Evolution of Disease-resistance Gene in Coffee Trees (*Coffea* L.)

S. NOIR[1], M.-C. COMBES[1], François ANTHONY[2] P. LASHERMES[1,*]

[1]IRD, GeneTrop, BP 5045, F-34032, Montpellier, France
[2]IRD/CATIE, 7170 Turrialba, AP59, Costa Rica
*to whom correspondence should be addressed - Fax: 33 4 67 41 63 30
E-mail: Philippe.Lashermes@mpl.ird.fr

SUMMARY

The ability of plant species to survive over evolutionary time depends to some instance on their capability to maintain and generate useful diversity at resistance loci. Therefore, analysis of diversity and evolution of resistance genes (R) in plant species could be of particular interest in the elaboration of strategies for developing durable resistance mediated by major genes. Molecular analysis of R-gene evolution concern so far mainly plants with short life cycle while perennial plants have retained little attention.

The majority of R-genes isolated so far encode a predicted nucleotide-binding site (NBS) domain. NBS domains related to R-genes show a highly conserved backbone of amino acid motifs offering the possibility of isolating resistance gene analogous sequences (RGAs) by polymerase chain reaction (PCR). Multiple combinations of primers with low degeneracy designed from two conserved motifs in the NBS regions of R-genes of various plants were used in coffee trees. Nine distinct classes of NBS-like resistance gene analogs (RGAs) representing a large diversity were isolated from *Coffea arabica* and *C. canephora* species. The analysis of one coffee RGA family suggested point mutations as the primary source of diversity. With one exception, coffee RGA families appeared closely related by sequence to at least one cloned R-gene. In addition, deduced amino acid sequences of coffee RGAs were identified showing strong sequence similarity with almost all known non-TIR type R-genes. The high similarity between particular coffee RGAs and R-genes isolated from other angiosperm species such as *Arabidopsis*, tomato and rice indicated an ancestral relationship and the existence of common ancestors. As revealed in coffee trees, the evolution of NBS-encoding sequences seems to involve accumulation and slow divergence mechanisms within distinct R-gene families rather than a fast-evolving process. Functional inference of the suggested NBS domain type of evolution is also discussed.

INTRODUCTION

In the gene-for-gene resistance, the rapid changes that occur in the virulence characteristics of pathogen populations raise a continuous threat to the effectiveness of individual major genes of resistance (R). Consequently, the ability of plant species to survive over evolutionary time depends to some instance on their ability to maintain and generate useful diversity at resistance loci. Therefore, analysis of diversity and evolution of R-genes in plant species could be of particular interest in the elaboration of strategies for developing durable resistance mediated by major genes.

