High AN1 variability and interaction with basic helix-loop-helix co-factors related to anthocyanin biosynthesis in potato leaves

Vincenzo D’Amelia1, Riccardo Aversano1, Giorgia Batelli2, Immacolata Caruso1, Mar Castellano Moreno3, Ana Beatriz Castro-Sanz2, Pasquale Chiaiese1, Carlo Fasano1, Francesca Palomba1 and Domenico Carputo1,*

1Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, Italy, 2CNR-IBBR Institute of Biosciences and Bioresources, Via Università’ 133, 80055 Portici, Italy, and 3Centro de Biotecnología y Genómica de Plantas, INIA-UPM, Campus de Montegancedo, 28223 Madrid, Spain

SUMMARY

AN1 is a regulatory gene that promotes anthocyanin biosynthesis in potato tubers and encodes a R2R3 MYB transcription factor. However, no clear evidence implicates AN1 in anthocyanin production in leaves, where these pigments might enhance environmental stress tolerance. In our study we found that AN1 displays intraspecific sequence variability in both coding/non-coding regions and in the promoter, and that its expression is associated with high anthocyanin content in leaves of commercial potatoes. Expression analysis provided evidence that leaf pigmentation is associated to AN1 expression and that StJAF13 acts as putative AN1 co-regulator for anthocyanin gene expression in leaves of the red leaf variety 'Magenta Love,' while a concomitant expression of StbHLH1 may contribute to anthocyanin accumulation in leaves of 'Double Fun.' Yeast two-hybrid experiments confirmed that AN1 interacts with StbHLH1 and StJAF13 and the latter interaction was verified and localized in the cell nucleus by bimolecular fluorescence complementation assays. In addition, transgenic tobacco (Nicotiana tabacum) overexpressing a combination of either AN1 with StJAF13 or AN1 with StbHLH1 showed deeper purple pigmentation with respect to AN1 alone. This further confirmed AN1/StJAF13 and AN1/StbHLH1 interactions. Our findings demonstrate that the classical loci identified for potato leaf anthocyanin accumulation correspond to AN1 and may represent an important step to expand our knowledge on the molecular mechanisms underlying anthocyanin biosynthesis in different plant tissues.

Keywords: Solanum tuberosum, MYB transcription factor, flavonoids, regulatory complex, allelic genotyping, bimolecular fluorescence complementation, yeast two hybrid assay, transient and stable transformation, Nicotiana tabacum.

INTRODUCTION

Three major classes of molecules confer colour to plants: anthocyanins, carotenoids and chlorophylls (Tanaka et al., 2008). Among them anthocyanins are of particular interest for their well documented beneficial effects on plant physiological processes and human health (Stintzing and Carle, 2004; De Pascual-Teresa and Sanchez-Ballesta, 2008). Anthocyanin biosynthesis is primarily controlled through regulation of genes encoding the structural enzymes of the phenylpropanoid metabolic pathway (Hichri et al., 2011a). Expression of the structural genes is tightly controlled by the ternary complex ‘MBW’ (Patra et al., 2013). This complex is composed of MYB and basic helix-loop-helix (bHLH) transcription factors, together with WD40 repeat proteins that regulate flavonoid spatiotemporal production in conjunction with the promoters of structural genes (Lin-Wang et al., 2010; Feller et al., 2011). Plant MYBs are a large gene family whose members have many different functions and represent key factors activating specific downstream genes (Takos et al., 2006). The subfamily R2R3 MYB is the largest group present in higher plants and the most extensively studied. It possesses a structurally conserved DNA-binding domain consisting of up to two imperfect repeats, R2 and R3 (Jin and Martin, 1999). The R3 repeat provides a platform for protein–protein interaction, especially with the bHLH co-factor (Grotewold et al., 2000). In this complex the affinity between MYB and
the cis-element of the target gene may be partly influenced by the bHLH partners (Hichri et al., 2011b). R2R3 MYBs play important roles in tissue-specific anthocyanin accumulation in many plants (Gao et al., 2013). They include AN2 in petunia (Petunia × hybrida), ROSEA1, ROSEA2, and VENOSA in snapdragon (Antirrhinum majus), C1 and P1 in maize (Zea mays) and PAP1 in Arabidopsis (Grotewold et al., 1991; Sainz et al., 1997; Quattrocchio et al., 1999; Borevitz et al., 2000; Schwinn et al., 2006). The effects of bHLH co-factors in different tissues of the same species remain unclear. Recently, there has been increasing interest in understanding the molecular mechanisms regulating phenylpropanoid production in potato (Solanum tuberosum) tubers. It has been reported that the production of red and purple anthocyanins is controlled by R and P loci, while D (the developer locus) is responsible for tissue-specific anthocyanin accumulation in tuber skin (Jung et al., 2009). The D, R and P loci have been mapped to chromosomes 10, 2 and 11, respectively (Jung et al., 2009). Furthermore, their structural and regulatory function in the anthocyanin biosynthetic pathway has been elucidated. It has been reported that R encodes a dihydroflavonol 4-reductase (DFR), P a flavonoid 3',5'-hydroxylase (F3'5'H). The D locus cosegregates with an ortholog of petunia AN2, an R2R3 MYB transcription factor (De Jong et al., 2004; Jung et al., 2009; Zhang et al., 2009a,b). Jung et al. (2009) designated this gene AN1 and identified two different allelic forms: AN1-777 and AN1-816. While our understanding of AN1’s involvement in anthocyanin regulation in tubers has expanded, the role in leaves has received scant attention. Here production of anthocyanins could be important for plant defence mechanisms (Gould, 2004). In potato the additional advantage of a high leaf anthocyanin content may be related to a protective role against herbivorous insects (Schaef er and Rolshausen, 2005). In the case of aphids, this would have important implications also for virus spread. In leaves, it has been shown that the potato loci conditioning anthocyanin accumulation are tightly linked to each other and to locus D (De Jong, 1991; Jung et al., 2009). Further studies showed that the constitutive expression of AN1 or StMFT1 (another potato MYB gene) causes increased accumulation of anthocyanins in foliage (Rommens et al., 2008; Jung et al., 2009). Recently, Payyavula et al. (2013) showed that AN1 expression was inducible by a sucrose treatment in plantlets of the cultivar Purple Majesty, suggesting that environmental conditions affect AN1 transcript abundance in vegetative tissues.

