Genome-wide analysis of the AP2/ERF superfamily in apple and transcriptional evidence of ERF involvement in scab pathogenesis
Cesar Luis Girardi, César Valmor Rombaldi, Joceani Dal Cero, Paula Macedo Nobile, François Laurent, Mondher Bouzayen, Vera Quecini

To cite this version:

HAL Id: hal-01071932
https://hal.archives-ouvertes.fr/hal-01071932
Submitted on 8 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://oatao.univ-toulouse.fr/
Eprints ID: 11480

To link to this article: DOI:10.1016/j.scienta.2012.12.017
URL: http://dx.doi.org/10.1016/j.scienta.2012.12.017

To cite this version:

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes.diff.inp-toulouse.fr
Genome-wide analysis of the AP2/ERF superfamily in apple and transcrip- tional evidence of ERF involvement in scab pathogenesis

César Luís Girardi A, César Valmor Rombaldi b, Joceani Dal Cero b, Paula M. Nobile c, François Laurens c, Mondher Bouzayan d, Vera Quecini a,∗

A B S T R A C T

The APETALA2 (AP2)/ETHYLENE RESPONSE FACTOR (ERF) superfamily of transcriptional regulators is involved in several growth, development and stress responses processes in higher plants. Currently, the available information on the biological roles of AP2/ERF genes is derived from Arabidopsis thaliana. In the present work, we have investigated genomic and transcriptional aspects of AP2/ERF genes in the economically important perennial species, Malus × domestica. We have identified 259 sequences containing at least one ERF domain in apple genome. The vast majority of the putative proteins display predicted nuclear localization, compatible with a biological role in transcription regulation. The AP2 and ERF families are greatly expanded in apple. Whole-genome analyses in other plant species have identified a single genomic sequence with divergent ERF, whereas in apple seven solosons are present. In the apple genome, the most noteworthy expansion occurred in sub-groups V, VIII and IX of the ERF family. Expression profiling analyses have revealed the association of ripening-involved ERF genes to scab (Venturia inequalis) pathogenesis in the susceptible Gala cultivar, indicating that gene expansion processes were accompanied by functional divergence. The present analyses of AP2/ERF genes in apple provide evidences of shared ethylene-mediated signaling pathways in ripening and disease responses.

1. Introduction

Hormones are responsible for the control of plant growth and development and are also involved in mediating responses to several biotic and abiotic stresses. Besides its important role in regulating distinct physiological processes; such as seed germination, cell elongation, woundind, senescence and abscission, ethylene also regulates the plant responses to biotic and abiotic stresses (Chen et al., 2002). Ethylene exerts its action through complex regulatory steps of its biosynthesis, perception and signal transduction, leading to dramatic changes in gene expression (Chen et al., 2002; Kendrick and Chang, 2008). The promoter sequences of several genes induced by ethylene were found to contain a cis-regulatory element known as the ethylene-responsive element (ERE) (Broglie et al., 1989). Sequence analysis of various ERE regions identified a short motif, rich in G/C nucleotides, labeled the GCC-box, essential for ethylene-mediated responses. The ERE motif is recognized by a family of transcription factors: the ERE binding factors (ERF) (Fujimoto et al., 2000).

The ERF is a large gene family of transcription factors that constitute a sub-group of the APETALA2 (AP2)/ERF superfamily, which also contains the AP2 and RAV families (Riechmann et al., 2000). The superfamily is defined by the presence of the AP2/ERF domain, a conserved DNA-binding sequence consisting of approximately 60–70 amino acids (Weigel, 1995). The AP2 family consists in proteins with two AP2/ERF domains, whereas the ERF family is constituted of proteins containing a single AP2/ERF domain, and the RAV family consists of proteins characterized by the presence of a B3 domain, a DNA-binding domain conserved in other plant-specific transcription factors, in addition to a single AP2/ERF domain. Previous studies have proposed a sub-division of the ERF family into two major sub-families; the ERF subfamily and the CBF/DREB subfamily (Sakuma et al., 2002). More recent data suggest the existence of ten groups containing distinct conserved motifs within the family rather than distinct sub-families (Nakano et al., 2006).

