



# Evolution of Structural Diversity of Triterpenoids

Pablo D. Cárdenas<sup>†</sup>, Aldo Almeida<sup>†</sup> and Søren Bak<sup>\*</sup>

Department of Plant and Environmental Science, University of Copenhagen, Frederiksberg, Denmark

## OPEN ACCESS

### Edited by:

Kazuki Saito,  
RIKEN Center for Sustainable  
Resource Science (CSRS),  
Japan

### Reviewed by:

Tetsuo Kushiro,  
Meiji University, Japan  
Tessa Moses,  
University of Edinburgh,  
United Kingdom

### \*Correspondence:

Søren Bak  
bak@plen.ku.dk

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Plant Metabolism and  
Chemodiversity,  
a section of the journal  
Frontiers in Plant Science

**Received:** 19 July 2019

**Accepted:** 01 November 2019

**Published:** 17 December 2019

### Citation:

Cárdenas PD, Almeida A and  
Bak S (2019) Evolution of Structural  
Diversity of Triterpenoids.  
*Front. Plant Sci.* 10:1523.  
doi: 10.3389/fpls.2019.01523

Plants have evolved to produce a blend of specialized metabolites that serve functional roles in plant adaptation. Among them, triterpenoids are one of the largest subclasses of such specialized metabolites, with more than 14,000 known structures. They play a role in plant defense and development and have potential applications within food and pharma. Triterpenoids are cyclized from oxidized squalene precursors by oxidosqualene cyclases, creating more than 100 different cyclical triterpene scaffolds. This limited number of scaffolds is the first step towards creating the vast structural diversity of triterpenoids followed by extensive diversification, in particular, by oxygenation and glycosylation. Gene duplication, divergence, and selection are major forces that drive triterpenoid structural diversification. The triterpenoid biosynthetic genes can be organized in non-homologous gene clusters, such as in *Avena* spp., Cucurbitaceae and *Solanum* spp., or scattered along plant chromosomes as in *Barbarea vulgaris*. Paralogous genes organized as tandem repeats reflect the extended gene duplication activities in the evolutionary history of the triterpenoid saponin pathways, as seen in *B. vulgaris*. We review and discuss examples of convergent and divergent evolution in triterpenoid biosynthesis, and the apparent mechanisms occurring in plants that drive their increasing structural diversity within and across species. Using *B. vulgaris*' saponins as examples, we discuss the impact a single structural modification can have on the structure of a triterpenoid and how this affect its biological properties. These examples provide insight into how plants continuously evolve their specialized metabolome, opening the way to study uncharacterized triterpenoid biosynthetic pathways.

**Keywords:** triterpenoid saponins, structural diversity, convergent evolution, plant specialized metabolism, unlinked versus clustered pathways

## INTRODUCTION

Plants have evolved the strategy of constantly diversifying the chemical structures they produce, evolving into an astonishing number of so-called plant specialized metabolites—of which the majority are thought to be involved in plant defense. The general conception has been that divergent evolutionary processes is the major driving force. However, biochemical and molecular knowledge of how these compounds have evolved has demonstrated that convergent evolution is surprisingly common (Pichersky and Lewinsohn, 2011).

Triterpenoids are one of the largest groups of plant specialized metabolites with over 14,000 structures known (Hamberger and Bak, 2013; <http://dnp.chemnetbase.com>). In the current model of triterpenoid biosynthesis, the isoprenoid precursor 2,3-oxidosqualene is cyclized by signature enzymes oxidosqualene cyclases (OSCs) to a number of cyclical triterpene scaffolds. Subsequently,

these structures are oxygenated by various cytochromes P450 (P450s) and finally glycosylated by UDP-glycosyltransferases (UGTs) (Figure 1). Plant genomes contain about a dozen OSC, about 250 P450s and about 110 UGTs, which are involved in multiple pathways. The sheer size of these gene families makes it a challenge to determine the specific genes coding for enzymes involved in generating the diversity of triterpenoids based on sequence alone. The OSC step constitutes the key branching point leading to biosynthesis of either sterols and steroidal saponins or to non-steroidal triterpenoids (e.g., triterpenoid saponins) (Thimmappa et al., 2014) (Figure 1). It is widely accepted that the cyclization of 2,3-oxidosqualene into sterols, occurs *via* the protosteryl cation and proceeds in a “chair-boat-chair” conformation. On the other hand, triterpenoid biosynthesis, *via* the dammarenyl cation, involves “chair-chair-chair” conformation of 2,3-oxidosqualene. In plants, more than 100 different triterpene scaffolds are derived from 2,3-oxidosqualene by OSCs, with lupeol and  $\beta$ -amyrin being among the most wide spread and studied regarding their biosynthesis and role in plants (Xu et al., 2004; Stephenson et al., 2019).

The increasing availability of genome sequencing of evolutionary distant plant species and in-depth studies of triterpenoid pathways in recent years, has revealed an emerging picture that triterpenoid pathways are evolving recurrently and thus represent an interesting case of both divergent and convergent evolution. Typically, it is the same gene families (OSCs, P450s, and UGTs) that are constantly recruited and afterwards expanded through gene duplications leading to tandem repeats. Here, we review the current knowledge about five triterpenoid scaffolds (and their derivatives) as examples of both divergent and convergent evolution in specialized metabolism and discuss the selection pressure on the genes, their organization, and how plants cope with the toxic triterpenoids they produce.

## CONVERGENT EVOLUTION IN BIOSYNTHESIS OF TRITERPENOID SCAFFOLDS

### Dammarenyl Cation-Derived Triterpenoids Have Arisen Multiple Times During Angiosperm Evolution

The majority of the known triterpenoids arise from the dammarenyl cation (Xiong et al., 2005). Resins from trees of the Dipterocarpaceae are known as dammar, hence the first triterpenoids isolated from dammar resins were coined dammaranediol I and II (Mills and Werner, 1955; Mills, 1956). Currently, the most commercially-valuable representatives of this class of compounds are the diverse ginsenosides. They can be obtained from mature six-year-old rhizomes of *Panax ginseng* and because of their anticancer activities some ginsenosides are used for chemotherapy treatment (Leung and Wong, 2010). Tansakul et al. (2006) were the first to clone a Dammarenyl-II synthase (PgDDS) and showed that this OSC from *P. ginseng* catalyzes the first committed step of ginsenoside biosynthesis. The only other DDS characterized thereafter

