhancement in Transgenic Arabidopsis under Heat Stress Environment. *PLoS ONE*, *8*(11), e79577. doi:10.1371/journal.pone.0079577

Choi, K., Reinhard, C., Serra, H., Ziolkowski, P. A., Underwood, C. J., Zhao, X., et al. (2016). Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes. *PLOS Genetics*, *12*(7), e1006179. doi:10.1371/journal.pgen.1006179

Coll, N. S., Epple, P., & Dangl, J. L. (2011). Programmed cell death in the plant immune

system. *Cell Death and Differentiation*, *18*(8), 1247–1256. doi:10.1038/cdd.2011.37

Combes, M.-C., Lashermes, P., Chalhoub, B., Noir, S., & Patheyron, S. (2004). Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (Coffea arabica L.). *TAG Theoretical and Applied Genetics*, *109*(1), 225–230. doi:10.1007/s00122-004-1604-1

CONAB. (2017). *Acompamento da safra brasileira : café*. Brasília. http://www.conab.gov.br

Cristancho, M. A., Botero-Rozo, D. O., Giraldo, W., Tabima, J., RiaÃ±o-PachÃ³n, D. M., Escobar, C., et al. (2014). Annotation of a hybrid partial genome of the coffee rust (Hemileia vastatrix) contributes to the gene repertoire catalog of the Pucciniales. *Frontiers in Plant Science*, *5*. doi:10.3389/fpls.2014.00594

Cruz, F., Kalaoun, S., Nobile, P., Colombo, C., Almeida, J., Barros, L. M. G., et al. (2009). Evaluation of coffee reference genes for relative expression studies by quantitative real-time RT-PCR. *Molecular Breeding*, *23*(4), 607–616. doi:10.1007/s11032-009-9259-x

Davis, A. P., Govaerts, R., Bridson, D. M., & Stoffelen, P. (2006). An annotated taxonomic conspectus of the genus Coffea (Rubiaceae). *Botanical Journal of the Linnean Society*, *152*(4), 465–512. doi:10.1111/j.1095-8339.2006.00584.x

De Bruyne, L., Höfte, M., & De Vleesschauwer, D. (2014). Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Molecular plant*, *7*(6), 943–59. doi:10.1093/mp/ssu050

Delledonne, M. (2005). NO news is good news for plants. *Current Opinion in Plant Biology*, *8*(4), 390–396. doi:10.1016/j.pbi.2005.05.002

Delprato, M. L., Krapp, A. R., & Carrillo, N. (2015). Green Light to Plant Responses to Pathogens: The Role of Chloroplast Light-Dependent Signaling in Biotic Stress. *Photochemistry and Photobiology*, *91*(5), 1004–1011. doi:10.1111/php.12466

Denancé, N., Sánchez-Vallet, A., Goffner, D., & Molina, A. (2013). Disease resistance or growth: the role of plant hormones in balancing immune responses and fitness costs. *Frontiers in Plant Science*, *4*. doi:10.3389/fpls.2013.00155

Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., et al. (2014). The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science*, *345*(6201), 1181–1184. doi:10.1126/science.1255274

Dereeper, A., Guyot, R., Tranchant-Dubreuil, C., Anthony, F., Argout, X., de Bellis, F., et al. (2013). BAC-end sequences analysis provides first insights into coffee (Coffea

canephora P.) genome composition and evolution. *Plant Molecular Biology*, *83*(3), 177–189. doi:10.1007/s11103-013-0077-5

DeYoung, B. J., & Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*, *7*(12), 1243–1249. doi:10.1038/ni1410

Diatchenko, L., Lau, Y. F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., et al. (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences*, *93*(12), 6025–6030. doi:10.1073/pnas.93.12.6025

Diatchenko, L., Lukyanov, S., Lau, Y. F., & Siebert, P. D. (1999). Suppression subtractive hybridization: a versatile method for identifying differentially expressed genes. *Methods in enzymology*, *303*, 349–80. doi:S0076-6879(99)03022-0 [pii]

Ding, B., & Wang, G.-L. (2015). Chromatin versus pathogens: the function of epigenetics in plant immunity. *Frontiers in plant science*, *6*, 675. doi:10.3389/fpls.2015.00675

Diniz, I., Talhinhas, P., Azinheira, H. G., Várzea, V., Medeira, C., Maia, I., et al. (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*(1), 141–157. doi:10.1007/s10658-011-9925-9

Diniz, L. E. C., Sakiyama, N. S., Lashermes, P., Caixeta, E. T., Oliveira, A. C. B., Zambolim, E. M., et al. (2005). Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. *Cropp Breeding and Applied Biotechnology*, *5*(4), 387–393. doi:10.12702/1984-7033.v05n04a03

Diola, V., Brito, G. G., Caixeta, E. T., Pereira, L. F. P., & Loureiro, M. E. (2013). A new set of differentially expressed signaling genes is early expressed in coffee leaf rust race II incompatible interaction. *Functional & integrative genomics*, *13*(3), 379–89. doi:10.1007/s10142-013-0330-7

Diola, V., de Brito, G. G., Caixeta, E. T., Maciel-Zambolim, E., Sakiyama, N. S., & Loureiro, M. E. (2011). High-density genetic mapping for coffee leaf rust resistance. *Tree Genetics & Genomes*, *7*(6), 1199–1208. doi:10.1007/s11295-011-0406-2

Djebbi, S., Bouktila, D., Makni, H., Makni, M., & Mezghani-Khemakhem, M. (2015). Identification and characterization of novel NBS-LRR resistance gene analogues from the pea. *Genetics and Molecular Research*, *14*(2), 6419–6428. doi:10.4238/2015.June.11.18

Dolatabadi, B., Ranjbar, G., Tohidfar, M., & Dehestani, A. (2014). Genetic transformation of Tomato with three pathogenesis-related protein genes for increased resistance to Fusarium oxysporum f . sp . lycopersici, *2*(1).

Durian, G., Rahikainen, M., Alegre, S., Brosché, M., & Kangasjärvi, S. (2016). Protein Phosphatase 2A in the Regulatory Network Underlying Biotic Stress Resistance in Plants. *Frontiers in Plant Science*, *7*. doi:10.3389/fpls.2016.00812

Ellis, J. (2006). Insights into Nonhost Disease Resistance: Can They Assist Disease Control in Agriculture? *THE PLANT CELL ONLINE*, *18*(3), 523–528. doi:10.1105/tpc.105.040584

Ellis, J., Dodds, P., & Pryor, T. (2000). Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology*, *3*(4), 278–284. doi:10.1016/S1369-5266(00)00080-7

Emanuelsson, O., Nielsen, H., Brunak, S., & von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of molecular biology*, *300*(4), 1005–16. doi:10.1006/jmbi.2000.3903

Eugenia Rivera, M., Codina, J. C., Olea, F., De Vicente, A., & Pérez-García, A. (2002). Differential expression of β -1,3-glucanase in susceptible and resistant melon cultivars in response to infection by Sphaerotheca fusca. *Physiological and Molecular Plant Pathology*, *61*(5), 257–265. doi:10.1006/pmpp.2002.0436

Farrar, K., & Donnison, I. S. (2007). Construction and screening of BAC libraries made from Brachypodium genomic DNA. *Nature protocols*, *2*(7), 1661–74. doi:10.1038/nprot.2007.204

Felsenstein, J. (1985). Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. *Evolution*. doi:Doi 10.2307/2408678

Feng, M., & Kim, J.-Y. (2015). Revisiting Apoplastic Auxin Signaling Mediated by AUXIN BINDING PROTEIN 1. *Molecules and cells*, *38*(10), 829–35. doi:10.14348/molcells.2015.0205

Fernandez, D., Santos, P., Agostini, C., Bon, M.-C., Petitot, A.-S., C Silva, M., et al. (2004). Coffee (Coffea arabica L.) genes early expressed during infection by the rust fungus (Hemileia vastatrix). *Molecular Plant Pathology*, *5*(6), 527–36. doi:10.1111/j.1364-3703.2004.00250.x

Figueiredo, S. A., Lashermes, P., & Aragão, F. J. L. (2011). Molecular characterization and functional analysis of the β-galactosidase gene during Coffea arabica (L.) fruit development. *Journal of experimental botany*, *62*(8), 2691–703.

doi:10.1093/jxb/erq440

Flajolet, M., Rakhilin, S., Wang, H., Starkova, N., Nuangchamnong, N., Nairn, A. C., & Greengard, P. (2003). Protein phosphatase 2C binds selectively to and dephosphorylates metabotropic glutamate receptor 3. *Proceedings of the National Academy of Sciences*, *100*(26), 16006–16011. doi:10.1073/pnas.2136600100

Flor, H. H. (1971). Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology*, *9*(1), 275–296. doi:10.1146/annurev.py.09.090171.001423

Freeman. (2008). An Overview of Plant Defenses against Pathogens and Herbivores. *The Plant Health Instructor*. doi:10.1094/PHI-I-2008-0226-01

Freitas, R. (2015). *Análise citológica e perfil de expressão gênica de Hemileia vastatrix (raça xxxiii) na interação com o cafeeiro*. Universidade Federal de Viçosa.

Ganesh, D., Petitot, A.-S., Silva, M. C., Alary, R., Lecouls, A.-C., & Fernandez, D. (2006). Monitoring of the early molecular resistance responses of coffee (Coffea arabica L.) to the rust fungus (Hemileia vastatrix) using real-time quantitative RT-PCR. *Plant Science*, *170*(6), 1045–1051. doi:10.1016/j.plantsci.2005.12.009

Gao, M., Niu, J., Zhao, S., Jiao, C., Xu, W., Fei, Z., & Wang, X. (2012). Characterization of Erysiphe necator-Responsive Genes in Chinese Wild Vitis quinquangularis. *International Journal of Molecular Sciences*, *13*(12), 11497–11519. doi:10.3390/ijms130911497

Gao, Y., Zhang, Y., Zhang, D., Dai, X., Estelle, M., & Zhao, Y. (2015). Auxin binding protein 1 (ABP1) is not required for either auxin signaling or Arabidopsis development. *Proceedings of the National Academy of Sciences*, *112*(7), 2275–2280. doi:10.1073/pnas.1500365112

Gaudet, D. A., Lu, Z.-X., Leggett, F., Puchalski, B., & Laroche, A. (2007). Compatible and Incompatible Interactions in Wheat Involving the Bt-10 Gene for Resistance to Tilletia tritici , the Common Bunt Pathogen. *Phytopathology*, *97*(11), 1397–1405. doi:10.1094/PHYTO-97-11-1397

Gichuru, E. K., Ithiru, J. M., Silva, M. C., Pereira, A. P., & Varzea, V. M. P. (2012a). Additional physiological races of coffee leaf rust (Hemileia vastatrix) identified in Kenya. *Tropical Plant Pathology*, *37*(6), 424–427. doi:10.1590/S1982-56762012000600008

Gichuru, E. K., Ithiru, J. M., Silva, M. C., Pereira, A. P., & Varzea, V. M. P. (2012b). Additional physiological races of coffee leaf rust (Hemileia vastatrix) identified in Kenya. *Tropical Plant Pathology*, *37*(December), 424–427. doi:10.1590/S1982-

56762012000600008

Gill, U. S., Lee, S., & Mysore, K. S. (2015). Host Versus Nonhost Resistance: Distinct Wars with Similar Arsenals. *Phytopathology*, *105*(5), 580–587. doi:10.1094/PHYTO-11-14-0298-RVW

Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual review of phytopathology*, *43*(1), 205–27. doi:10.1146/annurev.phyto.43.040204.135923

Goff, K. E., & Ramonell, K. M. (2007). The role and regulation of receptor-like kinases in plant defense. *Gene regulation and systems biology*, *1*, 167–75. http://www.ncbi.nlm.nih.gov/pubmed/19936086

Gomez-Cabrero, D., Abugessaisa, I., Maier, D., Teschendorff, A., Merkenschlager, M., Gisel, A., et al. (2014). Data integration in the era of omics: current and future challenges. *BMC systems biology*, *8 Suppl 2*, I1. doi:10.1186/1752-0509-8-S2-I1

González, V. M., Aventín, N., Centeno, E., & Puigdomènech, P. (2014). Interspecific and intraspecific gene variability in a 1-Mb region containing the highest density of NBS-LRR genes found in the melon genome. *BMC Genomics*, *15*(1), 1131. doi:10.1186/1471-2164-15-1131

Gou, X., He, K., Yang, H., Yuan, T., Lin, H., Clouse, S. D., & Li, J. (2010). Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in Arabidopsis thaliana. *BMC Genomics*, *11*(1), 19. doi:10.1186/1471-2164-11-19

Groszmann, M., Gonzalez-Bayon, R., Lyons, R. L., Greaves, I. K., Kazan, K., Peacock, W. J., & Dennis, E. S. (2015). Hormone-regulated defense and stress response networks contribute to heterosis in Arabidopsis F1 hybrids. *Proceedings of the National Academy of Sciences*, *112*(46), E6397–E6406. doi:10.1073/pnas.1519926112

Guerra-Guimarães, L., Tenente, R., Pinheiro, C., Chaves, I., Silva, M. do C., Cardoso, F. M. H., et al. (2015). Proteomic analysis of apoplastic fluid of Coffea arabica leaves highlights novel biomarkers for resistance against Hemileia vastatrix. *Frontiers in plant science*, *6*, 478. doi:10.3389/fpls.2015.00478

Guindalini, C., & Tufik, S. (2007). Uso de microarrays na busca de perfis de expressão gênica: aplicação no estudo de fenótipos complexos. *Revista Brasileira de Psiquiatria*, *29*(4), 370–374. doi:10.1590/S1516-44462007000400014

Guo, Y.-L., Fitz, J., Schneeberger, K., Ossowski, S., Cao, J., & Weigel, D. (2011).

Genome-Wide Comparison of Nucleotide-Binding Site-Leucine-Rich Repeat-Encoding Genes in Arabidopsis. *PLANT PHYSIOLOGY*, *157*(2), 757–769. doi:10.1104/pp.111.181990

Guttman, D. S., Vinatzer, B. A., Sarkar, S. F., Ranall, M. V, Kettler, G., & Greenberg, J. T. (2002). A functional screen for the type III (Hrp) secretome of the plant pathogen Pseudomonas syringae. *Science (New York, N.Y.)*, *295*(5560), 1722–6. doi:10.1126/science.295.5560.1722

Guzzo, S. D., Harakava, R., & Tsai, S. M. (2009). Identification of Coffee Genes Expressed During Systemic Acquired Resistance and Incompatible Interaction with Hemileia vastatrix. *Journal of Phytopathology*, *157*(10), 625–638. doi:10.1111/j.1439-0434.2008.01538.x

Hammond-Kosack, K. E., & Kanyuka, K. (2007). Resistance Genes (R Genes) in Plants. In *eLS*. Chichester, UK: John Wiley & Sons, Ltd. doi:10.1002/9780470015902.a0020119

Hao, W., Collier, S. M., Moffett, P., & Chai, J. (2013). Structural Basis for the Interaction between the Potato Virus X Resistance Protein (Rx) and Its Cofactor Ran GTPase-activating Protein 2 (RanGAP2). *Journal of Biological Chemistry*, *288*(50), 35868–35876. doi:10.1074/jbc.M113.517417

He, L., Du, C., Covaleda, L., Xu, Z., Robinson, A. F., Yu, J. Z., et al. (2004). Cloning, Characterization, and Evolution of the NBS-LRR-Encoding Resistance Gene Analogue Family in Polyploid Cotton (Gossypium hirsutum L.). *Molecular Plant-Microbe Interactions*, *17*(11), 1234–1241. doi:10.1094/MPMI.2004.17.11.1234

Helepciuc, F. E., Mitoi, M. E., Manole-P??unescu, A., Aldea, F., Brezeanu, A., & Cornea, C. P. (2014). Induction of plant antioxidant system by interaction with beneficial and/or pathogenic microorganisms. *Romanian Biotechnological Letters*, *19*(3), 9366–9375.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., & Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology*, *8*(2), R19. doi:10.1186/gb-2007-8-2-r19

Howald, C., Tanzer, A., Chrast, J., Kokocinski, F., Derrien, T., Walters, N., et al. (2012). Combining RT-PCR-seq and RNA-seq to catalog all genic elements encoded in the human genome. *Genome Research*, *22*(9), 1698–1710.

Hua-Van, A., Le Rouzic, A., Boutin, T. S., Filée, J., & Capy, P. (2011). The struggle for

life of the genome's selfish architects. *Biology Direct*, *6*(1), 19. doi:10.1186/1745-6150-6-19

Hulbert, S. H., Webb, C. A., Smith, S. M., & Sun, Q. (2001). Resistance gene complexes: evolution and utilization. *Annual review of phytopathology*, *39*, 285–312. doi:10.1146/annurev.phyto.39.1.285

Hunt, M. D., Neuenschwander, U. H., Delaney, T. P., Weymann, K. B., Friedrich, L. B., Lawton, K. A., et al. (1996). Recent advances in systemic acquired resistance research - A review. *Gene*, *179*(1), 89–95. doi:10.1016/S0378-1119(96)00429-5

Jach, G., Gornhardt, B., Mundy, J., Logemann, J., Pinsdorf, E., Leah, R., et al. (1995). Enhanced quantitative resistance against fungal disease by combinatorial expression of different barley antifungal proteins in transgenic tobacco. *The Plant Journal*, *8*(1), 97–109. doi:10.1046/j.1365-313X.1995.08010097.x

Jacob, F., Vernaldi, S., & Maekawa, T. (2013). Evolution and Conservation of Plant NLR Functions. *Frontiers in Immunology*, *4*. doi:10.3389/fimmu.2013.00297

Jashni, M. K., Mehrabi, R., Collemare, J., Mesarich, C. H., & de Wit, P. J. G. M. (2015). The battle in the apoplast: further insights into the roles of proteases and their inhibitors in plant–pathogen interactions. *Frontiers in Plant Science*, *6*. doi:10.3389/fpls.2015.00584

Jelenska, J., Yao, N., Vinatzer, B. A., Wright, C. M., Brodsky, J. L., & Greenberg, J. T. (2007). A J domain virulence effector of Pseudomonas syringae remodels host chloroplasts and suppresses defenses. *Current biology : CB*, *17*(6), 499–508. doi:10.1016/j.cub.2007.02.028

Jonak, C., Okrész, L., Bögre, L., & Hirt, H. (2002). Complexity, cross talk and integration of plant MAP kinase signalling. *Current opinion in plant biology*, *5*(5), 415–24. http://www.ncbi.nlm.nih.gov/pubmed/12183180

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323–329. doi:10.1038/nature05286

Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., et al. (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences*, *103*(29), 11086–11091. doi:10.1073/pnas.0508882103

Kang, Y., Kim, K., Shim, S., Yoon, M., Sun, S., Kim, M., et al. (2012). Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. *BMC Plant Biology*, *12*(1), 139. doi:10.1186/1471-2229-12-139

Kazan, K., & Lyons, R. (2014). Intervention of Phytohormone Pathways by Pathogen Effectors. *The Plant cell*, *26*(6), 2285–2309. doi:10.1105/tpc.114.125419

Keen, N. T. (1990). Gene-For-Gene Complementarity in Plant-Pathogen Interactions. *Annual Review of Genetics*, *24*(1), 447–463. doi:10.1146/annurev.ge.24.120190.002311

Kelley, J. M., Field, C. E., Craven, M. B., Bocskai, D., Kim, U. J., Rounsley, S. D., & Adams, M. D. (1999). High throughput direct end sequencing of BAC clones. *Nucleic Acids Research*, *27*(6), 1539–1546. doi:10.1093/nar/27.6.1539

Kierszniowska, S., Seiwert, B., & Schulze, W. X. (2009). Definition of Arabidopsis Sterol-rich Membrane Microdomains by Differential Treatment with Methyl- -cyclodextrin and Quantitative Proteomics. *Molecular & Cellular Proteomics*, *8*(4), 612–623. doi:10.1074/mcp.M800346-MCP200

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36. doi:10.1186/gb-2013-14-4-r36

Kim, S. H., Kwon, S. I., Saha, D., Anyanwu, N. C., & Gassmann, W. (2009). Resistance to the Pseudomonas syringae Effector HopA1 Is Governed by the TIR-NBS-LRR Protein RPS6 and Is Enhanced by Mutations in SRFR1. *PLANT PHYSIOLOGY*, *150*(4), 1723–1732. doi:10.1104/pp.109.139238

Kobe, B., & Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends in Biochemical Sciences*, *19*(10), 415–421. doi:10.1016/0968-0004(94)90090-6

Koç, E., & Üstün, A. S. (2012). Infl uence of Phytophthora capsici L . inoculation on disease severity , necrosis length , peroxidase and catalase activity , and phenolic content of resistant and susceptible pepper (Capsicum annuum L .) plants, *36*, 357–371. doi:10.3906/biy-1109-12

Korasick, D. A., Westfall, C. S., Lee, S. G., Nanao, M. H., Dumas, R., Hagen, G., et al. (2014). Molecular basis for AUXIN RESPONSE FACTOR protein interaction and the control of auxin response repression. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(14), 5427–32. doi:10.1073/pnas.1400074111

Kovtun, Y., Chiu, W.-L., Tena, G., & Sheen, J. (2000). Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. *Proceedings of*

the National Academy of Sciences, *97*(6), 2940–2945. doi:10.1073/pnas.97.6.2940

Kuang, H., Wei, F., Marano, M. R., Wirtz, U., Wang, X., Liu, J., et al. (2005). The R1 resistance gene cluster contains three groups of independently evolving, type I R1 homologues and shows substantial structural variation among haplotypes of Solanum demissum. *The Plant Journal*, *44*(1), 37–51. doi:10.1111/j.1365-313X.2005.02506.x

Kumar, D. (2012). Isolation of Nucleotide Binding Site (NBS)-Leucine Rich Repeat (LRR) Resistant Gene Analogs (Rgas) In Arabica Coffee (Coffea Arabica L. Cv S.288). *Journal of Biotechnology & Biomaterials*, *2*(5), 146. doi:10.4172/2155-952X.1000146

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. doi:10.1093/molbev/msw054

Kushalappa, A. C., Yogendra, K. N., & Karre, S. (2016). Plant Innate Immune Response: Qualitative and Quantitative Resistance. *Critical Reviews in Plant Sciences*, *35*(1), 38–55. doi:10.1080/07352689.2016.1148980

Lahaye, T. (2002). The Arabidopsis RRS1-R disease resistance gene--uncovering the plant's nucleus as the new battlefield of plant defense? *Trends in plant science*, *7*(10), 425–7. http://www.ncbi.nlm.nih.gov/pubmed/12399170

Lamb, C., & Dixon, R. A. (1997). THE OXIDATIVE BURST IN PLANT DISEASE RESISTANCE. *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*(1), 251–275. doi:10.1146/annurev.arplant.48.1.251

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, *11*(3), 204–220. doi:10.1038/nrg2719

Legay, G., Marouf, E., Berger, D., Neuhaus, J.-M., Mauch-Mani, B., & Slaughter, A. (2011). Identification of genes expressed during the compatible interaction of grapevine with Plasmopara viticola through suppression subtractive hybridization (SSH). *European Journal of Plant Pathology*, *129*(2), 281–301. doi:10.1007/s10658-010-9676-z

Leipe, D. D., Koonin, E. V., & Aravind, L. (2004). STAND, a Class of P-Loop NTPases Including Animal and Plant Regulators of Programmed Cell Death: Multiple, Complex Domain Architectures, Unusual Phyletic Patterns, and Evolution by Horizontal Gene Transfer. *Journal of Molecular Biology*, *343*(1), 1–28.

doi:10.1016/j.jmb.2004.08.023

Leroy, T., Marraccini, P., Dufour, M., Montagnon, C., Lashermes, P., Sabau, X., et al. (2005). Construction and characterization of a Coffea canephora BAC library to study the organization of sucrose biosynthesis genes. *Theoretical and Applied Genetics*, *111*(6), 1032–1041. doi:10.1007/s00122-005-0018-z

Leubner-Metzger, G. (2003). Functions and regulation of β-1,3-glucanases during seed germination, dormancy release and after-ripening. *Seed Science Research*, *13*(1), 17–34. doi:10.1079/SSR2002121

Levine, A., Tenhaken, R., Dixon, R., & Lamb, C. (1994). H2O2 from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell*, *79*(4), 583–593. doi:10.1016/0092-8674(94)90544-4

Lin, C., Mueller, L. A., Carthy, J. M., Crouzillat, D., Pétiard, V., & Tanksley, S. D. (2005). Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. *Theoretical and Applied Genetics*, *112*(1), 114–130. doi:10.1007/s00122-005-0112-2

Liu, J., Liu, X., Dai, L., & Wang, G. (2007). Recent Progress in Elucidating the Structure, Function and Evolution of Disease Resistance Genes in Plants. *Journal of Genetics and Genomics*, *34*(9), 765–776. doi:10.1016/S1673-8527(07)60087-3

Maeda, T., Tsai, A. Y., & Saito, H. (1993). Mutations in a protein tyrosine phosphatase gene (PTP2) and a protein serine/threonine phosphatase gene (PTC1) cause a synthetic growth defect in Saccharomyces cerevisiae. *Molecular and cellular biology*, *13*(9), 5408–17. http://www.ncbi.nlm.nih.gov/pubmed/8395005

Mahé, L., Combes, M. C., Várzea, V. M. P., Guilhaumon, C., & Lashermes, P. (2008). Development of sequence characterized DNA markers linked to leaf rust (Hemileia vastatrix) resistance in coffee (Coffea arabica L.). *Molecular Breeding*, *21*(1), 105–113.

