Utilization of Cobalamin Is Ubiquitous in Early-Branching Fungal Phyla

Małgorzata Orłowska (iD), Kamil Steczkiewicz (iD), and Anna Muszewska (iD)*

Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland

*Corresponding author: E-mail: musze@ibb.waw.pl.

Abstract

Cobalamin is a cofactor present in essential metabolic pathways in animals and one of the water-soluble vitamins. It is a complex compound synthesized solely by prokaryotes. Cobalamin dependence is scattered across the tree of life. In particular, fungi and plants were deemed devoid of cobalamin. We demonstrate that cobalamin is utilized by all non-Dikarya fungi lineages. This observation is supported by the genomic presence of both B12-dependent enzymes and cobalamin modifying enzymes. Fungal cobalamin-dependent enzymes are highly similar to their animal homologs. Phylogenetic analyses support a scenario of vertical inheritance of the cobalamin usage with several losses. Cobalamin usage was probably lost in Mucorinae and at the base of Dikarya which groups most of the model organisms and which hindered B12-dependent metabolism discovery in fungi. Our results indicate that cobalamin dependence was a widely distributed trait at least in Opisthokonta, across diverse microbial eukaryotes and was likely present in the LECA.

Key words: cobalamin, early-diverging fungi, metabolic traits, vitamin B12, fungal evolution.

Significance

Cobalamin dependence is scattered across the tree of life. In particular, fungi and plants have been deemed devoid of cobalamin. We demonstrate that cobalamin is utilized by all phyla of non-Dikarya fungal lineages. This observation is supported by the identification of B12-dependent enzymes and cobalamin modifying enzymes in published fungal genomes. We found that cobalamin dependence was a widely distributed trait at least in Opisthokonta, across diverse microbial eukaryotes and likely in the LECA (last eukaryotic common ancestor). Moreover, the genes identified are actively transcribed in many taxa.

Introduction

Cobalamin, also known as vitamin B12, is the most common cobalt-containing compound in nature and one of eight known water-soluble vitamins grouped into B class. Cobalamin is an organometallic complex compound that contains a cobalt atom placed within a corrin ring. Vitamin B12 is derived from uroporphyrinogen III, which is also the first macrocyclic intermediate in a common pathway of heme and chlorophyll biosynthesis (Frank et al. 2005; Chatthanawaree 2011; Dereven'kov et al. 2016). In animals, cobalamin is used as a cofactor in myelin formation and thus is crucial for the proper functioning of the nervous system. A deficit of this vitamin in the diet may lead to sensory or motor deficiencies and to degeneration of the spinal cord (Dardiotis et al. 2017).

Biosynthesis of cobalamin takes place only in bacteria and archaea which is quite unique for such a widely used vitamin. It is a very complex process involving more than 30 genes (Roth et al. 1993) collectively conserved only in B12-producing prokaryotes which suggest a common origin of the whole pathway. Nonetheless, animals and some protists, but not fungi, are known to utilize cobalamin in their metabolism so they have to intake this vitamin with food. Interestingly, several eukaryotic microorganisms, including

Phytophthora (Oomycota) and Dictyostelium (Amoebozoa), do possess B12-dependent enzymes (Grenville-Briggs et al. 2005; Crona et al. 2013). Some algae like *Porphyridium purpureum* and *Amphidinium operculatum*, can obtain the cobalamin cofactor from associated bacteria (Croft et al. 2005). Plants and fungi are believed to neither synthesize nor even have a need for the cobalamin (Duda et al. 1957; Martens et al. 2002). Even more, they are regarded as devoid of cobalt at all (Zhang et al. 2019).

In Eukaryotes, B12-dependent enzymes are used in diverse processes ranging from the regeneration of methionine from homocysteine, catabolic breakdown of some amino acids into succinyl-CoA (necessary for citric acid cycle) and proper myelin synthesis. Only seven enzymes from the above pathways seem to be uniquely present in B12-dependent organisms. They either modify cobalamin or use it as a cofactor. The former group contains methylmalonyl Co-A mutase-associated GTPase Cob (MeaB), cob(I)yrinic acid a, c-diamide adenosyltransferase (CblAdo transferase), cyanocobalamin reductase (CblC), and cobalamin trafficking protein D (CblD) proteins, whereas the latter includes methionine synthase (MetH), methylmalonyl-CoA epimerase (MM-CoA epimerase), and methylmalonyl-CoA mutase (MM-CoA mutase) (Kräutler 2012). All these proteins are present in animals, including Holozoa, for example, *Monosiga brevicollis*. There is one more cobalamin-dependent enzyme that is present uniquely in archaebacteria, eubacteria, and bacteriophages. This enzyme called ribonucleotide reductase class II (RNR class II) takes a part in DNA replication and repair (Larsson et al. 2004; Herrick and Sclavi 2007). For consistency and clarity, we will use the names of human representatives (given above) to tag the above seven enzymes and a bacterial representative for the last one.

RNR is an enzyme that catalyzes the formation of deoxyribonucleotides from ribonucleotides. It plays a pivotal role in the synthesis, reparation, and regulation of the total rate of DNA synthesis (Larsson et al. 2004; Herrick and Sclavi 2007). RNRs are divided into three classes that are working based on similar mechanisms but using a different compound to generate free radicals. Class I reductases are divided into IA and IB subclasses. These reductases generate tyrosyl free radicals from iron. Subclass IA is distributed in eukaryotes, eubacteria, and viruses. Subclass IB can be found only in eubacteria. Class II reductases use free radicals from cobalamin and are distributed in archaebacteria, eubacteria, and bacteriophages. The same distribution applies to class III, but this class uses a glycine radical (Kolberg 2004). Most eukaryotes, including animals, use class IA reductases, but surprisingly *Phytophthora* spp. uses cobalamin-dependent class II RNR.

Methionine synthetase (MetH) comes in two variants: cobalamin-dependent MetH (EC 2.1.1.13) and cobalamin-independent MetE (EC 2.1.1.14). MetH catalyzes the final step in the remethylation of homocysteine which explains increased levels of homocysteine upon vitamin B12 deficiency.

In animals, this may lead to blindness, neurological symptoms, and birth defects (Outteryck et al. 2012). MetH requires Cyanocobalamin reductase (CblC) and Cobalamin trafficking protein (CblD) for proper function. ClbC catalyzes the decyanation of cyanocobalamin and the dealkylation of alkylcobalamins. In bacteria, an analog of CblC/D, namely TonB, is involved in energy transduction for the uptake of cobalamin (Lerner-Ellis et al. 2006; Hannibal et al. 2009). CblD interacts with CblC and directs CblC-cob(II)alamin molecules to the mitochondrion. Consistently, ClbC localizes either to cytoplasm or mitochondria, whereas ClbD remains in the cytosol (Gherasim et al. 2013; Mah et al. 2013).

