the FEBS Journal

An unusual plant triterpene synthase with predominant α-amyrin-producing activity identified by characterizing oxidosqualene cyclases from *Malus* × *domestica*

Cyril Brendolise[1], Yar-Khing Yauk[1], Ellen D. Eberhard[1,*], Mindy Wang[1], David Chagne[1], Christelle Andre[1], David R. Greenwood[1,2] and Lesley L. Beuning[1,†]

1 The New Zealand Institute for Plant and Food Research Limited (Plant & Food Research), Auckland, New Zealand
2 School of Biological Sciences, University of Auckland, New Zealand

Keywords
apple; ursolic acid; triterpene synthase; α-amyrin; β-amyrin

Correspondence
D. R. Greenwood, Mt Albert Research Centre, Plant & Food Research, Private Bag 92 169, Auckland 1142, New Zealand
Fax: +64 9 925 7001
Tel: +64 9 925 7147
E-mail: dave.greenwood@plantandfood.co.nz

(Received 9 February 2009, revised 2 May 2011, accepted 10 May 2011)

doi:10.1111/j.1742-4658.2011.08175.x

The pentacyclic triterpenes, in particular ursolic acid and oleanolic acid and their derivatives, exist abundantly in the plant kingdom, where they are well known for their anti-inflammatory, antitumour and antimicrobial properties. α-Amyrin and β-amyrin are the precursors of ursolic and oleanolic acids, respectively, formed by concerted cyclization of squalene epoxide by a complex synthase reaction. We identified three full-length expressed sequence tag sequences in cDNA libraries constructed from apple (*Malus* × *domestica* 'Royal Gala') that were likely to encode triterpene synthases. Two of these expressed sequence tag sequences were essentially identical (> 99% amino acid similarity; *MdOSC1* and *MdOSC3*). MdOSC1 and MdOSC2 were expressed by transient expression in *Nicotiana benthamiana* leaves and by expression in the yeast *Pichia methanolica*. The resulting products were analysed by GC and GC-MS. MdOSC1 was shown to be a mixed amyrin synthase (a 5 : 1 ratio of α-amyrin to β-amyrin). MdOSC1 is the only triterpene synthase so far identified in which the level of α-amyrin produced is > 80% of the total product and is, therefore, primarily an α-amyrin synthase. No product was evident for MdOSC2 when expressed either transiently or in yeast, suggesting that this putative triterpene synthase is either encoded by a pseudogene or does not express well in these systems. Transcript expression analysis in Royal Gala indicated that the genes are mostly expressed in apple peel, and that the *MdOSC2* expression level was much lower than that of *MdOSC1* and *MdOSC3* in all the tissues tested. Amyrin content analysis was undertaken by LC-MS, and demonstrated that levels and ratios differ between tissues, but that the true consequence of synthase activity is reflected in the ursolic/oleanolic acid content and in further triterpenoids derived from them. Phylogenetic analysis placed the three triterpene synthase sequences with other triterpene synthases that encoded either α-amyrin and/or β-amyrin synthase. MdOSC1 and MdOSC3 clustered with the multifunctional triterpene synthases, whereas MdOSC2 was most similar to the β-amyrin synthases.

Database
The sequences reported in this article have been deposited in the DDBJ/EMBL/GenBank databases under the accession numbers FJ032006 (MdOSC1), FJ032007 (MdOSC2) and FJ032008 (MdOSC3)

Abbreviations
APCI, atmospheric pressure chemical ionization; EST, expressed sequence tag; FT, Fourier transform; LG, linkage group; OSC, oxidosqualene cyclase; qPCR, quantitative RT-PCR.

Introduction

The triterpenoids form a large group of structurally diverse natural compounds, many of which are widespread throughout the plant kingdom. Their biological role has not been clearly established; however, a potential antimicrobial activity of their glycosylated derivatives (saponins) suggests a role in protection against pathogens and pests [1–4]. Triterpenoids display a wide range of important medicinal activities, including anti-inflammatory [5,6], antitumour [7], anti-leukaemic [8], anti-HIV [9,10], antifungal [2,11] and antidiabetic [12] activities [13]. Over the years, these promising therapeutic properties have resulted in a great deal of interest in the triterpenoids, and well over 1000 of these natural compounds have been isolated from plants. However, low levels of production and difficulties in purifying these compounds have greatly hampered their commercial exploitation. A better understanding of triterpene biosynthesis is necessary to help to facilitate their biotechnological production and to take advantage of their natural properties.

The first step in the biosynthesis of all triterpenoids and sterols is the cyclization of a 30-carbon precursor, 2,3-oxidosqualene, arising from the isoprenoid pathway [14]. This reaction is catalysed by oxidosqualene cyclases (OSCs = triterpene synthases), and leads to the formation of tricyclic, tetracyclic or pentacyclic molecules in a complex series of concerted reaction steps catalysed by a single enzyme. At this point, the sterol and triterpenoid biosynthetic pathways diverge, depending on the type of OSC involved (Fig. 1). Cyclization of 2,3-oxidosqualene in the chair–boat–chair conformation leads to a protosteryl cation intermediate, the precursor of the sterols via the formation of lanosterol in animals and fungi, or via the formation of cycloartenol or lanosterol in plants [15,16]. In contrast, 2,3-oxidosqualene in the chair–chair–chair conformation is cyclized into a dammarenyl carbocation intermediate, which subsequently gives rise to diverse triterpenoid skeletons after further rearrangements. Many different types of OSC isolated from different species have been characterized in the last few years, including lanosterol synthase [17–20], cycloartenol synthase [21–24], lupeol synthase [25–28], and β-amyrin synthase [24,29–32]. In addition to these enzymes, multifunctional triterpene synthases producing more than one specific compound have also been isolated in plants [23,30,32–38]. More than 100 different carbon skeletons of naturally occurring triterpenes have now been described, suggesting that many other types of OSC have yet to be identified.

The ursane, oleanane and lupane series of triterpenes, derived from α-amyrin, β-amyrin, and lupeol, respectively, are the most widely distributed pentacyclic triterpenes in plants (Fig. 1). These compounds occur particularly in the waxy coating of leaves and on fruits such as apples and pears, where they may serve a protective function in repelling insect or microbial attack [39–41]. These triterpenes all derive from the dammarenyl cation intermediate by a concerted series of methyl and proton shifts. Interestingly, although several OSCs have been reported to produce β-amyrin or lupeol specifically, no enzyme producing α-amyrin

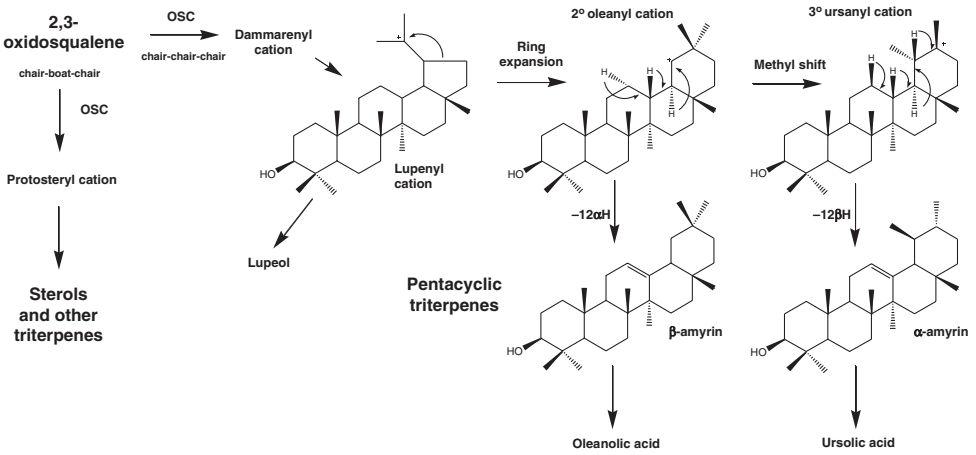


Fig. 1. Simplified scheme of triterpene biosynthesis, showing the concerted reaction sequence for OSCs producing the lupane, oleanane and/or ursane triterpene series. The sterols and other triterpene ring geometries are produced from different conformations of 2,3-oxidosqualene as it binds to the OSC surface template. The differential stability of the secondary oleanane and tertiary ursane cations bound to OSC is likely to affect the ratio of the resulting α/β-amyrin products.

as a sole product has yet been isolated. Multifunctional triterpene synthases accounting for α-amyrin production always appear to yield a combination of compounds including β-amyrin, lupeol or less broadly distributed products such as taraxasterol, butyrospermol and bauerenol in various proportions [27]. Also, as ursane-type triterpenes are always detected together with oleanane-type or lupane-type triterpenes, some authors have suggested that a specific α-amyrin synthase might actually not exist in nature [37]. Apples show a particularly high proportion of ursane-type triterpenes, with ursolic acid (typically 100 mg from a single fruit) and derivatives constituting the majority of the triterpenoid composition in apple peel, although oleanane-type triterpenes are also found [42]. In this study, we describe the identification and partial characterization of three new OSCs from *Malus* × *domestica*, including a novel mixed-amyrin synthase responsible for the production of α-amyrin and β-amyrin, with α-amyrin representing more than 80% of the enzyme product. Gene expression of the three OSCs was measured in various tissues and correlated with the contents of individual triterpenes, including their ultimate biosynthetic products. The molecular and functional evolution of this class of OSC is also discussed.