Marone, D., Russo, M., Laidò, G., De Leonardis, A., & Mastrangelo, A. (2013). Plant Nucleotide Binding Site–Leucine-Rich Repeat (NBS-LRR) Genes: Active Guardians in Host Defense Responses. *International Journal of Molecular Sciences*, *14*(4), 7302–7326. doi:10.3390/ijms14047302

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, *1*(1), vev003-vev003. doi:10.1093/ve/vev003

Martínez, C. P., Echeverri, C., Florez, J. C., Gaitan, A. L., & Góngora, C. E. (2012). In

vitro production of two chitinolytic proteins with an inhibiting effect on the insect coffee berry borer, Hypothenemus hampei (Ferrari) (Coleoptera: Curculionidae) and the fungus Hemileia vastatrix the most limiting pests of coffee crops. *AMB Express*, *2*(1), 22. doi:10.1186/2191-0855-2-22

Matschi, S., Hake, K., Herde, M., Hause, B., & Romeis, T. (2015). The calcium-dependent protein kinase CPK28 regulates development by inducing growth phase-specific, spatially restricted alterations in jasmonic acid levels independent of defense responses in Arabidopsis. *The Plant Cell*, *27*(3), 591–606. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4558673&tool=pmcentrez&rendertype=abstract

Mayne, W. W. (1932). Physiological Specialisation of Hemileia vastatrix B. and Br. *Nature*, *129*(3257), 510–510. doi:10.1038/129510a0

McGrann, G. R. D., Townsend, B. J., Antoniw, J. F., Asher, M. J. C., & Mutasa-Göttgens, E. S. (2009). Barley elicits a similar early basal defence response during host and non-host interactions with Polymyxa root parasites. *European Journal of Plant Pathology*, *123*(1), 5–15. doi:10.1007/s10658-008-9332-z

McHale, L., Tan, X., Koehl, P., & Michelmore, R. W. (2006). Plant NBS-LRR proteins: adaptable guards. *Genome biology*, *7*(4), 212. doi:10.1186/gb-2006-7-4-212

Medeiros, F. C. L., Resende, M. L. V., Medeiros, F. H. V., Zhang, H. M., & Paré, P. W. (2009). Defense gene expression induced by a coffee-leaf extract formulation in tomato. *Physiological and Molecular Plant Pathology*, *74*(2), 175–183. doi:10.1016/j.pmpp.2009.11.004

Melo, G. A., Shimizu, M. M., & Mazzafera, P. (2006). Polyphenoloxidase activity in coffee leaves and its role in resistance against the coffee leaf miner and coffee leaf rust. *Phytochemistry*, *67*(3), 277–285. doi:10.1016/j.phytochem.2005.11.003

Meyers, B. C., Kaushik, S., & Nandety, R. S. (2005). Evolving disease resistance genes. *Current Opinion in Plant Biology*, *8*(2), 129–134. doi:10.1016/j.pbi.2005.01.002

Michelmore, R. W., & Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome research*, *8*(11), 1113–30. http://www.ncbi.nlm.nih.gov/pubmed/9847076

Mondragon-Palomino, M. (2002). Patterns of Positive Selection in the Complete NBS-LRR Gene Family of Arabidopsis thaliana. *Genome Research*, *12*(9), 1305–1315. doi:10.1101/gr.159402

Mondragon-Palomino, M., & Gaut, B. S. (2005). Gene Conversion and the Evolution of

Three Leucine-Rich Repeat Gene Families in Arabidopsis thaliana. *Molecular Biology and Evolution*, *22*(12), 2444–2456. doi:10.1093/molbev/msi241

Moody, D. E. (2001). Genomics techniques: An overview of methods for the study of gene expression. *Journal of Animal Science*, *79*(E-Suppl), E128. doi:10.2527/jas2001.79E-SupplE128x

Morris, E. R., & Walker, J. C. (2003). Receptor-like protein kinases: the keys to response. *Current Opinion in Plant Biology*, *6*(4), 339–342. doi:10.1016/S1369-5266(03)00055-4

Mur, L. A. J., Kenton, P., Lloyd, A. J., Ougham, H., & Prats, E. (2008). The hypersensitive response; the centenary is upon us but how much do we know? *Journal of Experimental Botany*, *59*(3), 501–520. doi:10.1093/jxb/erm239

Mur, L. A. J., Prats, E., Pierre, S., Hall, M. A., & Hebelstrup, K. H. (2013). Integrating nitric oxide into salicylic acid and jasmonic acid/ ethylene plant defense pathways. *Frontiers in Plant Science*, *4*. doi:10.3389/fpls.2013.00215

Mussatto, S. I., Machado, E. M. S., Martins, S., & Teixeira, J. A. (2011). Production, Composition, and Application of Coffee and Its Industrial Residues. *Food and Bioprocess Technology*, *4*(5), 661–672. doi:10.1007/s11947-011-0565-z

Muthamilarasan, M., & Prasad, M. (2013). Plant innate immunity: An updated insight into defense mechanism. *Journal of Biosciences*, *38*(2), 433–449. doi:10.1007/s12038-013-9302-2

Ndimba, B. K., Chivasa, S., Hamilton, J. M., Simon, W. J., & Slabas, A. R. (2003). Proteomic analysis of changes in the extracellular matrix of Arabidopsis cell suspension cultures induced by fungal elicitors. *Proteomics*, *3*(6), 1047–59. doi:10.1002/pmic.200300413

Nei, M., & Kumar, S. (2000). *Molecular evolution and phylogenetics* (1st ed.). Oxford: Oxford University Press.

Nepal, M., & Benson, B. (2015). CNL Disease Resistance Genes in Soybean and Their Evolutionary Divergence. *Evolutionary Bioinformatics*, 49. doi:10.4137/EBO.S21782

Neuhaus, J.-M., Fritig, B., Linthorst, H. J. M., Meins, F., Mikkelsen, J. D., & Ryals, J. (1996). A revised nomenclature for chitinase genes. *Plant Molecular Biology Reporter*, *14*(2), 102–104. doi:10.1007/BF02684897

Nicot, N., Hausman, J.-F., Hoffmann, L., & Evers, D. (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic

stress. *Journal of experimental botany*, *56*(421), 2907–14. doi:10.1093/jxb/eri285

Nie, S., Yue, H., Zhou, J., & Xing, D. (2015). Mitochondrial-Derived Reactive Oxygen Species Play a Vital Role in the Salicylic Acid Signaling Pathway in Arabidopsis thaliana. *PLOS ONE*, *10*(3), e0119853. doi:10.1371/journal.pone.0119853

Niks, R. E., & Marcel, T. C. (2009). Nonhost and basal resistance: how to explain specificity? *New Phytologist*, *182*(4), 817–828. doi:10.1111/j.1469-8137.2009.02849.x

Nimchuk, Z., Eulgem, T., Holt III, B. F., Dangl, J. L., Holt, B. F., & Dangl, J. L. (2003). Recognition and response in the plant immune system. *Annual review of genetics*, *37*(1), 579–609. doi:10.1146/annurev.genet.37.110801.142628

Noir, S., Combes, M. C., Anthony, F., & Lashermes, P. (2001). Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (Coffea L.). *Molecular genetics and genomics : MGG*, *265*(4), 654–62. http://www.ncbi.nlm.nih.gov/pubmed/11459185

Noronha-Wagner and Bettencourt, A. (1967). Genetic study of the resistance of Coffea sp to leaf rust 1. Identification and behavior of four factors conditioning disease reaction in Coffea arabica to twelve physiologic races of Hemileia vastatrix. *Canadian Journal of Botany*, *45*, 2021–2031.

Nürnberger, T., & Scheel, D. (2001). Signal transmission in the plant immune response. *Trends in Plant Science*, *6*(8), 372–379. doi:10.1016/S1360-1385(01)02019-2

Padmanabhan, M., Cournoyer, P., & Dinesh-Kumar, S. P. (2009). The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities. *Cellular Microbiology*, *11*(2), 191–198. doi:10.1111/j.1462-5822.2008.01260.x

Padmanabhan, M. S., Kramer, S. R., Wang, X., & Culver, J. N. (2008). Tobacco mosaic virus replicase-auxin/indole acetic acid protein interactions: reprogramming the auxin response pathway to enhance virus infection. *Journal of virology*, *82*(5), 2477–85. doi:10.1128/JVI.01865-07

Panchy, N., Lehti-Shiu, M. D., & Shiu, S.-H. (2016). Evolution of gene duplication in plants. *Plant Physiology*, pp.00523.2016. doi:10.1104/pp.16.00523

Paquis, S., Mazeyrat-Gourbeyre, F., Fernandez, O., Crouzet, J., Clément, C., Baillieul, F., & Dorey, S. (2011). Characterization of a F-box gene up-regulated by phytohormones and upon biotic and abiotic stresses in grapevine. *Molecular Biology Reports*, *38*(5), 3327–3337. doi:10.1007/s11033-010-0438-y

Peck, S. C. (2003). Early phosphorylation events in biotic stress. *Current Opinion in Plant*

Biology, *6*(4), 334–338. doi:10.1016/S1369-5266(03)00056-6

Pereira, A. A., Sakiyama, N. S., Zambolim, L., Moura, W. M., Zambolim, E., & Caixeta, E. T. (2005). Identification and use of Sources of Durable Resistance to Coffee Leaf Rust in the UFV/EPAMIG Breeding Program. In L. Zambolim, E. M. Zambolim, & V. M. P. Várzea (Eds.), *Durable Resistance to Coffee Leaf Rust* (1st ed., pp. 215–232). Visconde do Rio Branco: Suprema Gráfica e Editora Ltda.

Petitot, A.-S., Lecouls, A.-C., & Fernandez, D. (2008). Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species Coffea arabica. *Tree Genetics & Genomes*, *4*(3), 379–390. doi:10.1007/s11295-007-0117-x

Petre, B., Joly, D. L., & Duplessis, S. (2014). Effector proteins of rust fungi. *Frontiers in Plant Science*, *5*. doi:10.3389/fpls.2014.00416

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*(9), 45e–45. doi:10.1093/nar/29.9.e45

Piedras, P., Rivas, S., Droge, S., Hillmer, S., & Jones, J. D. G. (2000). Functional, c-myc-tagged Cf-9 resistance gene products are plasma-membrane localized and glycosylated. *The Plant Journal*, *21*(6), 529–536. doi:10.1046/j.1365-313x.2000.00697.x

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J. P., & Hyvönen, J. (2013). Advances in plant gene-targeted and functional markers: a review. *Plant Methods*, *9*(1), 6. doi:10.1186/1746-4811-9-6

Qi, D., & Innes, R. W. (2013). Recent Advances in Plant NLR Structure, Function, Localization, and Signaling. *Frontiers in Immunology*, *4*. doi:10.3389/fimmu.2013.00348

Rairdan, G. J., Collier, S. M., Sacco, M. A., Baldwin, T. T., Boettrich, T., & Moffett, P. (2008). The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. *THE PLANT CELL ONLINE*, *20*(3), 739–751. doi:10.1105/tpc.107.056036

Rajesh, M. K., Rachana, K. E., Naganeeswaran, S. A., Shafeeq, R., Thomas, R. J., Shareefa, M., et al. (2015). Identification of expressed resistance gene analog sequences in coconut leaf transcriptome and their evolutionary analysis. *Turkish Journal of Agriculture and Forestry*, *39*, 489–502. doi:10.3906/tar-1409-75

Ratnaparkhe, M. B., Wang, X., Li, J., Compton, R. O., Rainville, L. K., Lemke, C., et al. (2011). Comparative analysis of peanut NBS-LRR gene clusters suggests

evolutionary innovation among duplicated domains and erosion of gene microsynteny. *New Phytologist*, *192*(1), 164–178. doi:10.1111/j.1469-8137.2011.03800.x

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. doi:10.1186/1471-2164-12-240

Rodrigues, A., Adamo, M., Crozet, P., Margalha, L., Confraria, A., Martinho, C., et al. (2013). ABI1 and PP2CA Phosphatases Are Negative Regulators of Snf1-Related Protein Kinase1 Signaling in Arabidopsis. *The Plant Cell*, *25*(10), 3871–3884. doi:10.1105/tpc.113.114066

Rodrigues, C. J., Bettencourt, A. J., & Rijo, L. (1975). Races of the Pathogen and Resistance to Coffee Rust. *Annual Review of Phytopathology*, *13*(1), 49–70. doi:10.1146/annurev.py.13.090175.000405

Rojas, C. M., Senthil-Kumar, M., Tzin, V., & Mysore, K. S. (2014). Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. *Frontiers in Plant Science*, *5*. doi:10.3389/fpls.2014.00017

Romeis, T. (2001). Calcium-dependent protein kinases play an essential role in a plant defence response. *The EMBO Journal*, *20*(20), 5556–5567. doi:10.1093/emboj/20.20.5556

Rozas, J. (2009). DNA sequence polymorphism analysis using DnaSP. *Methods in molecular biology (Clifton, N.J.)*. doi:10.1007/978-1-59745-251-9_17

Rutherford, M. a, & Phiri, N. (2006). Pests and Diseases of Coffee in Eastern Africa : A Technical and Advisory Manual compiled & edited by. *CABI African Regional Centre (Nairobi)*.

Rzhetsky, a, & Nei, M. (1992). A Simple Method for Estimating and Testing Minimum-Evolution Trees. *Molecular Biology and Evolution*, *9*(5), 945–967.

Saithong, T., Rongsirikul, O., Kalapanulak, S., Chiewchankaset, P., Siriwat, W., Netrphan, S., et al. (2013). Starch biosynthesis in cassava: a genome-based pathway reconstruction and its exploitation in data integration. *BMC Systems Biology*, *7*(1), 75. doi:10.1186/1752-0509-7-75

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, *4*(4), 406–25. http://www.ncbi.nlm.nih.gov/pubmed/3447015

Saminathan, T., Nimmakayala, P., Manohar, S., Malkaram, S., Almeida, A., Cantrell, R.,

et al. (2015). Differential gene expression and alternative splicing between diploid and tetraploid watermelon. *Journal of Experimental Botany*, *66*(5), 1369–1385. doi:10.1093/jxb/eru486

Sanabria, N., Goring, D., Nürnberger, T., & Dubery, I. (2008). Self/nonself perception and recognition mechanisms in plants: a comparison of self-incompatibility and innate immunity. *New Phytologist*, *178*(3), 503–514. doi:10.1111/j.1469-8137.2008.02403.x

Sanabria, N. M., Huang, J.-C., & Dubery, I. A. (2010). Self/non-self perception in plants in innate immunity and defense. *Self/Nonself*, *1*(1), 40–54. doi:10.4161/self.1.1.10442

Sanseverino, W., Hénaff, E., Vives, C., Pinosio, S., Burgos-Paz, W., Morgante, M., et al. (2015). Transposon Insertions, Structural Variations, and SNPs Contribute to the Evolution of the Melon Genome. *Molecular Biology and Evolution*, *32*(10), 2760–2774. doi:10.1093/molbev/msv152

Sara, M., Walter, S., Paola, C., Luigi, M., Luigi, F., & Raffalella, E. M. (2012). Solanaceae Evolutionary Dynamics of the <i>I</i>2-NBS Domain. *American Journal of Plant Sciences*, *3*(2), 283–294. doi:10.4236/ajps.2012.32034

Sasaki, A., Itoh, H., Gomi, K., Ueguchi-Tanaka, M., Ishiyama, K., Kobayashi, M., et al. (2003). Accumulation of phosphorylated repressor for gibberellin signaling in an F-box mutant. *Science (New York, N.Y.)*, *299*(5614), 1896–8. doi:10.1126/science.1081077

Saunders, D. G. O., Win, J., Cano, L. M., Szabo, L. J., Kamoun, S., & Raffaele, S. (2012). Using Hierarchical Clustering of Secreted Protein Families to Classify and Rank Candidate Effectors of Rust Fungi. *PLoS ONE*, *7*(1), e29847. doi:10.1371/journal.pone.0029847

Schneider, D. J., & Collmer, A. (2010). Studying plant-pathogen interactions in the genomics era: beyond molecular Koch's postulates to systems biology. *Annual review of phytopathology*, *48*, 457–79. doi:10.1146/annurev-phyto-073009-114411

Schwarzlander, M., Konig, A.-C., Sweetlove, L. J., & Finkemeier, I. (2012). The impact of impaired mitochondrial function on retrograde signalling: a meta-analysis of transcriptomic responses. *Journal of Experimental Botany*, *63*(4), 1735–1750. doi:10.1093/jxb/err374

Segonzac, C., Macho, A. P., Sanmartín, M., Ntoukakis, V., Sánchez-Serrano, J. J., &

Zipfel, C. (2014). Negative control of BAK1 by protein phosphatase 2A during plant innate immunity. *The EMBO journal*, *33*(18), 2069–79. doi:10.15252/embj.201488698

Seifi, A., Gao, D., Zheng, Z., Pavan, S., Faino, L., Visser, R. G. F., et al. (2014). Genetics and molecular mechanisms of resistance to powdery mildews in tomato (Solanum lycopersicum) and its wild relatives. *European Journal of Plant Pathology*, *138*(3), 641–665. doi:10.1007/s10658-013-0314-4

Sekhwal, M., Li, P., Lam, I., Wang, X., Cloutier, S., & You, F. (2015). Disease Resistance Gene Analogs (RGAs) in Plants. *International Journal of Molecular Sciences*, *16*(8), 19248–19290. doi:10.3390/ijms160819248

Sera, G. H., Sera, T., Ito, D. S., De Azevedo, J. A., Da Mata, J. S., Dói, D. S., et al. (2007). Resistance to leaf rust in coffee carrying SH3 gene and others SH genes. *Brazilian Archives of Biology and Technology*, *50*(5), 753–757.

Seybold, H., Trempel, F., Ranf, S., Scheel, D., Romeis, T., & Lee, J. (2014). Ca2+ signalling in plant immune response: from pattern recognition receptors to Ca2+ decoding mechanisms. *The New phytologist*, *204*(4), 782–90. http://www.ncbi.nlm.nih.gov/pubmed/25539002

Sheen, J. (1998). Mutational analysis of protein phosphatase 2C involved in abscisic acid signal transduction in higher plants. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 975–80. http://www.ncbi.nlm.nih.gov/pubmed/9448270

Shi, H., Shen, Q., Qi, Y., Yan, H., Nie, H., Chen, Y., et al. (2013). BR-SIGNALING KINASE1 Physically Associates with FLAGELLIN SENSING2 and Regulates Plant Innate Immunity in Arabidopsis. *The Plant Cell*, *25*(3), 1143–1157. doi:10.1105/tpc.112.107904

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., & Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8794–7. doi:10.1073/pnas.89.18.8794

Silva, M. C., Guerra-Guimarães, L., Loureiro, A., & Nicole, M. R. (2008). Involvement of peroxidases in the coffee resistance to orange rust (Hemileia vastatrix). *Physiological and Molecular Plant Pathology*, *72*(1–3), 29–38. doi:10.1016/j.pmpp.2008.04.004

Song, J., & Bent, A. F. (2014). Microbial Pathogens Trigger Host DNA Double-Strand Breaks Whose Abundance Is Reduced by Plant Defense Responses. *PLoS Pathogens*, *10*(4), e1004030. doi:10.1371/journal.ppat.1004030

Song, J., Niu, Z., Li, Q., Bao, Y., Ma, X., Wang, H., et al. (2015). Isolation and Identification of Differentially Expressed Genes from Wheat in Response to Blumeria graminis f. sp. tritici (Bgt). *Plant Molecular Biology Reporter*, *33*(5), 1371–1380. doi:10.1007/s11105-014-0838-6

Spoel, S. H., & Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nature Reviews Immunology*, *12*(2), 89–100. doi:10.1038/nri3141

Stanke, M., Steinkamp, R., Waack, S., & Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research*, *32*(Web Server), W309–W312. doi:10.1093/nar/gkh379

Stone, D. (2014). Fungus, Climate Change Threatening Big Part of Global Coffee Supply. *National Geographic*. http://news.nationalgeographic.com/news/2014/05/140531-coffee-rust-columbia-brazil-cost-problems/

Stotz, H. U., Mitrousia, G. K., de Wit, P. J. G. M., & Fitt, B. D. L. (2014). Effector-triggered defence against apoplastic fungal pathogens. *Trends in Plant Science*, *19*(8), 491–500. doi:10.1016/j.tplants.2014.04.009

Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., & Brown, J. W. S. (2012). Alternative splicing in plants – coming of age. *Trends in Plant Science*, *17*(10), 616–623. doi:10.1016/j.tplants.2012.06.001

Tan, S., & Wu, S. (2012). Genome Wide Analysis of Nucleotide-Binding Site Disease Resistance Genes in Brachypodium distachyon. *Comparative and Functional Genomics*, *2012*, 1–12. doi:10.1155/2012/418208

Thomma, B. P. H. J., Nürnberger, T., & Joosten, M. H. A. J. (2011). Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy. *The Plant Cell*, *23*(1), 4–15. doi:10.1105/tpc.110.082602

Tiwari, S. B. (2004). Aux/IAA Proteins Contain a Potent Transcriptional Repression Domain. *THE PLANT CELL ONLINE*, *16*(2), 533–543. doi:10.1105/tpc.017384

Ulmasov, T. (1997). Aux/IAA Proteins Repress Expression of Reporter Genes Containing Natural and Highly Active Synthetic Auxin Response Elements. *THE PLANT CELL ONLINE*, *9*(11), 1963–1971. doi:10.1105/tpc.9.11.1963

Umbrasaite, J., Schweighofer, A., & Meskiene, I. (2011). Substrate Analysis of

Arabidopsis PP2C-Type Protein Phosphatases (pp. 149–161). doi:10.1007/978-1-61779-264-9_8

van der Biezen, E. A., & Jones, J. D. (1998). The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Current biology : CB*, *8*(7), R226-7. doi:10.1016/S0960-9822(98)70145-9 showArticle Info

Van der Biezen, E. A., & Jones, J. D. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in biochemical sciences*, *23*(12), 454–6. http://www.ncbi.nlm.nih.gov/pubmed/9868361

Van der Hoorn, R. A. L. (2001). Identification of Distinct Specificity Determinants in Resistance Protein Cf-4 Allows Construction of a Cf-9 Mutant That Confers Recognition of Avirulence Protein AVR4. *THE PLANT CELL ONLINE*, *13*(2), 273–285. doi:10.1105/tpc.13.2.273

Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D. A., Cestaro, A., Pruss, D., et al. (2007). A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety. *PLoS ONE*, *2*(12), e1326. doi:10.1371/journal.pone.0001326

Verma, V., Ravindran, P., & Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC Plant Biology*, *16*(1), 86. doi:10.1186/s12870-016-0771-y

Vidal, R. O., Mondego, J. M. C., Pot, D., Ambrosio, A. B., Andrade, A. C., Pereira, L. F. P., et al. (2010). A High-Throughput Data Mining of Single Nucleotide Polymorphisms in Coffea Species Expressed Sequence Tags Suggests Differential Homeologous Gene Expression in the Allotetraploid Coffea arabica. *PLANT PHYSIOLOGY*, *154*(3), 1053–1066. doi:10.1104/pp.110.162438

Villalobos, D. P., Bautista, R., Cánovas, F. M., & Claros, M. G. (2004). Isolation of bacterial artificial chromosome DNA by means of improved alkaline lysis and double potassium acetate precipitation. *Plant Molecular Biology Reporter*, *22*(4), 419–425. doi:10.1007/BF02772684

Wang, D., Pajerowska-Mukhtar, K., Culler, A. H., & Dong, X. (2007). Salicylic Acid Inhibits Pathogen Growth in Plants through Repression of the Auxin Signaling Pathway. *Current Biology*, *17*(20), 1784–1790. doi:10.1016/j.cub.2007.09.025

Wang, G.-F., Ji, J., EI-Kasmi, F., Dangl, J. L., Johal, G., & Balint-Kurti, P. J. (2015). Molecular and Functional Analyses of a Maize Autoactive NB-LRR Protein Identify Precise Structural Requirements for Activity. *PLOS Pathogens*, *11*(2), e1004674. doi:10.1371/journal.ppat.1004674

Wang, X., Wu, F., Xie, Q., Wang, H., Wang, Y., Yue, Y., et al. (2012). SKIP Is a Component of the Spliceosome Linking Alternative Splicing and the Circadian Clock in Arabidopsis. *The Plant Cell*, *24*(8), 3278–3295. doi:10.1105/tpc.112.100081

Wang, Z., Gerstein, M., & Snyder, M. (2010). RNA-Seq: a revolutionary tool for transcriptomics, *10*(1), 57–63. doi:10.1038/nrg2484.RNA-Seq

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, N.J.)*, *112*, 531–52. http://www.ncbi.nlm.nih.gov/pubmed/10027275

Wu, S., Shan, L., & He, P. (2014). Microbial signature-triggered plant defense responses and early signaling mechanisms. *Plant Science*, *228*, 118–126. doi:10.1016/j.plantsci.2014.03.001

Wurzinger, B., Mair, A., Pfister, B., & Teige, M. (2011). Cross-talk of calcium-dependent protein kinase and MAP kinase signaling. *Plant Signaling & Behavior*, *6*(1), 8–12. doi:10.4161/psb.6.1.14012

Xue, R., Wu, J., Zhu, Z., Wang, L., Wang, X., Wang, S., & Blair, M. W. (2015). Differentially Expressed Genes in Resistant and Susceptible Common Bean (Phaseolus vulgaris L.) Genotypes in Response to Fusarium oxysporum f. sp. phaseoli. *PLOS ONE*, *10*(6), e0127698. doi:10.1371/journal.pone.0127698

Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., et al. (2014). PredictProtein--an open resource for online prediction of protein structural and functional features. *Nucleic Acids Research*, *42*(W1), W337–W343. doi:10.1093/nar/gku366

Yang, D.-L., Li, Q., Deng, Y.-W., Lou, Y.-G., Wang, M.-Y., Zhou, G.-X., et al. (2008). Altered disease development in the eui mutants and Eui overexpressors indicates that gibberellins negatively regulate rice basal disease resistance. *Molecular plant*, *1*(3), 528–37. doi:10.1093/mp/ssn021

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2014). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, *12*(1), 7–8. doi:10.1038/nmeth.3213

Yang, S., Zhang, X., Yue, J.-X., Tian, D., & Chen, J.-Q. (2008). Recent duplications dominate NBS-encoding gene expansion in two woody species. *Molecular Genetics and Genomics*, *280*(3), 187–198. doi:10.1007/s00438-008-0355-0

Yang, Y.-X., Ahammed, G. J., Wu, C., Fan, S., & Zhou, Y.-H. (2015). Crosstalk among Jasmonate, Salicylate and Ethylene Signaling Pathways in Plant Disease and Immune Responses. *Current protein & peptide science*, *16*(5), 450–61. http://www.ncbi.nlm.nih.gov/pubmed/25824390

Yu, C., Han, F., Zhang, J., Birchler, J., & Peterson, T. (2012). A transgenic system for generation of transposon Ac/Ds-induced chromosome rearrangements in rice. *Theoretical and Applied Genetics*, *125*(7), 1449–1462. doi:10.1007/s00122-012-1925-4

Yu, D., Chen, C., & Chen, Z. (2001). Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. *The Plant cell*, *13*(7), 1527–40. doi:10.1105/TPC.010115

Zambolim, L. (2016). Current status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology*, *41*(1), 1–8. doi:10.1007/s40858-016-0065-9

Zhang, X. S., Choi, J. H., Heinz, J., & Chetty, C. S. (2006). Domain-Specific Positive Selection Contributes to the Evolution of Arabidopsis Leucine-Rich Repeat Receptor-Like Kinase (LRR RLK) Genes. *Journal of Molecular Evolution*, *63*(5), 612–621. doi:10.1007/s00239-005-0187-z

Zhao, Y., Huang, J., Wang, Z., Jing, S., Wang, Y., Ouyang, Y., et al. (2016). Allelic diversity in an NLR gene BPH9 enables rice to combat planthopper variation. *Proceedings of the National Academy of Sciences*, *113*(45), 12850–12855. doi:10.1073/pnas.1614862113

Zhao, Y., Li, C., Ge, J., Xu, M., Zhu, Q., Wu, T., et al. (2014). Recessive mutation identifies auxin-repressed protein ARP1, which regulates growth and disease resistance in tobacco. *Molecular plant-microbe interactions : MPMI*, *27*(7), 638–54. doi:10.1094/MPMI-08-13-0250-R

Zhu, Z., & Lee, B. (2015). Friends or Foes: New Insights in Jasmonate and Ethylene Co-Actions. *Plant and Cell Physiology*, *56*(3), 414–420. doi:10.1093/pcp/pcu171

Zipfel, C. (2009). Early molecular events in PAMP-triggered immunity. *Current opinion in plant biology*, *12*(4), 414–20. doi:10.1016/j.pbi.2009.06.003

Zuckerkandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. *Evolving genes and proteins*, 97–166. doi:10.1209/epl/i1998-00224-x

CHAPTER 2

STRUCTURAL AND FUNCTIONAL ANALYSIS OF GENES WITH POTENTIAL INVOLVEMENT IN RESISTANCE TO COFFEE LEAF RUST: A FUNCTIONAL MARKER BASED APPROACH

ABSTRACT

Physiological based differentiation of S_H genes and *Hemileia vastatrix* races is one of the principal methods routinely employed in characterizing coffee leaf rust resistance. Molecular techniques like BAC clone sequencing and functional characterization of target gene is believed to enhance precisions to overcome the daunting limitations in classical breeding. In the present work, RGAs with resistance potential to coffee leaf rust were sequenced and characterized from a BAC library screened by functional marker. Among the 13 predicted ORFs, five RGAs were annotated and mapped to chromosome 0 of *C. canephora*. Four of the RGAs are actively expressed during *C. arabica-H. vastatrix* incompatible interaction. Based on the result, it could be inferred that at least one (gene 11) of the sequenced RGAs is a new S_H gene (S_H10) not yet identified. We also report an S_H gene (S_H10) in differential host clone 644/18 H. Kawisari for the first time. Moreover, comparative analysis of two RGAs belonging to the CC-NBS-LRR gene family showed intense diversifying selection due to nonsynonymous substitution and gene recombination/conversion. On top of that, phylogenetic analysis of orthologous genes showed high interaspecies variability among the two genes in related species than in coffee. The CC-NBS-LRR genes sequenced in this work are the largest and most complete sequence ever reported in Arabica coffee, making the work extremely important for molecular breeding of coffee rust resistance.

