

RESEARCH ARTICLE

The Origin and Evolution of Baeyer—Villiger Monooxygenases (BVMOs): An Ancestral Family of Flavin Monooxygenases

Maria Laura Mascotti, Walter Jesús Lapadula, Maximiliano Juri Ayub*

IMIBIO-SL CONICET, Facultad de Química Bioquímica y Farmacia, Universidad Nacional de San Luis, San Luis, Argentina

* mjuriayub@hotmail.com



CrossMark
click for updates

OPEN ACCESS

Citation: Mascotti ML, Lapadula WJ, Juri Ayub M (2015) The Origin and Evolution of Baeyer—Villiger Monooxygenases (BVMOs): An Ancestral Family of Flavin Monooxygenases. PLoS ONE 10(7): e0132689. doi:10.1371/journal.pone.0132689

Editor: Olivier Lespinet, Université Paris-Sud, FRANCE

Received: December 29, 2014

Accepted: June 18, 2015

Published: July 10, 2015

Copyright: © 2015 Mascotti et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: MJA is a member of the scientific career of the National Council for Research from Argentina (Consejo Nacional de Investigaciones Científicas y Técnicas; CONICET). MLM and WJL have post-doctoral and doctoral fellowships, respectively, from CONICET. Grants from CONICET (Grant Code PIP 0021) and the National Agency for Scientific and Technological Promotion (Agencia Nacional de Promoción Científica y Tecnológica); FONCyT (Grant Code PICT 2010-1468) to MJA are acknowledged. The funders had no role in study design, data

Abstract

The Baeyer—Villiger Monooxygenases (BVMOs) are enzymes belonging to the “Class B” of flavin monooxygenases and are capable of performing exquisite selective oxidations. These enzymes have been studied from a biotechnological perspective, but their physiological substrates and functional roles are widely unknown. Here, we investigated the origin, taxonomic distribution and evolutionary history of the BVMO genes. By using *in silico* approaches, 98 BVMO encoding genes were detected in the three domains of life: Archaea, Bacteria and Eukarya. We found evidence for the presence of these genes in Metazoa (*Hydra vulgaris*, *Oikopleura dioica* and *Adineta vaga*) and Haptophyta (*Emiliania huxleyi*) for the first time. Furthermore, a search for other “Class B” monooxygenases (flavoprotein monooxygenases –FMOs– and *N*-hydroxylating monooxygenases –NMOs) was conducted. These sequences were also found in the three domains of life. Phylogenetic analyses of all “Class B” monooxygenases revealed that NMOs and BVMOs are monophyletic, whereas FMOs form a paraphyletic group. Based on these results, we propose that BVMO genes were already present in the last universal common ancestor (LUCA) and their current taxonomic distribution is the result of differential duplication and loss of paralogous genes.

Introduction

Baeyer—Villiger monooxygenases are a class of NADPH, flavin dependent enzymes, which catalyze the insertion of an oxygen atom between a C-C bond in ketones and aldehydes, as well as the oxidization of heteroatom-containing molecules (*i.e.*: sulfides, amines and boron compounds) [1]. Extensive research has shown that these enzymes carry out the oxidation of a variety of substrates with exquisite chemo-, regio- and enantioselectivity [2]. Regarding BVMO’s physiological function, it has been demonstrated that some members of this family are involved in the synthesis of natural products, in the activation of prodrugs and in biodegradation processes. Some illustrative examples are the BVMOs from *Aspergillus fumigatus* Af293 [3], *Streptomyces argillaceus* ATCC 12956 [4, 5], *Pseudomonas* spp [6, 7], *Mycobacterium tuberculosis* [8] and *Acinetobacter radioresistens* [9]. Thus, it seems clear that BVMOs are not involved in

collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

primary cellular processes, but rather in a wide number of secondary metabolism pathways, meaning non-essential cellular processes [1, 10].

BVMOs are part of the so-called “Class B” (EC 1.14.13) monooxygenases along with flavo-protein monooxygenases (FMOs), *N*-hydroxylating monooxygenases (NMOs) and the YUCCAs [11]. All these enzymes share biochemical, mechanistic and sequence features. However, they present differences regarding substrate specificity and catalyzed reactions. FMOs have been described in all kingdoms of life. These enzymes use the electrophilic flavin C4a-hydroperoxide to oxygenate a wide range of carbon-bound nucleophilic nitrogens, sulfurs, and halides [12]. NMOs, which are more specific regarding the structure of substrates than other flavin monooxygenases, have been only detected in bacteria and fungi [13, 14]. Recently, the YUCCAs have been included as a subgroup of “Class B” monooxygenases. These are typically plant enzymes involved in the biosynthesis and metabolism of hormones and growth factors [15]. It is currently accepted that BVMOs occurrence is restricted to prokaryotic and some unicellular or modular eukaryotic organisms such as filamentous fungi [10]. Several bacterial BVMOs have been extensively studied from a catalytic perspective [16, 17]. Recently, a few fungal BVMOs have been also cloned and biochemically characterized. These are: the cycloalkane monooxygenase (CAMO) from *Cylindrocarpon radicolica* [18], the Baeyer–Villiger monooxygenase Af1 (BVMO_{Af1}) from *Aspergillus fumigatus* Af293 [19] and four BVMOs from *Aspergillus flavus* NRRL3357 [20]. Furthermore, novel BVMOs from photosynthetic organisms, the rhodophyta *Cyanidioschizon merolae* and the bryophyta *Physcomitrella patens* have been characterized [21]. Some research dealing with the functional importance of conserved residues has been performed [22–24]. However, the origin and the evolutionary history of these enzymes have not been explored in deep.

The restricted and fitful occurrence of BVMO encoding genes throughout the tree of life could be explained by two different, but not mutually exclusive, evolutionary mechanisms; (i) the horizontal gene transfer (HGT) among evolutionary distant species [25–27] or (ii) the differential duplication and loss of paralogous genes in species derived from ancestral taxa [28–30]. For non-essential genes, genetic drift is a major force that molds their evolution leading to its fixation or loss from genomes [31, 32].

In this work we have analyzed the taxonomic distribution of BVMOs across the tree of life and studied their phylogeny in the context of the whole “Class B” monooxygenases. Based on this, we have proposed an evolutionary model explaining the emergence and current distribution of BVMO genes.

Results

Distribution of BVMO encoding genes throughout the tree of life

In this work 98 BVMOs were newly identified: 69 sequences belong to eukaryotic organisms, 26 to Bacteria and 3 to Archaea. These sequences were identified by an iterative homology searching process of three consecutive steps aimed to detect both, close and remote homologues (see [Materials & Methods](#)). Distribution of sequence identity among the identified BVMOs is presented in [S1 Fig](#). All the sequences displayed the BVMO hallmarks: the two Rossmann motifs (GxGxx[G/A]) flanking the two BVMO fingerprints ([A/G]GxWxxxx[F/Y]P[G/M]xxxD and FxGxxxHxxxW[P/D] [22, 24]), confirming unambiguously the BVMO-nature of the selected sequences ([S2 Fig](#)). Recently Rebehmed *et al* described a refinement of the second BVMO fingerprint adding a Thr residue to the consensus (FxGxxxHTxxW[P/D]) [23]. We found that this residue was only partially conserved. This could be due to the fact that our dataset includes more divergent sequences.

Our research revealed a fitful taxonomic distribution, including the presence of BVMO encoding genes in unexpected taxa. All phyla of the domain Bacteria were evaluated for the presence of BVMO-encoding sequences. Already characterized BVMOs come from species belonging mainly to β and γ -Proteobacteria (*Pseudomonas* spp, *Acinetobacter* spp, *Comamonas* spp, among others) and Actinobacteria (e.g. *Rhodococcus* spp, *Mycobacterium* spp., *Gordonia* sp.). New BVMOs were found in almost all groups excepting Acidobacteria, Aquificae, Thermotogae, Fusobacteria, Chlamydiae, Chlorobi and Chloroflexi. Thus, only two or three sequences belonging to different species were selected as representatives of each phylum to be included in further phylogenetic analyses. In the domain Archaea, BVMOs were found only in the phylum Euryarchaeota, in three species belonging to the class Halobacteria, but were not found in the phyla Crenarchaeota or Korarchaeota (Fig 1A).

In the domain Eukarya, new BVMO sequences were detected in several fungal species belonging to phyla Basidiomycota and Ascomycota, as it was expected from previous reports [19, 20]. Interestingly, a high number of sequences were found in several fungal species such as *Aspergillus flavus* (25), *Aspergillus niger* (> 9) and *Aspergillus fumigatus* (8). Additionally, we found preliminary evidence *in silico* suggesting the presence of one BVMO gene in the monocot *Hordeum vulgare*. This sequence derived from a cDNA library and no matching sequences were found in the *H. vulgare* genomic database. Therefore, by performing PCR experiments, we determined that this sequence does not belong to *H. vulgare* genome (S3 Fig).

