**Functional characterization of a reproductive tissue specific promoter from Eucalyptus camaldulensis**

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>gen-2018-0106.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>29-Sep-2018</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Unnikrishnan, Boby; ITC Life Sciences and Technology Centre, Agrisciences Shankaranarayana, Gurumurthy; ITC Life Sciences and Technology Centre, Agrisciences</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Promoter, Pollen, Eucalyptus, SWEET gene, Transformation</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue?:</td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>
Functional characterization of a reproductive tissue specific promoter from *Eucalyptus camaldulensis*

Vattekkattu Unnikrishnan, Boby* and Demlapura Shankaranarayana, Gurumurthy

Authors and their affiliation
1. Boby V Unnikrishnan, ITC, LSTC, Bangalore
2. Gurumurthy DS, ITC, LSTC, Bangalore

Address of corresponding author
Boby V Unnikrishnan
Research Scientist
ITC Life Sciences and Technology Centre,
No.3, 1st Main, Peenya Industrial Area
Phase 1, Bangalore, 560058
INDIA
Phone: 919164753969
E-mail: boby.unnikrishnan@itc.in
Abstract

SWEET proteins are essential for the maintenance of nectar production, seed and pollen development in plants. Sequence similarity search in *Eucalyptus* genome resulted in 52 putative genes from SWEET gene family. Expression of two selected SWEET family genes namely *EcSWEET2* and *EcSWEET5* were analyzed in vegetative and reproductive tissues of *Eucalyptus camaldulensis*. Expression of *EcSWEET5* was specific to male reproductive tissues and transcripts were detected only at certain stages of flower development. Tobacco Rattle Virus-mediated suppression of *EcSWEET5* resulted in significant reduction in pollen germination percentage in *Nicotiana benthamiana* with no adverse effect on vegetative growth. Promoter sequence of one kb upstream to start codon of *EcSWEET5* contained many elements suggestive of pollen specificity of the promoter. Pollen specificity of *EcSWEET5* promoter was confirmed in transgenic tobacco lines harboring *GUS* gene with *EcSWEET5* promoter. *GUS* expression was limited to pollen alone in transgenic tobacco as evidenced by histochemical staining. Expression of cytotoxic gene, *barnase* under *EcSWEET5* promoter showed pollen ablation in transgenic tobacco with normal vegetative growth.

Key words: Promoter, Pollen, Eucalyptus, Expression, Transformation, SWEET gene
Introduction
Transgene flow from genetically engineered (GE) to non-GE plants and wild populations is mediated by dispersal of pollen and or seeds. Trees have long life cycles and produce copious amounts of pollen during their reproductive phase. Many of the commercially deployed tree species are highly cross pollinated. There is a potential for wide dispersal of pollen from GE trees and the ecological effects of transgenes would be difficult to predict in wild ecosystems. Therefore, escape and fitness of transgenes from GE trees and their effects on tree populations acquiring transgenes needs to be addressed before deployment of GE trees (Ahuja 2014).

There are several genetic containment strategies that have been discussed for forest trees which include fitness reduction, ablation, transgene excision, and floral gene suppression at RNA or protein levels (Brunner et al. 2007; Vining et al. 2012). Among these methods, pollen ablation is mostly studied and in this process, a cytotoxic gene would be expressed only in male reproductive tissues under the control of a tightly regulated promoter to render the pollen non-viable. Pollen specific promoters are used for pollen ablation in genetically engineered trees to address horizontal gene flow from transgenic trees (Elorriaga et al. 2014).

Pollen specific promoters also provide an opportunity to understand the regulatory mechanism in pollen development and to date many pollen specific promoters have been isolated from various plants (Chen et al. 2012a). Promoter of TA29 gene that is expressed only in tapetum cells was first employed to drive barnase to create male sterile tobacco and Brassica napus (Mariani et al. 1990). However, expression of cytotoxic gene such as barnase in non-target organs due to the non-specific basal activities of employed promoters often inhibited plant growth though sterility was achieved (Wei et al. 2007). Many researchers have reported that floral promoters were not expressed exclusively in flowers and low levels of unintended expression of cytotoxic gene under those promoters led to retarded growth (Rottmann et al. 2000; Skinner et al. 2000). Lannenpaa et al. (2005) reported that the birch tree transgenic lines with BpFULL promoter and barnase gene produced many bushy type phenotypes and suggested that phenotypes could be due to expression of barnase in early stages of shoot development. So, basal activity of promoter in non-target tissues could lead to negative effects on vegetative growth, especially in plant species with long regeneration times and long life span as in case of woody trees. Zhang et al. (2012) reported the need for the ablation efficiency of pollen ablation cassette should be high and stable without any negative effects on growth for commercial applications. Huang et al.
(2016) generated sterile plants in Arabidopsis and tobacco by expressing barnase fusion gene driven by solo dancers (SDS) promoter without affecting plant growth or modifying flower structure.

Eucalypts are one of the widely planted hardwood crops in the world and a number of transgenic eucalyptus trees have been generated for growth, biotic and abiotic stress tolerance (Girijashankar 2011). The amount of sucrose available for transport from source to sink tissues is very crucial for plant development (Lemoine et al. 2013). This process was controlled or mediated by a group of transport proteins for sucrose allocation across the membrane which are presently known as SWEET (Sugars Will Eventually be Exported Transporters) family of genes (Chen et. al. 2010). SWEET proteins are essential for plant nectar production and pollen development (Chen et al. 2012b). One of the SWEET family member, AtSWEET5 (At5g62850) was found to be expressed in flower tissues and the promoter was active in later stages of pollen development in arabidopsis (Engel et al. 2005). We analyzed expression of two members of SWEET family genes in E. camaldulensis namely EcSWEET2 and EcSWEET5 in different tissues for tissue specific expression patterns. The specificity of promoter of EcSWEET5 was studied in transgenic tobacco by GUS expression and basal levels of expression in unintended tissues if any were studied by barnase gene under EcSWEET5 promoter. This is to our knowledge, the first report of characterization of a pollen specific promoter from Eucalyptus camaldulensis.