∗ Corresponding author at: CNPUS (National Center for Grapevine and Wine Research), Embrapa (Brazilian Agricultural Corporation), Rua Livramento, 515, 95700-000 Bento Gonçalves, RS, Brazil. Tel.: +55 54 3455 8000; fax: +55 54 3451 2792.
E-mail address: vera.quecini@embrapa.br (V. Quecini).
The role of the ERF proteins in stress-related, important agronomical responses (Xu et al., 2011) and the accumulation of genomic and expressed sequence data have prompted in-depth bioinformatic investigation of the ERF sub-family in several plant species, including Arabidopsis, rice, Populus, soybean and grapevine (Nakano et al., 2006; Licausi et al., 2010; Zhang et al., 2008; Zhuang et al., 2008). In fruit species, ERFs are associated to flavor biosynthesis and texture modification during ripening (Alba et al., 2005; da Silva et al., 2005; Janssen et al., 2008; Licausi et al., 2010; Zilio et al., 2008). In apple (Malus domestica Borkh. cv. Golden Delicious), two ERF transcripts, MdERF1 and MdERF2, demonstrated to be predominantly and exclusively expressed in ripening fruits, respectively, were isolated from ripening fruit (Wang et al., 2007). Similarly, Newcomb et al. (2006) identified expressed sequences of ethylene response factors in the tag collection and the pipeline annotation of the draft apple genome has found 274 sequences corresponding to AP2/ERF factors (Velasco et al., 2010).

Due to its importance in plant reproductive development and abiotic and biotic stress responses, AP2/ERF transcription factors represent interesting gene pools to be investigated for breeding and genetic engineering purposes (Xu et al., 2011). Thus, we have performed an in-depth investigation of AP2/ERF superfamily in apple, demonstrating the presence of 259 sequences containing at least a single AP2 domain. The AP2, ERF and Arabidopsis soloist families have undergone differential expansion in apple. Groups V, VIII and IX of the ERF family were preferentially expanded. In apple, ERF genes are expressed in vegetative and reproductive tissues and an ERF gene, previously associated to ripening, has been demonstrated to be induced in Venturia inaequalis-infected tissue. Thus, suggesting a role for ethylene and ethylene response genes in fungal pathogenesis in apple.

2. Materials and methods

2.1. Database searches and alignments

Sequences of Arabidopsis thaliana AP2/ERF genes were retrieved from the TAIR website and the amino acid sequence of the members described by Nakano et al. (2006) were used as baits to search the apple genome at the Genome Database for Rosaceae (GDR) using the Batch BLAST search (Altschul et al., 1997). In order to retrieve all AP2-containing sequences, the consensus amino acid domain sequence was also employed to search the NCBI database for Malus. The recovered sequences were checked by reverse BLAST analyses against the Arabidopsis genome and sequences failing to retrieve the original bait were eliminated. The presence and the significance of AP2 domains in the recovered apple sequences were determined by SMART analyses (Letunic et al., 2008).

2.2. Phylogenetic analysis

The functionality of Malus genes in comparison to their Arabidopsis counterparts was assessed by genetic distance and phylogenetic studies. Amino acid sequence alignments were performed using ClustalX (Thompson et al., 1997). When necessary, alignments were manually adjusted using Lasergene MegAlign (DNASTAR, Madison, WI, USA). Phylogenetic analyses were performed using distance and parsimony methods in the software PAUP* 4.0b10 (http://paup.csit.fsu.edu/), using the software default parameters. Resampling bootstrap trees containing 1000 random samples were constructed using PISCANGIT software (http://www.bootstrap-software.org/). Modular functional domains were employed for genetic distance studies for genes previously characterized as having divergent regions and conserved blocks.