On the other hand, five BVMO encoding sequences were detected in the eukaryotic haptophyta *Emiliania huxleyi*. The presence of these sequences was checked by PCR. The paralogous sequences *Ehux1-2* and *Ehux3-4* were successfully amplified from *E. huxleyi* genomic DNA (S4 Fig). However, the sequence *Ehux5* could not be amplified, which could be related to the genomic intraspecific variation among *E. huxleyi* strains [33]. Likewise, BVMO sequences were identified in several Eumetazoa species: *Hydra vulgaris* (1) belonging to the phyla Cnidaria, and in the bilaterian animals *Oikopleura dioica* (2) and *Adineta vaga* (1) (Fig 1B). A surprising fact was the finding of four BVMO encoding genes in the draft genome of the Tibetan antelope *P. hodgsonii*. Therefore the genomic context of these genes was studied. The results strongly suggested that the BVMO-containing contigs came from bacterial contamination somewhere in the pipeline of the sequencing project [34] (see S1 File for details). Our observation is in agreement with the recent report of other bacterial contaminant sequences in the *P. hodgsonii* database [35].

In summary, we have demonstrated the presence of BVMO genes in Haptophyta and Metazoa. Moreover, during the revision process of this article, new BVMO sequences were detected in the stramenopiles *Saprolegnia diclina* (XP_008604020.1) and *S. parasitica* (KDO22379.1), the amoebozoa *Acytostelium subglobosum* (GAM26105.1) and the chlorophyta *Monoraphidium neglectum* (KIZ04091.1) (Fig 1B). Considering this new picture of BVMOs distribution, it can be stated that the presence of these enzymes is not restricted to a limited fraction of species as it was presumed before, displaying a wider dissemination along the three life domains.

Phylogenetic analysis of BVMOs family

To understand the origin and evolutionary relationships among BVMO genes from taxonomically distant organisms, a phylogenetic analysis was conducted employing the Maximum Likelihood (Fig 2) and Bayesian inference methods (S5 Fig). The obtained trees were equivalent regarding the distribution and the composition of clades. The branches were supported by high bootstrap (BS) and posterior probability (PP) values. Moreover, the addition or deletion of various sequences from the protein dataset did not affect the global topology.

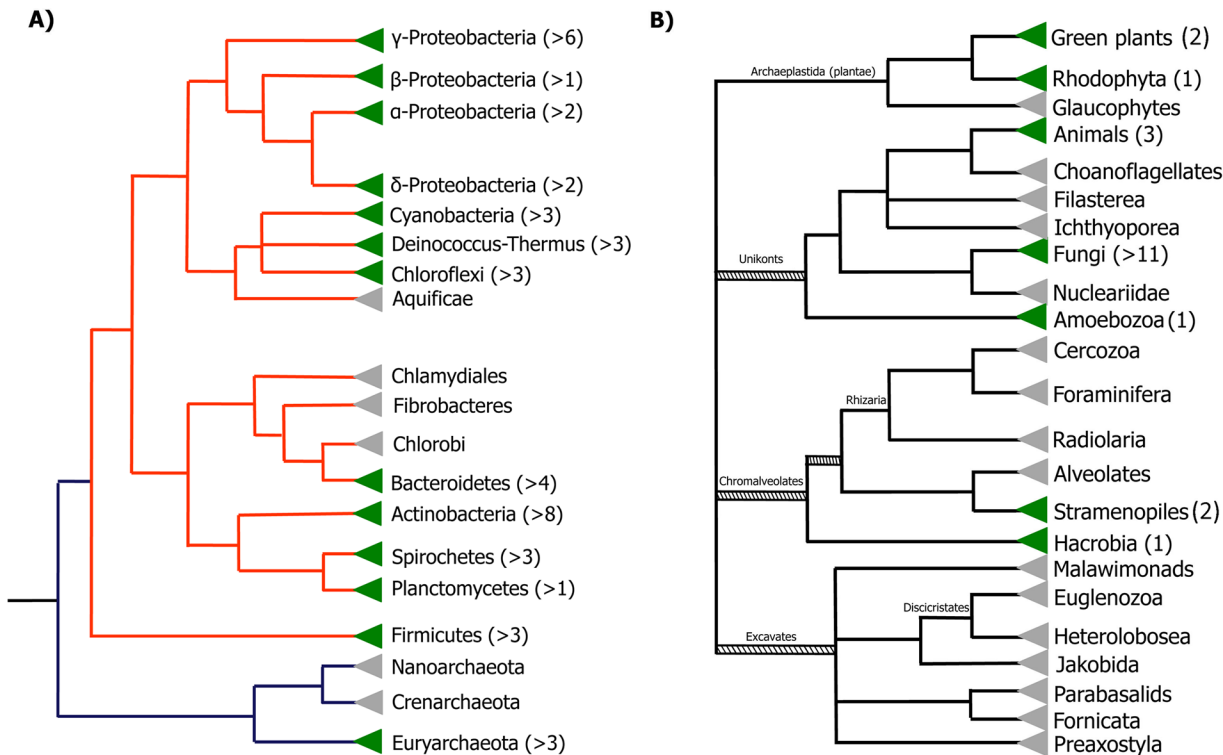


Fig 1. Taxonomic distribution of BVMO genes. **A.** Distribution of BVMO encoding genes in genomes from Bacteria and Archaea domains. **B.** Distribution of BVMO encoding genes in genomes from Eukarya domain. Green triangles show clades containing BVMOs; grey triangles show clades where BVMOs were not detected. Numbers in brackets show the number of different species where BVMOs were found. The tree topology is based on the tree of life web project (Maddison DR and Schulz KS (eds.) 2007. Available at: <http://tolweb.org> (Accessed 17 November 2014)).

doi:10.1371/journal.pone.0132689.g001

A very complex topology was noticed, being observed that closely phylogenetically related genes belong to genomes from evolutionary largely distant organisms. For instance, a clade of BVMOs with good statistic support (BS: 61, PP: 0.9582) includes sequences from gram positive (*Thermobifida fusca*, *Rhodococcus rhodochrous* and *R. palustris*) and gram negative (*Afipia* sp., *Bradyrhizobium* sp. and *Sorangium cellulosum*) bacteria, a haptophyta (*Emiliania huxleyi*), a bryophyta (*Physcomitrella patens*) and a metazoan (*Adineta vaga*). This implies that an ancestral gene gave rise to BVMO genes belonging to largely divergent species. As it can be clearly seen in the phylogenetic trees, BVMO genes from bacteria and eukaryotes are polyphyletic (Fig 2 and S5 Fig). In addition, there are many examples of organisms harboring several BVMO genes (e.g. *Yarrowia lipolitica*), whereas genomes of closely related species (e.g. *Saccharomyces cerevisiae*) are devoid of BVMO sequences, indicating the occurrence of gene loss events.

When analyzing metazoan BVMOs, it appears that in the common ancestor with fungi, meaning the Opisthokonta, at least two BVMO paralog genes were present. This can be inferred because two different clades harbor metazoan/fungal pair sequences, such as: TrkA (*Adineta vaga*)/Afla18 (*A. fumigatus*) and Odio1 (*Oikopleura dioica*)/BVMO_{Apar} (*A. parasiticus*). Likewise, in the haptophyta *Emiliania huxleyi* five different paralogs are observed, of which two pairs may have raised by recent duplications, rendering the paralogous couples *Ehux1-Ehux2* and *Ehux3-Ehux4*. In consequence in the Hacrobia ancestor at least three different BVMO genes may have been already present. This line of evidence can be extended to propose that in the common ancestor between Bacteria, Archaea and Eukarya there were at least

