The first complete mitochondrial genome of *Diadema antillarum* (Diadematoida, Diadematidae)

Audrey J. Majeske[a,b], Alejandro J. Mercado Capote[a], Aleksey Komissarov[c], Anna Bogdanova[c], Nikolaos V. Schizas[d], Stephanie O. Castro Márquez[a], Kenneth Hilkert[a], Walter Wolfsberger[a], and Tarás K. Oleksyk[a,b,*]

[a]University of Puerto Rico at Mayagüez, Department of Biology, Call Box 9000, Mayagüez, PR 00681, U.S.A.

[b]Oakland University, Department of Biological Sciences, 118 Library Drive, Rochester, MI 48309, U.S.A.

[c]Applied Genomics Laboratory, SCAMT Institute, ITMO University, Lomonosova 9 Str., Saint Petersburg 197101, Russia

[d]University of Puerto Rico at Mayagüez, Department of Marine Sciences, Call Box 9000, Mayagüez, PR 00681, U.S.A.

* Corresponding author.

E-mail address: amajeske@oakland.edu (A. J. Majeske).

ABSTRACT

The mitochondrial genome of the long-spined black sea urchin, *Diadema antillarum*, was sequenced using Illumina next generation sequencing technology. The complete mitogenome is 15,708 bp in length, containing 2 rRNA, 22 tRNA and 13 protein-coding genes, plus a non-

coding control region of 133 bp. The nucleotide composition includes 18.37% G, 23.79% C, 26.84% A and 30.99% T. The A + T bias is 57.84%. Phylogenetic analysis based on 12 complete mitochondrial genomes of sea urchins including four species of the family Diadematidae supported familial monophyly, however the two *Diadema* species, *D. antillarum* and *D. setosum* were not recovered as sister taxa.

1. Introduction

The long-spined black sea urchin, *Diadema antillarum* (Diadematoida, Diadematidae), is a marine benthic invertebrate that inhabits the shallow waters of the Western Atlantic Ocean and Caribbean Sea. It is an important herbivore and keystone species that helps maintain healthy coral reef systems along its coastal marine habitats (Liddell and Ohlhorst, 1986; Hughes et al., 1987; Carpenter, 1988, 1990; Hughes, 1994; Ferrari Legorreta, 2012; Vega Thurber et al., 2012; Burkepile et al., 2013; Steneck, 2020). After the human impact of overharvesting larger herbivorous fishes and the disappearance of larger vertebrates, *D. antillarum* were plentiful. Along with other smaller herbivore fishes they served as the primary grazers maintaining the health of a coral dominated reef system, up until the mid-1980s (Jackson, 2001; Pandolfi et al., 2003; Steneck, 2020). After the historic 1983–84 die-off event of this species presumably due to an unknown water borne pathogen, repeated die-off events occurred in the 1990s as well as in the current year (February - May, 2022, https://www.agrra.org/sea-urchin-die-off/; accessed May 24, 2022). These events have collectively decimated, as well as wiped out populations in some localities across the Caribbean (Lessios, 1995; Lessios, 1988; Lessios, 2016). Monitoring and recovery efforts have been promising, but future die-off events will likely occur as global climate change continues (Hylkema et al., 2022; Pilnick et al., 2021; Williams, 2022).

Until now, only a few genes have been described in the mitochondrial genome of *D. antillarum*, including partial sequences of *COI*, *COII* and *ATP6*, as well as assembled sequences of *ATP8* and *tRNA*^{Lys} (Lessios et al., 2001; Collin et al., 2020). While over 40 mitochondrial genomes of different sea urchin species have been completed so far, this includes only three species within the family Diadematidae, which contain 12 genera. Of these, both species of the genus *Echinothrix* have complete mitochondrial genomes in addition to one of the seven species within the genus *Diadema*, specifically *D. setosum*. In this study, we report the complete mitochondrial genome of *D. antillarum* and we explore the utility of whole mitochondrial genomes to place *D. antillarum* phylogenetically in the Diadematidae tree.

2. Data description

2.1 Sample collection and DNA extraction

An adult sea urchin was collected in Puerto Rico (18°20'35.2"N 67°15'40.5"W). A sample containing whole coelomocytes was withdrawn from the animal through the Aristotle's Lantern using a sterile 23-gauge needle connected to a 5 mL syringe. The animal was photographed and then returned to its habitat at the collection site, thus, it was not retained as a voucher specimen for this study. The sample was held on ice during transport to the University of Puerto Rico Mayagüez (UPRM). Total DNA was extracted from coelomocytes per company instructions using an a DNeasy® Blood and Tissue Kit (Qiagen, Inc.). The sample quality and concentration were assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) prior to shipment for sequencing at the Psomagen (Macrogen USA) laboratory in Rockville, Maryland. Approval for the sample collection was obtained from the Department of Natural and Environmental Resources of Puerto Rico (O-VS-PVS15-AG-00046-01082018).

