



Myxozoans (Cnidaria) do not Retain Key Oxygen-Sensing and Homeostasis Toolkit Genes

Allie M. Graham ^{1,2,*} and Felipe S. Barreto ¹

¹Department of Integrative Biology, Oregon State University, Corvallis, Oregon

²Current address: Department of Human Genetics, University of Utah, Salt Lake City, Utah

*Corresponding author: E-mail: graham.allie@gmail.com.

Accepted: 09 January 2023

Abstract

For aerobic organisms, both the hypoxia-inducible factor pathway and the mitochondrial genomes are key players in regulating oxygen homeostasis. Recent work has suggested that these mechanisms are not as highly conserved as previously thought, prompting more surveys across animal taxonomic levels, which would permit testing of hypotheses about the ecological conditions facilitating evolutionary loss of such genes. The Phylum Cnidaria is known to harbor wide variation in mitochondrial chromosome morphology, including an extreme example, in the Myxozoa, of mitochondrial genome loss. Because myxozoans are obligate endoparasites, frequently encountering hypoxic environments, we hypothesize that variation in environmental oxygen availability could be a key determinant in the evolution of metabolic gene networks associated with oxygen-sensing, hypoxia-response, and energy production. Here, we surveyed genomes and transcriptomes across 46 cnidarian species for the presence of HIF pathway members, as well as for an assortment of hypoxia, mitochondrial, and stress-response toolkit genes. We find that presence of the HIF pathway, as well as number of genes associated with mitochondria, hypoxia, and stress response, do not vary in parallel to mitochondrial genome morphology. More interestingly, we uncover evidence that myxozoans have lost the canonical HIF pathway repression machinery, potentially altering HIF pathway functionality to work under the specific conditions of their parasitic lifestyles. In addition, relative to other cnidarians, myxozoans show loss of large proportions of genes associated with the mitochondrion and involved in response to hypoxia and general stress. Our results provide additional evidence that the HIF regulatory machinery is evolutionarily labile and that variations in the canonical system have evolved in many animal groups.

Key words: HIF pathway, HIF α , EGLN, VHL, transcription factor, hypoxia.

Significance

Typically, in animals, two key molecular players control the response to cellular oxygen levels—the hypoxia-inducible factor (HIF) pathway and the mitochondria. Despite their importance, multiple instances of evolutionary loss of some of these components has occurred in certain invertebrate lineages. Cnidarians have a wide variety of mitochondrial chromosome phenotypes, including wholesale loss of this genome in the parasitic Myxozoa. Here we asked whether this variation was paralleled by loss of HIF pathway components. By mining genomic and transcriptomic data across 46 different cnidarian species, we found that the myxozoan lineage has an altered HIF pathway composition marked by the loss of its specific repression machinery, but cnidarian groups with a mitochondrial genome, regardless of shape, retained complete HIF toolkits.

© The Author(s) 2023. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

Metazoans have evolved a coordinated set of genes and pathways that control specialized mechanisms enhancing the maintenance of oxygen levels—the most prominent players are the hypoxia-inducible factor (HIF) pathway and the mitochondrion (Graham and Presnell 2017; Mills, et al. 2018). The roles of the HIF pathway and mitochondria are intertwined with numerous mechanisms by which HIF signaling is known to affect mitochondrial function, including aspects of mitochondrial metabolism, biogenesis, and mitophagy (Freeman, et al. 2006; Lin, et al. 2008; Schönenberger and Kovacs 2015). Specifically, the mitochondria and corresponding nuclear components of oxidative phosphorylation (OXPHOS) are responsible for energy production under aerobic conditions. While induction of HIF is known to be largely due to oxygen-sensing hydroxylases, mitochondrial processes themselves provide another path for sensing changes in oxygen levels (Bell, et al. 2005; Chandel and Schumacker 2000). This is possible via crosstalk between the HIF pathway and the mitochondria through the TCA cycle, electron transport system components, and mitochondrial respiration (Bao, et al. 2021; Poyton and McEwen 1996). In addition, the HIF pathway has direct effects on multiple mitochondrial functions, including mitochondrial oxidative capacity, OXPHOS, apoptosis, fission, and autophagy (Bao, et al. 2021). Mitochondrial oxygen consumption is thought to be a major determinant of HIF stability under hypoxia, and thus mitochondrial regulation of intracellular O₂ concentration is considered a more dominant mechanism regulating the cellular hypoxic response than previously thought (Chua, et al. 2010; Thomas and Ashcroft 2019).

The major transcription factor components of the HIF pathway (HIF α , and HIF β /ARNT) belong to a large gene family which contains 10 nuclear-encoded genes that have both bHLH (basic helix–loop–helix) and PAS (PER-ARNT-SIM) domains. In invertebrates, this family includes HIF α , Aryl hydrocarbon receptor nuclear translocator (ARNT), Nuclear Receptor Coactivator A (NCOA), Aryl Hydrocarbon Receptor (AhR), Neuronal PAS Domain Protein (NPAS1-4), Single-Minded (SIM), Circadian Locomotor Output Cycles Kaput (CLOCK) and methoprene-tolerant (Met) (Graham and Presnell 2017). The HIF pathway comprises a transcription factor (the heterodimer formed between HIF α and HIF β /ARNT), repression machinery and their downstream target genes. The repression machinery primarily comprises EGLNs (Egl-9 Family Hypoxia-Inducible Factor 1 & 3), and a E3 ubiquitin ligase complex that includes Von Hippel-Lindau Tumor Suppressor (VHL), Elongin B/C, and CUL2 (Cullin2) proteins. Prior work suggests that HIF α and EGLN co-occur in genomes (Rytönen et al., 2011), but that the presence of VHL is more evolutionarily labile—specifically if there is a loss of HIF α , a corresponding loss of VHL is not typically seen, because it has other protein

interactions beyond the HIF pathway (Graham and Barreto 2020). The beta member, HIF β /ARNT, dimerizes with other members of the bHLH-PAS gene family and is therefore not specific to the HIF pathway (fig. 1). However, only recently has genome-enabled studies examined the phylogenetic distribution of these genes and revealed that a few well-known animal lineages have lost the exclusive HIF pathway members HIF α and EGLN. These include tardigrades (Hashimoto, et al. 2016; Yoshida, et al. 2017), barnacles, and some copepod orders (Graham and Barreto 2020; Graham and Barreto 2019). These recent discoveries raise the question of what ecological and life-history conditions promote or permit the evolutionary loss of such essential cellular pathways. Yet, this question is unexplored and would benefit from more widespread sampling of animal taxa.

The mitochondrial component of oxygen homeostasis also shows wide genomic variation. Bilaterian mitochondrial DNA (mtDNA) is generally described as circular with a uniform size and gene content. However, there is considerable diversity in animal mitochondrial genome (mt-genome) organization, especially in nonbilaterian animal phyla like Cnidaria, Ctenophora, and Porifera (Lavrov and Pett 2016). Cnidarians have been shown to harbor a wide-range of mt-genome phenotypes, including the “traditional” single circular chromosome, multiple circular chromosomes, a single linear chromosome, or a highly fragmented genome (fig. 2) (Bridge, et al. 1992; Kayal, et al. 2012; Novosolov, et al. 2022; Osigus, et al. 2013; Smith, et al. 2012; Yahalomi, et al. 2017). The mitochondrion is also typically described as being essential and omnipresent in contemporary eukaryotes, but mitochondrial genomes in “early-branching” eukaryotes are frequently reduced (Santos, et al. 2018), or lost altogether (John, et al. 2019; Karnkowska, et al. 2016). Recently, a cnidarian clade (Myxozoa) was shown to harbor one species that completely lost their mitochondrial genome (Yahalomi, et al. 2020), perhaps mediated by portions of their life-stages occurring within anoxic tissues (Famme and Knudsen 1985; Fish 1939; Johnston, et al. 1975; Yahalomi, et al. 2020). Therefore, we hypothesize that this clade may have lost key metazoan genes involved in oxygen-sensing and response, including the master regulators of the HIF pathway. Documenting variation in mitochondrial and HIF pathway structure broadly across animal taxa is an important step for testing predictions about how environmental pressures affect genome evolution.

