Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes

B. Markus Lange*, Tamas Rujan†, William Martin†, and Rodney Croteau*‡

*Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340; and †Institut für Botanik III, Heinrich Heine Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany

Contributed by Rodney Croteau, September 22, 2000

Isopentenyl diphosphate (IPP) is the central intermediate in the biosynthesis of isoprenoids, the most ancient and diverse class of natural products. Two distinct routes of IPP biosynthesis occur in nature: the mevalonate pathway and the recently discovered deoxyxylulose 5-phosphate (DXP) pathway. The evolutionary history of the enzymes involved in both routes and the phylogenetic distribution of their genes across genomes suggest that the mevalonate pathway is germane to archaebacteria, that the DXP pathway is germane to eubacteria, and that eukaryotes have inherited their genes for IPP biosynthesis from prokaryotes. The occurrence of genes specific to the DXP pathway is restricted to plastid-bearing eukaryotes, indicating that these genes were acquired from the cyanobacterial ancestor of plastids. However, the individual phylogenies of these genes, with only one exception, do not provide evidence for a specific affinity between the plant genes and their cyanobacterial homologues. The results suggest that lateral gene transfer between eubacteria subsequent to the origin of plastids has played a major role in the evolution of this pathway.

chloroplast | deoxyxylulose 5-phosphate | endosymbiosis | mevalonate | phylogeny

Isoprenoids are the oldest known biomolecules, with hopanoids (membrane-associated triterpenoid derivatives) having been recovered from sediments as old as 2.5 billion years (1, 2). The isoprenoids are also the largest group of contemporary natural products, encompassing over 30,000 known compounds (3), and they serve numerous biochemical functions: as quinones in electron transport chains, as components of membranes (prenyllipids in archaebacteria and sterols in eubacteria and eukaryotes), in subcellular targeting and regulation (prenylation of proteins), as photosynthetic pigments (carotenoids, side chain of chlorophyll), as hormones (gibberellins, brassinosteroids, abscisic acid), and as plant defense compounds (monoterpenes, sesquiterpenes, diterpenes). Although isoprenoids are synthesized ubiquitously among eubacteria, archaebacteria and eukaryotes through condensations of the five-carbon compound isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate, two distinct and independent biosynthetic routes to IPP exist. The pathway to IPP in mammals and yeast starts from acetyl-CoA, proceeds through the intermediate mevalonic acid (MVA), and was previously thought to be ubiquitous in all organisms (4). More recently, eubacterial hopanoids and plastid-associated isoprenoids of algae and higher plants were found to derive from IPP that is synthesized by the condensation of pyruvate and glyceraldehyde-3-phosphate, via 1-deoxyxylulose-5-phosphate (DXP) as the first intermediate (5–8) (Fig. 1). The antiquity of isoprenoids and the disparity of their underlying biosynthetic routes suggest that the evolutionary history of these pathways may shed light on early cell evolution. We have investigated the occurrence and deduced evolution of genes and enzymes that constitute these pathways from prokaryotic to eukaryotic genomes.

Materials and Methods

Sequences for translated ORFs from genome projects and data from expressed sequence tag projects were extracted from selected websites (http://www.ncbi.nlm.nih.gov; http://sanger.ac.uk; http://www.tigr.org; http://www.arabidopsis.org; and http://rgp.dna.affrc.go.jp). Similarity searches were performed by using the BLAST (9), GAPPED-BLAST (10) and PSI-BLAST (10) algorithms and were retrieved from GenBank (http://www.ncbi.nlm.nih.gov). Genes were scored as putative homologues for e-values of $\leq 10^{-4}$ when compared with genes of established biochemical function.

Translated amino acid sequences were aligned by using PILEUP of the GCG package [Wisconsin Package Version 10.0, Genetics Computer Group, Madison, WI]. Alignments are available from the authors on request or from http://ibc.wsu.edu/faculty/rc.html. Phylogenies were inferred by using PROTML (11).

Results

Occurrence and Compartmentation of Isoprenoid Biosynthetic Pathways. The distribution of genes involved in isoprenoid biosynthesis across 35 genomes is summarized in supplementary Table 1 (which is published as supplemental data on the PNAS web site, www.pnas.org). In the six sequenced archaebacterial genomes, genes for the MVA pathway, but not for the DXP pathway, are found. The archaebacteria share a unique cell membrane composed of saturated isoprenoid side chains attached to a glycerol phosphate backbone by ether linkages (12, 13). This membrane composition is in contrast to eubacteria and eukaryotes, the membranes of which consist primarily of glycerol esters of fatty acids, which are not derived from IPP, although sterols derived from IPP are present. To define the origin of their isoprenoids, two archaebacteria (*Caldariella acidophilus* and *Halobacterium cutirubrum*) have been subjected to biosynthetic labeling experiments and were shown to use the MVA pathway (14, 15).