Keywords: Diversifying selection, CC-NBS-LRR, recombination, RGAs, 644/18 H. Kawisari

ANÁLISE ESTRUTURAL E FUNCIONAL DE GENES COM POTENCIAL EM ENVOLVEMENTO NA RESISTANCIA AO FERRUGEM DO CAFEEIRO: UMA ABORDAGEM BASEADA EM MARCADOR FUNCIONAL

RESUMO

A diferenciação fisiológica de genes S_H e raças de *H. vastatrix* é um dos principais métodos utilizados rotineiramente na caracterização da resistência à ferrugem do café. Técnicas moleculares como o sequenciamento e caracterização funcional de genes-alvo presentes em clones BAC aumentam a precisão e reduzem limitações encontradas nos programas de melhoramento clássicos. No presente trabalho, *resistance gene analogs* (RGAs) com potencial de resistência à ferrugem do café foram sequenciadas e caracterizadas a partir de uma biblioteca BAC *screened* por marcador funcional. Foram previstos 13 ORFs e a partir destes, cinco RGAs foram anotadas e mapeadas para o cromossomo 0 de *C. canephora*. Quatro RGAs são ativamente expressas em interação incompatível entre *C. arabica* e *H. vastatrix*. Os resultados obtidos no trabalho sugerem que um desses genes RGA sequenciado (gene 11) é um novo gene S_H (S_H10) ainda não identificado. Com base nesses dados, foi verificado pela primeira vez o novo gene S_H (S_H10) no clone diferenciador 644/18 H. Kawisari. Foi realizado a análise comparativa entre os cincos RGAs e verificado alta similaridade entre dois destes, os quais são pertencentes à família de genes CC-NBS-LRR. Foi verificado a ocorrência de intensa seleção diversificada promovida pela substituição não sinônima e pela recombinação genética. Foi feita a análise filogenética de genes ortólogos para as espécies de café, tomate e uva e observou-se alta variabilidade intraespecífica destes dois genes CC-NBS-LRR para as espécies, exceto para o café. Estes genes CC-NBS-LRR sequenciados são a maior e mais completa sequência disponível para o *C. arabica*. Estes resultados são de extrema importância para o melhoramento molecular visando a resistência à ferrugem do cafeeiro.

Palavras-chaves: Seleção diversificada, CC-NBS-LRR, recombinação, RGAs, 644/18 H. Kawisari

1.INTRODUCTION

Coffee is one of the most valuable cash crops in many developing economies as it provides employment opportunities in cultivation, processing and marketing; thereby sustaining the livelihoods of millions of people around the world (Mussatto *et al.*, 2011). *H. vastatrix*, the causative agent of coffee leaf rust, has been accounted for one of the major threats to coffee production in almost every coffee producing region. Despite the release of some resistant coffee cultivars in recent years, coffee leaf rust is still adversely affecting coffee production and undermining the incomes of many households (Zambolim, 2016). Currently, there are at least 49 physiological races of *H. vastatrix* characterized at CIFC (Centro de Investigação das Ferrugens do Cafeeiro, Portugal) (Gichuru *et al.*, 2012; Zambolim 2016). The constantly emerging new races and sporadic outbreak of the disease have imposed a challenging burden on resistance breeding. On top of that, the most pressing concern is the breakdown of resistance genes leading to the susceptibility of cultivars once praised for their source of genetic material for resistance breeding (Cristancho *et al.*, 2014).

Molecular components of coffee genes involved in different metabolic pathways, their evolution and annotation has been unveiled with the complete sequencing of *C. canephora* genome (Denoeud *et al.*, 2014). As *C. canephora* makes half of the arabica coffee genome, being a natural hybrid of *C. canephora* and *C. eugenioides*, the creation and open access to its genome has provided enormous insights in understanding the genome of *C. arabica* in the last two years. The discovery and successful introgression of S_H3 resistance gene locus to cultivated Arabica coffee from *C. liberica* was another landmark often considered as one of the greatest milestones in the development of coffee rust resistance (Prakash et al. 2004; Mahé et al. 2008; Ribas et al. 2011). Since then, molecular and physical mapping has enabled the sequencing and annotation of S_H3 region, resulting in the discovery of multiple resistance (R) genes (Cenci et al. 2010; Ribas et al. 2011). Dominantly inherited, the largest class of R-genes encode nucleotide binding site leucine-rich-repeat (NBS-LRR) proteins that directly recognize the corresponding pathogen virulence (v) protein or its effects (He *et al.*, 2004; Jones and Dangl 2006; McHale et al. 2006). These genes are believed to contain several hundred gene families, which are unevenly distributed in genomes of different plant species

(Hulbert *et al.*, 2001). Intracellular signaling domain similar to Drosophila toll/mammalian interleukin-1 receptor (TNL, Toll-NBS-LRR) and the coiled-coil (CNL, CC-NBS-LRR) are the two major N-terminal amino acid sequences preceding NBS domain involved in specific signal transduction (DeYoung and Innes 2006; Jones and Dangl 2006; Tan and Wu 2012). Yet the other N-terminal domain linked to LRR includes leucine-zipper (a transmembrane protein, TM), protein kinase (PK) and WRKY TIR (Liu *et al.*, 2007). These domains are predominantly involved in resistance signal transduction by conformational changes (Leipe *et al.*, 2004). On the carboxyl-terminal region is the LRR, mediating specific protein-protein interaction to recognize pathogen effectors (Van der Hoorn 2001; Kushalappa et al. 2016). Though these domains are few in number, nucleotide polymorphism and variability of LRR region is responsible for the perception of a specific pathogen effector (Ellis et al. 2000; McHale et al. 2006). Inter and intraspecific extreme variability of NBS-LRR has been attributed to gene duplication, unequal crossing over, recombination, deletion, point mutation and selection pressure due to continuous response to diverse pathogen races (Yang et al. 2008; Ribas et al. 2011).

The readily available Arabica coffee BAC libraries constructed from disease resistant cultivars at different laboratories have accelerated several studies involving resistance gene cloning (Combes et al. 2004; Cação et al. 2013). Furthermore, arbitrary DNA marker based and functional (gene) marker application in gene cloning has shown a huge applicability in crop improvement, either by map based cloning using the former or direct gene cloning using the latter or both (Poczai *et al.*, 2013). The advantage of direct cloning of gene of interest over map-based gene cloning is appealing as this method is more precise and straightforward for gene characterization.

In coffee, the origin and organization of disease resistance genes has begun to emerge in recent years as part of an effort in understanding the role of major resistance genes against coffee rust. One of such endeavors was the assembly of R genes spanning the S_H3 locus in order to trace the evolution and diversity of LRR domains in three coffee species (Ribas *et al.*, 2011). Despite the partial sequencing and annotation of several disease resistance genes in Arabica coffee (Noir *et al.*, 2001), completely sequenced and characterized candidate genes are not yet readily available. On the other hand, rust resistance conferred

by nine major genes (S_H1-9) and their corresponding v_{1-9} pathogen factors have been known for long in coffee rust pathosystem (Rodrigues *et al.*, 1975; Gichuru *et al.*, 2012; Alwora and Gichuru, 2014). Nonetheless, molecular and functional characterization of any of the S_H genes and associated regulatory elements is entirely obscure with immense potential in changing the perspective of rust resistance breeding. Likewise, the use of functional markers serving as a direct rust resistance screening amongst host differential coffee clones remains a timely breeding objective but a barely addressed issue. Moreover, the absence of characteristic candidate rust resistance gene to be used in coffee genetic transformation is one of the bottleneck problems in coffee breeding. Resistance gene analog (RGA) marker CARF005 was previously confirmed to share disease resistant ORF region in coffee (Noir et al. 2001; Alvarenga *et al.*, 2010). As this polymorphic RGA marker was confirmed to encode disease resistant protein domain NB-ARC (nucleotide binding site-ARC: ARC for APAF-1, R protein and CED-4 from Van der Biezen and Jones, 1998), exclusively conferred by *H. vastatrix* resistant coffee cultivars (Alvarenga *et al.*, 2010), its complete sequencing and molecular characterization could provide a potential candidate disease resistance gene. Therefore, the objective of the current study was to characterize resistance gene analog (RGA) (CARF005) and associated regulatory elements. We are also interested to investigate whether any of the completely sequenced genes are activated during *C. arabica*-*H. vastatrix* interaction and belong to the S_H gene series.

2. MATERIALS AND METHODS

2.1 Plant genetic materials

Greenhouse grown 22 differential coffee clones carrying at least one of the coffee rust resistance genes (S_H1-9) and 3 genotypes susceptible to all virulence factors (v1-9) of *H. vastatrix* were used for CARF005 screening. These differential plants for different physiological races of *H. vastatrix* were initially characterized by CIFC. All clones were vegetatively propagated at the Plant Pathology Department greenhouse of the Universidade Federal de Viçosa. Genomic DNA was extracted from young second pair leaves following Diniz *et al.* (2005). DNA integrity was checked by 1% gel electrophoresis stained with ethidium bromide (0.5 µg/ml) and Nanodrop (NanoDrop Technologies, Wilmingon, DE, USA) and stored at -20°C until use. RNA-Seq libraries

(here after referred as transcriptome) constructed at 12 and 24 hours after pathogen inoculation (hai) during *C. arabica* CIFC 832/2 –*H. vastatrix* (race XXXIII) incompatible interaction (unpublished) was used as reference in searching for candidate resistance genes.

2.3 PCR conditions

A Sigma made (Sigma-Aldrich, Belo Horizonte, Brazil) disease RGA primer, CARF005, (F: GGACATCAACACCAACCTC and R: ATCCCTACCATCCACTTCAAC) (Alvarenga *et al.*, 2010) was used to screen differential host clones. PCR reagents were 1x buffer, 0.2 mM dNTPs, 0.2 µM primers, 1 mM $MgCl_2$, 0.8 u Taq polymerase and 5 ng DNA in a reaction volume of 20 µl. PCR cycling parameters were as follows: DNA denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 30s, 60°C for 30s and 72°C for 1 min before an extension at 72°C for 10 min. PCR products were screened for target insert by running 1% UltraPure™ agarose (Invitrogen) gel electrophoresis stained with ethidium bromide (0.5 µg/ml). All PCR and gel electrophoresis conditions were maintained as such throughout this paper unless stated otherwise.

2.4 BAC clone screening

A renown rust resistant Hibrido de Timor clone CIFC 832/2 BAC library of 56,832 clones constructed by Cação et al. (2013) was used as a source of target RGA (CARF005). These clones were replicated on 384 well titer plates using plate replicator sterilized by 10% H_2O_2 for 2 min, rinsed in sterile water for seconds and socked in 70% ethanol under laminar airflow cabinet. After the alcohol was evaporated (3-5 min), old cultures were copied to a new 384 well titer plate with 70 µl fresh LB media (with 12.5 μgml^{-1} chloramphenicol) in each well. Culture multiplication was done by incubating in a temperature of 37°C for 18 h on a shaker at a velocity of 180 rpm. Identification of clones with CARF005 insert was done by grouping of clones and subsequent group decomposition of the 384 clones until a single clone was identified (Appendix II Figure 1A). BAC DNA was extracted using centrifugation protocol of Wizard® SV Plus Minipreps DNA Purification System (Promega).

2.5 Sequencing and contig assembly

Isolated single BAC clone with CARF005 fragment insert was sequenced by Illumina HiSeq2000/2500 100PE (paired-end reads) platform at Macrogen (Seoul, South Korea). Paired-end sequence processing and contig assembly was done using SPAdes software (Bankevich et al. 2012). Before any downstream sequence processing, contigs hit with bacterial genome *(E. coli)* and flanking vector sequence (pCC1BAC^{TM}) were duly excluded. The assembled BAC contigs were used to map against a transcriptome constructed from coffee genes activated in response to *H. vastatrix* infection by Tophat 2 (Kim et al. 2013) to locate the region of contig with active expression of genes.

2.6 Gene prediction and annotation

Contigs with ≥ 200 bp size and sharing ≥ 90% identity with *C. canephora* were taken to Augustus gene prediction (Stanke et al. 2004). Among the available genomes in the Augustus dataset, *Solanum lycopersicum* was used as a reference genome as they share common gene repertoires and similar in genome size (Lin et al. 2005). Predicted ORFs were annotated using different online annotation tools. First, NCBI BLASTp was launched to detect conserved domains and their description, followed by The Predict Protein Server (Yachdav et al. 2014) molecular analysis and associated GO search. Protein 3D structure and nucleotide (ATP/ADP/GTP/GDP) binding sites were predicted by I-TASSER suite online tool (Yang et al. 2014). As an annotation complement, predicted ORFs were BLASTed against coding sequences (CDS) of *S. lycopersicum* (Sol Genomics Network: https://solgenomics.net/tools/blast/) and *V. vinifera* (Phytozome 11: https://phytozome.jgi.doe.gov) genomes.

2.7 Motif search

Promoter motifs were searched in the upstream regions of all the predicted ORFs by using TSSP /Prediction of PLANT Promoters (http://www.softberry.com/berry). Theoretical isoelectric point (pI) and protein weight (kda) were estimated using the ExPASy proteomics server of the Swiss Institute of Bioinformatics (Wilkins et al. 1999) and Science Gateway online server (http://www.sciencegateway.org/tools/proteinmw.htm). Motif and feature detection was also confirmed by SMART motif analysis (http://smart.embl-heidelberg.de).

2.8 Sequence alignment and comparative analysis

Genes encoding resistance proteins were mapped to *C. canephora* genome (Denoeud et al. 2014) to trace back their probable origin and organization. BLASTn was launched against the *C. canephora* genome at coffee genome hub (http://coffee-genome.org/blast). The *H. vastatrix* defense transcriptome reads were aligned to contig 9 by using Tophat2 (-N 3 --read-gap-length 3 --read-edit-dist 6 --no-coverage-search --b2-very-sensitive) to locate the region of the contig encoding genes against the pathogen. Intergenic physical position, distance and orientation was analyzed for RGAs.

2.9 Point mutation analysis

RGAs were analyzed for indels and substitutions by EMBL MUSCLE multiple sequence alignment (http://www.ebi.ac.uk/Tools/msa/muscle/) and MEGA7 (Kumar et al. 2016). Gene duplication was exclusively analyzed by MEGA 7 while DNA polymorphism and non-synonymous/synonymous substitution rate (ka/ks) were analyzed by DnaSP v5.1 (Rozas, 2009). Recombination/conversion events and parental phylogeny among paralogous and orthologous genes were analyzed by RDP4 software (Martin et al. 2015).

2.10 Functional and phylogenetic analysis

Based on molecular evolution of protein domains, functional diversity between two NBS-LRR RGAs from coffee was analyzed. Additionally, protein sequence based comparative phylogenetic tree was constructed for the two genes and their orthologs from S. *lycopersicum* and *V. vinifera* genomes by MEGA7 software (Kumar et al. 2016). The evolutionary history was inferred using minimum evolutionary method (Rzhetsky and Nei 1992).

3. RESULTS

3.1 Resistance gene screening among differential coffee clones

Among the 22 differential coffee clones comprising different S_H genes, CARF005 (RGA marker) was detected in eight clones (Table 1 and Appendix II Figure 2). Based on PCR amplicon gel analysis, this RGA marker seemed to amplify S_H6 gene, with two

exceptions. One contradicting result is that the gene was detected in 128/2-Dilla & Alghe, which is supposed to have just S_H1 gene. And also, CARF005 was amplified in differential host clone 644/18 H. Kawisari for which no S_H gene was reported so far.

Table 1. Screening for CARF005 marker using 22 differential coffee clones currently used for S_H gene characterization.

No.	Differential clone*	Susceptible to (*H. vastatrix* physiological race)	S_H gene conferred**	CARF005 (+/-)
1	832/1-HT	-	6,7,8,9,?	+
2	HW17/12	XVI,XXIII	1,2,4,5	-
3	1343/269-HT	XXII,XXV,XXVI,XXVII,XXVIII,XXIX, XXXI,XXXII,XXXIII,XXXVII,XXXIX,XL	6	+
4	H153/2	XII, XVI	1,3,5	-
5	H420/10	XXIX	5,6,7,9	+
6	110/5-S 4 Agaro	X,XIV,XV, XVI,XXIII,XXIV,XXVI, XXVIII	4,5	-
7	128/2-Dilla & Alghe	III, X, XII, XVI, XVII, XIX,XXIII, XXVII	1	+
8	134/4-S12 Kaffa	X, XVI, XIX, XX, XXIII ,XXVII,	1,4	-
9	H419/20	XXIX, XXXI	5,6,9	+
10	635/3-S 12 Kaffa	X, XIV,XV,XVI,XIX, XXIII,XXIV,XXVI,XXVII,XXVIII	4	-
11	87/1-Geisha	III, X, XII, XVI, XVII, XXIII	1,5	-
12	1006/10-KP 532	XII,XVI,XVII, XXIII	1,2,5	-
13	7963/117-Catimor	XXXIII	5,7 or 5,7,9	-
14	HW17/12	XVI, XXIII	1,2,4,5	-
15	H420/2	XXIX, XXX	5,8	-
16	4106	-	5,6,7,8,9,?	+
17	644/18 H. Kawisari	XIII	?	+
18	832/2-HT	-	6,7,8,9,?	+
19	H147/1	XIV, XVI	2,3,4,5	-
20	32/1-DK1/6	I,VIII, XII, XIV, XVI, XVII, XXIII,XXIV, XXV, XXVIII, XXXI	2,5	-
21	H152/3	XIV,XVI, XXIII, XXIV, XXVII	2,4,5	-

22	33/1-S.288-23	VII, VIII, XII, XIV,XVI,	3,5	-
23	Caturra (c)	All	5	-
24	Catuaí 2143-236 (c)	All	5	-
25	Mundo Novo -376/4 (c)	All	5	-

*Differential clones were from CIFC (Centro de Investigação das Ferrugens do Cafeeiro, Portugal).
**SH_{1-9} genes as inferred by CIFC (Centro de Investigação das Ferrugens do Cafeeiro, Portugal).
Unknown race (-), coffee genotypes used as negative control (c), presence/absence of CARF005 band (+/-) and unknown S_H gene (s) (?).

3.2 BAC clone identification, sequence assembly and gene prediction

BAC clone 78-K-10 (with ~146 kb insert size) was identified (Appendix II Figure 1B and C), sequenced and assembled into contig 3 (16570 bp) and 9 (8285 bp). The sequences of the two contigs were joined and deposited at NCBI (accession number pending). These contigs shared ≥ 90% identity with *C. canephora* contigs at different chromosome regions with the highest identity (99% for contig 3 and 97% for contig 9) being on chromosome 0. All the 13 ORFs predicted had hits with different species in NCBI by BLASTp or at *C. canephora* genome hub by BLASTn with significant similarities (≤ 1e-05 e-value) (Appendix II Table 1). Among them, five genes (gene 5, 9, 10, 11 and 12) shared identities with RGAs from *C. canephora*. These genes are homologous to sequences in *C. canephora* genome with the highest query coverage being on chromosome 0 (Table 2).

Table 2. Size and structure of five resistance gene analogs and their mapping to chromosome 0 of *C. canephora* genome.

	Genes*				
	5	9	10	11	12
Contig	3	9	9	9	9
Exon 1	3393	113	121	1175	345
Intron 1	-	554	87	611	1786
Exon 2	-	118	112	2222	183
Intron 2	-	121	711	124	-
Exon 3	-	69	121	-	-
Intron 3	-	91	-	-	-
Exon 4	-	155	-	-	-
Intron 4	-	476	-	-	-
Exon 5	-	130	-	-	-
Query coverage (%)	99.94	72.68	30.48	99.46	97.33
Identity (%)	76.00	85.00	79.00	68.84	73.00
E-value	0.00	9,00E-30	5,00E-17	0.00	3,00E-48
Frame	N	N	P	N	P
Start hit-End hit	108638370-108641761	106998076-106999730	107000654-107000761	107000357-107003848	107000234-107004551
Protein (aa)	1130	194	117	1118	175

*Exon and intron sizes are in nucleotides.

N: negative reading frame and P: positive reading frame.

Gene prediction was performed by Augustus command-line version gene prediction (Stanke et al. 2004).

Gene 5 (intron-less, 1130 aa) and 11 (with two introns and two exons, 1118 aa) were the largest genes predicted. Both are positioned in the negative reading frame and belong to the CC-NBS-LRR gene family. Mapping to *C. canephora* genome showed that these genes are separated by 1,634,522 bp though they are delimited with far less distance (460 bp) in *C. arabica*. Contrarily, the other four RGAs retained the positions they are supposed to span in *C. arabica* (Figure 1).

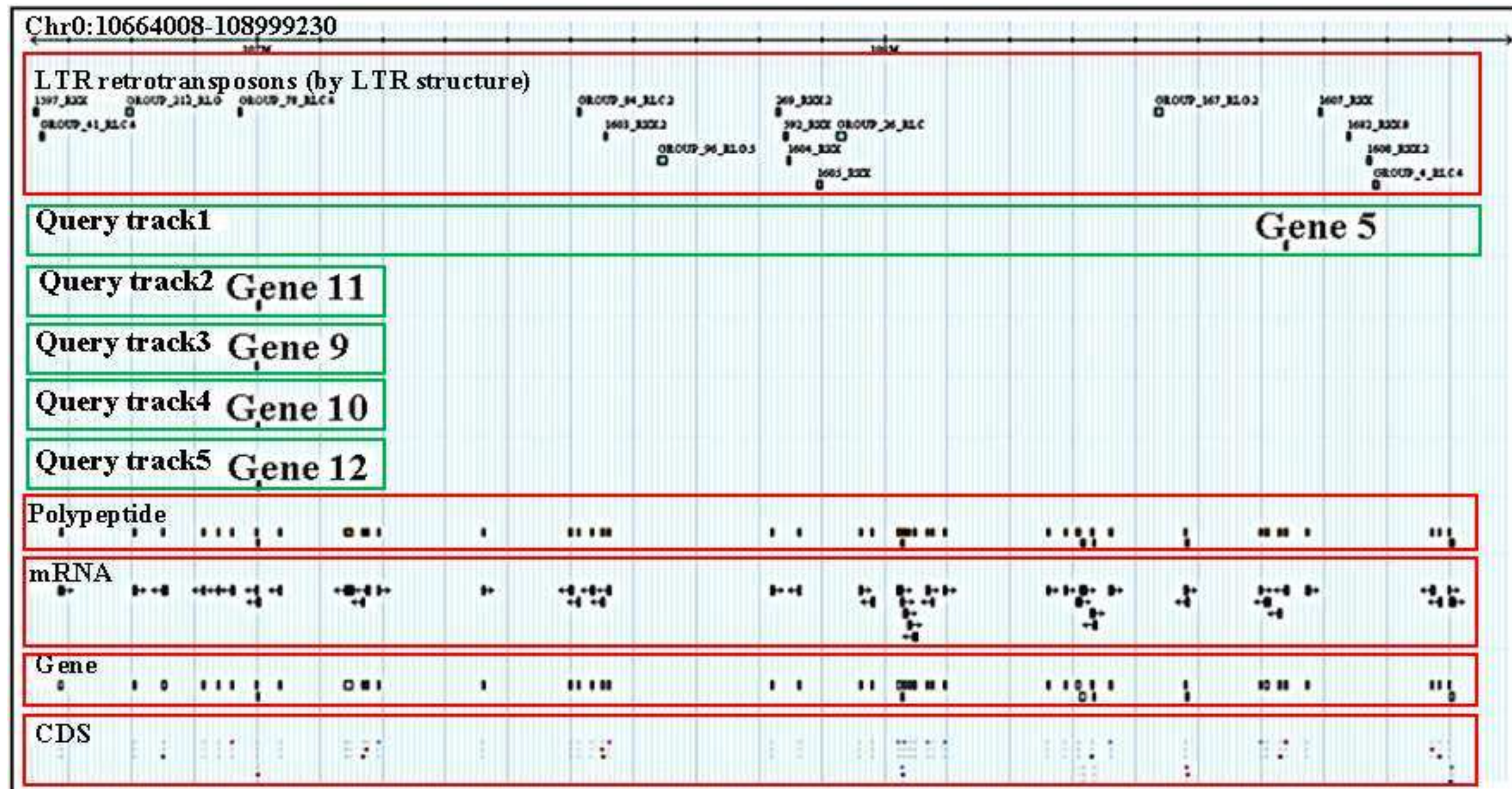


Figure 1. Mapping of five resistance gene analogs clustered on chromosome 0 of *C. canephora*. Putative transcription orientations are shown by black arrows. Green boxes are used to mark query positions relative to subject genes (gene products), all in red boxes. Note that gene 10 and 12 are in positive orientation with no matching transcript here, hence probably originated from different parent or attributed to mutation events in *C. arabica*. Mapping was carried out by CDS (coding sequence) BLASTn followed by track assembly on *C. canephora* genome hub server (Denoeud et al. 2014).

3.4 CARF005 amplicon verification

The size of CARF005 amplicon was 400 bp as shown by *in silico* PCR (http://www.bioinformatics.org/sms2/pcr_products.html) using contig 9 and gene 11 ORF as template strands. This size was confirmed by PCR using 78-K-10 clone as template (Appendix II Figure 1C). The PCR amplicon spans 6867 to 7266 on contig 9 (8285) in negative orientation and 2065 to 3115 on gene 11 ORF (3354 bp).

3.5 Motif search

The upstream sequences of the five RGAs were investigated for their regulatory sequence elements. TATA-box promoter motifs and transcription start sites (TSS) were detected for gene 5 and 10 at different distances upstream. Gene 5 5'UTR and possibly first intron spanned 1298 nts (12210-10913 on the negative strand) while TATA-box and TSS are positioned at 12239-12236 and 1210, respectively. Unlike gene 5, gene 10 ORF is as close as 663 nts to its predicted 5'UTR (2763-3425 on the positive strand) with TATA-box at 2734-2737 and TSS at 2763. The position of TATA-box was detected at the same distance (26 nts) from TSS for both genes.

3.6 Annotation

The annotation of 13 ORFs showed a range of protein arrays most of which have no resistance role and with no conserved domains. Among the five RGAs detected in either NCBI BLASTp, or BLASTn against *C. canephora* genome, gene 9 (unnamed protein product), 10 (putative resistance gene) and 12 (putative resistance gene) have shown similarity with RGAs as shown by mapping to *C. canephora* genome (Appendix II Table 1). Yet, gene 5 and 11 are the largest resistance proteins (Gene 5: 126.81 kDa and pi: 7.65; gene 11: 126.67 kDa and pi: 8.44) identified with several resistance associated GO terms characterizing their multiple functional domains (Table 3).

Table 3. Annotation and functional comparison of gene 5 and 11.