Here, we mined for the presence of key HIF pathway members (HIF α , EGLN, and VHL) across 46 publicly available Cnidaria genomic and transcriptomic resources (Table 1). We also sought to quantify how other hypoxia-related, mitochondria-related and stress-related nuclear genes were represented across Cnidaria. Ultimately, we found the HIF pathway to be fully retained in most cnidarian groups, including Anthozoa, Cubozoa, Hydrozoa, Scyphozoa and

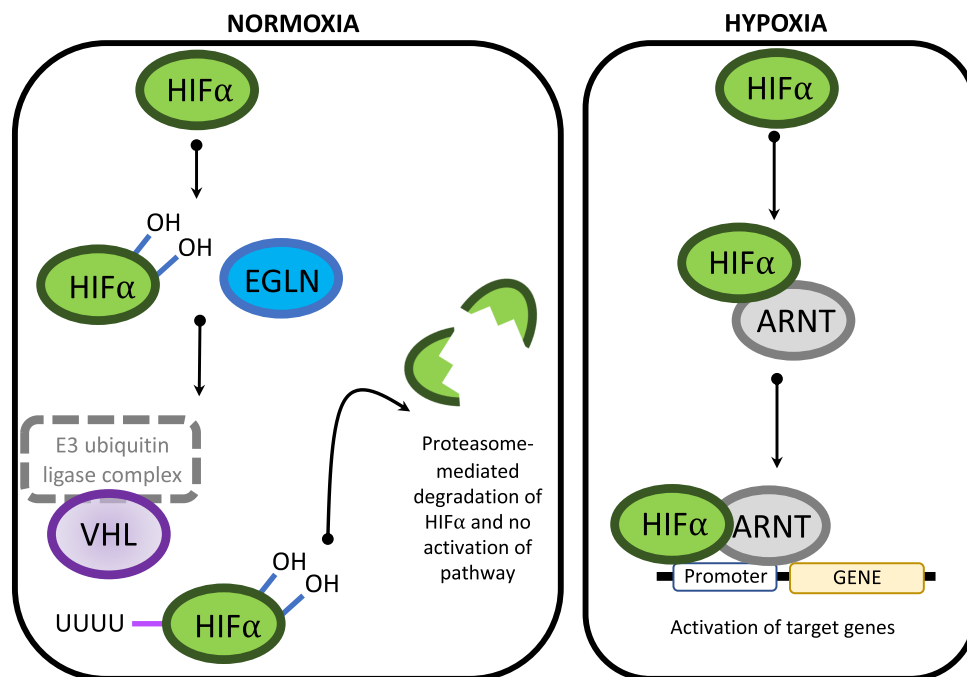


Fig. 1.—An overview of the HIF pathway under normoxia and hypoxia (adapted from Liao and Zhang 2020). Under normoxia, the HIF α subunit is hydroxylated by EGLN and ubiquitinated by VHL + E3 ubiquitinase ligase complex/ECV (dotted box). Under hypoxia, the HIF α subunit dimerizes with ARNT and binds to promoter regions affecting transcription. The proteins HIF α , EGLN, and VHL were surveyed in this study.

Staurozoa, but not in the Myxozoa, in which all representatives examined appear to have lost both oxygen-sensing HIF α repressors (EGLN and VHL). We also detected an apparent loss of EGLN but not HIF α or VHL in the parasitic *Polypodium hydriforme*, but this finding is more uncertain because this is the only one species in this class. In addition, myxozoans, regardless of mt-genome presence or absence, seem to have lost several other hypoxia and oxygen-related toolkit genes, as well as a significant subset of genes related to general mitochondrion function. Silhouette images are from Phylopic (CC BY-SA 3.0).

Results

De Novo Assembly Quality Assessment

Among the 46 protein sequence datasets we used in our analyses, 27 were assembled de novo from Illumina paired-read data obtained from NCBI. We assessed the quality of the transcriptomes that we assembled de novo to avoid biases in our power to detect loss or retention across groups. Based on BUSCO scores calculated for all 46 protein datasets, those based on transcriptomes we assembled showed no significant difference in proportions of complete genes (C) and of missing genes (M) compared to datasets from published genomes (supplementary Table S6, Supplementary Material online; Mann–Whitney U tests; C: $W=216$, $P=0.372$; means: transcriptomes = 85.5%, genomes = 76.6%; M: $W=281$, $P=0.592$; transcriptomes = 8.8%, genomes = 17.6%).

Myxozoa data sets had significantly more missing genes than the rest of cnidarians (supplementary Table S6, Supplementary Material online; M: $W=363$, $P=1.2 \times 10^{-5}$; means: Myxozoa = 36.9%, others = 4.8%). Nevertheless, myxozoan transcriptomes we assembled had better recovery of BUSCO gene content than those from myxozoan genomes (supplementary Table S6, Supplementary Material online; M: $W=28$, $P=0.0106$; means: transcriptomes = 22.6%, genomes = 61.9%).

Based on mapping of original paired-reads to their respective transcriptome assemblies, we observed a high level of concordant read-pair alignments across the 27 de novo assemblies (supplementary Table S2, Supplementary Material online; range: 70.4–97.9%; mean: 90.9%). Myxozoa assemblies had similar levels of read-pair alignments (range: 80–97.4%; mean: 89.6%) compared with those from the other taxonomic groups (range: 70.4–97.9%; mean: 91.5%; $W=28$, $P=0.39$). Overall, these assessments suggest that the transcriptome assembly pipeline we employed generated high-quality transcript sequences across all groups, and that BUSCO scores in the Myxozoa data are likely reflective of true widespread gene loss instead of poorer assemblies.

Removal of Host Contaminants From Parasite Transcriptomes

All 11 myxozoan and the *Polypodium* protein assemblies showed evidence of host contamination. Using the e-value

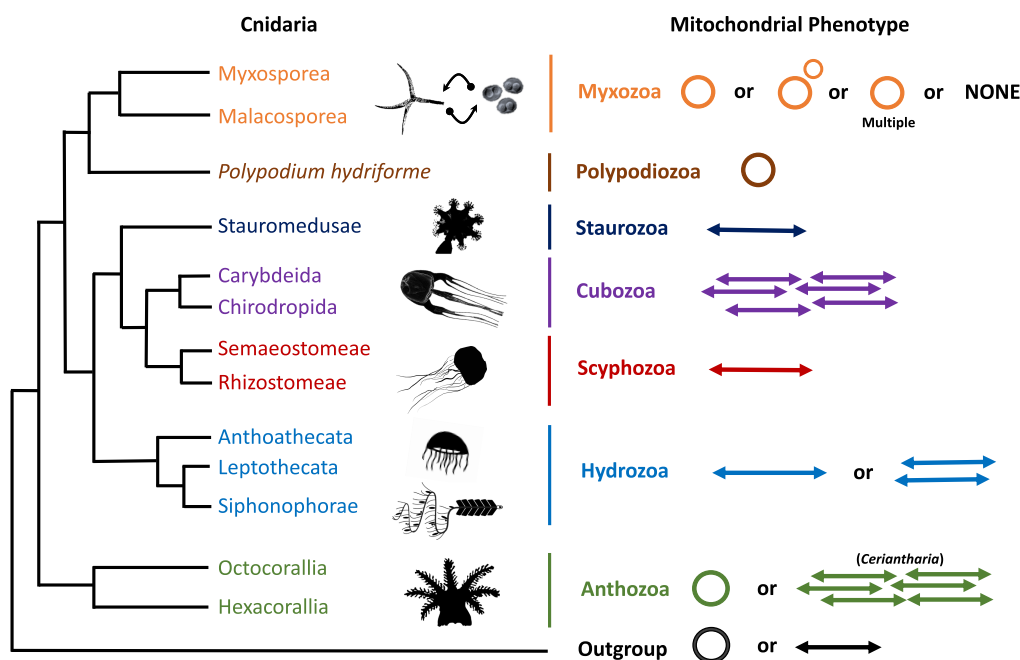


FIG. 2.—Variation in mitochondrial genome (mtDNA) phenotypes among Cnidaria, including linear, fragmented linear, single circular, or multiple circular. The Cnidaria phylogenetic tree is based on several sources (Kayal, et al. 2018; Okamura and Gruhl 2016) and the mtDNA phenotypes are from a variety of studies (Bridge, et al. 1992; Kayal, et al. 2012; Lavrov and Pett 2016; Novosolov, et al. 2022; Okamura and Gruhl 2016; Osigus, et al. 2013; Smith, et al. 2012; Yahalomi, et al. 2017).

cutoff of $1e^{-75}$, a portion of the transcriptomes matched sequences from hosts, including *Tetracapsensis bryosalmonae* (45.4% of its transcripts), *Sphaerospora molnari* (30.5%), *Ceratonova shasta* (12.9%), followed by *Buddenbrockia plumatellae* (10.2%), *Kudoa iwatai* (8.2%), the three *Myxolobus* species ($< 2.3\%$), *Henneguya salminicola* (2.3%), *Thelohanellus kitauei* (1.2%) and *Polypodium* (0.26%). These were removed from their respective gene assemblies before the HIF search analyses. However, even after this initial filter, there was evidence that some host proteins were still found in the assemblies; in the final phylogenetic trees of bHLH-PAS and P4HC proteins, some myxozoan proteins that grouped with HIF α or EGLN were subjected to additional BLAST interrogations, and all showed $>90\%$ identity with bony fishes—these were thus removed from the alignments and the trees estimated again. Future protocols for assembling transcriptomes from parasitic taxa should assess an appropriate e-value cutoff for efficient removal of contaminant host reads.

