FY23 Grantee Final Report

I. Overview

1. **Grant Number:** NA19OAR0110303
2. **Project Title:** Bioprospecting for industrial enzymes and drug lead compounds in an ancient submarine forest.
3. **Area of Operation (include a map and/or coordinates):**

Figure 1: Approximate site location. The site is located approximately 8 miles south of Orange Beach, AL. Due to the threat of commercial exploitation of the ancient wood found at this site, we are revealing the site co-ordinates only on a need-to-know basis. Please contact the lead PI to request more detailed site information.

4. **Co-PI(s), Significant Participants, Participating Institutions, and other affiliated personnel:**
    a. Northeastern University, Marine Science Center (NU)
    b. University of Utah (UU)
    c. Max Plank Institute, Bremen
    d. Johns Hopkins University
    e. Israel Oceanographic and Limnological Research Institute
    f. National Institute of Health/National Library of Medicine
    g. Dauphin Island Sea Lab
5. **Award Period:** From 09/01/2019 to 08/31/2023

II. Summary

1. Abstract—1-paragraph description of final report

In this project, we addressed Objective 1 of the 2019 NOFO—discovering microorganisms, sponges, corals, and other organisms with biopharmaceutical or biotechnical potential—by mapping, documenting, describing, archiving, and exploring the biodiversity of a prehistoric submerged bald cypress forest off the coast of Alabama. Our primary focus was on wood-associated animal species and their microbiomes. We

collected, identified, georeferenced, vouchered, and subsampled these animals, cultivated their associated bacterial symbionts, and sequenced the symbiont genomes and host metagenomes. We performed phylogenetic and bioinformatic analyses to identify these organisms and their associated antimicrobial compounds and hydrolytic enzymes capable of degrading the components of wood (lignocellulose). We then performed biochemical and growth inhibition assays to confirm the activities of these compounds. We discovered a new antibiotic with activity against *Acinetobacter*, an emerging multidrug-resistant pathogen of growing public health concern. We described and formally named one new bivalve genus and species and four new microbial species and submitted voucher specimens and isolates to appropriate public collections. We discovered and isolated at least 25 new bacterial species and hundreds of new bacterial strains and deposited them to the publicly accessible Ocean Genome Legacy biorepository. We showed that these bacteria, now available for further characterization, have great potential for drug and enzyme discovery. We determined the mechanism by which wood-boring bivalves (shipworms) produce and export cellulolytic enzymes and possibly antimicrobial compounds made by their symbiotic bacteria. Finally, we generated 3D maps of the site using photogrammetry and sidescan sonar. Our research demonstrated that marine wood deposits, like the Alabama undersea forest, are rich sources for discovering new animal and bacterial species and associated compounds of pharmaceutical and industrial value.

2. Purpose of Project

2.a. Describe topic and/or questions that were addressed:

The biopharmaceutical and biotechnological potential of marine environments present exciting, lucrative, and socially beneficial opportunities for the Nation’s Blue Economy. To explore and exploit this potential, we employed a comprehensive and systematic approach that spans spatial scales from ecosystems to molecules and used technological, ecological, physiological, biochemical, and evolutionary knowledge to guide and drive exploratory efforts. We aimed to systematically explore the diversity and biotechnological potential of a prehistoric undersea forest found off the coast of Alabama. Although common, such wood-powered benthic ecosystems are still largely unexplored. Our goal was to describe this novel ecosystem and assess the potential of its biodiversity as targets for the directed discovery of biomolecules of high biopharmaceutical and biotechnological value.

2.b. Describe/list the project objectives:

2.b.1. Biodiversity survey and site mapping

Perform a comprehensive survey of wood-associated taxa in the Alabama Undersea Forest (AUF) site and map patterns of biodiversity and species distribution to the three-dimensional physical structure of the forest.

2.b.2. Symbiont isolation and characterization

Isolate bacterial strains from these species and from wood substrates.

2.b.3. Gene and genome sequencing

Sequence the genes and genomes of bacterial isolates and microbial metagenomes of wood-associated animals and use these to identify bacterial species, genes, isolates, and pathways of interest.

2.b.4. Biochemical screening

Employ biochemical screens and growth inhibition assays to search for expression of lignocellulolytic enzymes and compounds with analgesic, antibiotic, and antimicrobial activities.

2.b.5. Compound discovery

Use a combination of bioactivity-guided purification and exogenous expression to obtain purified molecules for subsequent biochemical and physiological characterization.

2.b.6. Documentation

Describe, document, and disseminate animal and microbial taxa and associated data by:

- Depositing biological materials in and recording associated contextual data to the Ocean Genome Legacy biorepository and database.
- Reporting biogeographic occurrence data to public databases, e.g., GBIF and OBIS.
- Submitting sequence information to appropriate public repositories, e.g., GenBank NCBI.
- Formally describing selected bacterial isolates and microbiomes, animal species, genomic and metagenomic sequences, and potential drug and enzyme leads in appropriate journals.
- Submitting bacterial type strains to appropriate collections.

3. Approach

3.a. Describe the work that was performed:

3.a.1. Bacterial (symbiont) genome sequencing and annotation

During this project, 55 new shipworm symbiont genomes were sequenced, assembled, and annotated. Of these, 21 were "finished" (closed and polished circular genomes, Table 1), five are pending closure, and 29 were completed to "draft" or "improved quality draft" status.

#	16S clade	Isolate	Host species	Host location	Illumina assembly size, bp	Illumina assembly # contigs	Hybrid assembly size, bp	Hybrid assembly cc, closed circular, s, scaffolds	*nif* genes	antiSMASH clusters	CAZymes GH	CAZymes PL	CAZymes CE	CAZymes CBM	CAZymes AA10	Host voucher ID
1	1c	SR1834	Bankia gouldi	AUF, AL, USA	5,103,048	22	5,128,112	1, cc	yes	12	100	7	8	107	1	NOAA-5140L
2	3a	SR1855	Bankia gouldi	AUF, AL, USA	5,789,774	29	5,813,470	1, cc	yes	13	140	8	17	169	1	NOAA-5141M
3	10a	SR1925	Bankia gouldi	AUF, AL, USA	4,450,368	21	4,471,731	1, cc	no	5	94	5	9	81	1	NOAA-5312X
4	5	PMS-486K.S.1a.10	Dicyathifer mannii	Bohol, Philippines	5,728,280	105	5,760,468	1, cc	yes	11	148	17	18	193	3	n/a
5	11	SR1846	Jouannetia quillingi	AUF, AL, USA	8,169,842	66	8,270,762	1, cc	yes	13	228	29	12	304	1	NOAA-5075M
6	12	T1312	Lyrodus pedicellatus	Alamitos bay, CA, USA	6,244,385	56	5,771,624	1,cc	yes	12 + 1 mini on plasmid	-	-	-	-	-	n/a
7	3c	SR1920	Lyrodus pedicellatus	Merrit Isl., FL, USA	5,669,777	31	5,699,908	1, cc	yes	12	145	11	13	192	1	F0-FL03.01
8	10c	Lp-A-06	Lyrodus pedicellatus	Alamitos bay, CA, USA	4,999,590	72	5,079,498	1, cc	no	10	110	7	11	120	1	Lp-A-
9	3a	SR1906	Lyrodus pedicellatus	Merrit Isl., FL, USA	5,712,711	71	5,805,936	1, cc	yes	14	131	7	19	159	1	F0-FL03.01
10	8	SR1850	Nototeredo knoxi	AUF, AL, USA	3,643,507	195	3,949,449	1, cc	yes	3	81	2	8	72	2	NOAA-5025W
11	6	SR1874	Nototeredo knoxi	AUF, AL, USA	-	-	3,517,046	1, cc	yes	4	79	1	4	49	1	NOAA-5075M
12	8	SR1816	Nototeredo knoxi	AUF, AL, USA	4,012,486	14	3,648,677	1, cc	yes	3	84	2	10	85	3	NOAA-5025W
13	4	SR1818	Teredothyra matocotana	AUF, AL, USA	4,660,654	113	4,848,863	1, cc	yes	10	-	-	-	-	-	NOAA-5032P
14	10d	SI-064W.S.0b.05	Teredo somersi	Solomon Islands	4,631,015	19	4,652,372	1, cc	no	3	90	6	12	79	1	n/a
15	6	PMS-3231U.S.0c.02	Zachsia sp.	Bohol, Philippines	5,321,153	196	5,565,486	1, cc	yes	7	96	21	9	92	3	n/a
16	1a	SR1969	Lyrodus pedicellatus	Merrit Island, FL, USA	-	-	4,539,334	1, cc	-	-	-	-	-	-	-	-
17	1a	SR2026	Lyrodus pedicellatus	Merrit Island, FL, USA	5,150,607	37	5,150,607	1, cc	-	-	-	-	-	-	-	-
18	10	SR1967	Lyrodus pedicellatus	Merrit Island, FL, USA	-	-	4,539,232	1, cc	-	-	-	-	-	-	-	-
19	1a	T. turnerae T7901	Bankia gouldi	Beaufort, NC, USA	5,193,164	1	-	1, cc	yes		102	6	17	120	1	n/a
20	10b	PMS-1162T.S.0a.05	Lyrodus pedicellatus	Bohol, Philippines	4,404,964	1	-	1, s	no	-	-	-	-	-	-	n/a
21	6	T. haidensis Bs08	Bankia setacea	Puget Sound, WA, USA	4,915,572	1	-	1, cc	yes	-	-	-	-	-	-	n/a
22	6	T. purpureus Bs12	Bankia setacea	Puget Sound, WA, USA	-	-	4,643,613	1, cc	no	-	85	1	10	88	2	n/a
23	3b	PMS-1120W.S.0a.04	Teredo fulleri	Bohol, Philippines	5,699,307	1		1, s	yes	-	-	-	-	-	-	n/a
24	10	SR2048	Bankia gouldi	AUF, AL, USA	3,529,209	69	pending	pending	no	2	-	-	-	-	-	-
25	10	SR2011	Seawater	Nahant, MA, USA	3,541,351	7	pending	pending	no	1	-	-	-	-	-	-
26	4	SR1870	Nototeredo knoxi	AUF, AL, USA	4,609,851	197	pending	pending	no	8	-	-	-	-	-	-
27	10	SR2059	Bankia gouldi	AUF, AL, USA	4,464,691	22	pending	pending	no	2	-	-	-	-	-	-

Table 1. Progress toward closure and polishing of symbiont genomes. Twenty-one symbiont genomes are "finished" (closed and polished circular genomes). An additional 5 are pending closure, and 29 were completed to "draft" or "improved quality draft" status. One hundred and forty-five potential secondary metabolite pathways and genes encoding 2,027 carbohydrate-active enzymes have been identified. 16S clade # refers to clades identified in Figures 17-20. Hybrid refers to assemblies that combine Illumina short-read and Oxford Nanopore long-reads to improve contiguity. Cc, closed circular genome, single contig; s, scaffold; nif genes, genes for nitrogen fixation; antiSMASH clusters, genes encoding secondary metabolite pathways; CAZymes, genes encoding carbohydrate-active enzymes; GH, glycoside hydrolase; PL, pectin lyase; CE, carbohydrate esterase; CBM, carbohydrate-binding module; AA10, lytic polysaccharide monooxygenase. -, data not yet determined; n/a, data not available.

3.a.2. 16S rRNA gene sequencing

16S rRNA sequences were determined for 360 bacterial strains and analyzed phylogenetically.

3.a.3. Mitochondrial genome and 18 and 28S ribosomal gene sequencing

The complete mitochondrial genome and two nuclear genes (18S and 28S rRNA) were recovered from metagenome sequences, assembled, and annotated from five specimens of *Idas modioliformis*, and included as reference sequences in updated multigene phylogenies to describe the phylogenetic position of the new bathymodiolin mussel within Mytilidae.

3.a.4. Metagenome sequencing

Metagenome sequences were determined for the posterior gills of two specimens of *Teredothyra matocotana*, two specimens of *Lyrodus pedicellatus*, and one specimen of *Bankia gouldi*. Additionally, the anterior gills of two specimens and for the posterior gills of an additional two specimens of *Nototeredo knoxi*, inspired by the discovery in the prior performance period that, unlike other shipworm species examined to date, the symbiont community of this species is largely found in the anterior gill rather than the posterior gill.

3.b. Describe how the project was organized and managed (e.g., roles and responsibilities of participants):

The project was co-managed by the lead principal investigator from Northeastern University (NU) and co-PIs from NU and the University of Utah (UU). NU PIs were responsible for 2D and 3D site mapping. UU PIs were responsible for medicinal and analytical chemistry. Genome and metagenome sequencing were the joint responsibility of NU and UU labs. All PIs and members of their respective lab groups took part in field work, including both site mapping and biological collections. The lead PI at NU had responsibility for documentation, identification, dissection, preservation and storage of biological samples and distribution of samples to all participants. The lead PI at NU was primarily responsible for bacterial isolations, although the UU participants contributed to microbial isolations performed in the field and to batch cultivation and manipulations of strains for analytical chemistry. All PIs and participants met on a regular basis, every six weeks via video conferencing, with additional meetings called as needed. Publications were co-authored based on individual contribution irrespective of lab affiliation.

3.c. Describe how data were organized, processed, and archived to meet NOAA data management requirements:

3.c.1. The Ocean Genome Legacy Database

The OGL Database served as the primary organizational tool for tracing and organizing collections, taxonomy, sample storage and provenance. We designed this relational database in house using Filemaker Pro. The data is currently being migrated to the Arctos Collaborative Collection Management Solution https://arctos.database.museum/home.cfm and is expected to be publicly available there early in 2024.

3.c.2. Biogeographic occurrence and taxonomic data

Biogeographic occurrence and taxonomic data are being published to the Global Biodiversity Information Facility (GBIF; https://www.gbif.org) and the Ocean Biodiversity Information System (OBIS; https://obis.org).

3.c.3. Sequence data

Sequence data has been submitted to Genbank (https://www.ncbi.nlm.nih.gov/genbank), see publications.

3.c.4. Site mapping data

All processed models, including image data, will reside at NOAA National Centers for Environmental Information (NCEI) and be made available to the public. Processed models are captured as video files (.mp4) for outreach at this public YouTube channel of the Field Robotics Laboratory: https://www.youtube.com/channel/UCUUfgiSvrk5lQanwS_SqQyw.

4. Findings

4.a. Describe actual accomplishments and findings (provide maps, graphics, and images):

4.a.1. A newly discovered shallow water bathymodiolin genus and species

Among the most astonishing findings of this project has been the discovery of a new genus and species of the primarily deep-sea mytilid subfamily Bathymodiolinae living in vacated shipworm burrows within found wood and settlement panels at the AUF site (Figure 2). Mussels of the subfamily Bathymodiolinae encompass nine accepted genera (*Adipicola, Bathymodiolus, Benthomodiolus, Gigantidas, Idas, Nypamodiolus, Tamu, Terua*, and *Vulcanidas*.) Two more genera, *Lignomodiolus* and *Nipponiomodiolus,* have been proposed but not formally described and so are nomen nuda. All species described to date live in association with deep-sea reducing environments including hydrothermal vents, cold seeps, and decaying organic deposits such as wood and whale falls. All but one species harbors chemoautotrophic or methanotrophic bacterial endosymbionts within their gill tissues that contribute to host nutrition. Here we report the discovery of a new bathymodiolin mussel genus and species and describe it in context of biogeographic and bathymetric distribution, morphology, and taxonomy, molecular phylogenetics, mitochondrial genomics, and symbiont identity and localization.

Known depth distribution of Bathymodiolinae: Bathymodiolinae is reputed to be comprised exclusively of deep-water species. The most famous of these are the large chemoautotrophic mussels of the genera *Bathymodilus* and *Gigantidas* found at deep-sea hydrothermal vents and methane seeps. Less well known are the small to minute species of the genera *Idas* and *Adipicola*, which are associated with organic deposits, including wood and large vertebrate bones, in the deep sea. This new species demonstrates the first verified occurrence of this subfamily growing and reproducing in shallow water.

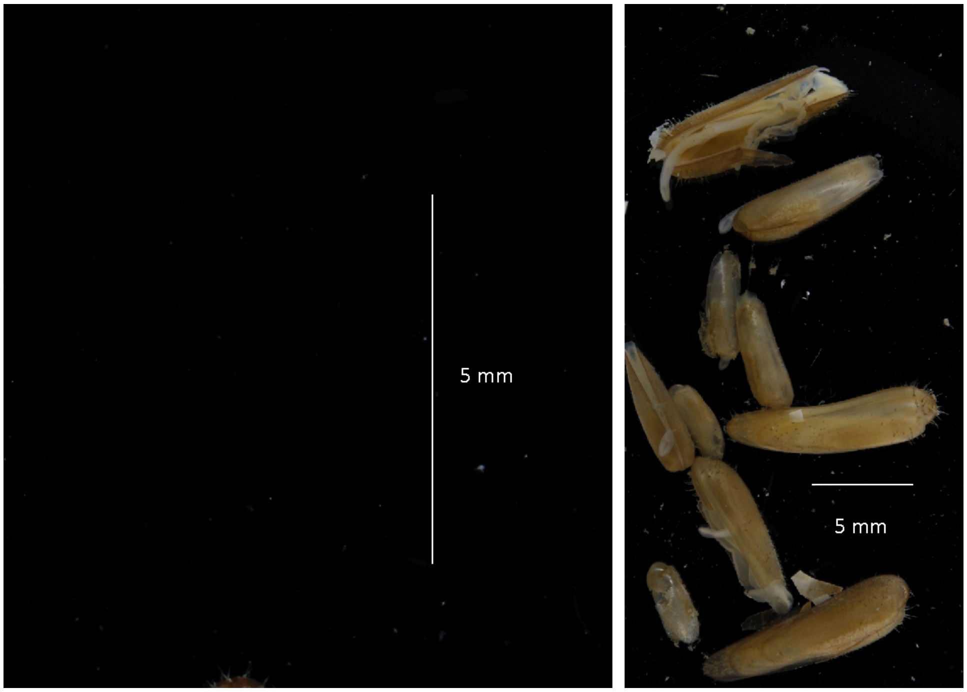

Figure 2: Newly discovered mussel species of the mytilid subfamily Bathymodiolinae. These specimens, likely represent a new species and genus and are the first members of this chemosymbiotic “deep-sea” subfamily found in shallow water.

To evaluate the known bathymetric range of Bathymodiolinae, we performed an extensive literature and database search. Of 6,478 observations of Bathymodiolinae reported in the Ocean Biodiversity Information System (OBIS), 49 occurrences are reported at depths 20 m or less (Figure 3). Ten of these records belong to the new AUF mussel described here and 39 are museum records attributed to the genus *Idas.* The latter specimens were collected off the coast of Papua New Guinea in 2012 but no identification or description has been published. Three additional museum records attributed to the genus *Idas* were collected between 78 and 120 m off the coast of Vanuatu. An additional 55 records attributed to various species of *Adipicola, Gigantidus, Idas,* and *Vulcanidas* appear in museum records at depths from 83–200 m (+/- 30 m). The remaining 6,370 records (98.3%) are at bathyal depths (200–4,000 m). Our literature review finds no peer-reviewed published record of an identified and described bathymodiolin species that grows and reproduces at depths less than 100 m, indicating that the new AUF mussel species, which we have named ***Vadumodiolus teredinicola,* is bathymodiolin mussel demonstrated to grow and reproduce at depths less than 100 m**.

