



Genome-wide analysis of the *AP2/ERF* superfamily in apple and transcriptional 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 Bouzayen^d, Vera Quecini^{a,*}

^a Embrapa Uva e Vinho, Caixa Postal 130, 95700-000 Bento Gonçalves, RS, Brazil

^b Universidade Federal de Pelotas, Faculdade de Agronomia Eliseu Maciel, Departamento de Ciência e Tecnologia Agroindustrial, s/n° caixa postal 354, 96010-900 Pelotas, RS, Brazil

^c UMR 1259 Génétique et Horticulture (GenHort), Angers, France

^d UMR 990, Génomique et Biotechnologie des Fruits, INP, ENSAT, INRA, F-31320 Castanet Tolosan, France

ARTICLE INFO

Article history:

Received 17 October 2012

Received in revised form

11 December 2012

Accepted 13 December 2012

Keywords:

Apple

Ethylene

Fungal pathogen

Fruit ripening

Physiological conditions

Venturia inaequalis

ABSTRACT

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 soloists 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 inaequalis*) pathogenesis in the susceptible Gala cultivar, indicating that gene expansion processes were accompanied by functional divergence. The presented analyses of *AP2/ERF* genes in apple provide evidences of shared ethylene-mediated signaling pathways in ripening and disease responses.

© 2012 Elsevier B.V. All rights reserved.

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, wounding, ripening, 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: CNPQV (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; Ziliotto 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 PSIGNFIT 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://bioportal.weizmann.ac.il/flddb/finindex>). 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://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.dfci.harvard.edu/tgi/gi/mdgi/searching/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://rana.lbl.gov/EisenSoftware.htm>) and presented in colorscale.

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>.

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.

Classification	Group	Species				
		<i>Arabidopsis thaliana</i>	<i>Malus × domestica</i>	<i>Vitis vinifera</i>	<i>Populus trichocarpa</i>	<i>Oryza sativa</i>
AP2 family		18	51	20	26	29
	Single AP2/ERF	4	15	0	0	5
	Double AP2/ERF	14	36	20	26	24
At4g13040	Divergent ERF	1	7	1	1	1
RAV family	AP2/ERF and B3	6	6	6	6	5
ERF family		122	195	122	169	145
	I	10	10	5	5	9
	II	15	13	8	20	16
	III	23	22	22	35	27
	IV	9	23	5	6	6
	V	5	19	11	10	8
	VI	8	6	5	11	6
	VI – L	4	4	2	4	3
	VII	5	8	3	6	15
	VIII	15	31	11	17	15
	IX	17	45	40	42	18
	X	8	12	10	9	12
	Xb – L	3	2	0	4	10
	Total	147	259	149	202	180

were transferred to rooting medium, containing MS salts and vitamins at pH 5.8 and supplemented with 5 μ M of indole-3-butyric acid, 15 gL⁻¹ sucrose and solidified with 6 gL⁻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 2.5×10^5 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.

