



Tansley Review

Bioengineering of plant (tri)terpenoids: from metabolic engineering of plants to synthetic biology *in vivo* and *in vitro*

Author for correspondence:

Alain Goossens

Tel: +32 9 3313851

Email: alain.goossens@psb.vib-ugent.be

Received: 7 March 2013

Accepted: 12 April 2013

Tessa Moses^{1,2,3,4}, Jacob Pollier^{1,2}, Johan M. Thevelein^{3,4} and Alain Goossens^{1,2}

¹Department of Plant Systems Biology, VIB, Technologiepark 927, B-9052, Gent, Belgium; ²Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, B-9052, Gent, Belgium; ³Department of Molecular Microbiology, VIB, Kasteelpark Arenberg 31, B-3001, Leuven, Heverlee, Belgium; ⁴Laboratory of Molecular Cell Biology, Institute of Botany and Microbiology, KU Leuven, Kasteelpark Arenberg 31, B-3001, Leuven, Heverlee, Belgium

Contents

Summary	27	IV. Perspectives: 'exploration of triterpenoids: the road ahead'	38
I. Introduction	27	Acknowledgements	40
II. 'Natural' terpenoid biology	28	References	40
III. 'Synthetic' terpenoid biology	32		

New Phytologist (2013) **200**: 27–43
doi: 10.1111/nph.12325

Key words: bioengineering, combinatorial biosynthesis, directed enzyme evolution, heterologous biosynthesis, secondary metabolism, synthetic biology, terpenoids, triterpenoids.

Summary

Terpenoids constitute a large and diverse class of natural products that serve many functions in nature. Most of the tens of thousands of the discovered terpenoids are synthesized by plants, where they function as primary metabolites involved in growth and development, or as secondary metabolites that optimize the interaction between the plant and its environment. Several plant terpenoids are economically important molecules that serve many applications as pharmaceuticals, pesticides, etc. Major challenges for the commercialization of plant-derived terpenoids include their low production levels *in planta* and the continuous demand of industry for novel molecules with new or superior biological activities. Here, we highlight several synthetic biology methods to enhance and diversify the production of plant terpenoids, with a foresight towards triterpenoid engineering, the least engineered class of bioactive terpenoids. Increased or cheaper production of valuable triterpenoids may be obtained by 'classic' metabolic engineering of plants or by heterologous production of the compounds in other plants or microbes. Novel triterpenoid structures can be generated through combinatorial biosynthesis or directed enzyme evolution approaches. In its ultimate form, synthetic biology may lead to the production of large amounts of plant triterpenoids in *in vitro* systems or custom-designed artificial biological systems.

I. Introduction

Plants synthesize and accumulate a wide range of small molecules or natural products that are involved in fundamental physiological and ecological processes. Some of these natural products have therapeutic potential which has been exploited by humans for thousands of years in the form of traditional herbal medicine. In

recent years, with our growing understanding of their biosynthesis, regulation and functioning, plant-derived natural products have emerged as high-value therapeutics, flavors and fragrances, colorants and health-promoting agents. Based on their structure and biosynthetic origin, plant natural products can be classified into different groups, such as the terpenoids, alkaloids and phenolic compounds (Croteau *et al.*, 2000). This review focuses on the

terpenoids, of which tens of thousands of compounds have been characterized from plants. The terpenoids or isoprenoids comprise structurally diverse compounds that are associated with primary as well as secondary metabolism. Gibberellin, abscisic acid and brassinosteroid phytohormones, phytosterols and carotenoid pigments are primary metabolic terpenoids involved in basic functions, such as the regulation of plant growth and development, photosynthesis, membrane permeability and fluidity (Bohlmann & Keeling, 2008; Vranová *et al.*, 2012). However, the majority of the plant terpenoids are secondary metabolites that play a crucial role in the interaction of the plant with its environment, for instance by serving as pollinator attractants, herbivore repellents, anti-feedants, toxins or antibiotics (Gershenzon & Dudareva, 2007).

The structural variety and inherent biological activities of many plant terpenoids have rendered them widely applicable. With an annual production of 10^7 tons, natural rubber is the most abundant terpenoid produced. Because of its unique properties, it serves as a biological material in the non-food industry for the production of heavy-duty tires, vibration dampers or latex products, such as surgical gloves (van Beilen & Poirier, 2007). Other examples of plant terpenoids with significant economic value include: menthol, a monoterpenoid extracted from peppermint and used in the flavor and fragrance industry; abietic acid, a diterpenoid isolated from conifer rosin that is used in lacquers, varnishes and soap; and the anti-malarial and anti-cancer drugs artemisinin and taxol, respectively (Bohlmann & Keeling, 2008).

A major hurdle in the commercialization of plant terpenoids is that they often accumulate in very low concentrations *in planta*, thereby hindering their purification in large amounts from the natural source. When the extraction of a natural product from its natural source is not sufficient, several alternative approaches can be explored, including: (1) plant breeding and genetic engineering to generate cultivars or transgenics accumulating higher levels of the desired compounds; (2) the development of scalable plant cell or root cultures; and (3) the engineering of microbial hosts to produce the compound. Commercially viable alternative production systems have already been established for some terpenoids, which is reflected in the emergence of companies, such as Phyton Biotech (<http://www.phytonbiotech.com/>), a global provider of chemotherapeutics, including paclitaxel extracted from *Taxus* cell cultures, and Amyris (<http://www.amyris.com/>), which uses a synthetic biology platform for the production of artemisinin in yeast.

Furthermore, the (pharmaceutical) industry is in constant search for novel molecules, primarily as a result of the discovery of new drug targets, the emergence of new diseases and, in the case of infectious diseases, the growing resistance of microbes to the currently marketed antibiotics (Pollier *et al.*, 2011). In addition, the business model of pharmaceutical companies is under threat, as leading blockbuster drugs will soon lose patent protection and become available for market competition, which often leads to lower market prices, thereby rendering the production of the drug non-profitable to the original developer. As traditional pharmacological screening of medicinal plants is time consuming and expensive, and the output of combinatorial chemistry libraries is low in terms of new drugs, alternative approaches to generate new

molecules or scaffolds are required (Koehn & Carter, 2005; Welsch *et al.*, 2010). Combinatorial biosynthesis accelerates the process of natural evolution and multiplies the natural diversity by generating novel enzyme–substrate combinations. Thereby, it can be rationally applied to custom design new compounds (Kirschning *et al.*, 2007; Pollier *et al.*, 2011).

In this review, we provide a futuristic view into the engineering of triterpenoids, the least engineered class of terpenoids with pharmaceutical potential, by drawing inspiration from the current status of engineering of other terpenoid classes in plants and microbial hosts. We highlight the latest approaches for enhancing the production and increasing the structural diversity of natural compounds, and frame the potential of the booming trends in synthetic biology in the perspective of triterpenoid production.

II. 'Natural' terpenoid biology

A basic understanding of the biosynthesis and regulation of a compound is strategic to any bioengineering initiative. Therefore, we set the base for triterpenoid biology by providing an insight into their synthesis and regulation in plants. A correct perception of their native production habitat and machinery permits the translation of this knowledge to the bioengineering of native plants or heterologous hosts.

1. Classification and biosynthesis of plant terpenoids

Despite their enormous structural diversity, terpenoids share a common biosynthetic origin and follow similar synthesis routes. All terpenoids are derived from the repetitive fusion of isoprene (C_5H_8) units, and the number of isoprene units determines their classification. In higher plants, the biosynthesis begins with the generation of isopentenyl pyrophosphate (IPP), the principal precursor, through the mevalonate (MVA)/3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) pathway or the 2-C-methyl-D-erythritol 4-phosphate (MEP)/1-deoxy-D-xylulose 5-phosphate (DOXP)/non-MVA pathway. The IPP is isomerized to its allylic isomer dimethylallyl pyrophosphate (DMAPP). The consecutive condensation of IPP and DMAPP units leads to the formation of prenylated pyrophosphates, the immediate precursors of the different terpenoid classes (Fig. 1). These condensation reactions are catalyzed by specific prenyltransferases which are named according to the product they generate. Specific terpenoid synthases then modify these precursors to terpenoid skeletons (Chen *et al.*, 2011), which are subsequently decorated by various enzymatic modifications to generate the structural and functional diversity of terpenoids. Plants also exhibit a clear compartmentalization for the generation of IPP and the synthesis of terpenoids (Croteau *et al.*, 2000; Vranová *et al.*, 2012; Fig. 1).

Here, we focus on plant triterpenoids, comprising primary metabolites, such as the phytosterols and the brassinosteroid hormones, and secondary metabolites, such as the saponins. The IPP for triterpenoid biosynthesis is generated through the cytosol, peroxisome and endoplasmic reticulum-localized MVA pathway. The 'head-to-tail' condensation of two IPP units with a DMAPP unit yields the C_{15} farnesyl pyrophosphate (FPP), two of which

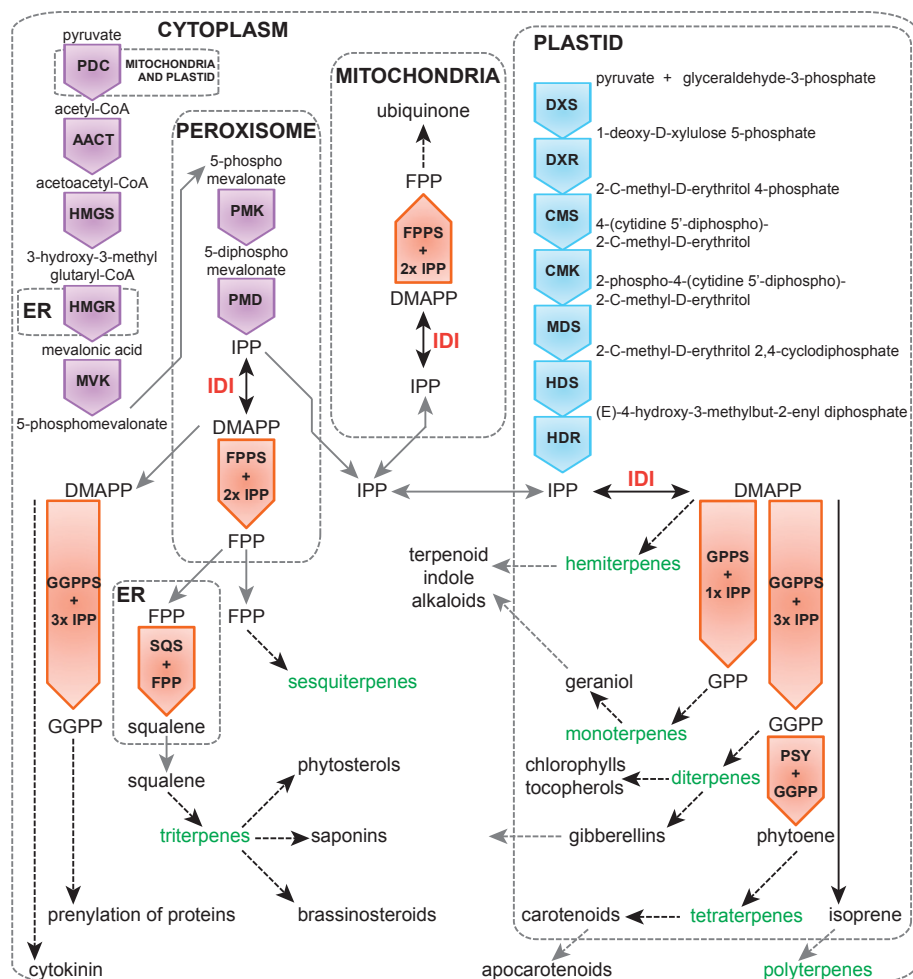


Fig. 1 Terpenoid biosynthesis in plants. Two distinct pathways for the synthesis of the universal precursors isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) exist in plants: the cytoplasm-, peroxisome-, mitochondria-, plastid- and endoplasmic reticulum (ER)-localized mevalonate (MVA) pathway (purple) and the plastid-localized methyl erythritol phosphate (MEP) pathway (blue). [Correction added after online publication 14 May 2013; replacement figure and text in preceding sentence correctly indicates that PDC enzyme is located in the mitochondria and plastids and not in the cytoplasm.] The prenyltransferases (orange) generate the immediate precursors for the different terpenoid classes (green). Dotted arrows indicate multiple reactions. Dotted grey boxes indicate the subcellular localization of the pathway. Grey arrows indicate metabolites that are transported between subcellular compartments. AACT, acetoacetyl-CoA thiolase; CMK, 4-diphosphocytidyl-methylerythritol kinase; CMS, 4-diphosphocytidyl-methylerythritol synthase; DMAPP, dimethylallyl pyrophosphate; DXR, deoxyxylulose 5-phosphate reductoisomerase; DXS, deoxyxylulose 5-phosphate synthase; FPP, farnesyl pyrophosphate; FPPS, FPP synthase; GGPP, geranylgeranyl pyrophosphate; GGPPS, GGPP synthase; GPP, geranyl pyrophosphate; GPPS, GPP synthase; HDR, hydroxymethylbutenyl 4-diphosphate reductase; HDS, hydroxymethylbutenyl 4-diphosphate synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; IDI, isopentenyl diphosphate isomerase; IPP, isopentenyl pyrophosphate; MDS, methylerythritol 2,4-cyclodiphosphate synthase; MVK, mevalonate kinase; PDC, pyruvate dehydrogenase complex; PMD, 5-diphosphomevalonate decarboxylase; PMK, 5-phosphomevalonate kinase; PSY, phytoene synthase; SQS, squalene synthase.

subsequently fuse ‘head-to-head’ to generate the linear C₃₀ triterpenoid precursor, squalene. This compound is further epoxidized to 2,3-oxidosqualene (Augustin *et al.*, 2011), which, in turn, is typically cyclized by specific oxidosqualene cyclases (OSCs) to tetra- or pentacyclic structures to form the dammarenes, tirucallanes and phytosterols, or the oleananes, ursanes, lupanes and taraxasteranes, respectively (Phillips *et al.*, 2006). In some plant species, 2,3-oxidosqualene can also be cyclized to mono- and tricyclic triterpenoid backbones (Xue *et al.*, 2012; Fig. 2).

