**Structure analyses andexpression pattern of the ERF transcription factor family in *Coffea arabica***

**Abstract**

Members of the ERF Family of Transcription Factors play an important role in plant development and gene expression that regulates responses to biotic and abiotic stress. This work identified 36 ERF family genes in *Coffea arabica* within the AP2/ERF full domain, using the EST-based genomic resource of the Brazilian Coffee Genome Project. The ERF family genes were classified into nine of the ten existing groups through phylogenetic analysis of the deduced amino acid sequences and comparison with the sequences of the ERF family genes in Arabidopsis*.* In addition to the AP2 domain, other conserved domains were identified, typical of members of each group. The *in silico* analysis and expression profiling showed high levels of expression for libraries derived from tissues of fruits, leaves and flowers as well as for libraries subjected to water stress. These results suggest the participation of the ERF family genes of *C. arabica* in distinct biological functions, such as control of development, maturation and responses to water stress. The results of this work imply in the selection of promising genes for further functional characterizations that will provide a better understanding of the complex regulatory networks related to plant development and responses to stress, opening up opportunities for coffee breeding programs.

**Keywords:** AP2/ERF, coffee, ethylene, transcription factor.

**Introduction**

With an annual worldwide production of 143.4 million sacks of grains in 2015/2016 (ICO 2016), coffee is an important agricultural commodity cultivated in more than 80 countries, which represents a significant source of income mainly for developing tropical countries (Pay, 2009). Brazil is the largest world producer and, together with Vietnam and Colombia, accounts for more than 50% of the world production (ICO 2016). Among the 124 identified species, only two are economically important: *Coffea arabica* and *Coffea canephora* (Hamon *et al*., 2015). The world market shares for these two species are 70% and 30%, respectively.

During their life cycle, crops are exposed to various biotic and abiotic stresses that limit their growth, development and production (Wang et al., 2003; Yu et al, 2016). To survive in stress conditions, plants have developed a complex molecular signaling network (Fujita et al., 2006; Lakhwani et al., 2016). Gene regulation by transcription is one of the main control points of biological processes in which Transcription Factors (TFs) play a central role (Sharma et al., 2010; Yamasaki et al., 2013).

The transcription factors belonging to the Ethylene Responsive Factor (ERF) family are part of the AP2/ERF superfamily, which also contains the AP2 and RAV families and consists of about 60-70 amino acids involved in DNA binding (Riechmann et al., 2000). The ERF family proteins contain a single AP2 domain and the AP2 family proteins contain two repeated AP2 domains (Jofuku et al., 1994). In addition to the single AP2 domain, the RAV family proteins contain a B3 domain that is a DNA-binding domain (Kagaya et al., 1999; Nakano et al., 2006). The ERF family is further divided into two subfamilies: the ERF and the CBF/DREB (Jofuku et al, 1994; Sakuma et al., 2002; Nakano et al., 2006).

Generally, the ERF family genes are partially involved in responses to biotic stress by recognizing the cis-acting sequence AGCCGCC, known as GCC-*box* (Hao et al., 1998; Licausi et al., 2010). The CBF/DREB subfamily genes play a crucial role on the plant responses to abiotic stress by recognizing the dehydration responsive element (DRE) with a central motif A/GTCGAC (Shinozaki, 1994; Thomashow, 1999; Shinozaki, 2006). The roles of the ERF and CBF/DREB proteins on the development and response to biotic and abiotic stresses in different plant species have been widely studied. Combining molecular genetic approaches, a series of ERF family regulatory genes involved in different metabolic pathways have been examined, including those related to drought (Dubouzet et al., 2003; Yamaguchi and Shinozaki, 2006; Mawlong et al., 2014; Yu et al., 2016), salinity (Dubouzet et al., 2003, Fu et al., 2007), low temperatures (Yang et al, 2005; Qin et al, 2007; Du et al, 2016), and diseases (Gutterson and Reuber, 2004; Charfeddine et al., 2015). In addition to responses to diverse stress types, the ERF family genes are also involved in the development of roots (Banno et al., 2001), germination (Pirrello et al., 2006; Yoong et al., 2016), and development and maturation of fruits (El-Sharkawy et al., 2009; Zhang et al., 2016).

Transcription Factors of the ERF subfamily and the CBF/DREB subfamily were identified in diverse plant species, including arabidopsis (Liu et al., 1998; Sakuma et al., 2002; Nakano et al., 2006), rice (Nakano et al., 2006; Sharoni et al., 2011), cotton (Jin and Liu, 2008; Champion et al., 2009), soybean (Zhang et al., 2008), poplar (Wang et al., 2016), grape (Zhuang et al., 2009; Licausi et al. 2010), corn (Zhuang et al., 2010), tomato (Sharma et al., 2010), apple (Zhuang et al., 2011), citrus (Ito et al., 2014) and banana (Lakhwani et al., 2016). Few studies on the characterization of the ERF family members in *Coffea* ssp were published (Lima et al., 2011). Bustamante-Porras et al. (2005) isolated the first ERF family member in *C. canephora*, whose expression is involved in processes of cell differentiation and fruit maturation. In *C. arabica*, no member of the ERF family was described until this moment.

Research on genomics and transcriptomics in coffee has gained more prominence. The Brazilian Coffee Genome Project (Vieira et al., 2006; Mondego et al., 2011) has been developed to investigate the coffee characteristics through complementary DNA sequencing (cDNA). This database has a set of 265,889 expressed sequence tags (ESTs) from different tissues, different developmental stages and different environmental conditions for *C. arabica, C. canephora* and *C. racemosa*. Therefore, the aim of this work was to identify and characterize possible ERF transcription factors in *C. arabica* from the ESTs database of the Brazilian Coffee Genome Project.

**Material and Methods**

*Identification of ERF family genes in Coffea arabica*

In order to isolate the ERF family genes in *C. arabica,* searches on the Brazilian Coffee Genome Database (<http://bioinfo03.ibi.unicamp.br/cafe/>) were performed using the AP2 domain of the *Solanum lycopersicon* ERF4 protein (GENBANK: AAO34706), with the BlastPsoftware (Altschul et al., 1997). More than 265,889 ESTs sequences are available in this database, which were obtained from forty three cDNA libraries, most of them of *C. arabica.* The cDNA was obtained from different tissues of the coffee plant (leaves, roots, flowers, seeds, fruits, among others) in different stages of development and subjected to various stress conditions (Vieira et al., 2006; Mondego et al., 2011). In order to increase the chances to identify new ESTs, searches were also performed using the following keywords: ERF, Ethylene Response Factor and EREBP. The sequences were compared with other sequences deposited in GenBank database using the BlastX and the BlastP programs (National Center for Biotechnology Information - NCBI - <http://www.ncbi.nlm.nih.gov>), in order to confirm their identity. The deduced sequence of amino acids of each *contig* was obtained by the ORF Finder *software* (Open Reading Frame Finder - NCBI -<https://www.ncbi.nlm.nih.gov/orffinder/>). The first collected sequences that presented incomplete AP2 domain or incorrect ORFs were excluded from analysis. The protein sequences of the AP2 domain were aligned by the Clustal Omega algorithm version 2.0.3 ([Sievers](http://europepmc.org/search;jsessionid=njdt1yzhUWoO8DFkfzyu.10?page=1&query=AUTH:%22Sievers+F%22)et al., 2011).

*Phylogenetic analysis*

The AP2 domain was aligned by the Clustal Omega algorithm version 2.0.3 ([Sievers](http://europepmc.org/search;jsessionid=njdt1yzhUWoO8DFkfzyu.10?page=1&query=AUTH:%22Sievers+F%22)et al, 2011). The phylogenetic tree was constructed using the neighbor-joining (NJ) method with *pair-wise deletion* by the MEGA *software* version 6.06 (Tamura et al., 2013). To test the reliability of the analysis 1,000 replicates of *bootstrap* were used.

*Determination of conserved motifs*

The identification of conserved motifs outside of the highly conserved AP2 domain was performed in protein sequences of *C. arabica* using the Clustal Omega algorithm version 2.0.3 ([Sievers](http://europepmc.org/search;jsessionid=njdt1yzhUWoO8DFkfzyu.10?page=1&query=AUTH:%22Sievers+F%22)et al., 2011) and Meme Suite version 4.11.2 (<http://meme-suite.org/tools/meme>), with the following parameters: ideal size: 6-80 amino acids; any number of repetitions for motifs and maximum number of motifs = 25. The resulting motifs were verified in the databases of NCBI (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and PROSITE (http: //www.expasy.org) to verify their significance.