CblAdo transferase, cob(I)yrinic acid a, c-diamide adenosyltransferase, converts cobalamin into adenosylcobalamin (AdoCbl). AdoCbl is a cofactor of multiple enzymes that catalyze unusual rearrangement or elimination reactions. Some of them are restricted to prokaryotes, for example, lysine-5,6-aminomutase, isobutyryl-CoA mutase, and glutamate mutase. Others are present also in Eukaryotes, for example, methylmalonyl-CoA mutase (Mera and Escalante-Semerena 2010; Marsh and Román Meléndez 2012).

In humans, MM-CoA epimerase and MM-CoA mutase are both involved in fatty acid catabolism. MM-CoA epimerase catalyzes the rearrangement of (S)-methylmalonyl-CoA to the (R) form and uses a vitamin B12 cofactor (Overath et al. 1962). MM-CoA mutase induces the formation of adenosyl radical from AdoCbl cofactor and subsequently initiates a free-radical rearrangement of its substrate, (R)-methylmalonyl-CoA to succinyl-CoA—a key molecule of the citric acid cycle (Mancia et al. 1996). Methylmalonyl Co-A mutase-associated GTPase Cob (MeaB) is crucial for the proper functioning of methylmalonyl-CoA mutase (Takahashi-Iñiguez et al. 2017). Mutational analysis of this protein performed in *Methylobacterium* sp. showed an inability to convert methylmalonyl-CoA to succinyl-CoA caused by an inactive form of methylmalonyl-CoA mutase (Froese et al. 2010).

Kingdom Fungi comprises several lineages of non-Dikarya which, in the order of divergence, are classified into Chytrydiomycota (Chang et al. 2015; Liu and Stajich 2015) and Blastocladiomycota (Tabima et al. 2020) grouping many aquatic organisms, fully terrestrial animal-related Zoopagomycotina (Ahrendt et al. 2018), Entomophthoromycotina, Kickxellomycotina (Chang et al. 2015) and plant/soil/dung-associated Mucoromycotina (Mondo et al. 2017), Mortierellomycotina (Uehling et al. 2017), and Glomeromycotina (Chen et al. 2018). We can also list Microsporidia, that are described as the earliest diverging clade of fungi (Capella-Gutiérrez et al. 2012). The remaining Dikarya include evolutionary youngest and best-studied fungal phyla—Ascomycota and Basidiomycota (Spatafora et al. 2016). None of the aforementioned B12-related enzymes has been reported from fungi. Yet non-Dikarya, early-diverging lineages of fungi share multiple ancestral traits with animals and microbial eukaryotes. Here, we show that all

B12-dependent eukaryotic pathways are present in non-Dikarya fungi as well.

Results

Our initial searches showed that only eight enzymes are uniquely present in B12-dependent organisms (table 1). All of them have their homologs within early-diverging fungal lineages (supplementary table S1, Supplementary Material online).

Distribution of B12-Dependent Enzymes in Fungi

Cobalamin-dependent enzymes were identified in 50 out of 59 analyzed non-Dikarya fungi (table 1; see supplementary table S2, Supplementary Material online, for detailed lists of all protein accessions). This data set contains all genome-derived protein predictions for all non-Dikarya isolates deposited in GenBank by October 2019, with representatives of all main lineages. The distribution of cobalamin-related enzymes among non-Dikarya fungi is shown in figure 1 and per organism occurrence of B12-related protein-coding genes is summarized in supplementary table S1, Supplementary Material online. The whole set of studied enzymes is present in five non-Dikarya fungal proteomes, four of them belonging to the Glomeromycotina (Mucoromycota). The occurrence of cobalamin-related enzymes is common for all Mucoromycota species, but worth noting are the differences between Glomeromycotina, Mortierellomycotina, and Mucoromycotina (the latter comprising saprotrophic Mucorales, Umbelopsidales, and plant symbionts Endogonales). In Mucorales only three families of cobalamin-dependent enzymes are conserved (CblD, MetH, and CblAdo transferase). For other Mortierallomycotina and Endogonales, it is common to retain four or more of the analyzed protein families. The whole set of enzymes can be found also in Blastocladiomycota. Other taxa with a high occurrence of cobalamin-dependent enzyme families are the animal-related Entomophthoromycotina, Kickxellomycotina, and Zoopagomycotina. All of them have homologs from four up to seven families.

Nine of the analyzed proteomes, all belonging to Chytridiomycota, do not contain any of the studied enzymes. The presence of six out of eight studied proteins in Chytridiales and Monoblepharidiales shows that not all Chytridiomycota are devoid of B12-dependent genes. Neocallimastigomycetes stand out especially here—none of the analyzed four proteomes from this taxon had any homologs of the cobalamin-related proteins family.

Conservation of B12 Processing Pathways

Cobalamin-dependent enzymes play roles in three pathways associated with RNR class II, MetH, and MM-CoA mutase. Obtained results suggest that among early-diverging fungi there is a tendency to conserve the key enzymes rather than whole pathways. This is especially true in Mucorales which retained only CblAdo transferase and part of MetH pathways. In other non-Dikarya fungi, MM-CoA mutase-associated pathway is also well conserved. The B12-dependent ribonucleotide reductase is least conserved but this might be associated with the presence of different RNR classes.

In order to ensure that all housekeeping functions provided by RNR class II, MM-CoA mutase, and MetH pathways are maintained in all of the studied isolates, even those devoid of B12-dependent enzymes, we searched for cobalamin-independent alternatives. We looked for MetE which can substitute MetH (González et al. 1996), RNR class I instead of class II (Jordan and Reichard 1998), and methylcitrate cycle (MCC) as an alternative to MM-CoA mutase pathway (Dubey et al. 2013). We found that all these enzymes involved in B12-independent metabolic tracks can be identified in Dikarya and some of the non-Dikarya fungi providing the required enzymatic activity without B12 cofactor. The B12-independent alternatives are differently distributed among studied genomes compared with their B12-dependent counterparts. RNR class I and MCC synthase were found in all analyzed genomes and in other early-diverging fungi that have not been included in our analysis but were present in the NCBI NR database. Some early-diverging fungi do not have B12-dependent enzymes at all for instance Batrachochytrium spp. have these two alternative B12-independent traits. Interestingly, both methionine synthases MetE and MetH can also co-occur in one genome—we observed co-occurrence of these two enzymes in 45 analyzed genomes. MetE which is an alternative to B12-dependent MetH is present in 51 studied genomes, mostly in Mortierellomycotina but also in genomes that do not have B12-dependent enzymes (e.g., Piromyces finnis). Seven Mortierellomycotina genomes, which were not included in our data sets, have also MetE copies. This might suggest that the metabolic pathways which first required B12 are still conserved among fungi, but have become independent of cobalamin for some organisms.

Since most of the identified homologs of eight B12-dependent enzymes are annotated as hypothetical unknown proteins without experimental characterization, we performed TBlastN searches on them against the NCBI EST database. This served as intermediate evidence that the predicted B12-related proteins in non-Dikarya fungi originate from active genes. TBlastN search results allow also to expect that genes encoding all identified proteins will be expressed.