The aim of our study was to identify the main factors regulating anthocyanin gene expression in potato leaves. Target genes believed to be involved in tuber pigmentation were selected for genetic and functional analysis. Sequence analysis of different potato genotypes showed an extensive intraspecific nucleotide sequence variation of AN1, both in the predicted promoter and coding sequence, where potential protein polymorphisms were identified. Expression analysis suggested that leaf pigmentation is associated with AN1 expression and that StJAF13 (previously named StbHLH2; Payyavula et al., 2013) acts as putative AN1 co-regulator for anthocyanin gene expression in leaves of the red leaf variety ‘Magenta Love,’ while a concomitant expression of StbHLH1 in ‘Double Fun’ may also contribute to anthocyanin accumulation in leaves of this cultivar. Protein interaction of AN1 with both StbHLH1 and StJAF13 was detected using yeast two-hybrid and, in the case of StJAF13, further confirmed using bimolecular fluorescence complementation (BiFC) assays. Stable co-transformation of AN1 and either StbHLH1 or StJAF13 in tobacco (Nicotiana tabacum) produced a stronger pigmentation with respect to single AN1 overexpression. These findings indicate that in potato leaves StJAF13 enhances AN1 activity in anthocyanin production.

RESULTS
High AN1 gene nucleotide variability and amino acid polymorphic sites

High variability was found in the AN1 genomic sequence of 17 potato varieties and several variants were identified (Table 1). In the coding sequence (CDS), the frequency of all sequence variants per bp was 7%, with 58 sequence variants identified. Exon 3 presented the highest number of sequence variants (38), 30 of which were due to SNPs (single nucleotide polymorphisms) and eight to indels. Analysis of exons 1 and 2 revealed a total of 8 and 12 sequence variants, respectively, all attributable to SNPs. Exon 2 displayed a frequency of polymorphic sites per bp (about 9%) higher than that of exons 1 and 3 (about 6%). As regards the analysis of the nucleotide coding for domains (NCD), R2 and R3 exhibited a similar number of sequence variants (10 and 11, respectively). All the variants were due to SNPs, with only one indel detected in R3. The variable region (VR) showed 37 sequence variants (30 SNPs and seven indels), with an average indel length of 10 bp. In the intronic region 72 sequence variants and the highest frequency of polymorphic sites per bp (15%) were found. In particular, intron 1 displayed 25 sequence variants, comprising 20 SNPs and five indels (on average 7.6 bp long), while intron 2 showed 39 SNPs and eight indels (on average 3.1 bp long) and a frequency of polymorphic sites per bp lower than that of intron 1 (14% versus 19%).

To investigate the variants found in CDS, in silico analyses were carried out. Sequence alignments between AN1 fragments of each potato genotype and the AN1 GenBank reference sequences (AN1-777 and AN1-816) provided evidence that indels within exon 3 were due to a deletion of 39 bp in AN1-777 allele, confirming a previous report (Jung et al., 2009). In the same site we found deletions in...
Table 1: Summary of sequence variants obtained analysing all fragments amplified from potato genotypes and using the alleles AN1-777 and AN1-816

<table>
<thead>
<tr>
<th>Exon 1</th>
<th>Exon 2</th>
<th>Exon 3</th>
<th>Total</th>
<th>R2</th>
<th>R3</th>
<th>VR</th>
<th>Intron 1</th>
<th>Intron 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of sequences (bp)</td>
<td>123</td>
<td>130</td>
<td>550</td>
<td>803</td>
<td>162</td>
<td>150</td>
<td>471</td>
<td>103</td>
<td>279</td>
</tr>
<tr>
<td>Frequency of all sequence variants (%)</td>
<td>6.5</td>
<td>9.2</td>
<td>6.9</td>
<td>7.2</td>
<td>6.2</td>
<td>7.3</td>
<td>7.9</td>
<td>24.3</td>
<td>16.8</td>
</tr>
<tr>
<td>All sequence variants (no.)</td>
<td>8</td>
<td>12</td>
<td>38</td>
<td>58</td>
<td>10</td>
<td>11</td>
<td>37</td>
<td>25</td>
<td>47</td>
</tr>
<tr>
<td>Frequency of polymorphic sites per bp (%)</td>
<td>6.5</td>
<td>9.2</td>
<td>5.5</td>
<td>6.2</td>
<td>6.2</td>
<td>6.7</td>
<td>6.4</td>
<td>19.4</td>
<td>14.0</td>
</tr>
<tr>
<td>Nucleotide substitutions per site (no.)</td>
<td>8</td>
<td>12</td>
<td>30</td>
<td>50</td>
<td>10</td>
<td>10</td>
<td>30</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td>Indels (no.)</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Average indel length (bp)</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>1</td>
<td>0</td>
<td>0.7</td>
<td>1.5</td>
<td>4.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Frequency of indels per bp (%)</td>
<td>6.5</td>
<td>9.2</td>
<td>6.9</td>
<td>7.2</td>
<td>6.2</td>
<td>7.3</td>
<td>7.9</td>
<td>24.3</td>
<td>16.8</td>
</tr>
</tbody>
</table>

*SNPs and indels.

NCD, nucleotide coding for domains; VR, variable region.

'Spunta' and 'Assergi' (Figure S1) consisting of 48 and 52 nucleotides, respectively. The latter overlapped the same region of a 39 bp deletion already described for allele AN1-777, but located 16 bp downstream. Table 2 summarises potential missense mutations found with respect to AN1-777 and AN1-816 sequences. Some of them were genotype-specific: in 'Adora', for example, isoleucin was substituted with threonine due to a single nucleotide mutation (T to C) at position 97 and isoleucin with valine as a consequence of a single mutation from A to G. 'Silvy', 'Double Fun,' 'Magenta Love' and 'Flamenco' displayed a common mutation at position 949 (C to G), producing the substitution of isoleucine with methionine. Within exon 3 we found three additional mutations causing amino acid substitutions on predicted protein sequence. In particular, two of them (at nucleotide positions 1095 and 1166) resulted in a substitution from polar amino acids (threonine and serine, respectively) to the apolar alanine. Figure 1 displays the consequences of the amino acid substitutions previously presented and their location on the R2R3 MYB predicted protein. The double consecutive substitution (GG to CA) on exon 2 found in 'Magenta Love' and 'Silvy' altered the conservative regularly spaced tryptophan repetition in the helix turn helix structure. In fact, this apolar residue was replaced with the amino acid proline. Amino acid motif analysis showed that the deletion of 48 nucleotides found in the first part of exon 3 in 'Spunta' caused a loss of 16 amino acids in the last part of R3 domain. This means that 'Spunta' has an allele without the conservative protein motif ANDV described by Lin-Wang et al. (2010). This identifier motif was found in all the other fragments analysed as well as in the annotated alleles.