2.3. Motif analysis and in silico characterization

Conserved motifs were further investigated by multiple alignment analyses using ClustalX and the MEME version 4.7 suite (Bailey and Elkan, 1994). The presence and sequence conservation of recognizable functional domains was studied employing protein analysis and gene function tools from databases (European Bioinformatics Institute-European Molecular Biology Laboratory – EMBL-EBI; Expert Protein Analysis System – ExPaSy from Swiss Institute of Bioinformatics – SIB; Gene Ontology database – GO; Protein Families database – Pfam).

2.4. Prediction of cellular and molecular parameters

The physical characteristics; isoelectric point (IP) and predicted molecular weight, of apple ERF sequences were calculated from the deduced amino acid using default parameter of the Lasergene MegAlign software (DNASTAR, Madison, WI, USA). The proteins folding state was predicted by the FoldIndex program (http://biportal.weizmann.ac.il/foldbin/findex). The prediction of the sub-cellular compartment location of the apple ERF-like proteins was performed using the plant algorithm of the software WoLF PSORT (http://www.wolfsort.org/) (Horton et al., 2007).

2.5. Gene expression analysis

Qualitative gene expression profiling was performed by in silico analyses of the Malus EST databases using virtual northern blot analyses. The gene of interest was used in queries against reference sequence databases, generating an alignment of the input gene to its orthologs. The resulting alignment was used to find sequences in the entire mRNA input that are specific to the gene (probe). The resulting alignments were collectively used to query the EST database again using BLAST. This heuristic was critical to avoid false-positives, or ESTs from a paralog of the input gene rather than the gene itself. The identity numbers of the ESTs matching the probes were recovered and the databases were used to find the names of the libraries from which those ESTs were derived. The description of the libraries used is available at The Malus × domestica Gene Index, at the Computational Biology and Functional Genomics Laboratory (Dana Farber Cancer Institute and Harvard School of Public Health, http://compbio.dfc.rockefeller.edu/cgi/gi/mdgi/search/xpress_search.html) and in Supplementary Table IV. The frequency of reads of each EST contig in a given library was calculated and normalized according to the total number of reads from the investigated library and the total number of reads in all libraries. A correlation matrix between EST contigs and libraries was then generated and gene expression patterns among ESTs and libraries were obtained by hierarchical clustering based on Spearman Rank correlation matrix using Cluster v.2.11 software (Eisen et al., 1998), by substituting the clusters by their average expression pattern. Graphic outputs were generated using Tree View v.1.6 software (http://www.ncbi.nlm.nih.gov/Trees/Cluster) and presented in cool scale.

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scientia.2012.12.017.

2.6. Plant material and V. inaequalis inoculation

Lateral shoots of Malus × domestica ‘Royal Gala’ were propagated in MS medium (Murashige and Skoog, 1962) pH 5.8 supplemented with 8.8 μM of 6-benzylaminopurine, 30 g·L−1 sucrose and solidified with 6 g·L−1 agar, under 16-h photoperiod and 22 ± 2°C temperature. After four weeks of culture, plants
Table 1
Summary of the AP2/ERF superfamily of Malus × domestica and model species Arabidopsis thaliana, Vitis vinifera, Populus trichocarpa and Oryza sativa. The total data for AP2, At4g13040, RAV and ERF families are presented in bold letters.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Group</th>
<th>Species</th>
<th>Arabidopsis thaliana</th>
<th>Malus × domestica</th>
<th>Vitis vinifera</th>
<th>Populus trichocarpa</th>
<th>Oryza sativa</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP2 family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single AP2/ERF</td>
<td>4</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double AP2/ERF</td>
<td>14</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divergent ERF</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAV family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP2/ERF and B3</td>
<td>6</td>
<td>122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERF family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>15</td>
<td>13</td>
<td>8</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>23</td>
<td>22</td>
<td>22</td>
<td>35</td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>9</td>
<td>23</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>5</td>
<td>19</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI – L</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIII</td>
<td>15</td>
<td>31</td>
<td>11</td>
<td>17</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td>17</td>
<td>45</td>
<td>40</td>
<td>42</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>8</td>
<td>12</td>
<td>10</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xb – L</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>147</td>
<td>259</td>
<td>149</td>
<td>202</td>
<td>180</td>
</tr>
</tbody>
</table>

