RNA sequencing on Solanum lycopersicum trichomes identifies transcription factors that activate terpene synthase promoters

Spyropoulou, E.A.; Haring, M.A.; Schuurink, R.C.

DOI

Publication date
2014

Document Version
Final published version

Published in
BMC Genomics

Citation for published version (APA):
RNA sequencing on *Solanum lycopersicum* trichomes identifies transcription factors that activate terpene synthase promoters

Eleni A Spyropoulou, Michel A Haring and Robert C Schuurink*

**Abstract**

**Background:** Glandular trichomes are production and storage organs of specialized metabolites such as terpenes, which play a role in the plant’s defense system. The present study aimed to shed light on the regulation of terpene biosynthesis in *Solanum lycopersicum* trichomes by identification of transcription factors (TFs) that control the expression of terpene synthases.

**Results:** A trichome transcriptome database was created with a total of 27,195 contigs that contained 743 annotated TFs. Furthermore a quantitative expression database was obtained of jasmonic acid-treated trichomes. Sixteen candidate TFs were selected for further analysis. One TF of the MYC bHLH class and one of the WRKY class were able to transiently transactivate *S. lycopersicum* terpene synthase promoters in *Nicotiana benthamiana* leaves. Strikingly, SlMYC1 was shown to act synergistically with a previously identified zinc finger-like TF, Expression of Terpenoids 1 (SlEOT1) in transactivating the *SlTPS5* promoter.

**Conclusions:** High-throughput sequencing of tomato stem trichomes led to the discovery of two transcription factors that activated several terpene synthase promoters. Our results identified new elements of the transcriptional regulation of tomato terpene biosynthesis in trichomes, a largely unexplored field.

**Keywords:** Tomato trichomes, Terpene biosynthesis, Transcription factor, High-throughput sequencing

**Background**

Specialized glandular trichomes can produce and accumulate large quantities of terpenoids, phenylpropanoids, flavonoids and alkaloids, which they can also secrete [1]. RNA sequencing in combination with metabolite profile analysis of glandular trichomes and proteomics have shed light on the biosynthesis of specialized metabolites in the trichomes of various plant species [2]. Through the production of EST libraries, micro-arrays and high-throughput sequencing of (glandular) trichome RNA, genes have been identified that are involved in the terpene, phenylpropanoid, alkaloid and flavonoid biosynthesis in various plant species, including tomato [3-5], sweet basil [6,7], tobacco [8,9], mint [10], alfalfa [11], *Artemisia annua* [12] and hop [13]. Although EST sequencing has been instrumental in the discovery of enzymes of trichome-specialized metabolism [4], next generation sequencing (NGS) can give a more in-depth picture of transcriptomes. NGS technologies (i.e. RNA sequencing) has been used for characterization of several trichome transcriptomes, for example from plants of medical importance like *Artemisia annua* (Asteraceae; [12]) or *Huperzia serrata* and *Phlegmaria var carinatus* (Huperziaceae; [14]). NGS has also been used for gene discovery, for example in combination with shotgun proteomics and metabolite analysis of tomato (*Solanum lycopersicum*) trichomes, leading to the discovery of the leaf-trichome-specific β-caryophyllene/α-humulene synthase (CAHS; [4]). NGS of trichomes RNA from wild and cultivated tomato varieties led to the discovery and characterization of various sesquiterpene synthases, providing insight into the evolution of terpene synthases [15].

Terpene biosynthesis in tomato plants is of major interest because terpenes play an important role in the plant’s defense [16-20]. The sequencing of the cultivated tomato genome has enabled the characterization of its terpene synthase (TPS) gene family [21,22], but not much...
is known about the regulation of the terpenoid pathway. Transcriptional control of biosynthetic genes is a major mechanism by which secondary metabolite production is regulated [23,24].

There are not many transcription factors (TFs) known to be involved in regulation of terpenoid pathways. ORCA3, a jasmonate-responsive APETALA2 (AP2)-domain transcription factor from Catharanthus roseus, has been shown to regulate expression of Strictosidine Synthase (STR) involved in terpene indole alkaloid biosynthesis [25]. Subsequently, a methyl-jasmonate (MeJA)-inducible transcription factor of the MYC family (CrMYC2) was shown to positively regulate ORCA3 [26]. CrWRKY1 was identified as being involved in the root-specific accumulation of serpentine in C. roseus plants and as being induced by phytohormones including JA [27]. This TF appeared to negatively regulate ORCA3 and to a lesser extend CrMYC2 [27]. A MeJA-inducible WRKY transcription factor from Gossypium arboreum that regulates the sesquiterpene synthase (+)-ß-cadinene synthase A in cotton fibers was identified by Xu et al. [28]. Ma et al. [29] demonstrated that a MeJA-inducible WRKY transcription factor from Artemisia annua is involved in the regulation of artemisinin biosynthesis. More recently two JA-responsive AP2 family transcription factors from A. annua (AaERF1 and 2) were found to regulate Amorpha-4,11-diene synthase (ADS), a sesquiterpene synthase involved in the biosynthesis of artemisinin [30] whereas Lu et al. [31] identified AaORA, a AP2/ERF TF, that regulates several genes in the artemisinin biosynthetic pathway including AaERF1. Most recently, the MeJA-inducible Arabidopsis thaliana MYC2 transcription factor [32] was shown to regulate sesquiterpene synthases AtTPS21 and AtTPS11 [33].

Here, we used NGS of tomato stem trichomes as a tool for gene discovery. First, a transcript database was created from normalized cDNA, which was mined for transcription factors. Then, in order to narrow down the number of TFs potentially involved in terpene biosynthesis, an expression profiling database was created using Illumina sequencing of trichome RNAs from plants treated with or without jasmonic acid (JA), since JA is known to induce terpene emission in tomato and to regulate several terpene synthases [16,21,34,35]. To identify TFs that regulate terpene biosynthesis we used a transient assay based on the transactivation of tomato terpene synthase promoters in planta.