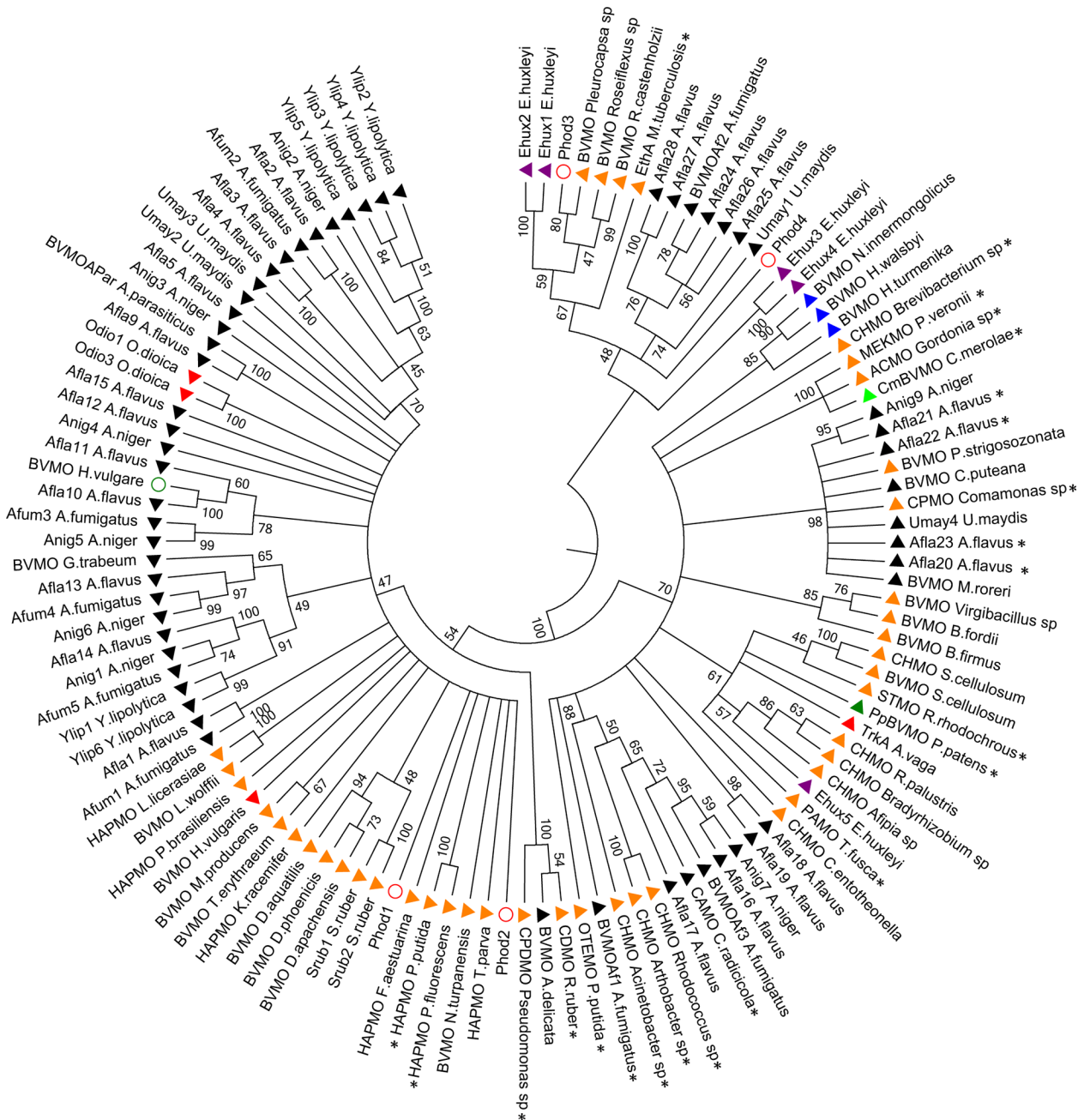


Fig 2. Phylogenetic analysis of BVMOs protein-family by Maximum Likelihood (ML) method. The tree was constructed by using the ML method, employing the alignment constructed with MAFFT 7 on-line tool and the best model parameters calculated with ProTest 3.4. Evolutionary analyses were conducted in PhyML 3.0 on-line server. Bootstrap values (> 45) are shown next to the branches. Colored triangles show the source of BVMO genes as follows: Fungi (black), Bacteria (orange), Green plants (dark green), Rhodophyta (light green), Metazoa (red), Haptophyta (purple) and Archaea (blue). Empty circles indicate BVMO encoding sequences where strong evidences of contamination artifacts were found: *Pantholops hodgsonii* (red) and *Hordeum vulgare* (dark green). Sequences marked with asterisks have been previously characterized.

doi:10.1371/journal.pone.0132689.g002

two paralogous BVMOs. Then, Bacteria and Eukarya conserved both of them. On the other hand in Archaea only one was retained, according to the current data. These observations are compatible with the occurrence of multiple gene loss and duplication events through different evolutionary lineages.

Evolutionary origin of metazoan BVMOs

The discovery of novel BVMOs encoded in metazoan genomes is a striking result. According to our analyses we propose that the origin of these genes comes from an Opisthokonta ancestor. This is also supported by the fact that a close neighbor of animals is the fungi lineage, where BVMOs are widely disseminated. It has been claimed BVMO *TrkA* from *Adineta vaga* was acquired by horizontal gene transfer from bacterial genomes [36]. We addressed this possibility by performing homology searching and genomic context analyses (See [S2 File](#) for details). We found that although *TrkA* resembles bacterial sequences, it also displays high similarity to sequences from the metazoans *Oikopleura dioica* and *Hydra vulgaris*, whose genomes have been deposited in 2010 and 2013, respectively. The phylogenetic reconstruction ([Fig 2](#)) depicts that, even when *TrkA* sequence shows high identity to bacterial proteins it is part of a clade including other eukaryotic genes (from *E. huxleyi* and *P. patens*). Moreover, a similar situation is observed when analyzing the phylogenetic relationships of the other identified BVMO genes from metazoans, e.g.: *H. vulgaris* and *O. dioica*. Therefore, the presence of a BVMO gene in *A. vaga* is an infrequent, but not unique, case among metazoans. Altogether, these data suggest that the evolutionary history of this gene involves an ancestral origin along with differential duplication and loss events through evolution, strongly challenging the hypothesis of its acquisition by HGT.

Phylogenetic relationships of BVMOs with other “Class B” flavin monooxygenases

BVMOs, along with FMOs, NMOs and the recently included YUCCAs [11], belong to the “Class B” of flavin monooxygenases. These proteins share some basic biochemical features and the presence of two Rossmann motifs on their sequences. To elucidate their evolutionary history, a survey of novel sequences in Bacteria, Archaea and Eukarya was conducted by homology searching. Then, phylogenetic analyses including the retrieved sequences as well as enzymes previously characterized at the biochemical level for each family [37, 38] were performed. “Class A” monooxygenases were included as outgroup in order to root the tree ([Fig 3](#) and [S6 Fig](#)). The topology of the tree proved to be highly robust, since by employing ML or Bayesian inference methods the clades were formed by the same sequences and the branches were supported by high BS and PP values.

The topology of the trees indicated that BVMOs and NMOs are monophyletic groups of proteins, while FMOs and YUCCAs form a paraphyletic group. The rooting of the tree using “Class A” monooxygenases as outgroup, suggests that the origin of BVMOs predates NMOs ([Fig 3](#) and [S6 Fig](#)). On the other hand, it was observed that YUCCAs are clustered with FMOs. Thus, although it has been recently proposed that YUCCAs and FMOs are different flavin monooxygenases groups [11], our results suggest that YUCCAs are part of FMOs group, in agreement with the report of Fraaije *et al* [22]. Additionally, NMOs were found in lineages other than Bacteria and Fungi [12], such as Metazoa and Archaea ([S7 Fig](#)). This is similar to the pattern observed for BVMOs and FMOs, suggesting that these three flavoproteins families (FMOs, NMOs and BVMOs) constitute ancient groups whose common ancestors must have been already present before the rising of the three life domains ([Fig 4](#) and [Discussion](#)).

Discussion

The evolution of enzyme families has been classically studied under the assumption that enzymes belonging to secondary metabolism have been originated in a late evolutionary stage “in order to cope with” specific problems of modern organisms [39, 40]. In the particular case

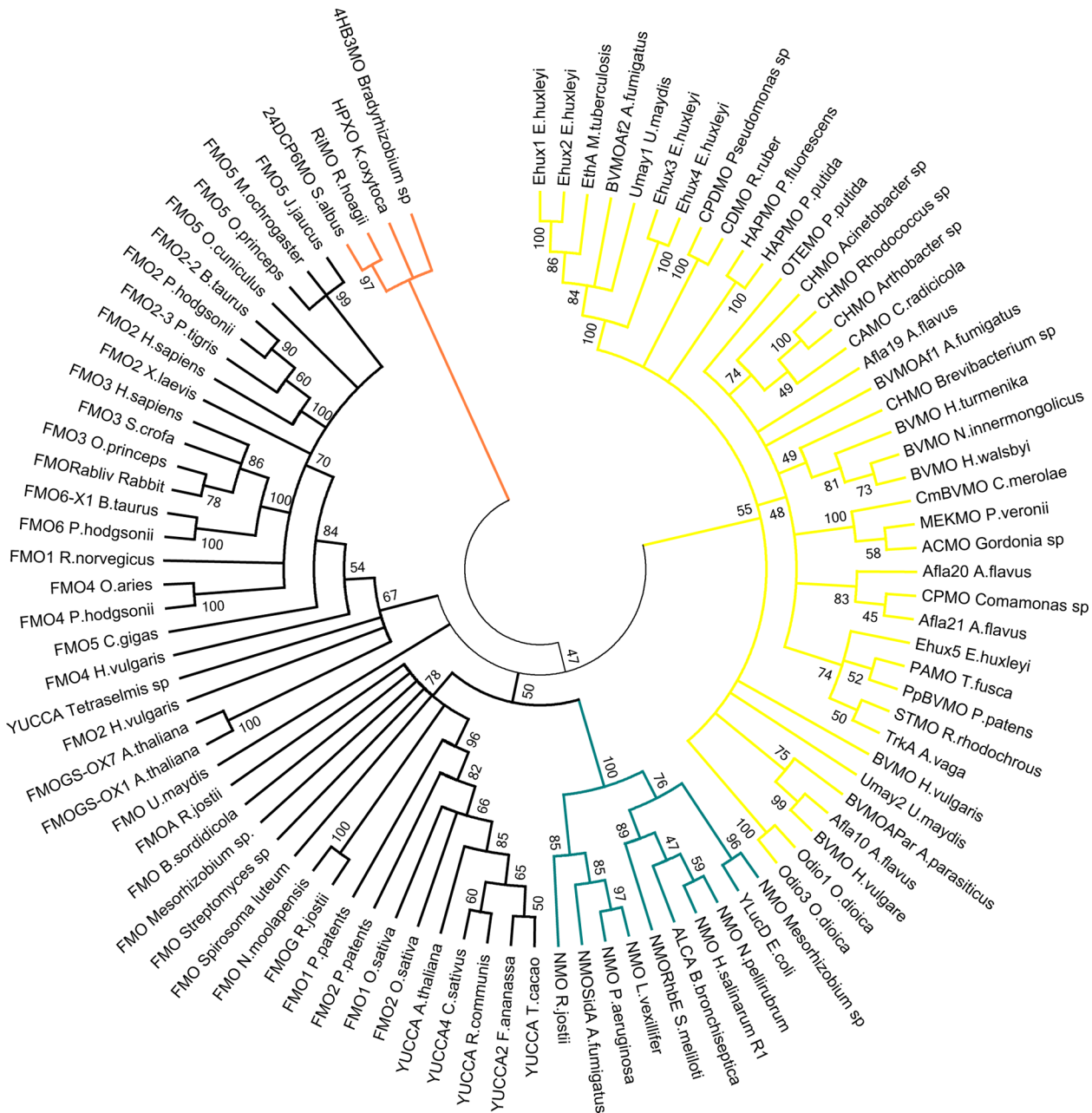


Fig 3. Rooted phylogenetic tree of flavin monooxygenases “Class B” by Maximum Likelihood (ML) method. The tree was constructed by using the ML method, employing the alignment constructed with MAFFT 7 on-line tool and the best model parameters calculated with ProTest 3.4. Evolutionary analyses were conducted in PhyML 3.0 on-line server. Bootstrap values (> 45) are shown next to the branches. Colored branches display: BVMOs (yellow), NMOs (blue) and FMOs (black). As outgroup, hydroxylases belonging to “Class A” flavin monooxygenases were employed (orange).