66 *2.2 Sequencing and bioinformatics analysis*

67 To prepare the DNA sample for sequencing, a library of sequences was generated using a
68 TruSeq® DNA PCR Free (350) library preparation kit (Illumina, Inc.). Quality and quantitation
69 verification of libraries was performed using a Bioanalyzer. Next, an Illumina HiSeq 2500
70 platform generated paired-end 151 bp sequence reads. The sequencer produced 60.8 GB of total
71 reads and 402,874,618 paired-end reads of which 88.21% had a quality score of ≥ Q30.

72 The raw read candidates for the mitochondrial genome were extracted bioinformatically from
73 the total Illumina data that represented both nuclear and mitochondrial genomes using
74 Cookiecutter software (Starostina et al., 2015). The reads were trimmed using a custom program
75 called v2trim. Next, the list of kmers was formed in two stages. In the first stage we used the
76 best-known reference genome of sea urchins (*Strongylocentrotus purpuratus*, NC_001453.1,
77 Jacobs et al., 1988) to assemble the draft mtDNA contigs of *D. antillarum* using SPAdes genome
78 assembler. In stage two, the assembled contigs were searched with BLAST NCBI web interface
79 to find the closest reference sequence, which matched to the species, *Echinothrix diadema*, with
80 the accession number KX385836.1. In the second stage, we constructed the kmer list using this
81 closest reference sequence. The extracted reads were then assembled with SPAdes v3.15.4 using
82 the default parameters (Bankevich et al., 2012). To verify assembly quality the extracted reads
83 were mapped back to the full length mtDNA assembly with bwa mem using the default
84 parameters (Li, 2013). The coverage plot was computed with bedtools genomecov with -d and -
85 split parameters (Aaron et al., 2010). In addition, a per-base coverage data table was generated
86 using the NGS data analysis tools provided in the Unipro UGENE software program
87 (Okonechnikov et al. 2012). A graph of this data table was generated in Microsoft Excel (2019)
88 and is presented as **Supplemental Figure 1**. Finally, the assembly start was rotated to tRNA-

Phe. Following assembly, an online annotation was performed using the MITOS web server (http://mitos.bioinf.uni-leipzig.de/index.py) with following manual verification of each predicted RNA and protein. The results from this annotation were used to generate a mitochondrial map in Geneious Prime v2022.1.1 (**Figure 1**). The extracted reads, bam,coverage files and the reproducible commands list are available in our GitHub repository: https://github.com/aglabx/mtDNA_assembly/tree/master/Diadema_anthilarum.

2.3 Phylogenetic analysis

Alignment and phylogenetic analysis were constructed in MEGA v11.0.11 (Tamura et al. 2021). Analysis included the complete mitochondrial genomes of 12 representative species from the orders Camarodonta (families Temnopleuridae, Echinometridae and Parechinidae), Arbacioida (family Arbaciidae), Temnopleuroida (family Toxopneustidae), Cidaroida (family Cidaridae), Echinoida (family Strongylocentrotidae), and Diadematoida (family Diadematidae) plus the sea cucumber, *Apostichopus* japonicus. A MUSCLE alignment method was implemented using the default parameters in the MEGA11 program. Another sequence alignment editor, BioEdit, was used to generate a sequence identity matrix of the aligned sequences (Hall, T.A., 1999). We performed an IQ-TREE analysis to determine the best-fit model of substitution among the mitogenome sequences (Nguyen et al 2015). A maximum likelihood tree (ML) was generated with the mitogenome nucleotide DNA sequences of the 12 sea urchin species and the sea cucumber that was included as an outgroup. We settled upon a comprehensive model of parameters for the phylogenetic analysis that resulted in branches containing the best supported bootstrap values (out of 500 bootstrap iterations) for the resulting consensus tree. The ML tree was generated using a Tamura-Nei model of evolution (Tamura and Nei, 1993). A discrete Gamma distribution was used to model evolutionary rate differences

among sites (5 categories (+*G*, parameter = 0.3129)). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then the ML tree was generated by selecting the topology with superior log likelihood value. Additional sequence alignments and phylogenetic analysis was performed on available sequences for *Echinothrix* spp. and *Diadema* spp. for each of the three genes: 16S ribosomal RNA (partial sequence), ATPase genes (including ATP synthase subunit 6 gene partial coding sequence and tRNA-Lys gene partial sequence ATP8 gene complete coding sequence and ATP6 gene partial coding sequence) and cytochrome oxidase subunit 1 (CO1) (partial coding sequence). For these additional analyses, a MUSCLE alignment was generated using the default parameters in the MEGA11 program. In addition, we extracted the portions of the mitogenome for *E. diadema* (KX385836), *D. setosum* (KX385835) and *D. antillarum* (present study) that corresponded to each gene, by generating an initial alignment using available sequences from the same genera for each of the three genes. A ML tree was obtained for each gene using a bootstrapped (500 iterations) general time reversible (GTR) model with invariant (I) substitution rates among the nucleotides. The initial trees were obtained by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then the ML tree was generated by selecting the topology with superior log likelihood value.