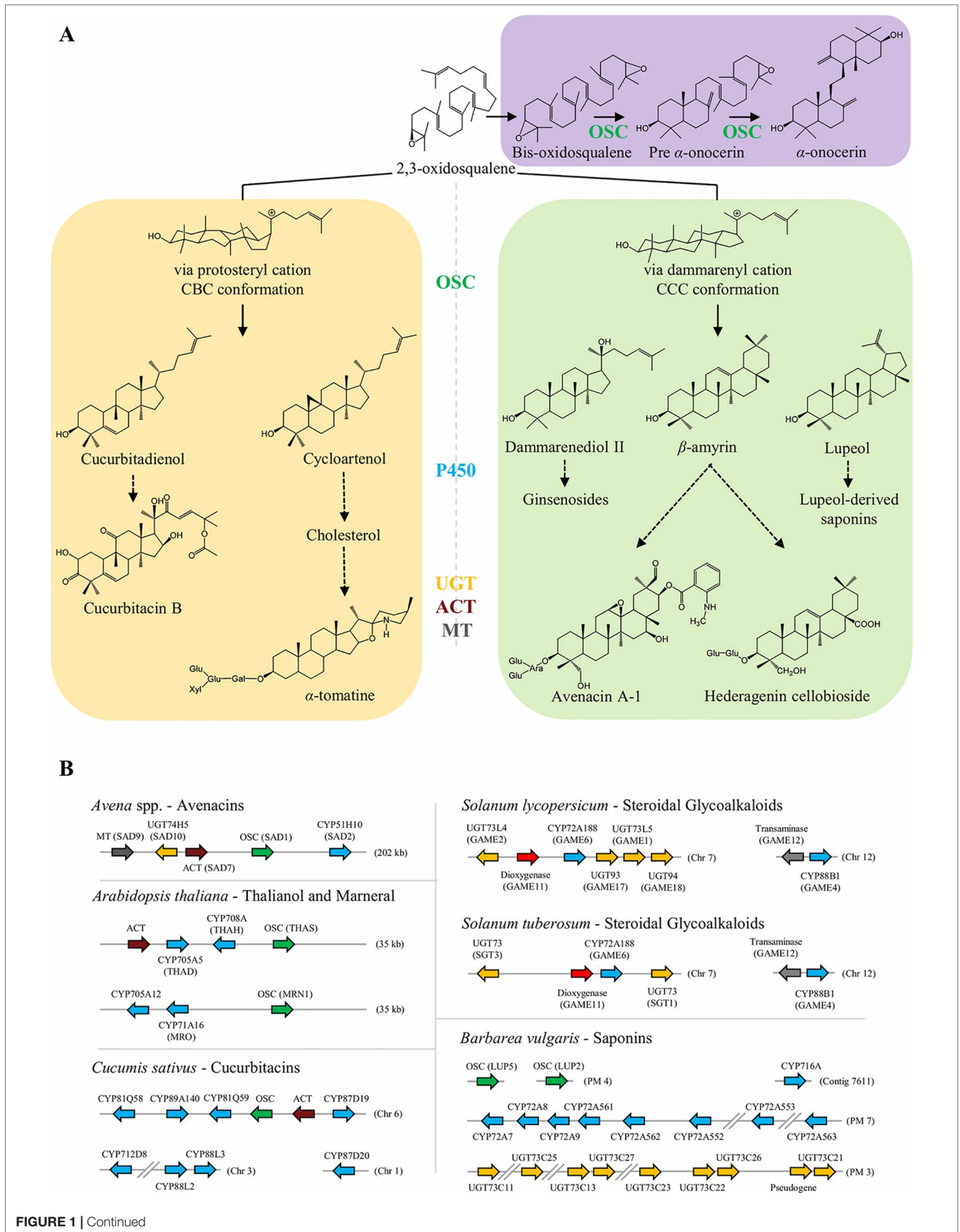
belongs to *Centella asiatica* (Kim et al., 2009). *P. ginseng* and *C. asiatica* belong to the Apiales and phylogenetic analysis showed that both DDS grouped in the same branch suggesting the DDS in these species evolved from a common ancestor. As DDSs have not been elucidated from the phylogenetically distant Dipterocarpaceae, future work on this family will shed some light on the evolutionary history of DDSs and could indicate if they arose from convergent or divergent evolution.

Lupeol and  $\beta$ -amyrin are prevalent pentacyclic triterpenoids derived from the dammarenyl cation and they are ubiquitously found in many different plant species (Figure 1). Nevertheless, phylogenetic analysis have shown that the genes producing these scaffolds group distinctively in different clades. Shibuya et al. (1999) first distinguished two clades of lupeol synthases in plants; one which is composed of specific lupeol synthases and another which is composed of multi-functional OSCs producing  $\alpha$ -,  $\beta$ -amyrin, and lupeol (Thimmappa et al., 2014; Khakimov et al., 2015). Site-directed mutagenesis experiments have shown that a single amino acid replacement could convert a lupeol synthase into a  $\beta$ -amyrin synthase (and conversely), indicating the noticeable role of specific residues may have played in the evolution of OSC product specificity and generation of triterpenoid diversity (Kushiro et al., 1999; Kushiro et al., 2000). Furthermore, phylogenetic analysis of both monocot and dicot OSCs by Xue et al. (2012) and Augustin et al. (2011), additionally distinguish two distinctive clades of  $\beta$ -amyrin synthase in monocots and dicots.

Lupeol and  $\beta$ -amyrin can be present in plants as unmodified compounds typically found in resins or waxes (Szakiel et al., 2012) or they have a major role as precursors for other specialized triterpenoid metabolites, usually involved in plant defense and development. Lupeol is involved in nodule formation in *Lotus japonicus* through regulation of *ENOD40* gene expression (Delis et al., 2011). Lupeol is also part of the cuticular wax surface of castor bean plant (*Ricinus communis*) where it was suggested to have a physiological role in protection against dehydration (Guhling et al., 2006). Additionally, lupeol is the precursor for betulinic acid, a triterpenoid that accumulates in the bark of birch tree and has a potent anticancer activity (Pisha et al., 1995; Kumar et al., 2018).  $\beta$ -amyrin is a precursor of glycyrrhizin, a triterpenoid saponin found in the roots and stolons of liquorice (*Glycyrrhiza glabra*) well known for its pharmaceutical properties and as a natural sweetener (Hayashi et al., 2002).  $\beta$ -amyrin is also precursor for the antifungal saponin avenacin A-1 in oat and for the potent insect-feeding deterrent hederagenin glycosides in the Brassicaceae *Barbarea vulgaris* (Kuzina et al., 2009; Nielsen et al., 2010; Khakimov et al., 2016; Liu et al., 2019). Additionally,  $\beta$ -amyrin seems to play a role in root development in oat (Kemen et al., 2014) and in *Lotus japonicus* (Krokida et al., 2013), suggesting that triterpenoids like lupeol and  $\beta$ -amyrin are not exclusively involved in plant defense.

### $\alpha$ -Onocerin—A Seco-Triterpenoid (*sensu lato*) That Evolved Convergently

Seco-triterpenoids are characterized by the absence of a C-C bond that would normally form one of the consecutive rings



**FIGURE 1** | Simplified representation of the biosynthesis of sterols and triterpenoids in plants. **(A)** OSC signature enzymes catalyze the cyclization of 2,3-oxidosqualene, and in more rare cases bis-oxidosqualene, into several triterpenoid scaffolds. These structures can be further modified by tailoring enzymes, including oxygenation by P450s, glycosylation by UGTs, acylation by ACT, and methylation by MT. Selected structures are depicted and discussed in more detail in the text. Dashed arrows represent multiple biosynthetic reactions whereas solid arrows represent a single step. **(B)** Biosynthesis of plant triterpenoids can be mediated by non-homologous clustered genes or through non-linked genes. In *Avena* spp., a cluster of five genes are involved in the biosynthesis of avenacin A-1. In *Arabidopsis thaliana*, two clusters have been reported: thalianol cluster with four genes (up) and marneral cluster with three genes (down). In *Cucumis sativus*, six genes associated with cucurbitacin biosynthesis are located in a cluster in chromosome 6, while four other genes are elsewhere in the genome. The core genes for biosynthesis of SGAs are clustered in chromosome 7 and 12 of *S. tuberosum* and *S. lycopersicum*, respectively. The key genes for biosynthesis of the insect-feeding deterrent hederagenin cellobioside are distributed along *B. vulgaris* genome in tandem repeats located at different pseudomolecules (PM). OSC, oxidosqualene cyclase; P450, cytochrome P450; UGT, UDP-glycosyltransferase; ACT, acyltransferase; MT, methyltransferase.

in a triterpenoid scaffold. They are known to be distributed across the plant kingdom. The *seco*-triterpenoid (*sensu lato*)  $\alpha$ -onocerin consists of two bicyclic systems connected by a two-carbon linkage; its occurrence appears to be limited to Lycopod species and some species within the *Ononis* genus in the Fabaceae. Lycopods and the Fabaceae originated in very distant evolutionary times (Garratt et al., 1984; Giraud and Lejal-Nicol, 1989), which implies that the  $\alpha$ -onocerin trait evolved convergently in Lycopods and in the *Ononis* genus. The biological function of  $\alpha$ -onocerin still remains unknown.

