Check for updates

Evolutionary flexibility and rigidity in the bacterial methylerythritol phosphate (MEP) pathway

Bailey Marshall[1,2]*, Kaustubh Amritkar[3], Michael Wolfe[1,2], Betül Kaçar[3] and Robert Landick[1,2,3]

[1]Department of Biochemistry, University of Wisconsin−Madison, Madison, WI, United States, [2]DOE Great Lakes Bioenergy Research Center, University of Wisconsin−Madison, Madison, WI, United States, [3]Department of Bacteriology, University of Wisconsin−Madison, Madison, WI, United States

Terpenoids are a diverse class of compounds with wide-ranging uses including as industrial solvents, pharmaceuticals, and fragrances. Efforts to produce terpenoids sustainably by engineering microbes for fermentation are ongoing, but industrial production still largely relies on nonrenewable sources. The methylerythritol phosphate (MEP) pathway generates terpenoid precursor molecules and includes the enzyme Dxs and two iron−sulfur cluster enzymes: IspG and IspH. IspG and IspH are rate limiting-enzymes of the MEP pathway but are challenging for metabolic engineering because they require iron−sulfur cluster biogenesis and an ongoing supply of reducing equivalents to function. Therefore, identifying novel alternatives to IspG and IspH has been an on-going effort to aid in metabolic engineering of terpenoid biosynthesis. We report here an analysis of the evolutionary diversity of terpenoid biosynthesis strategies as a resource for exploration of alternative terpenoid biosynthesis pathways. Using comparative genomics, we surveyed a database of 4,400 diverse bacterial species and found that some may have evolved alternatives to the first enzyme in the pathway, Dxs making it evolutionarily flexible. In contrast, we found that IspG and IspH are evolutionarily rigid because we could not identify any species that appear to have enzymatic routes that circumvent these enzymes. The ever-growing repository of sequenced bacterial genomes has great potential to provide metabolic engineers with alternative metabolic pathway solutions. With the current state of knowledge, we found that enzymes IspG and IspH are evolutionarily indispensable which informs both metabolic engineering efforts and our understanding of the evolution of terpenoid biosynthesis pathways.

KEYWORDS

isoprenoid biosynthesis, terpenoid biosynthesis, metabolic engineering, synthetic biology, comparative genomics

1. Introduction

Terpenoids are a diverse class of biomolecules with varied cellular functions, ranging from basic roles such as maintaining membrane fluidity and supporting electron transport to highly specialized roles such as hormone signaling and cellular defense (Sell, 2006; Avalos et al., 2022). All terpenoids are built from five carbon-building block precursors [dimethylallyl diphosphate (DMADP) and isopentenyl diphosphate (IDP)] and exhibit high structural diversity with over 95,000 different terpenoid compounds identified to date (Faylo, 2021). With applications as diverse as their molecular structures, terpenoids are used extensively in support of modern-day

life such as commodity solvents, pharmaceuticals, and fuels (Ajikumar et al., 2008; Tracy et al., 2009; Harvey et al., 2010; George et al., 2015) and for exploring life's earliest record on Earth (Brocks et al., 1999; Summons et al., 2006; Hillier and Lathe, 2019).

Terpenoids are currently harvested from plants, generated by microbial fermentation, or produced chemically from nonrenewable petroleum (Leavell et al., 2016). To date, variable yields, inefficiencies, and cost have made natural terpenoid production from plants or microbes unattractive leading to a global terpenoid supply chain that favors production from petroleum at the expense of sustainability and environmental impact (Leavell et al., 2016). Engineered microbes have the potential for improved yields and increased diversity in the types of terpenoids that are produced, with some important successes to date (George et al., 2015; Benjamin et al., 2016; Navale et al., 2021). Current efforts in engineering microbes for terpenoid production use one or both of two characterized "canonical" biosynthetic pathways (the mevalonate [MEV] and methylerythritol phosphate [MEP] pathways) and aim to produce higher yields and an ever-increasing range of terpenoid products as sustainably and economically as possible. However, knowledge about the evolutionary flexibility of terpenoid biosynthesis or possible alternative metabolic routes for terpenoid biosynthesis is lacking.

Terpenoids are used in membranes in all known domains of life (Lange et al., 2000; Holstein and Hohl, 2004) and are present in molecular fossils billions of years old (Ourisson and Albrecht, 1992; Brocks et al., 1999, 2005). Given their universality and presence in the fossil records, terpenoids are identified as biomarkers, providing insight into life's earliest evolution on Earth (Lange et al., 2000; Lombard and Moreira, 2011; Summons et al., 2022). Yet, important questions remain about the evolution of terpenoid biosynthesis, including the identity of the primordial pathway for terpenoid biosynthesis and how horizontal gene transfer may have contributed to the evolution of terpenoid biosynthetic pathways (Zeng and Dehesh, 2021; Hoshino and Villanueva, 2023). Greater understanding of the variations within terpenoid biosynthetic pathways may further elucidate the evolution of this important ancient pathway.

Of the two known modern metabolic routes that generate the five carbon building-block precursor molecules from which all terpenoids are derived, the MVA pathway is generally used by eukaryotes and archaea (Zhao et al., 2013) and the MEP pathway is generally used by bacteria and plastids (Figures 1A,B; Rohmer, 2008; Zhao et al., 2013). The MEP pathway can theoretically generate more terpenoid product from the same starting material as the MVA pathway (Figures 1A,B; Rude and Schirmer, 2009; Whited et al., 2010). As such, the MEP pathway may have greater potential for producing terpenoids efficiently, but engineering efforts have yet to surmount several important barriers to increasing flux through the pathway.

The first enzyme of the MEP pathway, Dxs, is an initial bottleneck to carbon flux (Figure 1A; Harker and Bramley, 1999; Li and Wang, 2016). This bottleneck can be relieved by overexpression of Dxs, at which point IspG which converts methylerythritol 2, 4-cyclodiphosphate (MEcPP) to hydroxymethylbutenyl 4-diphosphate (HMBPP) becomes rate limiting (Zhou et al., 2012; Liu et al., 2014; Li et al., 2017; Volke et al., 2019; Khana et al., 2023). IspG and the next enzyme downstream, IspH, are both iron- sulfur cluster containing [4Fe-4S] proteins (Wang and Oldfield, 2014). Iron–sulfur cluster biogenesis and insertion into proteins is a complex process (Xu and Møller, 2011). After IspG and IspH have been loaded with iron–sulfur

clusters, they require an external supply of reducing equivalents to function. Altogether, simple overexpression of IspG and IspH may not relieve the bottleneck because these proteins require extensive cellular processes in the generation and loading of iron–sulfur clusters and the maintenance of substrates for these proteins to function. Due to the barriers to increasing flux through IspG and IspH, identifying alternative enzymes to IspG and IspH would be advantageous for engineering of the MEP pathway.

Multiple routes that allow native or acquired circumvention of the first MEP pathway enzyme Dxs have been identified but to date no alternative routes circumventing IspG or IspH have been found (Figure 1A; Sauret-Güeto et al., 2006; Sangari et al., 2010; Erb et al., 2012; Perez-Gil et al., 2012; Warlick et al., 2012; Carretero-Paulet et al., 2013; Kirby et al., 2015). In contrast, multiple alternative routes for the MVA pathway for terpenoid precursor biosynthesis have been discovered, making it a pertinent demonstration of the potential for variations on a metabolic pathway to evolve using alternative enzymes and metabolites (Figure 1B; Hayakawa et al., 2018; Hoshino and Villanueva, 2023).

Here we define evolutionary flexibility as steps in metabolism for which multiple enzymatic solutions have evolved, as is seen in the MVA pathway (Figure 1B). In contrast, evolutionarily rigid metabolic steps are those for which alternative enzymes that would allow circumvention of that step have not evolved. To inform efforts to engineer the MEP pathway, we investigated the evolutionary diversity of terpenoid biosynthesis strategies in bacteria. We explored bacterial diversity by assessing an available database of bacterial genomes for terpenoid biosynthesis strategies and alternative enzymatic routes to steps in the MEP pathway. In particular, we asked if enzymes have evolved to circumvent IspG and IspH which pose known engineering constraints, or to Dxs, for which some alternatives exist. We identified bacteria with apparently incomplete MEP pathways as candidates for non-canonical terpenoid biosynthesis routes. Our study provides a comprehensive survey of the known bacterial species for non-canonical MEP pathways and underscores the evolutionary significance of IspG and IspH.