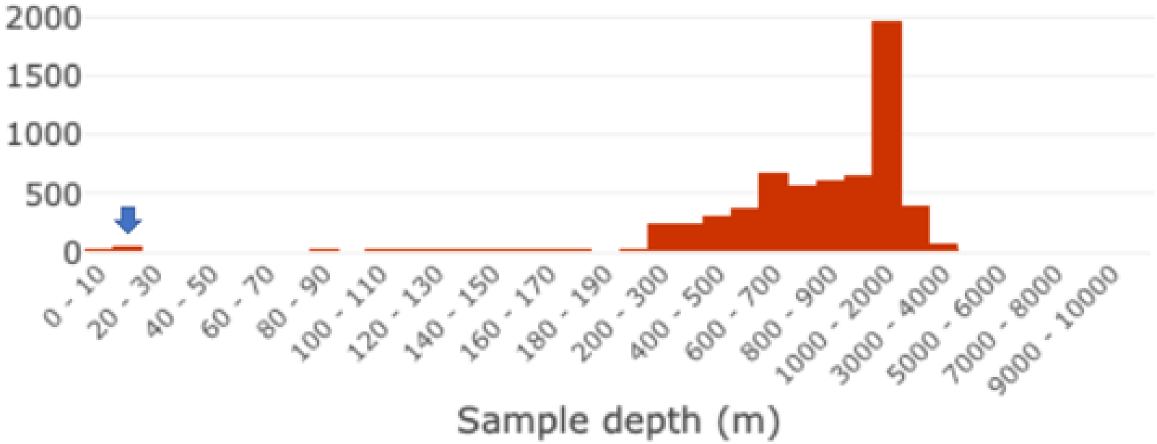

Figure 3. Bathymetric distribution of all occurrences of Bathymodiolinae in the Ocean Biodiversity Information system (OBIS) database as of March 25, 2022. The blue arrow indicates the specimens discovered at the Alabama Undersea Forest (AUF) site.

Morphology of the AUF mussel: The specimens collected at the AUF site display morphological features similar to the genus *Idas* of the subfamily Bathymodiolinae, including elongate modioliform shape, small size (<15 mm), glossy transparent golden-hued periostracum with abundant periostracal hairs, thin valves, mantle edges free except at the siphonal openings, and it is mobile via a long extensible foot (Figure 4).

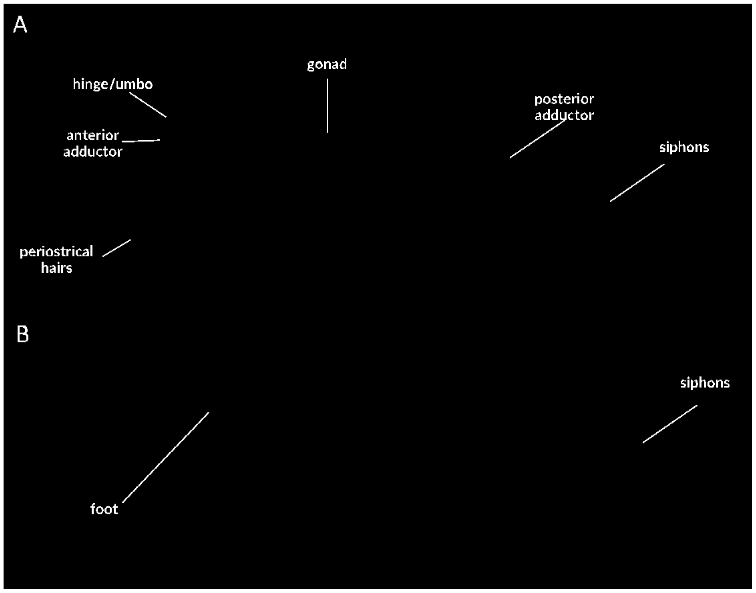

Figure 4. Anatomical features of *Vadumodiolus teredinicola*. A. Left side view. B. Ventral view.

We used microcomputed tomography (microCT) to construct a detailed 3D model of one specimen of *V. teredinicola* and used these results in combination with traditional sectioning and histochemical staining to produce a detailed description of the anatomy and morphology of this new species (Figure 5).

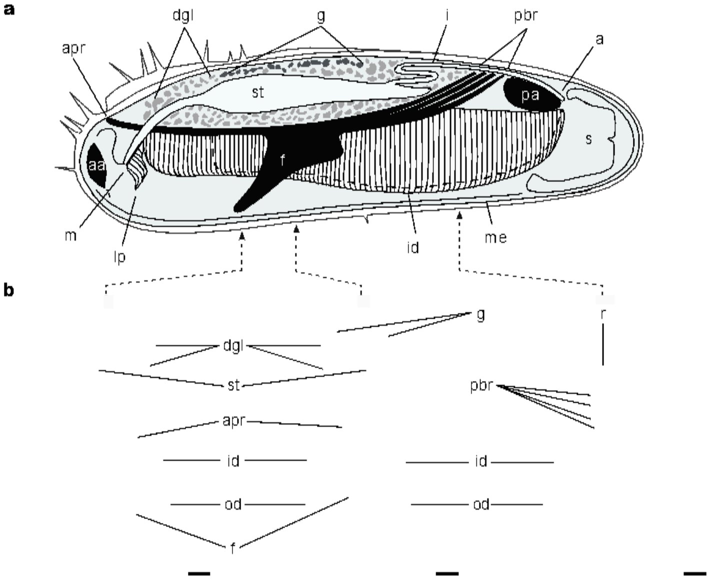

Figure 5. Anatomy of *Vadumodiolus teredinicola*. (a) Diagram depicts a sagittal view along the midline with the left half of the specimen removed. (b) Transverse sections from anterior, midbody and posterior of *Vadumodiolus teredinicola* stained with hematoxylin and eocin. a, anus; aa, anterior adductor; apr, anterior pedal retractor; dgl, digestive glands; f, foot; g, gonad; i, intestine; id, inner demibranch; lp, labial palps; m, mouth; me, mantle edge; od, outer demibranch; pbr, pedal byssal retractors; pa, posterior adductor; r; rectum s, siphon; st, stomach. Black, musculature, b; dark grey, gonadal tissue; medium grey,

digestive glands; light grey, mantle tissue. Dashed line in (a) indicates the location of the right outer demibranch, obscured from view by the right inner demibranch. Scale bars, 0.2 mm.

Metagenomic sequencing and analyses: We performed metagenomic sequencing in collaboration with collaborators at the Department of Symbiosis Research, Max Plank Institute, who are experts on the molecular biology of bathymodiolin mussels. We 1) sequenced and annotated the complete mitochondrial genome of this species, 2) performed metagenomic sequencing on the gill, foot, adductor muscle, mantle, and visceral mass tissues of the putative new species, and 3) performed phylogenetic analyses based on complete mitochondrial genomes and multiple nuclear and mitochondrial genes.

Mitochondrial genome: The complete closed circular mitochondrial genome of the AUF mussel is 20,572 bp in length and encodes the 13 canonical mitochondrial proteins found in animals (including *atp8*), 22 tRNAs, and 2 rRNAs (Figure 6). The *atp8* gene is present in all Bathymodiolinae examined to date but is absent from the closely related mytilid subfamily Modiolinae and most other mytilid taxa, including the reference species used in this study. The order of protein coding genes is uniformly conserved in examined Bathymodiolinae, however some variation is observed in the order of tRNA genes. The mitochondrial gene order observed in the AUF mussel is identical to the consensus gene order for Bathymodiolinae, except for the inversion of tRNAs L1 and K, and differs from those of other mytilid subfamilies (Table 2).

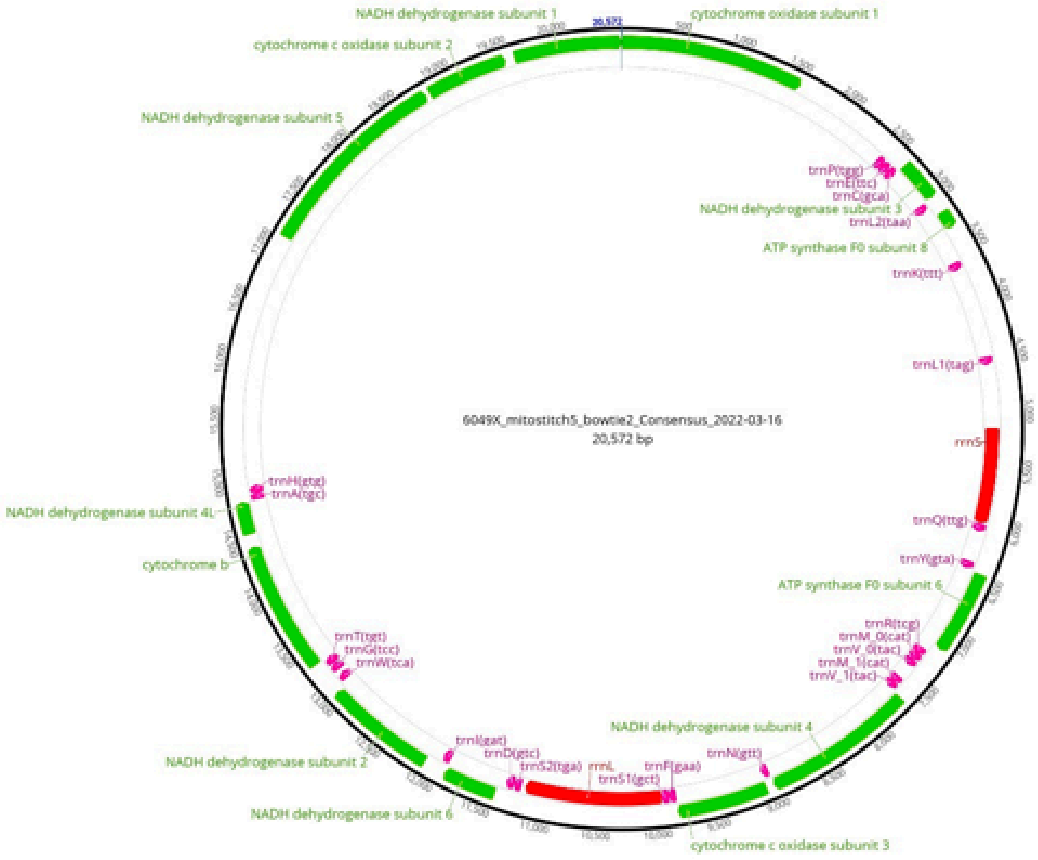

Figure 6. A map demonstrating the order and boundaries of genes within the complete closed circular mitochondrial genome of the Alabama Undersea Forest (AUF) mussel.

Subfamily Bathymodiolinae																Subfamily Modiolinae					Outgroups		
	Idas		Nipponiomodiolus		Bathymodiolus *sensu stricto*								Gigantidas			Modiolus							
Vadumodiolus teredinicola 6049X	*Idas modioliformis*	*Bathymodiolinae consensus*	*B. aduloides*	*B. manusensis*	*B. azoricus*	*B. brooksi*	*B. thermo*	*B. marisindicus*	*B. septemdierm*	*B. sp. 5 South*	*B. securiformis*	*B. childressi*	*G. haimaensis*	*G. platifrons*	*G. japonicus*	*M. kurilensis*	*M. comptus*	*M. modiolus*	*M. nipponicus*	*M. philippi*	*Vignadula atrata*	*Limnoperna fortunei*	*Lithophaga curta*
cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1	cox1
		-							Y												L1	cox3	L
		-						L	L	K											nad2	S	cox3
P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P					P	K	rrnL	F
E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E					E	atp6	S	S
C	cox1	C	C	C	C	C	C	C	C	C	C	C	C	C	C					L	G	D	T
nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	nad3	T	E	M
L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	L2	C	C	C	C	C	cob	nad3	L
	I	-				K		?		E	W	S		S	W	L1	P	L1	P	L1	L2	L	cob
atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	atp8	P	E	P	E	R	P	Q	rrnS
K	L1	L1	L1	L1	L1	L1	L1	L1	L1	L1	L1	L1	L1	L1	L1	E	L	E	L	S1	rrnS	atp6	atp6
L1	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	L	R	L	R		Q	R	R
		-														R	M	R	M		I	V	V
		-														S1	Q	S1	Q		R	nad4	nad4
		-														M	L	M	L		rrnL	T	K
		-														K	S1	K	S1		S	nad2	D
		-															K		K		nad1	nad1	P
rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	rrnS	A	N	E
Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q		M	H	M	C
Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y					Q	D	M	nad3
atp6	atp8	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	atp6	N	F	Q
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	Y	Y	Y	Y		nad5	W	N
M_0	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	V	V	V	V	V	cox2	K	rrnL
V_0	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V						Y	Y	S
M_1		-																			E	nad6	G
V_1		-																			V	rrnS	Y
nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	nad4	C	K	nad6
N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	W	cob	I
cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	cox3	nad6	nad4L	nad2
F	F	F	F		F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	nad4	A	W
S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1						cox3	H	nad4L
rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	rrnL	F	nad5	A
S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	S2	G	H
D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	nad3	P	nad5
nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	nad6	M	C	cox2
I		I	I		I	I		I	I	I	I	I	I	I	I	I	I	I	I	I	nad4l	cox2	nad1
nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2	nad2		L	
W	W	W	W	W	W	W	I	W	W	W	W	W	W	W	W	W	W	W	W	W		I	
G	G	G	G	G	G	G	W	G	G	G	G	G	G	T	G	G	G	G	G	G			
T	T	T	T	T	T	T	G	T	T	T	G	T	T	G	T	T	T	T	T	T			
		-					T				T	G	G	T									
		-										T	T	G									
		-										G		T									
		-										T		G									
		-												T									
cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob	cob			
nad4l	nad4l	nad4l	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4L	nad4l	nad4L	nad4L	nad4L			
A	A	A	A		A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
H	H	H	H		H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H			
		-	K		L	L	L			H													
		-	Y							L													
nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5	nad5			
cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2	cox2			
nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1	nad1			
																				K			
																				Y			

Table 2. Aligned gene orders for all available bathymodiolin and modiolin mitochondrial genomes and closely related reference taxa. Note that the gene order of the Alabama Undersea Forest (AUF) mussel (AUF-6049X) is nearly identical to the consensus gene order for Bathymodiolinae but differs from that of other mytilid subfamilies. Outgroup gene orders are too different for alignment. Note that the published mitogenome of *Bathymodiolus manusensis* is missing several tRNA genes.

Molecular phylogenetic analyses: Although few complete mitochondrial genomes are available for Mytilidae, phylogenetic analyses based on 12 concatenated mitochondrial protein coding genes from 14 bathymodiolin mussel species, 4 modiolin species, and 2 outgroup taxa show strong support (posterior probability = 1.0) for inclusion of the new AUF mussel as the basal member of a clade containing all Bathymodiolinae, consistent with assignment of the AUF mussel to this subfamily (Figure 7).

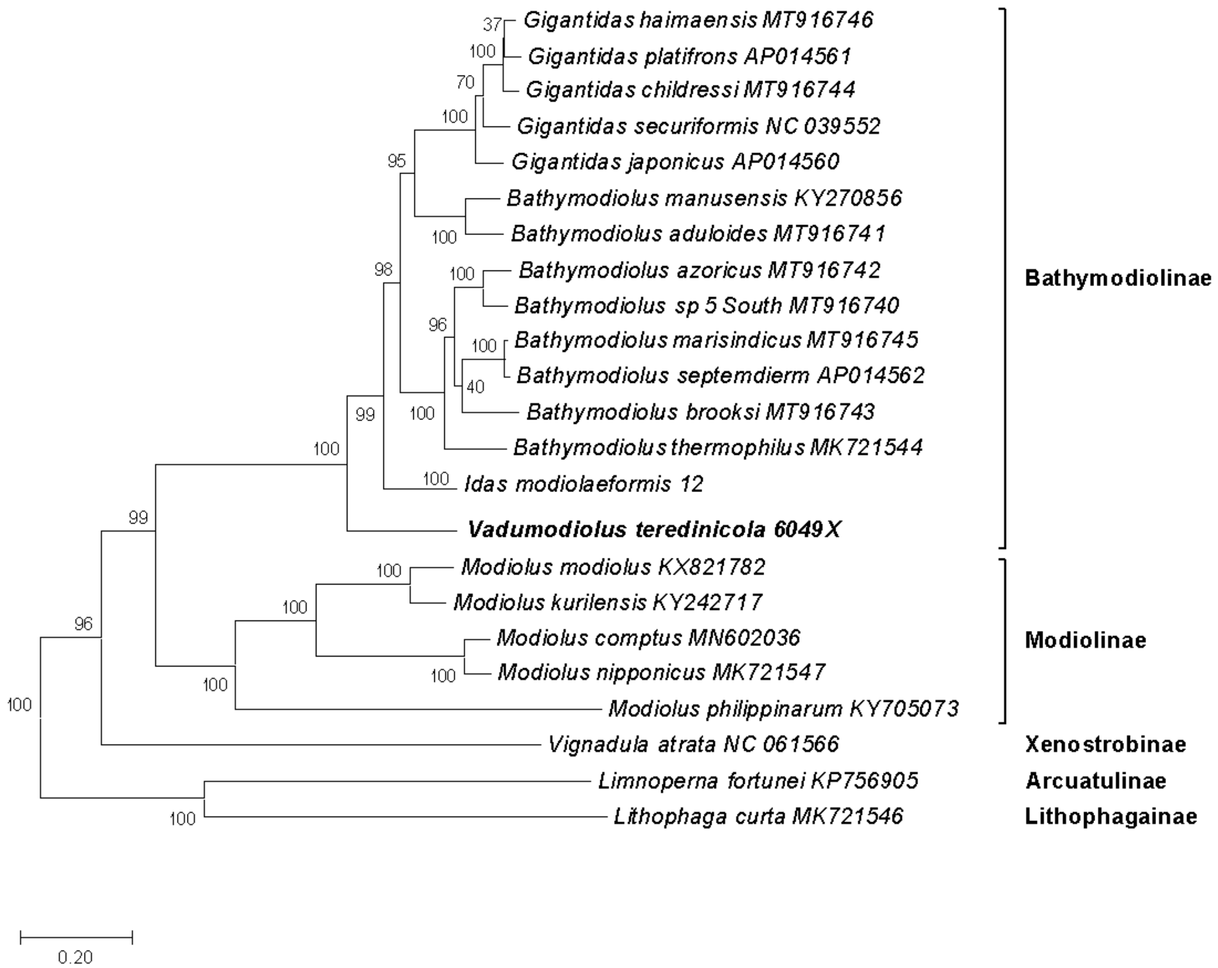

Figure 7. Phylogenetic tree inferred using by Maximum Likelihood (RAxML) based on 12 concatenated mitochondrial genes. The presented subtree is excerpted from a larger analysis including 49 species of mytilid reference taxa and 6 species of Ostreidae as outgroup taxa.

To confirm the assignment of the new AUF mussel to the subfamily Bathymodiolinae and to determine if it can be assigned to a known genus within this subfamily, we performed phylogenetic analyses using the nuclear encoded large and small subunit ribosomal RNA genes (18S and 28S), and the mitochondrial cytochrome oxidase1 (cox1) and mitochondrial 16S rRNA genes. These genes were selected because sequence is available for representatives of all accepted bathymodiolin genera. The tree shown in Figure 8 provides strong support (posterior probabilities > 0.99) for the monophyly of the subfamily Bathymodiolinae, for the accepted bathymodiolin genera *Idas*, *Terua*, *Gigantidas*, *Bathymodiolus*, *Vulcanidas*, *Nypamodiolus*, and *Benthomodiolus*, and the proposed genera *Lignomodiolus* and *Nipponiomodiolus*.