Results

3.1 Signatures of the mitogenome

The circular mitogenome of *D. antillarum* is 15,708 bp in length, which consists of 2 rRNA genes, 22 tRNA genes, a non-coding control region, and 13 protein-coding genes that are common in other echinoderms, as well as the order of the genes (Bronstein and Kroh 1999,

Ketchum et. al. 2018, De Giorgi et. al. 1996). Most of the genes are encoded on the H-strand except for one protein-coding gene, *ND6*, and five tRNA genes, including *tRNA*^{Gln}, *tRNA*^{Met}, *tRNA*^{Tyr}, *tRNA*^{Trp}, and *tRNA*^{Cys}, which are encoded on the L-strand (Bernt et al. 2013). Most of the protein-coding genes use the start codon ATG, with the exception of the *ATP8* gene, which starts with GTG. The length of the rRNA genes is 896 bp for *12S rRNA* and 1555 bp *16S rRNA*. The control region is 133 bp in length and contains the typical G repeat that is found in other echinoderms. This non-coding region is located at base positions 1111 to 1243, and is positioned in between the genes *tRNA*^{Ile} and *tRNA*^{Gln}.

The composition of the nucleotides for the mitogenome of *D. antillarum* was calculated in the MEGA11 program, and includes 18.37% G, 23.79% C, 26.84% A and 30.99% T. The A + T bias was also calculated in MEGA, which is 57.84% for *D. antillarum*, slightly lower but comparable to that of *D. setosum*, which is 58.19%. In comparison to the two other *Echinothrix* species in the same family as *D. antillarum*, the A + T bias in the mitogenome of *D. antillarum* is slightly higher than that of *E. diadema* (57.61%) and *E. calamaris* (56.43%). Nonetheless, when compared to species in different orders, the mitochondrial A + T bias of *D. antillarum* is generally lower than other species in the following orders: Cidaroida (*Stylocidaris reini*, 59.93%; *Prionocidaris baculosa*, 59.14%; *Eucidaris tribuloides*, 59.70%), Camarodonta (*Echinometra mathaei*, 59.21%; *Heterocentrotus mamillatus*, 58.91%; *Heliocidaris crassispina*, 58.89%), and Echinoida (*Strongylocentrotus droebachiensis*, 58.96%; *S. intermedius*, 58.92%; *S. purpuratus*, 58.98%).

3.2 Phylogenetic tree

The bootstrap consensus ML tree is shown in **Figure 2**, which was inferred from 500 replicates through clustering of the associated taxa (Felsenstein, 1985). The results of the IQ-

159 TREE analysis indicated that the best model of substitution was the most comprehensive model
160 tested, the GTR+F+R2 model (General time reversible model with unequal rates and unequal
161 base frequency with FreeRate rate heterogeneity model) (Tavaŕe 1986). The results of this test
162 included a ML tree lacking a bootstrap analysis, which is available on the GitHub repository site
163 listed above. In addition, a ML tree was generated in the MEGA11 program using the suggested
164 GTR model of evolution with invariant (or unequal) substitution rates among sites (tree not
165 shown). The two resulting ML consensus trees from the MEGA11 program plus the ML tree
166 generated by the IQ-TREE analysis had identical topologies with regards to the species within
167 the family Diadematidae. Of the two bootstrapped ML trees produced in MEGA, the tree
168 presented in this study generally showed higher bootstrap values for all branches within the tree.
169 Interestingly, the tree topology shown in this study indicated that *D. antillarum* was more
170 closely related to *E. diadema* than to *D. setosum*. The next most closely related taxon to these
171 three was *E. calamaris*. For *E. diadema*, *E. calamaris* and *D. setosum*, the length of these
172 mitochondrial genomes are 15,712 bp, 15,716 bp and 15,708 bp, respectively. A previous study
173 reporting the complete mitogenome of *D. setosum* included a maximum likelihood phylogenetic
174 tree with a highly supported sister clade (bootstrap value of 100) containing two taxa: *E.*
175 *diadema* and *D. setosum* (Li et al., 2016). Furthermore, the results of our sequence identity
176 matrix of mitogenomes between *D. antillarum* and *E. diadema* was 96.7% identical, whereas, *D.*
177 *antillarum* vs. *D. setosum* or *D. antillarum* vs. E. *calamaris* was 86.3% and 80% identity,
178 respectively. Additionally, the mitogenomes between *E. diadema* and *D. setosum* was 86.2%
179 identical in sequence. One may suggest that the overall higher similarity between *D. antillarum*
180 and *E. diadema*, as opposed to *D. setosum* may be due to the higher A + T bias in the
181 mitogenome. However, the mitogenome sequences of *E. diadema* and *D. setosum* were

