Plant-derived terpenes: A feedstock for specialty biofuels

Ritesh Mewalal[1], Durgesh K. Rai[2], David Kainer[3], Feng Chen[4], Carsten Külheim[3], Gary F. Peter[5] and Gerald A. Tuskan[1*]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, USA

[2]Biology and Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge Tennessee, 37831, USA

[3]Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia

[4]Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA,

[5]School of Forest Resources and Conservation, University of Florida, Gainesville, Florida, 32611, USA

*Correspondence: tuskanga@ornl.gov (G.A. Tuskan).

Keywords: specialty biofuels; *Eucalyptus*; plant secondary metabolites; systems biology; synthetic biology; terpenes

Abstract

Research toward renewable and sustainable energy has identified specific terpenes capable of supplementing or replacing current petroleum-derived fuels. Despite being naturally produced and stored by many plants, there are few examples of commercial recovery of terpenes from plants due to low yields. Plant terpene biosynthesis is regulated at multiple levels, leading to wide variability in terpene content and chemistry. Advances in the plant molecular toolkit, including annotated genomes, high-throughput omics profiling and genome editing, have begun to elucidate plant terpene metabolism, and such information is useful for bioengineering metabolic pathways for specific terpenes. Here, we review the status of terpenes as a specialty biofuel and discuss the potential of plants as a viable agronomic solution for future terpene-derived biofuels.

Energy-rich terpenes as specialty biofuels

In plants, terpenes (or terpenoids, isoprenoids) are a naturally occurring, chemically diverse set of metabolites associated with developmental physiology as well as mutualistic and antagonistic plant-herbivore and plant-environment interactions (**Box 1**) [1]. Terpenes, together with aromatic compounds, constitute the essential oils of plants, with the highest concentration usually found in the specialized storage cavities of leaves. Terpenes are hydrocarbons that are classified by the number of isoprene units (C_5H_8), with the most common being mono-, sesqui-, and di-terpenes (C_{10}, C_{15}, and C_{20}, respectively). Commercially, terpenes have industrial uses as agrichemicals, fragrances, nutraceuticals and pharmaceuticals [2]. From a growing repertoire of 40,000+ reported structures, specific terpenes have been identified as **specialty biofuels** (see Glossary) meeting current industrial and chemical requirements, including viscosities, flash points, and freezing points, high energy densities and high volumetric net heats of combustion (NHOC) [3, 4]. Table 1 provides a summary of specific biofuel-related terpenes identified.

Terpenes occur in both cyclic and acyclic forms; however, the cyclic forms have higher energy density and NHOC, and are therefore preferred as feedstocks for biofuels [3]. Despite this, high-density fuels have also been synthesized from acyclic linalool and farnesene via ring closure metathesis [4, 5]. As a biofuel, terpenes can be used directly or blended with existing jet fuel (e.g. Jet-A, JP-5 and JP-8), missile propellant (e.g. JP-10), gasoline or diesel fuels [6-9]. For example, hydrogenated monoterpenes myrcene and limonene were shown to be suitable additives to diesel [9], and β-pinene has been used to synthesize high-density biofuels comparable to JP-10 (although, the viscosity and freezing point were higher) [10]. In 2012, Meylemans *et al.* demonstrated that catalytic dimerization of α-pinene, camphene, limonene and crude turpentine (dominated by α-pinene) produces high-density biofuels comparable to JP-10 [6]. Importantly, in 2013, Meylemans *et al.* showed that blending terpene dimers with jet fuel could improve the NHOC and energy density while simultaneously avoiding the high viscosities associated with C_{20} molecules [7]. Also in 2013, Hellier *et al*. assessed the performance of twelve different terpenes in compression and spark ignition engines. The authors found that terpenes could serve as standalone fuels or in blends up to 65% (w/w) in diesel or gasoline engines [8]. In 2014, Harvey *et al.* showed that high density fuels could be

generated from hydrogenated valencene, premnaspirodiene and caryophyllene using a heterogeneous acid catalyst, Nafion SAC-13. These fuels could serve as diesel fuels or additives to jet fuels [11]. Traditionally, oxygenated terpenoids have been disregarded as a fuel source; however, using heterogeneous acid catalysts, Meylemans *et al*. (2014) demonstrated the conversion of oxygenated 1,4-cineole and 1,8-cineole to high-density hydrocarbons, which function as blend-in diesel fuel additives [12]. Several patents have been issued or are pending, covering the efficient conversion and use of terpenes as biofuels (U.S. Patent Number: US8227651B1, US9327279B2, US8975463B1 and US2015/0011810A1) [4, 13-15].

Plant-derived terpenes represent an alternate sustainable source of energy, potentially alleviating fossil fuel dependences and associated effects on atmospheric carbon dioxide and climate change [16]. However, naturally derived terpenes have not been commercially implemented as a biofuel due to the low yield of specific terpenes in plants. Over the last decade, plant research has experienced considerable growth in omics technologies (*i.e*., DNA sequencing, RNA sequencing, metabolic profiling and proteomics) and improvements in genome modification/editing tools such as CRISPR-Cas9 [17]. These developments, combined with the inherent ability of many plants to synthesize and store terpenes, provide a powerful platform for the commercial-scale production of terpenes for future biofuels. In this review, we assess synthetic routes towards producing these compounds and highlight plant-based systems as a viable source of terpenes. We further discuss approaches to understand the terpene biosynthetic machinery, and hypothesize future commercial recovery strategies. As an example, we outline the potential of an agronomic system based on coppiced *Eucalyptus* plantations as a source for terpene production.

Engineering terpene biosynthesis in non-plant systems

Thus far, efforts to produce terpenes as a specialty biofuel have been limited to microorganisms. The genetic tractability of microbial systems have facilitated significant progress in reconstruction of metabolic pathways within these organisms for several biofuels including terpenes [18]. In 2011, Peralta-Yahya *et al*. identified bisabolane as a suitable alternative to D2 diesel and showed that its precursor, bisabolene, could be

produced at high levels in both *Escherichia coli* (> 0.90 g l^{-1}) and *Saccharomyces cerevisiae* (> 0.90 g l^{-1}) by recombinant expression of the *Abies grandis* (E)-α-bisabolene synthase gene. Bisabolene was then converted to bisabolane via chemical hydrogenation [19]. In 2015, Kirby *et al*. identified a novel route within *E. coli* to synthesize the building blocks of terpenes, isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), using the pentose sugar, ribulose 5-phosphate (Ru5P) instead of pyruvate and glyceraldehyde 3-phosphate. Further co-expression with *A. grandis* bisabolene synthase resulted in an improved metabolic flux toward the synthesis of bisabolene and a yield of > 2.0 mg g^{-1} dry cell weight [20]. Zhang *et al.* (2014) demonstrated through **metabolic engineering** (see Glossary) of the terpene mevalonate (MVA) precursor pathway in *E. coli*, in combination with optimization of growth and culturing conditions, that sabinene could be synthesized at a concentration of 2.65 g l^{-1} [21]. Two recent reports demonstrated that the plasticity of the terpene biosynthetic pathway could be exploited for production of β-caryophyllene at high levels (1.05 g l^{-1} and 1.52 g l^{-1}) in engineered *E. coli* [22, 23]. With the application of **systems biology** (see Glossary) and **synthetic biology** (see Glossary), metabolic engineering of biofuels in microbial systems has further potential for commercialization including lower cost and high product yield [24, 25].