The genomes of the free-living eubacteria that are included in supplementary Table 1 possess genes of the DXP pathway, and related biosynthetic studies have established that the overwhelming majority of eubacteria exclusively use the DXP pathway for isoprenoid biosynthesis (16). Exceptions are the δ-proteobacterium *Myxococcus fulvus* (17) and the phototrophic eubacterium *Chloroflexus aurantiacus* (18), which both use the MVA pathway. The obligate parasitic eubacteria *Rickettsia prowazekii*, *Mycoplasma genitalium*, and *Borrelia burgdorferi* lack a complete DXP pathway and possess rather unusual distributions of enzymes of isoprenoid metabolism. *Rickettsia* lacks genes for IPP synthesis

Abbreviations: AACT, acetoacetyl-CoA thiolase; CMK, 4-(cytidine 5′-diphospho)-2-*C*-methylerythritol kinase; DXP, deoxyxylulose 5-phosphate; DXPS, deoxyxylulose 5-phosphate synthase; DXR, deoxyxylulose 5-phosphate reductoisomerase; HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; IPP, isopentenyl diphosphate; MCT, 2-*C*-methylerythritol 4-phosphate cytidyl transferase; MECPS, 2-*C*-methylerythritol 2,4-cyclodiphosphate synthase; MK, mevalonate kinase; MPDC, mevalonate 5-diphosphate decarboxylase; PMK, phosphomevalonate kinase; MVA, mevalonic acid.

‡To whom reprint requests should be addressed. E-mail: croteau@mail.wsu.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.240454797. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.240454797

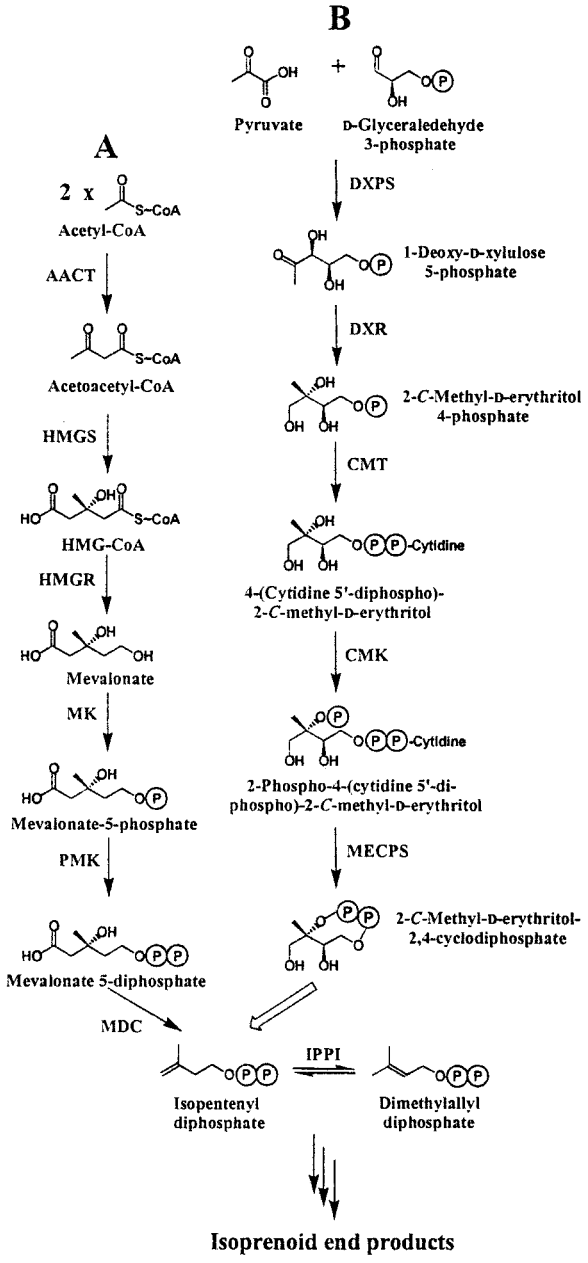

Fig. 1. Biosynthesis of IPP via the mevalonate pathway (*A*) and the DXP pathway (*B*). The circled P denotes the phosphate moiety. The large open arrow indicates several as yet unidentified steps. Isopentenyl diphosphate isomerase (EC 5.3.3.2) is abbreviated as IPPI.

but possesses enzymes for condensing IPP, a metabolite that it probably obtains from host cells, for the synthesis of quinones required for obligate aerobic respiration (19). *Mycoplasma* lacks even IPP-condensing enzymes, but this bacterium is a strict anaerobe that does not possess genes involved in membrane-associated electron transport (20), including genes for quinone synthesis, consistent with its fermentative lifestyle. *Borrelia* possesses a gene cluster with detectable similarity to MVA pathway enzymes, but these apparent homologues are highly divergent from orthologues found in other genomes, and their function has not been established. A noteworthy exception to the observation that eubacteria generally use the DXP pathway, or, alternatively, the MVA pathway, is a small group of actinomycetes that apparently employ both pathways (21). In *Streptomyces* sp. strain LS190, the MVA pathway genes form a gene cluster (22), the translated peptide sequences of which more closely resemble eukaryotic MVA pathway enzyme sequences than those from archaebacteria.

