Published in "Sexual Plant Reproduction 23(3): 173-186 , 2010"
which should be cited to refer to this work.

Characterisation of *sunflower-21* (*SF21*) genes expressed in pollen and pistil of *Senecio squalidus* (Asteraceae) and their relationship with other members of the *SF21* gene family

Alexandra M. Allen · Christian Lexer · Simon J. Hiscock

Abstract Two related flower-expressed gene copies belonging to the *SF21* (*sunflower-21*) gene family have been isolated from *Senecio squalidus* (Oxford Ragwort, Asteraceae). These gene copies are differentially expressed in pollen and pistil tissues; *ORSF21B* (*Oxford Ragwort SF21B*) is expressed exclusively in mature pollen, whereas *ORSF21A* (*Oxford Ragwort SF21A*) is expressed in the transmitting tissue of the style, where it is developmentally regulated. Despite differences in expression, the coding regions of *ORSF21A* and *ORSF21B* are highly similar. Amino acid sequence alignments of *SF21* genes from a number of angiosperm species indicate that this gene family is conserved in flowering plants and may play an important role in reproductive processes in a wide range of taxa. Phylogenetic analysis of *SF21* nucleotide sequence alignments supports this theory, and indicates a complicated history of evolution of this gene family in angiosperms. The putative roles of *SF21* genes in reproduction and pollen–pistil interactions are discussed.

Keywords Pollen · Pistil · *Senecio* · *SF21* · Negative differentiation regulator (NDR)

Communicated by Hugh Dickinson.

Electronic supplementary material see the online version of this article

A. M. Allen · S. J. Hiscock (✉)
School of Biological Sciences, University of Bristol,
Woodland Road, Bristol BS8 1UG, UK
e-mail: simon.hiscock@bristol.ac.uk

C. Lexer
Royal Botanic Gardens, Kew, Richmond,
Surrey TW9 3DS, UK

Present Address:
C. Lexer
Department of Biology, Unit of Ecology and Evolution,
University of Fribourg, Chemin du Musée 10,
1700 Fribourg, Switzerland

Introduction

The interplay between pollen and pistil tissues is vital for successful sexual reproduction in flowering plants. The tissue of the pistil acts as a physical and chemical interface between male and female gametophytes, from the stigma surface at pollen germination until successful fertilisation (Heslop-Harrison 1975; Hiscock and Allen 2008). Most research on pollen–pistil interactions has focussed on the self-incompatibility (SI) reaction; however, a diverse range of other molecular interactions take place between the male and female reproductive tissues, including species recognition, pollen hydration, pollen tube growth and defence mechanisms (Hiscock and Allen 2008). Despite the universal need for such mechanisms in angiosperms, a relatively low number of shared pistil-expressed genes have been identified in the different species studied so far (Hiscock and Allen 2008). This may indicate that different taxa have employed alternative molecular mechanisms for achieving successful reproduction, as has been illustrated by the diverse range of SI mechanisms identified to date (Hiscock and McInnis 2003; Allen and Hiscock 2008). It has been suggested that genes that are involved in regulating sexual reproduction are likely to evolve at a faster rate than those controlling housekeeping processes (Swanson and Vacquier 2002). Certainly, there are few examples of shared genes between species, despite the identification of an increasing number of molecules implicated in pollen–pistil interactions.

Senecio squalidus (Oxford Ragwort) has been used as a model to study pollen–pistil interactions and sporophytic SI (SSI) in the Asteraceae (Hiscock et al. 2002). To study pollen–pistil interactions in *S. squalidus,* several approaches have been implemented that have identified pistil- and/or pollen-specific genes. These include microarray analysis (Hegarty et al. 2005) and suppression subtractive hybridisation (SSH) (Allen et al. submitted). A screen of *Senecio* floral cDNA libraries through microarray analysis identified pistil- and pollen-specific genes that were up-regulated in self-incompatible *Senecio squalidus* compared to self-compatible *S. cambrensis* (Allen 2009). A second approach used SSH to create *S*-genotype-specific pistil-enriched cDNA libraries (Allen et al. submitted). Both of these techniques identified candidate genes potentially involved in pollen–pistil interactions. Here, we describe one gene that was common to both datasets, *Oxford Ragwort sunflower-21* (*ORSF21*). This gene showed significant sequence similarity to *SF21*, a gene of unknown function originally identified in *Helianthus annuus* (Sunflower) (Kräuter-Canham et al. 1997).

SF21 genes in angiosperms appear to belong to the *Negative differentiation regulator* (*Ndr*) gene family, found in mammals, suggesting an evolutionary link (Kräuter-Canham et al. 2001). Members of this gene family include the human protein *RTP* and the mouse protein *Ndrl*. *SF21* also shares significant sequence homology with an unidentified polypeptide from human cerebellum (N-terminus) and with the ligand-binding region of the vertebrate inositol 1,4,5 trisphosphate (IP3) receptor (C-terminus) (Kräuter-Canham et al. 1997). The precise function of members of this family is still unknown, although they are known to be involved in cellular differentiation events (Okuda and Kondoh 1999). The *SF21* and *Ndr* gene families also share sequence similarity with the alpha/beta hydrolases, suggesting an enzymatic function of these genes (Hotelier et al. 2004).

Homologous *SF21* genes have been identified in a number of angiosperms including *Arabidopsis thaliana, Nicotiana tabacum* and *Oryza sativa* (Kräuter-Canham et al. 1997, 2001; Nas et al. 2005). Despite this apparent widespread possession of *SF21* genes in the angiosperms, a function for this gene family is yet to be confirmed. Members of the *SF21* gene family have been detected in pollen and pistil tissue, where their expression is typically highly regulated (Kräuter-Canham et al. 1997, 2001; Lazarescu et al. 2006). This study analyses expression and sequence characteristics of *ORSF21* in *S. squalidus* and compares the findings with those for homologous genes in other species to assess possible functions of SF21 in pollen–pistil interactions. Consensus in expression and sequence data between diverse angiosperm species indicates that *SF21* may be a rare example of a conserved pollen–pistil gene, and hence play a vital role during reproduction in angiosperms.

Materials and methods

Plant material

All *Senecio* plants were grown under glasshouse conditions (Hiscock 2000). Individuals from an Oxford *S. squalidus* population used for DNA extraction were genotyped as per Brennan et al. (2006). Tissues used for other purposes were collected from a range of genotypes and pooled to ensure a representative sample of the population. Genomic DNA was extracted from leaf tissue using the cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987).

Northern hybridisation

Total RNA was extracted using TRIzol® reagent (Invitrogen) from *S. vulgaris, S. cambrensis* capitula and *S. squalidus* root, leaf, stem and various floral developmental stages: capitulum bud (2–3 mm), capitulum bud (5–6 mm), small flower buds, mature flower buds, mix of open florets and buds, florets all open, mature stigmas and pollen. Ten micrograms of total RNA was separated on a 1.2% formaldehyde agarose (FA) gel at 80 V for 3 h, blotted onto a Hybond-NX membrane and immobilised using ultraviolet (UV) cross-linking in a CL-1000 UV crosslinker (UVP), setting 70 mJ/cm^2. Blots were pre-hybridised for 3 h in northern hybridisation buffer (50% (v/v) formamide, 5× SSPE, 5× Denhardt's solution, 0.5% (w/v) SDS) at 42°C and hybridised overnight in fresh buffer at 42°C. Blots were probed with ^{32}P alpha-dCTP-labelled DNA. Membranes were washed in four changes of wash buffer (1 × SSC/0.1% SDS) at 42°C and exposed to BioMax MS-1 autoradiography film (Kodak).

Single nucleotide primer extension (SNuPe)

RNA was extracted from *S. squalidus* pollen and pistil tissues using TRIzol® reagent (Invitrogen) and used to synthesise cDNA using the SMART PCR cDNA synthesis kit (Clontech). Standard PCR was performed to amplify a region of cDNA containing a single nucleotide polymorphism (SNP). Primers were designed to flank a SNP between *ORSF21A* and *ORSF21B* in exon 3 (SNuPe forward, GACGTGTTTCCATGGACT; SNuPe reverse, CTGAATCGGGACTTATGAA). SNuPe genotyping was performed using the SNuPe Genotyping Kit (GE Healthcare), according to manufacturer's instructions. Briefly, the single nucleotide primer extension reactions contained 0.05 pmol template DNA, 1 pmol SNuPE primer, 4 µl SNuPe Premix in 10 µl sterile distilled water. The extension reactions were carried out in a G-Storm GS1 thermal cycler (GRI) using 25 cycles, each cycle consisting of a 10-s denaturing step at 96°C, a 5-s annealing step at 50°C

and a 10-s extension step at 60°C. After the SNuPE, the excess ddNTPs were removed by incubation for 30 min at 37°C with 0.5 U SAP (Roche), followed by deactivation of the enzyme at 72°C for 15 min. The reaction products were purified and diluted in a formamide loading solution (GE Healthcare) and run with the MegaBACE SNuPe Multiple Injection Marker. Control reactions and matrix standards were also loaded and run alongside the samples. Data were analysed using the MegaBACE DNA analysis instrument.