On a biochemical level,  $\alpha$ -onocerin biosynthesis differs from other triterpenoids as it is biosynthesized from 2,3;22,23-oxidosqualene (bis-oxidosqualene) instead of the typical triterpenoid precursor 2,3-oxidosqualene (Figure 1). In *Lycopodium clavatum*, biosynthesis of  $\alpha$ -onocerin is carried out in two sequential steps by two paralogous OSCs (Araki et al., 2016). Pre- $\alpha$ -onocerin synthase (LcLCC), initiates cyclization from one of the epoxide bonds of bis-oxidosqualene. The cyclization terminates after formation of the A and B rings through the generation of pre- $\alpha$ -onocerin. Subsequently  $\alpha$ -onocerin synthase (LcLCD) carries out the cyclization of the D and E rings from the remaining epoxide ring, to yield  $\alpha$ -onocerin. Our group demonstrated that *Ononis spinosa*  $\alpha$ -onocerin is biosynthesized by a single OSC (OsONS1) (Almeida et al., 2018). In *O. spinosa* a neofunctionalized squalene epoxidases (OsSQEs) provide the OSCs with the necessary bis-oxidosqualene. Fluorescence imaging microscopy experiments demonstrated protein-protein interactions between OsONS1 and the neofunctionalized OsSQEs (Almeida et al., 2018).

Phylogenetic analysis revealed that OsONS1 branches off from lupeol synthases, and is phylogenetically distant from the two *L. clavatum*  $\alpha$ -onocerin synthases which branch off directly from sterol biosynthesis; thus,  $\alpha$ -onocerin biosynthesis evolved convergently in these two plant species (Almeida et al., 2018). In addition, molecular docking simulation showed that OsONS1 alone produces  $\alpha$ -onocerin in two cyclization steps, as opposed to the specialized LcLCC and LcLCD in Lycopods.

## Protosteryl Cation-Derived Triterpenoids Are Found Across Angiosperm Phylogeny

Cucurbitacins are highly oxygenated tetracyclic triterpenes initially discovered in members of the Cucurbitaceae family and well known for their bitterness and toxicity (Metcalf et al., 1980). Cucurbitacins have been reported in 17 taxonomically distant related families in eudicots and some monocots. They consist of hundreds of derivatives from 20 main cucurbitacin molecules named from cucurbitacin A to T (Chen et al., 2005). At least 100 species in 30 genera of the Cucurbitaceae have been shown to contain cucurbitacins (Raemisch

and Turpin, 1984), with cucurbitacin B (Figure 1A) being present in 91% of the species (Chen et al., 2005).

Cucurbitacins are extremely toxic to mammals (David and Vallance, 1955) and function as feeding deterrents against several insects (Nielsen et al., 1977; Bigger and Chaney, 1998) but they have also been found to be feeding stimulant for leaf beetles belonging to the Luperini tribe of the Chrysomelidae (Metcalf, 1986). Cucurbitacin B can displace the insect steroidal hormone 20-hydroxyecdysone, thus affecting morphological changes in *Drosophila melanogaster* (Dinan et al., 1997).

The cucurbitacin biosynthesis pathway has mainly been studied in members of the Cucurbitaceae. The OSC cucurbitadienol synthase from *Cucurbita pepo* (CpCPQ) cyclizes 2,3-oxidosqualene to cucurbitadienol (Figure 1A) (Shibuya et al., 2004). The CpCPQ gene evolved after the divergence of dicots and monocots when the ancestral cycloartenol synthase gene duplicated, creating two clades of cycloartenol synthases (CASI and CASII) of which CpCPQ evolved from the *C. pepo* cycloartenol synthase (CpCPX) in the CASII clade (Xue et al., 2012).

Cucurbitacins are also present outside the Cucurbitaceae; for example *Iberis amara* belonging to the Brassicaceae contain cucurbitacins (Nielsen et al., 1977). The elucidation of the cucurbitacin biosynthetic pathway in *Iberis* will help to clarify the evolutionary history of Cucurbitacin biosynthesis.

Steroidal glycoalkaloids (SGAs) are triterpene-derived compounds found in major dicot Solanaceae crops such as tomato and potato. SGAs contain a nitrogen incorporated on their steroidal scaffold and they provide the plant with a barrier against a broad range of herbivores and pathogens, but are bitter and considered anti-nutritional compounds for humans (Cárdenas et al., 2015). SGAs and non-nitrogenous steroidal saponins are also present in distantly related monocot Liliaceae plant species. In these plant species, they have had special attention due to their pharmacological properties. For instance, the Liliaceae *Veratrum californicum* produces the potent anticancer molecule cyclopamine (Augustin et al., 2015). Biochemical and phylogenetic analysis of the enzymes involved in the biosynthesis of *Solanum* and *Veratrum* SGAs suggested their convergent origins and their partial recruitment from primary phytosterol metabolism (Augustin et al., 2015; Sonawane et al., 2016).

## GENOME ORGANIZATION OF TRITERPENOID PATHWAYS

The biosynthetic genes for triterpenoids display different genome organizations across plant species (Figure 1B). While

in some species they are arranged in clusters of non-homologous genes (e.g., avenacins in oat, Qi et al., 2004; cucurbitacins in cucumber, Shang et al., 2014); the biosynthesis of other triterpenoids is mediated by genes scattered along the plant genome organized typically in tandem repeats (e.g., saponins in *Barbarea vulgaris*, Khakimov et al., 2015; Erthmann et al., 2018; Liu et al., 2019; and mogrosides in *Siraitia grosvenorii*, Itkin et al., 2016).

## Clustered Genes Mediating Biosynthesis of Avenacins, Cucurbitacins and Steroidal Glycoalkaloids

Organization of genes in operon-like clusters has typically been associated with fungal genomes and operons present in bacteria. Nevertheless, nearly 20 metabolic gene clusters for plant specialized metabolites have been reported in multiple plant species (Boycheva et al., 2014). The avenacin, thalianol, and marneral triterpenoid biosynthetic pathways were among the first to be reported being organized in non-homologous gene clusters in oat and *Arabidopsis thaliana*, respectively (Figure 1B, Qi et al., 2004; Field and Osbourn, 2008; Field et al., 2011). The organization of metabolic genes in clusters is thought to provide both co-inheritance, co-regulation and avoiding accumulation of toxic intermediates and may thus be beneficial for securing stable inheritance of functional chemical defense pathways in a dynamic ecological context of natural populations (Osborn, 2010; Takos and Rook, 2012).