While microbial expression platforms are promising, they are not without shortcomings, including cellular toxicity associated with terpenes, accumulated toxic intermediates within the pathway, uncontrolled volatilization of terpenes from the culture media, and low titers [8, 18, 26-28]. For example, in 2013, Hellier *et al*. showed that even at low concentrations, geraniol and geranial can be toxic to the heterologous expression system [8]. Enzymes from the terpene biosynthetic pathway also require post-translational modifications (PTM) for functionality [29], which are absent in many prokaryotic expression systems and culminates in insoluble protein aggregates. Eukaryotic expression systems (e.g., mammalian cell expression), may address this particular disadvantage; however, the cost of batch fermentation can be high, especially for growth media and culturing conditions. Moreover, overexpression of foreign gene(s) can lead to imbalance of the host metabolic pathways in the case of enzymes, and exhaust host resources such as energy and amino acids, which is associated with inhibition of normal growth and stress responses [18]. Microbial expression platforms also do not have

storage systems innately found in several plants which may explain the toxicity associated with accumulation of terpenes within microbial cells. Finally, chemical synthesis may not provide a feasible route for low cost and high yield of terpenes due to the complexity of the chemical structures [30] and subsequent volatility.

Plant terpenes as a source for specialty biofuels

Considering that plants naturally produce and store thousands of terpenes, including those with a potential as biofuels (Table 1), the question remains: why have plants not been leveraged as a viable source of these compounds? Foliar tissue of many plants such as *Eucalyptus*, have large natural variation in **terpene content** (see Glossary) and composition [31]. For example, *Eucalyptus polybractea* showed approximately 20-fold natural variation in foliar terpene content (0.7-13.0% DW^{-1} in 170 seedlings), dominated by monoterpenes, specifically 1,8-cineole and smaller amounts of spathulenol (3.9%), a-pinene (2.3%), limonene (2.6%), α-terpinyl acetate (1.6%), β-pinene (1.6%), myrcene (1.2%), sabinene (1%), α-terpineol (1%), caryophellene oxide (1%) [32]. Gene and pathway perturbation for terpenes in several plants have resulted in anticipated changes in terpene content [33]. Lange *et al*. (2011) showed by candidate gene perturbation, consistent increase in monoterpene content in transgenic peppermint (>78%), grown in greenhouse and commercial-scale field trials [34]. The ability to synthesize terpenes suggest an appropriate physio-biochemical environment that facilitates the correct enzyme folding, subcellular targeting and PTM to synthesize the correct atomic configurations of the terpenes [29, 35]. Phytotoxicity associated with terpenes within the plant cellular environment is avoided by secretory cells which synthesize these metabolites and transport them into specialized storage cavities including resin ducts in gymnosperms, oil glands in *Eucalyptus* and glandular trichomes in *Mentha*. The availability of several plant genomic resources including Phytozome v11 (http://phytozome.jgi.doe.gov/pz/portal.html) and Plaza (http://bioinformatics.psb.ugent.be/plaza/) will facilitate comparative genomics to adopt knowledge learned from model plants to dedicated biomass crops such as *Eucalyptus*. These crops have marginal land requirements with little to no overlap with land for food crops and have an all-year-round harvesting potential in the case of short rotation coppice *Eucalyptus* [36]. The commercial-scale production of terpenes from these crops is

potentially carbon-neutral since plants sequester large amounts of above and belowground carbon thereby mitigating greenhouse gas emissions [37].

In the past, commercial-scale production of specific terpenes in plants was overlooked due to the relatively low yields from mature plants grown under non-managed, natural systems. Today, however, commercial scale recovery of plant terpenes occurs from pines, eucalypts, mints, and citrus. In the U.S., pine terpenes and fatty acids present in the wood are recovered as co-products during pulping [38]. The global industry annually recovers about 3 million tonnes. These hydrocarbon are fractionated into mono/diterpenoids and fatty acids which are used as renewable chemical feedstocks for a large number of products which compete on price and quality with petroleum-derived compounds. In the first 10 years of growth, planted pine trees with an average of 4% wood terpene content produce ~40 GJ ha^{-1} y^{-1} of energy in hydrocarbons in the woody stem (G. Peter, personal communication). In contrast to pines and citrus, eucalypt terpenes are extracted from leaves in short-rotation coppice agroforestry systems. This offers the potential to increase leaf biomass production per unit land area per unit time [39]. Furthermore, the availability of the *Eucalyptus* genome [40] will facilitate the commercial deployment of genetically improved plant with customized terpene production and maximized oil gland capacity [41, 42]. In recent year's significant research into the production of biofuels, including Jet-A1, from lignin biomass via pyrolysis methods has occurred. Eucalypts, grown and harvested for biomass, have the potential to provide biofuel with low ecological impact, but profitability may be questionable in the short to medium term [43]. The addition of a terpene extraction stage to the pipeline may improve the sustainability of such systems by maximizing the volume and variety of fuel obtained from a single harvest [44]. Engineering eucalypts for maximal production of both lignin and terpene presents a challenge as maximizing one may penalize the other, but an optimal balance may result in greater overall production of high value fuels per kg of harvested dry weight.