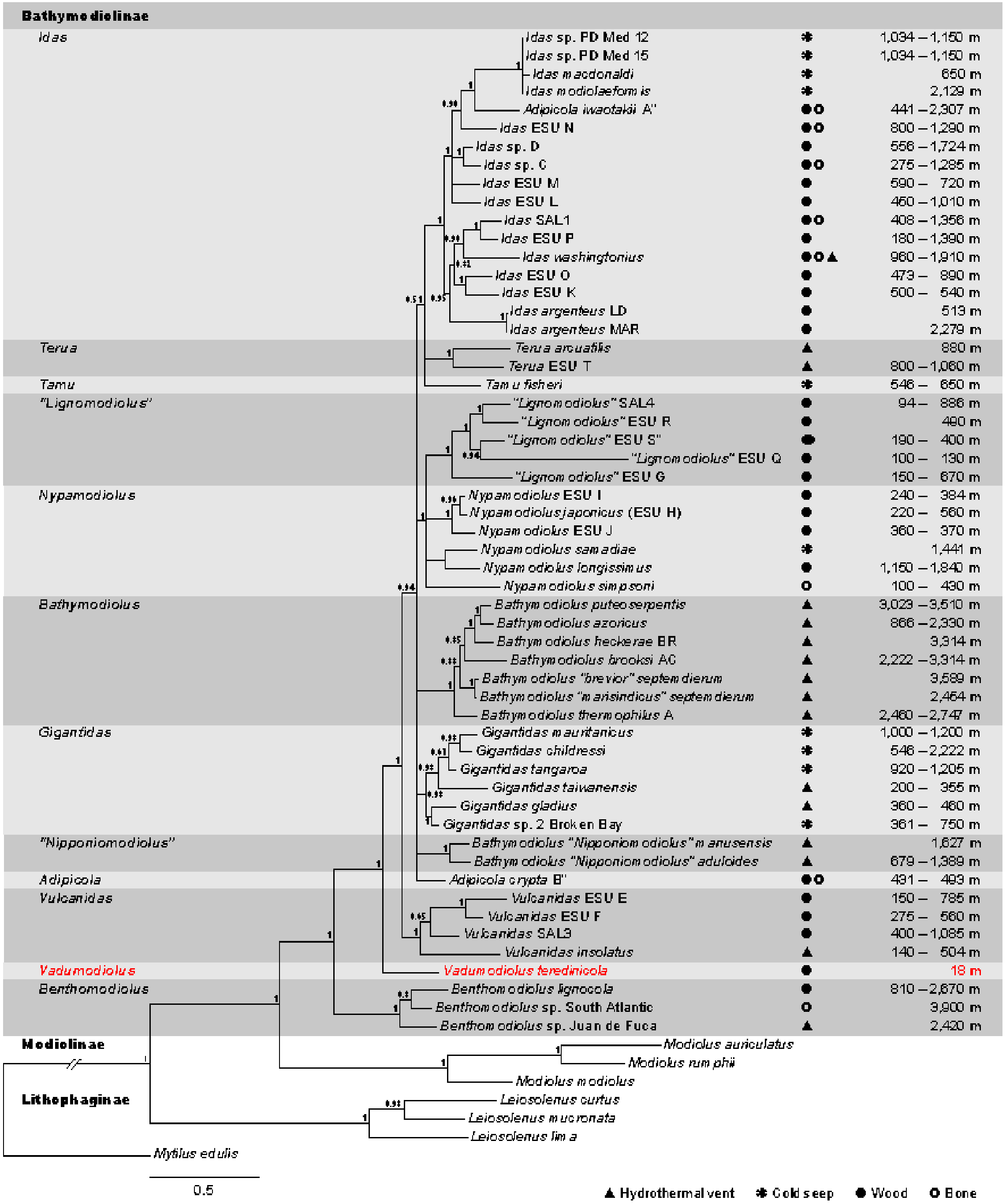

Figure 8. Phylogram depicting inferred phylogenetic relationships between *Vadumodiolus teredinicola* and representative Bathymodiolinae, showing observed habitat type and depth of occurrence. Bayesian inference (MrBayes) was used to create a phylogram based on an alignment of partial sequences of four concatenated gene loci, nuclear 18S rRNA (1,546 bp) and 28S rRNA (700 bp) and mitochondrial cox1(507 bp) and 16S rRNA (342 bp) genes. Taxon selection includes *V. teredinicola*, 54 bathymodiolin species from hydrothermal vents, cold seeps, and organic deposits, and 7 non-bathymodiolin reference taxa. Posterior probabilities greater than 0.90 are displayed at the associated

nodes. Scale bar = 0.1 substitutions per 100 base pairs. Depth ranges and habitat types for *V. teredinicola* and *Idas* sp. PD Med are from this study, others are from Thubaut et al. 2013.

Localization and identification of bacterial symbionts: To detect and localize symbiotic bacteria within the tissues of the AUF mussel, whole specimens were fixed, embedded in paraffin, sectioned, and subjected to fluorescent in situ hybridization using a bacteria-specific fluorescence-labeled oligonucleotide probe targeting the bacterial 16S rRNA (Figure 9). Sections were imaged using laser confocal microscopy.

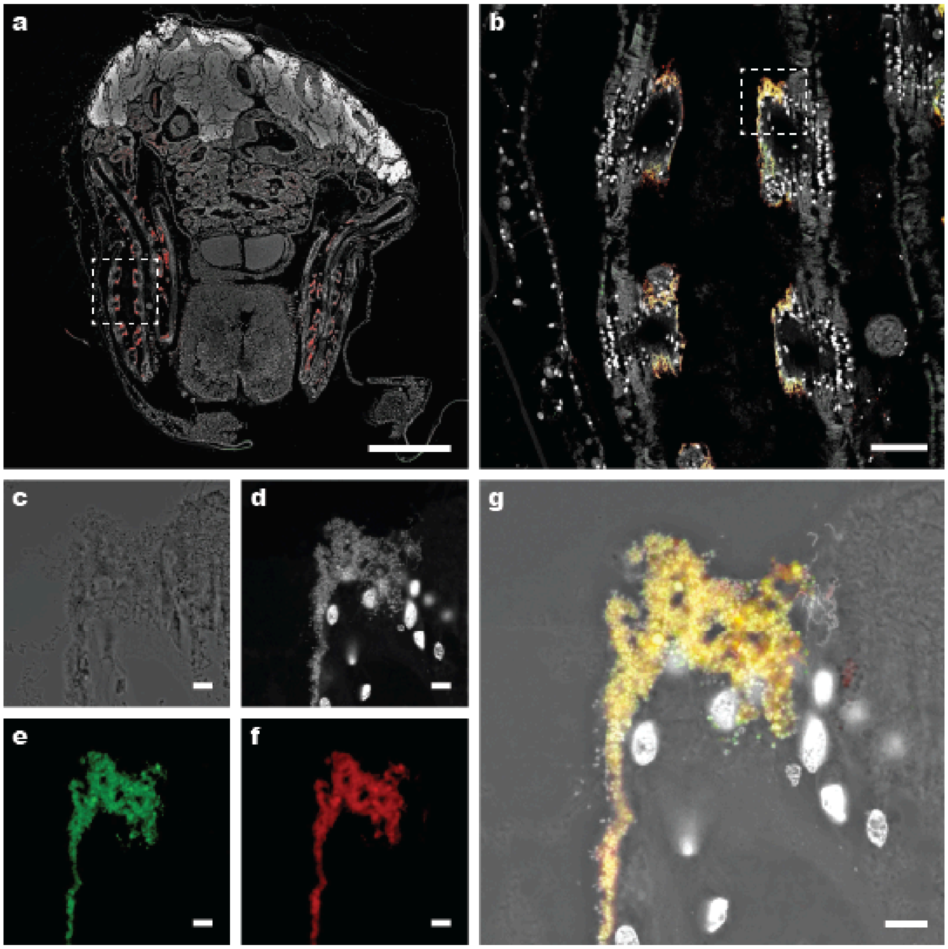

Figure 9. Confocal laser scanning microscopy overview of the distribution of symbionts in *Vadumodiolus teredinicola*. Dual-probe fluorescence *in situ* hybridization (FISH) using both a bacteria-domain-specific 16S rRNA probe Cy5-EUB338 (red) and a probe specifically targeting the 16S rRNAs of known bathymodiolin thioautotrophic symbionts, Cy3-Bthio-193 (green). DNA is stained with NucBlue Nuclear Stain (white). (a) transverse section through the gill and visceral mass of a mature male specimen of *V. teredinicola*, (b) detail of the boxed region in (a), (c–g) detail of the boxed region in (b). (c) differential interference contrast image with visible light illumination (d) NucBlue Nuclear Stain fluorescence channel only (e) red fluorescence channel only (f) green fluorescence channel only, and (g) overlay of DIC, NucBlue Nuclear Stain (white), and probes targeting the symbionts (green, Bthio-193; red, EUB338). Note that both EUB338 and Bthio-93 probes (combinatorial colors yellow to orange) hybridized to most bacterial cells within the abfrontal region of the ascending and descending filaments, indicating phylogenetic identity and location similar to the symbionts of other bathymodiolin mussels.

Phylogenetic identification of the symbionts of Vadumodiolus teredinicola: We examined the gill tissues of three specimens of *Vadumodiolus teredinicola* via Illumina shotgun metagenome sequencing and recovered the 16S rRNA gene of a putative symbiont closely related to the symbionts of other bathymodiolin mussels. We were able to identify and assemble a metagenome assembled genome (MAG) from the gill metagenomes of *V. teredinicola* using a variety of metagenome assembly and binning tools. The putative symbiont MAG is 1.06Mbp in size and is 87% complete as estimated by the CheckM tool. Using the Genome Taxonomy Database tool kit (GTDB-tk) 120 conserved single copy protein coding sequences of this putative symbiont MAG were aligned with those of other known thioautotrophic symbionts and of related free-living thioautotrophs. RAxML was used to generate a maximum likelihood tree based on this alignment demonstrating that the *V. teredinicola* putative symbiont is very closely related to the symbionts of *Bathymodiolus thermophilus*, *B. septemdierum* and *B. azoricus* (Figure 10).

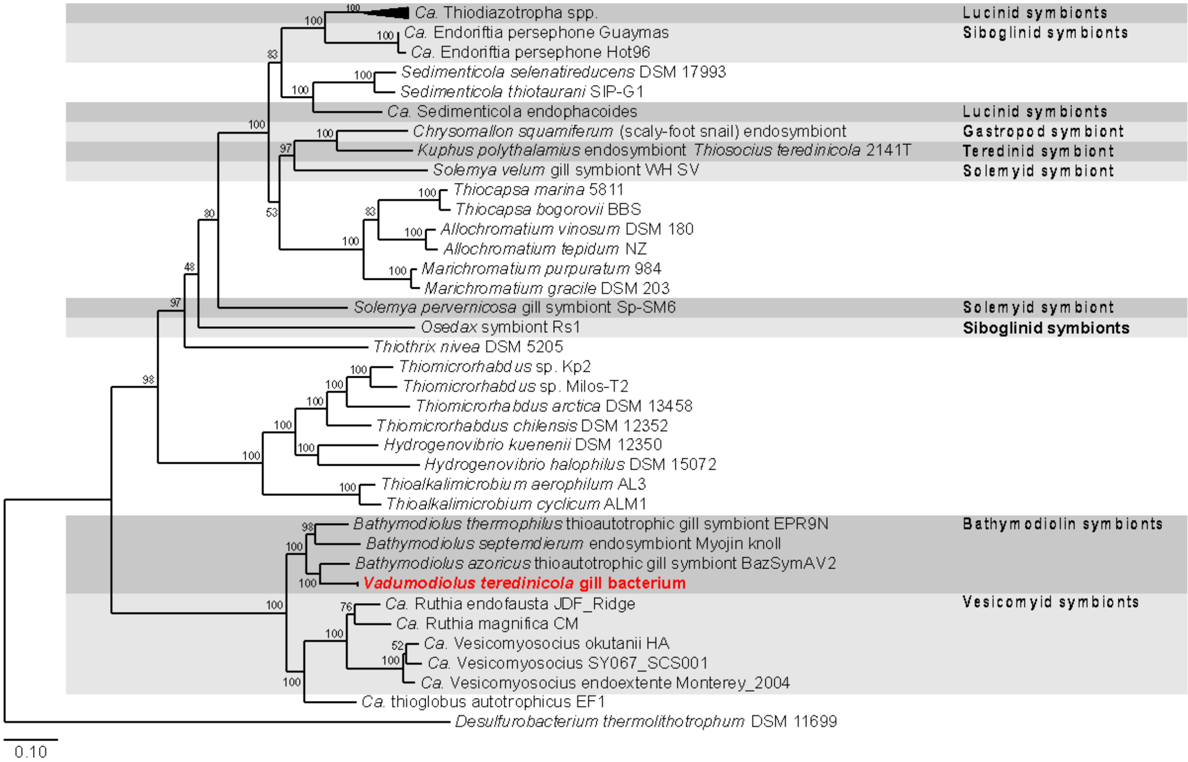

Figure 10. Phylogram depicting inferred phylogenetic relationships among the *Vadumodiolus teredinicola* gill bacterium and representative endosymbiotic and free-living thioautotrophic bacteria. Maximum likelihood analysis (RAxML) was used create a phylogram based on an alignment of sequences of 120 conserved bacterial genes identified in the partial metagenome assembled genome (MAG) of the *V. teredinicola* gill bacterium and representative thioautotrophic symbionts and free-living bacteria using the Genome Taxonomy Data Base tool kit (GTDB-tk). Bootstrap proportions of 1,000 replicates are displayed at associated nodes. The scale bar represents the substitution rate per site.

Manuscript: A manuscript describing the new species is in review and has been recommended for publication in Deep Sea Research Part 1 pending minor revision.

4.a.2. Contrasting evolutionary trajectory of Teredinid and Xylophagainid bivalves

We sequenced and annotated the first complete shipworm mitochondrial genomes. Forty-two mitochondrial genomes of woodborers from the AUF and other locations were sequenced, closed, annotated, and polished. A phylogeny of wood-feeding bivalves was generated based on complete mitochondrial genomes and dramatic differences in mitochondrial genome evolution were demonstrated between shipworms and their deep-sea sister taxon, Xylophagaidae (Figure 11). A manuscript has been published in Genome Biology and Evolution and 42 mitogenome sequences have been submitted to GenBank.

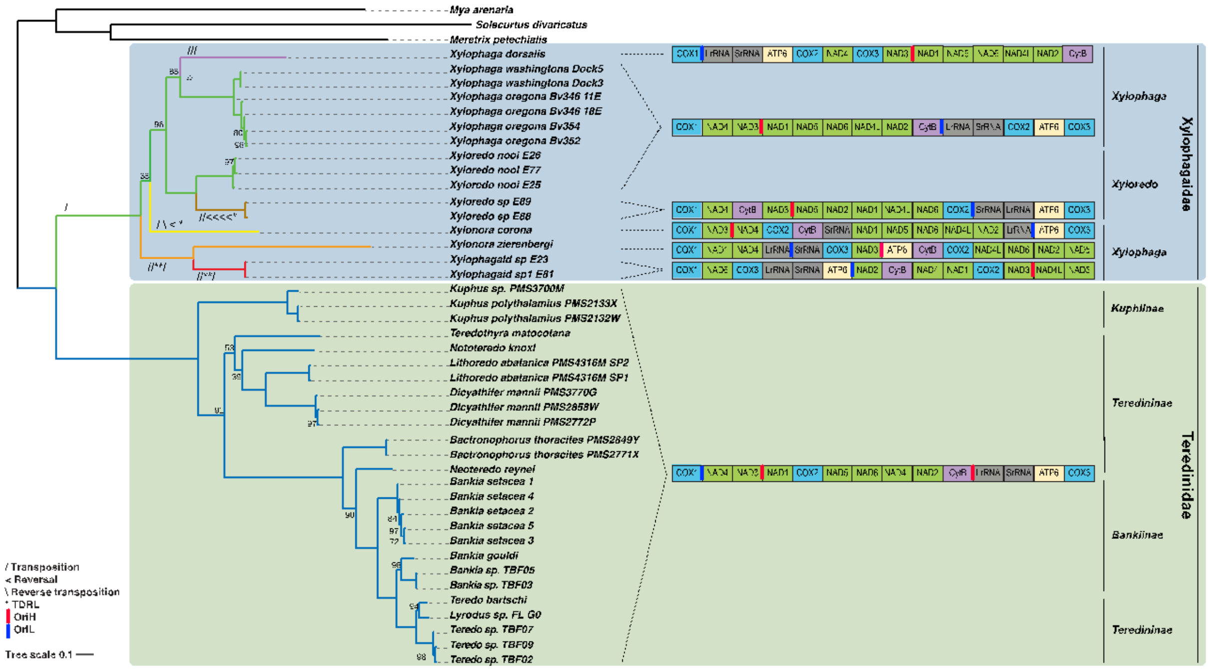

Figure 11. Phylogenetic relationships and variation in mitochondrial gene arrangement among species of Xylophagaidae and Teredinidae. Phylogenetic tree inferred by maximum likelihood (single site-homogeneous model, unpartitioned, IQ-Tree 1.6.7) based on 4,135 unambiguously aligned amino acid positions selected using Gblocks from an alignment of 12 concatenated mitochondrial protein coding genes. Bootstrap proportions, less than 100 are indicated at the nodes; where no numerical value is indicated, the bootstrap proportion = 100. Protein coding gene arrangements associated with each species are depicted on the right. Putative origins of replication are indicated by thick vertical bars. A proposed scenario for the evolutionary history of gene rearrangement, determined by analysis of common intervals using CRex, is superimposed on the phylogenetic tree with branches color-coded according to extant and hypothetical ancestral gene orders. Symbols indicate rearrangements required to transform each hypothetical ancestral gene order to the order observed in the extant species. / = transposition, \ = reverse transposition, < = reversal, * = tandem duplication with random loss (TDRL).

4.a.3. Symbionts distribution in *Nototeredo knoxi*

Another extremely surprising finding of this project was the unexpected location of the bacterial symbiont community in *Nototeredo knoxi*. In shipworms, the gill elongates as the animals develop from clam-like larvae to worm-like adults. As the gill grows, it separates into a tiny anterior gill, containing just a few filaments, and a large posterior gill containing hundreds to thousands of filaments. In all shipworm species examined to date, the large posterior gill holds the vast majority of symbiont cells. However, in *N. knoxi* we found that the posterior gill has few symbionts, while a large and well-developed symbiont community is found in the anterior gill (Figure 12).

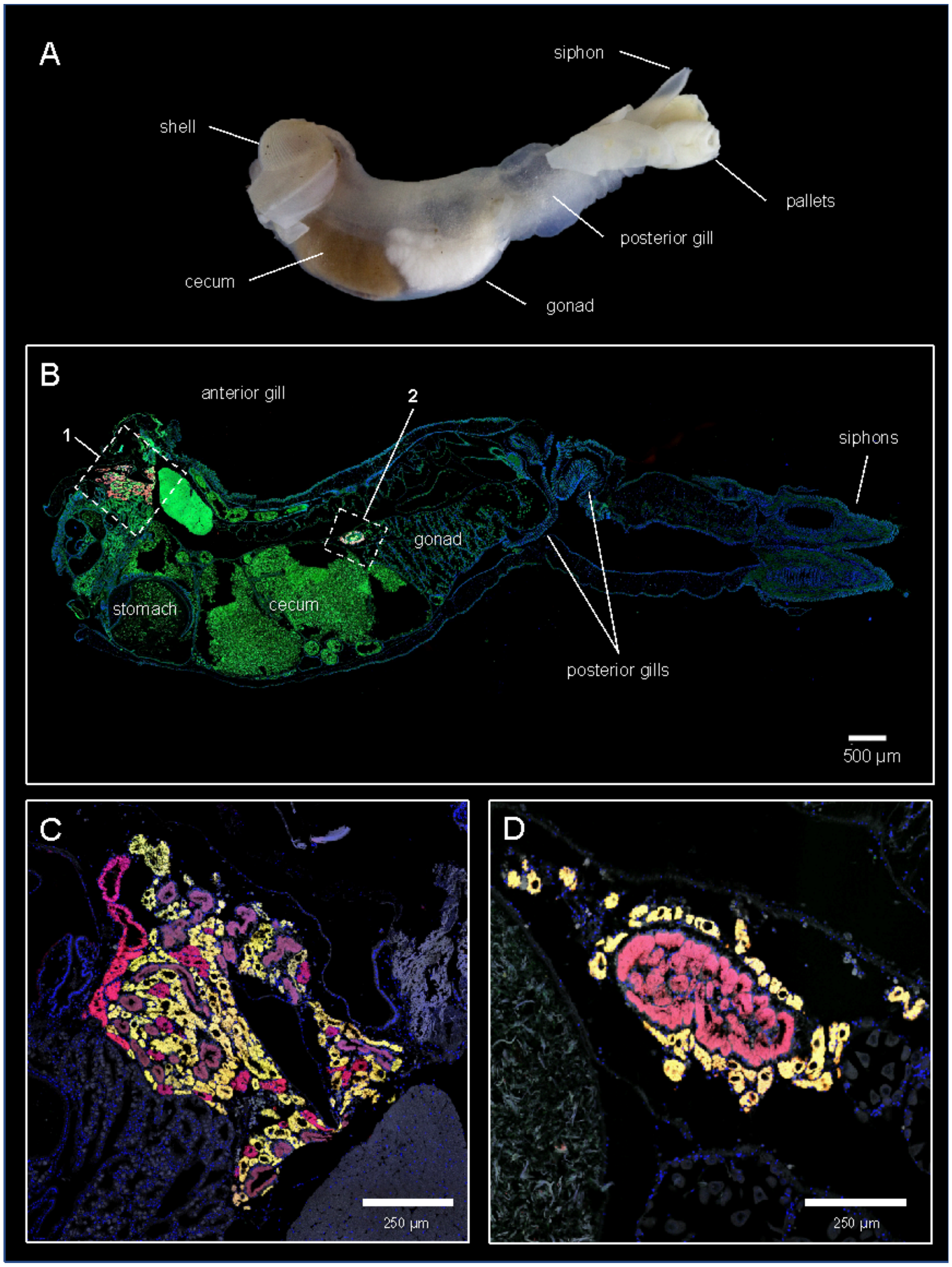

Figure 12. Localization of bacterial symbionts in a specimen of *Nototeredo knoxi* using fluorescent in situ hybridization (FISH). A. Photograph of a specimen of *N. knoxi* removed from wood, B. Section through a whole specimen of *N. knoxi,* C. Enlargement of the region of the anterior gill of *N. knoxi* indicated by dashed box 1 in B, D. Enlargement of a region near the anterior tip of the gonad indicated by the dashed box in B. Tissue sections are stained with Cy5-Eub338 probe (red, bacteria-specific), Cy3-Nk-cellulo-840 probe (green, specific to shipworm symbiont clade VIII), and NucBlue nuclear DNA stain (blue, labels host nuclei). Note that little or no bacteria- or symbiont-specific staining is observed in the posterior gill using bacteria-targeting probes, but significant staining is seen in the anterior gill and at the anterior tip of the gonad. Note also that due to probe colocalization, some bacteriocytes appear red (hybridizing only with the bacteria-specific Eub-338 probe) while others appear yellow, indicating hybridization with both the bacteria-specific and clade VIII symbiont-specific probes. These color

differences indicate that there are at least two distinctly different symbionts present within the symbiont community of *N. knoxi*, one of which belongs to symbiont clade VIII.