Phylogenetic Analysis of Protein Homologs Associated with B12 Utilization

To trace the evolution of the studied proteins, phylogenetic trees for each of the eight protein families were inferred using Bayesian (BA) and Maximum Likelihood (ML)

Table 1
B12-Specific Enzymes Used for the Identification of B12-Dependent Pathways in Fungal Proteomes with the Total Number of Homologs Identified in This Study

| Human Gene	EC	Enzyme Name	Reference	Pfam Name	Pfam ID	Structure ID	No. of Fungal Proteins	No. of Fungal Species
MUT	5.4.99.2	Methylmalonyl-CoA mutase, MM-CoA mutase	Forny et al. (2014)	MM_CoA_mutase	PF01642	2XIJ	39	36
MCEE	5.1.99.1	Methylmalonyl-CoA epimerase, MM-CoA epimerase	Bobik and Rasche (2001)	Glyoxalase_4	PF13669	1JC5	42	32
MMAA	3.6.5.	Methylmalonyl Co-A mutase-associated GTPase, MeaB	Bobik and Rasche (2001); Froese et al. (2010)	MeaB	PF03308	2WWW	30	24
MTR	2.1.1.13	Methionine synthase, MetH	Bobik and Rasche (2001); Froese et al. (2010); Bassila et al. (2017)	Met_synt_B12	PF02965	2O2K	65	53
MMAB	2.5.1.17	Methylmalonic aciduria and homocystinuria type B family, CblAdo transferase	Mera et al. (2007)	Cob_adeno_trans	PF01923	2R6X	71	50
MMACHC	1.16.1.6	Methylmalonic aciduria and homocystinuria type C family, CblC	Kim et al. (2009)	CblC	PF16690	3SBZ	23	20
MMADHC	—	Methylmalonic aciduria and homocystinuria type D family, CblD	Coelho et al. (2008)	CblD	PF10229	5CV0	55	47
—	1.17.4.2	Ribonucleoside-diphosphate reductase class II, RNR class II	Booker and Stubbe (1993)	RNR_Alpha	PF17975	1L1L	24	13
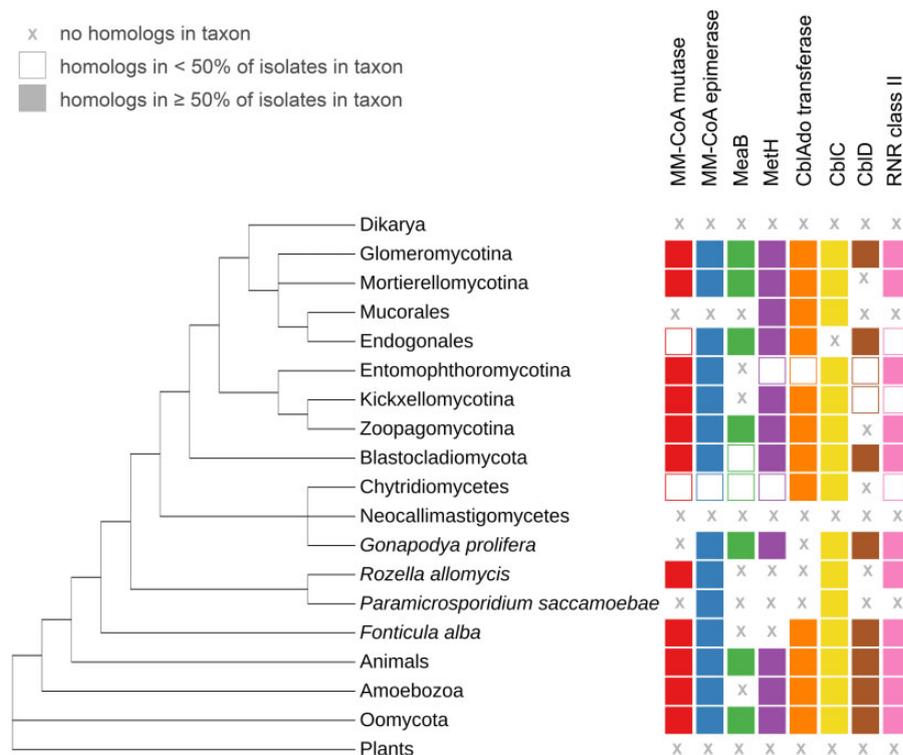
Fig. 1.—Distribution of B12-dependent protein families on a dendrogram showing a part of the eukaryotic tree of life, the schematic tree is based on Davis et al. (2019) and Spatafora et al. (2016) for fungi and on Burki et al. (2020) for remaining lineages. For each taxon, symbols on the right represent B12-dependent enzymes found in their proteome. X symbol means no identified homologs of the enzyme in the whole taxon; empty shape refers to the occurrence of the enzyme in less than half of studied representatives, filled shape means that half or more representatives have the enzyme in their proteomes.

approaches (supplementary data set DS1, Supplementary Material online), except for MeaB and CblAdo transferase (with highest numbers of identified homologs) where BA analyses did not converge to a reliable level of the standard deviation of split frequencies. We present an ML tree of MetH homologs as an example. All characteristic observations are common for all eight enzyme trees and are clearly visible on the MetH phylogenetic tree (fig. 2).

We noticed single bacterial sequences misannotated as fungal due to likely bacterial contamination of the fungal DNA samples. We also noticed single fungal sequences grouping within their bacterial relatives. In most cases these were proteins homologous to our enzyme yet with other function, for example, MeaB is similar to other GTPases (KAA6408927.1). Non-Dikarya fungal sequences rarely grouped with bacterial sequences with the exception of MM-CoA mutase from *Syncephalis pseudoplumigaleata* (RKP28319.1, RKP28318.1) which displayed a very high sequence identity reaching 100% with *Afipia* alphaproteobacteria which might indicate sample contamination. Notably, misannotated Dikarya sequences, like other Dikarya representatives, could not be found in Eukaryotic clades for the analyzed enzymes.

Exclusively fungal clades can be observed in five protein families (fig. 3). For the other three enzymes, there are clades composed mostly of fungal homologs and ones belonging to other eukaryotic microorganisms (Holozoa, Amoebozoa, and SAR). Observed topologies in the eukaryotic part of the trees generally are congruent with the species tree. Interestingly, in two cases (CblAdo transferase and RNR class II) Oomycota and Fungi clades are sisters to each other. Sequences identity of randomly chosen homologs is ~63% for RNR class II (ETI40368.1 and KNE69215.1) and ~52% for CblAdo transferase (XP_002997018.1 and KNE71581.1).