The cyclization of 2,3-oxidosqualene forms the branch point between primary and secondary triterpenoid metabolism. Cycloartenol, formed by the cycloartenol synthase (CAS)-mediated

cyclization of 2,3-oxidosqualene, is the committed precursor for phytosterol biosynthesis. Higher plants synthesize a mixture of various sterols from cycloartenol, which can accumulate in a free form or as esters or glycosides (Nes, 2011). In turn, the phytosterols cholesterol, campesterol and sitosterol are the precursors of the C₂₇, C₂₈ and C₂₉ brassinosteroid hormones, respectively (Fujioka & Yokota, 2003). In addition, cholesterol can also undergo a series of oxygenations and glycosylations to form C₂₇ secondary metabolites, the steroidal saponins (Dewick, 2001). The other cyclization products of 2,3-oxidosqualene form committed precursors for secondary metabolite biosynthesis (Fig. 2). These cyclized precursors are further oxidized by one or many cytochrome P450s (CytP450s) to form sapogenins. In some plants, such as the birch

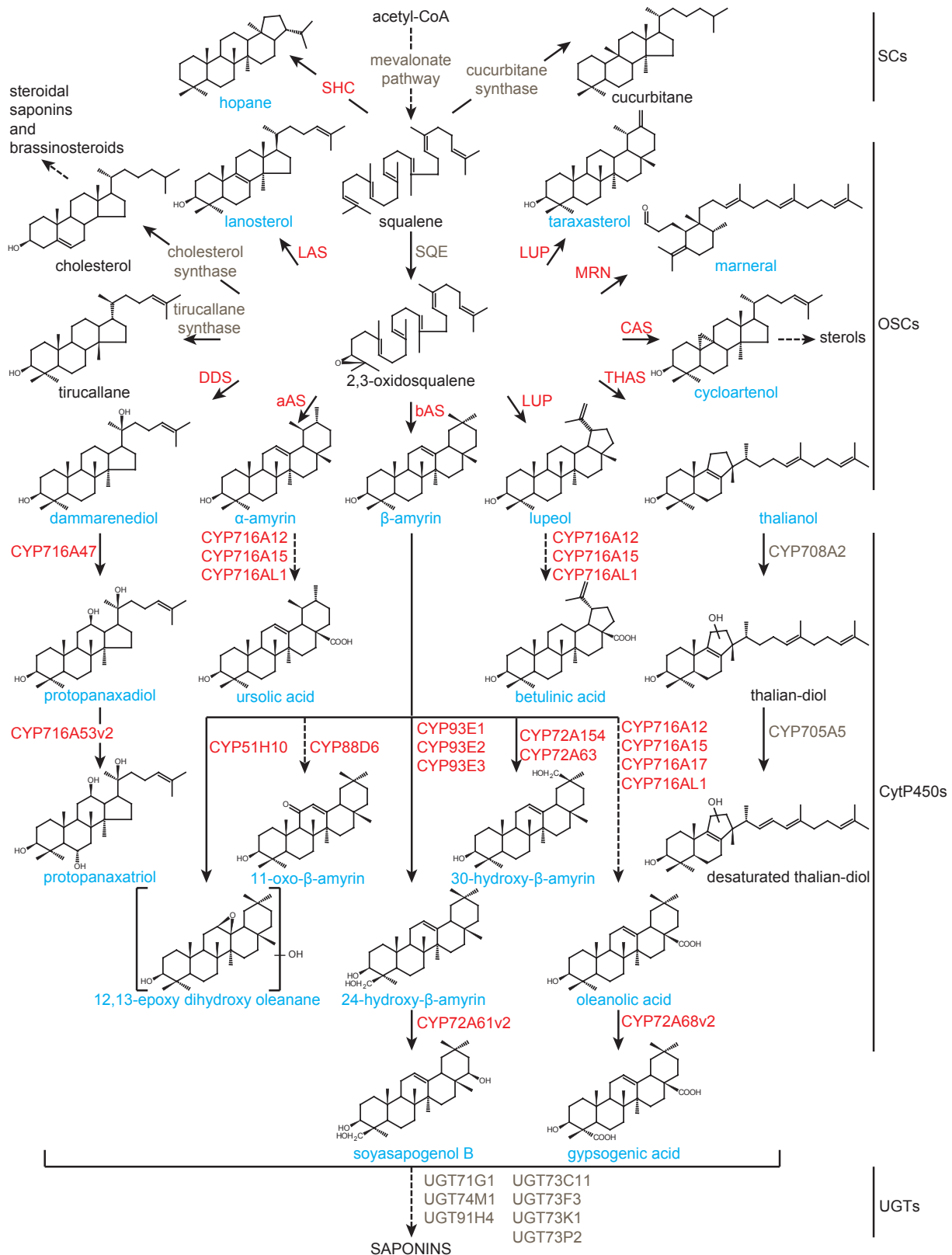


Fig. 2 A simplified scheme of triterpenoid saponin biosynthesis as expressed in *Saccharomyces cerevisiae*. Dotted arrows indicate multiple steps. Highlighted enzymes (red) and compounds (blue) were expressed and detected, respectively. aAS, α-amyrin synthase; bAS, β-amyrin synthase; CAS, cycloartenol synthase; CytP450s, cytochrome P450s; DDS, dammarenediol synthase; LAS, lanosterol synthase; LUP, lupeol synthase; MRN, marneral synthase; OSCs, oxidosqualene cyclases; SCs, squalene cyclases; SHC, squalene-hopane cyclase; SQE, squalene epoxidase; THAS, thalianol synthase; UGTs, UDP-dependent glycosyltransferases.

(*Betula pubescens*) and olive tree (*Olea europaea*), the saponins form the final accumulating secondary metabolite, whereas, in others, the saponins are glycosylated by UDP-dependent glycosyltransferases (UGTs) to generate amphipathic glycosides, the saponins (Augustin *et al.*, 2011).

2. Regulation of terpenoid biosynthesis in plants

The biosynthesis of terpenoids is tightly controlled in plants, as they serve many functions in plant growth, development and response to biotic and abiotic environmental factors (Tholl, 2006; Nagegowda, 2010; Vranová *et al.*, 2012). Terpenoid synthesis occurs within specific tissues or at specific plant developmental stages (Nagegowda, 2010). For instance, many plant species have glandular trichomes, specialized structures for the synthesis of secreted terpenoid natural products (Lange & Turner, 2013). The triterpenoid saponin glycyrrhizin accumulates only in the underground organs, stolons and roots of licorice (*Glycyrrhiza*) plants (Seki *et al.*, 2008). Avenacins, the bioactive saponins in oat (*Avena sativa*), accumulate only in the root epidermis, where they provide resistance to phytopathogenic fungi (Haralampidis *et al.*, 2001). Such specific terpenoid synthesis is mainly regulated at the transcriptional level. The avenacin biosynthesis genes are tightly co-regulated and expressed exclusively in the root epidermis in which the avenacins accumulate (Haralampidis *et al.*, 2001; Qi *et al.*, 2006; Field & Osbourn, 2008).

In addition to this spatiotemporal regulation, induced terpenoid biosynthesis is often observed in response to herbivore feeding, pathogen attack or various abiotic stresses (Nagegowda, 2010; Vranová *et al.*, 2012). For instance, 7 d after *Spodoptera littoralis* larvae fed on *Medicago sativa* leaves, the total saponin content of the damaged foliage increased by 84%, causing a deterrent effect on the larvae. Accordingly, larval performance was reduced when forced to feed on the damaged leaves (Agrell *et al.*, 2003, 2004). The increased accumulation or release of terpenoids in response to various (a)biotic stresses is often mediated by an increased transcriptional activity of the specific terpenoid biosynthetic genes (Tholl, 2006; Nagegowda, 2010). This transcriptional response is controlled by a complex signaling cascade in which jasmonate hormones (JAs) play a crucial role. Hence, the treatment of plants or plant cell cultures with JAs often causes transcriptional and metabolic changes comparable with pathogen or herbivore attack. The exposure of *Medicago truncatula* cell suspension cultures to methyl jasmonate (MeJA) leads to increased saponin accumulation, as a consequence of transcriptional activation of the saponin biosynthetic genes (Suzuki *et al.*, 2005).

The concerted transcriptional activation of entire secondary metabolic pathways by JAs is conserved across the plant kingdom. However, downstream of the conserved JA perception and initial signaling cascade, species-specific transcriptional machineries exist that regulate the transcriptional activity of the specific biosynthetic genes (Pauwels *et al.*, 2009; Pauwels & Goossens, 2011; De Geyter *et al.*, 2012). A few transcription factors regulated by the JA signaling cascade that activate the transcription of (sesqui)terpenoid biosynthetic genes have already been characterized (De Geyter *et al.*, 2012), but none for triterpenoids so far. It should be noted,

however, that JAs are not the only regulators of terpenoid metabolism in plants and that complex cross-talk between various stress- and development-related signaling cascades occurs (De Geyter *et al.*, 2012).

In addition to the transcriptional, developmental and spatio-temporal modulation of terpenoid biosynthetic genes, post-translational regulatory mechanisms also exist in terpenoid biosynthesis. The activity of HMGR, the enzyme that catalyzes the key regulatory step of the MVA pathway, is controlled at the protein level through the action of protein phosphatase 2A (Leivar *et al.*, 2011) or by the E3 ubiquitin ligase *SUDI* (Doblas *et al.*, 2013).

3. Bioengineering of terpenoids in planta

Because of their strict regulation, most terpenoids are produced in very small amounts in their natural sources. The low yield makes extraction expensive, which is eventually reflected in their market value. Consequently, there is a wide gap between demand and supply of terpenoids, which hampers their widespread application. The classical approach to ensure a constant or improved yield is the selection and propagation of high-producing cultivars or the production and/or elicitation of (transgenic) plant (cell) cultures (Zhao *et al.*, 2005; Georgiev *et al.*, 2009, 2012; Lambert *et al.*, 2011; Lim & Bowles, 2012; Wilson & Roberts, 2012). Our growing understanding of terpenoid biosynthesis, together with the development of functional genomics and systems biology toolkits, has enabled the metabolic engineering of whole plants and plant cultures to enhance productivity and alter terpenoid distribution *in planta* (Roberts, 2007; Dudareva *et al.*, 2013).

As terpenoid biosynthesis is strictly regulated and often controlled by specific transcription factors, one way to increase productivity is to modulate the expression of such or other regulatory factors (Broun, 2004; De Geyter *et al.*, 2012). However, despite the identification of transcription factors that steer the biosynthesis of terpenoids, the overexpression of a single transcription factor often does not lead to a higher production of the compounds. For instance, the overexpression of *ORCA3*, an APETALA2/ETHYLENE RESPONSE FACTOR (AP2/ERF) transcription factor that controls the expression of several terpenoid indole alkaloid (TIA) biosynthetic genes, is not sufficient to induce TIA production in *Catharanthus roseus* cell cultures, indicating that only a part of TIA biosynthesis is under the control of this transcription factor (van der Fits & Memelink, 2000). Hence, further elucidation of the complex signaling cascades that lead to an increased accumulation of terpenoids is mandatory for large-scale metabolic engineering of terpenoid production using transcription factors. To date, *in planta* triterpenoid engineering has been hampered by the lack of knowledge about the regulatory mechanisms controlling gene expression (Sawai & Saito, 2011). Hence, a challenge for future triterpenoid research will be to identify the transcription or other regulatory factors that steer their biosynthesis.

A second way to increase productivity is by the specific overexpression of rate-limiting enzymes in the pathway. The overexpression of genes encoding enzymes such as *HMGR*,

deoxyxylulose 5-phosphate synthase (*DXS*) and prenyltransferases, has been used to elevate terpenoid levels in plant tissue cultures (Degenhardt *et al.*, 2003). Enhanced terpenoid production has also been observed on alteration of the subcellular localization of enzymes, presumably resulting from the uncoupling of biosynthesis and regulation (Bouwmeester, 2006; Wu *et al.*, 2006; Farhi *et al.*, 2011; Kumar *et al.*, 2012). A single study has reported an attempt to engineer triterpenoid synthesis in tobacco (*Nicotiana tabacum*) by the heterologous expression of an avian FPP synthase (*FPPS*) and a yeast squalene synthase (*SQS*) gene targeted to the cytoplasm or plastid. No differences in squalene accumulation caused by specific targeting of the enzymes were observed. However, when the enzymes were directed to the trichomes through a trichome-specific promoter, higher squalene accumulation was accompanied by negative effects on plant growth and physiology. Remarkably, these additional effects were not observed when the same genes were expressed from a constitutive viral promoter (Wu *et al.*, 2012). Nonetheless, this study underscores the potential to engineer triterpenoids *in planta* by relocation of the biosynthetic pathway and enhancement of the precursor flux, and encourages future research on this terpenoid class.

In addition to enhancing terpenoid production yields, *in planta* engineering has also been used as a tool to modulate the terpenoid composition of plants for other purposes, such as β -carotene to engineer crop nutritional value (Farré *et al.*, 2011) and volatile terpenoid compounds to improve plant defense, pollinator attraction, scent or aroma (Dudareva *et al.*, 2013), amongst others.