In the entirely sequenced genomes of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, homologues for all genes of the MVA pathway are present, with no evidence for the occurrence of DXP pathway genes. It has been shown, by biosynthetic labeling studies, that isoprenoids of the yeast *Rhodotorula glutinis* and of four fungal species are synthesized exclusively via the MVA pathway (23–25). A central enzyme of the MVA pathway, 3-hydroxy-3-methylglutaryl-CoA reductase, has been cloned and characterized from the fungus *Gibberella fujikuroi* (26).

Animals also use the MVA pathway for the synthesis of more than a dozen classes of isoprenoids (27). Accordingly, homologues for MVA pathway genes, but not for any DXP pathway genes, are found in the human, *Caenorhabditis elegans* and *Drosophila melanogaster* genomes. In all animals studied to date, the biosynthetic pathway to cholesterol, the major end-product of MVA metabolism, is compartmentalized (28). The conversion of acetyl-CoA to 3-hydroxy-3-methylglutaryl (HMG)-CoA occurs in the cytosol and in peroxisomes, the reduction to MVA occurs both at the endoplasmatic reticulum and in peroxisomes, and the conversion of MVA to farnesyl diphosphate is predominantly, if not exclusively, localized to peroxisomes. The transformation of farnesyl diphosphate to squalene occurs at the endoplasmatic reticulum, whereas further conversions may also occur in peroxisomes. The capability of vertebrate mitochondria to convert acetyl-CoA to HMG-CoA is linked to ketogenesis, a catabolic pathway unrelated to isoprenoid biosynthesis (29).

Among photosynthetic eukaryotes, the chlorophytes *Scenedesmus obliquus*, *Chlamydomonas reinhardtii*, and *Chlorella fusca* have been shown to use exclusively the DXP pathway, whereas the rhodophyte *Cyanidium caldarum* and the heterokontophyte *Ochromonas danica* possess both the DXP pathway and the MVA pathway. *Euglena gracilis* is an exception among photosynthetic eukaryotes, in that it uses the MVA pathway for the synthesis of all of its isoprenoids (30). In higher plants, the cytosolic compartment contains all of the MVA pathway enzymes for sterol biosynthesis (31). Plastid-derived isoprenoids, however, including carotenoids, the prenyl side chains of chlorophyll and plastoquinone, as well as monoterpenes and diterpenes, are synthesized in plastids by the DXP pathway (7, 32, 33). IPP for sesquiterpene biosynthesis may be derived either from the MVA pathway (34) or from the DXP pathway (35), or may be of mixed origin (36). A peroxisomal (glyoxysomal) isoenzyme of the MVA pathway enzyme acetoacetyl-CoA thiolase (AACT) is involved in lipid degradation, which supplies the glyoxylate cycle, and, ultimately, through gluconeogenesis enables germinating seeds to convert storage triacylglycerols to glucose (37). Homologues for all known enzymes of both pathways, with only few exceptions, which are most likely due to incompletely sequenced genomes, are present in *Arabidopsis thaliana*, soybean (*Glycine max*), tomato (*Lycopersicon esculentum*), rice (*Oryza sativa*), and maize (*Zea mays*).

Phylogenetic Trees. Detailed phylogenetic analyses for the individual enzymes of the MVA and DXP pathways reveal patterns of similarity and distribution that are more complex than suggested by the simple presence or absence of these genes in the genome (Fig. 2). Biosynthetic acetoacetyl-CoA thiolase (acetyl-CoA: acetyl-CoA *C*-acetyltransferase; EC 2.3.1.9; AACT) catalyzes the first step of the MVA pathway, the condensation of two molecules of acetyl-CoA to acetoacetyl-CoA. This enzyme, which belongs to a larger family of acyl-CoA-metabolizing enzymes, provides an intermediate in the biosynthesis of membrane sterols in animals, plants, yeasts, and fungi, and of poly(3-hydroxybutyric acid), a carbon- and energy-storage com-

