Roa-Varon, Adela, and Stacy Pirro. 2022. "The Complete Genome Sequences of 18 Species of Gadiformes (Actinopterygii, Chordata)." *Biodiversity Genomes*, November.

GENOME SEQUENCING

The Complete Genome Sequences of 18 Species of Gadiformes (Actinopterygii, Chordata)

Adela Roa-Varon[1], Stacy Pirro[2]

[1] Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, [2] Iridian Genomes

Keywords: gadiformes, genomes

https://doi.org/10.56179/001c.56079

Biodiversity Genomes

We present the complete genome sequences of 18 species of Gadiformes from 9 genera. Illumina sequencing was performed on genetic material from single wild-caught individuals. The reads were assembled using a *de novo* method followed by a finishing step. The raw and assembled data is publicly available via Genbank.

Introduction

Gadiformes have been characterized as a monophyletic group within the supra-ordinal taxon Paracanthopterygii (Greenwood et al. 1966). Within the order different authors recognize between 11 and 17 families, about 84 genera, and more than 613 species based on morphological (Endo 2002) and molecular (Roa-Varón and Ortí 2009; Roa-Varón et al. 2021) data. They are distributed from the Arctic to Antarctic oceans, and occupy deep-sea to shallow marine waters, with a single fully freshwater species (Roa-Varón et al. 2021). The gadiforms include some of the most important commercially harvested fishes in the world (e.g., Alaskan Pollock, Atlantic and Pacific cods, Blue Whiting, Hake), accounting for more than a fifth of the world's catch of marine fishes (FAO 2016).

We present the assembled genome sequences from 18 species of Gadiformes. Tissue samples were obtained from museum specimens representing wild-caught individuals. More detailed information about the specimen used for each species can be found in the Biosample linked to each genome assembly in Genbank.

Methods

DNA extraction was performed using the Qiagen DNAeasy genomic extraction kit using the standard process. A paired-end sequencing library was constructed using the Illumina TruSeq kit according to the manufacturer's instructions. The library was sequenced on an Illumina Hi-Seq platform in paired-end, 2 × 150 bp format. The resulting fastq files were trimmed of adapter/primer sequence and low-quality regions with Trimmomatic v0.33 (Bolger, Lohse, and Usadel 2014). The trimmed sequence was assembled by SPAdes v2.5 (Bankevich et al. 2012) followed by a finishing step using Zanfona (Kieras, O'Neill, and Pirro 2021).

Results and Data Availability

All data, including raw reads and assembled genome sequence, is available via Genbank.

Eleginus gracilis	SRR12199771	JAOSLM000000000
Enchelyopus cimbrius	SRR12200187	JAOPJJ000000000
Gadus macrocephalus	SRR12200186	JAOSYZ000000000
Gaidropsarus macrophthalmus	SRR12200220	JAOTIX000000000
Merluccius bilinearis	SRR12199800	JAOEJO000000000
Merluccius hubbsi	SRR16638530	JAOEJN000000000
Merluccius paradoxus	SRR12200188	JAOSLR000000000
Merluccius polylepis	SRR12200221	JAOPJH000000000
Merluccius productus	SRR12200189	JAOPJI000000000
Molva macrophthalma	SRR12200191	JAOPJO000000000
Phycis chesteri	SRR12200192	JAOPJK000000000
Trisopterus esmarkii	SRR12200185	JAOPJL000000000
Trisopterus luscus	SRR12199799	JAOPUA000000000
Trisopterus minutus	SRR16369795	JAOPUD000000000
Urophycis chuss	SRR12200190	JAOYML000000000
Urophycis cirrata	SRR12200217	JAOPUH000000000
Urophycis floridana	SRR12200218	JAOPUI000000000
Urophycis regia	SRR12200219	JAOQGG000000000

……

Funding

Funding was provided by Iridian Genomes, grant# IRGEN_RG_2021-1345 Genomic Studies of Eukaryotic Taxa.

Submitted: November 21, 2022 EST, Accepted: November 21, 2022 EST

REFERENCES

Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, et al. 2012. “SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.” *Journal of Computational Biology* 19 (5): 455–77. https://doi.org/10.1089/cmb.2012.0021.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Endo, H. 2002. “Phylogeny of the Order Gadiformes (Teleostei, Paracanthopterygii).” *Mem. Grad Sch. Fish. Sci. Hokkaido Univ* 49: 75–149.

FAO. 2016. *The State of World Fisheries and Aquaculture. Contributing to Food Security and Nutrition for All*. Rome.

Greenwood, P.H., D.E. Rosen, S.H. Weitzman, and G.S. Myers. 1966. “Phyletic Studies of Teleostean Fishes, with a Provisional Classification of Living Forms.” *Bull. Am. Mus. Nat. Hist* 131: 341–455.

Kieras, M., K. O’Neill, and S. Pirro. 2021. *Zanfona, a genome assembly finishing tool for paired-end Illumina reads*. https://github.com/zanfona734/zanfona.

Roa-Varón, Adela, Rebecca B. Dikow, Giorgio Carnevale, Luke Tornabene, Carole C. Baldwin, Chenhong Li, and Eric J. Hilton. 2021. “Confronting Sources of Systematic Error to Resolve Historically Contentious Relationships: A Case Study Using Gadiform Fishes (Teleostei, Paracanthopterygii, Gadiformes).” *Systematic Biology* 70 (4): 739–55. https://doi.org/10.1093/sysbio/syaa095.

Roa-Varón, Adela, and Guillermo Ortí. 2009. “Phylogenetic Relationships among Families of Gadiformes (Teleostei, Paracanthopterygii) Based on Nuclear and Mitochondrial Data.” *Molecular Phylogenetics and Evolution* 52 (3): 688–704. https://doi.org/10.1016/j.ympev.2009.03.020.