Triterpenoid saponins are found mainly in dicotyledonous species. *Avena* spp. is the only known triterpenoid saponin producing monocotyledon. *Avena* spp. accumulates avenacins which are saponins produced in the roots of oat species and provide the plant with a potent barrier against soil-borne fungi (Papadopoulou et al., 2002) (Figure 1). SAD1, a  $\beta$ -amyrin synthase, catalyzes the first committed step in avenacin biosynthesis (Qi et al., 2004), and is present in the oat genome in an operon like cluster together with genes coding for the tailoring enzymes required for avenacin biosynthesis, including: SAD2, the P450 CYP51H10 that oxidizes  $\beta$ -amyrin; and three genes that act together in the acylation steps of avenacin: SAD9, a N-methyltransferase; SAD10, a UGT74H5 glycosyltransferase; and SAD7, a serine carboxypeptidase-like acyltransferase (Mugford et al., 2013). CYP51H10 (SAD2) evolved from a predisposition in sterol metabolism of the orthologous CYP51 catalyzing the conserved 14 $\alpha$ -demethylation reaction in sterol metabolism in eukaryotes (Geisler et al., 2013).

In Cucurbitaceae, part of the cucurbitacin pathway is similarly clustered; in cucumber the cucurbitadienol synthase is flanked by four P450s and an ACT gene out of which two P450s and the ACT have been functionally characterized (Shang et al., 2014). The P450s in this cluster belong to different subfamilies, indicating this cluster was not formed by tandem duplications but by genome reorganization. Later Zhou et al. (2016) uncovered the same conserved syntenic loci in melon and watermelon by comparative analyses of their genomes. While the core cluster is overall syntenically conserved, the CYP88A60 catalyzing the hydroxylation at position C19 and the CYP87D20 catalyzing

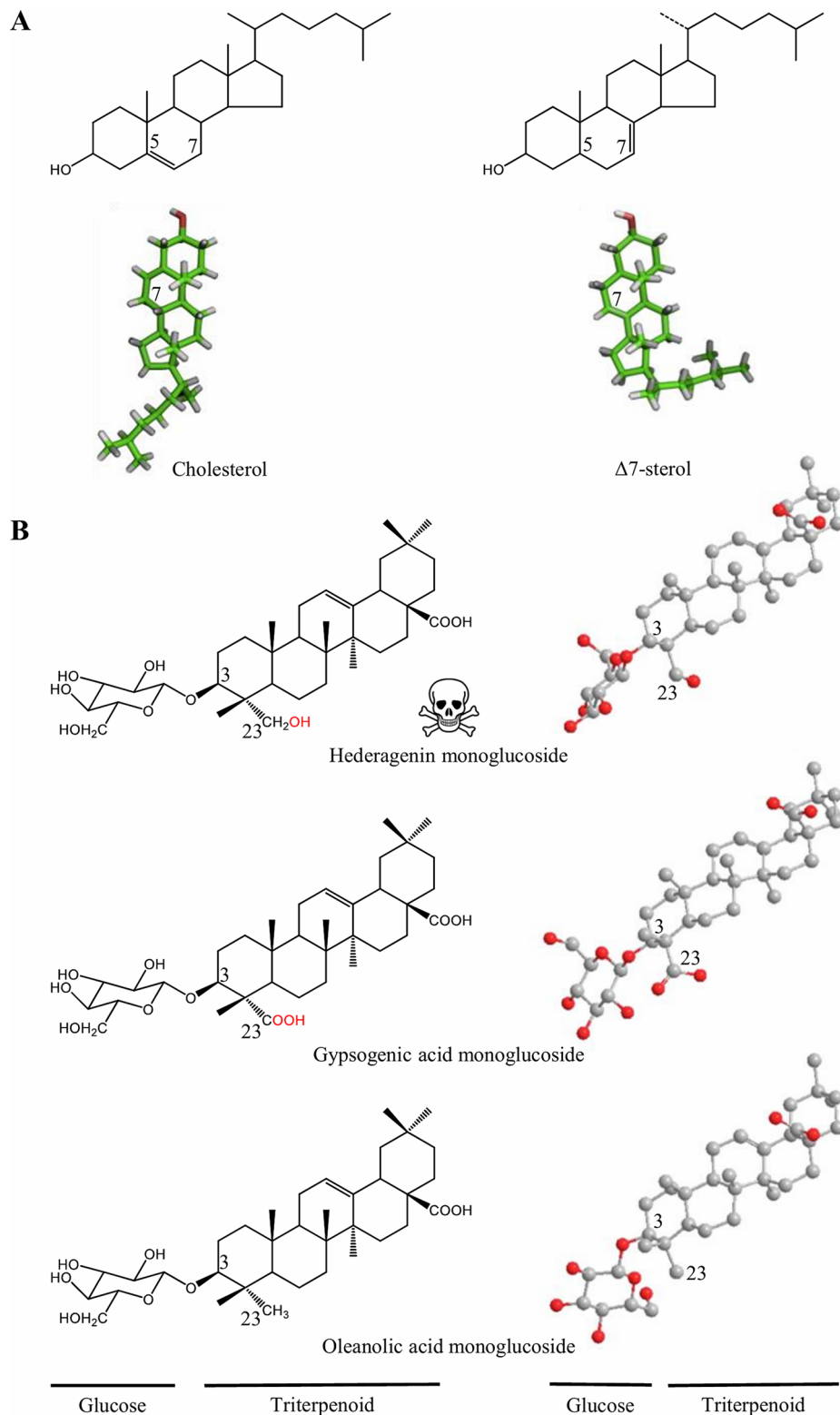
hydroxylations at C11 and C20 both lie outside of the non-homologous gene cluster.

In Solanaceae, the core genes required for the biosynthesis of steroidal glycoalkaloids are part of a metabolic operon-like gene cluster (Figure 1B). Itkin et al. (2013) showed that in tomato, six SGA genes are arranged in a cluster on chromosome 7 (UGT73L5: GAME1; UGT93: GAME17; UGT94: GAME18; UGT73L4: GAME2; Dioxygenase: GAME11; and CYP72A188: GAME6), whereas two other genes are located next to each other on chromosome 12 (CYP88B1: GAME4; and Transaminase: GAME12). Similarly in potato, four SGA-related genes are located in chromosome 7 (UGT73: SGT3; Dioxygenase: GAME11; CYP72A188: GAME6; and UGT73: SGT1) and two in chromosome 12 (Transaminase: GAME12; and CYP88B1: GAME4).

## Tandem Repeats of Triterpenoid-Biosynthetic Genes and Triterpenoid Diversity in *B. vulgaris*, *Medicago truncatula* and *Glycyrrhiza uralensis*

Triterpenoid saponin biosynthesis has evolved recurrently in evolution and thus the organization of the genes may not be conserved. In recent years, the *Barbarea* genus has appeared as a unique plant model as it is the only genus in the agronomically important cabbage family (Brassicaceae) known to accumulate saponins (Khakimov et al., 2016). Some of these saponins (e.g., hederagenin cellobioside) are highly deterrent to *Brassica* specialist herbivores such as flea beetles (*Phyllotreta nemorum*) and the diamondback moth (*Plutella xylostella*) (Kuzina et al., 2009; Kuzina et al., 2011; Liu et al., 2019).

