Phylogenetic Origins and Expression Analysis of a Duplicated *WRKY* Gene in the Polyploid Species *Coffea arabica*

A. S. PETITOT, A. C. LECOULS, AND D. FERNANDEZ

IRD - Institut de Recherche pour le Développement, UMR1097 DGPC (IRD/INRA/CIRAD), BP64501, 34032 Montpellier Cedex 5, France

SUMMARY

C. arabica is an allotetraploid species resulting from a recent hybridization between two wild diploids *Coffea* species. The merger of two genomes in a common nucleus may be accompanied by considerable genomic reorganization. In particular, gene redundancy may lead to gene silencing or to the functional divergence of duplicated genes. We recently identified a *C. arabica* gene (*CaWRKY1*) encoding a transcription factor of the WRKY superfamily associated with plant defense responses to pathogens. We describe here the isolation in the *C. arabica* genome of two *CaWRKY1* gene copies (*CaWRKY1a* and *CaWRKY1b*) that probably originate from each ancestral subgenome of *C. arabica*. *CaWRKY1a* and *CaWRKY1b* shared only 94.7% nucleotide identity and 98.5% aminoacids identity over the whole gene sequence but a strict conservation of the DNA-binding region (WRKY domain) was retained. To assess the phylogenetic relationships between *CaWRKY1a* and *CaWRKY1b*, a comparative sequence analysis of the *WRKY1* gene was conducted in 6 related diploid *Coffea* species, including the putative *C. arabica* ancestral progenitors, *C. canephora* and *C. eugenioides*. PCR cloning, sequencing and phylogenetic analysis of the *Coffea WRKY1* genes showed that *CaWRKY1a* grouped together with the *C. canephora* and *C. congensis WRKY1* genes whereas *CaWRKY1b* was genetically closer to the *C. eugenioides WRKY1* gene. To verify whether *CaWRKY1a* and *CaWRKY1b* were both functional in the *C. arabica* genome, gene expression studies were conducted using real-time quantitative reverse-transcription Polymerase Chain Reaction (qRT-PCR). Monitoring of the *CaWRKY1a* and *CaWRKY1b* mRNA levels during plant infection with the rust fungus *Hemileia vastatrix* and wounding or salicylic acid treatments showed that both genes were simultaneously and equally expressed. Overall, results evidenced that *CaWRKY1a* and *CaWRKY1b* are homoeologous sequences in the polyploid *C. arabica* genome, and are probably descents of the *C. canephora* and *C. eugenioides WRKY1* genes. Both genes contributed to the transcriptomic expression of plant defense responses to pathogens, suggesting that no functional divergence had occurred between the duplicated genes in the *C. arabica* genome.