Table 2: Summary of missense mutations found with respect to AN1-777 and AN1-816 consensus sequence

<table>
<thead>
<tr>
<th>Exon</th>
<th>Nucleotide positions</th>
<th>Missense mutations</th>
<th>Amino acid substitutions</th>
<th>Genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>71</td>
<td>T-&gt;C</td>
<td>I-&gt;T</td>
<td>Adora</td>
</tr>
<tr>
<td>88</td>
<td>A-&gt;G</td>
<td>I-&gt;V</td>
<td></td>
<td>Adora</td>
</tr>
<tr>
<td>2</td>
<td>323-326</td>
<td>TTGG-&gt;ACCA</td>
<td>DW-&gt;EP</td>
<td>Silvy</td>
</tr>
<tr>
<td>3</td>
<td>323-326</td>
<td>TTGG-&gt;ACCA</td>
<td>DW-&gt;EP</td>
<td>Magenta Love</td>
</tr>
<tr>
<td>858</td>
<td>G&gt;A</td>
<td>C-&gt;Y</td>
<td></td>
<td>Adora</td>
</tr>
<tr>
<td>949</td>
<td>C-&gt;G</td>
<td>I-&gt;M</td>
<td></td>
<td>Silvy</td>
</tr>
<tr>
<td>949</td>
<td>C-&gt;G</td>
<td>I-&gt;M</td>
<td>Double Fun</td>
<td>Magenta Love</td>
</tr>
<tr>
<td>949</td>
<td>C-&gt;G</td>
<td>I-&gt;M</td>
<td>Magenta Love</td>
<td></td>
</tr>
<tr>
<td>949</td>
<td>C-&gt;G</td>
<td>I-&gt;M</td>
<td>Flamenco</td>
<td></td>
</tr>
<tr>
<td>1095</td>
<td>A-&gt;G</td>
<td>T-&gt;A</td>
<td></td>
<td>Spunta</td>
</tr>
<tr>
<td>1095</td>
<td>A-&gt;G</td>
<td>T-&gt;A</td>
<td>Magneta Love</td>
<td></td>
</tr>
<tr>
<td>1114</td>
<td>T-&gt;G</td>
<td>D-&gt;E</td>
<td>Magneta Love</td>
<td></td>
</tr>
<tr>
<td>1114</td>
<td>T-&gt;G</td>
<td>D-&gt;E</td>
<td>Biosa</td>
<td></td>
</tr>
<tr>
<td>1114</td>
<td>T-&gt;G</td>
<td>D-&gt;E</td>
<td>Double Fun</td>
<td></td>
</tr>
<tr>
<td>1114</td>
<td>T-&gt;G</td>
<td>D-&gt;E</td>
<td>Spunta</td>
<td></td>
</tr>
<tr>
<td>1116</td>
<td>T-&gt;G</td>
<td>S-&gt;A</td>
<td>Double Fun</td>
<td></td>
</tr>
<tr>
<td>1116</td>
<td>T-&gt;G</td>
<td>S-&gt;A</td>
<td>Silvy</td>
<td></td>
</tr>
<tr>
<td>1166</td>
<td>T-&gt;G</td>
<td>S-&gt;A</td>
<td>Magneta Love</td>
<td></td>
</tr>
</tbody>
</table>
additional potato MYBs grouped with the tomato SIAN2. In particular, MYBA1 shared a similarity of 90% with AN1 cloned from 'Magenta Love'. MYBA1 corresponds to the translated sequence of StAN3 indicated as a possible AN1 pseudogene (Jung et al., 2009).

Regulation of AN1 gene expression is fundamental towards anthocyanin accumulation (Payyavula et al., 2013). Sequencing of the 5’ upstream region was therefore carried out to identify potential polymorphisms between red and green leaf varieties. Sequence analysis revealed the presence in the candidate AN1 promoter of a simple motif repeat (TA)36 at 1/969 bp from the ATG. BLASTn analysis of sequence immediately surrounding the TA motif showed an identity of 95–96% with the retrotransposons of the potato SINE (short interspersed elements) family named SolSV (Wenke et al., 2011). A region between nucleotides 2077 and 1895 of the AN1 promoter was aligned with the SINE SolSV_St3. The TA motif showed high similarity with the tail of the putative SINE retrotransposons (Figure S1). Through sequence analysis we found that in 'Silvy' there were 8 TA repeats, as in SolS-V_St3. By contrast, in red leaf 'Magenta Love' and 'Double Fun' the number of TA repeats was 16.