2. Materials and methods

2.1. Database of species and HMMER profiles

Fasta files of the NCBI database of reference and representative bacterial genomes available at the assembly level "complete chromosome" were downloaded (access date 6/19/2023). Profile Hidden Markov Models (pHMM) from the KOfam database (Aramaki et al., 2020) were downloaded from KEGG using genome.jp/ftp/db/kofam/ (access date 6/19/2023, KEGG release 106.0) as well as the "ko_list" file which contains adaptive score thresholds that define an ortholog match for each pHMM as determined by KOfam and described in more detail in section 2.3. The "ko_list" file was trimmed to only include the MEP and MVA genes of interest which increased computational analysis speed by limiting ortholog detection to only these genes. DxrII did not have a pHMM from KOfam, so we used hmmbuild to construct a pHMM model (Eddy, 2011) using the set of 7 genes predicted to be DxrII as annotated in NCBI's conserved protein domain family in COG4091.
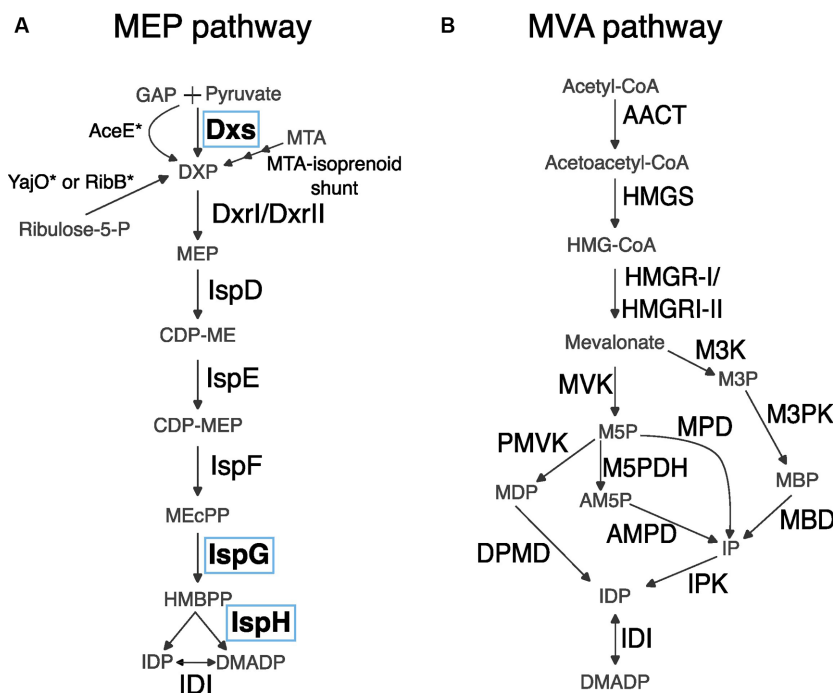
FIGURE 1

Reactions of MEP and MVA metabolic pathways. **(A)** MEP pathway. Asterisks indicate a mutant version of an enzyme is necessary for completing a reaction. Enzymes circled in blue, Dxs, IspG, and IspH, are the focus of this report. GAP, glyceraldehyde 3-phosphate; Dxs, DXP synthase; DXP, 1-deoxyxylulose 5-phosphate; Dxr, DXS reductoisomerase; MEP, 2-C-methyl-erythritol 4-phosphate; CDP-ME, 4-(cytidine 5′-diphospho)-2-C-methyl-erythritol; CDP-MEP, 2-phospho-4-(cytidine 5′-diphospho)-2-C-methyl-erythritol; MEcPP, 2-C-methyl-erythritol-2,4-cyclodiphosphate; HMBPP, 4-hydroxy-3-methyl 2-butenyl diphosphate; IDP, isopentenyl diphosphate; DMADP, dimethylallyl diphosphate. **(B)** MVA pathway. AACT, Acetoacetyl-CoA thiolase; HMGS, HMG synthase; HMG-CoA, hydroxymethylglutaryl-CoA; HMGR, HMG reductase; MVK, mevalonate kinase; M5P, mevalonate 5-phosphate; M3K, mevalonate 3-kinase; M3P, mevalonate 3-phosphate; PMVK, phosphomevalonate kinase; MDP, mevalonate 5-diphosphate; M5PDH, M5P dehydratase; AM5P, trans-anhydromevalonate 5-phosphate; AMPD, AM5P decarboxylase; MPD, M5P decarboxylase; IP, isopentenyl phosphate; M3PK, M3P 5-kinase; MBP, mevalonate 3,5-bisphosphate; MBD, MBP decarboxylase; DPMD, diphosphomevalonate decarboxylase.

Full bioinformatic pipeline is available at https://github.com/bamarshall2/evolutionary_flexibility.

2.2. Open reading frame (ORF) and KofamScan for ortholog identification

For each genome, orfipy was used to identify ORFs of at least 150 nucleotides in length and generate an output of the peptide translation using the specifications "--min 150 --pep $file_name.fa" (Singh and Wurtele, 2021). All genomes were translated using standard codon table 1 except genera *Spiroplasma*, *Mycoplasma*, *Mycoplasmoides*, *Malacoplasma*, and *Ureaplasma* which were translated with codon table 4 according to their taxonomy using the additional orfipy argument "--table 4." Kofam_scan tool was run on each orfipy output using the curated KO list of MEP and MVA genes of interest and pHMM profiles from the KOfam database. The kofam_scan command used was "./exec_annotation --format=detail-tsv -o~path_to_output/$output_file_name ~path_to_input/$input_file_name."

2.3. Identification of orthologous genes

For Dxs, DxrI, IspE, IspDF, IspG, IspH, HMGS, HMGRI, HMGRII, MVK, DPMD, Rubisco-like protein MTRu-1P isomerase,

and cupin type I MTXu-5P methylsulferase a gene ortholog was determined to be present if the similarity score from alignment exceeded the KOfam score threshold for identifying similarity (Aramaki et al., 2020). This task was completed using custom scripts, all code is available at GitHub.[1]

The computation of the KOfam score thresholds is described in detail in Aramaki et al. (2020). In brief, to compute the score threshold for a protein family in the KEGG Orthology database, the manually curated set of proteins is randomly divided into three groups, one for use as a positive training set and the other two for generation of the pHMM. Proteins that are not part of the protein family are used as a negative training dataset. Bit scores of alignments between the pHMM and the training data sets allows determination of a score threshold that maximizes the F measure which is the harmonic mean of precision and recall. The process of determining a threshold score is repeated three times by swapping the positive training set with one used to train the pHMM until each group of data has been used to both train and test the pHMMs. A final threshold score is then determined as the average of each iteration's threshold score. For the proteins for which we used the score thresholds, the average F measure was 0.979 with MVK having

---

the smallest F measure of 0.897 and DxrI having the highest F measure of 1. IspD and IspF occasionally exist as a bifunctional fusion protein called IspDF. We observed that presence/absence calls for these two proteins were sometimes incorrect when using the KOfam adaptive thresholds which is reflected by lower F measures for IspD (F measure = 0.720) and IspF (F measure = 0.704). These inaccurate calls may have been due to attempts to distinguish whether a monofunctional or bifunctional protein is present, which requires high thresholds for alignment scores and causes reduced accuracy. For our purposes, identification of an IspD or IspF domain is of key relevance, but whether it is in a monofunctional or bifunctional version of the protein is not. To best identify IspD and IspF, we use the KOfam adaptive threshold for identification of bifunctional IspDF (F measure = 0.997), and for IspD or IspF we set a cutoff threshold of ORFs with alignment e-values of less than 10^{-10}. The KOfam database did not have an adaptive threshold cutoff for DxrII for determination of orthology. For DxrII, we set the score cutoff at alignments with e-values of less than 10^{-10} as well.

To determine a species' MEP and MVA gene status, we sorted for presence and absence using Excel. Full table output of gene presence or absence is available in Supplementary Table S1.