Results

Assembly of RNAseq data and Genome Analyzer II transcript profiling

We created a tomato trichome EST database by sequencing a mixture of glandular and non-glandular trichome RNAs, derived from stems of Solanum lycopersicum cv. Moneymaker plants. The resulting cDNA was normalized prior to being used as input for 454 GS FLX Titanium pyrosequencing. A full plate was sequenced consisting of two halves: one with cDNAs originating from control plants and the other half with cDNAs originating from plants treated with JA. In total we obtained 979,076 high-quality reads with an average length of 337 bp. The reads from control and JA-treated samples were assembled de novo resulting in 27,195 contigs with an average length of 931 bp, leaving 24,187 reads unmapped (singletons), with an average length of 241 bp. Nucleotide sequences of the contigs were blasted against the Solanaceae Genomics Network (SGN) tomato database for annotation, using a local E-Blast tool; 3,295 contigs were not annotated.

For creating the transcript profiling databases with Genome Analyzer II, the same RNA material as for the 454 sequencing was used, but this time the cDNA derived from control and JA-treated stem trichomes was not normalized before being processed. We specifically obtained 5,631,975 3’ sequences from the Control sample and 5,882,547 from the JA-treated sample. 4,840,738 and 5,169,891 reads from the Control and JA-samples, respectively, were mapped to one unique contig of the trichome database. In addition, 38,699 (C) and 45,375 (JA) reads were mapped to multiple contigs and 791,237 (C) and 712,656 (JA) remained unmapped.

Both the 454 GS FLX Titanium reads and the Genome Analyzer II reads can be found in the Sequence Read Archive of NCBI (http://www.ncbi.nlm.nih.gov/sra) under accession number SRP041373.

Annotation, gene ontology and protein families

In order to characterize the S. lycopersicum stem trichome transcriptome the unique contigs (27,195 ESTs) were submitted to homology searches (BLASTX) in the National Center for Biotechnology Information (NCBI) non-redundant protein database using Blast2GO [36], 4,733 contigs did not return a BLASTX hit. The majority of the top hits were to protein sequences of Vitis vinifera, followed by Populus trichocarpa, Ricinus communis and Solanum lycopersicum.

Next, gene ontology (GO) and enzyme classifications (EC) were performed in order to classify the ESTs. It must be noted that one sequence could be assigned to more than one GO term. For the cellular component class the assignments were mostly given to cell and organelle (54,82% and 29,35% respectively; Additional file 1: Figure S1a). The highest percentage of molecular function GO terms were in binding and catalytic activity (42,96% and 41,38% respectively; Additional file 1: Figure S1c). In the biological processes, the majority of the GO terms was grouped into two categories- those of metabolic and cellular process (36,55% and 32,79% respectively; Additional file 1: Figure S1b). Finally, within the predicted ECs, the prevailing categories of enzymes were transferases.
and oxidoreductases (31.38% and 29.65% respectively; Additional file 1: Figure S1d).

The search of additional databases for protein families, domains, regions and sites was performed from Blast2GO via the InterPro EBI web server. The 30 top InterPro entries obtained are presented in Table 1. The most dominant class of enzymes was protein kinases. Abundantly represented were also cytochrome P450s.

Finally, within Blast2GO, the EC numbers were classified in KEGG pathways, enabling the presentation of enzymatic functions in the context of the metabolic pathways in which they are part of (Blast2GO Tutorial, [37]). Among the pathways identified, the ones related to secondary metabolism are shown in Table 2. Lipid transfer proteins represented 0.19% of the tomato stem trichome transcripts.

Transcripts of enzymes involved in jasmonic acid biosynthesis and signaling pathways were also identified in the trichome database. Data for a selection of enzymes are presented in Table 3, including known JA marker genes such as LOXA (U09026), AOC (AW624058; [38]), JAZ1 (EF591123; [39]) and JAZ3 (EU194561; [40]).

A closer look was taken at the terpene biosynthesis pathway (Figure 1) in order to see if the precursor pathways were up-regulated by JA. As shown in Table 4, expression of some precursor genes in tomato was induced by JA although not strongly (max induction ~2.5-fold for HDS). As in other plants [41], genes encoding enzymes of the precursor pathways can belong to small gene families and it appears that expression levels and JA-inducibility of these members can vary. Transcript abundance of precursor genes is presented in Table 4 for comparison with the

<table>
<thead>
<tr>
<th>Table 1 Summary of the most common InterPro entries found in the <em>S. lycopersicum</em> stem trichome transcriptome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>InterPro</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>IPR011009</td>
</tr>
<tr>
<td>IPR000719</td>
</tr>
<tr>
<td>IPR002290</td>
</tr>
<tr>
<td>IPR008271</td>
</tr>
<tr>
<td>IPR020635</td>
</tr>
<tr>
<td>IPR016040</td>
</tr>
<tr>
<td>IPR013083</td>
</tr>
<tr>
<td>IPR017441</td>
</tr>
<tr>
<td>IPR002885</td>
</tr>
<tr>
<td>IPR015943</td>
</tr>
<tr>
<td>IPR001841</td>
</tr>
<tr>
<td>IPR012677</td>
</tr>
<tr>
<td>IPR016024</td>
</tr>
<tr>
<td>IPR001245</td>
</tr>
<tr>
<td>IPR000504</td>
</tr>
<tr>
<td>IPR001680</td>
</tr>
<tr>
<td>IPR011046</td>
</tr>
<tr>
<td>IPR011990</td>
</tr>
<tr>
<td>IPR011989</td>
</tr>
<tr>
<td>IPR001128</td>
</tr>
<tr>
<td>IPR017986</td>
</tr>
<tr>
<td>IPR017853</td>
</tr>
<tr>
<td>IPR012287</td>
</tr>
<tr>
<td>IPR001611</td>
</tr>
<tr>
<td>IPR009057</td>
</tr>
<tr>
<td>IPR016196</td>
</tr>
<tr>
<td>IPR012336</td>
</tr>
<tr>
<td>IPR013781</td>
</tr>
<tr>
<td>IPR002213</td>
</tr>
<tr>
<td>IPR002401</td>
</tr>
<tr>
<td>KEGG pathway</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td><strong>Terpenoid biosynthesis</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Phenylpropanoid biosynthesis</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Flavonoid biosynthesis</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Alkaloid biosynthesis</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Steroid biosynthesis</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
expression levels of the 13 terpene synthases (TPSs) found in stem trichomes Table 5).