In situ hybridisation

In situ hybridisation was performed on partially and fully mature *S. squalidus* capitula according to Langdale (1993), with minor modifications. Capitulum bud tissue at various stages of development were picked from plants and immediately fixed in FAA (5% formaldehyde, 5% acetic acid, 45% ethanol). Fixed samples were then dehydrated through an ethanol series (50, 75, 80, 100 and 100%; 2 h each step) and then through a Histoclear (Raymond A Lamb)—ethanol series (30, 50, 80, 100 and 100%; 30 min each step). Samples were then placed in paraffin wax (VWR international). Sections were cut at 10 μm using a rotary microtome (Leica) and mounted on Superfrost Plus slides (BDH). Samples were rehydrated through an ethanol–water series (100, 100, 80, 70, 50 and 30%, ddH_20; 5 min each step).

The probe was amplified from *S. squalidus* pistil cDNA (in situ forward, CAATGGCGGGTGCTTATATC; in situ reverse, GCTTGAAGATACCGCCACAC). Probe cDNA fragments were amplified via PCR and cloned into the pSC-A-amp/kan vector using the Strataclone kit (Stratagene). Colony PCR was performed using the M13 reverse primer and each of the gene-specific PCR primers to determine insert orientation. Colonies containing inserts were cultured overnight and the plasmid DNA was purified. Inserts were amplified via standard PCR using the M13 forward and reverse primers and purified using the MinElute PCR purification kit (Qiagen). Antisense and sense probes were transcribed from the T3 and T7 promoters and labelled with DIG-UTP (Roche). Probes were diluted in 50% formamide and stored at −80°C. The labelling efficiency of the probes was checked against control-labelled RNA (Roche) by applying and fixing a dilution series of each to a Hybond-X nylon membrane (GE Healthcare). After hybridisation, sections were mounted in Entellan (Merck) and viewed using a light microscope (Leica) under magnification ×100 to ×400. Images were captured using a digital camera (Nikon Coolpix 995).

Rapid amplification of cDNA ends (RACE)

RNA ligase-mediated RACE was performed on the *ORSF21* cDNA clone using the FirstChoice RLM-RACE kit (Ambion) to obtain the full-length sequence. Gene-specific primers were designed to *S. squalidus ORSF21* cDNA (gene-specific RACE 5′ outer, GAAGCTGAATCG GGACTTATG; gene-specific RACE 5′ inner, CACGT CGTGTGGTTTAGAGC; gene-specific RACE 3′, GTACC AGCGGGAGCTAACAC). Amplified products were gel-purified using the QIAquick® gel extraction kit (Qiagen) and cloned directly into the TOPO 2.1 vector using the TOPO TA cloning kit (Invitrogen) according to the manufacturer's instructions. Clones were sequenced by University of Oxford, Department of Biochemistry and sequences were aligned in ChromasPro (Technelysium).

Restriction digestion analysis

ORSF21 genomic sequences were digested using the restriction enzymes *Rsa*1 and *Hind*III (supplied by New England Biolabs). Approximately 5 μg of PCR product was digested, a 50 μl volume, with 10U restriction enzyme and 5 μl recommended digestion buffer (10×). Digests were allowed to proceed overnight at 37°C in a proportional temperature controller (LEEL). The digested DNA was purified via phenol/chloroform extraction and ethanol precipitation and resuspended in 10 μl sterile distilled water. The product was viewed via standard agarose gel electrophoresis.

Sequencing of full-length gene

The complete cDNA coding region was amplified by PCR using primers designed to anneal within the 3′ and 5′ untranslated regions (cDNA forward, GAAAATAACAT CAAATAAAGCAAGGAC; cDNA reverse, GAAAGAGA ACG TTATGAACACGAAC). The corresponding genomic sequences for *ORSF21* were PCR amplified using the same primer pair and Phusion Hot Start High-Fidelity DNA polymerase (NEB). PCR products were cloned into the TOPO 2.1 vector using the TOPO TA cloning kit (Invitrogen) according to the manufacturer's instructions. Clones were sequenced by the University of Oxford. Sequencher software (Gene Codes Corporation) was used to create contigs and align sequences. The complete cDNA sequences of ORSF21A and ORSF21B are deposited in the Genbank database under accession numbers GQ227732 and GQ227733, respectively.

Sequence analysis

Annotation of transcripts was performed using the BLASTX algorithm against the non-redundant protein database (http://www.ncbi.nlm.nih.-gov/BLAST/), using default parameters. Functional annotation was assigned according to the Gene Ontology database. Nucleotide

sequence diversity was analysed using the Proseq (Filatov 2002) and DNAsp (Rozas et al. 2003) packages. Sequences were translated, and secondary structure prediction performed using DNAMAN (Lynnon Corporation). Translated sequences were assessed for conserved domains using the Conserved Domain database and Conserved Domain Architecture Retrieval Tool (CDART) (http://www.ncbi.nlm.nih.gov/Structure/). Transmembrane helix prediction was performed using the TMHMM Server v. 2.0 (Krogh et al. 2001) and modelled using the TMRPres2D program (Spyropoulos et al. 2004). Prediction of protein 3D structure was performed using the Phyre protein fold recognition server (Kelley and Sternberg 2009).

Phylogenetic analysis

Phylogenetically related sequences were identified using the BLAST-x and BLAST-n algorithms (Altschul et al. 1997). Gene and protein sequences were aligned using DNAMAN (Lynnon Corporation). Phylogenetic trees were generated from alignments using PAUP v 4.0b 10 (Swofford 2003). Parsimonius trees were generated via a heuristic search with branch swapping set at 1,000 rearrangements. Bootstrap calculations were based on 1,000 replicates.

Results

ORSF21 is expressed exclusively in pistil and pollen tissues

Northern blot analysis indicated that *ORSF21* is expressed at a higher level in the SI species *Senecio squalidus* in comparison with the SC species *S. vulgaris* and *S. cambrensis* (Fig. 1a). *ORSF21* is expressed specifically in the floral tissues of *S. squalidus,* where transcripts are localised to pistil and pollen tissues (Fig. 1b). Within the pistil, there is developmental regulation of expression of *ORSF21*, with maximal expression occurring in the mature flower bud (Fig. 1b). In situ hybridisation of *S. squalidus* mature flowers with a probe designed to the *ORSF21* sequence revealed a complex expression pattern with developmental regulation. In maturing capitula, the gene was expressed in the transmitting tissue of the upper pistil at stigma stages 3 and 4 only (Figs. 2e, f, 3a). In fully mature capitula, expression of the gene was detected in the transmitting tissue immediately above the ovaries and in the integument of the ovules of mature pistils (ovary stages 2 and 3, Figs. 2h, i, 4a, c); this expression was not present at earlier stages of development (ovary stage 1, Figs. 2g, 3e). *ORSF21* was also expressed at high levels in mature pollen grains (Fig. 4e).