*Gene Expression profiling by electronic Northern Blot*

The analysis of *in silico* qualitative gene expression profiling was performed by the Northern Blot technique. The frequency of *reads* in each library was normalized according to Lima et al. (2011). The e-Northern Blot was performed using Genesis *software*, version 1.7.5.

**Results**

*Identification of the ERF Family Genes and Phylogenetic Relationships between the ERF Family in C. arabica and Arabidopsis*

The *in silico* analysis comprised a data mining process within the Coffee Genome Project Database to identify the ERF family members in *C. arabica.* For this, the *reads* that possibly encode the AP2 domain in *C. arabica* were selected. The search by keywords and through the AP2 domain of ESTs enabled the identification of 38 Transcription Factors encoding ERF proteins. Among these 38 possible ERFs, it was observed that only 36 ERF proteins contained a full AP2 domain, while the other 2 ERFs contained only a part of the AP2 domain and, therefore, were excluded from analysis. The identity of the ERF proteins in *C. arabica,* with their homologs in Arabidopsis, varied from 63-90% (Table 1). The conservation of the sequence in comparison with arabidopsis was higher in the Group X members, varying from 75-90%. Lower values were observed in the Groups III (65%) and VI (63%).

A phylogenetic reconstruction was obtained from the identification of 122 ERFs in Arabidopsis, previously described by Sakuma et al., (2002). The phylogenetic tree was constructed by the *neighbor-joining* technique using the 58-59 sequences of amino acids of the AP2 domain in *C. arabica* and Arabidopsis. The ERF sequences are highly conserved between species. It favored the distinction of the 10 main groups, named as Groups I-X by Nakano et al., (2006). According to Figure 1, the comparative analysis of the phylogenetic tree for Arabidopsis and *C. Arabica* grouped a high number of identified sequences in *C. arabica* (22.22%, 8 sequences of 36) together with the arabidopsis sequences belonging to the Groups VII and IX. In *C. arabica,* a lower amount of proteins was grouped into the ERFs of the Arabidopsis belonging to the Groups I and VIII (13.89%, 5 sequences), followed by the Groups III and IV (8.33%, 3 sequences). The ERF groups with fewer members in *C. arabica* were X, II and VI (5.56%, 2.78% and 2.78%; related to the sequences 2, 1 and 1, respectively). No member of *C. arabica* was found in the Group V and in the Sub-Groups VI-L and Xb-L. These two sub-groups are characterized in Arabidopsis by a low homology in the C-terminal region of the AP2/ERF domain (Champion et al., 2009).

The ERFs were divided into two subfamilies based on similarity of the amino acids sequences. It was identified 12 putative ERFs as members of the DREB subfamily (Groups I, II, III and IV), in comparison with 57, 40, 75 and 58 in *Arabidopsis*, grape, poplar and rice, respectively. It was identified 24 encoding genes within the ERF subfamily (Groups VI, VII, VIII, IX, X) in comparison with 65, 82, 134 and 87 in Arabidopsis, grape, poplar and rice, respectively. The organization of the ERF family genes in *C. arabica* is showed in Table 2 along with the comparative distribution of *Arabidopsis*, grape, poplar and rice.

In order to study the phylogenetic relationship between the ERF family genes of *C. arabica* and *Arabidopsis*, multiple analyses were realized with the alignment of the deduced sequences of amino acids of the AP2 domain. The alignment indicated that the residues Gly-4, Arg-6, Glu-16, Trp-28 and Gly-30 are completely conserved among all proteins within the ERF family in *C. arabica* and Arabidopsis. In addition, more than 95% of the ERF family members contain the conserved residues Arg-8, Gly-11, Ile-17, Arg-18, Arg-26, Leu-29, Ala-38, Ala-39, Asp-43 and Asn-56. As previously demonstrated by Sakuma et al. (2002), the ERF gene subfamily include two main residues of amino acids, the alanine (A) at position 14 and the aspartate (D) at position 19, which possibly contribute to a functional activity of binding to GCC-*box* in many ERFs. The CBF/DREB family contains a valine (V) and a glutamic acid (E) at positions 14 and 19, respectively. In the DREB subfamily, all genes present the conserved residue Valine14 and, at position 19, the genes *CaERF01-05* contain 1 Leucine (L) and the genes *CaERF06-12* contain 1 Glutamic Acid (E). The C-terminal region of the AP2/ERF domain of the *CaERF34* protein showed low homology with the region of consensus with other genes (Figure 2). This region corresponds to the half of terminal α-*helix* (Allen et al, 1998), which includes the highly conserved residues Asp-43 and Asn-57. In general, the ERF family showed significant similarity to the remaining domain.

*Distribution of conserved motifs*

In general, the regions in Transcription Factors outside the DNA-binding domain contain important domains that are involved in transcription activities as the protein-protein interactions, which may be involved in nuclear localization (Liu et al., 1999). Such functional domains are often conserved among members of a subgroup within large families of transcription factors in plants. Probably these motifs are sharing the same functions (Kranz et al., 1998; Reyes et al., 2004; Nakano et al., 2006).

In order to relate the putative ERFs in *C. arabica* to biological functions, other conserved motifs (CM) outside the AP2/ERF region were investigated on the deduced sequences of amino acids. Most members of the same group shared one or more motifs outside the AP2 domain (Figure 3). For instance, the Group I comprises 5 ERFs *(CaERF01-05)* and contain 5 conserved motifs (Figure 3). Except *CaERF05*, the members of this group contain the CMI-1 and CMI-2 motifs in the C-terminal region. The *CaERF06* gene belonging to the Group II, unique member identified in this work, contains the CMII-1motif in the C-terminal region, adjacent to the AP2 domain. Belonging to the Group III are the *CaERF07-09* genes*,* which contain the CMIII-1 motif in the C-terminal region. In addition to the CMIII-1 motif, the *CaERF08* gene contains CMIII-2 and CMIII-4, the last identified as the LWSY conserved motif in the *OsDREB1A/B/C* and *AtCBF3/DREB1A* (Dubouzet et al., 2003). The *CaERF09* represents the CMIII-6 and the CMIII-7 motifs. In the Group IV *(CaERF10-12)* only the CMIV-2 motif was found in the *CaERF10* and *CaERF11* genes. The CMIV-2 motif includes a putative nuclear localization signal (Liu et al., 1998).

The *CaERF13* gene, unique of the Group VI, has two proteins that share the CMVI-1 and CMVI-2 conserved motif on the N-terminal region. The Group VII was firstly described by Tournier et al., (2003) and is characterized by the presence of one highly conserved motif in the N-terminal region (MCGGAII/L). Within this group, 8 members were found in *C. arabica.* The EAR motif (CMVIII-1) (ERF associated with amphiphilic repression) was found in members of the Group VIII, in the ERFs *CaERF22* and *CaERF23*, which also contain the CMVIII-2 motif. The Group IX is composed by 8 genes (*CaERF27-34).* The *CaERF27-30* genes contain only the CMIX-3 motif while the *CaERF32* gene contains only the CMIX-2 motif. The CMIX-2 and CMIX-3 motifs are putative acidic regions that may function as a transcription activation domain (Fuijimoto et al, 2000). The CMIX-3 motif corresponds to a conserved sequence that was referred previously to a DMLV motif (Gutterson and Reuber, 2004). In addition to the CMIX-3, the *CaERF32* gene contains the CMIX-5 and CMIX-6 motifs that are probably a MAP kinase phosphorylation site, located at the C-terminal region of the protein (Fujimoto et al., 2000). The Group X is represented by the *CaERF35* and *CaERF36* genes. The members of this group contain one CMX-1 conserved motif in the N-terminal region, such as the *CaERF35*, whereas the *CaERF36* presents no conserved motif.

*In silico gene expression profiling of the transcription factors of the ERF family in Coffea arabica*

In order to assess the differences among transcripts of different tissues or organs, the ERF expression profiling was obtained *in silico* by e-Northen in the *C. arabica* libraries. High levels of expression were observed in libraries derived from tissues of fruits, leaves and flowers (Figure 4). The ERFs of the cDNA libraries from the tissues subjected to different types of stresses were also detected, however, with fewer transcripts than those from the tissues of diverse parts of the plant and different stages of development. Transcripts were detected in the majority of the evaluated libraries for the ERFs *CaERF20*, *CaERF21* and *CaERF2*3. However, the majority of the 36 transcription factors of the ERF family were detected in specific libraries, showing that they are tissue-specific. This is the case of *CaERF03, CaERF08, CaERF12, CaERF18, CaERF22, CaERF26, CaERF31* and *CaERF36*, which are expressed only in fruits. On the other hand, the *CaERF03, CaERF08, CaERF12, CaERF18, CaERF22, CaERF26, CaERF31* and *CaERF36* genes are expressed only in flowers, leaves and roots. Expression profiling in libraries subjected to water stress were observed for *CaERF06, CaERF07, CaERF11, CaERF21* and *CaERF23*; the first three ERFs showed a higher expression in this library.