generated from specimens collected from the South China Sea from the same research group (Li, Wu, Fu, and Zeng, 2016 unpublished, Li et. al., 2016). Given that there are other *Diadema* species in the South China Sea that are generally similar morphologically to *E. diadema* (including *D. setosum* and *D. savignyi*) the reported mitogenome of *E. diadema* may have actually been sampled from a *Diadema* species, namely *D. savignyi*. Unfortunately, the mitogenome has not been completed or otherwise available for *D. savignyi*. Nonetheless, to provide evidence for this claim we performed sequence analysis for three additional genes (*16S, ATPase* and *CO1*) for available data from species of the genera *Echinothrix* and *Diadema*. Results of the tree topologies for all three genes indicate that the specific gene sequences extracted from the mitogenome of *E. diadema* were more closely related to the particular gene sequences of *Diadema* spp. as opposed to those from *Echinothrix* spp. (**Supplemental Figures 2, 3, and 4**). In particular, the specific gene sequences that were extracted from the mitogenome of *E. diadema* were placed as a sister clade with *D. savignyi* (see Supplemental Figures 2, 3, and 4), which was placed next to a larger group of clades that included *D. africanum* (see Supplemental Figure 2), *D. antillarum* (see Supplemental Figures 2, 3, and 4) *D. mexicanum* (see Supplemental Figure 2) and *D. setosum* (see Supplemental Figures 2, 3, and 4). This was more distantly related to groupings of clades that included *E. diadema* (see Supplemental Figures 2) and *E. calamaris* (see Supplemental Figures 2, 3, and 4).

Furthermore, it is important to note that the habitat ranges of both *E. diadema* and *D. setosum* are in the Indo-Pacific region which is distinct from the habitat range of *D. antillarum*, which are typically found in the western Atlantic Ocean, the Caribbean Sea, the tropical coasts of South America down to Brazil, and from Bermuda to Florida. The authors are aware that a new subspecies of *D. antillarum*, called *D. antillarum ascensionis* has been recently identified in the

eastern Atlantic, which is similar to the mid-Atlantic species of *D. antillarum* (Rodriguez, et al. 2013). This new subspecies is also similar to *D. africanum* that are found in the Eastern Atlantic islands, from Madeira Islands to the Guinean Gulf including Salvage Islands, Canary Islands, Cape Verde Islands (Hernández et al. 2008), and Sâo Tome Island (Lessios et al. 2001). While *D. antillarum* have been repopulating areas, which have extended their previous habitat range, we are not aware of any *D. africanum* that has migrated as far west as Puerto Rico, in the Caribbean Sea. In addition, when identifying the specimen prior to collection for this study, it lacked the typical iridophores that are identifiable in the sunlight of *D. africanum.* While the *ATPase* sequence analysis presented in this study does reflect that there are similarities in these sequence between *D. antillarum, D. africanum and D. mexicanum,* the *ATPase* gene sequence that was extracted from the mitogenome used in this analysis most closely resembles that of another *D. antillarum* (see Supplemental Figure 2). Thus, the authors think that the species used in this study has been correctly identified. Yet, as additional mitogenome sequences become available, the phylogenetic tree describing the relationships between genera within Diadematidae should become more thorough.

In conclusion, these results may provide some insight into the dispersal and speciation events of the family Diadematidae. However, more data is needed from other genera to draw further conclusions about the evolution and adaptation of this family lineage. The mitogenome sequence produced through this study can serve as a reference sequence for this species of sea urchin. In addition, the nuclear sequences generated in this study will be included in a larger study to assembly the whole genome for this species. Moreover, given the necessity of *D. antillarum* for maintaining the health and current structure of the remaining coral dominated

coastlines in the face of global climate change, the genetics of this species will be likely important to include in future conservation and population genetics studies.

3. Data accessibility

The mitochondrial genome sequence has been deposited at DDBJ/ENA/GenBank under the accession number ON725136 (pending; submitted on May 19, 2022). The data is linked to the NCBI BioProject and BioSample numbers PRJNA839760 and SAMN28553754, respectively. All intermediate files are available in our GitHub repository: https://github.com/aglabx/mtDNA_assembly/tree/master/Diadema_anthilarum.

Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this article.

Acknowledgements

Funding for the project was provided by start-up funds granted to Tarás Oleksyk by Oakland University.

References

1. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V., Nikolenko, S., Pham, S., Prjibelski, A., Pyshkin, A., Sirotkin, A., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P. A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*. 2012. 19(5): 455-477.

247 2. Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J.,
248 Middendorf, M. and Stadler, P. F. MITOS: Improved de novo Metazoan Mitochondrial
249 Genome Annotation. Molecular Phylogenetics and Evolution. 2013. 69(2):313-319
250 3. Bronstein, O. and Kroh, A. The first mitochondrial genome of the model echinoid
251 Lytechinus variegatus and insights into Odontophoran phylogenetics. Genomics. 2018.
252 111 (4) DOI:10.1016/j.ygeno.2018.04.008
253 4. Burkepile, D.E., Allgeier, J.E., Shantz, A.A., Pritchard, C.E., Lemoine, N.P., Bhatti,
254 L.H., and Layman, C.A. Nutrient supply from fishes facilitates macroalgae and
255 suppresses corals in a Caribbean coral reef ecosystem. *Scientific Reports*. 2013. 3(1493):
256 1–9.
257 5. Cantatore,P., Roberti,M., Rainaldi,G., Gadaleta,M.N. and Saccone,C. The complete
258 nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of
259 Paracentrotus lividus. J. Biol. Chem. 264 (19), 10965-10975 (1989)
260 6. Carpenter, R.C. Mass mortality of a Caribbean sea urchin: Immediate effects on
261 community metabolism and other herbivores. *Proceedings of the National Academy of*
262 *Sciences*. 1988. 85: 511–514.
263 7. Carpenter, R.C. Mass mortality of *Diadema antillarum*—I. Long-term effects on sea
264 urchin population dynamics and coral reef algal communities. *Marine Biology*. 1990.104:
265 67–77
266 8. Collin, R., Venera-Ponton, D.E., Driskell, A.Сю, Boyle M.J., Lessios, H.A. DNA
267 barcoding of echinopluteus larvae uncovers cryptic diversity in neotropical echinoids.
268 *Invertebrate Biology*. 2020. 139 (2) https://doi.org/10.1111/ivb.12292