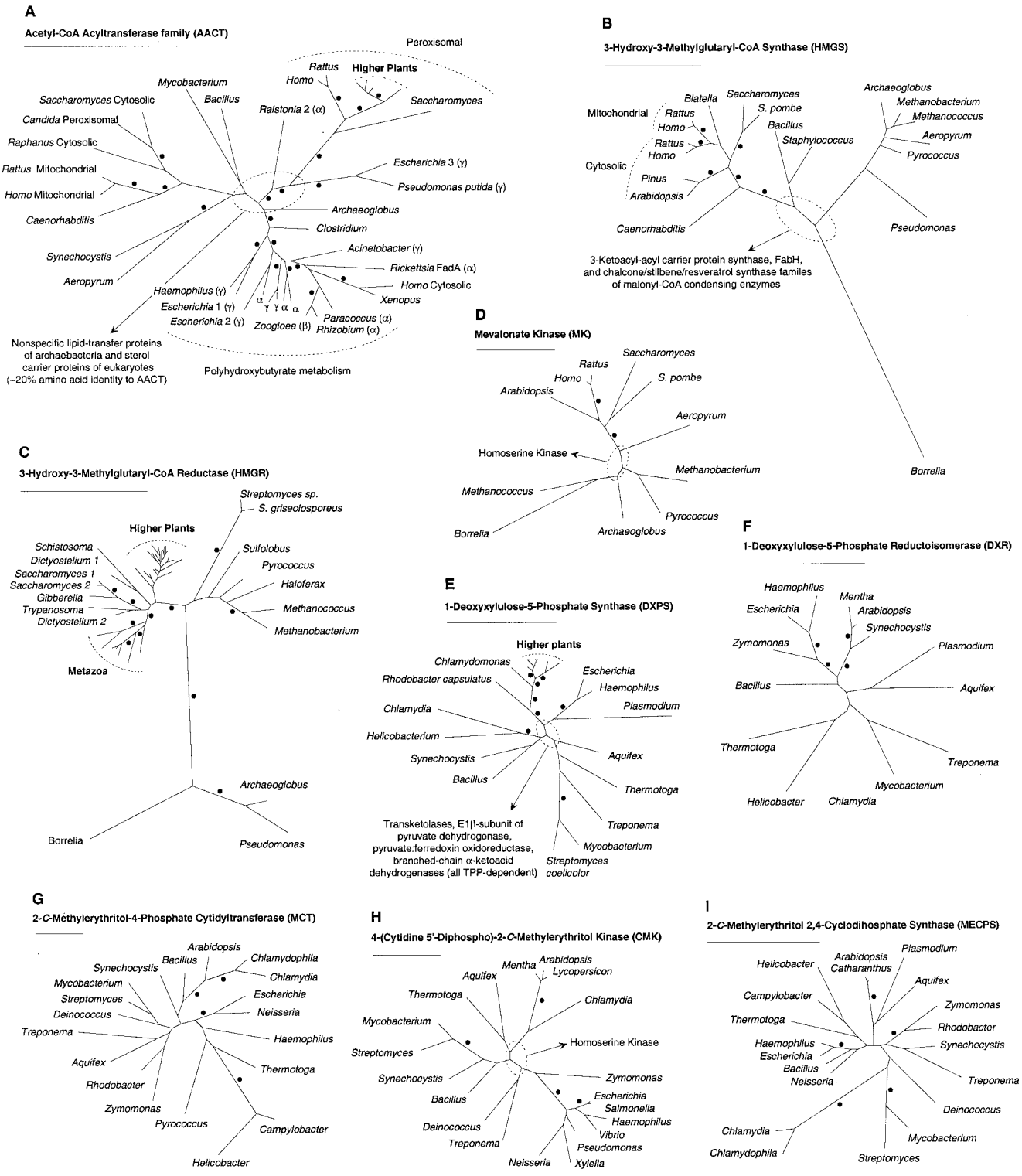

Fig. 2. Phylogenetic relationships of the enzymes of IPP biosynthesis. The trees were constructed by using the PROTML algorithm (11). The scale bar indicates 100 substitutions for each tree. Dotted ovals indicate that the sequences shown are related to other proteins, but that the positions of the branches by which the families are connected are uncertain. Branches with RELL bootstrap proportions ≥ 0.98 are indicated by a dot. Some of the genes that were detected in supplementary Table 1 are not included in the figure because of discontinuous reading frames.

pound in many eubacteria. One of the isoenzymes of this thiolase, referred to as degradative thiolase (EC 2.3.1.16), shows broad specificity for CoA-initiated thiolysis of β-ketoacyl-CoAs of chain-length from C_4 to C_{16}, and is involved in the β-oxidation of fatty acids (38, 39). A second isoenzyme has strict substrate specificity for acetoacetyl-CoA and plays a role in ketogenesis (38). Homologues with a high level of sequence similarity to the biosynthetic thiolase are not found in five of the six archaebac-

terial genomes sampled. However, these genomes do harbor distantly related proteins annotated as "hypothetical nonspecific lipid-transfer protein (acetyl CoA synthetase)," suggesting an alternative, but related, means of synthesizing acetoacetyl-CoA for the subsequent step of the MVA pathway. Distinct groups of thiolase isoenzmyes encoded in some eukaryotic nuclei (cytosolic human AACT and *Xenopus laevis* AACT) appear as tips on the branches of a tree of prokaryotic, primarily eubacterial, gene diversity (Fig. 2*A*), suggesting that they are acquisitions from eubacteria. The human cytosolic enzyme is very similar to homologues encoded in the genomes of α-proteobacteria, suggesting that this enzyme was probably acquired from the antecedents of mitochondria and was recruited for the MVA pathway, by inference from an original role in poly(3-hydroxybutyric acid) biosynthesis. The skew distribution of poly(3-hydroxybutyric acid)-related AACT genes among proteobacteria, in addition to the odd position of AACT from the β-proteobacterium *Zoogloea ramigera*, suggest that these genes have been subject to a number of horizontal transfers. The separation of *Escherichia coli* isoenzymes thiolase 1 and thiolase 3 could conceivably be attributed to ancient gene duplication events followed by massive differential loss. Human peroxisomal (degradative) thiolase has homologues in higher plants and yeast that also tend to branch with proteobacterial homologues. That thiolase from human mitochondria branches with cytosolic homologues from a plant (*Raphanus sativus*) and yeast, and with the peroxisomal enzyme of *Candida albicans*, indicates that there is no strict correlation between subcellular compartmentation and phylogeny for this enzyme, as has been observed in previous studies of pathway evolution (40).