Selection of transcription factors potentially involved in regulating terpene synthases

Based on the annotated contigs 743 transcription factors of different classes were found in the trichome database: 69 WRKY, 151 MYB, 8 MYC, 52 bZIP, 9 AFE, 71 ERF, 17 ZnF, 28 bHLH, 12 MADS, 1 NAC and 325 of unknown function/class. Out of those, 151 were up-regulated (>1.5x) by the treatment with JA, 119 were down-regulated (<0.67x) and expression of 473 TFs remained unaltered. Since JA is known to play a role in the plant’s direct and indirect defenses we were interested in those transcription factors that were induced by JA and could therefore potentially be involved in up-regulating terpene biosynthesis. 56 of the TFs that were up-regulated by JA showed an induction higher than 2-fold. The sequence of these 56 TFs was blasted against the tomato genomic sequence (Solanaceae Genomics Network, SGN) and complete ORFs were constructed when possible (GENSCAN, [46]), if not provided by the RNAseq. These sequences were submitted to homology search after translation against the NCBI database for identifying conserved domains. From this analysis 16 TFs (Table 6) were selected for further investigation as follows: we focused on classes of TFs involved in the regulation of terpenoids identified so far in other plant species—namely TFs of the APETALA2 class [25,30,31], WRKY class [27-29] and MYC class [26,33]. In total eleven transcription factors of the AP2 class, four of the WRKY class and one of the MYC class, although it only showed a 1.4-fold induction, were selected for further investigation of their potential involvement in regulating expression of terpene synthases.

Table 3: List of selected enzymes involved in the jasmonic acid biosynthesis and signaling

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Contig Nr</th>
<th>SGN Nr</th>
<th>Annotation</th>
<th>Transcript length (bases)</th>
<th>Expression values JA*</th>
<th>Expression values C*</th>
<th>Fold JA induction</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOXA⁴</td>
<td>6402</td>
<td>SGN-U592535</td>
<td>Lipoygenase A</td>
<td>2837</td>
<td>35.2</td>
<td>1.6</td>
<td>22</td>
</tr>
<tr>
<td>AOC⁵</td>
<td>24817</td>
<td>SGN-U562649</td>
<td>Allene oxide cyclase</td>
<td>1645</td>
<td>1898.8</td>
<td>417.6</td>
<td>4.5</td>
</tr>
<tr>
<td>JAZ1ᵇ</td>
<td>6863</td>
<td>SGN-U579837</td>
<td>Jasmonate ZIM-domain 1</td>
<td>1156</td>
<td>12.2</td>
<td>0.7</td>
<td>17.4</td>
</tr>
<tr>
<td>JAZ3⁰</td>
<td>20751</td>
<td>SGN-U564446</td>
<td>Jasmonate ZIM-domain 3</td>
<td>986</td>
<td>18.87</td>
<td>2.3</td>
<td>8.2</td>
</tr>
<tr>
<td>COI²</td>
<td>24353</td>
<td>SGN-U568988</td>
<td>Coronatine-insensitive 1</td>
<td>2260</td>
<td>11</td>
<td>8.5</td>
<td>1.29</td>
</tr>
</tbody>
</table>

*Expression values are in RPKM (reads per kilobase of transcript per million mapped reads).

Figure 1: Enzymes involved in the precursor biosynthesis for mono-(C10), sesqui-(C15) and di-(C20) terpenes. An explanation of the abbreviations used in the pathways and the GAII reads for each enzyme are shown in Tables 4 and 5, respectively. ER, endoplasmic reticulum, TPS; terpene synthase, SPS; solanesyl diphosphate synthase. The grey oval circles between organelles and at the cell membrane represent putative transporter systems. Peroxisomal localization of precursor enzymes of the MVA pathway has been previously reported in Arabidopsis thaliana and Catharanthus roseus [42-44].
The sixteen candidate TFs should ideally be trichome-specifically expressed and possibly induced by jasmonic acid. In order to investigate the expression pattern of these genes, cDNA was synthesized from different *S. lycopersicum* cv. Moneymaker organs and tissues: leaves, stems, isolated stem trichomes and roots from 4-week-old plants, as well as flowers and fruit of mature plants. In Figure 2 transcript levels, as determined by Q-RT-PCR, are presented for four of the sixteen selected transcription factors. For the other twelve candidate TFs expression in the trichomes was much lower than that in the other organs/tissues and these were excluded from further analysis. TF *SlMYC1* (KF430611) was predominately expressed in trichomes, but also in leaves and flowers (Figure 2a). *SlWRKY78* was expressed in leaves, trichomes, roots and flowers (Figure 2b). *SlWRKY28* was a trichome-specific
gene (Figure 2c) and SIWRKY73 was expressed in trichomes, roots and fruit (Figure 2d). Q-RT-PCR analyses indicated that none of the selected transcription factors was significantly induced by JA according (Figure 2). SIWRKY73 expression appeared to be approximately 1.7-fold reduced in JA treated plants (p = 0.07).

**SIMYC1 and SIWRKY73 can transactivate terpene synthase promoters in Nicotiana benthamiana leaves**

In order to investigate whether these TFs could activate a selection of terpene synthase promoters, a transient assay in *Nicotiana benthamiana* leaves was used, which has been previously shown to work for the interaction between the zinc finger-like transcription factor Expression of Terpenoids 1 (SlEOT1) and the SlTPS5 promoter [47]. In the reporter construct, expression of β-glucuronidase (*uidA, GUS*) is driven by the glandular trichome-specific promoter of *SlTPS5*. Co-infiltration with the 35S:SlEOT1 effector construct resulted in transactivation of the *SlTPS5* promoter, leading to GUS expression in this heterologous system (Figure 3). As negative control for the effector, a 35S:RFP construct was used. Various other reporter constructs with promoters of other terpene synthases-SlTPS3, SlTPS7 and SlTPS8-driving expression of GUS or a GUSsYFP1 fusion (*SlTPS9*) were included in the analyses.