Results and Discussion

Isolation of apple triterpene synthases and comparison of their amino acid sequences with those of other OSCs

The Plant & Food Research expressed sequence tag (EST) database [43] was searched for putative 2,3-oxidosqualene cyclases by similarity to known triterpene synthases. Three candidates – named *MdOSC1*, *MdOSC2*, and *MdOSC3* – were identified and fully sequenced (Genbank accession nos. FJ032006, FJ032007, and FJ032008, respectively). They were isolated from apple fruit libraries (*MdOSC1* and *MdOSC3*) and a seedling leaf (infected with *Venturia inaequalis*) library (*MdOSC2*). The corresponding cDNAs contained ORFs encoding 760-residue, 762-residue and 760-residue proteins (Fig. 2; *MdOSC1*, *MdOSC2*, and *MdOSC3*, respectively). With 99% similarity at the amino acid level and 95% identity at the DNA level within the coding sequences, MdOSC1 and MdOSC3 seem to be encoded by two different alleles of the same gene. However, mapping analysis of the two sequences using the high-resolution melting technique over a reference 'Malling 9′ × 'Robusta 5′ genetic map (see Experimental procedures) revealed that the markers for *MdOSC1* and *MdOSC3* were located close to simple sequence repeats (SSR) markers CH05c07 and CH02g04, on linkage groups (LGs) 9 and 17, respectively. Recent publication of the apple genome [44] subsequently confirmed that *MdOSC1* and *MdOSC3* are paralogous genes with duplication of loci on LG9 and LG17.

MdOSC2 shares high similarity with β-amyrin synthases (93% with BPY from *Betula platyphylla*, 92% with BgbAS from *Bruguiera gymnorhiza*, and 91% with EtAS from *Euphorbia tirucalli*) and only 78% similarity with MdOSC1 and MdOSC3. The closest homologs of MdOSC1 and MdOSC3 are the lupeol synthases BgLUS (*B. gymnorhiza*) and RcLUS (*Ricinus communis*), with 79% and 78% similarity respectively, and the multifunctional triterpene synthase KcMS from *Kandelia candel*, with 79% similarity. Like other OSCs, MdOSC1, MdOSC2 and MdOSC3 contain the highly conserved SDCTAE motif, which is implicated in substrate binding [45,46] (Fig. 2), and six repeats of the QW motifs [47,48]. It has been suggested that the QW motifs may strengthen the structure of the enzyme and stabilize the carbocation intermediates during cyclization [49]. All of these data suggest that MdOSC1, MdOSC2 and MdOSC3 belong to the triterpene synthase superfamily.

Functional expression of MdOSC1 and MdOSC2

To identify the product specificity of these enzymes, functional expression was carried out using transient expression in *Nicotiana benthamiana* leaves. Because of the strong similarity between MdOSC1 and MdOSC3 at the amino acid level, we decided to focus on MdOSC1 as a potential multifunctional triterpene synthase and the potential β-amyrin synthase (MdOSC2). Triterpene products from transiently expressed MdOSC1 and MdOSC2 were extracted 7 days after *Agrobacterium tumefaciens* infiltration, and analysed by GC. As shown in Fig. 3, the MdOSC1 extract contained two compounds that were not detected in the control plants transformed by the empty vector. These compounds had the same retention times as authentic α-amyrin and β-amyrin on capillary GC. This result was confirmed by coinjection experiments with standard α-amyrin and β-amyrin. Interestingly, α-amyrin was the major compound, produced, with a 5 : 1 ratio to β-amyrin. Under the same conditions, no products were detected in samples extracted from cells transiently expressing MdOSC2. To enhance the α-amyrin and β-amyrin production, the leaf patches previously infiltrated by *Agrobacterium* were infiltrated with either squalene or farnesyl pyrophosphate 4–5 h before extraction. However, no significant improvement in α-amyrin

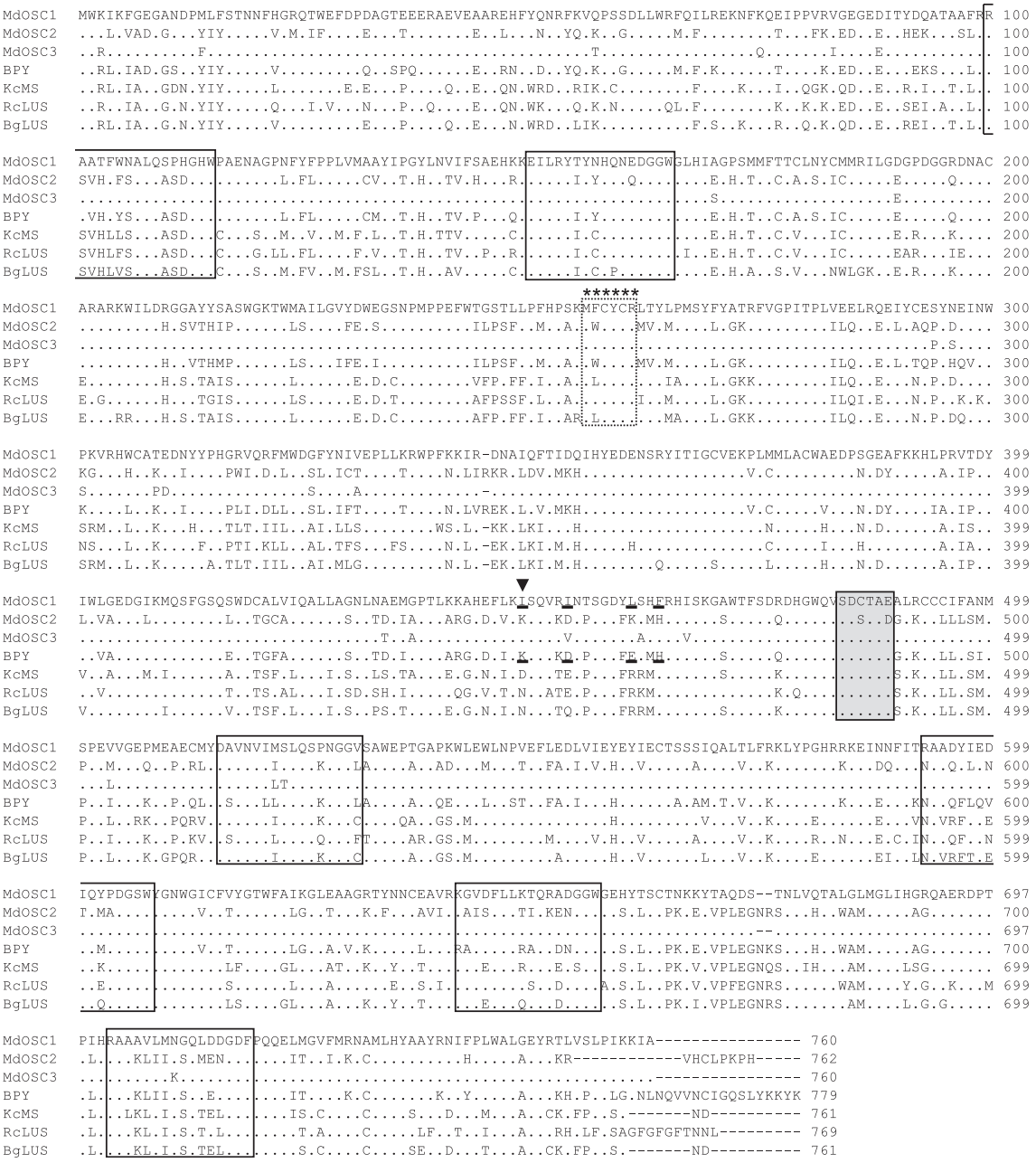


Fig. 2. Comparison of deduced amino acid sequences of MdOSC1, MdOSC2 and MdOSC3 and other plant OSCs [BPY (AB055512), KcMS (AB257507), RcLUS (DQ268869), BgLUS (AB289586)]. Motifs are indicated as follows: QW repeats (clear boxes), SDTAE motif (grey box), MFCYCR motif (stars), Lys449 (arrow), and nonpolar substitutions in MdOSC1 to BPY (bars). Dots represent amino acids that are identical to those in the MdOSC1 sequence.

or β-amyrin production was observed (data not shown). To confirm the identity of the MdOSC1 products, the extracts were further purified and analysed by GC-MS. On the basis of the intensity of ion *m/z* 218, two peaks were detected with the same retention times and the same MS fragmentation patterns as with authentic α-amyrin and β-amyrin (Fig. 4).