doi:10.1371/journal.pone.0132689.g003

of BVMOs, it has been proposed that its distribution is limited to Fungi and Bacteria and, even more, its evolution would be linked to its taxonomic distribution [20]. In contrast to this, we have demonstrated that these enzymes are widely distributed across the tree of life, finding novel sequences in Haptophyceae, Metazoa, Archaea, Stramenopiles and Amoebozoa. Therefore, it can be predicted that, as new genomic information of phylogenetically distant species

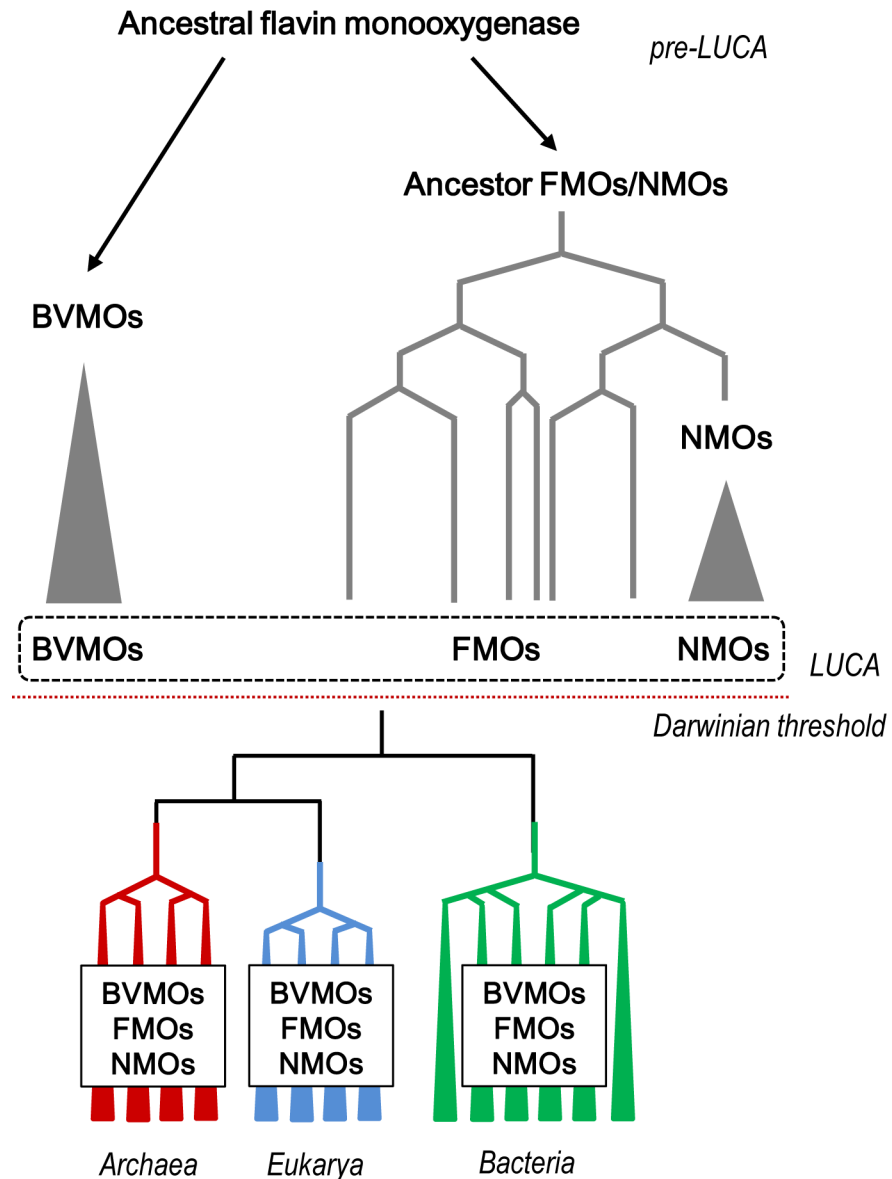


Fig 4. Scheme of the proposed model for the evolutionary history of “Class B” flavin monooxygenases. Hypothetical ancestral flavin monooxygenases genes were present in a pre-LUCA time. These genes gave origin to the monophyletic group of BVMOs and to a group of FMOs/NMOs-like genes, which later originated the monophyletic group of NMOs and a paraphyletic group of modern FMOs. We propose that BVMOs, FMOs and NMOs were already present in LUCA. After crossing the Darwinian threshold different paralogous of flavin monooxygenase genes were inherited by the three life domains: Bacteria, Archaea and Eukarya.

doi:10.1371/journal.pone.0132689.g004

becomes available, our knowledge about the distribution of BVMOs across extant organisms will be further modified.

On the basis of the evidences obtained in this work, we propose that BVMOs constitute a very ancient group of enzymes, whose evolution has been mainly shaped by genetic drift, yielding gain and loss of paralogs at different stages of the evolution of organisms. This hypothesis explains the unexpected associations among largely divergent species in the BVMOs tree. Besides, this argument takes into account that inside some lineages, BVMOs may have become

useful resources (e.g. in defense metabolite synthesis) preventing gene loss events by genetic drift during the evolution of these species. Such could be the case of the genus *Aspergillus* or *Rhodococcus*, where a significant number of BVMO encoding genes have been detected [3, 41–43].

The occurrence of HGT events requires several filters (the cellular and nuclear membranes, being incorporated to the genome, the germline barrier) to be passed. In addition, phylogenetic anomalies can be explained by HGT as well as by differential segregation of paralogous. Therefore, the real impact of HGT on the evolution of eukaryotic genomes is strongly debated [27, 44, 45]. Regarding to BVMO genes, although HGT cannot be completely excluded, the evidences found in our work suggest that the more parsimonious scenario explaining the occurrence of BVMOs in metazoa involves gene duplication and differential loss of paralogous genes. This is supported by: (i) the low sequence identity of metazoan sequences to BVMOs from other taxa thus being difficult to postulate the source organisms of HGT, (ii) the non-similar genomic contexts among species grouped in the same clade of BVMOs tree and (iii) the phylogenetic relationships among metazoan/non-metazoan BVMOs which are not consistent with HGT model.

Concerning the biochemical features, it has been largely proposed that BVMOs sharing biocatalytic properties cluster together [46, 47]. However, we have not found further evidence to support this idea. This is because only a few BVMOs possess known physiological substrates [3, 48, 49] (S3 File). Additionally, we have not found proofs suggesting that BVMOs may cluster together according to their domain architecture (S4 File).

Based on our results, we propose a model explaining the evolution of BVMOs in the context of a *last universal common ancestor* (LUCA) conceived as a genetically redundant complex community of organisms [50–52]. In a pre-LUCA time, there may have been ancestral flavin monooxygenase genes which gave origin to a monophyletic group of BVMOs and a paraphyletic group of FMOs. NMOs arose later from FMOs as a monophyletic group and, until LUCA time, FMOs remained as a paraphyletic group. After crossing the “Darwinian threshold” and the consequent crystallization of the three life domains, multiple paralogous genes of FMOs, NMOs and BVMOs were inherited by each domain (Fig 4). Although we do not know how refined these ancestral enzymes were, based on the biochemical features of modern FMOs, BVMOs and NMOs, it is tempting to hypothesize that FMOs are widespread among organisms due to their wider reaction scope as well as substrate acceptance [53]. Therefore, FMOs may display more universal functions such as detoxification of xenobiotic compounds, while NMOs and BVMOs would be more specific enzymes, since they are stricter regarding both its substrate and reaction scope [54].

Conclusions

1. 98 newly-identified BVMOs of Bacteria, Archaea and Eukarya are described in this work.
2. BVMOs constitute an ancestral family of proteins, irregularly and widely spread among the tree of life.
3. BVMOs and NMOs form monophyletic groups whereas FMOs form a paraphyletic group of genes including the so-called YUCCAs.
4. An evolutionary model, in which FMOs, NMOs and BVMOs genes were already present in LUCA, is proposed here.