3-Hydroxy-3-methylglutaryl-CoA synthase (EC 4.1.3.5; HMGS), which catalyzes the condensation of acetoacetyl-CoA with acetyl-CoA to yield HMG-CoA, belongs to a larger protein family comprising other acetyl-CoA condensing enzymes, such as acyl carrier protein synthase of fatty acid biosynthesis and chalcone synthase of plant phenylpropanoid metabolism. HMGS is readily detectable in several sequenced archaebacterial genomes but not, with the exception of *Borrelia* and *Streptomyces* homologues, in eubacterial genomes. However, eubacteria contain genes coding for a relative of HMGS, β-ketoacyl-ACP synthase III, which catalyzes a similar condensation reaction to produce the fatty acid precursor acetoacetyl-ACP from acetyl-ACP and malonyl-CoA as substrates. This finding suggests diversification from a common ancestral gene very early in evolution (Fig. 2*B*). The interleaving of mitochondrial and cytosolic isoforms of HMGS among eukaryotes indicates that compartment-specific isoforms have arisen relatively recently through gene duplications.

3-Hydroxy-3-methylglutaryl-CoA reductase [(S)-mevalonate: NAD^+ oxidoreductase (CoA-acylating); EC 1.1.1.34; HMGR] catalyzes the reduction of HMG-CoA to mevalonate. The carboxyl-terminal region of this enzyme, containing the active site, exhibits extensive sequence identity among different organisms. The N-terminal domain, however, is highly divergent. The significance of the divergent architecture of the N-terminal region, and the presence of multiple copies in plants, yeast, and the slime mold *Dictyostelium discoideum*, are still matters of debate (41). HMGR is frequently found among archaebacteria, but only few eubacterial genes are known to encode proteins similar to HMGR, i.e., two *Streptomyces* species, *Borrelia*, and the unclassified proteobacterium *Pseudomonas mevalonii*, in which it serves a strictly biodegradative function. The paucity of this enzyme among eubacteria and its prevalence among archaebacteria tend to suggest that the former have acquired their HMGR genes from the latter (Fig. 2*C*).

Mevalonate kinase (EC 2.7.1.36; MK), which catalyzes the phosphorylation of mevalonate at C5, is part of a larger gene family that encompasses galactokinase, homoserine kinase, mevalonate kinase, and phosphomevalonate kinase (the GHMP family) (ref. 42; for details see http://www.expasy.ch/prosite). The distribution of this gene across sequenced archaebacterial and eubacterial genomes is similar to that of HMGR, indicating that, as in the case of HMGR, *Borrelia* acquired its MK gene from archaebacteria (Fig. 2*D*). The last two steps of the MVA pathway, catalyzed by phosphomevalonate kinase (EC 2.7.4.2; PMK) and mevalonate 5-diphosphate decarboxylase (EC 4.1.1.33; MDC), lead to the conversion of mevalonate phosphate to IPP. These two enzymes are poorly conserved across genomes, and too few homologues have been defined for phylogenetic analysis.

The five enzymes of the DXP pathway that have been characterized to date are ubiquitous among the genomes of free-living eubacteria evaluated thus far. 1-Deoxyxylulose-5-phosphate synthase (DXPS) catalyzes the condensation of glyceraldehyde-3-phosphate and "activated acetaldehyde" generated from pyruvate (43–46). Like transketolase (EC 2.2.1.1) and the E1 subunit of pyruvate dehydrogenase (EC 1.2.4.1), DXPS performs a two-carbon-transfer with thiamin diphosphate as a cofactor. A high level of similarity is observed in the alignments of these proteins, with 50 invariant residues and an extremely well-conserved stretch of amino acids around the cofactor-binding site. The plant enzymes tend to branch with the homologue from the α-proteobacterium *Rhodobacter capsulatus* (Fig. 2*E*). DXPS from the cyanobacterium *Synechocystis* tends to branch with the homologue from *Bacillus subtilis*. Enzymatic activity and a cDNA for DXPS have been detected in the causal agent of malaria, the apicomplexan *Plasmodium falciparum*, and this sequence bears an N-terminal extension, suggesting that it might be localized to the apicoplast (47). The tree location of *Plasmodium* DXPS indicates a eubacterial origin, but the long branch bearing this sequence suggests that its position is unstable.

1-Deoxyxylulose-5-phosphate reductoisomerase (DXR) catalyzes the rearrangement and subsequent reduction of DXP to 2-*C*-methylerythritol-4-phosphate (MEP) (48, 49). Like DXPS, DXR is very common among sequenced eubacterial genomes but is not detectable in archaebacterial genomes. The plant enzymes share the greatest similarity with the homologue from *Synechocystis*, providing a reasonably straightforward argument that this nuclear encoded enzyme was acquired through gene transfer to the nucleus in the process of the endosymbiotic origin of plastids (Fig. 2*F*). As in the case of DXPS, the *Plasmodium* gene appears to be an acquisition from eubacteria but does not branch specifically with the plant homologues.