III. 'Synthetic' terpenoid biology

1. Definition of synthetic biology concepts

Metabolic engineering was defined in 1991 as 'the directed improvement of production, formation, or cellular properties through the modification of specific biochemical reactions or the introduction of new ones with the use of recombinant DNA technology' (Bailey, 1991). Since then, metabolic engineering has enabled spectacular advances in the production of a myriad of small compounds, including terpenoids, particularly in microbes. Now, demands have increased and biological functions that do not exist in nature are also desired. This can be achieved through synthetic biology, which can be defined as 'the design and construction of new biological components, such as enzymes, genetic circuits, and cells, or the redesign of existing biological systems' (Keasling, 2008). More elaborately, synthetic biology refers to the redesign of complex natural living systems in a rational and systematic way to simplified, predictable and controllable modules that can be modeled and manipulated to generate industrially scalable systems with a defined purpose. For many years, the term 'synthetic biology' was used to describe concepts that would be classified today as metabolic engineering. However, the definitions are not sharp-edged, and hence metabolic engineering might still be considered as the simplest form of synthetic biology (Channon *et al.*, 2008).

Based on the approach used for synthetic biology, two main branches, commonly referred to as 'top-down' and 'bottom-up' synthetic biology, can be recognized. The top-down approach

involves the introduction of exogenous genes into a host and the engineering of its native metabolic networks to reprogram cellular behavior by employing engineering and mathematical modeling toolkits. The bottom-up approach utilizes the biochemical toolkit for the *de novo* construction of synthetic genomes and unnatural components that behave in an analogous manner to their natural counterparts, and thereby allows the genesis of artificial living systems. The top-down approach of metabolic engineering for the production of useful products pertains to one of the most established concepts in the field of synthetic biology. Metabolic engineering combines transgene expression with the analysis of metabolic networks to optimize genetic and regulatory processes within cells for the production of a desired product. Metabolic engineering in a heterologous host may also involve the mathematical modeling of the host's native metabolic networks to calculate the yield of the desired product, the measurement of metabolic fluxes to pinpoint parts of the network that constrain production, genetic engineering of the host network to relieve these constraints and modeling of the modified network to calculate the product yield until an industrially applicable level is obtained (Koffas *et al.*, 1999).

Contrary to cell-based synthetic biology, in which the cell's growth and survival objectives might interfere with the engineering objective, that is, the production of a desired compound, cell-free '*in vitro* synthetic biology' provides a bottom-up platform, in which all available resources are concentrated on a user-defined objective, which could eventually result in improved production systems (Harris & Jewett, 2012). A cell-free environment is highly flexible and devoid of genetic regulation or transport barriers, facilitating substrate addition and product purification.

Alongside the engineering of organisms for enhanced production, synthetic biology also aims to create novel compounds with useful properties. One way to achieve this is by 'combinatorial biosynthesis', which allows the generation of new-to-nature compounds through the assembly of genes from different organisms, but catalyzing reactions in related pathways in a native or heterologous host, thereby establishing new enzyme-substrate combinations *in vivo* (Julsing *et al.*, 2006). An alternative way to create novel compounds is by 'directed evolution' or 'enzyme engineering'. The concept of directed enzyme evolution mimics the process of natural evolution and employs a set of methodologies to enhance or modify the function of a progenitor enzyme to accept an unnatural substrate or to catalyze a new biosynthetic reaction, thereby resulting in the formation of novel products (Dalby, 2011). Obviously, this concept can also be used in metabolic engineering for enhanced production by improving enzyme performance with its natural substrates.

2. Metabolic engineering and microbial biosynthesis of plant terpenoids

Compared with plant production systems, microorganisms are attractive alternatives as heterologous hosts because of their rapid doubling time, robustness under process conditions, ease of scalability, simplicity of product purification because of the absence of competing contaminants and cost-effectiveness

resulting from the conversion of inexpensive feedstock to valuable compounds (Zhang *et al.*, 2011). The choice of a suitable host (or 'chassis') is critical and should be based on multiple factors, including the chemical nature and complexity of the product to be synthesized, the genetic amenability of the host, the intrinsic availability of precursors for product biosynthesis, the codon usage bias of the host, the need for post-translational modifications and the feasibility to metabolically engineer the host to boost productivity (Keasling, 2010). Microbial synthesis of any plant natural product can be achieved by 'precursor-mediated product synthesis', in which an existing host pathway is altered to incorporate a heterologous pathway, or by 'de novo synthesis', in which new-to-host biosynthetic routes are imported, thereby avoiding feedback regulation (Chang & Keasling, 2006). After the establishment of heterologous synthesis, it is usually imperative to metabolically engineer the host to optimize the production yield and rate (Chemler & Koffas, 2008).

The colloquial hosts *Escherichia coli* and *Saccharomyces cerevisiae* have been employed for both precursor-mediated and *de novo* synthesis of mono-, di-, sesqui-, tri- and tetraterpenoids (Misawa, 2011), with artemisinic acid, the precursor of the antimalarial drug artemisinin, as the showcase for plant-derived terpenoids (Keasling, 2012). The prokaryotic *E. coli* has an inherent MEP pathway and the eukaryotic *S. cerevisiae* has the MVA pathway to produce IPP and its isomer DMAPP. Theoretically, terpenoid biosynthesis can be incorporated into these hosts by expressing the corresponding genes, but low yields may be obtained because of the limited intracellular IPP pool. The IPP and subsequent precursor levels have been supplemented by metabolic engineering of: (1) the MVA pathway in *E. coli* (Campos *et al.*, 2001); (2) the MEP pathway and prenyltransferases in *E. coli* (Kajiwar *et al.*, 1997); (3) the MVA pathway by a feedback regulation-deficient *HMGR* in *S. cerevisiae* (Ro *et al.*, 2006); (4) the MVA pathway by decreasing downstream enzymes to accumulate precursors in *S. cerevisiae* (Paradise *et al.*, 2008); (5) the global transcription factor regulating sterol biosynthesis in *S. cerevisiae* (Davies *et al.*, 2005); and (6) protein scaffolds for the MVA pathway in *S. cerevisiae* (Dueber *et al.*, 2009; Fig. 3).

Alongside targeted engineering, global approaches have been applied to improve the terpenoid pathway flux in microbial hosts. A 'chromosomal promoter engineering' strategy was used to express some of the endogenous MEP genes from a strong bacteriophage T5 promoter in an *E. coli* strain harboring β -carotene biosynthetic genes, resulting in the enhanced production of β -carotene relative to the parental strain (Yuan *et al.*, 2006). Similarly, a 'global transcription machinery engineering' on the *rpoD* gene encoding σ^{70} , the primary sigma factor, resulted in increased lycopene production in *E. coli* (Alper & Stephanopoulos, 2007).

Once precursor synthesis has been optimized, another major hurdle to overcome is to achieve functional expression of the pathway genes downstream of the precursor, particularly CytP450s. Plant CytP450s are endoplasmic reticulum-localized enzymes with a prerequisite for a CytP450 reductase (CPR) partner for efficient functioning (Podust & Sherman, 2012). In this regard, *S. cerevisiae*, with its native CytP450s and CPR, has an advantage over *E. coli* for the expression of complex terpenoid pathways (Hamann & Møller, 2007). Nevertheless, plant CytP450s

supplemented with a plant CPR have been successfully expressed in both *E. coli* and *S. cerevisiae* (Arsenault *et al.*, 2008).

Saccharomyces cerevisiae has already been employed for the expression of triterpenoid saponin biosynthetic genes. Through its native ergosterol biosynthesis, *S. cerevisiae* produces oxidosqualene, the precursor of saponins. In engineered strains optimized to accumulate oxidosqualene, different OSCs and CytP450s have been expressed, mainly for their functional characterization (Augustin *et al.*, 2011; Fig. 2). Engineering efforts have been limited to the production of β -amyrin only. Through a conventional pathway engineering approach, a final titer of 6 mg l^{-1} was demonstrated in an *S. cerevisiae* strain expressing a β -amyrin synthase (*bAS*) from *Artemisia annua* (Kirby *et al.*, 2008). Subsequent to a genotype-to-phenotype linking study, a 500% improvement in β -amyrin production was achieved by overexpression of the native genes, *ERG8*, *ERG9* and *HFA1*, in an *S. cerevisiae* strain expressing a *Pisum sativum* *bAS*, resulting in a final titer of 3.93 mg l^{-1} (Madsen *et al.*, 2011). The β -amyrin levels produced by the parent strains in the above reports reflect the cyclization efficiency of the enzymes employed. Therefore, by employing a more efficient *bAS* (or any other saponin biosynthetic gene), followed by targeted and/or global engineering, it should be possible to further enhance β -amyrin (or triterpenoid) levels.

The β -amyrin-producing *S. cerevisiae* strains have been utilized as a tool for the *in vivo* expression and characterization of novel CytP450s. The co-expression of a CytP450 with a plant-derived CPR resulted in the generation of yeast strains producing different sapogenins. The expression of *M. truncatula* *CYP716A12*, together with the *Lotus japonicus* *bAS* and the *L. japonicus* CPR, resulted in the production of oleanolic acid in yeast (Fukushima *et al.*, 2011; Fig. 2). β -Amyrin has also been modified to natural and rare triterpenoids by the combination of multiple CytP450s in yeast. The expression of *M. truncatula* *CYP72A68v2* and *CYP93E2* in the oleanolic acid-producing strain resulted in the production of gypsogenic acid and 4-*epi*-hederagenin, respectively (Fukushima *et al.*, 2013). In addition to β -amyrin-producing strains, α -amyrin-, lupeol- and dammarenediol-producing yeasts have been employed for the functional characterization of CytP450s. The *M. truncatula* *CYP716A12* also catalyzes the C-28 oxidations of α -amyrin to ursolic acid and lupeol to betulinic acid in yeast (Fukushima *et al.*, 2011). Similarly, the C-6 and C-12 hydroxylations of dammarenediol by *CYP716A53v2* and *CYP716A47*, respectively, have been demonstrated in yeast (Han *et al.*, 2011, 2012).

To complement metabolic engineering, synthetic biology offers a plethora of tools through the generation of minimal hosts, standard biological parts, regulatory elements, vectors, assembly methods and *in silico* computer-aided design tools (Keasling, 2012). The first and main requirement for the production of any natural product is the availability of a robust host. Synthetic biology facilitates the generation of 'minimal hosts' that contain only the genes essential for their growth to synthesize macromolecules from simple and inexpensive feedstock. Minimal hosts of *E. coli* have been generated with *c.* 15% genome reduction by the deletion of non-essential genes (Pósfai *et al.*, 2006). For *S. cerevisiae*, the synthetic yeast genome project Sc2.0 aims to design fully synthetic

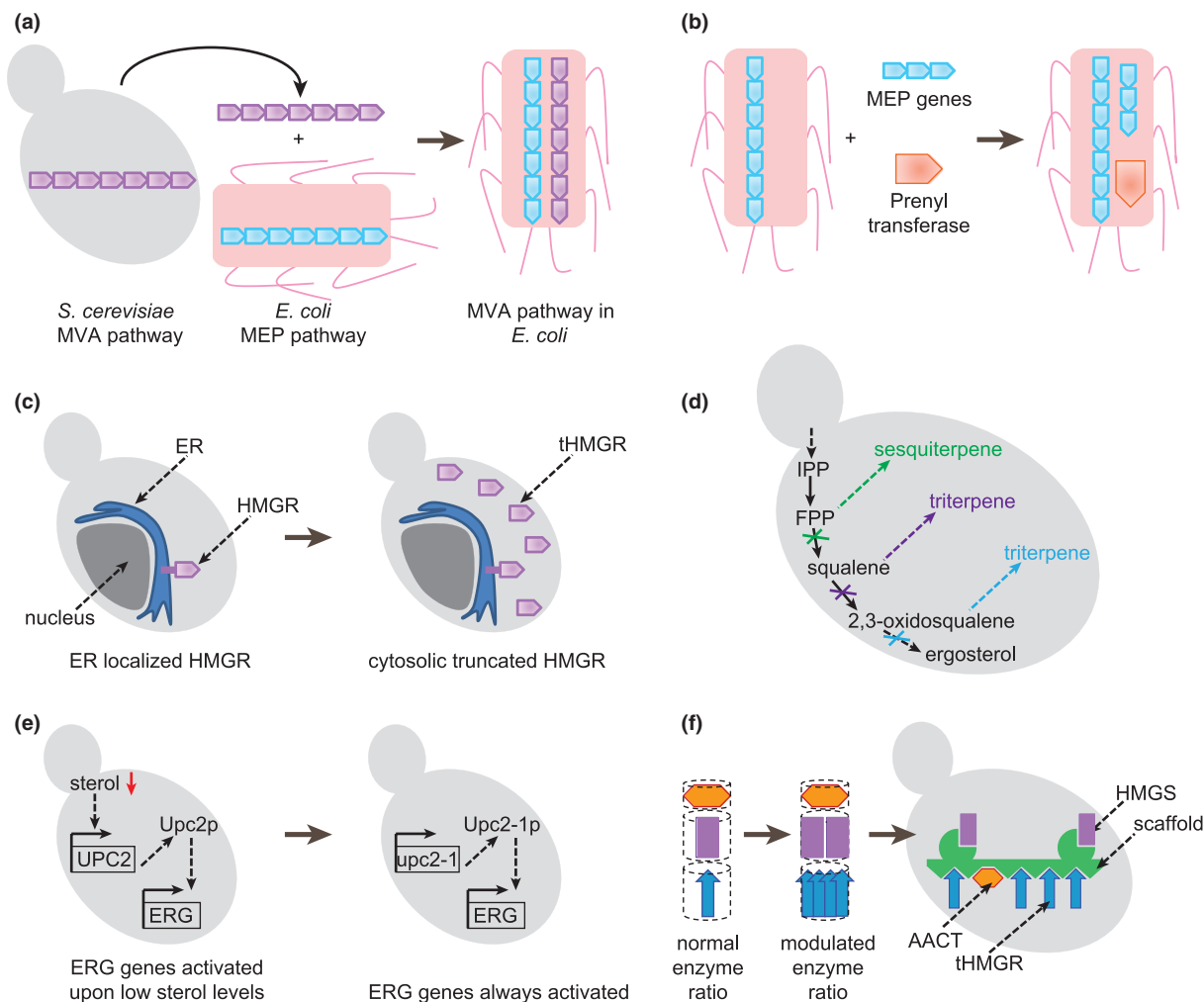


Fig. 3 Strategies employed to enhance the production of isopentenyl pyrophosphate (IPP) and terpenoids in *Escherichia coli* and *Saccharomyces cerevisiae*. (a) Expression of the *S. cerevisiae* mevalonate (MVA) pathway in *E. coli*. (b) Expression of rate-limiting 2-C-methyl-D-erythritol 4-phosphate (MEP) enzymes in *E. coli*. (c) Expression of a truncated form of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) in *S. cerevisiae*. ER, endoplasmic reticulum. (d) Downregulation of endogenous sterol biosynthesis to accumulate terpenoid precursors in *S. cerevisiae*. FPP, farnesyl pyrophosphate. (e) Expression of a mutant version (*upc2-1*) of the global transcription factor (UPC2) upregulates the expression of the native sterol biosynthesis genes in *S. cerevisiae*. (f) Protein scaffolding to prevent rate limitation in *S. cerevisiae* by the spatial organization of rate-limiting sterol biosynthetic enzymes in a modulated ratio. AACT, acetoacetyl-CoA thiolase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase.