To validate these results further, full-length cDNAs (*MdOSC1* and *MdOSC2*) were cloned into a yeast expression vector and transformed into *Pichia methanolica*. Transformants were induced for protein expression and extracted. MdOSC1 and MdOSC2 expression was monitored by SDS/PAGE and staining the gels with colloidal Coomassie Blue. Both proteins

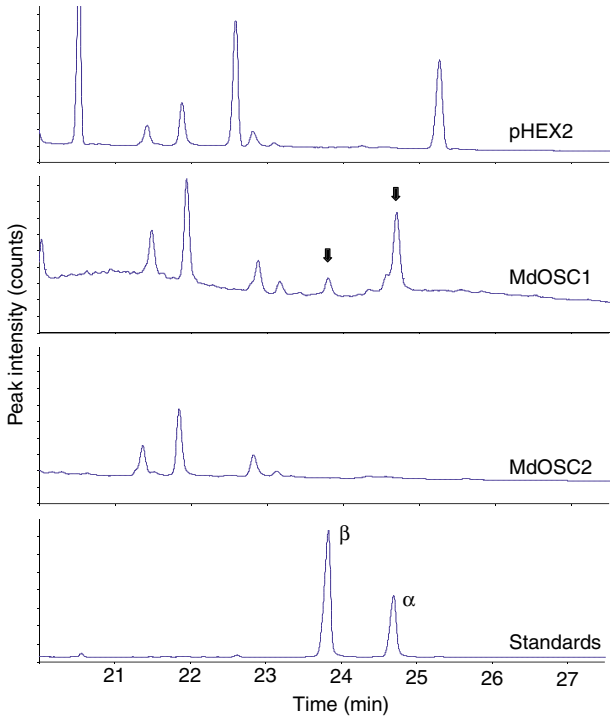


Fig. 3. GC Analysis of MdOSC1 and MdOSC2 transient expression. Products were monitored by flame ionization detector (FID), with pHEX2 empty vector as a negative control and a mixture of α-amyrin and β-amyrin as standards. Arrows indicate peaks with the same retention time as α-amyrin and β-amyrin standards.

were detected from 24 h to 72 h after induction, with no significant increase over time (data not shown). It is noteworthy that the expression level of MdOSC2 was significantly higher than that of MdOSC1 under the same induction conditions. Triterpene products were extracted 48 h and 72 h after induction, and analysed by GC-MS. Consistent with the plant transient expression results, the introduction of *MdOSC1* into yeast resulted in the production of two compounds not detected in the empty vector control extract (Fig. 5). These compounds were identified as α-amyrin and β-amyrin by comparison of their GC retention times and MS fragmentation patterns with authentic standards. These results also confirm that the major compound produced by MdOSC1 is α-amyrin, with a ratio to β-amyrin of 5 : 1, establishing that, unlike other multifunctional triterpene synthases described so far, MdOSC1 has a unique product specificity, with α-amyrin representing more than 80% of the enzyme product. No lupeol was detected under our experimental conditions. Although the expression of MdOSC2 was higher than that of MdOSC1, no triterpene products were detected in the MdOSC2 extracts (data not shown), supporting the transient expression results. This suggests that, although MdOSC2 is strongly related to β-amyrin synthases, this enzyme might actually be involved in the production of terpenes that were not detected under our experimental conditions. Another explanation for the absence of product is that *MdOSC2* could be a pseudogene producing an inactive enzyme. This would imply that MdOSC1 and/or MdOSC3 may account for the entire production of β-amyrin, as no other β-amyrin synthase has been identified yet in apple.

Expression analysis of the apple OSCs

MdOSC1, *MdOSC2* and *MdOSC3* gene expression was analysed by quantitative PCR (qPCR) in root, leaf, apple peel and apple flesh tissues. The data indicated that the three genes have a very similar expression patterns (Fig. 6A–C). The highest level of expression for the three OSCs was measured in apple peel, being up to 40-fold higher than in apple flesh. The levels of expression measured in root were eight-fold, 30-fold and five-fold lower than in the peel for *MdOSC1*, *MdOSC2*, and *MdOSC3*, respectively. Finally, levels of expression in leaf were extremely low for *MdOSC1* and *MdOSC3*, and not even detectable for *MdOSC2*. Relative levels of expression of *MdOSC2* were overall very low in all the tissues tested as compared with *MdOSC1* and *MdOSC3* (Fig. 6D). It is noteworthy that the *MdOSC2* EST was isolated from a *V. inaequalis*-infected seedling leaf library, and yet its transcript could not be detected in healthy leaf tissue, suggesting that this gene could be involved in a defence mechanism against pathogen attack, which triggers its expression. However, such low levels of expression would also be consistent with our hypothesis of it being a pseudogene. The differential expression of *MdOSC1* in apple peel as compared with flesh is consistent with the high level of ursane-type triterpenes present in apple peel, as previously described [42].

Amyrin and other triterpenoid content

Chemical analysis by LC-MS of extracts of Royal Gala apple tissues (Fig. 7A) showed that α-amyrin predominates as the major amyrin form in all tissues except leaves, where it is essentially identical in concentration to β-amyrin. The low expression level of *MdOSC1*, *MdOSC2* and *MdOSC3* in leaves suggest that other, unknown, OSCs might be present in this tissue to account in particular for the β-amyrin production. This is supported by the observation that several additional triterpene skeleton products are detected with accurate mass LC-MS (data not shown)

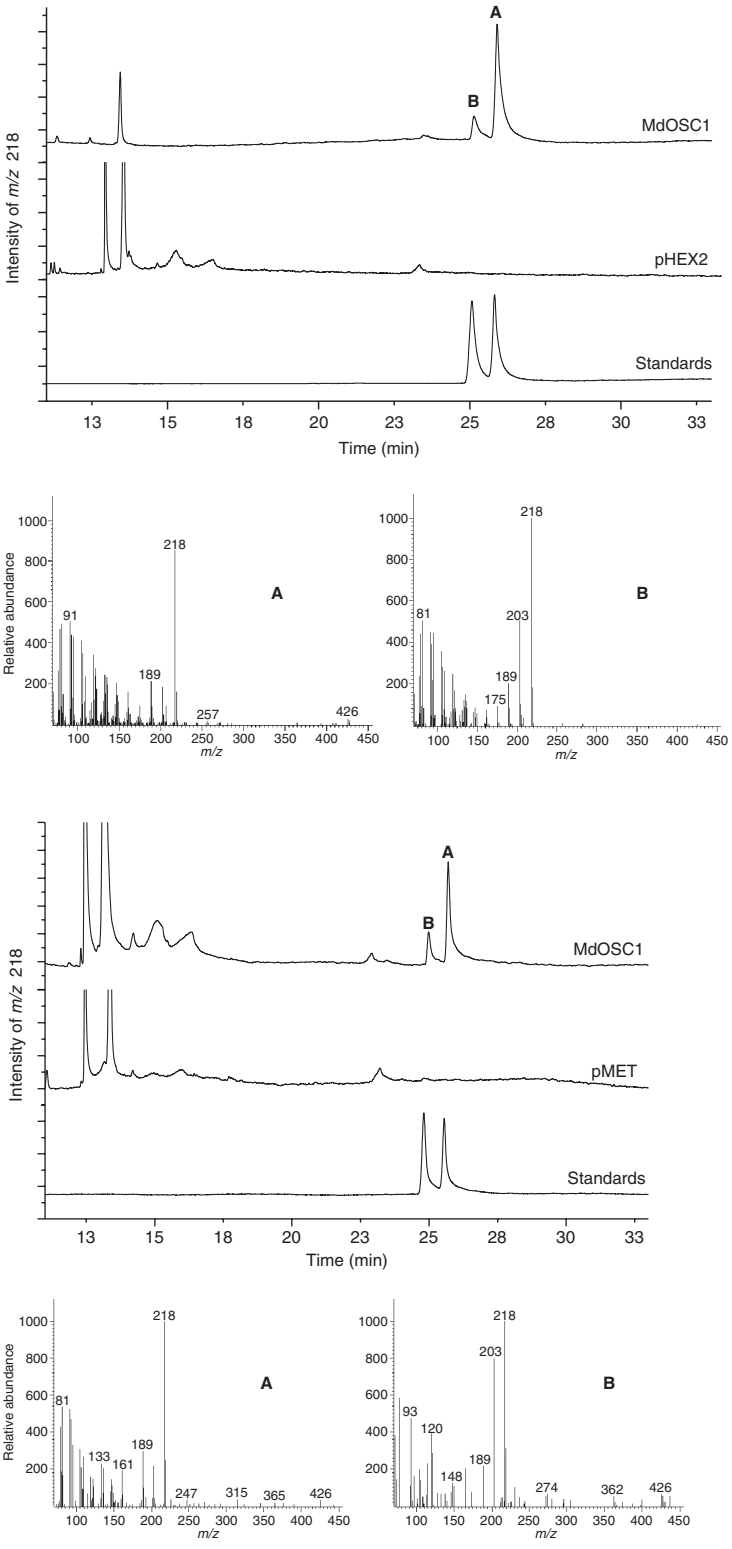


Fig. 4. GC-ToF-MS analysis of MdOSC1 transient expression. Products were monitored on the basis of the intensity of the base peak (*m/z* 218), with pHEX2 empty vector as a negative control and a mixture of α-amyrin and β-amyrin as standard. MS fragmentations of peaks A and B (lower panel) were identical to those of authentic α-amyrin and β-amyrin (data not shown).