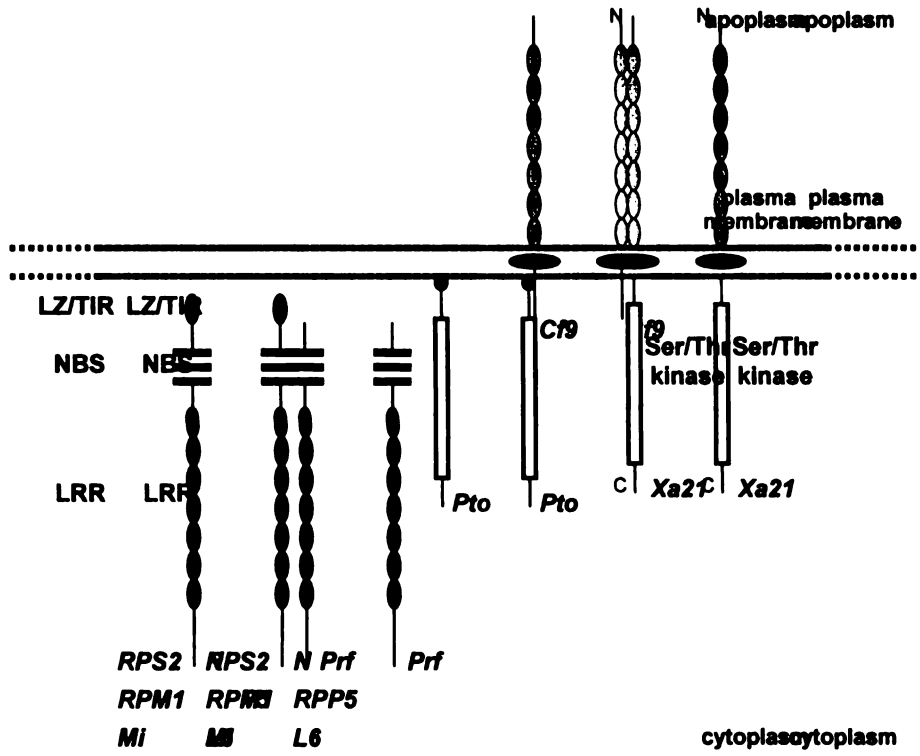

Figure 1. Schematic representation and putative location of conserved protein domains of cloned R-genes. Five mains peptidic conserved domains were distinguished: NBS (Nucleotide Binding Site), LRR (Leucine Rich Repeat), TIR (Toll-Interleukin Receptor), LZ (Leucine Zipper) and Serine/Threonine Kinase

A growing number of disease resistance genes conferring resistance to a wide range of pathogens have recently been isolated from several plant species. Analysis of the sequences of some R-genes revealed conservation of specific amino acid domains in the putative products (Figure 1). Otherwise, remarkable similarities in the general structure of R loci have also been observed. They are generally members of multigene families and show a complex physical organization of repeated sequences (i.e. cluster). These recent data have shed light on the molecular evolution of R-genes. The organization of these R-gene families suggests indeed that novel sequences and therefore novel specificities are generated by various evolutionary events such as substitutions, different mechanisms of recombination (i.e. unequal crossing-over, gene conversion) and more exceptionally, transposable elements (2). However, these analysis concern mainly plants with short life cycle while perennial plants have retained so far little attention.

Coffea arabica, an important tropical crop, is characterized by a low genetic diversity. This allotetraploid species shows a high susceptibility to many pests and diseases, and the diploid species *C. canephora* constitutes the main resistance source for breeding purposes. Our project displays two principal objectives:

- to precise the molecular organization and the evolution of R-genes in this perennial plant
- to improve the coffee tree for disease resistance, especially against *Meloidogyne* root-knot nematodes and *Hemileia vastatrix* rust fungi.

MATERIALS AND METHODS

Plant material

Plant material involved the accessions of both species *C. arabica* and *C. canephora*. Genomic DNA and mRNA extracted from leaves were used.

Amplification with degenerate primers

NBS domains related to R-genes show a highly conserved backbone of amino acid motifs offering the possibility of isolating resistance gene analogous sequences (RGAs) by polymerase chain reaction (PCR) with degenerate primers. Multiple combinations of primers with low and without degeneracy were designed from two conserved motifs (i.e. P-Loop and GLPL) in the NBS regions of R-genes of various plants. These primers were used in PCR amplification from coffee genomic DNA. The amplified products were cloned and sequenced.

Amplification with coffee RGA family-specific primers

Primers specific to the identified coffee RGA families were tested on mRNA samples by RT-PCR analysis.

RESULTS AND DISCUSSION

Isolation of coffee RGAs

Twenty five combinations of degenerate or non-degenerate primers were tested. From the amplified products showing the expected size considering the NBS domain length of known R-genes (~500 bp), 120 clones were isolated and sequenced. On the base of several features, 40 PCR-derived coffee NBS sequences were identified as RGAs. These sequences contained uninterrupted open reading frames. Moreover, all sequences contained the characteristic conserved motifs of NBS R-genes. Most coffee RGAs were closely related by sequence to at least one known R-gene.

A high coffee RGA diversity

The NBS-encoding RGAs isolated from coffee trees showed considerable sequence variation.