were transferred to rooting medium, containing MS salts and vitamins at pH 5.8 and supplemented with 5 μM of indole-3-butyric acid, 15 g L⁻¹ sucrose and solidified with 6 g L⁻¹ agar to induce rooting. Rooted plants were gradually ex vitro acclimated and four weeks later, the plants were inoculated with the fungus. V. inequalis (apple scab) inocula were prepared from monoconidial isolates cultivated in Potato-Dextrose-Agar (PDA) medium. The plates were incubated at 16 °C under continuous light for 15 days and sporulating colonies were scraped for inocula preparation.

Conidia were resuspended in saline solution and the concentration was adjusted to 2.5 × 10⁵ conidia/mL before plant spraying. The plants were sprayed with approximately 50 mL of conidia suspension until the formation of droplets. Negative control plants were sprayed with sterile saline solution without conidia. After inoculation, the plants were kept for 30 days under 20 °C, 90–100% relative humidity and 12-h photoperiod and symptom development was visually evaluated. Leaf samples were harvested after 30 days, immediately frozen in liquid nitrogen and stored at −80 °C until processed.

Fig. 1. Phylogenetic analyses of AP2/ERF superfamily in Malus × domestica. Neighbor-joining tree for full length sequences were aligned with ClustalX. Sub-families AP2, RAV and the Arabidopsis-like soloist sequences in apple genome are represented by gray shading. Roman numerals represent ERF family sub-groups.
2.7. RNA extraction and cDNA synthesis

Total RNA was extracted as described by Boss et al. (1996) and the samples were treated with 100 U of RNase-free DNase (NEB Biolabs, USA) and further purified with chlorophorm wash and ethanol precipitation. The purity and concentration of the isolated RNAs were checked on 1% (w/v) agarose gels and by spectrophotometry. Reverse transcription was carried out using 2 μL of total RNA and M-MLV Reverse Transcriptase RT-PCR (Promega) and oligo d(T)25 primers, as recommended by the manufacturer. The concentration of the cDNA was determined spectrophotometrically.

2.8. Semi-quantitative PCR analyses

Polymerase chain reactions were carried out in a final volume of 25 μL containing 2.5 μL of 10× PCR buffer, 2.5 μL of MgCl2 (25 mM), 4.0 μL of dNTP mix (2.5 mM), 1 μL of each primer (10 mM), 1 μL cDNA and 2.0 μL Taq DNA polymerase. The sequence of the primers used is as follows: 5′-ATGACCTGGTGCGATATCAG-3′ and 5′-CACCGTAGCAGAAACACACAC-3′ for MDP0000128979, 5′-TATGCTGGCAATTGGCGAGC-3′ and 5′-ATGACCAATCCCGCAC-3′ for MDP0000128979, 5′-GCCTGGATTTGCTGGTGATG-3′ and 5′-GTGCCTTGCAGTTCTGAGAT-3′ and 5′-TGTCCTATGCGGTGCTCA-3′ for a normalizer gene MdAct (Wang et al., 2007). The reaction linear range was determined by serial tests of cycle lengths and was determined to be linear at 27 amplification cycles. Digital analyses of band intensity on agarose gels were performed to quantify gene expression.