minimized hosts without transposable elements and telomeric sequences, with relocated tRNAs and with site-specific recombination sites incorporated into the genome. Two partially synthetic *S. cerevisiae* chromosomes with genome reductions of 15–20% have been generated and successfully reincorporated (Dymond *et al.*, 2011). The *Streptomyces avermitilis* linear chromosome was reduced to 81.46% of the wild-type chromosome by stepwise deletion of a region of > 1.4 Mb, including genes coding for the synthesis of all endogenous secondary metabolites. The minimized strain was able to produce artemisin precursors on expression of a synthetic codon-optimized *A. annua* amorphaadiene synthase gene (Komatsu *et al.*, 2010). In addition, the feasibility of generating completely artificial synthetic hosts with a desired set of genes has been demonstrated by the cloning of a chemically synthesized and assembled *Mycoplasma genitalium* genome in *S. cerevisiae* (Gibson *et al.*, 2008).

Most often metabolic engineering focuses on the maximization of the production of a final compound with less attention to the behavior of intermediates. Contrary to this, the bottom-up synthetic biology approach allows the deconvolution of metabolic pathways to independent parts that are optimized for host-specific expression, and are subsequently incorporated rationally to build production modules. The repositories of functional parts (promoters, ribosomal binding sites, protein domains, terminators, etc.), generated within synthetic biology initiatives, facilitate the assembly of metabolic pathways (Boyle & Silver, 2012). Two depositories with codon-optimized parts for pathway engineering in *E. coli* (The Registry of Standard Biological Parts, partsregistry.org/Main_Page) and terpenoid engineering in *S. cerevisiae* (Serber *et al.*, 2012) have been described. Synthetic biology also promotes the variable expression of related biosynthetic genes to avoid metabolic bottlenecks. Robust synthetic promoter libraries

with defined promoter strengths enable modular gene expression in bacteria and yeast (Hammer *et al.*, 2006; Nevoigt *et al.*, 2006). Tunable intergenic regions that generate mRNA secondary structures and RNase recognition sites have been employed for the differential stabilization of segments of mRNA encoding multiple enzymes in the form of operons (Pfleger *et al.*, 2006). Synthetic protein scaffolds that are particularly efficient in overcoming rate-limiting steps have been generated to increase flux through metabolic pathways by tethering enzymes together (Dueber *et al.*, 2009).

Natural product biosynthesis typically involves multigene pathways, thus implementing the necessity for the simultaneous expression of multiple genes in a microbial chassis. Both *in vitro* and *in vivo* methods facilitate multigene assembly in *E. coli* and *S. cerevisiae* (Ellis *et al.*, 2011; Wang *et al.*, 2012), some of which have been employed to assemble carotenoid biosynthetic pathways (Shao *et al.*, 2009; Lemuth *et al.*, 2011). In parallel, viral mechanisms, such as internal ribosome entry sites and 2A oligopeptide sequences, have been adapted for polycistronic gene expression (de Felipe, 2002). However, the latter tools have not yet been implemented for the expression of terpenoid pathway enzymes.

3. Combinatorial biosynthesis of plant terpenoids

Combinatorial biosynthesis-based reconstitution of pathways is a useful tool to generate known and novel natural products, which can be further modified by semi-synthesis. In its simplest form, combinatorial biosynthesis is the process of generating different, but structurally related, molecules through the assembly of genes from different organisms in a single host (Kirschning *et al.*, 2007; Fig. 4a). Plants possess an immense potential for combinatorial biosynthesis (Pollier *et al.*, 2011). However, apart from a pioneering study, in which the expression of a bacterial halogenase in *C. roseus* resulted in the generation of novel chlorinated TIAs (Runguphan *et al.*, 2010), there have been no reports on a directed combinatorial biosynthesis approach for any other terpenoid or metabolite in plants to date. Nonetheless, the existing chemical diversity, together with our growing understanding of their biosynthesis, renders (tri)terpenoids appealing compounds for the combinatorial generation of novel analogs. For instance, the screening of a synthetic triterpenoid combinatorial library derived from betulinic and ursolic acid led to the identification of compounds with an enhanced anti-malarial activity relative to the parent compounds (Pathak *et al.*, 2002).

Combinatorial biosynthesis of triterpenoid saponins holds great potential, as they exhibit a plethora of biological activities. Bardoxolone methyl, a semi-synthetic derivative of oleanolic acid, has been clinically evaluated for the treatment of chronic kidney disease. The synthesis of bardoxolone methyl occurs through chemical modifications of the three active portions of oleanolic acid that render the derivative biologically more potent than the starter molecule (Sporn *et al.*, 2011). The enzymatic addition of extra functionalities to the triterpenoid backbone through combinatorial biosynthesis could increase the number of sites that can be accessed for further synthetic modifications (Pollier & Goossens, 2012).

A major drawback of the generation of novel molecules *in planta* lies in the complexity of plant metabolite extracts and the complications of purifying a compound of interest from a large pool of different molecules, including compounds with similar structures and physicochemical properties. Therefore, combinatorial biosynthesis of plant secondary metabolites has also been performed in microorganisms, which lack the production of compounds similar to the target compound (Fig. 4b). Novel carotenoid structures with an enhanced antioxidative activity have been generated in *E. coli* by the combinatorial expression of bacterial and plant genes (Sandmann, 2002). Recently, rare triterpenoids have been combinatorially produced in *S. cerevisiae* (Fukushima *et al.*, 2013). A current obstacle to the wider utilization of combinatorial biosynthesis for plant-derived compounds is the limited availability of plant genes encoding the enzymes that catalyze the biosynthetic reactions. In the future, these bottlenecks may be solved by gene discovery in (medicinal) plants or, alternatively, by directed evolution of enzymes towards novel functions (Kwon *et al.*, 2012).

4. Enzyme engineering or directed evolution of terpenoid biosynthetic enzymes

Small-molecule drugs, considered to be relevant as lead molecules, often have a high degree of chemical complexity with multiple functional groups and defined stereochemistry (Nannemann *et al.*, 2011). In their natural source, these small molecules are most often synthesized by enzymes that have a high regio- and stereoselectivity, high catalysis rate and relaxed substrate specificity. Nonetheless, natural enzymes often cannot meet the requirements of industrial chemists in terms of substrate tolerance, efficiency, process tolerance and economic viability. Hence, enzymes have been engineered by directed evolution to improve one or more of their properties under defined conditions (Dalby, 2011; Fig. 4c). Directed enzyme evolution has progressed tremendously lately, and it is now feasible to engineer enzymes to accept unnatural substrates and to catalyze regio- and stereospecific reactions with an efficiency comparable with that of the natural enzymes (Goldsmith & Tawfik, 2012). The promiscuous nature of proteins gives them an inherent ability to generate novel or altered functions with a small number of amino acid substitutions (Aharoni *et al.*, 2005), and computational methods, such as catalytic active site prediction (CLASP) and directed evolution using CLASP: an automated flow (Chakraborty *et al.*, 2011; Chakraborty, 2012), utilize virtual screening for spatial, electrostatic and scaffold matching to identify target progenitor proteins. Enzymes catalyzing branch-point reactions in multi-branched pathways, in which a substrate is converted to multiple products, have a high evolvability. In addition, evolvable enzymes exhibit multiple mutational residues and are 'locally specific' as they recognize a common motif on structurally diverse substrates (Umeno *et al.*, 2005).

Oxidosqualene, the immediate precursor of triterpenoid biosynthesis, is a versatile molecule that is cyclized into multiple products by different OSCs. Several of these OSCs are multifunctional in nature and generate multiple products in a single reaction (Phillips *et al.*, 2006), highlighting the promiscuity, and thus

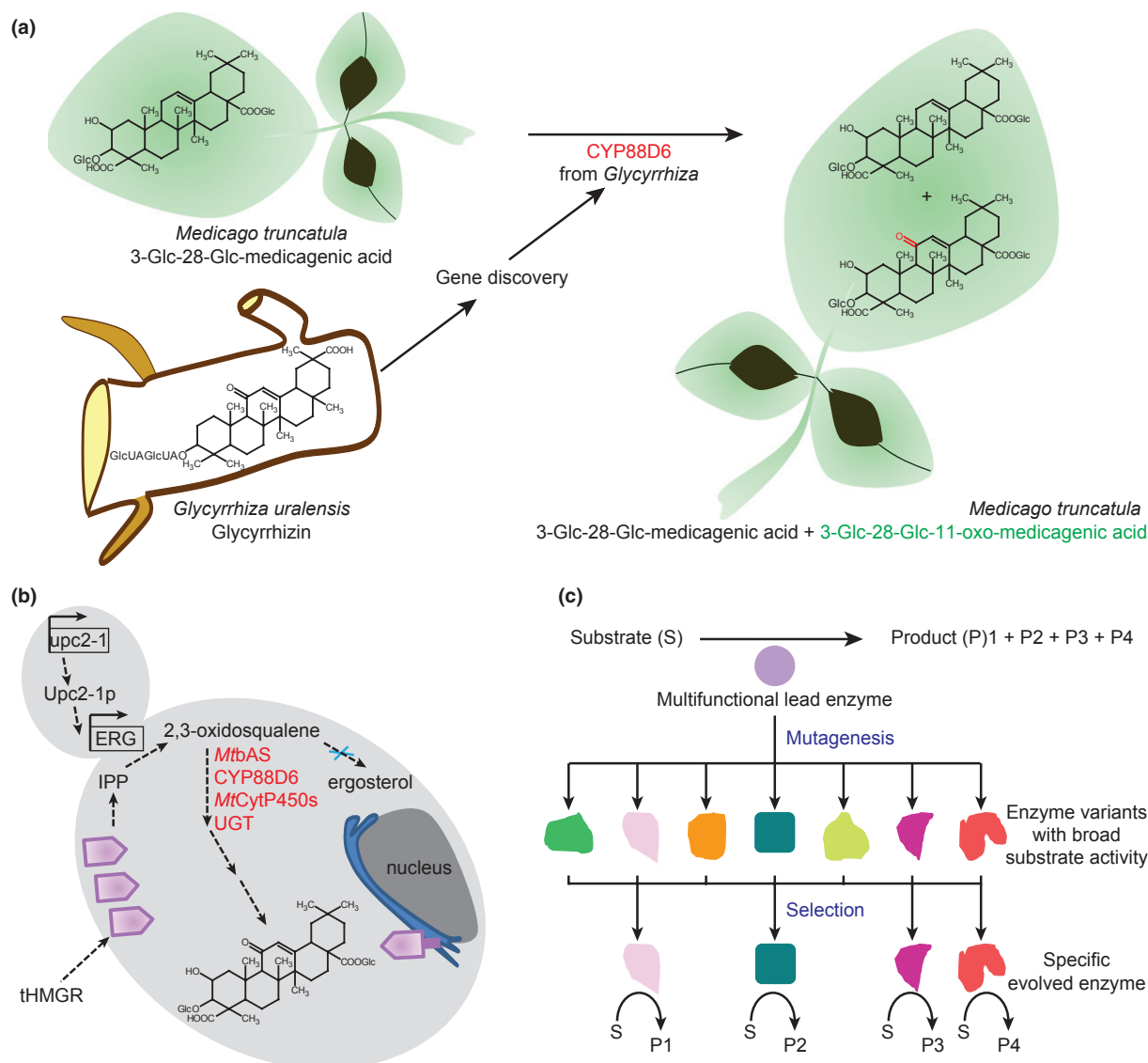


Fig. 4 Strategies to generate novel triterpenoid saponins. (a) Combinatorial biosynthesis in the model legume *Medicago truncatula* which produces 3-Glc-28-Glc-medicagenic acid endogenously. The overexpression of CYP88D6, a CytP450 from *Glycyrrhiza uralensis* roots that produces glycyrrhizin, in *M. truncatula* could lead to the formation of a combinatorial product together with the naturally occurring saponins. (b) Combinatorial biosynthesis of saponins in a sterol-reduced *Saccharomyces cerevisiae* strain by the heterologous expression of saponin biosynthetic genes from *M. truncatula* and *G. uralensis*. (c) The process of directed enzyme evolution involves mutagenesis and selection for desired enzyme properties. Here, the evolution of a multifunctional enzyme with an increased reaction specificity is depicted. Glc, glucose; GlcUA, glucuronic acid; IPP, isopentenyl pyrophosphate; MtbAS, *M. truncatula* β -amyrin synthase; MtCytP450s, *M. truncatula* cytochrome P450 monooxygenases; UGT, UDP-glucosyltransferase.

evolvability, of the enzymes. Through directed evolution, the major cyclization product of a multifunctional OSC could be redefined to a specific or novel product. For other terpene synthases, this has already been successfully attempted. Following a site-saturation mutagenesis, the specificity of a carotenoid synthase was altered to generate unnatural C_{45} and C_{50} backbones in *E. coli* (Umeno & Arnold, 2004). The product specificity of a γ -humulene synthase from *Abies grandis* that cyclizes FPP to 52 different sesquiterpenoids was evolved by site-saturation mutagenesis to generate independent synthases, each producing one or a few products derived from a predominant reaction pathway (Yoshikuni *et al.*, 2006).