2.4. Determining MVA pathway status

Diversity exists in the lower part of the MVA pathway (Figure 1B) but complete characterization of the MVA pathway was not the goal of this research so determination of MVA pathway presence was more relaxed than analyses for the MEP pathway. For the MVA pathway, if 2 out of 4 genes (encoding HMGS, HMGRI/II, MVK, and DPMD) were present, the organism was considered to possess the MVA pathway.

2.5. Phylogenetic tree building

Species trees were constructed using Genome Taxonomy Database (GTDB) master tree (Parks et al., 2018) and only the species included in this study were retained. The GTDB master tree included 2,210 of the 4,400 species (~50%) in the presented analyses, so the phylogenetic trees show half of the species analyzed.

For protein trees, the protein sequences corresponding to the entries from the species tree were extracted and aligned by MAFFT (Katoh et al., 2002) using default parameters with iteration refinement over 1,000 cycles. Columns with >95% gaps were removed using TrimAL (Capella-Gutiérrez et al., 2009). These sequences were used to build the phylogenetic tree for the MEP pathway proteins. A maximum-likelihood phylogenetic tree was constructed for each of the protein using IQ-Tree 2 (Nguyen et al., 2015; Minh et al., 2020) and ModelFinder plus (Kalyaanamoorthy et al., 2017) was implemented for selecting the best-fit evolutionary model. Branch support values were calculated using the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT) with 1,000 bootstrap replicates and 1,000 ultrafast bootstrap (UFBoot) replicates optimized by the nearest neighbor interchange (NNI). The phylogenetic trees were visualized using FigTree v1.4.4 (Rambaut, 2018). The reconstructed trees were rooted using midpoint rooting and annotated in R studio (Farris, 1972).

3. Results

3.1. Some bacteria use a non-canonical MEP pathway lacking Dxs

To find candidate species for a non-canonical MEP pathway, we started with an available database of representative bacterial genomes (O'Leary et al., 2016) and extracted all open reading frames (ORFs) from these genomes (Singh and Wurtele, 2021). Next, we used a HMMER based search tool called KofamScan for identification of orthologs to MEP and MVA genes (Figure 2A; Potter et al., 2018; Aramaki et al., 2020).

Of 4,400 species in the database, 3,411 (78%) had all 7 MEP genes detected (Dxs, DxrI/II, IspD, IspE, IspF, IspG, and IspH) but lack the MVA pathway according to our criteria (see Methods) (Figure 2B). Some species (541; 12%) had the MVA pathway and no MEP pathway genes or an incomplete MEP pathway detected whereas others (83; 1.8%) had all genes for the MEP pathway and at least two MVA pathway genes. Of 136 species (3.1%) that lacked a complete set of identifiable orthologs for both the MEP and MVA pathways, many had small genomes with a median size of 843,495 base pairs and included known endosymbiotic species *Mycoplasma*, *Rickettsia*, and *Spiroplasma* (Eberl et al., 2004). The remaining 229 species (5.2%) had at most one MVA gene and an incomplete MEP pathway where at least one gene of Dxs, DxrI/II, IspD, IspE, IspF, IspG, or IspH appeared to be missing making them candidates for species using a non-canonical MEP pathway (Figure 2B).

Of these 229 candidates, most were missing Dxs only (181, 79%) (Figure 2C) and thus might use an alternative to Dxs. One such potential alternative is the 5-methylthioadenosine (MTA) isoprenoid shunt first identified in *Rhodospirillum rubrum* (Figure 1; Erb et al., 2012). The MTA-isoprenoid shunt uses the enzymes Rubisco-Like protein MTRu-1P isomerase and Cupin type I protein MTXu-5P methylsulferase to convert polyamine biosynthesis intermediate MTA to isoprenoid biosynthesis intermediate 1-deoxy-D-xylulose 5-phosphate (DXP) (Warlick et al., 2012). For the candidate non-canonical MEP genomes lacking Dxs, 4 (2%) had only the MTXu-5P methylsulfurase and 5 (3%) had both the MTXu-5P methylsulfurase and the MTRu-1P isomerase (Supplementary Table S2). In these species, it is possible the MTA-isoprenoid shunt may substitute for Dxs, but this hypothesis will require further testing.

For the remaining 172 candidate species (95%) missing Dxs, the specific mechanism that allows the absence of Dxs requires further investigation to determine. Regardless of the mechanism, the identification of candidate species that contain all other MEP enzymes except Dxs suggest that the MEP pathway has evolutionary flexibility for the enzymatic activity that Dxs provides.

3.2. IspG and IspH are evolutionarily rigid

Dxs is the only MEP pathway enzyme for which we identified a subset of candidate species that lack Dxs suggesting that an alternative strategy that circumvents Dxs has evolved. For IspG and IspH, there were no promising candidates for identification of alternative MEP pathways that circumvent one or both enzymes (Figure 2C). Thus, we define IspG and IspH as evolutionary rigid.
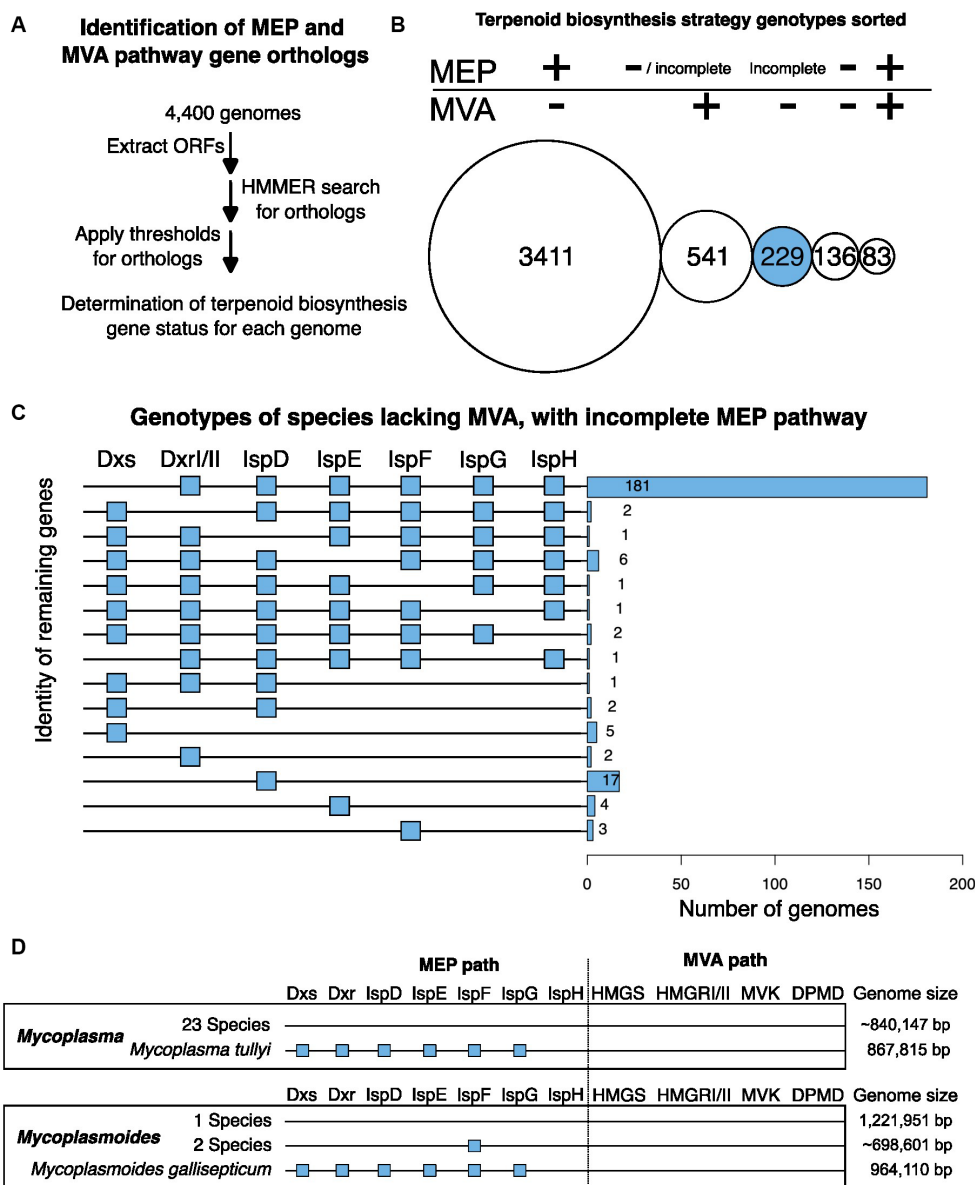
FIGURE 2
In silico identification of terpenoid biosynthesis metabolic pathways. **(A)** Bioinformatic analysis pipeline. The HMMER search for orthologs used profile Hidden Markov Models (pHMM) from the KOfam database (Aramaki et al., 2020). For ortholog identification cutoff values of genes Dxs, DxrI, IspE, IspDF, IspG, and IspH, we relied on the KOfam database of adaptive score thresholds determined for each gene (Aramaki et al., 2020). For DxrII, IspD, and IspF we defined orthologs as alignments to the pHMM with e-values lower than 1^{-10}. See Methods for more details. **(B)** Terpenoid biosynthesis strategy genotypes sorted. Circle areas are scaled to the total number of species for each genotype. "+" indicates all enzymes of the MEP pathway or at least two enzymes of the MVA pathway were detected. "−"indicates none of the enzymes of the MEP pathway or fewer than two enzymes of the MVA pathway were detected. Incomplete indicates that only some of the enzymes of the MEP pathway were detected. For the 541 MVA+ species, some have no detectable MEP orthologs while others do. This is indicated by the "−/incomplete" description of the MEP pathway. **(C)** Bar chart of 229 candidate species for alternative MEP pathways with visual descriptor of genotype on the left. A blue box indicates the presence of the gene whereas no box indicates a lack of that gene. **(D)** Visual representation of MEP and MVA genes present in *Mycoplasma* and *Mycoplasmoides*. The presence of a blue square indicates the identification of a gene ortholog. The lack of a square indicates that no ortholog was identified. Genome sizes are displayed to the right, with average genome sizes indicated with "~".

