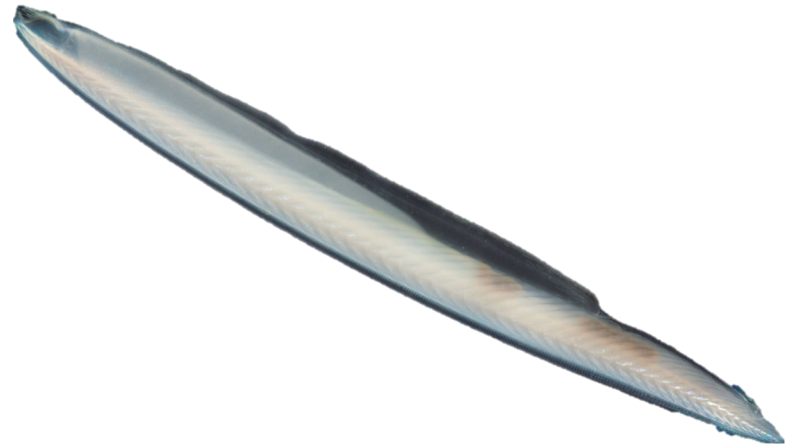


The Genome of Lancelets

Abdulatif Alfulaij

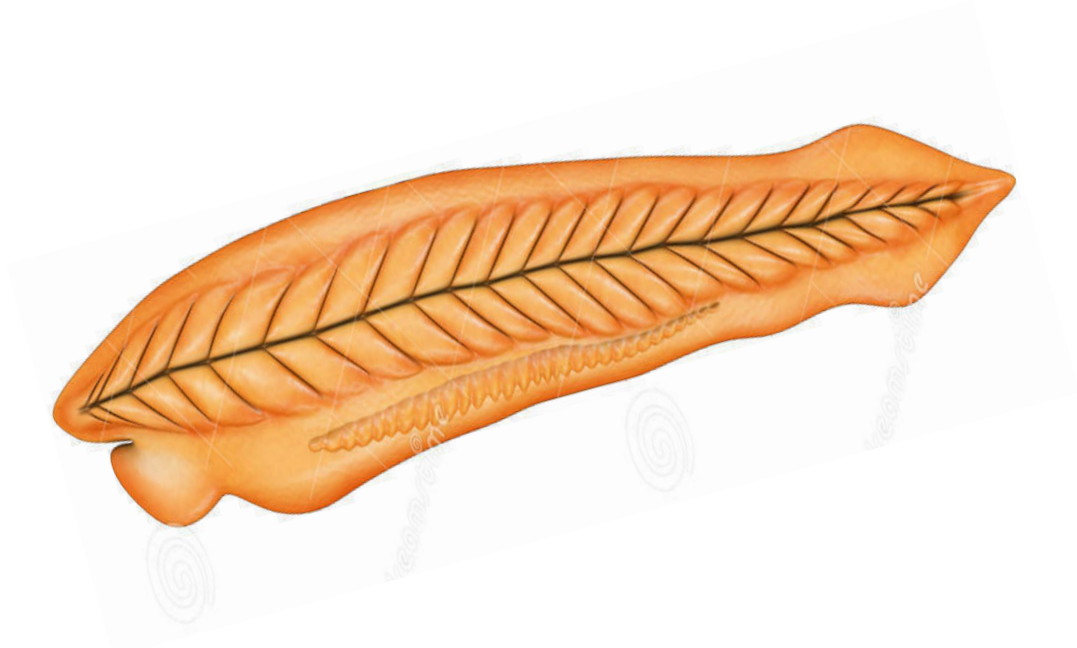
Intro. To Genomics 485

Fall 2020



Oultine

- General information
- Genome paper
- General genome info
- Sequencing strategy
- Sequencing methods
- Genome assembly
- Genome outcome