Isolation of StJAF13 coding sequence

Two potato bHLH genes (StJAF13 and StbHLH1) were identified on Potato Genome Sequencing Consortium (PGSC) database based on homology with other species (Payyavula et al., 2013). These genes are putatively involved in regulation of anthocyanin biosynthesis in potato. The potato StJAF13 sequence was blasted onto the NCBI database. We found an identity of 97% with S. lycopersicum GLABRA3-like, 90% with N. tabacum JAF13-like a and b and 87% with that of P. x hybrida PhJAF13. Transcript alignments showed that annotated potato StJAF13 CDS (1083 bp) was 800 bp shorter than its homolog (1883 bp). To investigate the real dissimilarity of the coding sequence length between potato and other species, StJAF13 CDS sequence was cloned from complementary DNA (cDNA) of 'Magenta Love' beginning from the common start codon annotated for petunia PhJAF13 and tomato GLABRA3-like CDS. The obtained StJAF13 sequence was aligned against PGSC JAF13 transcript and genomic sequences. We found that the CDS of StJAF13 presented a deletion of 381 nucleotides with respect to the annotated sequence. This caused the elimination of a premature stop codon and the production of
of a CDS of 1,881 bp. Therefore, a revised genomic structure for potato StJAF13 is proposed, as reported in Figure S3. No alternative splicing was found in the flanking regions of the deletion. In all genotypes a single amplification product of about 181 bp was amplified from their respective cDNAs using primers flanking the deletion. No products of 562 bp (indicating the presence of a 381 bp insertion) were found (Figure S4). Phylogenetic analysis showed that the StJAF13 translated sequence is grouped with tomato GLABRA3-like, tobacco JAF13-like a and b and petunia PhJAF13 (Figure S5). By contrast, in the other branch of the same subclade we found many putative anthocyanin regulators of the genus *Ipomoea*. The same relation with genus *Ipomoea* was also found for the bHLH1 translated sequence (Figure S6). StJAF13 sequence was aligned with phylogenetically related proteins (Figure 2): GLABRA3-like of tomato, PhJAF13 of petunia, JAF13-like b of tobacco, MYC-RP of *Perilla frutescens*, DELILA2 of *A. majus* and bHLH of *I. nil*. The amino terminal part, comprising approximately the first 200 amino acids is involved in the interaction with MYB partners (Hichri *et al.*, 2011a). Our data showed that this region is highly conserved among the seven different species. JAF13 predicted protein annotated on PGSC database lacks this portion. The region located between MIR (MYB-interacting region) and bHLH domain was less conserved among the sequences aligned, except for tomato GLABRA3-like and StJAF13, which shared similar sequences. In addition, GLABRA3-like and StJAF13 showed an identical bHLH domain.

**High AN1 and StJAF13 expression in red leaf genotypes**

Anthocyanin content was analysed to determine the amount of these pigments in leaves of potato genotypes. Red leaf ‘Double Fun’ and ‘Magenta Love’ showed the highest total anthocyanin content (17 mg g⁻¹ and 13 mg g⁻¹ FW, respectively) (Figure 3). On the same genotypes expression analysis of anthocyanin structural and regulatory genes was carried out. Absolute quantification of dihydroflavonol 4-reductase (*DFR*) and anthocyanin synthase (*ANS*) genes expression is reported in Figure 4(a). Compared with the other genotypes, ‘Double Fun’ and ‘Magenta Love’ showed a significantly higher expression of the two genes. The former genotype displayed the highest value for *DFR* and *ANS*: about 1750 copies/µl and 238 copies/µl, respectively. In ‘Magenta Love’ the number of copies/µl detected was about 612 for *DFR* and 15 for *ANS*. In all the other genotypes, expression of both genes was close to 0 copies/µl. In ‘Violet Queen,’ ‘Flamenco’ and ‘Blue Star’ the expression of *DFR* was between 20 and 100 copies/µl. A similar trend of gene expression was observed for AN1 and StJAF13 genes (Figure 4b). As for AN1, about 3000 copies/µl and 1200 copies/µl were detected in ‘Double Fun’ and ‘Magenta Love,’ respectively. As for StJAF13, 48 copies/µl were found in ‘Double Fun’ and 17 in ‘Magenta Love.’ Different results were found for StbHLH1 (Figure 4b). Indeed, no expression was detected in red leaf ‘Magenta Love,’ whereas estimated copies/µl in ‘Double Fun’ was 57.

**AN1 interacts with StbHLH1 and StJAF13**

AN1 and StJAF13 showed largely overlapping expression profiles, with higher levels detected in genotypes with red leaves. It is therefore conceivable that StJAF13 could be an AN1 interacting partner in leaves. The interaction between AN1 and StJAF13 was therefore tested in the yeast two-hybrid system (Figure 5b). StbHLH1, previously shown to be strongly expressed in red- and purple-tuber varieties (Payyavula *et al.*, 2013) was also tested. The coding sequences of StJAF13 and StbHLH1 were fused with the binding domain of GAL4 and transformed into yeast strain AH109 together with the prey AN1 fused to GAL4 activation domain. Yeast growth on selective media lacking histidine and adenine was observed, indicative of an interaction between AN1 and both StJAF13 and StbHLH1. No interaction was detected when the truncated versions of StJAF13 and StbHLH1, amplified using primers on the predicted translation initiation site as indicated in PGSC, were used (StJAF13Δ264, StbHLH1Δ58). This provided evidence that amino acids 1–264 of StJAF13 and 1–58 of StbHLH1 are required for the interaction with AN1 in yeast. To confirm the results obtained with the yeast two-hybrid, BiFC studies were performed. The protein AN1 fused upstream of the N-term fragment of YFP and StJAF13 fused to the C-terminal fragment of YFP were expressed by agroinfiltration in *N. benthamiana* leaves in combination or along with the respective controls. As shown in Figure 5(c), fluorescence corresponding to reconstituted YFP was observed in the nucleus of co-transformed cells with the AN1 and StJAF13 constructs. No fluorescence was detected when these proteins were expressed along with only the N-term or the C-term of the YFP moieties nor when StJAF13Δ264 was used in combination with AN1. Together, our results confirmed that AN1 interacts with StbHLH1 and StJAF13 and that amino acids at the N-term of bHLHs might be required for the interaction to take place.

**Co-expression of AN1, StbHLH1 and StJAF13 genes enhances anthocyanin accumulation in tobacco plants**

To validate the interaction of AN1 with StJAF13 and StbHLH1 and to assess the impact of this interaction on anthocyanin biosynthesis, both transient and stable nuclear genetic transformations were performed. Transient expression affected tissue pigmentation in cells agroinfiltrated with the expression vectors p35s::AN1, and p35s::AN1 plus p35s::StJAF13 or p35s::StbHLH1. Anthocyanin accumulation was visibly enhanced in leaves co-infiltrated with AN1 and StJAF13 (Figure 6a). Stable transgenic
tobacco plants overexpressing AN1, bHLHs or a combination of AN1 and bHLHs were generated. Transgenic plants showed differential phenotypes depending on the transgenes inserted (Figure 6b). Overexpression of either StJAF13 or StbHLH1 alone produced shoots with no pigmentation variation, while the constitutive expression of...
AN1 positively affected the pigmentation (Figure 6c). Furthermore, co-expression of AN1 with both bHLHs in tobacco cells produced shoots with an enhanced pigmentation. Measurements of anthocyanin levels of transgenic plants demonstrated that AN1 overexpression increased the anthocyanin content in shoots and that this effect is potentiated by the co-expression with StJAF13 and StbHLH1 (Figure 6d). These assays confirmed the important role of AN1 along with StJAF13 and StbHLH1 in regulating leaf anthocyanin pathway.