Genome (Byrne et al., 2017) and genetic (Kuzina et al., 2009; Kuzina et al., 2011) analysis showed that the genes mediating the biosynthesis of the deterrent triterpenoid saponins in *B. vulgaris* are not linked, but are present in tandem repeats spread along the genome (Figure 1B). The OSCs, P450s, and UGTs in the pathway are generally characterized to be rather substrate and product promiscuous enzymes which may facilitate that more than 49 different saponin structures can be generated in *B. vulgaris* with a limited number of genes (Khakimov et al., 2016). QTLs for flea beetle resistance and accumulation of saponins co-localize in *B. vulgaris* (Kuzina et al., 2009; Kuzina et al., 2011; Khakimov et al., 2015). Two unlinked QTLs containing two OSC (i.e., LUP2 and LUP5) and a tandem repeat of eight P450s (CYP72As), respectively, were identified. *In vitro* and *in planta* assays shown that LUP5 was preferentially expressed in the insect-deterrent *B. vulgaris* G-type. LUP5 leads to predominant accumulation of  $\beta$ -amyrin, the precursor for the deterrent hederagenin cellobioside. Conversely, LUP2 was preferentially expressed in the P-type plants (insect-susceptible), and produces mainly lupeol, the precursor for lupeol-derived saponins that appear not to be deterrent. Of the eight CYP72As, only CYP72A552 oxidizes oleanolic acid at the C23 position leading to the formation of insect deterrent hederagenin glycosides (Liu et al., 2019). CYP716A80 and CYP716A81 catalyze C28 carboxylation and two additional hydroxylation (Khakimov et al., 2015) while



**FIGURE 2 |** Structural changes may affect the biological activity of triterpenoids. **(A)** Structural formulae (up) and 3D models (down) of cholesterol and  $\Delta 7$  sterol. In the  $\Delta 7$  sterol the side chain is bent as compared to cholesterol. In sea cucumber cell membrane systems, cholesterol are replaced by  $\Delta 7$  sterols to possible modulate the lytic action of saponins as the 3D structure is altered (Claereboudt et al., 2018). **(B)** Structural formulae (left) and 3D models (right) of *B. vulgaris* saponins. In the deterrent hederagenin monoglucoside, the glucose (C3) is twisted in respect to the triterpenoid backbone when a hydroxyl group is added at position C23 (Liu et al., 2019). 3D models are derived from Claereboudt et al. (2018) and Liu et al. (2019), respectively.

a tandem repeat of at least five UGT73Cs (Augustin et al., 2012; Erthmann et al., 2018) are also involved in the pathway. Both the CYP716As and the UGT73Cs are not linked to QTLs for neither insect resistance nor saponin accumulation. All members of the UGT73C tandem repeat accepted both hederagenin and oleanolic acid as substrates, but generated different glycosylated products, creating an spectrum of mono and bisdesmosidic saponins (Erthmann et al., 2018). If the *B. vulgaris* saponin non-linked gene organization is part of facilitating the opportunity to create multiple end products with a number of limited genes is an open question, but interestingly other pathways (e.g., anthocyanins and glucosinolates) also producing a multiple range of end products, are also known to be unlinked.

As in *B. vulgaris*, the genome of *Medicago truncatula* and the draft genome of *Glycyrrhiza uralensis* did not reveal operon-like gene cluster of OSCs, P450s, and UGTs involved in neither soyasaponins nor glycyrrhizin biosynthesis, but rather the gene candidates are organized as tandem repeats of each of these gene families (Naoumkina et al., 2010; Mochida et al., 2017). However, while most of the genes in these tandem repeats remain functionally uncharacterized, UGT73K1 was proven to glycosylate soyasapogenol B and E at the C3 position (Achnine et al., 2005).

## WHAT DETERMINES TOXICITY VERSUS AUTOTOXICITY OF PLANT TRITERPENOID SAPONINS?

Saponins are generally assumed to target and disrupt vital functions common among organisms, such as the integrity of the cell membrane. In this regard, a fundamental paradox in plant evolution is that many defense metabolites may be harmful for the plant themselves. Therefore, plants have had to evolve strategies to both biosynthesize and accumulate defense molecules without causing autotoxicity. Some of the strategies used by plants to avoid autotoxicity are the compartmentalization of defense compounds in specialized structures and chemical modifications leading to decreased toxicity (i.e., glycosylation). For instance, avenacin A-1 is normally localized in the vacuoles of root epidermal cells. Mutants for the *sad3* glucosyltransferase accumulate partially glycosylated avenacin A-1, having root epidermis defects and affected saponin cellular distribution (Mylona et al., 2008), illustrating the detrimental effects of pathway intermediates.

Even though saponins are among the largest classes of natural products, their mode of action is not really understood—this may partly be attributed to the often complex mixtures of saponins in plants. Saponins have been shown to cause increased insect mortality, lower food intake, weight reduction, and developmental problems among others (Agerbirk et al., 2003; Christensen et al., 2019); and due to their amphiphilic nature, their most studied effect is regarding their membrane permeabilizing properties. Early studies using artificial lipid bilayers have shown that avenacin A-1 induces permeabilization in a cholesterol-dependent manner and requires the presence of an intact glycoside moiety in the C3 position of the

triterpenoid scaffold (Armah et al., 1999). The dependence of the permeabilizing activity on the presence of cholesterol in the membrane and the sugar moiety also was shown for the steroidal saponin digitonin and glycoalkaloids. For example,  $\alpha$ -chaconine had strong lytic activity of cholesterol-containing lipid vesicles compared to  $\alpha$ -solanine, which has the same aglycone structure but different types of glycosides (Nishikawa et al., 1984; Keukens et al., 1995; Keukens et al., 1996).

Consequently, a prominent role in the avoidance of autotoxicity can be attributed to modifications in the cell membrane composition and further modifications on the saponin structure. Saponins are also present in marine organisms such as sea stars and sea cucumbers (Yendo et al., 2014). Interestingly, recent studies on how sea cucumbers tolerate the toxic saponins they produce, suggest sea cucumbers have replaced cholesterol by other sterols (e.g.,  $\Delta 7$  and  $\Delta 9$  sterols) in their membranes, presumably in order to modulate the lytic action of its own saponins (Popov, 2003). A detailed analysis of the sterols present in sea cucumbers showed that although cholesterol and  $\Delta 7$  sterols have essentially the same chemical formula and molecular weight (Figure 2), the double bond present in  $\Delta 7$  sterols has a dramatic effect on its 3D structure, possibly affecting its molecular interaction with sea cucumbers own saponins (Claereboudt et al., 2018). A similar effect on the 3D structure has been speculated to play a role in the biological activity of saponins from *B. vulgaris* (Figure 2). The hydroxylation of oleanolic acid at position C23 catalyzed by CYP72A552, leading to the formation of the highly-deterrent hederagenin causes a rotation of the monoglucoside at C3 of c. 90°, relative to the plane of the aglycone. When a carboxyl group is introduced instead as in gypsogenic acid, the glucose is situated about the same plane as in oleanolic acid. Gypsogenic and oleanolic acid glycosides are much less toxic to larvae of both *Manduca sexta* and *Plutella xylostella* than they are compared to hederagenin monoglucoside (Liu et al., 2019). Accordingly, the rotation of the sugar moiety in hederagenin glycosides may be responsible for changes on its physicochemical properties with membrane sterols, and this could lead to alter its biological role.