The single candidate genome lacking IspG only was *Lelliotia steviae* (Figure 2C). When investigated, we found that the fasta file for this species was 1 million base pairs shorter than the sequence originally published and that other closely related *Lelliottia* species all had IspG present (Supplementary Table S4). Thus, we suggest the full *Lelliotia steviae* genome likely encodes IspG but that it may be in the 1 million missing base pairs.

We interpret this single genome lacking IspG to be indicative of the level of inaccuracy within a large, collaborative database of species.

The two species lacking IspH only were *Mycoplasma tullyi* and *Mycoplasmoides gallisepticum* (Figures 2C,D). Both species have genomes smaller than 1 million base pairs and all other *Mycoplasma* and *Mycoplasmodies* included in this database had

neither the MVA nor the MEP pathway consistent with all being endosymbionts and lacking terpenoid biosynthesis metabolisms (Figure 2D; Eberl et al., 2004). It appears that though closely related *Mycoplasma* and *Mycoplasmoides* species have evolved to rely on host isoprenoid biosynthesis, *Mycoplasma tullyi* and *Mycoplasmoides gallisepticum* have retained a near complete MEP pathway despite having comparably small genomes. *Mycoplasmoides gallisepticum* has been demonstrated to generate MEP pathway intermediate HMBPP (Eberl et al., 2004) though the essentiality of the MEP pathway in either organism remains to be confirmed. With only two species that are outliers in their genera, they do not constitute a compelling lead for identifying an alternative enzyme to IspH. It is possible that these species are making use of a near complete MEP pathway that ends with the generation of HMBPP and it is also possible that they do possess an alternative enzyme that circumvents IspH. These two species may have their own interesting evolutionary rationale for maintaining a nearly complete MEP pathway, but it will require further study to understand. The remaining 45 candidate species have no compelling patterns that suggest the presence of a non-canonical MEP pathway that would aid in the engineering of the MEP pathway (Figure 2B).
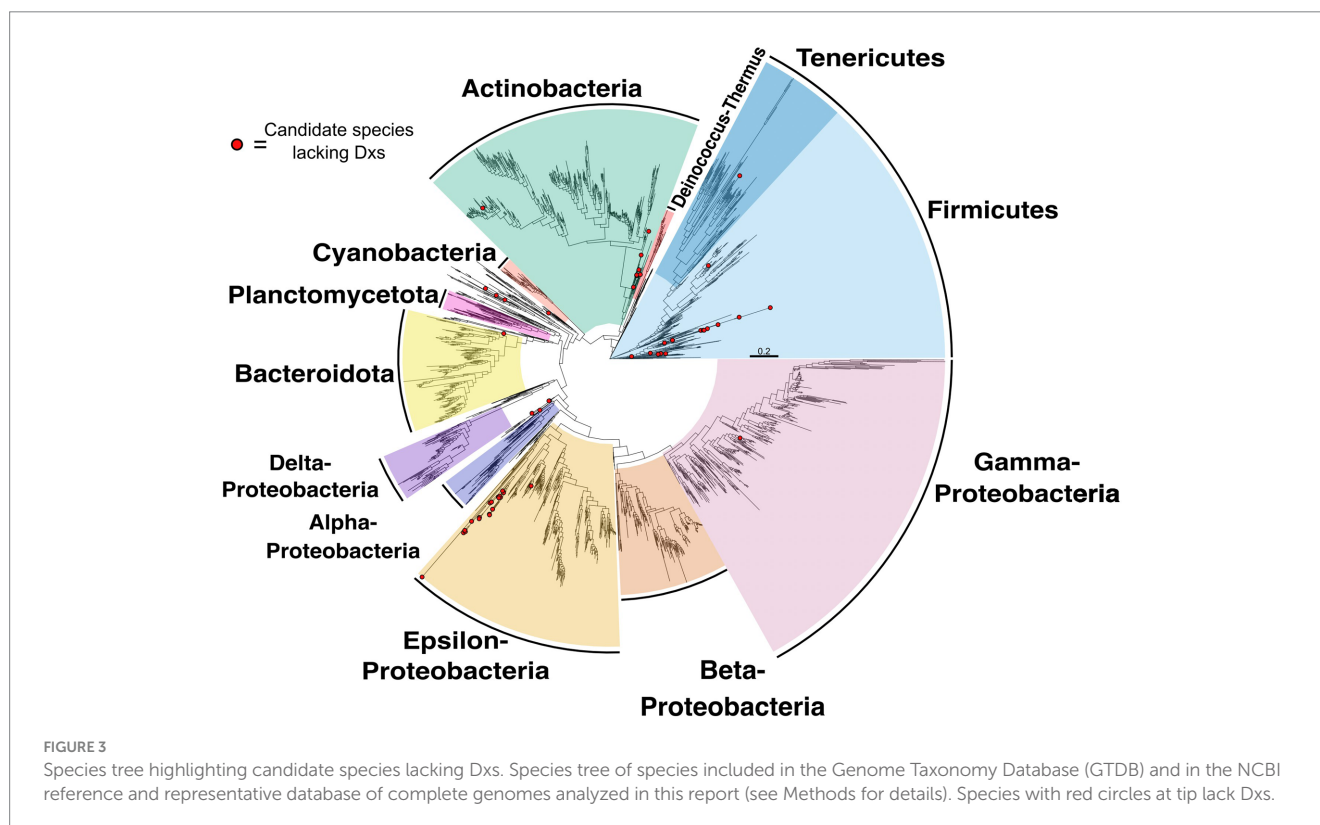
We conclude that an alternative enzyme to Dxs probably exists and is in use in some of the 181 candidate species lacking Dxs. We interpret this to mean that an alternative to Dxs has probably evolved making it evolutionarily flexible. Meanwhile, given our approach, we cannot identify any species that may be using alternative enzymes to IspG or IspH.
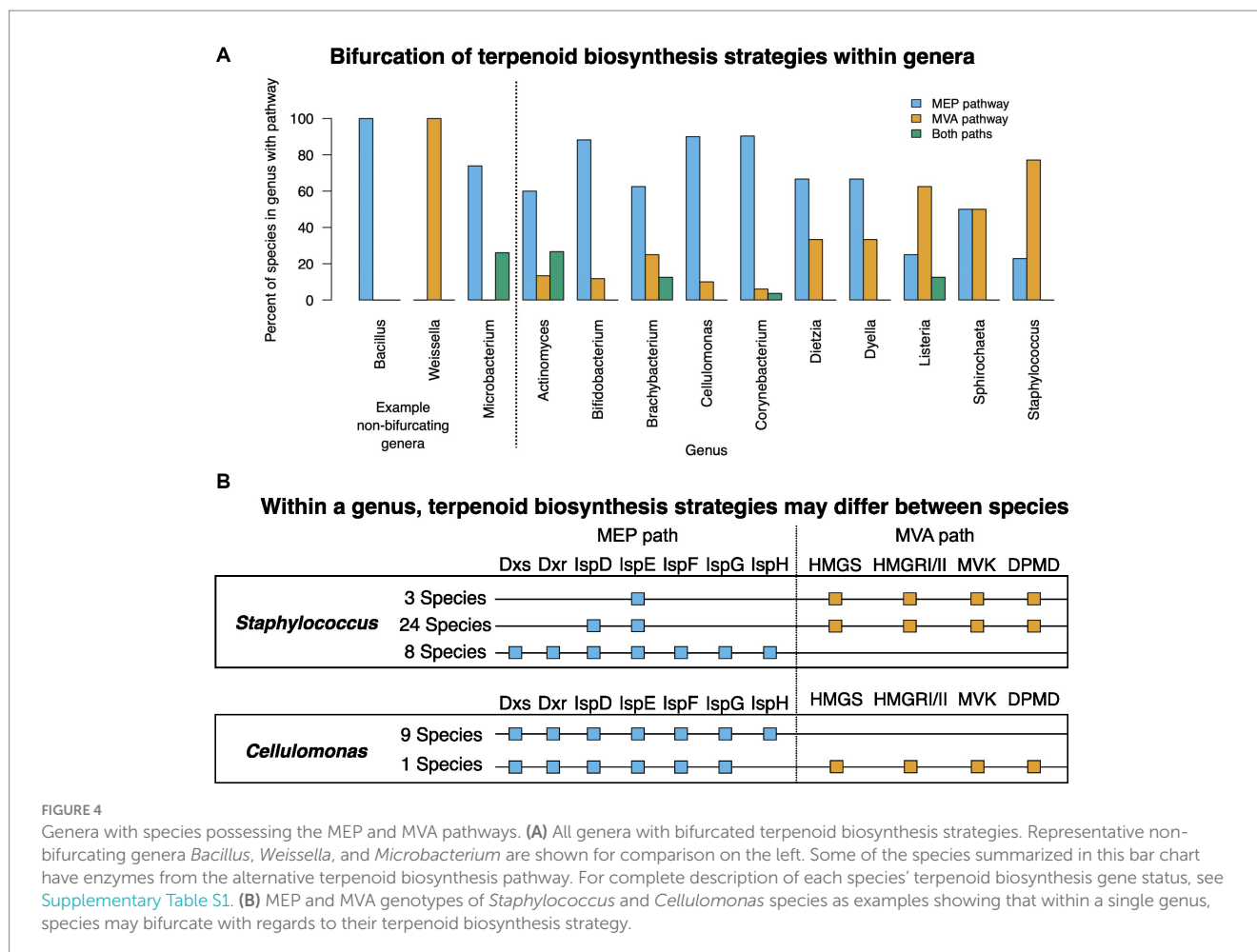
3.3. Phylogenetic distribution of species lacking Dxs