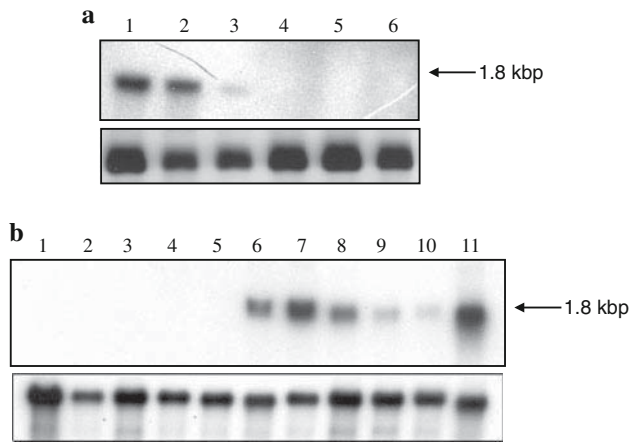

Fig. 1 RNA gel blot analysis of *ORSF21* expression. **a** *upper panel*: Expression of *ORSF21* in the floral tissues of different *Senecio* species. Each lane contains 10 µg total RNA. *Lane 1*: *S. squalidus* mature flowers, *2*: *S. cambrensis* mature flowers, *3*: *S. vulgaris* mature flowers, *4*: *S. squalidus* capitulum bud, *5*: *S. squalidus* leaf, *6*: *S. squalidus* stem. *Lower panel*: actin loading control. **b** *upper panel*: Expression of *ORSF21* in various tissue types and floral developmental stages of *S. squalidus*. Each lane contains 10 µg total RNA. *Lane 1*: root, *2*: leaf, *3*: stem, *4*: capitulum bud (2–3 mm), *5*: capitulum bud (5–6 mm), *6*: small flower buds, *7*: mature flower buds, *8*: mix of open florets and buds, *9*: florets all open, *10*: mature stigmas, *11*: pollen. *Lower panel*: as above but membrane stripped and re-probed with actin gene fragment as a loading control

The *S. squalidus* genome contains multiple copies of *ORSF21*

Two distinct sequences were obtained for *ORSF21* in *Senecio squalidus* in multiple genotypes. Genomic sequences were obtained from four different *S. squalidus* *S*-genotypes (S_1S_2, S_1S_3, S_1S_4 and S_1S_6), via long range PCR. Two genotypes yielded a double PCR product, indicating the possible presence of at least two *SF21* sequences of different lengths in the *S. squalidus* genome (Fig. S1a). Restriction digestion analysis of the two PCR products indicated sequence polymorphism (Fig. S1b). Both gene copies were isolated and sequenced within one *S*-genotype, confirming that the variation is not an *S*-genotype-specific polymorphism. Gene copy 1 is hereafter referred to as *ORSF21A* (Oxford ragwort *SF21*), whereas gene copy 2 is referred to as *ORSF21B*. Most nucleotide diversity between *ORSF21A* and *ORSF21B* occurred within introns, with large insertions and deletions accounting for the size differences seen in the PCR products (Table 1; Fig. S1). Thirteen SNPs were present in the coding region; two of these are non-synonymous, resulting in the production of a different amino acid during translation (Fig. 5). Analysis of the secondary structure of the proteins predicts that the mutation at position 264 has no overall effect. However, a mutation at position 36 alters the predicted secondary structure from a beta-strand to a coil.

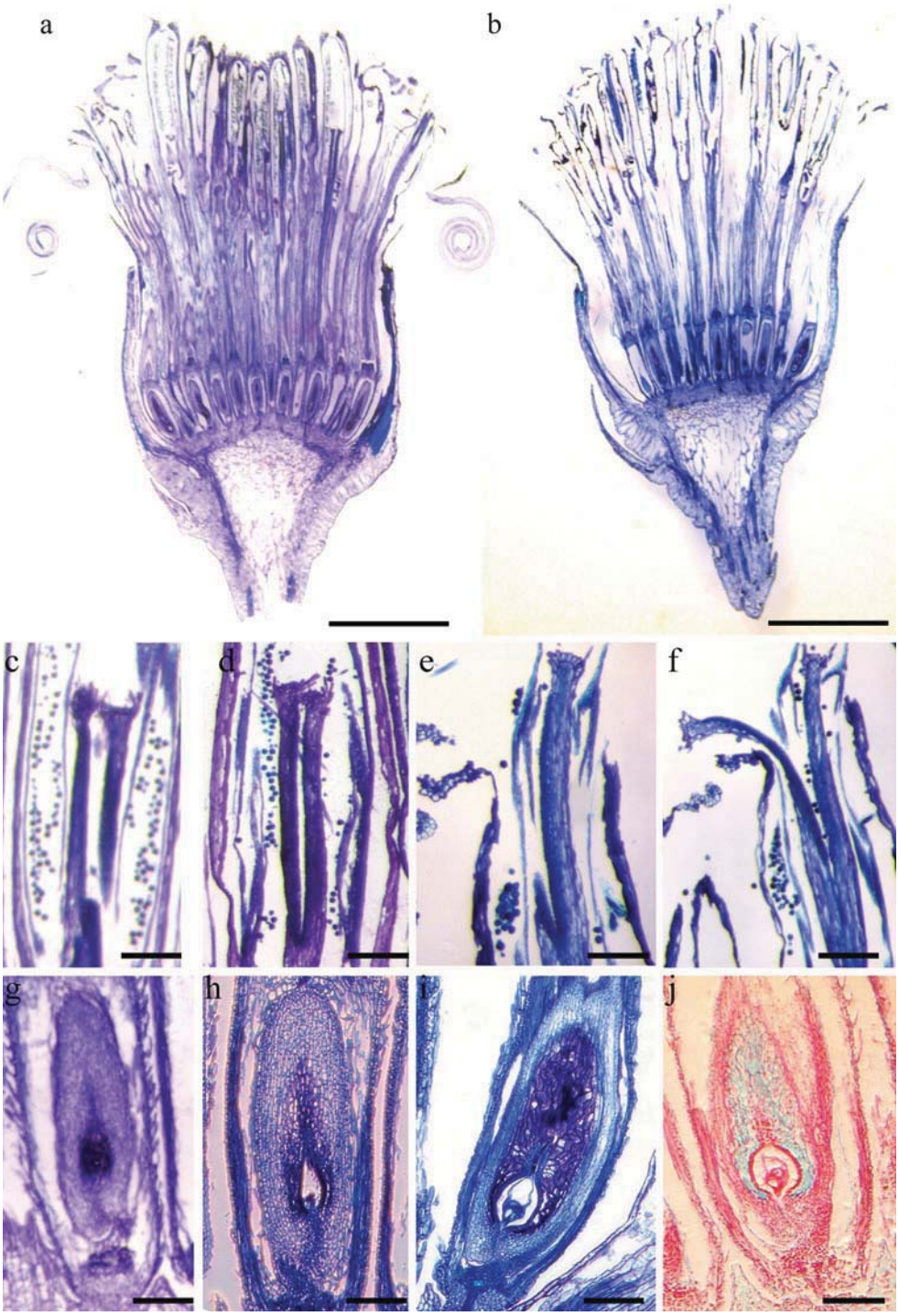

Fig. 2 Section through *Senecio squalidus* capitula, showing individual disc florets in different stages of development. **a** Maturing capitulum (Fig. 1b, stage 8); **b** fully mature capitulum (Fig. 1b, stage 9). *Bars* = 5 mm. **c–j** Detail of pistil developmental stages: **c** stigma stage 1, surrounded by developing anthers; **d** stigma stage 2, growing through anthers; **e** stigma stage 3, emerging from floret; **f** stigma stage 4, fully mature; **g** ovary stage 1, immature; **h** ovary stage 2, maturing; **i** ovary stage 3, fully mature and receptive. Sections stained with toluidine blue according to Hiscock et al. (2002). **j** Mature ovary stained with AGS, highlighting the carbohydrate-rich integumentary tapetum (*blue*) and lignified/cutinised cell walls and lipids (*red*)

