methods has explored the combinatorial space of different sources of the heterologous genes comprising the pathway of interest.[32] Still dependent on a high throughput assay, these methods generally ignore the need for determining an optimal level of expression for the individual pathway genes and, as such, have proven less effective in structuring an optimal pathway.

In the present work, as an example of aspects of the invention, we focus on the optimal balancing between the upstream, IPP-forming pathway with the downstream terpenoid pathway of taxadiene synthesis. This is achieved by grouping the nine-enzyme pathway into two modules - a four-gene, upstream, native (MEP) pathway module and a two-gene, downstream, heterologous pathway to taxadiene (Fig. 1). Using this basic configuration, parameters such as the effect of plasmid copy number on cell physiology, gene order and promoter strength in an expression cassette, and chromosomal integration are evaluated with respect to their effect on taxadiene production. This modular and multivariable combinatorial approach allows us to efficiently sample the main parameters affecting pathway flux without the need for a high throughput screen. The multivariate search across multiple promoters and copy numbers for each pathway module reveals a highly non-linear taxadiene flux landscape with a global maximum exhibiting a 15,000 fold increase in taxadiene production over the control, yielding 300 mg/L production of taxadiene in small-scale fermentations. Further, we have engineered the P450 based oxidation chemistry in Taxol biosynthesis in *E. coli*, with our engineered strains improving the taxadien-5α-ol production 2400-fold over the state of the art. These improvements unlock the potential for the large scale production of thousands of valuable terpenoids by well-established microbial systems.

In one aspect, the present invention refers to a method for increasing terpenoid production in a cell that produces one or more terpenoids, comprising

providing an *Escherichia coli* (*E. coli*) that produces isopentyl pyrophosphate (IPP) and dimethylallyl pyrophospate (DMAPP) through an upstream methylerythritol pathway (MEP) and converts the IPP and DMAPP to a terpenoid through a recombinantly expressed downstream terpenoid synthesis pathway; and

controlling the accumulation of indole in the *E. coli* or in a culture of the *E. coli*, thereby increasing terpenoid production in the *E. coli*, wherein the indole accumulates in the *E. coli* culture at less than 100 mg/L, wherein the step of controlling the accumulation of indole in the *E. coli* or in a culture of the *E. coli* comprises balancing the upstream MEP pathway with

the downstream terpenoid synthesis pathway, modifying or regulating an indole pathway, and/or removing the accumulated indole from the cell culture through chemical methods, optionally using absorbents or scavengers.

~~Aspects of the invention relate to~~Also described herein are methods involving recombinantly expressing a taxadiene synthase enzyme and a geranylgeranyl diphosphate synthase (GGPPS) enzyme in a cell that overexpresses one or more components of the non-mevalonate (MEP) pathway. In some ~~embodiments~~examples the cell is a bacterial cell such as an *Escherichia coli* cell. In some ~~embodiments~~examples, the bacterial cell is a Gram-positive cell such as a *Bacillus* cell. In some ~~embodiments~~examples, the cell is a yeast cell such as a *Saccharomyces* cell or a *Yarrowia* cell. In some ~~embodiments~~examples, the cell is an algal cell or a plant cell.

In some ~~embodiments~~examples, the taxadiene synthase enzyme is a *Taxus* enzyme such as a *Taxus brevifolia* enzyme. In some ~~embodiments~~examples, the GGPPS enzyme is a *Taxus* enzyme such

as a *Taxus canadenis* enzyme. In some ~~embodiments~~examples, the gene encoding for the taxadiene synthase enzyme and/or the gene encoding for the GGPPS enzyme and/or the genes encoding for the one or more components of the MEP pathway is expressed from one or more plasmids. In some ~~embodiments~~examples, the gene encoding for the taxadiene synthase enzyme and/or the gene encoding for the GGPPS enzyme and/or the genes encoding for the one or more components of the MEP is incorporated into the genome of the cell.

In some ~~embodiments~~examples, one or more components of the non-mevalonate (MEP) pathway are selected from the group consisting of *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* and *ispB*. In certain ~~embodiments~~examples, *dxs*, *idi*, *ispD* and *ispF* are overexpressed. For example, *dxs*, *idi*, *ispD* and *ispF* can be overexpressed on the operon dxs-idi-idpDF. In some ~~embodiments~~examples, the gene encoding for the taxadiene synthase enzyme and the gene encoding for the GGPPS enzyme are expressed together on an operon.

In some ~~embodiments~~examples, the cell further expresses a taxadiene 5α-hydroxylase (T5αOH) or a catalytically active portion thereof. In certain ~~embodiments~~examples, the T5αOH enzyme or a catalytically active portion thereof is fused to a cytochrome P450 reductase enzyme or a catalytically active portion thereof. For example, the T5αOH enzyme can be At24T5αOH-tTCPR.