This evolution approach could also be extended to downstream triterpenoid biosynthetic enzymes, in particular the CytP450s. Triterpenoid saponin backbones are made up of 30 carbons, *c.* 20 of which are accessible for CytP450-mediated modifications, as deduced from known saponins (Dinda *et al.*, 2010). In addition, diverse functional groups are observed at the modifiable carbons, pointing to the existence of specific CytP450s that catalyze these specific reactions. For instance, the C-11 position of many triterpenoid backbones can be oxidized with an α - or β -hydroxy group, and a CytP450 that specifically catalyzes the α -hydroxylation has already been characterized (Seki *et al.*, 2008). To date, only a few CytP450 families involved in triterpenoid modifications

have been identified (Fig. 2). Through directed CytP450 evolution, it should be possible to: broaden their substrate acceptance to divergent backbones, target specific carbon positions, and specify the functional group to be added to the triterpenoid skeleton. Such approaches have been implemented on carotenoid desaturases that have been evolved by random mutagenesis to accept unnatural C₃₅ carotenoid backbones in *E. coli* (Umeno & Arnold, 2003).

Protein engineering based on molecular evolution also serves as a tool to enhance enzyme efficiency or to abolish feedback regulation on enzymes. Through adaptive evolution, the unfavorable *in vivo* properties of truncated yeast HMGR were minimized for optimal functioning in *E. coli*, thereby also enhancing the final product yield by *c.* 1000-fold (Yoshikuni *et al.*, 2008). Key to directed evolution studies is a profound understanding of sequence-to-structure-to-function relationships of a protein. Integrated databases of triterpenoid cyclases (TTCED; Racolta *et al.*, 2012) and CytP450s (CYPED; Sirim *et al.*, 2009) facilitate the identification of functionally relevant and selectivity-determining amino acid residues within members of a protein family by extensive sequence analysis. Therefore, the boosting of protein engineering efforts could enhance synthetic biology efforts in triterpenoid engineering in the future.

5. *In vitro* synthetic biology: an evolving tool

In vitro synthetic biology systems can comprise 'synthetic enzymatic pathways' (SEPs), in which purified enzymes are combined in an aqueous environment to convert a substrate to a product through a series of reactions. Alternatively, 'crude extract cell-free' (CECF) systems, in which cells are grown, harvested and lysed to obtain a crude extract, can be utilized for the conversion of a substrate to a product (Hodgman & Jewett, 2012; Fig. 5a). The choice between SEP and CECF is influenced by time, cost and the need for cellular reinforcement to support the desired network. A CECF approach, for instance, is more suited for a reaction requiring a constant supply of energy, such as protein synthesis (Carlson *et al.*, 2012); however, unlike SEP, CECF reactions can exhibit undesirable activities because of the crude nature of the cellular extract.

The multireaction nature of biochemical networks, low protein concentrations, enhanced substrate diffusion, low enzyme proximity and low reaction rates as a result of unbalanced enzyme activity still hamper the efficiency of cell-free synthetic biology. *In vitro* compartmentalization (IVC) is one way of achieving proximity of reaction components. In IVC, genes are coupled to a substrate and encapsulated in water-in-oil emulsions, together with

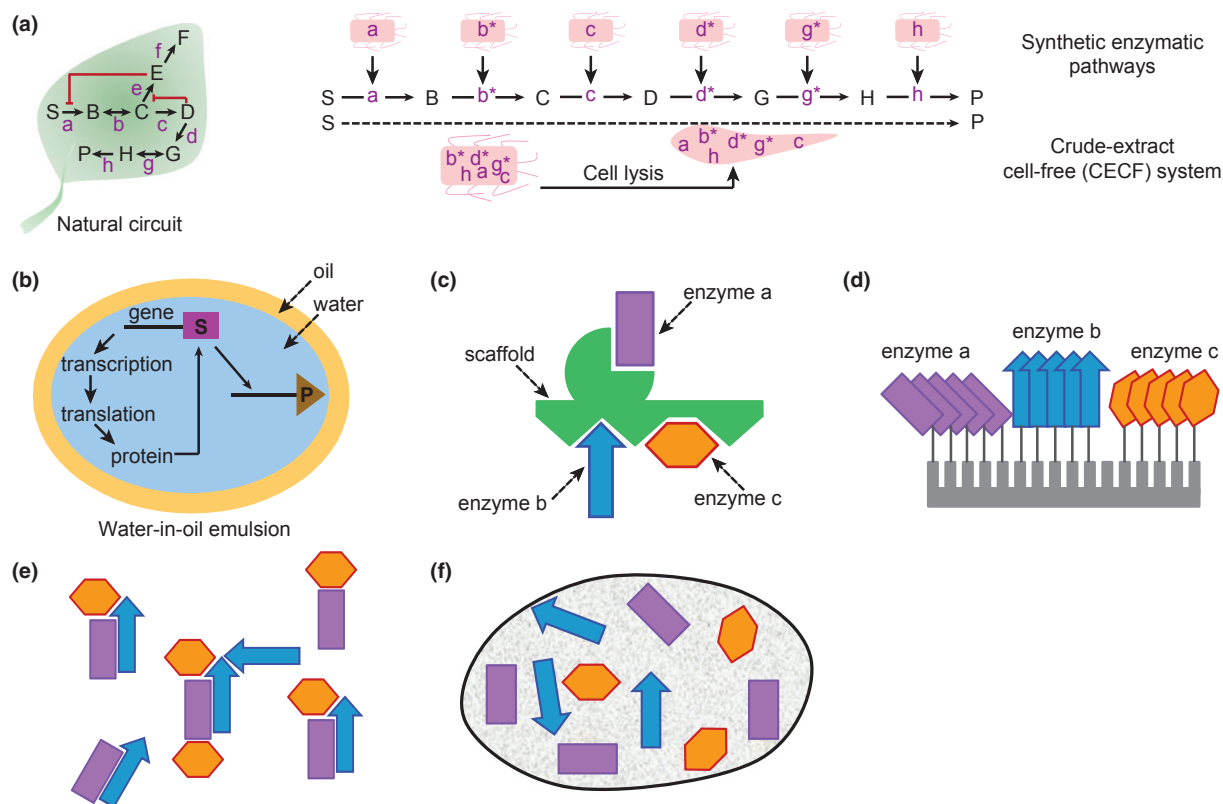


Fig. 5 *In vitro* synthetic biology platforms. (a) Synthetic enzymatic pathways in which purified enzymes are combined with reaction components in an aqueous environment to convert a substrate to a product through a series of reactions, and crude extract cell-free systems in which resources from the cell convert an exogenously provided substrate to a product. (b) *In vitro* compartmentalization using water-in-oil emulsions. The encapsulated water phase consists of a substrate coupled to a gene which is transcribed and translated *in vitro* to generate an enzyme that can convert the substrate to the product. (c–f) Metabolic channeling brings enzymes in close proximity with their substrate by (c) protein scaffolding, (d) tethering enzymes to a surface, (e) covalently linking related enzymes into aggregates and (f) foam dispersion techniques in which the enzymes are encapsulated using surfactants. a, b, c, d, e, f, g, h, native enzymes; b*, d*, g*, synthetically modified enzymes; B, C, D, E, F, G, H, intermediates; P, product; S, substrate.

transcription and translation machineries, to facilitate enzyme synthesis and consequent product formation (Fig. 5b). Novel enzymes have been uncovered by linking product formation to genes in a confined microenvironment through IVC (Rothe *et al.*, 2006). IVC is also being employed as a screening approach for the directed evolution of enzymes (Arnold & Volkov, 1999; Forster & Church, 2007). In addition to enclosing reaction components in a defined environment through IVC, metabolic channeling has been employed as an alternative to reduce substrate diffusion lengths (Idan & Hess, 2013). Protein scaffolding (Fig. 5c), surface tethering of enzymes (Fig. 5d), covalently linked enzyme aggregates (Fig. 5e) and foam dispersion of enzymes with liposomes using surfactants (Fig. 5f) have been employed to facilitate the spatial organization of pathway components (Hodgman & Jewett, 2012).

Current applications of *in vitro* synthetic biology are limited to proteins, nucleic acids and small-molecule ligands. Nonetheless, these tools can undoubtedly be extended to natural product or (tri) terpenoid engineering in the future. For instance, IVC could be employed as a tool for the directed evolution of CytP450s. A potential hurdle is the membranous nature of CytP450s, which prevents their solubilization in the aqueous reaction environment, but which may be overcome by the utilization of nanodisc membranes (Denisov & Sligar, 2011). A great advantage of using *in vitro* synthetic biology in triterpenoid engineering is the simplicity and ease of catalysis of precise regio- and stereospecific reactions with a high efficiency in a relatively pure form, which may

overcome the drawbacks of chemical synthesis, metabolic engineering and product purification.

IV. Perspectives: 'exploration of triterpenoids: the road ahead'

Triterpenoid saponins comprise a wide range of bioactive compounds, some of which (mainly pentacyclic triterpenoids) can be readily isolated from plant sources in considerable amounts for pharmacological studies or to serve as scaffolds for the semi-synthesis of new lead bioactive agents. Semi-synthetic derivatives of the natural pentacyclic triterpenoids oleanolic, ursolic and betulinic acid (Fig. 6) are a thousand-fold more active than the parent compound, and have been utilized in *in vitro* and *in vivo* studies for a broad range of clinical applications (Liby *et al.*, 2007a; Liby & Sporn, 2012; Salvador *et al.*, 2012). Such compounds certainly hold great potential, but many challenges remain. In this concluding section, we address some of the most prominent.

1. Triterpenoids that have entered clinical trials

Two types of pentacyclic triterpenoid derivatives have recently been clinically evaluated. First in class was bardoxolone (CDDO, Fig. 6), an intravenously administered semi-synthetic derivative of oleanolic acid, which was evaluated as an anti-cancer agent in patients with metastatic disease (Tsao *et al.*, 2010). Following this

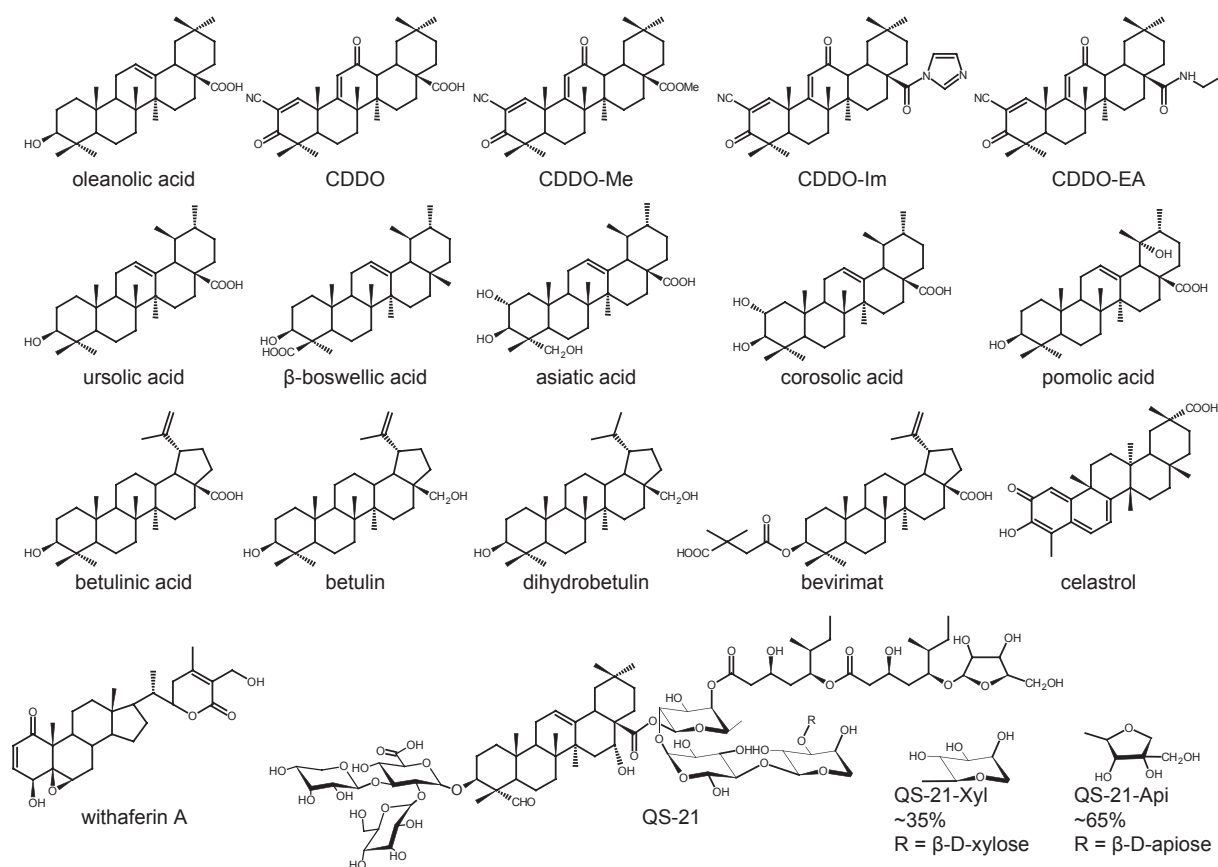


Fig. 6 Overview of the chemical structures of pharmacologically relevant triterpenoids. The *Quillaja* saponin fraction (QS-21) is composed of c. 35% QS-21-Xyl and c. 65% QS-21-Api saponins. Api, apiose; CDDO, bardoxolone; EA, ethylamide; Im, imidazolidine; Me, methyl, Xyl, xylose.