MEP is conjugated with CDP by MEP cytidyltransferase (MCT) to form 4-(cytidine 5′-diphospho)-2-*C*-methylerythritol (50–52). MCT sequences share a noticeable sequence homology with other pyrophosphorylases. The MCT gene occurs in only one archaebacterial genome studied to date, that of *Pyrococcus horikoshii*, where it is the sole representative of the typically eubacterial DXP pathway (see supplementary Table 1), strongly suggesting a lateral transfer from eubacteria. The only full-length eukaryotic homologue available, that from *Arabidopsis*, branches close to its cyanobacterial counterpart, which would be consistent with a cyanobacterial origin of the plant gene, but it branches even more closely to the homolgues from *Chlamydia* and *Chlamydophila* (Fig. 2*F*).

4-(Cytidine 5′-diphospho)-2-*C*-methylerythritol kinase (CMK), which catalyzes the phosphorylation of 4-(cytidine 5′-diphospho)-2-*C*-methylerythritol (53–55) is, like MK and PMK of the MVA pathway, a member of the GHMP family of metabolite kinases (42). This gene product was previously misidentified as isopentenyl monophosphate kinase, which was thought to operate as the last step of the DXP pathway (56). Homologues of CMK have been detected only in eubacteria and plastid-bearing eukaryotes. As with DXPS and DXR, the *Syn-*

echocystis CMK is most similar to its homologues from Gram-positive eubacteria. However, it shares the greatest similarity with the homologue from *Aquifex aeolicus* (Fig. 2*G*).

4-(Cytidine 5′-diphospho)-2-*C*-methylerythritol 2-phosphate, the product of the reaction catalyzed by CMK, is then converted to 2-*C*-methylerythritol 2,4-cyclodiphosphate by the action of 2-*C*-methylerythritol 2,4-cyclodiphosphate synthase (MECPS) (57, 58). No homologues of this gene were found among archaebacteria. As in the case of CMK, the plant and *Plasmodium* forms tend to branch with the homologue from the *Aquifex* genome (Fig. 2*I*).

Conclusions

At the level of gene distribution across genomes for enzymes of isoprenoid biosynthesis, the data indicate that the MVA pathway is widespread among archaebacteria. The MVA pathway thus appears to represent the ancestral pathway of IPP biosynthesis in archaebacteria, the prime function of which would appear to be the synthesis of ether-linked prenyl-lipids that constitute their plasma membrane. This suggestion is consistent with biosynthetic labeling experiments. Similarly, the data indicate that the ancestral route of IPP formation in eubacteria is the DXP pathway, which serves the biosynthesis of quinones, carotenoids, and sterols, and, additionally, produces the precursor (DXP) for the synthesis of the essential cofactors thiamin diphosphate and pyridoxal phosphate. Some enzymes from both pathways can be traced at the level of sequence similarity to larger superfamiles with similar catalytic properties (AACT, HMGS, MK, PMK, DXPS, MCT, and CMK), suggesting that several steps of these pathways share common ancestral genes that underwent functional diversification during the earliest stages of evolution. There is a discernable correlation between the presence of these pathways and some types of ecological specialization, notably in the lack of complete pathways for IPP biosynthesis in the parasitic eubacteria *Rickettsia* and *Mycoplasma*, which are able to obtain this intermediate from their hosts.

At the level of individual gene phylogenies, patterns of sequence similarity for IPP biosynthetic genes are complex, especially for the DXP pathway (Fig. 2 *E–I*). Taken strictly at face value, the phylogenies of the currently available sequence sample would suggest that plants have assembled the DXP pathway through lateral acquisitions from several independent eubacterial sources, including α-proteobacteria (DXPS), cyanobacteria (DXR), chlamydias (MCT and CMK), and *Aquifex* (MECPS). This simple interpretation is unlikely to be correct for two reasons. First, the phylogenies for eubacterial DXP pathway genes neither resemble rRNA systematics for the same species, nor do they strongly resemble one another. This lack of internal phylogenetic consistency is most easily attributed to two well-known factors, the limited degree of phylogenetic resolution that can be achieved with individual proteins (59) and lateral gene transfer between prokaryotes (60). Second, there is strong evidence that many plant nuclear genes are acquisitions from cyanobacteria, having been transferred to the nucleus subsequent to the origins of plastids (59, 61). The lack of DXP genes in non-plastid-bearing eukaryotes suggests that plants acquired these genes from the cyanobacterial ancestor of plastids (62). Given these considerations, the finding that four of the five known plant DXP pathway enzymes (except DXR) do not branch with their cyanobacterial homologues suggests that lateral transfer of DXP pathway genes between eubacteria has occurred subsequent to the origin of plastids (40).

Overall, horizontal gene transfer appears to have contributed substantially to the distribution across prokaryotic genomes of genes for IPP biosynthesis. The individual phylogenies, and the skew and highly sporadic distribution of genes of the MVA pathway among eubacteria (see supplementary Table 1), provide evidence in support of this conclusion. Taken together, these findings suggest that selection for maintenance of isoprenoid biosynthesis acts at the level of the pathway as a whole, rather than at the level of individual genes, which apparently are easily exchanged. For all enzymes of the MVA and DXP pathways, the eukaryotic homologues tend to constitute a distinct and specific subset of prokaryotic gene diversity, indicating that eukaryotes inherited these genes from prokaryotes.