General information

Small fish less than 5cm
in length



General information

Live in sandy bottom



General information

members of the invertebrate

found widely in the coastal waters
of the warmer parts

General information

Lancelets or **amphioxus**, translates to
“both ends pointed,

ARTICLES

The amphioxus genome and the evolution of the chordate karyotype

Nicholas H. Putnam^{1,2}, Thomas Butts³, David E. K. Ferrier⁴, Rebecca F. Furlong³, Uffe Hellsten¹, Takeshi Kawashima^{2,†}, Marc Robinson-Rechavi^{5,6}, Eiichi Shoguchi^{7,†}, Astrid Terry¹, Jr-Kai Yu⁸, Èlia Benito-Gutiérrez⁹, Inna Dubchak¹, Jordi Garcia-Fernández¹⁰, Jeremy J. Gibson-Brown^{11,†}, Igor V. Grigoriev¹, Amy C. Horton^{11,†}, Pieter J. de Jong¹², Jerzy Jurka¹³, Vladimir V. Kapitonov¹³, Yuji Kohara¹⁴, Yoko Kuroki¹⁵, Erika Lindquist¹, Susan Lucas¹, Kazutoyo Osoegawa¹², Len A. Pennacchio¹, Asaf A. Salamov¹, Yutaka Satou⁷, Tatjana Sauka-Spengler⁶, Jeremy Schmutz¹⁶, Tadasu Shin-Itani¹⁴, Atsushi Toyoda¹⁵, Marianne Bronner-Fraser⁸, Asao Fujiyama^{15,17}, Linda Z. Holland¹⁸, Peter W. H. Holland³, Nori Satoh^{2,†} & Daniel S. Rokhsar^{1,2}

Lancelets ('amphioxus') are the modern survivors of an ancient chordate lineage, with a fossil record dating back to the Cambrian period. Here we describe the structure and gene content of the highly polymorphic ~520-megabase genome of the Florida lancelet *Branchiostoma floridae*, and analyse it in the context of chordate evolution. Whole-genome comparisons illuminate the murky relationships among the three chordate groups (tunicates, lancelets and vertebrates), and allow not only reconstruction of the gene complement of the last common chordate ancestor but also partial reconstruction of its genomic organization, as well as a description of two genome-wide duplications and subsequent reorganizations in the vertebrate lineage. These genome-scale events shaped the vertebrate genome and provided additional genetic variation for exploitation during vertebrate evolution.

Lancelets, or amphioxus, are small worm-like marine animals that spend most of their lives buried in the sea floor, filter-feeding through jawless, ciliated mouths. The vertebrate affinities of these modest creatures were first noted in the early part of the nineteenth century^{1,2}, and were further clarified by the embryologist Alexander Kowalevsky³. In particular, Kowalevsky observed that, unlike other invertebrates, amphioxus shares key anatomical and developmental features with vertebrates and tunicates (also known as urochordates). These include a hollow dorsal neural tube, a notochord, a perforated pharyngeal region, a segmented body musculature (embryologically derived from somites) and a post-anal tail. Together, the vertebrates, urochordates and lancelets (also known as cephalochordates) constitute the phylum Chordata, descended from a last common ancestor that lived perhaps 550 million years ago.

Although Kowalevsky, Darwin and others recognized the evolutionary relationship between chordate groups, the greater morphological, physiological and neural complexity of vertebrates posed a puzzle: how did the chordate ancestor—presumably a simple creature that resembled a modern amphioxus or ascidian larva—make such a transition?

Perhaps the most prevalent hypothesis for the origins of vertebrate complexity is founded on the ideas of Susumu Ohno (1970)⁴, who proposed that vertebrate genomes were shaped by a series of ancient

genome-wide duplications. In Ohno's original proposal, lancelet and vertebrate genomes were enlarged relative to the basic invertebrate complement by one or two rounds of genome doubling, although subsequent work suggested that these events occurred on the vertebrate stem after divergence of the lancelet lineage^{5,6}.

Although the sequencing of the human and other vertebrate genomes has shown that the gene number in vertebrates is comparable to, or only modestly greater than, that of invertebrates^{7,8}, evidence for large-scale segmental or whole-genome duplications on the vertebrate stem has mounted, with the parallel realization that most gene duplicates from such events are rapidly lost (reviewed in ref. 9). The relatively few surviving gene duplicates from the vertebrate stem provide evidence for ancient paralogous relationships between groups of human chromosomes^{10–14} that plausibly arose from multiple rounds of whole-genome duplication before the emergence of modern vertebrates. However, the number, the timing and even the genomic scale of the duplication events, and their consequences for subsequent genome evolution, are poorly understood (for a review, see ref. 15), in part because the tunicate genomes are highly rearranged relative to the unduplicated early chordate karyotype (see below).

The Florida lancelet *B. floridae* (the generic name *Branchiostoma* refers to the characteristic perforated branchial arches) provides a critical point of reference for these studies¹⁶. This species and its