GO ID	GO term	Reliability (%)	Gene 5	Gene 11
Molecular function ontology				
GO:1901363	Heterocyclic compound binding	49	✓	✓
GO:0000166	Nucleotide binding	49	✓	✓
GO:0005488	Binding	49	✓	✓
GO:1901265	Nucleoside phosphate binding	49	✓	✓
GO:0097159	Organic cyclic compound binding	49	✓	✓
GO:0036094	Small molecule binding	49	✓	✓
GO:0097367	Carbohydrate derivative binding	41	✓	✓
GO:0017076	Purine nucleotide binding	41	✓	✓
GO:0032559	Adenyl ribonucleotide binding	41	✓	✓
GO:0032555	Purine ribonucleotide binding	41	✓	✓
Biological process ontology				
GO:0006952	Defense response	36	✓	✓
GO:0006950	Response to stress	36	✓	✓
GO:0050896	Response to stimulus	36	✓	✓
GO:0002376	Immune system process	16	✓	✓
GO:0006955	Immune response	16	✓	✓
GO:0045087	Innate immune response	16	✓	✓
GO:0044699	Single-organism process	14	✓	✓
GO:0009987	Cellular process	14	✓	✓
GO:0044763	Single-organism cellular process	14	✓	✓
GO:0033554	Cellular response to stress	12	✓	✓
GO:0016265	Death	12	✓	✓
GO:0051716	Cellular response to stimulus	12	✓	✓
GO:0012501	Programmed cell death	12	✓	✓
GO:0008219	Cell death	12	✓	✓
GO:0034050	Host programmed cell death induced by symbiont	12	✓	✓
GO:0009626	Plant-type hypersensitive response	12	✓	✓
GO:0009814	Defense response, incompatible interaction	7	✓	✓
Cellular component ontology				

GO:0016020	Membrane	33	✓	✓
GO:0044464	Cell part	33	✓	✓
GO:0005623	Cell	33	✓	✓
GO:0005737	Cytoplasm	32	✓	✓
GO:0044424	Intracellular part	32	✓	✓
GO:0005886	Plasma membrane	31	✓	✓
GO:0071944	Cell periphery	31	✓	✓
GO:0043227	Membrane-bounded organelle	24	✓	✓
GO:0043226	Organelle	24	✓	✓
GO:0005634	Nucleus	24	✓	✓

Annotation was performed by Predict Protein online server (Yachdav et al. 2014) (URL: https://www.predictprotein.org).

3.7 Gene characterization

The two contigs were BLASTed against the transcriptome to identify the candidate R genes activated against *H. vastatrix* incursion. Gene 9, 10, 11 and 12 were mapped to transcripts differentially expressed at 12 and 24 h.a.i. during *C. arabica -H. vastatrix* (race XXXIII) incompatible interaction (Figure 2).

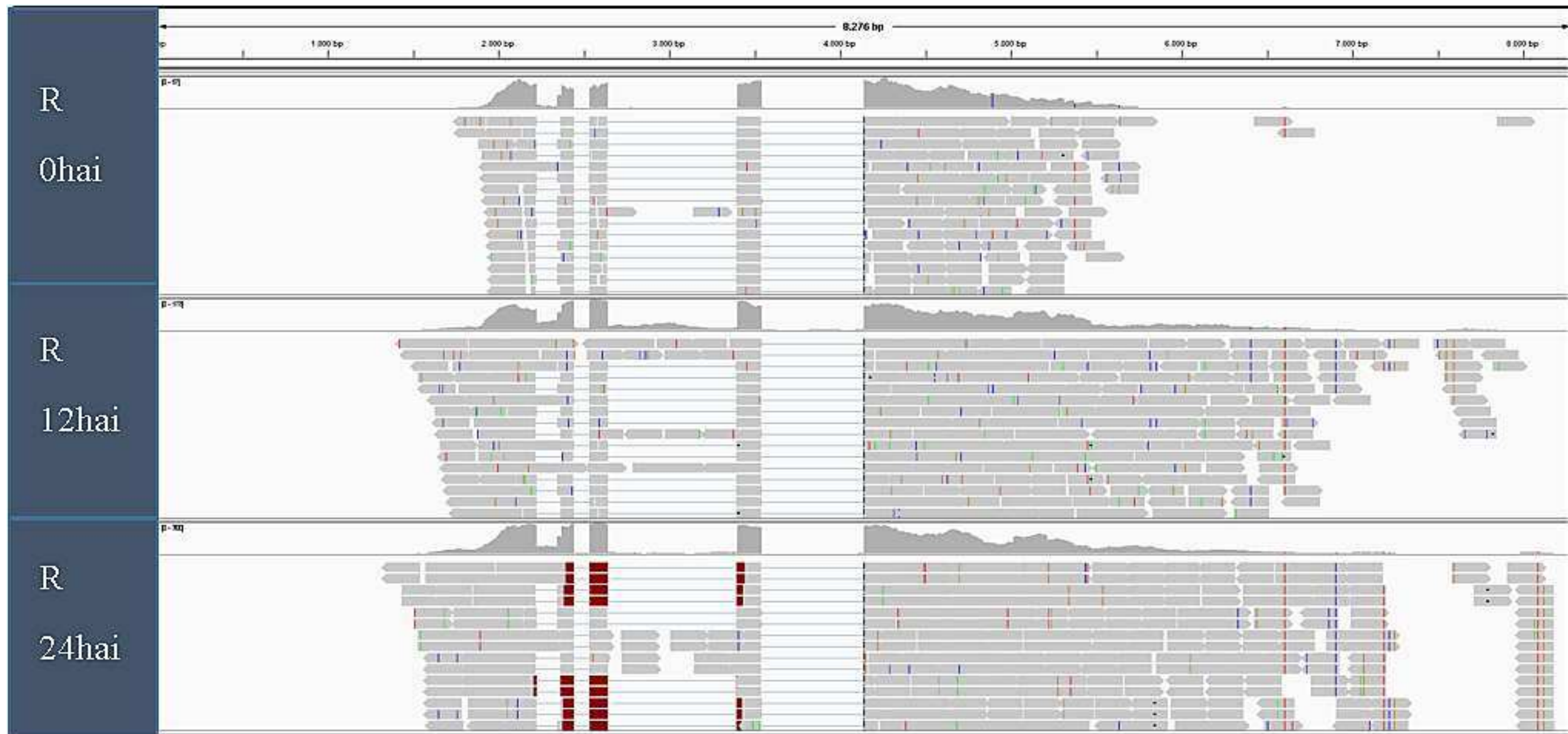


Figure 2. Mapping of contig 9 to transcriptome of differentially expressed genes during *C. arabica-H. vastatrix* (race XXXIII) incompatible interaction to show the region of active gene (gene 9, 10, 11 and 12) expression. Note the three expression profiles (three rows) corresponding to control (uninoculated at 0 hour, top row), 12 (middle row) and 24 h.a.i. (bottom row) of *H. vastatrix*. R (resistant cultivar, CIFC HDT 832/2), hai (hours after pathogen inoculation). Gray shades indicate matching transcriptome reads while nucleotide substitutions (mismatches) were shown by colored strips (yellow: G, green: A, red: R and blue: C). Large red shades indicate deletions. Contig mapping was performed by Tophat 2 (D. Kim et al. 2013) and visualized with Integrative Genomics Viewer (IGV) v. 2.3 (Robinson *et al.*, 2011).

We did an estimation of contig 9 region where most R genes are positioned. The result showed that ~81.58% of the contig encodes rust resistance variant transcripts, which are activated at 12 and 24 h.a.i. in response to *H. vastatrix* inoculation. Further analysis of gene 5 and 11 showed that these genes belong to NBS-LRR gene family. Both have the Rx-cc-like coiled-coil potato virus x resistance protein domain and four more multi-domains, featuring the entire protein sequence (Figure 3). Both genes can be referred as CC-NBS-LRR, as they are composed of the N-terminal CC and LRR C-terminal domains flanking the NBS in either side.

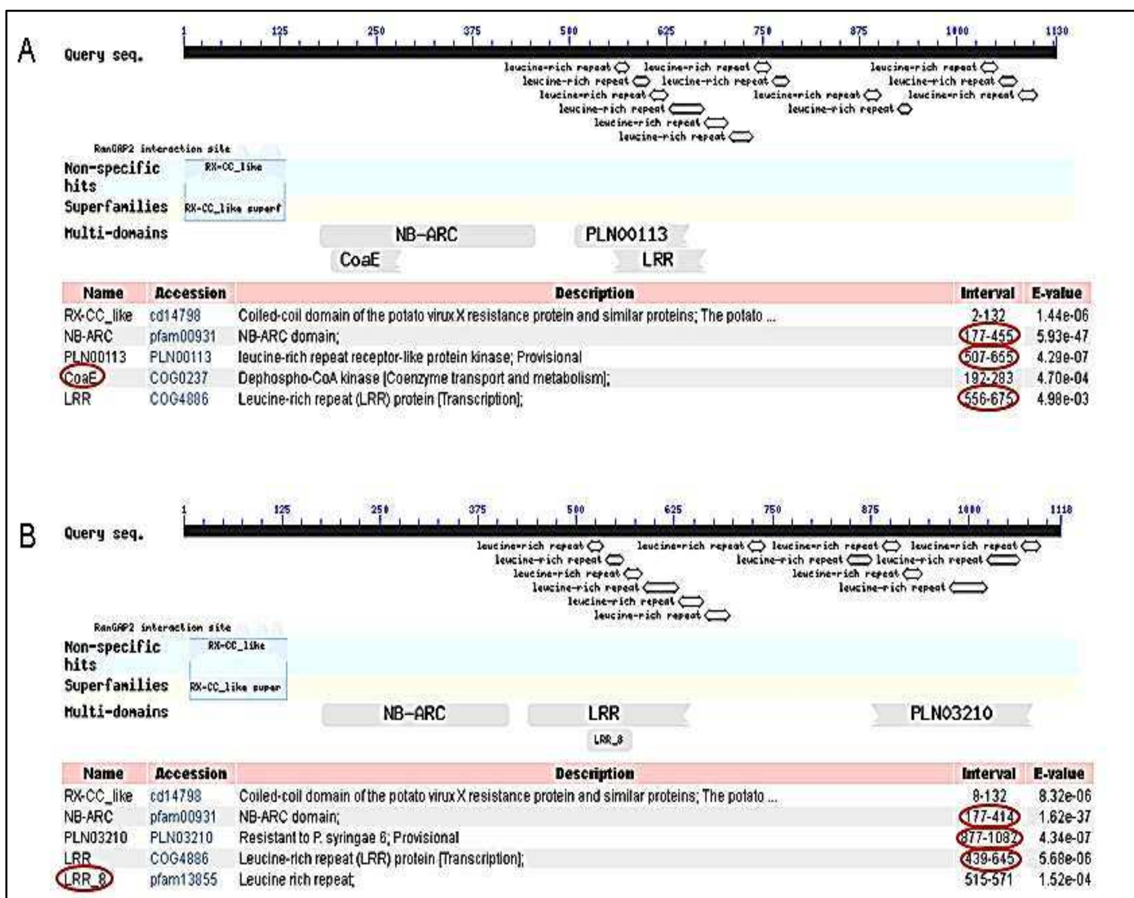


Figure 3. Conserved domain and motif architecture comparative illustration in gene 5 and 11. Note the polymorphism of four domains in both genes. Conserved domains were detected by NCBI BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins).

Annotation of both genes showed that they encode defense proteins involved in various types of defense biological processes (Table 3). Both are intracellular resistance proteins that directly or indirectly recognize pathogen effector proteins and subsequently trigger a

response that may be as severe as localized cell death. Though these genes share 90.24% nucleotide identity, their amino acid sequence identity is just 80.03% (Figure 4). The attribution of indel, gene duplication and substitution mutation events were also investigated for their possible role in diverging the genes. Gene duplication was not detected while indel events had no major role but overall non-synonymous substitution events (non-synonymous/synonymous ratio, ka/ks = 1.5913) were detected in both genes. Further analysis of LRR region showed higher rate of non-synonymous substitution mutation (ka/ks, non-synonymous/synonymous substitution ratio = 1.9660).

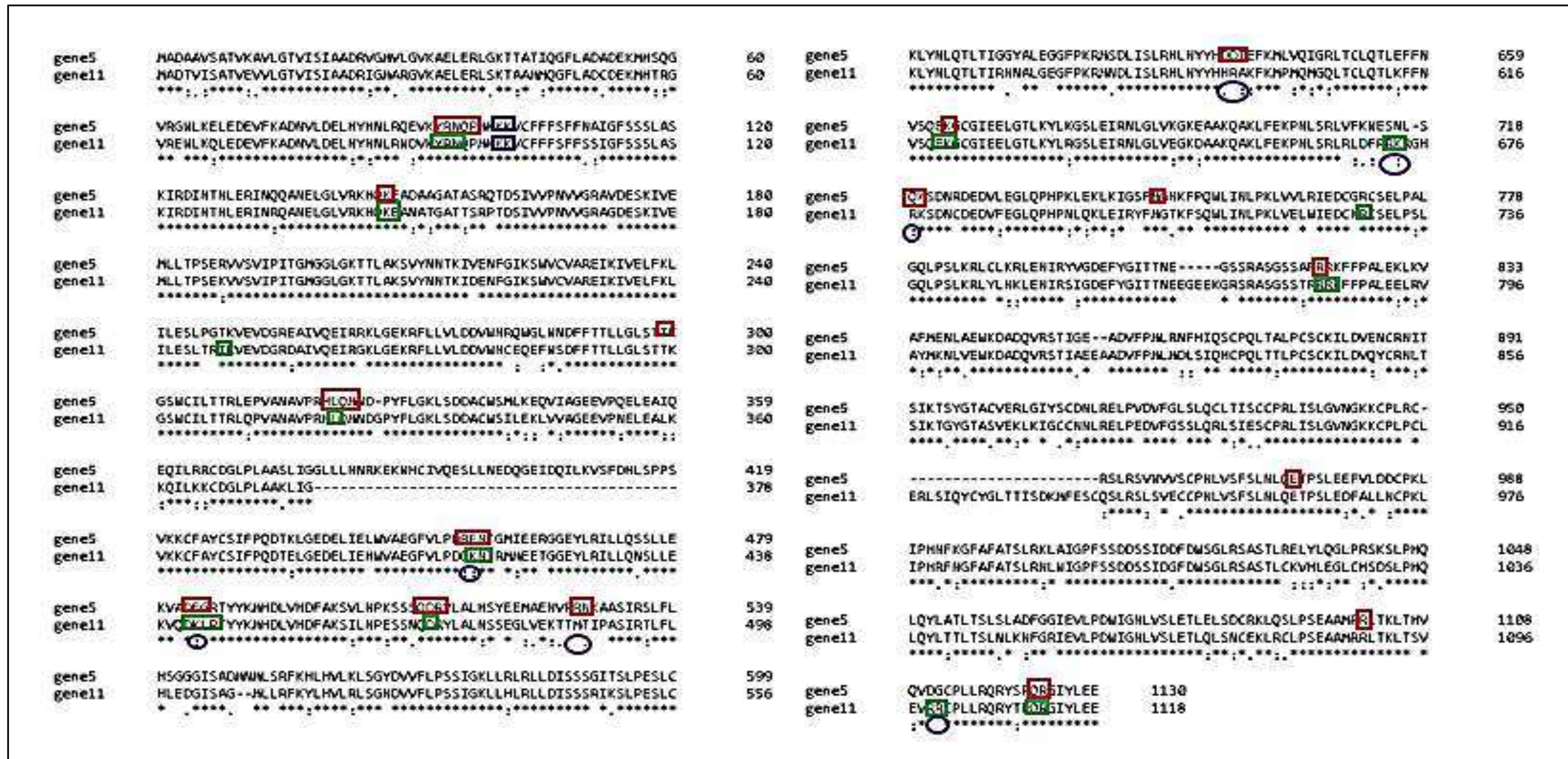


Fig. 4. Alignment of gene 5 and 11 encoded proteins and their detected protein binding regions. *In silico* prediction of protein binding regions of gene 5 (boxed in red), protein binding regions of gene 11 (boxed in green) and conserved protein binding region (boxed in dark blue). Amino acid substitution: unrelated amino acid substitution (space), weakly similar substitution (period), strongly similar substitution (colon) and conserved amino acids (star, unmarked). Note the seven substitution mutations resulting in seven protein binding site polymorphisms (purple encircled) in either of the sequences. Sequence alignment was carried out by Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).

3.8 Structural and functional site comparative analysis

The number of LRR domains in gene 5 and 11 are conserved (13 repeats in both) but differently arranged, with the introduction of coenzyme domain (CoaE, dephospho-CoA kinase) in gene 5 while an LRR variant (LRR_8) is introduced into gene 11(Figure 3). Despite sharing similar protein multi-domains, two transmembrane motifs (spanning 5-22 and 102-119 regions) were exclusively detected in the coiled-coil domain of gene 11 (not shown in Figure 3). The amino acid sequences of gene 5 and 11 were further analyzed for their protein and nucleotide binding site polymorphisms. Protein binding sites of the two genes showed different sensitivity to substitution mutation. Few sites specific to each gene are highly affected while most sites show moderate to no effect (Appendix II Table 2. There are 17 protein binding sites in gene 5, whereas 14 sites are featuring gene 11. Similarly, their secondary structure and solvent accessibility show shared features (Figure 5AI-IV and CI-IV). Nevertheless, the amino acid residues forming the protein binding sites of the two genes showed high variability in the LRR regions (Figure 4 and Figure 5A and C). On the other hand, ADP binding sites of the NBS domain show some shared sites while most of the residues are not conserved (Figure 5B-II and D-II).

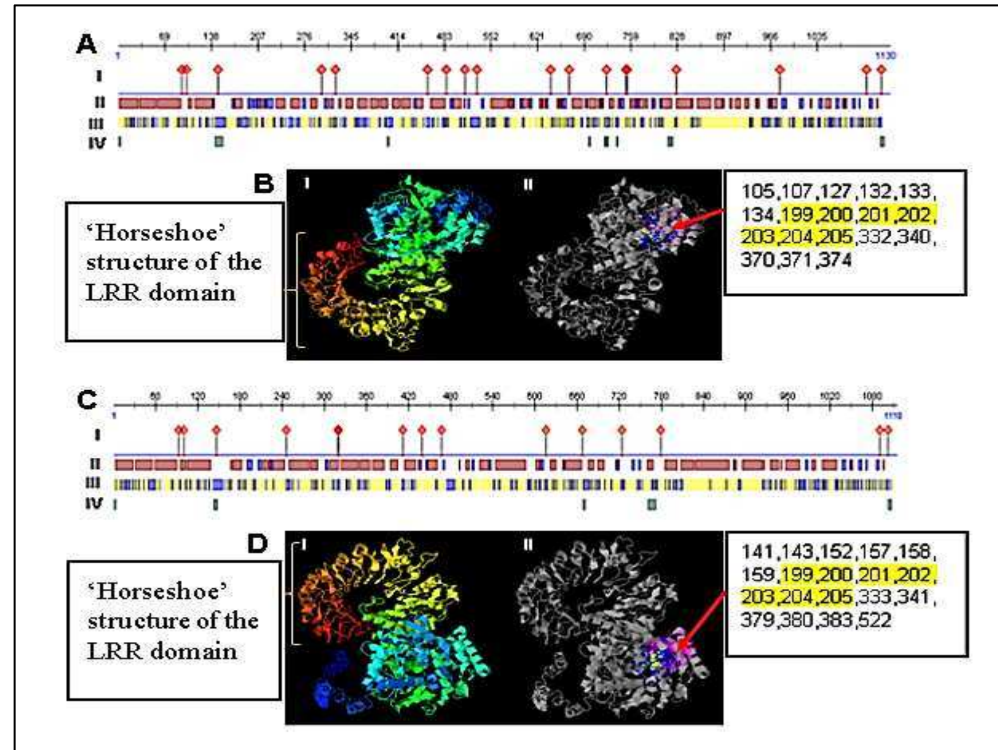


Figure 5. *In silico* 3D structure, protein and nucleotide binding site prediction for gene 5 (A and B) and 11 (C and D). Protein binding and secondary structure (A and C): Protein binding sites (I), the three types of secondary structures assumed at different regions (helical: red boxes, strand: blue boxes and loop: intervening white spaces) (II), solvent accessibility (exposed: blue boxes, buried: yellow boxes and intermediate: white spaces) (III) and high protein disorder and flexibility: green boxes (IV). 3D structure and nucleotide binding sites (B and D): 3D structures with Rx-CC-like (blue) to LRR (red to light: orange forming the 'horseshoe' structure) domains as shown in Figure 3B-I and D-I and colored residues (NBS) forming the nucleotide (ATP/GTP/ADP/GDP) binding site (BII and DII). Nucleotide binding site residues with the highest C-score were listed in the right box (conserved residues highlighted in yellow) with the red arrow pointing the sites. I-TASSER modeling C-score (Yang *et al*., 2014) was -1.73 and -1.75 (C-score ranging from -5 to 2, where 2 is with the highest confidence) and 0.29 and 0.17 (C-score ranging from 0-1, where higher score indicates reliable prediction) for nucleotide binding prediction for the two proteins, respectively.

3.9 Interlocus comparison of S_H genes

To investigate the shared identities between other S_H gene locus and any of the five RGAs (gene 5, 9, 10, 11 and 12), contig 3 and 9 were BLASTed against three *C. canephora* and 10 *C. arabica* specific BAC clones spanning S_H3 locus from the work of Ribas *et al.* (2011). All the ten S_H3 contigs had hits with contig 3 with different alignment length and identities (Appendix II Table 3). Contig GU123898 and HQ696508 (both specific to *C. arabica*) produced hits against contig 9 (from which four clustered RGAs were predicted) with alignment lengths of 170 nts (77.647% identity and 1.57e-31 e-value) and 179 nts (76.536% identity and 1.21e-26 e-value), respectively. The closest contig (HQ696508) is positioned on the complementary strand to gene 11, 505 bp away from where CARF005 forward primer annealed to gene 11.

3.10 Phylogenetic analysis

Two resistance gene families (the NBS-LRR and non-NBS-LRR) were identified, completely sequenced and mapped to chromosome 0 of *C. canephora* genome with query coverage of 99.94% for genes 5 and 11, 72.68% for gene 9, 33.05% for gene 10 and 97.52% for gene 12. To analyze the occurrence of gene recombination/conversion among the NBS-LRR families since their last divergence from their common ancestor (*C. canephora*), the best two hits (hit1 and hit2) against gene 5 and 11 (all from *C. canephora*) were analyzed. Multi-site gene recombination was detected in all the six genes by RDP4 detection method (Figure 6A).

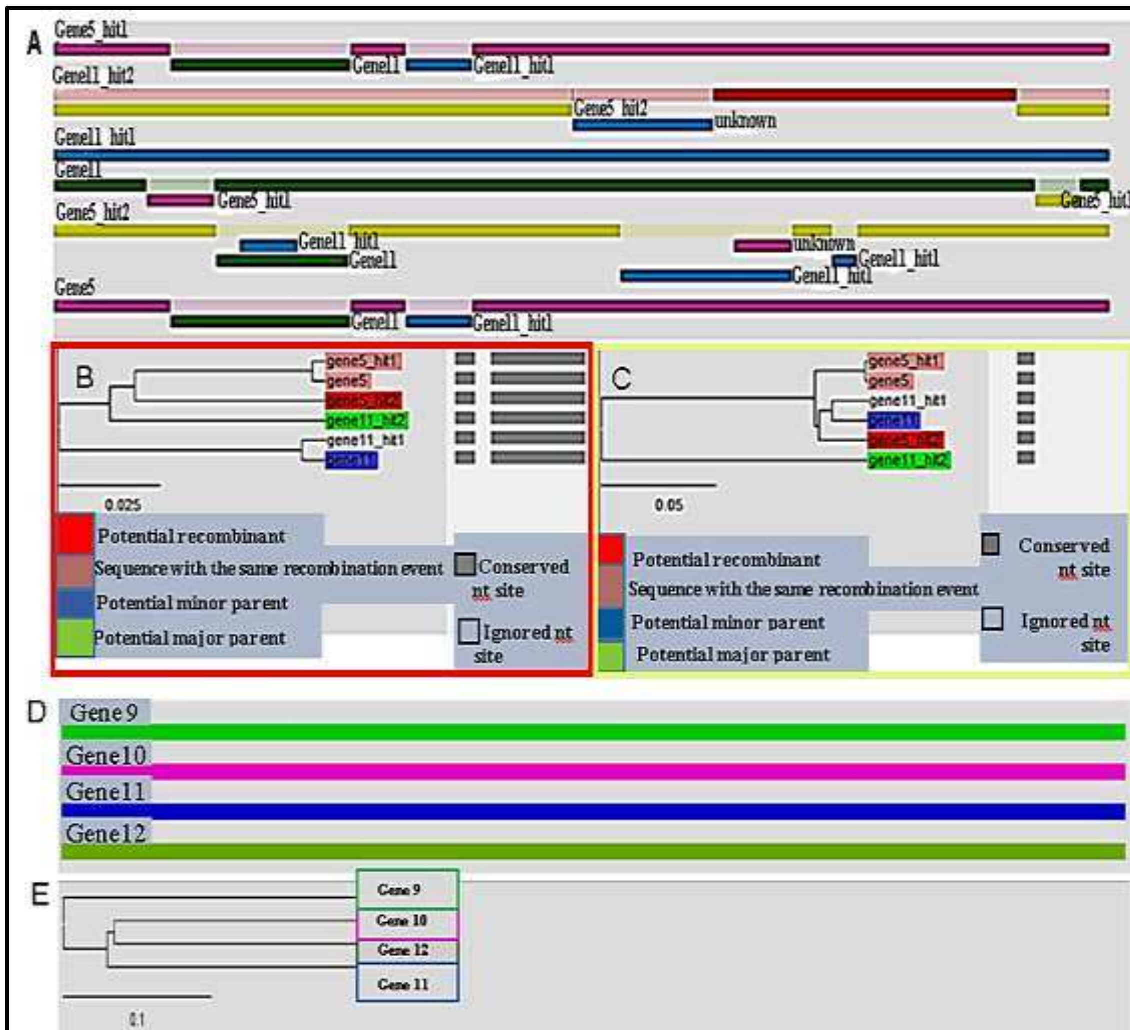


Figure 6. Gene conversion/recombination analysis among homologous genes. Gene recombination events were detected among gene 5 and 11 and their best two hits from *C. canephora* shown with sites of gene recombination (A), UPGM of regions derived from major parent (1-568 and 1033) (B), UPGM of regions derived from minor parent (569-1032) (C). Phylogenetic tree of four RGAs (9,10,11 & 12) (D) and their intact gene integrity with no recombination detected (E). Recombinant detection methods employed are the different versions of RDP software, all complementing one another at $p<0.05$.

A remarkable event was the introgression of a reasonable portion of gene 11 into gene 5 and the recombination/conversion event in the LRR region of gene 11 (Figure 6A). Parental lineage analysis indicated that gene 11 formed potential minor parent while gene11_hit2 attributed to major parent in both major and minor parent analysis (Figure 6B and C). Contrarily, there was no gene recombination/conversion event among the three non NBS-LRR RGAs (gene 9, 10, 12) and gene 11, even though they are tightly linked (Figure 6D and E). Another supporting evidence for the diversity of the NBS-LRR family was detected by ka/ks ratio analysis (Table 4). The analysis showed that non-synonymous substitution event is intensively operating in CDS, in all pairwise analysis. When LRR region is exclusively considered, non-synonymous substitution of CDS is more intensively operating (ka/ks in almost all pairwise comparisons) (Table 4).

Table 4. Pair-wise synonymous and non-synonymous nucleotide substitution analysis among the six resistance gene analogs (gene 5 and 11 and their respective two top hits as mined by BLASTn in NCBI).