Candidate species that might encode a non-canonical MEP pathway that lack Dxs are distributed across different phylogenetic groups (Figure 3). Species lacking Dxs can be found in the phyla Firmicutes, Actinobacteria, Alpha-proteobacteria and others. A complete list of these species is presented in Supplementary Figure S3. These species may use the same or different strategies to circumvent Dxs, but additional analysis will be required to identify these pathways.

3.4. Within a genus terpenoid biosynthesis strategies vary at the species level

Generally, all species in a genus use either the MEP or MVA pathway (e.g., *Bacillus* and *Weissella*; Figure 4A). Another common pattern is for species to possess one complete pathway, and some of the enzymes of the other pathway. In other genera, most species use the same isoprenoid biosynthesis pathway although a subset encode enzymes for both pathways (e.g., *Microbacterium*; Figure 4A). For 10 genera with multiple represented species, terpenoid biosynthesis strategies are bifurcated (Figure 4A). In these genera, some species use one terpenoid biosynthesis strategy whereas other species use the other terpenoid biosynthesis strategy.

FIGURE 3
Species tree highlighting candidate species lacking Dxs. Species tree of species included in the Genome Taxonomy Database (GTDB) and in the NCBI reference and representative database of complete genomes analyzed in this report (see Methods for details). Species with red circles at tip lack Dxs.

FIGURE 4
Genera with species possessing the MEP and MVA pathways. **(A)** All genera with bifurcated terpenoid biosynthesis strategies. Representative non-bifurcating genera *Bacillus*, *Weissella*, and *Microbacterium* are shown for comparison on the left. Some of the species summarized in this bar chart have enzymes from the alternative terpenoid biosynthesis pathway. For complete description of each species' terpenoid biosynthesis gene status, see Supplementary Table S1. **(B)** MEP and MVA genotypes of *Staphylococcus* and *Cellulomonas* species as examples showing that within a single genus, species may bifurcate with regards to their terpenoid biosynthesis strategy.

For example, *Staphylococcus* was represented by 35 species in the database, 27 of which (77%) used the MVA pathway and possess one or two MEP pathway enzymes whereas the remaining 8 species (23%) encode complete MEP pathways but no MVA pathway enzymes (Figure 4B) (Misic et al., 2016). In *Cellulomonas*, 9 of the species encode a complete MEP pathway with no MVA enzymes but a single species, *Cellulomonas taurus,* encodes an incomplete MEP pathway and a complete MVA pathway (Figure 4B). The MVA pathway genes in *Cellulomonas taurus* are in two operons in the genome, roughly 1 million base pairs apart.
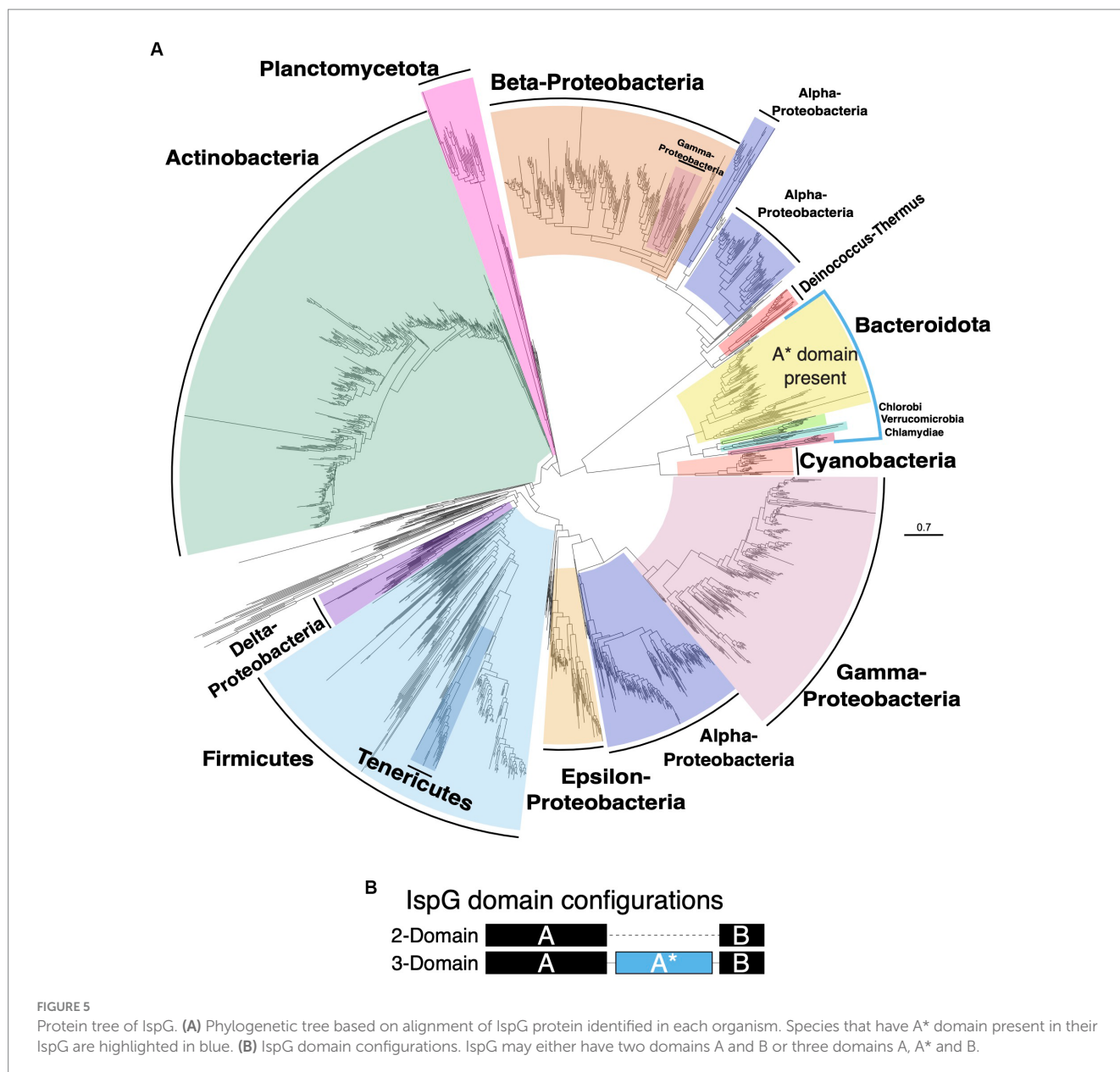
These data suggest that selective pressure exists to retain a terpenoid biosynthesis strategy, but that a species may use MEP, MVA, or both based on evolutionary trajectory.