The expression of the taxadiene synthase enzyme, the GGPPS enzyme and the one or more components of the MEP pathway can be balanced to maximize production of the taxadiene. Methods ~~associated with the invention~~described herein can further encompass culturing a cell to produce taxadiene or taxadiene-5α-ol. In some ~~embodiments~~examples, at least 10 mg L^{-1} of taxadiene is produced. In certain ~~embodiments~~examples, at least 250 mg L^{-1} of taxadiene is produced. In some ~~embodiments~~examples, at least 10 mg L^{-1} of taxadiene-5α-ol is produced. In certain ~~embodiments~~examples, at least 50 mg L^{-1} of taxadiene-5α-ol is produced. In some ~~embodiments~~examples, the percentage of taxadiene conversion to taxadiene-5α-ol and the byproduct 5(12)-Oxa-3(11)-cyclotaxane is at least 50%, at least 75% or at least 95%.

Methods ~~associated with the invention~~described herein can further comprise recovering the taxadiene or taxadiene-5α-ol from the cell culture. In some ~~embodiments~~examples, the taxadiene or taxadiene-5α-ol is recovered from the gas phase while in other ~~embodiments~~examples, an organic layer is added to the cell culture, and the taxadiene or taxadiene-5α-ol is recovered from the organic layer.

~~Aspects of the invention relate to~~Also described herein are cells that overexpress one or more components of the non-mevalonate (MEP) pathway, and that recombinantly expresses a taxadiene synthase

enzyme and a geranylgeranyl diphosphate synthase (GGPPS) enzyme. In some ~~embodiments~~examples the cell is a bacterial cell such as an *Escherichia coli* cell. In some ~~embodiments~~examples, the bacterial cell is a Gram-positive cell such as a *Bacillus* cell. In some ~~embodiments~~examples, the cell is a yeast cell such as a *Saccharomyces* cell or a *Yarrowia* cell. In some ~~embodiments~~examples, the cell is an algal cell or a plant cell.

In some ~~embodiments~~examples, the taxadiene synthase enzyme is a *Taxus* enzyme such as a *Taxus brevifolia* enzyme. In some ~~embodiments~~examples, the GGPPS enzyme is a *Taxus* enzyme such as a *Taxus canadenis* enzyme. In some ~~embodiments~~examples, the gene encoding for the taxadiene synthase enzyme and/or the gene encoding for the GGPPS enzyme and/or the genes encoding for the one or more components of the MEP pathway is expressed from one or more plasmids. In some ~~embodiments~~examples, the gene encoding for the taxadiene synthase enzyme and/or the gene encoding for the GGPPS enzyme and/or the genes encoding for the one or more components of the MEP is incorporated into the genome of the cell.

In some ~~embodiments~~examples, the one or more components of the non-mevalonate (MEP) pathway is selected from the group consisting of *dxs*, *ispC*, *ispD*, *ispE*, *ispF*, *ispG*, *ispH*, *idi*, *ispA* and *ispB*. In certain ~~embodiments~~examples, *dxs*, *idi*, *ispD* and *ispF* are overexpressed. For example, *dxs*, *idi*, *ispD* and *ispF* can be overexpressed on the operon dxs-idi-idpDF. In some ~~embodiments~~examples, the gene encoding for the taxadiene synthase enzyme and the gene encoding for the GGPPS enzyme are expressed together on an operon. In some ~~embodiments~~examples, the expression of the taxadiene synthase enzyme, the GGPPS enzyme and the one or more components of the MEP pathway are balanced to maximize production of the taxadiene.

In some ~~embodiments~~examples, the cell further expresses a taxadiene 5α-hydroxylase (T5αOH) or a catalytically active portion thereof. In certain ~~embodiments~~examples, the T5αOH enzyme or a catalytically active portion thereof is fused to a cytochrome P450 reductase enzyme or a catalytically active portion thereof. For example, the T5αOH enzyme can be At24T5αOH-tTCPR. In some ~~embodiments~~examples, the cell produces taxadiene and/or taxadiene-5α-ol.

~~Aspects of the invention relate to~~Also described herein are methods for selecting a cell that exhibits enhanced production of a terpenoid, including creating or obtaining a cell that overexpresses one or more components of the non-mevalonate (MEP) pathway, producing terpenoid from the cell, comparing the amount of terpenoid produced from the cell to the amount of terpenoid produced in a control cell, and selecting a first improved cell that produces a higher amount

of terpenoid than a control cell, wherein a first improved cell that produces a higher amount of terpenoid than the control cell is a cell that exhibits enhanced production of terpenoid.