Materials and methods

Cloning of promoter constructs and sequence analysis

The promoter region of EcSWEET5 (~1kb) was amplified by gene specific primers as given in Table 1. The PCR product was cloned in to pJET 1.2 (MBI Fermentas) followed by transformation in to DH5α E.coli cells (New England Biolabs) by heat shock method. The transformed cells were plated on LB agar with ampicillin (100 ug/ml) and the colonies were confirmed by PCR. The recombinant plasmid was subjected to sequencing reaction with vector specific primers using ABI 3730 at Eurofin India Pvt Ltd (Bangalore, India). The promoter sequence from eucalyptus was analysed online using PLACE (Plant Cis-acting Regulatory DNA Elements) database (http://www.dna.affrc.go.jp/PLACE) for the motifs and possible functions of the promoter region.
To prepare EcSWEET5p::barnase construct barnase gene was amplified from construct pPZP200bsnp::TA29-bn. The PCR product was cloned in to E. coli by infusion cloning method and the recombinant clones were confirmed by PCR. Agrobacterium strains LBA 4404 were transformed with recombinant plasmids by freeze thaw method (Holsters et al 1978). The cells were plated on YM agar with streptomycin (100 µg/ml) and kanamycin (50 µg/ml). After two days, positive colonies were picked up and confirmed by PCR.

Sequence comparisons and phylogenetic analysis

Multiple sequence alignments were carried on using ClustalW2 (available in http://www.ebi.ac.uk/Tools/msa/clustalw2/) with default setting. SWEET homologous gene sequences from *Eucalyptus globulus* and *A. thaliana* were retrieved in Phytozome Database (http://www.phytozome.net/). A total of 52 putative SWEET family genes from *E. grandis* and 17 SWEET genes from arabidopsis were aligned using ClustalW2. A phylogenetic tree was constructed by the MEGA 7.0 using the neighbor-joining method with 1,000 replicates for bootstrap analysis.

Expression analysis of *EcSWEET2* and *EcSWEET5* in eucalyptus

Different stages of flower development (S1 to S6) were collected for expression analysis. The features of each stage collected are as follows, Stage 1: small bud (5x2mm) operculum and receptacle are green in color and continuous; Stage 2: medium size bud (6x 3mm) operculum and receptacle are differentiated; Stage 3: normal bud (8x5mm) operculum is loose and about to fall; Stage 4: Opened flower, operculum is fallen, stamens and carpel developed and intact; Stage 5: Stamens are fallen and carpel is intact on receptacle; Stage 6: Partially dried receptacle post fertilization.

Total RNA from eucalyptus tissues were isolated by RNA easy kit (Fermentas) as per manufacture’s instruction. RNA samples were treated with RNase- free DNase I (MBI Fermentas) at the concentration of 1 U/µg of sample to remove genomic DNA. Total RNA was quantified by spectrophotometer ND 8000 (Nano Drop Technologies Inc, DE USA) and equal quantity of RNA was taken for cDNA synthesis for all the samples. cDNA was synthesized from 1 µg of RNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) with random hexamer primers according to manufacturer’s instruction.
Diluted cDNA template was used as template for PCR reactions and semi-quantitative RT-PCR was performed as described before (Liu et al. 2002) and 18S gene was used as an internal control.

**Histochemical GUS staining**
Tissues were transferred to fixing solution (10 mM MES (pH 5.6), 300 mM mannitol and 0.3% formaldehyde) for 45 min at room temperature. After three washes with distilled water, tissues were transferred to staining solution (500 mM sodium phosphate buffer, 0.1 mM potassium ferrocyanide, 0.1 mM potassium ferricyanide and X-gluc (1 mg/ml)). The samples were incubated at 37°C for up to 24 h until a blue stain developed. The tissue clearing was performed with 70% ethanol for 1-3 h (Yu et al. 2005).

**Tobacco transformation with promoter construct**
The transformation of tobacco with various constructs was carried out as described by Horsch et al. (1985) with minor modifications. Briefly, leaf discs from a month old in-vitro grown tobacco plants were incubated with agrobacterium cell suspension at 28°C with gentle shaking for 30 min. After incubation, leaf discs were transferred to MS plates with BAP (2 mg/L) and NAA (0.2 mg/L) and kept at 25°C for four days without antibiotics. Excess growth of agrobacteria was removed by washing with cefotaxime (250 ug/ml) and explants were transferred to MS media with kanamycin (100 ug/ml) and cefotaxime (250 ug/ml). Full grown shoots were transferred to 0.5X MS media with kanamycin (100 ug/ml) for rooting. Transgenic lines were confirmed by PCR with transgene specific primers. Confirmed lines were grown in glass house conditions till flowering.