**Discussion**

Transcription factors are the principal regulators of biological factors and emerged as a powerful tool to manipulate complex metabolic pathways (Grotewolda, 2008). Using these proteins on plant breeding programs involves knowledge on their role in gene regulatory networks. The ERF family of transcription factors presents a highly conserved element including the AP2/ERF domain, responsible by the DNA binding activity and important to plant development (Ohta et al., 2001; Sakuma et al., 2002, Cao et al., 2006). Nakano et al., (2006) have systematically studied the phylogeny, the structures and the conserved motifs of the ERF family in Arabidopsis and rice. In order to obtain more information on this family in *C. arabica,* this present work identified and analyzed 36 possible ERF proteins from the EST database in *C. arabica,* which is available at the Brazilian Coffee Genome Project (Vieira et al., 2006; Mondego et al., 2011).

The AP2 conserved domains of Arabidopsis were used to group their homologs in *C. arabica.* The majority of the sequences in *C. arabica* are grouped into Groups VII and IX followed by the Groups I and VIII. Only few sequences were grouped into Groups III, VI, X, II and VI. However, the Group VII represents about 4.1% of the family in Arabidopsis (Nakano et al, 2006), 2.46% in grape (Licausi et al., 2010), 9.56% in Poplar (Wang et al., 2016) and 10.34% in rice (Nakano et al., 2006). This group represents more than 22% of the proteins containing the single AP2 domain, found in the Coffee Genome database. In this work, from the ten main groups identified in the ERF family in Arabidopsis, nine occurred in *C. arabica.* Therefore, the methodology used by Nakano et al, (2006) is applicable in this species. The presence of the majority of groups and subgroups in the two dicot species, as well as in monocot species suggests that many of the genes predate the divergence between monocotyledonous and dicotyledonous (Sharma et al., 2010). On the same way, some groups and subgroups are present in only one specie, for instance, the Groups XI-XIV occur only in the ERF family in rice, excluding Arabidopsis and other dicot species. It suggests that these groups have evolved or were lost in a certain species after divergence (Zhang et al., 2008).

Despite the great number of ESTs already analyzed, the number of ERFs found in *C. arabica* was lower than in Arabidopsis and rice, which contain 122 and 129 ERF family members, respectively (Nakano et al., 2006). Zhang et al, (2008) showed that although soybean has a large genome (1.115 Mb), the structure and phylogeny of the AP2/ERF superfamily was similar to Arabidopsis and rice. The same was observed in complete genomes as poplar (Wang et al, 2016) and grape (Zhuang et al., 2009). *C. arabica* is an allotetraploid species (2n=4x=44) formed by a recent (about 1 million years) natural hybridization between the diploid species: *C.* *eugenioides* and *C. canephora* (Lashermes et al., 1999; Clarindo and Carvalho, 2009). Because the sequencing and annotation of the *C. arabica* genome is still incomplete and the sequences containing only a part of the AP2 domain are excluded, it is difficult to predict the size of the ERF gene family in this species. However, it tends to be similar in size to other plant organisms.

The structural analysis revealed that all EFR proteins contain conserved Ala-14 and Asp-19, whereas the DREB proteins contain Val-14 and Glu/Leu-19. The comparative analysis of the amino acids residues of the ERF/AP2 domain in *C. arabica* with the ERF family proteins in *Arabidopsis* showed that the AP2/ERF domains were well conserved between the two species. These conserved amino acids probably play an important role in the ERF gene family, where they can be involved with different ways of contact with DNA. According to Allen et al (1998), the AP2/ERF domain recognizes the DNA by the conserved residues arginine and tryptophan, located into *β-sheets.* The Ala-37 in the ERF domain plays an important role in the stability of the ERF domain or DNA binding to the DRE element or GCC-*box* (Liu et al. 2006).

The transcription factors generally contain conserved domains that are outside the DNA binding domain, but functionally important (De Bodt et al., 2003; Arora et al., 2007). The distribution of the specific motifs into proteins belonging to the specific groups of the phylogenic tree was also observed for the ERF proteins in *C. arabica,* which demonstrated structural similarities within the same subgroup. The majority of the ERF sequences identified in *C. arabica* share one or more motifs outside the AP2 domain with their homologs in Arabidopsis, such as in rice and soybean (Nakano et al., 2006; Zhang et al., 2008). For instance, Ohta et al., (2001) identified an EAR motif (ERF related to amphiphilic repression), which works as a repression domain. The EAR motif of conserved sequence, (L/F)DLN(L/F)xP, identified in this present work as CMVIII-1, is found in the C-terminal regions of the Group VIII. This motif was already identified in various repressors, including ZAT7, 10, 12, ERF3, AUX/IAA, NIMIN1, HSI2, SUPERMAN (Arabidopsis), NRR (rice), ZFT1 (tobacco) and ZPT2-3 (petunia), which play different roles - from the plant development to stress tolerance (Tiwari et al., 2004; Hiratsu et al., 2004; Kazan, 2006). Currently the DEAR1, a DREB protein containing the EAR motif, appeared as a protein repressor of binding to dehydration responsive element, which mediates responses to biotic and abiotic stresses (Tsutsui et al., 2009). The CMIV-2 motif in the N-terminal region could work as a nuclear localization signal (NLS) (El Kayal et al., 2006). Recently, it has been considered essential in Arabidopsis that CBF1 bind to DNA, since it is indispensable for transcriptional activity (Canella et al., 2009). A putative nuclear localization signal (KRKRK) was identified in ERF proteins (Van Raemdonck et al., 2005, El-Sharkawy et al., 2009). The comparative analysis of the conserved motifs in *C. arabica* and Arabidopsis suggests that the protein functions were conserved and diverged during the evolution of the ERF gene family. Sharma et al., (2010) showed that some motifs are specific in spermatophytes whereas many motifs have been identified in non-vascular plants, bacteria, fungi and animals. The presence of these conserved motifs in evolutionarily different organisms indicates that they play an important functional role, while specific motifs in spermatophytes may have later evolved to fulfill specific functions.

In this present work, 8 ERFs belonging to the Groups I, III, VI, VII, VIII, IX and X were expressed only in fruit libraries. Although the ERF transcription factors are regulated by a series of physical and chemical stimuli, many ERFs are responsive to ethylene, and therefore they may be involved in the maturation process of climacteric fruits. Tournier et al., (2003) demonstrated that the *SlERF2,* an ERF that binds to the GCC-*box,* plays an important role during the tomato maturation process. The same was observed by Yin et al., (2010) for different ERFs expressed during the kiwi maturation process. Pereira et al., (2005) showed that the autocatalytic production of ethylene in fruits of green *C. arabica* is very low; however, it increases considerably during the initial stage of ripening. Such observations demonstrate the climacteric nature of the maturation of *C. arabica* fruits and the importance of ethylene in this process. Bustamante-Porras et al., (2005) isolated an ERF gene (*CoERF*) in *C. canephora,* with expression during fruit differentiation and maturation. Comparing this *CoERF* (GENBANK: AY522505) with the *CaERF17* in *C. arabica,* it shared 97% of identity and 98% of similarity. Given that *C. canephora* is one of the ancestors of *C. arabica*, the CaERF17 was expected to be expressed in fruits. However, *reads* were not found in fruits libraries for CaERF17. In allotetraploids, genes are expected to be present in two homologous forms, highly similar, but not identical (Petitot et al., 2008). The redundancy of genes can lead to gene silencing or functional divergence of duplicated genes (Chen and Ni 2006). Vidal et al., (2010) found that, in some cases, apparently a homolog of *C. canephora* is recruited to be expressed in certain tissues, while *C. eugenioides* homologs are silenced. On this way, differences in expressions in *C. arabica* can be attributed to different sub-genomes of the ancestors of *C. canephora* or *C. eugenioides*. These genes may be good candidates for future characterizations that would help to understand regulation processes during development and maturation of fruits in *Coffea* ssp*.*

The majority of the ERFs have demonstrated an increase in plant tolerance to biotic and abiotic stresses (Park et al., 2001; Guo et al., 2004; Zhang et al., 2009). In this work, the ERFs *CaERF06, CaERF07, CaERF11* presented high expression in libraries subjected to water stress. These genes belong to the DREB subfamily, which play an important role in plant tolerance to abiotic stress, recognizing the Dehydration Responsive Element (DRE), the *core motif* A/GCCGAC (Liu et al, 1998). Studies have showed that the overexpression of DREB genes in Arabidopsis activate the expression of several genes related to stress, thus improving the tolerance to drought, salinity and low temperature (Liu et al. 1998; Sakuma et al., 2006; Bouaziz et al., 2015). For example, the overexpression of the *AoDREB* gene of *Asparagus officinalis* L in transgenic Arabidopsis induced the expression of genes rd29A and COR15A, resulting in higher tolerance of transgenic plants to drought and salinity (Liu et al., 2010). On this way, these genes are promising for further studies that will help to understand the regulation mechanisms of the ERF family related to responses of *C. arabica* to different stresses.