Fig. 5. GC-ToF-MS analysis of MdOSC1 expression in yeast. Products were monitored on the basis of the intensity of the base peak (*m/z* 218) with pHEX2 empty vector as a negative control and a mixture of α-amyrin and β-amyrin as standard. MS fragmentations of peaks A and B (lower panel) were identical to those of authentic α-amyrin and β-amyrin (data not shown).

by monitoring the single ion at *m/z* 409.3820–409.3830, which is characteristic of most, if not all, C-30 OSC products. Water is lost on atmospheric pressure chemical ionization (APCI) [50] in positive ion mode, generating a $C_{30}H_{49}^{+}$ ion corresponding to $[M + H\text{-}18]^{+}$ (requiring *m/z* 409.38288) and representing the

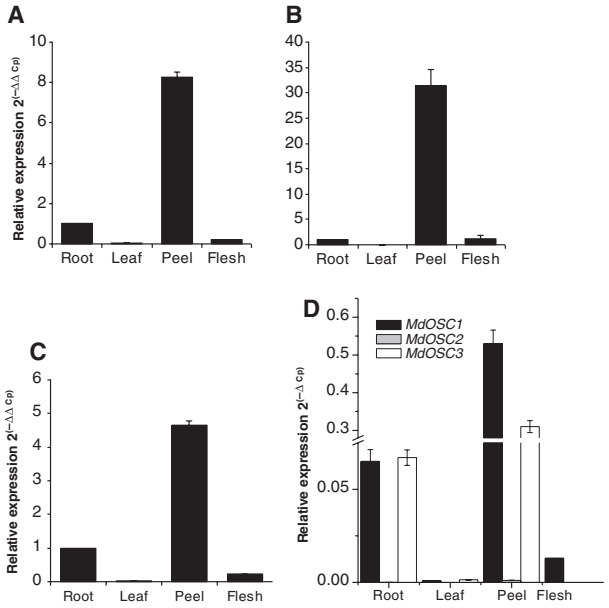


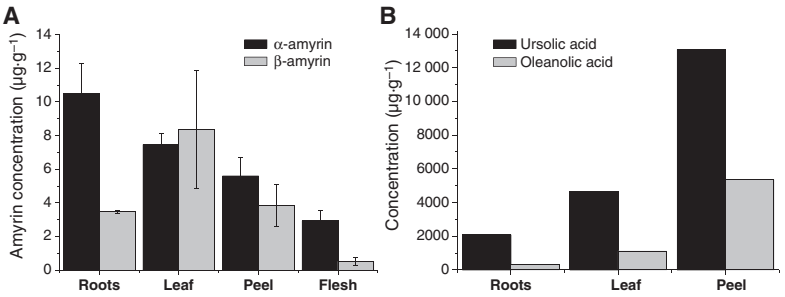
Fig. 6. Expression analysis of the transcripts of apple OSCs in various tissues by qPCR. Primers specific for *MdOSC1* (A), *MdOSC2* (B) and *MdOSC3* (C) were used to measure the levels of transcripts in root, leaf, fruit skin and fruit flesh tissues. Expression is given relative to the apple actin and normalized to the root sample (A, B, C) or not normalized to any sample (D). Error bars represent the standard errors of the means calculated from four technical replicates.

predominant OSC mass-to-charge species detected by Fourier transform (FT) MS. MS^n fragmentation confirmed that these additional 409 ions were related to the amyrins, but without NMR or standards it is not possible to ascribe structural formulae. These additional triterpene synthase products were largely confined to the leaves.

The levels of α-amyrin and β-amyrin in peel do not correlate well with the very high level of expression of *MdOSC1* and *MdOSC3* measured in this tissue. Quantitative analysis of the downstream biosynthetic products of both amyrins indicated a substantial flux of carbon directed into these more polar forms, especially in peel (Fig. 7B,C). The hydroxylated and progressively oxidized (aldehyde, and then acid) products are present at significantly higher levels in peel than in other tissues, although the conditions used for LC-MS did not separate the individual ring E isomers (Fig. 7C). In support of this carbon flux argument is the finding that the ursolic acid level analysed by HPLC is much higher in peel than in any other tissues (Fig. 7B), reflecting the high expression level of *MdOSC1* and *MdOSC3* in this tissue. Interestingly, whereas no ursolic acid could be detected in flesh, the levels measured in all of the other tissues (roots, leaf, and peel) were consistently higher than that of oleanolic acid, confirming that ursane (α-amyrin-derived) products predominate (Fig. 7B). Not shown are further hydroxylated and cinnamate ester derivatives [42] that provide a further sink for amyrin-derived carbon. Overall, the HPLC and LC-MS results agree in relative terms, although the magnitude of the tissue variations is somewhat different.

Phylogenetic analysis of apple OSCs

A phylogenetic tree has been generated on the basis of the deduced amino acid sequences of these proteins



C

Compounds (μg·g $^{-1}$ fresh weight)	Root	Leaf	Peel	Flesh
Uvaol + Oleanol	24.0	22.6	60.1	0.45
Standard deviation	3.4	9.1	21.2	0.29
Uvaal + Oleanal	45.2	40.0	63.9	0.11
Standard deviation	7.7	4.8	24.9	0.07
Ursolic + Oleanolic acids	382.9	1787.5	4509.5	1.48
Standard deviation	63.0	381.8	1763.2	0.43

Fig. 7. Quantitative analysis of amyrins and consequent biosynthetic products in apple root leaf, peel, and flesh. α-Amyrin and β-amyrin were separated and analysed by LC-MS (A); ursolic and oleanolic acids were measured by HPLC (B) and LC-MS together with coeluting biosynthetic intermediates of the ursane and oleanane families (C) expressed as $\mu g.g^{-1}$ of fresh tissue.

along with other members of the OSC superfamily with known function (GenBank). Three main branches can be distinguished, represented by cycloartenol synthases, lupeol synthases, and the dicot β-amyrin synthase-like group (Fig. 8). However, in addition to authentic β-amyrin synthases, this later branch contains other types of triterpene synthase with different product specificities. It includes, in particular, most of the multifunctional triterpene synthases that have been characterized to date, together with several monofunctional enzymes. As most of the enzymes with the same specificity cluster together, some authors have suggested a molecular evolution mechanism for lupeol synthases and the β-amyrin synthase-like group arising from a common ancestral cycloartenol synthase [28,51]. The increasing diversification of the cyclization reaction sequence from the dammarenyl to the oleanyl cation via the lupenyl cation is consistent with this evolutionary scheme.

MdOSC1, MdOSC2 and MdOSC3 are located within the group of enzymes that produce a dammarenyl cation intermediate. MdOSC2 clusters within the authentic β-amyrin synthase subgroup, whereas MdOSC1 and MdOSC3 align with the multifunctional synthase subgroup. In particular, MdOSC1 and MdOSC3 cluster next to the recently described new class of lupeol synthases, which are more related to β-amyrin synthases than to authentic lupeol synthases [52,53]. This new class of lupeol synthases includes BgLUS, RcLUS and the multifunctional triterpene synthase KcMS, and another putative OSC (EtOSC) for which no triterpene synthase activity has been detected when it is expressed in yeast [54]. Although closely related to this group, MdOSC1 and MdOSC3 sit on a distinct branch, and no traces of lupeol could be detected for MdOSC1 in our heterologous expression experiments. This suggests that MdOSC1 has already diverged sufficiently to acquire a different specificity.

Several examples of subtle changes responsible for drastic modifications of OSC specificities have been reported [55]. For instance, Kushiro *et al.* [31] have demonstrated, using site-directed mutagenesis experiments on the β-amyrin synthase PNY, that the Trp residue in the MWCYCR(256–261) motif is crucial for β-amyrin specificity, and that, instead, a Leu at this position is characteristic of all functional lupeol synthases. More recently, RcLUS, which belongs to the new class of lupeol synthases, has been shown to harbour a Phe instead of Leu at this position [MFCYCR(256–261)]. Interestingly, MdOSC1 and MdOSC3 also have a Phe, whereas MdOSC2 has conserved the intact MWCYCR(246–261) motif, which is characteristic of β-amyrin synthases (Fig. 2). Lys449 (in BPW) is another example of a key residue that has been reported to be present in all specific β-amyrin synthase sequences, whereas it is replaced by an Ala or Asn in all specific lupeol synthases [52]. This rule, however, becomes more questionable with respect to multifunctional synthases, for which some exceptions occur. For instance, in *Arabidopsis*, At1g78500 produces lupeol as a main product, despite having a Lys at position 449. Also, the MdOSC1 sequence has a hydrophobic residue at the corresponding position (Ile448), and yet is able to proceed into the E-ring expansion towards synthesis of α-amyrin and β-amyrin. Interestingly, the region in the vicinity of Ile448 in MdOSC1 contains several other radical amino acid changes as compared with monofunctional β-amyrin synthases; these include the replacement of basic and acidic amino acids with nonpolar residues (Fig. 2), which would probably have a drastic effect on the enzyme specificity. Additional amino acid substitutions are scattered along the sequence of MdOSC1 as compared with β-amyrin synthases; however, addressing their significance will require further studies using, for instance, site-directed mutagenesis and/or domain swapping approaches.