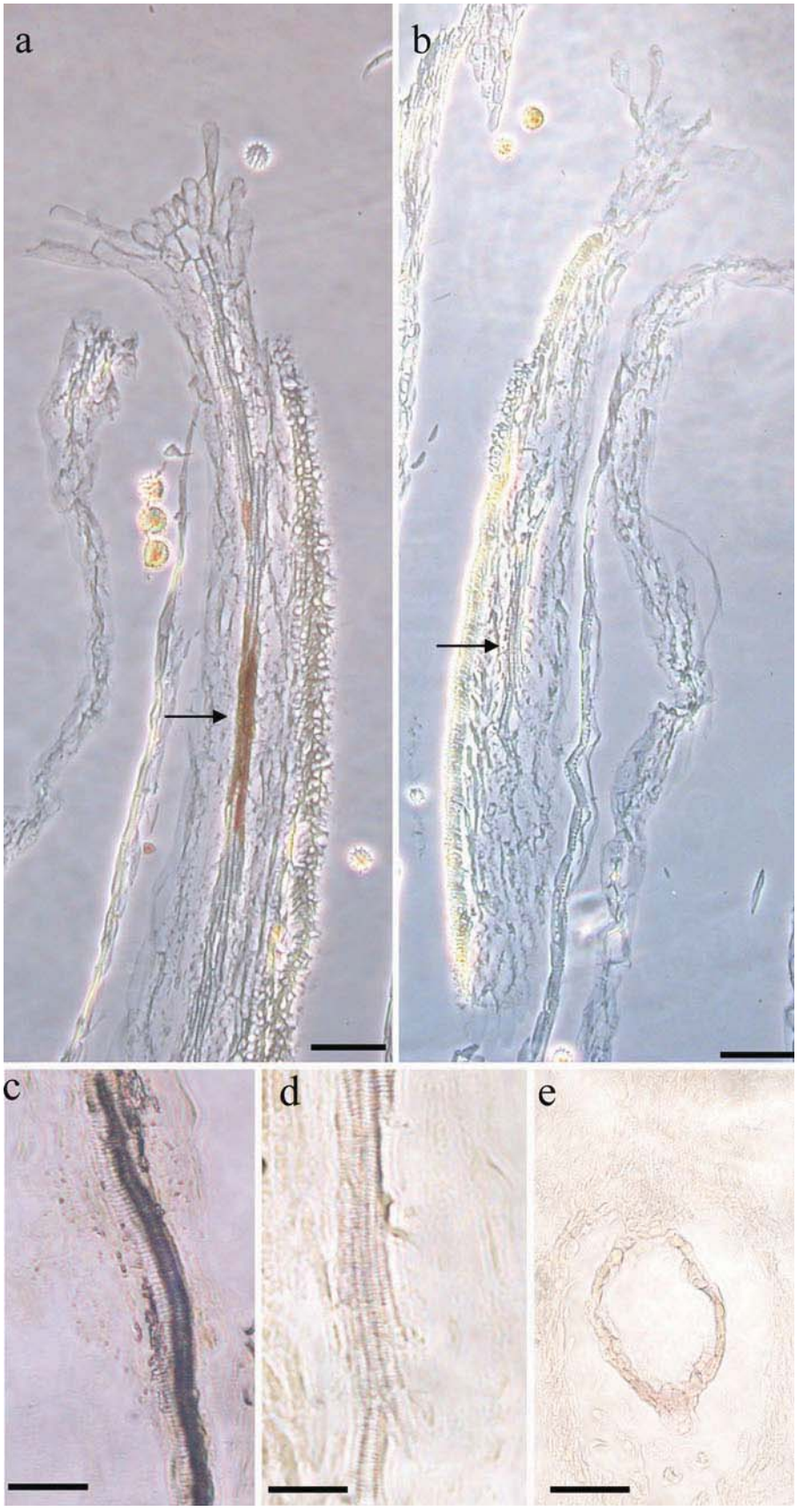

Fig. 3 In situ hybridisations on longitudinal sections of maturing *Senecio squalidus* pistils (stigma developmental stages 3 and 4). **a**, **c**, **e** hybridised with SF21 antisense probe; **b**, **d**, hybridised with sense probe. **a** Upper section of pistil showing expression in transmitting tissue (*black arrow*); **b** corresponding sense control showing no expression in transmitting tissue (*black arrow*). **c**, **d** Detail of transmitting tissue hybridised with antisense and sense probe, respectively. **e** Maturing ovary, stage 1, showing no expression. **a**, **b** ,*bars* = 50 μm; **c**, **d**, **e** *bars* = 25 μm

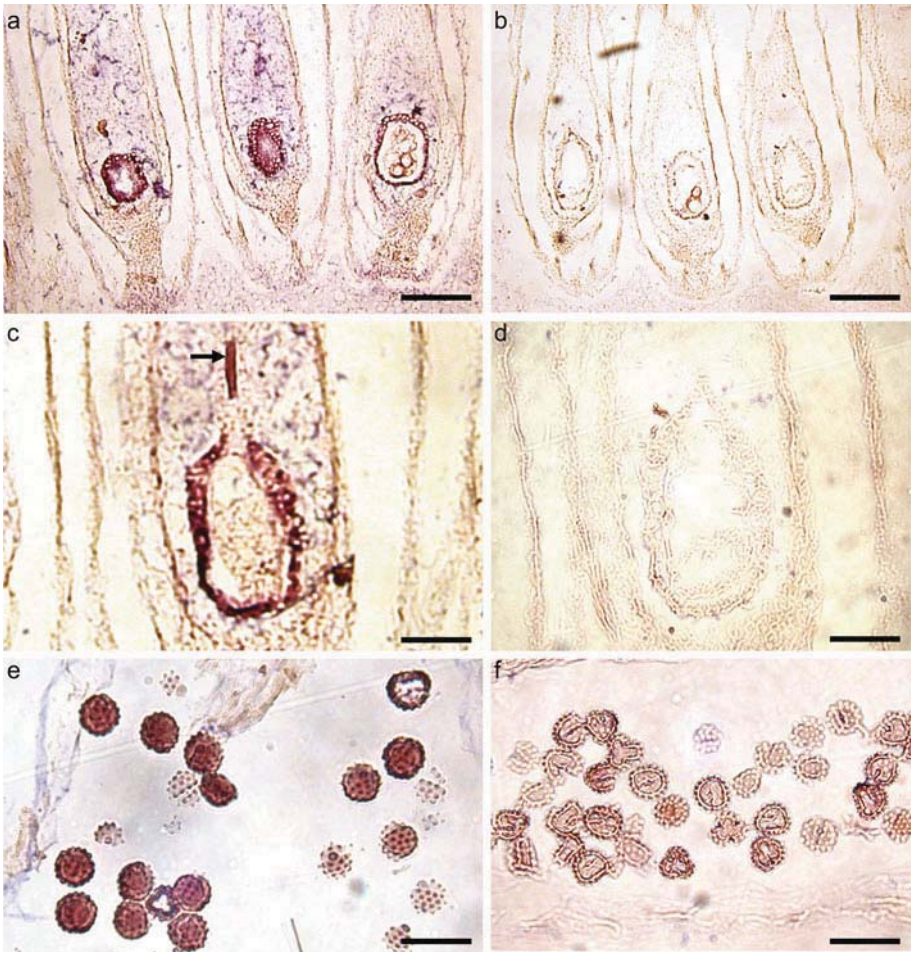

Fig. 4 In situ hybridisations on longitudinal sections of *Senecio squalidus* pistils. **a**, **c**, **e** Pistils at ovule developmental stages 2 and 3 hybridised with SF21 antisense probe; **b**, **d**, **f** hybridised with sense probe. **a** Base of pistil with staining in ovules. **c** Expression is localised to the integument cells surrounding embryo sac and transmitting tissue immediately above ovule (*black arrow*). **e** Expression was also detected in mature pollen grains. **a**–**b** *bars* = 50 μm; **c**–**f** *bars* = 25 μm

ORSF21A and *ORSF21* have different expression patterns in reproductive tissues

The SNuPe Genotyping Kit (GE Healthcare) was used to identify in which tissues each gene copy of the *SF21* gene was expressed. SNuPe primers were designed to either side of a SNP in exon 3 of the *ORSF21* gene (Fig. 6a). The SNuPe reaction was performed using each primer combined with either pollen or pistil cDNA. The forward primer was designed to amplify cytosine (C) or thymine (T), whereas the reverse primer amplified the complimentary nucleotide base; guanine (G) or adenine (A), respectively. The SNuPe reaction using the forward primer produced a clear cytosine peak from pollen

Table 1 Nucleotide diversity between *ORSF21A* and *ORSF21B*

Coding regions				Non-coding regions			
Genomic region (bp)	Exon	SNP's	Indels	Genomic region (bp)	Intron	SNP's	Indels
0–162	1	2	0	163–914	1	39	1 (137 bp)
915–1,017	2	0	0	1,018–1,115	2	0	0
1,116–1,215	3	1	0	1,216–1,292	3	1	0
1,293–1,442	4	1	0	1,443–1,591	4	1	1 (73 bp)
1,592–1,676	5	2	0	1,677–2,651	5	3	1 (652 bp)
2,652–2,729	6	0	0	2,730–2,829	6	3	0
2,830–2,886	7	0	0	2,887–3,112	7	3	1 (36 bp)
3,113–3,175	8	0	0	3,176–3,257	8	3	1 (1 bp)
3,258–3,393	9	4	0	3,394–3,596	9	2	1 (106 bp)
3,597–4,101	10	3	0				