Materials & Methods

Data retrieval and multiple sequence alignments

To carry out a survey for the presence of BVMOs encoding sequences across species, a deep *in silico* homology search was performed using genomic (refseq genomic, gss, wgs), transcriptomic (refseq rna, ESTs) and protein (refseq protein) databases at NCBI. The search was conducted as an iterative process in each life domain as follows; (i) BLASTp: PHI-BLAST using as a query PAMO sequence (Q47PU3.1) and the two previously described BVMO canonical motifs: [A/G]GxWxxxx[F/Y]P[G/M]xxxD [24] and FxGxxxHxxxW[P/D] [22], (ii) BLASTp using as a query each sequence of the proteins identified in the previous step and (iii) tBLASTn in nucleotide databases using as a query each sequence of the proteins identified in the two previous steps. The complete set of sequences is available in [S1 Table](#). Multiple sequence alignments (MSA) were performed using the online version of MAFFT 7 program and the BLOSUM 62 scoring matrix [55]. For alignment visualization MEGA 6.0 software was employed. Additionally, a search was conducted using HMMER (<http://hmmer.janelia.org/>) with the hmsearch tool (protein alignment vs. protein sequence database) and the UniProtKB database [56]. However no additional sequences were collected. This analysis provided the HMM logos ([S2 Fig](#)) for the MSA used as input. All the retrieved sequences were curated by confirming the presence of the two BVMOs fingerprints [22, 24] and the two Rossmann motifs (GxGxx[G/A]). The MSA was edited excluding N- and C-terminal regions flanking the two Rossmann motifs. A dataset of 121 BVMOs was obtained ([S1 Dataset](#)).

To construct the dataset of FMOs, YUCCAs and NMOs, BLASTp searches in protein databases of Archaea, Bacteria and Eukarya domains were performed. For FMOs, sequences of FMO 3 from *Homo sapiens* (NP_008825.4), FMO from Rabbit (AAB19844.1) and FMO GS-OX1 from *Arabidopsis thaliana* (NP_176761.1) were employed as queries in different searches. Besides the FMO typifying motif FxGxxxHxxxY[K/R] [42] was used to revise the retrieved sequences. For NMOs, the sequences of L-ornithine N5-oxygenase SidA from *Aspergillus fumigatus* (XP_755103.1) and the alcaligin biosynthesis enzyme from *Bordetella bronchiseptica* (Q44740.2) were used as queries. The YUCCAs sequences were obtained from the literature [15]. The complete set of sequences is available in [S2 Table](#). MSAs were performed for each group of flavoproteins to analyze the presence of conserved motifs, meaning the two Rossmann motifs for all “Class B” monooxygenases and the specific fingerprint in the case of FMOs. The alignment was edited excluding N- and C-terminal regions flanking the two Rossmann fingerprints. A complete alignment, including BVMOs, FMOs, YUCCAs and NMOs sequences, was constructed ([S2 Dataset](#)). Besides, a MSA including all collected “Class B” flavin monooxygenases (BVMOs, NMOs, FMOs and YUCCAs) plus a group of “Class A” flavin monooxygenases (to be used as an outgroup) was constructed ([S3 Dataset](#)). “Class A” monooxygenases were obtained from previous reports [11].

Construction of phylogenetic trees

Two independent approaches were used to construct the phylogenetic trees of BVMOs and of “Class B” monooxygenases. The first approach was using the Maximum likelihood (ML) method. For this, MSAs previously constructed were manually edited to eliminate gaps and the best fit model parameters were calculated employing ProtTest 3.4 version [57]. To build the tree PhyML 3.0 on-line tool was used [58]. To estimate the robustness of the phylogenetic inference 100 bootstraps were run. To visualize and edit the consensus tree MEGA 6.0 software was used.

The other approach consisted in employing Bayesian inference method. The edited MSAs employed for ML trees, were used as input to build the tree using Mr. Bayes 3.2 software [59]. A mixed amino acid substitution model was set up, 4 gamma categories and a proportion of invariable sites were considered. The analyses were concluded after 1,000,000 generations when the split frequency was < 0.02 . To visualize and edit the tree, FigTree 1.4.2 software was used.

Genomic context analyses

To analyze the genomic context of selected BVMOs (*e.g.* BVMOs detected in *Pantholops hodgsonii*, ACOMO from *Gordonia* sp. Odio1 and Odio3 from *Oikopleura doica*, etc.) whole contigs were downloaded. Initially these fragments were subjected to BLASTn searches in reference genomic sequences (refseq_genome) and genomic survey sequences (gss) databases. Then, individual genes upstream and downstream from the BVMO open reading frame (ORF) were analyzed by tBLASTn and the retrieved sequences were subjected to reciprocal tBLASTn in order to determine possible orthologous/xenologous sequences. Also, the contigs were subjected to progressive alignment using MAUVE software [60] with comparison purposes.

PCR experiments

PCR experiments were conducted to confirm the presence of BVMO encoding sequences in selected organisms. Genomic DNA of *Hordeum vulgare* subsp. *vulgare* was obtained from Instituto Nacional de Tecnología Agropecuaria (INTA-Castelar). Genomic DNA from *Emiliana huxleyi* was obtained from Provasoli-Guillard National Center for Culture of Marine Phytoplankton strain CCMP #373. Genomic DNA from *Aspergillus flavus* was purified from a growing culture, employing the Wizard Genomic DNA extraction Kit (Promega). Primers were designed to amplify each BVMO sequence and also housekeeping genes (S3 Table). Fragments were amplified using *Taq*-UBA DNA polymerase (gifted by Dr. Mauro Morgenfeld). PCR conditions were: initial denaturation for 2 min at 94°C, followed by 35 cycles of denaturation (30s at 94°C), annealing (30s at Ta°C) and extension (30-45s at 72°C), final extension of 10 min at 72°C.

Supporting Information

S1 Dataset. BVMOs multiple sequence alignment.

(FASTA)

S2 Dataset. BVMOs, FMOs, YUCCAs and NMOs multiple sequence alignment.

(FASTA)

S3 Dataset. BVMOs, FMOs, YUCCAs, NMOs and “Class A” flavin monooxygenases multiple sequence alignment.

(FASTA)

S1 Fig. Distribution of sequence identities among the BVMOs dataset. The histogram displays the ranges of sequence identity *vs.* the amount (as percentage) of sequences. The identity matrix was constructed employing CLUSTALW program and the BLOSUM62 scoring matrix. (TIF)

S2 Fig. HMM logos of BVMOs sequences. The diagram was obtained from S1 Dataset by performing an analysis employing HMMER tool (available at <http://hmmer.janelia.org/>), as depicted in Materials & Methods section. The two Rossmann motifs (GxGxx[G/A]) and the two BVMO fingerprints ([A/G]GxWxxxx[F/Y]P[G/M]xxxD and FxGxxxHxxxW[P/D]) are highlighted. Arrow indicates the conserved Thr residue detected by Rebehmed *et al* [23]. (TIF)

S3 Fig. PCR experiments to detect the BVMO gene in *Hordeum vulgare*.
(PDF)

S4 Fig. PCR experiments to detect BVMO genes in *Emiliania huxleyi*.
(PDF)

S5 Fig. Phylogenetic analysis of BVMOs protein-family employing Bayesian method.
(PDF)

S6 Fig. Rooted phylogenetic tree of “Class B” flavin monooxygenases employing Bayesian method. The tree was constructed by using the Bayesian method, employing the alignment constructed with MAFFT 7 on-line tool. Evolutionary analyses were conducted using Mr.Bayes 3.2 software. Posterior probabilities values are shown next to the branches. Colored branches display: BVMOs (yellow), NMOs (blue) and FMOs (black). As outgroup, hydroxylases belonging to “Class A” flavin monooxygenases were employed (orange).
(PDF)

S7 Fig. Phylogenetic analysis of selected “Class B” monooxygenases by Maximum Likelihood (ML) method. The tree was constructed by using the ML method, employing the alignment constructed with MAFFT 7 on-line tool and the best model parameters calculated with ProTest 3.4. Evolutionary analyses were conducted in PhyML 3.0 on-line server. Bootstrap values (> 45) are shown next to the branches. Colored branches show: BVMOs (yellow), NMOs (blue), FMOs & YUCCAs (black). Colored triangles display the source of BVMO encoding genes as follows: Fungi (black), Bacteria (orange), Green plants (dark green), Rhodophyta (light green), Metazoa (red), Haptophyta (purple), Archaea (blue), Chlorophyta (light blue).
(PDF)

S1 File. Detection of contaminating sequences in *Pantholops hodgsonii* genome draft.
(PDF)

S2 File. Analyses of *Adineta vaga* genes proposed to be acquired by HGT.
(PDF)

S3 File. Analysis of the phylogenetic distribution of BVMOs and its substrate specificity.
(PDF)

S4 File. Conserved domain architecture analysis of identified BVMO sequences.
(PDF)

S1 Table. BVMO sequences employed in this work. All BVMOs protein sequences employed in this work are depicted. Colors are used to display the origin of each sequence, as follows: Bacteria (white), Fungi (grey), Green plants and Rhodophyta (green), Haptophyta (light purple), Metazoa (orange) and Archaea (blue). Sequences marked as (*) have been previously characterized. Sequences marked as (#) have been previously named as AFL838, AFL619, AFL210 and AFL456, respectively [20]. Sequences without marks are reported here for the first time as putative BVMOs. Most of these sequences are derived from automated annotation of genomic data.
(PDF)

S2 Table. Other “Class B” flavin monooxygenases encoding sequences employed in this work.
(PDF)

S3 Table. Primer sequences and PCR conditions.
(PDF)

Acknowledgments

We are grateful to Dra. Maria Virginia Sanchez Puerta for her helpful comments and criticisms about the manuscript and for her contribution with the *Emiliania huxleyi* genomic DNA. We also would like to thank to Dra. Alejandra Landau who kindly provided the genomic DNA from barley. We also thank to two anonymous reviewers for their comments and suggestions about the manuscript.