The evolution of a genome is the sum of the evolutionary histories of the individual genes encoded therein. The distribution and case-by-case phylogeny of genes for isoprenoid biosynthesis suggest that, within isoprenoid biosynthetic pathways, individual enzymes are easily replaced by intruders, particularly in prokaryotes. When gene transfer between organisms occurs, it can confer new combinations of functions that are selectable. Between the level of individual genes and complete genomes, biochemical pathways may emerge as intermediate units of function on which selection acts, independent of the evolutionary histories of individual, functionally equivalent enzymes that catalyze the steps of the pathway.

Note Added in Proof. A paper (63) has recently appeared that surveys the distribution of genes for the DXP and MVA pathways across a number of completely and partially sequenced eubacterial genomes, the phylogeny of HMGR genes, and the biochemical evidence for the distribution of these pathways among eubacteria. The salient conclusion of this paper, that lateral gene transfer has played a substantial role in the evolution of genes for isopentenyl disphosphate biosynthetic pathways, is in agreement with the findings and conclusions presented here and in the Supplementary Material.

This investigation was supported by a grant from the U.S. Department of Energy.

1. Summons, R. E., Jahnke, L. L., Hope, J. M. & Logan, G. A. (1999) *Nature (London)* **400,** 554–557.
2. Brocks, J. J., Logan, G. A., Buick, R. & Summons, R. E. (1999) *Science* **285,** 1033–1036.
3. Buckingham, J. (1998) in *Dictionary of Natural Products on CD-ROM* (Chapman & Hall, London), Version 6.1.
4. Spurgeon, S. L. & Porter, J. W. (1981) in *Biosynthesis of Isoprenoid Compounds*, eds. Porter, J. W. & Spurgeon, S. L., (Wiley, New York), Vol. 1, pp. 1–46.
5. Rohmer, M., Knani, M., Simonin, P., Sutter, B. & Sahm, H. (1993) *Biochem. J.* **295,** 517–524.
6. Broers, S. T. J. (1994) Ph.D. thesis (Eidgenössische Technische Hochschule, Zürich, Switzerland).
7. Schwarz, M. C. (1994) Ph.D. thesis (Eidgenössische Technische Hochschule, Zürich, Switzerland).
8. Schwender, J., Seemann, M., Lichtenthaler, H. K. & Rohmer, M. (1996) *Biochem. J.* **316,** 73–80.
9. Altschul, S. F., Gis, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
10. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
11. Adachi, J. & Hasegawa, M. (1996) in *Computer Science Monographs* (Institute of Statistical Mathematics, Tokyo), No. 28, Version 2.3.
12. Heathcock, C. H., Finkelstein, B. L., Aoki, T. & Poulter, C. D. (1985) *Science* **229,** 862–864.
13. Zillig, W. (1991) *Curr. Opin. Genet. Dev.* **1,** 544–551.
14. de Rosa, M., Gambacorta, A. & Nicolaus, B. (1980) *Phytochemistry* **19,** 791–793.
15. Horbach, S., Sahm, H. & Welle, R. (1993) *FEMS Microbiol. Lett.* **111,** 135–140.
16. Eisenreich, W., Schwarz, M., Cartayrade, A., Arigoni, D., Zenk, M. H. & Bacher, A. (1998) *Chem. Biol.* **5,** R211–R233.
17. Rosa Putra, S., Disch, A., Bravo, J.-M. & Rohmer, M. (1998) *FEMS Microbiol. Lett.* **164,** 169–175.
18. Rieder, C., Strauss, G., Fuchs, G., Arigoni, D., Bacher, A. & Eisenreich, W. (1998) *J. Biol. Chem.* **273,** 18099–18108.
19. Andersson, S. G. E., Zomorodipour, A. & Andersson, J. O. (1998) *Nature (London)* **396,** 133–140.