## CONCLUDING REMARKS

Increasing characterization of biosynthetic pathways for plant specialized metabolites is revealing that convergent evolution is surprisingly common (Pichersky and Lewinsohn, 2011). Nevertheless, we lack a biological understanding of how identical classes of specialized metabolites evolved recurrently across lineages and even across Kingdoms of Life. Biochemical analysis of the enzymes combined with molecular phylogenetic analysis have until now given detailed insights of how the enzymes for these compounds might have evolved. Further studies are needed to elucidate which biological pressures drive the recurrent emergence of the same or similar specialized metabolites across lineages of plants.

Genome analysis have revealed that some triterpenoid pathways are unlinked while others are organized in operon like clusters. The driving forces behind the genome organization of triterpenoid pathways is currently not understood and difficult

to address experimentally. However, as more plant genomes become readily available this key question in evolution of triterpenoid diversity might become addressable.

## AUTHOR CONTRIBUTIONS

All the authors conceived and wrote the manuscript.

## REFERENCES

- Achnine, L., Huhman, D. V., Farag, M. A., Sumner, L. W., Blount, J. W., and Dixon, R. A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant J.* 41, 875–887. doi: 10.1111/j.1365-313X.2005.02344.x
- Agerbirk, N., Olsen, C. E., Bibby, B. M., Frandsen, H. O., Brown, L. D., Nielsen, J. K., et al. (2003). A saponin correlated with variable resistance of *Barbarea vulgaris* to the diamondback moth *Plutella xylostella*. *J. Chem. Ecol.* 29, 1417–1433. doi: 10.1023/a:1024217504445
- Almeida, A., Dong, L., Khakimov, B., Bassard, J. E., Moses, T., Lota, F., et al. (2018). A single oxidosqualene cyclase produces the seco-triterpenoid  $\alpha$ -Onocerin. *Plant Physiol.* 176, 1469–1484. doi: 10.1104/pp.17.01369
- Araki, T., Saga, Y., Marugami, M., Otaka, J., Araya, H., Saito, K., et al. (2016). Onocerin biosynthesis requires two highly dedicated triterpene cyclases in a fern *Lycopodium clavatum*. *ChemBioChem* 17, 288–290. doi: 10.1002/cbic.201500663
- Armah, C. N., Mackie, A. R., Roy, C., Price, K., Osbourn, A. E., Bowyer, P., et al. (1999). The membrane-permeabilizing effect of avenacin A-1 involves the reorganization of bilayer cholesterol. *Biophys. J.* 76, 281–290. doi: 10.1016/S0006-3495(99)77196-1
- Augustin, J. M., Kuzina, V., Andersen, S. B., and Bak, S. (2011). Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochem.* 72, 435–457. doi: 10.1016/j.phytochem.2011.01.015
- Augustin, J. M., Drok, S., Shinoda, T., Sanmiya, K., Nielsen, J. K., Khakimov, B., et al. (2012). UDP-glycosyltransferases from the UGT73C subfamily in *Barbarea vulgaris* catalyze saponin 3-O-glucosylation in saponin-mediated insect resistance. *Plant Physiol.* 160, 1881–1895. doi: 10.1104/pp.112.202747
- Augustin, M. M., Ruzicka, D. R., Shukla, A. K., Augustin, J. M., Starks, C. M., O’Neil-Johnson, M., et al. (2015). Elucidating steroid alkaloid biosynthesis in *Veratrum californicum*: production of verazine in Sf9 cells. *Plant J.* 82, 991–1003. doi: 10.1111/tpj.12871
- Bigger, D. S., and Chaney, W. E. (1998). Effects of *Iberis umbellata* (Brassicaceae) on insect pests of cabbage and on potential biological control agents. *Environ. Entomol.* 27, 161–167. doi: 10.1093/ee/27.1.161
- Boycheva, S., Daviet, L., Wolfender, J. L., and Fitzpatrick, T. B. (2014). The rise of operon-like gene clusters in plants. *Trends Plant Sci.* 19, 447–459. doi: 10.1016/j.tplants.2014.01.013
- Byrne, S. L., Erthmann, P. Ø., Agerbirk, N., Bak, S., Hauser, T. P., Nagy, I., et al. (2017). The genome sequence of *Barbarea vulgaris* facilitates the study of ecological biochemistry. *Sci. Rep.* 7, 40728. doi: 10.1038/srep40728
- Cárdenas, P. D., Sonawane, P. D., Heinig, U., Bocobza, S. E., Burdman, S., and Aharoni, A. (2015). The bitter side of the nightshades: genomics drives discovery in Solanaceae steroidal alkaloid metabolism. *Phytochem.* 113, 24–32. doi: 10.1016/j.phytochem.2014.12.010
- Chen, J. C., Chiu, M. H., Nie, R. L., Cordell, G. A., and Qiu, S. X. (2005). Cucurbitacins and cucurbitane glycosides: structures and biological activities. *Nat. Prod. Rep.* 22, 386–399. doi: 10.1039/b418841c
- Christensen, S., Enge, S., Jensen, K. R., Müller, C., Kiær, L. P., Agerbirk, N., et al. (2019). Different herbivore responses to two co-occurring chemotypes of the wild crucifer *Barbarea vulgaris*. *Arthropod. Plant Interact.* 13, 19–30. doi: 10.1007/s11829-018-9633-x
- Claereboudt, E. J. S., Eeckhaut, I., Lins, L., and Deleu, M. (2018). How different sterols contribute to saponin tolerant plasma membranes in sea cucumbers. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-29223-x
- David, A., and Vallance, D. K. (1955). Bitter principles of Cucurbitaceae. *J. Pharm. Pharmacol.* 7, 295–296. doi: 10.1097/00006247-197402000-00002
- Delis, C., Krokida, A., Georgiou, S., Pena-Rodriguez, L. M., Kavroulakis, N., Ioannou, E., et al. (2011). Role of luteol synthase in *Lotus japonicus* nodule formation. *New Phytol.*, 335–346. doi: 10.1111/j.1469-8137.2010.03463.x
- Dinan, L., Whiting, P., Girault, J. P., Lafont, R., Dhadialla, T. S., Cress, D. E., et al. (1997). Cucurbitacins are insect steroid hormone antagonists acting at the ecdysteroid receptor. *Biochem. J.* 327, 643–650. doi: 10.1042/bj3270643
- Erthmann, P. Ø., Agerbirk, N., and Bak, S. (2018). A tandem array of UDP-glycosyltransferases from the UGT73C subfamily glycosylate saponins, forming a spectrum of mono- and bisdesmosidic saponins. *Plant Mol. Biol.* 97, 37–55. doi: 10.1007/s11103-018-0723-z
- Field, B., and Osbourn, A. E. (2008). Metabolic diversification-independent assembly of operon-like gene clusters in different plants. *Sci.* 320, 543–547. doi: 10.1126/science.1154990
- Field, B., Fiston-Lavier, A. S., Kemen, A., Geisler, K., Quesneville, H., and Osbourn, A. E. (2011). Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc. Natl. Acad. Sci.* 108, 16116–16121. doi: 10.1111/j.1095-8339.1984.tb02566.x
- Garratt, M. J., Tims, J. D., Rickards, R. B., Chambers, T. C., and Douglas, J. G. (1984). The appearance of Baragwanathia (Lycophytina) in the Silurian. *Bot. J. Linn. Soc.* 89, 355–358. doi: 10.1111/j.1095-8339.1984.tb02566.x
- Geisler, K., Hughes, R. K., Sainsbury, F., Lomonosoff, G. P., Rejzek, M., Fairhurst, S., et al. (2013). Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proc. Natl. Acad. Sci.* 110, E3360–E3367. doi: 10.1073/pnas.1309157110
- Giraud, K., and Lejal-Nicol, A. (1989). Cassinium dongolense n. sp. bois fossile de caesalpinaceae du nubien du soudan septentrional. *Rev. Palaeobot. Palynol.* 59, 37–50. doi: 10.1016/0034-6667(89)90004-3
- Guhling, O., Hobl, B., Yeats, T., and Jetter, R. (2006). Cloning and characterization of a luteol synthase involved in the synthesis of epicuticular wax crystals on stem and hypocotyl surfaces of *Ricinus communis*. *Arch. Biochem. Biophys.* 448, 60–72. doi: 10.1016/j.abb.2005.12.013
- Hamberger, B., and Bak, S. (2013). Plant P450s as versatile drivers for evolution of species-specific chemical diversity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20120426. doi: 10.1098/rstb.2012.0426
- Hayashi, H., Huang, P., Kirakosyan, A., Inoue, K., Hiraoka, N., Ikeshiro, Y., et al. (2002). Cloning and characterization of a cDNA encoding  $\beta$ -amyrin synthase involved in glycyrrhizin and soyasaponin biosyntheses in licorice. *Biol. Pharm. Bull.* 24, 912–916. doi: 10.1248/bpb.24.912
- Itkin, M., Heinig, U., Tzfadia, O., Bhidi, A. J., Shinde, B., Cardenas, P. D., et al. (2013). Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Sci.* 341, 175–179. doi: 10.1126/science.1240230
- Itkin, M., Davidovich-Rikanati, R., Cohen, S., Portnoy, V., Doron-Faigenboim, A., Oren, E., et al. (2016). The biosynthetic pathway of the nonsugar, high-intensity sweetener mogrosin V from *Siraitia grosvenorii*. *Proc. Natl. Acad. Sci.* 113, E7619–E7628. doi: 10.1073/pnas.1604828113
- Kemen, A. C., Honkanen, S., Melton, R. E., Findlay, K. C., Mugford, S. T., Hayashi, K., et al. (2014). Investigation of triterpene synthesis and regulation in oats reveals a role for  $\beta$ -amyrin in determining root epidermal cell patterning. *Proc. Natl. Acad. Sci.* 111, 8679–8684. doi: 10.1073/pnas.1401553111
- Keukens, E. A. J., de Vrije, T., van den Boom, C., de Waard, P., Plasman, H. H., Thiel, F., et al. (1995). Molecular basis of glycoalkaloid induced membrane disruption. *Biochim. Biophys. Acta* 1240, 216–228. doi: 10.1016/0005-2736(95)00186-7