Given that 1) we have good knowledge about the quantitative and qualitative variation of terpenes in eucalypts [45, 46] and 2) a commercial coppice system for pharmaceutical terpene production exists [47], we are able to estimate commercial-scale production of biofuels from specific terpenes. At an oil content of 48 mg g^{-1} fresh weight of leaves and

94% 1,8-cineole, Iqbal *et al*. suggests *E. polybractea* is capable of producing commercial scale quantities of *Eucalyptus* oils which consist of terpenes and aromatics [48]. Wu *et al*. estimates that mallee eucalypts grown under coppice rotations reach an energy productivity of 206 GJ $ha^{-1}y^{-1}$, 35.5% of which of is from leaves for a production period [49]. In 2007, Goodger *et al*. measured the oil yield produced from *E. polybractea* during a 12 month coppicing rotation, comparing 20 randomly selected saplings out of 1000, sourced from seeds of 30 open pollinated maternal lines. The individual with the highest oil yield produced 137 g of total oil from 715.6 g of leaf (dry weight, DW; 19.1% oil concentration DW^{-1}). At the recommended planting density of 5000 plants ha^{-1} [50], clones of this elite individual have the potential to produce 686 kg of essential oil ha^{-1} y^{-1} [51].

Crude steam-distilled oil from pine stems is made up mostly of α-pinene (75-85%) with smaller amounts of β-pinene (0-3%), camphene (4-15%) and limonene (5-15%), with traces of terpinolene and 3-carene [52]. The refining steps of these terpenes include hydrotreatment and fractionation to produce biodiesel. The composition of crude oil extracted by steam-distillation from eucalypt leaves is highly variable both between and within species [45]. For example, current populations of *E. polybractea* contain between 8-98% 1,8-cineole (Kainer *et al*. submitted). Other eucalypt species, however, are dominated by α-, or β-pinene, which can be used directly (**Table 2**). With sufficient selection or genetic engineering, oils from eucalypts can contain over 95% of usable terpenes, which may be separated by fractionation distillation and dimerized with high efficiency. At 500 kg ha^{-1} y^{-1} useable terpene production, 20 million ha (current global area planted with eucalypts) would produce 10 million tons of high energy jet fuel. According to the US States Energy Information Administration (http://www.eia.gov/), daily global jet fuel consumption in 2012 was 5.381 million barrels - an annual consumption of 181 million tons. Therefore 20 million ha could provide over 5% of the fuel as additive, which is less than current production of bioethanol (225 million tons annually [53].

Challenges in understanding the plant terpene biosynthetic machinery

Despite the potential of economic and environmental benefits of terpenes derived from plants, key biological challenges must be overcome. Not all terpenes are represented in all plant lineages, and individuals of the same species may vary in terpene content [54].

There have also been reports of variations in terpene content in response to biotic stress [55], seasonal changes [56] and endogenous hormone levels [57]. Interestingly, terpene diversity is not paralleled by an equivalent number of encoded enzymes. In fact, a single terpene synthase (TPS) may catalyzes the synthesis of multiple terpenes [42, 58]. Perturbation of terpene-related genes have been accompanied by inadvertent pleiotropic changes in plant development and terpene content (see Lange *et al.* (2013) for review). Phillips *et al*., showed that different TPS homologs, (−)-α-pinene synthase and (+)-α-pinene synthase, are required to synthesize enantiomers of α-pinene [59], while Peralta-Yahya *et al.* (2011) demonstrated that orthologous TPS vary in their activity to synthesize specific terpenes [19]. In 2015, Irmisch *et al*. reported that a single amino acid difference in the active site of *Populus trichocarpa* PtTPS19 and PtTPS20 produced either ent-kaurene or 16α-hydroxy-ent-kaurane in *E. coli*, highlighting the importance of enzyme specificity at the level of gene variants [60]. It is evident from the aforementioned studies that terpene biosynthesis is a dynamic process with a complex underlying genetic circuitry, and before the economic and environmental benefit can be recognized, a comprehensive understanding of the biological process is necessary.

A systems-level understanding of terpene biosynthesis

In plants, genetic and environmental factors regulate terpene biosynthesis at multiple levels, including transcription, protein accumulation, and post-transcriptional/translation regulation [29, 61]. Understanding the molecular underpinnings of terpene metabolism in plants will benefit from a systems biology approach, where the goal is to quantitatively describe the cellular processes through global modeling of interactions and dynamics of the molecular components. Systems biology integrates multiple omics datasets through mathematical and probabilistic modeling including correlation networks, graphs and statistical models, to understand the emergent phenotype of a cell or organism [62, 63]. While transcriptomic and proteomic studies have catalogued many genes putatively involved in terpene biosynthesis [64-67]; these studies in isolation provide only a snapshot of this biological process and does not reflect *in vivo* conditions. For example, transcriptional profiles of putative homologs of *E. grandis* MVA and MEP pathway (**Figure 1**) revealed multiple paralogous genes from the MEP pathway exist, and are co-expressed across tissues [68]. This will complicate metabolic engineering due to genetic

redundancy, and possible unintended pleiotropic effects e.g., Pasoreck *et al.* showed broad off-target transcriptomic and metabolomic changes caused by perturbation of squalene biosynthetic genes [69].

Systems biology network outputs are correlations and do not imply causation. However, the approach does allow for the identification of gene modules, *i.e.*, groups of highly correlated gene expression patterns, to refine a candidate gene list. Furthermore, positive and negative correlations within networks, across developmental stages, different plant organs (such as whole leaf and oil glands), or from genetically modified plants in comparison to wildtype, will yield predictive models of pathway regulation and flux [70, 71]. Otero *et al*. showed in *S. cerevisiae* that systems biology could be used to rationally select a suite of genes to perturb in order to redirect carbon flux for improved succinate yield [72]. Because metabolic engineering for specific terpenes will have to be tailored and given the gaps in our understanding of this biological process **(Figure 2),** application of systems biology will be practical to catalog the genetic determinants. These determinants include: preferentially expressed enzyme variants, transporters for precursors and metabolites, signaling and regulators proteins, regulatory RNAs, etc. Furthermore, depending on the initial experimental design, network analysis can reveal gene-gene relationships, feedback loops, pathway flux, spatial-temporal regulation and relationships to different biochemical pathways (which may circumvent unwanted pleiotropy during gene perturbation) to predictively maximize terpene biosynthesis in metabolic engineering strategies.

The system-level understanding of terpene biosynthesis may be extended beyond the traditional molecular dissection to include physical characterization of storage cavities and terpene-terpene dynamics. Storage cavities contain a complex mixture of nonvolatile terpenes, volatile terpenes, and other secondary metabolites such as phloroglucinol compounds. The phase behavior of the entrapped volatile and nonvolatile terpenes is poorly understood yet may have important implications for terpene-terpene interaction and stability, relating the structure of oil glands to their function and understanding the transport of terpenes into the cavity. Scanning electron microscopy revealed modified epidermal cells surrounding the storage cavities [73], while confocal fluorescence

microscopy has revealed spatial organization of cavity metabolites with the nonvolatile components forming a layer between secretary cells living in the lumen and the volatile terpenes [74]. Based on SEM and transmission electron microscopy analysis, it has been proposed that the volatile terpenes in the cavities pass through modified epidermal cells, which cap the oil glands allowing a gradual loss of essential oils from the oil glands, when exposed to vacuum conditions [75]. Neutron scattering techniques, such as small-angle neutron scattering and ultra-small angle neutron scattering are emerging complementary techniques that can be used to characterize the structural properties of oil glands. The highly-penetrating neutrons and the ability to selectively highlight specific regions in a macromolecular complex by deuterium/hydrogen atom exchange, is advantageous for studying complex biological systems. These technologies are able to resolve structural details on scales from ~1 nm to 100 μm in size [76-78] and have the potential to provide insight into the internal structure of oil glands.