**DISCUSSION**

The cultivated potato *S. tuberosum* is the fourth most important crop worldwide. It is also classified as sensitive to environmental stresses, which affect tuber yield and quality. The presence of high levels of anthocyanins in tissues exposed to stress conditions could be an important advantage for plant resistance: in leaves these pigments can act as UV-B filters, protect DNA from oxidative damage, increase resistance to pathogens thanks to their antimicrobial activity (Solfanelli *et al.*, 2006; Van Oosten *et al.*, 2013). In addition, since insects may show preferences for green leaves for food or oviposition and do not possess red colour receptors, leaf anthocyanins participate in defence mechanisms against herbivores (Schaefer and Rolshausen, 2005; Karageorgou and Manetas, 2006; Chittka and Doring, 2007). Lovdal *et al.* (2010) pointed out that a high level of anthocyanins and flavonoids in plants may reduce the need for pesticide treatments, and hence can be an interesting target for plant breeding. So far in potato

---

*Figure 4. Gene expression analysis in leaves of potato genotypes as monitored by absolute qRT-PCR. (a) Expression of DFR and ANS anthocyanin structural genes. (b) Expression of AN1, StJAF13 and StbHLH1 transcription factors. Each value represents the mean of three determinations (± SD). Means denoted by the same letter did not differ significantly at *P* ≤ 0.05 according to Duncan’s multiple range test. Genotypes denoted by * showed no detectable levels of expression.*
only the molecular mechanisms and genes that control anthocyanin accumulation or biosynthesis in the tubers have received much attention (Jung et al., 2009; Zhang et al., 2009a,b; Stushnoff et al., 2010; Payyavula et al., 2013; Tai et al., 2013). All these works emphasised the role of the AN1 gene, coding for a R2R3 MYB, in controlling the
expression of structural genes involved in the anthocyanin pathway, especially in the tuber periderm (Jung et al., 2009). As there are few studies on other potato tissues, the research we carried out focused on the leaves. We analysed the AN1 genomic sequence in a set of potato genotypes displaying different leaf pigmentations. Indeed, as reported by Schwinn et al. (2006), striking effects on the phenotype could be caused by small changes in MYB sequence. We found a high variability in both coding and non-coding sequences of AN1. Our results were consistent with the AN1 sequence variability already reported by Jung et al. (2009) but also allowed us to explore its allelic

![Image](image_url)

Figure 5. AN1 interacts with StJAF13 and StbHLH1.
(a) Diagrams of the proteins used for the yeast two-hybrid and BiFC assay. R2R3, MYB repeat domains; MIR, MYB-interacting region; bHLH, basic helix-loop-helix domain.
(b, c) AN1 interacts with StJAF13 (b) and StbHLH1 (c) in a yeast two-hybrid assay. StJAF13 and StbHLH1 were cloned in the bait plasmid pGBK7 and co-transformed in yeast with the prey AN1. The pGBK7/AN1pGADT7 and pGBK7bHLHpGADT7 combinations were used as negative controls. Truncated fragments of StJAF13 and StbHLH1 lacking 264 and 58 amino acids at the N-term respectively (StJAF13264 and StbHLH1588), did not interact with AN1. Yeast cells grown for three days on synthetic complete media lacking tryptophan and leucine (–W/–L) and on selective media lacking tryptophan, leucine, histidine and adenine (–W/–L/–H/–A) are shown.
(d) Analysis of the AN1 and StJAF13 interaction by BiFC in N. benthamiana leaves. Constructs expressing different combinations of AN1 fused to N-terminal and StJAF13 fused to C-terminal YFP fragments were agroinfiltrated together or along with control constructs. DAPI staining, reconstituted YFP fluorescence and merge images for a representative field for each combination are shown.

Figure 6. The effects of overexpression of AN1, StbHLH1 and StJAF13 in Nicotiana spp.
(a) Leaves of N. benthamiana after agroinfiltration with pGWB411 empty vector, AN1, StJAF13, StbHLH1 or a combination of AN1 with StJAF13 or StbHLH1.
(b) Stable transgenic plants transformed with pGWB411 empty vector, AN1, StJAF13, StbHLH1 or a combination of AN1 with StJAF13 or StbHLH1.
(c) Roots of transgenic plants overexpressing AN1.
(d) Total anthocyanin content (mg g–1 FW) in transformed tobacco plants. Each value represents the mean of three determinations (±SD). Means denoted by the same letter did not differ significantly at P ≤ 0.05 according to Duncan’s multiple range test. Genotypes denoted by * showed no detectable level of anthocyanin content.

© 2014 The Authors
diversity. In particular, a number of indels in the second intron and the third exon were found. These indels could explain results of Jung et al. (2009), who found bands of different sizes when they amplified a portion of these regions from different tetraploid potatoes. Most of the variants we found were either silent mutations or were located in intronic regions. By contrast, others may have a functional meaning. The deletion found in 'Spunta', for example, caused the loss of important residues for strong (K-N) and weak (T-N) interaction with DNA bases (Hichri et al., 2011b). Similarly, the single amino acid substitutions we detected may have a potentially functional effect, producing structural protein modification. In grape, Hichri et al. (2011b) found that a single residue mutation in the R2 domain modified the protein interaction properties of MYB (2011b) finding structural protein modification. In grape, Hichri et al. (2011b) detected may have a potentially functional effect, producing effects on leaf and fruit flesh colour (Espley et al., 2009). Research on MYB expression is correlated with anthocyanin production in leaves. Genotypes with high leaf anthocyanin content had the highest number of copies of both AN1 mRNA and DFR/ANS mRNA. The presence of a putative SINE retrotransposon element in AN1 promoter may help to explain differences in AN1 expression between our green and red leaf varieties. In fact, it is known that the presence of transposable elements affects MYB expression (Kobayashi et al., 2004; Walker et al., 2007; Telias et al., 2011; Butelli et al., 2012; Lisch, 2013). Therefore it is possible that the different length of TA motifs in 'Magenta Love' and 'Double Fun' with respect to 'Silvy' may influence AN1 transcription levels. In apple it has been demonstrated that the presence of a 23-bp repeat motif causes an increase in MYB10 transcription levels, producing effects on leaf and fruit flesh colour (Espley et al., 2009). Research on AN1 promoter is in progress to clarify the effects of TA motif on AN1 expression level.