**Cloning and agro infiltration of TRV constructs**
The construction of TRV based vectors were carried out by the protocol suggested by Senthil kumar and Mysore (2014). A fragment of *EcSWEET5* (~150bp) and *PDS* were amplified from eucalyptus genomic DNA by gene specific primers. A total of 50 ng of purified PCR product was treated with T4 DNA polymerase (New England Biolabs) in 1X reaction buffer containing 5 mM dATP and dithiothreitol at 22°C for 30 min followed by 20 min of inactivation of T4 DNA polymerase at 70°C. The TRV2-LIC vector was then digested with PstI and similarly treated with T4 DNA polymerase with dTTP. A total of 50 ng of treated PCR product and TRV2-LIC
vector were mixed and incubated at 65°C for 2 min and then 22°C for 10 min. The recombinants were elected on LB plate with kanamycin (50 ug/ml). Agrobacterium strains were made by transforming GV2260 with recombinant plasmids by freeze thaw method. Overnight agrobacterium cultures (2 ml) were grown at 28°C in LB broth containing kanamycin (50 ug/ml) and rifampicin (25 ug/ml). Cultures were spun down and cells were resuspended in infiltration medium (10 mM MES, 10 mM MgCl₂, 200 mM acetosyringone), adjusted to OD600 of 0.6 and incubated at room temperature for 3 h. *N. benthamiana* infiltration was performed as previously reported (Liu et al. 2002) and plants were observed for 8 weeks in glass house maintained with 22°C and 70% relative humidity. Each construct had minimum of five replicates for growth measurements.

**Pollen germination and viability**

For germination studies, fresh pollen grains were collected from plants grown in glass house. The pollen grains were immersed with BK medium (0.01% H₃BO₄, 0.02% Ca (NO₃)₂, 0.01% KNO₃, 0.02% MgSO₄.7H₂O, and 15% sucrose, pH 5.9). The tubes were incubated at 28°C under dark condition for 2–12 h (Brewbaker and Kwack, 1963). Pollen grains were stained with Lugol staining solution as described by Mulugeta et al. (1994).

**Data Archiving Statement**

The present work is a part of work covered under patent application WO2015056070A1 titled “Tissue specific plant promoter and uses thereof” and the sequences are disclosed in the same.

**Results**

**Phylogenetic analysis of SWEET genes from Eucalyptus**

Search for SWEET family of genes in *E. grandis* genome revealed 52 putative members. The protein sequences of all eucalyptus orthologs were aligned with SWEET genes from arabidopsis. The sequences of SWEET proteins from arabidopsis and eucalyptus are presented in supplementary data (Table S1). There were four distinct clades for SWEET genes from eucalyptus on comparison with arabidopsis homologues. In clade IV, *AtSWEET5* (*At5g62850*), one of the characterized SWEET gene for pollen specific expression pattern in arabidopsis was found to be in same clade with Eucgr. B00360.1, Eucgr.C01371.1, Eucgr. B00362.1 and Eucgr.
Among the orthologs in eucalyptus, Eucgr. B00360.1 was found to be nearest to the AtSWEET5 (Fig. 1). Hence Eucgr. B00360.1 was selected for further characterization and ortholog from E. camaldulensis was referred as EcSWEET5. Another ortholog, Eucgr. H04549 in clade II was also selected for studying the expression pattern in different tissues of E. camaldulensis and was named as EcSWEET2.

**Tissue specific expression of EcSWEET genes in E. camaldulensis**

Expression of transcripts, EcSWEET2 and EcSWEET5 were carried out in various vegetative and reproductive tissues of Eucalyptus camaldulensis. Expression of EcSWEET5 was limited to flower alone when various tissue samples like young leaf, mature leaf, xylem, roots and flowers were analyzed. EcSWEET2 transcripts were detected in leaves in addition to flower samples (Fig. 2a). As the EcSWEET5 transcript was specific to flower tissue, subsequent analysis of EcSWEET5 was continued.

Samples were collected from different stages of flower development starting from young bud to partially dried receptacle stage (S1 to S6) as given in Fig. 2b. EcSWEET5 transcript was found to be accumulating from S2 onwards and reached maximum at S4 followed by a drop in the expression levels. High transcript abundance was observed for EcSWEET5 in S3 and S4 stages of flower development. Three developmental stages of flowers namely S2, S3 and S4 showed highest accumulation of EcSWEET5 transcript. We further analyzed EcSWEET5 transcripts in stamen and pistil tissues separately. The EcSWEET5 transcript was detected in S2, S3 and S4 stages of male reproductive tissues while no expression was observed in female organs (Fig. 3).

**Suppression of EcSWEET5 gene in N. benthamiana.**

We studied the function of EcSWEET5 gene using N. benthamiana as model plant through Tobacco Rattle Virus (TRV) mediated suppression. A fragment of EcSWEET5 gene was selected from +152 to +298 (relative to translation start site) with in coding region of EcSWEET5. The fragment was searched against N. benthamiana database to confirm homologue region in N. benthamiana. The alignment between EcSWEET5 and NbSWEET5 mRNA sequences are presented in Fig 4. Similar exercise was carried out for phytene destaturase gene (PDS) to serve as control. The fragment of EcSWEET5 gene was cloned in to pTRV2 vector to generate pTRV2::EcSWEET5 and gene insert was confirmed by sequencing. pTRV2::EcSWEET5
construct in agrobacterium strain GV2260 was inoculated to one month old *N. benthamiana* plants to evaluate the phenotypes. Plants inoculated with pTRV2::PDS construct served as control. Minimum of five plants were evaluated from EcSWEET5 and EcPDS suppressed lines in *N. benthamiana*. The plants were evaluated for growth in terms of height, stem and root dry weight. The dry weights of stem and roots from EcSWEET5 suppressed lines did not show significant variation compared to plants suppressed for EcPDS or untransformed control plants. TRV mediated suppression of EcSWEET5 indicated that suppression of EcSWEET5 does not lead to any vegetative abnormalities (Fig. 5a)