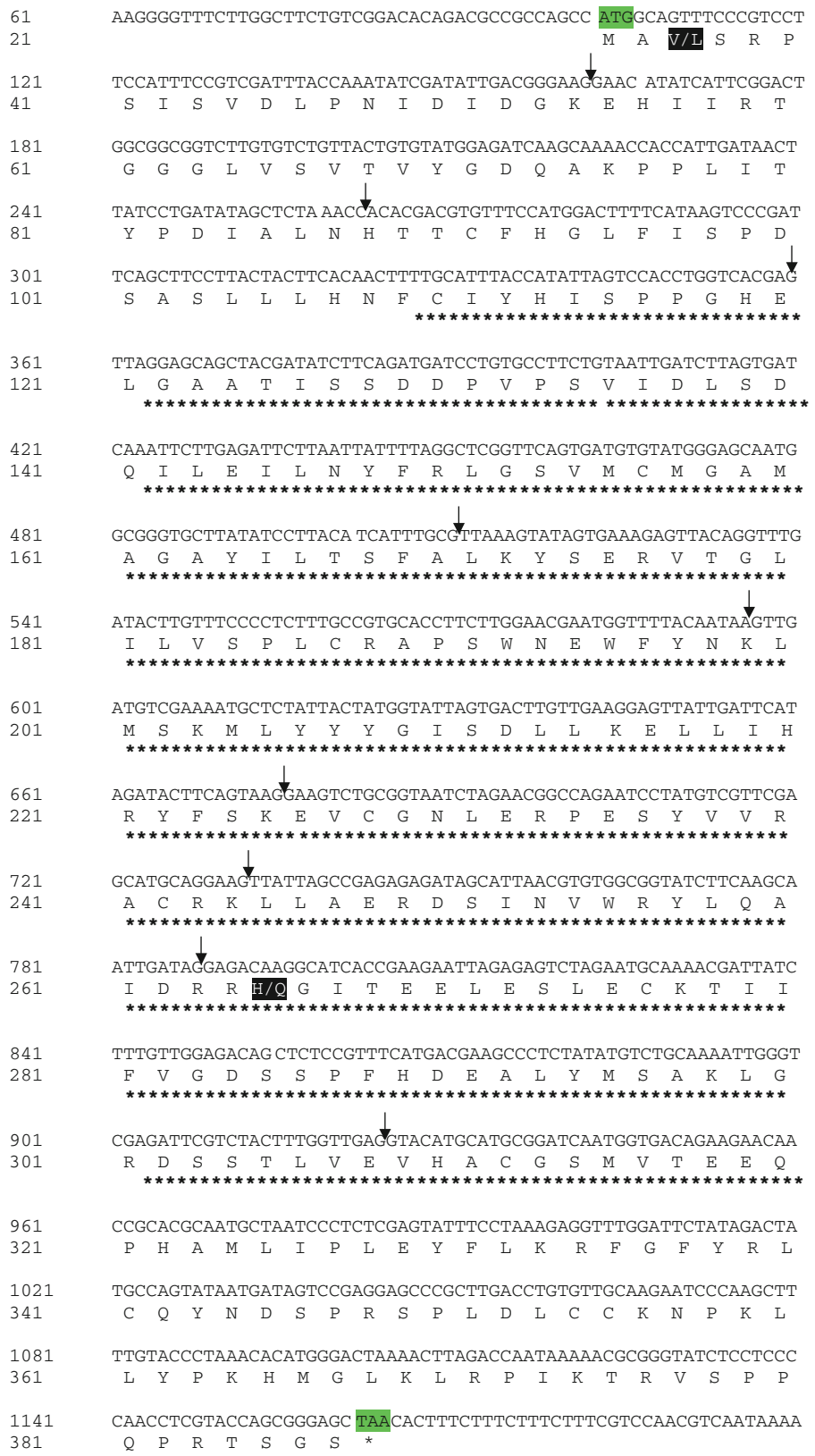

Fig. 5 Complete cDNA sequences (*top line*) and translations (*bottom line*) of the *S. squalidus ORSF21* genes. The start and stop codons are highlighted green, intron positions marked by *arrows*. The section highlighted by *stars* shows homology to the amino acid sequence of an alpha/beta hydrolase fold, a catalytic domain found in a wide range of enzymes (Hotelier et al. 2004). Non-synonymous mutations between *ORSF21A* and *ORSF21B* are highlighted in black

cDNA, indicating that *ORSF21B* is expressed in pollen tissues (Fig. 6b). This was confirmed by the reverse reaction, where the complimentary base, guanine was produced (Fig. 6c). Both results reflected the peaks seen in the control cytosine and guanine reactions (Figs. 6f, 7g, respectively).

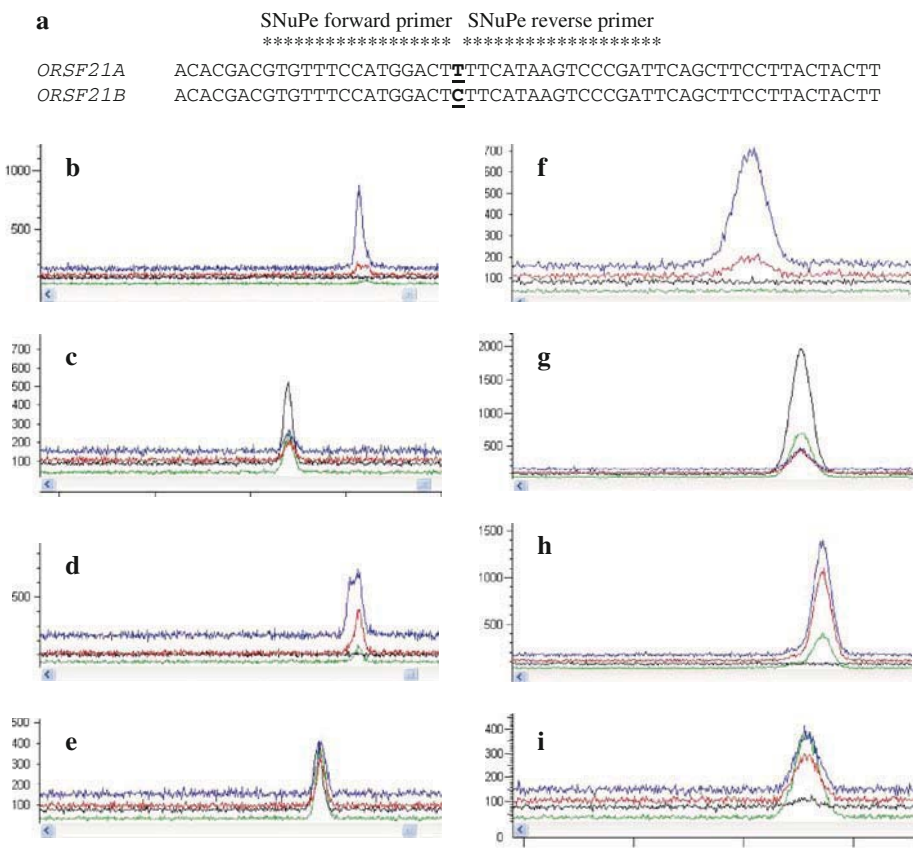

Fig. 6 SNuPe genotyping was performed to determine the site of gene expression of *ORSF21A* and *ORSF21B*. **a** SNuPe primers were designed to either side of a SNP (*underlined*) in exon 3 of the *ORSF21* gene. The SNuPe reaction was performed on stigma and pollen cDNA. **b** Pollen cDNA, forward primer; **c** Pollen cDNA, reverse primer; **d** stigma cDNA, forward primer; **e** stigma cDNA, reverse primer. Control nucleic acids were supplied with the SNuPe genotyping kit and run alongside experimental samples for validation of data; **f** control cytosine; **g** control guanine; **h** control thymine; **I** control adenine

The results from pistil cDNA indicate several possibilities. The reaction performed with the forward primer produced a peak in cytosine and thymine (Fig. 6d), whereas the reverse primer produced a peak in all four bases (Fig. 6e). SNuPe software was used to filter background noise; the results for the forward and reverse reactions were calculated as thymine and adenine, respectively. This result strongly indicates that *ORSF21A* is expressed in pistil tissues. However, due to the presence of more than one peak in the unfiltered results, it is possible that both gene copy transcripts are present in pistil cDNA. This may be due to expression of both gene copies in stigma tissues or may be the result of contamination of pistil cDNA with pollen tissue. It is clear that the pollen tissues express only one copy of the *ORSF21* gene, *ORSF21B*. It appears that the pistil tissues express *ORSF21A*, possibly together with *ORSF21B*. These results indicate differential expression of the two *ORSF21* gene copies in *S. squalidus* male and female reproductive tissues, with expression of at least one gene copy (*ORSF21A*) being restricted to one tissue type (pistil), and the other (*ORSF21B*) being expressed either exclusively in pollen (and expression detected in pistil is due to pollen contamination) or co-expressed in pollen and pistil tissues.