These observations suggest that the branch point between lupeol, α-amyrin and β-amyrin synthesis involves several regions along the protein backbone; and although a point mutation can radically modify the specificity, it is likely that several sequence modifications counterbalance each other without modifying enzyme specificity (for instance, members of the two different groups of lupeol synthases have the same product specificity despite being phylogenetically distant; likewise, OEA and PSM within the β-amyrin synthase-like group have an identical product pattern while sharing only 74% similarity). Consequently, sequence comparisons and phylogenetic analysis, although providing information on enzyme relationships, cannot accurately predict the enzyme specificity within this particular subfamily of OSCs that produce the damarenyl cation intermediate. This OSC subfamily shows significant postspeciation expansion that leads to a large diversity of triterpene skeletons. This is consistent with the postulated role in pathogen or disease resistance of several of its members, as it would be advantageous for plants producing new compounds with enhanced efficiency to select for such beneficial traits. In this context, multifunctional triterpene synthases may represent ongoing evolutionary mechanisms for transition of one specificity to another. In contrast, members of the CAS subfamily that have more core housekeeping functions as precursors of sterols and plant hormones have undergone very little postspeciation expansion, and remain very similar to one another.

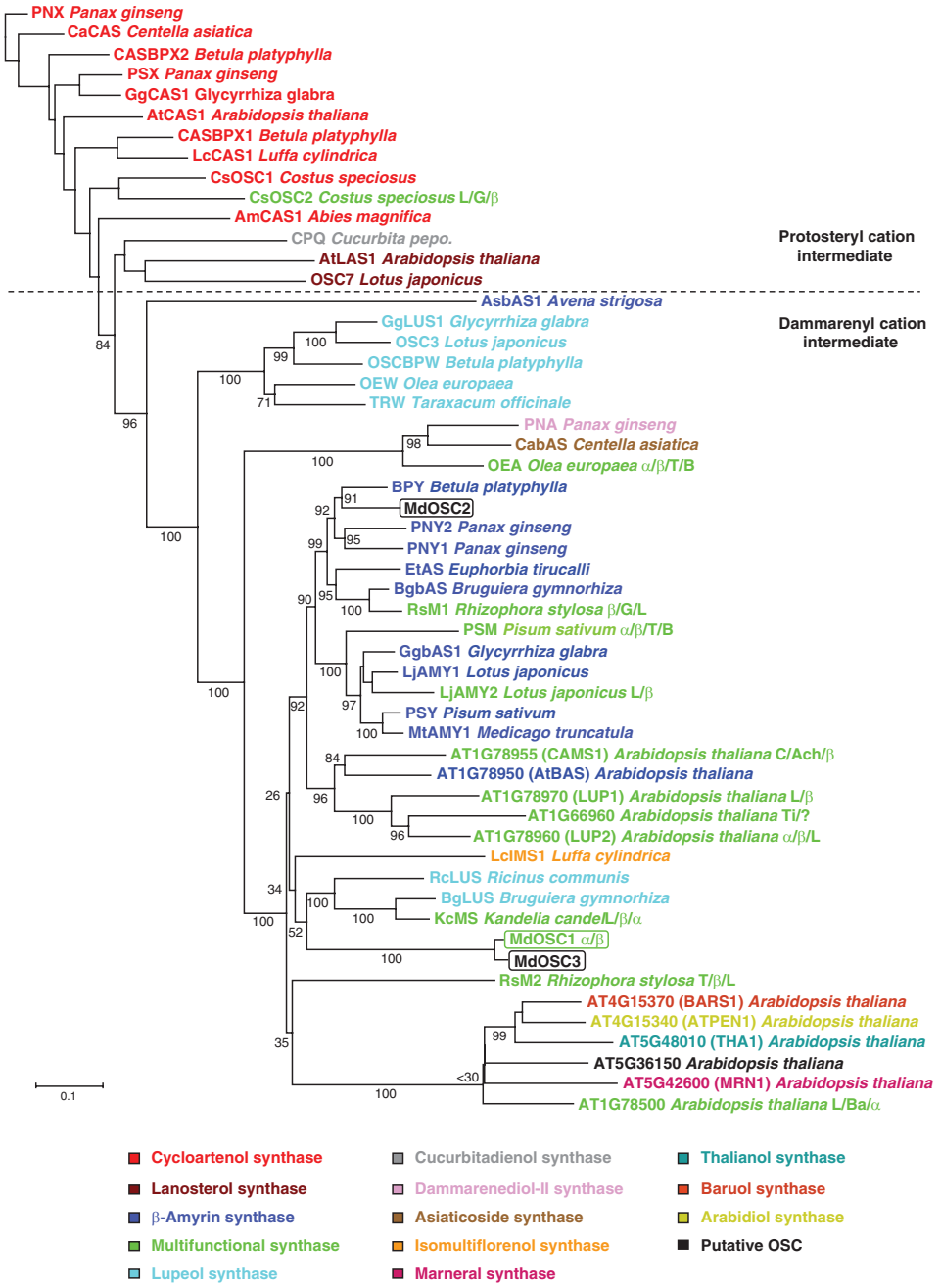


Fig. 8. Phylogenetic tree of plant OSCs. Deduced amino acid sequences were aligned with CLUSTALX. Protein distances were calculated with PROTDIST and the Jones–Taylor–Thornton matrix of the PHYLIP package. The tree was constructed by the neighbour-joining method, and visualized in TREEVIEW (version 1.6.6). Numbers indicate the bootstrap support for each node (1000 replicates). The scale represents 0.1 amino acid substitutions per site. The catalytic specificities of the OSCs are indicated by different colours. Compounds produced by the multifunctional OSCs are indicated as follows: α, α-amyrin; β, β-amyrin; Ba, bauerenol; B, butyrospermol; G, germanicol; L, lupeol; T, taraxosterol; Ti, tirucalla-7,21-diene-3β-ol; ?, unknown. DDBJ/GenBank/EMBL accession numbers used in this analysis are indicated in Experimental procedures.

Concluding remarks

No monofunctional α-amyrin synthase has been identified to date in the plant kingdom. As apple (*Malus* × *domestica*) contains a high level of ursane-type triterpenes [42], we decided to isolate and characterize the triterpene synthases present in this organism. Three new OSC genes were identified from the Plant & Food Research apple EST database. MdOSC1 and MdOSC3, sharing more than 99% similarity, cluster

within the multifunctional triterpene synthase subgroup. Using two different expression systems, we have shown that MdOSC1 is a mixed amyrin synthase responsible for the synthesis of α-amyrin and β-amyrin, with a ratio of 5 : 1, and therefore is unusual in favouring α-amyrin synthesis. Unfortunately, sequence comparison of MdOSC1 with other known OSCs did not provide clear evidence of the catalytically important residues responsible for the higher level of production of α-amyrin. Further functional and structural analysis will be needed to identify these regions.

The phylogenetic analysis of MdOSC2 suggested that this enzyme could be a β-amyrin synthase. However, as no product was detected in either of our two expression systems, we hypothesized that this enzyme has already evolved into either a form that is inactive or that it produces compounds that were not detected under our experimental conditions. This hypothesis would imply either that MdOSC1 and MdOSC3 account on their own for the production of β-amyrin and α-amyrin in apple, or that there are other triterpene synthases yet to be identified in this plant species. Recent publication of the *Malus* genome should facilitate the identification of new candidate genes [44]. The additional triterpene compounds detected by FT-MS and detailed by He and Liu [42] are likely to be formed by less specific OSC regioselectivity, but why these products should be confined principally to the leaves is unknown. Surprisingly, expression levels of all MdOSCs are low in leaves, suggesting that either further triterpene synthases remain to be identified or transport may be occurring between tissues. The high levels of OSC gene expression and corresponding biosynthetic products in apple peel, particularly of the ursane series, are particularly significant for the purported effects of ursane terpenoids in producing a range of health benefits [42]. Leaving the skin on before consumption of apples would ensure that appreciable amounts of ursane terpenoids are consumed.