		Entire protein			LRR region		
Seq. 1	Seq. 2	Ks	Ka	ka/ks	Ks	Ka	ka/ks
gene5_hit1	gene11_hit1	0.0786	0.1302	1.6565	0.0702	0.1383	1.9701
gene5_hit1	gene11_hit2	0.0899	0.1614	1.7953	0.0536	0.1408	2.6269
gene5_hit1	gene11	0.0723	0.1177	1.6279	0.0622	0.1233	1.9823
gene5_hit1	gene5_hit2	0.0635	0.0999	1.5732	0.0583	0.1029	1.7650
gene5_hit1	gene5	0.0009	0.0039	4.3333	0.0015	0.0045	3.0000
gene11_hit1	gene11_hit2	0.1092	0.1854	1.6978	0.0756	0.1602	2.1190
gene11_hit1	gene11	0.0061	0.0164	2.6885	0.0095	0.0170	1.7895
gene11_hit1	gene5_hit2	0.0761	0.1309	1.7201	0.0736	0.1369	1.8601
gene11_hit1	gene5	0.0786	0.1288	1.6387	0.0701	0.1383	1.9729
gene11_hit2	gene11	0.1074	0.1742	1.6220	0.0686	0.1445	2.1064
gene11_hit2	gene5_hit2	0.0846	0.5829	6.8901	0.0607	0.1440	2.3723
gene11_hit2	gene5	0.0902	0.1620	1.7960	0.0540	0.1430	2.6481
gene11	gene5_hit2	0.0704	0.1155	1.6406	0.0616	0.1199	1.9464
gene11	gene5	0.0723	0.1164	1.6100	0.0622	0.1234	1.9839
gene5_hit2	gene5	0.0635	0.0997	1.5701	0.0583	0.1052	1.8045

Seq.: sequence, non-synonymous/synonymous substitution rate was computed by DnaSP v5.1 (Rozas 2009).

Moreover, phylogenetic analysis showed that tomato gene 5 was closely related to gene 5 and 11 of coffee than gene 11 of both tomato and grape (Figure 7). Within coffee itself,

a significant diversity between gene 5 and 11 was detected by MEGA 7 bootstrap method test of phylogeny.

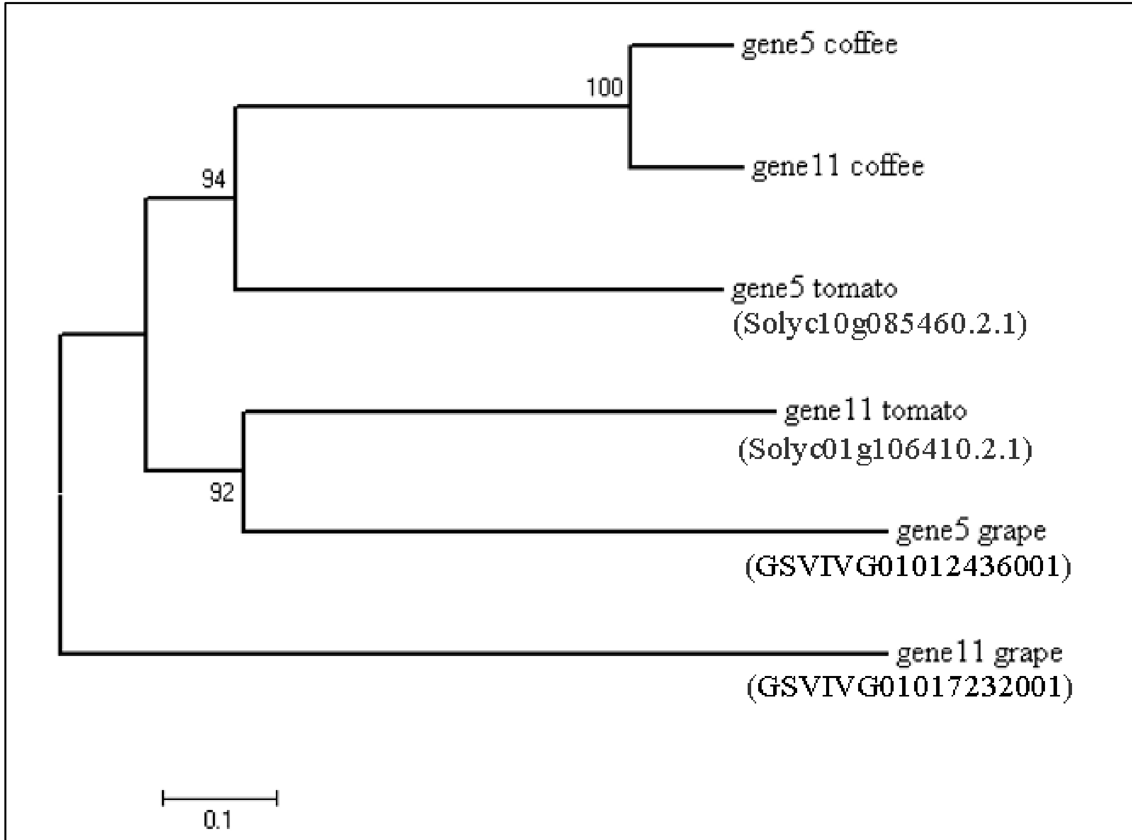


Figure 7. Phylogenetic history of gene 5 and 11 in three related genomes. The evolutionary history was inferred using the Minimum Evolution method (Rzhetsky and Nei 1992). The optimal tree with the sum of branch length = 2.98805978 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965) and are in the units of the number of amino acid substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm (Nei and Kumar 2000) at a search level of 1. The Neighbor-joining algorithm (Saitou and Nei 1987) was used to generate the initial tree. The analysis involved 6 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 554 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al. 2016). Subject IDs are indicated in parenthesis for the corresponding two homologous sequences in tomato (Sol Genomics Network) and grape (Phytozome) genome databases.

4. DISCUSSION

It has been understood for long that the majority of NBS-LRR encoding genes are clustered and unevenly distributed in plant genomes (Hulbert *et al.*, 2001; Meyers *et al.*, 2005; Hammond-Kosack and Kanyuka, 2007). NBS-ARC domain is known to be involved in directly blocking biotrophic pathogens by the activation of hypersensitive response (HR) (Mur et al. 2008). HR begins with a programmed cell death of affected and surrounding cells and ends with the activation systemic acquired resistance (SAR), in which defense is induced in distal non-infected cells of the host under attack (Sanabria *et al.*, 2008; Qi and Innes, 2013). By recognizing the corresponding pathogen virulence (vr) factors or their effects, NBS-LRR proteins are sufficient to induce HR (McHale et al. 2006; Jones and Dangl 2006; Rairdan *et al.*, 2008; Qi and Innes, 2013). We report a cluster of two different classes of RGAs resistant to coffee rust, the NBS-LRRs linked to non-NBS-LRR genes. The size of the two NBS-LRR genes (gene 5 and 11) is the largest non-TNL genes sequenced in Arabica coffee and most other plants investigated to date (Kim *et al.*, 2009; Ratnaparkhe *et al.*, 2011; Ribas *et al.*, 2011; Djebbi *et al.*, 2015).

Here, we report a completely sequenced and characterized novel RGA (gene 11), probably a major component of an S_H gene. Looking into compatibility data, this gene seemed to be activated against v6 protein of *H. vastatrix* (with exception in one differential clone). The discovery of Híbrido de Timor (HDT) (*C. arabica* x *C. canephora*), immune to all known virulence factors of *H. vastatrix* physiological races (Bettencourt, 1973), is an extremely important source of resistance as far as coffee leaf rust is concerned. Since then, many differential hosts of coffee species were characterized for their susceptibility to specific physiological races (Bettencourt AJ, 1981; Bettencourt AJ, 1988). In line with the gene-for-gene theory (Mayne 1932), every differential host with a single or multiple S_H genes is compatible to the corresponding virulence factor of a pathogen race for susceptibility to be revealed. Additionally, three more RGAs (gene 9, 10 and 12), along with several resistance gene variants, which are clustered and co-expressed against *H. vastatrix* were unveiled as seen by contig 9 mapping to the reference transcriptome. Also, mapping showed that, there are reads exclusively mapped to transcripts of pathogen infected plant at 12 and 24 hai. These variant transcripts are probably due to alternative splicing to cope up to the evolving pathogen effectors. The

prevalence and importance of alternative splicing to create multiple mRNA transcripts in response to biotic stresses has been reported in other plants (Gou *et al.* 2010; Syed *et al.*, 2012; Wang *et al.*, 2012; Saminathan *et al.*, 2015).

The main rust resistance genes, S_H3 in *C. liberica* (Noronha-Wagner and Bettencourt 1967), S_H6, 7, 8 and 9 in *C. canephora* (Bettencourt and Rodrigues 1988) and S_H1, 2, 4 and 5 genes in *C. arabica* are understood as dominantly inherited genes (Bettencourt and Coronha-Wagner 1971). One of the fundamental questions to be investigated is how distinct are these 9 S_H genes as they belong to different coffee species. The comparative analysis of contigs from S_H3 locus of *C. arabica* and *C. canephora* (Ribas et al. 2011) showed different levels of shared identity with our two contigs, contig 3 and 9. The results indicate that RGAs may share large conserved regions but few and highly polymorphic regions encoding specific protein motifs of critical role. This characteristic conserved domain among S_H gene loci was once more confirmed based on comparative analysis of the cloned gene (S_H10) with S_H6 using differential coffee clones currently used for S_H gene identification. The PCR amplification of S_H10 gene (gene 11) indicates that this new gene may share a conserved domain (on which CARF005 primer was designed) with S_H6. Also, we report this same new S_H gene (S_H10) in 644/18 H. Kawisari differential clone for the first time. Overall, we leave the following hypothesis to be extensively and rigorously investigated: the discovered gene (S_H10, gene 11) is one of the unidentified and not yet supplanted (at least in Brazil) S_H gene in HDT with a conserved domain (CARF005) shared with S_H6 gene.

Mapping of the RGAs to *C. canephora*, the result from differential clone screening and annotation altogether confirm that the new gene (S_H10) locus is descended from *C. canephora*. The gene (S_H10) presented in this work is a sibling of S_H6-9 originating from *C. canephora* (Bettencourt and Rodrigues,1988). The disparity of gene 5 position in relation to the other genes (gene 9, 10, 11 and 12) could be attributed to linked LTR-retrotransposons and transposase gene (gene 1). Multiple transposable elements linked to NBS-LRR regions were reported in other plants (Kang et al. 2012; Ratnaparkhe et al. 2011). Transposition of genes and gene fragments are some of the mechanisms that generate variability and positional change among the NBS-LRR genes in different plants

(Ratnaparkhe *et al.*, 2011; Kang *et al.*, 2012; González *et al.*, 2014; Sanseverino *et al.*, 2015; Panchy *et al.*, 2016).

Rx-CC, PLN and NB-ARC domains are conserved in NBS-LRR genes across diverse plant species (van der Biezen and Jones, 1998; Kim *et al.*, 2009; Wang *et al.*, 2015). The potato virus x resistance (Rx) protein-like N-terminal coiled-coil domain mediates intramolecular interaction with NB-ARC and intermolecular interaction through RanGAP2 (Ran-GTPase-activating protein-2) in potato (Rairdan *et al.*, 2008; Hao *et al.*, 2013). Rx-CC, RanGAP2 interaction site and NB-ARC were detected in gene 5 and 11, suggesting their similar role in coffee. However, unlike the Rx-CC domain with four helical structures, five helical structures are conserved in gene 5 and 11, indicating their interspecies polymorphisms. PLN00113 domain in gene 5 and PLN03210 in gene 11, spanning the LRR region were initially reported in *A. thaliana* (Kim et al. 2009). The distinct position of these domains in gene 5 and 11 indicates the high variability of LRR region in both genes. Given SMART motif analysis detection of TM motif in the Rx-CC domain of gene 11 protein, in addition to the functional motif prediction, the PLN03210 (LRR domain) is likely engaged in direct effector interaction while the corresponding PLN00113 of gene 5 is engaged in LRR-reception and downstream kinase mediated signaling. This result is in accordance with functional and structural analysis of LRR proteins in *A. thaliana* (Lahaye, 2002; Kierszniowska *et al.*, 2009; Kim *et al.*, 2009; Gou *et al.*, 2010).

Different selection pressures are shaping the evolution of each domain in the NBS-LRR encoding genes. The NBS domain was assumed to be under purifying selection (a negative selection in which variation is minimized by stablizing selection) than diversifying selection, which is acting on the LRR domain (Michelmore and Meyers, 1998; McHale *et al.*, 2006). By contrary, diversifying selection (positive selection) is acting on all domains of gene 5 and 11 (ka/ks >1). This result is contrary to the general assumption that diversifying selection is diluted when overall non-synonymous substitution is taken into account (Ribas et al. 2011), indicating a fierce diversifying selection action on both genes. Further investigation of four more orthologous genes also resulted in similar results, indicating these NBS-LRR genes are highly variable due to

substitution mutation. As the LRR domains are involved in direct ligand binding, their variability due to non-synonymous substitution is more than the other domains, resulting in super polymorphic region to cope up to the continosly evolving pathogen effectors. Similar findings (from different plants including coffee) and reviews were made on diversifying selection (DeYoung and Innes, 2006; McHale *et al.*, 2006; Hammond-Kosack and Kanyuka, 2007; Padmanabhan *et al.*, 2009; Ratnaparkhe *et al.*, 2011; Ribas *et al.*, 2011; Zhao *et al.*, 2016). Diversifying selection by nonsynonymous substitution was also detected in non-NBS-LRR genes (gene 10 and 12) (data not shown), reiterating the importance of substitution mutation in these clustered R genes. Synergistic activation of the two groups (NBS-LRR and non-NBS-LRR) may enhance resistance durability, hence their expression pattern needs further investigation.

While NBS-LRR encoding genes are considered as a reservoir for the continuous evolution of R gene variants due to the coevolving pathogen effectors, the specific evolutionary route taken is mostly unknown (Hulbert *et al.*, 2001; Ribas *et al.*, 2011). Sequence exchange, gene duplication and gene conversions were reported to generate variations among the NBS-LRR encoding genes, noticeably in the LRR domain (Mondragon-Palomino, 2002; McHale *et al.*, 2006; Ratnaparkhe *et al.*, 2011; Ribas *et al.*, 2011). However, sequence duplication was not detected in any of the six orthologous genes analyzed in the present work, while gene conversion/sequence exchange was detected in all. Of intriguing result was the gene recombination/conversion event in the LRR of gene 11, which might have led to the functional divergence from gene 5, with which the highest CDS identity is shared. Gene recombination/conversion in the LRR domain combined with point mutation and positive selection are the main deversifying events for gene 5 and 11, similar to other reports on other genes (Guo *et al.*, 2011; Ratnaparkhe *et al.*, 2011; Ribas *et al.*, 2011; Jacob *et al.*, 2013; Nepal and Benson, 2015; Choi *et al.*, 2016). According to Michelmore and Meyers (1998), R genes tend to be organized in clusters and their high recombination rate with the component genes is the key to determine resistance specificity. In the present work, unequal segment exchanges were also detected among orthologous genes from *C. canephora* and the two genes, indicating the importance of gene recombination/conversion event before they assorted into different species and specific resistance functions. Sequence exchange by haplotype gene conversion and non-homologous recombination between different genomes of

clustered CNL genes is a common phenomenon in plants (Kuang *et al.*, 2005; Mondragon-Palomino and Gaut, 2005; Ribas *et al.*, 2011). In the present work, the gene conversion/recombination event detected in all the homologs indicates that it might have occurred before speciation and conserved across the three genomes.

From the phylogenetic tree of orthologous genes from related genomes, the six genes could be divided into two groups. Gene 5 from tomato is more related to gene 5 and 11 from coffee, making the first group, whereas grape gene 5 and 11 from the second group are highly diversified. Intraspecies diversity of non-TIR-NBS-LRR due to substitution and genetic recombination were reported in grape (Velasco et al. 2007) and tomato (Sara et al. 2012) while gene duplication and conversion events were inferred in coffee (Ribas et al. 2011). In general, the phylogenetic tree showed that gene 5 and 11 have recently diverged in coffee while the divergences in the other species were earlier events.

We conclude that the two groups of RGAs, NBS-LRR and non-NBS-LRR, are clustered on a single locus from which multiple variants of resistance genes are expressed to confer a specific resistance function. The two CC-NBS-LRR protein encoding genes are under strong diversifying selection acting on all the component domains. More intense LRR region diversification indicates that effector binding site variability is the main force for the divergence of resistance specificity. The four cloned, sequenced and characterized RGAs span a new S_H gene locus (likely SH_{10}) descended from *C. canephora*. Therefore, this work provides a practical application of functional marker for marker assisted selection breeding by developing new sets of markers from the R gene locus identified by mapping to rust resistance transcriptome. Such work can also be applied in molecular breeding as it has a potential in replacing arbitrary DNA marker assisted breeding at least for two reasons. The first and straight forward is that there is no probability to loss due to segregation, which is the case even for finely saturated markers. Secondly, four of the RGAs (gene 9, 10, 11 and 12) are stacked at a locus, from which different primers can be designed to screen genotypes for co-segregation analysis of these genes.

5. REFERENCES

Alvarenga, M.S., Caixeta, T.E., Hufnagel, B., Thiebaut, F., Maciel-Zambolim, E., Zambolim, L. and Sakiyama, S.N. (2010) In silico identification of coffee genome expressed sequences potentially associated with resistance to diseases. *Genet. Mol. Biol.* **33**, 795–806.

Alwora, G.O. and Gichuru, E.K. (2014) Advances in the Management of Coffee Berry Disease and Coffee Leaf Rust in Kenya. *J. Renew. Agric.* **2**, 5.

Bankevich, A., Nurk, S., Antipov, D., et al. (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **19**, 455–477.

Bettencourt, A.J. and Coronha-Wagner (1971) Genetic factors conditioning resistance of Coffea arabica L. to Hemileia vastatrix Berk et Br. *Agron. Lusit.* **31**, 285–292.

Bettencourt, J.. (1973) Considerações gerais sobre o “Híbrido de Timor” 1st ed., Sao Paulo: Instituto Agronômico.

Bettencourt AJ, R.C. (1981) Melhoramento genético do cafeeiro - Transferência de fatores de resistência à Hemileia vastatrix Berk et Br. para as principais cultivares de Coffea arabica L 1st ed., Lisboa.

Bettencourt and Rodrigues (1988) Principles and practice of coffee breeding for resistance to rust and other diseases. *Elsevier Appl. Sci.* **3**, 199–234.

Biezen, E.A. van der and Jones, J.D. (1998) The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* **8**, R226-7.

Biezen, E.A. Van der and Jones, J.D. (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* **23**, 454–6.

Cação, S.M.B., Silva, N. V., Domingues, D.S., et al. (2013) Construction and characterization of a BAC library from the Coffea arabica genotype Timor Hybrid CIFC 832/2. *Genetica* **141**, 217–226.

Cenci, A., Combes, M.-C. and Lashermes, P. (2010) Comparative sequence analyses

indicate that Coffea (Asterids) and Vitis (Rosids) derive from the same paleo-hexaploid ancestral genome. *Mol. Genet. Genomics* **283**, 493–501.

Choi, K., Reinhard, C., Serra, H., et al. (2016) Recombination Rate Heterogeneity within Arabidopsis Disease Resistance Genes Michelmore, R., ed. *PLOS Genet.* **12**, e1006179.

Combes, M.-C., Lashermes, P., Chalhoub, B., Noir, S. and Patheyron, S. (2004) Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (Coffea arabica L.). *TAG Theor. Appl. Genet.* **109**, 225–230.

Cristancho, M.A., Botero-Rozo, D.O., Giraldo, W., et al. (2014) Annotation of a hybrid partial genome of the coffee rust (Hemileia vastatrix) contributes to the gene repertoire catalog of the Pucciniales. *Front. Plant Sci.* **5**.

Denoeud, F., Carretero-Paulet, L., Dereeper, A., et al. (2014) The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science (80-.).* **345**, 1181–1184.

DeYoung, B.J. and Innes, R.W. (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat. Immunol.* **7**, 1243–1249.

Diniz, L.E.C., Sakiyama, N.S., Lashermes, P., Caixeta, E.T., Oliveira, A.C.B., Zambolim, E.M., Loureiro, M.E., Pereira, A.A. and Zambolim, L. (2005) Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. *Crop. Breed. Appl. Biotechnol.* **5**, 387–393.

Djebbi, S., Bouktila, D., Makni, H., Makni, M. and Mezghani-Khemakhem, M. (2015) Identification and characterization of novel NBS-LRR resistance gene analogues from the pea. *Genet. Mol. Res.* **14**, 6419–6428.

Ellis, J., Dodds, P. and Pryor, T. (2000) Structure, function and evolution of plant disease resistance genes. *Curr. Opin. Plant Biol.* **3**, 278–284.

Felsenstein, J. (1985) Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. *Evolution (N. Y).*

Gichuru, E.K., Ithiru, J.M., Silva, M.C., Pereira, A.P. and Varzea, V.M.P. (2012) Additional physiological races of coffee leaf rust (Hemileia vastatrix) identified in

Kenya. *Trop. Plant Pathol.* **37**, 424–427.

Gichuru, E.K., Ithiru, J.M., Silva, M.C., Pereira, A.P. and Varzea, V.M.P. (2012) Additional physiological races of coffee leaf rust (Hemileia vastatrix) identified in Kenya. *Trop. Plant Pathol.* **37**, 424–427.

González, V.M., Aventín, N., Centeno, E. and Puigdomènech, P. (2014) Interspecific and intraspecific gene variability in a 1-Mb region containing the highest density of NBS-LRR genes found in the melon genome. *BMC Genomics* **15**, 1131.

Gou, X., He, K., Yang, H., Yuan, T., Lin, H., Clouse, S.D. and Li, J. (2010) Genome-wide cloning and sequence analysis of leucine-rich repeat receptor-like protein kinase genes in Arabidopsis thaliana. *BMC Genomics* **11**, 19.

Guo, Y.-L., Fitz, J., Schneeberger, K., Ossowski, S., Cao, J. and Weigel, D. (2011) Genome-Wide Comparison of Nucleotide-Binding Site-Leucine-Rich Repeat-Encoding Genes in Arabidopsis. *PLANT Physiol.* **157**, 757–769.

Hammond-Kosack, K.E. and Kanyuka, K. (2007) Resistance Genes (R Genes) in Plants. In eLS., Chichester, UK: John Wiley & Sons, Ltd.

Hao, W., Collier, S.M., Moffett, P. and Chai, J. (2013) Structural Basis for the Interaction between the Potato Virus X Resistance Protein (Rx) and Its Cofactor Ran GTPase-activating Protein 2 (RanGAP2). *J. Biol. Chem.* **288**, 35868–35876.

He, L., Du, C., Covaleda, L., Xu, Z., Robinson, A.F., Yu, J.Z., Kohel, R.J. and Zhang, H.-B. (2004) Cloning, Characterization, and Evolution of the NBS-LRR-Encoding Resistance Gene Analogue Family in Polyploid Cotton (Gossypium hirsutum L.). *Mol. Plant-Microbe Interact.* **17**, 1234–1241.

Hoorn, R.A.L. Van der (2001) Identification of Distinct Specificity Determinants in Resistance Protein Cf-4 Allows Construction of a Cf-9 Mutant That Confers Recognition of Avirulence Protein AVR4. *PLANT CELL ONLINE* **13**, 273–285.

Hulbert, S.H., Webb, C.A., Smith, S.M. and Sun, Q. (2001) Resistance gene complexes: evolution and utilization. *Annu. Rev. Phytopathol.* **39**, 285–312.

Jacob, F., Vernaldi, S. and Maekawa, T. (2013) Evolution and Conservation of Plant NLR Functions. *Front. Immunol.* **4**.

Jones, J.D.G. and Dangl, J.L. (2006) The plant immune system. *Nature* **444**, 323–329.

Kang, Y., Kim, K., Shim, S., Yoon, M., Sun, S., Kim, M., Van, K. and Lee, S.-H. (2012) Genome-wide mapping of NBS-LRR genes and their association with disease resistance in soybean. *BMC Plant Biol.* **12**, 139.

Kierszniowska, S., Seiwert, B. and Schulze, W.X. (2009) Definition of Arabidopsis Sterol-rich Membrane Microdomains by Differential Treatment with Methyl- -cyclodextrin and Quantitative Proteomics. *Mol. Cell. Proteomics* **8**, 612–623.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36.

Kim, S.H., Kwon, S.I., Saha, D., Anyanwu, N.C. and Gassmann, W. (2009) Resistance to the Pseudomonas syringae Effector HopA1 Is Governed by the TIR-NBS-LRR Protein RPS6 and Is Enhanced by Mutations in SRFR1. *PLANT Physiol.* **150**, 1723–1732.

Kuang, H., Wei, F., Marano, M.R., et al. (2005) The R1 resistance gene cluster contains three groups of independently evolving, type I R1 homologues and shows substantial structural variation among haplotypes of Solanum demissum. *Plant J.* **44**, 37–51.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874.

Kushalappa, A.C., Yogendra, K.N. and Karre, S. (2016) Plant Innate Immune Response: Qualitative and Quantitative Resistance. *CRC. Crit. Rev. Plant Sci.* **35**, 38–55.

Lahaye, T. (2002) The Arabidopsis RRS1-R disease resistance gene--uncovering the plant's nucleus as the new battlefield of plant defense? *Trends Plant Sci.* **7**, 425–7.

Leipe, D.D., Koonin, E. V. and Aravind, L. (2004) STAND, a Class of P-Loop NTPases Including Animal and Plant Regulators of Programmed Cell Death: Multiple, Complex Domain Architectures, Unusual Phyletic Patterns, and Evolution by Horizontal Gene Transfer. *J. Mol. Biol.* **343**, 1–28.

Lin, C., Mueller, L.A., Carthy, J.M., Crouzillat, D., Pétiard, V. and Tanksley, S.D. (2005) Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. *Theor. Appl. Genet.* **112**, 114–130.

Liu, J., Liu, X., Dai, L. and Wang, G. (2007) Recent Progress in Elucidating the Structure, Function and Evolution of Disease Resistance Genes in Plants. *J. Genet. Genomics* **34**, 765–776.

Mahé, L., Combes, M.C., Várzea, V.M.P., Guilhaumon, C. and Lashermes, P. (2008) Development of sequence characterized DNA markers linked to leaf rust (Hemileia vastatrix) resistance in coffee (Coffea arabica L.). *Mol. Breed.* **21**, 105–113.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B. (2015) RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **1**, vev003-vev003.

MAYNE, W.W. (1932) Physiological Specialisation of Hemileia vastatrix B. and Br. *Nature* **129**, 510–510.

McHale, L., Tan, X., Koehl, P. and Michelmore, R.W. (2006) Plant NBS-LRR proteins: adaptable guards. *Genome Biol.* **7**, 212.

Meyers, B.C., Kaushik, S. and Nandety, R.S. (2005) Evolving disease resistance genes. *Curr. Opin. Plant Biol.* **8**, 129–134.

Michelmore, R.W. and Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* **8**, 1113–30.

Mondragon-Palomino, M. (2002) Patterns of Positive Selection in the Complete NBS-LRR Gene Family of Arabidopsis thaliana. *Genome Res.* **12**, 1305–1315.

Mondragon-Palomino, M. and Gaut, B.S. (2005) Gene Conversion and the Evolution of Three Leucine-Rich Repeat Gene Families in Arabidopsis thaliana. *Mol. Biol. Evol.* **22**, 2444–2456.

Mur, L.A.J., Kenton, P., Lloyd, A.J., Ougham, H. and Prats, E. (2008) The hypersensitive response; the centenary is upon us but how much do we know? *J. Exp. Bot.* **59**, 501–520.

Mussatto, S.I., Machado, E.M.S., Martins, S. and Teixeira, J.A. (2011) Production, Composition, and Application of Coffee and Its Industrial Residues. *Food Bioprocess Technol.* **4**, 661–672.

Nei, M. and Kumar, S. (2000) Molecular evolution and phylogenetics 1st ed., Oxford: Oxford University Press.

Nepal, M. and Benson, B. (2015) CNL Disease Resistance Genes in Soybean and Their Evolutionary Divergence. *Evol. Bioinforma.*, 49.