Retention of HIF Pathway Members

The HIF pathway was identified via the presence of the three regulatory gene members (HIF α , EGLN, and VHL) with a complementary combination of homology and motif searches as well as phylogenetic clustering. We detected the clear presence of HIF α protein sequence and at least one of the

repressors across all representative species examined of Anthozoa, Cubozoa, Hydrozoa, Scyphozoa, and Staurozoa (supplementary Table S6, supplementary figs. S1–S8, Supplementary Material online). Therefore, the canonical mechanism of the HIF pathway appears to be conserved in these groups. Within the Myxozoa, in contrast, we detected the presence only of HIF α . Eight of the eleven species examined exhibited HIF α (fig. 3; supplementary Table S6, Supplementary Material online), so this gene may be considered present in this group. However, we found no evidence for the presence of either EGLN or VHL in any of the eleven myxozoan species examined through any of the methods we used, except for the malacosporeans (fig. 4, supplementary Table S6, Supplementary Material online). The eight myxozoans in which HIF α was identified were *Henneguya salminicola*, *Sphaerospora molnari*, *Tetracapsuloides bryosalmonae*, *Thelohanellus kitauei*, *Ceratonova shasta*, *Myxobolus cerebralis*, *M. pendula*, and *M. honghuensis* (fig. 3; supplementary Table S6, Supplementary Material online). These were not the result of host contamination, since the taxonomy of significant BLAST hits were cnidarian in origin (e.g., top hit for *H. salminicola* was *Exaiptasia diaphana* HIF α ; for *S. molnari* was *Hydra vulgaris* HIF α ; for *M. cerebralis* was *E. diaphana* HIF α). The protein alignments and phylogenetic clustering of Myxozoa also include proteins from *Polypodium hydriforme*, which is the only species in the class Polypodiozoa

Table 1

List of Taxa Used in This Study. “G” and “T” Next to Species Names Denote, Respectively, Whether Protein Sequences Were Obtained From Publicly Available Genome Resources or Were Assembled *de Novo* in This Study Using raw Data Available From NCBI

Taxonomic classification	Species	Number of proteins in assembly
Anthozoa; Hexacorallia	<i>Exaiptasia pallida</i> (G)	27 753
Anthozoa; Hexacorallia	<i>Actinia tenebrosa</i> (G)	27 037
Anthozoa; Hexacorallia	<i>Acropora millepora</i> (G)	35 355
Anthozoa; Hexacorallia	<i>Pocillopora damicornis</i> (G)	25 183
Anthozoa; Hexacorallia	<i>Montipora capitata</i> (G)	41 863
Anthozoa; Hexacorallia	<i>Stylophora pistillata</i> (G)	33 252
Anthozoa; Hexacorallia	<i>Orbicella faveolata</i> (G)	32 587
Anthozoa; Hexacorallia	<i>Acropora digitifera</i> (G)	33 878
Anthozoa; Octocorallia	<i>Dendronephthya gigantea</i> (G)	28 741
Cubozoa; Carybdeida	<i>Alatina alata</i> (T)	77 006
Cubozoa; Carybdeida	<i>Morbakka virulenta</i> (G)	28 983
Cubozoa; Carybdeida	<i>Tripedalia cystophora</i> (T)	48 205
Cubozoa; Carybdeida	<i>Copula sivickisi</i> (T)	113 030
Cubozoa; Chirodropida	<i>Chironex yamaguchii</i> (T)	44 573
Hydrozoa; Anthoathecata	<i>Podocoryna carnea</i> (T)	70 805
Hydrozoa; Anthoathecata	<i>Ectopleura larynx</i> (T)	22 753
Hydrozoa; Anthoathecata	<i>Hydra vulgaris</i> (G)	21 990
Hydrozoa; Anthoathecata	<i>Hydractinia polyclina</i> (T)	94 528
Hydrozoa; Anthoathecata	<i>Hydractinia symbiolongicarpus</i> (T)	49 635
Hydrozoa; Leptothecata	<i>Dynamena pumila</i> (T)	28 302
Hydrozoa; Leptothecata	<i>Aequorea victoria</i> (T)	29 698
Hydrozoa; Leptothecata	<i>Clytia hemisphaerica</i> (G)	25 087
Hydrozoa; Siphonophorae	<i>Athorybia rosacea</i> (T)	56 888
Hydrozoa; Siphonophorae	<i>Craseoa lathetica</i> (T)	80 100
Hydrozoa; Siphonophorae	<i>Nanomia bijuga</i> (T)	94 103
Myxozoa; Myxosporea	<i>Myxobolus cerebralis</i> (T)	31 114
Myxozoa; Myxosporea	<i>Myxobolus pendula</i> (T)	80 650
Myxozoa; Myxosporea	<i>Myxobolus squamalis</i> (G)	5 589
Myxozoa; Myxosporea	<i>Myxobolus honghuensis</i> (G)	11 071
Myxozoa; Myxosporea	<i>Henneguya salminicola</i> (G)	7 999
Myxozoa; Myxosporea	<i>Thelohanellus kitauei</i> (G)	14 848
Myxozoa; Myxosporea	<i>Ceratonova shasta</i> (T)	74 774
Myxozoa; Myxosporea	<i>Sphaerospora molnari</i> (T)	54 847
Myxozoa; Myxosporea	<i>Kudoa iwatai</i> (T)	74 070
Myxozoa; Malacosporea	<i>Buddenbrockia plumatellae</i> (T)	22 810
Myxozoa; Malacosporea	<i>Tetracapsuloides bryosalmonae</i> (T)	46 784
Polypodiozoa; Polypodiidea	<i>Polypodium hydriforme</i> (T)	34 620
Scyphozoa; Rhizostomeae	<i>Cassiopea xamachana</i> (G)	337 695
Scyphozoa; Rhizostomeae	<i>Nemopilema nomurai</i> (T)	77 952
Scyphozoa; Rhizostomeae	<i>Stomolophus meleagris</i> (T)	19 942
Scyphozoa; Semaestomeae	<i>Pelagia noctiluca</i> (T)	35 998
Scyphozoa; Semaestomeae	<i>Sanderia malayensis</i> (T)	26 147
Scyphozoa; Semaestomeae	<i>Aurelia aurita</i> (G)	38 007
Staurozoa; Stauromedusae	<i>Calvadosia cruxmelitensis</i> (G)	74 577
Staurozoa; Stauromedusae	<i>Craterolophus convolvulus</i> (T)	40 661
Staurozoa; Stauromedusae	<i>Lucernaria quadricornis</i> (T)	38 721

and is considered a sister clade to Myxozoa (Novosolov et al. 2022). This species showed no evidence for the presence of EGLN, but there were sequences which we identified as likely HIF α and VHL.

Within bHLH-PAS-containing families, gene loss within Myxozoa was not restricted to the HIF pathway. Additional

losses, based on phylogenetic clustering, include *Ahr/AHRR* (Aryl hydrocarbon receptor), *ARNTL* (Aryl Hydrocarbon Receptor Nuclear Translocator), *CLOCK* (circadian rhythm), *NCOA1/2/3* (Nuclear Receptor Coactivator), *NPAS1/2/3/4* (Neuronal PAS Domain Protein), *SIM1/2* (Single-Minded), and *MET* (Methoprene-tolerant) (fig. 3).