Previous works suggest the hypothesis that a group-specific expression profile are occurring. For example, from 8 genes belonging to the Group VII, 5 are expressed in fruits, where the ERFs *CaERF14*, *CaERF15* and *CaERF18* present a high relative expression. This group has been associated to fruit maturation. *LeERF2* in tomato (Tournier et al., 2003), *MdERF1* in apple (Wang et al., 2007), *PsERF2a* and *PsERF2b* in plum (El-Sharkawy et al, 2009) and *AdERF10* and *AdERF14* in kiwi (Yin et al, 2010) are proteins that are expressed during the maturation of fruits belonging to Group VII. The ripening induction was also related in the Group VIII in plum (El-Sharkawy et al., 2007) and grape (Licausi et al., 2010). Other works have demonstrated that members of the Group IX present induction of expression when subjected to diverse pathogen attacks. Constitutive overexpression of the AtERF2 of the IXa subgroup probably induced the gene expression of the PDF1.2 gene (Brown et al, 2003). On the same way, the overexpression of the *AtERF1* gene in Arabidopsis, a homolog next to *AtERF2*, gives resistance to *Botrytis cinerea, Sclerotinia sclerotiorum* and *Erysiphe orontii* in Arabidopsis (Gutterson and Reuber, 2004). Anderson et al., (2010) demonstrated that the overexpression of the *MtERF1-*1 gene in roots of *Medicago truncatula* increased the resistance to *Rhizoctonia solani,* as well as to *Phytophthora medicaginis*. Thus, the genetic profile expression suggests a functional level of specialization for the investigated ERF Groups, although it is expected a high degree of overlapping functions in large plant genes families (Soltis et al., 2008; Semon and Wolfe, 2007). On this way, the presence of distinct expression profiles of the ERFs observed in *C. arabica* by *in silico* analysis may be associated to the phylogenetic distance among sequences, that is, the ERF phylogenetically related proteins have more similar patterns of expression than the divergent sequences.

The ERF gene family plays a crucial role in the development regulation, as well as in the responses to environmental stresses. With the sequenced transcriptome of *C. arabica* by a Brazilian consortium (Brazilian Coffee Genome Project), 36 ERFs were identified in *C. arabica* in this work, where 12 ERFs belong to the DREB subfamily and 24 to the ERF subfamily. The gene expression profiling showed high levels of expression in libraries derived from tissues of fruits, leaves and flowers as well as in libraries subjected to water stress. From the comparison of the homologs with other species, whose genome was sequenced together with expression profiles, it is suggested that the ERFs of *C. arabica* is involved in different biological functions that are mediated by ethylene, such as control of development, maturation and responses to water stress. *C. arabica* is a perennial species whose fruits have commercial value. Knowledge on the role of the ERF transcription factors in processes of development and maturation of this species opens opportunities for investments in plant breeding programs to increase the production and the coffee bean quality.

**Acknowledgements**

We thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Brazil) and CNPq (Conselho Nacional de Desenvolvimento Cientifico- Brazil) for the PhD Fellowship to Silvia G. H. de Souza.

**References**

Allen, M. D.; Yamasaki, K.; Ohme-Takagi, M.; Tateno, M.; Suzuki, M. A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. EMBO J. 1998;(17): 5484-5496.

Altschul, S. F., Madden, T. L., Scha¨Ffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research.1997;(25):3389–3402.

Anderson, J. P.; Lichtenzveig, J.; Gleason, C.; Oliver, R. P.; Singh, K. B. The B-3 Ethylene Response Factor MtERF1-1 Mediates Resistance to a Subset of Root Pathogens in Medicago truncatula without Adversely Affecting Symbiosis with Rhizobia. Plant Physiology. 2010; (154):861-873

Arora, R.; Agarwal, P.; Ray, S.; Singh, A. K.; Singh, V. P.; Tyagi, A. K.; Kapoor, S. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. BMC Genomics. 2007; (8):242.

Banno, H.; Ikeda, Y.; Niu, Q. W.; Chua, N. H. Overexpression of arabidopsis ESR1 induces initiation of shoot regeneration. The Plant Cell. 2001; (13):2609-2618.

Brown, R. L.; Kazan, K.; Mcgrath, K. C.; Maclean, D. J.; Manners, J. M. A role for the GCC-box in jasmonate-mediated activation of the *PDF1.2* gene of *Arabidopsis*. Plant Physiology.2003.(132):1020-1032.

Bouaziz, D.; Jbir, R.; Charfeddine, S.; Saidi, M.N.; Gargouri-Bouzid, R. The StDREB1 transcription factor is involved in oxidative stress response and enhances tolerance to salt stress. Plant Cell Tissue Organ Cult. 2015121):237–248.

Bustamante-Porras, J.; Noirot, M.; Campa, C.; Hamon, S.; Kochko. A. Isolation and characterization of a *Coffea canephora* ERF-like cDNA. African Journal of Biotechnology. 2005. 2 (4):157-159, 2005.

Canella, D.; Gilmour, S. J.; Kuhn, L. A.; Thomashow, M. F. DNA binding by the Arabidopsis CBF1 transcription factor requires the PKKP/RAGRxKFxETRHP signature sequence. Biochim Biophys Acta. 2010 (1799):454-462.

Cao, Y.; Song, F.; Goodman, R. M.; Zheng, Z. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. Journal of Plant Physiology. 2006;11.(163): 1167-1178

Champion, A.; Hebrard, E.; Parra, B.; Bournaud, C.; Marmey, P.; Tranchant, C.; Nicole, M. Molecular diversity and gene expression of cotton ERF transcription factors reveal that group IXa members are responsive to jasmonate, ethylene and *Xanthomonas*. Molecular Plant Pathology; 2009(10): 471-485.

Charfeddine, M.; Bouaziz, D.; Charfeddine, S.; Hammami, A.; Nouri-Ellouz, O.; Bouzid G.R. Overexpression of dehydration responsive element binding 1 protein (DREB1) in transgenic *Solanum tuberosum* enhances tolerance to biotic stress. Plant Biotechnol Rep.; 2015(9):79-88.

Chen, Z. J.; Ni, Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. BioEssays; 2006(28): 240-252.

Clarindo, W. R.; Carvalho, C. R. Comparison of the Coffea canephora and C. arabica karyotype based on chromosomal DNA content. Plant Cell Rep. 2009;1(28), :73-81.

De Bodt, S.; Raes, J.; Florquin, K.; Rombauts, S.; Rouze, P.; Theissen, G.; Van De Peer, Y. Genome-wide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. J. Mol. Evol., 2003 (56):573-586. 2003.

Du, C.; Hu, K.; Xian, S.;Liu, C.; Fan, J.; Tu, J. Dynamic transcriptome analysis reveals AP2/ERF transcription factors responsible for cold stress in rapeseed (*Brassica napus* L.). 2016(291):1053–1067.

Dubouzet, J. G.; Sakuma, Y.; Ito, Y.; Kasuga, M.; Dubouzet, E. G.; Miura, S.; Seki, M.; Shinozaki, K.; Yamaguchi, S. K. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. Plant J., 2003(33): 751-763.

EL Kayal, W.; Navarro, M.; Marque, G.; Keller, G.; Marque, C. Teulieres, C. Expression profile of CBF-like transcriptional factor genes from *Eucalyptus* in response to cold. J. Exp. Bot.; 2006(57):2455-2469.

El-Sharkawy, I.; Sherif, S.; Mila, I.; Bouzayen, M.; Jayasankar, S. Molecular characterization of seven genes encoding ethylene-responsive transcriptional factors during plum fruit development and ripening. J. Exp. Bot., 2009 (60):907-922.