Noir, S., Combes, M.C., Anthony, F. and Lashermes, P. (2001) Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (Coffea L.). *Mol. Genet. Genomics* **265**, 654–62.

Noronha-Wagner and Bettencourt, A. (1967) Genetic study of the resistance of Coffea sp to leaf rust 1. Identification and behavior of four factors conditioning disease reaction in Coffea arabica to twelve physiologic races of Hemileia vastatrix. *Can. J. Bot.* **45**, 2021–2031.

Padmanabhan, M., Cournoyer, P. and Dinesh-Kumar, S.P. (2009) The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities. *Cell. Microbiol.* **11**, 191–198.

Panchy, N., Lehti-Shiu, M.D. and Shiu, S.-H. (2016) Evolution of gene duplication in plants. *Plant Physiol.*, pp.00523.2016.

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J.P. and Hyvönen, J. (2013) Advances in plant gene-targeted and functional markers: a review. *Plant Methods* **9**, 6.

Qi, D. and Innes, R.W. (2013) Recent Advances in Plant NLR Structure, Function, Localization, and Signaling. *Front. Immunol.* **4**.

Rairdan, G.J., Collier, S.M., Sacco, M.A., Baldwin, T.T., Boettrich, T. and Moffett, P. (2008) The Coiled-Coil and Nucleotide Binding Domains of the Potato Rx Disease Resistance Protein Function in Pathogen Recognition and Signaling. *PLANT CELL ONLINE* **20**, 739–751.

Ratnaparkhe, M.B., Wang, X., Li, J., Compton, R.O., Rainville, L.K., Lemke, C., Kim, C., Tang, H. and Paterson, A.H. (2011) Comparative analysis of peanut

NBS-LRR gene clusters suggests evolutionary innovation among duplicated domains and erosion of gene microsynteny. *New Phytol.* **192**, 164–178.

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

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26.

Rodrigues, C.J., Bettencourt, A.J. and Rijo, L. (1975) Races of the Pathogen and Resistance to Coffee Rust. *Annu. Rev. Phytopathol.* **13**, 49–70.

Rozas, J. (2009) DNA sequence polymorphism analysis using DnaSP. *Methods Mol. Biol.* **537**, 337–350.

Rzhetsky, a and Nei, M. (1992) A Simple Method for Estimating and Testing Minimum-Evolution Trees. *Mol. Biol. Evol.* **9**, 945–967.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–25.

Saminathan, T., Nimmakayala, P., Manohar, S., et al. (2015) Differential gene expression and alternative splicing between diploid and tetraploid watermelon. *J. Exp. Bot.* **66**, 1369–1385.

Sanabria, N., Goring, D., Nürnberger, T. and Dubery, I. (2008) Self/nonself perception and recognition mechanisms in plants: a comparison of self-incompatibility and innate immunity. *New Phytol.* **178**, 503–514.

Sanseverino, W., Hénaff, E., Vives, C., Pinosio, S., Burgos-Paz, W., Morgante, M., Ramos-Onsins, S.E., Garcia-Mas, J. and Casacuberta, J.M. (2015) Transposon Insertions, Structural Variations, and SNPs Contribute to the Evolution of the Melon Genome. *Mol. Biol. Evol.* **32**, 2760–2774.

Sara, M., Walter, S., Paola, C., Luigi, M., Luigi, F. and Raffalella, E.M. (2012) Solanaceae Evolutionary Dynamics of the <i>I</i>2-NBS Domain. *Am. J. Plant Sci.* **3**, 283–294.

Stanke, M., Steinkamp, R., Waack, S. and Morgenstern, B. (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* **32**, W309–W312.

Syed, N.H., Kalyna, M., Marquez, Y., Barta, A. and Brown, J.W.S. (2012) Alternative splicing in plants – coming of age. *Trends Plant Sci.* **17**, 616–623.

Tan, S. and Wu, S. (2012) Genome Wide Analysis of Nucleotide-Binding Site Disease Resistance Genes in Brachypodium distachyon. *Comp. Funct. Genomics* **2012**, 1–12.

Velasco, R., Zharkikh, A., Troggio, M., et al. (2007) A High Quality Draft Consensus Sequence of the Genome of a Heterozygous Grapevine Variety Dilkes, B., ed. *PLoS One* **2**, e1326.

Wang, G.-F., Ji, J., EI-Kasmi, F., Dangl, J.L., Johal, G. and Balint-Kurti, P.J. (2015) Molecular and Functional Analyses of a Maize Autoactive NB-LRR Protein Identify Precise Structural Requirements for Activity Mackey, D., ed. *PLOS Pathog.* **11**, e1004674.

Wang, X., Wu, F., Xie, Q., et al. (2012) SKIP Is a Component of the Spliceosome Linking Alternative Splicing and the Circadian Clock in Arabidopsis. *Plant Cell* **24**, 3278–3295.

Wang, Z., Gerstein, M. and Snyder, M. (2010) RNA-Seq: a revolutionary tool for transcriptomics. **10**, 57–63.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D. and Hochstrasser, D.F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **112**, 531–52.

Yachdav, G., Kloppmann, E., Kajan, L., et al. (2014) PredictProtein--an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res.* **42**, W337–W343.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2014) The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8.

Yang, S., Zhang, X., Yue, J.-X., Tian, D. and Chen, J.-Q. (2008) Recent duplications dominate NBS-encoding gene expansion in two woody species. *Mol. Genet. Genomics* **280**, 187–198.

Zambolim, L. (2016) Current status and management of coffee leaf rust in Brazil. *Trop. Plant Pathol.* **41**, 1–8.

Zhao, Y., Huang, J., Wang, Z., et al. (2016) Allelic diversity in an NLR gene BPH9 enables rice to combat planthopper variation. *Proc. Natl. Acad. Sci.* **113**, 12850–12855.

Zuckerkandl, E. and Pauling, L. (1965) Evolutionary divergence and convergence in proteins. *Evol. genes proteins*, 97–166.

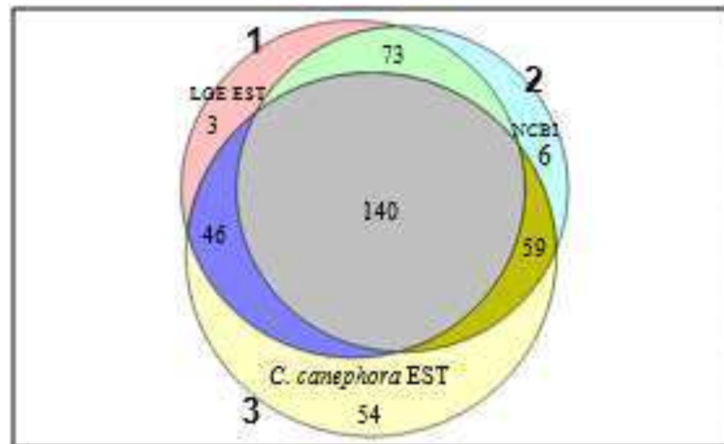
6. GENERAL CONCLUSIONS AND RECOMMENDATIONS

The first chapter of this thesis is a continuation of earlier work, in which SSH library was constructed during compatible and incompatible interactions. The objective of this chapter was identification of genes differentially expressed during incompatible and compatible interactions at 12 and 24 hours following *H. vastatrix* race II urediniospore inoculation. By SSH, we identified and sequenced 433 clones from which 352 non-redundant EST clones were annotated. Comparative analysis of genes upregulated and downregulated in both interactions showed different number of genes at 12 and 24 h.a.i. RT-qPCR analysis of seven genes (at 12, 24, 48 and 72 h.a.i.), showed that most of these genes are involved in basal defense in both interactions. The work provides a compiled expression profile of genes upregulated and downregulated at 12 and 24 h. a. i. during incompatible and compatible interactions.

The second chapter was intended to identify and sequence BAC clone with RGA (CARF005) functional marker belonging to NBS-LRR gene family, confirmed to be expressed in coffee. At the same time, we were also interested to screen coffee differential host clones for CARF005 and assess the differentiation of S_H genes. 13 ORFs were identified from a BAC clone with CAR005 insert. Four RGAs (gene 9, 10, 11 and 12) are clustered genes spanning a locus in either orientations (two genes in each orientation). Out of the total 5 RGAs identified, two genes (gene 5 and 11) encoding NBS-LRR proteins share 90.24% nucleotide (CDS) and 80.03% amino acid identities. These two genes are under intense diversifying selection resulting in functional diversity. Gene 11 correspond to a new S_H gene (S_H10) resistance gene (sharing a conserved domain CARF005 with S_H6 gene), as confirmed by screening 22 coffee differential host clones. This same gene (S_H10) was also detected in differential clone 644/18 H. Kawisari for the first time. All the five RGAs are positiond on chromosome 0 of *C. canephora* with high query coverage. Structural, functional and phylogenetic analysis was carried out for the two NBS-LRR (gene 5 and 11) in order to have indepth undersatnding of their structural and functional variability. The NBS-LRR genes sequenced in this work are the largest and most complete gene sequences reported in Arabica coffee to date, which makes the work very important for molecular breeding of coffee rust resistance.

As a recommendation, we suggest further research on 128/2-Dilla & Alghe to reaffirm that it confers S_H1, S_H6, a new S_H gene (S_H10) or all. Future work should also pursue the development of functional markers from sequenced RGAs and analyze their polymorphism in differential host clones conferring the gene. As S_H gene function is generally believed to be either independent or joint, further investigation should be sought to understand their functional cross-talk. We also recommend coffee genetic transformation with gene 11 to attest the practical applicability of this gene in developing transgenic coffee resistant to coffee rust.

7. APPENDIX I



Appendix I Figure 1. BLASTx homology search output of 352 ESTs at different databases. The number of significant matches found at three databases; LGE EST (1), NCBI (**2**) and *C. canephora* EST (3). Figures in the overlapping sections show the number of ESTs shared in respective databases as mined by BLAST2GO and/or BLASTx.

Appendix I Table 1. Differentially expressed genes in response to *H. vastatrix* in incompatible and compatible interactions as identified by suppression subtractive hybridization (SSH).

Clone	Putative sequence identification[a]	Clone size	Putative sequence description[b]	Number of hits[c]	E-value
Incompatible Interaction - Upregulated at 12 h.a.i.					
A. Resistance and anti-microbial Functions					
HT12F-08	emb|CAJ43737.1	532	class III chitinase [*Coffea arabica*]	20	1.00E-113
HT12F-12	ref|XM_002527501.1	1657	*Ricinus communis* Xyloglucan endotransglucosylase/hydrolase protein 2 precursor, putative, mRNA	17	6.00E-110
HT12F-31	dbj|BAA96501.1	697	cysteine protease [*Nicotiana tabacum*]	1	1.00E-154
HT12F-36	XM_008800454.1	910	PREDICTED: Acidic endochitinase-like	3	2.90E-07
HT12F-81	gb|KJ201790.1	699	*Populus trichocarpa* class III peroxidase (PRX25) mRNA, complete cds	8	6.00E-47
HT12F-82	gb|DQ993351.1	647	TransId-81849 CACAT36FLBUD *Coffea arabica* cDNA clone CACAT36FLBUD_Q4_12_P02_032.F	6	3.00E-114
HT12F-88	AY924306.1	1067	*Catharanthus roseus* secretory peroxidase (prx) mRNA, complete cds	52	4.3E-89
HT12F-94	AB678719.1	882	Petunia x hybrida PhCHS-A1, PhCHS-A2 genes for chalcone synthase, chalcone synthase, complete cds, cultivar: Baccara Rose Picotee	20	4.00E-72
HT12F-104	XM_006466850.1	428	PREDICTED: *Citrus sinensis* lactoylglutathione lyase-like (LOC102609048), mRNA	20	4.00E-22

B. Signal induction and transduction					
HT12F-23	emb|CAO21633.1	1000	unnamed protein product [*Vitis vinifera*]	3	1.00E-17
HT12F-24	emb|CAO22226.1	678	unnamed protein product [*Vitis vinifera*]	1	6.00E-20
HT12F-27	gb|DQ123920.1	631	*Coffea arabica* x *Coffea canephora* clone HT-SSH1-A01 mRNA sequence	25	7.00E-66
HT12F-46	gb|DQ124082.1	637	*Coffea arabica* clone MN-SSH1-C02 mRNA sequence	111	0.00
HT12F-47	gb|AAP03420.1	987	unknown protein [*Oryza sativa* (japonica cultivar-group)]	2	4.00E-28
HT12F-50	EM_EST:GT003697	886	TransId-204764 CACATN1 *Coffea arabica* cDNA clone CACATN1-1B26TVB similar to Auxin-binding protein ABP20 precursor - *Prunus persica* (Peach), mRNA sequence	50	0.00
HT12F-53	XM_002267183.3	653	PREDICTED: *Vitis vinifera* MLP-like protein 423 (LOC100243241), mRNA	20	1.00E-16
HT12F-76	dbj|BAC22124.1	1059	t-complex polypeptide 1 [*Bruguiera sexangula*]	1	0.00
HT12F-79	EM_EST:GT008407	1453	TransId-237846 CAET42MIX *Coffea arabica* cDNA clone CAET42MIX-CFEY012TV similar to Putative ethylene responsive element binding protein 2 - *Gossypium hirsutum* (Upland cotton), mRNA sequence	2	3.00E-103
HT12F-100	gb|ABR18801.1	261	kinase-associated protein phosphatase 1 [*Solanum peruvianum*]	1	1.00E-178
HT12F-102	gb|EF147243.1	495	*Populus trichocarpa* clone WS01229_M17 unknown mRNA	30	8.00E-114
C. Cell maintenance and homeostasis					
HT12F-01	ref|XM_009106879.1	639	PREDICTED: *Brassica rapa* DNA-directed RNA polymerase subunit 10-like protein (LOC103831031), mRNA	11	4.00E-20

HT12F-02	gb|GQ497218.1	427	*Coffea arabica* galactinol synthase (GolS1) mRNA, complete cds	1	4.00E-151
HT12F-05	dbj|AB032245.1	561	*Cucumis sativus* psbP mRNA for 23kDa polypeptide of the oxygen-evolving complex of photosystem II, complete cds	7	5.00E-42
HT12F-13	gb|AAK30204.1	887	endoxyloglucan transferase [*Daucus carota*]	20	1.00E-142
HT12F-14	ref|XM_007132605.1	514	*Phaseolus vulgaris* hypothetical protein (PHAVU_011G114500g) mRNA, partial cds	1	2.00E-20
HT12F-16	emb|CAO48519.1	815	unnamed protein product [*Vitis vinifera*]	1	4.00E-39
HT12F-18	emb|CAO16005.1	608	unnamed protein product [*Vitis vinifera*]	1	1.00E-58
HT12F-19	gb|ABK95508.1	796	unknown [*Populus trichocarpa*]	20	1.00E-41
HT12F-21	EM_EST:GR991879	758	TransId-111665 CarCTFrHh24FL *Coffea arabica* cDNA clone CarCTFrHh24FL_Q3_01_J11_170.F Gene encoding ADP-ribosylation factor and similar to other ARFs and ARF-like proteins.	50	0.00
HT12F-25	gb|AAA33107.1	618	CYC02_CATRO CYC02 protein	5	1.00E-14
HT12F-26	dbj|AK286901.1	690	Glycine max cDNA, clone: GMFL01-39-N15	6	3.00E-65
HT12F-30	gb|AY461597.1	1040	Synthetic construct arsenic-like protein gene, complete cds	2	7.00E-18
HT12F-38	XM_003634815.2	1381	*Vitis vinifera* omega-6 fatty acid desaturase, chloroplastic (LOC100248377), mRNA	9	1.00E-135
HT12F-39	emb|CAA36249.1	387	metallothionein [*Mimulus guttatus*]	10	2.00E-15
HT12F-40	emb|CAD11990.1	1040	rubisco small subunit [*Coffea arabica*]	6	6.00E-39
HT12F-41	emb|CAO44191.1	1040	unnamed protein product [*Vitis vinifera*]	1	1.00E-43

HT12F-44	dbj|BAA31582.1	331	mitochondrial phosphate transporter [Glycine max]	13	5.00E-16
HT12F-49	EM_EST:GR992352	705	TransId-112624 CarCTFrHh24FL *Coffea arabica* cDNA clone CarCTFrHh24FL_Q4_04_H22_344.F lipid transfer protein 3 (LTP3), identical to lipid transfer protein 3 from *Arabidopsis thaliana* (gi:8571921); contains Pfam protease inhibitor/seed storage/LTP family domain	50	0.00
HT12F-64	emb|CAA66109.3	1027	specific tissue protein 2 [*Cicer arietinum*]	6	2.00E-12
HT12F-66	dbj|AB043960.2	999	*Bruguiera gymnorhiza* psbO mRNA for oxygen evolving enhancer protein 1 precursor, complete cds	7	2.00E-98
HT12F-68	gb|ABK23160.1	728	unknown [*Picea sitchensis*]	2	2.00E-36
HT12F-70	DV685708	702	CGN-27126 Leaf *Coffea canephora* cDNA clone cccl5p22 5', mRNA sequence	9	6.00E-19
HT12F-71	emb|CAO47331.1	750	unnamed protein product [*Vitis vinifera*]	5	3.00E-52
HT12F-75	gb|ABK93168.1	978	unknown [*Populus trichocarpa*]	2	9.00E-17
HT12F-83	gb|EF044213.1	1189	*Coffea arabica* chloroplast, complete genome	2	9.00E-152
HT12F-84	HQ288064.1	939	*Hottentotta judaicus* clone Hj0135 beta-buthitoxin-Hj2a mRNA, complete cds	20	8.21E-11
HT12F-92	gb|KF467245.1	605	*Dendrocalamus latiflorus* actin (act) mRNA, complete cds	45	2.00E-61
HT12F-101	XM_011089750.1	521	PREDICTED: *Sesamum indicum* 50S ribosomal protein L1, chloroplastic (LOC105169360), mRNA	20	8.00E-120

D. No GO term associated

HT12F-29	EG329044	492	CR-EST271 Leaf EST Library of robusta coffee cultivar 'CxR' *Coffea congensis* x *Coffea canephora* cDNA clone C2F12R 3', mRNA sequence	20	6.00E-52
HT12F-32	DV665453	424	CGN-2958 Pericarp *Coffea canephora* cDNA clone cccp23d23 5', mRNA sequence	2	2.00E-110
HT12F-35	EM_EST:GT650787	1184	CC00-XX-SH3-039-B05-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-039-B05-EM, mRNA sequence.	50	0.00
HT12F-37	DV688739	801	CGN-31537 Leaf *Coffea canephora* cDNA clone cccl14k10 5', mRNA sequence	9	5.00E-65
HT12F-55	EE195364	646	CC-L01_058_H02 Coffea young leaves *Coffea canephora* cDNA clone CC-L01_058_H02, mRNA sequence	2	8.00E-107
HT12F-59	GT647699	433	CC00-XX-SH3-009-C03-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-009-C03-EM,mRNA sequence	34	4.00E-63
HT12F-65	emb|CAO65816.1	321	unnamed protein product [*Vitis vinifera*]	1	4.00E-15
HT12F-74	XM_009799371.1	872	PREDICTED: *Nicotiana sylvestris* uncharacterized LOC104243800 (LOC104243800), mRNA	19	1.94E-07
HT12F-77	emb|CAN73948.1	460	hypothetical protein [*Vitis vinifera*]	2	8.00E-36
HT12F-78	emb|HG964666.1	673	*Coffea arabica* clone BAC 126-3E, cultivar IAPAR 59	1	6.00E-22
HT12F-80	EM_EST:DN478953	1601	altr013xa03 A. *brassicicola mycelial* culture infecting *B. oleracea* Alternaria brassicicola cDNA clone altr013xa03, mRNA sequence.	50	8.4E-23

HT12F-90	GT647948	485	CC00-XX-SH3-047-E10-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-047-E10-EM,mRNA sequence	90	0.00
HT12F-91	DV665145	766	CGN-2541 Pericarp *Coffea canephora* cDNA clone cccp30g1 5', mRNA sequence	60	0.00
HT12F-98	GT647942	517	CC00-XX-SH3-020-F01-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-020-F01-EM,mRNA sequence	91	0.00
Incompatible Interaction - Downregulated at 12 h.a.i.					
B. Signal induction and transduction					
HT12R-15	gb|ABK96401.1	1214	unknown [*Populus trichocarpa* x *Populus deltoides*]	2	1.00E-127
HT12R-18	gb|JF897606.1	1004	*Nicotiana benthamiana* chloroplast PsbO4 precursor (psbO4) mRNA, complete cds; nuclear gene for chloroplast product	7	3.00E-156
HT12R-28	dbj|BAC22124.1	1048	t-complex polypeptide 1 [*Bruguiera sexangula*]	1	0.00
HT12R-44	emb|FP099325.1	454	*Phyllostachys edulis* cDNA clone: bphylf003j16, full insert sequence	45	6.00E-65
HT12R-77	XM_002276888.3	1059	PREDICTED: *Vitis vinifera* T-complex protein 1 subunit alpha (LOC100249575), mRNA	20	1.00E-142
HT12R-109	KF588660.1	357	*Actinidia deliciosa* GA signaling F-Box (SLY1_12) gene, complete cds	9	7.00E-20
HT12R-123	gb|DQ459385.1	521	*Nicotiana tabacum* serine/threonine kinase mRNA, partial cds	2	2.00E-11
HT12R-124	emb|CAA58701.1	549	inorganic pyrophosphatase [*Nicotiana tabacum*], CA00-XX-EA1-044-G02-EC.F Coffea arabica EA1 Coffea arabica cDNA clone CA00-XX-EA1-044-G02-EC, mRNA sequence	1	0.00

C. Cell maintenance and homeostasis

HT12R-01	dbj|BAC77694.1	517	non-specific lipid-transfer protein 1-like	20	1.62E-31
HT12R-16	emb|CAA66109.3	1041	specific tissue protein 2 [*Cicer arietinum*]	6	2.00E-12
HT12R-20	AK322608.1	741	*Solanum lycopersicum* cDNA, clone: LEFL1039DE01, HTC in leaf	2	1.00E-60
HT12R-26	emb|CAO15699.1	545	unnamed protein product [*Vitis vinifera*]	1	5.00E-18
HT12R-36	emb|CAN65487.1	899	hypothetical protein [*Vitis vinifera*]	20	6.00E-39
HT12R-50	EM_EST:GR994152	345	TransId-227175 CACATN1 *Coffea arabica* cDNA clone CACATN1-8I04TV similar to Alcohol dehydrogenase - *Solanum tuberosum*, mRNA sequence.	50	2.9E-28
HT12R-71	gb|ABK94769.1	702	unknown [*Populus trichocarpa*]	2	2.00E-06
HT12R-99	gb|KC008722.1	351	*Camellia sinensis* aquaporin protein 7 mRNA, complete cds	108	2.00E-69
HT12R-107	XM_002280122.3	435	PREDICTED: *Vitis vinifera* probable chalcone--flavonone isomerase 3 (LOC100255217), mRNA	20	3.00E-34
HT12R-110	ref|XM_007028432.1	689	*Theobroma cacao* Haloacid dehalogenase-like hydrolase superfamily protein isoform 2 (TCM_024330) mRNA, complete cds	14	1.00E-112
HT12R-114	emb|FQ380233.1	501	*Vitis vinifera* clone SS0ADG1YA21	56	7.00E-45
HT12R-115	emb|CAO60899.1	199	unnamed protein product [*Vitis vinifera*]	1	3.00E-15
HT12R-120	emb|CAO15686.1	475	unnamed protein product [*Vitis vinifera*]	3	0.00

HT12R-127	XM_009601794.1	770	PREDICTED: *Nicotiana tomentosiformis* 3-hydroxyisobutyryl-CoA hydrolase-like protein 5 (LOC104095630), transcript variant X2, mRNA	20	1.00E-45
HT12R-132	emb|CAO70990.1	674	probable isoleucine--trna mitochondrial isoform x1	20	2.16E-84
D. No GO term associated					
HT12R-05	emb|CAO40936.1	346	unnamed protein product [*Vitis vinifera*]	3	2.00E-40
HT12R-11	GT647699	262	CC00-XX-SH3-009-C03-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-009-C03-EM,mRNA sequence	34	2.00E-63
HT12R-22	DV685708	177	CGN-27126 Leaf *Coffea canephora* cDNA clone cccl5p22 5', mRNA sequence	1	1.00E-19
HT12R-29	emb|CAN73948.1	297	hypothetical protein [*Vitis vinifera*]	2	8.00E-36
HT12R-32	gb|BT009458.1	1608	*Triticum aestivum* clone wlsu2.pk0001.h3:fis, full insert mRNA sequence	1	2.00E-19
HT12R-53	DV674191	841	CGN-10605 Seed of Middle Development Stage *Coffea canephora* cDNA clone cccs30w14j9 5',mRNA sequence	2	0.00
HT12R-54	EM_EST:CK909834	453	hasp018xi08 *Heterobasidion annosum* - Scots pine infection stage subtraction cDNA library (hasp) *Pinus sylvestris*/*Heterobasidion annosum* cDNA clone hasp018xi08, mRNA sequence.	50	1.6E-30
HT12R-55	emb|CAO42517.1	497	unnamed protein product [*Vitis vinifera*]	4	7.00E-39
HT12R-60	EM_EST:DR957107	909	ZM_BFb0055H20.f ZM_BFb *Zea mays* cDNA 3', mRNA sequence.	50	1.5E-7

HT12R-61	EM_EST:GW485541	528	CA00-XX-FB2-067-H11-AC.F *Coffea arabica* FB2 *Coffea arabica* cDNA clone CA00-XX-FB2-067-H11-AC, mRNA sequence.	50	1.1E-28
HT12R-67	XM_004230670.2	858	PREDICTED: *Solanum lycopersicum* uncharacterized LOC101246097 (LOC101246097), mRNA	20	5.00E-27
HT12R-78	gb|EU022152.1	499	*Montipora capitata* clone DCMU1L5 unknown mRNA	1	2.00E-11
HT12R-88	EM_EST:GT010382	411	TransId-242754 CACAT36FR26 *Coffea arabica* cDNA clone CACAT36FR26-CAFR746TVC similar to Secretory peroxidase - *Catharanthus roseus* (Rosy periwinkle) (Madagascar periwinkle), mRNA sequence.	100	0.00
HT12R-96	EM_EST:GR992840	444	TransId-200615 CarCatBudEnri *Coffea arabica* cDNA clone CarCatBudEnri_14-F07, mRNA sequence.	100	1.2E-120
HT12R-111	EM_EST:GW485783	154	CA00-XX-FB2-034-F08-BM.F *Coffea arabica* FB2 *Coffea arabica* cDNA clone CA00-XX-FB2-034-F08-BM, mRNA sequence.	50	5.0E-71
HT12R-119	DV673184	421	CGN-9127 Pericarp *Coffea canephora* cDNA clone cccp8l24 5', mRNA sequence	193	1.00E-153
HT12R-121	XM_006605921.1	287	PREDICTED: Glycine max alpha-L-fucosidase 2-like (LOC100808977), transcript variant X3, mRNA	20	1.00E-13
HT12R-126	EM_EST:GW463905	234	CA00-XX-FB2-062-H03-SB.F *Coffea arabica* FB2 *Coffea arabica* cDNA clone CA00-XX-FB2-062-H03-SB, mRNA sequence.	50	2.5E-86
HT12R-133	DV691634	408	CGN-35366 Leaf *Coffea canephora* cDNA clone cccl16d15 5', mRNA sequence	121	3.00E-169
Incompatible Interaction - Upregulated at 24 h.a.i.					