3.5. IspG protein tree differs from species tree and the three-domain IspG proteins cluster together

To understand the impact of horizontal gene transfer on the MEP pathway in bacteria, we generated evolutionary trees from protein sequences for Dxs, DxrI, IspE, IspG, and IspH (Supplementary Figure S5). The Dxs, DxrI, IspE, IspG, and IspH trees differ from the species tree in that Actinobacteria have clustered with Proteobacteria in the protein trees instead of

Bacteroidota as was the case of the species tree (Figures 3, 5A and Supplementary Figure S5). Also, a notable feature for these protein trees is the paraphyletic nature of the Proteobacteria phylum, which is most clearly observed in the IspG tree. In the IspG protein tree, the Proteobacterial phylum is split into two groups; one Proteobacterial group clusters with Bacteroidota and the other with Firmicutes. Further, the Proteobacterial phylum is more widely distributed in the IspG protein tree with some γ-Proteobacteria clustering within β-Proteobacteria and the α-Proteobacteria being split.

IspG is generally a two-domain protein in Bacteria, but a third domain called A* may be present (Figure 5B; Liu et al., 2012). The A* domain has a TIM barrel fold, but its exact role has yet to be elucidated (Liu et al., 2012). The phyla Bacteroidota, Chlorobi, Verrucomicrobia, and Chlamydiae all possess the A* domain and cluster together (Figures 5A,B). Planctomycetota, Verrucomicrobia, and Chlamydiae are the major phyla that constitute superphylum PVC (Wagner and Horn, 2006), however only Verrucomicrobia and Chlamydiae possess the three-domain form of IspG whereas Planctomycetota possess the 2-domain form. This splitting of the PVC superphylum with regards to IspG has been illustrated previously (Zeng and Dehesh, 2021). Absence of the three-domain form of IspG only in Planctomycetota of the PVC superphylum could be due to horizontal gene transfer as the phylogenetic placement of the Plancomycetota IspG sequence in the protein

FIGURE 5
Protein tree of IspG. **(A)** Phylogenetic tree based on alignment of IspG protein identified in each organism. Species that have A* domain present in their IspG are highlighted in blue. **(B)** IspG domain configurations. IspG may either have two domains A and B or three domains A, A* and B.

tree is not congruent with the PVC clade in the species tree. An alternative possibility is that the A* domain evolved in a common ancestor of Bacteroidota, Plancomycetota, Verrucomicrobia, and Chlamydiae that was then lost in Planctomycetota. With the data presented here, these two possibilities cannot be distinguished.

Based on the differences between the MEP protein trees and the species tree, MEP pathway inheritance is not strictly vertical. Therefore, we suggest that horizontal gene transfer may have played a role in the evolution of this metabolic pathway.

4. Discussion

In this study, we assessed the MEP and MVA pathways and protein diversity across bacteria. Our goal was to inform options for engineering the precursor pathways for microbial terpenoid biomanufacturing by determining if alternatives exist for some steps in precursor pathways. In particular, we were interested in alternatives to Dxs and whether non-canonical MEP pathways can circumvent the IspG and IspH enzymes, which are known to pose major engineering challenges. To achieve this goal, we identified genomes lacking the MVA pathway, with an incomplete MEP pathway referred to as candidate species for a non-canonical MEP pathway. We found that Dxs is evolutionarily flexible and that alternatives may exist whereas for IspG and IspH we could not identify any candidate species that appear to circumvent these enzymes. These results may aid the ongoing quest to understand early evolution of terpenoid biosynthetic pathways.

Identified alternatives to the canonical MEP enzymes include DxrII, the MTA-isoprenoid shunt, and mutations that arise when Dxs is knocked out (Figure 1A; Sauret-Güeto et al., 2006; Sangari et al., 2010; Erb et al., 2012; Perez-Gil et al., 2012; Warlick et al., 2012; Carretero-Paulet et al., 2013; Kirby et al., 2015). The suppressor mutations that have been documented to arise when

Dxs is knocked out include mutations in AceE, RibB, or overexpression of YajO. Altogether, there are multiple routes that circumvent Dxs but alternatives pathways to circumvent IspG or IspH have not yet been identified.

We identified 181 candidate species for a non-canonical MEP pathway missing Dxs. In contrast, we identified fewer than 40 species missing IspG, IspH, or both in combination with other MEP enzymes and no compelling patterns among these species (Figure 2C). We conclude that within this database of 4,400 species, all MEP utilizing species possess IspG and IspH as part of their MEP pathways. On the other hand, some of the 181 candidate species missing Dxs likely have an alternative MEP pathway strategy. This result would align with previous research where alternative routes for circumventing Dxs have been identified (see above).

Taken together, our results suggest that within the MEP pathway, evolutionary flexibility exists for Dxs. Alternative strategies to circumvent Dxs have likely evolved whereas IspG and IspH are evolutionarily rigid and no alternatives appear to have evolved or can be identified at this time. DxrI/II, IspD, IspE and IspF are also evolutionarily rigid as they appear to be retained by essentially all species using the MEP pathway. Whether the evolutionary rigidity of IspG, IspH, DxrI/II, IspD, IspE and IspF is due to a functional constraint hindering biochemical ability or due to the lack of sequence diversity available for the enzyme to explore under an evolving landscape remains an open question.

The evolutionary rigidity of enzymes in the MEP pathway may be valuable for unraveling the evolution of terpenoid biosynthesis. Analysis of the protein phylogenetic trees for the MEP enzymes was consistent with the possibility that horizontal gene transfer has played a role in the evolution of these enzymes (Figure 5; Supplementary Figure S5). The role of horizontal gene transfer in extant species is highlighted in Figure 4 where within some genera bifurcation of the terpenoid biosynthesis strategies has occurred.

Species retaining all MEP enzymes except Dxs, are distributed across different phylogenetic groups (Figure 3). It is possible that these species acquired the same alternative to Dxs via horizontal gene transfer. Another possibility is that multiple strategies have evolved. Further work will be required to determine how these species circumvent Dxs as well as the relationship of these metabolic routes to each other. Possible explanations for the lack of Dxs include incomplete sequencing, promiscuous activity of related enzymes at levels adequate to support loss of Dxs, or an alternative enzyme evolved that is functionally equivalent to Dxs which should be sought.

