Year 1 Executive Summary, FIA-FSP project Y092092

Development of molecular markers to aid in the identification of western redcedar (*Thuja plicata*) populations that are resistant to deer browsing and heartwood rot fungi

Project start date: 1 April 2007
Length of project: year 1, multi-year

Submitted by Jim Mattsson, Project leader, on April 24th 2008

Contact information:
Jim Mattsson, PhD
Assistant professor
Department of Biological Sciences
Simon Fraser University
8888 University Drive
Burnaby BC V5A 1S6
Canada
phone: 778-782-4291
jmattsso@sfu.ca

Project purpose and management implications

Based on total logs harvested, western redcedar (*Thuja Plicata*; hereafter abbreviated as WRC) ranks as the fifth most important species in BC. WRC wood is highly valued for its high dimension stability and its natural durability and is used in many exterior applications. The utilization of WRC is not without problems however. In BC, reforestation with WRC is expensive and inefficient due to extensive herbivory by ungulates, i.e. deer and elk. Currently, the industry spends up to $6 per tree in deer/elk areas, which corresponds to most of the BC coast and part of the interior, in order to establish WRC. Although regarded as highly durable, second growth WRC heartwood succumbs to early rot, especially in the interior cedar/hemlock biogeoclimatic zone, where complete trees are culled. However, substantial tree-to-tree variation among second-growth trees with respect to both ungulate herbivory and heartwood rot resistance has been noted and the BC Ministry of Forests (MoF) is currently breeding for browsing resistance and intends to breed for heartwood rot resistance. Here we are proposing research to develop markers that correlate with browsing resistance and heartwood rot resistance. The intended use of these markers is to substantially shorten the time and cost for each breeding cycle with respect to browsing resistance and, for the first time, enable efficient breeding for rot resistance in WRC. The deployment of such markers could have considerable short-term benefits for reforestation with WRC and invaluable long-term benefit with regard to the quality and reputation of WRC heartwood from BC.
Methodology overview

There is a strong positive correlation between monoterpenoid content in WRC and resistance to ungulate browsing (Vourc’h et al., 2002 a,b). It is known that induced synthesis of monoterpenoids following attack by stem-boring insects is accompanied by a significant up-regulation of monoterpen synthase (mono-TPS) gene expression (Byun-McKay et al., 2003; Miller et al., 2005). Thus, it is possible to obtain an indirect measure of monoterpenoid content and thereby resistance by quantifying the expression of mono-TPS genes. One of the classes of monoterpenoids implicated in plant defense are the tropolones (Reviewed in Zhao, 2007). Tropolones are natural fungicides found in the heartwood of mature trees where they are involved in WRC resistance to fungal degradation DeBell et al., 1999; Haluk et al., 2000, 2003). An early screening tool would be invaluable in rapidly identifying durable, rot-resistant trees for reforestation. Previous research has shown that there is good correlation between the production of transcripts (mRNAs) from genes encoding monoterpen synthases (mono-TPS) and production of the corresponding monoterpenes (Reviewed by Keeling and Bohlmann, 2006). Our initial approach is to clone the most abundant mono-TPS mRNAs from trees producing high levels of monoterpenes. The isolated mono-TPS cDNAs will be used to correlate expression levels with monoterpenoid levels so as to identify potential expression markers for predicting ungulate resistance.

Results

Plant material

More than 700 provenance clones with known monoterpenoid profiles have been transferred from BC MoF to the SFU greenhouse and have been used in the experiments described below. In addition, material has been cloned from trees with high and low levels of tropolones in adult heartwood and established as calli cultures at SFU.