A. Resistance and anti-microbial functions					
HT24F-04	gb|HM051339.1	532	*Coffea arabica* class III chitinase precursor, mRNA, partial cds	2	0.00
HT24F-21	ref|NP_193383.1	400	cysteine proteinase inhibitor like protein [*Arabidopsis thaliana*]	6	5.00E-06
HT24F-31	gb|JQ013437.1	335	*Coffea arabica* x *Coffea canephora* putative class III peroxidase (POX6) mRNA, partial cds	1	6.00E-168
HT24F-64	sp|Q9C7Z9	800	SCP18_ARATH Serine carboxypeptidase-like 18 precursor	10	1.00E-133
HT24F-115	gb|DQ993351.1	435	TransId-81849 CACAT36FLBUD *Coffea arabica* cDNA clone CACAT36FLBUD_Q4_12_P02_032.F similar to trypsin and protease inhibitor family protein / Kunitz family protein, similar to LeMir (miraculin homolog) GI:2654440 from (*Lycopersicon esculentum*)	1	5.00E-145
HT24F-142	XM_011082582.1	910	probable glutathione peroxidase 2	20	1.00E-57
HT24F-144	emb|AJ749800.1	523	Populus x canadensis mRNA for ferulate-5-hydroxylase (f5h gene)	500	5.00E-17
B. Signal induction and transduction					
HT24F-10	XM_006469498.1	313	PREDICTED: *Citrus sinensis* glycine-rich protein-like (LOC102625227), mRNA	2	0.014
HT24F-27	emb|CAA88841.1	897	phosphoglycerate kinase [*Nicotiana tabacum*]	4	7.00E-25
HT24F-30	dbj|BAF37542.1	377	cell wall glycine-rich protein [*Cucumis sativus*]	8	1.00E-06
HT24F-32	gb|ABC87760.1	629	jasmonic acid-amino acid-conjugating enzyme [*Nicotiana attenuata*]	4	3.00E-71

HT24F-36	gb|EAZ18741.1	289	hypothetical protein OsJ_032950 [*Oryza sativa* (japonica cultivar-group)]	1	1.00E-106
HT24F-40	XM_011093019.1	503	PREDICTED: *Sesamum indicum* ATPase GET3 (LOC105171786), mRNA	20	2.00E-61
HT24F-54	gb|ABL67651.1	823	putative auxin-repressed/dormancy-associated protein [Citrus cv. Shiranuhi]	4	5.00E-51
HT24F-85	EM_EST:GR990444	627	TransId-100452 CACAT45FR *Coffea arabica* cDNA clone CACAT45FR_32_B02_018.F similar to auxin-responsive protein / indoleacetic acid-induced protein 9 (IAA9), identical to SP:Q38827 Auxin-responsive protein IAA9 (Indoleacetic acid-induced protein 9)	61	0.00
HT24F-123	ref|XM_004232861.1	826	PREDICTED: *Solanum lycopersicum* auxin-repressed 12.5 kDa protein-like, transcript variant 2 (LOC101258429), mRNA	5	2.00E-22
HT24F-133	ref|XM_006338244.1	691	PREDICTED: LRR receptor-like serine/threonine-protein kinase NIK1/protein NSP-INTERACTING KINASE 1-like (LOC102591832), mRNA	494	1.00E-147
C. Cell maintenance and homeostasis					
HT24F-02	emb|CAO21633.1	472	unnamed protein product [*Vitis vinifera*]	4	1.00E-17
HT24F-14	ref|XM_006436891.1	863	*Citrus clementina* hypothetical protein (CICLE_v10032123mg) mRNA, complete cds	2	9.00E-111
HT24F-12	gb|ABS84825.1	309	thioredoxin [*Limonium bicolor*]	2	5.00E-50
HT24F-26	emb|HF571519.1	1173	*Coffea arabica* mRNA for polyubiquitin 10 (ubq10 gene)	3	7.00E-93
HT24F-38	emb|CAN64392.1	933	hypothetical protein [*Vitis vinifera*]	1	2.00E-70

HT24F-42	XM_004291729.2	515	PREDICTED: *Fragaria vesca* subsp. vesca photosystem II 5 kDa protein, chloroplastic (LOC101292766), mRNA	5	3.00E-21
HT24F-44	gb|AAF03675.1	562	|AF149311_1 raucaffricine-O-beta-D-glucosidase [*Rauvolfia serpentina*]	1	5.00E-63
HT24F-48	gb|KC570455.1	445	*Coffea arabica* Guatemalan metallothionein I (CAMETAL1) mRNA, complete cds	2	0.00
HT24F-49	gb|AAX84672.1	1216	aldo/keto reductase AKR [*Manihot esculenta*]	11	2.00E-60
HT24F-51	gb|ABL63124.1	497	MYB transcription factor [*Catharanthus roseus*]	1	1.00E-59
HT24F-58	gb|HQ599861.1	415	*Gardenia jasminoides* aquaporin-like protein (AQP) mRNA, complete cds	7	7.00E-114
HT24F-59	gb|JX134620.1	1015	Neolamarckia cadamba xyloglucan endotransglycosylase/hydrolase 2 (XTH2) mRNA, complete cds	6	3.00E-146
HT24F-62	gb|AAF66242.1	169	AF243180_1 dicyanin [*Lycopersicon esculentum*]	2	4.00E-27
HT24F-76	gb|ABY57764.1	675	extracellular Ca^{2+} sensing receptor [*Nicotiana tabacum*]	2	2.00E-85
HT24F-79	ref|XM_008224900.1	655	*Prunus persica* hypothetical protein (PRUPE_ppa005087mg) mRNA, complete cds	22	2.00E-90
HT24F-80	sp|O80363	1099	RK17_TOBAC 50S ribosomal protein L17, chloroplast precursor	2	2.00E-74
HT24F-81	ref|XM_002275947.2	622	*Vitis vinifera* clone SS0ABG47YM12	3	2.00E-85
HT24F-88	AJ310148.1	367	hydroquinone glucosyltransferase-like	20	1.00E-43
HT24F-89	emb|CAC35167.1	438	arbutin synthase [*Rauvolfia serpentina*]	1	0.00

HT24F-92	XM_009774342.1	436	PREDICTED: *Nicotiana sylvestris* sugar transporter ERD6-like 8 (LOC104222990), transcript variant X3, mRNA	20	7.00E-06
HT24F-94	ref|XM_004241360.1	280	*Solanum lycopersicum* cDNA, clone: LEFL1087BG04, HTC in leaf	4	1.00E-19
HT24F-95	ref|XM_007052335.1	562	*Theobroma cacao* Zinc-binding ribosomal protein family protein (TCM_005784) mRNA, complete cds	29	4.00E-73
HT24F-109	ref|XM_007202630.1	290	PREDICTED: *Prunus mume* eukaryotic translation initiation factor 5A-2-like (LOC103341263), mRNA	25	8.00E-77
HT24F-111	ref|NM_001288153.1	250	*Solanum tuberosum* eukaryotic translation initiation factor 5A-1/2-like (LOC102594671), mRNA	6	9.00E-61
HT24F-121	ref|XM_002521115.1	271	*Ricinus communis* photosystem I reaction center subunit IV A, chloroplast precursor, mRNA	2	5.00E-34
HT24F-138	KJ159913.1	808	plant cadmium resistance 2-like protein	20	5.31E-84
HT24F-146	XM_011075154.1	893	F130dihydroxy-3-keto-5-methylthiopentene dioxygenase 2	20	3.00E-98
D. No GO term associated					
HT24F-01	DV701419	356	CGN-47604 Seed of Late Development Stage *Coffea canephora* cDNA clone cccs46w18c22 5',mRNA sequence	30	5.00E-146
HT24F-08	DV675664	548	CGN-12719 Pericarp *Coffea canephora* cDNA clone cccp24f2 5', mRNA sequence	247	0.00
HT24F-23	DV709338	728	CGN-57616 Cherry of Early Development Stage *Coffea canephora* cDNA clone cccwc22w20d6 5',mRNA sequence	3	0.00

HT24F-24	GW346015	608	CC_09_UAS296 Coffee drought stressed leaf cDNA library *Coffea canephora* cDNA, mRNA sequence	36	5.00E-113
HT24F-25	DV665453	426	CGN-2958 Pericarp *Coffea canephora* cDNA clone cccp23d23 5', mRNA sequence	1	7.00E-66
HT24F-33	ref|XP_001642067.1	369	predicted protein [*Nematostella vectensis*]	5	1.00E-08
HT24F-39	EM_EST:GT679570	911	CA00-XX-CL2-122-D11-EQ.F *Coffea arabica* CL2 *Coffea arabica* cDNA clone CA00-XX-CL2-122-D11-EQ, mRNA sequence	50	7.0E-104
HT24F-46	DV691787	766	CGN-35555 Leaf *Coffea canephora* cDNA clone cccl20g17 5', mRNA sequence	278	0.00
HT24F-47	DV692053	800	CGN-35888 Leaf *Coffea canephora* cDNA clone cccl19k11 5', mRNA sequence	275	0.00
HT24F-72	DV691805	873	CGN-35586 Leaf *Coffea canephora* cDNA clone cccl26g8 5', mRNA sequence	2	9.00E-28
HT24F-74	DV677463	366	CGN-14856 Pericarp *Coffea canephora* cDNA clone cccp5g2 5', mRNA sequence	74	3.00E-178
HT24F-82	DV710153	973	CGN-58630 Cherry of Early Development Stage *Coffea canephora* cDNA clone cccwc22w18h19 5', mRNA sequence	278	0.00
HT24F-91	DV666702	213	CGN-4669 Pericarp *Coffea canephora* cDNA clone cccp16n1 5', mRNA sequence	184	7.00E-93
HT24F-100	DV699983	762	CGN-45912 Seed of Late Development Stage *Coffea canephora* cDNA clone cccs46w21l18 5',mRNA sequence	13	0.00
HT24F-105	DV687755	569	CGN-29859 Leaf *Coffea canephora* cDNA clone cccl6d20 5', mRNA sequence	2	2.00E-152

HT24F-106	DV684384	538	CGN-25272 Seed of Middle Development Stage *Coffea canephora* cDNA clone cccs30w32b2 5',mRNA sequence	22	0.00
HT24F-107	DV704049	365	CGN-50738 Seed of Late Development Stage *Coffea canephora* cDNA clone cccs46w30m17 5',mRNA sequence	9	1.00E-162
HT24F-110	gb|DQ123920.1	277	*Coffea arabica* x *Coffea canephora* clone HT-SSH1-A01 mRNA sequence	14	6.00E-88
HT24F-113	gb|DQ123993.1	532	*Coffea arabica* x *Coffea canephora* clone HT-SSH4-E08 mRNA sequence	55	0.00
HT24F-114	DV692025	623	CGN-35851 Leaf *Coffea canephora* cDNA clone cccl19h10 5', mRNA sequence	85	0.00
HT24F-116	EM_EST:GW477134	480	CA00-XX-AR1-017-A11-EB.F *Coffea arabica* AR1 *Coffea arabica* cDNA clone CA00-XX-AR1-017-A11-EB, mRNA sequence.	50	0.00
HT24F-119	EE200494	238	CC-F01_019_M17 Cherry of different development stage *Coffea canephora* cDNA clone CC-F01_019_M17,mRNA sequence	238	2.00E-79
HT24F-120	EM_EST:GT007255	222	TransId-234692 CAET42MIX *Coffea arabica* cDNA clone CAET42MIX-CFEWX92TV similar to Eukaryotic peptide chain release factor subunit 1-3-like - *Solanum tuberosum*, mRNA sequence.	50	6.6E-90
HT24F-124	ref|XM_007133793.1	587	PREDICTED: Glycine max 60S ribosomal protein L3-like (LOC100806256), mRNA	34	1.00E-72
HT24F-126	EM_EST:GR986918	367	TransId-89006 CACAT36L *Coffea arabica* cDNA clone CACAT36L_32_A08_113.F similar to ribulose bisphosphate carboxylase/oxygenase activase / RuBisCO activase	50	0.00

HT24F-127	EM_EST:GW486895	396	CA00-XX-LV9-002-C06-JE.F *Coffea arabica* LV9 *Coffea arabica* cDNA clone CA00-XX-LV9-002-C06-JE, mRNA sequence.	50	0.00
HT24F-131	emb|CU223716.1	545	*Camellia sinensis* clone U10BcDNA 3161 acyl-CoA-binding protein mRNA, complete cds	13	2.00E-60
HT24F-132	EM_EST:GT011431	1007	TransId-244980 CACAT36FR26 *Coffea arabica* cDNA clone CACAT36FR26_Q2_03_O22_351.F similar to Epicotyl-specific tissue protein - *Striga asiatica*, mRNA sequence.	50	5.7E-103
HT24F-136	EM_EST:FL035568	404	13091293 CERES-CL13 *Zea mays* cDNA clone 1384678 5', mRNA sequence.	50	1.0E-6
HT24F-141	EE194521	557	CC-L01_040_A06 Coffea young leaves Coffea canephora cDNA clone CC-L01_040_A06, mRNA sequence	263	4.00E-164
HT24F-143	EM_EST:GW484397	582	CA00-XX-AR1-007-E03-EB.F *Coffea arabica* AR1 *Coffea arabica* cDNA clone CA00-XX-AR1-007-E03-EB, mRNA sequence	50	2.7E-125
HT24F-145	EM_EST:GR992327	605	TransId-112571 CarCTFrHh24FL *Coffea arabica* cDNA clone CarCTFrHh24FL_Q4_03_P06_096.F similar to acidic endochitinase (CHIB1)	50	8.8E-107
HT24F-151	emb|AM706411.1	859	Eristalis tenax partial mRNA for hypothetical protein (ORF1), isolate 3	2	2.00E-18
Incompatible Interaction - Downregulated 24 h.a.i.					
B. Signal induction and transductions					
HT24R-05	ref|NP_193606.1	503	BIN2 (BRASSINOSTEROID-INSENSITIVE 2); kinase [*Arabidopsis thaliana*]	1	0.00

HT24R-75	gb|AAD12209.1	596	Ac-like transposase [*Arabidopsis thaliana*],similar to Histone H1 - *Plantago major* (Common plantain), mRNA sequence	4	5.00E-11
HT24R-77	ref|XM_006475097.1	426	PREDICTED: *Citrus sinensis* polyadenylate-binding protein RBP45C-like (LOC102613263), mRNA	20	5.00E-58
HT24R-94	gb|AAC49651.1	663	actin [*Striga asiatica*]	9	1.00E-100
C. Cell maintenance and homeostasis					
HT24R-02	emb|CAO48896.1	1006	unnamed protein product [*Vitis vinifera*]	2	2.00E-29
HT24R-10	emb|CAC01237.1	640	RNA Binding Protein 45 [*Nicotiana plumbaginifolia*]	3	4.00E-34
HT24R-24	gb|AAB33480.2	898	alcohol dehydrogenase ADH [*Lycopersicon esculentum*]	6	5.00E-10
HT24R-25	ref|XM_006583692.1	432	PREDICTED: Glycine max probable xyloglucan endotransglucosylase/hydrolase protein 28-like (LOC100778185), transcript variant X3, mRNA	15	1.00E-81
HT24R-28	emb|CAO38811.1	458	unnamed protein product [*Vitis vinifera*]	2	1.00E-72
HT24R-29	gb|JQ678763.1	775	*Camellia sinensis* cell division cycle protein 48 (CDC48) mRNA, complete cds	11	7.00E-171
HT24R-30	ref|XM_004298176.1	687	PREDICTED: *Fragaria vesca* subsp. vesca brassinosteroid-regulated protein BRU1-like (LOC101306522), mRNA	39	1.00E-112
HT24R-32	ref|XM_006452953.1	840	*Citrus clementina* hypothetical protein (CICLE_v10009915mg) mRNA, complete cds	5	1.00E-59
HT24R-33	dbj|BAA04633.1	679	PSI-H precursor [*Nicotiana sylvestris*]	1	3.00E-41
HT24R-34	emb|CAN69723.1	442	hypothetical protein [*Vitis vinifera*]	20	6.00E-85

HT24R-37	gb|EF044213.1	877	*Coffea arabica* chloroplast, complete genome	486	0.00
HT24R-42	ref|NM_001288557.1	506	*Solanum tuberosum* serine hydroxymethyltransferase, mitochondrial-like (SHMT), mRNA	46	6.00E-120
HT24R-43	gb|JQ693578.1	316	*Corchorus capsularis* voucher 890 xyloglucan endotransglycosylase/hydrolase (XTH1) gene, partial cds	15	3.00E-32
HT24R-45	ref|XM_006356093.1	701	PREDICTED: *Solanum tuberosum* photosystem I reaction center subunit VI-2, chloroplastic-like (LOC102578347), mRNA	10	3.00E-84
HT24R-50	emb|AM087674.1	475	*Coffea arabica* mRNA for sucrose synthase (sus1 gene)	3	0.00
HT24R-53	emb|FQ394537.1	806	*Vitis vinifera* clone SS0AFA20YB24	9	1.00E-149
HT24R-55	ref|XM_002307327.2	774	*Populus trichocarpa* 14-3-3 protein 32kDa endonuclease (POPTR_0005s17300g) mRNA, complete cds	15	3.00E-120
HT24R-58	ref|XM_003631941.1	359	PREDICTED: *Vitis vinifera* photosystem I reaction center subunit II, chloroplastic-like, transcript variant 4 (LOC100240928),mRNA	25	5.00E-45
HT24R-63	ref|XM_007216922.1	257	*Prunus persica* hypothetical protein (PRUPE_ppa002714mg) mRNA, complete cds	20	1.00E-39
HT24R-66	ref|XM_007039716.1	374	*Theobroma cacao* Acyl-CoA oxidase 1 (TCM_015928) mRNA, complete cds	1	8.00E-48
HT24R-76	emb|CAO22714.1	440	nadp-dependent malic enzyme-like	234	1.12E-24
HT24R-91	emb|CAO47717.1	577	unnamed protein product [*Vitis vinifera*]	2	1.00E-41
HT24R-97	emb|CAJ38366.1	659	histone H1 [*Plantago major*]	4	5.00E-23

HT24R-116	emb|CAN68564.1	919	hypothetical protein [*Vitis vinifera*]	1	4.00E-80
HT24R-127	emb|CAB87415.1	550	putative protein [*Arabidopsis thaliana*]	1	2.00E-15
HT24R-128	emb|CAO43530.1	632	unnamed protein product [*Vitis vinifera*]	1	0.00
D. No GO term associated					
HT24R-07	gb|AAM34773.1	1770	nam-like protein 10 [Petunia x hybrida]	1	1.00E-125
HT24R-18	DV668167	633	CGN-6703 Pericarp *Coffea canephora* cDNA clone cccp12k13 5', mRNA sequence	1	1.00E-99
HT24R-21	gb|DQ124083.1	855	*Coffea arabica* clone MN-SSH3-E02 mRNA sequence	112	0.00
HT24R-35	GT649239	363	CC00-XX-SH3-075-F10-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-075-F10-EM,mRNA sequence	1	1.00E-87
HT24R-56	EM_EST:GT670561	350	CA00-XX-PC1-002-B08-EC.F *Coffea arabica* PC1 *Coffea arabica* cDNA clone CA00-XX-PC1-002-B08-EC, mRNA sequence.	50	3.9E-15
HT24R-57	EM_EST:GR982023	400	TransId-73831 CACAT36FL *Coffea arabica* cDNA clone CACAT36FL_43_M21_333.F similar to shaggy-related protein kinase eta / ASK-eta (ASK7), identical to shaggy-related protein kinase eta (ASK-eta) (*Arabidopsis thaliana*)	50	3.5E-68
HT24R-59	EM_EST:GR993804	1330	TransId-227980 CACATN1 *Coffea arabica* cDNA clone CACATN1-8Q94TV similar to Hypothetical protein ORF1 - *Catharanthus roseus* (Rosy periwinkle) (Madagascar periwinkle), mRNA sequence.	50	1.5E-119
HT24R-64	EM_EST:GT164709	1212	M82T1_07_C05_M13R LA3475 Type I trichomes *Solanum lycopersicum* cDNA, mRNA sequence.	50	2.1E-7

HT24R-60	emb|CAO65613.1	421	unnamed protein product [*Vitis vinifera*]	2	2.00E-24
HT24R-62	DV697501	409	CGN-42922 Leaf *Coffea canephora* cDNA clone cccl30f13 5', mRNA sequence	9	1.00E-102
HT24R-78	ref|XM_011094698.1	752	PREDICTED: uncharacterized protein LOC105173052	20	2.36E-44
HT24R-79	XR_847884.1	722	PREDICTED: *Sesamum indicum* uncharacterized LOC105162974 (LOC105162974), transcript variant X2, misc_RNA	20	1.00E-05
HT24R-83	XM_003522508.2	406	PREDICTED: Glycine max ER membrane protein complex subunit 7 homolog (LOC100816780), mRNA	20	3.00E-23
HT24R-86	DV679297	584	CGN-17448 Pericarp *Coffea canephora* cDNA clone cccp21o13 5', mRNA sequence	89	0.00
HT24R-87	gb|DQ834312.1	399	*Coffea canephora* sucrose synthase (SS2) mRNA, complete cds	15	8.00E-98
HT24R-90	EM_EST:DN573574	920	93841390 Sea Urchin primary mesenchyme cell cDNA library *Strongylocentrotus purpuratus* cDNA clone PMCSPR2-182C23 3', mRNA sequence.	50	4.1E-11
HT24R-92	DV707123	433	CGN-54749 Cherry of Early Development Stage *Coffea canephora* cDNA clone cccwc22w8i2 5',mRNA sequence	5	0.00
HT24R-98	DV705414	316	CGN-52483 Cherry of Early Development Stage *Coffea canephora* cDNA clone cccwc22w4b18 5',mRNA sequence	1	2.00E-65
HT24R-100	gb|ABK93877.1	289	unknown [*Populus trichocarpa*]	1	2.00E-06
HT24R-102	EE198120	778	CC-F01_011_J14 Cherry of different development stage *Coffea canephora* cDNA clone CC-F01_011_J14,mRNA sequence	18	0.00

HT24R-04	EM_EST:GR991938	249	TransId-111780 CarCTFrHh24FL *Coffea arabica* cDNA clone CarCTFrHh24FL_Q3_02_N21_334.F similar to 60S acidic ribosomal protein P2 (RPP2D), acidic ribosomal protein P2, maize	50	4.0E-8
HT24R-103	emb|CAO22157.1	317	unnamed protein product [*Vitis vinifera*]	1	1.00E-53
HT24R-106	DV667630	395	CGN-5938 Pericarp *Coffea canephora* cDNA clone cccp17a6 5', mRNA sequence	1	4.00E-73
HT24R-111	EM_EST:FL640959	425	PhSFL 13 Podophyllum hexandrum Royle suppression subtractive hybridization (SSH) library *Sinopodophyllum hexandrum* cDNA similar to Late embryogenesis abundant protein 1, mRNA sequence.	50	3.9E-13
HT24R-112	EE193325	615	CC-L01_021_G10 Coffea young leaves *Coffea canephora* cDNA clone CC-L01_021_G10, mRNA sequence	12	0.00
HT24R-113	EG328965	370	CR-EST327 Leaf EST Library of robusta coffee cultivar 'CxR' *Coffea congensis x Coffea canephora* cDNA clone C2J17R 3', mRNA sequence	114	1.00E-162
HT24R-114	EM_EST:EC887053	1132	ZM_BFc0026D13.f ZM_BFc *Zea mays* cDNA clone ZM_BFc0026D13 3', mRNA sequence.	50	1.2E-8
HT24R-115	EM_EST:EY871696	932	CL06-C4-500-064-F11-CT.F Rangpur lime root, greenhouse plant *Citrus limonia* cDNA, mRNA sequence.	50	4.0E-8
HT24R-120	DV672601	357	CGN-8297 Pericarp *Coffea canephora* cDNA clone cccp21p24 5', mRNA sequence	64	4.00E-132
HT24R-123	EM_EST:EV545608	718	RR3AU25JQ RR3(NY) *Raphanus raphanistrum* subsp. raphanistrum cDNA 3', mRNA sequence.	50	1.3E-13

HT24R-125	EG329030	330	CR-EST782 Leaf EST Library of robusta coffee cultivar 'CxR' *Coffea congensis* x *Coffea canephora* cDNA clone C5H5R 3', mRNA sequence	62	3.00E-153
HT24R-129	gb|ABA40468.1	822	Drm3-like protein [*Solanum tuberosum*]	4	4.00E-23
HT24R-130	GT647563	566	CC00-XX-SH3-007-C03-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-007-C03-EM,mRNA sequence	4	1.00E-174
HT24R-131	DV683630	554	CGN-24170 Seed of Middle Development Stage *Coffea canephora* cDNA clone cccs30w34j16 5',mRNA sequence	4	1.00E-174
Compatible Interaction - Upregulated at 12 h.a.i.					
A. Resistance and anti-microbial functions					
Cat12F-13	ref|XR_078082.1	994	PREDICTED: *Vitis vinifera* probable pectinesterase/pectinesterase inhibitor 6-like (LOC100252606), miscRNA	2	3.00E-81
Cat12F-15	EM_GT:008845.1	779	CAET42MIX *Coffea arabica* cDNA clone CAET42MIX-CFEYC88TV similar to Stress and pathogenesis-related protein - Fagus sylvatica (Beechnut), mRNA sequence	50	0.00
Cat12F-22	emb|CAN63832.1	1068	hypothetical protein [*Vitis vinifera*], response to bacteria	3	2.00E-12
Cat12F-25	dbj|BAD46374.1	437	protease II -like [*Oryza sativa* (japonica cultivar-group)]	1	2.00E-42
B. Signal induction and transduction					
Cat12F-04	emb|CAN82027.1	156	hypothetical protein [*Vitis vinifera*]	7	5.00E-55
C. Cell maintenance and homeostasis					

Cat12F-02	gb|KC570456.1	504	*Coffea arabica* metallothionein type 3 mRNA, complete cds	111	2.21E-17
Cat12F-07	emb|CAO46372.1	267	unnamed protein product [*Vitis vinifera*]	20	1.00E-138
Cat12F-08	gb|ABY59947.1	764	nectar protein 1 [*Jacaranda mimosifolia*]	2	1.00E-165
Cat12F-09	emb|CAN80552.1	523	hypothetical protein [*Vitis vinifera*]	3	6.00E-80
Cat12F-12	XM_010315925.1	345	PREDICTED: *Solanum lycopersicum* probable ADP-ribosylation factor GTPase-activating protein AGD11 (LOC101247408), mRNA	20	5.81E-28
Cat12F-14	emb|FQ392600.1	422	*Vitis vinifera* clone SS0AFA5YA07	8	2.00E-103
Cat12F-18	dbj|AB236867.1	681	*Panax ginseng* cab mRNA for chlorophyll a/b binding protein, complete cds	63	0.00
Cat12F-21	ref|XM_006484938.1	1007	PREDICTED: *Citrus sinensis* oxygen-evolving enhancer protein 2, chloroplastic-like (LOC102625749), mRNA	300	4.00E-115
Cat12F-26	gb|JQ693578.1	480	*Corchorus capsularis* voucher 890 xyloglucan endotransglycosylase/hydrolase (XTH1) gene, partial cds	15	3.00E-53
Cat12F-27	gb|AF343968.1	1050	*Coffea arabica* phosphoglycerate kinase erki 8 mRNA, partial cds	1	7.00E-93
Cat12F-28	emb|CAN75197.1	603	hypothetical protein [*Vitis vinifera*]	4	1.00E-80
Cat12F-29	gb|JF747157.1	1190	*Dimocarpus longan* chloroplast chlorophyll A/B binding protein (LCAB3) mRNA, complete cds; nuclear gene for chloroplast product	1	6.00E-24
Cat12F-30	emb|CAO21633.1	642	unnamed protein product [*Vitis vinifera*]	3	1.00E-17
Cat12F-32	emb|CAO63570.1	1073	unnamed protein product [*Vitis vinifera*]	4	0.00