Author Contributions

Conceived and designed the experiments: MLM WJL MJA. Performed the experiments: MLM WJL MJA. Analyzed the data: MLM WJL MJA. Contributed reagents/materials/analysis tools: MLM WJL MJA. Wrote the paper: MLM WJL MJA.

References

1. Leisch H, Morley K, Lau PC. Baeyer-Villiger monooxygenases: more than just green chemistry. *Chem Rev*. 2011; 111(7):4165–222. Epub 2011/05/06. doi: [10.1021/cr1003437](https://doi.org/10.1021/cr1003437) PMID: [21542563](https://pubmed.ncbi.nlm.nih.gov/21542563/).
2. Balke K, Kadow M, Mallin H, Sass S, Bornscheuer UT. Discovery, application and protein engineering of Baeyer-Villiger monooxygenases for organic synthesis. *Org Biomol Chem*. 2012; 10(31):6249–65. Epub 2012/06/27. doi: [10.1039/c2ob25704a](https://doi.org/10.1039/c2ob25704a) PMID: [22733152](https://pubmed.ncbi.nlm.nih.gov/22733152/).
3. Wiemann P, Guo CJ, Palmer JM, Sekonyela R, Wang CC, Keller NP. Prototype of an intertwined secondary-metabolite supercluster. *Proc Natl Acad Sci U S A*. 2013; 110(42):17065–70. Epub 2013/10/02. [1313258110 \[pii\]](https://pubmed.ncbi.nlm.nih.gov/24082142/) doi: [10.1073/pnas.1313258110](https://doi.org/10.1073/pnas.1313258110) PMID: [24082142](https://pubmed.ncbi.nlm.nih.gov/24082142/); PubMed Central PMCID: [PMC3801025](https://pubmed.ncbi.nlm.nih.gov/PMC3801025/).
4. Gibson M, Nur-e-alam M, Lipata F, Oliveira MA, Rohr J. Characterization of kinetics and products of the Baeyer-Villiger oxygenase MtmOIV, the key enzyme of the biosynthetic pathway toward the natural product anticancer drug mithramycin from *Streptomyces argillaceus*. *J Am Chem Soc*. 2005; 127(50):17594–5. Epub 2005/12/15. doi: [10.1021/ja055750t](https://doi.org/10.1021/ja055750t) PMID: [16351075](https://pubmed.ncbi.nlm.nih.gov/16351075/).
5. Bosserman MA, Downey T, Noinaj N, Buchanan SK, Rohr J. Molecular insight into substrate recognition and catalysis of Baeyer-Villiger monooxygenase MtmOIV, the key frame-modifying enzyme in the biosynthesis of anticancer agent mithramycin. *ACS Chem Biol*. 2013; 8(11):2466–77. Epub 2013/09/03. doi: [10.1021/cb400399b](https://doi.org/10.1021/cb400399b) PMID: [23992662](https://pubmed.ncbi.nlm.nih.gov/23992662/); PubMed Central PMCID: [PMC3830731](https://pubmed.ncbi.nlm.nih.gov/PMC3830731/).
6. Tang M-C, He H-Y, Zhang F, Tang G-L. Baeyer-Villiger Oxidation of Acyl Carrier Protein-Tethered Thioester to Acyl Carrier Protein-Linked Thiocarbonate Catalyzed by a Monooxygenase Domain in FR901464 Biosynthesis. *ACS Catalysis*. 2013; 3(3):444–7. doi: [10.1021/cs300819e](https://doi.org/10.1021/cs300819e)
7. Leisch H, Shi R, Grosse S, Morley K, Bergeron H, Cygler M, et al. Cloning, Baeyer-Villiger biooxidations, and structures of the camphor pathway 2-oxo-Delta(3)-4,5,5-trimethylcyclopentylacetyl-coenzyme A monooxygenase of *Pseudomonas putida* ATCC 17453. *Appl Environ Microbiol*. 2012; 78(7):2200–12. Epub 2012/01/24. AEM.07694-11 [pii] doi: [10.1128/AEM.07694-11](https://doi.org/10.1128/AEM.07694-11). PMID: [22267661](https://pubmed.ncbi.nlm.nih.gov/22267661/); PubMed Central PMCID: [PMC3302634](https://pubmed.ncbi.nlm.nih.gov/PMC3302634/).
8. Fraaije MW, Kamerbeek NM, Heidekamp AJ, Fortin R, Janssen DB. The prodrug activator EtaA from *Mycobacterium tuberculosis* is a Baeyer-Villiger monooxygenase. *J Biol Chem*. 2004; 279(5):3354–60. Epub 2003/11/12. doi: [10.1074/jbc.M307770200](https://doi.org/10.1074/jbc.M307770200) M307770200 [pii]. PMID: [14610090](https://pubmed.ncbi.nlm.nih.gov/14610090/).
9. Minerdi D, Zgrablic I, Sadeghi SJ, Gilardi G. Identification of a novel Baeyer-Villiger monooxygenase from *Acinetobacter radioresistens*: close relationship to the *Mycobacterium tuberculosis* prodrug activator EtaA. *Microb Biotechnol*. 2012; 5(6):700–16. Epub 2012/08/07. doi: [10.1111/j.1751-7915.2012.00356.x](https://doi.org/10.1111/j.1751-7915.2012.00356.x) PMID: [22862894](https://pubmed.ncbi.nlm.nih.gov/22862894/); PubMed Central PMCID: [PMC3815892](https://pubmed.ncbi.nlm.nih.gov/PMC3815892/).
10. Torres Pazmino DE, Dudek HM, Fraaije MW. Baeyer-Villiger monooxygenases: recent advances and future challenges. *Curr Opin Chem Biol*. 2010; 14(2):138–44. doi: [10.1016/j.cbpa.2009.11.017](https://doi.org/10.1016/j.cbpa.2009.11.017) PMID: [20015679](https://pubmed.ncbi.nlm.nih.gov/20015679/).
11. Huijbers MM, Montersino S, Westphal AH, Tischler D, van Berkel WJ. Flavin dependent monooxygenases. *Arch Biochem Biophys*. 2014; 544:2–17. Epub 2013/12/24. S0003-9861(13)00370-6 [pii] doi: [10.1016/j.abb.2013.12.005](https://doi.org/10.1016/j.abb.2013.12.005) PMID: [24361254](https://pubmed.ncbi.nlm.nih.gov/24361254/).
12. Schlenk D. Occurrence of flavin-containing monooxygenases in non-mammalian eukaryotic organisms. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol*. 1998; 121(1–3):185–95. Epub 1999/02/11. PMID: [9972460](https://pubmed.ncbi.nlm.nih.gov/9972460/).