El-Sharkawy, I.; Kim, W. S.; El-Kereamy, A.; Jayasankar, S.; Svircev Amd.; Brown, C. W. Isolation and characterization of four ethylene signal transduction elements in plums (*Prunus salicina* L.). J. Exp. Bot., 2007(58): 3631-3643.

Fu, X. Y.; Zhang, Z.; Peng, R. H.; Xiong, A. S.; Liu, J. G.; Wu, L. J.; Gao, F.; Zhu, H.; Guo, Z. K.; Yao, Q. H. Isolation and characterization of a novel cDNA encoding ERF/AP2-type transcription factor OsAP25 from *Oryza sativa* L. Biotechnol Lett., 2007 (29):1293-1299.

Fujimoto S. Y.; Ohta, M.; Usui, A.; Shinshi, H.; Ohme-Takagi, M. *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell., 2000 (12):393-404.

Fujita, M.; Fujita, Y.; Noutoshi, Y.; Takahashi, F.; Narusaka, Y.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. Current Opinion in Plant Biology., 2006(9):436-442.

Grotewolda, E. Transcription factors for predictive plant metabolic engineering: are we there yet? Curr Opin Biotechnol., 2008(19):138-144.

Guo, Z. J.; Chen, X. J.; Wu, X. L.; Ling, J. Q.; Xu, P. Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. Plant Mol. Biol., 2004 (55):607-618.

Gutterson, N.; Reuber, T. L. Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr. Opin. Plant Biol.; 2004 (7):465–471.

Hamon, P.; Hamon, S.; [1](http://www.sciencedirect.com/science/article/pii/B9780124095175000048#aff1) Razafinarivo, N.J.; Guyot, R.; [Sonja Siljak-Yakovlev](http://www.sciencedirect.com/science/article/pii/B9780124095175000048), S.; Couturon, E.; [Crouzillat](http://www.sciencedirect.com/science/article/pii/B9780124095175000048), D.; Rigoreau, M.; Akaffou, S.; Rakotomalala, J.J.; Kochko A. Chapter 4 - Coffea Genome Organization and Evolution, In: Coffee in Health and Disease Prevention, edited by Victor R. Preedy, Academic Press, San Diego, 2015, Pages 29-37.

Hao, D.; Ohme-Takagi, M.; Sarai, A. Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. J. Biol. Chem.;1998 (273): 26857-26861.

Hiratsu, K.; Mitsuda, N.; Matsui, K.; Ohme-Takagi, M. Identification of the minimal repression domain of SUPERMAN shows that the DLELRL hexapeptide is both necessary and sufficient for repression of transcription in *Arabidopsis*. Biochem Biophys Res Commun.; 2004 (321): 172-178.

ICO. International Coffee Organization.

Available<http://www.ico.org/monthly\_coffee\_trade\_stats.asp.> Accessed 08 jun 2016.

Ito, T.M.; Polido, P.B.; Rampim, M.C.; Kaschuk, G.; Souza, S.G.H. Genome-wide identification and phylogenetic analysis of the AP2/ERF gene superfamily in sweet orange (*Citrus sinensis*). Genet Mol Res.; 2014(13): 7839–7851.

Jin, L. G.; Liu, J. Y. Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene GhERF4 from cotton. Plant Physiology and Biochemistry.; 2008 (46):46–53.

Jofuku, K. D.; Den Boer, B. G.; Van Montagu, M.; Okamuro, J. K. Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. Plant Cell; 1994 (6):1211–1225.

Kazan, K. Negative regulation of defense and stress genes by EAR-motif-containing repressors. Trends Plant Sci., 2006 (11):109-112.

Kagaya, Y.; Ohmiya, K.; Hattori, T. Rav1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. Nucleic Acids Res., 1999 (27): 470-478.

Kranz, H. D.; Denekamp, M.; Greco, R.; Jin, H.; Leyva, A.; Meissner, R. C, Petroni, K.; Urzainqui, A.; Bevan, M.; Martin, C. et al. Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. Plant J.,1998 (16): 263-276.

Lakhwani, D., Pandey, A.; Dhar, Y.V.; Bag, S.K.; Trivedi, P.K.; Asif, M.H. Genome-wide analysis of the AP2/ERF family in Musa species reveals divergence and neofunctionalisation during evolution. Scientific Reports. 2016 (6): 18878.

Lashermes, P.; Combes, M. C.; Robert, J.; Trouslot, P.; D'hont, A.; Anthony, F.; Charrier, A. Molecular characterization and origin of the *Coffea arabica* L. genome. Mol. Gen. Genet., 1999; 261(2): 259-266.

Licausi, F.; Giorgi, F. M.; Zenon, S.; Osti, F.; Pezzotti, M.; Perata, P. Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera.* BMC Genomics. 2010 (11): p. 719.

Lima, A.A.; Ságio, S.A.; Chalfun-Júnior, A.; Paiva, L.V. In silico characterization of putative members of the coffee (*Coffea arabica*) ethylene signaling pathway. Genet Mol Res.; 2011 (10):1277-1289.

Liu, Y.; Zhao, T. J.; Liu, J. M.; Liu, W. Q.; Liu, Q.; Yan, Y. B.; Zhou, H. M. The conserved Ala37 in the ERF/AP2 domain is essential for binding with the DRE element and the GCC box. FEBS Lett.; 2006 (580): 1303-1308.

Liu, Y.; Chen, H.; Zhuang, D.; Jiang, D.; Liu, J.; Wu, G.; Yang, M.; Shen, S.Characterization of a DRE-binding transcription factor from asparagus (*Asparagus officinalis* L.) and its overexpression in *Arabidopsis* resulting in salt- and drought-resistant transgenic plants. Journal of Plant Sciences.; 2010 (171):12-23.

Liu, L, White, M. J.; Macrae, T. H. Transcription factors and their genes in higher plants functional domains, evolution and regulation. Eur. J. Biochem., 1999 (262): 247-257.

Liu Q.; Kasuga, M.; Sakuma, Y.; Abe H.; Miura, S.; Yamaguchi-Shinozaki, K.;Shinozaki, K. Two transcription factors, DREB1 and DREB2, withan EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis.* Plant Cell., 1998 (10): 1391-1406.

**Mawlong, I.; Ali, K.; Kurup, D.;Yadav, S.; Aruna, T. Isolation and characterization of an AP2/ERF-type drought stress inducible transcription factor encoding gene from rice. Plant Biochem. Biotechnol, 2014 (23): 42-51.**

**Mondego, J. M. C.**; **Vidal, R. O.**; **Carazzolle,** **M. F.**; **Tokuda,** **E. K.**; **Parizzi, L. P.**; **Costa, G. G. L.**; **Pereira,** **L. F. P**.; **Andrade, A. C.**; **Colombo, C. A.**.; **Vieira, L. G. E.; Pereira,** **G. A. G. Brazilian Coffee Genome Project Consortium. An EST-based analysis identifies new genes and reveals distinctive gene expression features of *Coffea arabica* and *Coffea canephora*.** BMC Plant Biology.,2011(**11)p**30.

Nakano, T.; Suzuki, K.; Fujimura, T.; Shinshi, H. Genomewide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiology., 2006 (140):411-432.

Ohta, M.; Matsui, K.; Hiratsu, K.; Shinshi, H.; Ohme-Takagi, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. Plant Cell., 2001 (13): 1959–1968.

Park, J. M.; Park, C. J.; Lee, S. B.; Ham, B. K.; Shin, R.; Paek, K. H. Overexpression of the tobacco Tsi1 gene encoding an EREBP/ AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. Plant Cell., 2001 (13):1035–1046.

Pay, E. The market for organic and fair-trade coffee. FAO Rome 2009. Available in http://<[www.fao.org/.../organicexports/.../Market\_Organic\_FT\_Coffee.pdf](http://www.fao.org/.../organicexports/.../Market_Organic_FT_Coffee.pdf)>accessed 31 October 2016

Pereira, L. F. P.; Galvao, R. M.; Kobayashi, A. K.; Cação, S. M. B.; Esteves, V. L. G. Ethylene production and acc oxidase gene expression during fruit ripening of Coffea arabica L. Braz. J. Plant Physiol., 2005(17):283-289.

Petitot, A. S.; Lecouls, A. C.; Fernandez, D. Sub-genomic origin and regulation patterns of a duplicated WRKY gene in the allotetraploid species Coffea Arabica. Tree Genetics & Genomes. 2008; 3(4): :379-390.