Cat12F-34	ref|XM_007136498.1	695	*Phaseolus vulgaris* hypothetical protein (PHAVU_009G055100g) mRNA, complete cds	1	1.00E-39
D. No GO term associated					
Cat12F-01	DV675973	216	CGN-13170 Pericarp *Coffea canephora* cDNA clone cccp1b1 5', mRNA sequence	1	7.00E-103
Cat12F-03	dbj|AK371956.1	310	*Hordeum vulgare* subsp. vulgare mRNA for predicted protein, complete cds, clone: NIASHv2143O24	1	3.00E-12
Cat12F-05	ref|XM_007144935.1	745	*Phaseolus vulgaris* hypothetical protein (PHAVU_007G201000g) mRNA, complete cds	2	1.00E-59
Cat12F-06	EM_EST:GT726529	445	CA00-XX-EA1-006-F03-EC.F *Coffea arabica* EA1 *Coffea arabica* cDNA clone CA00-XX-EA1-006-F03-EC,	50	0.00
Cat12F-10	ref|XM_004246195.1	617	PREDICTED: *Solanum lycopersicum* auxilin-related protein 1-like (LOC101246463), mRNA	496	1.45E-70
Cat12F-11	XM_011072309.1	387	probable adp-ribosylation factor gtpase-activating protein agd11	20	7.19E-08
Cat12F-17	DV693056	465	CGN-37125 Leaf *Coffea canephora* cDNA clone cccl26l22 5', mRNA sequence	22	0.00
Cat12F-19	ref|XM_003577956.1	1088	PREDICTED: *Brachypodium distachyon* ubiquitin domain-containing protein 2-like (LOC100827721), mRNA	6	1.96E-27
Cat12F-20	emb|CAO39237.1	415	unnamed protein product [*Vitis vinifera*]	2	9.00E-31
Cat12F-23	XM_009784895.1	622	PREDICTED: *Nicotiana tomentosiformis* uncharacterized LOC104098147 (LOC104098147), mRNA	20	7.21E-20

Cat12F-24	emb|CAO47910.1	1061	unnamed protein product [*Vitis vinifera*]	41	4.00E-69
Cat12F-33	ref|XM_008246618.1	660	PREDICTED: *Prunus mume* haloacid dehalogenase-like hydrolase domain-containing protein 3 (LOC103342953), mRNA	8	4.65E-38
Cat12F-35	ref|XM_008232898.1	1071	PREDICTED: *Prunus mume* B2 protein (LOC103330336), transcript variant X4, mRNA	30	1.00E-75
Compatible Interaction - Downregulated 12 h.a.i.					
B. Signal induction and transduction					
Cat12R-07	emb|CAO65122.1	332	unnamed protein product [*Vitis vinifera*]	1	1.00E-79
Cat12R-13	gb|KF743541.1	368	*Coffea arabica* ERF (Ethylene Response Factor) transcription factor 02 mRNA, complete cds	1	3.00E-60
Cat12R-15	emb|CAO18274.1	294	unnamed protein product [*Vitis vinifera*]	2	2.00E-98
Cat12R-18	emb|CAO66017.1	399	unnamed protein product [*Vitis vinifera]*	5	1.00E-177
Cat12R-19	emb|CAO67538.1	471	unnamed protein product [*Vitis vinifera*]	5	1.00E-110
C. Cell maintenance and homeostasis					
Cat12R-01	gb|ABW35320.1	363	photosystem ii 10 kda chloroplastic	20	2.94E-32
Cat12R-05	emb|FQ379540.1	1115	unnamed protein product [*Vitis vinifera*]	4	0.00
Cat12R-06	emb|CAN70526.1	908	d-glycerate 3- chloroplastic	3	4.38E-169
Cat12R-09	emb|CAO62015.1	306	atp-dependent zinc metalloprotease chloroplastic-like	20	5.50E-12
Cat12R-11	gb|DQ124032.1	668	*Coffea arabica* clone MN-SSH3-H07 mRNA sequence	1	4.73E-16

Cat12R-12	emb|CAA29123.1	269	histone h1-like	20	4.00E-14
Cat12R-16	gb|JF410859.1	1041	*Coffea arabica* cultivar Catuai Vermelho chitinase-like xylanase inhibitor protein (clxip) mRNA, partial cds	2	0.00
Cat12R-17	ref|XM_008221130.1	837	PREDICTED: *Prunus mume* 29 kDa ribonucleoprotein B, chloroplastic (LOC103319578), mRNA	2	7.00E-22
Cat12R-23	gb|DQ401313.1	407	*Coffea arabica* ATPase alpha subunit (atp1) gene, partial cds; mitochondrial	494	2.00E-178
Cat12R-26	emb|CAO65599.1	942	unnamed protein product [*Vitis vinifera*]	1	7.4E-70
Cat12R-28	emb|CAO69143.1	1040	btb poz and math domain-containing protein 3-like isoform x1	20	1.66E-56
Cat12R-29	emb|CAO46360.1	977	unnamed protein product [*Vitis vinifera*]	4	2.00E-85
Cat12R-30	gb|DQ124083.1	581	Coffea arabica clone MN-SSH3-E02 mRNA sequence (metallothioneine-like protein type-3, in other spps)	110	0.00
Cat12R-31	emb|CAN82190.1	693	hypothetical protein [*Vitis vinifera*]	2	1.00E-38
Cat12R-32	sp|P54767	453	DCE_LYCES Glutamate decarboxylase (GAD) (ERT D1)	2	3.00E-17
Cat12R-37	emb|CAO44932.1	801	unnamed protein product [*Vitis vinifera*]	1	5.00E-29
Cat12R-39	emb|CAN76185.1	885	hypothetical protein [*Vitis vinifera*]	5	0.00
D. No GO term associated					
Cat12R-03	EM_EST:GT677672	575	CA00-XX-CL2-012-F04-AC.F *Coffea arabica* CL2 *Coffea arabica* cDNA clone CA00-XX-CL2-012-F04-AC, mRNA sequence	50	1.0E-47

Cat12R-08	EM_EST:GT682795	385	CA00-XX-CL2-104-B11-JE.F *Coffea arabica* CL2 *Coffea arabica* cDNA clone CA00-XX-CL2-104-B11-JE, mRNA sequence.	50	9.6E-115
Cat12R-14	emb|CAO67543.1	741	PREDICTED: uncharacterized protein LOC104227401	20	7.29E-86
Cat12R-20	EM_EST:JK614459	305	GUCCSH1014F06.b ESTs from a SSH Library for drought stress tolerance in *Phaseolus vulgaris* L. roots cv. GUCCSH1014F06 5', mRNA sequence.	50	3.5E-18
Cat12R-21	gb|KF278731.1	154	*Coffea arabica* ethylene response factor 5 mRNA, complete cds	1	1.00E-12
Cat12R-25	EM_EST:GT716107	510	CA00-XX-SS1-008-E06-EF.F *Coffea arabica* SS1 *Coffea arabica* cDNA clone CA00-XX-SS1-008-E06-EF, mRNA sequence.	50	0.00
Cat12R-33	DV666302	311	CGN-4132 Pericarp *Coffea canephora* cDNA clone cccp15c20 5', mRNA sequence	1	4.00E-97
Cat12R-36	emb|CAO61050.1	449	unnamed protein product [*Vitis vinifera*]	3	8.00E-63
Compatible Interaction - Upregulated at 24 h.a.i.					
A. Resistance and anti-microbial functions					
Cat24F-01	XM_002280758.3	715	pyroglutamyl-peptidase 1-like	20	4.81E-115
Cat24F-13	ref|XM_002526142.1	474	*Ricinus communis* protease C56, putative, mRNA	3	4.00E-27
Cat24F-30	dbj|BAF37542.1	198	cell wall glycine-rich protein [*Cucumis sativus*]	7	1.00E-06
Cat24F-44	ref|XM_007035684.1	617	*Theobroma cacao* Chitinase-like protein 2 (TCM_021326) mRNA, complete cds	8	2.00E-51
Cat24F-49	gb|DQ123920.1	330	catalase isozyme 1	20	1.21E-27

Cat24F-50	emb|CAO66235.1	312	unnamed protein product [*Vitis vinifera*]	5	4.00E-35
B. Signal induction and transduction					
Cat24F-28	gb|KF743542.1	637	*Coffea arabica* ERF transcription factor 03 mRNA, complete cds	32	4.00E-163
Cat24F-41	ref|XM_009152715.1	719	serine threonine-protein kinase endoribonuclease ire1a-like isoform x2	20	3.45E-127
C. Cell maintenance and homeostasis					
Cat24F-02	gb|KC570456.1	606	*Coffea arabica* metallothionein type 3 mRNA, complete cds	111	0.00
Cat24F-03	ref|XM_009764010.1	416	PREDICTED: *Nicotiana sylvestris* beta-galactosidase-like (LOC104214355), mRNA	2	1.00E-17
Cat24F-15	emb|CU231910.1	343	Populus EST from mild drought-stressed leaves	14	2.00E-64
Cat24F-16	emb|CAO62888.1	458	polyketide cyclase dehydrase and lipid transport superfamily protein	20	1.10E-11
Cat24F-21	ref|XR_182964.1	781	PREDICTED: *Solanum lycopersicum* beta-galactosidase-like (LOC101263334), misc_RNA	51	8.00E-91
Cat24F-23	gb|DQ401313.1	409	*Coffea arabica* ATPase alpha subunit (atp1) gene, partial cds; mitochondrial	495	3.00E-157
Cat24F-24	ref|XM_011100146.1	423	*Sesamum indicum* glutaredoxin (LOC105177111), transcript variant X4, mRNA	20	5.68E-28
Cat24F-29	emb|CAC01237.1	326	RNA Binding Protein 45 [*Nicotiana plumbaginifolia*]	2	4.00E-34
Cat24F-32	gb|DQ834313.1	399	*Coffea canephora* sucrose phosphatase (SP1) mRNA, complete cds	1	8.00E-138
Cat24F-37	XM_006441085.1	614	ring u-box superfamily protein isoform 1	20	4.00E-21

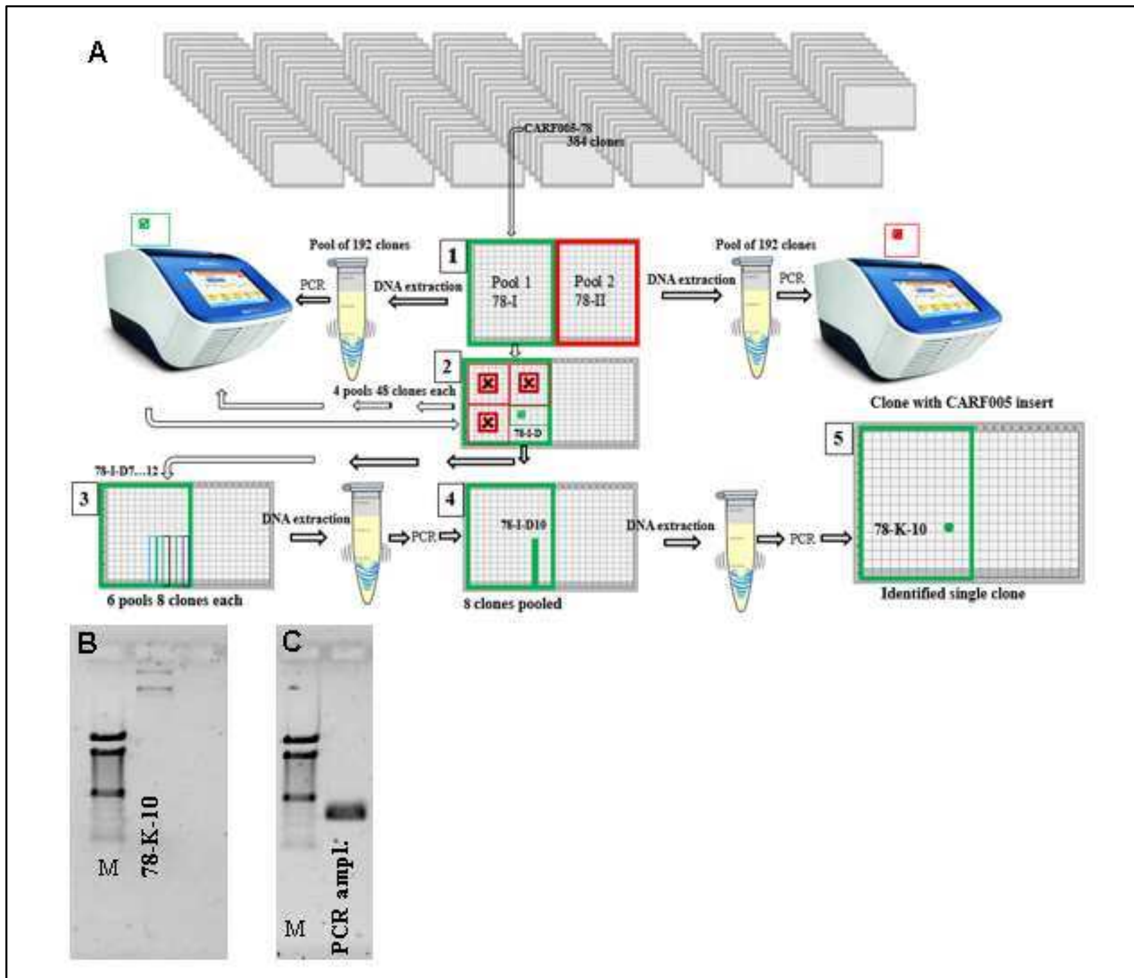
Cat24F-42	gb|AAF76226.1	166	AF272572_1 14-3-3 protein [Populus x canescens]	1	1.00E-125
Cat24F-43	gb|ABN08775.1	613	Glycoside hydrolase, family 19	2	1.00E-133
Cat24F-47	KJ796402.1	446	*Vitis vinifera* RING-type E3 ubiquitin ligase (GW2) mRNA, complete cds	20	1.00E-19
Cat24F-59	gb|HQ130481.1	890	*Coffea arabica* cultivar Catuai chloroplast chlorophyll a/b-binding photosystem II 22kDa subunit S (PsbS) mRNA, partial cds; nuclear gene for chloroplast product	20	2.00E-173
Cat24F-60	XM_011073867.1	644	f-box protein pp2-a12-like	20	5.60E-44
Cat24F-61	dbj|BAF98176.1	680	trehalose-6-phosphate synthase [*Solanum lycopersicum*]	2	2.00E-29
D. No GO term associated					
Cat24F-04	gb|DQ124082.1	552	*Coffea arabica* clone MN-SSH1-C02 mRNA sequence	111	0.00
Cat24F-09	GT655832	641	CC00-XX-EC1-029-D07-EC.F *Coffea canephora* EC1 *Coffea canephora* cDNA clone CC00-XX-EC1-029-D07-EC,mRNA sequence	89	0.00
Cat24F-27	EM_EST:GT014225	340	TransId-254506 CAET1425FR28 *Coffea arabica* cDNA clone CAET1425FR28-CAETM53TV similar to Glutaredoxin - Tilia platyphyllos (Large-leaved lime), mRNA sequence.	50	2.2E-174
Cat24F-34	gb|AAM34773.1	299	nam-like protein 10 [Petunia x hybrida]	1	1.00E-125
Cat24F-48	emb|CAO61141.1	447	unnamed protein product [*Vitis vinifera*]	4	8.00E-33
Cat24F-51	DV709635	404	CGN-57990 Cherry of Early Development Stage *Coffea canephora* cDNA clone cccwc22w20d1 5',mRNA sequence	2	1.00E-92

Cat24F-53	gb|JQ693578.1	604	*Corchorus capsularis* voucher 890 xyloglucan endotransglycosylase/hydrolase (XTH1) gene, partial cds	16	4.00E-53
Cat24F-56	GT654547	532	CC00-XX-EC1-042-D11-EC.F *Coffea canephora* EC1 *Coffea canephora* cDNA clone CC00-XX-EC1-042-D11-EC,mRNA sequence	1	5.00E-153
Cat24F-65	GT649010	841	CC00-XX-SH3-044-B12-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-044-B12-EM,mRNA sequence	41	0.00
Compatible Interaction - Downregulated 24 h.a.i.					
C. Cell maintenance and homeostasis					
Cat24R-01	ref|XM_006353619.1	795	PREDICTED: *Solanum tuberosum* ATP synthase subunit b', chloroplastic-like (LOC102584964), mRNA	9	8.00E-101
Cat24R-02	ref|XM_008240028.1	1040	PREDICTED: *Prunus mume* peroxidase 73-like (LOC103336895), mRNA	3	7.00E-88
Cat24R-05	emb|AJ419826.1	242	*Coffea arabica* mRNA for rubisco small subunit (rbcs1 gene)	1	1.00E-113
Cat24R-06	ref|XM_002305653.2	781	*Populus trichocarpa* hypothetical protein (POPTR_0004s06790g) mRNA, complete cds	2	1.00E-89
Cat24R-08	ref|NP_177742.2	468	maoC-like dehydratase domain-containing protein [*Arabidopsis thaliana*]	1	1.00E-109
Cat24R-10	gb|DQ124036.1	814	*Coffea arabica* clone MN-SSH1-A01 mRNA sequence	1	1.00E-24
Cat24R-14	gb|KP074964.1	593	*Vitis vinifera* serine acetyltransferase 2;2 (SERAT) mRNA, complete cds	20	1.00E-52
Cat24R-15	gb|EF044213.1	560	*Coffea arabica* chloroplast, complete genome	20	0.00

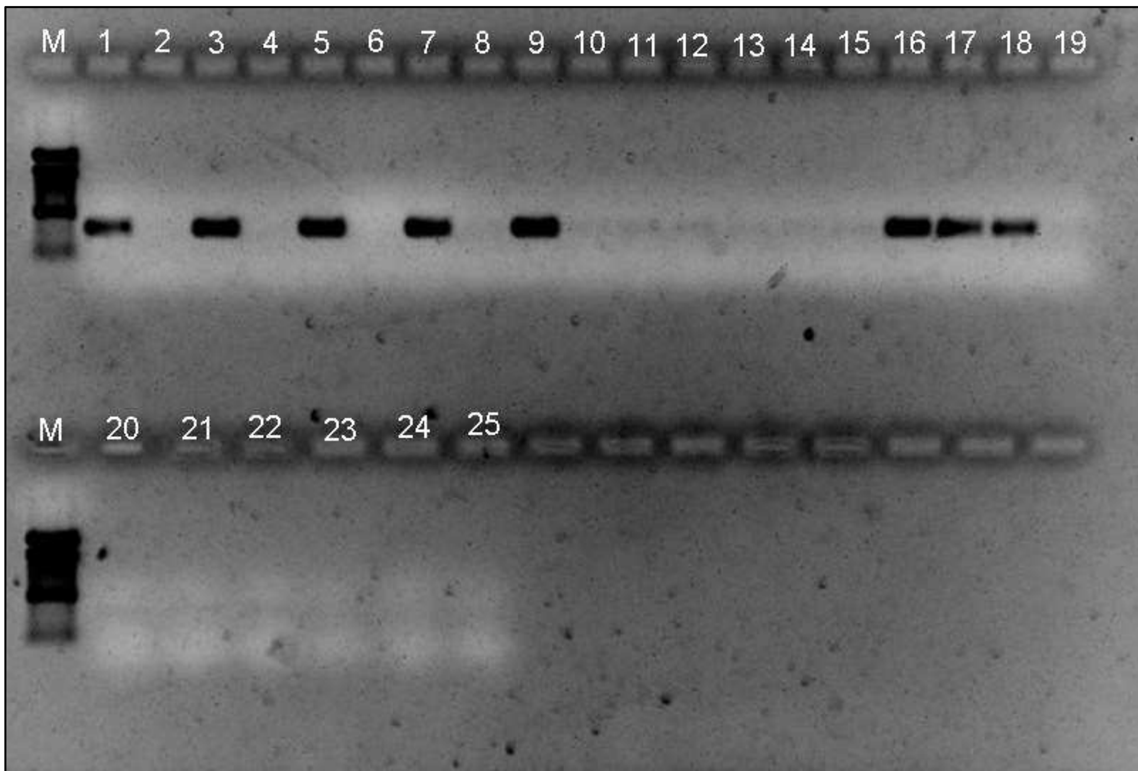
Cat24R-16	dbj|D26578.1	738	*Daucus carota* mRNA for DNA-binding protein (transcriptional regulator), partial cds, CHB6	2	4.00E-74
Cat24R-17	ref|XM_006584572.1	739	PREDICTED: Glycine max homeodomain-leucine zipper protein 56 (HDL56), transcript variant X1, mRNA	2	4.00E-44
D. No GO term associated					
Cat24R-12	GT647948	489	CC00-XX-SH3-047-E10-EM.F *Coffea canephora* SH3 *Coffea canephora* cDNA clone CC00-XX-SH3-047-E10-EM,mRNA sequence	90	0.00

[a]Gene bank identification number .
[b]Description of a sequence with the highest similarity (lowest E-value).
[c]Number of hits with minimum e-value was taken from the databases with significant match.

8. APPENDIX II



Appendix II Figure 1. Work flow in BAC clone screening. Clone pooling and subsequent group decomposition to isolate a single clone with CARF005 insert (A), DNA of isolated clone 78-K-10 (B) and CARF005 PCR amplicon (C) as revealed by 1% UltraPure^{TM} agarose gel electrophoresis. M is 100 bp DNA size marker.



Appendix II Figure 2. The 22 differential coffee clones screened for CARF005 marker (listed in order as in Table 1). Clones with CARF005 were 1 (832/1-HT), 3 (1343/269-HT), 5 (H420/10), 7 (128/2-Dilla & Alghe), 9 (H419/20), 16 (4106), 17 (644/18 H. Kawisan, a new report) and 18 (832/2-HT). M: DNA weight marker ladder (the lightest band being 100 bp).

Appendix II Table 1. Top hits for the 13 ORFs as found in NCBI by BLASTp or at *C. canephora* genome by BLASTn.

Gene (Query)	Top hit (species)	Subject ID/Accession numbers*	Description	E-value	Identity (%)
1	*Coffea arabica*	-	no description	3.00E-176	79
1	*Phyllostachys edulis*	ADB85290.1	putative retrotransposon protein	9.00E-94	54
2	*Coffea canephora*	-	no description	1.00E-26	76
3	*Solanum lycopersicum*	XP_010322277.1	PREDICTED: uncharacterized protein LOC104647985	1.00E-72	50
3	*Coffea arabica*	-	no description	1.00E-60	68
4	*Coffea canephora*	CDP13079.1	unnamed protein product	9.00E-05	91
5	*Coffea canephora*	CDP20136.1	unnamed protein product	0.0	91
5	*Solanum tuberosum*	XP_015160818.1	PREDICTED: putative disease resistance protein RGA4	0.0	38
6	*Coffea canephora*	-	no description	2.00E-73	68
7	*Coffea canephora*	CDP13085.1	unnamed protein product	0.0	98
7	*Cynara cardunculus*	KVI07273.1	Chloramphenicol acetyltransferase-like domain-containing protein	2.00E-99	40
8	*Coffea canephora*	-	No description	2.00E-34	84
9	*Coffea canephora*	CDP13077.1	unnamed protein product	1.00E-35	67
9	*Capsicum annuum*	XP_016542041.1	PREDICTED: protein ALWAYS EARLY 3-like isoform X2	2.00E-10	58
10	*Coffea canephora*	GAQ44625.1	Unnamed RGA fragment	5.00E-11	74
11	*Coffea canephora*	CDP20093.1	unnamed protein product	0.0	91
11	*Solanum tuberosum*	XP_015160818.1	PREDICTED: putative disease resistance protein RGA4	0.0	37
12	*Coffea canephora*	-	Putative disease resistance protein RGA3 complete	1.00E-45	72
13	*Coffea canephora*	-	no description	4e-29	77

*Homologous sequences for which no ID/Accession number has been assigned are indicated in hyphen.
BLASTp was performed by NCBI online server (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins).

Appendix II Table 2. Mutation (substitution) effect on protein binding regions of gene 5 and 11 indicated by amino acid sequence in respective genes.

Protein binding sites*	High effect	Moderate effect	No effect	Amino acid residues in the protein binding sites
Gene 5				
93-97			X	YRNQP
100-101		X		KK
147			X	K
299		X		T
320		X		HLQM
455-457			X	REN
483-485			X	DEG
510-512	X			QDR
528-529			X	RN
636-638	X			DDT
664			X	K
719-720	X			QK
747	X			M
821			X	R
973		X		E
1101		X		R
1123-1124	X			QR
Gene 11				
93-95	X			YRN
100-101	X			KK
147-148			X	KE
248			X	T
321			X	L
415-416		X		KN
442-445		X		DKLR
470		X		D
620-621	X			EK
672-673			X	RK
729			X	R
784-785		X		RR
1099-110			X	RR
1111-1112		X		QR

*Hyphen indicates range of amino acids constituting the binding site.
Yellow highlighted residues are conserved residues in both genes while purple highlighted residues are specific protein binding sites in respective gene. Substitution mutation effect analysis was performed by The Predict Protein Server (Yachdav et al. 2014).

Appendix II Table 3. Output of the two contigs BLASTed against S_H3 locus contigs specific to *C. arabica* and *C. canephora**.

Query	Subject accession number	Identity (%)	Alignment length	Mismatch	Gap open	Query start	Query end	Subject start	Subject end	e-value	Bit score	Contig serial number
Contig3	gb|HQ696509.1|	88.247	485	56	1	11920	12403	124032	124516	4.33e-175	614	*C. arabica*_7
Contig3	gb|GU123899.1|	88.795	473	52	1	11929	12400	208422	208894	5.28e-174	610	*C. arabica*_6
Contig3	gb|HQ696513.1|	87.243	486	59	3	11920	12403	186601	187085	6.02e-167	587	*C. canephora*_3
Contig3	gb|HQ696507.1|	87.243	486	59	3	11920	12403	13356	13840	6.02e-167	587	*C. canephora*_1
Contig3	gb|GU123895.1|	72.468	385	69	15	12583	12953	101631	101992	2.41e-45	183	*C. arabica*_2
Contig3	gb|GU123894.1|	70.698	430	63	18	12583	12965	47408	47821	2.41e-45	183	*C. arabica*_1
Contig3	gb|HQ696512.1|	81.915	188	25	4	12583	12769	92911	92732	3.58e-43	176	*C. canephora*_2
Contig3	gb|GU123898.1|	68.343	338	53	13	12583	12907	84162	83866	8.43e-26	118	*C. arabica*_5
Contig3	gb|HQ696510.1|	73.913	207	36	7	12580	12786	120484	120296	1.03e-24	114	*C. arabica*_8
Contig3	gb|GU123897.1|	64.972	354	64	11	12600	12931	93064	92749	1.74e-15	84.2	*C. arabica*_4
Contig3	gb|HQ696511.1|	75.410	122	28	2	12283	12404	15692	15573	7.39e-14	78.8	*C. arabica*_9
Contig3	gb|HQ696508.1|	87.755	49	6	0	12583	12631	45396	45444	5.68e-09	62.6	*C. arabica*_10
Contig3	gb|GU123896.1|	80.303	66	13	0	12583	12648	100998	101063	1.98e-08	60.8	*C. arabica*_3
Contig9	gb|GU123898.1|	77.647	170	38	0	7794	7963	258864	258695	1.57e-31	136	*C. arabica*_5
Contig9	gb|HQ696508.1|	76.536	179	38	4	7771	7946	68305	68482	1.21e-26	120	*C. arabica*_10

*Ten contigs specific to *C. arabica* and three contigs specific to *C. canephora*, all assembled from BAC clones with S_H3 locus were taken from the work of Ribas et al. (2011).