Commercialization potential of terpenes from plants

To achieve terpene-based fuel production at commercially competitive scales, gains in terpene yield per hectare could be made by producing elite lines with consistently high foliar terpene content, establishing silviculture techniques and decreasing terpene loss during harvesting and extraction. A crucial step is to identify species with high foliar content of specific terpenes that grow well on marginal lands with minimal impact on food production and native biodiversity. Padovan *et al*. (2014) reviewed the dominant terpenes from 1393 species of Myrtaceae, including 648 species of eucalypts (genera of *Eucalyptus*, *Angophora* and *Corymbia*) [45]. Of these eucalypts, fifteen species have oil profiles dominated by α-pinene, β-pinene, or *p*-cymene (Table 2), most have good ability to re-sprout after coppicing and are able to grow in the semi-arid sub-tropic climate zones.

Having a systems-level understanding of terpene biosynthesis forms an umbrella over two foreseeable translational strategies for the development of elite lines *i.e.* synthetic biology and/or traditional **selective breeding** (see Glossary) strategies. Synthetic biology in plants will integrate predictive modeling from systems biology with a high-precision genome editing toolkit, including assembly of multigene constructs for gene stacking [79], high-throughput genome modification (CRISPR-Cas9) [9] and spatial-temporal regulation

[80] to design metabolic pathways and fluxes for optimized production of specific terpenes. A number of genetic tools are available or being optimized for plant modification which will serve as a foundation for biotechnological engineering applications (refer to Baltes and Voytas (2015) for a comprehensive review). Moreover, studies have reported successful *Agrobacterium*-mediated genetic transformation of many plants including *Eucalyptus* species [41, 82]. In addition to affecting terpene content, synthetic design should also target terpene stability and storage capacity in plants. For example, monoterpenes have been reported to exist in nonvolatile glycosylated forms, therefore identifying specific glycosyltransferases catalyzing such sugar conjugations to increase terpene stability and water solubility will be important for future biotechnology [83, 84]. Utility genetic promoter sequences, specific to the storage organs, such as those reported by Kortbeek *et al*., will be useful to prevent undesired pleiotropy or phytotoxicity associated with terpenes in non-native cells [85].

Eucalyptus oil yield is a complex quantitative trait that is influenced by foliar oil concentration, oil composition, leafy biomass accumulation and leaf morphology [86]. Terpene yield, in particular, displays large variance and high heritability, hence, selective breeding for these traits is likely to generate significant gains. For example, in related tea tree, three cycles of phenotypic selection since 1993 resulted in doubled oil yield, from 148 to 300 kg ha^{-1} [87]. This process could be accelerated in un-domesticated eucalypt species through molecular breeding techniques such as genomic selection [61]. Previous studies reported that the terpene content is limited by the storage gland capacity, which in turn is regulated by the leaf area and thickness [46]. Furthermore, Goodger and Woodrow showed that oil concentration in *E. polybractea* is tightly correlated with the total volume of secretory oil cavities (r^2=0.96), which in turn is constrained by leaf size and metabolic support requirements [88]. Therefore, improving other traits such as leafy biomass accumulation, crown form, post-coppice regrowth rate and even herbivory resistance, may provide additional paths to higher overall yield (Kainer et al, (2016) submitted). *E. polybractea* is typically harvested on a 2 year rotation and may do so for up to 100 years, but shorter rotation times (e.g. 1 year) result in diminishing returns as the trees are unable to fully recover lignotuber mass between harvests [89]. The

development of viable short-rotation genotypes has the potential for very large annualized yield gains in short to medium term.

Intensive silvicultural management will also play an important role in maintaining productivity from elite plant lines. **Silviculture** (see Glossary) considerations include plantation sites, planting density, crop rotation (harvesting), coppicing ability, fertilizer requirements and pest management. For terpene production in the foliage of *Eucalyptus*, shorter rotations will be advantageous and as such, densities can be in the range of ~3000-5000 trees ha^{-1} to maximize per ha productivities [50, 90]. Silviculture techniques will have to be established, if they have not been already, for specific plant genera. For example, plantation sites will be dictated by cold tolerance of the crop species. Currently, ~50 000 ha of *Eucalyptus* are planted in California, Hawaii and Florida [90], however, expanding this distribution, particularly into the south of the US, will require frost-tolerant species [91].

Finally, with current short-rotation coppice practices, a proportion of the essential oil present in *Eucalyptus* foliage is lost during harvesting and up to 25% is left un-extracted due to steam distillation inefficiency [92]. Improvements in harvesting and distillation efficiency will be important factors for net energy gain. Currently, studies are ongoing and have demonstrated improvements in extraction of specific terpenes from complex mixtures in plants [93, 94].

Concluding remarks

Plants are attractive systems for the commercial-scale production of specific terpenes due to their inherent ability to synthesize, transport, accumulate and store these compounds. To leverage plant systems for terpenes requires addressing several challenges (**Box 1**). A roadmap for the commercial recovery of terpenes synthesized in plant is proposed in **Figure 3**. Exploiting the existing variation in plant terpene content and using high-throughput omics profiling technology should lead to a fundamental understanding of this biological process and facilitate predictive engineering of plants for specific terpenes using synthetic biology (single gene effects or additive effects of gene stacking to rewire the terpene metabolic pathway), selective breeding approaches, and

their combination. Together with well-established silviculture and harvesting practices, *in planta*-synthesized terpenes hold tremendous potential as a sustainable specialty biofuel.

Acknowledgements

This review is based on work supported by the Department of Energy (DOE) and funding from the LDRD project LOIS ID: 7428. The contents of this review are solely the responsibility of the authors and do not necessarily represent the official views of the DOE. The authors acknowledge Lee E. Gunter, Aparna Annamraju, Sai Venkatesh Pingali, Hugh M. O'Neill, Daniel Jacobson and Timothy J. Tschaplinski for their critical review of the manuscript. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US DOE under Contract Number DE–AC05–00OR22725.