Pirrello, J.; Jaimes-Miranda, F.; Sanchez-Ballesta, M. T.; Tournier, B.; Khalil-Ahmad, Q.; Regad, F.; Latche, A.; Pech, J. C.; Bouzayen, M. Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. Plant Cell Physiol., 2006; 9 (47):1195-1205.

Qin, Q. L.; Liu, J. G.; Zhang, Z.; Peng, R. H.; Xiong, A. S.; Yao, Q. H.; Chen, J. M. Isolation, optimization, and functional analysis of the cDNA encoding transcription factor RdreB1 in *Oryza Sativa* L. Mol. Breed., 2007 (19):329-340.

Reyes, J. C.; Muro-Pastor M. I.; Florencio, F. J. The GATA family of transcription factors in Arabidopsis and rice. Plant Physiol., 2004 (134):1718-1732.

Riechmann, J. L.; Heard, J.; Martin, G.; Reuber, L.; Jiang, C.; Keddie, J.; Adam, L.; Pineda, O.; Ratcliffe, O. J.; Samaha, R. R, Et Al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science. 2000 ( 290): 2105-2110.

Sakuma Y.; Liu Q.; Dubouzet J. G.; Abe H.; Shinozaki, K. DNA binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration and cold inducible gene expression. Biochem Biophys Res Commun., 2002 (290): 998-1009.

Sakuma, Y.; Maruyama, K.; Osakabe, Y.; Qin, F.; Seki, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. Plant Cell. 2006 (18):1292–1309.

Semon M.; Wolfe, K. H. Consequences of genome duplication. Curr. Opin. Gen. Dev.; 2007 (17): 505-512.

Sharma, M. K.; Kumar, R.; Solanke, A. U.; Sharma, R.; Tyagi, A. K.; Sharma A. K. Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. Mol. Genet. Genomics, 2010. 284: 455. doi:10.1007/s00438-0100580-1

Sharoni, A. M.; Nuruzzaman, M.; Satoh, K.; Shimizu, T.; Kondoh, H, Sasaya, T.; Choi, Ir.; Omura, T.; Kikuchi, S. Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice Plant Cell Physiol., 2011;2 (52): 344–360.

Shinozaki, K. Y.; Shinozaki, K. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. The Plant Cell., 1994 (6):251-264.

Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; Mcwilliam, H.; Remmert, M.; Söding, J.; Thompson, J.D.; Higgins, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 2011 (7): 539.

Soltis, D. E.; Bell, C. D.; Kim, S.; Soltis, P. S. Origin and early evolution of angiosperms. Ann. New York Acad. Sci., 2008 (1133): 3-25.

Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. Mega6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol., 2013 (30), p.2725-2729.

Thomashow, M. F. Plant Cold Acclimation: freezing tolerance genes and regulatory mechanism. Annual Review of Plant Physiology and Plant Molecular Biology, 1999 (50):571–599.

Tiwari, S. B.; Hagen, G.; Guilfoyle, T. J. Aux/Iaa Proteins Contain A Potent Transcriptional Repression Domain. Plant Cell, 2004 (16): 533-543.

Tournier, B.; Sanchez-Ballesta, M. T.; Jones, B.; Pesquet, E.; Regad, F.; Latche, A.; Pech, J. C.; Bouzayen, M. New Members Of The Tomato Erf Family Show Specific Expression Pattern And Diverse Dna-Binding Capacity To The Gcc Box Element. Febs Lett, 2003 (550):149-154.

Tsutsui, T.; Kato, W.; Asada, Y.; Sako, K.; Sato, T.; Sonoda, Y.; Kidokoro, S.; Yamaguchi-Shinozaki, K.; Tamaoki, M.; Arakawa, K.; Ichikawa, T.; Nakazawa, M.; Seki, M.; Shinozaki, K.; Matsui, M.; Ikeda, A.; Yamaguchi, J. Dear1, A Transcriptional Repressor Of Dreb Protein That Mediates Plant Defense And Freezing Stress Responses In Arabidopsis. J. Plant. Res., 2009 (122):633-643.

Van Raemdonck, D.; Pesquet, E.; Cloquet, S.; Beeckman, H.; Boerjan, W.; Goffner, D.; El, Jaziri, M.; Baucher, M. Molecular Changes Associated With The Setting Up Of Secondary Growth In Aspen. J. Exp. Bot., 2005 (56):2211–2227.

Vidal, R. O.; Mondego, J. M.; Pot, D.; Ambrosio, A. B.; Andrade, A. C.; Pereira, L. F.; Colombo, C. A.; Vieira, L. G.; Carazzolle, M. F.; Pereira, G. A. A High-Throughput Data Mining Of Snps In Coffea Spp Ests Suggests Differential Homeologous Gene Expression In The Allotetraploid Coffea Arabica. Plant Physiology, 2010; 3(154): 1053-1066.

Vieira, L. G. E.; Andrade, A. C.; Colombo, C. A.; Moraes, A. H. A.; Metha, A.; Oliveira A. C.; Labate, C. A.; Marino, C. L.; Monteiro-Vitorello, C. B,.; Monte, D. C. Et Al: Brazilian Coffee Genome Project: An Est-Based Genomic Resource. Brazil J Plant Physiol. 2006 (18): 95-108.

Wang, A.; Tan D. M.; Takahashi, A.; Li, T. Z.; Harada, T. Mderfs, Two Ethylene-Response Factors Involved In Apple Fruit Ripening. J. Exp. Bot., 2007 (58): 3743-3748.

Wang, S.;Yao, W.;Zhou, B.; Jiang, T.Structure Analysis And Expression Pattern Of The Erf Transcription Factor Family In Poplar. Acta Physiol Plant, 2016 (38) p.239. doi:10.1007/s11738-016-2242-3

Wang, W.; Vinocur, B.; Altman, A. Plant Responses To Drought, Salinity And Extreme Temperatures: Towards Genetic Engineering For Stress Tolerance. Planta; 2003 (218): 1-14.

Yamaguchi, S. K.; Shinozaki, K. Transcriptional Regulatory Networks In Cellular Responses And Tolerance To Dehydration And Cold Stresses. Annu. Rev. Plant Biol. 2006 (57): 781-803.

Yamasaki, K.; Kigawa, T.; Seki, M.; Shinozaki, K.; Yokoyama, S. Dna-Binding Domains Of Plant-Specific Transcription Factors: Structure, Function, And Evolution.Trends Plant Sci. 2013 (18): 267-76.

Yang, T. W.; Zhang, L. J.; Zhang, T. G.; Zhang, H.; Xu, S. J.; An, L. Z. Transcriptional Regulation Network Of Cold-Responsive Genes In Higher Plants. Plant Science, 2005 (169):987-995.

Yin, X-R.; Allan, A. C.; Chen, K-S.; Ferguson, I. B. Kiwifruit Eil And Erf Genes Involved In Regulating Fruit Ripening. Plant Physiology, 2010, 3(153):1280-1292.

Yoong, F.Y.; O'brien, L.K.; Truco, M.J.; Huo, H.; Sideman, R.; Hayes, R.; Michelmore, R.W.; Bradford, K.J. Genetic Variation For Thermotolerance In Lettuce Seed Germination Is Associated With Temperature-Sensitive Regulation Of Ethylene Response Factor1 (Erf1). Plant Physiol. 2016 (170):472-488.

Yu, Y.; Yang, D.; Zhou, S.; Gu, J.; Wang, F.; Dong, J.; Huang, R. The Ethylene Response Factor Oserf109 Negatively Affects Ethylene Biosynthesis And Drought Tolerance In Rice. Protoplasma, Pp.1-8. 2016. Doi:10.1007/S00709-016-0960-4

Zhang, A.D.; Hu, X.; Kuanga, S.; Ge, H.; Yin, X.R.; Chen, K.S. Isolation, Classification And Transcription Profiles Of The Ethylene Response Factors (Erfs) In Ripening Kiwifruit. Scientia Horticulturae, 2016 (199):209-215.

Zhang, G.; Chen, M.; Chen, X.; Xu, Z.; Guan, S, Li Lc, Li A, Guo J, Mao L, Ma Y. Phylogeny, Gene Structures, And Expression Patterns Of The Erf Gene Family In Soybean (*Glycine Max* L). J. Exp. Bot. ;2008 (59): 4095-4107.

Zhang, G.; Chen, M.; Li, L.; Xu, Z.; Chen, X.; Guo, J.; Ma, Y. Overexpression Of The Soybean Gmerf3 Gene, An Ap2/Erf Type Transcription Factor For Increased Tolerances To Salt, Drought, And Diseases In Transgenic Tobacco. J. Exp. Bot., 2009; 13 (60):3781-3796.