3. Results

3.1. The family of ethylene response factors in apple

Sequences containing AP2 domains were extensively searched in the apple genome and 259 were confirmed by reverse BLAST and SMART domain analyses to be members of the AP2/ERF superfamily (Table 1, Table S1). The number of identified sequences is smaller than that found by automated annotation of the apple genome (Velasco et al., 2010) (274) due to the elimination of short sequences and of those with introns longer than 10 kb, which are likely resultant from assembly mistakes. The AP2 family in apple consists of 51 genes, from which 36 contain the characteristic double AP2 domain. The RAV family of proteins, containing the B3 domain in addition to the family characteristic AP2, consists of six members in apple as observed in other investigated plant species and its member number appears to be highly conserved throughout evolution. One hundred and ninety-five genes encoding proteins with a single AP2/ERF domain were assigned to the ERF family. In Arabidopsis and other higher plant species, a single locus (At4g13040) codes for a protein exhibiting divergent AP2 domain was identified, thus it was labeled the soloist (Nakano et al., 2006; Zhuang et al., 2008; Licausi et al., 2010). In apple genome the group consists of seven closely related proteins containing the divergent AP2 domain, similar to that of At4g13040 (Figs. 1 and 2). Apple ERF proteins were classified into 11 groups according to their similarity to Arabidopsis ERF sequences and conserved motifs outside the DNA-binding
AP2 domain were identified (Figs. 1 and 2, Table SII). Although the vast majority of the Arabidopsis motifs were present in apple sequences, specific conserved regions were identified in the proteins identified in Malus × domestica (Figs. S1–S10). The apple ERF family (195 members) is expanded in comparison to the Arabidopsis (147), poplar (202) and rice (180) families. The expansion of the apple family is significant for ERF proteins clustered in group V, VIII and IX (Table 1, Figs. S1–S10). Interestingly, groups V and IX are also expanded although no functional significance has been attributed to those groups (Zhuang et al., 2008). The Xb-like group of ERF proteins is absent from the grapevine genome, however, in apple it consists of two proteins, similar to the numbers found in Arabidopsis (3) and Populus (4) (Nakano et al., 2006; Zhuang et al., 2008; Licausi et al., 2010).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2012.12.017.

3.2. Accuracy of protein AP2/ERF predictions

Genome databases employ ab initio protein prediction methods, which are prone to errors. In order to confirm the existence of the identified AP2/ERF proteins, we have checked intron junctions in comparison to the homologs in Arabidopsis and poplar. Subsequently, the predicted mRNAs of all identified proteins were used to query apple EST databases. We found no unrealistic introns and EST sequences matching (e-value lower than 1e−25) all predicted mRNA were found in the databases, except for gene model MDP0000876858.

3.3. Cellular and molecular characteristics of apple AP2/ERF proteins

We have employed bioinformatic tools to determine the physical properties of the ERF-like sequences from apple, such as the molecular weight (MW), pH value of isoelectric point (pI) and folding index. The vast majority of the apple ERF-like sequences were predicted to have low molecular weight, ranging from 1.5 to 4.48 kDa and an average of 233 amino acids, going from 111 to 413 amino acids (Table SIII). In apple, most of the sequences sharing sequence similarity to Arabidopsis ERF proteins (74.2%, 49) were predicted to be basic to neutral, whereas only 25.8% (17) displayed predominance of acidic amino acids (Table SIII).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2012.12.017.

The folding states of ERF family proteins in apple were predicted by FoldIndex program (Prilusky et al., 2005). The majority
of the identified apple sequences sharing sequence similarity to ERFs were predicted to be unfolded (59.1%, 39) under physiological conditions (Table III). The predicted sub-cellular localization of the ERF proteins identified in Malus was investigated using the software WoLF PSORT (Horton et al., 2007). Most of the ERF-like sequences from apple (83.3%, 55) was predicted to be exclusively nuclear or with ambiguous sub-cellular location for the nucleus and organelle or cytoplasm (Table III) (Fig. 3).

3.4. Expression profiling of ERF genes

The expression profile of 195 ERF genes was preliminarily investigated by in silico analyses for reproductive and vegetative libraries. Moreover, the relative expression of the 10 groups constituting the sub-family was also investigated in scab infected tissues. The expression patterns of the genes encoding for members of the ERF family, presented as groups, are shown in Figs. 4 and 5. Transcripts corresponding to the vast majority of the identified genes were expressed in apple leaves, flowers and fruits, except for gene models MDP0000451365 (group III), group V sequences MDP000031419, MDP0000525933, MDP000026536, MDP000051969, MDP0000246184, MDP0000127054, MDP0000297646 and MDP0000343375, MDP0000467753 and MDP0000778140 (group VIII) and group IX sequences MDP0000445770, MDP0000313854, MDP0000829925, MDP0000155543, MDP00002589227, MDP0000880063, MDP0000087369 that were expressed in buds, roots and vascular vessel tissues. The expression profile of most ERF genes is developmentally and/or environmentally controlled in apple (Figs. 4 and 5). Within ERF groups, the expression patterns were also divergent, especially for groups I, VII, VIII, and IX (Figs. 4 and 5). Transcripts corresponding to the sequences MDP0000128979 and MDP000316843, from group Xb, were exclusively present in fruits (Figs. 4 and 5). Most ERF genes from groups VII and IX expressed in the fruits were also highly expressed in V. inequalis infected tissues (Figs. 4 and 5).

Thus, we have further investigated the transcriptional profile of ripening-associated ERFs in scab pathogenesis using monosporic V. inequalis cultures inoculated to in vitro propagated plants and semi-quantitative PCR.

Thirty days after inoculation, plant leaves exhibited scab symptoms in the leaves and stem and a three-fold induction of MDP0000128979 (group VII) expression (Fig. 6). Phylogenetic analyses demonstrated that MDP0000128979 correspond to fruit-ripening associated MdERF1 (Wang et al., 2007). In silico analyses suggested the association of the induced transcription of group IX genes to scab response and fruit ripening (Fig. 5), however transcription of MDP0000226115 was only slightly altered in scab infected tissues (Fig. 6).

4. Discussion

4.1. Identification of the AP2/ERF family transcription factors in apple

A comprehensive search for AP2/ERF transcription factors in Malus × domestica genome has revealed the presence of 259
AP2-motif containing sequences. The number is smaller than the 274 AP2/ERF found by the automated annotation of the apple genome (Velasco et al., 2010) due to the elimination of short sequences and of those containing unrealistic introns. The apple superfamily is larger than the rice and Populus ones, with 180 and 202 genes, respectively (Table 1). The larger size of the superfamily in apple is mainly due to expansion of the AP2, ERF and At4g13040-like families, since the number of genes in the RAV family is conserved throughout evolution (Table 1). In apple genome, recent gene duplication appears to be involved in the expansion of the At4g13040-like genes, since it is represented by a single sequence with divergent AP2 domain in all investigated plant species, whereas in apple, it is represented by seven sequences. The sub-set of AP2 proteins containing a doubled AP2 domain in apple is more than two times larger than it is in Arabidopsis and has approximately 40% more members than the one in Populus trichocarpa. Interestingly, the expansion pattern of the families is conserved in the woody perennial species investigated; poplar, grapevine and apple (Zhuang et al., 2008; Licausi et al., 2010).

The ERF family in apple consists of 195 sequences phylogenically classified into the ten groups (I–X) identified in Arabidopsis by Nakano et al. (2006). The distribution of apple ERF sequences in groups is similar to that of other investigated species. As observed in woody perennial Vitis vinifera and Populus trichocarpa, groups V and IX are expanded in apple genome, although both also display a lower number of genes in group I, which does not occur in apple. A recent study in Vitis vinifera has proposed that the lower number of ERF genes in group I woody perennial is likely to be due to its functional redundancy in controlling wax accumulation (Aharoni et al., 2004) to the genes found in expanded group V (Licausi et al., 2010). The differential expression and the expansion of both groups I and V suggest that in apple it does not occur. In contrast, the low levels of expression of group IV genes in apple may indicate functional redundancy to other ERF gene group. The group XB in apple is exclusively expressed in fruits and has a number of genes similar to that found in Arabidopsis and poplar. The group is absent from grapevine. The largest ERF group in apple is group IX, as observed in poplar and grape genomes, it is almost twice the size of the group in Arabidopsis and rice (Table 1). As observed in grapevine (Licausi et al., 2010), apple ERF-IX sequences exhibit species-specific protein domains (Supplementary Table SII) and distinct expression patterns. Interestingly, members of group IX displayed a particular clustered organization along chromosomes 7 and 16 (Fig. 2), similar to the one observed in the poplar and grapevine genome (Zhuang et al., 2008; Licausi et al., 2010), although in this species the tandem repetition does not exceed four genes. The coding sequences of proteins involved in signal transduction and transcriptional regulation are frequently duplicated in higher plant genomes, due to their role in plant fitness responses (Paterson et al., 2010).

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2012.12.017.

Transcription factors often exhibit regions outside the DNA-binding domain that are involved in regulation of the transcriptional activity, protein-protein interactions, and nuclear localization (Liu et al., 1998). These motifs, frequently associated to specific functions or regulation patterns, are characteristic of large transcription factor families in plants, such as MYB, WRKY, NAC, Dof, GATA, and GRAS (Eulgem et al., 2000; Kranz et al., 1998; Ljivetzky et al., 2003; Ooka et al., 2003; Reyes et al., 2004; Tian et al., 2004). The vast majority of the ERF-like sequences identified in apple shared one or more motifs outside the AP2/ERF domain with their
Arabidopsis counterparts, as observed in rice, Populus, soybean and grapevine (Nakano et al., 2006; Licausi et al., 2010; Zhang et al., 2008; Zhuang et al., 2008). The functionally characterized gene MdERF1 (A1LM1), associated to fruit ripening in apple (Wang et al., 2007) clustered in the most abundant group VII, whereas MdERF2 (A1LM2), which is exclusively expressed in ripening fruits was clustered in group IX.

4.2. Cellular and molecular characteristics of apple ERF-like proteins

Proteins biochemical and physical properties are responsible for several characteristics associated to their biological function, such as molecular structure, ligand binding and subcellular location. Similar to Arabidopsis, rice, Populus, soybean and grapevine (Nakano et al., 2006; Licausi et al., 2010; Zhang et al., 2008; Zhuang et al., 2008), apple ERF-like proteins are predicted to have low molecular weight. In apple, most of the ERF-like sequences were predicted to be basic to neutral, as observed for Arabidopsis and Populus, although the proteins from the annual and perennial model plants exhibit slightly lower pl values. These observations are consistent with the presence of a basic region mediating sequence-specific DNA-binding in families of transcriptional regulators from evolutionarily distant organisms (Miller et al., 2003; Panne et al., 2004). Thus, the predicted chemical nature of the proteins identified in apple is compatible with a functional role in transcriptional control.

The predicted sub-cellular localization of the ERF proteins identified in Malus was investigated using the software WoLF pSORT (Horton et al., 2007). The sub-cellular location is an important clue for the biological function of identified sequences, especially for those hypothesized to be involved in gene expression regulation. The majority of the ERF-like sequences from apple were predicted to be exclusively nuclear or exhibiting ambiguous sub-cellular location for the nucleus and organelle or cytoplasm. At this point, the incomplete nature of some of the sequences and the limitations of bioinformatic analyses prevent us from establishing how accurate the predictions are in vivo. Further functional analyses will be necessary to determine the sub-cellular location of the apple ERF-like sequences.

4.3. Expression analyses of ERF genes

The expression profiling of ERF genes in apple along with the knowledge on their molecular function in model plant species is an important tool to uncover the biological role of these transcriptional regulators from evolutionarily distant organisms.
Acknowledgments

The authors would like to acknowledge the excellent technical assistance of Daniela Dal Bosco and Iraci Sinski in plant propagation and acclimation. We also thank Renata Gava for fungus cultivation and plant inoculation. This research is funded by an Embrapa grant number 02.07.05.001.00 Macroprograma 2 – Agrofuturo to C.L.G. VQ is recipient of a productivity research grant from CNPq (307031/2010-1).

References