For MeaB, CblD, MetH, MM-CoA mutase, and RNR class II, fungal sequences form a sister clade to *Fonticula alba* (Holomycota), the closest relative of fungi belonging to Nucleariida. This pattern was observed for class II with the following score of aLTR support of the fungal clade: MeaB—1.00, CblD—1.00, MetH—1.00, MM-CoA mutase—1.00, and RNR class II—0.98. In the case of CblAdo transferase, CblC and MM-CoA epimerase fungal sequences group together with either ancient Metazoa representatives (CblAdo transferase) or with microbial eukaryotes from Holomycota, Amoebozoa, and SAR groups (MM-CoA epimerase). Importantly, non-Dikarya fungal sequences are always
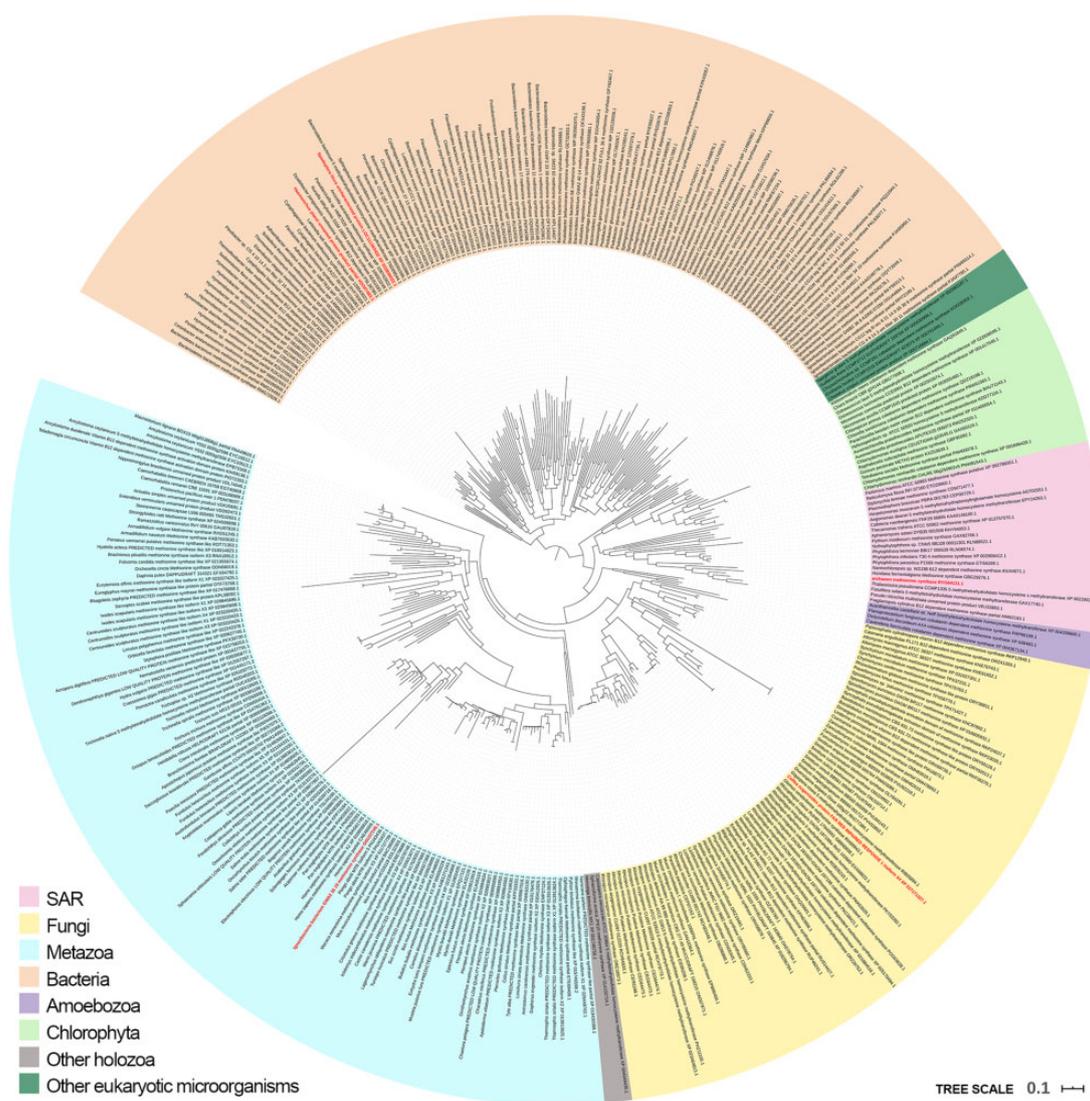
Fig. 2.—Phylogenetic tree of methionine synthase MetH homologs. The tree was built based on 72 sequences from non-Dikaryal proteomes, aligned with their homologs from NCBI nonredundant database 291 (see Materials and Methods). Sequences marked with red labels do not belong to organisms to which they were assigned.

sister to other Eukaryotic sequences which rules out bacterial contamination. Sequences from model organisms belonging to diverse lineages of microbial eukaryotes, not only Opisthokonta, were represented in these clades, including representatives from *Polysphondylium pallidum* and *Dictyostelium* spp. (Amebozoa), *Thecamonas trahens* (Apusozoa), *Chlamydomonas reinhardtii* (Chlorophyta), *Stentor coeruleus* (Alveolata), *Emiliania huxleyi* (Haptophyta), *Thalassiosira pseudonona*, and *Blastocystis* spp. (Heterokonta), *Naegleria gruberi* and *Euglena gracilis* (Excavata).

Discussion

Our discoveries contradict the current opinion that fungi neither synthesize nor use cobalamin (Duda et al. 1957; Martens et al. 2002) and do not have cobalt at all (Zhang et al. 2019). This claim remains true for Dikarya, but we demonstrate that the early-diverging fungi do have proteins that either process or use cobalamin as a cofactor.