### Table 5 Terpene synthases (TPS) found in *S. lycopersicum* stem trichomes

<table>
<thead>
<tr>
<th>TPS</th>
<th>Transcript length (bases)</th>
<th>Expression values JA</th>
<th>Expression values C</th>
<th>Fold JA induction</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2099</td>
<td>15.8</td>
<td>2.3</td>
<td>6.87</td>
</tr>
<tr>
<td>5</td>
<td>2186</td>
<td>103.18</td>
<td>46.7</td>
<td>2.2</td>
</tr>
<tr>
<td>7</td>
<td>1069</td>
<td>0.5</td>
<td>0.18</td>
<td>2.78</td>
</tr>
<tr>
<td>9</td>
<td>2011</td>
<td>3060.75</td>
<td>2359.03</td>
<td>1.3</td>
</tr>
<tr>
<td>12</td>
<td>407</td>
<td>4.08</td>
<td>2.66</td>
<td>1.53</td>
</tr>
<tr>
<td>16</td>
<td>1868</td>
<td>18.36</td>
<td>16.87</td>
<td>1.09</td>
</tr>
<tr>
<td>17</td>
<td>1190</td>
<td>4.25</td>
<td>3.55</td>
<td>1.2</td>
</tr>
<tr>
<td>19</td>
<td>776</td>
<td>44.7</td>
<td>37.46</td>
<td>1.19</td>
</tr>
<tr>
<td>20</td>
<td>1148</td>
<td>142.79</td>
<td>106.89</td>
<td>1.33</td>
</tr>
<tr>
<td>24</td>
<td>854</td>
<td>0.17</td>
<td>0.18</td>
<td>0.94</td>
</tr>
<tr>
<td>31</td>
<td>1991</td>
<td>0.68</td>
<td>0.18</td>
<td>3.78</td>
</tr>
<tr>
<td>39</td>
<td>1131</td>
<td>12.07</td>
<td>5.68</td>
<td>2.12</td>
</tr>
<tr>
<td>41</td>
<td>2368</td>
<td>71.57</td>
<td>66.58</td>
<td>1.07</td>
</tr>
</tbody>
</table>

*Expression values are in RPKM (reads per kilobase of transcript per million mapped reads). For *SITPS8* no transcripts were identified in the stem trichome database although by Q-RT-PCR minimal expression has been observed [45].

### Table 6 List of selected *S. lycopersicum* transcription factors (TF) potentially involved in terpene biosynthesis

<table>
<thead>
<tr>
<th>Name</th>
<th>Contig nr</th>
<th>SGN nr</th>
<th>Annotation</th>
<th>Transcript length (bases)</th>
<th>Expression values JA</th>
<th>Expression values C</th>
<th>Fold JA induction</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIAP2_9</td>
<td>83</td>
<td>SGN-US72361</td>
<td>ERF (ethylene response factor) subfamily B-3 of ERF/AP2 TF family</td>
<td>371</td>
<td>9.69</td>
<td>4.08</td>
<td>2.37</td>
</tr>
<tr>
<td>SIAP2_2</td>
<td>1719</td>
<td>SGN-US96590</td>
<td>DREB subfamily A-1 of ERF/AP2 TF family</td>
<td>422</td>
<td>13.6</td>
<td>1.95</td>
<td>6.97</td>
</tr>
<tr>
<td>SIAP2_6</td>
<td>5289</td>
<td>SGN-US63871</td>
<td>AP2 domain-containing TF</td>
<td>933</td>
<td>45.39</td>
<td>12.78</td>
<td>3.55</td>
</tr>
<tr>
<td>SIAP2_3</td>
<td>7031</td>
<td>SGN-US63215</td>
<td>DREB subfamily A-1 of ERF/AP2 TF family</td>
<td>981</td>
<td>5.95</td>
<td>1.42</td>
<td>4.19</td>
</tr>
<tr>
<td>SIAP2_7</td>
<td>7865</td>
<td>SGN-US87768</td>
<td>DREB subfamily A-4 of ERF/AP2 TF family</td>
<td>775</td>
<td>2.2</td>
<td>0.71</td>
<td>3.09</td>
</tr>
<tr>
<td>SIAP2_4</td>
<td>10714</td>
<td>SGN-US86437</td>
<td>ERF (ethylene response factor) subfamily B-3 of ERF/AP2 TF family</td>
<td>611</td>
<td>2.72</td>
<td>0.71</td>
<td>3.83</td>
</tr>
<tr>
<td>SIAP2_10</td>
<td>14672</td>
<td>SGN-US85539</td>
<td>AP2 domain-containing TF</td>
<td>297</td>
<td>18.87</td>
<td>7.99</td>
<td>2.36</td>
</tr>
<tr>
<td>SIAP2_5</td>
<td>16204</td>
<td>SGN-US77088</td>
<td>ERF (ethylene response factor) subfamily B-4 of ERF/AP2 TF family</td>
<td>523</td>
<td>0.68</td>
<td>0.18</td>
<td>3.78</td>
</tr>
<tr>
<td>SIAP2_11</td>
<td>25582</td>
<td>SGN-US84756</td>
<td>ERF (ethylene response factor) subfamily B-2 of ERF/AP2 TF</td>
<td>660</td>
<td>1.19</td>
<td>0.53</td>
<td>2.24</td>
</tr>
<tr>
<td>SIAP2_1</td>
<td>25985</td>
<td>SGN-US86438</td>
<td>ERF (ethylene response factor) subfamily B-3 of ERF/AP2 TF family</td>
<td>942</td>
<td>16.32</td>
<td>0.53</td>
<td>30.8</td>
</tr>
<tr>
<td>SIAP2_8</td>
<td>26482</td>
<td>SGN-US81852</td>
<td>Ethylene-responsive element-binding factor 4 Homolog</td>
<td>788</td>
<td>58.81</td>
<td>24.32</td>
<td>2.42</td>
</tr>
<tr>
<td>SIWRKY22</td>
<td>9827</td>
<td>SGN-US565154</td>
<td>WRKY family TF</td>
<td>1366</td>
<td>0.68</td>
<td>0.18</td>
<td>3.78</td>
</tr>
<tr>
<td>SIWRKY28</td>
<td>10561</td>
<td>SGN-US84367</td>
<td>WRKY family TF</td>
<td>1326</td>
<td>0.5</td>
<td>0.18</td>
<td>2.78</td>
</tr>
<tr>
<td>SIWRKY78</td>
<td>13200</td>
<td>SGN-US565174</td>
<td>WRKY family TF</td>
<td>1099</td>
<td>1.19</td>
<td>0.18</td>
<td>6.61</td>
</tr>
<tr>
<td>SIWRKY73</td>
<td>20918</td>
<td>SGN-US71278</td>
<td>WRKY family TF</td>
<td>1453</td>
<td>0.5</td>
<td>0.18</td>
<td>2.78</td>
</tr>
<tr>
<td>SIMYC1</td>
<td>24332</td>
<td>SGN-US76396</td>
<td>MYC TF</td>
<td>2174</td>
<td>32.47</td>
<td>22.7</td>
<td>1.43</td>
</tr>
</tbody>
</table>

*Expression values are in RPKM (reads per kilobase of transcript per million mapped reads).
Figure 2 (See legend on next page.)
As shown in Figure 3a, SlWRKY73 could transactivate the SITPS5 promoter, albeit to a lower extent than SlEOT1. SlWRKY73 transactivated the SITPS3 and SITPS7 promoters only weakly, and the SITPS8 and SITPS9 promoters not at all (35S:RFP negative controls shown in Additional file 1: Figure S2). SlWRKY78 or SlWRKY28 did not transactivate any of the terpene synthase promoters (Additional file 1: Figure S3).