9. De Giorgi, C., Martiradonna, A., Lanave, C. and Saccone, C. Complete sequence of the mitochondrial DNA in the sea urchin *Arbacia lixula*: conserved features of the echinoid mitochondrial genome. Molecular Phylogenetics and Evolution. 1996. 5: 323-332.
10. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*. 1985. 39:783-791.
11. Ferrari Legorreta, R. Building resilience of Caribbean coral reefs to macroalgal phase shifts: Identifying key habitat features. PhD Thesis, School of Biological Sciences, The University of Queensland, Australia. 2012Hughes, T.P. Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. *Science*. 1994. 265: 1547–1551.
12. Hernandez, J.C., Clemente, S., Sangil, C., and Brito, A. 2008. Actual status of the sea urchin *Diadema aff. Antillarum* populations and macroalgal cover in the Marine Protected Areas comparing to a highly fished area (Canary Islands-Eastern Atlantic Ocean). 2008. Aquatic Conservation: Marine and Freshwater Research. 66. 259-270.
13. Hughes, T.P., Reed, D.C., and Boyle, M.-J. Herbivory on coral reefs: Community structure following mass mortalities of sea urchins. *Journal of Experimental Marine Biology and Ecology*. 1987. 113: 39–59
14. Hylkema, A., Debrot, A.O., Pistor, M., Postma, E., Williams, S.M. and Kitson-Walters, K., High peak settlement of *Diadema antillarum* on different artificial collectors in the Eastern Caribbean. *Journal of Experimental Marine Biology and Ecology*. 2022. 549 p.151693.
15. Jackson, J.B.C. What was natural in the coastal oceans? *Proceedings of the National Academy of Sciences*. 2001. 98(10): 5411–5418

16. Jacobs, H. T., Elliott, D. J., Math, V. B., Farquharson, A. Nucleotide sequence and gene organization of sea urchin mitochondrial DNA. *Journal of molecular biology*. 1988. 202(2): 185-217.

17. Ketchum, R.N, DeBiasse, M.B., Ryan, J.F., Burt, J.A. and Reitzel, A.M. The complete mitochondrial genome of the sea urchin, *Echinometra* sp. EZ. Mitochondrial DNA B, Resour. 2018. 3(2) 1225-1227.

18. Laruson,A.J. Rates and relations of mitochondrial genome evolution across the Echinoidea, with special focus on the superfamily Odontophora Ecol Evol 7 (13), 4543-4551 (2017)

19. Lessios, H.A., Kessing, B.D., Pearse, J.S. Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution*. 2001. 55(5):955-75.

20. Lessios, H.A., Kessing, B.D., Robertson, D.R. Massive gene flow across the world's most potent marine biogeographic barrier. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 1998. 265(1396):583-8.

21. Lessios, H.A. *Diadema antillarum* 10 years after mass mortality: Still rare, despite help from a competitor. *Proceedings of the Royal Society B: Biological Sciences, B.* 1995. (259) 331-337.

22. Lessios, H.A. The Great *Diadema antillarum* Die-Off: 30 Years Later. *Annual Review of Marine Science*. 2016. 8:267–83

23. Li, C., Wu, G., Fu, W. and Zeng, X. The complete mitochondrial genome of *Diadema setosum* (Aulodonta: diadematidae). Mitochondrial DNA Part B, Resources. 2016. 1 (1): 873-874

24. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint*. 2013. arXiv: 1303.3997.

25. Liddell, W.D. and Ohlhorst, S.L. Changes in benthic community composition following the mass mortality of *Diadema* at Jamaica. *Journal of Experimental Marine Biology and Ecology*. 1986. 95 (3): 271–278

26. Nguyen, L-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*. 2015. 32(1) 268–274.

27. Okonechnikov, K., Golosova, O., Fursov, M., and the UGENE team. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012. 28(8) 1166-1167.

28. Pandolfi, J.M., Bradbury, R.H, Sala, E., Hughes, T.P., Bjorndal, K.A., Cooke, R.G., McArdle, D., McClenachan, L., Newman, M.J., Paredes, G., and Warner, R.R. Global trajectories of the long-term decline of coral reef ecosystems. *Science*. 2003. 301(5635): 955–958

29. Pilnick, A.R., O'Neil, K.L., Moe, M. and Patterson, J.T. A novel system for intensive *Diadema antillarum* propagation as a step towards population enhancement. *Scientific Reports*. 2021. 11(1), pp.1-13. https://www.nature.com/articles/s41598-021-90564-1.pdf

30. Quinlan, A. R., Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010. 26(6): 841-842.

31. Qureshi,S.A. and Jacobs,H.T. Two distinct, sequence-specific DNA-binding proteins interact inddependently with the major replication pause region of sea urchin mtDNA Nucleic Acids Res. 21 (12), 2801-2808 (1993)

32. Rodriguez A., Hernandez, J.C., Clement, S., Coppard, S.E. A new species of *Diadema* (Echinodermata: Echinoidea: Diadematidae) from the eastern Atlantic Ocean and a neotype designation of *Diadema antillarum* (Philippi, 1845). 2013. Zootaxa 3636 (1) 144-170.

33. Starostina, E., Tamazian, G., Dobrynin, P., O'Brien, S., Komissarov, A. Cookiecutter: a tool for kmer-based read filtering and extraction. *BioRxiv*. 2015. 024679. https://doi.org/10.1101/024679

34. Steneck, R.S. Sea urchins as drivers of shallow water benthic community structure. In Sea Urchins: Biology and Ecology, Lawrence, J.M. (ed.). 2020. pp. 195–212. Academic Press, San Diego, California

35. Sun,X.-J., Li,Q. and Kong,L.-F. Comparative mitochondrial genomics within sea cucumber(Apostichopus japonicus): Provide new insights into relationships among color variants. Aquaculture 309 (1-4), 280-285 (2010)

36. Tamura K. and Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*. 1993. 10:512-526.

37. Tamura, K., Stecher, G., Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*. 2021. 38 (7) 3022-3027.

38. Tavaŕe, S. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences. 1986. 17:57-86

39. Valverde,J.R., Marco,R. and Garesse,R. A conserved heptamer motif for ribosomal RNA transcription termination in animal mitochondria Proc. Natl. Acad. Sci. U.S.A. 91 (12), 5368-5371 (1994)

40. Vega Thurber, R., Burkepile, D.E., Correa, A.M.S., Thurber, A.R, Shantz, A.A., Welsh, R., Pritchard, C., and Rosales, S. Macroalgae decrease growth and alter microbial community structure of the reef-building coral, Porites astreoides. *PLoS One*. 2012. 7(9): e44246

41. Wakayama,N., Kiyono,Y., Matsumoto,N., Saitoh,M. and Kanazawa,K. Effective DNA extraction methods for mitochondrial phylogenomics of the sea urchins. Zoosymposia 15 (1), 192-202 (2019)

42. Williams, S.M., The reduction of harmful algae on Caribbean coral reefs through the reintroduction of a keystone herbivore, the long□spined sea urchin *Diadema antillarum*. Restoration Ecology. 2022. 30(1), p.e13475.

The first complete mitochondrial genome of *Diadema antillarum* (Diadematoida, Diadematidae)

Audrey J. Majeske[a,b], Alejandro J. Mercado Capote[a], Aleksey Komissarov[c], Anna Bogdanova[c], Nikolaos V. Schizas[d], Stephanie O. Castro Márquez[a], Kenneth Hilkert[a], Walter Wolfsberger[a], and Tarás K. Oleksyk[a,b,*]

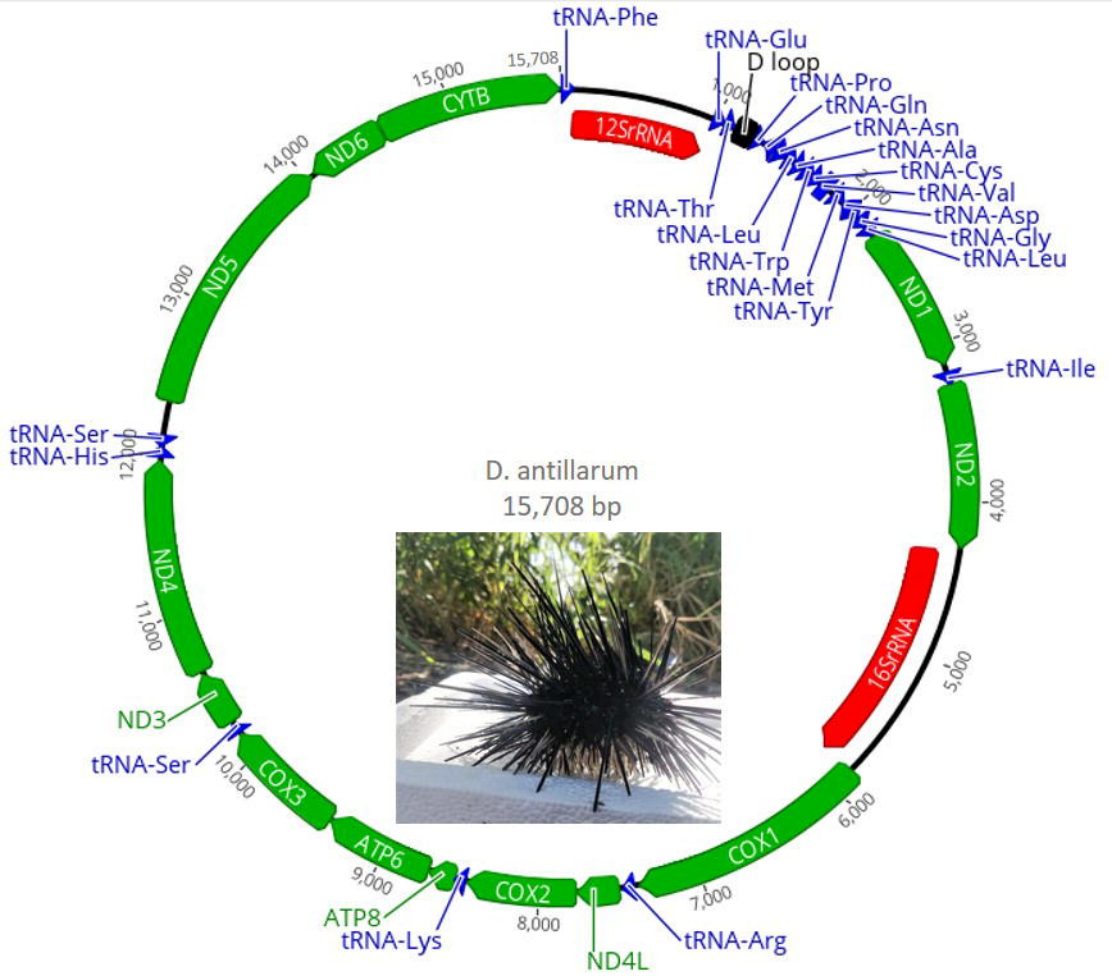
[a]University of Puerto Rico at Mayagüez, Department of Biology, Call Box 9000, Mayagüez, PR 00681, U.S.A.

[b]Oakland University, Department of Biological Sciences, 118 Library Drive, Rochester, MI 48309, U.S.A.

[c]Applied Genomics Laboratory, SCAMT Institute, ITMO University, Lomonosova 9 Str., Saint Petersburg 197101, Russia

[d]University of Puerto Rico at Mayagüez, Department of Marine Sciences, Call Box 9000, Mayagüez, PR 00681, U.S.A.

Figure 1. Mitochondrial genome map of *Diadema antillarum*. Arrow shapes are colored to indicate tRNA genes (blue), rRNA genes (red) and protein coding genes (green). The black rectangular shape refers to the non-coding control region or D loop. The directions of the arrows indicate direction of transcription on the H-strand (arrow points to the right) and L-strand (arrow points to the left). The map was generated in Geneious Prime v2022.1.1. The animal image is the specimen used for sampling in this study. Image taken by Alejandro Mercado Capote.


tRNA-Phe
15,708
tRNA-Glu
D loop
tRNA-Pro
tRNA-Gln
tRNA-Asn
tRNA-Ala
tRNA-Cys
tRNA-Val
tRNA-Asp
tRNA-Gly
tRNA-Leu
tRNA-Ile
15,000
CYTB
12SrRNA
1,000
2,000
tRNA-Thr
tRNA-Leu
tRNA-Trp
tRNA-Met
tRNA-Tyr
ND1
3,000
14,000
ND6
13,000
ND5
tRNA-Ser
tRNA-His
12,000
D. antillarum
15,708 bp
ND2
4,000
ND4
11,000
16SrRNA
5,000
ND3
tRNA-Ser
10,000
COX3
ATP6
9,000
COX1
6,000
ATP8
tRNA-Lys
COX2
8,000
7,000
ND4L
tRNA-Arg

The first complete mitochondrial genome of *Diadema antillarum* (Diadematoida, Diadematidae)

Audrey J. Majeske[a,b], Alejandro J. Mercado Capote[a], Aleksey Komissarov[c], Anna Bogdanova[c], Nikolaos V. Schizas[d], Stephanie O. Castro Márquez[a], Kenneth Hilkert[a], Walter Wolfsberger[a], and Tarás K. Oleksyk[a,b,*]

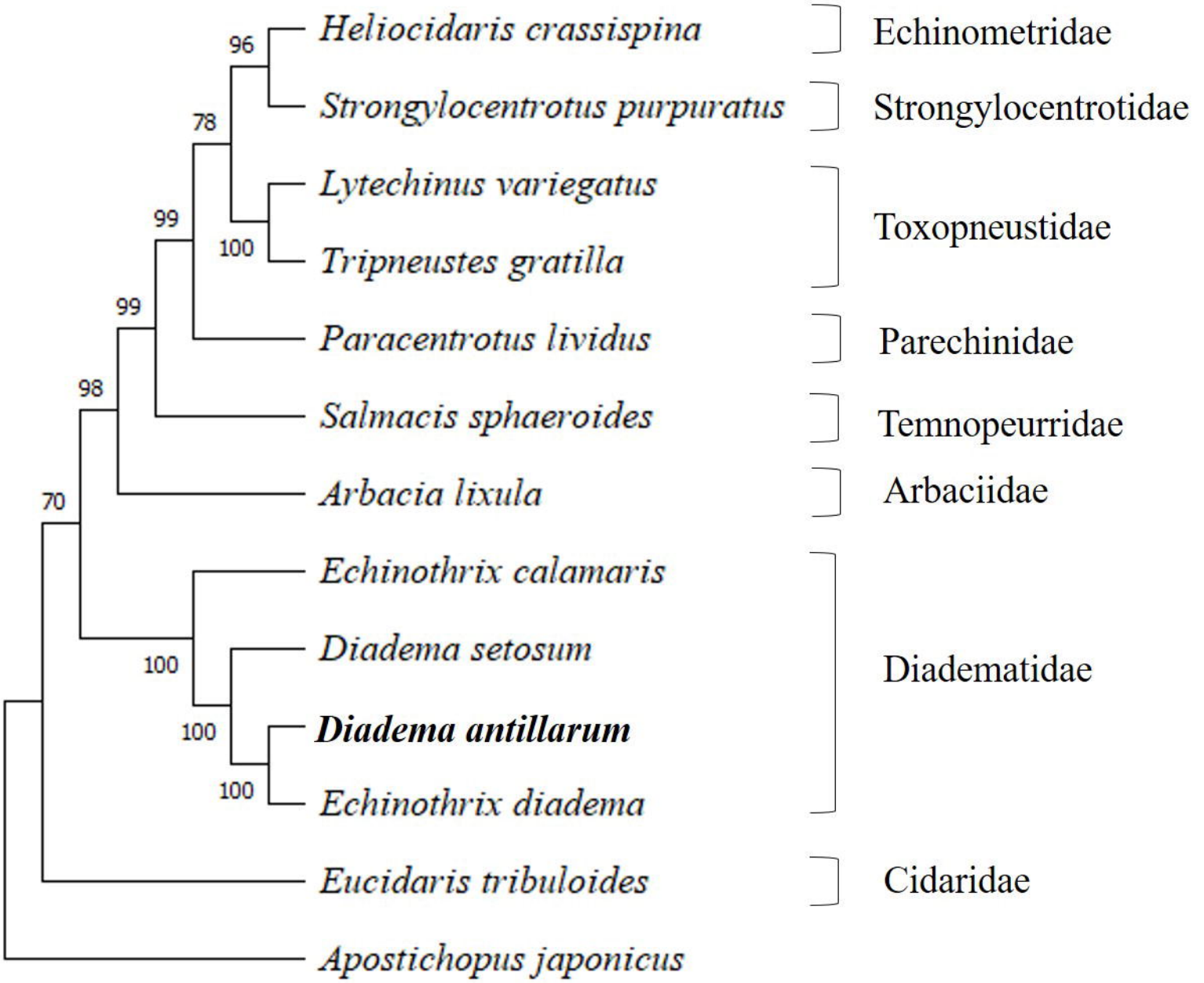
[a]University of Puerto Rico at Mayagüez, Department of Biology, Call Box 9000, Mayagüez, PR 00681, U.S.A.

[b]Oakland University, Department of Biological Sciences, 118 Library Drive, Rochester, MI 48309, U.S.A.

[c]Applied Genomics Laboratory, SCAMT Institute, ITMO University, Lomonosova 9 Str., Saint Petersburg 197101, Russia

[d]University of Puerto Rico at Mayagüez, Department of Marine Sciences, Call Box 9000, Mayagüez, PR 00681, U.S.A.

Figure 2. Phylogenetic analysis of mitochondrial genomes indicates that *Diadema antillarum* is more closely related to *Echinothrix diadema* than to *D. setosum*. A maximum likelihood consensus tree of sequences representing 12 sea urchin species and a sea cucumber (outgroup). Genbank accession numbers for taxa on the tree include: *H. crassispina* (KC479025.1, Jung, G. and Lee, Y.H., 2013, unpublished, direct submission), *S. purpuratus* (NC_001453.1, Valverde et al., 1996, Qureshi, S.A., and Jacobs, H.T. 1993, Jacobs et al., 1988), *L. variegatus* (NC_037785.1, Bronstein and Kroh 2018), *T. gratilla* (KY268294.1, Laruson 2017), *P. lividus* (J04815.1, Cantatore et al. 1989), *S. sphaeroides* (KU302103.1, Jung, G. and Lee, Y.H., 2013, unpublished, direct submission), *A. lixula* (X80396.1, De Giorgi et al., 1996), *E. calamaris* (NC_050274.1, Wakayama et al., 2019), *D. setosum* (KX385835.1, Li et al., 2016), *E. diadema* (KX385836.1, Li et al., 2016, unpublished direct submission), *E. tribuloides* (MH614962.1, Kroh A., and Bronstein, O., 2018, unpublished direct submission) and *A. japonicus* (NC_012616, Sun et al, 2010).


Heliocidaris crassispina
Strongylocentrotus purpuratus
Lytechinus variegatus
Tripneustes gratilla
Paracentrotus lividus
Salmacis sphaeroides
Arbacia lixula
Echinothrix calamaris
Diadema setosum
Diadema antillarum
Echinothrix diadema
Eucidaris tribuloides
Apostichopus japonicus
96
78
99
100
99
98
70
100
100
100
Echinometridae
Strongylocentrotidae
Toxopneustidae
Parechinidae
Temnopeurridae
Arbaciidae
Diadematidae
Cidaridae