References

1 Gershenzon, J. and Dudareva, N. (2007) The function of terpene natural products in the natural world. *Nat. Chem. Biol.* 3, 408-414

2 Kirby, J. and Keasling, J.D. (2009) Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annu. Rev. Plant Biol.* 60, 335-355

3 Filley, J.*, et al.* (2001) Energetics of the 2+ 2 cyclization of limonene. *J. Photochem. Photobiol. A: Chem.* 139, 17-21

4 Harvey, B.G. United States of America as Represented by the Secretary of the navy, Arlington, VA (US). High density cyclic fuels derived from linear sesquiterpenes, US2015/0011810A1

5 Meylemans, H.A.*, et al.* (2011) Solvent-free conversion of linalool to methylcyclopentadiene dimers: A route to renewable high-density fuels. *ChemSusChem* 4, 465-469

6 Meylemans, H.A.*, et al.* (2012) Efficient conversion of pure and mixed terpene feedstocks to high density fuels. *Fuel* 97, 560-568

7 Meylemans, H.A.*, et al.* (2013) Low-temperature properties of renewable high-density fuel blends. *Energy Fuels* 27, 883-888

8 Hellier, P.*, et al.* (2013) Combustion and emissions characterization of terpenes with a view to their biological production in cyanobacteria. *Fuel* 111, 670-688

9 Tracy, N.I., *et al.* (2009) Hydrogenated monoterpenes as diesel fuel additives. *Fuel* 88, 2238-2240

10 Harvey, B.G., *et al.* (2009) High-density renewable fuels based on the selective dimerization of pinenes. *Energy Fuels* 24, 267-273

11 Harvey, B.G., *et al.* (2014) High-density biosynthetic fuels: the intersection of heterogeneous catalysis and metabolic engineering. *Phys. Chem. Chem. Phys.* 16, 9448-9457

12 Meylemans, H.A., *et al.* (2014) Low-temperature, solvent-free dehydration of cineoles with heterogeneous acid catalysts for the production of high-density biofuels. *J. Chem. Technol. Biotechnol.* 89, 957-962

13 Harvey, B.G., *et al.* The United States of America as Represented by the Secretary of the Navy, Washington, DC (US). High density renewable fuels based on the selective dimerization of pinenes, US8227651B1

14 Harvey, B.G., *et al.* The United States of America as Represented by the Secretary of the Navy, Washington, DC (US). Methods for the production of renewable dimethyl JP10, US9327279B2

15 Harvey, B.G., *et al.* The United States of America as Represented by the Secretary Navy, Washington, DC (US). Efficient conversion of pure and mixed terpene feedstocks to high density fuels, US8975463B1

16 Solomon, S., *et al.* (2009) Irreversible climate change due to carbon dioxide emissions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1704–1709

17 Liu, D., *et al.* (2016) Advances and perspectives on the use of CRISPR/Cas9 systems in plant genomics research. *Curr. Opin. Plant Biol.* 30, 70-77

18 Peralta-Yahya, P.P. and Keasling, J.D. (2010) Advanced biofuel production in microbes. *Biotechnol. J* 5, 147-162

19 Peralta-Yahya, P.P., *et al.* (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nat. Commun.* 2, 483

20 Kirby, J., *et al.* (2014) Enhancing terpene yield from sugars via novel routes to 1-deoxy-D-xylulose 5-phosphate. *Appl. Environ. Microbiol.* 81, 130-138

21 Zhang, H., *et al.* (2014) Microbial production of sabinene—a new terpene-based precursor of advanced biofuel. *Microb Cell Fact* 13, 3826-3831

22 Yang, J.*, et al.* (2016) Biosynthesis of β-caryophyllene, a novel terpene-based high-density biofuel precursor, using engineered *Escherichia coli*. *Renewable Energy* 99, 216-223

23 Yang, J. and Nie, Q. (2016) Engineering *Escherichia coli* to convert acetic acid to β-caryophyllene. *Microb Cell Fact* 15, 1

24 Peralta-Yahya, P.P.*, et al.* (2012) Microbial engineering for the production of advanced biofuels. *Nature* 488, 320-328

25 Erickson, B. and Winters, P. (2012) Perspective on opportunities in industrial biotechnology in renewable chemicals. *Biotechnol. J* 7, 176-185

26 Dunlop, M.J. (2011) Engineering microbes for tolerance to next-generation biofuels. *Biotechnology Biofuels* 4, 32

27 Arendt, P.*, et al.* (2016) Synthetic biology for production of natural and new-to-nature terpenoids in photosynthetic organisms. *Plant J.* DOI: 10.1111/tpj.13138

28 Pitera, D.J.*, et al.* (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metab. Eng.* 9, 193-207

29 Hemmerlin, A. (2013) Post-translational events and modifications regulating plant enzymes involved in isoprenoid precursor biosynthesis. *Plant Sci.* 203, 41-54

30 Chang, M.C. and Keasling, J.D. (2006) Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* 2, 674-681

31 Goodger, J.Q.*, et al.* (2010) Isolation of intact sub-dermal secretory cavities from *Eucalyptus*. *Plant Methods* 6, 20

32 King, D.J.*, et al.* (2006) The accumulation of terpenoid oils does not incur a growth cost in *Eucalyptus polybractea* seedlings. *Funct. Plant Biol.* 33, 497-505

33 Lange, B.M. and Ahkami, A. (2013) Metabolic engineering of plant monoterpenes, sesquiterpenes and diterpenes—current status and future opportunities. *Plant Biotechnol. J.* 11, 169-196

34 Lange, B.M.*, et al.* (2011) Improving peppermint essential oil yield and composition by metabolic engineering. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16944-16949

35 Davies, F.K.*, et al.* (2014) Toward a photosynthetic microbial platform for terpenoid engineering. *Photosynthesis Res.* 123, 265-284