ORSF21A and *ORSF21B* belong to the *Ndr* (*Negative differentiation regulator*) gene family

A BLAST-x homology search identified sequence similarity (52%) between ORSF21 and the human protein Ndr1**,** which belongs to a large gene family containing members from a wide range of eukaryotes (Kräuter-Canham et al. 2001; Okuda and Kondoh 1999). The precise molecular and cellular functions of these proteins are still unknown, but they have been implicated in cellular differentiation events. Sequence and structural similarities suggest that members of this family are related to the alpha/beta hydrolases and that these proteins may have an enzymatic function (Fig. S2). Alignment of homologous SF21 proteins from *Senecio, Arabidopsis, Brassica,* sunflower (*Helianthus*) and lettuce (*Lactuca*) highlighted highly conserved regions, including a putative transmembrane helix, predicted from a hydrophobic segment of 18 amino acids (Fig. 7). The transmembrane and cytoplasmic regions of the proteins also appear to be highly conserved, whereas the extracellular region shows significantly more sequence diversity (Fig. 7a). Overall, the protein sequences show a high degree of sequence identity (>64%), particularly the *Senecio* and sunflower sequences (75–79%).

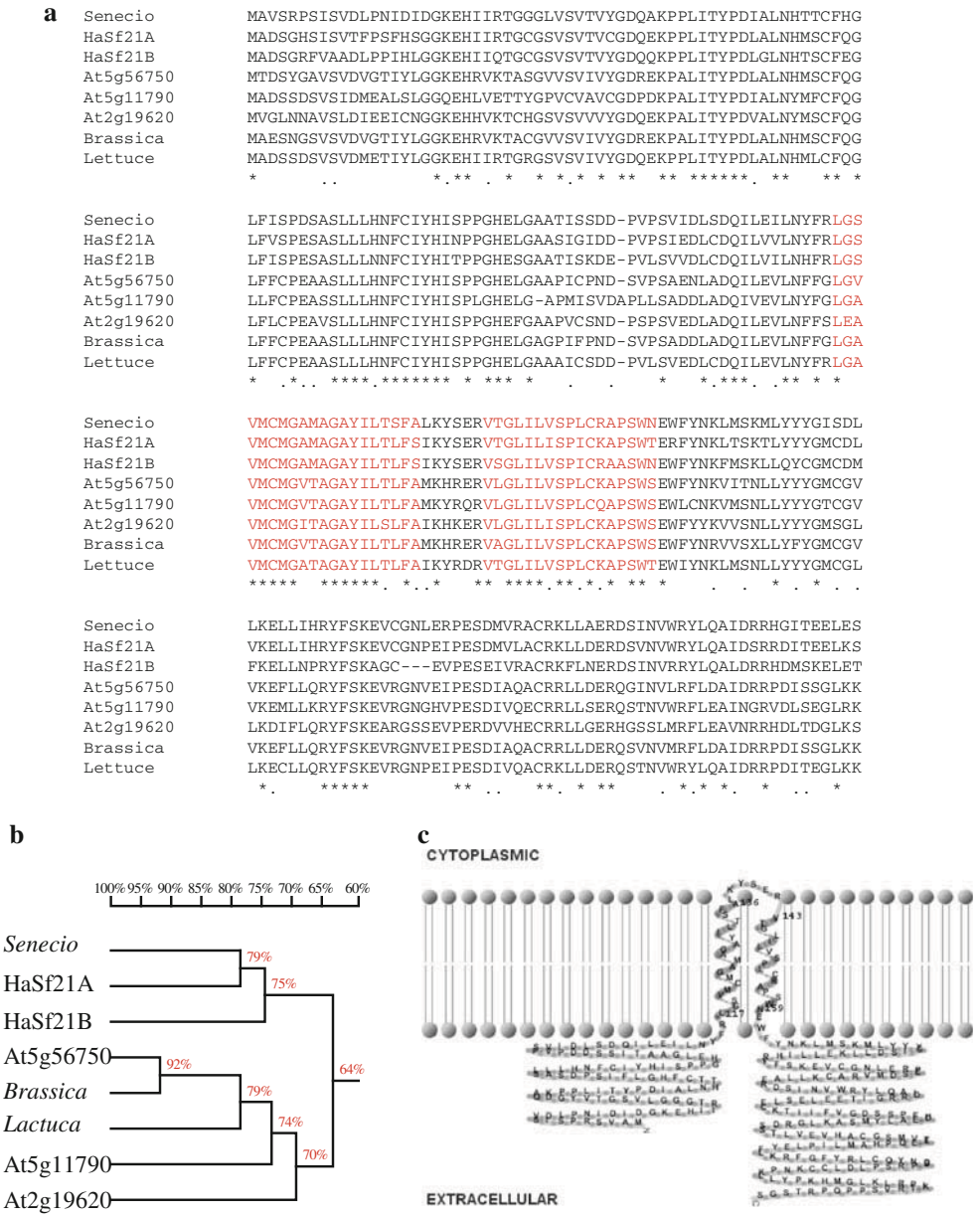

Fig. 7 **a** Predicted amino acid translation of *Senecio ORSF21* genes aligned with homologous sequences from *Helianthus annuus, Arabidopsis thaliana, Brassica napus* and *Lactuca sativa* (*asterisk* marks identical residues, *dot* marks conservative amino acid substitutions). Predicted transmembrane regions are coloured *red*. **b** Homology tree of aligned protein sequences. **c** Model of predicted transmembrane location of ORSF21, showing cytoplasmic and extracellular regions and predicted transmembrane helices

Discussion

SF21 sequences have been identified in all the major angiosperm clades and also in the gymnosperm taxa *Pinus* and *Picea* (Cairney et al. 2006). Phylogenetic analysis revealed that the nucleotide sequences of *SF21* genes appear to be remarkably conserved between diverse angiosperm and gymnosperm species, suggesting that they belong to an ancient gene family (Fig. 8). This analysis indicates that the *SF21* gene family has a complicated history of evolution, with multiple losses and gains of gene copies. For instance, in the monocotyledon species *Oryza sativa,* four gene copies exist that are more closely related to each other than to the genes in any other species (Jantasuriyarat et al. 2005; Zhang et al. 2005). This suggests that the *SF21* genes in *O. sativa* have undergone diversification from a single gene, but to determine whether this event happened at the base of the monocot clade or after speciation in *Oryza* would require further analysis with sequences from more monocot species. A similar