At the time of flowering pollen were collected from different lines of *N. benthamiana* suppressed for EcSWEET5 and were evaluated for pollen germination and viability. Pollen germination indicated that a significant reduction in pollen germination percentage in EcSWEET5 suppressed lines on comparison with pollen from PDS suppressed lines or untransformed control lines (Fig. 5b). There was approximately 50 % reduction in germination percentage in EcSWEET5 suppressed lines on comparison with untransformed control plants. Pollen viability staining by iodine blue indicated a reduced pollen viability in EcSWEET5 suppressed lines

**Sequence analysis of EcSWEET5 promoter from *E. camaldulensis***

The expression data of EcSWEET5 indicated tissue specificity of transcripts in male reproductive tissues. The promoter of EcSWEET5 was cloned from *E. camaldulensis* using a sequence 976 bp upstream to start codon of Eucgr.B00360.1 from *Eucalyptus grandis* v2.0. The cloned region was sequenced and analyzed by SIGNALSCAN in Plant Cis-acting Regulatory DNA Elements (PLACE) database to find motifs and possible functions in the promoter region (http://www.dna.affrc.go.jp/PLACE) (Prestridge 1991; Higo et al. 1999). The list of various elements in promoter region are presented in Table 2. Various putative cis-acting regulatory elements required for pollen-specific expression or involved in enhancing gene expression in pollen were predicted. Three copies of AGAAA, seven copies of GTGA and two copies of GATA elements were found in the EcSWEET5 promoter. Gupta et al. (2007) reported presence of three copies of the AGAAA and seven copies of the GTGA in pollen specific promoter of OSIPK and GATABOX was known to be an element required for tissue specific gene expression (Reyes et al. 2004). Presence of pollen specific elements such as GTGA and
AGAAA in the sequence indicated pollen specificity of the promoter under study. Presence of putative Transcription Start Sites (TSS) and TATA BOX were analyzed through NNPP database and results were presented in Fig. 6. Though there were two TSS detected (Positions -112 and -173), TSS at -173 position could be the true TSS as it was 25 bases away from a TATA-box positioned at -197. However this need further studies to confirm the assumption.

**Pollen specific expression of EcSWEET5 promoter in tobacco**

Analysis of promoter sequences showed presence of pollen specific elements in promoter region of EcSWEET5. Construct was made with promoter of EcSWEET5 to drive the expression of GUS in pBI121 background. Tobacco was used as model system to study the specificity of promoter. Plants were regenerated on kanamycin selection medium and were confirmed by PCR by using EcSWEET5 specific and GUS specific junction primers. The plants were allowed to grow till flowering in glass house. Multiple tissues from transgenic tobacco such as leaf, stem, root, shoot tips and flower tissues were collected at the time of flowering. When tissues were analyzed for promoter activity by histochemical staining for GUS no activity was noticed in stem, root, shoot tip, and leaves (Data not shown). However, pollen grains were stained in transgenic tobacco lines harboring promoter and GUS while pollen in untransformed tobacco did not show any staining. It was also observed that not all pollen grains were stained in transgenic tobacco, indicating a probable stage specific expression during pollen development process (Fig. 7). As it is a time consuming process to generate transgenic in trees, a transient expression of GUS under EcSWEET5 promoter was carried out by infiltrating EcSWEET5p::GUS construct in to anthers of eucalyptus and expression of GUS was noticed in pollen grains indicating the promoter activity in eucalyptus tissues (Data not shown).

**Expression of barnase under EcSWEET5 promoter and pollen ablation in tobacco**

EcSWEET5 promoter was found to be specific to pollen as evidenced by histochemical staining for GUS gene. In order to use the promoter for any genetic containment strategy we need to be sure that the promoter activity of EcSWEET5 is restricted only to reproductive tissues and there is no basal level expression in other tissues. Expression of a cytotoxic gene like barnase under the promoter would help to confirm the tissue specificity of expression. Constructs containing
barnase gene linked to EcSWEET5 promoter were generated in base vector pBI121 and construct were verified by sequencing. Transgenic tobacco lines were generated harboring EcSWEET5p::barnase construct by agrobacterium mediated method. Transgenic tobacco lines were generated on kamamycin selection media (100 µg/ml) and plants were confirmed by PCR with NPTII primers. The presence of barnase was confirmed by barnase gene-specific primers in kanamycin-selected lines. The confirmed lines were grown in glass house till flowering. The plants were monitored for height and number of leaves throughout the growth period and no growth defect was observed in transformed plants on comparison with untransformed control plants. Flower morphology was also comparable among transformed and control plants except for anther features. The anthers in transformed lines were shrunken and no pollen was visible when anther was opened. In case of untransformed tobacco, anthers were stout and bulky and copious amounts of pollen was visible when opened (Fig. 8).

Discussion

SWEETs are an important superfamily of sugar transporters with critical roles in plants, animals and potentially a variety of microorganisms (Xuan et al. 2013). Putative orthologs of SWEET genes from eucalyptus were aligned with members from Arabidopsis and cladogram was generated. There were four different clades namely Clade I, II, III and IV as shown in Fig.1. The fourth clade contained a well-studied SWEET family member from arabidopsis AtSWEET5 (At5g62850) which had shown to be involved in pollen specific expression (Engel et al. 2005). The nearest ortholog in eucalyptus was Eucgr.B00360.1 and hence was used a possible candidate for further studies. There were three more members in the same clade viz Eucgr. CO1371, Eucgr.B00362.1, Eucgr.B00363.1 and hence we analyzed the promoter region (1kb) of all those members for pollen specific elements. Eucgr.B00360.1 contained three copies of AGAAA, six copies of GTGA (Table S2). Promoter region of Eucgr.CO1371 had many elements of AGAAA, GTGA, AAATGA, AGGTCA and GATA which all were known to be involved in pollen specific expression. As the other members also had same elements it suggests the probability of having same promoter functions for the promoters discussed. But this need further experiments.
AtSWEET5 is a member of MtN3 family found to be preferentially expressed in the vegetative

cell of pollen by transcriptomic approaches (Borges et al. 2008). EcSWEET5 and 2 had 62 and