SlMYC1 could transactivate all terpene synthase promoters tested except SITPS8. Transactivation of the trichome-specific SITPS5 and SITPS9 promoters was strongest (Figure 3b; 35S:RFP negative control shown in Additional file 1: Figure S2). However, it should be noted that GUS activity of a promoter driving the GUS-sYFP1 fusion was lower than when the same promoter driving GUS alone was transactivated by an effector construct (data not shown), possibly because the fusion protein was less stable or produced. Therefore, transactivation by SlMYC1 of the trichome-specific SITPS9 promoter was potentially stronger than that detected here.

SlMYC1 and SlEOT1 act synergistically on the SITPS5 promoter in N. benthamiana leaves

Since SlEOT1, SlMYC1 and SlWRKY73 were shown in separate experiments to be able to transactivate the SITPS5 promoter (Figure 3), we investigated what effect a combination of these transcription factors would have on the transactivation of this promoter. To this end, Agrobacterium cultures carrying the CaMV 35S-driven effector constructs were mixed in pairs or all three together and combined with the SITPS5p:GUS reporter construct and infiltrated in N. benthamiana leaves.

![Graph](image-url)
Interestingly, co-expression of SlEOT1 and SlMYC1 almost tripled the transactivation of SlTPS5 promoter compared to the effect of each TF alone. Adding SlWRKY73 did not have an additional effect, but rather seemed to have a negative effect on the combinatorial action of the other two TFs, although not at a statistically significant level (Figure 4).

Discussion and conclusions

RNA sequencing of S. lycopersicum stem trichomes led to the identification of one MYC bHLH and one WRKY transcription factor that can transactivate several terpene synthase promoters. The observation that SlMYC1 acts synergistically with SlEOT1 in the transactivation of the SlTPS5 promoter suggests a complex regulatory network for terpene biosynthesis.

High-throughput sequencing of Solanum lycopersicum stem trichomes

We used massive parallel pyrosequencing on the 454 GS FLX Titanium platform to sequence S. lycopersicum stem trichome RNAs with the goal to identify transcription factors involved in terpene biosynthesis. We used normalized cDNA to maximize representation of low abundant transcripts and reduce representation of highly abundant transcripts. Attempts to map the obtained reads to the publicly available mixed tissue SGN database led to a high percentage of unmapped reads and assignment of the same reads to multiple unigenes and therefore the reads were assembled de novo. 2.5% of the reads could not be matched and were not used in further analysis. 87.9% of the resulting contigs were subsequently annotated after blasting against the SGN tomato database using a local E-Blast tool. In this database we identified annotated enzymes involved in several metabolic pathways (Additional file 1: Table S1). In short, compared to the study published by McDowell and colleagues [3] on S. lycopersicum cv. M82 trichomes, we identified in Moneymaker trichomes cDNAs encoding enzymes involved in for example the TCA cycle, starch and sucrose metabolism (Additional file 1: Table S1), as well as secondary metabolite biosynthesis (Table 2). Photosynthesis related genes were also identified but were not as prevalent (Additional file 1: Table S1) as in M82 trichomes. Such differences could originate from the fact that in our study we used a mix of Moneymaker trichome types, including stalks, whereas McDowell and colleagues focused on comparing different types of trichomes between Solanum species and so clipped off and analyzed only the secretory cells of glandular trichomes [3]. Furthermore we created an expression profiling database using Illumina sequencing in order to obtain genes regulated by JA. The success of the JA treatment is evident by the high induction of known JA markers, some of which are presented in Table 3 (LOXA, AOC [38]; JAZ1 [39]; JAZ3 [40]).

Jasmonic acid regulation of the terpene biosynthesis pathway in tomato trichomes

In order to investigate whether in stem trichomes of tomato Moneymaker plants, regulation of terpene biosynthesis by JA is also on the precursor level besides on the level of individual TPSs [21], the quantitative database was mined for enzymes of the precursor pathways. The copy number of these genes varies between different plant species [41] and, as shown in Table 4, different family members can vary in their expression levels and/or JA-inducibility. For example 1-deoxy-d-xylulose 5-phosphate synthase (DXS), in contrast with Arabidopsis, which contains a single functional gene, has diversified into two isogenes in other plant species including tomato [48]. Whereas SlDXS1 is ubiquitously expressed, SlDXS2 is expressed only in a few tissues and in leaf trichomes its transcript abundance is much higher than that of SlDXS1 [49], although this is not the case in stem trichomes (Table 4). Furthermore, SlDXS2 is moderately induced by

![Figure 4](https://example.com/screenshot.png)

**Figure 4** Transactivation of SlTPS5 promoter by SlEOT1, SlMYC1, SlWRKY73 or combination thereof in N. benthamiana leaves. Normalized GUS activity after co-infiltration with A. tumefaciens harboring the 35S:EOT1, 35S:MYC1, 35S:WRKY73 effector constructs or combination thereof and the SlTPS5p:GUS reporter construct. The 35S:RFP effector construct was used as negative control. The bars represent the obtained mean values and the error bars the standard error (n = 4). RFP, red fluorescent protein. Letters indicate significant differences (ANOVA, P < 0.05 according to Tuckey’s B posthoc test). Representative results from two experiments are shown.
wounding in the cultivar Moneymaker [49], which correlates with the observed moderate induction of *SIDXS2* by JA (~1.6-fold, Table 4). *SIDXS2* expression is also approximately threefold upregulated in the tomato cultivar Castlemart upon feeding by *Manduca sexta* larvae [50].