In Eukaryotes three main metabolic pathways use cobalamin—RNR class II, MM-CoA mutase, and MetH pathways. Functions provided by these pathways are needed for the independent functioning of a living cell and can be lost in parasites (Zhang et al. 2009). Transport and trafficking of cobalamin in the cell is described in animals but homologs of the proteins responsible for the cobalamin transport, for example, LMBR1-like membrane protein transporters have a universal distribution in the Opisthokonta. This means that non-Dikarya fungi also have other components that are necessary for processing of cobalamin. Many of the enzymes involved in the
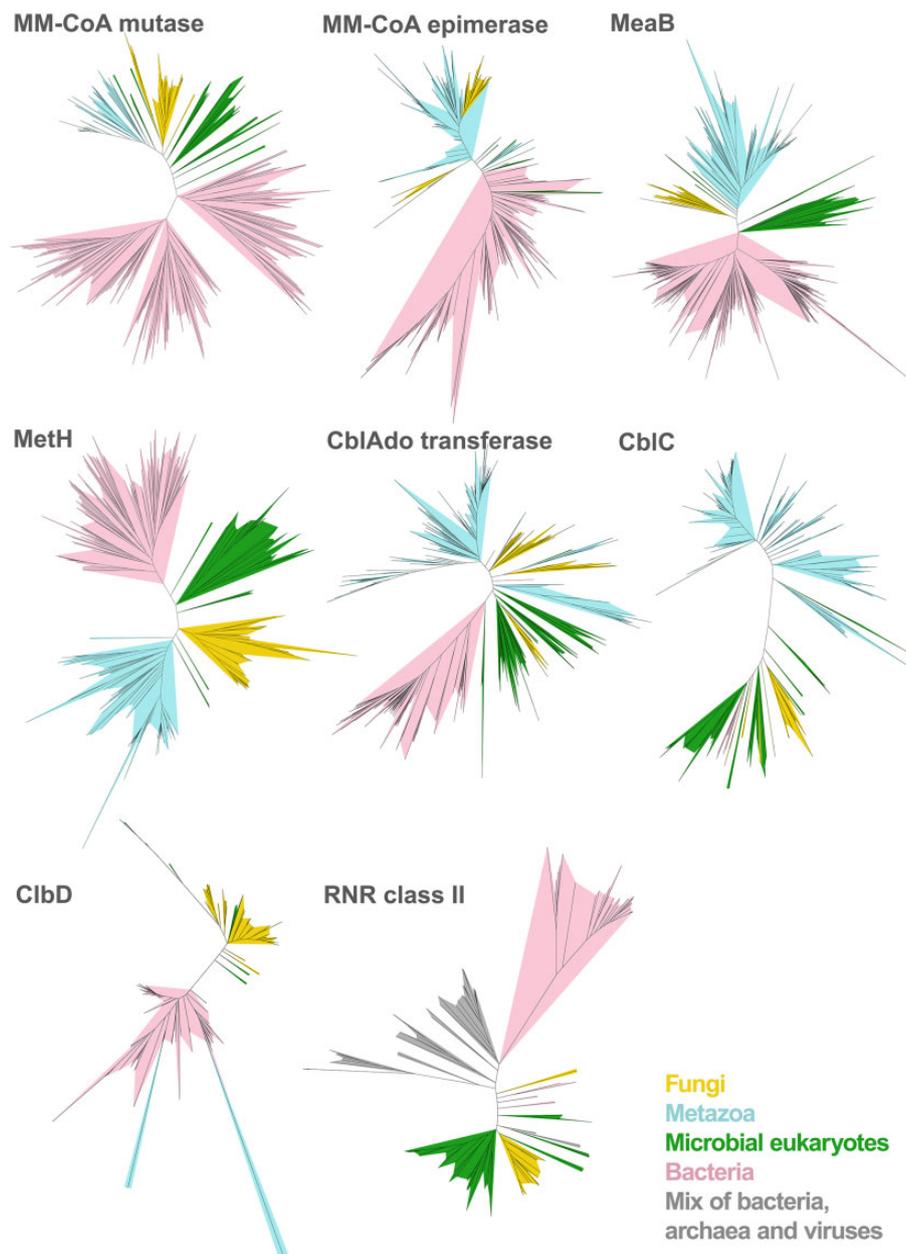
Fıɢ. 3.—Unrooted ML trees of eight B12-related protein family representatives.

MM-CoA mutase and MetH pathways, like mevalonate ki-nase and methionine synthase reductase, respectively, are conserved independently of B12 usage.

We found traces of all of these pathways among all line-ages of early-diverging fungi. The distribution of the genes encoding the above-mentioned enzymes is not uniform across the analyzed organisms. For non-Dikarya fungi, it is common to either have two out of three pathways or to have them incomplete.

Only Glomeromycotina and Blastocladiomycota have all three complete B12-dependent pathways. These two taxo-nomic groups are evolutionary and ecologically distant, they share only a few characteristics among others some of these fungi possess relatively big genomes. The latter may be a highlight of relaxed pressure on genome compactness.

The least conserved among fungal lineages is the RNR class II pathway. Organisms missing this class use cobalamin-independent RNR class I, which is the RNR used by animals. It is worth noticing that organisms are not limited to having only one class of RNRs at once (Jordan and Reichard 1998). Cobalamin-dependent RNR class II appears mostly in bacteria and, according to our results, also in non-Dikarya fungi and Oomycota. Additionally, RNR class II sequences from fungi and Oomycota form sister clades in phylogenetic trees (this

is also true for CblAdo transferase homologs). This may suggest an ancient horizontal gene transfer between Oomycota and fungi resulting in nonidentical but highly similar sequences. This is yet another parallel molecular trait that groups fungi and other filamentous fungi-like organisms together, next to similarities in weaponry to attack plants (Latijnhouwers et al. 2003), the evolution of the nitrate assimilation pathway (Ocaña-Pallarès et al. 2019), and the role of horizontal gene transfer (Soanes and Richards 2014; Rosewich and Corby Kistler 2000). This trait is exquisitely interesting because it is shared by eukaryotic microorganisms but is absent from big multicellular forms.

The best-conserved pathway in non-Dikarya fungi—MetH—can be substituted with a cobalamin-independent enzyme called MetE (González et al. 1996). We checked if this enzyme variant also can occur in non-Dikarya fungi proteomes. MetE is present in all non-Dikarya fungi phyla, even in Neocallimastigomycetes, which do not have any other cobalamin-dependent or independent alternatives of studied pathways. For some of the non-Dikarya fungi, lack of CblC protein can be observed. We did not look for substitutes for this protein, because the cooperation of CblC and CblD in the MetH pathway was described only for animals—outside this group, the exact function of CblD protein is not documented, and perhaps in other organisms, CblC is sufficient to perform its function by itself. One might speculate that other proteins are recruited to catalyze decyanation of cyanocobalamin and dealkylation of alkylcobalamins in nonanimal organisms.

MM-CoA mutase pathway is more or less conserved among early-diverging fungal lineages. Interestingly, all Mucorales members lack all three B12-dependent enzymes of that pathway. We checked for alternatives for this metabolism track and it turned up to be more complex than in the other two cases. In Dikarya propionate metabolism is carried out in the methylcitrate cycle (MCC). Three key enzymes for this track are methylcitrate synthase (MCS), methylcitrate dehydrogenase (MCD), and methylisocitrate lyase (MCL) (Dubey et al. 2013). All of them are present in Dikarya and, interestingly also in Choanoflagellida and Metazoa, but not in early-diverging fungal and other ancient lineages like Ichthyosporea and in Fonticula. MCS and MCL are conserved as well in old fungal phyla as in Dikarya, but that does not apply to MCD. Following information about the MCC gene cluster (Santos et al. 2020) genomic context of this pathway was checked for Batrachochytrium and Mucor representatives showing no synteny. Moreover, no candidate dehydrogenases were found upstream or downstream of MCS and MCL genes. We assume that the function of MCD can be taken over by other dehydrogenases.

