GigaByte

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

DRR-202205-01 | Data Release

Submitted by: Audrey Majeske, Alejandro Castro Marquez, Aleksey Komissarov, Anna Bogdanova, Nikolaos Schizas, Stephanie Castro Marquez, Kenneth Hilkert, Walter Wolfsberger, Taras Oleksyk

Keywords: MITOGENOME, ECHINODERM, INVERTEBRATE, PHYLOGENY, BLACK SEA URCHIN, LONG-SPINED SEA URCHIN

PDF auto-generated using **ReView**
from
RIVER VALLEY
TECHNOLOGIES

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 next generation sequencing technology. The complete mitogenome is 15,707 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 freely living 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, and 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 that 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) have decimated and in some localities wiped out populations 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 and causes for these occurrences remain unknown (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 complete 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 plus 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 western Puerto Rico. A sample containing coelomocytes was withdrawn from the animal through the Aristotle's Lantern which 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).

2.2 Sequencing and bioinformatics analysis

To prepare the DNA sample for sequencing, a library of sequences was generated using a TruSeq® DNA PCR Free library preparation kit (Illumina, Inc.). Following quality and quantitation analysis of libraries, paired-end 150 bp sequence reads were generated on an Illumina HiSeq next generation sequencing platform. For this sample, the sequencer produced 60.8 GB of total reads and 402,874,618 paired-end reads of which 88.21% had a quality score of ≥ Q30.

The mitochondrial genome was extracted bioinformatically from the genomic sequences using a unique analysis pipeline. A custom approach was implemented with a supervised reference-assisted de Bruijn graph walkthrough that was performed in g language. First, Jellyfish 2 (Marcais et al., 2012) was used to compute 23-mer frequencies for the generated Illumina reads. Next, the k-mers were mapped to another sea urchin genome, *Strongylocentrotus purpuratus*, that provided the most comprehensive and highest quality assembly available to use as a reference (GenBank accession number: GCF_000002235.5). This was implemented with a Python script called kmer_cov_for_fasta.py. The first 23-mer reads with an exact match to the 12S rRNA genes in the *S. purpuratus* mitogenome was then used as a starting point for extension in both directions. Next, Python scripts called kmer_continue_left.py and kmer_continue_right.py extended a given k-mer with respect to the difference in autosomal and mitogenomic 23-mer coverages of the reads. These Python scripts are publicly available in the GitHub repository (https://github. com/Applied-Genomics-Laboratory/mtDNA_assembly). Following the implementation of these scripts, the resulting full-length mitochondrial genome sequence was generated for this sample, and this assembly was verified using the Blasr tool (Chaisson et al., 2012). Following assembly, an online annotation was performed using the MITOS web server (http://mitos.bioinf.uni-leipzig.de/index.py). The results from this annotation were used to generate a mitochondrial map in Geneious Prime v2022.1.1 (**Figure 1**).

The circular mitogenome of *D. antillarum* is 15,707 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 other vertebrates in general. The gene order is identical to those of other echinoderms. Most of the genes are encoded on the H-strand except for one protein-coding gene, *ND6*, and five tRNA genes, including $tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Val}$, $tRNA^{Asp}$, and $tRNA^{Ser}$, which is encoded on the L-strand. 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 $tRNA^{Thr}$ and $tRNA^{Pro}$.

The composition of the nucleotides for the mitogenome of *D. antillarum* includes 18.37% G, 23.79% C, 26.84% A and 30.99% T. The A + T bias is 57.84%, which is slightly lower but comparable to that of *D. setosum* (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%).

2.3 Phylogenetic analysis

Alignment and phylogenetic analysis were constructed in MEGA v11.0.11 (Tamura et al. 2021). Analyses was completed for the whole 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 implement using the default parameters in the MEGA11 program. Next, a maximum likelihood analysis was completed on the 12 sea urchin species and the sea cucumber that was included as an outgroup. The 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)). This analysis involved 13 nucleotide sequences. 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 selecting the topology with superior log likelihood value. The bootstrap consensus tree is shown in **Figure 2**, which was inferred from 500 replicates through clustering of the associated taxa (Felsenstein, 1985). Interestingly, the tree topology indicated that *D. antillarum* was more closely related to *E. diadema* than to *D. setosum*. The next most closely related taxon to these three was *E. calamaris*. For *E. diadema*, *E. calamaris* and *D. setosum*, the length of the mitochondrial genomes are 15,712 bp, 15,716 bp and 15,708 bp, respectively. The overall higher similarity between *D. antillarum* and *E. diadema*, as opposed to *D. setosum* may be due to the higher A + T bias in the mitogenome. Given these findings, 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*.

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. 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 (pending; submitted on May 19, 2022). The data is linked to the NCBI BioProject and BioSample numbers PRJNA839760 and SAMN28553754, respectively.

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

Burkepile, D.E., Allgeier, J.E., Shantz, A.A., Pritchard, C.E., Lemoine, N.P., Bhatti, L.H., and Layman, C.A. Nutrient supply from fishes facilitates macroalgae and suppresses corals in a Caribbean coral reef ecosystem. Scientific Reports. 2013. 3(1493): 1–9.

Carpenter, R.C. Mass mortality of a Caribbean sea urchin: Immediate effects on community metabolism and other herbivores. Proceedings of the National Academy of Sciences. 1988. 85: 511–514.

Carpenter, R.C. Mass mortality of *Diadema antillarum*—I. Long-term effects on sea urchin population dynamics and coral reef algal communities. Marine Biology. 1990.104: 67–77

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. 2012.

Hughes, T.P. Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. 1994. Science 265: 1547–1551.

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

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

Lessios HA, Kessing BD, Robertson DR. 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.

Lessios HA, Kessing BD, Pearse JS. Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. Evolution. 2001. 55(5):955-75.

Lessios HA. *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.

Lessios HA. The Great *Diadema antillarum* Die-Off: 30 Years Later. Annu. Rev. Mar. Sci. 2016. 8:267–83

Collin R, Venera-Ponton DE, Driskell AC, Boyle MJ and Lessios HA. DNA barcoding of echinopluteus larvae uncovers cryptic diversity in neotropical echinoids. Invertebrate Biology. 2020. 139 (2) https://doi.org/10.1111/ivb.12292

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

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.

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

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.

Jackson, J.B.C. What was natural in the coastal oceans? Proceedings of the National Academy of Sciences. 2001. 98(10): 5411–5418

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

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

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.

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

Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 1985. 39:783-791.

Marcais, G.; Kingsford, C. Jellyfish: A Fast k-Mer Counter, version 1.1.4. Tutor. E Manuais: 2012. Available online: http: //www.cbcb.umd.edu/software/jellyfish/ (accessed on 5 April 2019)

Chaisson, M.J.; Tesler, G. Mapping Single Molecule Sequencing Reads Using Basic Local Alignment with Successive Refinement (BLASR): Application and Theory. BMC Bioinformatics. 2012. 13:238.

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,*]

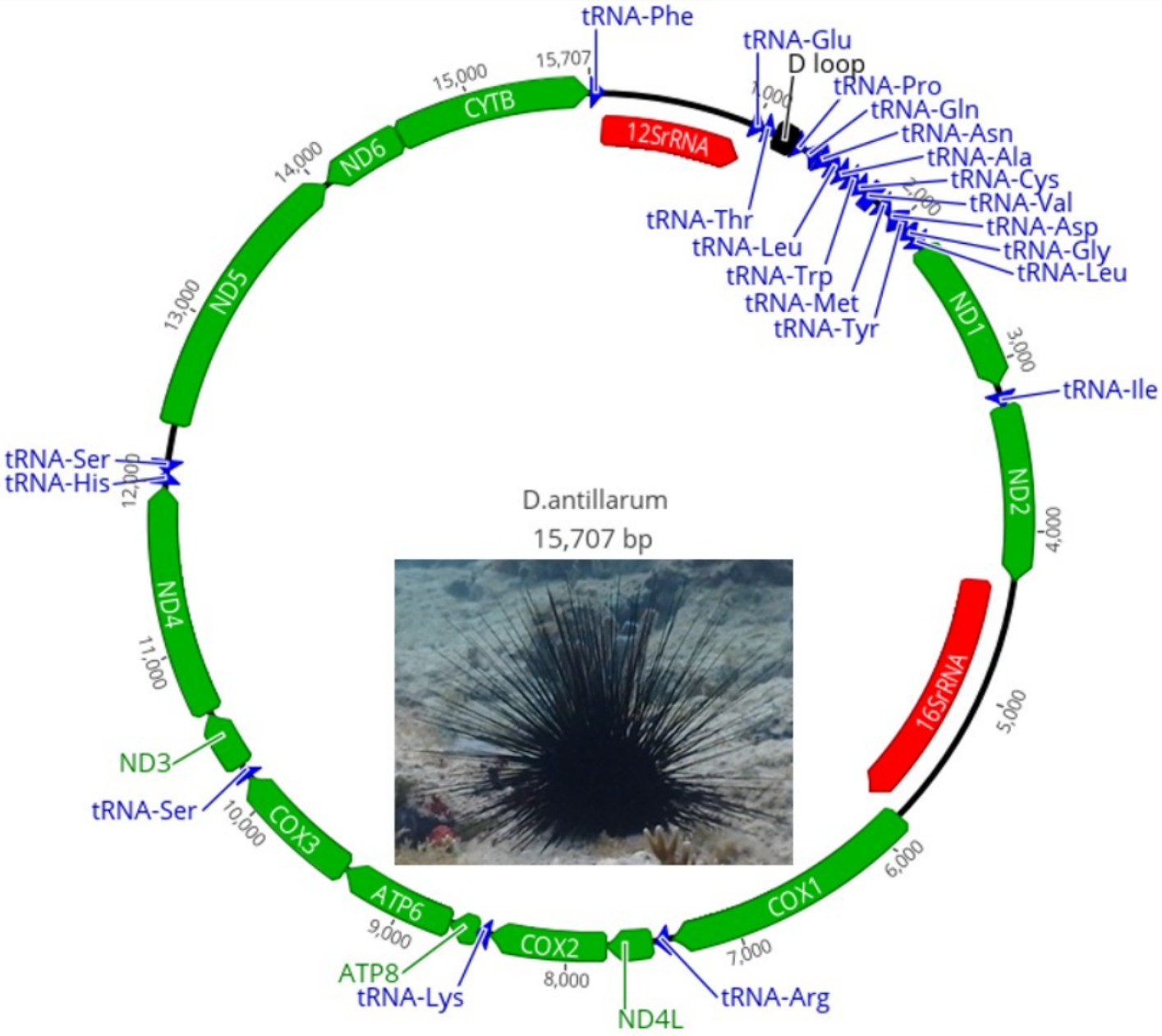
[a]University of Puerto Rico at Mayagüez, Department of Biology, Call Box 9000, Mayagüez, PR 00681, U.S.A.

[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.


tRNA-Phe
tRNA-Glu
D loop
tRNA-Pro
tRNA-Gln
tRNA-Asn
tRNA-Ala
tRNA-Cys
tRNA-Val
tRNA-Asp
tRNA-Gly
tRNA-Leu
tRNA-Thr
tRNA-Leu
tRNA-Trp
tRNA-Met
tRNA-Tyr
12SrRNA
15,707
15,000
CYTB
14,000
ND6
13,000
ND5
ND1
3,000
tRNA-Ile
ND2
4,000
16SrRNA
5,000
6,000
COX1
7,000
tRNA-Arg
ND4L
COX2
8,000
ATP8
tRNA-Lys
ATP6
9,000
COX3
10,000
ND3
tRNA-Ser
ND4
11,000
tRNA-Ser
tRNA-His
12,000
D.antillarum
15,707 bp

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,*]

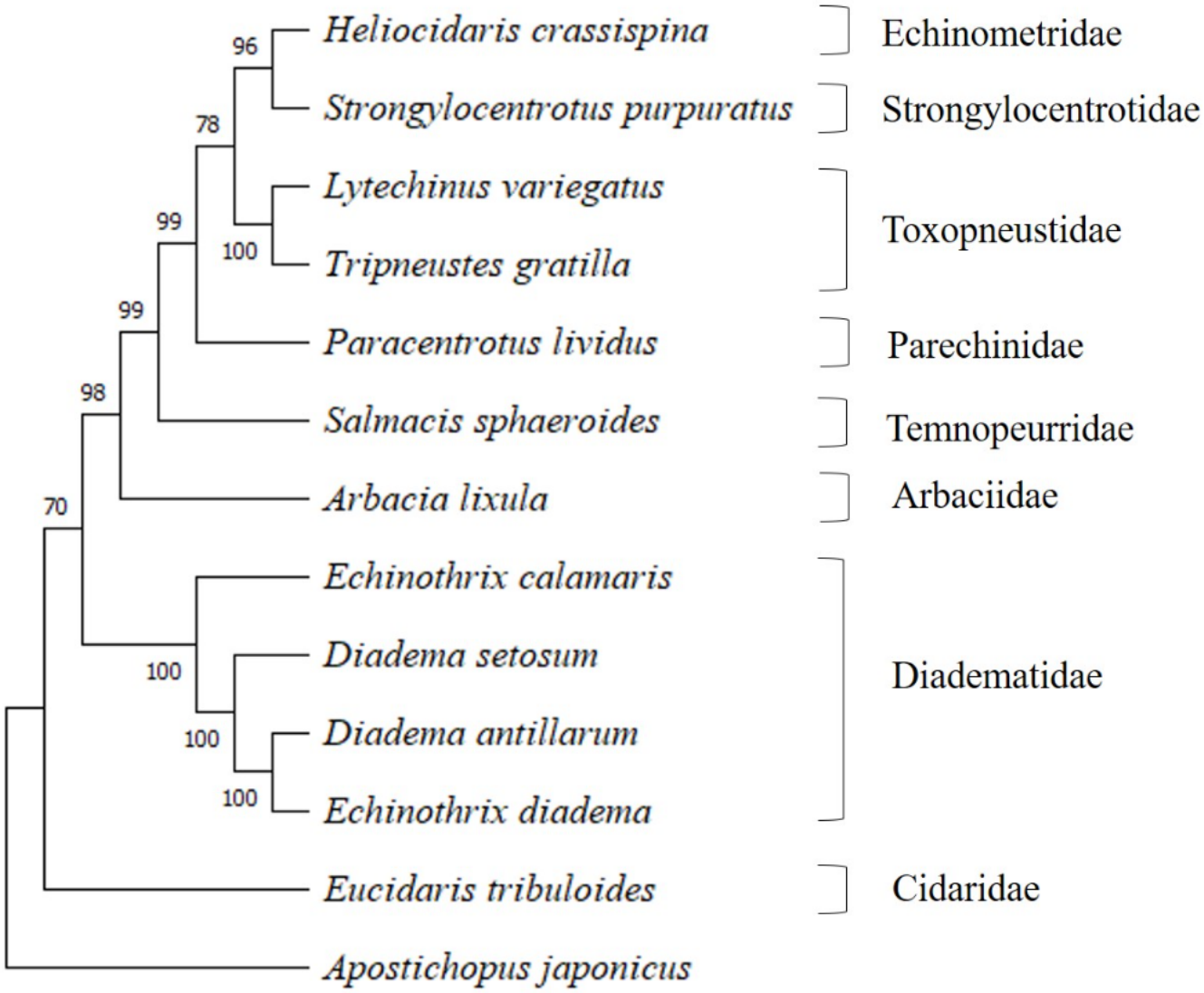
[a]University of Puerto Rico at Mayagüez, Department of Biology, Call Box 9000, Mayagüez, PR 00681, U.S.A.

[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), *S. purpuratus* (NC_001453.1), *L. variegatus* (NC_037785.1), *T. gratilla* (KY268294.1), *P. lividus* (JO4815.1), *S. sphaeroides* (KU302103.1), *A. lixula* (X80396.1), *E. calamaris* (NC_050274.1), *D. setosum* (KX385835.1), *E. diadema* (KX385836.1), *E. tribuloides* (MH614962.1) and *A. japonicus* (NC_012616).


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
Echinometridae
Strongylocentrotidae
Toxopneustidae
Parechinidae
Temnopeurridae
Arbaciidae
Diadematidae
Cidaridae
96
78
99
100
99
98
70
100
100
100