A limitation of our bioinformatic approach is that it will only identify alternative MEP pathways in species lacking redundancy. Species with a complete MEP or MVA pathway in addition to an alternative non-canonical MEP pathway would not be identified. Additionally, we only surveyed the current database of bacterial species, which is inherently biased toward culturable bacteria and does not capture the full diversity of life (Hugenholtz and Pace, 1996; Hug et al., 2016). It is possible that in organisms not yet sequenced, alternative MEP pathways may exist. Finally, we note that sequence-based identification of orthologs could miss functional enzymes that lack sequence similarity. Regardless, our results provide valuable insights into MEP pathway constraints and identify a set of species

that are promising candidates for an alternative MEP pathway that circumvents Dxs.

Our findings also inform metabolic engineering of the MEP pathway in two ways. First regarding Dxs, we find that alternative enzymes to Dxs may exist. As Dxs governs MEP pathway flux via feedback regulation (Banerjee et al., 2013) and requires pyruvate and glyceraldehyde-3-phosphate as the MEP pathway substrates, it would be of value to metabolic engineers to identify Dxs alternatives. An alternative enzyme to Dxs may consume different substrates or avoid regulation. For IspG and IspH, ongoing efforts to engineer these enzymes to increase their flux are justified, as no alternatives to their role in the MEP pathway were apparent.

The development of sequencing technologies has enabled the determination of unprecedented numbers of bacterial genomes. This database of bacterial genomes is a repository of knowledge with astounding potential. Here, we leveraged the current repository of bacterial genomes as a resource for identifying alternative enzymes to enable engineering of a key metabolic pathway. Future efforts in microbial engineering should also utilize the database of bacterial genomes to search for naturally evolved alternatives for cellular metabolism that go beyond currently characterized literature.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

BM: Conceptualization, Data curation, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. KA: Software, Visualization, Writing – review & editing. MW: Software, Supervision, Writing – review & editing. BK: Conceptualization, Funding acquisition, Supervision, Writing – review & editing. RL: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1286626/full#supplementary-material

References

Ajikumar, P. K., Tyo, K., Carlsen, S., Mucha, O., Phon, T. H., and Stephanopoulos, G. (2008). Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Mol. Pharm.* 5, 167–190. doi: 10.1021/mp700151b

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). Kofam KOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36, 2251–2252. doi: 10.1093/bioinformatics/btz859

Avalos, M., Garbeva, P., Vader, L., Van Wezel, G. P., Dickschat, J. S., and Ulanova, D. (2022). Biosynthesis, evolution and ecology of microbial terpenoids. *Nat. Prod. Rep.* 39, 249–272. doi: 10.1039/D1NP00047K

Banerjee, A., Wu, Y., Banerjee, R., Li, Y., Yan, H., and Sharkey, T. D. (2013). Feedback inhibition of deoxy-d-xylulose-5-phosphate synthase regulates the methylerythritol 4-phosphate pathway. *J. Biol. Chem.* 288, 16926–16936. doi: 10.1074/jbc.M113.464636

Benjamin, K. R., Silva, I. R., Cherubim, J. P., McPhee, D., and Paddon, C. J. (2016). Developing commercial production of semi-synthetic artemisinin, and of β-Farnesene, an isoprenoid produced by fermentation of Brazilian sugar. *J. Braz. Chem. Soc.* 27, 1339–1345. doi: 10.5935/0103-5053.20160119

Brocks, J. J., Logan, G. A., Buick, R., and Summons, R. E. (1999). Archean molecular fossils and the early rise of eukaryotes. *Science* 285, 1033–1036. doi: 10.1126/science.285.5430.1033

Brocks, J. J., Love, G. D., Summons, R. E., Knoll, A. H., Logan, G. A., and Bowden, S. A. (2005). Biomarker evidence for green and purple sulphur bacteria in a stratified palaeoproterozoic sea. *Nature* 437, 866–870. doi: 10.1038/nature04068

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). Trim Al: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Carretero-Paulet, L., Lipska, A., Pérez-Gil, J., Sangari, F. J., Albert, V. A., and Rodríguez-Concepción, M. (2013). Evolutionary diversification and characterization of the eubacterial gene family encoding DXR type II, an alternative isoprenoid biosynthetic enzyme. *BMC Evol. Biol.* 13:180. doi: 10.1186/1471-2148-13-180

Eberl, M., Hintz, M., Jamba, Z., Beck, E., Jomaa, H., and Christiansen, G. (2004). *Mycoplasma penetrans* is capable of activating Vγ9/Vδ2 T cells while other human pathogenic mycoplasmas fail to do so. *Infect. Immun.* 72, 4881–4883. doi: 10.1128/IAI.72.8.4881-4883.2004

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195

Erb, T. J., Evans, B. S., Cho, K., Warlick, B. P., Sriram, J., Wood, B. M., et al. (2012). A rubis CO-like protein links SAM metabolism with isoprenoid biosynthesis. *Nat. Chem. Biol.* 8, 926–932. doi: 10.1038/nchembio.1087

Farris, J. S. (1972). Estimating phylogenetic trees from distance matrices. *Am. Nat.* 106, 645–668. doi: 10.1086/282802

Faylo, J. (2021). Assembly-line catalysis in bifunctional terpene synthases. *Acc. Chem. Res.* 54, 3780–3791. doi: 10.1021/acs.accounts.1c00296

George, K. W., Alonso-Gutierrez, J., Keasling, J. D., and Lee, T. S. (2015). Isoprenoid drugs, biofuels, and chemicals--artemisinin, farnesene, and beyond. *Adv. Biochem. Eng. Biotechnol.* 148, 355–389. doi: 10.1007/10_2014_288

Harker, M., and Bramley, P. M. (1999). Expression of prokaryotic 1-deoxy-d-xylulose-5-phosphatases in *Escherichia coli* increases carotenoid and ubiquinone biosynthesis. *FEBS Lett.* 448, 115–119. doi: 10.1016/S0014-5793(99)00360-9

Harvey, B. G., Wright, M. E., and Quintana, R. L. (2010). High-density renewable fuels based on the selective dimerization of Pinenes. *Energy Fuel* 24, 267–273. doi: 10.1021/ef900799c

Hayakawa, H., Motoyama, K., Sobue, F., Ito, T., Kawaide, H., Yoshimura, T., et al. (2018). Modified mevalonate pathway of the archaeon *Aeropyrum pernix* proceeds via trans-anhydromevalonate 5-phosphate. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10034–10039. doi: 10.1073/pnas.1809154115

Hillier, S. G., and Lathe, R. (2019). Terpenes, hormones and life: isoprene rule revisited. *J. Endocrinol.* 242, R9–R22. doi: 10.1530/JOE-19-0084

Holstein, S. A., and Hohl, R. J. (2004). Isoprenoids: remarkable diversity of form and function. *Lipids* 39, 293–309. doi: 10.1007/s11745-004-1233-3

Hoshino, Y., and Villanueva, L. (2023). Four billion years of microbial terpenome evolution. *FEMS Microbiol. Rev.* 47:fuad008. doi: 10.1093/femsre/fuad008

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree of life. *Nat. Microbiol.* 1, 1–6. doi: 10.1038/nmicrobiol.2016.48

Hugenholtz, P., and Pace, N. R. (1996). Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. *Trends Biotechnol.* 14, 190–197. doi: 10.1016/0167-7799(96)10025-1

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., Misawa, K., Kuma, K. I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/NAR/GKF436

Khana, D. B., Tatli, M., Rivera Vazquez, J., Weraduwage, S. M., Stern, N., Hebert, A. S., et al. (2023). Systematic analysis of metabolic bottlenecks in the Methylerythritol 4-phosphate (MEP) pathway of *Zymomonas mobilis*. *mSystems* 8:e0009223. doi: 10.1128/msystems.00092-23