The regulation of precursor genes of the MEP pathway by wounding, hormones or elicitors has been demonstrated in various plant species [49-54]. Similarly, evidence for the regulation of precursor biosynthesis of the mevalonate (MVA) pathway is also abundant [55-60]. For example, HMGR enzyme activity and protein level were shown to increase by fungal infection in potato tubers and sweet potato root [59]. Furthermore, *HMGR1* expression was induced by treatment with MeJA in potato, whereas *HMGR2* expression was reduced [56]. In response to caterpillar herbivory, transcripts of *HMGR1* were reduced in alfalfa [60]. Our results show that in tomato stem trichomes *HMGR1* and *HMGR3* were induced by JA treatment approximately 2-fold, whereas expression of *HMGR2* remained unaltered (Table 4). None of the prenyl diphosphate synthases were induced in tomato trichomes by JA treatment, whereas two seemed to be downregulated (*FPS*, SGN-U578686; and *GGPS*, SGN-U573348; Table 4). We did not find any transcripts for *GGPS1* (SGN-U574849) in our stem trichome database, although it has been shown to be induced in tomato leaves by JA-treatment [61]. Finally, from the very recently identified *cis*-prenyltransferases only *CPTS*, that produces medium-length chain polyisoprenoids [62], was upregulated by JA, 1.7-fold (Table 4).

**Identification of transcription factors involved in regulating terpene synthases in tomato trichomes**

Our primary aim was to identify transcription factor(s) that regulate terpene biosynthesis. Based on the annotated contigs, 2.7% of the transcripts in the tomato stem trichomes encode transcription factors. For comparison, in *Arabidopsis thaliana* ~6% of the genes in all tissues encode TFs (TAIR10 genome release, [63]). Since JA is essential for establishing indirect defense responses in tomato [34,35] and the induction of terpene synthases in trichomes [16,21], we hypothesized that TFs involved in the regulation of terpene biosynthesis would also be JA-inducible genes. Most of the transcription factors known to be involved in regulation of terpenoid pathways are jasmonate-inducible and of the APETALA2, WRKY or MYC class [25-30,33]. However, in Arabidopsis it was recently shown that two MYC transcription factors (*AtMYC3* and *AtMYC4*), which act additively with *AtMYC2* in the activation of JA responses, are, in contrast to *AtMYC2*, only marginally induced by JA treatment [64]. Based on all the above, the initial selection of transcription factors to be analyzed from our quantitative stem trichome database was limited to TFs of the AP2, WRKY and MYC class that showed a 2-fold or higher induction by JA treatment (2.2-fold was the induction rate of control gene *SIMTS1*; [16], renamed *SITPS8* [21]; Table 5). None of the MYC transcription factors of our database showed induction higher than 2, so for further analysis the closest homolog of *AtMYC2* [32] was selected, as it has been shown to activate the *AtTPS11* and *AtTPS21* promoters [33]. After discarding TFs that were not trichome-specific or did not show highest expression in trichomes, the list was narrowed down to four candidate transcription factors. According to the Q-RT-PCR data however, none of these TFs was significantly induced by JA treatment (Figure 2). Since the numbers of sequence reads of these genes is very low both in the Control and JA samples (Table 6), the fold-induction in the Illuma experiments must have been overestimated.

**SIMYC1 and SIWRKY73 transactivate terpene synthase promoters in planta**

A specific indication of whether any of these TFs are involved in regulating terpene biosynthesis would be the activation of terpene synthase promoters by the transcription factor. In transient activation assays in *N. benthamiana* leaves two of the four selected transcription factors were able to transactivate at least one terpene synthase promoter. SIWRKY73 showed strongest transactivation of the *SITPS5* promoter and in lesser extent of the *SITPS3* and *SITPS7* promoters (Figure 3a). Although SIWRKY73 is expressed highly in roots (Figure 2), SIWRKY73 could not transactivate the promoter of *SITPS8* that is mainly expressed in roots. It could also not transactivate the trichome-specific sesquiterpene synthase *SITPS9* promoter so it is possible that SIWRKY73 can transactivate only monoterpenes synthases or at least not the sesquiterpene synthase tested here (Figure 3a). As shown in Figure 5 *SIWRKY73* and the respective TPSs that it can transactivate are co-expressed in various tissues where the regulation could take place in the plant.

SIMYC1 showed strongest transactivation of *SITPS5* and *SITPS3* and to a lesser extent of *SITPS7* and *SITPS9* but no transactivation of *SITPS8* promoter (Figure 3b), although SIMYC1 is also expressed in the root, albeit not strongly (Figure 2). As shown in Figure 5 **SIMYC1** is expressed (at different levels) in every plant tissue and SIMYC1 is able to activate all the terpene synthase promoters tested except one, so it seems to be a regulator of multiple TPSs, in contrast to *SIEOT1* that is only expressed in the glandular trichomes and can specifically transactivate the *SITPS5* promoter and none of the other TPS promoters tested (Figure 5, [47]). The other two selected TFs (SIWRKY78 and SIWRKY28; Additional file 1: Figure S3) were not able to significantly transactivate any of the tested terpene synthase promoters. However it cannot be excluded that these TFs were not expressed in the transient assay.
One question that arises is, of course, where SlWRKY73 and SlMYC1 bind on these terpene synthase promoters. In the promoter sequence of SlTPS5, SlTPS3 and SlTPS7 [47] there are five, four and one W-boxes (TGAC(C/T)) respectively (PLACE; [65], Additional file 1: Table S2), which could serve as potential binding site(s) for SlWRKY73. Furthermore, SlTPS5 promoter contains two G-box-like elements (CACCA TG instead of the canonical CACGTG), one T/G-box element (AACGTG) and one T/G-box-like element (TACGTG) (Additional file 1: Table S2), which could potentially be the binding site(s) of SlMYC1. The promoter of SlTPS3, with which SlMYC1 interacts less strongly, contains one G-box-like element and one T/G-box element (Additional file 1: Table S2). The SlTPS7 promoter, which SIMYC can also activate, contains one T/G-box element (Additional file 1: Table S2). The SlTPS9 promoter [47] however, does not contain any of these elements, which could indicate the existence of an uncharacterized motif to which SlMYC1 binds. When using the motif search program MEME [66] with all four promoters that SIMYC1 can activate, one 8 bp motif was identified in the plus or minus (for SlTPS9) orientation: CTAGG(T/A) (A/G)G. The validation of a (putative) regulatory element as the binding site for these TFs would require extensive further experimentation. However, since our transactivation assays do not indicate direct binding, the TF-TPS promoter interactions observed in the ATTAs, could take place through an additional protein. To address the issue of which terpene synthases (and possibly other genes as well) these TFs regulate, we are currently starting the more laborious but more conclusive approach of creating stably transformed silenced and overexpressing plants.