35 % similarity with AtSWEET5 at amino acid level. However, Vining et al. (2015) analyzed the

transcriptomes of early and late floral buds from Eucalyptus grandis on comparison with

vegetative tissues. However, in this study no transcripts homologous to EcSWEET5 were

detected in flower tissues. The floral tissues used in the study were early buds with operculum

intact where in male and female reproductive structures were actively developing, could be the

probable reason why SWEET5 transcripts were not detected in the flower transcriptome of

eucalyptus. This indicated that EcSWEET5 could be expressing in later stages of flower

development and our data clearly indicated the expression of EcSWEET5 from S2 onwards and

reaching the peak at S4. We found that EcSWEET5 transcript was specific to male reproductive

tissues only. At the same time, EcSWEET2 transcript was found to be expressed in other

eucalyptus tissues. The function of EcSWEET5 gene was characterized by TRV mediated

suppression in N. benthamaiana. There were no abnormal phenotypes associated with

EcSWEET5 suppression as evidenced by stem and root dry weight. However, pollen germination

was significantly affected in EcSWEET5 suppressed lines confirming its role in pollen

development. Presence of pollen specific elements and data on EcSWEET5 suppression indicated

nature of promoter and the function of gene. As we were interested in characterization promoter

for pollen ablation purposes, promoter of EcSWEET5 was studied in detail.

Promoter analysis of EcSWEET5 promoter showed presence of elements such as GTGA and

AGAAA indicating pollen specificity of the promoter under study. AGAAA is a regulatory

element responsible for pollen specific activation of tomato LAT52 gene. The AGAAA element

was required for pollen-specific expression with another co-dependent element TCCACCATA in

the LAT52 promoter (Bate and Twell 1998). This particular element had also been found in the

promoter of tomato endo-beta-mannanase gene expressed in late stages of anther development

(Filichkin et al. 2004). The GTGA motif was identified as a quantitative element in pollen-
specific expression of the LAT56, LAT59 and g10 promoters (Twell et al. 1990; Rogers et al.

2001). Though many pollen/anther-specific promoters have these cis-elements, (Gupta et al.

2007; Cook and Thilmony 2012) their functionality is still unknown and hence mutation or gain

of function analyses of these elements is required to understand their functionality (Chen et al.

2012a).
To provide direct evidence of reproductive tissue specificity of expression driven by EcSWEET5 promoter, we generated transgenic tobacco plants expressing the GUS gene under EcSWEET5 promoter. GUS expression was studied in vegetative and reproductive tissues of the transgenic tobacco. There was no GUS expression detected in any of the vegetative tissues and staining was noticed only for pollen grains. Not all pollen grains were stained indicating the stage specific expression of promoter during pollen development as reported in case of promoter of tobacco late pollen gene g10. Although GUS expression pattern indicated specificity of pollen, we were interested in understanding whether expression from promoter is tightly regulated. This information is very important especially when a toxic gene or any excision mechanisms are to be expressed under this promoter. A cytotoxic gene namely barnase was expressed under this promoter in tobacco followed by evaluation of vegetative and reproductive characteristics of the plant. There was no apparent expression of barnase in vegetative tissues as indicated by normal vegetative growth indicating promoter is not active in vegetative tissues. At the same time, anthers were shrunken and pollen ablation was observed in transgenic tobacco confirming expression from EcSWEET5 promoter limited to pollen only. Flowers from the confirmed transgenic plants were bagged to allow self-pollination in glass house. Seeds collected from the plants did not show any germination indicating no seed set in transgenic line expressing EcSWEET5p:barnase construct. The data showed activity of EcSWEET5 promoter is different from PrMC2 (Zhang et al. 2012) where in RNase activity of barnase had to be reduced to achieve non adverse effect on vegetative features in Pine and Eucalyptus.

Genetic engineering to eliminate pollen production to facilitate breeding, hybrid seed production, and reducing out-crossing has been widely studied in various crops (Strauss et al. 1995; Brunner et al. 2007). A major obstacle in deployment of genetically engineered trees is concern over transgene dispersal to the environment (Strauss et al. 2009). There has been long term interest in containing the pollen from transgenic trees as pollen from many trees could travel from hundreds of meters to several kilometers. The gene flow from transgenic to wild relatives is of particular concern for forest trees because they are virtually undomesticated (Difazio et al. 2004). Pollen ablation or use of gene excision mechanisms are used to eliminate pollen production under a tightly regulated reproductive specific promoter. Elorriaga et al. (2014) reported that transgenic poplar transformed with TA29:barnase resulted in slow growth in terms of stem volume and suggested tobacco TA29 promoter had imperfect tissue fidelity especially in taxonomically
distant dicots like Populus. A promoter with tight expression in male reproductive tissues would be an ideal candidate for pollen ablation programmes in forest trees including eucalyptus. Several promoters have been employed to bring pollen ablation in trees which include BpFULL1, PrMC2 and TA29 (Lannenpaa et al. 2005; Zhang et al. 2012; Elorriaga et al. 2014). However, there have been several reports of vegetative abnormalities associated with cytotoxic gene expression under male reproductive specific promoters. Zhang et al. (2012) hypothesized that the strong RNase activity of native barnase, when used for flower ablation, often damaged non targeted tissues due to non-flower-specific activities by many floral promoters. They could achieve pollen ablation by using a modified barnase cassette PrMC2-barnaseH102E where RNase activity of barnase was reduced to avoid any effects on unintended tissues. This basically warrants the importance of having a tissue specific promoter like EcSWEET5 promoter with no detectable basal activity in non-target tissues as evidenced by barnase expression. In conclusion, EcSWEET5 promoter is a candidate promoter to regulate gene expression during the process of pollen development and could be useful for pollen ablation programmes. Overall, the data suggest that the promoter fragment derived from eucalyptus is functional and can drive the expression in tobacco and eucalyptus in a tissue specific manner. Our data showed EcSWEET5p is tightly regulated promoter in pollen grains and could be useful in generating male sterile lines in eucalyptus.