Kirby, J., Nishimoto, M., Chow, R. W. N., Baidoo, E. E. K., Wang, G., Martin, J., et al. (2015). Enhancing terpene yield from sugars via novel routes to 1-deoxy-d-Xylulose 5-phosphate. *Appl. Environ. Microbiol.* 81, 130–138. doi: 10.1128/AEM.02920-14

Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci.* 97, 13172–13177. doi: 10.1073/pnas.240454797

Leavell, M. D., McPhee, D. J., and Paddon, C. J. (2016). Developing fermentative terpenoid production for commercial usage. *Curr. Opin. Biotechnol.* 37, 114–119. doi: 10.1016/j.copbio.2015.10.007

Li, Q., Fan, F., Gao, X., Yang, C., Bi, C., Tang, J., et al. (2017). Balanced activation of IspG and IspH to eliminate MEP intermediate accumulation and improve isoprenoids production in *Escherichia coli*. *Metab. Eng.* 44, 13–21. doi: 10.1016/j.ymben.2017.08.005

Li, Y., and Wang, G. (2016). Strategies of isoprenoids production in engineered bacteria. *J. Appl. Microbiol.* 121, 932–940. doi: 10.1111/jam.13237

Liu, Y.-L., Guerra, F., Wang, K., Wang, W., Li, J., Huang, C., et al. (2012). Structure, function and inhibition of the two- and three-domain 4Fe-4S IspG proteins. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8558–8563. doi: 10.1073/pnas.1121107109

Liu, H., Wang, Y., Tang, Q., Kong, W., Chung, W.-J., and Lu, T. (2014). MEP pathway-mediated isopentenol production in metabolically engineered *Escherichia coli*. *Microb. Cell Factories* 13:135. doi: 10.1186/s12934-014-0135-y

Lombard, J., and Moreira, D. (2011). Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol. Biol. Evol.* 28, 87–99. doi: 10.1093/molbev/msq177

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Misic, A. M., Cain, C. L., Morris, D. O., Rankin, S. C., and Beiting, D. P. (2016). Divergent isoprenoid biosynthesis pathways in Staphylococcus species constitute a drug target for treating infections in companion animals. *mSphere* 1, e00258–e00216. doi: 10.1128/msphere.00258-16

Navale, G. R., Dharne, M. S., and Shinde, S. S. (2021). Metabolic engineering and synthetic biology for isoprenoid production in Escherichia coli and *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 105, 457–475. doi: 10.1007/s00253-020-11040-w

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/MOLBEV/MSU300

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189

Ourisson, G., and Albrecht, P. (1992). Hopanoids. 1. Geohopanoids: the most abundant natural products on earth? *Acc. Chem. Res.* 25, 398–402. doi: 10.1021/ar00021a003

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996–1004. doi: 10.1038/nbt.4229

Perez-Gil, J., Uros, E. M., Sauret-Güeto, S., Lois, L. M., Kirby, J., Nishimoto, M., et al. (2012). Mutations in *Escherichia coli* aceE and ribB genes allow survival of strains defective in the first step of the isoprenoid biosynthesis pathway. *PLoS One* 7:e43775. doi: 10.1371/journal.pone.0043775

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204. doi: 10.1093/nar/gky448

Rambaut, A., (2018). *FigTree v1.4.4*. Institute of Evolutionary Biology University of Edinburgh. Available at: http://tree.bio.ed.ac.uk/software/figtree/ Accessed July 8.2023.

Rohmer, M. (2008). From molecular fossils of bacterial Hopanoids to the formation of isoprene units: discovery and elucidation of the methylerythritol phosphate pathway. *Lipids* 43, 1095–1107. doi: 10.1007/s11745-008-3261-7

Rude, M. A., and Schirmer, A. (2009). New microbial fuels: a biotech perspective. *Curr. Opin. Microbiol.* 12, 274–281. doi: 10.1016/j.mib.2009.04.004

Sangari, F. J., Pérez-Gil, J., Carretero-Paulet, L., García-Lobo, J. M., and Rodríguez-Concepción, M. (2010). A new family of enzymes catalyzing the first committed step of the methylerythritol 4-phosphate (MEP) pathway for isoprenoid biosynthesis in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14081–14086. doi: 10.1073/pnas.1001962107

Sauret-Güeto, S., Urós, E. M., Ibáñez, E., Boronat, A., and Rodríguez-Concepción, M. (2006). A mutant pyruvate dehydrogenase E1 subunit allows survival of *Escherichia coli* strains defective in 1-deoxy-d-xylulose 5-phosphate synthase. *FEBS Lett.* 580, 736–740. doi: 10.1016/j.febslet.2005.12.092

Sell, C. S. (2006). "Terpenoids" in *Kirk-Othmer encyclopedia of chemical technology* (John Wiley and Sons, Ltd.).

Singh, U., and Wurtele, E. S. (2021). Orfipy: a fast and flexible tool for extracting ORFs. *Bioinformatics* 37, 3019–3020. doi: 10.1093/bioinformatics/btab090

Summons, R. E., Bradley, A. S., Jahnke, L. L., and Waldbauer, J. R. (2006). Steroids, triterpenoids and molecular oxygen. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 361, 951–968. doi: 10.1098/rstb.2006.1837

Summons, R. E., Welander, P. V., and Gold, D. A. (2022). Lipid biomarkers: molecular tools for illuminating the history of microbial life. *Nat. Rev. Microbiol.* 20, 174–185. doi: 10.1038/s41579-021-00636-2

Tracy, N. I., Chen, D., Crunkleton, D. W., and Price, G. L. (2009). Hydrogenated monoterpenes as diesel fuel additives. *Fuel* 88, 2238–2240. doi: 10.1016/j.fuel.2009.02.002

Volke, D. C., Rohwer, J., Fischer, R., and Jennewein, S. (2019). Investigation of the methylerythritol 4-phosphate pathway for microbial terpenoid production through metabolic control analysis. *Microb. Cell Factories* 18:192. doi: 10.1186/s12934-019-1235-5

Wagner, M., and Horn, M. (2006). The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 17, 241–249. doi: 10.1016/j.copbio.2006.05.005

Wang, W., and Oldfield, E. (2014). Bioorganometallic chemistry with IspG and IspH: structure, function, and inhibition of the [Fe(4)S(4)] proteins involved in isoprenoid biosynthesis. *Angew. Chem. Int. Ed. Engl.* 53, 4294–4310. doi: 10.1002/anie.201306712

Warlick, B. P., Evans, B. S., Erb, T. J., Ramagopal, U. A., Sriram, J., Imker, H. J., et al. (2012). 1-methylthio-D-xylulose 5-phosphate methylsulfurylase: a novel route to 1-deoxy-D-xylulose 5-phosphate in *Rhodospirillum rubrum*. *Biochemistry* 51, 8324–8326. doi: 10.1021/bi301215g

Whited, G. M., Feher, F. J., Benko, D. A., Cervin, M. A., Chotani, G. K., McAuliffe, J. C., et al. (2010). TECHNOLOGY UPDATE: development of a gas-phase bioprocess for isoprene-monomer production using metabolic pathway engineering. *Ind. Biotechnol.* 6, 152–163. doi: 10.1089/ind.2010.6.152

Xu, X. M., and Møller, S. G. (2011). Iron–Sulfur Clusters: Biogenesis, Molecular Mechanisms, and Their Functional Significance. *Antioxid. Redox Signal.* 15, 271–307. doi: 10.1089/ars.2010.3259

Zeng, L., and Dehesh, K. (2021). The eukaryotic MEP-pathway genes are evolutionarily conserved and originated from Chlaymidia and cyanobacteria. *BMC Genomics* 22:137. doi: 10.1186/s12864-021-07448-x

Zhao, L., Chang, W., Xiao, Y., Liu, H., and Liu, P. (2013). Methylerythritol phosphate pathway of isoprenoid biosynthesis. *Annu. Rev. Biochem.* 82, 497–530. doi: 10.1146/annurev-biochem-052010-100934

Zhou, K., Zou, R., Stephanopoulos, G., and Too, H.-P. (2012). Metabolite profiling identified Methylerythritol Cyclodiphosphate efflux as a limiting step in microbial isoprenoid production. *PLoS One* 7:e47513. doi: 10.1371/journal.pone.0047513