13. Olucha J, Lamb AL. Mechanistic and structural studies of the N-hydroxylating flavoprotein monooxygenases. *Bioorg Chem.* 2011; 39(5–6):171–7. Epub 2011/08/30. S0045-2068(11)00049-6 [pii] doi: [10.1016/j.bioorg.2011.07.006](https://doi.org/10.1016/j.bioorg.2011.07.006) PMID: [21871647](https://pubmed.ncbi.nlm.nih.gov/21871647/); PubMed Central PMCID: PMC3188341.
14. Franceschini S, Fedkenheuer M, Vogelaar NJ, Robinson HH, Sobrado P, Mattevi A. Structural insight into the mechanism of oxygen activation and substrate selectivity of flavin-dependent N-hydroxylating monooxygenases. *Biochemistry.* 2012; 51(36):7043–5. Epub 2012/08/30. doi: [10.1021/bi301072w](https://doi.org/10.1021/bi301072w) PMID: [22928747](https://pubmed.ncbi.nlm.nih.gov/22928747/).
15. Schlaich NL. Flavin-containing monooxygenases in plants: looking beyond detox. *Trends Plant Sci.* 2007; 12(9):412–8. Epub 2007/09/04. S1360-1385(07)00189-6 [pii] doi: [10.1016/j.tplants.2007.08.009](https://doi.org/10.1016/j.tplants.2007.08.009) PMID: [17765596](https://pubmed.ncbi.nlm.nih.gov/17765596/).
16. Bisagni S, Hatti-Kaul R, Mamo G. Cloning, expression and characterization of a versatile Baeyer-Villiger monooxygenase from *Dietzia* sp. D5. *AMB Express.* 2014; 4:23. doi: [10.1186/s13568-014-0023-1](https://doi.org/10.1186/s13568-014-0023-1). PMID: [24949258](https://pubmed.ncbi.nlm.nih.gov/24949258/); PubMed Central PMCID: PMC4052671.
17. Kamerbeek NM, Olsthoorn AJ, Fraaije MW, Janssen DB. Substrate specificity and enantioselectivity of 4-hydroxyacetophenone monooxygenase. *Appl Environ Microbiol.* 2003; 69(1):419–26. PMID: [12514023](https://pubmed.ncbi.nlm.nih.gov/12514023/); PubMed Central PMCID: PMC152415.
18. Leipold F, Wardenga R, Bornscheuer UT. Cloning, expression and characterization of a eukaryotic cycloalkanone monooxygenase from *Cylindrocarpum radicolica* ATCC 11011. *Appl Microbiol Biotechnol.* 2012; 94(3):705–17. doi: [10.1007/s00253-011-3670-z](https://doi.org/10.1007/s00253-011-3670-z) PMID: [22075635](https://pubmed.ncbi.nlm.nih.gov/22075635/).
19. Mascotti ML, Juri Ayub M, Dudek H, Sanz MK, Fraaije MW. Cloning, overexpression and biocatalytic exploration of a novel Baeyer-Villiger monooxygenase from *Aspergillus fumigatus* Af293. *AMB Express.* 2013; 3(1):33. doi: [10.1186/2191-0855-3-33](https://doi.org/10.1186/2191-0855-3-33) PMID: [23767684](https://pubmed.ncbi.nlm.nih.gov/23767684/); PubMed Central PMCID: PMC3762062.
20. Ferroni FM, Smit MS, Opperman DJ. Functional divergence between closely related Baeyer-Villiger monooxygenases from *Aspergillus flavus*. *Journal of Molecular Catalysis B: Enzymatic.* 2014; 107(0):47–54. <http://dx.doi.org/10.1016/j.molcatb.2014.05.015>.
21. Beneventi E, Niero M, Motterle R, Fraaije M, Bergantino E. Discovery of Baeyer—Villiger monooxygenases from photosynthetic eukaryotes. *Journal of Molecular Catalysis B: Enzymatic.* 2013; 98(0):145–54. <http://dx.doi.org/10.1016/j.molcatb.2013.10.006>.
22. Fraaije MW, Kamerbeek NM, van Berkel WJ, Janssen DB. Identification of a Baeyer-Villiger monooxygenase sequence motif. *FEBS Lett.* 2002; 518(1–3):43–7. PMID: [11997015](https://pubmed.ncbi.nlm.nih.gov/11997015/).
23. Rebehmed J, Alphand V, de Berardinis V, de Brevérn AG. Evolution study of the Baeyer-Villiger monooxygenases enzyme family: functional importance of the highly conserved residues. *Biochimie.* 2013; 95(7):1394–402. doi: [10.1016/j.biochi.2013.03.005](https://doi.org/10.1016/j.biochi.2013.03.005) PMID: [23523772](https://pubmed.ncbi.nlm.nih.gov/23523772/).
24. Riebel A, de Gonzalo G, Fraaije MW. Expanding the biocatalytic toolbox of flavoprotein monooxygenases from *Rhodococcus jostii* RHA1. *Journal of Molecular Catalysis B: Enzymatic.* 2013; 88(0):20–5. <http://dx.doi.org/10.1016/j.molcatb.2012.11.009>.
25. Nasir A, Kim KM, Caetano-Anolles G. Global patterns of protein domain gain and loss in superkingdoms. *PLoS Comput Biol.* 2014; 10(1):e1003452. Epub 2014/02/07. doi: [10.1371/journal.pcbi.1003452](https://doi.org/10.1371/journal.pcbi.1003452) PCOMPBIOL-D-13-01258 [pii]. PMID: [24499935](https://pubmed.ncbi.nlm.nih.gov/24499935/); PubMed Central PMCID: PMC3907288.
26. Noda-Garcia L, Camacho-Zarco AR, Medina-Ruiz S, Gaytan P, Carrillo-Tripp M, Fulop V, et al. Evolution of substrate specificity in a recipient's enzyme following horizontal gene transfer. *Mol Biol Evol.* 2013; 30(9):2024–34. Epub 2013/06/27. mst115 [pii] doi: [10.1093/molbev/mst115](https://doi.org/10.1093/molbev/mst115) PMID: [23800623](https://pubmed.ncbi.nlm.nih.gov/23800623/).
27. Choi IG, Kim SH. Global extent of horizontal gene transfer. *Proc Natl Acad Sci U S A.* 2007; 104(11):4489–94. Epub 2007/03/16. 0611557104 [pii] doi: [10.1073/pnas.0611557104](https://doi.org/10.1073/pnas.0611557104) PMID: [17360551](https://pubmed.ncbi.nlm.nih.gov/17360551/); PubMed Central PMCID: PMC1815472.
28. Pace NR, Sapp J, Goldenfeld N. Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. *Proc Natl Acad Sci U S A.* 2012; 109(4):1011–8. Epub 2012/02/07. 1109716109 [pii] doi: [10.1073/pnas.1109716109](https://doi.org/10.1073/pnas.1109716109) PMID: [22308526](https://pubmed.ncbi.nlm.nih.gov/22308526/); PubMed Central PMCID: PMC3268332.
29. Koonin EV, Wolf YI. The common ancestry of life. *Biol Direct.* 2010; 5:64. Epub 2010/11/23. 1745-6150-5-64 [pii] doi: [10.1186/1745-6150-5-64](https://doi.org/10.1186/1745-6150-5-64) PMID: [21087490](https://pubmed.ncbi.nlm.nih.gov/21087490/); PubMed Central PMCID: PMC2993666.
30. Furnham N, Sillitoe I, Holliday GL, Cuff AL, Laskowski RA, Orengo CA, et al. Exploring the evolution of novel enzyme functions within structurally defined protein superfamilies. *PLoS Comput Biol.* 2012; 8(3): e1002403. Epub 2012/03/08. doi: [10.1371/journal.pcbi.1002403](https://doi.org/10.1371/journal.pcbi.1002403) PCOMPBIOL-D-11-01458 [pii]. PMID: [22396634](https://pubmed.ncbi.nlm.nih.gov/22396634/); PubMed Central PMCID: PMC3291543.
31. Krylov DM, Wolf YI, Rogozin IB, Koonin EV. Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. *Genome Res.* 2003; 13(10):2229–35. Epub 2003/10/04. [pii]. PMID: [14525925](https://pubmed.ncbi.nlm.nih.gov/14525925/); PubMed Central PMCID: PMC403683.