Acknowledgement
The authors are grateful to Dr. Deepak Pental, University of Delhi for providing the source of barnase gene. We are thankful to Mr. Suresh Kumar for generating the cladogram for SWEET proteins.

References

tion is developmentally regulated by multiple stage-specific and co-dependent activator elements. Plant Mol. Biol. 37: 859–869.


https://mc06.manuscriptcentral.com/genome-pubs
Vining, K.J., Romanel, E., Jones, R.C., Klocko, A., Alves-Ferreira, M., Hefer, C.A.,
Amarasinghe, V., Dharmawardhana, P., Naithani, S., Ranik, M., Wesley-Smith, J.,

promoter of the pollen-specific gene NTP303 reveals a novel pollen-specific, and

trial detects incomplete barstar attenuation of vegetative cytotoxicity in Populus trees

Functional role of oligomerization for bacterial and plant SWEET sugar transporter family.


Zhang, C., Norris-Caneda, K.H., Rottmann, W.H., Gulledge, J.E., Chang, S., Kwan, B.Y.,

**Figure captions**

**Fig. 1**

Phylogenetic analysis of SWEET proteins from Arabidopsis and Eucalyptus.
SWEET proteins from Eucalyptus are grouped into Clade I, II, III and IV. Protein sequences from arabidopsis and eucalyptus are available in Supplementary file.

**Fig. 2 a and b**
Expression profile of EcSWEET2 and EcSWEET5 from various tissues of *E. camaldulensis*.

YL-young leaf, ML- Mature leaf, X- xylem, F – flower, R- roots; S1-S6: different stages of flower development in *E. camaldulensis*

**Fig. 3**
Expression of EcSWEET5 in stamen and pistil of *E. camaldulensis*.

S2, S3 and S4 are selected stages of flower development

**Fig. 4**
Alignment between mRNA sequences EcSWEET5 and NbSWEET5

The sequences used for TRV mediated suppression are indicated in box.

**Fig. 5**
a. Effect of EcSWEET5 suppression on stem and root dry weight in *N. benthamiana*

WT-Wild type, Data represent the means ± standard error of minimum three independent replicates of plants infected with TRV:EcPDS and TRV:EcSWEET5 constructs

b. Effect of EcSWEET5 suppression on pollen germination percentage in *N. benthamiana*

n- number of pollen analyzed for germination

**Fig. 6**
DNA sequence of EcSWEET5 promoter.
The translation initiation code is represented by ATG. The predicted putative transcription start site (TSS) and TATA box are boxed and coloured in green. The putative cis elements GTGA, AGAAA and GATA are colored in red.

**Fig. 7**
Histochemical analysis of transgenic tobacco pollen harboring EcSWEET5 promoter

**Fig. 8**
Phenotypes associated with *barnase* expression under EcSWEET5 promoter in transgenic tobacco

a. Transgenic tobacco harboring EcSWEET5p::*barnase* showing normal growth b and c. anthers and opened anther lobe from transgenic plant d and e. anthers and opened anther lobe from wild type tobacco.
<table>
<thead>
<tr>
<th>No</th>
<th>Sequence (5’-3’ direction)</th>
<th>Description</th>
<th>Amplicon (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G GAG AGA GTT GAT GTG AAG</td>
<td>Promoter of EcSWEET5</td>
<td>960</td>
</tr>
<tr>
<td>2</td>
<td>GACAATGCC TAC AAT CGT TC</td>
<td>Infusion primers for EcSWEET5p</td>
<td>960</td>
</tr>
<tr>
<td>3</td>
<td>TGATTACGCAAGCTT G GAG AGA GTT GAT GTG AAG</td>
<td>EcSWEET5 RT-PCR</td>
<td>120</td>
</tr>
<tr>
<td>4</td>
<td>CCGGGGATCTCTCTAGA GACAATGCC TAC AAT CGT TC</td>
<td>EcSWEET2 RT-PCR</td>
<td>130</td>
</tr>
<tr>
<td>5</td>
<td>TCCCCATCCCCGCATCATCACCATAT</td>
<td>NPTII specific</td>
<td>120</td>
</tr>
<tr>
<td>6</td>
<td>AGGGCATGTACTTGGACTTTGT</td>
<td>Cloning of TRV:: NbPDS construct</td>
<td>140</td>
</tr>
<tr>
<td>7</td>
<td>GCTCAACGCCTACTTCTGGACTTACTAT</td>
<td>Cloning of TRV:: EcSWEET5 construct</td>
<td>200</td>
</tr>
<tr>
<td>8</td>
<td>AGGTGACCCGAGTATTTGCAGTTC</td>
<td>Infusion primers for barnase</td>
<td>350</td>
</tr>
</tbody>
</table>