Previous reports have demonstrated that potato loci affecting tissue-specific accumulation of anthocyanins, including locus Pw, Pd and Pv (Kessel and Rowe, 1974; Garg et al., 1981; De Jong, 1991), are strongly associated to locus D (coding for AN1). Based on analysis of PCR fragments obtained through the amplification of AN1 in several cultivars and breeding clones, Jung et al. (2009) suggested that two or more of the classical loci may correspond to AN1 sequence variants. In light of this and considering the high variation both in AN1 expression and nucleotide sequence here reported, we can speculate that AN1 is a key gene responsible for differences in anthocyanin biosynthesis not only in the tuber but also in the leaves. In sweet potato a mechanism of anthocyanin biosynthesis common to different tissues was also suggested (Mano et al., 2007). A similar hypothesis has been formulated for the apple (Malus x domestica) Rni locus that controls the red flesh phenotype but may be an allelic variant at the MYB10 locus that cosegregates with the red foliage phenotype (Chagné et al., 2007; Espley et al., 2009). We found a strong association of StJAF13 expression with anthocyanin production in leaves: red leaf genotypes 'Double Fun' and 'Magenta Love' showed a high expression of AN1 and StJAF13. This was correlated with the expression levels of late anthocyanin structural genes (ANS and DFR). In tomato leaves a positive correlation between some secondary metabolites, including anthocyanins, and the expression of AN1 and StJAF13 (corresponding to GLABRA3-like) was also observed (Lvdval et al., 2010). The expression of StbHLH1, proposed as AN1 co-regulator candidate in tubers (De Jong et al., 2004; Payyavula et al., 2013), was virtually undetectable in leaves of 'Magenta Love,' suggesting that in this genotype StJAF13 is the main bHLH-type AN1 co-factor in leaves. By contrast, in the other red leaf genotype 'Double Fun' here analysed, the levels of expression found for StbHLH1 may still be enough to trigger or contribute to anthocyanin production. It is therefore possible that the relative contribution to anthocyanin accumulation of StbHLH1 and StJAF13 could vary depending on the tissue (e.g. leaves or tubers), the genotype and/or environmental conditions. A similar mechanism was described in snapdragon, where R2R3 MYB ROSEA determines the pattern and the level of pigmentation in both lobes and tubes, while bHLH DELILA is required in both corolla tubes and lobes of the flowers, whereas bHLH MUTABILIS is required in lobes if DELILA is not functional (Schwinn et al., 2006; Petroni and Tonelli, 2011).

StbHLH1 and StJAF13 interact with AN1 and the N-terminal portion of bHLHs is required for the interaction to take place. Our results are in accordance to those reported in petunia by Quattrocchio et al. (2006). The authors suggested that the N-terminal of PhJAF13 was sufficient for interaction with AN2 (homolog of potato AN1). They also observed that the PhJAF13 expression pattern perfectly overlapped with that of DFR and AN2 (Quattrocchio et al., 1998). The same authors found that the co-bombardment of AN2 and PhJAF13 in leaf cells induced activity of the DFR promoter, while AN2 alone was less efficient and PhJAF13 alone was insufficient for its activation. We can speculate that, depending on the potato transcription factor transformed, the expression of tobacco anthocyanin structural genes was differentially controlled. We observed an increase in anthocyanin accumulation in the co-presence of the over-expressed bHLHs and AN1, and no pigmentation variation with only StJAF13 or StbHLH1. Similarly, Butelli et al. (2012) observed that orange (Citrus sinensis) RUBY promoted a stronger pigmentation of transformed tobacco plants when co-expressed with snapdragon bHLHs. As postulated for tuber flesh (De Jong et al., 2004), AN1 may be fundamental but not sufficient for the complete pigmentation of the leaves, and the interaction with bHLHs can have a crucial importance to improve the affinity with the promoter cis-element of the structural genes. Based on the
results reported herein, it seems that the production of anthocyanins is associated to two combined mechanisms. One is linked to AN1 expression, that is correlated with pigmentation intensity. The other to AN1 allelic sequences, that could influence the mechanism of specific tissue production. It is also possible that, as hypothesised by Jung et al. (2009) for tuber flesh, the allelic configuration of different loci may influence the phenotype when AN1 is constitutively expressed. In this scenario, a single amino acid substitution could cause an alteration of the interaction between MYB proteins with the co-partners, resulting in a variety of different pigment accumulation.

In conclusion, we found that high sequence variation characterizes AN1 in potato, both in the gene body and in the promoter, and that high leaf anthocyanin content is associated to a high expression of AN1 and StJAF13 in ‘Magenta Love.’ We also demonstrated that AN1 protein physically interacts with StbHLH1 and StJAF13, and we located this latter interaction in the cell nucleus. Overexpression of AN1 together with either StJAF13 or StbHLH1 in tobacco led to a stronger pigmentation as compared to plants where only bHLHs or AN1 where expressed. All together, our results suggest that in leaves of ‘Magenta plants where only bHLHs or AN1 where expressed. All together, our results suggest that in leaves of ‘Magenta Love’ AN1 and StJAF13 can form a functional complex that drives anthocyanin biosynthesis. Future studies would be usefully spent further investigating the MBW complex, to characterize the potential pleiotropic functions of AN1 and StJAF13 and to better clarify the role of StbHLH1 in potato leaves. In the tetraploid cultivated potato, the introgression of traits like high leaf pigmentation may require several generations of crosses and selection. For this reason, our results may provide the basis to identify genes responsible for anthocyanin biosynthesis, facilitating the selection of progeny with a high level of anthocyanins in leaves.