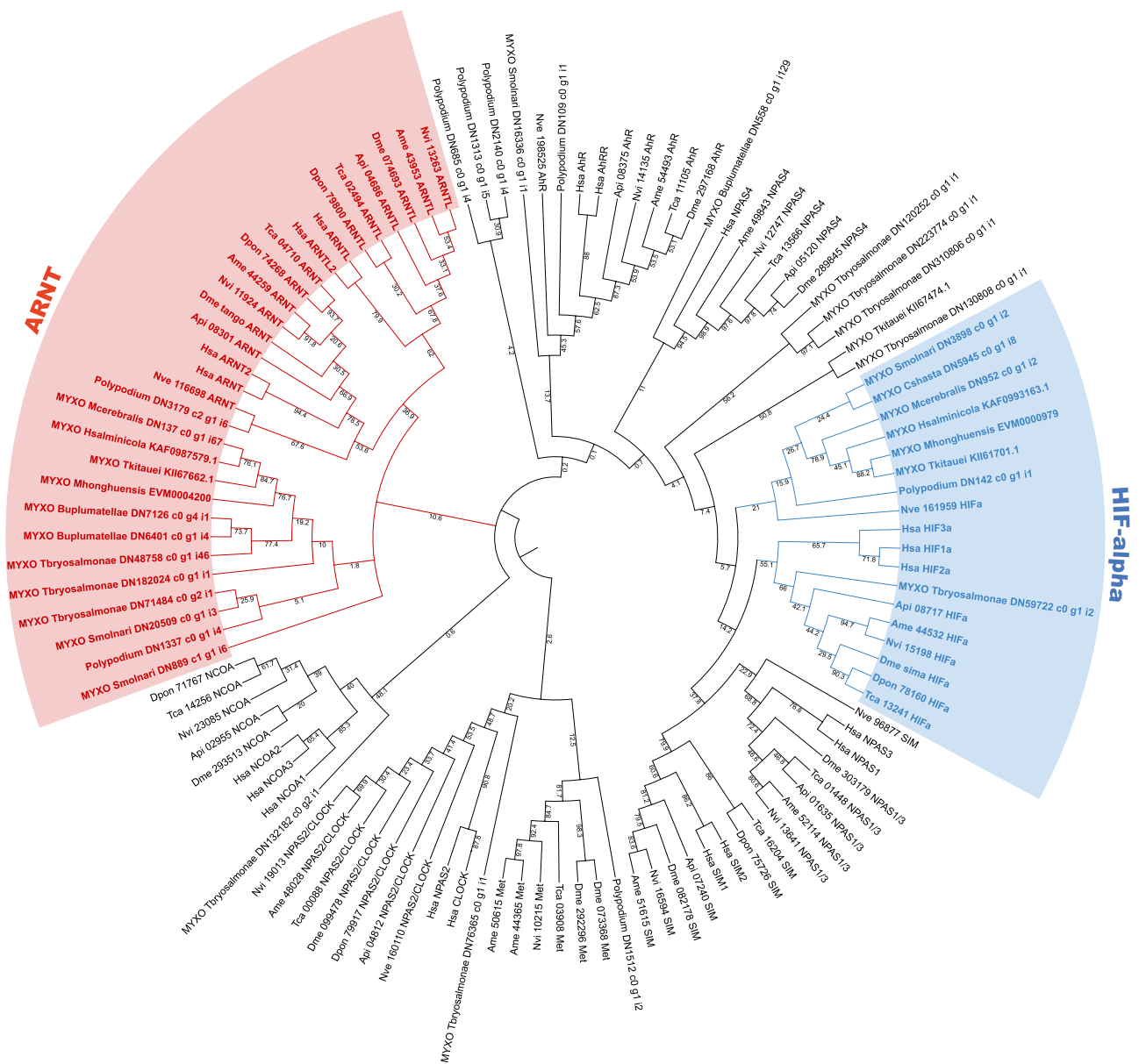


Fig. 3.—Maximum-likelihood phylogenetic tree of bHLH-PAS proteins from Myxozoa ($n = 11$ species) and *Polypodium hydriforme* (RAxML; 1 000 bootstraps). Myxozoa taxa are labeled with “MYXO” and abbreviated with the first letter of the genus followed by species epithet. Additional “anchor” sequences (in blue or red font) include known bHLH-PAS-containing proteins from insects (Api = *Acyrtosiphon pisum*, Dme = *Drosophila melanogaster*, Ame = *Apis mellifera*, Nvi = *Nasonia vitripennis*, Tca = *Tribolium castaneum*, Dpon = *Dendroctonus ponderosae*), the cnidarian *Nematostella vectensis*, and human. Grouping of HIF α and ARNT protein are shown in boxes. Numbers next to branches are bootstrap support values. See Table 1 for complete list of cnidarian species sampled.

Abundance of Stress and Mitochondrion-related Genes

Of the 50 “response to hypoxia” genes enumerated from the *N. vectensis* genome, myxozoans have retained on average 72% (52–88%) compared to an average of 98% (90%—100%) in the rest of the Cnidaria groups (supplementary Table S6, Supplementary Material online). Four of those genes were lost in all myxozoan species examined, while two of

those were present in *Polypodium*. These genes encode mitochondrial import inner membrane translocase subunit (*TIM44*; Q8IHE3_DROME), and complex I intermediate-associated protein 30 (*CIA30*; CIA30_DROME) (supplementary Table S3, Supplementary Material online).

Of the 692 “mitochondrion” genes, myxozoans have retained on average 47.8% (16.9–69.7%) compared to

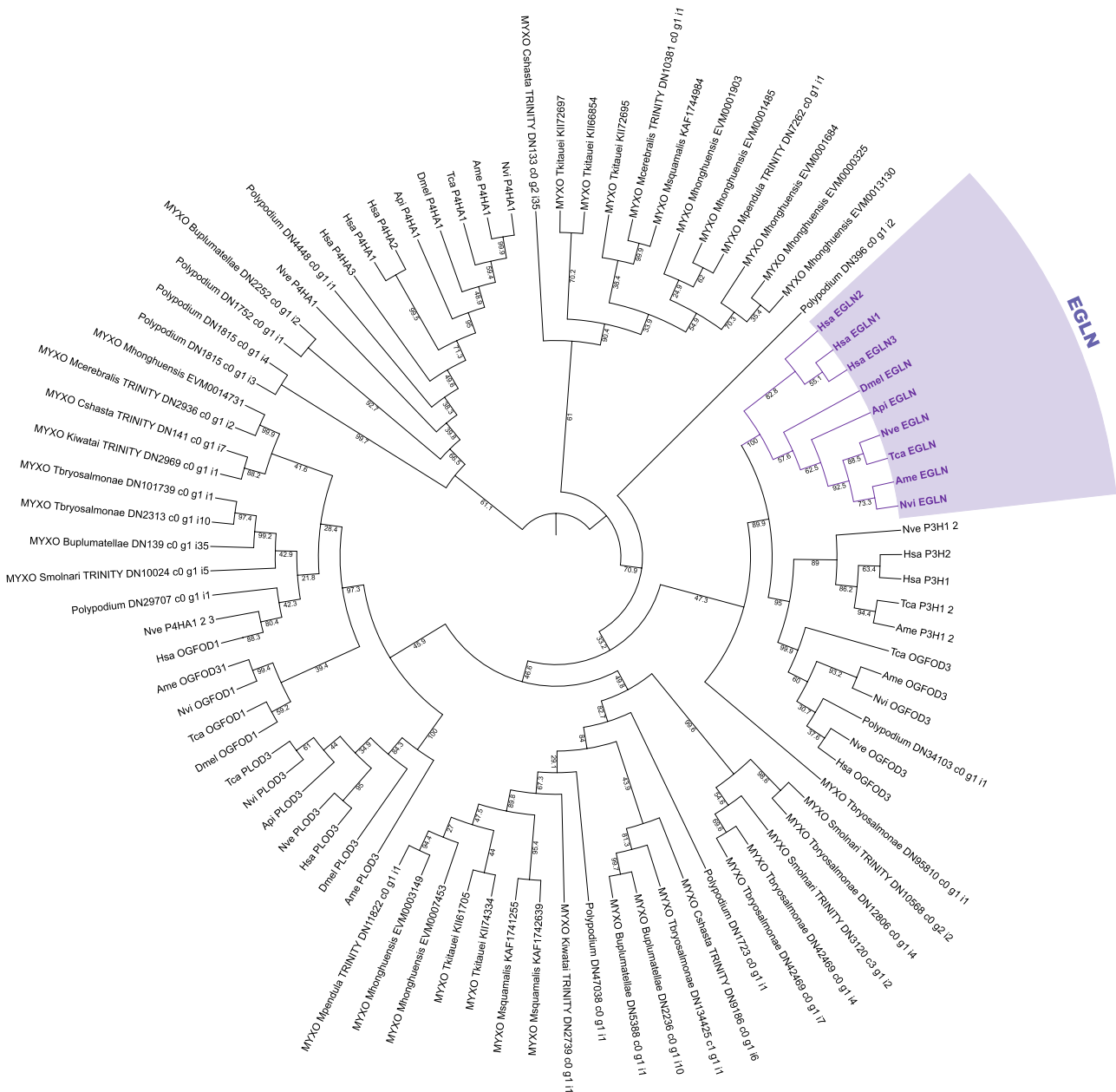


Fig. 4.—Maximum-likelihood phylogenetic tree of P4HC proteins from Myxozoa ($n = 11$ species) and *Polypodium hydriforme* (RAxML; 1 000 bootstraps). Myxozoa taxa are labeled with “MYXO” and abbreviated with the first letter of the genus followed by species epithet. Additional “anchor” sequences (in purple font) include known P4HC-containing proteins from insects (Api=*Acyrtosiphon pisum*, Dme=*Drosophila melanogaster*, Ame=*Apis mellifera*, Nvi=*Nasonia vitripennis*, Tca=*Tribolium castaneum*, Dpon=*Dendroctonus ponderosae*), the cnidarian *Nematostella vectensis*, and human. Grouping of EGLN is shown in the box. Numbers next to branches are bootstrap support values. See Table 1 online for complete list of cnidarian species sampled.

an average of 92.5% (79.3–96.7%) in the rest of the cnidarian groups (supplementary Table S6, Supplementary Material online). Seventy-five of those genes were lost in all of the Myxozoa species examined. Of those missing in myxozoans, 53 were missing from *Polypodium*; there were also 40 instances of loss in *Polypodium* which were found at least once in Myxozoa, although this

may be due to only one *Polypodium* representative being available. The GO terms associated with the missing genes were mitochondrial fission, mitochondrial DNA replication, protein insertion into mitochondrial inner membrane, and protein import into mitochondrial matrix (supplementary Table S4, Supplementary Material online).

Finally, of the 576 “response to stress” genes, myxozoans retained on average 46% (20–83%) of genes while the other cnidarian groups retained an average of 97% (86–99%) of genes (supplementary Table S6, Supplementary Material online). Thirty-six of those genes were not present in any representative myxozoan we examined but were present in *Polypodium*. While 102 were lost in both Myxozoans and *Polypodium* (supplementary Table S5, Supplementary Material online). The groups of genes lost in Myxozoa includes genes encoding various KICSTOR complex proteins, GATOR complex proteins, INSIG2 (insulin-induced gene 2 protein), NBN (Nibrin), among others.

Discussion

The key mechanisms associated with oxygen-sensing and homeostasis have largely been thought to be conserved within Metazoa. At the center of these cellular processes is the HIF pathway. Recent work has shown various instances of organismal lineages which have lost this crucial pathway yet are still able to modulate response to oxygen levels (Graham and Barreto 2020; Graham and Barreto 2019). Another key, but parallel cornerstone of these cellular processes is the mitochondrion and its associated machinery. Loss of mitochondrial genomes had previously been documented in specific protist groups (Santos, et al. 2018), with no Metazoan examples until recently (fig. 2) —a myxozoan species, *H. salminicola*, was shown to have lost the mitochondrial genome and portions of the nuclear genome associated with mitochondrial functions (Yahalom, et al. 2020). In lineages that have reduced or lost their mitochondrial genome, these evolutionary events have been attributed to pressures from their environment,

which is largely anoxic (Maguire and Richards 2014; Stairs, et al. 2015). For *H. salminicola* specifically, their parasitic life-stages occur within anoxic tissues (Fish 1939; Johnston, et al. 1975; Yahalom, et al. 2020).

Using genomes and transcriptomes of 46 different cnidarians across the 7 major classes of the phylum, our results suggest that alteration of the HIF pathway and loss of associated toolkit genes is restricted to the Myxozoa (fig. 5). We show evidence that all cnidarian groups have retained HIF α and at least one of its repressors, but that the Myxozoa appear to have lost both repressors EGLN and VHL as well as a large proportion of genes associated with stress response and oxygen homeostasis. Given the incomplete nature of and variation across transcriptomic data, “missing” genes in some of these species are artifacts of potential data incompleteness; therefore, our approach was to deem evolutionary loss of a HIF pathway gene only when all sampled species of a subgroup failed to reveal a gene sequence. This was the case for *EGLN* and *VHL* in Myxozoa, and we hence conclude that these repressors were evolutionarily lost in this clade. The singular member of the monotypic genus *Polypodium* (*P. hydriforme*) is also missing *EGLN*. *Polypodium* has both a free-living and parasitic life stage, which ultimately infects eggs within mature fish species of Acipenseriformes (Raikova, et al. 1979; Raikova 1994). Hence our finding may suggest that switching to a parasitic lifestyle facilitated initial alteration of the HIF pathway, which was then further modified in the Myxozoa.

Our results also indicate that HIF α is encoded in Myxozoa genomes, including the species known to have lost its full mitochondrial genome (*H. salminicola*). Although three sampled myxozoan species had a large amount of host contamination, the putative HIF α protein sequence from each were not of host origin and were most similar to myxozoan and other cnidarian sequences found in public databases. All of the putative HIF α sequences were confirmed through InterProScan annotations and phylogenetic grouping with invertebrate HIF α sequences (including that of the model anemone *N. vectensis*). This suggests that HIF α and HIF β /ARNT, which is also present in Myxozoa, may still be able to dimerize. However, without its associated repression machinery (EGLN and VHL), it is uncertain how active or functional the canonical HIF pathway is in the clade.

We also assessed the degree to which gene groups functionally associated with the HIF pathway were retained across groups, including genes involved in “response to hypoxia”, “mitochondrion” and “response to stress” (supplementary Tables S3–S6, Supplementary Material online). In myxozoan genomes, prior work has documented depletion of genes associated with development, cell differentiation, and cell-to-cell communication, including hallmark pathways like the *Wnt*, *Hedge* and *TGF β* (Chang, et al. 2015). In addition, *H. salminicola* showed wholesale loss of mtDNA-encoded subunits, as well as major reductions in corresponding nuclear

		HIF α	EGLN	VHL
Hydrozoa	Anthoathecata	Y	Y	Y
	Leptothecata	Y	Y	Y
	Siphonophorae	Y	Y	Y
Scyphozoa	Semaeostomeae	Y	Y	Y
	Rhizostomeae	Y	Y	Y
Staurozoa	Stauromedusae	Y	Y	Y
Cubozoa	Carybdeida	Y	Y	Y
	Chirodropida	Y	Y	Y
Polypodiozoa	<i>Polypodium</i>	Y	N	Y
	Malacosporae	Y	N	Y
Myxozoa	Myxosporea	Y	N	N
	Octocorallia	Y	Y	Y
Anthozoa	Hexacorallia	Y	Y	Y

Fig. 5.—Summary of results of our scan for HIF pathway genetic elements based on HMMER, BLAST, and phylogenetic grouping. Presence and absence of these genes are denoted with Y (Yes) or N (No), with results for Myxozoa + *Polypodium* in bold font.

subunits for most OXPHOS complexes (Complex I, III, IV), but largely retained associated metabolic mitochondrial pathways, including those involved in amino acid, carbohydrate, nucleotide metabolism (Yahalomi, et al. 2020). In our assessment of eleven myxozoan species for “mitochondrial” genes, most were mitochondrially-embedded by nuclearly encoded members and a variety of mitochondrial ribosomes (supplementary Table S4, Supplementary Material online). On average, only 72% genes typically associated with “hypoxia” were retained in myxozoans compared to other cnidarians. In this dataset, across myxozoans, we found evidence of potential wholesale loss of four genes, among which both *TIM44* and *NDUFAF1* are directly associated with the mitochondria by providing protein transport into the mitochondrial matrix (Chacinska, et al. 2002; Williamson, et al. 2008) and assembly of NADH dehydrogenase, respectively (Cho, et al. 2012; Gaudet, et al. 2011). The other three genes are more indirectly related to the mitochondria and are largely involved in regulating other pathways related to hypoxia and stress (DeYoung, et al. 2008; Jha, et al. 2016; Mossman, et al. 2017; Papaconstantinou, et al. 2005; Woodford, et al. 2017). However, for these sets of genes, we did not employ the same set of stringency metrics as we did for *HIF α* , *EGLN* or *VHL* (See “Abundance of hypoxia, mitochondrion and stress-related gene groups” in Methods). In addition, myxozoans are known to have rapidly evolving genes which may hamper our ability to properly identify their presence or absence (Chang, et al. 2015; Evans, et al. 2008; Lavrov and Pett 2016; Takeuchi, et al. 2015). Thus, additional analyses may need to be performed to reveal the extent to which specific “hypoxia”, “mitochondrion”, or “stress” genes aren’t present.

Because there is significant variation in mitochondrial genome phenotype within Cnidaria, we initially hypothesized that such variation, especially reduction or loss of mtDNA, could be useful to predict differential loss or alteration of *HIF α* and other HIF pathway-associated genes. However, there was no reduction or change in their HIF pathway composition across Cnidaria regardless of mitochondrial phenotype, with the exception of Myxozoa. Thus, with putative alteration of HIF pathway functionality restricted to one class, it is clear that such losses do not occur in parallel to changes in mitochondrial chromosome morphology. Our results also revealed wholesale losses of other major transcription factors and enzyme groups in Myxozoa, including most members of the bHLH-PAS domain family and P4HC domain containing proteins. The potential losses of these other genes (*CLOCK*, *Met*, *NCOA*, *NPAS*, *AhR*, *SIM*) are not the focus of this paper, but potentially warrant additional study. Overall, these losses are likely associated with a reduction in genome size, potentially connected to typical genome contractions when evolving a parasitic lifestyle (Jackson 2015; Lu, et al. 2019; Slyusarev, et al. 2020). However, the retention of *HIF α*

and *HIF β /ARNT*, despite the loss of all other bHLH-PAS-containing protein family members, might represent evidence of their importance.

In addition, the putative myxozoan *HIF α* sequences are not monophyletic, while the other cnidarian groups consistently are (fig. 3, supplementary figs. S1–S4, Supplementary Material online). This may be due to lineage specific sequence divergence of *HIF α* , which is largely consistent with known high rates of substitution (Chang, et al. 2015; Evans, et al. 2008; Lavrov and Pett 2016; Takeuchi, et al. 2015). Nevertheless, the loss of HIF-specific repression machinery suggests two scenarios: (1) that both *HIF α* and *HIF β /ARNT* are nonfunctional and were retained by chance, or (2) that *HIF α* and *HIF β /ARNT* are functional, can dimerize and affect transcription unabated by normal feedback from repression machinery that would occur during normoxia. Regardless of scenario, without *EGLN*, *HIF α* is not hydroxylated and thus not degraded through usual mechanisms. The activity of the HIF dimer could easily be ascertained by assessing transcription factor binding activity through CHIP-seq (Schödel et al. 2011). Our findings are potentially consistent with the general hypothesis that constant hypoxia or anoxia stress in their environment (i.e., within host tissues) may have contributed to dramatic changes in mitochondrial and nuclear genome composition, although it is important to point out that not all myxozoans reside solely in hypoxic tissues (Hartigan, et al. 2020).

Ultimately, our results strengthen the notion that loss (or alteration) of a once-thought crucial pathway is not unusual, and likely has occurred many times within invertebrate lineages. Such assessments can only be performed with an expansion of sequence data across a wider spectrum of organisms. Although we were unable to reconcile consistent changes to the HIF pathway correlated with mitochondrial phenotypes because there was little variation in the former, there are numerous questions which our results raise. Specifically, given how we understand the role of the HIF pathway and mitochondrial metabolism, it is unclear if myxozoans are using the HIF pathway in the expected manner, or whether their hypoxia-response mechanism is more similar to that of lineages that lost HIF completely (tardigrades, barnacles, and some copepods (Graham and Barreto 2020). The discovery of additional taxonomic groups lacking key metazoan genes will also permit inquiries regarding ecological or life-history correlates that may have facilitated such major evolutionary events.

Materials and Methods

Sequence Resources and Transcriptome Assemblies

Of the 35 genome assemblies publicly available for Cnidaria, 19 had annotated protein sequence files that were used in this study. These included classes Anthozoa

($n=9$), Cubozoa (1), Hydrozoa (2), Myxozoa (4), Scyphozoa (2), and Staurozoa (1) and were downloaded directly from their respective repositories (supplementary Table S1, Supplementary Material online). To increase taxon sampling, transcriptomic data (Illumina paired-end mRNA reads) from an additional 27 species were downloaded from the NCBI Sequence Read Archive (SRA) and assembled de novo. Among these were species in the classes Cubozoa (4), Hydrozoa (9), Polypodiozoa (1), Myxozoa (7), and Scyphozoa (4) and Staurozoa (2) (supplementary Table S2, Supplementary Material online). Thus, the final “groups” included in this study were Anthozoa, Cubozoa, Hydrozoa, Scyphozoa, Staurozoa, Myxozoa, and Polypodiozoa. Transcriptome data available in the SRA often varied in tissue source and developmental stage. Whenever possible, we combined reads from a variety of these sources in order to improve the breadth of transcripts assembled for each taxon. Full list of species and their classification is found on Table 1.

For de novo transcriptome assemblies, downloaded reads were trimmed for base quality ($>Q30$) and adapters using BBDuk (Bushnell, et al. 2017), and transcripts assembled with Trinity v2.9.1 (Haas, et al. 2013) using default parameters. The assembled transcripts were then translated to protein sequences using Transdecoder v5.5.0, which included steps for predicting all open reading frames (ORFs), BLASTing of ORFs against the Uniprot/Swissprot database (release date March 2021), and retaining the best protein model for each transcript. For transcriptome assemblies from the endoparasitic myxozoan species and *Polypodium*, contamination from host tissue (fish, annelid or bryozoan) is likely to be present in the pool of reads, as observed for mRNA-based assemblies from parasitic copepods (Graham and Barreto 2020). We hence filtered the myxozoan and *Polypodium* assemblies by identifying proteins assembled from potential host tissue. We prepared a custom BLAST database containing protein sequences from three fish, one bryozoan and one annelid obtained from Ensembl. These were *Oncorhynchus mykiss* (assembly Omyk_1.0, accession GCA_002163495.1), *Cyprinus carpio* (common_carp_genome, GCA_000951615.2), *Sparus aurata* (SpaAur1.1, GCA_900880675.1), *Capitella teleta* (Capitella telata v1.0, GCA_000328365.1) and *Bugula neritina* (ASM1079987v2, GCA_010799875.2). We then performed BLASTP (blast + v2.10) searches of our assembled myxozoan sequences against this database. Hits that had a match within these potential host sequences, with $e\text{-value} \leq 1e\text{-}75$ and percent identity ≥ 85 , were considered contaminant and removed from the assemblies, as done in Yahalomi, et al. (2020).

To examine overall quality of de novo assembled transcriptomes, we used two complementary approaches. To assess completeness, we used a benchmarking universal single-copy orthologs (BUSCO) analysis using BUSCO v3

(Simão et al. 2015; Waterhouse, et al. 2017). The Metazoan database provided by the software website (<http://busco.ezlab.org>; metazoa_odb9) was used for the assessment with default settings ($e\text{-value}$ of $1e\text{-}3$). The BUSCO analysis was performed on all 46 protein sequence data sets. Finally, we quantified read representation and pairing by mapping the read pairs to their respective assemblies using Bowtie2 (Langmead and Salzberg 2012) and counting the proportion of reads that map in proper pairs and those that are “broken” or orphan during alignment. Assemblies with high proportion of broken read pairs tend to be highly fragmented. The read mapping analysis was performed for the 27 transcriptomes we assembled de novo from SRA data.

Identification of HIF Complex Members

The sequence data for the 46 species used in our analyses were generated by multiple research groups under variable conditions, and hence likely vary in the completeness of their gene content. Our approach to minimize false negative findings (i.e., false gene losses) was to examine multiple species within each group when possible, and to conservatively claim loss of a HIF gene only when it was not detected in all species examined within the respective group (Graham and Barreto 2020). In all groups, the presence of three key elements of the HIF pathway were assessed—the transcription factor subunit, HIF α , and its repression machinery (the proteins EGLN and VHL)—as a way to evaluate the extent to which the canonical oxygen-sensing and -responding HIF pathway was retained or lost.

We used the hmmssearch command from the HMMER 3.0 program (Eddy 2011, 1998) to identify proteins that contained [1] the PAS domain (PF00989.24) to identify HIF α , [2] P4HC domain (PF13640.5), to identify EGLN, [3] and VHL domain (PF01847) to identify VHL. Proteins that had a positive match in HMMER motif searches were then interrogated via BLASTP (blast + v2.10) to a custom database of previously identified invertebrate gene members (Graham and Barreto, 2020), and via InterProScan annotation v81 (Blum, et al. 2021; Jones, et al. 2014); this method has been used in prior studies of this pathway (Graham and Barreto 2020; Graham and Barreto 2019; Graham and Presnell 2017). For InterProScan, multiple levels of annotation needed to be present for HIF α verification, including “Hypoxia-Inducible Factor 1 α ” and “Similar (sima)-like”, as well as appropriate combinations of canonical domains (bHLH, PAS, NTAD, CTAD, ODDD, depending on protein “completeness”). For the myxozoans, due to the potential from host contamination even after the filtering steps employed in preprocessing, any putative HIF α , EGLN or VHL proteins passing filters above were then subjected to a BLASTP (blast + v2.10) comparison to the full GenBank non-redundant “nr” database and identified top 10 candidates

—if a candidate had all top hits matching a known host (i.e., fish, annelid), then they were not considered a myxozoan protein.

Ultimately, our identification of a HIF α was contingent upon the protein sequence having a (I) significant BLAST hit, plus InterProScan identification as HIF α and/or (II) grouping with HIF α “anchors” in a phylogenetic analysis. At both steps, InterProScan was used to annotate protein domains for detection of HIF α by having consistent annotations to “Hypoxia-Inducible Factor 1 α ” and “Similar (sima)-like”, as well as appropriate combinations of canonical domains (bHLH, PAS, NTAD, CTAD, ODDD, depending on protein completeness). Any sequence with InterProScan annotation to “Hypoxia-Inducible Factor 1 α ”, in combination with equally supported annotation to other members of the bHLH-PAS gene families (e.g., single-minded/SIM) were ultimately not classified as a HIF α . We classified a sequence as a HIF α member if it met either criteria I or II above.

For P4HC domain containing proteins, our identification of an EGLN was contingent upon the protein sequence having a (I) significant BLAST hit, plus InterProScan identification as EGLN and/or (II) grouping with EGLN “anchors” in a phylogenetic analysis. The specific annotation classification necessary to be considered an EGLN is “EGL Nine Homolog” and/or “HIF proyl hydroxylase” from InterProScan, but not the general annotation of “Proyl hydroxylase related”. For VHL sequences, our identification was based on a sequence having (I) a significant VHL domain via HMMR and (II) a significant BLAST hit, plus InterProScan identification as VHL. The specific accepted annotation classifications to be classified as a VHL were “VHL”, “pVHL”, and/or “Von Hippel-Lindau Protein”.

As mentioned above, variation across genome or transcriptome sequencing efforts likely affected downstream assembly and annotation. Thus, we took a conservative approach and called presence/absence of genes of interest based on collective results across species within a particular group. Ultimately, we classified presence or absence of the HIF pathway largely based on HIF α ; thus, if HIF α was absent, we would classify the pathway as having been lost, and conversely if HIF α was present (regardless of repressor machinery presence or absence) we would classify the pathway as having been retained. Nevertheless, results indicating a loss of both repression machinery genes consistently within a group, when HIF α is present, would suggest that the HIF pathway, while present by our criteria above, may not function in the canonical manner.

Phylogenetic Analysis of Cnidarian HIF α and EGLN

After screening based on functional motifs and BLAST, we performed phylogenetic analyses with the identified protein sequences containing bHLH-PAS and P4HC domains, with a separate analysis for each domain. These analyses

provide an additional level of assessment, permitting possibly incomplete or fragmented sequences to still be detected as HIF α (in the bHLH-PAS tree) or EGLN (in the P4HC tree) even if there were missed in our first set of searches above.

Part of our gene-identification pipeline for the HIF pathway uses the concept of phylogeny-based orthology (Altenhoff and Dessimoz 2009; Fang, et al. 2010). In order to facilitate further identification, putative sequences are aligned with already known priors (i.e., anchors) for each gene family, which in our case includes several insects (Api = *Acyrtosiphon pisum*, Dme = *Drosophila melanogaster*, Ame = *Apis mellifera*, Nvi = *Nasonia vitripennis*, Tca = *Tribolium castaneum*, Dpon = *Dendroctonus ponderosae*), the cnidarian *Nematostella vectensis* (Nve), and human (Hsa). These have been included as unambiguous indicators of gene identify during phylogenetic analyses of this gene family in previous work (e.g., Graham and Presnell, 2017; Graham and Barreto, 2019, 2020).

Of those sequences with significant HMMR hits, redundant sequences were purged using CD-HIT v4.8.1 (Fu, et al. 2012; Huang, et al. 2010). Due to the large number of sequences, a separate phylogenetic analysis was performed for smaller sets of cnidarian taxa, with one tree each for anthozoans (9 species), hydrozoans (11 species), cubozoans and staurozoans (8 species combined), scyphozoans (6 species) and myxozoans + *Polypodium* (12 species). For each domain set in each cnidarian group, a multiple sequence alignment was built with MUSCLE v3.8.425 using the default parameters (Edgar 2004), and no additional edits performed by hand. Maximum-likelihood analyses were performed using RAXML-VI-HPC (version 2.2.3) with bootstraps (1,000 replicates), using the model VT + G for both bHLH-PAS and P4HC trees, after this models was determined to be best fit for the alignment via Modeltest in IQtree v1.6.12 (Nguyen, et al. 2014; Stamatakis 2014). Final phylogenetic trees were visualized using iTOL v6 (Letunic and Bork 2006; Letunic and Bork 2016) through their online portal (<https://itol.embl.de/>). No phylogenetic trees were created for VHL sequences because the VHL domain is present only in VHL sequences and is not present in any other gene family.

Abundance of Hypoxia, Mitochondrion and Stress-related Gene Groups

With the potential loss of the HIF regulatory machinery, we also assessed the retention of genes involved in response to stress (including hypoxia) as well as those targeted to the mitochondrion. While the HIF pathway is controlled by the few genes targeted above, response to hypoxia stress includes many genes that are downstream targets of HIF pathway activation.

From Uniprot, we downloaded all genes from *Nematostella vectensis* that are associated with hypoxia, including the gene ontology term “mitochondrion” (Cellular Component, GO:5739, $n = 692$), and “response to stress” (Biological Process, GO:6950, $n = 576$). There are currently few annotated genes associated with any hypoxia GO term category— thus, we next used the *D. melanogaster* sequences to help create a cnidarian-specific database using proteins from the model cnidarian *Nematostella vectensis* —“response to hypoxia” (Biological Process, GO:1666, $n = 96$ genes). For this, we performed a batch search of *D. melanogaster* sequences against a custom database of all predicted protein sequences from *N. vectensis*—genes that received a significant BLASTP (blast + v2.10, e-value > 1e-10) hit were retained. This resulted in 50 “response to hypoxia” proteins which had counterparts in *N. vectensis* (supplementary Tables S3–S5, Supplementary Material online).

These *N. vectensis* proteins were compiled into BLAST-able databases, and all cnidarian protein sequences from our target taxa were then compared to the *N. vectensis* database lists using BLAST for group-level identification.

Acknowledgements

We thank the core facilities at the Center for Genome Research and Biocomputing at Oregon State University, where computing was performed. This work was supported by a National Science Foundation Postdoctoral Research Fellowship in Biology to A.M.G. (award no. 1812103), and an NSF-IOS grant to F.S.B. (award no. 2037574).

Data Availability

Genome-based sequences and raw RNA-seq Illumina reads are already available through the NCBI SRA or other repositories, and their accession numbers can be found on supplementary Tables S1 and S2, Supplementary Material online. The transcriptome assemblies generated during this study are deposited on Dryad (<https://doi.org/10.5061/dryad.sbcc2fr72>), and scripts/commands used are available as a github repository (https://github.com/amgraham07/myxozoa_hypoxia).

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online (<http://www.gbe.oxfordjournals.org/>).

Literature Cited

Altenhoff AM, Dessimoz C. 2009. Phylogenetic and functional assessment of orthologs inference projects and methods. *PLoS Comput Biol* 5:e1000262.

- Bao X, et al. 2021. The crosstalk between HIFs and mitochondrial dysfunctions in cancer development. *Cell Death Dis* 12:1–13.
- Bell EL, Emerling BM, Chandel NS. 2005. Mitochondrial regulation of oxygen sensing. *Mitochondrion* 5:322–332.
- Blum M, et al. 2021. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* 49:D344–D354.
- Bridge D, Cunningham CW, Schierwater B, Desalle R, Buss LW. 1992. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. *Proc Natl Acad Sci USA* 89: 8750–8753.
- Bushnell B, Rood J, Singer E. 2017. BBMerge—accurate paired shotgun read merging via overlap. *PLoS One* 12:e0185056.
- Chacinska A, Pfanner N, Meisinger C. 2002. How mitochondria import hydrophilic and hydrophobic proteins. *Trends Cell Biol* 12: 299–303.
- Chandel NS, Schumacker PT. 2000. Cellular oxygen sensing by mitochondria: old questions, new insight. *J Appl Physiol* 88: 1880–1889.
- Chang ES, et al. 2015. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. *Proc Natl Acad Sci USA* 112: 14912–14917.
- Cho J, Hur JH, Graniel J, Benzer S, Walker DW. 2012. Expression of yeast NDI1 rescues a *Drosophila* complex I assembly defect. *PLoS One* 7:e50644.
- Chua YL, et al. 2010. Stabilization of hypoxia-inducible factor-1 α protein in hypoxia occurs independently of mitochondrial reactive oxygen species production. *J Biol Chem* 285:31277–31284.
- DeYoung MP, Horak P, Sofer A, Sgroi D, Ellisen LW. 2008. Hypoxia regulates TSC1/2–mTOR signaling and tumor suppression through REDD1-mediated 14–3–3 shuttling. *Genes Dev* 22: 239–251.
- Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* 14: 755–763.
- Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* 7:e1002195.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797.
- Evans NM, Lindner A, Raikova EV, Collins AG, Cartwright P. 2008. Phylogenetic placement of the enigmatic parasite, *Polypodium hydriforme*, within the Phylum Cnidaria. *BMC Evol Biol* 8: 1–12.
- Famme P, Knudsen J. 1985. Anoxic survival, growth and reproduction by the freshwater annelid, *Tubifex* sp., demonstrated using a new simple anoxic chemostat. *Comp Biochem Physiol A Mol Integr Physiol* 81:251–253.
- Fang G, Bhardwaj N, Robilotto R, Gerstein MB. 2010. Getting started in gene orthology and functional analysis. *PLoS Comput Biol* 6: e1000703.
- Fish FF. 1939. Observations on *Henneguya salminicola* Ward, a myxosporidian parasitic in Pacific salmon. *J Parasitol* 25:169–172.
- Freeman H, Shimomura K, Cox R, Ashcroft F. 2006. Nicotinamide nucleotide transhydrogenase: a link between insulin secretion, glucose metabolism and oxidative stress. *Biochem Soc Trans* 34: 806–810.
- Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28: 3150–3152.
- Gaudet P, Livstone MS, Lewis SE, Thomas PD. 2011. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform* 12:449–462.
- Graham AM, Barreto FS. 2019. Loss of the HIF pathway in a widely distributed intertidal crustacean, the copepod *Tigriopus californicus*. *Proc Natl Acad Sci USA* 116:12913–12918.

- Graham AM, Barreto FS. 2020. Independent losses of the Hypoxia-Inducible Factor (HIF) Pathway within Crustacea. *Mol Biol Evol* 37:1342–1349.
- Graham AM, Presnell JS. 2017. Hypoxia Inducible Factor (HIF) transcription factor family expansion, diversification, divergence and selection in eukaryotes. *PLoS One* 12:e0179545.
- Haas BJ, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8:1494.
- Hartigan A, Kosakyan A, Pecková H, Eszterbauer E, Holzer AS. 2020. Transcriptome of *Sphaerospora molnari* (Cnidaria, Myxosporaea) blood stages provides proteolytic arsenal as potential therapeutic targets against sphaerosporosis in common carp. *BMC Genomics* 21:1–17.
- Hashimoto T, et al. 2016. Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade-unique protein. *Nat Comm* 7:12808.
- Huang Y, Niu B, Gao Y, Fu L, Li W. 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26: 680–682.
- Jackson AP. 2015. The evolution of parasite genomes and the origins of parasitism. *Parasitology* 142:S1–S5.
- Jha AR, et al. 2016. Shared genetic signals of hypoxia adaptation in *Drosophila* and in high-altitude human populations. *Mol Biol Evol* 33:501–517.
- John U, et al. 2019. An aerobic eukaryotic parasite with functional mitochondria that likely lacks a mitochondrial genome. *Sci Adv* 5:eaav1110.
- Johnston I, Ward P, Goldspink G. 1975. Studies on the swimming musculature of the rainbow trout I. Fibre types. *J Fish Biol* 7: 451–458.
- Jones P, et al. 2014. Interproscan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240.
- Karnkowska A, et al. 2016. A eukaryote without a mitochondrial organelle. *Curr Biol* 26:1274–1284.
- Kayal E, et al. 2012. Evolution of linear mitochondrial genomes in mesozoan cnidarians. *Genome Biol Evol* 4:1–12.
- Kayal E, et al. 2018. Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol Biol* 18:1–18.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
- Lavrov DV, Pett W. 2016. Animal mitochondrial DNA as we do not know it: mt-genome organization and evolution in nonbilaterian lineages. *Genome Biol Evol* 8:2896–2913.
- Letunic I, Bork P. 2006. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23: 127–128.
- Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245.
- Liao C, Zhang Q. 2020. Understanding the oxygen-sensing pathway and its therapeutic implications in diseases. *Am J Pathol* 190:1584–1595.
- Lin X, et al. 2008. A chemical genomics screen highlights the essential role of mitochondria in HIF-1 regulation. *Proc Natl Acad Sci USA* 105:174–179.
- Lu T-M, Kanda M, Furuya H, Satoh N. 2019. Dicyemid Mesozoans: a unique parasitic lifestyle and a reduced genome. *Genome Biol Evol* 11:2232–2243.
- Maguire F, Richards TA. 2014. Organelle evolution: a mosaic of ‘mitochondrial’ functions. *Curr Biol* 24:R518–R520.
- Mills DB, et al. 2018. The last common ancestor of animals lacked the HIF pathway and respired in low-oxygen environments. *Elife* 7: e31176.
- Mossman JA, et al. 2017. Mitonuclear interactions mediate transcriptional responses to hypoxia in *Drosophila*. *Mol Biol Evol* 34: 447–466.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274.
- Novosolov M, et al. 2022. The Phylogenetic position of the enigmatic, *Polypodium hydriforme* (Cnidaria, Polypodiozoa): insights from mitochondrial genomes. *Genome Biol Evol* 14:evac112.
- Okamura B, Gruhl A. 2016. Myxozoa + Polypodium: a common route to endoparasitism. *Trends Parasitol* 32:268–271.
- Osigus H-J, Eitel M, Bernt M, Donath A, Schierwater B. 2013. Mitogenomics at the base of Metazoa. *Mol Phylogenet Evol* 69: 339–351.
- Papaconstantinou M, et al. 2005. Menin is a regulator of the stress response in *Drosophila melanogaster*. *Mol Cell Biol* 25:9960–9972.
- Poyton RO, McEwen JE. 1996. Crosstalk between nuclear and mitochondrial genomes. *Annu Rev Biochem* 65:563–607.
- Raikova EV. 1994. Life cycle, cytology, and morphology of *Polypodium hydriforme*, a coelenterate parasite of the eggs of acipenseriform fishes. *J Parasit* 80:1–22.
- Raikova E, Suppes VC, Hoffman G. 1979. The parasitic coelenterate, *Polypodium hydriforme* Ussov, from the eggs of the American acipenseriform *Polyodon spathula*. *J Parasit* 65:804–810.
- Rytönen KT, Williams TA, Renshaw GM, Primmer CR, Nikinmaa M. 2011. Molecular evolution of the metazoan PHD–HIF oxygen-sensing system. *Mol Biol Evol* 28(6):1913–1926.
- Santos HJ, Makiuchi T, Nozaki T. 2018. Reinventing an organelle: the reduced mitochondrion in parasitic protists. *Trends Parasitol* 34: 1038–1055.
- Schödel J, et al. 2011. High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* 117:e207–e217.
- Schönenberger MJ, Kovacs WJ. 2015. Hypoxia signaling pathways: modulators of oxygen-related organelles. *Front Cell Dev Biol* 3: 42.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210–3212.
- Slyusarev GS, Starunov VV, Bondarenko AS, Zorina NA, Bondarenko NI. 2020. Extreme genome and nervous system streamlining in the invertebrate parasite *Intoshia variabilis*. *Curr Biol* 30: 1292–1298.
- Smith DR, et al. 2012. First complete mitochondrial genome sequence from a box jellyfish reveals a highly fragmented linear architecture and insights into telomere evolution. *Genome Biol Evol* 4:52–58.
- Stairs CW, Leger MM, Roger AJ. 2015. Diversity and origins of anaerobic metabolism in mitochondria and related organelles. *Philos Trans R Soc Lond B Biol Sci* 370:20140326.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Takeuchi F, et al. 2015. The mitochondrial genomes of a myxozoan genus *Kudoa* are extremely divergent in Metazoa. *PLoS One* 10: e0132030.
- Thomas LW, Ashcroft M. 2019. Exploring the molecular interface between hypoxia-inducible factor signalling and mitochondria. *Cell Mol Life Sci* 76:1759–1777.
- Waterhouse RM, et al. 2017. BUSCO Applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol* 35: 543–548.
- Williamson CL, Dabkowski ER, Dillmann WH, Hollander JM. 2008. Mitochondria protection from hypoxia/reoxygenation injury with

- mitochondria heat shock protein 70 overexpression. *Am J Physiol Heart Circ Physiol* 294:H249–H256.
- Woodford MR, et al. 2017. Tumor suppressor Tsc1 is a new Hsp90 co-chaperone that facilitates folding of kinase and non-kinase clients. *EMBO J* 36:3650–3665.
- Yahalomi D, et al. 2017. The multipartite mitochondrial genome of *Enteromyxum leei* (Myxozoa): eight fast-evolving megacircles. *Mol Biol Evol* 34:1551–1556.
- Yahalomi D, et al. 2020. A cnidarian parasite of salmon (Myxozoa: *Henneguya*) lacks a mitochondrial genome. *Proc Natl Acad Sci USA* 117:5358–5363.
- Yoshida Y, et al. 2017. Comparative genomics of the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus*. *PLoS Biol.* 15: e2002266.

Associate editor: Dr. Ellen Pritham