Experimental procedures

Isolation and cloning of the apple triterpene synthase cDNAs

OSC candidates were identified from the Plant & Food Research apple 'Royal Gala' EST database [43] by a search (BLAST) for known similarity with known triterpene synthases from Genbank. *MdOSC1*, *MdOSC2* and *MdOSC3* originate from the AASB, ABEA and ABCA cDNA libraries respectively, as described in [43].

qPCR analysis

Total RNA was isolated from apple tissues, *Malus* × *domestica* 'Royal Gala', following a method adapted from that described by Chang *et al*. [56]. Following DNase treatment, reverse transcription was performed in 20-μL reactions with 750 ng of RNA, oligod(T) primers and SuperScript III RNase H-reverse transcriptase, according to the manufacturer's instructions (Invitrogen, Auckland, New Zealand). qPCR amplifications were carried out with a LightCycler 480 (Roche Diagnostics, Mannheim, Germany). Reactions were performed four times, with 1.25 μL of 50-fold diluted cDNA, 2.5 μL of 2 × LightCycler 480 SYBR Green Master Mix (Roche Diagnostics) and 0.5 μM specific primers to a final volume of 5 μL. The specific primers used were as follows: *MdOSC1* forward, 5′-TTGT ACTACTAATCCAGTGATCAAGATGTGG-3′; *MdOSC1* reverse, 5′-CTCTCTTAGTATCTGAAAACGCCATAGG AG-3′; *MdOSC2* forward, 5′-CGCAGATGGTGGCAATG ATCCATACATC-3′; *MdOSC2* reverse, 5′-TGAAGTTCT TCTCCCTTAAGAACTGCATTC-3′; *MdOSC3* forward, 5′-GCAATCGTGATCAAAGAAGATGTGGAGG-3′; and *MdOSC3* reverse, 5′-TTCTCTTAAAATCTGAAAACGCC ATAGG-3′. Amplification conditions included an initial denaturation step of 95 ºC for 5 min, followed by 45 cycles of 95 ºC for 10 s, 60 ºC for 10 s, and 72 ºC for 12 s. Fluorescence was measured at the end of each annealing step, and this was followed by a melting curve analysis with continual fluorescence acquisition from 65 to 95 °C to check for single product amplification. Negative water controls were included in each run for each set of primers. Data were analysed with LIGHTCYCLER software version 1.5.0.39. For each gene, a standard curve was generated with serial dilutions of the initial cDNA reaction, and the resultant primer efficiencies were used in the relative expression analysis. Expression was calculated relative to *Malus* × *domestica* actin (*MdActin*, accession number CN938023) to minimize variations in cDNA template levels. Figure 6A–C shows relative quantification using the root values as calibrator and set to a nominal value of 1, whereas the data shown in Fig. 6D have not been calibrated to compare the differential expression between the three OSCs. Error bars shown in the qPCR data represent the standard errors of the means calculated from the four technical replicates.

Mapping analysis

PCR primers were designed within *MdOSC1* (forward, 5′-GGACTGCACATAGCGGGG-3′; reverse, 5′-CCACGGT CAAGAATCCACTT-3′) and *MdOSC3* (forward, 5′-GGA CTGCACATATCAGGC-3′; reverse, 5′-AGTTTTTCCCC ATGATGCAG-3′) sequences to amplify a 137-bp and a 171-bp fragment, respectively. The high-resolution melting technique was used to detect sequence polymorphisms

within the amplified fragments, as described by Chagné *et al.* [57]. The new markers were screened over the bin mapping set of the reference 'Malling 9' × 'Robusta 5' genetic map [58].

Phylogenetic analysis

Phylogenetic analysis was conducted with programs from the PHYLIP package [59]. Deduced amino acid sequences of MdOSC1, MdOSC2 and MdOSC3, together with other members of the OSC superfamily, were aligned with CLUSTALX (version 1.83) [60]. The DDBJ/GenBank/EMBL accession numbers of the sequences used are as follows: AB055512 (BPY, *Be. platyphylla*), AB206469 (EtAS, *E. tirucalli*), AB014057 (PNY2, *Panax ginseng*), AB009030 (PNY1, *P. ginseng*), AB257507 (KcMS, *K. candel*), NM_106545 (ATLUP2, *Arabidopsis thaliana*), NM_106546 (ATLUP1, *A. thaliana*), NM_148667 (CAMS1/AT1G78955, *A. thaliana*), AB374428 (AtBAS/AT1G78950, *A. thaliana*), AB265170 (PNA, *P. ginseng*), AB291240 (OEA, *Olea europaea*), AB034803 (PSM, *Pisum sativum*), DQ268869 (RcLUS, *R. communis*), NM_126681 (AtCAS1, *A. thaliana*), AB037203 (GgbAS1, *Glycyrrhiza glabra*), AB009029 (PNX, *P. ginseng*), AB034802 (PSY *Pi. sativum*), D89619 (PSX, *Pi. sativum*), AF489920 (AT1G66960, *A. thaliana*), BT020312 (AT5G42600, *A. thaliana*), NM_117625 (AT4G15370/BARS1, *A. thaliana*), NM_106497 (AT1G78500, *A. thaliana*), NM_124175 (AT5G48010, *A. thaliana*), NM_117622 (AT4G15340/ATPEN1, *A. thaliana*), NM_123006 (AT5G36150/ATPEN3, *A. thaliana*), AF478455 (LjAMY2, *Lotus japonicus*), AB181244 (LjAMY1, *L. japonicus*), AF478453 (MtAMY1, *Medicago truncatula*), AF216755 (AmCAS1, *Abies magnifica*), AB055509 (CASBPX1, *Be. platyphylla*), AB055510 (CASBPX2, *Be. platyphylla*), AB025968 (GgCAS1, *G. glabra*), AB058507 (CsOSC1, *Costus speciosus*), AY520819 (CaCAS, *Centella asiatica*), AB033334 (LcCAS1, *Luffa cylindrica*), AJ311789 (AsbAS1, *Avena strigosa*), AB058508 (CsOSC2, *C. speciosus*), AB055511 (OSCBPW, *Be. platyphylla*), AB116228 (GgLUS1, *G. glabra*), AB025343 (OEW, *O. europaea*), AB025345 (TRW, *Taraxacum officinale*), AB058643 (LcIMS1, *Lu. cylindrica*), NM_114382 (AtLAS1, *A. thaliana*), AB244671 (OSC7, *L. japonicus*), AB116238 (CPQ, *Cucurbita pepo*), AY520818 (CabAS, *Ce. asiatica*), AB263203 (RsM1, *Rhizophora stylosa*), AB263204 (RsM2, *Rh. stylosa*), AB289585 (BgbAS, *B. gymnorhiza*), AB289586 (BgLUS, *B. gymnorhiza*), and AB181245 (OSC3, *L. japonicus*). Protein distances were calculated with the Jones–Taylor–Thornton matrix (PROTDIST). The tree was generated by the neighbour-joining method, with 1000 bootstrap replications, and visualized in TREEVIEW (version 1.6.6) [61]. Very similar trees were obtained with the parsimony program (PROTPARS) or the Kitsch and the Fitch construction methods.

Transient expression in *N. benthamiana*

Full-length cDNAs were cloned into the pHEX2 binary vector and transformed into *Ag. tumefaciens* strain GV3101(MP90) [62]. Freshly grown transformed cells were resuspended in 10 mL of infiltration medium (10 mM $MgCl_2$, 10 μM acetosyringone) to a $D_{600\ nm}$ of 0.5–0.6, and incubated at room temperature for 2 h without shaking before infiltration [62]. After 7 days of plant growth under standard greenhouse conditions, infiltrated leaves were removed and soaked in 15 mL of methanol overnight. When mentioned, squalene (emulsified in water) and farnesyl pyrophosphate (250 μM solution) were infiltrated 4–5 h prior to leaf removal. Plant material was removed from the methanol, which was evaporated down to 5 mL and applied to a C-18 LC column (Accubond II; Agilent, Auckland, New Zealand). The column was washed with 20 mL of methanol, and the resulting eluent was evaporated and resuspended in 0.5 mL of ethyl acetate for GC analysis. For the GC-MS analysis, 0.2-mL aliquots were evaporated down to 10 μL, dissolved in 400 μL of hexane, and applied to an LC-Si SPE cartridge (Supelco; Sigma-Aldrich, Auckland, New Zealand). The cartridge was washed successively with 1 mL of hexane/1% acetic acid with increasing amounts of ethyl acetate (ranging from 5% to 100%) in acidified hexane. TLC analysis of the fractions indicated that α-amyrin and β-amyrin compounds coeluted in the hexane/10% ethyl acetate/1% acetic acid fraction (data not shown). Fractions containing α-amyrin and β-amyrin were evaporated to dryness, redissolved in 100 μL of ethyl acetate, and analysed by GC-MS.

Yeast expression

Full-length cDNAs were cloned in the pMET A yeast expression vector (Invitrogen), under control of the AUG1 promoter, and transformed into *Pic. methanolica* PMAD16. Cells were grown in MDY medium (1% yeast extract, 2% peptone, 1.34% Yeast Nitrogen Base, 4 × 10e-5% biotin, 2% dextrose) at 30 °C with shaking (~ 2 ***g***) until saturation, and then collected and resuspended in MMY medium (1% yeast extract, 2% peptone, 1.34% YNB, 4 × 10e-5% biotin, 0.5% methanol) for induction. Upon induction, cells were harvested at different time points over a period of up to 72 h, and extracted overnight with a solvent mix of 1 : 2 : 1 methanol/dichloromethane/water. Extracts were evaporated to dryness, resuspended in 10 μL of ethyl acetate, mixed with 400 μL of hexane, applied to an LC-Si SPE column (Supelco) as described above, and analysed by GC-MS.