**EXPERIMENTAL PROCEDURES**

**Plant materials and growth conditions**

Used in this study were tetraploid (2n = 4x = 48) potato (Solanum tuberosum) commercial varieties employed as parents in the breeding programs carried out in Portici. They included ‘Double Fun,’ ‘Magenta Love’ (both with purple leaves), ‘Blue Star,’ ‘Violet Queen,’ ‘Flamenco,’ ‘Briosa’ ‘Assergi,’ ‘Adora,’ ‘Blondy,’ ‘Carmine,’ ‘Desirée,’ ‘Pukara,’ ‘Ilaria,’ ‘Silvy’ and ‘Spunta.’ Two S. tuberosum haploids (2n = 2x = 24) named ‘CO125’ and ‘DE123’ were also analysed. Plants were micropropagated in vitro on Murashige and Skoog (MS) medium (Sigma-Aldrich, http://www.sigmaaldrich.com) with 1% (w/v) sucrose and 0.8% (w/v) agar, and incubated at 24°C, exposed to an irradiance of 200 μmol m⁻² sec⁻¹ and under a 16 h/8 h (light/dark) photoperiod. Three plants of each genotype were transplanted to styrofoam trays filled with sterile soil. Plants were maintained in a growth chamber at 25°C ± 2 under a 16 h/8 h (light/dark) photoperiod at 220 μmol m⁻² sec⁻¹ irradiance provided by a cool, white-fluorescent tube (Philips, http://www.philips.com).

**Gene expression analyses**

Conventional PCR was carried out on ‘Double Fun’ cDNA to amplify fragments of anthocyanin synthase (ANS), dihydroflavono- nol 4-reductase (DFR), StbHLH1, StJAF13 and AN1. The fragments obtained were cloned in pGEM-T Easy (Promega, http://www.promega.com). Plasmids obtained were normalized to a concentration of 25 ng/μl and 10-fold serial dilutions (ranging from 10⁻² to 10⁻⁷) were used to construct standard curves. The concentration of each plasmid dilution was measured using the Qubit 2.0 Fluorometer (Invitrogen) and the corresponding copy number for each concentration (copies/μl) was calculated as reported by Whe- lan et al. (2003). The synthesised cDNA of each potato genotype was then used to construct a standard curve for quantification. The PCR reactions were performed using GoTaq DNA Polymerase (Promega, http://www.promega.com) as reported by manufacturer. PCR products were analysed on 2% (w/v) agarose gel and amplicons obtained were gel-purified with QIAquick Gel Extraction Kit (Qiagen). Purified products were sequenced and aligned against alleles AN1-777 and AN1-876. A region between nucleotides −1545 and −2154 from AN1 start codon was amplified using three pairs of primers designed on PGSC (Table S1). Total DNA-free RNA was purified using the Spectrum Plant Total RNA Kit (Geneaid, On-Column DNase I Digestion Set (Sigma-Aldrich), following manufacturer's instructions. One μg of total RNA was reverse transcribed to complementary DNA (cDNA) using oligo-dT(20) (Invitrogen) and SuperScript III reverse transcriptase (Invitrogen, http://www.invitrogen.com) in 20 μl of final reaction according to the manufacturer’s instructions.

**Cloning of AN1 and basic helix-loop-helix (bHLH) genes from potato leaves**

The DNA sequences (cDS) of potato AN1 and JAF13 were amplified from cDNA of ‘Magenta Love,’ while bHLH1 from ‘Blue Star’ using Phusion High-Fidelity DNA Polymerase (Thermo-Scientific, www.thermoscientific.com). Gateway attB primers were used to obtain attB-flanked PCR. All primer pairs used are listed in Table S1. AttB-flanked PCR products were cloned in pDONR207 (Invitrogen) to obtain entry clones. Two different cDSs were cloned for both StJAF13 and StbHLH1 using different start codons of the transcript sequence annotated on Potato Genome Sequencing Consortium (PGSC). They were named StJAF13, StJAF13264, StbHLH1 and StbHLH1358. StJAF13 cloned sequence started from nucleotide 298 to nucleotide 2559. StbHLH1 corresponded to sequence annotated by Payyavula et al. (2013). StJAF13264 and StbHLH1358 corresponded to cDSs predicted on the PGSC database. All the obtained clones were sequenced.

© 2014 The Authors

---

**Total anthocyanin analysis**

Total anthocyanin content was estimated with the pH-differential spectrum method as described by Zhang et al. (2012). One hundred mg of powdered samples of tobacco (Nicotiana tabacum) shoots and potato leaves were used for this analysis.

**Nucleic acid extraction and molecular analysis of AN1**

Genomic DNA was extracted using DNeasy Plant Mini Kit (Qiagen, http://www.qiagen.com) according to the manufacturer’s instructions. Four overlapping primer pairs were designed based on the consensus sequence of AN1-777, AN1-816 alleles and their respective mRNA sequences. Primer pairs were used to amplify a region from nucleotide 14 to nucleotide 1045 of the AN1 genomic sequence (Table S1). PCR reactions were performed using GoTaq DNA Polymerase (Promega, http://www.promega.com) as reported by manufacturer. PCR products were analysed on 2% (w/v) agarose gel and amplicons obtained were gel-purified with QIAquick Gel Extraction Kit (Qiagen). Purified products were sequenced and aligned against alleles AN1-777 and AN1-876. A region between nucleotides −1545 and −2154 from AN1 start codon was amplified using three pairs of primers designed on PGSC (Table S1). Total DNA-free RNA was purified using the Spectrum Plant Total RNA Kit (Geneaid, On-Column DNase I Digestion Set (Sigma-Aldrich), following manufacturer’s instructions. One μg of total RNA was reverse transcribed to complementary DNA (cDNA) using oligo-dT(20) (Invitrogen) and SuperScript III reverse transcriptase (Invitrogen, http://www.invitrogen.com) in 20 μl of final reaction according to the manufacturer’s instructions.
was diluted five times in sterile water. One microlitre of cDNA from each genotype and of plasmid dilution were used for quantitative RT-PCR (qRT-PCR) analysis using 0.3 μM of each primer pair. All reactions were run in triplicate using QuantFast SYBR Green PCR Kit (Qiagen) in a final volume of 20 μl of reaction. qRT-PCR was carried out using the Rotor-Gene 6000 (Corbett, http://www.corbettlifeScience.com) and cycle conditions indicated by QuantFast SYBR Green PCR Kit handbook (Qiagen). Gene expression analysis was carried out using Rotor-Gene 6000 software. The standard curves were used to calculate the copy number of molecules per μl of the corresponding target genes in each potato genotype. Primer pairs used for qRT-PCR analysis are listed in Table S1.