36 Hinchee, M.*, et al.* (2011) Short-rotation woody crops for bioenergy and biofuels applications. In *Biofuels*, pp. 139-156, Springer
37 Jansson, C.*, et al.* (2010) Phytosequestration: carbon biosequestration by plants and the prospects of genetic engineering. *Bioscience* 60, 685-696
38 Johansson, A. (1982) By-product recovery and valorisation in the kraft industry: A review of current trends in the recovery and use of turpentine and tall oil derivatives. *Biomass* 2, 103-113
39 Tuskan, G. (1998) Short-rotation woody crop supply systems in the United States: what do we know and what do we need to know? *Biomass and Bioenerg* 14, 307-315
40 Myburg, A.A.*, et al.* (2014) The genome of *Eucalyptus grandis*. *Nature* 510, 356-362
41 Fernando, S.C.*, et al.* (2016) Plant regeneration through indirect organogenesis and genetic transformation of *Eucalyptus polybractea* RT Baker. *Industrial Crops and Products* 86, 73-78
42 Külheim, C.*, et al.* (2015) The *Eucalyptus* terpene synthase gene family. *BMC Genomics* 16, 450
43 Goss, K.*, et al.* (2014) *Sustainable mallee jet fuel: Sustainability and life cycle assessment for supply to Perth airport, western Australia*. Future Farm Industries CRC
44 Joyce, B. and Stewart, C. (2012) Designing the perfect plant feedstock for biofuel production: using the whole buffalo to diversify fuels and products. *Biotechnol. Adv.* 30, 1011-1022
45 Padovan, A.*, et al.* (2014) The evolution of foliar terpene diversity in Myrtaceae. *Phytochem. Rev.* 13, 695-716
46 King, D.J.*, et al.* (2006) Regulation of oil accumulation in single glands of *Eucalyptus polybractea*. *New Phytol.* 172, 440-451
47 Goodger, J.Q. and Woodrow, I.E. (2010) The influence of micropropagation on growth and coppicing ability of *Eucalyptus polybractea*. *Tree Physiol* 30, 285-296
48 Iqbal, Z.*, et al.* (2011) Variation in composition and yield of foliage oil of *Eucalyptus polybractea*. *J. Chem. Soc. Pak* 33, 183-187
49 Wu, H.*, et al.* (2008) Production of mallee biomass in western Australia: Energy balance analysis. *Energy Fuels* 22, 190-198

50 Milthorpe, P.*, et al.* (1998) Optimum planting densities for the production of eucalyptus oil from blue mallee (*Eucalyptus polybractea*) and oil mallee (*E. kochii*). *Ind Crops Prod* 8, 219-227

51 Goodger, J.Q.*, et al.* (2007) Examination of the consistency of plant traits driving oil yield and quality in short-rotation coppice cultivation of *Eucalyptus polybractea*. *For. Ecol. Manage.* 250, 196-205

52 Gscheidmeier, M. and Fleig, H. (2000) *Turpentines*. Wiley-VCH Verlag GmbH & Co. KGaA

53 Yuan, J.S.*, et al.* (2008) Plants to power: Bioenergy to fuel the future. *Trends in Plant Sci* 13, 421-429

54 Andrew, R.L.*, et al.* (2013) Intensive sampling identifies previously unknown chemotypes, population divergence and biosynthetic connections among terpenoids in *Eucalyptus tricarpa*. *Phytochemistry* 94, 148-158

55 Oates, C.N.*, et al.* (2015) The transcriptome and terpene profile of *Eucalyptus grandis* reveals mechanisms of defence against the insect pest, *Leptocybe invasa*. *Plant Cell Physiol.* 56, 1418-1428

56 Llusià, J. and Peñuelas, J. (2000) Seasonal patterns of terpene content and emission from seven Mediterranean woody species in field conditions. *Am. J. Bot.* 87, 133-140

57 Zulak, K.G.*, et al.* (2009) Targeted proteomics using selected reaction monitoring reveals the induction of specific terpene synthases in a multi-level study of methyl jasmonate-treated Norway spruce (*Picea abies*). *Plant J.* 60, 1015-1030

58 Jones, C.G.*, et al.* (2008) Isolation of cDNAs and functional characterisation of two multi-product terpene synthase enzymes from sandalwood, *Santalum album* L. *Arch. Biochem. Biophys.* 477, 121-130

59 Phillips, M.A.*, et al.* (2003) cDNA isolation, functional expression, and characterization of (+)-α-pinene synthase and (−)-α-pinene synthase from loblolly pine (*Pinus taeda*): Stereocontrol in pinene biosynthesis. *Arch. Biochem. Biophys.* 411, 267-276

60 Irmisch, S.*, et al.* (2015) One amino acid makes the difference: the formation of ent-kaurene and 16α-hydroxy-ent-kaurane by diterpene synthases in poplar. *BMC Plant Biol.* 15, 262

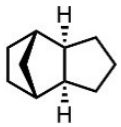
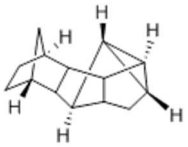
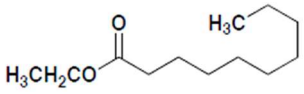
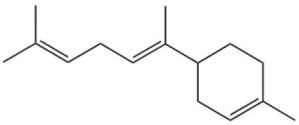
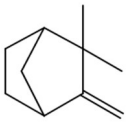
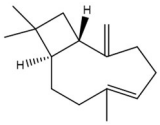
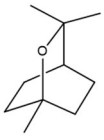
61 Kainer, D.*, et al.* (2015) Genomic approaches to selection in outcrossing perennials: focus on essential oil crops. *Theor. Appl. Genet.* 128, 2351-2365
62 Kitano, H. (2002) Systems biology: a brief overview. *Science* 295, 1662-1664
63 Nielsen, J. and Jewett, M.C. (2008) Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 8, 122-131
64 Champagne, A. and Boutry, M. (2013) Proteomic snapshot of spearmint (*Mentha spicata* L.) leaf trichomes: a genuine terpenoid factory. *Proteomics* 13, 3327-3332
65 Voo, S.S.*, et al.* (2012) Assessing the biosynthetic capabilities of secretory glands in citrus peel. *Plant Physiol.* 159, 81-94
66 Champagne, A.*, et al.* (2012) In-depth proteome mining of cultured *Catharanthus roseus* cells identifies candidate proteins involved in the synthesis and transport of secondary metabolites. *Proteomics* 12, 3536-3547
67 Wang, W.*, et al.* (2009) Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. *BMC Genomics* 10, 465
68 Hefer, C.*, et al.* (2011) The *Eucalyptus* genome integrative explorer (EucGenIE): a resource for *Eucalyptus* genomics and transcriptomics. In *BMC Proceedings*, pp. 1-1, Springer
69 Pasoreck, E.K.*, et al.* (2016) Terpene metabolic engineering via nuclear or chloroplast genomes profoundly and globally impacts off-target pathways through metabolite signalling. *Plant Biotechnol. J.*
70 Mizrachi, E. and Myburg, A.A. (2016) Systems genetics of wood formation. *Curr. Opin. Plant Biol.* 30, 94-100
71 Fukushima, A.*, et al.* (2009) Integrated omics approaches in plant systems biology. *Curr. Opin. Chem. Biol.* 13, 532-538
72 Otero, J.M.*, et al.* (2013) Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory. *PloS One* 8, e54144
73 Kalachanis, D. and Psaras, G. (2005) Structure and development of the secretory cavities of *Myrtus communis* leaves. *Biol. Plant.* 49, 105-110
74 Heskes, A.*, et al.* (2012) Multiphoton fluorescence lifetime imaging shows spatial segregation of secondary metabolites in *Eucalyptus* secretory cavities. *J. Microsc* 247, 33-42