Identification of candidate genes

Presently the genome of WRC is virtually unknown with only a handful of EST sequences in Genbank. Under this study, we generated more than 15 cDNA libraries. The libraries stemmed from various treatments such as methyl jasmonate, fungal extract, and other stresses to induce genes involved in terpene and tropolone production, in whole seedlings and foliage and cambium of adult trees. The cDNA libraries were pooled and subjected to a transcriptome sequencing by novel next generation pyrosequencing technology provided by Roche/454 Lifesciences (FLX machine). This sequencing generated a total of 48.6 million bases of sequences from 243,101 reads of an average read length of 199.8 bases. Other than inclusion of more than 50% of mRNA from non-green cambium tissues to avoid excess amount of photosynthesis-related sequences, the material was not normalized to enrich for rare sequences. Nevertheless, the data was assembled into 16,076 unique contigs, with an average 7-fold sequencing depth. Among these sequences, we have identified 14 contigs corresponding to putative TPS-encoding genes, with 12 of them summarized in Table 1. The length of contigs varies between 1047 and 145 basepairs (bp) and generally correlate with the number of sequence reads that each contig consists of. Two of the contigs consists of a much larger than average number of reads, which is interesting as it provides an indication that the corresponding transcript may be more abundant than transcripts corresponding to the other contigs. Seven of the 12 contigs match the same
Limonene/borneol synthase from *Chamaecyparis obtusa*. This similarity cannot be taken for evidence though that they all encode Limonene/borneol synthases. Instead, since *Chamaecyparis obtusa* is by far the most closely related species with identified TPS genes, the high similarity may reflect strong evolutionary relationship between these species rather than strong functional relationship between these proteins. The degree of similarity as indicated by the e-value provides good evidence that the identified sequences correspond to transcripts for terpene synthases. The e-values varies from relatively low for TpTPS-L12 indicating distantly related sequences, to extremely high for TpTPS-7, that if they were from the same species, could be taken as evidence for allelic variants of the same gene in the scored region. The degree of sequence similarity in a region of overlap of the deduced amino acid sequences is illustrated in Figure 1. The position of 12 of the translated sequences in relation to typical mono-, sesqui- and di-terpene synthases is shown in Fig. 1. This figure shows that (1) the contigs do not cover complete open reading frames, (2) that the majority are biased towards the C-terminus of proteins or the 3’ end of transcripts. The bias towards the 3’ ends is not surprising considering that the first strand cDNA synthesis was primed with oligo (dT) matching the 3’ poly (A) tail of mRNA molecules. Nevertheless, 9 of the contigs overlap and have considerable differences in sequence, providing evidence that we have identified at least 9 different TPS-encoding genes. In the dataset generated by the large-scale pyrosequencing of WRC cDNAs we have identified additional genes of interest. One of the few, if not the only, gene with a know function in the conversion of sabinene into thujone is a menthol dehydrogenase from sage (Dehal and Croteau, 1987). This is of particular interest in this project as thujones are the most abundant monoterpenoids in WRC. Similarity search using BLASTn identified one WRC sequence with substantial similarity to the sage menthol dehydrogenase. An alignment of deduced amino acid sequences show regions of particularly high similarity (Fig. 3), providing evidence that the similarity based on nucleotide sequence translates into relevant homology at the amino acid level. While the enzymatic functions that are responsible for the conversion of monoterpenes into tropolones are unknown, the molecular structure of tropolones suggests that cytochrome P450 proteins play important roles in this process. We have identified a large number of contigs corresponding to putative cytochrome P450-encoding genes in our 454 EST population. However, since cytochrome P450s have a vast number of activities, additional experimental support is required before we can link any of them to tropolone biosynthesis.

**Assessment of Tp-TPS gene expression in clones with established variation in monoterpenic content**

In line with the goal of this project we have initiated studies to identify potential positive correlations between the expression of individual m Tp-TPS genes and the actual levels of individual monoterpenes in clones with established variation in monoterpenic content. To date, we have identified a correlation between the expression of Tp-TPS-L7 and the levels of β-thujone in a limited number of provenance clones (Fig. 4). The correlations hold particularly well for levels of gaseous β-thujone quantified at the time of extraction of RNA. Experiments to test the correlations in enlarged provenance populations are under way.
Table 1. Summary of 12 of the identified contigs corresponding to putative *Thuja plicata* terpene synthase (Tp-TPS) encoding genes. The “reads” column show the number of 454 sequence reads that each contig consists of. The length of contigs is shown in basepairs, followed by most similar match in GenBank Blast searches, and obtained bit scores and e-values.