Metagenome sequencing demonstrated that the symbiont community of the anterior gill is among the most phylogenetically complex examined to date. A manuscript is in preparation.

4.a.4. Ecology of the AUF site

In December 2019, we conducted the first biodiversity survey of the Alabama ancient undersea forest site. Our results indicate that this site supports a diverse community of organisms specifically adapted for life in association with deposited wood on the sea floor (Figures 13 and 14). This includes both wood specialists and opportunistic organisms taking advantage of niches created by the former (Figure 15).

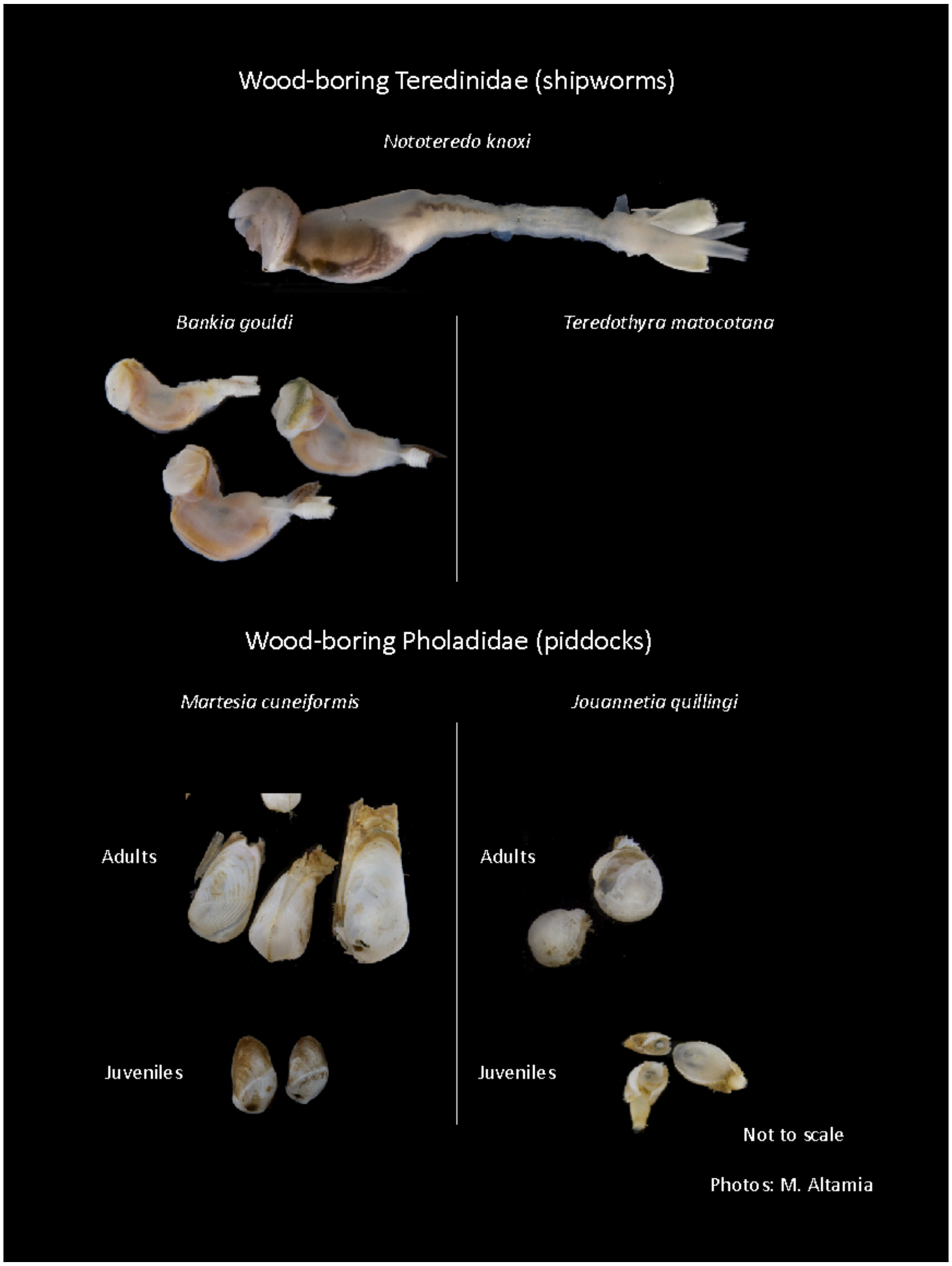

Figure 13. Wood-boring Teredinidae (shipworms) and Pholadidae (piddocks) collected at the Alabama Undersea Forest (AUF).

Figure 14. Some additional wood-associated invertebrates collected at the Alabama Undersea Forest (AUF) site.

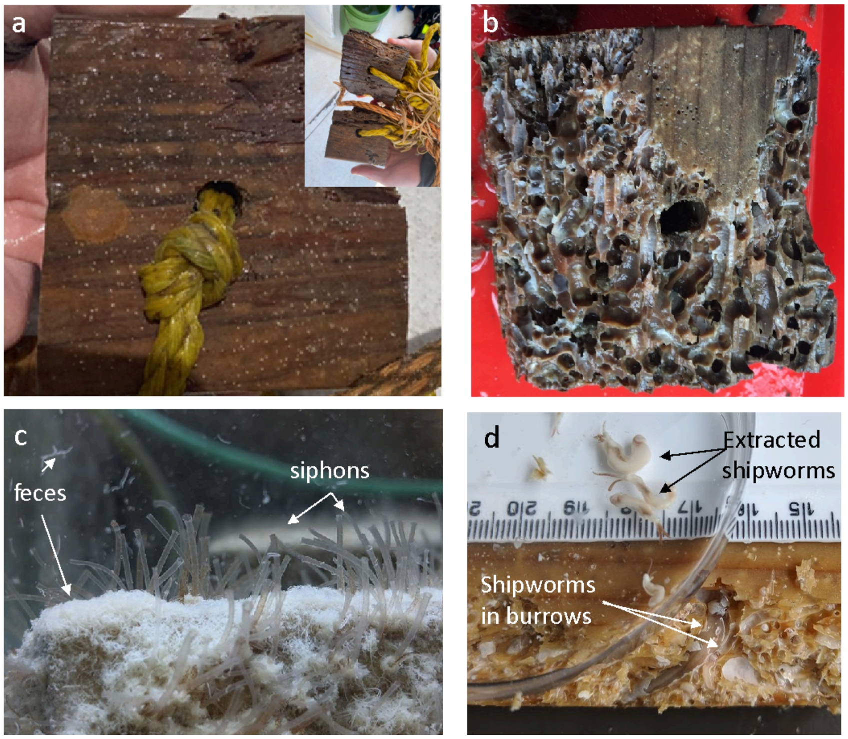

Figure 15. Settlement panels recovered after six months (a, b) and two months (c, d) of exposure at the Alabama site. Note that after six months, the surface of the panel appears intact (a) but removing the thin venire of surface wood reveals a network of empty shipworm burrows (b). After two months, a high density of shipworm siphons and massive fecal production can be seen at the wood surface (c) and a dense community of living shipworms is found within the test panel (d).

Observations of wooden collection panels deployed at the AUF site for periods ranging from two to 12 months provide valuable insight into wood colonization processes. The first of three sets of panels deployed at the site were examined after six months of exposure. Although the external surfaces of the two retrieved settlement panels were intact, extensive damage to the panels was observed internally (Figure 15a, b). Within panels, shipworms (Bivalvia: Teredinidae) consumed nearly all of the internal wood and subsequently died, leaving behind a network of empty burrows.

However, small living shipworms were found in the interstitial wood between the empty burrows, indicating a second round of infestation following the death of the first invaders. It is presumed that the third panel, which was no longer attached to its intact tether line, had completely disintegrated, and so could not be recovered.

While most burrows in the six-month recovered panels contained no living shipworms, hundreds of shipworm pallets were recovered from the empty burrows. Pallets are calcareous structures used by shipworms to seal the burrow entrance. Pallet morphology is used by taxonomists as primary evidence for species diagnosis. Examination of the living specimens and pallet remains in these panels demonstrated that the observed shipworm infestation was due to a single species, *Bankia gouldi*. Other shipworm species previously observed in naturally occurring wood at the site (*Nototeredo knoxi* and *Teredothyra matocotana*) were not observed in these settlement panels.

In addition to shipworms, many minute bivalves, crabs, and annelid worms were recovered from the empty burrows of the six-month panels. No wood-boring piddocks (Bivalvia: Pholadidae) were found, although these were previously observed to be abundant in naturally occurring wood at the site.

To examine earlier stages of wood colonization, a second set of settlement panels were deployed at the site and recovered after two months of exposure (Figure 15c, d). The two-month panels were infested and damaged to an extent similar to that observed in the six-month panels, with most internal volume excavated. However, unlike the six-month panels, most burrows in two-month panels were still occupied by living shipworms. As in the six-month panels, all shipworms collected from two-month panels were a single species, *B. gouldi*, and several annelid worms and small crabs, but no piddocks, were found. Natural wood collected on the same date and at the same location as the two-month settlement panels contained no living shipworms, although empty shipworm burrows were observed. Instead, many small piddocks, several crabs, annelid worms, and encrusting bryozoan colonies were collected from this material.

These observations of settlement panels yield important insight into the dynamics of wood colonization at the site, indicating that: 1) The shipworm *B. gouldi* is the main pioneer wood-invading species at this site. 2) As new wood is exposed, shipworms may consume its internal volume within months, rendering it unsuitable to support further shipworm growth. 3) Upon death, shipworms are consumed by, and their burrows are utilized by, a diverse variety of worms, crustaceans, and bivalves.

While we cannot directly extrapolate from the test panel data to natural wood at the site, these data are consistent with our observations of natural wood. Together, the two types of data suggest that newly exposed wood at the site is very rapidly infested, and its internal contents are consumed by shipworms, likely on the time scale of months. However, shipworm wood degradation is incomplete and self-limiting, as shipworms are neither able to efficiently degrade the surface of the wood nor the thin layers of remaining interstitial wood surrounding the empty burrows left by prior infestations.

As a result, the remaining intact surfaces of the wood and the hollow burrow-riddled interiors may persist long after initial infestation, providing substrate for anemones, sponges, and encrusting organisms such as bryozoans, as well as shelter for organisms that may otherwise burrow in sediments or nestle within the crevices of hard substrates. Excellent examples are ghost shrimp (Decapoda: Thalassinidea) and snapping shrimp (Decapoda: Alpheidae), which are typically found burrowing in soft sediment, but that are abundant in the abandoned burrows of shipworms at the site. Polychaete worms and crabs are also abundant in shipworm burrows, often observed feeding on shipworm remains. Interestingly, piddocks (Bivalvia: Pholadidae) were abundant in older softer natural wood, but were not found in settlement panels, suggesting that they may not be primary wood invaders, but only appear later in the “succession” of wood colonization.

Interestingly, wood-boring isopod crustaceans (gribbles), which often compete with bivalves for wood substrates, were not abundant at the site. Gribbles are capable of rapidly degrading surface and interstitial wood and, in combination with shipworms, can completely disintegrate wood substrates. The relative scarcity of gribbles may help explain the long-term persistence of skeletal wood remains and elevated biodiversity at the site.

4.a.5. Diversity of shipworm symbionts and symbiont-encoded molecules

16S rRNA diversity: These wood-associated organisms have also yielded a diverse selection of bacterial isolates that are phylogenetically related to previously identified shipworm symbionts. Figure 16 shows a phylogenetic analysis of shipworm symbionts including 95 AUF isolates plus 116 previously isolated shipworm symbionts and 31 reference species.

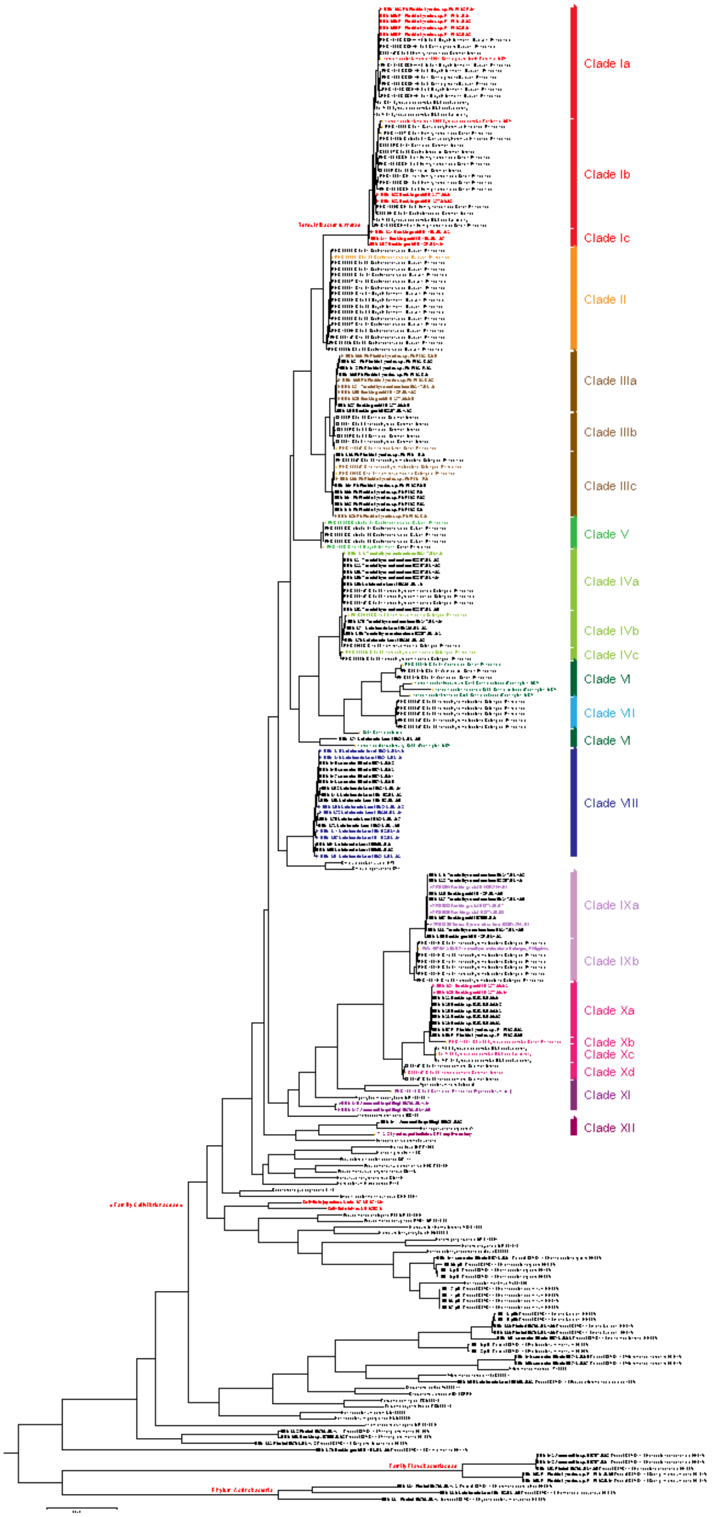

Figure 16. Phylogenetic relationships among 95 unique symbiotic bacterial isolates from the Alabama Undersea Forest shipworms and piddocks (bold), 116 shipworm symbiont isolates from other locations, and 31 reference taxa. The tree was inferred using a 1,219 bp MAAFT alignment of 16S rRNA using FastTree v2.1.11, GTR model, rate categories = 20, Gamma20. Twelve putative symbiont clades are identified by colored bars.

These analyses indicate that:

- These cellulolytic strains isolated from wood-boring bivalves are distributed among at least twelve distinct bacterial clades.
- Some of these symbiont species and clades occur in diverse host taxa and locations, e.g., the Gulf of Mexico, the Philippines, and the Solomon Islands, whereas other symbiont clades appear to have narrower host and geographic ranges.
- Several related cellulolytic isolates have been cultivated from the pholadid bivalve *Jouannetia quillingi*, to our knowledge the first isolated from a pholad. Conversely, none have been isolated from the related pholad, *Martesia cuneiformis,* found at the same site.
- Two of these isolates from *J. quillingi,* SR01846 and SR01847, contain some of the largest genomes (8.1 and 8.3 Gb, respectively) ever detected in the family *Cellvibrionacea* and are rich in both secondary metabolite pathways and lignocellulolytic enzymes.
- One symbiont species, represented by strains SR01945–SR01948, was repeatedly isolated both from seawater and from the gills of the shipworm *N. knoxi.* This is potentially the first isolation of a shipworm symbiont from ambient seawater and may support the hypothesis that shipworm symbionts are acquired from the environment.

Genome and Metagenome Sequence Analyses: In performance period 3, we began analyses of complete bacterial symbiont genomes and metagenomes. To date we have obtained genome sequences from 24 unique symbiont isolates and metagenome sequences from the gills of three shipworm species found at the AUF site.

Isolate genomes: To evaluate species and genus delimitations among the shipworm symbiont isolates, average nucleotide identity (ANI) and average amino acid identity (AAI) were calculated for all pairwise comparisons and matrices were constructed including the 24 AUF isolate genomes and 23 shipworm isolate genomes collected by our group in other locations. We used the commonly cited thresholds of ≤ 95% ANI and ≤ 95% AAI for species delimitation and ≤ 60% AAI for genus delimitation. These results support the division of these isolates into at least four genera comprising 25 distinct species (Figures 17 and 18). These results show strong agreement with the phylogeny inferred based on 16S rRNA sequences as visually demonstrated by the agreement between color-coded clade designations in Figures 17 and 18.