GC and GC-MS analysis

GC analysis was performed on an HP 5890 Series II gas chromatograph fitted with a flame ionization detector (FID)

and a Zebron ZB-5 column (30 m × 0.2 mm and 0.25-μm-thick bonded coating; Phenomenex, Auckland, New Zealand). The column temperature was set at 40 °C for 1 min, raised to 320 °C at a rate of 20 °C·min^{-1}, and held for 15 min. Hydrogen total flow rate was set to 50 mL·min^{-1} and 12 psi, producing a linear velocity of 30 cm·s^{-1} through the column. Injector and detector temperatures were set at 200 and 225 °C, respectively. GC-MS analysis was performed on a LECO Pegasus III GC-ToF mass spectrometer with the same column and conditions as described above, and helium as the carrier gas. Peaks were identified by similarity search in mass-spectrum libraries (NIST and Wiley). α-Amyrin and β-amyrin (Extrasynthèse, Genay, France) were used as authentic standards.

Triterpene content analysis

Snap-frozen tissue samples (0.5 g of root, leaf, and peel, and 2 g of flesh, all in triplicate) were ground in a pestle and mortar with liquid nitrogen, and extracted in 95% ethanol with a CEM Discover microwave. Samples were refluxed for 20 min in 50-mL RB flasks, with the temperature maintained at 85 °C and a power setting of 100 W. Extracts were rotary evaporated to dryness, dissolved in 20% aqueous methanol, and applied to pre-equilibrated Maxi-Clean 300-mg large-pore 100-Å C-18 SPE cartridges (Alltech Associates, Grace Davison Discovery Sciences, Auckland, New Zealand), and washed with 4 mL of 20% aqueous methanol. Compounds were eluted with 4 mL of 80% methanol. Eluates were evaporated under nitrogen and redissolved in 80% methanol for HPLC and LC-MS.

HPLC

Aliquots of each tissue extract were applied to an Agilent 1200 HPLC, using a Hypersil C-18 column (25 × 4.6 mm, 5μm particle size; Agilent) with an isocratic solvent mix of 88% methanol and 0.05% phosphoric acid, and monitoring of elution at a wavelength of 210 nm. Elution times were consistent with authentic α-amyrin and β-amyrin standards, which were used to calibrate the quantitation of both compounds.

High-resolution accurate mass tandem MS

Extract samples (10 μL) were subjected to LC APCI tandem MS with a ThermoFinnigan LTQ-FT mass spectrometer coupled to a Surveyor HPLC system. Compounds were separated by reversed-phase chromatography on a Phenomenex Luna (150 × 2 mm, 5-μm particle size) with a Security Guard cartridge, and elution over 60 min with a 20–95% aqueous acetonitrile gradient containing 20 mM ammonium acetate and 0.1% formic acid. The eluant entered the APCI source of the LTQ-FT at a flow rate of 100 μL·min^{-1}. The vaporizator was set at 480 °C with a nitrogen sheath gas flow of 40 units, an auxiliary gas flow of 5 units, and sweep gas flow of 1 unit. The mass spectrometer was operated in the positive ion mode, with helium as the collision gas and a mass range acquired over *m/z* 100–700. The capillary temperature was set at 175 °C, the capillary voltage was 9 V, the tube lens voltage was 100 V, and the source voltage was set at 4.2 kV. Data were acquired in centroid mode in a top 3 × 2 experiment [one full scan in the ICR cell (parallel mode) followed by two averaged MS^2 scans of each of the top three ions recorded in the ion trap, followed by MS^3 scans of the two most intense scans from each MS^2 spectrum] in data-dependent mode with dynamic exclusion enabled. Full-scan FT data were obtained at a resolution of 50 000 at *m/z* 400. An isolation width of 2 a.m.u. and a normalized collision energy of 35 were used for both MS^2 and MS^3 fragmentation stages. A maximum ion time of 500 ms was used for collecting ions in the ion cyclotron resonance cell.

Extracted ion chromatograms and associated MS^n spectra were examined to determine the elution profile, atomic composition, and quantitation and fragmentation patterns of the eluting species as required. The choice of column and elution conditions enabled separation of α-amyrin and β-amyrin and resolution of uvaol, uvaal, ursolic and its higher hydroxylated cinnamate ester derivatives [42] from one another, but not from the corresponding oleanane series. Extracted ion chromatograms were centred on the monoisotopic mass of the compound with a mass window of < 5 p.p.m. Quantitation was performed as the area under each peak by reference to cholesterol as an internal standard.

Acknowledgements

We would like to thank all members of the Plant & Food Research Genomics Programme, including the bioinformatics, cloning and sequencing teams. Our thanks go to W. Laing and D. Rowan for reviewing the manuscript. This work was funded by Plant & Food Research internal capability funding.

References

1 Agrell J, Oleszek W, Stochmal A, Olsen M & Anderson P (2003) Herbivore-induced responses in alfalfa (Medicago sativa). *J Chem Ecol* **29**, 303–320.
2 Johann S, Oliveira VL, Pizzolatti MG, Schripsema J, Braz R, Branco A & Smania A (2007) Antimicrobial activity of wax and hexane extracts from Citrus spp. peels. *Mem Inst Oswaldo Cruz* **102**, 681–685.
3 Osbourn A (1996) Saponins and plant defence – a soap story. *Trends Plant Sci* **1**, 4–9.

4 Papadopoulou K, Melton RE, Leggett M, Daniels MJ & Osbourn AE (1999) Compromised disease resistance in saponin-deficient plants. *Proc Natl Acad Sci USA* **96**, 12923–12928.
5 Ikeda Y, Murakami A & Ohigashi H (2008) Ursolic acid: an anti- and pro-inflammatory, triterpenoid. *Mol Nutr Food Res* **52**, 26–42.
6 Safayhi H & Sailer ER (1997) Anti-inflammatory actions of pentacyclic triterpenes. *Planta Med* **63**, 487–493.
7 Banno N, Akihisa T, Tokuda H, Yasukawa K, Taguchi Y, Akazawa H, Ukiya M, Kimura Y, Suzuki T & Nishino H (2005) Anti-inflammatory and antitumor-promoting effects of the triterpene acids from the leaves of Eriobotrya japonica. *Biol Pharm Bull* **28**, 1995–1999.
8 Kokpol U, Chittawong V & Miles DH (1986) Chemical constituents of the roots of Acanthus illicifolius. *J Nat Prod* **49**, 355–356.
9 Kashiwada Y, Takanaka K, Tsukada H, Miwa Y, Taga T, Tanaka S & Ikeshiro Y (2001) Sesquiterpene glucosides from anti-leukotriene B-4 release fraction of Taraxacum officinale. *J Asian Nat Prod Res* **3**, 191–197.
10 Kashiwada Y, Wang HK, Nagao T, Kitanaka S, Yasuda I, Fujioka T, Yamagishi T, Cosentino LM, Kozuka M, Okabe H *et al.* (1998) Anti-AIDS agents. 30. Anti-HIV activity of oleanolic acid, pomolic acid, and structurally related triterpenoids. *J Nat Prod* **61**, 1090–1095.
11 Johann S, Soldi C, Lyon JP, Pizzolatti MG & Resende MA (2007) Antifungal activity of the amyrin derivatives and in vitro inhibition of Candida albicans adhesion to human epithelial cells. *Lett Appl Microbiol* **45**, 148–153.
12 Guerrero-Analco J, Medina-Campos O, Brindis F, Bye R, Pedraza-Chaverri J, Navarrete A & Mata R (2007) Antidiabetic properties of selected Mexican copalchis of the Rubiaceae family. *Phytochemistry* **68**, 2087–2095.
13 Liu J (1995) Pharmacology of oleanolic acid and ursolic acid. *J Ethnopharmacol* **49**, 57–68.
14 Abe I, Rohmer M & Prestwich GD (1993) Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. *Chem Rev* **93**, 2189–2206.
15 Kolesnikova MD, Xiong QB, Lodeiro S, Hua L & Matsuda SPT (2006) Lanosterol biosynthesis in plants. *Arch Biochem Biophys* **447**, 87–95.
16 Suzuki M, Xiang T, Ohyama K, Seki H, Saito K, Muranaka T, Hayashi H, Katsube Y, Kushiro T, Shibuya M *et al.* (2006) Lanosterol synthase in dicotyledonous plants. *Plant Cell Physiol* **47**, 565–571.
17 Baker CH, Matsuda SPT, Liu DR & Corey EJ (1995) Molecular cloning of the human gene encoding lanosterol synthase from a liver cDNA library. *Biochem Biophys Res Commun* **213**, 154–160.
18 Corey EJ, Lee JM & Liu DR (1994) First demonstration of a carbocation–olefin cyclization route to the lanosterol series. *Tetrahedron Lett* **35**, 9149–9152.
19 Shi Z, Buntel CJ & Griffin JH (1994) Isolation and characterization of the gene encoding 2,3-oxidosqualene lanosterol cyclase from *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* **91**, 7370–7374.
20 Sung CK, Shibuya M, Sankawa U & Ebizuka Y (1995) Molecular cloning of cDNA encoding human lanosterol synthase. *Biol Pharm Bull* **18**, 1459–1461.
21 Bach TJ & Benveniste P (1997) Cloning of cDNAs or genes encoding enzymes of sterol biosynthesis from plants and other eukaryotes: heterologous expression and complementation analysis of mutations for functional characterization. *Prog Lipid Res* **36**, 197–226.
22 Corey EJ, Matsuda SPT & Bartel B (1993) Isolation of an Arabidopsis thaliana gene encoding cycloartenol synthase by functional expression in a yeast mutant lacking lanosterol synthase by the use of a chromatographic screen. *Proc Natl Acad Sci USA* **90**, 11628–11632.
23 Kawano N, Ichinose K & Ebizuka Y (2002) Molecular cloning and functional expression of cDNAs encoding oxidosqualene cyclases from Costus speciosus. *Biol Pharm Bull* **25**, 477–482.
24 Kushiro T, Shibuya M & Ebizuka Y (1998) beta-Amyrin synthase – cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *Eur J Biochem* **256**, 238–244.
25 Hayashi H, Huang P, Takada S, Obinata M, Inoue K, Shibuya M & Ebizuka Y (2004) Differential expression of three oxidosqualene cyclase mRNAs in Glycyrrhiza glabra. *Biol Pharm Bull* **27**, 1086–1092.
26 Herrera JB, Bartel B, Wilson WK & Matsuda SP (1998) Cloning and characterization of the Arabidopsis thaliana lupeol synthase gene. *Phytochemistry* **49**, 1905–1911.
27 Segura MJR, Meyer MM & Matsuda SPT (2000) Arabidopsis thaliana LUP1 converts oxidosqualene to multiple triterpene alcohols and a triterpene diol. *Org Lett* **2**, 2257–2259.
28 Shibuya M, Zhang H, Endo A, Shishikura K, Kushiro T & Ebizuka Y (1999) Two branches of the lupeol synthase gene in the molecular evolution of plant oxidosqualene cyclases. *Eur J Biochem* **266**, 302–307.
29 Hayashi H, Huang PY, Kirakosyan A, Inoue K, Hiraoka N, Ikeshiro Y, Kushiro T, Shibuya M & Ebizuka Y (2001) Cloning and characterization of a cDNA encoding beta-amyrin synthase involved in glycyrrhizin and soyasaponin biosyntheses in licorice. *Biol Pharm Bull* **24**, 912–916.
30 Iturbe-Ormaetxe I, Haralampidis K, Papadopoulou K & Osbourn AE (2003) Molecular cloning and characterization of triterpene synthases from Medicago truncatula and Lotus japonicus. *Plant Mol Biol* **51**, 731–743.
31 Kushiro T, Shibuya M, Masuda K & Ebizuka Y (2000) Mutational studies on triterpene syntheses: engineering