According to our results, we can speculate the best-conserved elements of cobalamin-dependent pathways are key enzymes. For example, in the MetH pathway, the best-conserved element is MetH protein. On the contrary, it is quite common to lose CblC and CblD proteins from the proteome

(supplementary table S1, Supplementary Material online). The question is why in some organisms only the main part of pathways is conserved and how it is possible for these pathways to work without helper protein. We speculate that our results may be biased toward the main enzymes because they are well-known, especially have a well-known active site what allows for more rational data curation. Because our selection of potentially active homologs heavily relied on identified active site residues, it could have resulted in an underestimate of helper protein identification. Additionally, the MetH pathway is well described only in animal metabolism, so we cannot be sure about the role of CblC and CblD in fungal metabolism and about the necessity of having these proteins. On the other hand, the best-conserved protein in the MM-CoA mutase pathway is CblAdo transferase. For Mucorales, it is common to have only this one protein from the whole MM-CoA mutase pathway. It is worth noticing that this protein is responsible for synthesizing AdoCbl cofactor for MM-CoA mutase which is the only protein in fungal metabolism that is known to require the AdoCbl cofactor. The question is why in Mucorales proteomes there is still pressure to conserve CblAdo transferase while it is common to lose MM-CoA mutase.

Literature suggests that host-associated organisms have a tendency for the loss of cobalt utilization pathways (Zhang et al. 2009). Our results suggest that non-Dikarya fungi comply with this assumption. Chytridiomycota phylum combines amphibian parasites Batrachochytrium sp. and herbivorous mammals symbionts from class Neocallimastigomycetes. For these organisms, no cobalamin-dependent enzyme was found. These organisms may obtain the required resources from the host. However, our observations for plant-associated fungi are different. Mycorrhizal fungi from Glomeromycotina and Endogonales, despite maintaining extensive symbiotic relationships with 80% of plant species (Smith and Read 2010), retain well-conserved cobalamin-dependent pathways. It is possible that plant-associated non-Dikarya fungi kept these pathways simply because plant metabolism lacks cobalamin. The difference between plant and animal associated fungi may be a consequence of different pressures in such diverse ecological niches. Generally, parasites and obligate symbionts are biotrophs characterized by reduced genome size. However, in the case of mycorrhizal fungi, for a yet unknown reason, the pressure to reduce the genome seems to be relaxed (Lynch and Conery 2003; James et al. 2020).

The question that still remains is what is the source of cobalamin for fungi. We speculate that fungi are able to accumulate B12 acquired from bacterial sources. B12 cofactor supply for at least some of the fungi with B12-dependent enzymes may be mediated by endohyphal bacteria with an intact B12 synthesis pathway. All crucial components of the B12 de novo synthesis pathway required for such a relationship were found in the case of symbiosis between Glomeromycotina fungus *Gigaspora margarita* and β-proteobacterium

CandidatusGlomeribacter gigasporarum (Ghignone et al. 2012). Further experimental verification will be required to validate this hypothesis and confirm bacterial contribution to fungal B12 metabolism. There is also another open question of how fungi acquire the essential cofactor when growing in pure culture or in nature.

During the analysis of obtained results, we tried to understand the evolution of cobalamin-dependent metabolic pathways among kingdom Fungi. To widen the picture we checked studied proteomes for cobalamin-independent alternative metabolic pathways, and we confirmed their occurrence. Based on current knowledge we hypothesized that B12-dependent pathways are replaced by B12-independent alternatives in course of the evolution, and finally disappear in Dikarya lineages. In fact, the ability to utilize cobalamin is either retained or lost independently from the time of phyla divergence. Surprisingly, we observed a correlation between the preservation of this ability and fungal ecology. As we observe, cobalamin-dependent pathways are more common in fungi associated with plants, than in species associated with animals and living as soil saprophytes. Correlation like this is unclear for mycoparasites. In our data set, we have three fungi representing such a lifestyle and they have different enzyme distribution. In this case, we observed a correlation that mycoparasites which have B12-dependent enzymes infect fungi that also possess such enzymes (e.g., *Rozella allomycis*—*Allomyces* sp.). Similarly, parasites of organisms without B12-dependent genes do not have any of these enzymes themselves (*Caulochytrium protostelioides*—*Sordaria* sp.). This remark needs further investigation because we do not have enough data to form a reliable conclusion. Our discovery challenges the current view that fungi can neither synthesize nor utilize cobalamin. We proved that non-Dikarya fungal proteomes contain three metabolic pathways utilizing vitamin B12. We speculate these organisms have the possibility to accumulate cobalamin. Yet, these genomic observations need to be tested experimentally. Our discoveries may open the way for the selection of B12 over accumulating strains of food fermenting fungi without the need for genetic material manipulation.

B12 Enzymes in Other Eukaryotes

We also confirmed the occurrence of B12-related enzymes in other Holomycota taxa like Cryptomycota (Rozellida and Microsporidia) and Fonticulida. These organisms retain a maximum of only six out of the eight enzymes, but it is worth noting they are not independent, free-living organisms.

Studied enzymes are also present in Amoebozoa and Oomycota. Some other species from the SAR supergroup, to which Oomycota belongs (Burki et al. 2020), are known to have cobalamin-dependent methionine synthase (Boudouresque 2015). The matter is not clear about B12 utilization in Amoebozoa. There is contradictory information on

the necessity to supplement the culture of *Dictyostelium discoideum* with that vitamin (Stephan et al. 2003). In addition, class II RNR has been observed in *D. discoideum* previously (Crona et al. 2013). B12-dependent enzymes are encountered also in green algae (Chlorophyta), red algae (Rhodophyta) (Croft et al. 2005; Thi Vu et al. 2013), and Excavata (Helliwell et al. 2016). Green algae are known to acquire vitamin B12 through a symbiotic relationship with bacteria (Croft et al. 2005; Thi Vu et al. 2013).

Taken together, B12 dependence seems to be a widely distributed trait in Eukaryotes and was likely present in the last common ancestor of Eukaryotes. Several multicellular lineages including vascular plants and Dikarya developed B12-independent alternative pathways and, eventually, lost the B12 metabolism completely. The main question that remains about our discovery is the actual role of conserving B12-dependent enzymes in early-diverging lineages of fungi—is this dependency on cobalamin operative or is it just a relic from shared ancestry with animals and amoebae?

Materials and Methods

Fifty-nine predicted non-Dikarya proteomes were downloaded from NCBI in October 2019 (Sayers et al. 2020) (supplementary table S3, Supplementary Material online). Next, a pfam_scan.pl (default settings) (Mistry et al. 2007) search of all protein sequences against a library of Pfam HMMs was performed. To expand our data set NCBI NR database was searched for homologs of those non-Dikarya fungal B12-dependent proteins and additionally for homologs of proteins from model eukaryotic organisms with known B12-dependent enzymes (*Homo sapiens*, *Dictyostelium discoideum*, *Fonticula alba*, *Phytophthora infestans*) using PSI-BLAST (evalue = 0.001, num_iterations = 3) (Altschul et al. 1997). The data set was unified and clustered with CD-HIT ($n = 4$, $c = 0.7$, aS = 0.95, aL = 0.95), all fungal hits were retained regardless of their sequence similarity. To get only homologs of a protein of our interest, there was a need to discard homologs from related protein families. To do this we visualized protein pairwise similarity using CLANS (Frickey and Lupas 2004; Mistry et al. 2007) and selected separated groups of sequences. Dikarya sequences did not group together with the non-Dikarya-animal-protist clusters, except for single cases of clear contamination.