In some ~~embodiments~~examples, the cell recombinantly expresses a terpenoid synthase enzyme and/or a geranylgeranyl diphosphate synthase (GGPPS) enzyme. Methods can further comprise altering the level of expression of one or more of the components of the non-mevalonate (MEP) pathway, the terpenoid synthase enzyme and/or the geranylgeranyl diphosphate synthase (GGPPS) enzyme in the first improved cell to produce a second improved cell, and comparing the amount of terpenoid produced from the second improved cell to the amount of terpenoid produced in the first improved cell, wherein a second improved cell that produces a higher amount of terpenoid than the first improved cell is a cell that exhibits enhanced production of terpenoid. In some ~~embodiments~~examples, the terpenoid synthase enzyme is a taxadiene synthase enzyme. The cell can further recombinantly express any of the polypeptides ~~associated with the invention~~described herein.

~~Aspects of the invention relate to~~Also described herein are isolated polypeptides comprising a taxadiene 5α-hydroxylase (T5αOH) enzyme or a catalytically active portion thereof fused to a cytochrome P450 reductase enzyme or a catalytically active portion thereof. In some ~~embodiments~~examples, the cytochrome P450 reductase enzyme is a Taxus cytochrome P450 reductase (TCPR). In certain ~~embodiments~~examples, the taxadiene 5α-hydroxylase and TCPR are joined by a linker such as GSTGS (SEQ ID NO:50). In some ~~embodiments~~examples, the taxadiene 5α-hydroxylase and/or TCPR are truncated to remove all or part of the transmembrane region. In certain ~~embodiments~~examples, 8, 24, or 42 N-terminal amino acids of taxadiene 5α-hydroxylase are truncated. In certain ~~embodiments~~examples, 74 amino acids of TCPR are truncated. In some ~~embodiments~~examples, an additional peptide is fused to taxadiene 5α-hydroxylase. In certain ~~embodiments~~examples, the additional peptide is from bovine 17α hydroxylase. In certain ~~embodiments~~examples, the peptide is MALLLAVF (SEQ ID NO:51). In certain ~~embodiments~~examples, the isolated polypeptide is At24T5αOH-tTCPR. Also disclosed herein are~~Aspects of the invention also encompass~~ nucleic acid molecules that encode for any of the polypeptides ~~associated with the invention~~described herein and cells that recombinantly express any of the polypeptides ~~associated with the invention~~described herein.

~~Aspects of the invention relate to~~Also disclosed herein are methods for increasing terpenoid production in a cell that produces one or more terpenoids. The methods include controlling the accumulation of indole in the cell or in a culture of the cells, thereby increasing terpenoid production in a cell. Any of the cells described herein can be used in the methods, including bacterial cells,

such as *Escherichia coli* cells; Gram-positive cells, such as *Bacillus* cells; yeast cells, such as *Saccharomyces* cells or *Yarrowia* cells; algal cells; plant cell; and any of the engineered cells described herein.

In some ~~embodiments~~examples, the step of controlling the accumulation of indole in the cell or in a culture of the cells includes balancing the upstream non-mevalonate isoprenoid pathway with the downstream product synthesis pathways and/or modifying or regulating the indole pathway. In other ~~embodiments~~examples, the step of controlling the accumulation of indole in the cell or in a culture of the cells includes or further includes removing the accumulated indole from the fermentation through chemical methods, such as by using absorbents or scavengers.

The one or more terpenoids produced by the cell(s) or in the culture can be a monoterpenoid, a sesquiterpenoid, a diterpenoid, a triterpenoid or a tetraterpenoid. In certain ~~embodiments~~examples, the terpenoids is taxadiene or any taxol precursor.

~~Aspects of the invention relate to~~Also disclosed herein are methods that include measuring the amount or concentration of indole in a cell that produces one or more terpenoids or in a culture of the cells that produce one or more terpenoids. The methods can include measuring the amount or concentration of indole two or more times. In some ~~embodiments~~examples, the measured amount or concentration of indole is used to guide a process of producing one or more terpenoids. In some ~~embodiments~~examples, the measured amount or concentration of indole is used to guide strain construction.

These and other aspects of the invention, as well as various embodiments thereof, will become more apparent in reference to the drawings and detailed description of the invention.

Brief Description of the Drawings

The accompanying drawings are not intended to be drawn to scale. In the drawings, each identical or nearly identical component that is illustrated in various figures is represented by a like numeral. For purposes of clarity, not every component may be labeled in every drawing. In the drawings:

Figure 1. Multivariate-modular isoprenoid pathway engineering reveals strong non-linear response in terpenoid accumulation. To increase the flux through the upstream MEP pathway, we targeted reported bottleneck enzymatic steps (dxs, idi, ispD and ispF) for overexpression by an operon (dxs-idi-ispDF).[28] To channel the overflow flux from the