study, further efforts focused on the more potent, orally administered derivative bardoxolone methyl (CDDO-Me, Fig. 6) in patients with advanced solid tumors and lymphomas. Interestingly, 90% of the patients showed significant improvements in kidney function, without developing any serious adverse drug effects (Hong *et al.*, 2012), which prompted a phase II trial in patients with moderate to severe chronic kidney disease and type 2 diabetes. Unfortunately, the improvements in kidney function were accompanied by adverse drug effects (Pergola *et al.*, 2011). Nonetheless, a worldwide phase III trial was initiated to access the long-term clinical benefit of CDDO-Me in slowing the progression of end-stage renal disease and lessening cardiovascular death in patients with advanced chronic kidney disease and type 2 diabetes. This trial was halted in October 2012 as a result of severe adverse effects and mortality in patients taking the drug (<http://www.clinicaltrials.gov/show/NCT01351675>). Synthetic oleanane triterpenoids, such as CDDO and CDDO-Me, are multifunctional drugs with potent anti-inflammatory, anti-oxidative, anti-proliferative, pro-apoptotic and differentiating effects (Liby *et al.*, 2007b). They probably interact with multiple targets or entire regulatory networks, rather than with single molecular targets; hence, they might be most effective in the early stages of disease when a homeostatic agent is desired, contrary to an application as treatment for late-stage disease when irreversible tissue damage and cell death have occurred (Sporn *et al.*, 2007; Liby & Sporn, 2012).

The second synthetic triterpenoid to be clinically evaluated was bevirimat (Fig. 6), a betulinic acid derivative and an orally administered, novel inhibitor of human immunodeficiency virus (HIV) maturation. Bevirimat inhibits HIV type 1 (HIV-1) replication by binding to the Gag polyprotein, thereby blocking its processing and resulting in the production of non-infectious virions (Zhou *et al.*, 2005). Phase I and II clinical studies with bevirimat showed dose-proportional pharmacokinetics and no serious adverse events in HIV-1-infected adults (Smith *et al.*, 2007). However, the clinical development of bevirimat was halted in June 2010 (<http://www.clinicaltrials.gov/show/NCT01026727>). Bevirimat has been questioned with respect to its effectiveness when used in a combined therapeutic regimen with other drugs and with regard to the ability of HIV to evolve resistance (Malet *et al.*, 2007; Nijhuis *et al.*, 2007; Martínez-Cajas *et al.*, 2008; Verheyen *et al.*, 2010).

2. Is there a future for bioactive triterpenoids in therapeutics?

Many triterpenoids still hold great potential as future therapeutics in myriad applications. The synthetic oleanane triterpenoids bardoxolone imidazolide (CDDO-Im, Fig. 6) and bardoxolone ethylamide (CDDO-EA, Fig. 6) are being studied for their ability to induce chondrogenic differentiation, which, together with their potent anti-inflammatory effect, could serve to prevent or treat osteoarthritis (Suh *et al.*, 2012). CDDO-Me has the potential to be developed as a chemopreventive drug, as demonstrated by the delayed tumorigenesis in mouse cancer models (Tran *et al.*, 2013). Celastrol (Fig. 6), another oleanane triterpenoid, could be of therapeutic value for the treatment of chronic diseases, such as

asthma, arthritis, neurodegenerative diseases and cancer (Kannaiyan *et al.*, 2011).

The natural and semi-synthetic derivatives of ursane triterpenoids, such as ursolic, β -boswellic, asiatic, corosolic and pomolic acid (Fig. 6), have been investigated in cancer research for their anti-proliferative and apoptotic effects (Salvador *et al.*, 2012). A phase I study with intravenously administered ursolic acid nanoliposomes showed a linear pharmacokinetic profile and good tolerance in healthy volunteers and patients with advanced solid tumors (Zhu *et al.*, 2013).

The betulin scaffold is also still being explored for the development of new anti-HIV agents. Betulin derivatives have been recently conjugated to other anti-HIV agents to generate multi-target single agents which could simplify treatment regimens and reduce risks caused by drug–drug interactions. Hybrid conjugates of betulin and dihydrobetulin (Fig. 6) with the nucleoside reverse transcriptase inhibitor 3'-azido-3'-deoxythymidine (AZT) have been found to be more potent than bevirimat (Xiong *et al.*, 2010). Furthermore, an ointment containing the natural triterpenoid betulinic acid is being evaluated in a phase II study for the treatment of dysplastic melanocytic nevus, a likely precursor to melanoma (<http://www.clinicaltrials.gov/show/NCT00346502>).

Currently, the most promising immunological adjuvant undergoing clinical investigation is QS-21 (Fig. 6), a fraction of soluble triterpenoid glycosides from the soap bark tree (*Quillaja saponaria*; Sun *et al.*, 2009). It can augment antibody and T-cell response to a variety of antigens involved in infectious diseases, degenerative disorders and cancers. Adjuvant systems containing QS-21 in combination with other immunostimulants have been formulated to promote protective immune responses following vaccination (Garçon & Van Mechelen, 2011). Clinical studies utilizing a QS-21 adjuvant system for a candidate malaria vaccine have advanced to phase III trials, where modest protection against clinical and severe malaria was observed in African infants (RTS *et al.*, 2012). Another QS-21 adjuvant system has been employed in a phase I/II study for a candidate HIV-1 vaccine which induced T-cell response in seronegative volunteers, thus supporting further clinical investigation (Van Braeckel *et al.*, 2011).

Tetracyclic triterpenoids have been hitherto less explored, but also exhibit great therapeutic potential. Withanolides, such as withaferin A (Fig. 6), display anti-inflammatory, immunoregulatory, anti-tumor, anti-angiogenic and chemopreventive activities (Mirjalili *et al.*, 2009; Mayola *et al.*, 2011; Zhang *et al.*, 2012). Cucurbitacins have been studied for their ability to induce apoptosis in cancer cell lines (Chen *et al.*, 2012) and, like most triterpenoids, target multiple signaling networks, highlighting their usefulness as cytostatic agents (Ríos *et al.*, 2012). The ginsenosides, the tetracyclic triterpenoid glycosides from Ginseng (*Panax spp.*), have been demonstrated to possess anti-cancer activities through the modulation of diverse molecular mechanisms in various pre-clinical and clinical studies (Nag *et al.*, 2012).

The multiple mechanisms by which triterpenoids can instigate cell death impede the development of resistance against them and maintain their status as attractive candidates for drug development. Nonetheless, true proof-of-concept for their utility as effective

drugs, and ultimately market blockbusters, can only be brought about via a series of well-designed pre-clinical studies that use triterpenoid compounds in well-characterized models to unambiguously establish structure-to-activity relationships. Such information can then be exploited further to semi-synthesize even more efficacious derivatives with superior ADMET (absorption, distribution, metabolism, excretion, toxicity) properties. In addition, an in-depth understanding of the molecular mechanisms that underlie their biological activities will be necessary to harness their full potential.

3. The need for more 'plant' knowledge

In addition to the cost and effort involved in the drug discovery and development process itself, pharmaceutical companies often face another major challenge, which is to be able to scale up the production of the active principle and make the process cost-efficient, and, last but not least in the case of natural products, sustainable!

Although some triterpenoids, such as oleanolic acid, can be extracted from by-products of the olive (oil) industry, and thus are available in ample amounts (Pollier & Goossens, 2012), many others, such as the ginsenosides, are scarce, and extraction from plants alone is insufficient. In addition, the triterpenoid profiles of plants are variable and often influenced by environmental factors, which may affect the quantity and quality of the bioactive principle that can be extracted from the same biomass. Furthermore, triterpenoid-producing plants may have a slow growth rate or be difficult to grow, which makes cultivation non-profitable to farmers. Even when a natural product drug can be produced in large amounts *in planta*, there can be supply and demand imbalances, which may feed back to fluctuations in cultivation acreages and yields. A metabolic engineering or synthetic biology platform may provide an alternative and sustainable prospect to agricultural supply by creating a complementary non-seasonal, high-quality source for valuable bioactive triterpenoids. Clearly, the development of alternative performing heterologous production platforms will be accompanied by multiple challenges, which need to be balanced against the concerns about the (stability of the) market value of the drug to be produced.

The artemisinin case has shown that synthetic biology can reach industrial-scale deployment for drug production (Keasling, 2012; <http://www.nature.com/news/malaria-drug-made-in-yeast-causes-market-ferment-1.12417>). In the case of triterpenoids, bioengineering may follow the beaten track established for semi-synthetic artemisinin. However, it may also involve distinct host optimization for large-scale triterpenoid production. Obviously, yeasts will remain potent vehicles, but microalgae or plants amenable to culture in bioreactors and engineering technologies certainly represent attractive alternative hosts for a triterpenoid-oriented synthetic biology program.

A major restraint to the successful bioengineering of plant-derived triterpenoids is the scarcity of indispensable knowledge about their biosynthesis, which hampers both plant and microbial engineering. Most triterpenoid saponins are known to accumulate in a tissue-, organ- or signal-specific manner in plants, but there is

virtually no insight into the mechanisms responsible for this pattern. Multiple OSCs catalyzing the cyclization of 2,3-oxidosqualene to different triterpenoid precursor backbones have been isolated already, but only a handful of genes corresponding to the 'decorating' enzymes have been identified, whereas hundreds must exist when considering the structural diversity of triterpenoids in the plant kingdom. Similarly, although the biosynthetic enzymes are mostly microsomal in nature, triterpenoids typically localize to the epidermal wax layer or the vacuoles, suggesting the existence of yet undiscovered transporter systems. Hence, there is a great need to unravel the molecular mechanisms involved in triterpenoid saponin production *in planta* to assist their exogenous engineering.

Fortunately, the booming number of functional genomics technologies with ever-increasing resolution and coverage of the genome, transcriptome, proteome, interactome and metabolome will offer the necessary power to list all the possible elements involved in the synthesis of plant terpenoids in the near future. In particular, the linking of signal- and tissue-dependent metabolome and transcriptome analysis will remain a powerful principle to pinpoint biosynthetic genes, transporters and transcription factors. If successful triterpenoid-related gene discovery can profit from the numerous tools and platforms that are meanwhile being developed in the field of synthetic biology to reduce the cost and time required to engineer biological systems, triterpenoid bioengineering awaits a bright future.

Acknowledgements

We thank Annick Bleys for help in preparing the manuscript. This work was supported by the European Union Seventh Framework Programme FP7/2007–2013 under grant agreement number 222716 – SMARTCELL. T.M. is indebted to the VIB International PhD Fellowship Program for a predoctoral fellowship. J.P. is a postdoctoral fellow of the Research Foundation-Flanders.

References

- Agrell J, Anderson P, Oleszek W, Stochmal A, Agrell C. 2004. Combined effects of elevated CO₂ and herbivore damage on alfalfa and cotton. *Journal of Chemical Ecology* 30: 2309–2324.
- Agrell J, Oleszek W, Stochmal A, Olsen M, Anderson P. 2003. Herbivore-induced responses in alfalfa (*Medicago sativa*). *Journal of Chemical Ecology* 29: 303–320.
- Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C, Tawfik DS. 2005. The 'evolvability' of promiscuous protein functions. *Nature Genetics* 37: 73–76.
- Alper H, Stephanopoulos G. 2007. Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metabolic Engineering* 9: 258–267.
- Arnold FH, Volkov AA. 1999. Directed evolution of biocatalysts. *Current Opinion in Chemical Biology* 3: 54–59.
- Arsenault PR, Wobbe KK, Weathers PJ. 2008. Recent advances in artemisinin production through heterologous expression. *Current Medicinal Chemistry* 15: 2886–2896.
- Augustin JM, Kuzina V, Andersen SB, Bak S. 2011. Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochemistry* 72: 435–457.
- Bailey JE. 1991. Toward a science of metabolic engineering. *Science* 252: 1668–1675.