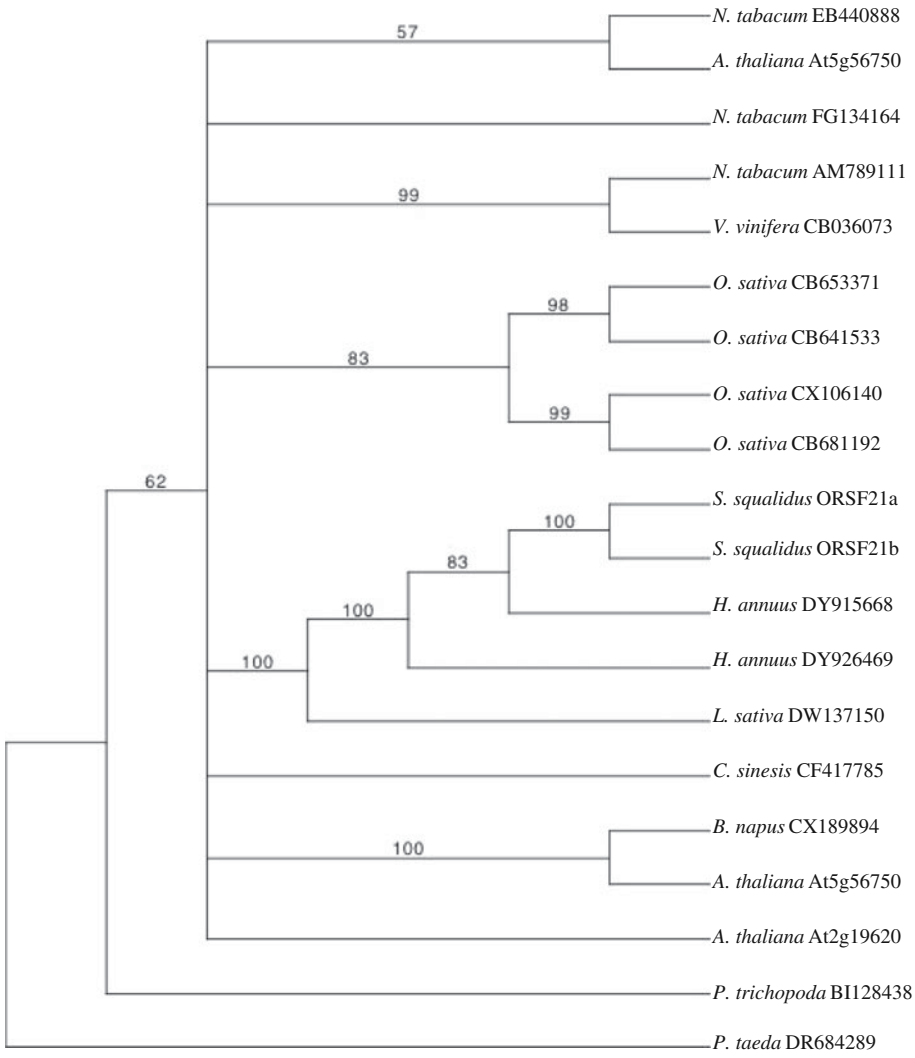

Fig. 8 Phylogenetic relationship between SF21 genes from all major clades of the angiosperms. The parsimonious tree was constructed in PAUP v4.0b 10 using a heuristic search method executed on a nucleotide sequence alignment. Nucleotide sequences used were from the following taxa: Asterids, *Senecio, Helianthus, Lactuca, Nicotiana;* Rosids, *Arabidopsis, Brassica, Citrus, Vitis, Populus;* Monocots, *Oryza* (Genbank accession numbers indicated next to species name). The tree is rooted with a sequence from *Pinus taeda.* Bootstrap values are indicated above branches

pattern is seen in the *Helianthus annuus SF21* gene family (Fig. 9), as sequences from the Asteraceae form a monophyletic clade within the phylogenetic tree (Fig. 8). Interestingly, sequences from other species of the Asterids and Rosids appear to show a more complex evolutionary history and do not segregate into clades consistent with the evolutionary relationships between the taxa. In particular, the tobacco sequences are strongly correlated with those from *Arabidopsis* and *Vitis,* a surprising result considering their evolutionary relationships (tobacco lies in the Asterids, while *Arabidopsis* and *Vitis* belong to the Rosids).

The high resolution of the Asteraceae clade indicates that these gene copies have evolved from one common ancestral gene (Fig. 8). Both sequences isolated from *S. squalidus* are more similar to the sunflower sequences *HaSF21B1* and *HaSF21A* than any of the other gene copies/splice variants. Furthermore, these four gene copies all show specialisation of expression in reproductive tissues (Fig. 9). The two putative gene copies in *S. squalidus* are much less diverged from each other than those found in *H. annuus*. Additionally, there are more gene copies and splice variants in *H. annuus* than so far detected in *S. squalidus.* This suggests that a common ancestor had one or two gene copies that were subjected to different selection pressures in each taxa. The palaeopolyploidy event that occurred at the base of the Heliantheae (Barker et al. 2008) may have contributed to the increase in *SF21* gene copy number in *Helianthus.* The predicted high degree of conservation of function of *SF21* genes is reflected in *Senecio*, where a low rate of evolution is indicated by the low intra- and interspecific nucleotide diversity, and low rate of non-synonymous to synonymous mutations (Table 2). A high positive Tajima's D value suggests that balancing selection has acted to maintain multiple gene copies in populations.

Phylogenetic and expression analyses of *SF21* genes have together suggested that these genes play specialised roles in reproductive tissues and that these roles may have

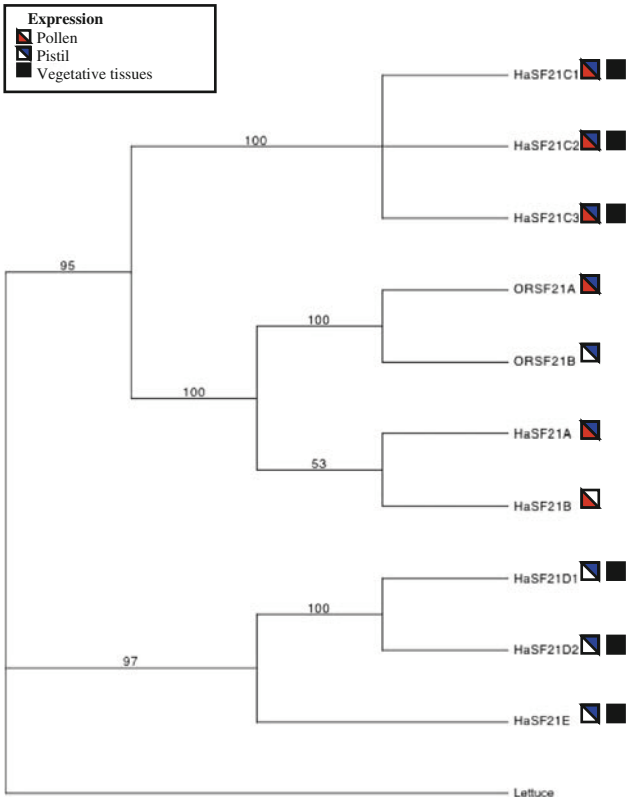

Fig. 9 Phylogenetic relationship between *S. squalidus* and *H. annuus SF21* genes. The parsimonious tree was constructed in PAUP v4.0b 10 using a heuristic search method executed on a nucleotide sequence alignment. Expression pattern of each gene copy is indicated in key; *H. annuus* expression data from Lazarescu et al. (2006)

evolved independently several times in the angiosperms. In the eudicots, pistil- and pollen-expressed *SF21* genes have been identified in *Arabidopsis* and tobacco. In *Arabidopsis,* all three of the *SF21* orthologues are expressed in reproductive tissues (At5g11790 and At5g56759, in pollen; At2g19620, in pistil) and at least one gene copy is pollen-specific (At5g56750, Winter et al. 2007). The expression pattern of *SF21* in tobacco pistil and pollen tissues has been well studied (Kräuter-Canham et al. 1997; Kräuter-Canham et al. 2001). An orthologue in poplar is expressed exclusively in male and female catkins (Ptp.4059.1, Winter et al. 2007). In species from the Asteraceae, two gene copies are expressed exclusively in reproductive tissues (Fig. 9, Lazarescu et al. 2006) and in *Senecio,* expression of these genes is highly regulated within these tissues (Figs. 3, 4).

In *Senecio,* the two gene copies have a high degree of sequence similarity, particularly in the coding region, but despite this, show differential expression in reproductive tissues, with *ORSF21B* localised to both pollen and pistil (potentially pollen-specific) and *ORSF21A* expressed exclusively in pistils (Fig. 6). The presence of ORSF21 in both male and female *S. squalidus* reproductive tissues has interesting implications for the function of the proteins. In the pistil, expression of *ORSF21* appears to be developmentally regulated. In maturing pistils (stigma stages 3 and 4, Fig. 2), expression is detected in the transmitting tissues in the upper part of the pistil, but not at the base of the pistil or in the ovary. In fully mature pistils (ovary stages 2 and 3), the gene is expressed in integument cells surrounding the embryo sac of the ovule and in the transmitting tissue immediately above the ovule. These results suggest a putative role for the protein in facilitating pollen tube elongation in *Senecio*, as the sites and timings of expression may be correlated with the expected passage of pollen tubes in the pistil. Studies in *Nicotiana tabacum* have shown that there is a gradient of expression of *SF21* transcripts towards the ovary, and both the transcript and protein have been shown to gradually disappear from pistil tissues as pollen tubes grow through the transmitting tissue (Krauter-Canham et al. 2001). Furthermore, hydropathy plots of homologous *SF21* sequences in a number of diverse species have revealed the presence of two transmembrane domains (Fig. 6), suggesting a role in transmembrane transport or signalling.