## FUNDING

PC work was supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 752437. SB and AA were supported by grants from the Independent Research Fund Denmark grant No. 7017-00275B and Novo Nordisk Foundation grant No. NNF17OC0027646.



- Keukens, E. A., de Vrije, T., Jansen, L. A., de Boer, H., Janssen, M., de Kroon, A. I., et al. (1996). Glycoalkaloids selectively permeabilize cholesterol containing biomembranes. *Biochim. Biophys. Acta* 1279, 243–250. doi: 10.1016/0005-2736(95)00253-7
- Khakimov, B., Kuzina, V., Erthmann, P. Ø., Fukushima, E. O., Augustin, J. M., Olsen, C. E., et al. (2015). Identification and genome organization of saponin pathway genes from a wild crucifer, and their use for transient production of saponins in *Nicotiana benthamiana*. *Plant J.* 84, 478–490. doi: 10.1111/tpj.13012
- Khakimov, B., Tseng, L. H., Godejohann, M., Bak, S., and Engelsens, S. B. (2016). Screening for triterpenoid saponins in plants using hyphenated analytical platforms. *Mol.* 21, 1–19. doi: 10.3390/molecules21121614
- Kim, O. T., Lee, J. W., Bang, K. H., Kim, Y. C., Hyun, D. Y., Cha, S. W., et al. (2009). Characterization of a dammarenyliol synthase in *Centella asiatica* (L.) Urban. *Plant Physiol. Biochem.* 47, 998–1002. doi: 10.1016/j.plaphy.2009.08.001
- Krokida, A., Delis, C., Geisler, K., Garagounis, C., Tsikou, D., Peña-Rodríguez, L. M., et al. (2013). A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytol.* 200, 675–690. doi: 10.1111/nph.12414
- Kumar, P., Bhadauria, A. S., Singh, A. K., and Saha, S. (2018). Betulinic acid as apoptosis activator: molecular mechanisms, mathematical modeling and chemical modifications. *Life Sci.* 209, 24–33. doi: 10.1016/j.lfs.2018.07.056
- Kushiro, T., Shibuya, M., and Ebizuka, Y. (1999). Chimeric triterpene synthase. A possible model for multifunctional triterpene synthase. *J. Am. Chem. Soc.* 121, 1208–1216. doi: 10.1021/ja983012h
- Kushiro, T., Shibuya, M., Masuda, K., and Ebizuka, Y. (2000). Mutational studies on triterpene synthases: engineering lupeol synthase into  $\beta$ -amyrin synthase. *J. Am. Chem. Soc.* 122, 6816–6824. doi: 10.1021/ja0010709
- Kuzina, V., Ekstrom, C. T., Andersen, S. B., Nielsen, J. K., Olsen, C. E., and Bak, S. (2009). Identification of defense compounds in *Barbarea vulgaris* against the herbivore *Phyllotreta nemorum* by an ecometabolomic approach. *Plant Physiol.* 151, 1977–1990. doi: 10.1104/pp.109.136952
- Kuzina, V., Nielsen, J. K., Augustin, J. M., Torp, A. M., Bak, S., and Andersen, S. B. (2011). *Barbarea vulgaris* linkage map and quantitative trait loci for saponins, glucosinolates, hairiness and resistance to the herbivore *Phyllotreta nemorum*. *Phytochem.* 72, 188–198. doi: 10.1016/j.phytochem.2010.11.007
- Leung, K. W., and Wong, A. S.-T. (2010). Pharmacology of ginsenosides: a literature review. *Chin. Med.* 5, 20. doi: 10.1186/1749-8546-5-20
- Liu, Q., Khakimov, B., Cardenas, P. D., Federico, C., Olsen, C. E., Jensen, K. R., et al. (2019). The cytochrome P450 CYP72A552 is key to production of hederagenin-based saponins that mediate plant defense against herbivores. *New Phytol.* 222, 1599–1609. doi: 10.1111/nph.15689
- Metcalfe, R. L., Metcalfe, R. A., and Rhodes, A. M. (1980). Cucurbitacins as kairomones for diabroticite beetles. *Proc. Natl. Acad. Sci. U. S. A.* 77, 3769–3772. doi: 10.1073/pnas.77.7.3769
- Metcalfe, R. L. (1986). Coevolutionary adaptations of rootworm beetles (Coleoptera: Chrysomelidae) to cucurbitacins. *J. Chem. Ecol.* 12, 1109–1124. doi: 10.1007/BF01638999
- Mills, J. S., and Werner, A. E. A. (1955). The chemistry of dammar resin. *J. Chem. Soc.* 3132–3140. doi: 10.1039/JR9550003132
- Mills, J. S. (1956). The constitution of the neutral, tetracyclic triterpenes of dammar resin. *J. Chem. Soc.* 2196–2202. doi: 10.1039/JR9560002196
- Mochida, K., Sakurai, T., Seki, H., Yoshida, T., Takahagi, K., Sawai, S., et al. (2017). Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. *Plant J.* 89, 181–194. doi: 10.1111/tpj.13385
- Mugford, S. T., Louveau, T., Melton, R., Qi, X., Bakht, S., Hill, L., et al. (2013). Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. *Plant Cell* 25, 1078–1092. doi: 10.1105/tpc.113.110551
- Mylona, P., Owatworakit, A., Papadopoulou, K., Jenner, H., Qin, B., Findlay, K., et al. (2008). Sad3 and Sad4 are required for saponin biosynthesis and root development in oat. *Plant Cell* 20, 201–212. doi: 10.1105/tpc.107.056531
- Naoumkina, M. A., Modolo, L. V., Huhman, D. V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L. W., et al. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Cell.* 22, 850–866. doi: 10.1105/tpc.109.073270
- Nielsen, J. K., Larsen, L. M., and Sørensen, H. (1977). Cucurbitacin E and I in *Iberis amara*: Feeding inhibitors for *Phyllotreta nemorum*. *Phytochem.* 16, 1519–1522. doi: 10.1016/0031-9422(77)84014-4