Zhuang, J.; Cai, B.; Peng, R-H.; Zhu, B.; Jin, X. F.; Xue Y.; Gao, F.; Fu, X. Y.; Tian, Y. S, Zhao, W.; Qiao, Y. S.; Zhang, Z.; Xiong, A. S.; Yao, Q. H. Genome-Wide Analysis Of The Ap2/Erf Gene Family In *Populus Trichocarpa*. Biochem. Biophys. Res. Commun., 2008; 3 (371): 468-474.

Zhuang, J.; Peng, R-H.; Cheng, Z-M.; Zhang, J.; Cai, B.; Zhang, Z.; Gao, F.; Zhu, B.; Fu, X-Y.; Jin, X-F.; Chen, J-M, Qiao, Y-S.; Xiong, A-S.; Yao, Q-H. Genome-Wide Analysis Of The Putative Ap2/Erf Family Genes In *Vitis Vinifera.* Sci Hortic., 2009; 1(123): 73-81.

Zhuang, J.; Deng, De-X.; Yao.; Q-H.; Zhang, J.; Xiong, F.; Chen, J-M.; Xiong, Ai- S. Discovery, Phylogeny And Expression Patterns Of Ap2-Like Genes In Maize. Plant Growth Regul., 2010; 1( 62): 51-58.

Zhuang, J.; Yao, Q-H.; Xiong, A. S.; Zhang, J. Isolation, Phylogeny And Expression Patterns Of Ap2-Like Genes In Apple (Malus × Domestica Borkh). Plant Mol. Biol. Rep., 2011; 1(29):209–216. doi:10.1007/s11105-010-0227-8

**Table 1** - *Coffea arabica* sequences with identity to ERF gene family in *Arabidopsis thaliana.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Arabidopsis thaliana* | |  | *Coffea arabica* | | | |
| Gene | Gene |  | ERF | Coverage (%) | e-value | Identity (%) |
| AT1G78080 | [RAP2.4](http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?maps=blast_set&db=at_refg&na=1&gnl=ref%7CNC_003070.9%7C&gi=240254421&term=240254421%5Bgi%5D&taxid=3702&RID=STS1N9XP01S&QUERY_NUMBER=1&log$=nuclalign) |  | CaERF01 | 99 | 1,00E-68 | 81 |
| AT1G78080 | [RAP2.4](http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?maps=blast_set&db=at_refg&na=1&gnl=ref%7CNC_003070.9%7C&gi=240254421&term=240254421%5Bgi%5D&taxid=3702&RID=STS1N9XP01S&QUERY_NUMBER=1&log$=nuclalign) |  | CaERF02 | 97 | 2,00E-63 | 81 |
| AT1G78080 | [RAP2.4](http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?maps=blast_set&db=at_refg&na=1&gnl=ref%7CNC_003070.9%7C&gi=240254421&term=240254421%5Bgi%5D&taxid=3702&RID=STS1N9XP01S&QUERY_NUMBER=1&log$=nuclalign) |  | CaERF03 | 100 | 2,00E-58 | 80 |
| AT1G78080 | [RAP2.4](http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?maps=blast_set&db=at_refg&na=1&gnl=ref%7CNC_003070.9%7C&gi=240254421&term=240254421%5Bgi%5D&taxid=3702&RID=STS1N9XP01S&QUERY_NUMBER=1&log$=nuclalign) |  | CaERF04 | 100 | 5,00E-62 | 80 |
| AT1G78080 | [RAP2.4](http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?maps=blast_set&db=at_refg&na=1&gnl=ref%7CNC_003070.9%7C&gi=240254421&term=240254421%5Bgi%5D&taxid=3702&RID=STS1N9XP01S&QUERY_NUMBER=1&log$=nuclalign) |  | CaERF05 | 100 | 2,00E-34 | 81 |
| AT5G67190 | AtERF10 |  | CaERF06 | 100 | 3,00E-39 | 74 |
| AT5G52020 | AtERF25 |  | CaERF07 | 63 | 8,00E-30 | 71 |
| AT5G52020 | AtERF25 |  | CaERF08 | 85 | 6,00E-45 | 65 |
| AT244940 | AtERF34 |  | CaERF09 | 65 | 6,00E-48 | 77 |
| AT240340 | DREB2C |  | CaERF10 | 84 | 9,00E-49 | 68 |
| AT240340 | DREB2C |  | CaERF11 | 78 | 8,00E-38 | 70 |
| AT1G75490 | DREB2D |  | CaERF12 | 87 | 3,00E-37 | 69 |
| AT4G27950 | CRF4 |  | CaERF13 | 71 | 3,00E-30 | 63 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF14 | 96 | 8,00E-34 | 72 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF15 | 93 | 2,00E-29 | 73 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF16 | 96 | 2,00E-29 | 68 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF17 | 62 | 1,00E-28 | 81 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF18 | 90 | 6,00E-27 | 81 |
| AT3G14230 | RAP2.2 |  | CaERF19 | 99 | 8,00E-33 | 77 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF20 | 85 | 8,00E-28 | 77 |
| AT3G16770 | ATEBP/RAP2.3 |  | CaERF21 | 96 | 1,00E-27 | 84 |
| AT3G15210 | AtERF4/RAP2.5 |  | CaERF22 | 42 | 3,00E-30 | 74 |
| AT3G15210 | AtERF4/RAP2.5 |  | CaERF23 | 41 | 5,00E-30 | 74 |
| AT1G50640 | AtERF3 |  | CaERF24 | 82 | 2,00E-42 | 76 |
| AT5G44210 | ATERF9 |  | CaERF25 | 97 | 4,00E-30 | 76 |
| AT1G28360 | ATERF12 |  | CaERF26 | 68 | 1,00E-26 | 85 |
| AT4G17500 | AtERF1 |  | CaERF27 | 75 | 9,00E-53 | 75 |
| AT4G17500 | AtERF1 |  | CaERF28 | 86 | 1,00E-44 | 73 |
| AT4G17500 | AtERF1 |  | CaERF29 | 79 | 5,00E-52 | 74 |
| AT3G23240 | ERF1 |  | CaERF30 | 82 | 4,00E-31 | 73 |
| AT4G17490 | AtERF-6 |  | CaERF31 | 97 | 5,00E-42 | 77 |
| AT4G17490 | AtERF-6 |  | CaERF32 | 84 | 4,00E-31 | 78 |
| AT4G17490 | AtERF-6 |  | CaERF33 | 93 | 7,00E-34 | 74 |
| AT3G23240 | ERF1 |  | CaERF34 | 54 | 2,00E-15 | 73 |
| AT5G61890 | ABR1 |  | CaERF35 | 94 | 2,00E-33 | 90 |
| At2G33710 | AtERF112 |  | CaERF36 | 49 | 4,00E-28 | 75 |

**Table 2 –** Number of genes in each Group of the ERF Family in *C. arabica* and in species whose genome was completely sequenced and size of the respective genomes.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Family Subfamily | Group | Arabidopsisa | Vitisb | Poplarc | Ricea | Coffea |
| ERF | I | 10 | 5 | 8 | 9 | 5 |
|  | II | 15 | 8 | 51 | 16 | 1 |
| DREB | III | 23 | 22 | 6 | 27 | 3 |
|  | IV | 9 | 5 | 10 | 6 | 3 |
|  | V | 5 | 11 | 42 | 8 | 0 |
|  | VI | 8 | 2 | 20 | 6 | 1 |
|  | VI-L | 4 | 5 | - | 10 | 0 |
| ERF | VII | 5 | 3 | 12 | 15 | 8 |
|  | VIII | 15 | 11 | 39 | 15 | 5 |
|  | IX | 17 | 40 | 19 | 18 | 8 |
|  | X | 8 | 10 | 2 | 12 | 2 |
|  | Xb-L | 3 | 0 | - | 3 | 0 |
| Total ERFs |  | 122 | 122 | 209 | 145 | 36 |
| Genome Size (Mb) | | 125 | 487 | 465 | 430 | 1.300 |

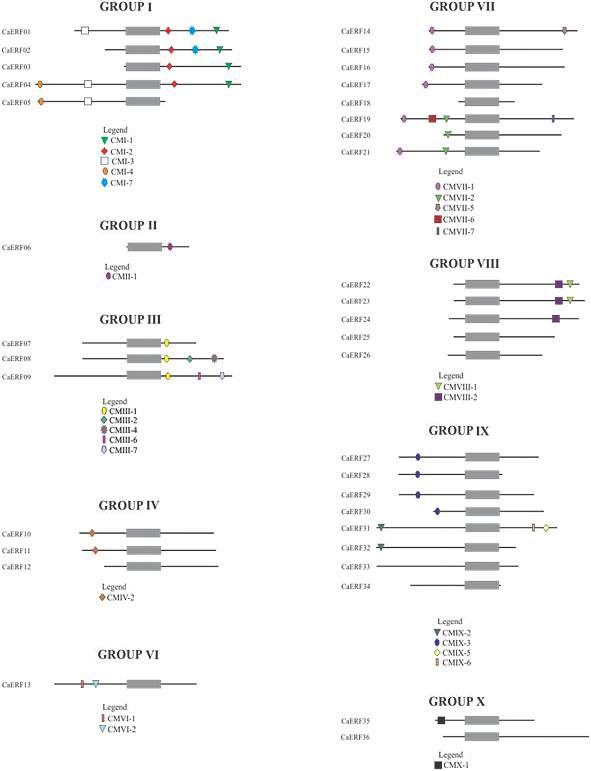
aNakano et al, (2006), bLicausi et al, (2010) and c Wang et al, (2016).