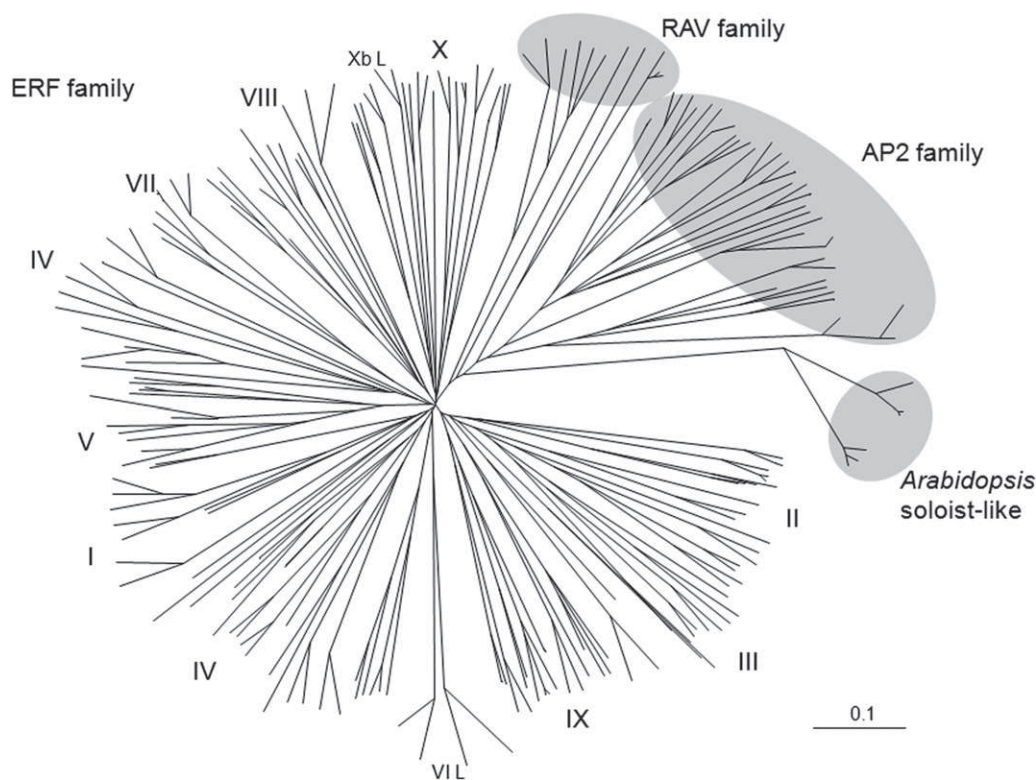


Fig. 1. Phylogenetic analyses of AP2/ERF superfamily in *Malus × domestica*. Neighbor-joining tree for full length sequences were aligned with ClustalX. Sub-families AP2, RAV and the *Arabidopsis*-like soloist sequences in apple genome are represented by gray shading. Roman numerals represent ERF family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

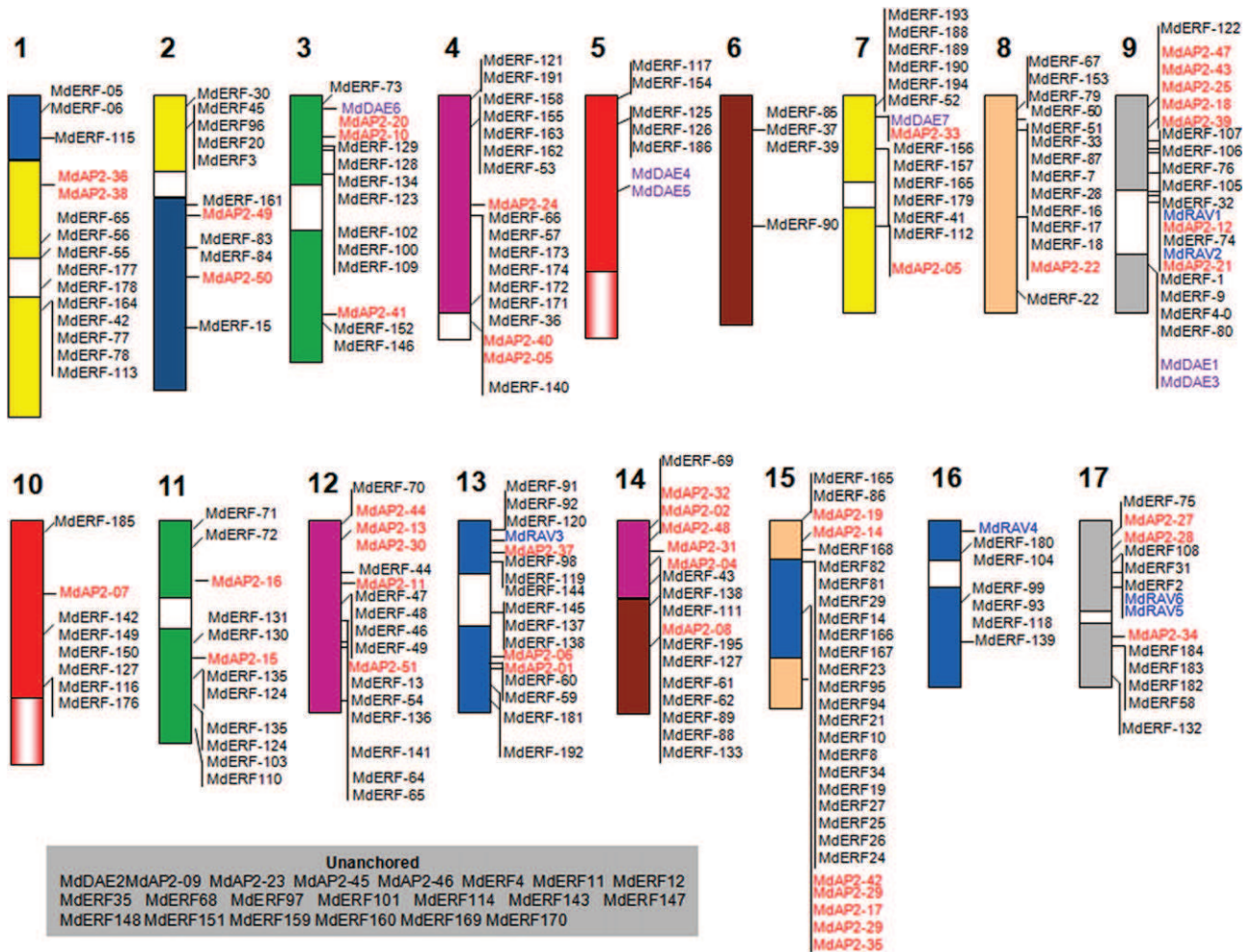


Fig. 2. Chromosome location of ERF genes in apple genome. *M. × domestica* chromosomes are represented according to the evolution model proposed by Velasco et al. (2010).

