

The genome sequence of the common sole, *Solea solea*

Volckaert Filip¹, Geslain Enora², Mc Cartney Ann³, Formenti Giulio⁴, Mouton Alice⁵, Wellcome Sanger Institute Tree of Life Management⁶, Wellcome Sanger Institute Scientific Operations: Sequencing Operations⁷, Wellcome Sanger Institute Tree of Life Core Informatics Team⁸, Tree of Life Core Informatics Collective⁹ and Darwin Tree of Life Consortium¹⁰

- ¹ Biology, Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven
E-mail: filip.volckaert@kuleuven.be
- ² Biology, KU Leuven, Charles Deberiotstraat 32, B-3000 Leuven, Belgium
- ³ Genomics Institute, University of California, Santa Cruz, California, Santa Cruz, CA 95060, USA
- ⁴ The Rockefeller University, The Vertebrate Genome Laboratory, New York, NY 10065, USA
- ⁵ InBios-Conservation Genetics Laboratory, University of Liège, B-4000, Liège, Belgium
- ⁶ <https://doi.org/10.5281/zenodo.12162482>, Wellcome Sanger Institute, Cambridge, UK
- ⁷ <https://doi.org/10.5281/zenodo.12165051>, Wellcome Sanger Institute, Cambridge, UK
- ⁸ <https://doi.org/10.5281/zenodo.12160324>, Wellcome Sanger Institute, Cambridge, UK
- ⁹ <https://doi.org/10.5281/zenodo.12205391>, Tree of Life, Cambridge, UK
- ¹⁰ <https://doi.org/10.5281/zenodo.4783558>, Darwin Tree of Life, Cambridge, UK

We present a genome assembly from an individual female *Solea solea* (Linnaeus, 1758) (the common sole; Chordata; Actinopterygii; Pleuronectiformes; Soleidae). The genome sequence spans 643.80 megabases. Most of the assembly (97.81 %) is scaffolded into 21 chromosomal pseudomolecules. The mitochondrial genome has also been assembled and is 17.03 kilobases in length. Gene annotation of this assembly on Ensembl identified 21,646 protein-coding genes. We compare the common sole genome with Senegal sole *Solea senegalensis* and thickback sole *Microchirus variegatus* in a synteny, structural variation and phylogenetic analysis. This reference genome represents a valuable resource to understand the biology of common sole and teleost fishes, support conservation, fisheries management and a budding aquaculture industry.

Keywords

Evolution; Pleuronectiformes; Reference Genome; Structural Variation