**Table 1**: Primer sequences used in the study
<table>
<thead>
<tr>
<th>No</th>
<th>Factor/Site Name</th>
<th>Signal Sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2SSEEDPROTBANAPA</td>
<td>(+) CAAACAC</td>
<td>Conserved in storage-protein gene promoters and important for high activity of the napA promoter</td>
</tr>
<tr>
<td>2</td>
<td>ABRELATERD1</td>
<td>(+) ACGTG</td>
<td>required for etiolation-induced expression of erd1 (early responsive to dehydration) in Arabidopsis</td>
</tr>
<tr>
<td>3</td>
<td>ABREOSRAB21</td>
<td>(+) ACGTSSSC</td>
<td>ABA responsive element (ABRE) of wheat and rice rab21 genes</td>
</tr>
<tr>
<td>4</td>
<td>ACGTATERD1</td>
<td>(+) ACGT</td>
<td>required for etiolation-induced expression of erd1 (early responsive to dehydration) in Arabidopsis</td>
</tr>
<tr>
<td>5</td>
<td>ARFAT</td>
<td>(-) TGTCTC</td>
<td>ARF (auxin response factor) binding site found in the promoters of primary/early auxin response genes of A.thaliana.</td>
</tr>
<tr>
<td>6</td>
<td>ARR1AT</td>
<td>(+) NGATT</td>
<td>ARR1-binding element found in Arabidopsis; ARR1 is a response regulator</td>
</tr>
<tr>
<td>7</td>
<td>ASF1MOTIFCAMV</td>
<td>(+) TGACG</td>
<td>Involved in transcriptional activation of several genes by auxin and/or salicylic acid</td>
</tr>
<tr>
<td>8</td>
<td>AUXRETGA2GMGH3</td>
<td>(+) TGACGTGGC</td>
<td>putative auxin-responsive element (AUXRE) E1 of soybean GH3 promoter</td>
</tr>
<tr>
<td>9</td>
<td>BOXIIINTPATPB</td>
<td>(+) ATAGAA</td>
<td>Important for the activity of NCII promoter</td>
</tr>
<tr>
<td>10</td>
<td>BOXIIIPCCHS</td>
<td>(+) ACGTGGC</td>
<td>Found in the parsley chs genes, Essential for light regulation.</td>
</tr>
<tr>
<td>11</td>
<td>CAATBOX1</td>
<td>(+) CAAT</td>
<td>CAAT promoter consensus sequence found in legA gene of pea</td>
</tr>
<tr>
<td>12</td>
<td>CANBNNAPA</td>
<td>(+) CNAACAC</td>
<td>Elements found in Embryo- and endosperm-specific transcription of napin (storage protein) gene.</td>
</tr>
<tr>
<td>13</td>
<td>CAREOSREP</td>
<td>(-) CAACCTC</td>
<td>CAREs (CAACTC regulatory elements) found in the promoter region of a cystein proteinase (REP-1) gene in rice</td>
</tr>
<tr>
<td>14</td>
<td>CBFHV</td>
<td>(+) RYCGAC</td>
<td>Binding site of barley CBF1, and CBF2; Also known as dehydration-responsive element (DRE) binding proteins (DREBs)</td>
</tr>
<tr>
<td>15</td>
<td>CIACADIANLELHC</td>
<td>(+) CAANNNNATC</td>
<td>Region necessary for circadian expression of tomato Lhc gene</td>
</tr>
<tr>
<td>16</td>
<td>DPBFCOREDCDC3</td>
<td>(+) ACACNNNG</td>
<td>bZIP transcription factors binding core sequence, embryo-specific expression</td>
</tr>
<tr>
<td>17</td>
<td>DRE1COREZMRAB17</td>
<td>(+) ACCGAGA</td>
<td>DRE1 core found in maize rab17 gene promoter; rab17 is expressed during late embryogenesis</td>
</tr>
<tr>
<td>18</td>
<td>EBOXBNNAPA</td>
<td>(+) CANNTG</td>
<td>E-box of napA storage-protein gene of Brassica napus</td>
</tr>
<tr>
<td>19</td>
<td>GATABOX</td>
<td>(+) GATA</td>
<td>Required for high level, light regulated, and tissue specific expression</td>
</tr>
<tr>
<td></td>
<td>Symbol</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>---</td>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>20</td>
<td>GT1GMSCAM4</td>
<td>(+) GAAAAA</td>
<td>Plays a role in pathogen- and salt-induced SCaM-4 gene expression</td>
</tr>
<tr>
<td>21</td>
<td>GTGANTG10</td>
<td>(+) GTGA</td>
<td>GTGA motif found in the promoter of the tobacco late pollen gene g10.</td>
</tr>
<tr>
<td>22</td>
<td>HEXAT</td>
<td>(+) TGACGTGG</td>
<td>Binding site of Arabidopsis bZIP protein TGA1 and G box binding factor GBF1.</td>
</tr>
<tr>
<td>23</td>
<td>IBOX</td>
<td>(-) GATAAG</td>
<td>Conserved sequence upstream of light-regulated genes.</td>
</tr>
<tr>
<td>24</td>
<td>L1BOXATPDF1</td>
<td>(+) TAAATGYA</td>
<td>L1 box found in promoter of A. thaliana PROTODERMAL FACTOR1 (PDF1) gene, Involved in L1 layer-specific expression</td>
</tr>
<tr>
<td>25</td>
<td>LECPLEACS2</td>
<td>(+) TAAATAT</td>
<td>Core element in LeCp (tomato Cys protease) binding cis-element in LeAcs2 gene</td>
</tr>
<tr>
<td>26</td>
<td>LTRE1HVBLT49</td>
<td>(+) CCGAAA</td>
<td>LTRE-1 (low-temperature-responsive element) in barley blt4.9 gene promoter.</td>