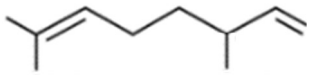
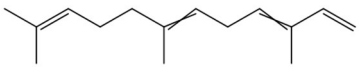
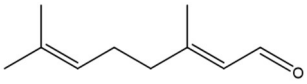
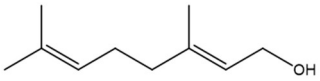
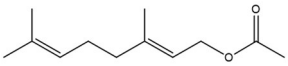
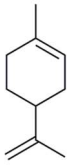
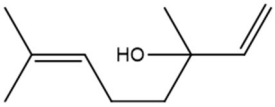
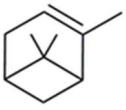
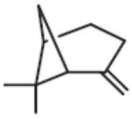
75 List, S.*, et al.* (1995) Functional anatomy of the oil glands of *Melaleuca alternifolia* (Myrtaceae). *Aust. J. Bot.* 43, 629-641
76 Schaefer, D.W. and Agamalian, M.M. (2004) Ultra-small-angle neutron scattering: a new tool for materials research. *Curr. Opin. Solid State Mater. Sci.* 8, 39-47
77 Nishiyama, Y.*, et al.* (2014) Structural coarsening of aspen wood by hydrothermal pretreatment monitored by small-and wide-angle scattering of X-rays and neutrons on oriented specimens. *Cellulose* 21, 1015-1024
78 Pingali, S.V.*, et al.* (2014) Morphological changes in the cellulose and lignin components of biomass occur at different stages during steam pretreatment. *Cellulose* 21, 873-878
79 De Paoli, H.C.*, et al.* (2016) An innovative platform for quick and flexible joining of assorted DNA fragments. *Sci Rep* 6, 19278
80 Yang, F.*, et al.* (2013) Engineering secondary cell wall deposition in plants. *Plant Biotechnol. J.* 11, 325-335
81 Baltes, N.J. and Voytas, D.F. (2015) Enabling plant synthetic biology through genome engineering. *Trends Biotechnol.* 33, 120-131
82 Aggarwal, D.*, et al.* (2011) Agrobacterium tumefaciens mediated genetic transformation of selected elite clone (s) of *Eucalyptus tereticornis*. *Acta Physiol Plant* 33, 1603-1611
83 Bönisch, F.*, et al.* (2014) A UDP-glucose: monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome. *Plant Physiol.* 165, 561-581
84 Luan, F.*, et al.* (2004) Enantioselective analysis of free and glycosidically bound monoterpene polyols in *Vitis vinifera* L. cvs. Morio Muscat and Muscat Ottonel: evidence for an oxidative monoterpene metabolism in grapes. *J. Agric. Food Chem.* 52, 2036-2041
85 Kortbeek, R.*, et al.* (2016) Engineering of tomato glandular trichomes for the production of specialized metabolites. *Methods Enzymol.*
86 Doran, J.*, et al.* (1998) Variation in first-harvest oil production in *Eucalyptus radiata*. *Australian Forestry* 61, 27-33
87 Baker, G.*, et al.* (2014) Highly improved tea tree varieties to maximise profit. *Rural Industries Research and Development Corporation* 13, RIRDC Project No PRJ - 003689

88 Goodger, J.Q. and Woodrow, I.E. (2012) Genetic determinants of oil yield in *Eucalyptus polybractea* RT Baker. *Trees* 26, 1951-1956
89 Wildy, D. and Pate, J. (2002) Quantifying above-and below-ground growth responses of the western Australian oil mallee, *Eucalyptus kochii* subsp. plenissima, to contrasting decapitation regimes. *Ann. Bot.* 90, 185-197
90 Zalesny, R.J.*, et al.* (2011) Woody biomass from short rotation energy crops. *American Chemical Society: Washington, DC* 1067, 27-63
91 Stanturf, J.A.*, et al.* (2013) Eucalyptus beyond its native range: Environmental issues in exotic bioenergy plantations. *International Journal of Forestry Research* 2013, 1-5
92 Denny, E. (2003) *Distillation of Eucalyptus leaf oils: Theory and practice*. CRC Press
93 Jiang, Z.*, et al.* (2016) Extraction and analysis of terpenes/terpenoids. *Curr Protoc Plant Biol.*, 345-358
94 Ormeno, E.*, et al.* (2011) Extracting and trapping biogenic volatile organic compounds stored in plant species. *Trends Anal. Chem.* 30, 978-989
95 Hellier, P.*, et al.* (2013) The importance of double bond position and cis–trans isomerisation in diesel combustion and emissions. *Fuel* 105, 477-489
96 Saad, N.Y.*, et al.* (2013) Major bioactivities and mechanism of action of essential oils and their components. *Flavour Fragrance J.* 28, 269-279
97 Harvey, B.G.*, et al.* (2015) High-density renewable diesel and jet fuels prepared from multicyclic sesquiterpanes and a 1-hexene-derived synthetic paraffinic kerosene. *Energy Fuels* 29, 2431-2436
98 Beller, H.R.*, et al.* (2015) Natural products as biofuels and bio-based chemicals: fatty acids and isoprenoids. *Nat. Prod. Rep.* 32, 1508-1526
99 Peralta-Yahya, P.P.*, et al.* (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nature Communications* 2, 483
100 Coppen, J.J. (2003) *Eucalyptus: the genus Eucalyptus*. CRC Press
101 Brophy, J.J. and Southwell, I.A. (2002) *Eucalyptus* chemistry. *Eucalyptus: the genus Eucalyptus*, 102
102 Elaissi, A.*, et al.* (2007) Contribution to the qualitative and quantitative study of seven *Eucalyptus* species essential oil harvested of Hajeb's Layoun arboreta (Tunisia). *J Essent Oil Bear Pl* 10, 15-25

103 Bignell, C.*, et al.* (1998) Volatile leaf oils of some south-western and southern Australian species of the genus *Eucalyptus*. *Flavour Fragrance J.* 13, 43–47

Table 1. Specific terpenes identified that are a suitable replacement or blend-in for current petroleum-derived fuels and comparison with known fuels. Density values were extracted from the cited literature and is based on the conditions outlined therein.