SIMYC1 acts synergistically with SIEOT1 in the transactivation of the SlTPS5 promoter
Interestingly, SIEOT1 and SIMYC1 acted synergistically in the transactivation of the SlTPS5 promoter (Figure 4). Combinatorial control of transcriptional regulation is commonly found in plants and other eukaryotes [67]. For example, in abscisic acid (ABA) signaling, the 67 bp promoter region of the dehydration-responsive gene rd22 contains a MYC and a MYB recognition site, where
AtMYC2 and AtMYB2 can bind, respectively. In *Arabidopsis* leaf protoplasts it was shown that these TFs could individually activate transcription of *β-glucuronidase* driven by this 67 bp promoter region of *rd22* and that the transient activation was stronger when AtMYC2 and AtMYB2 were combined [68]. Transgenic plants overexpressing these TFs each showed ABA hypersensitivity but the effect was more profound in plants overexpressing both TFs [69].

Given the fact that *SIMYCI* and *SIEOTI* are not induced by JA (Figure 2, [47]) and yet the proteins can transactivate the JA-inducible *SITPSS* promoter indicates that they could be regulating the steady-state transcription of *SITPSS*. These TFs might however also be involved in the enhanced *SITPSS* expression by interacting with other, inducible TF(s). From the well-studied cases of transcriptional regulation in *Catharanthus roseus* [25-27] and *Arabidopsis* [32,33,64] it has becomes clear that it usually involves a network of TFs. In *Solanum lycopersicum* we are only just starting to unravel the complexity of transcriptional regulation of terpene biosynthesis.

**Methods**

**Hormone treatment and RNA isolation**

Tomato plants (*Solanum lycopersicum* cultivar Money-maker) were grown in soil in a greenhouse with day/night temperatures of 23°C/18°C and a 16/8 h light/dark regime for four weeks. They were then sprayed either with JA solution (1 mM JA; Duchefa, NL, in tap water + 0.05% SilwetL-77; GE Silicones, VA, USA) or with control solution (0.05% SilwetL-77 in tap water). Stem pieces were collected 30 min, 2 h, 8 h and 24 h later for pyrosequencing or 24 h later for expression analyses and trichomes were isolated by shaking the stems in liquid nitrogen. Total RNA was isolated using TRizol (Invitrogen, Paisley, UK) according to the manufacturer’s instructions. Equal amount of trichome RNA from the different time points was pooled creating the control (C) and JA samples. RNA used for pyrosequencing was then purified on a RNeasy Plant column (Qiagen, Valencia, CA, USA).

**Transcriptome database construction**

RNA quality was determined with the Agilent RNA pico chip (Agilent Technologies, Waldbronn, Germany). Synthesis of cDNA was performed using the MessageAmp II aRNA Amplification kit (Applied Biosystems/Ambion, CA, USA) according to manufacturer’s instructions. RNA quality was determined with the Agilent RNA pico chip (Agilent Technologies, Waldbronn, Germany). Shearing and ligation was carried out using TRizol (Invitrogen, Paisley, UK) according to the manufacturer’s instructions. Equal amount of trichome RNA from the different time points was pooled creating the control (C) and JA samples. RNA used for pyrosequencing was then purified on a RNeasy Plant column (Qiagen, Valencia, CA, USA).

The normalization efficiency was determined both on an agarose/EtBr gel (1%) and with an Agilent DNA 7500 chip. The cDNA was purified and concentrated using the Qiaquick PCR purification kit (Qiagen, Valencia, CA, USA). cDNA shearing and FLX Titanium library preparation was carried out using the Roche GS FLX Titanium General Library Preparation Method kit (Roche Diagnostics, Mannheim, Germany) according to the manufacturer’s protocol. The size range of the fragments was determined with an Agilent DNA 1000 chip (Agilent Technologies, Waldbronn, Germany). Exclusion of smaller-sized fragments was performed using the double SPRI method as described in the Roche GS FLX Titanium General Library Preparation protocol (Roche Diagnostics, Mannheim, Germany). End-polishing, small fragment removal, library immobilization, fill-in reaction and single-stranded library isolation was performed using the GS FLX Titanium General Library Preparation Method kit (454 Life Sciences, Roche Diagnostics, Mannheim, Germany) according to manufacturer’s instructions.

**Expression profiling database construction**

Starting from the same total RNA samples (C and JA, see above), mRNA was amplified and purified using the MessageAmp II aRNA Amplification kit (Applied Biosystems/Ambion, CA, USA) according to manufacturer’s instructions. RNA quality was determined with the Agilent RNA pico chip (Agilent Technologies, Waldbronn, Germany). Synthesis of cDNA was performed using the MessageAmp II aRNA Amplification kit (Applied Biosystems/Ambion, CA, USA) according to manufacturer’s instructions with modifications of the adapters to enable sequencing of 3’ cDNA ends. cDNA was purified with the Qiaquick PCR purification kit (Qiagen, Valencia, CA, USA). cDNA quality was determined with the Agilent DNA 7500 chip (Agilent Technologies, Waldbronn, Germany) or on an 1% agarose/EtBr gel. Shearing and ligation was carried out using standard Illumina PE adapters containing a specific sample ID tag. Adapter-ligated cDNA fragments were column purified with the Qiaquick PCR purification kit (Qiagen, Valencia, CA, USA). The size range of the fragments was determined with an Agilent DNA 1000 chip (Agilent Technologies, Waldbronn, Germany). Exclusion of smaller-sized fragments was performed using a single SPRI procedure as described in the Agencourt Ampure PCR Purification protocol (Agencourt Bioscience Corporation, MA, USA). The size range of single-stranded fragments was determined with an Agilent RNA pico 6000 chip (Agilent Technologies, Waldbronn, Germany). Exclusion profiling was performed using the Illumina Genome Analyzer II System (Illumina, USA).

**Databases assembly, EST annotation and homology searches**

The 454 sequencing reads (Control and JA combined) were assembled into contigs *de novo* by Vertis Biotechnologie.
AG, Germany using the CLCbio software [70]. Nucleotide sequences of the contigs were then blasted against the SGN unigenes v2 tomato database (ftp.solgenomics.net/unigene_builds/combined_species_assemblies/tomato_species) for annotation, using a local Eblast tool (E value 1e–9). The GA II reads (Control and JA separately) were mapped to the annotated contigs of the 454 sequencing trichome database by Vertis Biotechnologie AG, Germany.