</tr>
<tr>
<td>27</td>
<td>MARTBOX</td>
<td>(+) TTWTWTTWTT</td>
<td>found in SAR (scaffold attachment region) or matrix attachment region, (MAR).</td>
</tr>
<tr>
<td>28</td>
<td>MYB1AT</td>
<td>(+) WAACCA</td>
<td>MYB recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in Arabidopsis</td>
</tr>
<tr>
<td>29</td>
<td>MYBCORE</td>
<td>(+) CNGTTR</td>
<td>Binding site for a ATMYB2 involved in regulation of genes that are responsive to water stress in Arabidopsis</td>
</tr>
<tr>
<td>30</td>
<td>MYBCOREATCYCB1</td>
<td>(+) AACGG</td>
<td>found in the promoter of Arabidopsis thaliana cyclin B1:1 gene</td>
</tr>
<tr>
<td>31</td>
<td>MYBGAHV</td>
<td>(+) TAACAAA</td>
<td>Central element of gibberellin (GA) response complex (GARC) in high-pI alpha-amylase gene in barley</td>
</tr>
<tr>
<td>32</td>
<td>MYCCONSENSUSAT</td>
<td>(+) CANNTG</td>
<td>MYC recognition site found in the promoters of the dehydration-responsive gene rd22</td>
</tr>
<tr>
<td>33</td>
<td>OSE2ROOTNODULE</td>
<td>(+) CTCTT</td>
<td>One of the consensus sequence motifs of organ-specific elements of the promoters activated in infected cells of root nodules</td>
</tr>
<tr>
<td>34</td>
<td>POLLEN1LELAT52</td>
<td>(+) AGAAA</td>
<td>One of two co-dependent regulatory elements responsible for pollen specific activation of tomato lat52 gene</td>
</tr>
<tr>
<td>35</td>
<td>PREATPRODH</td>
<td>(-) ACTCAT</td>
<td>Necessary for the efficient expression of ProDH in response to hypoosmolarity</td>
</tr>
<tr>
<td>36</td>
<td>PROLAMINBOXOSGLUB1</td>
<td>(-) TGCAAAG</td>
<td>Involved in quantitative regulation of the GluB-1 gene</td>
</tr>
<tr>
<td>37</td>
<td>PYRIMIDINEBOXOSRAMY1A</td>
<td>(-) CCTTTT</td>
<td>pyrimidine box is partially involved in sugar repression</td>
</tr>
<tr>
<td>38</td>
<td>RYREPEATBNNAPA</td>
<td>(+) CATGCA</td>
<td>Required for seed specific expression</td>
</tr>
<tr>
<td>Element Code</td>
<td>Sequence/Effect</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>RYREPEATLEGUMINBOX (+) CATGCAY</td>
<td>+</td>
<td>Legumin box found in seed-storage protein genes in legumes</td>
<td></td>
</tr>
<tr>
<td>SEBFCONSSTPR10A (-) YTGTWC</td>
<td>-</td>
<td>Binding site of the potato silencing element binding factor (SEBF) gene found in promoter of pathogenesis-related gene</td>
<td></td>
</tr>
<tr>
<td>SREATMSD (+) TTATCC</td>
<td>+</td>
<td>Sugar-repressive element (SRE); found in down-regulated genes after main stem decapitation in Arabidopsis</td>
<td></td>
</tr>
<tr>
<td>SURECOREATSULTR11 (+) GAGAC</td>
<td>+</td>
<td>SURE contains auxin response factor (ARF) binding sequence.</td>
<td></td>
</tr>
<tr>
<td>TAAAGSTKST1 (+) TAAAG</td>
<td>+</td>
<td>TAAAG motif found in promoter of KST1 gene encoding for a K+ influx channel of guard cells</td>
<td></td>
</tr>
<tr>
<td>TATABOX5 (+) TTATTT</td>
<td>+</td>
<td>TATA box found in the 5' upstream region of pea glutamine synthetase gene</td>
<td></td>
</tr>
<tr>
<td>TATCCAOASAMY (+) TATCCA</td>
<td>+</td>
<td>Mediate sugar and hormone regulation of alpha-amylase gene expression.</td>
<td></td>
</tr>
<tr>
<td>TATCCAYMOTIFOSRAMY3D (+) TATCCAY</td>
<td>+</td>
<td>Responsible for sugar repression.</td>
<td></td>
</tr>
<tr>
<td>TGACGTVMAMY (+) TGACGT</td>
<td>+</td>
<td>Required for high level expression of alpha-Amylase in the cotyledons of the germinated seeds</td>
<td></td>
</tr>
<tr>
<td>WBOXATNPR1 (-) TTGAC</td>
<td>-</td>
<td>Recognized specifically by salicylic acid-induced WRKY DNA binding proteins</td>
<td></td>
</tr>
<tr>
<td>WBOXNTERF3 (+) TGACY</td>
<td>+</td>
<td>Involved in activation of ERF3 gene by wounding.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2**: List of elements identified in EcSWEET5 promoter by PLACE.
Fig. 2 a and b

Fig. 3

EcSWEET5
EcSWEET2
18S

Stamen

Pistil

EcSWEET5
18S

https://mc06.manuscriptcentral.com/genome-pubs
Fig. 4
**Fig. 5a and b**

(a) Stem dry wt (g) and Root dry wt (g)

(b) Pollen germination %

- **WT**
  - Stem dry wt: n=350
  - Root dry wt: n=350
- **TRV::EcPDS**
  - Stem dry wt: n=220
  - Root dry wt: n=220
- **TRV::EcSWEET5**
  - Stem dry wt: n=206
  - Root dry wt: n=206

https://mc06.manuscriptcentral.com/genome-pubs
Fig. 6

https://mc06.manuscriptcentral.com/genome-pubs