20. The *Mycoplasma genitalium* Sequencing Consortium (1995) *Science* **270,** 397–403.
21. Seto, H., Watanabe, H. & Furihata, K. (1996) *Tetrahedron Lett.* **37,** 7979–7982.
22. Takagi, M., Kuzuyama, T., Takahashi, S. & Seto, H. (2000) *J. Bacteriol.* **182,** 4153–4157.
23. Disch, A. &. Rohmer, M. (1998) *FEMS Microbiol. Lett.* **168,** 201–208.
24. Wang, Y., Dreyfuss, M., Ponelle, M., Oberer, L. & Riezman, H. (1998) *Tetrahedron* **54,** 6415–6426.
25. Mühlbauer, A., Beyer, J & Steglich, W. (1998) *Tetrahedron Lett.* **39,** 5167–5170.
26. Woitek, S., Unkles, S. E. & Kinghorn, J. R. (1997) *Curr. Genet.* **31,** 38–47.
27. Goldstein, J. L. & Brown, M. S. (1990) *Science* **343,** 425–430.
28. Aboushadi, N., Engfelt, W. H., Paton, V. G. & Krisans, S. K. (1999) *J. Histochem. Cytochem.* **47,** 1127–1132.
29. Madsen, L., Garras, A., Asins, G., Serra, D., Hegardt, F. G. & Berge, R. K. (1999) *Biochem. Pharmacol.* **57,** 1011–1019.
30. Disch, A., Schwender, J., Müller, C., Lichtenthaler, H. K. & Rohmer, M. (1998) *Biochem. J.* **333,** 381–388.
31. Bach, T. J., Boronat, A., Campos, N., Ferrer, A. & Vollack, K.-U. (1999) *Crit. Rev. Biochem. Mol. Biol.* **34,** 107–122.
32. Lichtenthaler, H. K., Schwender, J., Disch, A. & Rohmer, M. (1997) *FEBS Lett.* **400,** 271–274.
33. Eisenreich, W., Sagner, S., Zenk, M. H. & Bacher, A. (1997) *Tetrahedron Lett.* **38,** 3889–3892.
34. Facchini, P. J. & Chappell, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 11088–11092.
35. McCaskill, D. & Croteau, R. (1995) *Planta* **197,** 49–56.
36. Adam, K.-P. & Zapp, J. (1998) *Phytochemistry* **48,** 953–959.
37. Ap Rees, T. (1980) in *The Biochemistry of Plants*, ed. Preiss, J. (Academic, New York), pp. 1–42.
38. Staack, H., Binstock, J. F. & Schulz, H. (1978) *J. Biol. Chem.* **253,** 1827–1831.
39. Raaka, B. M. & Lowenstein, J. M. (1979) *J. Biol. Chem.* **254,** 6755–6762.
40. Martin, W. & Schnarrenberger, K. (1997) *Curr. Genet.* **32,** 1–18.
41. Hampton, R., Dimster-Denk, D. & Rine, J. (1996) *Trends Biochem. Sci.* **21,** 140–145.
42. Tsay, Y. H. & Robinson, G. W. (1991) *Mol. Cell. Biol.* **33,** 483–492.
43. Sprenger, G. A., Schörken, U., Wiegert, T., Grolle, S., De Graaf, A. A., Taylor, A. V., Begley, T. P., Bringer-Meyer, S. & Sahm, H. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 12857–12862.
44. Lois, L. M., Campos, N., Rosa Putra, S., Danielsen, K., Rohmer, M. & Boronat, A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2105–2110.
45. Lange, B. M., Wildung, M. R., McCaskill, D. & Croteau, R. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2100–2104.
46. Bouvier, F., d'Harlingue, A., Suire, C., Backhaus, R. A. & Camara, B. (1998) *Plant Physiol.* **117,** 1423–1431.
47. Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Türbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H. K., Soldati, D. & Beck, E. (1999) *Science* **285,** 1573–1576.
48. Takahashi, S., Kuzuyama, T., Watanabe, H. & Seto, H. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9879–9884.
49. Lange, B. M. & Croteau, R. (1999) *Arch. Biochem. Biophys.* **365,** 170–174.
50. Rohdich, F., Wungsintaweekul, J., Fellermeier, M., Sagner, S., Herz, S., Kis, K., Eisenreich, W., Bacher, A. & Zenk, M. H. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11758–11763.
51. Kuzuyama, T., Takagi, M., Kaneda, K., Dairi, T. & Seto, H. (2000) *Tetrahedron Lett.* **41,** 703–706.
52. Rohdich, F., Wungsintaweekul, J., Eisenreich, W., Richter, G., Schuhr, C. A., Hecht, S., Zenk, M. H. & Bacher, A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6451–6456.
53. Lüttgen, H., Rohdich, F., Herz, S., Wungsintaweekul, J., Hecht, S., Schuhr, C. A., Fellermeier, M., Sagner, S., Zenk, M. H., Bacher, A. & Eisenreich, W. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 1062–1067.
54. Kuzuyama, T., Takagi, M., Kaneda, K., Watanabe, H., Dairi, T. & Seto, H. (2000) *Tetrahedron Lett.* **41,** 2925–2928.
55. Rohdich, F., Wungsintaweekul, J., Lüttgen, H., Fischer, M., Eisenreich, W., Schuhr, C. A., Fellermeier, M., Schramek, N., Zenk, M. H. & Bacher, A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 8251–8256.
56. Lange, B. M. & Croteau, R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13714–13719.
57. Herz, S., Wungsintaweekul, J., Schuhr, C. A., Hecht, S., Lüttgen, H., Sagner, S., Fellermeier, M., Eisenreich, W., Zenk, M. H., Bacher, A. & Rohdich, F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 2486–2490.
58. Takagi, M., Kuzuyama, T., Kaneda, K., Watanabe, H., Dairi, T. & Seto, H. (2000) *Tetrahedron Lett.* **41,** 3395–3398.
59. Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M. & Kowallik, K. V. (1998) *Nature (London)* **393,** 162–165.
60. Doolittle, W. F. (1999) *Science* **284,** 2124–2128.
61. Abdallah, F., Salamini, F. & Leister, D. (2000) *Trends Plant Sci.* **5,** 141–142.
62. Lichtenthaler, H. K. (1999) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50,** 47–65.
63. Boucher, Y. & Doolittle, W.F. (2000) *Mol. Microbiol.* **37,** 703–716.