In the next step, sequences were aligned using local iterative mode in Mafft v. 3.7 (localpair, maxiterate = 100) (Katoh et al. 2002). The alignment was additionally cleared manually from potential inactive homologs. All sequences that showed a lack of amino acids crucial for enzyme activity or substitution of them with amino acids that are not able to maintain enzyme activity, were discarded from the set.

All alignments were trimmed with TrimAl (model = gappy-out) (Capella-Gutiérrez et al. 2009) to remove poorly conserved regions. Then, by using ProtTest (all-matrices,

all-distributions) (Abascal et al. 2005; Capella-Gutiérrez et al. 2009), we appointed the best amino-acid substitution models based on Akaike Information Criterion (AIC). Phylogenetic trees were built using LG model for each of the B12 metabolism-related enzymes with Bayesian (BA) and maximum likelihood (ML) approaches using MrBayes 3.2.7a x86_64 (Huelsenbeck and Ronquist 2001) and PhyML (Guindon et al. 2010) respectively. ML trees were estimated with a gamma distribution of rates between sites (four categories and alpha parameter estimated by PhyML) and aLRT χ^2-based parametric branch supports. In the course of BA inference, four Markov chains were run for three runs from random starting trees for 10^7 generations, and trees were sampled every 2.5×10^2 generations. The first one-fourth of generations were discarded as burn-in. Then, we used the remaining samples to calculate the tree of maximum clade credibility.

Expression of representatives of each of the protein sets (three randomly chosen homologs from each family) was confirmed by TBlastN (Altschul et al. 1990) (default settings) searches against the EST database at NCBI website.

Created phylogenetic trees were visualized and edited by iTOL v4 (Letunic and Bork 2019). Some bacterial sequences were misannotated as fungal due to likely bacterial contamination of the fungal DNA samples. A similar situation applies to some fungal sequences grouping within their bacterial relatives. In most cases, these proteins belong to another subfamily of homologous proteins with different substrate specificity. For instance, MeaB protein family groups diverse GTPases processing different substrates and only one of the subfamilies interacts with methylmalonyl-CoA mutase.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Author Contributions

A.M. designed the study. M.O. and A.M. prepared the data set and performed sequence analyses. M.O., K.S., and A.M. interpreted the data and wrote the manuscript.

Data Availability

The analyses are based on publicly available sequences. All identifiers of analyzed proteins are listed in Supplementary Material online. All trees built from those sequences are also listed in the Supplementary Material online.

Literature Cited

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21(9):2104–2105.

Ahrendt SR, et al. 2018. Leveraging single-cell genomics to expand the fungal tree of life. Nat Microbiol. 3(12):1417–1428.

Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389–3402.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410.

Bassila C, et al. 2017. Methionine synthase and methionine synthase reductase interact with MMACHC and with MMADHC. Biochim Biophys Acta. 1863(1):103–112.

Bobik TA, Rasche ME. 2001. Identification of the human methylmalonyl-CoA racemase gene based on the analysis of prokaryotic gene arrangements. Implications for decoding the human genome. J Biol Chem. 276(40):37194–37198.

Booker S, Stubbe J. 1993. Cloning, sequencing, and expression of the adenosylcobalamin-dependent ribonucleotide reductase from *Lactobacillus leichmannii*. Proc Natl Acad Sci U S A. 90(18):8352–8356.

Boudouresque C-F. 2015. Taxonomy and phylogeny of unicellular eukaryotes. In: Bertrand J-C, et al., editors. Environmental Microbiology: Fundamentals and Applications. Dordrecht: Springer. p. 191–257.

Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The new tree of eukaryotes. Trends Ecol Evol. 35(1):43–55.

Capella-Gutiérrez S, Marcet-Houben M, Gabaldón T. 2012. Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. BMC Biol. 10:47.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25(15):1972–1973.

Chang Y, et al. 2015. Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. Genome Biol Evol. 7(6):1590–1601.

Chatthanawaree W. 2011. Biomarkers of cobalamin (vitamin B12) deficiency and its application. J Nutr Health Aging. 15(3):227–231.

Chen ECH, et al. 2018. High intraspecific genome diversity in the model arbuscular mycorrhizal symbiont *Rhizophagus irregularis*. N Phytol. 220(4):1161–1171.

Coelho D, et al. 2008. Gene identification for the cblD defect of vitamin B12 metabolism. N Engl J Med. 358(14):1454–1464.

Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. 2005. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature 438(7064):90–93.

Crona M, et al. 2013. A rare combination of ribonucleotide reductases in the social amoeba *Dictyostelium discoideum*. J Biol Chem. 288(12):8198–8208.

Dardiotis E, et al. 2017. Vitamin B12, folate, and homocysteine levels and multiple sclerosis: A meta-analysis. Mult Scler Relat Disord. 17:190–197.

Davis WJ, et al. 2019. Genome-scale phylogenetics reveals a monophyletic Zoopagales (Zoopagomycota, Fungi). Mol Phylogenet Evol. 133:152–163.

Dereven'kov IA, Salnikov DS, Silaghi-Dumitrescu R, Makarov SV, Koifman OI. 2016. Redox chemistry of cobalamin and its derivatives. Coord Chem Rev. 309:68–83.

Dubey MK, Anders B, Jensen DF, Magnus K. 2013. Role of the methylcitrate cycle in growth, antagonism and induction of systemic defence responses in the fungal biocontrol agent trichoderma atroviride. Microbiology 159(Pt 12):2492–2500.

Duda J, Pedziwilk Z, Zodrow K. 1957. Studies on the appearance of vitamin B12 in leguminous plants. Acta Microbiol Pol. 6(3):233–238.

Forny P, Froese DS, Suormala T, Yue WW, Baumgartner MR. 2014. Functional characterization and categorization of missense mutations that cause methylmalonyl-coa mutase (MUT) deficiency. Hum Mutat. 35(12):1449–1458.

Frank S, et al. 2005. Anaerobic synthesis of vitamin B12: characterization of the early steps in the pathway. Biochem Soc Trans. 33(Pt 4):811–814.

Frickey T, Lupas A. 2004. CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics 20(18):3702–3704.

Froese DS, et al. 2010. Structures of the human GTPase MMAA and vitamin B12-dependent methylmalonyl-CoA mutase and insight into their complex formation. J Biol Chem. 285(49):38204–38213.

Gherasim C, Hannibal L, Rajagopalan D, Jacobsen DW, Banerjee R. 2013. The C-terminal domain of CblD interacts with CblC and influences intracellular cobalamin partitioning. Biochimie 95(5):1023–1032.

Ghignone S, et al. 2012. The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. ISME J. 6(1):136–145.

González JC, Peariso K, Penner-Hahn JE, Matthews RG. 1996. Cobalamin-independent methionine synthase from Escherichia coli: a zinc metalloenzyme. Biochemistry 35(38):12228–12234.

Grenville-Briggs LJ, et al. 2005. Elevated amino acid biosynthesis in phytophthora infestans during appressorium formation and potato infection. Fungal Genet Biol. 42(3):244–256.