Nine distinct families of NBS-like RGAs were identified in both *C. arabica* and *C. canephora* species (Figure 2). More particularly, the analysis of one coffee RGA family suggested point mutations as the primary source of diversity. Moreover, by transcription analysis based on RT-PCR experiments, cDNAs corresponding to 8 different coffee RGA families were detected in coffee leaves of both studied species. Otherwise, RGAs belonging to the 9 distinguished families were observed in a same individual.

Coffee RGA phylogenetic analysis

Phylogenetic relationships between deduced amino acid coffee RGA sequences and a representative set of NBS domains of known R-gene products (isolated from Arabidopsis, tomato, potato, pepper, lettuce, maize and rice) were investigated. According to the reported distinction between the TIR class and the non-TIR class of R-genes (3), all isolated coffee RGAs seemed to belong to the non-TIR class type of R-genes (Figure 3). In addition, all non-TIR type NBS domains of R-genes considered in this study (excepted Dm3) were associated to one of the isolated coffee RGA families. Lastly, the alignment between coffee RGA peptidic sequences of a particular family and the NBS domain of R-genes related to this

family shows that similarities were shared beyond characteristic conserved NBS motifs of R-genes (Figure 4).

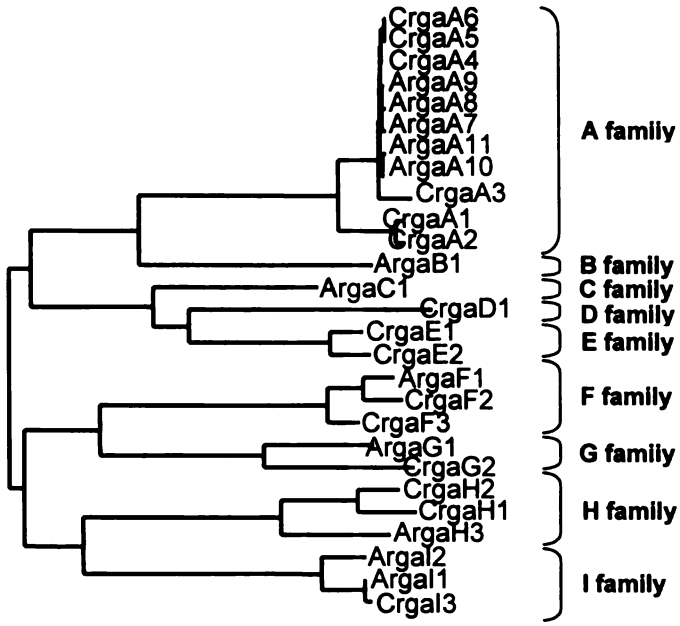

Figure 2. Phylogenetic tree for nucleotide RGA sequences isolated from *C. arabica* (A) and *C. canephora* (C) species. This tree was constructed by Neighbor-Joining method. Coffee RGA families (high sequence identity) are labeled A to I

CONCLUSIONS

In the coffee tree, the genetic diversity analysis between RGA families suggested an independent evolution of these different families. In this perennial plant, the evolution of NBS-encoding sequences seems to involve accumulation and slow divergence mechanisms within distinct R-gene families rather than a fast-evolving process. Otherwise, the high similarity between particular coffee RGAs and R-genes isolated from other angiosperm species such as *Arabidopsis*, tomato and rice indicated a common ancestral origin. This also supports a duplication and primary diversification of NBS-LRR R-genes more ancient than the divergence of monocots and dicots. The maintenance since dicot differentiation of different R-gene families showing sequence divergence and specific signature might result in fitness superiority and suggests a functional family-specificity.