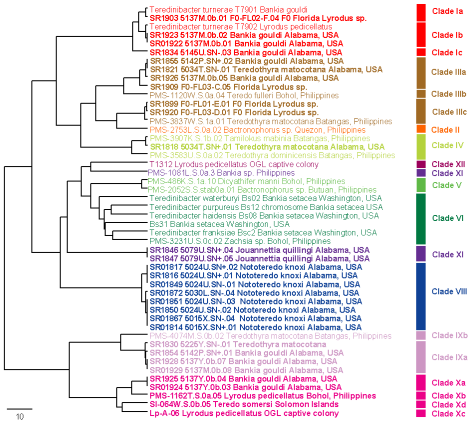

Figure 17. Average amino acid identity (AAI) tree for symbiont isolate genomes. AAI calculated for all pairwise comparisons among 24 shipworm symbiont genomes from the Alabama Undersea Forest (AUF) site (bold) and 23 shipworm symbiont isolates from other locations (Ward clustering). This analysis demonstrates strong congruence between clades inferred by 16S rRNA phylogeny and genome similarity as estimated by AAI.

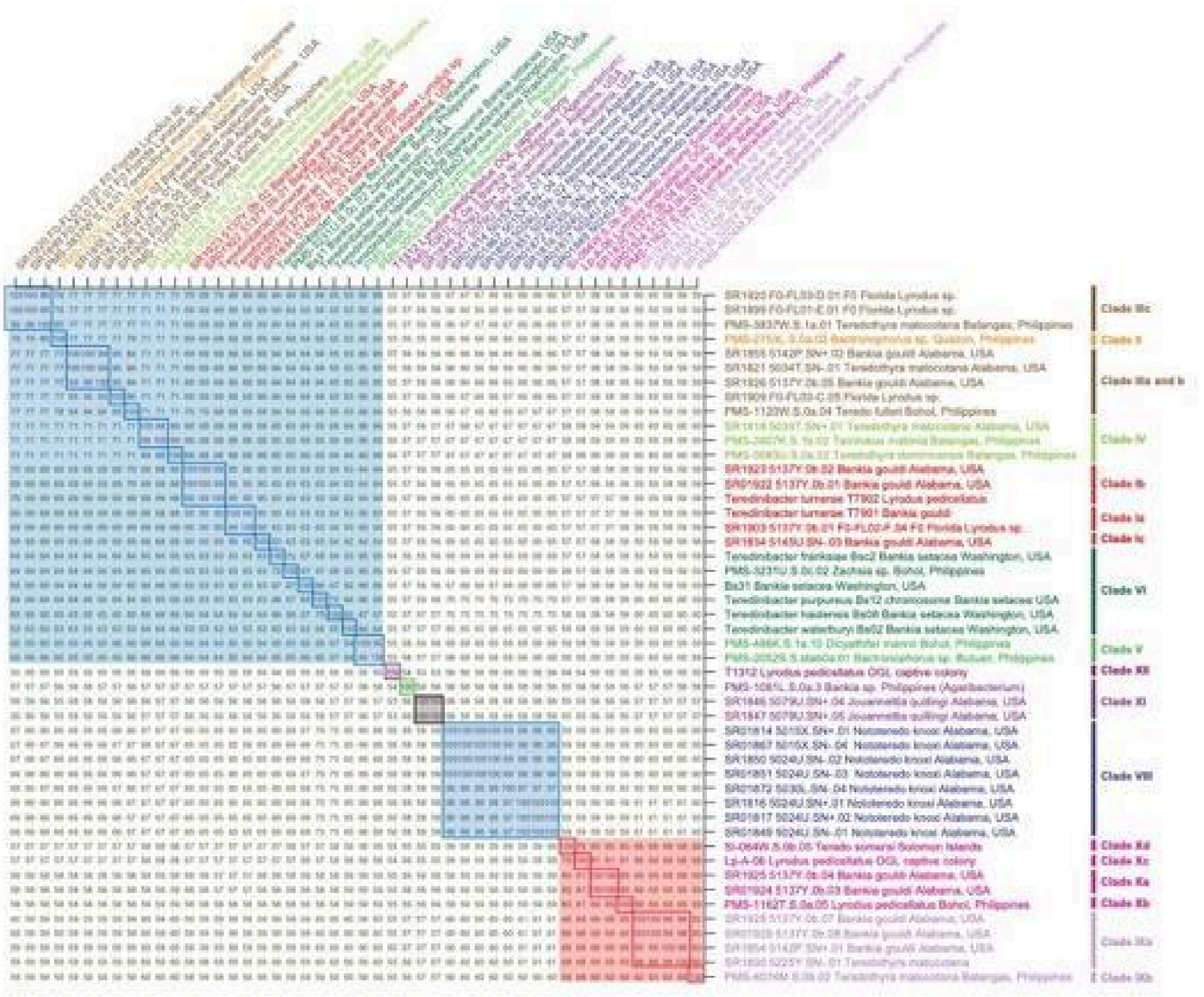

Figure 18. Average amino acid identity (AAI) matrix for symbiont isolate genomes. AAI calculated for all pairwise comparisons among 24 shipworm symbiont genomes from the Alabama Undersea Forest (AUF) site (bold) and 23 shipworm symbiont isolates from other locations (Ward clustering). Strain names are color coded according to the symbiont clades identified by 16S rRNA phylogeny. Putative species are indicated by boxes with solid lines at AAI threshold ≥ 0.95. Shading indicates putative genera at AAI threshold ≥ 0.60; blue, Teredinibacter; red, green, grey, purple represent putative new genera.

Comparison of Metagenome Assembled Genomes (MAGs) and Symbiont Isolate Genomes: An important goal of this project was to compare the metagenomes of wood associated organisms to the genomes of putative symbionts isolated in pure culture. The main purpose of this effort was to determine the phylogenetic composition and genome content of the symbiont communities of shipworms and to assess our success in bringing the members of these communities into pure culture. Figure 19 summarizes progress toward these aims. Here, putative symbiont isolates are compared to metagenome assembled genomes (MAGs) based on 120 conserved protein coding genes. Closed circles indicate MAGs that are >99% identical to a cultivated symbiont isolate. Closed squares indicate isolates that are >99% identical to a MAG. These results show that the majority of identified MAGs correspond to isolates brought into pure culture. In other words, most of the diversity of these shipworm symbiont communities has been successfully cultured. A major aim for the coming performance period is to obtain high-quality closed circular genome sequences for representatives of most symbiont community members to facilitate genomic analyses. This is being done by combining Oxford Nanopore long reads and Illumina short reads in hybrid assemblies. Black stars in Figure 19 indicate completed high quality closed circular isolate genomes, open stars indicate genome improvement in progress, and blue stars indicate high quality genomes with fewer than 5 scaffolds.

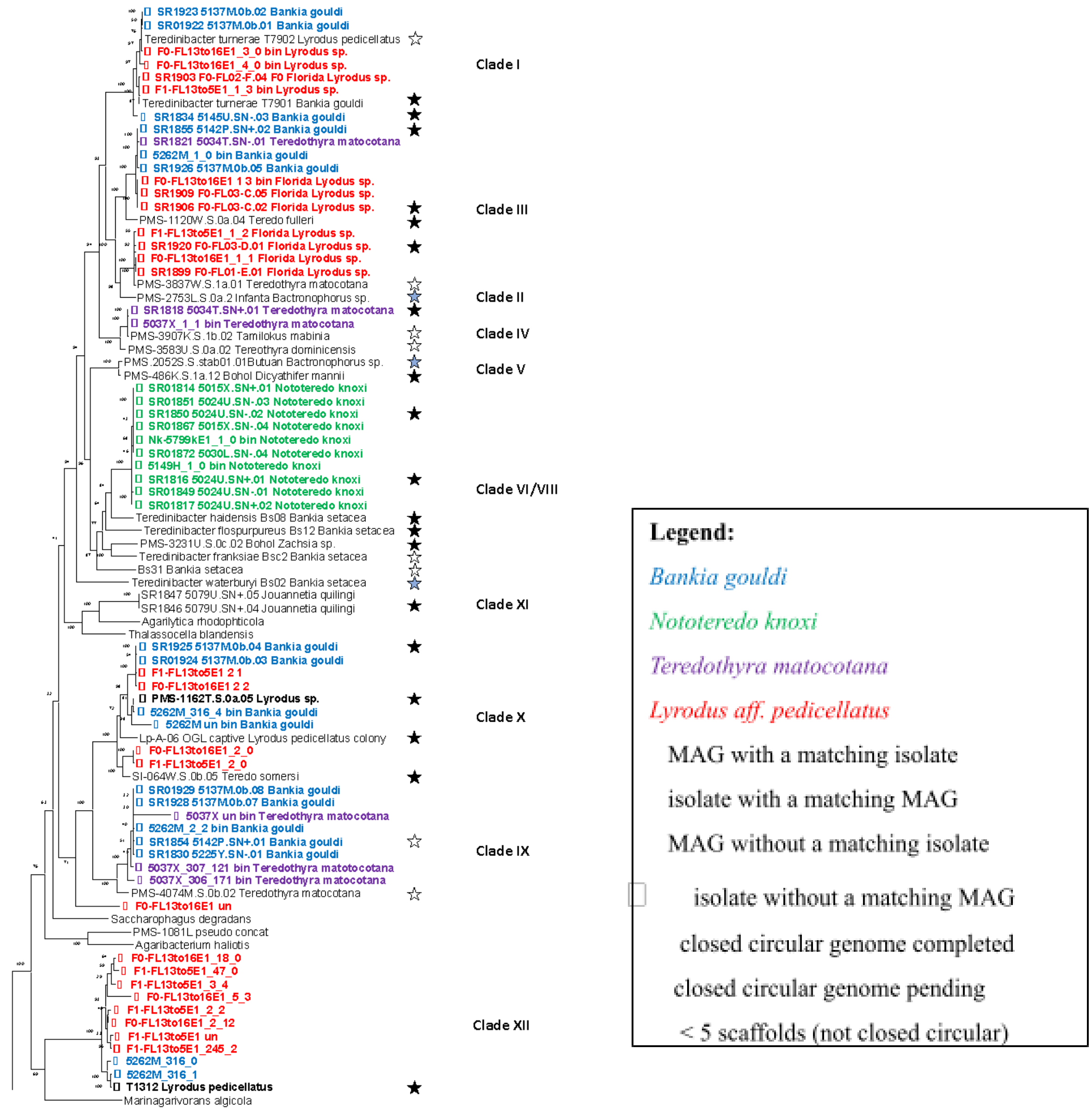

Figure 19. Combined phylogenetic tree for putative symbiont isolates and metagenome assembled genomes (MAGs). Conserved bacterial marker genes present in bacterial genomes and metagenomic bins were identified using GTDB-Tk software. Amino acid sequences of marker genes were automatically trimmed, aligned, and concatenated. The final concatenated dataset has 5,037 amino acid characters, which represent a total of 120 markers. The number of marker genes in bacterial genomes ranges from 96 to 119; while metagenomic bins range, from 15 to 118. The displayed figure is a subtree from a larger phylogenetic tree inferred by maximum-likelihood using RAxML v.8 using PROTGAMMABLOSUM62 substitution matrix and 250 bootstrap replicates on an alignment of 146 OTUs including 44 shipworm isolates, 37 metagenome bins, and 65 bacterial reference species.

4.a.6. Analysis of secondary metabolite pathways of shipworm symbionts

Phylogenetic analyses: Biosynthetic gene clusters (BGCs) are groups of genes identified as encoding secondary metabolite biosynthetic pathways, including those producing polyketides, non-ribosomal peptides, and bacteriocins. This diverse category of pathways has historically been the source of many successful antibiotics, anticancer agents, and analgesics drugs. Our group recently categorized the BGCs found in shipworm symbionts into biosynthetic gene cluster families (GCFs) based on sequence homology. Figure 20 summarizes the occurrence of 172 GCFs in the 24 AUF genomes and 23 additional cellulolytic shipworm symbiont genomes.

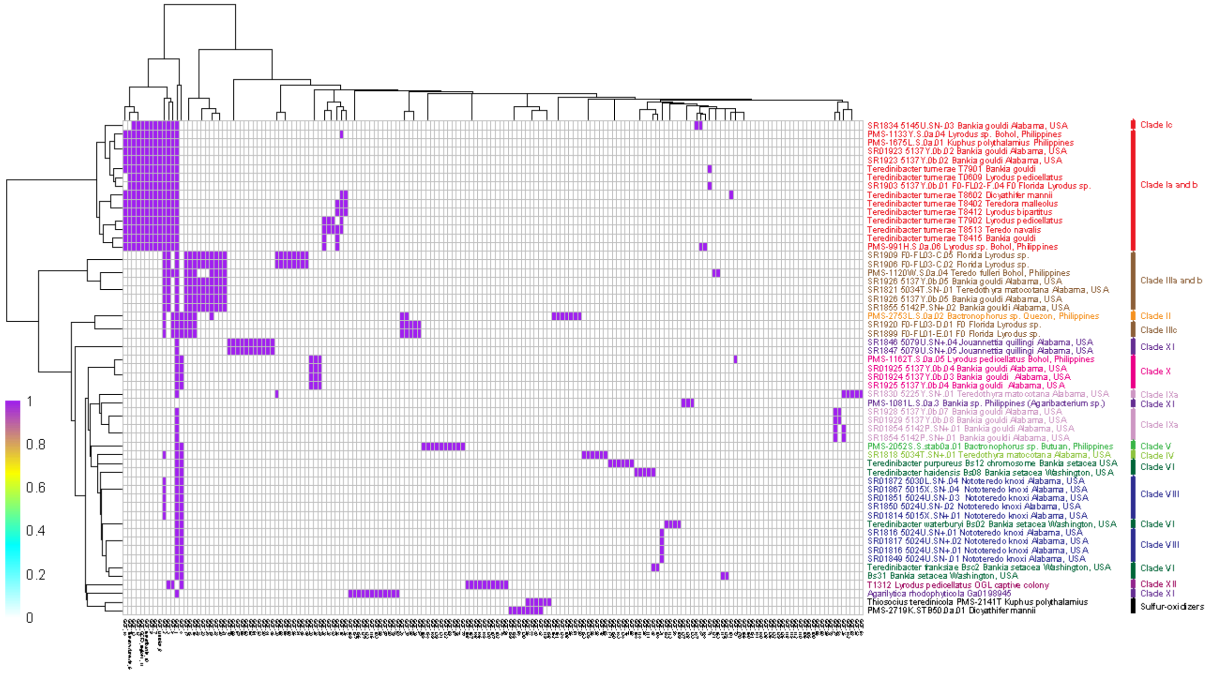

Figure 20. Cluster diagram demonstrating distribution of secondary metabolite biosynthetic gene cluster families (GCFs) among shipworm symbiont isolate genomes and clades. Distribution of 172 GCFs identified in the genomes of 24 shipworm symbiont genomes from the AUF site (bold) and 23 shipworm symbiont isolates from other locations (Ward clustering). Strain names are color coded according to the symbiont clades identified by 16S rRNA phylogeny.

These results show:

- As a group, shipworm symbiont genomes are rich in biosynthetic pathways.
- The occurrence of individual GCFs is strongly correlated with host phylogeny, i.e., strains within a single clade or species show strong similarity in the GCFs that they encode.
- Host phylogeny is therefore a good predictor of secondary metabolic potential.
- The occurrence of similar GCFs within clades and distinctly different GCFs between clades suggests that GCFs are predominantly inherited vertically, rather than by horizontal gene transfer. This suggests that examination of phylogenetically distinct symbiont genomes will likely yield distinct sets of secondary metabolites with potentially useful characteristics.
- One GCF (GCF_1) is present in all but one of the 47 cellulolytic shipworm symbiont genomes, suggesting a potentially critical role for this gene product in the formation or maintenance of shipworm symbioses. This GCF is predicted to be an uncharacterized bacteriocin.

Biochemical screening and compound discovery: To date, 66 isolates were grown, extracted and are currently being tested in five assays described below. These are constellation pharmacology (analgesic; 45 isolates), HEKOETRP (analgesic; 21 isolates), M^{pro} (SARS-COV-2 main protease; 22 isolates), antimicrobial susceptibility; *Staphylococcus aureus* (antimicrobial: 30 isolates) and *Escherichia coli* (43 isolates) assays.

Constellation pharmacology: Constellation pharmacology is a calcium imaging cell-based high-content phenotypic screening platform that is used to assess the molecular targeting profiles of samples/compounds on dorsal root ganglion (DRG; neurons). A primary culture of ~2,000 mouse DRG neurons are plated on wells, which are then monitored for intracellular Ca^{2+} concentration changes using fura2AM dye. Initial screening of samples includes periodic depolarizations of the DRGs with extracellular application of either ATP or KCl to activate purinergic receptors or voltage-gated calcium channels, respectively. The responses of individual neurons to sample/compound application are scored based on their direct, inhibitory, and excitatory effects. Different cell types are then divided into 16 subclasses based on their responsiveness to different pharmacological identifiers. This method is useful for discovering neuroactive compounds.

Fluorometric calcium flux assay with Human Embryonic Kidney Cells-293 Overexpressing Transient Receptor Potential Channels (HEKOETRP): This is a cell-based calcium imaging assay that is used to identify compounds/samples that activate, inhibit, or functionally modulate different human transient receptor potential channels (i.e., TRP-Melastatin 8, TRP-vanilloid receptor 1, TRP ankyrin 1). For this assay, human embryonic kidney (HEK-293) cells that are stably over expressing human TRP channels are cultured, seeded, and incubated with the calcium indicator Fluo4-AM on a 96-well plate. Changes in cellular fluorescence are monitored in response to incubation of sample/compounds. This method is useful for discovering neuroactive compounds.

Antimicrobial Susceptibility Testing: The antimicrobial potential of samples is initially assessed against pathogenic strains of *Staphylococcus aureus* and *Escherichia coli* following a standard colorimetric assay using [3-(4,5-dimethyl-2thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide] (MTT). This method is useful for discovering antimicrobial compounds.

SARS-CoV-2 main protease M^{pro} Screening: The main protease of the SARS-CoV-2 coronavirus has emerged as a potential drug target for the development of pharmacological agents against this and other coronaviruses. The Schmidt lab has expressed and purified this enzyme and optimized a fluorogenic assay to screen for inhibitors. This method is useful for discovering antiviral compounds, specifically targeted to SARS-CoV-2, the causative agent for COVID-19.

Compound discovery: A manuscript describing new antibiotic compounds Turnercyclamycins A and B, discovered in a prior reporting period and active against a variety of drug resistant bacterial pathogens including colistin-resistant Acinetobacter, has been published and is now in print. These compounds not only show activity at low minimum inhibitory concentrations (MICs) ranging from 1-8 µg/mL and no apparent cytotoxicity against the human kidney-derived HEK-293 cells but also are non-hemolytic up to 64 µg/mL. While these compounds belong to well-known structural classes of antibiotics, they include unique substituents and substructures, and so may be considered a new family of compounds. Preliminary data indicates that these, or closely related compounds, are not only expressed by this bacterial strain when grown in culture but can also be detected in tissues of shipworm specimens collected at the Alabama site. Due to the conserved nature of the biosynthetic genes responsible for producing these compounds, it will be possible to search in tissues, genomes, and metagenome sequences of additional strains from additional sources for new variants of these exciting new molecules.

Lipopeptide antibiotics from Clade I: One of the biosynthetic gene clusters that is highly conserved in shipworm symbiont Clade I was further investigated. This gene cluster encodes a 13 module NRPS, and is present across all sequenced *T. turnerae* strains, and is detected in the gill tissue metagenomes of *T. turnerae* containing shipworms including strains SR01923, SR01922, and SR01903 isolated from AUF specimens. Due to the previous extensive work on optimizing growth conditions in the type strain T7901, this lineage was used to investigate the chemical product of this BGC.

Chemistry-guided fractionation and isolation efforts, using DAD-HPLC and UPLC-MS methods to target lipopeptides, led to the purification of a series of new cyclic lipodepsipeptide compounds named the Turnercyclamycins. These compounds include a number of interesting structural features including three beta-hydroxylated amino acids and five epimerized alpha carbons (Figure 21). The domain architecture of the NRPS modules maps well to the elucidated structure, and three additional genes just upstream of the cluster are predicted to encode enzymes capable of installing the hydroxyl groups (Figure 22).