Guindon S, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3):307–321.

Hannibal L, et al. 2009. Processing of alkylcobalamins in mammalian cells: a role for the MMACHC (cblC) gene product. Mol Genet Metab. 97(4):260–266.

Helliwell KE, et al. 2016. Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B12. Curr Biol. 26(8):999–1008.

Herrick J, Sclavi B. 2007. Ribonucleotide reductase and the regulation of DNA replication: an old story and an ancient heritage. Mol Microbiol. 63(1):22–34.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17(8):754–755.

James TY, Stajich JE, Hittinger CT, Rokas A. 2020. Toward a fully resolved fungal tree of life. Annu Rev Microbiol. 74(Sept):291–313.

Jordan A, Reichard P. 1998. Ribonucleotide reductases. Annu Rev Biochem. 67:71–98.

Katoh K, Misawa K, Kuma K-I, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30(14):3059–3066.

Kim J, Hannibal L, Gherasim C, Jacobsen DW, Banerjee R. 2009. A human vitamin B12 trafficking protein uses glutathione transferase activity for processing alkylcobalamins. J Biol Chem. 284(48):33418–33424.

Kolberg M. 2004. Structure, function, and mechanism of ribonucleotide reductases. Biochim Biophys Acta. 1699(1–2):1–34.

Kräutler B. 2012. Biochemistry of B12-cofactors in human metabolism. Subcell Biochem. 56:323–346.

Larsson K-M, et al. 2004. Structural mechanism of allosteric substrate specificity regulation in a ribonucleotide reductase. Nat Struct Mol Biol. 11(11):1142–1149.

Latijnhouwers M, de Wit PJGM, Govers F. 2003. Oomycetes and fungi: similar weaponry to attack plants. Trends Microbiol. 11(10):462–469.

Lerner-Ellis JP, et al. 2006. Identification of the gene responsible for methylmalonic aciduria and homocystinuria, cblC type. Nat Genet. 38(1):93–100.

Letunic I, Bork P. 2019. Interactive tree of life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47(W1):W256–W259.

Liu P, Stajich JE. 2015. Characterization of the carbohydrate binding module 18 gene family in the amphibian pathogen Batrachochytrium dendrobatidis. Fungal Genet Biol. 77(April):31–39.

Lynch M, Conery JS. 2003. The origins of genome complexity. Science 302(5649):1401–1404.

Mah W, et al. 2013. Subcellular location of MMACHC and MMADHC, two human proteins central to intracellular vitamin B(12) metabolism. Mol Genet Metab. 108(2):112–118.

Mancia F, et al. 1996. How coenzyme B12 radicals are generated: the crystal structure of methylmalonyl-coenzyme A mutase at 2 å resolution. Structure 4(3):339–350.

Marsh ENG, Román Meléndez GD. 2012. Adenosylcobalamin enzymes: theory and experiment begin to converge. Biochim Biophys Acta. 1824(11):1154–1164.

Martens J-H, Barg H, Warren MJ, Jahn D. 2002. Microbial production of vitamin B12. Appl Microbiol Biotechnol.

Mera PE, Escalante-Semerena JC. 2010. Multiple roles of ATP:cob(I)alamin adenosyltransferases in the conversion of B12 to coenzyme B12. Appl Microbiol Biotechnol. 88(1):41–48.

Mera PE, St Maurice M, Rayment I, Escalante-Semerena JC. 2007. Structural and functional analyses of the human-type corrinoid adenosyltransferase (PduO) from Lactobacillus reuteri. Biochemistry 46(48):13829–13836.

Mistry J, Bateman A, Finn RD. 2007. Predicting active site residue annotations in the Pfam database. BMC Bioinformatics 8(1):298.

Mondo SJ, et al. 2017. Widespread adenine N6-methylation of active genes in fungi. Nat Genet. 49(6):964–968.

Ocaña-Pallarès E, Najle SR, Scazzocchio C, Ruiz-Trillo I. 2019. Reticulate evolution in eukaryotes: origin and evolution of the nitrate assimilation pathway. PLoS Genet. 15(2):e1007986.

Outteryck O, et al. 2012. Methionine synthase deficiency: a rare cause of adult-onset leukoencephalopathy. Neurology 79(4):386–388.

Overath P, Kellerman GM, Lynen F, Fritz HP, Keller HJ. 1962. [On the mechanism of the rearrangement of methylmalonyl-Co A into succinyl-Co A. II. Experiments on the mechanism of action of methylmalonyl-Co A isomerase and methylmalonyl-Co A racemase]. Biochem Z. 335:500–518.

Rosewich UL, Corby Kistler H. 2000. Role of horizontal gene transfer in the evolution of fungi. Annu Rev Phytopathol. 38(Sept):325–363.

Roth JR, Lawrence JG, Rubenfield M, Kieffer-Higgins S, Church GM. 1993. Characterization of the cobalamin (vitamin B12) biosynthetic genes of Salmonella typhimurium. J Bacteriol. 175(11):3303–3316.

Santos LPA, et al. 2020. Propionate metabolism in a human pathogenic fungus: proteomic and biochemical analyses. IMA Fungus. 11:9.

Sayers EW, et al. 2020. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 48(D1):D9–D16.

Smith SE, Read DJ. 2010. Mycorrhizal symbiosis. New York: Academic Press.

Soanes D, Richards TA. 2014. Horizontal gene transfer in eukaryotic plant pathogens. Annu Rev Phytopathol. 52:583–614.

Spatafora JW, et al. 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia 108(5):1028–1046.

Stephan M, Beshay U, Friehs K, Flaschel E. 2003. Influence of medium composition on growth behaviour of Dictyostelium discoideum for cultivation on axenic media. Process Biochem. 39(3):333–343.

Tabima JF, et al. 2020. Phylogenomic analyses of non-Dikarya fungi supports horizontal gene transfer driving diversification of secondary metabolism in the amphibian gastrointestinal symbiont. G3 (Bethesda). 10(9):3417–3433.

Takahashi-Iñiguez T, González-Noriega A, Michalak C, Flores ME. 2017. Human MMAA induces the release of inactive cofactor and restores methylmalonyl-CoA mutase activity through their complex formation. Biochimie 142(Nov):191–196.

Thi Vu H, Itoh H, Ishii S, Senoo K, Otsuka S. 2013. Identification and phylogenetic characterization of cobalamin biosynthetic genes of *Ensifer adhaerens*. Microbes Environ. 28(1):153–155.

Uehling J, et al. 2017. Comparative genomics of *Mortierella elongata* and its bacterial endosymbiont *Mycoavidus cysteinexigens*. Environ Microbiol. 19(8):2964–2983.

Zhang Y, Rodionov DA, Gelfand MS, Gladyshev VN. 2009. Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. BMC Genomics 10(Feb):78.

Zhang Y, Ying H, Xu Y. 2019. Comparative genomics and metagenomics of the metallomes. Metallomics 11(6):1026–1043.

Associate editor: Jason Stajich