32. Jordan IK, Rogozin IB, Wolf YI, Koonin EV. Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome Res.* 2002; 12(6):962–8. Epub 2002/06/05. doi: [10.1101/gr.87702](https://doi.org/10.1101/gr.87702). Article published online before print in May 2002. PMID: [12045149](https://pubmed.ncbi.nlm.nih.gov/12045149/); PubMed Central PMCID: PMC1383730.
33. Read BA, Kegel J, Klute MJ, Kuo A, Lefebvre SC, Maumus F, et al. Pan genome of the phytoplankton *Emiliania underpins its global distribution*. *Nature.* 2013; 499(7457):209–13. Epub 2013/06/14. nature12221 [pii] doi: [10.1038/nature12221](https://doi.org/10.1038/nature12221) PMID: [23760476](https://pubmed.ncbi.nlm.nih.gov/23760476/).
34. Ge RL, Cai Q, Shen YY, San A, Ma L, Zhang Y, et al. Draft genome sequence of the Tibetan antelope. *Nat Commun.* 2013; 4:1858. Epub 2013/05/16. ncomms2860 [pii] doi: [10.1038/ncomms2860](https://doi.org/10.1038/ncomms2860) PMID: [23673643](https://pubmed.ncbi.nlm.nih.gov/23673643/); PubMed Central PMCID: PMC3674232.
35. Laurence M, Hatzis C, Brash DE. Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLoS One.* 2014; 9(5):e97876. Epub 2014/05/20. doi: [10.1371/journal.pone.0097876](https://doi.org/10.1371/journal.pone.0097876) PONE-D-14-10042 [pii]. PMID: [24837716](https://pubmed.ncbi.nlm.nih.gov/24837716/); PubMed Central PMCID: PMC4023998.
36. Gladyshev EA, Meselson M, Arkipova IR. Massive horizontal gene transfer in bdelloid rotifers. *Science.* 2008; 320(5880):1210–3. Epub 2008/05/31. 320/5880/1210 [pii] doi: [10.1126/science.1156407](https://doi.org/10.1126/science.1156407) PMID: [18511688](https://pubmed.ncbi.nlm.nih.gov/18511688/).
37. Mei B, Budde AD, Leong SA. sid1, a gene initiating siderophore biosynthesis in *Ustilago maydis*: molecular characterization, regulation by iron, and role in phytopathogenicity. *Proc Natl Acad Sci U S A.* 1993; 90(3):903–7. Epub 1993/02/01. PMID: [8430103](https://pubmed.ncbi.nlm.nih.gov/8430103/); PubMed Central PMCID: PMC45778.
38. Orru R, Pazmino DE, Fraaije MW, Mattevi A. Joint functions of protein residues and NADP(H) in oxygen activation by flavin-containing monooxygenase. *J Biol Chem.* 2010; 285(45):35021–8. Epub 2010/09/03. M110.161372 [pii] doi: [10.1074/jbc.M110.161372](https://doi.org/10.1074/jbc.M110.161372) PMID: [20807767](https://pubmed.ncbi.nlm.nih.gov/20807767/); PubMed Central PMCID: PMC2966116.
39. Diminic J, Starcevic A, Lisfi M, Baranasic D, Gacesa R, Hranueli D, et al. Evolutionary concepts in natural products discovery: what actinomycetes have taught us. *J Ind Microbiol Biotechnol.* 2014; 41(2):211–7. Epub 2013/09/26. doi: [10.1007/s10295-013-1337-8](https://doi.org/10.1007/s10295-013-1337-8) PMID: [24061567](https://pubmed.ncbi.nlm.nih.gov/24061567/).
40. Chen H, Li G, Kollner TG, Jia Q, Gershenzon J, Chen F. Positive Darwinian selection is a driving force for the diversification of terpenoid biosynthesis in the genus *Oryza*. *BMC Plant Biol.* 2014; 14:239. Epub 2014/09/17. s12870-014-0239-x [pii] doi: [10.1186/s12870-014-0239-x](https://doi.org/10.1186/s12870-014-0239-x) PMID: [25224158](https://pubmed.ncbi.nlm.nih.gov/25224158/); PubMed Central PMCID: PMC4172859.
41. Mascotti ML, Kurina-Sanz M, Juri Ayub M, Fraaije MW. Insights in the kinetic mechanism of the eukaryotic Baeyer-Villiger monooxygenase BVMOAf1 from *Aspergillus fumigatus* Af293. *Biochimie.* 2014; 107 Pt B:270–6. doi: [10.1016/j.biochi.2014.09.005](https://doi.org/10.1016/j.biochi.2014.09.005) PMID: [25230086](https://pubmed.ncbi.nlm.nih.gov/25230086/).
42. Riebel A, Dudek HM, de Gonzalo G, Stepniak P, Rychlewski L, Fraaije MW. Expanding the set of rhodococcal Baeyer-Villiger monooxygenases by high-throughput cloning, expression and substrate screening. *Appl Microbiol Biotechnol.* 2012; 95(6):1479–89. doi: [10.1007/s00253-011-3823-0](https://doi.org/10.1007/s00253-011-3823-0) PMID: [22218769](https://pubmed.ncbi.nlm.nih.gov/22218769/); PubMed Central PMCID: PMC3427485.
43. Szolkowy C, Eltis LD, Bruce NC, Grogan G. Insights into sequence-activity relationships amongst Baeyer-Villiger monooxygenases as revealed by the intragenomic complement of enzymes from *Rhodococcus jostii* RHA1. *Chembiochem.* 2009; 10(7):1208–17. doi: [10.1002/cbic.200900011](https://doi.org/10.1002/cbic.200900011) PMID: [19360806](https://pubmed.ncbi.nlm.nih.gov/19360806/).
44. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet.* 2008; 9(8):605–18. Epub 2008/07/02. nrg2386 [pii] doi: [10.1038/nrg2386](https://doi.org/10.1038/nrg2386) PMID: [18591983](https://pubmed.ncbi.nlm.nih.gov/18591983/).
45. Kurland CG, Canback B, Berg OG. Horizontal gene transfer: a critical view. *Proc Natl Acad Sci U S A.* 2003; 100(17):9658–62. Epub 2003/08/07. doi: [10.1073/pnas.1632870100](https://doi.org/10.1073/pnas.1632870100) 1632870100 [pii]. PMID: [12902542](https://pubmed.ncbi.nlm.nih.gov/12902542/); PubMed Central PMCID: PMC187805.
46. Mihovilovic MD, Rudroff F, Grotzl B, Kapitan P, Snajdrova R, Rydz J, et al. Family clustering of Baeyer-Villiger monooxygenases based on protein sequence and stereopreference. *Angew Chem Int Ed Engl.* 2005; 44(23):3609–13. Epub 2005/04/30. doi: [10.1002/anie.200462964](https://doi.org/10.1002/anie.200462964) PMID: [15861400](https://pubmed.ncbi.nlm.nih.gov/15861400/).
47. Fink MJ, Fischer TC, Rudroff F, Dudek H, Fraaije MW, Mihovilovic MD. Extensive substrate profiling of cyclopentadecanone monooxygenase as Baeyer–Villiger biocatalyst reveals novel regiodivergent oxidations. *Journal of Molecular Catalysis B: Enzymatic.* 2011; 73(1–4):9–16. <http://dx.doi.org/10.1016/j.molcatb.2011.07.003>.
48. Hu Y, Dietrich D, Xu W, Patel A, Thuss JAJ, Wang J, et al. A carbonate-forming Baeyer-Villiger monooxygenase. *Nat Chem Biol.* 2014; 10(7):552–4. doi: [10.1038/nchembio.1527](https://doi.org/10.1038/nchembio.1527) <http://www.nature.com/nchembio/journal/v10/n7/abs/nchembio.1527.html#supplementary-information>. PMID: [24838010](https://pubmed.ncbi.nlm.nih.gov/24838010/)
49. Weiss M, Denger K, Huhn T, Schleheck D. Two enzymes of a complete degradation pathway for linear alkylbenzenesulfonate (LAS) surfactants: 4-sulfoacetophenone Baeyer-Villiger monooxygenase and 4-sulfophenylacetate esterase in *Comamonas testosteroni* KF-1. *Appl Environ Microbiol.* 2012; 78

- (23):8254–63. Epub 2012/09/25. AEM.02412-12 [pii] doi: [10.1128/AEM.02412-12](https://doi.org/10.1128/AEM.02412-12) PMID: [23001656](https://pubmed.ncbi.nlm.nih.gov/23001656/); PubMed Central PMCID: [PMC3497377](https://pubmed.ncbi.nlm.nih.gov/PMC3497377/).
50. Woese C. The universal ancestor. *Proc Natl Acad Sci U S A*. 1998; 95(12):6854–9. Epub 1998/06/17. PMID: [9618502](https://pubmed.ncbi.nlm.nih.gov/9618502/); PubMed Central PMCID: [PMC22660](https://pubmed.ncbi.nlm.nih.gov/PMC22660/).
 51. Woese CR. On the evolution of cells. *Proc Natl Acad Sci U S A*. 2002; 99(13):8742–7. Epub 2002/06/22. doi: [10.1073/pnas.132266999](https://doi.org/10.1073/pnas.132266999) 132266999 [pii]. PMID: [12077305](https://pubmed.ncbi.nlm.nih.gov/12077305/); PubMed Central PMCID: [PMC124369](https://pubmed.ncbi.nlm.nih.gov/PMC124369/).
 52. Forterre P. The common ancestor of archaea and eukarya was not an archaeon. *Archaea*. 2013; 2013:372396. Epub 2013/12/19. doi: [10.1155/2013/372396](https://doi.org/10.1155/2013/372396) PMID: [24348094](https://pubmed.ncbi.nlm.nih.gov/24348094/); PubMed Central PMCID: [PMC3855935](https://pubmed.ncbi.nlm.nih.gov/PMC3855935/).
 53. Ceccoli RD, Bianchi DA, Rial DV. Flavoprotein monooxygenases for oxidative biocatalysis: recombinant expression in microbial hosts and applications. *Front Microbiol*. 2014; 5:25. Epub 2014/02/26. doi: [10.3389/fmicb.2014.00025](https://doi.org/10.3389/fmicb.2014.00025) PMID: [24567729](https://pubmed.ncbi.nlm.nih.gov/24567729/); PubMed Central PMCID: [PMC3915288](https://pubmed.ncbi.nlm.nih.gov/PMC3915288/).
 54. Frederick RE, Ojha S, Lamb A, Dubois JL. How pH modulates the reactivity and selectivity of a siderophore-associated flavin monooxygenase. *Biochemistry*. 2014; 53(12):2007–16. Epub 2014/02/05. doi: [10.1021/bi401256b](https://doi.org/10.1021/bi401256b) PMID: [24490904](https://pubmed.ncbi.nlm.nih.gov/24490904/); PubMed Central PMCID: [PMC3985866](https://pubmed.ncbi.nlm.nih.gov/PMC3985866/).
 55. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*. 2013; 30(4):772–80. doi: [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010) PMID: [23329690](https://pubmed.ncbi.nlm.nih.gov/23329690/)
 56. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res*. 2011; 39(Web Server issue):W29–37. doi: [10.1093/nar/gkr367](https://doi.org/10.1093/nar/gkr367) PMID: [21593126](https://pubmed.ncbi.nlm.nih.gov/21593126/); PubMed Central PMCID: [PMC3125773](https://pubmed.ncbi.nlm.nih.gov/PMC3125773/).
 57. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*. 2011; 27(8):1164–5. Epub 2011/02/22. btr088 [pii] doi: [10.1093/bioinformatics/btr088](https://doi.org/10.1093/bioinformatics/btr088) PMID: [21335321](https://pubmed.ncbi.nlm.nih.gov/21335321/).
 58. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 2010; 59(3):307–21. Epub 2010/06/09. syq010 [pii] doi: [10.1093/sysbio/syq010](https://doi.org/10.1093/sysbio/syq010) PMID: [20525638](https://pubmed.ncbi.nlm.nih.gov/20525638/).
 59. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012; 61(3):539–42. Epub 2012/02/24. sys029 [pii] doi: [10.1093/sysbio/sys029](https://doi.org/10.1093/sysbio/sys029) PMID: [22357727](https://pubmed.ncbi.nlm.nih.gov/22357727/); PubMed Central PMCID: [PMC3329765](https://pubmed.ncbi.nlm.nih.gov/PMC3329765/).
 60. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*. 2010; 5(6):e11147. Epub 2010/07/02. doi: [10.1371/journal.pone.0011147](https://doi.org/10.1371/journal.pone.0011147) PMID: [20593022](https://pubmed.ncbi.nlm.nih.gov/20593022/); PubMed Central PMCID: [PMC2892488](https://pubmed.ncbi.nlm.nih.gov/PMC2892488/).