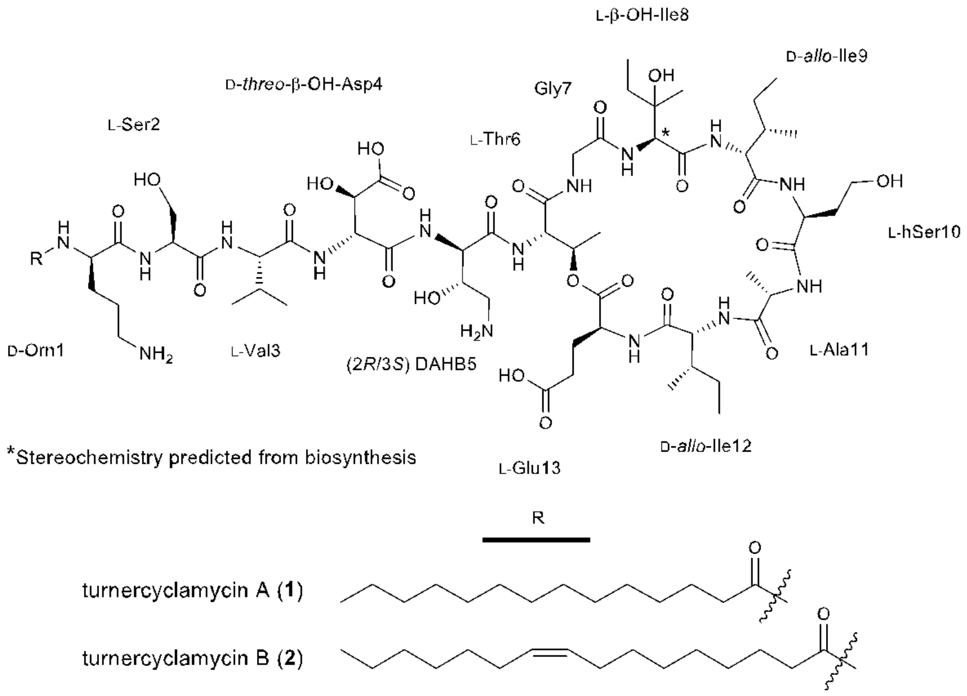

Figure 21: Structures of Turnercyclamycins A and B.

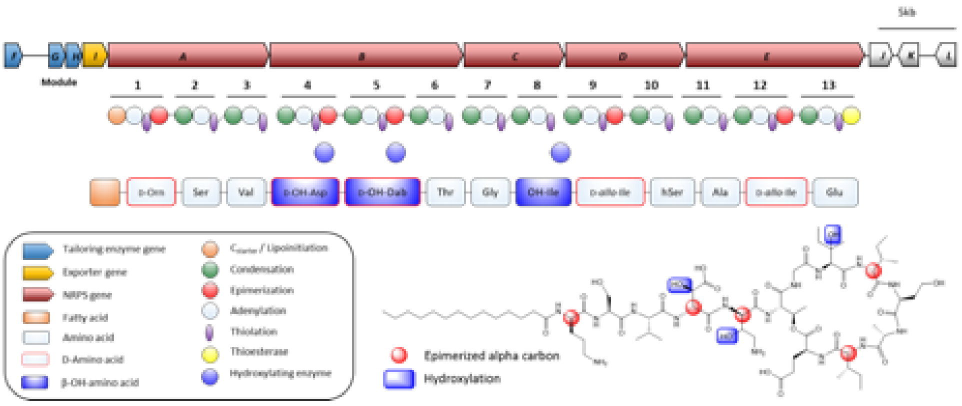

Figure 22: Proposed biogenesis of Turnercyclamycins, highlighting epimerization and hydroxylation events.

Broth microdilution assays show these compounds to have selective activity against gram-negative pathogens, including *Acinetobacter baumannii*, at concentrations that are non-cytotoxic against mammalian cells (Table 3) and do not lyse murine erythrocytes (data not shown). This profile is extremely promising for antibiotic drug discovery, where gram-negative active compounds are severely limited. Furthermore, the compounds retained activity against a panel of colistin-resistant *Acinetobacter* complex strains (Table 3), and phenotypic assays have identified the outer-membrane as the likely target.

Pathogen or Cell Line	MIC (µg/mL)	
	1	**2**
Acinetobacter baumannii	8	8
Escherichia coli	1	1
Enterococcus faecalis	>64	>64
Klebsiella pneumoniae	2	2
Staphylococcus aureus	>64	>64
HEK-293	>64	>64

Table 3: Turnercyclamycin activity against a panel of Gram-negative pathogens, Gram-positive pathogens, and mammalian HEK-293 cells.

In performance period 6, two intriguing biosynthetic gene clusters (BGCs) that are linked to homoserine lactone synthases and LuxR transcriptional regulators have been identified in isolates from the Alabama collections. The first of these, which is found in strain SR01818, was previously reported as being highly conserved in shipworm symbionts and the HSL signaling molecule that regulates its expression was identified (Robes et al. 2022). Despite the interest in this cluster, however, the final molecular product has not been isolated and described. The cluster is a hybrid TransAT-PKS/NRPS and includes an efflux pump, an enzyme related to cell membrane biogenesis, and an outer membrane lipoprotein sorting protein, which could all reasonably be associated with resistance mechanisms to an antibiotic product. Isolate and structure elucidation of this molecule is currently underway using the HSL signaling molecule to stimulate production in a number of strains.

The second QS-linked pathway is found in the *Jouanettia*-associated SR01864 and consists of a hybrid Type 1 PKS/NRPS. This cluster is quite unique, and has not been identified in other strains. Interestingly, the cluster includes a number of transport related proteins, as well as a swarming motility protein swrC which is linked to self-resistance to surfactin production in *B. subtilis* (Kearns et al. 2004). Currently, investigations are ongoing to identify the HSL signaling molecule that activates the LuxR regulator in order to upregulate the pathway and isolate the final molecular product.

Total synthesis of turnercyclamycin analogs: Previously, a series of lipopeptide antibiotics called the turnercyclamycins were isolated from several strains of *Teredinibacter* isolated from shipworm gills. In order to fully explore the structure-activity relationship of this family of compounds, a total synthesis is currently underway. The general synthetic scheme involves solid phase peptide synthesis (SPPS) starting by anchoring the side chain carboxylic acid of the final glutamic acid residue to the resin bead. The peptide and lipid tail can all be synthesized through SPPS, then orthogonal protecting groups can be cleaved from the C-terminal carboxylate and the threonine side chain hydroxyl, allowing for the on-bead cyclization reaction. Finally, global deprotection can be done simultaneously with cleavage from the resin using acid. Currently, two starting materials, which are not commercially available, are being synthesized and prepared for use in the SPPS methodology.

4.a.7. Carbohydrate Active Enzyme (CAZyme) Analyses

Enzymes responsible for the breakdown of carbohydrates are classified as Carbohydrate-Active Enzymes or CAZymes. The CAZymes primarily responsible for wood (lignocellulose) decomposition belong to a class of CAZymes called glycoside hydrolases (GH). Glycoside hydrolases are currently classified into 171 families (GH families) based on amino acid sequence homology. Figure 23 shows a preliminary summary of the number and identity of GH family genes identified in the genomes of 29 shipworm isolates, including 12 from this study, categorized according to the substrates against which they are active. Note that although the highest numbers of GH family genes known in a single bacterial genome belong to the human gut symbionts *Bacteroides cellulosyliticus* and *B. thetaiotamicron,* isolates from wood boring animals from the AUF site show both the highest number and the highest fraction of total GH modules dedicated to the degradation of cellulose and hemicellulose (Figure 24). This indicates that bacteria associated with shipworms and other wood boring marine invertebrates may be the most highly specialized bacteria known for the breakdown of the wood components cellulose and hemicellulose.

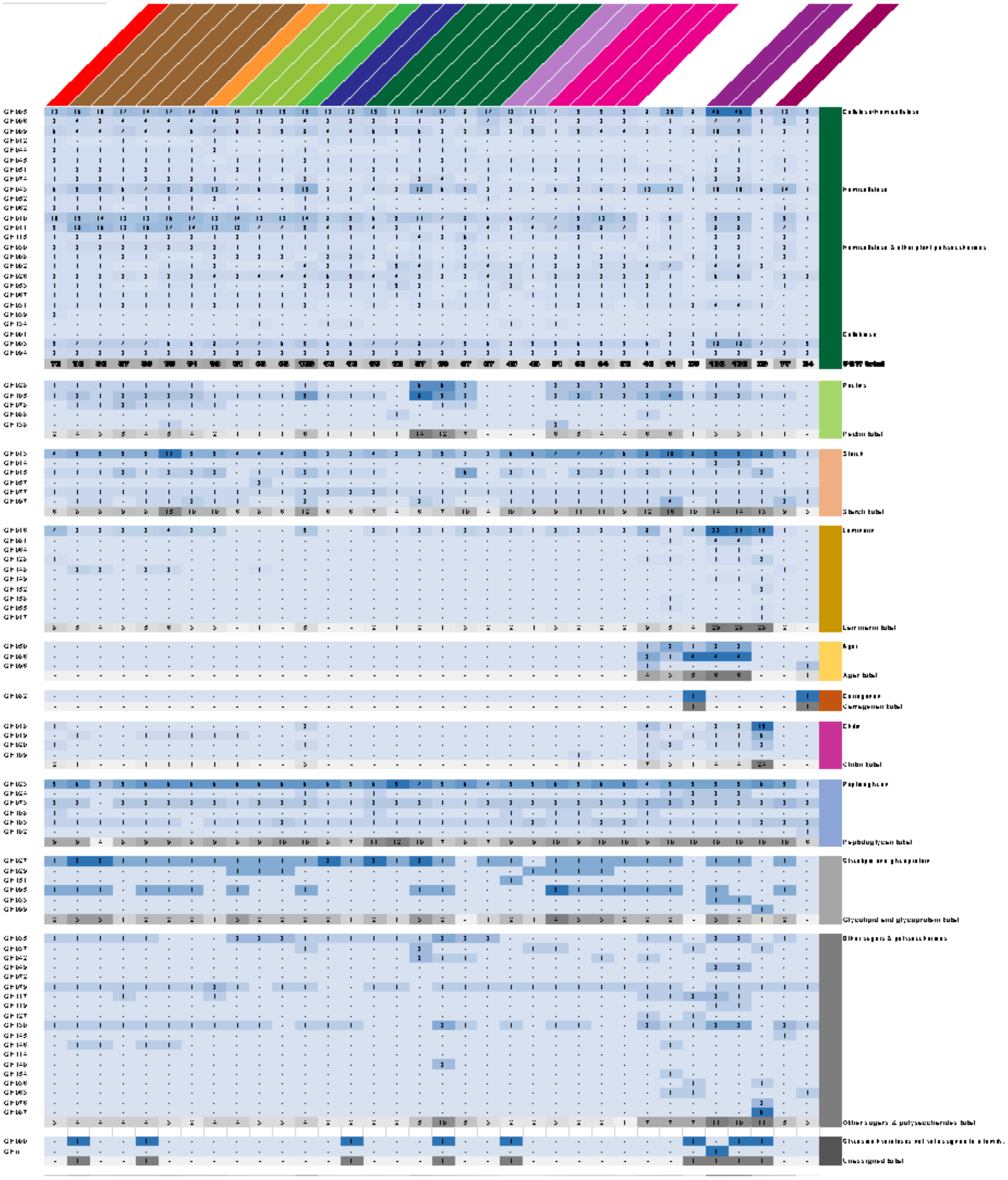

Figure 23. Heat map demonstrating the comparative representation of genes from all glycoside hydrolase families in cellulolytic isolates from 29 wood boring bivalves and 5 representative free-living cellulolytic bacteria. Dark blue indicates the greatest and light blue the least number of GH family genes in each species. Species names are color coded according to clades as in Figure 22.

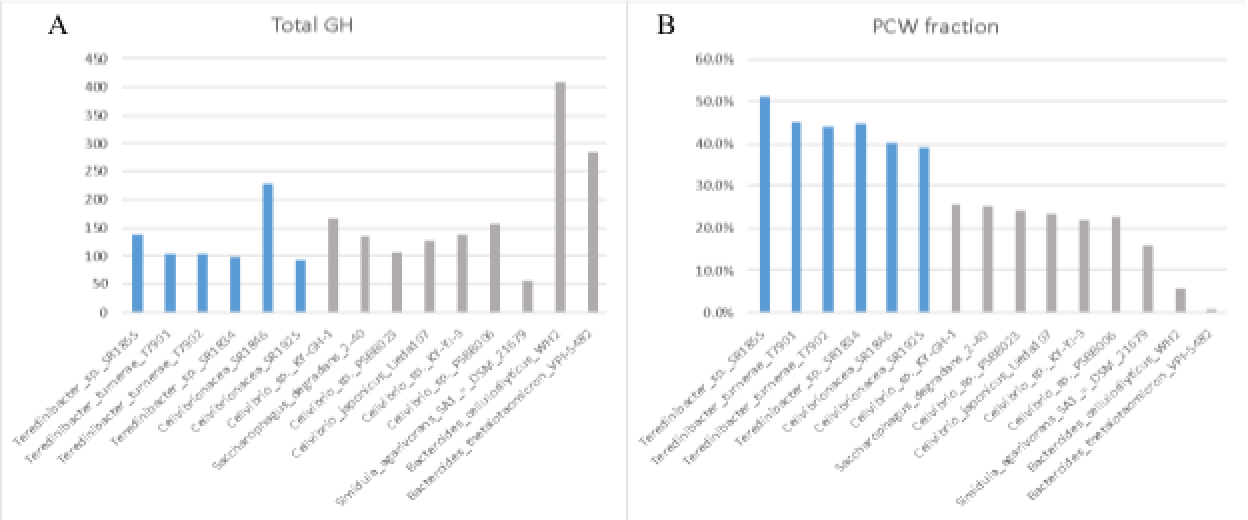

Figure 24. Glycoside hydrolase genes in genomes of isolates from wood boring animals (blue bars) and highly CAZyme-rich reference bacteria (grey bars). (A) Total number of GH family genes per genome. (B) Percent of GH modules that act on plant cell wall (PCW) components (cellulose and hemicellulose). Note that although the highest numbers of GH family genes known in a single genome belong to the human gut symbionts *Bacteroides cellulosilyticus* and *B. thetaiotamicron,* the highest fraction of GH modules dedicated to the degradation of cellulose and hemicellulose is found in the isolates from wood boring animals from the Alabama Undersea Forest (AUF) site. This indicates that these bacteria are among the most highly specialized for the breakdown of the wood components cellulose and hemicellulose.

4.b. Inventory of activities (number of submersible/ROV dives, CTD, net tows, square kilometers mapped in the U.S. Exclusive Economic Zone, technology development milestones, etc.), including coordinates and/or maps:

4.b.1. Dive activity

During our December 2019 expedition, 5 dives were conducted from the DISL research vessel *E.O. Wilson* over 2 days. Nine wood settlement panels were deployed to compare colonization patterns at the site to those in coastal locations remote from the site. Over 200 pounds of wood was collected from the site and transported to DISL for documentation and analysis. On 06/18/2020 and 08/14/2020, DISL's Dive Safety Officer lead teams of divers on 2 dive trips, conducting an additional 4 dives and deploying 18 additional settlement panels. On 08/14/2020, an additional ~50 pounds of naturally occurring wood were collected by DISL divers. This wood was maintained in flowing seawater at DISL and shipped to the Marine Science Center at Northeastern University in batches on 08/18/2020, 08/26/2020, and 10/5/2020 for subsequent dissection, sample processing, and analyses over the period from 09/3/2020 to 11/20/2020. A subsequent collection of wood was made on 10/15/2020 and was dissected and processed over the period from 11/12/2020 to 3/8/2021. On 08/13/2021, DISL divers collected naturally occurring wood and 9 settlement panels, which had been deployed since 08/14/2020 and 10/15/2020. This wood and panels were processed between 08/19/2021 and 08/30/2021. Six bald cypress settlement panels and 4 laminated pine settlement panels were deployed during the 08/13/2021 dive. On 11/8/2021, DISL divers collected another batch of found wood and deployed 2 strings of cypress baits (each string had 3 "logs" of wood). From 06/11-18/2022 co-PIs conducted 9 dives over 5 days, recovered the 3 bait strings deployed on 11/8/2021 deployed six additional bait strings. Bait strings from the June expedition were recovered by DISL divers on 08/22/2022 and 04/26/2023.

4.b.2. Site mapping

From June 11-18, 2022, researchers were able to fully document the site using high resolution and still imagery. In total, approximately 1,000 high-resolution still images and ~50 ultra high resolution (5K) videos were made using a GoPro 9. They additionally conducted transects using a Canon EOS 6D Mark II digital camera with a Canon 1:4 Zoom wide angle (17-40 mm) lens in an Ikelite DL200 housing shooting at high speed (7 frames/sec; 10 Mb/image). These transects consisted of 4 areas that measured 60 m x 5 m each, as well as a high resolution transect of the edge of ancient creek bed exposed at this site measuring 75 m along its perimeter, with a width of 2 m. These 5 transects add up to > 40,000 color-corrected images. These were used to generate 3D surfaces and orthorectified photomosaics using a workflow developed by NOAA (Suka et al. 2019) to map the sunken forest along an ancient creek bed at a depth of 18 m.

The dive team also used an Insta360 camera to take approximately 30 minutes of immersive 360-degree footage during the dives. This footage is being color-corrected and stitched together using the manufacturer’s software with further processing using the Unity framework to enable distribution on VR platforms. This video will allow the public to visit the site as if they were the divers who made the film. This VR experience can occur using either a web browser, or a virtual reality headset like an Oculus.

The team also deployed and recovered a YSI Eco 2 water quality sonde equipped with sensors for dissolved oxygen, pH, conductivity and temperature, turbidity, chlorophyll fluorescence, and depth near exposed wood at the end of the first transect. The sonde sampled for 5 days at 1-minute intervals to gather data on how variable water quality conditions are around the sunken forest (Figure 25). These data show typical offshore conditions at 18m depth with diel periodicity in dissolved oxygen, but surprisingly stable but lower than typical seawater pH over diel cycles. This may result from the location of the pH probe being close to the sediment/water interface. Temperature variations on the bottom were not influenced by internal wave activity. Here is the time series of the data collected. Note that there are two y-axes so that all data can appear on the same graph. Also, note that turbidity spikes often coincide with increased dissolved oxygen (ODO). We suspect that these spikes are driven by current mixing that includes a vertical flux bringing down more oxygenated water from higher in the water column while re-suspending the very flocculent sediment that accrues at this site from the exposure of ancient Quaternary sediments.

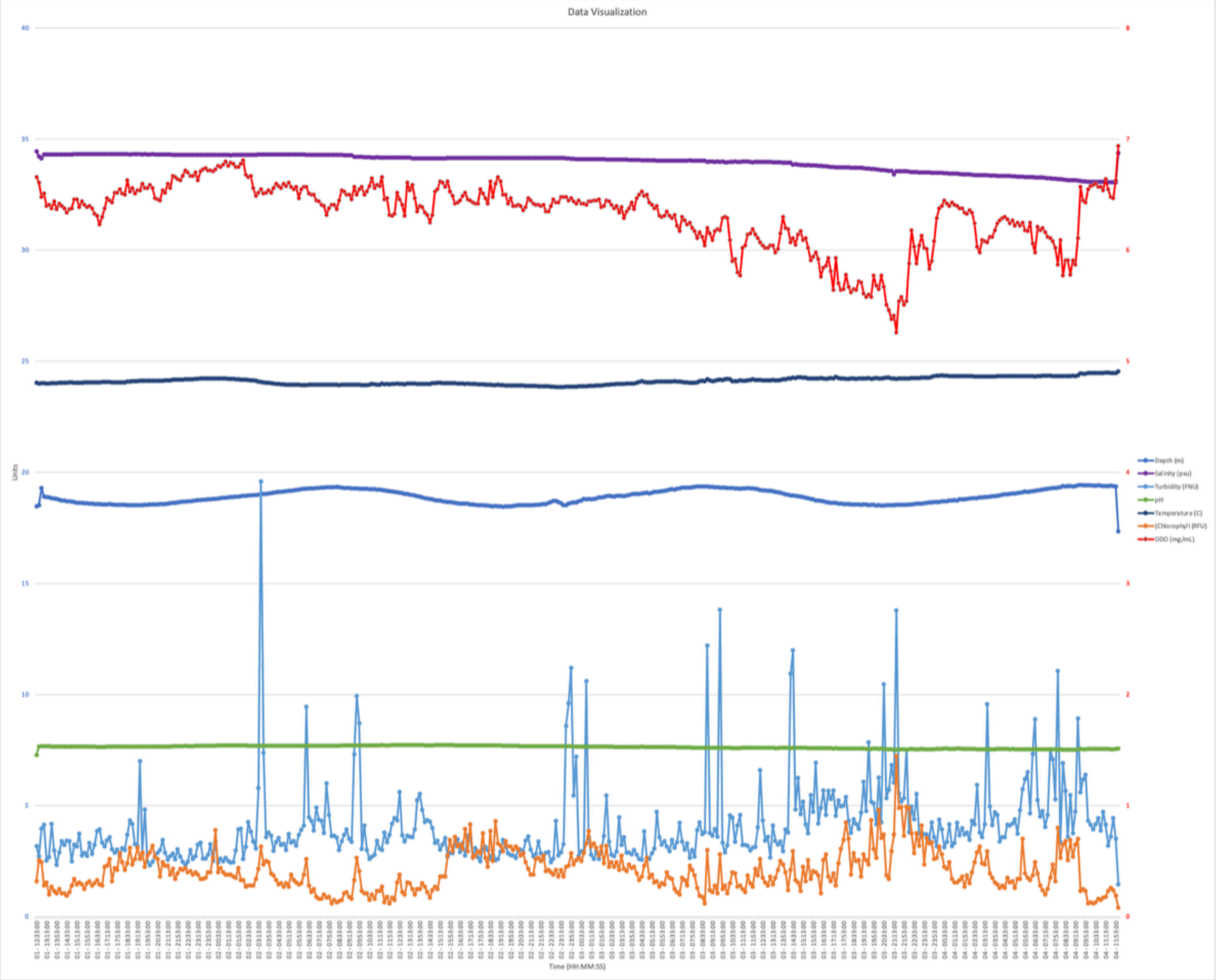

Figure 25: Time series of data logged by a YSI Eco 2 water quality sonde equipped with sensors for dissolved oxygen, pH, conductivity and temperature, turbidity, chlorophyll fluorescence, and depth deployed near exposed wood at the end of the first transect. Note that that turbidity spikes often coincide with increased dissolved oxygen (ODO).

In addition to conducting surveys using divers, the team also used the Dauphin Island Sea Lab's Simrad side scan and 3D imaging sonar system recently installed on the R/V *E.O. Wilson* to explore potential new sites suggested by a collaborator at Louisiana State University. At two of the suggested sites, the team saw objects proud of the bottom with clear acoustic shadows, suggestive of exposed wood. The presence of wood was confirmed by sequentially deploying a Trident Remotely Operated Vehicle (ROV) to find structures looking like wood on the video feed with a close approach by the ROV, followed by a quick SCUBA dive to confirm that the images seen on the ROV feed were wood. These efforts resulted in wood being positively confirmed at two new locations. One of the new locations appeared to correspond to an interior section of the ancient forest, where the trunks were much larger and more widely spaced than at our main location.

We initiated construction of Structure-from-Motion (SfM) models from the site based on data collected in June 2022. Following best practices suggested by Suka et al. (2019) and Sandin et al. (2019), we established a workflow for very high-resolution 3D model, and orthomosaic creation using images we gathered using a Canon D6 camera in an Ikelite underwater housing. SfM models are being constructed for 4 areas surveyed in June, each approximately 20 m x 3 m. SfM used either Metashape 2.0.1 run locally using an Apple Mac Pro workstation, CPU: 3.2 GHz 16-Core Intel Xeon W, Graphics: AMD Radeon Pro 580X Compute Engine with 36 compute units @ 1200 MHz, 8192 MB, Memory: 96 GB 2933 MHz DDR34, running MacOS Ventura 13.2.1, or using Metashape 1.8.3 running locally on the same workstation but with batch processing in the agisoft.com cloud. Compute times for local processing for c. 3,000 images per line (6 passes) are c. 24 hours. Compute times in the cloud are on the order of 12 hours. Four lines are in process, and we are experimenting with building a 3D model of the stream bank, which has high structural complexity, and appears to have high biodiversity based on qualitative assessment of species observed.

A key performance goal was to produce models and orthomosaic capable of showing features at sub-millimeter scale so that biological features such as shipworms and other invertebrates that use the exposed wood as habitat can be identified visually through model manipulation. Our results show that this goal can be achieved but requires models to have c. 200M points in the dense cloud, derived from stereo pairing of images, with a multitude of settings in the Metashape workflow set to very specific values. Digital Elevation Models (DEM) were also derived (see Figure 26). We found that a 20 m x 3 m benthic survey was doable on a single dive, resulting in 2560-4,000 images/per dive. The resolution obtained was < 1 mm and was resolved to 0.25 mm when visibility was excellent. Major results: trees, bark, limbs, and trunks are all easily discernible in these maps, and the volume of wood can be estimated from the underlying volumes created from the mesh model generated in Metashape (Figure 27). Siphons of shipworms that have colonized exposed wood are visible, and Mathematica code suitable for estimating wood abundance and shipworm infestation, using EfficientNet V2 and Yolo V8 libraries, shows good promise for automatic classification of the 300 m^2 of the seafloor that was mapped.

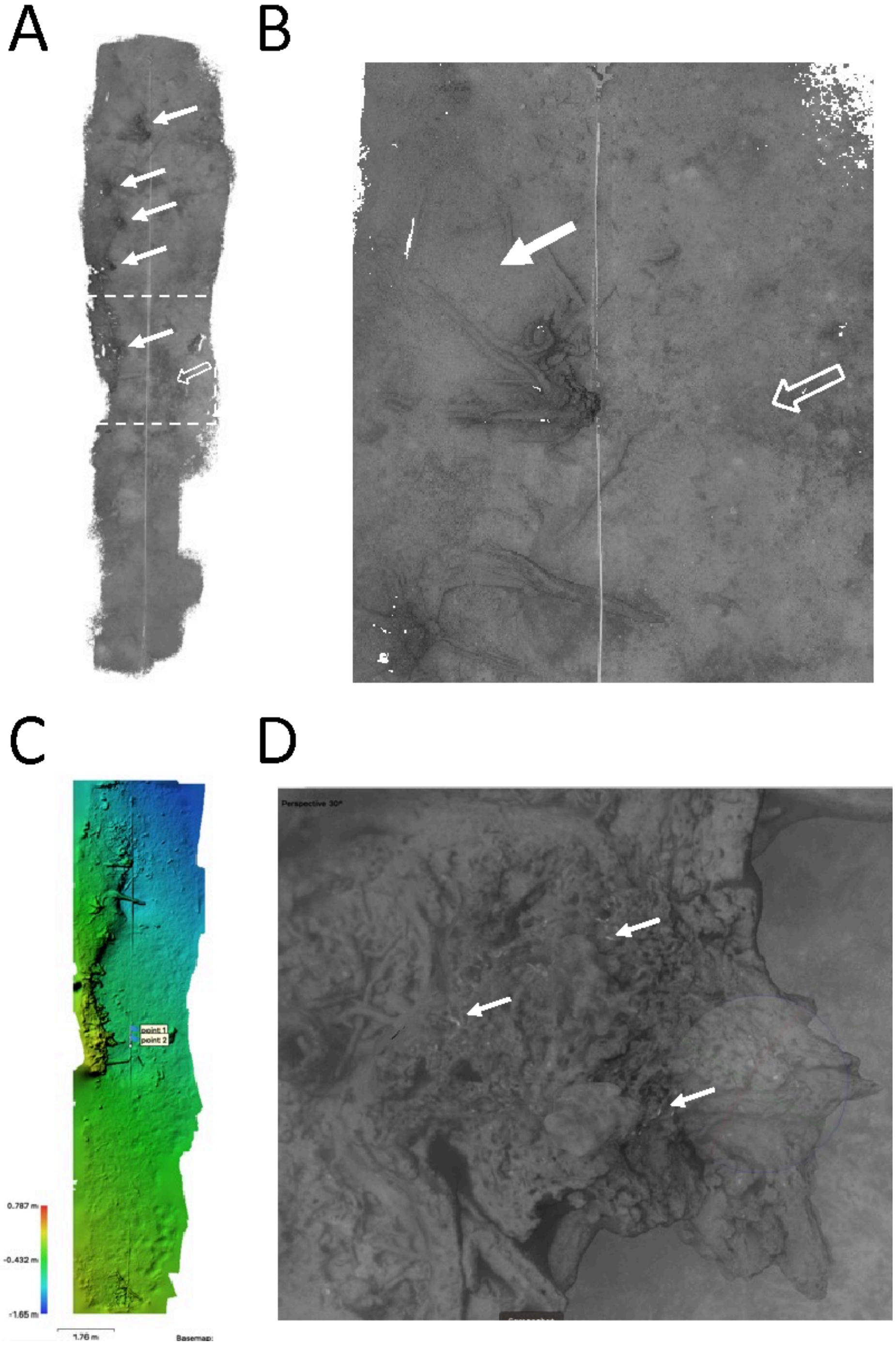

Figure 26: Structure from motion model of ancient riverbed. This image shows the first of 4 transect line surveys to be completed and is constructed from almost 200M points derived from stereo pairs of the 2,671 images gathered during the first 2 dives of June 2022 expedition. Vertical white line is the survey tape. (A) Orthomosaic of 46.3 m^2 of sunken forest along edge of an ancient stream bank. Solid arrows indicate bald cypress stumps, estimated age 60,000 years BP. Open arrow shows extensive bed of wood debris including branches and bark eroded from the reiver bank. (B) Detail of the region between the dashed lines in (A). Resolution of orthomosaic and 3D model is c. 0.2 mm/pixel. (C) Digital elevation model (DEM) of Line 1, sunken forest. Depths relative to mean depth at site of 18.6 m.

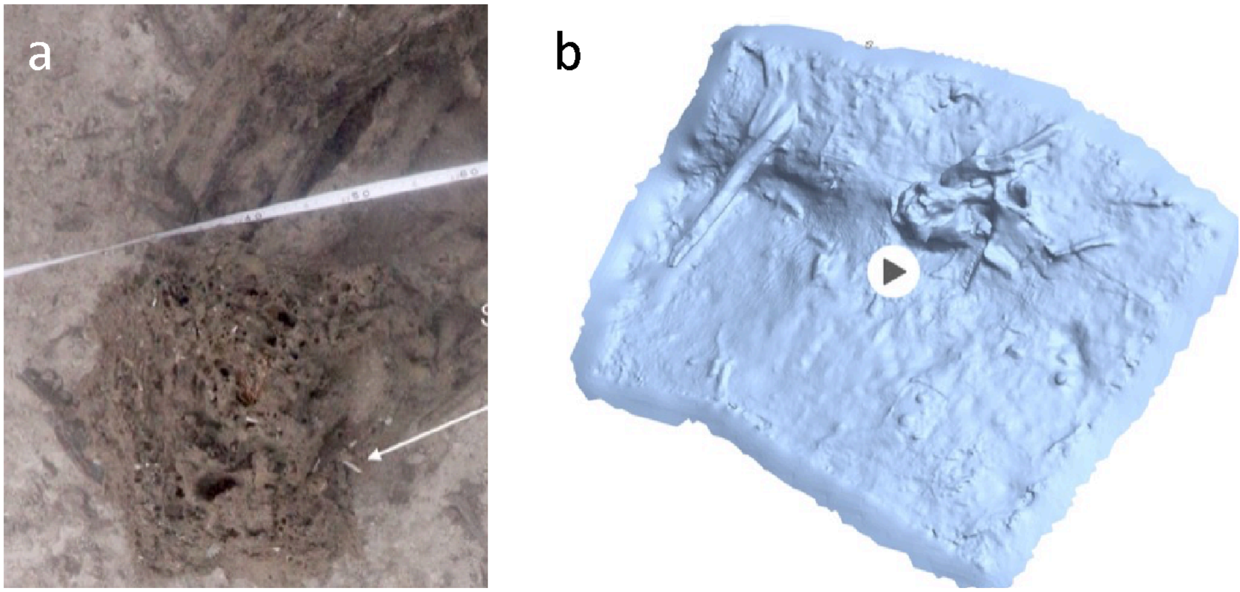

Figure 27: Detail of structure from motion models. (a) Example of structure-from-motion surface showing details allowing detection of individual shipworm tubes (arrow). Resolution of zoomed images is 0.25-1 mm. (b) Output from Mathematica showing surface mesh to two stumps at edge of creek bed that allows computation of volume of exposed wood, using the voxel count proud of the seabed. Voxels are calibrated to the underlying structure-from-motion imagery.

Note: although orientations and depths and scale bars have been used to generate these maps, they have not been georeferenced to protect the integrity of the site, which is proposed for protection by a bipartisan group led by Rep. Jerry Carl (R-AL) as H.R.6148—Alabama Underwater Forest National Marine Sanctuary and Protection Act.

4.c. Inventory of samples collected:

4.c.1. Specimen collections

1,163 individual organisms (animal specimens) have been collected from wood and baits representing 49 taxa across 10 phyla. These have been sorted, photographed, identified, and preserved for this project. A detailed inventory can be found here: https://arctos.database.museum/search.cfm?project_id=10004701&collection_id=362.

4.c.2. Samples

From these specimens, 3,588 tissue samples and bacterial stocks have been deposited to the OGL collection. 991 samples were distributed to partners for analysis. A detailed inventory can be found here: https://arctos.database.museum/search.cfm?project_id=10004701&collection_id=362.

4.c.3. Bacterial strains

260 bacterial strains and 2,352 bacterial freezer stocks have been added to the OGL collection A total of 777 bacterial stocks extracts have been distributed to partners for analysis.

4.c.4. DNA extracts

173 DNA extracts have been added to the OGL collection. A detailed inventory can be found here: https://arctos.database.museum/project/10004701.

4.d. List resulting publications, presentations, websites, etc. All publications must acknowledge NOAA Ocean Exploration funding and be submitted to the NOAA Institutional Repository. Abstracts, publications, and other materials must be appended or linked to this report:

4.d.1. Publications (7 peer-reviewed publications)

1. Altamia MA, Shipway JR, Stein D, Betcher MA, Fung JM, Jospin G, et al. *Teredinibacter waterburyi* sp. nov., a marine, cellulolytic endosymbiotic bacterium isolated from the gills of the wood-boring mollusc *Bankia setacea* (Bivalvia: Teredinidae) and emended description of the genus *Teredinibacter*. International Journal of Systematic and Evolutionary Microbiology. 2020;70(4):2388-94. https://www.ncbi.nlm.nih.gov/pubmed/32100688.
2. Altamia MA, Lin Z, Trindade-Silva AE, Uy ID, Shipway JR, Wilke DV, et al. Secondary metabolism in the gill microbiota of shipworms (Teredinidae) as revealed by comparison of metagenomes and nearly complete symbiont genomes. mSystems. 2020;5(3).
3. Altamia MA, Shipway JR, Stein D, Betcher MA, Fung JM, Jospin G, Eisen J, Haygood MG and Distel DL. *Teredinibacter haidensis* sp. nov., *Teredinibacter purpureus* sp. nov. and *Teredinibacter franksiae* sp. nov., marine, cellulolytic endosymbiotic bacteria isolated from the gills of the wood-boring mollusc *Bankia setacea* (Bivalvia: Teredinidae) and emended description of the genus *Teredinibacter*. International Journal of Systematic and Evolutionary Microbiology. 2021;71:004627 [DOI 10.1099/ijsem.0.004627].
4. Miller BW, Lim AL, Lin Z, Bailey J, Barrows LR, Manoil C, Schmidt EW, Haygood MG, Shipworm symbiosis ecology-guided discovery of an antibiotic that kills colistin-resistant *Acinetobacter.* Cell Chem Biol. 2021.
5. Li Y, Altamia MA, Shipway JR, Brugler MR, Bernardino AF, de Brito TL, Lin Z, da Silva Oliviera FA, Sumida, P, Smith CR, Tindade-Silva, A, Halanych KM, and Distel DL. Contrasting modes of mitochondrial genome evolution in sister taxa of wood-eating marine bivalves (Teredinidae and Xylophagaidae). Genome Biol Evol. 2022.
6. Altamia MA and Distel DL. Transport of symbiont-encoded cellulases from the gill to the gut of shipworms via the enigmatic ducts of Deshayes: a 174-year mystery solved. Proc Biol Sci. 2022; 289(1986): 20221478.
7. Altamia MA, Appiah-Madson HJ, Falco Poulin R, Rubin-Blum M, Huettel B, Dubilier N, Gruber-Vodicka H, Leisch N, and Distel DL. Wooden steps to shallow depths: a new bathymodiolin mussel, *Vadumodiolus teredinicola* gen. nov. sp. nov. inhabits shipworm burrows in an ancient submarine forest. Deep Sea Research Part I, (Accepted for publication).

4.d.2. Presentations and abstracts (8 invited talks, 3 posters)

1. Poster presentation: Altamia MA, Lin Z, Schmidt EW, Haygood MG, Distel, DL, Genomic, Metagenomic and Cultivation-based Analyses of Symbiont Phylogeny, Host Range and Community Composition in Shipworm Hosts *(Teredinidae: Bivalvia),* ASM FEMS World Microbe Forum June 20-24, 2021, abstract #6603.
2. Invited lecture: Patterson MR, and Hulver A. Oxygen dynamics: scaling up from the colony to the reef using hand-held profilers and Autonomous Underwater Vehicles (AUVs). 14th International Coral Reef Symposium, Bremen, Germany. 2021.
3. Poster presentation: Altamia MA, Lin Z, Schmidt EW, Haygood MG, Distel DL. Probing the global diversity of shipworm symbionts using genomic, metagenomic and cultivation-based methods. Boston Bacterial Meeting; 2022-06-13; Boston, MA, USA.
4. Invited keynote lecture: Patterson M. Reducing the cost of data acquisition in extreme environments like sunken forests: machine learning, structure-from-motion, and novel sensing

platforms. Réunion Scientifique Sentinelle Nord: Repousser les frontières du savoir. Centre des congrès de Québec, Québec, Canada. 2022.

5. Invited poster: Patterson M. Fetch-class Autonomous Underwater Vehicle (Northeastern University) and Ascension microplastics sampler (Ocean Diagnostics Inc): hardware and software tools to overcome the aliasing problem. Réunion Scientifique Sentinelle Nord: Repousser les frontières du savoir. Centre des congrès de Québec, Québec, Canada. 2022.
6. Invited lecture: Patterson MR, Helmuth B, and Distel DL. Structure-from-motion and marine robotics: examples from a sunken forest in the Gulf of Mexico. Northeastern University-Woods Hole Oceanographic Institution Workshop. 2023.
7. Invited lecture: Patterson M. AUVs, ROVs, and advanced sensors for understanding water column and benthic processes in Lake Manicouagan: a freshwater sunken forest with parallels to the Gulf of Mexico sunken forest. Workshop on MANIC-5 (Lac Manicouagan). Département de geographie, Université Laval, Université do Québec à Chicoutimi, Institut national de la researche scientifique. 2023.
8. Invited lecture: Distel DL. Life in an ancient undersea forest: the secrets of a marine ecosystem powered by wood. Darwin Festival, Salem State University, Salem, MA, USA. February 15, 2023. A recording is available here: https://www.youtube.com/watch?v=Qp1Cbz6_Xm8&t=2s.
9. Invited lecture: Distel DL. Marine symbiosis powered by wood. Center for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria. June 15, 2023.
10. Invited seminar: Distel DL. Shipworms: collaborative opportunities for protocol development in genomics and proteomics. New England Biolabs, Beverly, MA, USA. August 15, 2023.
11. Invited talk: Distel DL. Symbiosis research at OGL. Gloucester Marine Genomics Institute, Gloucester, MA, USA. September 12, 2023.