lupeol synthase into beta-amyrin synthase. *J Am Chem Soc* **122**, 6816–6824.
32 Morita M, Shibuya M, Kushiro T, Masuda K & Ebizuka Y (2000) Molecular cloning and functional expression of triterpene synthases from pea (Pisum sativum) – new alpha-amyrin-producing enzyme is a multifunctional triterpene synthase. *Eur J Biochem* **267**, 3453–3460.
33 Basyuni M, Oku H, Inafuku M, Baba S, Iwasaki H, Oshiro K, Okabe T, Shibuya M & Ebizuka Y (2006) Molecular cloning and functional expression of a multifunctional triterpene synthase cDNA from a mangrove species Kandelia candel (L.) Druce. *Phytochemistry* **67**, 2517–2524.
34 Ebizuka Y, Katsube Y, Tsutsumi T, Kushiro T & Shibuya M (2003) Functional genomics approach to the study of triterpene biosynthesis. *Pure Appl Chem* **75**, 369–374.
35 Husselstein-Muller T, Schaller H & Benveniste P (2001) Molecular cloning and expression in yeast of 2,3-oxidosqualene-triterpenoid cyclases from Arabidopsis thaliana. *Plant Mol Biol* **45**, 75–92.
36 Kushiro T, Shibuya M, Masuda K & Ebizuka Y (2000) A novel multifunctional triterpene synthase from Arabidopsis thaliana. *Tetrahedron Lett* **41**, 7705–7710.
37 Saimaru H, Orihara Y, Tansakul P, Kang YH, Shibuya M & Ebizuka Y (2007) Production of triterpene acids by cell suspension cultures of Olea europaea. *Chem Pharm Bull* **55**, 784–788.
38 Shibuya M, Xiang T, Katsube Y, Otsuka M, Zhang H & Ebizuka Y (2007) Origin of structural diversity in natural triterpenes: direct synthesis of seco-triterpene skeletons by oxidosqualene cyclase. *J Am Chem Soc* **129**, 1450–1455.
39 Baker EA (1982) Chemistry and morphology of plant epicuticular waxes. In *The Plant Cuticle* (Cutler DF, Alvin KL & Price CE eds), pp. 139–165. Academic Press, London.
40 Belding RD, Blankenship SM, Young E & Leidy RB (1998) Composition and variability of epicuticular waxes in apple cultivars. *J Am Soc Hortic Sci* **123**, 348–356.
41 Bringe K, Schumacher CFA, Schmitz-Eiberger M, Steiner U & Oerke EC (2006) Ontogenetic variation in chemical and physical characteristics of adaxial apple leaf surfaces. *Phytochemistry* **67**, 161–170.
42 He XJ & Liu RH (2007) Triterpenoids isolated from apple peels have potent antiproliferative activity and may be partially responsible for apple's anticancer activity. *J Agric Food Chem* **55**, 4366–4370.
43 Newcomb RD, Crowhurst RN, Gleave AP, Rikkerink EHA, Allan AC, Beuning LL, Bowen JH, Gera E, Jamieson KR, Janssen BJ *et al.* (2006) Analyses of expressed sequence tags from apple. *Plant Physiol* **141**, 147–166.
44 Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D *et al.* (2010) The genome of the domesticated apple (Malus × domestica Borkh.). *Nat Genet* **42**, 833–839.
45 Abe I & Prestwich GD (1994) Active site mapping of affinity-labeled rat oxidosqualene cyclase. *J Biol Chem* **269**, 802–804.
46 Abe I & Prestwich GD (1995) Identification of the active site of vertebrate oxidosqualene cyclase. *Lipids* **30**, 231–234.
47 Poralla K (1994) The possible role of a repetitive amino acid motif in evolution of triterpenoid cyclases. *Bioorg Med Chem Lett* **4**, 285–290.
48 Poralla K, Hewelt A, Prestwich GD, Abe I, Reipen I & Sprenger G (1994) A specific amino acid repeat in squalene and oxidosqualene cyclases. *Trends Biochem Sci* **19**, 157–158.
49 Wendt KU, Lenhart A & Schulz GE (1999) The structure of the membrane protein squalene-hopene cyclase at 2.0 angstrom resolution. *J Mol Biol* **286**, 175–187.
50 Rhourri-Frih B, Chaimbault P, Claude B, Lamy C, Andre P & Lafosse M (2009) Analysis of pentacyclic triterpenes by LC-MS. A comparative study between APCI and APPI. *J Mass Spectrom* **44**, 71–80.
51 Zhang H, Shibuya M, Yokota S & Ebizuka Y (2003) Oxidosqualene cyclases from cell suspension cultures of Betula platyphylla var. japonica: molecular evolution of oxidosqualene cyclases in higher plants. *Biol Pharm Bull* **26**, 642–650.
52 Basyuni M, Oku H, Tsujimoto E, Kinjo K, Baba S & Takara K (2007) Triterpene synthases from the Okinawan mangrove tribe, Rhizophoraceae. *FEBS J* **274**, 5028–5042.
53 Guhling O, Hobl B, Yeats T & Jetter R (2006) Cloning and characterization of a lupeol synthase involved in the synthesis of epicuticular wax crystals on stem and hypocotyl surfaces of Ricinus communis. *Arch Biochem Biophys* **448**, 60–72.
54 Kajikawa M, Yamato KT, Fukuzawa H, Sakai Y, Uchida H & Ohyama K (2005) Cloning and characterization of a cDNA encoding beta-amyrin synthase from petroleum plant Euphorbia tirucalli L. *Phytochemistry* **66**, 1759–1766.
55 Segura MJR, Jackson BE & Matsuda SPT (2003) Mutagenesis approaches to deduce structure–function relationships in terpene synthases. *Nat Prod Rep* **20**, 304–317.
56 Chang S, Puryear J & Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* **11**, 113–116.
57 Chagné D, Gasic K, Crowhurst RN, Han Y, Bassett HC, Bowatte DR, Lawrence TJ, Rikkerink EHA, Gardiner SE & Korban SS (2008) Development of a set of

SNP markers present in expressed genes of the apple. *Genomics* **92**, 353–358.
58 Celton JM, Tustin D, Chagné D & Gardiner S (2009) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from Malus ESTs and Pyrus genomic sequences. *Tree Genet Genomes* **5**, 93–107.
59 Felsenstein J (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166.
60 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
61 Page RDM (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comp Appl Biosci* **12**, 357–358.
62 Voinnet O, Rivas S, Mestre P & Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J* **33**, 949–956.