The resulting contigs were also imported in the bioinformatics tool Blast2GO v.2.5.0 [37] and were blasted against the National Center for Biotechnology Information (NCBI) non-redundant protein database BLASTX (E value 1e–3). Further analyses with this tool included functional annotation by Gene Ontology (GO) terms and Enzyme Commission numbers (EC code), InterPro terms (InterProScan; [71]) and metabolic pathways (Kyoto Encyclopedia of Genes and Genomes, KEGG; [72]).

cDNA synthesis and quantitative-real time-PCR
DNA was removed from RNA with DNase (Ambion, Huntingdon, UK) according to the manufacturer’s instructions and cDNA was synthesized from 1.5 μg RNA using M-MuLV H Reverse Transcriptase (Fermentas, St. Leon-Rot, Germany). For Q-RT-PCR, cDNA equivalent to 100 ng total RNA was used as template in 20 μl volume and reactions were performed in the ABI 7500 Real-Time PCR System (Applied Biosystems) using the Platinum SYBR Green qPCR SuperMix-UDG kit (Invitrogen, Paisley, UK) with the following cycling program: 2 min 50°C, 7 min 95°C, 45 cycles of 15 sec at 95°C and 1 min at 60°C, followed by a melting curve analysis. Primer pairs were tested for amplification kinetics and linearity with a standard cDNA dilution curve and new primers were designed if necessary. Expression levels were normalized using ACTIN (SGN-U579547) mRNA levels. Effects of JA on gene expression were analyzed in three biological replicates by T-test using PASW Statistics 17.0 [73]. The homogeneity of variance was tested by Levene’s test.

Cloning, construct design and ATTAs
TFs SIMYCI (KF430611; sequence of the full-length ORF obtained from the 454 trichome database), SIWRKY28 and SIWRKY73 ([74]; Additional file 1: Figure S4) were cloned between restriction sites NcoI (at the ATG) and SacI (at the 3’ end of the sequence) in front of the Nos terminator in vector pGKi1662 (KeyGene, Wageningen, NL; for a map of the vector see patent nr US2011/013512A1) driven by the CaMV 35S promoter. TF SIWRKY78 ([74]; Additional file 1: Figure S4) was cloned downstream of the CaMV 35S promoter in vector pJVII, a pMON9999-based vector (Monsanto, St. Louis, MO) with a modified multiple cloning site (MCS), between restriction sites XbaI (at the ATG) and BsrGI (at the 3’ end of the sequence). All constructs were verified by sequencing and then the expression cassettes containing 35S promoter, cDNA of interest and nos terminator were transferred to the MCS of the binary vector pBlNplus [75] between HindIII and SmaI restriction sites. The final constructs were transformed to Agrobacterium tumefaciens GV3101 (pMP90). The promoter:GUS constructs used in the transient transactivation assay have been described elsewhere [47]. The A. tumefaciens transient transactivation assay (ATTA) was performed as described in Spyropoulou et al., [47].

Additional file

**Additional file 1: Figure S1.** Gene ontology (GO) and enzyme classifications (EC) for S. lycopersicum stem trichome transcriptome at level 2 (a) Cellular component GO terms, (b) biological process GO terms, (c) molecular function GO terms and (d) general EC terms. Figure S2. Transactivation of terpene synthase promoters by SS5RFP in N. benthamiana leaves. Letters indicate significant differences (n = 4, ANOVA, P < 0.05 according to Tuckey’s B posthoc test). The normalized GUS activity of the STIP53, 7, 8 and 9 reporter constructs with the RFP effector construct is not significantly higher from the STIP53 reporter construct with the RFP effector construct, indicating that any relevant activation of an effector construct (in Figures 3, 4, and Additional file 1: Figure S3) must be significantly higher than that of the STIP53pGUS reporter-35SRFP effector combination.

**Figure S3.** Transactivation of terpene synthase promoters by SIRKYY28 and SIRKYY28 in N. benthamiana leaves. Letters indicate significant differences (n = 3, ANOVA, P < 0.05 according to Tuckey’s B posthoc test). Representative results from two experiments are shown. The normalized GUS activity of the S35/SIRKYY78 effector-SITP53pGUS reporter construct combination was only marginally higher than that of the negative control (35SRFP effector-SITP53pGUS reporter constructs) and was not further investigated.

**Figure S4.** Nucleotide sequence of transcription factors SIRKYY78 (Soyc07g035250.2.1), SIRKYY28 (Soyc12g011200.1.1), SIRKYY73 (Soyc03g13120.2.1) and SIMYCI (KF430611)). The predicted coding sequences are in capital letters, 5’ and 3’ UTRs are in small letter type. Start and stop codons are in bold. Table S1. KEGG pathways found in the S. lycopersicum stem trichome transcriptome. Table S2. Selected regulatory motifs in the sequence of STIP53, 3 and 7 promoters analyzed by PLACE [65]. Table S3. List of primers used.

**Abbreviations**

TPS: Terpene synthase; TF: Transcription factor; EOT1: Expression of terpenoids 1; SGN: Solanaceae genomics network; JA: Jasmonic acid.

**Competing interests**

The authors declare no competing interest. Regulation of terpenoid production through the use of the EOT1 gene and its application for enhanced insect resistance are protected by patents and/or patent applications owned by Keygene N.V. (WO2012169893).

**Authors’ contributions**

EAS performed the experiments and wrote the manuscript. RCS coordinated the research and corrected the manuscript. MAH revised the manuscript. All authors read and approved the final manuscript.

**Acknowledgments**

We thank Raymond Hulzink (KeyGene, Wageningen, NL) for the preparation of the transcriptome and expression profiling libraries, Carlos A. Villarroel Figueroa for help with the bioinformatics, Juan M. Alba Cano for help with the statistical analyses and Harold Lemeris, Ludek Tikovsky and Thijs Hendrix for the excellent care of the plants. This work was supported by grant 1C0029P of the Technological Top Institute-Green Genetics with KeyGene as project leader.
### References


105. PLACE. http://www.dna.aflrc.go.jp/PLACE/


Cite this article as: Spyropoulou et al.: RNA sequencing on Solanum lycopersicum trichomes identifies transcription factors that activate terpenne synthase promoters. BMC Genomics 2014 15:402.

Submit your next manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit