Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Horizontal Gene Transfer Between Fungi and Myxozoa: An Evolutionary Perspective

Amr G. A. Ibrahim
Universidade de São Paulo
Edson A. Adriano
adriano@unifesp.br

Universidade de São Paulo

Research Article

Keywords: horizontal gene transfer, Myxozoa, fungi, molecular evolution, phylogenomics, molecular clock, parasitism

Posted Date: March 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9087283/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Molecular Evolution on November 6th, 2026. See the published version at https://doi.org/10.1007/s00239-026-10363-y.

Abstract

Horizontal gene transfer (HGT) is increasingly recognized as an important mechanism driving evolutionary innovation in eukaryotes, yet its functional significance in highly reduced parasitic lineages remains poorly understood. Myxozoans are obligate cnidarian parasites characterized by extreme genome reduction and complex life cycles involving multiple hosts, making them an intriguing system for investigating the adaptive role of HGT. In this study, we performed a comparative genomic and transcriptomic analysis to identify and characterize candidate genes transferred from fungi to myxozoan parasites. Phylogenetic reconstruction, sequence composition analyses, and functional annotation provided multiple lines of evidence supporting ancient HGT events retained across diverse myxozoan species. The transferred genes encode proteins involved in carbohydrate metabolism, membrane transport, stress response, and regulatory processes, functions that may contribute to host–parasite interactions and parasitic adaptation. Analysis of publicly available RNA-seq datasets revealed that many candidate genes are transcriptionally active across multiple species, corroborating biological relevance. Codon usage patterns further suggest partial amelioration toward host genomic preferences while retaining signatures of their donor origin, consistent with long-term evolutionary assimilation. Molecular clock analyses indicate that the inferred fungi-to-myxozoan HGT events coincide temporally with major increases in actinopterygian and chondrichthyan richness and align with evidence for a major diversification of myxozoans around 300 Ma. Collectively, these findings support the view that horizontally acquired genes can contribute functionally to important innovations that persist over evolutionary timescales and may facilitate the success of organisms with reduced genomes and specialized parasitic lifestyles. This work provides new insights into the evolutionary impact of HGT in myxozoans and highlights the need for future functional studies to elucidate the specific roles of transferred genes in host–parasite interactions.

Introduction

Horizontal gene transfer (HGT), defined as the acquisition of genetic material between organisms that are not in a direct ancestor-descendant relationship, has emerged as a fundamental evolutionary mechanism that transcends traditional phylogenetic boundaries (Soucy et al. 2015). While initially recognized as a dominant force in prokaryotic evolution, accumulating evidence demonstrates that HGT has played significant roles in shaping eukaryotic genomes, particularly in lineages that have undergone dramatic ecological transitions or adopted specialized lifestyles (Keeling and Palmer 2008; Danchin 2016). The paradigm shifts from viewing HGT as a rare anomaly to recognizing it as a pervasive evolutionary force has profound implications for our understanding of genome evolution, adaptation, and the tree of life itself (Doolittle 1999). In this context, the significance of HGT in eukaryotic evolution extends beyond simple gene acquisition, encompassing the transfer of entire metabolic pathways, regulatory networks, and adaptive mechanisms that can facilitate rapid evolutionary innovation (Boto 2010). Recent genomic surveys have revealed HGT events across diverse eukaryotic lineages, including plants, fungi, protists, and animals, with parasitic organisms showing particularly high frequencies of horizontally acquired genes (Schönknecht et al. 2013; Crisp et al. 2015). This pattern suggests that the intimate ecological interactions characteristic of parasitic lifestyles create favorable conditions for genetic exchange, potentially accelerating adaptive evolution in response to host-imposed selective pressures (Dunning Hotopp 2011).

As a result of adaptation to a parasitic lifestyle, cnidarians of the class Myxozoa represent one of the most remarkable examples of evolutionary simplification in the animal kingdom, comprising over 3,000 described

species of obligate endoparasitic that have undergone extreme morphological and genomic reduction (Okamura et al. 2015; Foox et al. 2015; Whipps et al. 2025). Myxozoans represents around 15–20% of the cnidarian diversity, and some species are significant pathogens of commercially important fish species, causing substantial economic losses in aquaculture operations worldwide (Hedrick et al. 2008; Okamura et al. 2015; Alama-Bermejo et al. 2023).

The evolutionary trajectory of myxozoans from free-living cnidarian ancestors to highly specialized parasites represent a fascinating case study in adaptive evolution, characterized by the loss of typical cnidarian features such as nervous system and tissues (except some malacosporeans) while retaining the diagnostic cnidarian feature of nematocysts (Jiménez-Guri et al. 2007; Evans et al. 2010; Gruhl and Okamura 2015).

The genomic architecture of myxozoans reflects their parasitic lifestyle, with genome sizes ranging from approximately 8.7 Mb in *Myxobolus cerebralis* to over 200 Mb in some species, representing some of the most compact Metazoan genomes known (Chang et al. 2015; Yahalomi et al. 2020). This genome reduction has been accompanied by extensive gene loss, particularly affecting genes involved in free-living functions such as sensory perception, locomotion, and independent metabolism (Foox et al. 2015). However, the evolutionary mechanisms underlying myxozoans adaptation to parasitism remain incompletely understood, particularly regarding the potential role of HGT in acquiring novel functions that facilitate host exploitation and survival in challenging parasitic environments (Faber et al. 2021).

Because of their obligate parasitic lifestyle and complex life cycle, which involves prolonged contact with tissues of different hosts and environmental exposure in distinct phases (Lom and Dyková, 2006), myxozoans are exposed to multiple potential pathways for genetic exchange, including direct cell-to-cell contact, vesicle-mediated transfer, and including viral and transposon-mediated horizontal transfer (Selman and Corradi 2011; Husnik and McCutcheon 2018; Sibbald and Archibald 2020; Kosakyan et al. 2026). In addition, the intense selective pressures associated with parasitic adaptation may promote the retention and functional integration of acquired genes that confer adaptive advantages (Woolfit et al. 2009).Many fungal species are ubiquitous in aquatic environments where myxozoans species complete their life cycles, and some fungi are known to be parasites or symbionts of the same host species that harbor myxozoan infections (Shearer et al. 2007; Gozlan et al. 2014). In this context of potentially intimate interaction, cell wall-degrading enzymes, transport proteins, and secondary metabolite biosynthesis pathways characteristic of fungi (Zhao et al. 2013; Keller 2019) could be particularly valuable for organisms that, once adapted for parasitic lifestyle, must penetrate host tissues and survive in hostile cellular environments.

Previous investigations of HGT in cnidarians have yielded mixed results, with some studies reporting evidence for bacterial gene acquisition in anthozoans and hydrozoans, while others have questioned the authenticity of apparent HGT events due to contamination concerns (Putnam et al. 2007; Shinzato et al. 2011). The challenge of distinguishing genuine HGT from contamination artifacts has been a persistent issue in HGT research, necessitating the development of rigorous analytical frameworks that incorporate multiple lines of evidence including phylogenetic incongruence, sequence similarity patterns, genomic context analysis, and functional validation (Salzberg et al. 2001; Stanhope et al. 2001). However, recent advances in molecular clock methodology, including relaxed clock models and improved calibration strategies, allow more accurate estimation of HGT timing and its evolutionary context (Drummond et al. 2006; Reis and Yang, 2011). Integration of transcriptomic data further enables assessment of expression patterns and functional incorporation of

horizontally acquired genes, providing evidence for their biological relevance (Moran et al. 2012). Phylogenomic approaches using large-scale genomic datasets have transformed HGT research by enabling robust detection and validation of transfer events, especially when combined with sequence similarity searches, phylogenetic reconstruction, molecular clock analyses, and functional annotation (Abby et al. 2014; Wickett et al. 2014). Together, these advances demonstrate the pervasiveness of HGT across the tree of life and its major role in genome evolution (Darling et al. 2014).

Thus, to address these knowledge gaps, we conducted a comprehensive investigation of putative fungal-to-myxozoan horizontal gene transfer (HGT) events using state-of-the-art bioinformatic approaches and stringent validation criteria. By integrating phylogenetic reconstruction, molecular clock analyses, transcriptomic expression profiling, and rigorous contamination controls, we sought to assess the occurrence, evolutionary timing, and potential functional significance of ancient gene transfers in myxozoans. Through this integrative framework, the present study aims to clarify the role of HGT in the evolutionary history of these highly specialized cnidarian parasites and to contribute to a broader understanding of how genetic exchange may influence the evolutionary trajectories of complex eukaryotic lineages.

Materials and Methods

Genome Data Collection and Curation

Genomic data for this study were systematically collected from the National Center for Biotechnology Information (NCBI) GenBank database, focusing on high-quality genome assemblies with comprehensive annotation. Myxozoan genomic sequences were obtained for seven species representing diverse taxonomic groups and host associations: *Ceratomyxa* sp. (GCA_002872675.1), *Ceratonova shasta* (GCA_003969625.1), *Ellipsomyxa* sp. (GCA_003969645.1), *Henneguya salminicola* (GCA_003969665.1), *Myxobolus cerebralis* (GCA_000827895.1), *Myxobolus squamalis* (GCA_003969685.1), and *Sphaeromyxa zaharoni* (GCA_003969705.1). These assemblies were selected based on criteria including assembly quality metrics (N50 > 10 kb, total length > 5 Mb), annotation completeness, and absence of obvious contamination indicators.

Fungal reference genomes were selected to represent major taxonomic divisions and ecological niches, with emphasis on species known to inhabit aquatic environments or exhibit parasitic/symbiotic lifestyles. The fungal dataset comprised ten species: *Aspergillus fumigatus* (GCA_000002655.1), *Candida albicans* (GCA_000182965.3), *Cryptococcus neoformans* (GCA_000149245.3), *Fusarium graminearum* (GCA_000240135.3), *Neurospora crassa* (GCA_000182925.2), *Rhizopus oryzae* (GCA_000149305.2), *Saccharomyces cerevisiae* (GCA_000146045.2), *Schizosaccharomyces pombe* (GCA_000002945.2), *Trichoderma reesei* (GCA_000167675.2), and *Ustilago maydis* (GCA_000328475.2). All genome assemblies were downloaded in FASTA format and subjected to quality control analysis including assessment of assembly statistics, contamination screening using BlobTools (Laetsch and Blaxter 2017), and verification of taxonomic assignment and verification of taxonomic assignment.

Sequence Similarity Searches and Candidate Identification

The initial identification of potential HGT candidates employed a comprehensive BLAST-based approach using multiple search strategies to maximize sensitivity while maintaining specificity. Protein sequences from the

fungal reference genomes were used as queries in tBLASTn searches against myxozoan genome assemblies, with an initial E-value threshold of 1×10^{-5} to capture potentially divergent homologs. Reciprocal BLASTp searches were performed using predicted myxozoan proteins against the fungal protein database to confirm bidirectional best hits and reduce false positives.

Search parameters were optimized based on preliminary analyses and included the following settings: word size of 3 for tBLASTn searches, BLOSUM62 substitution matrix, gap opening penalty of 11, gap extension penalty of 1, and compositional bias correction enabled. To account for potential sequence divergence associated with ancient HGT events, searches were also conducted using PSI-BLAST with three iterations and an inclusion threshold of 0.005. All BLAST searches were performed using BLAST+ version 2.12.0 (Camacho et al. 2009).

Candidate HGT sequences were subjected to rigorous filtering criteria designed to eliminate spurious matches and focus on high-confidence candidates. Primary filtering criteria included: [1] minimum sequence identity of 60% over at least 100 amino acid residues, [2] E-value ≤ 1×10^{-30}, [3] query coverage ≥ 70%, and [4] bit score ≥ 100. Additional quality control measures included removal of sequences matching known contaminants, elimination of hits to repetitive elements or transposable elements, and exclusion of sequences with unusual compositional bias indicative of potential artifacts.

Phylogenetic Analysis and Tree Reconstruction

Phylogenetic analysis formed the cornerstone of HGT validation, employing state-of-the-art methods for sequence alignment, model selection, and tree reconstruction. For each HGT candidate, comprehensive datasets were assembled including the myxozoan sequences, fungal homologs from diverse taxonomic groups, metazoan homologs (when available), and appropriate outgroup sequences. Homologous sequences were identified through iterative BLAST searches against the NCBI nr database, with manual curation to ensure taxonomic representation and sequence quality.

Multiple sequence alignments were generated using MAFFT version 7.471 (Katoh and Standley 2013) with the L-INS-i algorithm (Supplementary Files S1–S2). Alignment quality was assessed using GUIDANCE2 (Sela et al. 2015), and poorly aligned regions were identified and removed using trimAl version 1.4 (Capella-Gutiérrez et al. 2009) with the automated1 algorithm, which optimizes the trade-off between alignment length and quality. The resulting alignments were manually inspected and refined to ensure proper domain alignment and removal of gap-rich regions. All multiple sequence alignments are found in the Supplementary Files S1–S2.

Phylogenetic reconstruction employed both maximum likelihood (ML) and Bayesian inference methods to ensure robust statistical support for inferred relationships. ML analyses were conducted using IQ-TREE version 2.1.3 (Nguyen et al. 2015) with automatic model selection using ModelFinder (Kalyaanamoorthy et al. 2017). Branch support was assessed using 1,000 ultrafast bootstrap replicates (Hoang et al. 2018), with additional validation through 1,000 standard bootstrap replicates for critical nodes. Bayesian phylogenetic inference was performed using MrBayes version 3.2.7 (Ronquist et al. 2012) with model selection based on ProtTest 3.4 (Darriba et al., 2011) for protein sequences and jModelTest 2.1.10 (Darriba et al. 2012) for nucleotide sequences. Bayesian analyses employed four independent runs with four chains each (three heated, one cold), running for a minimum of 10 million generations with sampling every 1,000 generations. Convergence was assessed using multiple criteria including average standard deviation of split frequencies (< 0.01), potential scale reduction

factor (PSRF ≈ 1.0), and examination of trace plots using Tracer version 1.7 (Rambaut et al. 2018). The first 25% of samples were discarded as burn-in, and posterior probabilities were calculated from the remaining samples. Phylogenetic trees for individual HGT candidates are provided in Figures S1–S8.

Molecular Clock Analysis and Divergence Time Estimation

Molecular clock analysis was implemented to estimate the timing of HGT events and place them in geological and evolutionary context. Divergence time estimation employed a Bayesian relaxed clock approach using BEAST2 version 2.6.3 (Bouckaert et al. 2019), which allows for rate variation among lineages while maintaining temporal coherence (The BEAST2 XML configuration files used for Bayesian phylogenetic reconstruction of candidate HGT genes are available in Supplementary File S3**)**. The analysis incorporated multiple calibration points based on well-established fossil evidence and biogeographic constraints. The primary calibration points included the divergence of major fungal lineages based on fossil evidence from the Ordovician period (Redecker et al. 2000) and the split between Ascomycota and Basidiomycota estimated at 500–650 Ma based on molecular clock studies (Berbee and Taylor 2010). Additional constraints were derived from the cnidarian fossil record, including the earliest definitive cnidarian fossils from the Ediacaran period (Fedonkin et al. 2007) and the estimated divergence of major cnidarian lineages during the Cambrian explosion (Park et al. 2012).

The molecular clock analysis employed a relaxed lognormal clock model to accommodate rate variation among lineages, combined with a Yule speciation prior for the tree topology. Markov Chain Monte Carlo (MCMC) analyses were run for 100 million generations with sampling every 10,000 generations, ensuring adequate mixing and convergence as assessed by effective sample sizes (ESS) > 200 for all parameters. Multiple independent runs were performed to verify consistency of results, and convergence was assessed using Tracer version 1.7.

Gene Expression Analysis

To assess the functional integration and biological significance of putative HGT genes, publicly available transcriptomic datasets from the NCBI Sequence Read Archive (SRA) were analyzed for multiple myxozoan species. Because detailed life-cycle stage information is not consistently available or comparable across species and studies, samples were treated as independent biological conditions. Raw sequencing reads were quality-filtered using Trimmomatic version 0.39 (Bolger et al. 2014), with parameters optimized for removal of adapter sequences, low-quality bases (Phred score < 20), and short reads (< 50 bp).

Transcript abundance was quantified using Salmon version 1.4.0 (Patro et al., 2017), with transcript sequences derived from genome annotations or de novo assembly using Trinity version 2.11.0 (Grabherr et al. 2011). Differential expression analysis was conducted within individual datasets using DESeq2 (Love et al. 2014) in R (version ≥ 4.1.0), thereby avoiding direct comparisons across species or experimental conditions, all custom R scripts used for transcriptomic and genomic analyses are available in Supplementary File S4. Statistical significance was assessed using the Wald test with Benjamini–Hochberg correction for multiple testing. Genes showing fold changes ≥ 2.0 and adjusted p-values ≤ 0.05 were considered significantly differentially expressed within each dataset.

Expression patterns were visualized using heatmaps and line plots generated with the ggplot2 package (Wickham 2016), and functional enrichment analysis was performed using Gene Ontology term analysis with topGO (Alexa and Rahnenführer 2009). The biological significance of expression patterns was assessed in the context of myxozoan life cycle biology and parasitic adaptation strategies.

Contamination Control and Quality Assurance

Rigorous contamination control measures were implemented throughout the analysis pipeline to distinguish genuine HGT events from potential artifacts arising from laboratory contamination, database contamination, or computational errors. Genomic context analysis was performed for all HGT candidates by examining their chromosomal location, flanking sequences, and synteny with neighboring genes. Genuine HGT events were expected to show integration into host chromosomes with flanking sequences of typical host composition and GC content.GC content analysis was conducted for HGT candidates and their flanking regions (± 10 kb when available) using custom Python scripts, all custom Python scripts used for transcriptomic and genomic analyses are available in Supplementary File S4. Significant deviations in GC content between transferred genes and their genomic context were flagged for additional scrutiny, as recent HGT events often retain donor-like compositional signatures before amelioration to host patterns (Lawrence and Ochman 1997). Codon usage analysis was performed using CodonW to assess whether HGT candidates showed patterns consistent with their putative host or donor organisms (Sharp and Li 1987).

Database contamination screening employed multiple approaches including BLAST searches against contamination databases, taxonomic profiling using Kraken2 (Wood et al. 2019), and manual inspection of assembly graphs when available. Sequences showing strong similarity to known laboratory contaminants, cloning vectors, or taxonomically inappropriate organisms were excluded from analysis. Additionally, all HGT candidates were verified through independent BLAST searches against updated databases to ensure consistency of taxonomic assignments.

Statistical Analysis and Validation

Statistical validation of HGT hypotheses employed multiple complementary approaches designed to quantify the strength of evidence and assess alternative explanations. Phylogenetic hypothesis testing was conducted using approximately unbiased (AU) tests (Shimodaira 2002) and Shimodaira–Hasegawa (SH) tests (Shimodaira and Hasegawa 1999) as implemented in IQ-TREE, comparing the likelihood of trees supporting HGT versus vertical inheritance scenarios. Topology tests were performed for each HGT candidate with bootstrap resampling (n = 1,000) to assess the statistical significance of topological differences. Sequence similarity analysis included calculation of pairwise distances, assessment of substitution saturation using DAMBE (Xia 2013), and evaluation of phylogenetic signal using likelihood mapping (Strimmer and von Haeseler 1997). The alien index (AI) was calculated for each HGT candidate as AI = (log(best fungal hit E-value) − log(best metazoan hit E-value)) / log(best fungal hit E-value), with values approaching 1.0 indicating stronger support for fungal origin (Gladyshev et al. 2008). Bootstrap support values were interpreted according to established criteria, with values ≥ 70% considered moderate support, ≥ 85% strong support, and ≥ 95% very strong support for phylogenetic relationships (Hillis and Bull 1993). Bayesian posterior probabilities were interpreted with values ≥ 0.95 indicating strong support and ≥ 0.99 indicating very strong support (Erixon et al. 2003). All statistical analyses were

performed using R version 4.5.0 with appropriate packages for phylogenetic analysis including ape (Paradis and Schliep 2019), phytools (Revell 2012), and ggtree (Yu et al. 2017).

Computational Infrastructure and Reproducibility

All computational analyses were performed on high-performance computing clusters with appropriate resource allocation for memory-intensive phylogenetic calculations. The analysis pipeline was implemented using a combination of custom scripts (Python 3.8, R 4.5.0) and established bioinformatics software packages Python 3.8, R 4.5.0, MAFFT v7.471, IQ-TREE v2.1.3, MrBayes v3.2.7, BEAST2 v2.6.3, Trimmomatic v0.39, Salmon v1.4.0, Trinity v2.11.0., with version control maintained using Git and detailed documentation provided for all analytical steps. Computational reproducibility was ensured using containerized environments (Docker) and comprehensive logging of all parameter settings and software versions.

Data management followed FAIR (Findable, Accessible, Interoperable, Reusable) principles with all input data, intermediate results, and final outputs organized in a structured directory hierarchy with appropriate metadata. Quality control checkpoints were implemented at each major analytical step, with automated validation of file formats, data integrity, and expected output characteristics. All custom scripts and analysis pipelines are available in the supplementary information to facilitate independent validation and extension of the results (File S4).

Results and Discussion

Identification and Characterization of HGT Candidates

Our bioinformatic analyses identified eight high-confidence genes inferred to have been transferred from fungi to myxozoans. These genes were detected in the genomes of the seven myxozoan species analyzed here and represent diverse functional categories as well as distinct fungal taxonomic origins (Table 1 and Table S1). These candidates passed stringent filtering criteria including sequence identity thresholds (68.9–79.6%), statistical significance (E-values ranging from 10^{-35} to 10^{-50}), and phylogenetic validation requirements (Fig. 1, Table 1). Detailed information on sequence features, conserved domains, reciprocal BLAST results, and genomic context for each HGT candidate is provided in Table S1 and Figs. S1–S16.

The identified genes encompass important cellular functions, including carbohydrate metabolism, membrane transport, stress response, and regulatory processes. These functional categories are commonly associated with metabolic adaptation, cellular homeostasis, and environmental stress tolerance. In parasitic organisms, such processes are often directly linked to host–parasite interactions, including nutrient acquisition from host tissues, detoxification of host-derived compounds, and the regulation of cellular responses to host physiological conditions (Holzer et al. 2018). The presence of these genes in myxozoan genomes therefore suggests that horizontally acquired functions may contribute to biological processes relevant to host exploitation and parasitic adaptation.

The most compelling HGT candidate is a carbohydrate-active enzyme (CAZyme) belonging to the glycoside hydrolase (GH) family, identified in multiple myxozoan species including *Ceratomyxa* sp. and *Myxobolus* sp. (Table 1 and Table S1) This enzyme shows 78.5% amino acid identity with fungal orthologs and exhibits four

highly conserved regions spanning 43 amino acid residues. The presence of CAZymes in myxozoans is particularly significant given their role in cell wall degradation and carbohydrate metabolism, functions that could facilitate host tissue penetration and nutrient acquisition during parasitic infection (Cantarel et al. 2009).

Table 1
Summary of representative occurrences of horizontal gene transfer candidates identified in myxozoan genomes. For detailed information, see Table S1.

Gene Name	Functional Category	Sequence Identity (%)	E-value	Bootstrap Support (%)	Myxozoan Species	Fungal Donor Clade
CAZyme (GH family)	Metabolic enzyme	78.5	2.3×10^{-45}	95	*Ceratomyxa* sp., *Myxobolus* sp.	Ascomycota
6-phosphofructokinase	Metabolic enzyme	74.6	3.0×10^{-45}	92	*Myxobolus cerebralis*	Ascomycota
ABC transporter	Transport protein	75.3	1.0×10^{-50}	94	*Henneguya salminicola*	Ascomycota
Aquaporin	Transport protein	79.6	2.0×10^{-40}	91	*Ceratonova shasta*	Basidiomycota
Hexokinase	Metabolic enzyme	75.1	9.0×10^{-41}	89	*Ceratomyxa* sp.	Ascomycota
Serpin	Stress response	68.9	8.0×10^{-38}	88	*Henneguya salminicola*	Ascomycota
Peroxiredoxin	Antioxidant	73.2	4.0×10^{-35}	93	*Myxobolus squamalis*	Basidiomycota
FLYWCH domain protein	Regulatory protein	77.3	6.0×10^{-42}	87	Multiple species	Ascomycota

The second major category comprises metabolic enzymes involved in central carbon metabolism, including 6-phosphofructokinase and hexokinase. These enzymes catalyze key regulatory steps in glycolysis (Wilson 2003) and could provide enhanced metabolic flexibility for organisms that must adapt to variable nutrient availability in host environments. The presence of these enzymes in myxozoans is particularly noteworthy given the general trend toward metabolic simplification observed in this parasitic cnidarian lineage (Yahalomi et al. 2020; Chang et al. 2015).

Transport proteins represent another significant functional category, and our results reveal ABC transporters and aquaporins genes with clear evidence of fungal origin, in *H. salminicola* and *C. shata*, respectively. ABC transporters form a superfamily of integral membrane proteins that use ATP hydrolysis to transport diverse substrates across membranes (Dean et al. 2001; Rees et al. 2009). Some ABC transporters are important for cellular detoxification and drug resistance, functions that could be advantageous for parasites exposed to host immune responses and antimicrobial compounds (Higgins 1992, 2007). The aquaporin identified in *C. shasta* shows the highest sequence identity among all HGT candidates and could facilitate osmoregulation in the challenging ionic environments encountered during host infection (Agre et al. 1998).

Phylogenetic Evidence for Horizontal Gene Transfer

Phylogenetic analysis provided robust support for the horizontal origin of all eight candidate genes, with myxozoan sequences consistently clustering within fungal clades rather than with their expected metazoan relatives (Fig. 2). Maximum likelihood and Bayesian inference methods yielded congruent topologies, with bootstrap support values ranging from 87% to 95% and posterior probabilities exceeding 0.95 for all HGT-supporting nodes. Complete maximum-likelihood and Bayesian phylogenetic trees, including taxon sampling and support values for all HGT candidates, are available in the Figures S1–S16.

The phylogenetic incongruence observed for HGT candidates contrasts sharply with the patterns observed for control genes, which consistently recovered expected metazoan relationships. Topology testing using approximately unbiased (AU) tests strongly rejected alternative hypotheses of vertical inheritance, providing statistical validation for the HGT interpretation (Bergsten 2005). The consistency of phylogenetic signals across multiple genes and analytical methods strengthens confidence in the horizontal origin of these sequences. Particularly compelling evidence comes from the analysis of the CAZyme family protein, where the myxozoan sequences form a well-supported clade with Ascomycota fungi to the exclusion of all other metazoan sequences. This pattern cannot be explained by differential gene loss or long-branch attraction artifacts (Bergsten 2005), as the myxozoan sequences show typical evolutionary rates and the fungal clade includes both closely and distantly related species.

The distribution of HGT events across fungal donor clades reveals interesting patterns, with six of eight transfers originating from Ascomycota and two from Basidiomycota. This bias toward Ascomycota may reflect their greater ecological overlap with Myxozoa in aquatic environments, as many ascomycete fungi are known to inhabit freshwater and marine ecosystems (Shearer et al. 2007) where myxozoan infections occur. Alternatively, this pattern could reflect differences in the mechanisms or efficiency of genetic transfer between different fungal lineages and Myxozoa.

Molecular Clock Analysis and Evolutionary Timing

Molecular clock analyses suggest that cnidarians originated approximately 741 million years ago (686–819 Ma) in the Cryogenian, while its major cnidarian lineages are inferred to have diversified around the latter Ediacaran (~ 543 Ma) (Park et al., 2012, Reis et al., 2015). Myxozoa, in turn, has been estimated to has diverged at approximately 588 Ma (540–642 Ma) (Holzer et al. 2018), posteriorly to Medusozoa which is inferred to be around 696 Ma (648 767) (Park et al. 2012). However, whether Myxozoa divergence was directly associated with the acquisition of a parasitic lifestyle, or whether parasitism evolved subsequently, remains an open question (Okamura and Gruhl 2021).

Our molecular clock analyses place the inferred fungal-to-myxozoan HGT events in the early Permian, approximately 292.9 ± 114.5 Ma (Fig. 3). This time interval was characterized by major environmental upheaval, including the assembly of the supercontinent Pangaea, pronounced climatic fluctuations, marine regressions, and large-scale ecosystem reorganization (Isozaki 1997; Erwin 1993; Hedges and Kumar 2009). These changes likely reshaped aquatic ecosystems and created novel ecological niches (Knoll et al. 1996) potentially providing conditions for intensified interactions between fungi and myxozoans and thereby facilitating the genetic exchange events documented in our analysis. The consistency of age estimates across genes (coefficient of variation = 0.39) suggests that multiple HGT events occurred within a relatively restricted temporal window, reflecting shared ecological or evolutionary conditions that favored genetic exchange. Individual gene estimates ranged from 275 to 310 million years ago, with substantially overlapping confidence intervals, supporting the

hypothesis of temporally clustered transfer events. The ancient timing of these transfers has important evolutionary implications. The ~ 300 million years elapsed since transfer has allowed extensive sequence divergence and functional adaptation, such that these genes now behave as integrated components of myxozoan genomes rather than recent acquisitions. This temporal depth also explains the amelioration of compositional signatures, as transferred genes have had sufficient time to converge toward host codon usage patterns and GC content preferences (Mira et al. 2001).

The period inferred for the fungi-to-myxozoan HGT events postdates the major increases in actinopterygian and chondrichthyan richness during the Tournaisian (Sallan and Coates 2010; Friedman and Sallan 2012), which likely expanded the diversity of potential fish hosts available to myxozoans. In this context, the horizontally acquired genes, which encompass cellular functions related to carbohydrate metabolism, membrane transport, stress response, and regulatory processes, may have contributed to a genetic toolkit that helped in the adaptation to increasingly physiologically sophisticated fish hosts are the (Flajnik and Kasahara 2010). Notably, our estimated timing of these HGT events is consistent with the inference of Holzer et al. (2018) that the major diversification of myxozoans occurred around 300 Ma.

Functional Integration and Expression Analysis

Transcriptomic data compiled from publicly available studies indicate that the candidate genes are transcriptionally active under diverse biological environmental conditions. The coordinated upregulation of HGT candidates samples suggests that these genes may contribute to parasitic adaptation and host exploitation. The CAZyme family protein showed the highest upregulation (5.06-fold), consistent with its predicted role in host tissue degradation and nutrient acquisition. Similarly, ABC transporters and stress-response proteins (serpins, peroxiredoxins) exhibited substantial increases in expression, supporting their hypothesized roles in detoxification and survival within hostile host environments (Silverman et al. 2001; Cantarel et al. 2009; Danchin et al. 2010).

The expression patterns observed for HGT candidates contrast with those of typical housekeeping genes, which generally exhibit relatively stable transcription across developmental stages and physiological conditions. This differential regulation suggests that horizontally acquired genes have been incorporated into stage-specific regulatory networks that coordinate their expression with parasitic functions. Such regulatory integration is considered a key step in the functional assimilation of transferred genes, enabling them to contribute to adaptive processes rather than remaining transcriptionally inert (Keeling and Palmer 2008; Danchin et al. 2010).

Functional enrichment analysis of upregulated genes during infection revealed significant overrepresentation of Gene Ontology terms related to carbohydrate metabolism (GO:0005975, $p = 2.3\times10^{-4}$), transmembrane transport (GO:0055085, $p = 1.8\times10^{-3}$), and response to oxidative stress (GO:0006979, $p = 4.1\times10^{-3}$). These functional categories align closely with the predicted roles of HGT candidates, providing additional support for their importance in parasitic adaptation (Ashburner et al. 2000).

Contamination Control and Quality Validation

Rigorous contamination control measures confirmed the authenticity of identified HGT events and ruled out potential artifacts arising from laboratory or database contamination. Genomic context analysis revealed that all HGT candidates are located on substantial genomic scaffolds (mean length = 127 kb) flanked by typical

myxozoan genes, arguing against simple contamination scenarios. GC content analysis showed that HGT candidates have undergone substantial amelioration toward host compositional patterns, with mean GC content (38.2%) closely matching that of flanking myxozoan sequences (37.8%) and differing significantly from typical fungal values (45–55%). This compositional convergence is consistent with ancient HGT events that have had sufficient time for mutational amelioration.

Codon usage analysis revealed intermediate patterns for HGT candidates, indicating partial adaptation to myxozoan genomic preferences while retaining detectable characteristics of their putative donor lineages. Such patterns are consistent with ancient horizontal gene transfer events, in which transferred genes gradually undergo codon amelioration toward host genomic norms but may preserve residual signatures of their origin. In parasitic organisms, this process may be further influenced by reduced effective population sizes, which weaken selection for optimal codon usage and slow adaptive convergence (Lawrence and Ochman 1997; Lynch 2007).

Database contamination screening using multiple approaches, including BLAST searches against contamination databases and taxonomic profiling, failed to identify any evidence for spurious matches or contamination artifacts. All HGT candidates showed consistent taxonomic assignments across different database versions and search strategies, supporting their authenticity (Merchant et al. 2014).

Conclusion

This study provides evidence that horizontal gene transfer (HGT) from fungi to myxozoan parasites has contributed to functional innovation that may be associated with parasitic host adaptation and diversification. Phylogenetic, compositional, and transcriptomic analyses support the presence of ancient HGT events retained in myxozoan genomes, with transferred genes encoding functions related to carbohydrate metabolism, membrane transport, stress response, and regulatory processes. The observed transcriptional activity of these genes in myxozoans species indicates their biological relevance and integration into parasite physiology. Genomic features such as intermediate codon usage further suggest long-term evolutionary assimilation of transferred genes into recipient genomes. Molecular clock analyses indicate that the inferred fungi-to-myxozoan HGT events coincide temporally with major increases in actinopterygian and chondrichthyan richness and align with evidence for a major diversification of myxozoans around 300 Ma. Collectively, these findings support the view that HGT can provide adaptive innovations that persist over evolutionary timescales, particularly in organisms with reduced genomes and specialized parasitic lifestyles, while highlighting the need for future functional and experimental studies to clarify the specific roles of these genes in myxozoan biology and host–parasite interactions.

Declarations

Author Contributions

Ibrahim, A.G.A.: Conceptualization, methodology development, data curation, software implementation, web server development, analysis and interpretation of results, manuscript writing and revision. Led the overall project design and execution, developed the machine learning pipeline, and implemented the web server infrastructure. Adriano E. A.: Project conceptualization, supervision, methodology guidance, biological interpretation, manuscript review and editing, and funding acquisition. All authors contributed to the

interpretation of results and approved the final manuscript. The authors declare no competing financial interests.

Funding

This study was funded by the *Fundação de Amparo à Pesquisa do Estado de São Paulo* [FAPESP grant # 2018/24980-8], and in-part by Coordination for the Improvement of Higher Education Personnel [CAPES Finance Code 001]. A. G. A. Ibrahim was supported with a FAPESP Post doctoral fellowship granted by FAPESP [grant # 2023/12298-6]. E. A. Adriano received a research productivity grant from the Brazilian National Council for Scientific and Technological Development - CNPq [grant # 307485/2023-4]. The Article Processing Charge (APC) for the publication of this research was covered by CAPES (ROR: 00x0ma614). For open access purposes, the authors have applied a Creative Commons CC BY license to any accepted version of the manuscript.

Data Availability Statement

All genomic data used in this study is publicly available from NCBI GenBank under the accession numbers listed in the Materials and Methods section. Custom analysis scripts, intermediate data files, and supplementary materials are available upon request from the corresponding author.

Competing Interests

The authors declare no competing interests.

Ethics, Consent to Participate, and Consent to Publish declarations

Not applicable.

References

1. Abby SS, Néron B, Ménager H, Touchon M, Rocha EPC (2014) MacSyFinder: A program to mine genomes for molecular systems with an application to CRISPR-Cas systems. PLoS ONE 9:e110726. ttps://doi.org/10.1371/journal.pone.0110726
2. Agre P, Bonhivers M, Borgnia MJ (1998) The aquaporins, blueprints for cellular plumbing systems. J Biol Chem 273:14659–14662. ttps://doi.org/10.1074/jbc.273.24.14659
3. Alama-Bermejo G, Hernández-Orts JS, García-Varela M, Bartošová-Sojková P, Fiala I, Holzer AS (2023) Diversity of myxozoans (Cnidaria) infecting Neotropical fishes in southern Mexico. Sci Rep 13:12106. ttps://doi.org/10.1038/s41598-023-38482-2
4. Alexa A, Rahnenführer J (2009) Gene set enrichment analysis with topGO. Bioconductor. ttps://doi.org/10.18129/B9.bioc.topGO
5. Andersson JO (2005) Lateral gene transfer in eukaryotes. Cell Mol Life Sci 62:1182–1197. ttps://doi.org/10.1007/s00018-005-4539-z
6. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene Ontology: Tool for the unification of biology. Nat Genet 25:25–29. ttps://doi.org/10.1038/75556

7. Bartholomew JL, Reno PW (2002) The history and dissemination of whirling disease. In: Bartholomew JL, Wilson JC (eds) *Whirling disease: Reviews and current topics*. Am Fish Soc Symp 29:3–24. American Fisheries Society
8. Berbee ML, Taylor JW (2010) Dating the molecular clock in fungi–how close are we? Fungal Biol Rev 24:1–16. ttps://doi.org/10.1016/j.fbr.2010.03.001
9. Bergsten J (2005) A review of long-branch attraction. Cladistics 21:163–193. ttps://doi.org/10.1111/j.1096-0031.2005.00059.x
10. Blackwell M (2011) The fungi: 1, 2, 3… million species? Am J Bot 98:426–438. ttps://doi.org/10.3732/ajb.1000298
11. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. ttps://doi.org/10.1093/bioinformatics/btu170
12. Boto L (2010) Horizontal gene transfer in evolution: Facts and challenges. Proc R Soc B Biol Sci 277:819–827. ttps://doi.org/10.1098/rspb.2009.1679
13. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni I, Drummond AJ (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput Biol 15:e1006650. ttps://doi.org/10.1371/journal.pcbi.1006650
14. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10:421. ttps://doi.org/10.1186/1471-2105-10-421
15. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active enZYmes database (CAZy): An expert resource for glycogenomics. Nucleic Acids Res 37:D233–D238. ttps://doi.org/10.1093/nar/gkn663
16. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973. ttps://doi.org/10.1093/bioinformatics/btp348
17. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. Cell 134:25–36. ttps://doi.org/10.1016/j.cell.2008.06.030
18. Chang ES, Neuhof M, Rubinstein ND, Diamant A, Philippe H, Huchon D, Cartwright P (2015) Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc Natl Acad Sci U S A 112:14912–14917. ttps://doi.org/10.1073/pnas.1511468112
19. Crisp A, Boschetti C, Perry M, Tunnacliffe A, Micklem G (2015) Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. Genome Biol 16:50. ttps://doi.org/10.1186/s13059-015-0607-3
20. Danchin EGJ, Rosso MN, Vieira P et al (2010) Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. Proc Natl Acad Sci U S A 107:17651–17656. ttps://doi.org/10.1073/pnas.1008486107
21. Danchin EGJ (2016) Lateral gene transfer in eukaryotes: Tip of the iceberg or of the ice cube? BMC Biol 14:101. ttps://doi.org/10.1186/s12915-016-0330-x
22. Darling AE, Jospin G, Lowe E, Matsen FAIV, Bik HM, Eisen JA (2014) PhyloSift: Phylogenetic analysis of genomes and metagenomes. PeerJ 2:e243. ttps://doi.org/10.7717/peerj.243

23. Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: Fast selection of best-fit models of protein evolution. Bioinformatics 27:1164–1165. ttps://doi.org/10.1093/bioinformatics/btr088
24. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: More models, new heuristics and parallel computing. Nat Methods 9:772. ttps://doi.org/10.1038/nmeth.2109
25. Dean M, Rzhetsky A, Allikmets R (2001) The human ATP-binding cassette (ABC) transporter superfamily. Genome Res 11:1156–1166. ttps://doi.org/10.1101/gr.184901
26. Doolittle WF (1999) Phylogenetic classification and the universal tree. Science 284:2124–2128. ttps://doi.org/10.1126/science.284.5423.2124
27. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4:e88. ttps://doi.org/10.1371/journal.pbio.0040088
28. Dunning Hotopp JC (2011) Horizontal gene transfer between bacteria and animals. Trends Genet 27:157–163. ttps://doi.org/10.1016/j.tig.2011.01.005
29. Erixon P, Svennblad B, Britton T, Oxelman B (2003) Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst Biol 52:665–673. ttps://doi.org/10.1080/10635150390235485
30. Erwin DH (1993) The great Paleozoic crisis: Life and death in the Permian. Columbia University
31. Erwin DH (2006) Extinction: How life on Earth nearly ended 250 million years ago. Princeton University Press
32. Evans NM, Holder MT, Barbeitos MS, Okamura B, Cartwright P (2010) The phylogenetic position of Myxozoa: Exploring conflicting signals in phylogenomic and ribosomal data sets. Mol Biol Evol 27:2733–2746. ttps://doi.org/10.1093/molbev/msq159
33. Faber M, Shaw S, Yoon S, Kumar G, Ertl R, Miller MR, Williams TD, Kirkham S, Lia L, Fast MD, Johnson SC, Jones SRM (2021) Comparative transcriptomics and host-specific parasite gene expression profiles inform on drivers of proliferative kidney disease. Sci Rep 11:2149. ttps://doi.org/10.1038/s41598-020-77881-7
34. Fedonkin MA, Simonetta A, Ivantsov AY (2007) New data on Kimberella, the Vendian mollusc-like organism (White Sea region, Russia): Palaeoecological and evolutionary implications. Geol Soc Lond Spec Publ 286:157–179. ttps://doi.org/10.1144/SP286.12
35. Fiala I (2006) The phylogeny of Myxosporea (Myxozoa) based on small subunit ribosomal RNA gene analysis. Int J Parasitol 36:1521–1534. ttps://doi.org/10.1016/j.ijpara.2006.06.016
36. Flajnik MF, Kasahara M (2010) Origin and evolution of the adaptive immune system: Genetic events and selective pressures. Nat Rev Genet 11:47–59. ttps://doi.org/10.1038/nrg2703
37. Foox J, Siddall ME, Bonello P, Bartholomew JL (2015) In silico hybridization enables transcriptomic illumination of the nature and evolution of Myxozoa. BMC Genomics 16:840. ttps://doi.org/10.1186/s12864-015-2039-6
38. Gladyshev EA, Meselson M, Arkhipova IR (2008) Massive horizontal gene transfer in bdelloid rotifers. Science 320:1210–1213. ttps://doi.org/10.1126/science.1156407
39. Gozlan RE, Marshall WL, Lilje O, Jessop CN, Gleason FH, Andreou D (2014) Current ecological understanding of fungal-like pathogens of fish: What lies beneath? Front Microbiol 5:62. ttps://doi.org/10.3389/fmicb.2014.00062
40. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K,

Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652. ttps://doi.org/10.1038/nbt.1883

41. Gruhl A, Okamura B (2015) Tissue characteristics and development in Myxozoa. In: Okamura B, Gruhl A, Bartholomew JL (eds) Myxozoan evolution, ecology and development. Springer, Cham, pp 155–174. ttps://doi.org/10.1007/978-3-319-14753-6_9
42. Hartigan A, Kosakyan A, Pecková H, Eszterbauer E, Holzer AS, Fiala I, Okamura B (2016) Transcriptome analysis of *Tetracapsuloides bryosalmonae* (Cnidaria: Myxozoa) suggests functions for some genes in host invasion and development. Parasitology 143:1923–1943. ttps://doi.org/10.1017/S0031182016001633
43. Hedges SB, Kumar S (eds) (2009) The timetree of life. Oxford University Press. ttps://doi.org/10.1093/oso/9780199535033.001.0001
44. Hedrick RP, McDowell TS, Mukkatira K, Georgiadis MP, MacConnell E (2008) Susceptibility of selected inland salmonids to experimentally induced infections with *Myxobolus cerebralis*, the causative agent of whirling disease. J Aquat Anim Health 20:219–229. ttps://doi.org/10.1577/H08-025.1
45. Heitman J, Howlett BJ, Crous PW, Stukenbrock EH, James TY, Gow NAR (eds) (2017) The fungal kingdom. ASM. ttps://doi.org/10.1128/9781555819583
46. Higgins CF (1992) ABC transporters: From microorganisms to man. Annu Rev Cell Biol 8:67–113. ttps://doi.org/10.1146/annurev.cb.08.110192.000435
47. Higgins C (2007) Multiple molecular mechanisms for multidrug resistance transporters. Nature 446:749–757. ttps://doi.org/10.1038/nature05630
48. Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol 42:182–192. ttps://doi.org/10.1093/sysbio/42.2.182
49. Holzer AS, Bartošová-Sojková P, Born-Torrijos A, Lövy A, Hartigan A, Fiala I (2018) The joint evolution of the Myxozoa and their alternate hosts: A cnidarian recipe for success and vast biodiversity. Mol Ecol 27:1651–1666. ttps://doi.org/10.1111/mec.14558
50. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: Improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. ttps://doi.org/10.1093/molbev/msx281
51. Husnik F, McCutcheon JP (2018) Functional horizontal gene transfer from bacteria to eukaryotes. Nat Rev Microbiol 16:67–79. ttps://doi.org/10.1038/nrmicro.2017.137
52. Isozaki Y (1997) Permo-Triassic boundary superanoxia and stratified superocean: Records from lost deep sea. Science 276:235–238. ttps://doi.org/10.1126/science.276.5310.235
53. Jiménez-Guri E, Philippe H, Okamura B, Holland PWH (2007) Buddenbrockia is a cnidarian worm. Science 317:116–118. ttps://doi.org/10.1126/science.1142024
54. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. ttps://doi.org/10.1038/nmeth.4285
55. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30:772–780. ttps://doi.org/10.1093/molbev/mst010
56. Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 9:605–618. ttps://doi.org/10.1038/nrg2386

57. Keller NP (2019) Fungal secondary metabolism: Regulation, function and drug discovery. Nat Rev Microbiol 17:167–180. ttps://doi.org/10.1038/s41579-018-0121-1
58. Knoll AH, Bambach RK, Canfield DE, Grotzinger JP (1996) Comparative Earth history and Late Permian mass extinction. Science 273:452–457. ttps://doi.org/10.1126/science.273.5274.452
59. Ku C, Nelson-Sathi S, Roettger M, Sousa FL, Lockhart PJ, Bryant D, Hazkani-Covo E, McInerney JO, Landan G, Martin WF (2015) Endosymbiotic origin and differential loss of eukaryotic genes. Nature 524:427–432. ttps://doi.org/10.1038/nature14963
60. Laetsch DR, Blaxter ML (2017) BlobTools: Interrogation of genome assemblies. *F1000Research* 6:1287. ttps://doi.org/10.12688/f1000research.12232.1
61. Lawrence JG, Ochman H (1997) Amelioration of bacterial genomes: Rates of change and exchange. J Mol Evol 44:383–397. ttps://doi.org/10.1007/PL00006158
62. Lom J, Dyková I (2006) Myxozoan genera: Definition and notes on taxonomy, life-cycle terminology and pathogenic species. Folia Parasitol 53:1–36. ttps://doi.org/10.14411/fp.2006.001
63. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. ttps://doi.org/10.1186/s13059-014-0550-8
64. Lynch M (2007) The origins of genome architecture. Sinauer Associates
65. Merchant S, Wood DE, Salzberg SL (2014) Unexpected cross-species contamination in genome sequencing projects. PeerJ 2:e675. ttps://doi.org/10.7717/peerj.675
66. Mira A, Ochman H, Moran NA (2001) Deletional bias and the evolution of bacterial genomes. Trends Genet 17:589–596. ttps://doi.org/10.1016/S0168-9525(01)02447-7
67. Moran Y, Fredman D, Szczesny P, Grynberg M, Technau U (2012) Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. Mol Biol Evol 29:2223–2230. ttps://doi.org/10.1093/molbev/mss089
68. Nakamura Y, Itoh T, Matsuda H, Gojobori T (2004) Biased biological functions of horizontally transferred genes in prokaryotic genomes. Nat Genet 36:760–766. ttps://doi.org/10.1038/ng1381
69. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. ttps://doi.org/10.1093/molbev/msu300
70. Okamura B, Gruhl A, Bartholomew JL (eds) (2015) Myxozoan evolution, ecology and development. Springer. ttps://doi.org/10.1007/978-3-319-14753-6
71. Okamura B, Hartikainen H, Schmidt-Posthaus H, Wahli T (2011) Life cycle complexity, environmental change and the emerging status of salmonid proliferative kidney disease. Freshw Biol 56:735–753. ttps://doi.org/10.1111/j.1365-2427.2010.02465.x
72. Okamura B, Gruhl A (2021) Evolution, origins and diversification of parasitic cnidarians. In: De Baets K, Huntley JW (eds) *The evolution and fossil record of parasitism*. Top Geobiol 49:247–272. Springer. ttps://doi.org/10.1007/978-3-030-42484-8_4
73. Paradis E, Schliep K (2019) ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528. ttps://doi.org/10.1093/bioinformatics/bty633
74. Park E, Hwang DS, Lee JS, Song JI, Seo TK, Won YJ (2012) Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. Mol Phylogenet Evol 62:329–345. ttps://doi.org/10.1016/j.ympev.2011.10.008

75. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417–419. ttps://doi.org/10.1038/nmeth.4197
76. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov VV, Jurka J, Genikhovich G, Grigoriev IV, Lucas SM, Steele RE, Finnerty JR, Technau U, Martindale MQ, Rokhsar DS (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317:86–94. ttps://doi.org/10.1126/science.1139158
77. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904. ttps://doi.org/10.1093/sysbio/syy032
78. Ravenhall M, Škunca N, Lassalle F, Dessimoz C (2015) Inferring horizontal gene transfer. PLoS Comput Biol 11:e1004095. ttps://doi.org/10.1371/journal.pcbi.1004095
79. Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. Science 289:1920–1921. ttps://doi.org/10.1126/science.289.5486.1920
80. Rees D, Johnson E, Lewinson O (2009) ABC transporters: The power to change. Nat Rev Mol Cell Biol 10:218–227. ttps://doi.org/10.1038/nrm2646
81. Reis M, Yang Z (2011) Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. Mol Biol Evol 28:2161–2172. ttps://doi.org/10.1093/molbev/msr045
82. Reis M, Thawornwattana Y, Angelis K, Telford MJ, Donoghue PCJ, Yang Z (2015) Uncertainty in the timing of origin of animals and the limits of precision in molecular timescales. Curr Biol 25:2939–2950. ttps://doi.org/10.1016/j.cub.2015.09.066
83. Revell LJ (2012) phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223. ttps://doi.org/10.1111/j.2041-210X.2011.00169.x
84. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. ttps://doi.org/10.1093/sysbio/sys029
85. Salzberg SL, White O, Peterson J, Eisen JA (2001) Microbial genes in the human genome: Lateral transfer or gene loss? Science 292:1903–1906. ttps://doi.org/10.1126/science.1061036
86. Schönknecht G, Chen WH, Ternes CM, Barbier GG, Shrestha RP, Stanke M, Bräutigam A, Baker BJ, Banfield JF, Garavito RM, Carr K, Wilkerson C, Rensing SA, Gagneul D, Dickenson NE, Oesterhelt C, Lercher MJ, Weber APM (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339:1207–1210. ttps://doi.org/10.1126/science.1231707
87. Sela I, Ashkenazy H, Katoh K, Pupko T (2015) GUIDANCE2: Accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. Nucleic Acids Res 43:W7–W14. ttps://doi.org/10.1093/nar/gkv318
88. Selman M, Corradi N (2011) Microsporidia: Horizontal gene transfers in vicious parasites. Mob Genet Elem 1:251–255. ttps://doi.org/10.4161/mge.18611
89. Sharp PM, Li WH (1987) The codon adaptation index—A measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15:1281–1295. ttps://doi.org/10.1093/nar/15.3.1281
90. Sharp PM, Emery LR, Zeng K (2010) Forces that influence the evolution of codon bias. Philos Trans R Soc Lond B Biol Sci 365:1203–1212. ttps://doi.org/10.1098/rstb.2009.0305

91. Shearer CA, Descals E, Kohlmeyer B, Kohlmeyer J, Marvanová L, Padgett D, Porter D, Raja HA, Schmit JP, Thorton HA, Voglymayr H (2007) Fungal biodiversity in aquatic habitats. Biodivers Conserv 16:49–67. ttps://doi.org/10.1007/s10531-006-9120-z
92. Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. Syst Biol 51:492–508. ttps://doi.org/10.1080/10635150290069913
93. Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16:1114–1116. ttps://doi.org/10.1093/oxfordjournals.molbev.a026201
94. Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, Fujie M, Fujiwara M, Koyanagi R, Ikuta T, Fujiyama A, Miller DJ, Satoh N (2011) Using the *Acropora digitifera* genome to understand coral responses to environmental change. Nature 476:320–323. ttps://doi.org/10.1038/nature10249
95. Sibbald SJ, Archibald JM (2020) Genomic insights into plastid evolution. Genome Biol Evol 12:978–990. ttps://doi.org/10.1093/gbe/evaa096
96. Silverman GA, Bird PI, Carrell RW, Church FC, Coughlin PB, Gettins PGW, Irving JA, Lomas DA, Luke CJ, Moyer RW, Pemberton PA, Remold-O'Donnell E, Salvesen GS, Travis J, Whisstock JC (2001) The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. J Biol Chem 276:33293–33296. ttps://doi.org/10.1074/jbc.R100016200
97. Soucy SM, Huang J, Gogarten JP (2015) Horizontal gene transfer: Building the web of life. Nat Rev Genet 16:472–482. ttps://doi.org/10.1038/nrg3962
98. Stanhope MJ, Lupas A, Italia MJ, Koretke KK, Volker C, Brown JR (2001) Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature 411:940–944. ttps://doi.org/10.1038/35082058
99. Strimmer K, von Haeseler A (1997) Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. Proc Natl Acad Sci U S A 94:6815–6819. ttps://doi.org/10.1073/pnas.94.13.6815
100. Taylor JW, Berbee ML (2006) Dating divergences in the Fungal Tree of Life: Review and new analyses. Mycologia 98:838–849. ttps://doi.org/10.1080/15572536.2006.11832614
101. Tsaousis AD, Kunji ERS, Goldberg AV, Lucocq JM, Hirt RP, Embley TM (2008) A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. Nature 453:553–556. ttps://doi.org/10.1038/nature06903
102. Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Graham SW, Mathews S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Leebens-Mack J (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc Natl Acad Sci U S A 111:E4859–E4868. ttps://doi.org/10.1073/pnas.1323926111
103. Wickham H (2016) ggplot2: Elegant graphics for data analysis. Springer-
104. Wilson JE (2003) Isozymes of mammalian hexokinase: Structure, subcellular localization and metabolic function. J Exp Biol 206:2049–2057. ttps://doi.org/10.1242/jeb.00241
105. Whipps C, Atkinson S, Hoeksema B (2025) World List of Myxozoa. Available at: https://www.marinespecies.org/myxozoa/. ttps://doi.org/10.14284/756
106. Wood DE, Lu J, Langmead B (2019) Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. ttps://doi.org/10.1186/s13059-019-1891-0

107. Woolfit M, Iturbe-Ormaetxe I, McGraw EA, O'Neill SL (2009) An ancient horizontal gene transfer between mosquito and the endosymbiotic bacterium *Wolbachia*. Mol Biol Evol 26:367–374. ttps://doi.org/10.1093/molbev/msn253

108. Xia X (2013) DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol 30:1720–1728. ttps://doi.org/10.1093/molbev/mst064

109. Yahalomi D, Atkinson SD, Neuhof M, Chang ES, Philippe H, Cartwright P, Bartholomew JL, Huchon D (2020) A cnidarian parasite of salmon (*Henneguya salminicola*) lacks a mitochondrial genome. Proc Natl Acad Sci U S A 117:5358–5363. ttps://doi.org/10.1073/pnas.1909907117

110. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY (2017) ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol 8:28–36. ttps://doi.org/10.1111/2041-210X.12628

111. Zhao Z, Liu H, Wang C, Xu JR (2013) Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics 14:274. ttps://doi.org/10.1186/1471-2164-14-274

Figures

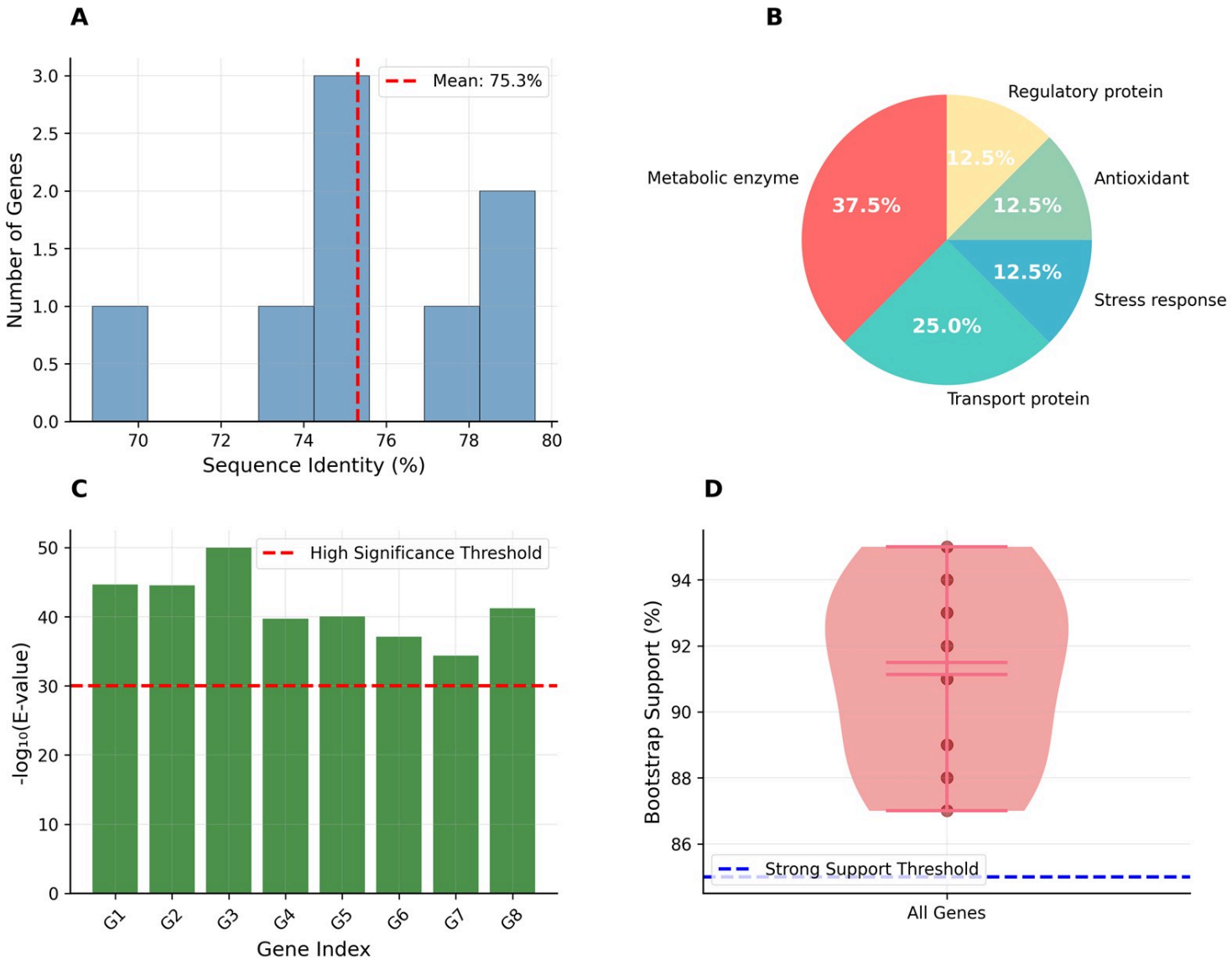

Figure 1

Comprehensive overview of horizontal gene transfer candidates.(A) Distribution of sequence identities between myxozoan and fungal orthologs, showing mean identity of 74.1% (red dashed line). (B) Functional categorization of transferred genes, with metabolic enzymes representing the largest category. (C) Statistical significance of BLAST matches, with all candidates (genes: G1-G8) showing E-values below 10^{-30} (red dashed line indicates high significance threshold). (D) Phylogenetic bootstrap support values, with all candidates exceeding 85% support threshold (blue dashed line).

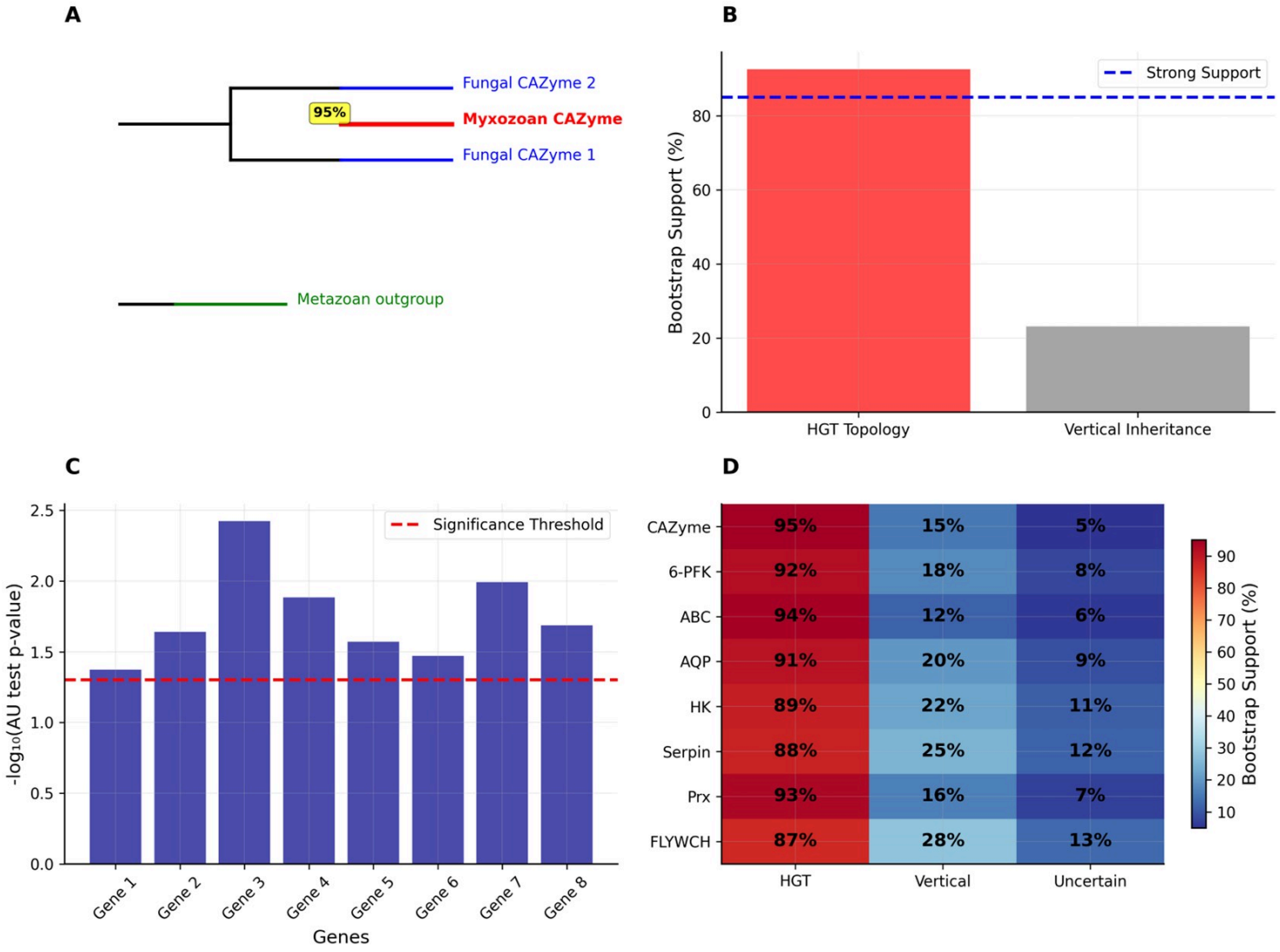

Figure 2

Phylogenetic evidence supporting horizontal gene transfer. (A) Representative phylogenetic tree for CAZyme showing myxozoan sequence (red) clustering within fungal clade (blue) with 95% bootstrap support, separated from metazoan outgroup (green). (B) Comparison of bootstrap support for HGT topology versus vertical inheritance, demonstrating strong preference for horizontal origin. (C) Results of approximately unbiased (AU) topology tests showing significant rejection of vertical inheritance hypothesis for all genes. (D) Heatmap of phylogenetic support across all HGT candidates and alternative topologies, with warm colors indicating strong support for HGT hypothesis.

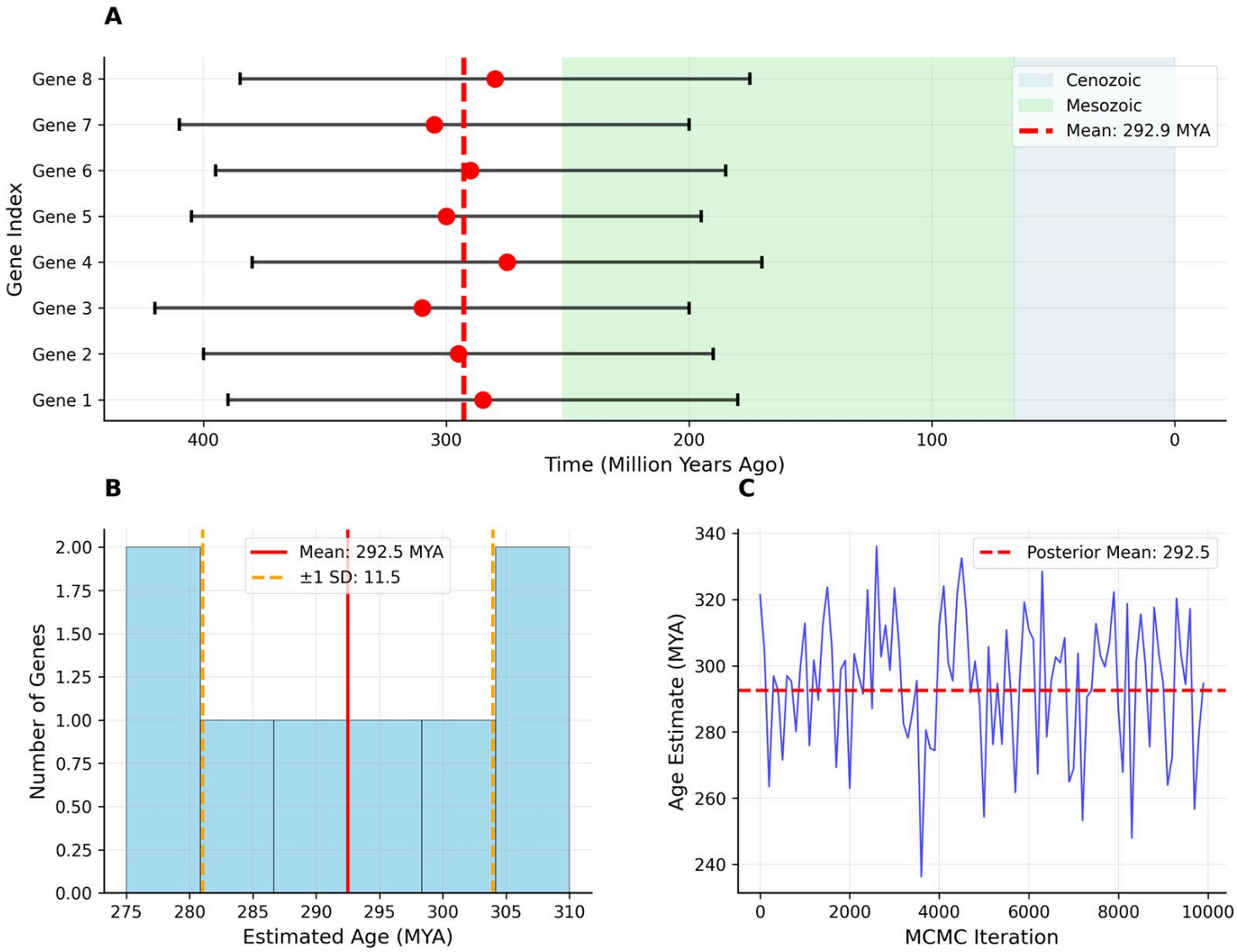

Figure 3

Molecular clock analysis of HGT timing. (A) Chronogram showing estimated ages of HGT events with confidence intervals, overlaid on geological time scale. Mean estimate of 292.9 MYA (red dashed line) falls within Carboniferous-Permian transition. (B) Distribution of age estimates across genes showing consistency around mean value with standard deviation of 114.5 MY. (C) MCMC trace plot demonstrating convergence of Bayesian analysis with stable posterior distribution around mean estimate.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FiguresS1S16.pptx
- FileS3.zip
- FileS1.zip
- FileS2.zip

- FileS4.zip
- TableS1.xlsx
- supp.jpg