- Nielsen, J. K., Nagao, T., Okabe, H., and Shinoda, T. (2010). Resistance in the plant, *Barbarea vulgaris*, and counter-adaptations in flea beetles mediated by saponins. *J. Chem. Ecol.* 36, 277–285. doi: 10.1007/s10886-010-9758-6
- Nishikawa, M., Nojima, S., Akiyama, T., and Inoue, K. (1984). Interaction of digitonin and its analogs with membrane cholesterol. *J. Biochem.* 96, 1231–1239. doi: 10.1093/oxfordjournals.jbchem.a134941
- Osborn, A. (2010). Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. *Plant Physiol.* 154, 531–535. doi: 10.1104/pp.110.161315
- Papadopoulou, K., Melton, R. E., Leggett, M., Daniels, M. J., and Osborn, A. E. (2002). Compromised disease resistance in saponin-deficient plants. *Proc. Natl. Acad. Sci.* 99, 12923–12928. doi: 10.1073/pnas.99.22.12923
- Pichersky, E., and Lewinsohn, E. (2011). Convergent evolution in plant specialized metabolism. *Annu. Rev. Plant Biol.* 62, 549–566. doi: 10.1146/annurev-arplant-042110-103814
- Pisha, E., Chai, H., Lee, I. S., Chagwedera, T. E., Farnsworth, N. R., Cordell, G. A., et al. (1995). Discovery of betulinic acid as a selective inhibitor of human melanoma that functions by induction of apoptosis. *Nat. Med.* 1, 369–375. doi: 10.1038/nm1095-1046
- Popov, A. M. (2003). Comparative study of effects of various sterols and triterpenoids on permeability of model lipid membranes. *J. Evol. Biochem. Physiol.* 39, 314–320. doi: 10.1023/A:1026147925642
- Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R., and Osborn, A. (2004). A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc. Natl. Acad. Sci. U. S. A.* 101, 8233–8238. doi: 10.1073/pnas.0401301101
- Raemisch, D. R., and Turpin, F. T. (1984). Field tests for an adult western corn rootworm aggregation pheromone associated with the phagostimulatory characteristic of bitter *Cucurbita* spp. *J. Agric. Entomol.* 1, 339–344.
- Shang, Y., Ma, Y., Zhou, Y., Zhang, H., Duan, L., Chen, H., et al. (2014). Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science* 346, 1084–1088. doi: 10.1126/science.1259215
- Shibuya, M., Zhang, H., Endo, A., Shishikura, K., Kushiro, T., and Ebizuka, Y. (1999). Two branches of the lupeol synthase gene in the molecular evolution of plant oxidosqualene cyclases. *Eur. J. Biochem.* 266, 302–307. doi: 10.1046/j.1432-1327.1999.00875.x
- Shibuya, M., Adachi, S., and Ebizuka, Y. (2004). Cucurbitadienol synthase, the first committed enzyme for cucurbitacin biosynthesis, is a distinct enzyme from cycloartenol synthase for phytosterol biosynthesis. *Tetrahedron* 60, 6995–7003. doi: 10.1016/j.tet.2004.04.088
- Sonawane, P. D., Pollier, J., Panda, S., Szymanski, J., Massalha, H., Yona, M., et al. (2016). Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. *Nat. Plants* 3, 16205. doi: 10.1038/nplants.2016.205
- Stephenson, M. J., Field, R. A., and Osborn, A. (2019). The protosteryl and dammarenyliol cation dichotomy in polycyclic triterpene biosynthesis revisited: has this ‘rule’ finally been broken? *Nat. Prod. Rep.* 36, 1044–1052. doi: 10.1039/c8np00096d
- Szakiel, A., Pączkowski, C., Pensec, F., and Bertsch, C. (2012). Fruit cuticular waxes as a source of biologically active triterpenoids. *Phytochem. Rev.* 11, 263–284. doi: 10.1007/s11101-012-9241-9
- Takos, A. M., and Rook, F. (2012). Why biosynthetic genes for chemical defense compounds cluster. *Trends Plant Sci.* 17, 383–388. doi: 10.1016/j.tplants.2012.04.004
- Tansakul, P., Shibuya, M., Kushiro, T., and Ebizuka, Y. (2006). Dammarenyliol-II synthase, the first dedicated enzyme for ginsenoside biosynthesis, in *Panax ginseng*. *FEBS Letters* 580, 5143–5149. doi: 10.1016/j.febslet.2006.08.044
- Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P., and Osborn, A. (2014). Triterpene biosynthesis in plants. *Annu. Rev. Plant Biol.* 65, 225–257. doi: 10.1146/annurev-arplant-050312-120229
- Xiong, Q., Rocco, F., Wilson, W. K., Xu, R., Ceruti, M., and Matsuda, S. P. R. (2005). Structure and reactivity of the dammarenyliol cation: configurational transmission in triterpene synthesis. *J. Org. Chem.* 70, 5362–5375. doi: 10.1021/jo050147e

- Xu, R., Fazio, G. C., and Matsuda, S. P. T. (2004). On the origins of triterpenoid skeletal diversity. *Phytochem.* 65, 261–291. doi: 10.1016/j.phytochem.2003.11.014
- Xue, Z., Duan, L., Liu, D., Guo, J., Ge, S., Dicks, J., et al. (2012). Divergent evolution of oxidosqualene cyclases in plants. *New Phytol.* 193, 1022–1038. doi: 10.1111/j.1469-8137.2011.03997.x
- Yendo, A., Costa, F., Da Costa, C. T., Colling, L., Gosmann, G., and Fett-Neto, A. (2014). Biosynthesis of plant triterpenoid saponins: genes, enzymes and their regulation. *Mini Rev. Org. Chem.* 11, 292–306. doi: 10.2174/1570193x1103140915111425
- Zhou, Y., Ma, Y., Zeng, J., Duan, L., Xue, X., Wang, H., et al. (2016). Convergence and divergence of bitterness biosynthesis and regulation in Cucurbitaceae. *Nat. Plants* 2, 16183. doi: 10.1038/nplants.2016.183

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Cárdenas, Almeida and Bak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.