**Yeast two-hybrid assay**

The CDS of AN1 was cloned in frame in pGADT7 (Clontech, http://clontech.com) between EcoRI and XhoI restriction sites. StJAF13, StbHLH1, StJAF13::StbHLH1 and StbHLH358 were inserted into pGBK7 (Clontech) using XmaI and SalI. All plasmids were sequenced to ensure that no mutations had been introduced. Bait and prey plasmids were co-transformed into yeast strain AH109 according to Bai and Elledge (1997). Co-transformants were grown overnight in liquid culture and an equal amount of cells for all co-transformations was spotted on media with and without histidine and adenine to check for bait and prey interaction.

**Bimolecular Fluorescent Complementation (BiFC) assay**

The p35S::AN1::Nt-YFP, p35S::StJAF13::Ct-YFP, p35S::Nt-YFP or p35S::Ct-YFP constructs were obtained from the corresponding entry clones using Gateway recombination technology (Invitrogen). BiFC experiments were performed by co-transforming different combinations of the constructs into 2-week-old Nicotiana benthamiana leaves via Agrobacterium tumefaciens and setting the final OD (optical density) of each transformation as described in Payyavula et al. (2013) using 150 μM acetosyringone and setting the final OD (optical density) of each Agrobacterium suspension to 0.3. Imaging was conducted three days after infiltration with a Leica TCS SP8 confocal laser scanning microscope. The software package provided by the manufacturer was used for projections of serial optical sections and image processing.

**Overexpression of potato transcription factors in tobacco plants**

The StJAF13, StbHLH1 and AN1 genes from entry clones were cloned in the 3SSCaMV expression cassette of pGWB411 (Nakagawa et al., 2007) using Gateway recombination technology (Invitrogen). A. tumefaciens strain LBA4404 transformed with each expression vector was used for agroinfiltration in fully expanded leaves of N. benthamiana as reported by Payyavula et al. (2013). Stable genetic transformation was carried out by co-cultivation of N. tabacum leaf explants with A. tumefaciens in accordance with Horsch et al. (1985).

**Bioinformatics and statistical analysis**

Amino acid and nucleotide sequences were collected from the National Center for Biotechnology Information (NCBI) and PGSC database. Alignments and phylogenetic trees were performed with ClustalW (Geneious software v6.0.6) (Biomatters, http://www.geneious.com/). BLASTP and BLASTX programs (http://www.ncbi.nlm.nih.gov/blast) were used to perform homology researches in GenBank. Analysis of variance (ANOVA) on qPCR data was carried out using XLSTAT-PRO 7.5.3 software (Addinsoft, http://www.xlstat.com). Duncan’s test was performed to compare mean values.

**Accession numbers**

The cloned CDS of StJAF13 (HG783386) and StbHLH1 (HG783383) were submitted to the GenBank/EMBL database. Sequence data from this article can be found at the NCBI and PGSC database. Nucleotide sequences can be found under the following accession numbers: potato, AN1-816 genomic DNA (AY841128), AN1-816 mRNA (AY841127), AN1-777 genomic DNA (AY841130), AN1-777 mRNA (AY841129), bHLH1 (JX848660), ANS (HG071028), DFR (AY289921); tomato (S. lycopersicum), GLABRA3-like (XM_004245552); tobacco (N. tabacum), JAF13-like a (KF305768) and b (KF305769); petunia (Petunia x hybrida), PhJAF13 (AF020545); SINe SoISv_Sis (HE583477). Protein sequences have the following accession numbers: potato, AN1 (AGG31676), CAI1 (ABY40370), MYBA1 (AF031616); tomato, GLABRA2-like (XP_004245660), AN1 (ACT398604); petunia, PhJAF13 (AAC29459); tobacco, JAF13-like a (AGX01001) and b (AGX01002); Perilla frutescens, MYC-RP (AAV75513); snapper (Antirrhinum majus); DELIL2 (AEM3394); Ipomea nil, bHLH (BAE94395). Transcript and genomic sequence of JAF13 can be found at PGSC database (PGSC0003DMT400032139 and chr08:54833821..54833836, respectively). The genomic localization of the analysed AN1 promoter region was chr10:51745200..51749200 on the PGSC database.

**ACKNOWLEDGEMENTS**

This work was funded by the Italian Ministry of University and Research (MIUR)-PON02 R&C 2007-2013 PON02_00395_3215002 GenHORT (D.D. n. 813/Ric.). We thank M. Walters and M. Van Osst for editing the manuscript, M. Mazzei for providing tubers of commercial varieties and two anonymous reviewers for helpful comments. Technical assistance was provided by R. Nocerino and R. Garramone.

**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article.

**Figure S1.** Single nucleotide substitutions and deletions in AN1 exons with respect to reference alleles (AN1-816 and AN1-777) present in GenBank database and AN1 promoter analysis.

**Figure S2.** Phylogenetic analysis of MYB transcription factors which displayed similarity with the predicted protein of AN1 cloned from “Magenta Love.”

**Figure S3.** Genomic structure proposed for StJAF13; JAF13 genomic sequence annotated on PGSC database; genomic sequence of GLABRA3-like of S. lycopersicum.

**Figure S4.** Agarose gel showing the products obtained from the amplification of exon 4 of JAF13.

**Figure S5.** Phylogenetic analysis of bHLH transcription factors that displayed similarity with the predicted protein of StbHLH1 cloned from “Magenta Love”.

**Figure S6.** Phylogenetic analysis of bHLH transcription factors that displayed similarity with the predicted protein of StbHLH1 cloned from “Blue Star”.

**Table S1** List of primers used in this study.

**REFERENCES**

Anthocyanin accumulation in potato leaves


© 2014 The Authors