Imagemárvore

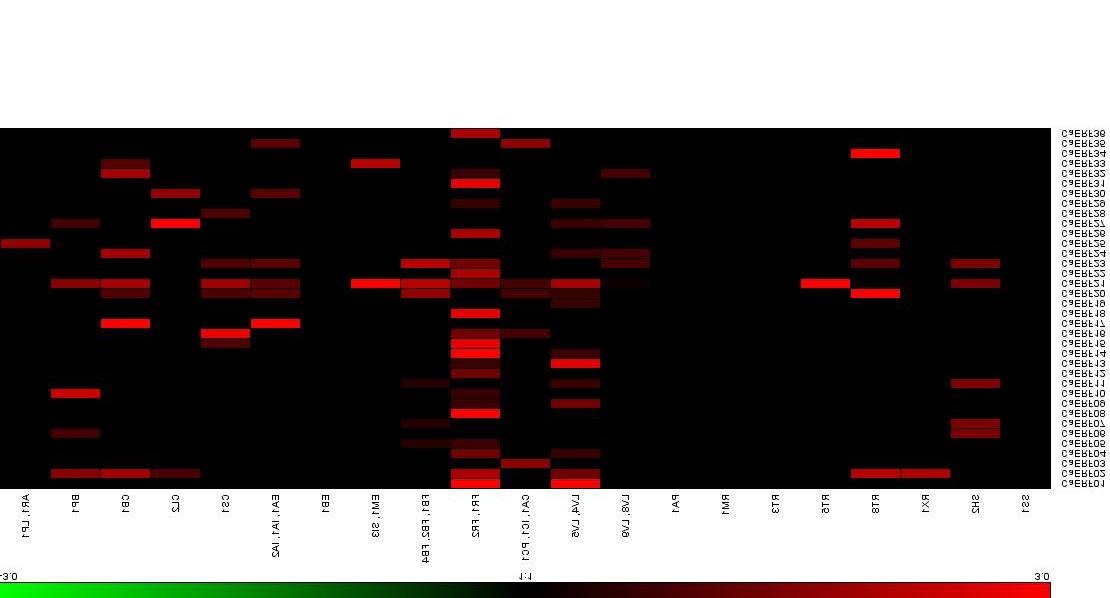
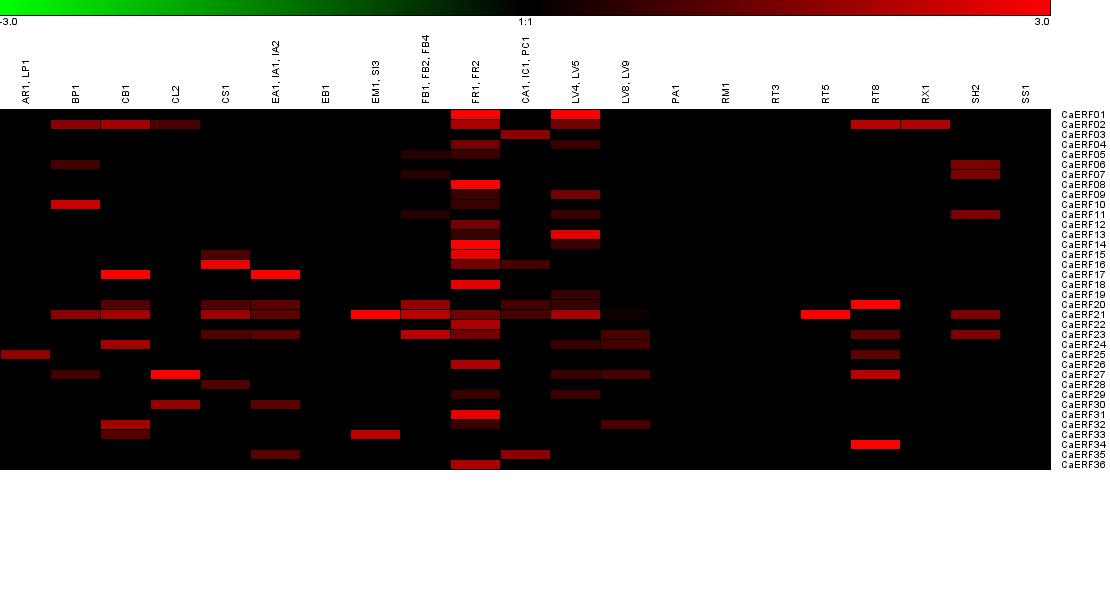
**Figure 1 –** Phylogenetic tree of the ERFs in *C. arabica* and arabidopsis. The amino acids sequences of the AP2 domain of 36 ERF genes of *C. arabica* and arabidopsis were aligned by Clustal Omega and the phylogenetic tree was constructed with MEGA 6.06 using the NJ method. (DREB subfamily: ○ - Group I, □ - Group II, Imagem1 - Group III, Imagem2 - Group IV; ERF subfamily: ▼- Group VI, ● – Group VII, ■ - Group VIII, ▲ – Group IX, Imagem3 - Group X).

Imagem12

**Figure 2 –** Alignmentof the AP2 domains from the ERF proteins of *C. arabica*. Dark and light blue shading indicates identical and conserved amino acids residues, respectively. The black bar and arrows represent the predicted *α-helix* and *β-sheets* regions, respectively. *CaERF01-12* belongs to the DREB subfamily; *CaERF13*-*36* belongs tothe ERF subfamily.



**Figure 3 –** Distribution of the conserved motifs in the ERF family in *Coffea arabica.* The motifs were identified in *C. arabica* and classified according to the classification proposed by Nakano et al, (2006). The position and order of the symbols corresponds to the actual conserved region in the protein sequence.



3

0

Relative expression

**Figure 4 –** *In-silico* expression profiling of the ERFs in *Coffea arabica*. The number of *reads* was normalized in each library and values were represented by the relative expression scale. Libraries: AR1, LP1-Plantlets and leaves treated with araquidonic acid; BP1-Suspension cells treated with acibenzolar-S-methyl; CB1- Suspension cells treated with acibenzolar-S-methyl and brassinoesteroids; CL2-Hypocotyls treated with acibenzolar-S-methyl; CS1- Suspension cells treated with NaCl; EA1, IA1, IA2-Embryogenic calli; EB1-Zygotic embryo (immature fruits); EC1- Embryogenic calli from *Coffea canephora*; EM1, S13- Germinating seeds (whole seeds and zygotic embryos); FB1. FB2, FB4- Flower buds in different developmental stages; FR1, FR2- Flower buds + pinhead fruits + fruits at different stages; FR4- Fruits (*Coffea racemosa*); FV2- Fruits, stages 1, 2 and 2 (*Coffea racemosa*); CAI, ICI, PCI-Non embryogenic calli with and without 2, 4 D; LV4, LV5- Young leaves from orthotropic branch; LV8, LV8- Mature leaves from plagiotropic branches; NS1- Roots infected with nematodes; PAI-Primary embryogenic calli; RMI-Leaves infected with leaf miner and coffee leaf rust; RT3-Roots; RT5- Roots with acibenzolar-S-methyl; RT8- Suspension cells with stressed with aluminum; RX1- Stems infected with *Xylella ssp.*; SH1-Leaves from water deficit stresses plants (*Coffea canephora*); SH2-Water deficit stresses field (pool of tissues); SS1-Well-watered field (pool of tissues).