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 μ L 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 μ L, containing 2.5 μ L of 10 \times PCR buffer, 2.5 μ L of MgCl₂ (25 mM), 4.0 μ L of dNTP mix (2.5 mM), 1 μ L of each primer (10 mM), 1 μ L cDNA and 2.0 μ L Taq DNA polymerase. The sequence of the primers used is as follows: 5'ATGACCTGGTGGCATATCAG3' and 5'CACCGTAGCAAACAACACAC3' for MDP0000128979, 5'TATGCTGGCAATTGGCGAGC3' and 5'ATGACCAATCCCGCAC-TCAC3' for MDP0000226115 and 5'GGCTGGATTGCTGGTGATG3' and 5'TGCTCACTATGCCGTCTCA3' 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

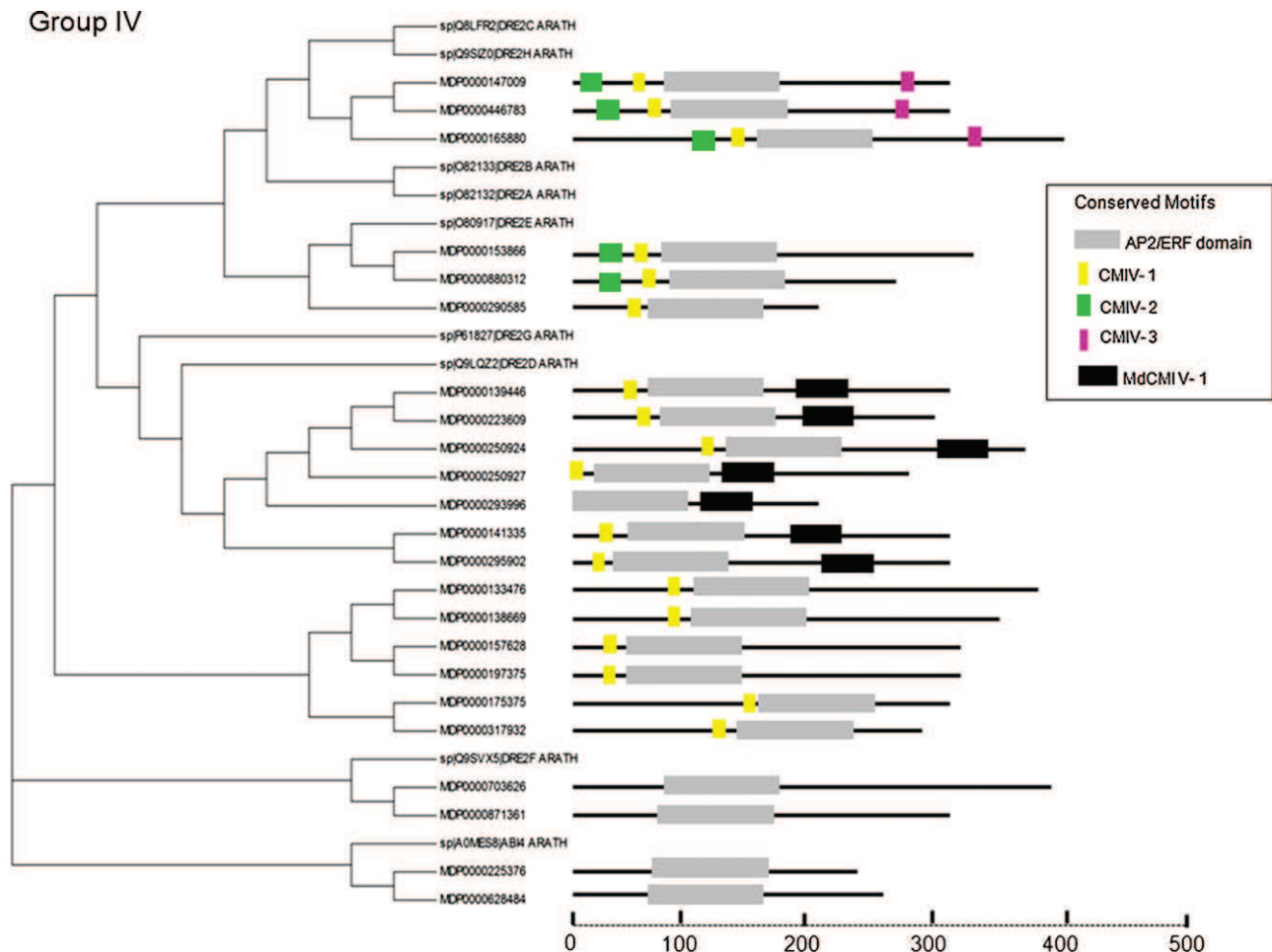


Fig. 3. Phylogenetic relationships among the *Arabidopsis* ERF genes, from group IV (D) in *Arabidopsis* and *Malus* ERF families. Bootstrap values from 1000 replicates are shown above the branches. The phylogenetic tree and a schematic diagram of the protein structures of groups I–X are shown as Supplementary Figs. S1–S10. Each colored box represents the AP2/ERF domain and conserved motifs, as indicated in legends below the trees. The amino acid sequences of the conserved motifs are summarized in Supplementary Table IV. Classification by Nakano et al. (2006) is shown above each tree and by Sakuma et al. (2002), indicated in parentheses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