- van Beilen JB, Poirier Y. 2007. Establishment of new crops for the production of natural rubber. *Trends in Biotechnology* 25: 522–529.
- Bohlmann J, Keeling CI. 2008. Terpenoid biomaterials. *Plant Journal* 54: 656–669.
- Bouwmeester HJ. 2006. Engineering the essence of plants. *Nature Biotechnology* 24: 1359–1361.
- Boyle PM, Silver PA. 2012. Parts plus pipes: synthetic biology approaches to metabolic engineering. *Metabolic Engineering* 14: 223–232.
- Broun P. 2004. Transcription factors as tools for metabolic engineering in plants. *Current Opinion in Biotechnology* 7: 202–209.
- Campos N, Rodríguez-Concepción M, Sauret-Güeto S, Gallego F, Lois LM, Boronat A. 2001. *Escherichia coli* engineered to synthesize isopentenyl diphosphate and dimethylallyl diphosphate from mevalonate: a novel system for the genetic analysis of the 2-C-methyl-d-erythritol 4-phosphate pathway for isoprenoid biosynthesis. *The Biochemical Journal* 353: 59–67.
- Carlson ED, Gan R, Hodgman CE, Jewett MC. 2012. Cell-free protein synthesis: applications come of age. *Biotechnology Advances* 30: 1185–1194.
- Chakraborty S. 2012. An automated flow for directed evolution based on detection of promiscuous scaffolds using spatial and electrostatic properties of catalytic residues. *PLoS ONE* 7: e40408.
- Chakraborty S, Minda R, Salaye L, Bhattacharjee SK, Rao BJ. 2011. Active site detection by spatial conformity and electrostatic analysis—unravelling a proteolytic function in shrimp alkaline phosphatase. *PLoS ONE* 6: e28470.
- Chang MCY, Keasling JD. 2006. Production of isoprenoid pharmaceuticals by engineered microbes. *Nature Chemical Biology* 2: 674–681.
- Channon K, Bromley EH, Woolfson DN. 2008. Synthetic biology through biomolecular design and engineering. *Current Opinion in Structural Biology* 18: 491–498.
- Chemler JA, Koffas MAG. 2008. Metabolic engineering for plant natural product biosynthesis in microbes. *Current Opinion in Biotechnology* 19: 597–605.
- Chen F, Tholl D, Bohlmann J, Pichersky E. 2011. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *The Plant Journal* 66: 212–229.
- Chen X, Bao J, Guo J, Ding Q, Lu J, Huang M, Wang Y. 2012. Biological activities and potential molecular targets of cucurbitacins: a focus on cancer. *Anti-cancer Drugs* 23: 777–787.
- Croteau R, Kutchan TM, Lewis NG. 2000. Natural products (secondary metabolites). In: Buchanan B, Gruissem W, Jones R, eds. *Biochemistry & molecular biology of plants*. Rockville, MD, USA: American Society of Plant Biologists, 1250–1318.
- Dalby PA. 2011. Strategy and success for the directed evolution of enzymes. *Current Opinion in Structural Biology* 21: 473–480.
- Davies BS, Wang HS, Rine J. 2005. Dual activators of the sterol biosynthetic pathway of *Saccharomyces cerevisiae*: similar activation/regulatory domains but different response mechanisms. *Molecular and Cellular Biology* 25: 7375–7385.
- De Geyter N, Gholami A, Goormachtig S, Goossens A. 2012. Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends in Plant Science* 17: 349–359.
- Degenhardt J, Gershenzon J, Baldwin IT, Kessler A. 2003. Attracting friends to feast on foes: engineering terpene emission to make crop plants more attractive to herbivore enemies. *Current Opinion in Biotechnology* 14: 169–176.
- Denisov IG, Sligar SG. 2011. Cytochromes P450 in nanodiscs. *Biochimica et Biophysica Acta* 1814: 223–229.
- Dewick PM. 2001. *Medicinal natural products: a biosynthetic approach*. Chichester, UK: John Wiley & Sons Ltd.
- Dinda B, Debnath S, Mohanta BC, Harigaya Y. 2010. Naturally occurring triterpenoid saponins. *Chemistry & Biodiversity* 7: 2327–2580.
- Doblas VG, Amorim-Silva V, Posé D, Rosado A, Esteban A, Arró M, Azevedo H, Bombarely A, Borsani O, Valpuesta V *et al.* 2013. The *SUD1* gene encodes a putative E3 ubiquitin ligase and is a positive regulator of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity in *Arabidopsis*. *Plant Cell* 25: 728–743.
- Dudareva N, Klempien A, Muhlemann JK, Kaplan I. 2013. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist* 198: 16–32.
- Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KL, Keasling JD. 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology* 27: 753–759.
- Dymond JS, Richardson SM, Coombes CE, Babatz T, Muller H, Annaluru N, Blake WJ, Schwerzmann JW, Dai J, Lindstrom DL *et al.* 2011. Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* 477: 471–476.
- Ellis T, Adie T, Baldwin GS. 2011. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integrative Biology* 3: 109–118.
- Farhi M, Marheva E, Ben-Ari J, Algamas-Dimantov A, Liang Z, Zeevi V, Edelbaum O, Spitzer-Rimon B, Abeliovich H, Schwartz B *et al.* 2011. Generation of the potent anti-malarial drug artemisinin in tobacco. *Nature Biotechnology* 29: 1072–1074.
- Farré G, Bai C, Twyman RM, Capell T, Christou P, Zhu C. 2011. Nutritious crops producing multiple carotenoids—a metabolic balancing act. *Trends in Plant Science* 16: 532–540.
- de Felipe P. 2002. Polycistronic viral vectors. *Current Gene Therapy* 2: 355–378.
- Field B, Osbourn AE. 2008. Metabolic diversification—-independent assembly of operon-like gene clusters in different plants. *Science* 320: 543–547.
- van der Fits L, Memelink J. 2000. ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* 289: 295–297.
- Forster AC, Church GM. 2007. Synthetic biology projects *in vitro*. *Genome Research* 17: 1–6.
- Fujioka S, Yokota T. 2003. Biosynthesis and metabolism of brassinosteroids. *Annual Reviews of Plant Biology* 54: 137–164.
- Fukushima EO, Seki H, Ohya K, Ono E, Umemoto N, Mizutani M, Saito K, Muranaka T. 2011. CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant and Cell Physiology* 52: 2050–2061.
- Fukushima EO, Seki H, Sawai S, Suzuki M, Ohya K, Saito K, Muranaka T. 2013. Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. *Plant and Cell Physiology*. doi: 10.1093/pcp/pct015.
- Garçon N, Van Mechelen M. 2011. Recent clinical experience with vaccines using MPL- and QS-21-containing adjuvant systems. *Expert Review of Vaccines* 10: 471–486.
- Georgiev MI, Agostini E, Ludwig-Müller J, Xu J. 2012. Genetically transformed roots: from plant disease to biotechnological resource. *Trends in Biotechnology* 30: 528–537.
- Georgiev MI, Weber J, Maciuk A. 2009. Bioprocessing of plant cell cultures for mass production of targeted compounds. *Applied Microbiology and Biotechnology* 83: 809–823.
- Gershenzon J, Dudareva N. 2007. The function of terpene natural products in the natural world. *Nature Chemical Biology* 3: 408–414.
- Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA *et al.* 2008. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319: 1215–1220.
- Goldsmith M, Tawfik DS. 2012. Directed enzyme evolution: beyond the low-hanging fruit. *Current Opinion in Structural Biology* 22: 1–7.
- Hamann T, Möller BL. 2007. Improved cloning and expression of cytochrome P450s and cytochrome P450 reductase in yeast. *Protein Expression and Purification* 56: 121–127.
- Hammer K, Mijakovic I, Jensen PR. 2006. Synthetic promoter libraries—tuning of gene expression. *Trends in Biotechnology* 24: 53–55.
- Han JY, Hwang HS, Choi SW, Kim HJ, Choi YE. 2012. Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*. *Plant and Cell Physiology* 53: 1535–1545.
- Han JY, Kim HJ, Kwon YS, Choi YE. 2011. The Cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammaradiol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant and Cell Physiology* 52: 2062–2073.
- Haralampidis K, Bryan G, Qi X, Papadopoulou K, Bakht S, Melton R, Osbourn A. 2001. A new class of oxidosqualene cyclases directs synthesis of antimicrobial phytoprotectants in monocots. *Proceedings of the National Academy of Sciences, USA* 98: 13431–13436.
- Harris DC, Jewett MC. 2012. Cell-free biology: exploiting the interface between synthetic biology and synthetic chemistry. *Current Opinion in Biotechnology* 23: 672–678.
- Hodgman CE, Jewett MC. 2012. Cell-free synthetic biology: thinking outside the cell. *Metabolic Engineering* 14: 261–269.

- Hong DS, Kurzrock R, Supko JG, He X, Naing A, Wheler J, Lawrence D, Eder JP, Meyer CJ, Ferguson DA *et al.* 2012. A phase I first-in-human trial of bardoxolone methyl in patients with advanced solid tumors and lymphomas. *Clinical Cancer Research* 18: 3396–3406.
- Idan O, Hess H. 2013. Engineering enzymatic cascades on nanoscale scaffolds. *Current Opinion in Biotechnology*. doi: 10.1016/j.copbio.2013.01.003.
- Julsing MK, Koulman A, Woerdenbag HJ, Quax WJ, Kayser O. 2006. Combinatorial biosynthesis of medicinal plant secondary metabolites. *Biomolecular Engineering* 23: 265–279.
- Kajiwarra S, Fraser PD, Kondo K, Misawa N. 1997. Expression of an exogenous isopentenyl diphosphate isomerase gene enhances isoprenoid biosynthesis in *Escherichia coli*. *The Biochemical Journal* 324: 421–426.
- Kannaiyan R, Shanmugam MK, Sethi G. 2011. Molecular targets of celastrol derived from Thunder of God Vine: potential role in the treatment of inflammatory disorders and cancer. *Cancer Letters* 303: 9–20.
- Keasling JD. 2008. Synthetic biology for synthetic chemistry. *ACS Chemical Biology* 3: 64–76.
- Keasling JD. 2010. Manufacturing molecules through metabolic engineering. *Science* 330: 1355–1358.
- Keasling JD. 2012. Synthetic biology and the development of tools for metabolic engineering. *Metabolic Engineering* 14: 189–195.
- Kirby J, Romanini DW, Paradise EM, Keasling JD. 2008. Engineering triterpene production in *Saccharomyces cerevisiae*-beta-amyrin synthase from *Artemisia annua*. *FEBS Journal* 275: 1852–1859.
- Kirschning A, Taft F, Knobloch T. 2007. Total synthesis approaches to natural product derivatives based on the combination of chemical synthesis and metabolic engineering. *Organic and Biomolecular Chemistry* 5: 3245–3259.
- Koehn FE, Carter GT. 2005. The evolving role of natural products in drug discovery. *Nature Reviews: Drug Discovery* 4: 206–220.
- Koffas M, Roberge C, Lee K, Stephanopoulos G. 1999. Metabolic engineering. *Annual Review of Biomedical Engineering* 1: 535–557.
- Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H. 2010. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proceedings of the National Academy of Sciences, USA* 107: 2646–2651.
- Kumar S, Hahn FM, Baidoo E, Kahlon TS, Wood DF, McMahan CM, Cornish K, Keasling JD, Daniell H, Whalen MC. 2012. Remodeling the isoprenoid pathway in tobacco by expressing the cytoplasmic mevalonate pathway in chloroplasts. *Metabolic Engineering* 14: 19–28.
- Kwon SJ, Mora-Pale M, Lee MY, Dordick JS. 2012. Expanding nature's small molecule diversity via *in vitro* biosynthetic pathway engineering. *Current Opinion in Chemical Biology* 16: 186–195.
- Lambert E, Faizal A, Geelen D. 2011. Modulation of triterpene saponin production: *in vitro* cultures, elicitation, and metabolic engineering. *Applied Biochemistry and Biotechnology* 164: 220–237.
- Lange BM, Turner GW. 2013. Terpenoid biosynthesis in trichomes—current status and future opportunities. *Plant Biotechnology Journal* 11: 2–22.
- Leivar P, Antolín-Llovera M, Ferrero S, Closa M, Arró M, Ferrer A, Boronat A, Campos N. 2011. Multilevel control of Arabidopsis 3-hydroxy-3-methylglutaryl coenzyme A reductase by protein phosphatase 2A. *Plant Cell* 23: 1494–1511.
- Lemuth K, Steuer K, Albermann C. 2011. Engineering of a plasmid-free *Escherichia coli* strain for improved *in vivo* biosynthesis of astaxanthin. *Microbial Cell Factories* 10: 29.
- Liby K, Honda T, Williams CR, Risingsong R, Royce DB, Suh N, Dinkova-Kostova AT, Stephenson KK, Talalay P, Sundararajan C *et al.* 2007a. Novel semisynthetic analogues of betulinic acid with diverse cytoprotective, antiproliferative, and proapoptotic activities. *Molecular Cancer Therapeutics* 6: 2113–2119.
- Liby KT, Sporn MB. 2012. Synthetic oleanane triterpenoids: multifunctional drugs with a broad range of applications for prevention and treatment of chronic disease. *Pharmacological Reviews* 64: 972–1003.
- Liby KT, Yore MM, Sporn MB. 2007b. Triterpenoids and rexinoids as multifunctional agents for the prevention and treatment of cancer. *Nature Reviews Cancer* 7: 357–369.
- Lim EK, Bowles D. 2012. Plant production systems for bioactive small molecules. *Current Opinion in Biotechnology* 23: 271–277.
- Madsen KM, Udatha GD, Semba S, Otero JM, Koetter P, Nielsen J, Ebizuka Y, Kushi T, Panagiotou G. 2011. Linking genotype and phenotype of *Saccharomyces cerevisiae* strains reveals metabolic engineering targets and leads to triterpene hyper-producers. *PLoS ONE* 6: e14763.
- Malet I, Wiriden M, Derache A, Simon A, Katlama C, Calvez V, Marcelin AG. 2007. Primary genotypic resistance of HIV-1 to the maturation inhibitor PA-457 in protease inhibitor-experienced patients. *AIDS* 21: 871–873.
- Martínez-Cajas JL, Pant-Pai N, Klein MB, Wainberg MA. 2008. Role of genetic diversity amongst HIV-1 non-B subtypes in drug resistance: a systematic review of virologic and biochemical evidence. *AIDS Reviews* 10: 212–223.
- Mayola E, Gallerne C, Esposti DD, Martel C, Pervais S, Larue L, Debuire B, Lemoine A, Brenner C, Lemaire C. 2011. Withaferin A induces apoptosis in human melanoma cells through generation of reactive oxygen species and down-regulation of Bcl-2. *Apoptosis* 16: 1014–1027.
- Mirjalili MH, Moyano E, Bonfill M, Cusido RM, Palazón J. 2009. Steroidal lactones from *Withania somnifera*, an ancient plant for novel medicine. *Molecules* 14: 2373–2393.
- Misawa N. 2011. Pathway engineering for functional isoprenoids. *Current Opinion in Biotechnology* 22: 627–633.
- Nag SA, Qin JJ, Wang W, Wang MH, Wang H, Zhang R. 2012. Ginsenosides as anticancer agents: *in vitro* and *in vivo* activities, structure–activity relationships, and molecular mechanisms of action. *Frontiers in Pharmacology* 3: 25.
- Nagegowda DA. 2010. Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. *FEBS Letters* 584: 2965–2973.
- Nannemann DP, Birmingham WR, Scism RA, Bachmann BO. 2011. Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. *Future Medicinal Chemistry* 3: 809–819.
- Nes WD. 2011. Biosynthesis of cholesterol and other sterols. *Chemical Reviews* 111: 6423–6451.
- Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. 2006. Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* 72: 5266–5273.
- Nijhuis M, van Maarseveen NM, Laster S, Schipper P, Coakley E, Glass B, Rovenska M, de Jong D, Chappay C, Goedegebuure IW *et al.* 2007. A novel substrate-based HIV-1 protease inhibitor drug resistance mechanism. *PLoS Medicine* 4: e36.
- Paradise EM, Kirby J, Chan R, Keasling JD. 2008. Redirection of flux through the FPP branch-point in *Saccharomyces cerevisiae* by down-regulating squalene synthase. *Biotechnology and Bioengineering* 100: 371–378.
- Pathak A, Singh SK, Biabani MA, Kulshreshtha DK, Puri SK, Srivastava S, Kundu B. 2002. Synthesis of combinatorial libraries based on terpenoid scaffolds. *Combinatorial Chemistry and High Throughput Screening* 5: 241–248.
- Pauwels L, Goossens A. 2011. The JAZ proteins: a crucial interface in the jasmonate signaling cascade. *Plant Cell* 23: 3089–3100.
- Pauwels L, Inzé D, Goossens A. 2009. Jasmonate-inducible gene: what does it mean? *Trends in Plant Science* 14: 87–91.
- Pergola PE, Raskin P, Toto RD, Meyer CJ, Huff JW, Grossman EB, Krauth M, Ruiz S, Audhya P, Christ-Schmidt H *et al.* 2011. Bardoxolone methyl and kidney function in CKD with type 2 diabetes. *New England Journal of Medicine* 365: 327–336.
- Pflegler BF, Pitera DJ, Smolke CD, Keasling JD. 2006. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nature Biotechnology* 24: 1027–1032.
- Phillips DR, Rasbery JM, Bartel B, Matsuda SP. 2006. Biosynthetic diversity in plant triterpene cyclization. *Current Opinion in Plant Biology* 9: 305–314.
- Podust LM, Sherman DH. 2012. Diversity of P450 enzymes in the biosynthesis of natural products. *Natural Product Reports* 29: 1251–1266.
- Pollier J, Goossens A. 2012. Oleanolic acid. *Phytochemistry* 77: 10–15.
- Pollier J, Moses T, Goossens A. 2011. Combinatorial biosynthesis in plants: a (p) review on its potential and future exploitation. *Natural Product Reports* 28: 1897–1916.
- Pósfai G, Plunkett G, Fehér T, Frisch D, Keil GM, Umenhoffer K, Kolisnichenko V, Stahl B, Sharma SS, de Arruda M *et al.* 2006. Emergent properties of reduced-genome *Escherichia coli*. *Science* 312: 1044–1046.
- Qi X, Bakht S, Qin B, Leggett M, Hemmings A, Mellon F, Eagles J, Werck-Reichhart D, Schaller H, Lesot A *et al.* 2006. A different function for a member of