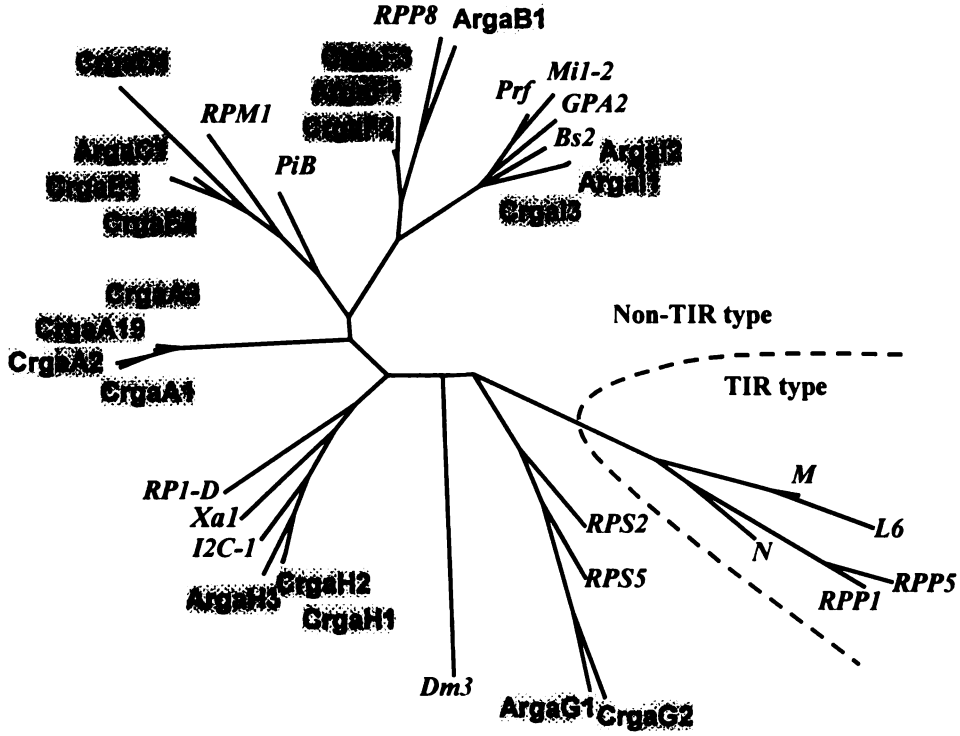

Figure 3. Neighbor-Joining tree based on alignment of amino acid sequences of representative coffee RGAs and NBS domains of cloned R-genes. Coffee RGAs are shaded gray. The dotted line distinguishes TIR and non-TIR type sequences

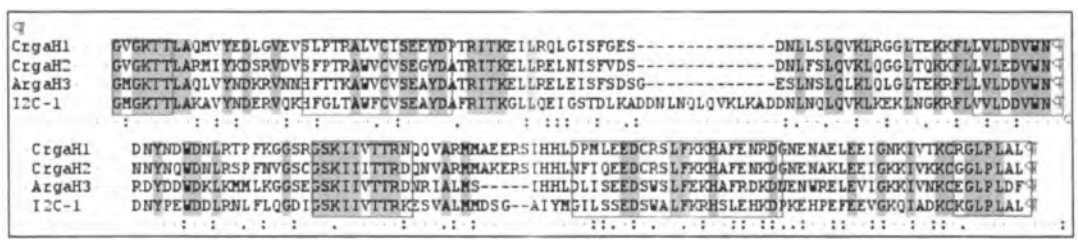

Figure 4. Multiple amino acid sequence alignment of coffee RGAs of the H family and the NBS domain of the closely related R-gene, *I2C-1*. Strict consensus residues are shown by shading. Residues sharing high (:) or low (.) physico-chemical properties are specified. Sequence blocks marked with boxes correspond to conserved motifs of NBS-encoding regregion R-genes

REFERENCES

Ellis J., Dodds P., Pryor T. (2000) Current Opinion in Plant Biology, **3**: 278-284.

Noir S., Combes M.-C., Anthony F., Lashermes P. (2001) Molecular Genetics and Genomics, DOI 10.1007/s004380100459.

Pan Q., Wendel J., Fluhr R. (2000) Journal of Molecular Evolution, **50**: 203-213.