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 in grapevine and are transcriptionally associated to berry ripening processes (Licausi et al., 2010). In poplar, ERF 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 $1 \cdot e^{-25}$) 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

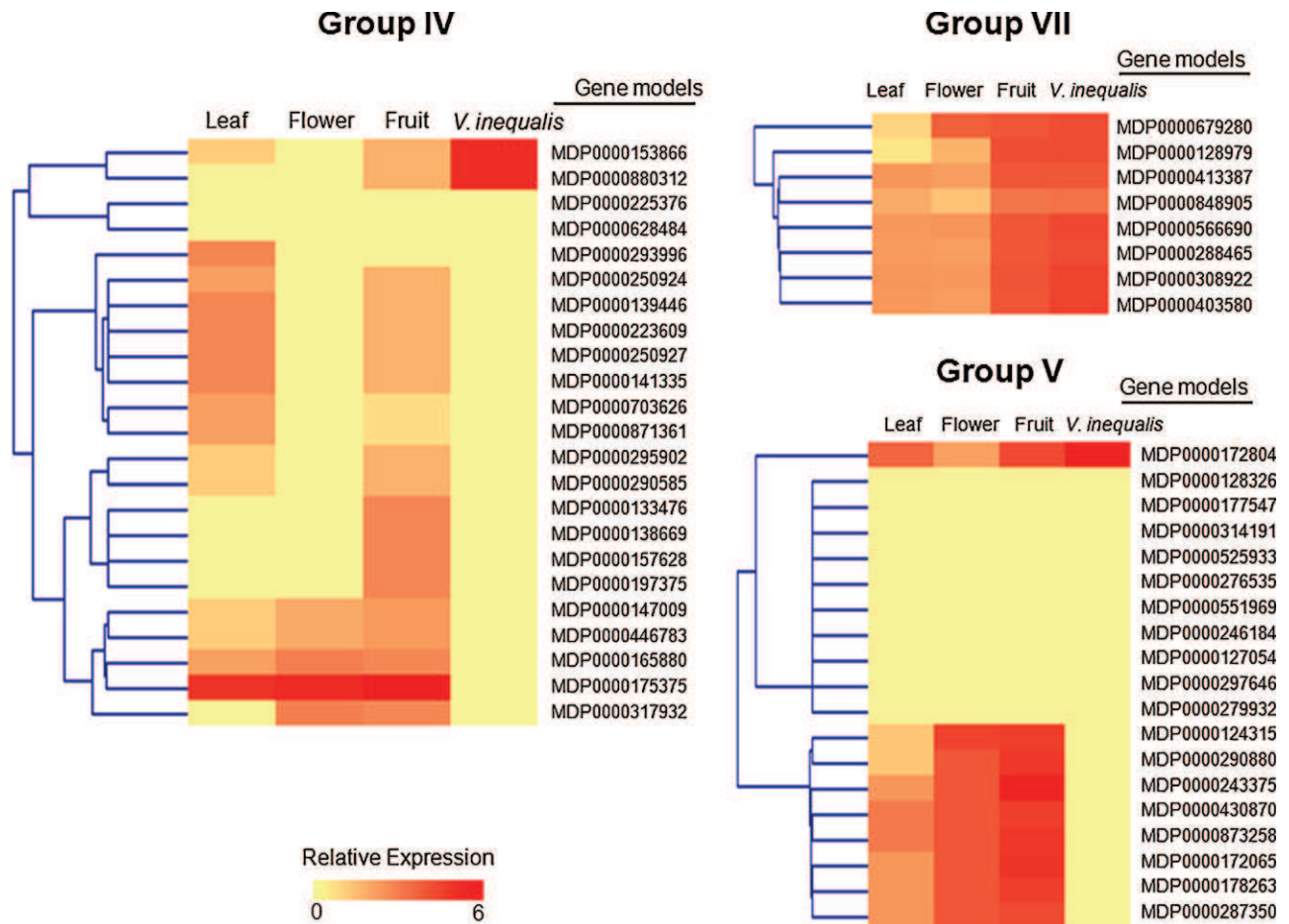


Fig. 4. Expression profile of ERF genes from groups IV, VII and V in leaf, flower, fruit and scab-infected tissues. Apple gene models are listed as available at Genome Database for Rosaceae (GDR). Hierarchical clustering is based on Spearman Rank correlation. Graphic outputs are represented as color scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

of the identified apple sequences sharing sequence similarity to ERFs were predicted to be unfolded (59.1%, 39) under physiological conditions (Table SIII). 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 SIII) (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, MDP0000276536, MDP0000551969, MDP0000246184, MDP0000127054, MDP0000297646 and MDP0000243375, MDP0000467753 and MDP0000778140 (group VIII) and group IX sequences MDP0000447570, MDP0000133854, MDP0000829925, MDP0000155543, MDP0000789227, MDP0000880063, MDP0000187369 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, VI L, VII, VIII and IX (Figs. 4 and 5). Transcripts corresponding to the sequences MDP0000127123 and MDP0000316843, 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. inaequalis* infected tissues (Figs. 4 and 5). Thus, we have further investigated the transcriptional profile of ripening-associated ERFs in scab pathogenesis using monospore *V. inaequalis* 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