4.d.3. Videos

1. Fischer A, and Eames M. (2022-10-31). Experience the wonder of the Alabama Undersea Forest. YouTube. https://www.youtube.com/watch?v=Q9x_pcO5AxU
2. Choi F, Helmuth B, and Patterson M. (2023-04-03). Line 1 Sunken Forest. YouTube. https://www.youtube.com/watch?v=TcugT3ZRCeI
    a. This 3D model was created in Agisoft Metashape of part of the AUF site. The animation zooms in on some tree stumps where the tubes from shipworms (little white cylinders) who are feasting on this wood are visible. The model was derived from 2,671 images.

4.d.4 Websites

1. Alabama Undersea Forest Project Website (Northeastern University) https://undersea-forest.sites.northeastern.edu
2. Alabama Undersea Forest Project Website (NOAA Ocean Exploration) https://oceanexplorer.noaa.gov/explorations/20ancient-forest/welcome.html
3. Alabama Underwater Forest https://storymaps.arcgis.com/stories/ecaa5c2252cd411e932c08e8a6ba5c37

4.d.5 Events

1. Alabama Undersea Forest Project ham radio event. (August 29 - September 11, 2021).
    a. For details, see the description in section II.5 below.

4.d.6. Books

1. Swanson, J, “LEGO Amazing Earth: Fantastic Building Ideas and Facts About Our Planet,” Penguin Random House Books, 2023.

https://www.penguinrandomhouse.com/books/723872/lego-amazing-earth-by-jennifer-swanson/ (Provides plans to build a model of the Alabama Undersea Forest using LEGO™ bricks).

2. Swanson, J, "The Lost Forest: An Unexpected Discovery beneath the Waves," Lerner Publishing Group, 2024. ISBN: 9781728493015, https://lernerbooks.com/shop/show/23173, (Describes our research at the Alabama Undersea Forest.) April 2024 (Junior Library Guild Selection; Now available for preorder).

4.d.7. News media (39 news media publications)

1. Solomon, N., Appiah-Madson, H. J., & Distel, D. L. (2019-11-27). Diving into an ancient forest. *OGL News Briefs*. Retrieved from https://ogl.northeastern.edu/2019/11/27/diving-into-an-ancient-forest/
2. Crum, E. (2020-03-31). Bioprospecting for industrial enzymes and drug compounds in an ancient submarine forest. *NOAA Ocean Exploration and Research*. Retrieved from https://oceanexplorer.noaa.gov/explorations/20ancient-forest/welcome.html
3. Klein, J. (2020-03-31). A forest submerged 60,000 years ago could save your life one day. *The New York Times*. Retrieved from https://www.nytimes.com/2020/03/31/science/underwater-forest-shipworms.html?smtyp=cur&smid=tw-nytimesscience
4. Klein, J. (2020-04-04). Creatures in this underwater forest could save your life one day. *Chicago Tribune*. Retrieved from https://www.chicagotribune.com/featured/sns-nyt-underwater-forest-creautres-provide-valuable-information-20200404-yszqa4otordmdjddxpvxshpxnq-story.html
5. Ferreira, B. (2020-04-06). Scientists unearth an ancient lost forest 'entombed' off Alabama's coast. *VICE*. Retrieved from https://www.vice.com/en_us/article/7kz3wy/scientists-unearth-an-ancient-lost-forest-entombed-off-alabamas-coast
6. Liberatore, S. (2020-04-06). Shipworms pulled from underwater forest that grew 60,000 years ago could be used to make new life-saving antibiotics. *Daily Mail*. Retrieved from https://www.dailymail.co.uk/sciencetech/article-8192977/Shipworms-pulled-underwater-forest-grew-60-000-years-ago-used-make-antibiotics.html
7. Price, M. (2020-04-06). Prehistoric forest of massive trees found buried in sea floor off Alabama, NOAA says. *Miami Herald*. Retrieved from https://www.miamiherald.com/news/nation-world/national/article241683666.html
8. Shipworms from 60,000-year-old underwater trees promise new effective antibiotics-research. (2020-04-07). *Sputnik News*. Retrieved from https://sputniknews.com/science/202004071078868544-shipworms-from-60000-year-old-underwater-trees-promise-new-effective-antibiotics---research/
9. Graham, K. (2020-04-07). An ancient forest could help scientists to develop new medicines. *Digital Journal*. Retrieved from http://www.digitaljournal.com/tech-and-science/science/an-ancient-forest-could-help-scientists-to-develop-new-medicines/article/569937
10. A 60,000-year-old underwater forest may help pioneer new medicines. (2020-04-08). *NBC-2*. Retrieved from https://www.nbc-2.com/story/41985642/a-60000yearold-underwater-forest-may-help-pioneer-new-medicines

11. Krajcsik, J. (2020-04-08). A 60,000-year-old underwater forest found off Alabama’s coast may pioneer new medicines. *Pennsylvania Real-Time News*. Retrieved from https://www.pennlive.com/news/2020/04/a-60000-year-old-underwater-forest-found-off-alabamas-coast-may-pioneer-new-medicines.html
12. Lee, A. (2020-04-08). Scientists uncover a 60,000-year-old forest underwater and think its preserved trees may help pioneer new medicines. *CNN*. Retrieved from https://www.cnn.com/2020/04/07/us/ancient-underwater-forest-alabama-scn-trnd/index.html
13. Lieser, E. K. (2020-04-08). 60,000-year-old preserved wood has ancient bacteria that could create new medicines. *The National Interest*. Retrieved from https://nationalinterest.org/blog/buzz/60000-year-old-preserved-wood-has-ancient-bacteria-could-create-new-medicines-142392
14. Nagaraj, R. (2020-04-08). Scientists discover an ancient underwater forest that may help produce new medicines. *News Landed*. Retrieved from https://newslanded.com/2020/04/08/underwater-rainforest/
15. Scientists study 60,000-year-old underwater forest. (2020-04-10). *Daily News*. Retrieved from http://www.dailynews.lk/2020/04/10/world/216212/scientists-study-60000-year-old-underwater-forest?page=1
16. Distel, D. L. (2020-04-10). OGL in the New York Times. *OGL News Briefs*. Retrieved from https://ogl.northeastern.edu/2020/04/10/ogl-in-the-nytimes/
17. Lee, A. (2020-04-10). Ancient underwater forest could help pioneer new medicine. *Napa Valley Register*. Retrieved from https://napavalleyregister.com/news/science/ancient-underwater-forest-could-help-pioneer-new-medicine/article_cf96470a-75ab-5865-beb9-b9dab7f77a3f.html
18. The underwater forest and the future of medicine. (2020-04-11). *Reuters*. Retrieved from https://www.reuters.com/video/watch/idRCV008334?fbclid=IwAR1MGKyVTgy8K3yKfZZqtCLOXBR6PpfpoDXZ6bS8mnxY3xTh_V50Wij7vcc
19. Manson, L. (2020-04-11). Forest submerged for 60 thousand years may be a source of new medicines. *Somag News,* p. 4. Retrieved from https://www.somagnews.com/forest-submerged-60-thousand-years-may-source-new-medicines/
20. O’Malley, I. (2020-04-13). 60,000-year-old underwater forest could make new drugs, scientists say. *The Weather Network*. Retrieved from https://www.theweathernetwork.com/ca/news/article/60000-year-old-underwater-forest-could-make-new-drugs-scientists-say
21. Rayne, E. (2020-04-13). Submerged 60,000-year-old fossilized forest could be hiding medicines of the future. *SYFY Wire*. Retrieved from https://www.syfy.com/syfywire/fossilized-forest-hiding-medicines-future-coronavirus
22. Dalheim, R. (2020-04-14). Scientists discover 60,000-year-old underwater cypress forest. *Woodworking Industry News*. Retrieved from https://www.woodworkingnetwork.com/news/woodworking-industry-news/scientists-discover-60000-year-old-underwater-cypress-forest
23. (2020-04-14). *Underwater forest could help in providing new medicines* [Retrieved from http://www.byuradio.org/episode/7b035e43-0dd7-4b33-8b59-017e7ab7463d/top-of-mind-with-julie-rose-pandemic-business-stimulus-victory-gardens-pandemic-voices-part-ii?playhead=4258&autoplay=true&fbclid=IwAR1csr3POCshclLw1ouIJuv68FgsGjtPx7fnDyzBZ4oLABCsog05Gy

24. Romano, A. (2020-05-05). Scientists found a 60,000-year-old forest underwater — and it could improve people's lives in the long run. *Travel + Leisure*. Retrieved from https://www.travelandleisure.com/travel-news/scientists-find-prehistoric-bald-cypress-forest-underwater

25. Mitra Das, S. (2020-05-16). An ancient sunken forest can give us vital medicines — and climate lessons. *The Times of India*. Retrieved from https://timesofindia.indiatimes.com/an-ancient-sunken-forest-can-give-us-vital-medicines-and-climate-lessons/articleshow/75765410.cms

26. (2020-05-17). *Episode 1: How do you map an underwater forest?* Retrieved from https://solveitsciencepodcastforkids.com/podcast/episode-1-mapping-underwater

27. OceanX (Producer). (2020-06-30). Researching the ocean is one of the most promising sources of new medicines. *Ocean News*. Retrieved from https://twitter.com/oceanx/status/1277692249116196864 https://www.instagram.com/p/CCCNa_5DI9E/

28. Turner, E., Appiah-Madson, H. J., & Distel, D. L. (2020-08-18). You've got a friend in me! Promising anti-parasitic compound derived from shipworms' bacteria. *OGL News Briefs*. Retrieved from https://www.northeastern.edu/ogl/you've-got-a-friend-in-me-promising-anti-parasitic-compound-derived-from-shipworms-bacteria/

29. (2020-09-15). *How do you explore an underwater forest?* [Retrieved from https://www.mos.org/mos-at-home/pulsar/how-do-you-explore-an-underwater-forest

30. (2021-06-01). Exploring the scientific potential of an ancient underwater forest [Retrieved from https://open.spotify.com/episode/4Lqz5HaVwH5DyzGAoqZz9N]

31. Kutenplon, N., Poulin, R., Appiah-Madson, H. J., Patterson, M., & Distel, D. L. (2021-09-01). Federal Communications Commission special event station (W4U) celebrating the Alabama Undersea Forest. *OGL News Briefs*.

32. Kutenplon, N., Poulin, R., Appiah-Madson, H. J., Patterson, M., & Distel, D. L. (2021-09-02). Federal Communications Commission special event station (W4U) celebrating the Alabama Undersea Forest. *NOAA Ocean Exploration and Research*. https://oceanexplorer.noaa.gov/news/oer-updates/2021/underwater-forest-ham-radio.html

33. Piselli, G. (2021-09-15). Ocean waves meet radio waves: HAM radio at the MSC. *Between Northeastern* Tides. https://cosresearch.wordpress.com/2021/09/15/ocean-waves-meet-radio-waves-ham-radio-at-the-msc/

34. Castañón, L. (2022-04-25). What Lurks in a Drowned Forest in Alabama? *Nautilus*. Retrieved from https://nautil.us/what-lurks-in-a-drowned-forest-in-alabama-16508/

35. Distel, D. L. (2022-07-11). Going Deep on the Mitochondrial Genome. *OGL News Briefs.* Retrieved from https://ogl.northeastern.edu/2022/07/11/going-deep-on-the-mitochondrial-genome/

36. McCormick Hibbert, C. (2022, 2022-11-23). Shipworms can chew through a boat. Northeastern researchers explain their digestive process. News@Northeastern. https://news.northeastern.edu/2022/11/23/how-do-shipworms-eat/

37. Distel, D. L. (2022, 2022-11-30). A new paper from OGL solves an old mystery. OGL News Briefs. https://ogl.northeastern.edu/2022/11/30/a-new-paper-from-ogl-solves-an-old-mystery/

38. Appiah-Madson, H. J., & Distel, D. L. (2022, 2022-10-31). Experience the wonder of Alabama's undersea forest! OGL News Briefs. https://ogl.northeastern.edu/2022/11/01/experience-the-wonder-of-alabamas-undersea-forest/
39. Fischer, A., & Eames, M. (2022, 2022-10-31). Experience the wonder of the Alabama Undersea Forest. YouTube. https://www.youtube.com/watch?v=Q9x_pcO5AxU

4.e. List the final data inventory, including a complete list of all data types collected (e.g., CTD, MBES, images). Describe the location and status of the data archive and/or sample storage and the plan for timely public access. If the data are/will be archived at an approved facility outside of NCEI, the URL link(s) to the data should be provided to NOAA Ocean Exploration:

4.e.1. Site mapping data

All processed models, including image data, will reside at NOAA National Centers for Environmental Information (NCEI) and be made available to the public.

4.e.2. Images

2,784 image records have been created, and 1,544 have been published to Arctos: https://arctos.database.museum/search.cfm?media_type=any&project_id=10004701&collection_id=362.

4.e.3. Biogeographic location records

1,163 occurrence records have been created, 908 have been served to the Global Biodiversity Information Facility (GBIF) and the Ocean Biodiversity Information System (OBIS).

4.e.4. Genomic resources made discoverable via GGBN

2,456 animal and bacterial samples, and 153 DNA extracts have been made discoverable to the research community via the Global Genome Biodiversity Network (GGBN).

4.f. Note any major changes/adjustments to activities, expenditures, results, etc. reported in previously-submitted documents (e.g., Cruise Report, Semiannual Report):

Due to unforeseen delays resulting from COVID-19 and extreme weather events, we requested and received approval for two no-cost extensions, which included a re-budgeting of the funds for the remaining reporting periods. The approved budget modifications allowed us to complete the original project goals over a more extended period while keeping spending in line with the interim re-budgets and the final cumulative spending for the project in line with the originally budgeted amount.

4.g. Equipment inventory procured with grant funds and final disposition by NOAA on ownership:

No capital equipment was procured.

5. Highlights from outreach and education and diversity and inclusion activities

1. More than 200 area high school students received presentations based on this project during the "School to Sea" outreach event that took place at the Northeastern University Marine Science Center in the Fall of 2019.
2. Between 2019 and 2021, OGL staff presented the AUF project to students during the Northeastern University Marine Science Center "School to Sea" program, which focuses on high school students from local public schools. Over the course of 10 events, this program reached more than 1,500 area students.
3. Children's book author and collaborator Jennifer Swanson featured a co-PI from this project in an educational packet on TeachersPayTeachers.com, an online marketplace for original educational resources. The learning packet entitled "STEM Career Lesson, Challenge & Podcast: How do you map an underwater forest?" is part of the "Solve It For Kids" series produced by Swanson. It is

available here: https://www.teacherspayteachers.com/Store/Jennifer-Swanson-Solve-It-For-Kids. The podcast accompanying this packet has been downloaded over 2,500 times.

4. The Federal Communications Commission authorized special event callsign W4U to operate August 29 to September 11, 2021. On September 7 and 8, 2021, from 10 AM - 4 PM EDT, a ham radio station was set up on the lawn in front of the Marine Science Center, 430 Nahant Rd., Nahant, MA 01908. A co-PI, Ocean Genome Legacy staff, and the Northeastern Marine Science Center community, were on hand to answer any questions about the underwater forest research and to discuss the remarkable animals that live in the Alabama Undersea Forest in the Gulf of Mexico. Members of the public who have never seen a ham radio station in operation were able to contact ham radio operators thousands of miles away. A former NOAA Corps officer was on hand to talk about the NOAA Corps and her adventures at sea conducting marine research. The NOAA Corps officer was profiled in the September 2021 issue of QST, the official publication of the American Radio Relay League. We continued to receive responses (e.g., QSL cards, see attached document W4U.pdf) to our FCC authorized amateur radio event, call sign W4U for a long time. This event, which reached hundreds of ham radio operators around the world (see W4U.mov), made the public aware of the Alabama Undersea Forest, its unique ecology, and its potential contributions to science and medicine.
5. To facilitate public access to the AUF site, we have hosted processed 3D site models that are captured as video files (.mp4) to the Field Robotics Laboratory YouTube channel: https://www.youtube.com/channel/UCUUfgiSvrk5lQanwS_SqQyw.
6. Two new books by children's author Jennifer Swanson, featuring the work of this project at the Alabama Undersea Forest, are available for purchase or preorder:
   a. Swanson, J, "LEGO Amazing Earth: Fantastic Building Ideas and Facts About Our Planet," Penguin Random House Books, 2023. https://www.penguinrandomhouse.com/books/723872/lego-amazing-earth-by-jennifer-swanson/
      i. Provides plans to build a model of the Alabama Undersea Forest using LEGO™ bricks.
   b. Swanson, J, "The Lost Forest: An Unexpected Discovery beneath the Waves," Lerner Publishing Group, 2024. ISBN: 9781728493015, https://lernerbooks.com/shop/show/23173, April 2024.
      i. Describes our research at the Alabama Undersea Forest.
      ii. Junior Library Guild Selection.
      iii. Now available for preorder.
7. The AUF project was featured in two sessions of the Northeastern University Coastal Ocean Science Academy Marine Science Summer Day Program for middle school and high school students on July 27 and August 9, 2023, respectively, reaching a total of 39 students.

III. Evaluation

1. Accomplishments—Explain special problems, differences between scheduled and accomplished work

Two no-cost extensions were requested and received. These were required by interruptions due to COVID-19 and two hurricanes that hit our field site coinciding with planned field expeditions. With the no-cost extensions, we were able to meet or exceed all project goals without substantial changes in approach.

2. Next Steps

2.a. Planned or expected reports (professional papers, presentations, etc.):

This project has spawned many important research questions leading to manuscripts and projects that will be completed as new funding becomes available:

1. Altamia, MA, Appiah-Madson, HJ, Falco Poulin, R, Flatau, R, Gasser, M, Haygood, M, Lin, Z, Schmidt, EW, and Distel, DL. Symbiont distribution and community composition in the shipworm *Nototeredo knoxi* (Bivalvia: Teredinidae).
2. Altamia, MA and Distel, DL. The complete mitochondrial genome of *Vadumodiolus teredinicola*: insights into the evolution of Bathymodiolinae.
3. Altamia, MA, Appiah-Madson, HJ, Falco Poulin, R, Flatau, R, Gasser, M, Haygood, M, Lin, Z, Schmidt, EW, and Distel, DL. Symbiont community composition in the shipworms *Bankia gouldi*, *Nototeredo knoxi*, and *Teredothyra matocotana*.
4. Altamia, MA, Appiah-Madson, HJ, Falco Poulin, R, Flatau, R, Gasser, M, Haygood, M, Lin, Z, Schmidt, EW, and Distel, DL. Diversity and distribution of carbohydrate active enzymes of cultivated shipworm symbionts.
5. Patterson, MR, Helmuth, B, and Distel, DL. Structure-from-motion and marine robotics: examples from a sunken forest in the Gulf of Mexico.
6. Altamia, MA, Appiah-Madson, HJ, Falco Poulin, R, Flatau, R, Gasser, M, Haygood, M, Lin, Z, Schmidt, EW, and Distel, DL. Symbiont distribution and community composition in the shipworm *Nototeredo knoxi* (Bivalvia: Teredinidae).

2.b. Brief description of how project deliverables and outcomes contribute to societal and/or ecosystem well-being:

The project has been featured in 39 media mentions to date, helping to bring attention to this unique marine ecosystem, and to raise awareness of the importance of protecting the marine environment. The project has also influenced efforts to attain protection for the site as a National Marine Sanctuary under H.R.6148—Alabama Underwater Forest National Marine Sanctuary and Protection Act. Also please see section II.5 for educational impacts.

2.c. Brief description of the need for additional work, if any (next project phase, new research questions, unaccomplished work, etc.):

This project was intended to investigate the diversity of organisms and useful biochemical compounds in the Alabama undersea forest. As expected, the work produced many new as yet undescribed bacterial species, genomes, and candidate molecules for further investigation. The work also highlights the importance of submerged wood in supporting diverse marine communities, calling for further investigation of the ecological impact of naturally occurring and deployed wood on marine ecosystems and the potential of wood-associated communities for species and natural product discovery.

Additionally, this project produced a large quantity of image data that will allow estimates of percent coverage of exposed wood, fraction of wood in various categories (limbs, small branches, bark fragments, stumps, trunks), rugosity, spatial frequency of wood exposure, and percent observed to be colonized by various organisms. Other metrics useful to understand the structural complexity of this unique habitat, and its impacts on biodiversity at the site will be developed.