<table>
<thead>
<tr>
<th>name</th>
<th>Reads</th>
<th>Length (bp)</th>
<th>Most Similar Match</th>
<th>Bit Score</th>
<th>E- Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TpTPS-L1</td>
<td>12</td>
<td>985</td>
<td><em>Abies grandis</em> (-)-4S-limonene synthase</td>
<td>300</td>
<td>7.00E-78</td>
</tr>
<tr>
<td>TpTPS-L2</td>
<td>32</td>
<td>794</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>235</td>
<td>2.00E-58</td>
</tr>
<tr>
<td>TpTPS-L3</td>
<td>2</td>
<td>229</td>
<td><em>Pseudotsuga menziesii</em> (E)-beta-farnesene synthase</td>
<td>78.8</td>
<td>4.00E-11</td>
</tr>
<tr>
<td>TpTPS-L4</td>
<td>3</td>
<td>445</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>214</td>
<td>8.00E-52</td>
</tr>
<tr>
<td>TpTPS-L5</td>
<td>5</td>
<td>440</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>192</td>
<td>3.00E-45</td>
</tr>
<tr>
<td>TpTPS-L6</td>
<td>8</td>
<td>488</td>
<td><em>Taxus canadensis</em> taxa-4(5),11(12)-diene synthase</td>
<td>70</td>
<td>2.00E-08</td>
</tr>
<tr>
<td>TpTPS-L7</td>
<td>26</td>
<td>1047</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>383</td>
<td>7.0E-103</td>
</tr>
<tr>
<td>TpTPS-L8</td>
<td>1</td>
<td>201</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>147</td>
<td>9.00E-32</td>
</tr>
<tr>
<td>TpTPS-L9</td>
<td>1</td>
<td>269</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>224</td>
<td>4.00E-55</td>
</tr>
<tr>
<td>TpTPS-L10</td>
<td>2</td>
<td>224</td>
<td><em>Picea abies</em> E,E-alpha-farnesene synthase</td>
<td>46.4</td>
<td>2.40E-01</td>
</tr>
<tr>
<td>TpTPS-L11</td>
<td>3</td>
<td>145</td>
<td><em>Chamaecyparis obtusa</em> Limonene/borneol synthase</td>
<td>123</td>
<td>1.00E-24</td>
</tr>
<tr>
<td>TpTPS-L12</td>
<td>2</td>
<td>205</td>
<td><em>Abies grandis</em> myrcene synthase</td>
<td>50</td>
<td>1.90E-02</td>
</tr>
</tbody>
</table>

Figure 1. Alignment of the deduced amino acid sequences of seven mono-TPS from other conifers and five putative mono-TPS EST’s from Western redcedar (*Thuja plicata*). All EST’s from Western redcedar are labeled TpTPS1-5. Alignment was conducted using Clustalx 1.83. AgTPSPhe is (-)-β phellandrene synthase from *Abies grandis*; AgTPSCam is (-)-camphene synthase from *Abies grandis*; CoLim-Bor is limonene/borneol synthase from *Chamaecyparis obtusa*; PaTPS-Myr is myrcene synthase from *Picea abies*; PsTPSPin is (-)-pinene synthase from *Picea sitchensis*; PsTPSLin is (-)-linalool synthase from *Picea sitchensis*; PtTPSTer is α-terpineol synthase from *Pinus taeda*. Black shading represents a >90% consensus, dark grey shading represents a >60% consensus and light grey shading represents >30% consensus.
**Figure 2.** Illustration showing the position of identified contigs (TptPS-L1-12) in relation to typical terpene synthases. The domain structure on the top of the figure is adapted from Keeling and Bohlmann (2007).
Contig8817: ----------------------------------------: -
Menthol : MADTFTQRYALVTGANKGIGFEIRCQLASKGMKVLASRNKEKRGIEARERLLKESRISDDDDVFHQLVDAPAS : 75

*  80  *  100  *  120  *  140  *
Contig8817: ------------------------KLDILVNNAGIMLVPIVVLKADADPRQSTS-----------------YKVEKSYVNMNEQEC- : 53
Menthol : AVAVAHFIETKPIRLDLVNNAGTFTVAEGDLVYQVCLEANAAAQGQAHPFHPGGTRGTTILTSPRCECE : 150

  160  *  180  *  200  *  220  *
Contig8817: ---------NYIGSKTLGALLPLKAKAGDGGRI-------------------------------: 80
Menthol : TVYVCTKFGFEPPLKXSPTIVNVSSTFSTLLLQPNEMAKGVFSSNLNEGKVEELHEFLKDIFGKLOQ : 225

  *  240  *  260  *  280  *  300  *
Contig8817: ----------------------------------------: -
Menthol : NHWPPNFAYKVSKAAVNAYTRI1ARKYFSFCINSVCPGVRDTICYNLGVSEAEAGAEAPVKLALLFDGPSSGS : 300

*  311  *
Contig8817: ----------------------------------------: -
Menthol : FFSREEALSLSY : 311

**Figure 3.** Alignment of deduced amino acid sequence between menthol dehydrogenase from sage (Dehal and Croteau, 1987; labeled Menthol) and the WRC contig 8817.
Figure 4. Results from correlative studies between Tp-TPS gene expression and actual β-thujone levels in different provenance clones. The expression levels were quantified using quantitative Reverse Transcription Polymerase Chain Reaction (Q-RT-PCR). Headspace (gaseous) levels of thujone was measured by R. Gries, at SFU. Extracted levels of thujone was provided by J. Russell, BC MoF.
Cited references