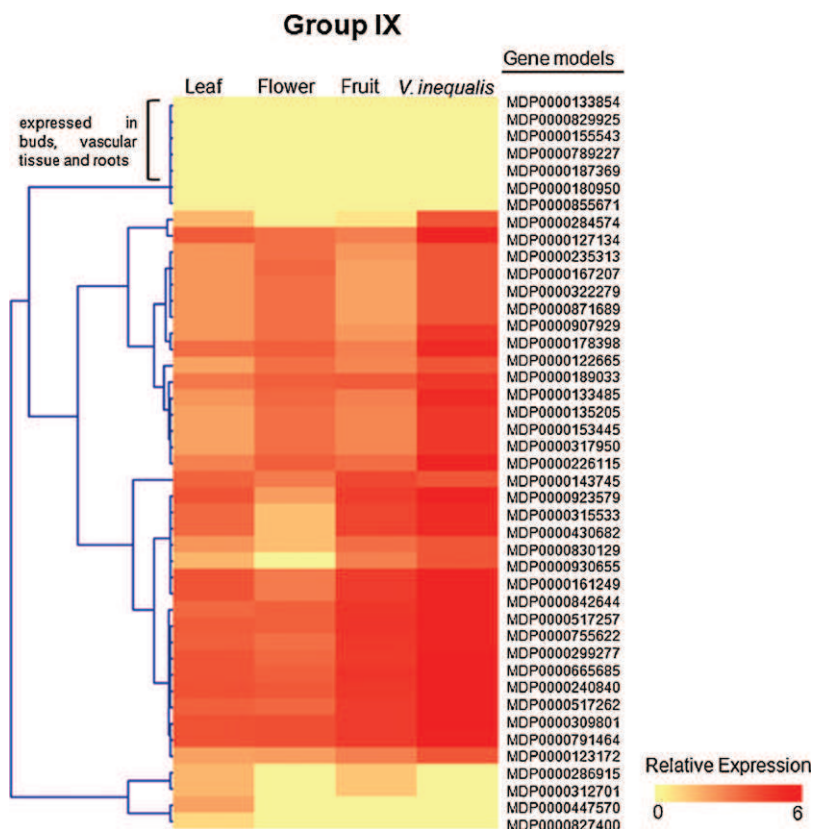


Fig. 5. Expression profile of *ERF* genes from group IX in leaf, flower, fruit and scab-infected tissues. Apple gene models are listed as available at Genome Database for Rosaceae (GDR). Hierarchical clustering is based on Spearman Rank correlation. Graphic outputs are represented as color scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

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 in 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 SIII) 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; Lijavetzky 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

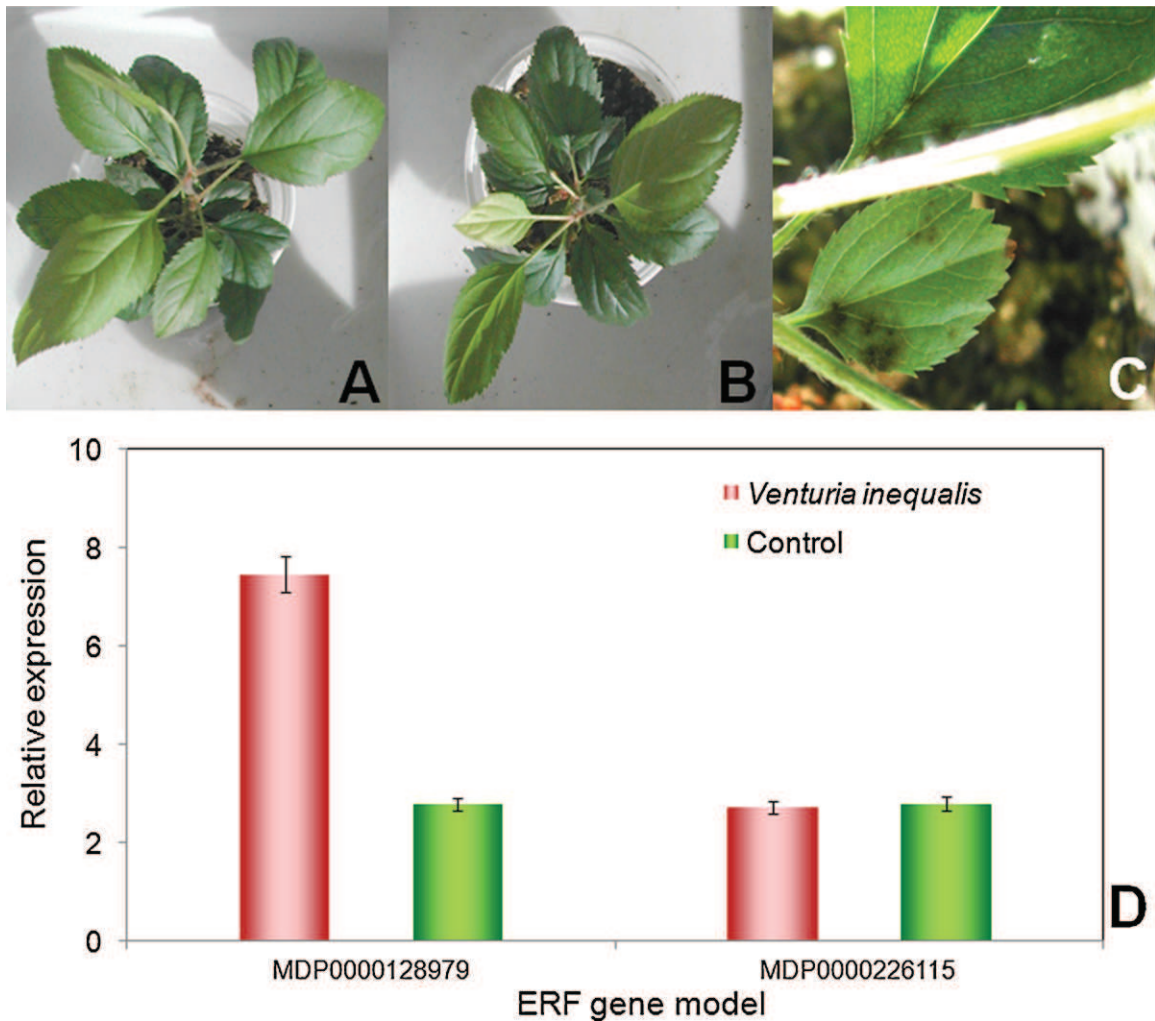


Fig. 6. Expression profile of apple ERF sequences from groups VII and IX in *Venturia inequalis*-infected vegetative tissue. (A) Mock-inoculated control; (B) plant sprayed with scab conidia; (C) scab symptoms on leaves and petioles 30 days after inoculation and (D) relative expression of MDP0000128979 (group VII) and MDP0000226115 (group IX) in fungus-infected and control plants. The bars represent percentage standard error.

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 pI 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 transcription

factors in developmental and stress-response processes. Traditionally, the expression of ERF genes has been associated with the molecular response to ethylene (Dietz et al., 2010). However, genome-wide transcriptional profiling of AP2/ERF genes in several plant species has demonstrated that they are regulated by a number of physical-chemical stimuli. A sub-set of ERF genes exhibit a classical ethylene responsive pattern and have been shown to be involved in the ripening process in climacteric fruits. Although several of the investigated ERF genes were highly expressed in fruits, only those from group Xb were exclusively found in fruits. In *Arabidopsis*, *Nicotiana tabacum* and *Solanum lycopersicum* transcripts of group II and III ERFs have been associated to abiotic stress responses (Gilmour et al., 1998; Tsutsui et al., 2009), whereas in *Vitis vinifera*, their expression is constitutively high, suggesting a role in a default defense system (Licausi et al., 2010). In apple, genes from these groups exhibit an expression pattern more closely related to stress responses, as observed in *Arabidopsis*, tobacco and tomato. The genes from ERF group IX, the larger group in apple, exhibit distinct regulation patterns, with a sub-set being highly induced in fruits and in scab infected tissues, thus indicating that the ripening and the *V. inequalis*-induced stress pathway share common transcriptional regulators but whether the targets of these transcription factors is also shared or independent remains to be elucidated. In a previous study to characterize ripening-associated genes in apple, the expression of MDP0000128979 (*MdERF1*) and MDP0000226115 (*MdERF2*), analyzed by northern blot and RT-PCR, was demonstrated to occur exclusively in fruit tissues (Wang et al., 2007).

Our results indicate that ERF genes are also expressed in vegetative tissue and that MDP0000128979 (*MdERF1*) transcription is induced in plants exhibiting scab symptoms, confirming the transcriptional association between the ripening factor and *V. inequalis* pathogenesis in apple. These results are in accordance to other reports demonstrating a role for ethylene in biotic stress responses in apple (Akagi et al., 2011) and other plant species (Gutterson and Reuber, 2004; Broekaert et al., 2006). In general, analyses of ERF expression patterns suggest some extent of functional specialization in apple.