- an ancient and highly conserved cytochrome P450 family: from essential sterols to plant defense. *Proceedings of the National Academy of Sciences, USA* 103: 18843–18853.
- Racolta S, Juhl PB, Sirim D, Pleiss J. 2012. The triterpene cyclase protein family: a systematic analysis. *Proteins* 80: 2009–2019.
- Ríos JL, Andújar I, Escandell JM, Giner RM, Recio MC. 2012. Cucurbitacins as inducers of cell death and a rich source of potential anticancer compounds. *Current Pharmaceutical Design* 18: 1663–1676.
- Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J *et al.* 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440: 940–943.
- Roberts SC. 2007. Production and engineering of terpenoids in plant cell culture. *Nature Chemical Biology* 3: 387–395.
- Rothe A, Surjadi RN, Power BE. 2006. Novel proteins in emulsions using *in vitro* compartmentalization. *Trends in Biotechnology* 24: 587–592.
- RTS,S Clinical Trials Partnership, Agnandji ST, Lell B, Fernandes JF, Abossolo BP, Methogo BG, Kabwende AL, Adegnik AA, Mordmüller B, Issifou S *et al.* 2012. A phase 3 trial of RTS,S/AS01 malaria vaccine in African infants. *New England Journal of Medicine* 367: 2284–2295.
- Runguphan W, Qu X, O'Connor SE. 2010. Integrating carbon–halogen bond formation into medicinal plant metabolism. *Nature* 468: 461–464.
- Salvador JA, Moreira VM, Gonçalves BM, Leal AS, Jing Y. 2012. Ursane-type pentacyclic triterpenoids as useful platforms to discover anticancer drugs. *Natural Product Reports* 29: 1463–1479.
- Sandmann G. 2002. Combinatorial biosynthesis of carotenoids in a heterologous host: a powerful approach for the biosynthesis of novel structures. *ChemBioChem: A European Journal of Chemical Biology* 3: 629–635.
- Sawai S, Saito K. 2011. Triterpenoid biosynthesis and engineering in plants. *Frontiers in Plant Science* 2: 25.
- Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K, Muranaka T. 2008. Licorice beta-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proceedings of the National Academy of Sciences, USA* 105: 14204–14209.
- Serber Z, Lowe R, Ubersaz JA, Chandran SS. 2012. *Compositions and methods for the assembly of polynucleotides*. USA: Amyris, Inc. [US patent 8,110,360].
- Shao Z, Zhao H, Zhao H. 2009. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Research* 37: e16.
- Sirim D, Wagner F, Lisitsa A, Pleiss J. 2009. The cytochrome P450 engineering database: integration of biochemical properties. *BMC Biochemistry* 10: 27.
- Smith PF, Ogundele A, Forrest A, Wilton J, Salzwedel K, Doto J, Allaway GP, Martin DE. 2007. Phase I and II study of the safety, virologic effect, and pharmacokinetics/pharmacodynamics of single-dose 3-o-(3',3'-dimethylsuccinyl)betulinic acid (bevirimat) against human immunodeficiency virus infection. *Antimicrobial Agents and Chemotherapy* 51: 3574–3581.
- Sporn MB, Liby KT, Yore MM, Fu L, Lopchuk JM, Gribble GW. 2011. New synthetic triterpenoids: potent agents for prevention and treatment of tissue injury caused by inflammatory and oxidative stress. *Journal of Natural Products* 74: 537–545.
- Sporn MB, Liby K, Yore MM, Suh N, Albini A, Honda T, Sundararajan C, Gribble GW. 2007. Platforms and networks in triterpenoid pharmacology. *Drug Development Research* 68: 174–182.
- Suh N, Paul S, Lee HJ, Yoon T, Shah N, Son AI, Reddi AH, Medici D, Sporn MB. 2012. Synthetic triterpenoids, CDDO-Imidazolide and CDDO-Ethyl amide, induce chondrogenesis. *Osteoarthritis Cartilage* 20: 446–450.
- Sun HX, Xie Y, Ye YP. 2009. Advances in saponin-based adjuvants. *Vaccine* 27: 1787–1796.
- Suzuki H, Reddy MS, Naoumkina M, Aziz N, May GD, Huhman DV, Sumner LW, Blount JW, Mendes P, Dixon RA. 2005. Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. *Planta* 220: 696–707.
- Tholl D. 2006. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Current Opinion in Plant Biology* 9: 297–304.
- Tran K, Risingsong R, Royce DB, Williams CR, Sporn MB, Pioli PA, Gediya LK, Njar VC, Liby KT. 2013. The combination of the histone deacetylase inhibitor vorinostat and synthetic triterpenoids reduces tumorigenesis in mouse models of cancer. *Carcinogenesis* 34: 199–210.
- Tsao T, Kornblau S, Safe S, Watt JC, Ruvo V, Chen W, Qiu Y, Coombes KR, Ju Z, Abdelrahman M *et al.* 2010. Role of peroxisome proliferator-activated receptor-gamma and its coactivator DRIP205 in cellular responses to CDDO (RTA-401) in acute myelogenous leukemia. *Cancer Research* 70: 4949–4960.
- Umeno D, Arnold FH. 2003. A C35 carotenoid biosynthetic pathway. *Applied and Environmental Microbiology* 69: 3573–3579.
- Umeno D, Arnold FH. 2004. Evolution of a pathway to novel long-chain carotenoids. *Journal of Bacteriology* 186: 1531–1536.
- Umeno D, Tobias AV, Arnold FH. 2005. Diversifying carotenoid biosynthetic pathways by directed evolution. *Microbiology and Molecular Biology Reviews* 69: 51–78.
- Van Braeckel E, Bourguignon P, Koutsoukos M, Clement F, Janssens M, Carletti I, Collard A, Demoitie MA, Voss G, Leroux-Roels G *et al.* 2011. An adjuvanted polyprotein HIV-1 vaccine induces polyfunctional cross-reactive CD4+ T cell responses in seronegative volunteers. *Clinical Infectious Diseases* 52: 522–531.
- Verheyen J, Verhofstede C, Knops E, Vandekerckhove L, Fun A, Brunen D, Dauwe K, Wensing AM, Pfister H, Kaiser R *et al.* 2010. High prevalence of bevirimat resistance mutations in protease inhibitor-resistant HIV isolates. *AIDS* 24: 669–673.
- Vranová E, Coman D, Gruijssem W. 2012. Structure and dynamics of the isoprenoid pathway network. *Molecular Plant* 5: 318–333.
- Wang T, Ma X, Zhu H, Li A, Du G, Chen J. 2012. Available methods for assembling expression cassettes for synthetic biology. *Applied Microbiology and Biotechnology* 93: 1853–1863.
- Welsch ME, Snyder SA, Stockwell BR. 2010. Privileged scaffolds for library design and drug discovery. *Current Opinion in Chemical Biology* 14: 347–361.
- Wilson SA, Roberts SC. 2012. Recent advances towards development and commercialization of plant cell culture processes for the synthesis of biomolecules. *Plant Biotechnology Journal* 10: 249–268.
- Wu S, Jiang Z, Kempinski C, Eric Nybo S, Husodo S, Williams R, Chappell J. 2012. Engineering triterpene metabolism in tobacco. *Planta* 236: 867–877.
- Wu S, Schalk M, Clark A, Miles RB, Coates R, Chappell J. 2006. Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nature Biotechnology* 24: 1441–1447.
- Xiong J, Kashiwada Y, Chen CH, Qian K, Morris-Natschke SL, Lee KH, Takaishi Y. 2010. Conjugates of betulin derivatives with AZT as potent anti-HIV agents. *Bioorganic and Medicinal Chemistry* 18: 6451–6469.
- Xue Z, Duan L, Liu D, Guo J, Ge S, Dicks J, Ó Máille P, Osbourn A, Qi X. 2012. Divergent evolution of oxidosqualene cyclases in plants. *New Phytologist* 193: 1022–1038.
- Yoshikuni Y, Dietrich JA, Nowroozi FF, Babbitt PC, Keasling JD. 2008. Redesigning enzymes based on adaptive evolution for optimal function in synthetic metabolic pathways. *Chemistry & Biology* 15: 607–618.
- Yoshikuni Y, Ferrin TE, Keasling JD. 2006. Designed divergent evolution of enzyme function. *Nature* 440: 1078–1082.
- Yuan LZ, Rouvière PE, Larossa RA, Suh W. 2006. Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metabolic Engineering* 8: 79–90.
- Zhang H, Boghigian BA, Armando J, Pfeifer BA. 2011. Methods and options for the heterologous production of complex natural products. *Natural Product Reports* 28: 125–151.
- Zhang X, Samadi AK, Roby KF, Timmermann B, Cohen MS. 2012. Inhibition of cell growth and induction of apoptosis in ovarian carcinoma cell lines CaOV3 and SKOV3 by natural withanolide Withaferin A. *Gynecologic Oncology* 124: 606–612.
- Zhao J, Davis LC, Verpoorte R. 2005. Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnology Advances* 23: 283–333.
- Zhou J, Huang L, Hachey DL, Chen CH, Aiken C. 2005. Inhibition of HIV-1 maturation via drug association with the viral Gag protein in immature HIV-1 particles. *Journal of Biological Chemistry* 280: 42149–42155.
- Zhu Z, Qian Z, Yan Z, Zhao C, Wang H, Ying G. 2013. A phase I pharmacokinetic study of ursolic acid nanoliposomes in healthy volunteers and patients with advanced solid tumors. *International Journal of Nanomedicine* 8: 129–136.