Compound	Chemical structure	Density (g ml^{-1})	References
JP-10		0.94	Meylemans et al. (2012)
RJ-5		1.08	Meylemans et al. (2012)
Gasoline		0.74	Meylemans et al. (2012)
Biodiesel		0.88	Meylemans et al. (2012)
Bisabolene		0.82	Peralta-Yahya et al. (2011)
Camphene		0.84	Meylemans et al. (2012)
Caryophyllene		0.90	Harvey et al. (2015)
1,8-Cineole		1.01	Saad et al. (2013) & Meylemans et al. (2014)

(-)-β-Citronellene		0.76	Hellier et al. (2013)
Farnesene		0.84-0.88	Hellier et al. (2013)
Geranial		0.89	Hellier et al. (2013)
Geraniol		0.88	Saad et al. (2013)
Geranyl acetate		0.92	Hellier et al. (2013)
Limonene		0.85	Filley et al. (2001)
Linalool		0.87	Hellier et al. (2013)
α-Pinene		0.86	Harvey et al. (2009) & Meylemans et al. (2012)
β-Pinene		0.86	Harvey et al. (2009) & Meylemans et al. (2012)

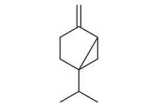
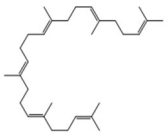
Sabinene		0.84	Beller et al. (2015)
Squalene		0.86	Hellier et al. (2013)

Table 2. List of candidate species for production of high energy terpenes in *Eucalyptus*.

Species	Terpene yield	Geographic range	Climate classification	Dominant terpene	References
E. eremicola	3.1% DM	Cent WA, SA	BWh, BWk, BSh, BSk	β-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. eremophila	1.8-5.2% DM	S WA	BSh, BSk, Csa	α-pinene	Coppen *et al*. (2003), Elaissi *et al*. (2007) & Brophy and Southwell (2002)
E. forrestiana	3.3% DM	S WA	BSk, Csb	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. macrocarpa	2.5% DM	W WA	Csa	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. pimpiniana	1.9% DM	S Cent SA, WA	BWh	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. platypus	0.6-1.2% FW	S WA	BSk, Csb	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. suggrandis	1.1% DM	SW WA	BSk, Csb	α-pinene	Coppen *et al.* (2003) & Bignell *et al.* (1998)
E. striaticalyx	0.7-4.4% DM	Cent WA, SW SA	BWh	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. camphora	1.3-4.0% DM	SE Aus	Cfa, Cfb	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)

E. morrisii	1.7% FW	Cent NSW	BSh, BSk, Cfa	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. ebbanoensis	1.8% DM	W, Cent WA	BWh, BSh, Csa	α-pinene	Coppen *et al.* (2003) & Bignell *et al.* (1998)
E. longicornis	1.2% FW	SW WA	BSk, CSb, Csa, BSh, BWh	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. bosistoana	1-1.5% FW	SE NSW, Vic	Cfb, Cfa	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)
E. calcareana	1.4% DW	S SA	BSk, BWk, Csb, Csa	α-pinene	Coppen *et al*. (2003) & Brophy and Southwell (2002)

Cent, Central; S, South; W, West; S Cent, South Central; SW, South West; SE, South East; E, East; WA, Western Australia; SA, South Australia; Aus, Australia; NSW, New South Wales; Vic, Victoria; DM, dry matter; FW, fresh weight.

Figure 1. Expression profiles of putative functional homologs of the terpene precursor mevalonate (MVA) and methyl-D-erythritol 4-phosphate (MEP) pathway in *Eucalyptus grandis*. Pathway was adapted from Vranová et al. (2013) and putative *Eucalyptus* orthologs are based on the expect-value (*E*-value) of sequence alignment ≤ 1.0E-10. Expression data was extracted from EucGenIE [68] and is expressed as fragments per kilobase of transcript per million mapped reads (FPKM). Data is ordered as follows: YL, young leaf; ML, mature leaf; ST, shoot tips; F, flowers; R, roots; Ph, phloem and IX, immature xylem with yellow indicating lowest expression and red the highest expression between tissue. Enzyme abbreviated names: AACT, acetyl-CoA *C*-acetyltransferase; HMGS, 3-Hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-Hydroxy-3-methylglutaryl-CoA reductase; MK, mevalonate kinase; PMK, Phospho-MVA kinase; MPDC, Diphospho-MVA decarboxylase; IPPI, Isopentenyl diphosphate Δ-isomerase; DXS, 1-Deoxy-$_D$-xylulose 5-phosphate synthase; DXR, 1-Deoxy-$_D$-xylulose 5-phosphate reductoisomerase; MCT, 2-C-Methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK, 4-(Cytidine 5'-diphosphate)-2-*C*-methyl-$_D$-erythritol kinase; MDS, 2-*C*-methyl-$_D$-erythritol 2,4-cyclodiphosphate synthase; HDS, 4-Hydroxy-3-methylbut-2-enyl-diphosphate synthase; HDR, 4-Hydroxy-3-methylbut-2-enyl-diphosphate reductase.

Figure 2. Gaps in our understanding of proteins involved in the biosynthesis of specific terpenes. User-designed plants for biofuel terpenes, using synthetic biology, will require a comprehensive inventory of proteins specific to this process. Examples of these unknown proteins (represented by question marks) include signaling or receptor proteins such as kinases that elicit specific terpene biosynthesis; preferentially expressed enzyme variants; transporter proteins for precursor metabolites (from the mesophyll to the secretory cells) and secondary metabolites (from the secretory cells to the storage cavity) and regulatory proteins such as transcription factors. While not illustrated, modifying proteins such as glycosyltransferases which increases the stability of terpenes within the storage cavities at ambient temperatures is also an important consideration.

Figure 3. Proposed roadmap for commercialization of specific terpenes from plants. We propose a system biology approach using omics profiling of different plant organs e.g. oil glands (Strategy 1) or across developmental stages (Strategy 2). Generated omics data including transcriptome, proteome and metabolome profiling will be integrated using appropriate mathematical and statistical models to produce a candidate gene list, reveal co-expression patterning across tissues and gene-gene correlations within networks. This information will be used for strategic metabolic engineering via synthetic biology or selective breeding strategies to customize plants for specific terpenes. In combination with established silviculture techniques for maintaining elite lines, and optimized extraction protocols to minimize terpene lost during harvesting, will fortify these tailored plants as a viable source of terpenes for specialty biofuels.