ORSF21B is expressed at high levels in pollen (Fig. 1), where it is detected in mature pollen grains (Fig. 4) and may also be expressed in pollen tubes, as is seen in other species (Kräuter-Canham et al. 1997). The putative role for

Table 2 Measures of polymorphism levels in *ORSF21* between *S. squalidus* and related *Senecio* species

Species comparison	Sequence length	*N*	*S*	Indel	Nucleotide diversity	Ka/Ks	Tajima's D
Intraspecific							
S. squalidus	1,245	4	15	0	0.0069	0.439	2.052*
Interspecific I							
S. chrysthanemifolius	1,393	2	41	0	0.0293	0.696	–
Interspecific I							
S. cambrensis	740	2	0	0	0	–	–

N number of sequences; *S* number of segregating sites; *Indel* insertion/deletion mutation; *Ka* non-synonymous substitutions; *Ks* synonymous substitution

* $P < 0.05$

SF21 genes in pollen has been studied in several other species. In rice, an *SF21* gene has been identified as a candidate thermosensitive male sterility gene, which is predicted to be essential for pollen development (Nas et al. 2005). Additionally, in sunflower, *SF21* transcripts were absent in anthers from a male-sterile line, but present in the wild type (Kräuter-Canham et al. 1997). In tobacco, the SF21 protein was detected in the tip region of the pollen tube, suggesting a role in pollen tube elongation (Lazarescu et al. 2006). The SF21 protein also shows homology (46% sequence similarity) with the ligand-binding region of the inositol 1,4,5 triphosphate (IP3) receptor protein from vertebrates and therefore may respond to signals involving the second messenger inositol 1,4,5 triphosphate (Kräuter-Canham et al. 1997). In *Papaver rhoeas*, it has been shown that pollen tube elongation is regulated by a calcium wave propagated by inositol 1,4,5 triphosphate (Franklin-Tong et al. 1996), suggesting that SF21 proteins may interact with IP3 to regulate pollen tube growth.

The diverse expression patterns of members of the *SF21* gene family suggest a number of different roles for these proteins in different plant tissues. This is illustrated particularly in the reproductive tissues, where there is evidence of adaptation of the gene to perform alternative functions in the male and female tissues. Research so far into *ORSF21* supports this hypothesis, with one gene copy expressed in pollen and the other in pistil tissues. The *SF21* gene may have undergone subfunctionalisation, in which the two duplicated gene copies specialised to perform separately localised complementary functions. The occurrence of different pollen and pistil-expressed gene copies of *SF21* may indicate these genes have a role in pollen–pistil interactions. A classic model of the SI mechanism proposed by Lewis (1965) involved the expression of an *S*-gene in both the pollen and pistil tissues. The rejection of self-pollen was hypothesised to occur when the identical proteins from each tissue interacted to form a dimer repressor. However, no evidence has been found for polymorphism between the *S*-genotypes, as would be predicted for a candidate *S*-gene. A more likely function for *SF21* is in guidance of compatible pollen tubes from the stigma to the ovary, a theory supported by the expression characteristics of *ORSF21A* in this study and *SF21* genes in tobacco. If *SF21* does function in pollen tube guidance, this important role may explain the high degree of conservation of gene sequence across the angiosperms. Certainly, the *SF21* gene family represents a rare example of a family of highly conserved pistil- and pollen-expressed genes.

Acknowledgments We thank Matthew Hegarty, Tom Batstone and Chris Thorogood for constructive comments on earlier drafts of this manuscript, Andrew Hughes for technical support and plant maintenance, Liliana Costa and Chris Thorogood for help with in situ hybridisations. This work was funded by a NERC-CASE studentship to SJH and CL, NERC research grant: NE/D005353/1 to SJH and The Lady Emily Smyth Research Station, University of Bristol.

References

Allen AM (2009) Identification of genes involved in pollen-pistil interactions and self-incompatibility in *Senecio squalidus* L. (Asteraceae). PhD Thesis, University of Bristol

Allen AM, Hiscock SJ (2008) Evolution and phylogeny of self-incompatibility systems in Angiosperms. In: Franklin-Tong VE (ed) Self-incompatibility in flowering plants. Evolution, diversity and mechanisms. Springer, Berlin, pp 73–101

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389–3402

Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knapp SJ, Rieseberg LH (2008) Multiple paleopolyploidizations during the evolution of the compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol Biol Evol 25:2445–2455

Brennan AC, Harris SA, Hiscock SJ (2006) The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae): *S* allele diversity across the British range. Evolution 60:213–224

Cairney J, Zheng L, Cowles A, Hsiao J, Zismann V, Liu J, Ouyang S, Thibaud-Nissen F, Hamilton J, Childs K, Pullman GS, Zhang Y, Oh T, Buell CR (2006) Expressed sequence tags from loblolly pine embryos reveal similarities with angiosperm embryogenesis. Plant Mol Biol 62:485–501

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Filatov DA (2002) Proseq: a software for the preparation and evolutionary analysis of DNA sequence datasets. Mol Ecol Notes 2:621–624

Franklin-Tong VE, Drobak BK, Allan AC, Watkins P, Trewavas AJ (1996) Growth of pollen tubes of *Papaver rhoeas* is regulated by a slow-moving calcium wave propagated by inositol 1, 4, 5-trisphosphate. Plant Cell 8:1305–1321

Hegarty M, Jones J, Wilson I, Barker G, Sanchez-Baracaldo P, Buggs RJA, Liu G, Abbott RJ, Edwards KJ, Hiscock SJ (2005) Development of anonymous cDNA microarrays to study changes to the *Senecio* floral transcriptome during hybrid speciation. Mol Ecol 14:2493–2510

Heslop-Harrison J (1975) Incompatibility and the pollen-stigma interaction. Annu Rev Plant Physiol 26:403–425

Hiscock SJ (2000) Genetic control of self-incompatibility in *Senecio squalidus* L. (Asteraceae): a successful colonizing species. Heredity 85:10–19

Hiscock SJ, Allen AM (2008) Diverse cell signalling pathways regulate pollen–stigma interactions: the search for consensus. New Phytol 179:286–317

Hiscock SJ, McInnis SM (2003) The diversity of self-incompatibility systems in flowering plants. Plant Biol 5:23–32

Hiscock SJ, Hoedemaekers K, Friedman WE, Dickinson HG (2002) The stigma surface and pollen-stigma interactions in *Senecio squalidus* L. (Asteraceae) following cross (compatible) and self (incompatible) pollinations. Int J Plant Sci 163:1–16

Hotelier T, Renault L, Cousin X, Negre V, Marchot P, Chatonnet A (2004) ESTHER, the database of the α/ß-hydrolase fold superfamily of proteins. Nucl Acids Res 32:D145–D147

Jantasuriyarat C, Gowda M, Haller K, Hatfield J, Lu G, Stahlberg E, Zhou B, Li H, Kim H, Yu Y, Dean RA, Wing RA, Soderlund C,

Wang GL (2005) Large scale identification of expressed sequence tags involved in rice and rice blast fungus interaction. Plant Physiol 138:105–115

Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. Nat Protoc 4:363–371

Kräuter-Canham R, Bronner R, Evrad JL, Hahne G, Friedt W, Steinmetz A (1997) A transmitting tissue- and pollen-expressed protein from sunflower with sequence similarity to the human RTP protein. Plant Sci 129:191–202

Kräuter-Canham R, Bronner R, Steinmetz A (2001) SF21 is a protein which exhibits a dual nuclear and cytoplasmic localization in developing pistils of sunflower and tobacco. Ann Bot 87:241–249

Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J Mol Biol 305:567–580

Langdale JA (1993) In situ hybridization. In: Walbot V, Freeling M (eds) The maize handbook. Springer, New York, pp 165–180

Lazarescu E, Friedt W, Horn R, Steinmetz A (2006) Expression analysis of the sunflower SF21 gene family reveals multiple alternative and organ-specific splicing of transcripts. Gene 374:77–86

Nas TMS, Sanchez DL, Diaz MGQ, Mendioro MS, Virmani SS (2005) Pyramiding of thermosensitive genetic male sterility (TGMS) genes and identification of a candidate tms5 gene in rice. Euphytica 145:67–75

Okuda T, Kondoh H (1999) Identification of new genes Ndr2 and Ndr3 which are related to Ndr1/RTP/Drg1 but show distinct tissue specificity and response to N-myc. Biochem Biophys Res Comm 266:208–215

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Spyropoulos IC, Liakopoulos TD, Bagos PG, Hamodrakas SJ (2004) TMRPres2D: high quality visual representation of transmembrane protein models. Bioinformatics 20:3258–3260

Swanson WJ, Vacquier VD (2002) The rapid evolution of reproductive proteins. Nat Rev Genet 3:137–144

Swofford DL (2003) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE 2:e718

Zhang J, Feng Q, Jin C, Qiu D, Zhang L, Xie K, Yuan D, Han B, Zhang Q, Wang S (2005) Features of the expressed sequences revealed by a large-scale analysis of ESTs from a normalized cDNA library of the elite *indica* rice cultivar Minghui 63. Plant J 42:772–780