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

- Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., van Arkel, G., Pereira, A., 2004. The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis*. *Plant Cell* 16, 2463–2480.
- Akagi, A., Dandekar, A., Stotz, H., 2011. Resistance of *Malus domestica* fruits to *Botrytis cinerea* depends on endogenous ethylene biosynthesis. *Phytopathology* 101, 1311–1321.
- Alba, R., Payton, P., Fei, Z., McQuinn, R., Debbie, P., Martin, G.B., Tanksley, S.D., Giovannoni, J.J., 2005. Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. *Plant Cell* 17, 2954–2965.
- Altschul, S.F., Madden, L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25, 3389–3402.
- Bailey, T.L., Elkan, C., 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In: Altman, R., Brutlog, D., Karp, P., Lathrop, R., Searls, D. (Eds.), *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*. American Association for Artificial Intelligence Press, Menlo Park, CA, pp. 28–36.
- Boss, P.K., Davies, C., Robinson, S.P., 1996. Analysis of the expression of anthocyanin pathway genes in developing *Vitis vinifera* L. cv. Shiraz grape berries and the implications for pathway regulation. *Plant Physiol.* 111, 1059–1066.
- Broekaert, W.F., Delauré, S.L., De Bolle, M.F., Cammue, B.P., 2006. The role of ethylene in host–pathogen interactions. *Ann. Rev. Phytopathol.* 44, 393–416.
- Brogliè, K.E., Biddle, P., Cressman, R., Brogliè, R., 1989. Functional analysis of DNA sequences responsible for ethylene regulation of a bean chitinase gene in transgenic tobacco. *Plant Cell* 1, 599–607.
- Chen, W., Provart, N.J., Glazebrook, J., Katagiri, F., Chang, H.S., Eulgem, T., Mauch, F., Luan, S., Zou, G., Whitham, S.A., Budworth, P.R., Tao, Y., Xie, Z., Chen, X., Lam, S., Kreps, J.A., Harper, J.F., Si-Ammour, A., Mauch-Mani, B., Heinlein, M., Kobayashi, K., Hohn, T., Dangl, J.L., Wang, X., Zhu, T., 2002. Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* 14, 559–574.
- da Silva, F.G., Iandolino, A., Al-Kayal, F., Bohlmann, M.C., Cushman, M.A., Lim, H., Ergul, A., Figueroa, R., Kabuloglu, E.K., Osborne, C., Rowe, J., Tattersall, E., Leslie, A., Xu, J., Baek, J., Cramer, G.R., Cushman, J.C., Cook, D.R., 2005. Characterizing the grape transcriptome. Analysis of expressed sequence tags from multiple *Vitis* species and development of a compendium of gene expression during berry development. *Plant Physiol.* 139, 574–597.
- Dietz, K.J., Vogel, M.O., Viehhauser, A., 2010. AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. *Protoplasma* 245, 3–14.
- Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D., 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863–14868.
- Eulgem, T., Rushton, P.J., Robatzek, S., Somssich, I.E., 2000. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5, 199–206.
- Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., Ohme-Takagi, M., 2000. *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell* 12, 393–404.
- Gilmour, S.J., Zarka, D.G., Stockinger, E.J., Salazar, M.P., Houghton, J.M., Thomashow, M.F., 1998. Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced gene expression. *Plant J.* 16, 433–442.
- Gutterson, N., Reuber, T.L., 2004. Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr. Opin. Plant Biol.* 7, 465–471.
- Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: protein localization predictor. *Nucl. Acids Res.* 35, W585–W587.
- Janssen, B.J., Thodey, K., Schaffer, R.J., Alba, R., Balakrishnan, L., Bishop, R., Bowen, J.H., Crowhurst, R.N., Gleave, A.P., Ledger, S., McArtney, S., Pichler, F.B., Snowden, K.C., Ward, S., 2008. Global gene expression analysis of apple fruit development from the floral bud to ripe fruit. *BMC Plant Biol.* 8, 16.
- Kendrick, M.D., Chang, C., 2008. Ethylene signaling: new levels of complexity and regulation. *Curr. Opin. Plant Biol.* 11, 479–485.
- Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R.C., Petroni, K., Urzainqui, A., Bevan, M., Martin, C., Smeekens, S., Tonelli, C., Paz-Ares, J., Weishaar, B., 1998. Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J.* 16, 263–276.
- Letunic, I., Doerks, T., Bork, P., 2008. SMART 6: recent updates and new developments. *Nucl. Acids Res.* 37, D229–D232.
- Licausi, F., Giorgi, F.M., Zenoni, S., Osti, F., Pezzotti, M., Perata, P., 2010. Genomic and transcriptomic analysis of the AP2/ERF superfamily in *Vitis vinifera*. *BMC Genom.* 11, 719.
- Lijavetzky, D., Carbonero, P., Vicente-Carbajosa, J., 2003. Genome-wide comparative phylogenetic analysis of the rice and *Arabidopsis* Dof gene families. *BMC Evol. Biol.* 3, 17.
- Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., Shinozaki, K., 1998. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10, 1391–1406.
- Miller, M., Shuman, J.D., Sebastian, T., Dauter, Z., Johnson, P.F., 2003. Structural basis for DNA recognition by the basic region leucine zipper transcription factor CCAAT/enhancer-binding protein alpha. *J. Biol. Chem.* 278, 15178–15184.
- Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497.
- Nakano, T., Suzuki, K., Fujimura, T., Shinshi, H., 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* 140, 411–432.
- Newcomb, R.D., Crowhurst, R.N., Gleave, A.P., Rikkerink, E.H.A., Allan, A.C., Beuning, L.L., Bowen, J.H., Gera, E., Jamieson, K.R., Janssen, B.J., Laing, W.A., McArtney, S., Nain, B., Ross, G.S., Snowden, K.C., Souleyre, E.J., Walton, E.F., Yauk, Y.K., 2006. Analyses of expressed sequence tags from apple. *Plant Physiol.* 141, 147–166.
- Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., Matsubara, K., Osato, N., Kawai, J., Carninci, P., Hayashizaki, Y., Suzuki, K., Kojima, K., Takahara, Y., Yamamoto, K., Kikuchi, S., 2003. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Res.* 10, 239–247.
- Panne, D., Maniatis, T., Harrison, S.C., 2004. Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer. *EMBO J.* 23, 4384–4393.
- Paterson, A.H., Freeling, M., Tang, H., Wang, X., 2010. Insights from the comparison of plant genome sequences. *Ann. Rev. Plant Biol.* 61, 349–372.

- Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E., Man, O., Beckmann, J.S., Silman, I., Sussman, J.L., 2005. FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 21, 3435–3438.
- Reyes, J.C., Muro-Pastor, M.I., Florencio, F.J., 2004. The GATA family of transcription factors in Arabidopsis and rice. *Plant Physiol.* 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., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K., Yu, G., 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110.
- Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K., Yamaguchi-Shinozaki, K., 2002. 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.* 290, 998–1009.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* 24, 4876–4882.
- Tian, C., Wan, P., Sun, S., Li, J., Chen, M., 2004. Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. *Plant Mol. Biol.* 54, 519–532.
- 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., 2009. DEAR1, a transcriptional repressor of DREB protein that mediates plant defense and freezing stress responses in Arabidopsis. *J. Plant Res.* 122, 633–643.
- Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S.K., Troglio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavaiuolo, M., Coppola, G., Costa, F., Cova, V., Dal Ri, A., Goremykin, V., Komjanc, M., Longhi, S., Magnago, P., Malacarne, G., Malnoy, M., Micheletti, D., Moretto, M., Perazzolli, M., Si-Ammour, A., Vezzulli, S., Zini, E., Eldredge, G., Fitzgerald, L.M., Gutin, N., Lanchbury, J., Macalma, T., Mitchell, J.T., Reid, J., Wardell, B., Kodira, C., Chen, Z., Desany, B., Niazi, F., Palmer, M., Koepke, T., Jiwan, D., Schaeffer, S., Krishnan, V., Wu, C., Chu, V.T., King, S.T., Vick, J., Tao, Q., Mraz, A., Stormo, A., Stormo, K., Bogden, R., Ederle, D., Stella, A., Vecchiotti, A., Kater, M.M., Masiero, S., Lasserre, P., Lespinasse, Y., Allan, A.C., Bus, V., Chagné, D., Crowhurst, R.N., Gleave, A.P., Lavezzo, E., Fawcett, J.A., Proost, S., Rouzé, P., Sterck, L., Toppo, S., Lazzari, B., Hellens, R.P., Durel, C.E., Gutin, A., Bumgarner, R.E., Gardiner, S.E., Skolnick, M., Egholm, M., van de Peer, Y., Salamini, F., Viola, R., 2010. The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat. Genet.* 42, 833–839.
- Wang, A., Tan, D., Takahashi, A., Li, T.Z., Harada, T., 2007. MdERFs, two ethylene-response factors involved in apple fruit ripening. *J. Exp. Bot.* 58, 3743–3748.
- Weigel, D., 1995. The APETALA2 domain is related to a novel type of DNA binding domain. *Plant Cell* 7, 388–389.
- Xu, Z.S., Chen, M., Li, L.C., Ma, Y.Z., 2011. Functions and application of the AP2/ERF transcription factor family in crop improvement. *J. Integr. Plant Biol.* 53, 570–585.
- Zhang, G., Chen, M., Chen, X., Xu, Z., Guan, S., Li, L.-C., Li, A., Guo, J., Mao, L., Ma, Y., 2008. Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). *J. Exp. Bot.* 9, 4095–4107.
- 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., 2008. Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem. Biophys. Res. Commun.* 371, 468–474.
- Ziliotto, F., Begheldo, M., Rasori, A., Bonghi, C., Tonutti, P., 2008. Transcriptome profiling of ripening nectarine (*Prunus persica* L. Batsch) fruit treated with 1-MCP. *J. Exp. Bot.* 59, 2781–2791.