¹Department of Energy Joint Genome Institute, Walnut Creek California 94598, USA. ²Center for Integrative Genomics, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA. ³Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK. ⁴The Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife KY16 9LB, UK. ⁵Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland. ⁶Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. ⁷Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan. ⁸Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. ⁹National Institute for Medical Research, Mill Hill, London NW7 1AA, UK. ¹⁰Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avinguda Diagonal, 645, Barcelona 08028, Spain. ¹¹Department of Biology, Washington University in St Louis, St Louis, Missouri 63130, USA. ¹²Children's Hospital of Oakland Research Institute, Oakland, California 94609, USA. ¹³Genetic Information Research Institute, 1925 Landings Drive, Mountain View, California 94043, USA. ¹⁴National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. ¹⁵RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. ¹⁶JGI Stanford Human Genome Center, 975 California Avenue, Palo Alto, California 94304, USA. ¹⁷National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan. ¹⁸Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093-0202, USA. [†]Present addresses: Okinawa Institute of Science and Technology (OIST), Uruma, Okinawa 904-2234, Japan (T.K., E.S. and N.S.); Genome Sequencing Center, Department of Genetics, Washington University in St Louis School of Medicine, St Louis, Missouri 63108, USA (A.C.H.); Institute for Evolutionary Discovery, 909 Hiawatha Drive, Mount Pleasant, Michigan 48858, USA (J.J.G.-B.).

General genome info

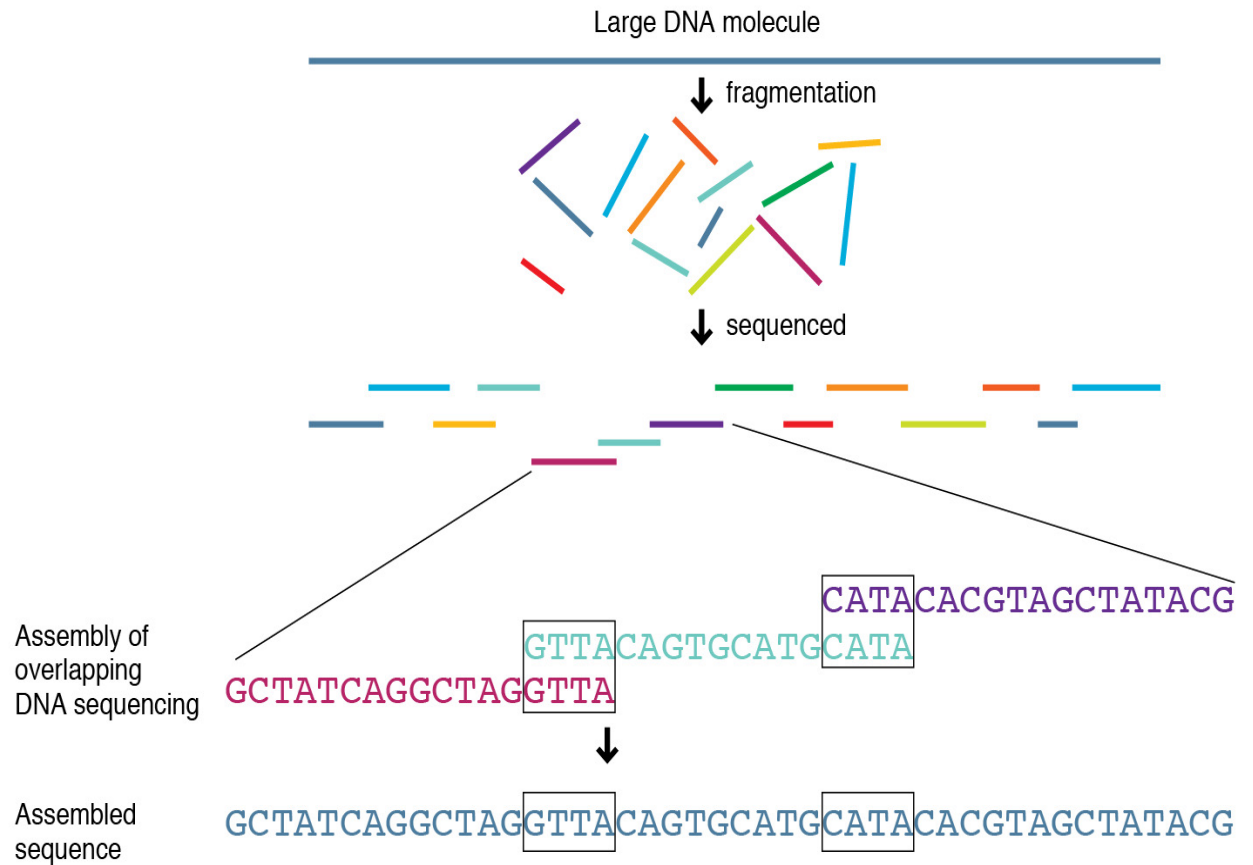
Scitefic name : Cephalochordata

Phylum : Chordata

Genome source : from gonads of a
single gravid male

Sequencing strategy

whole-genome shotgun strategy

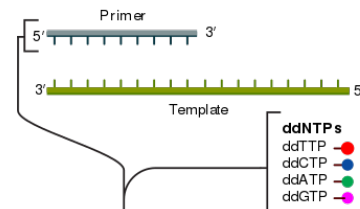


Sequencing methods

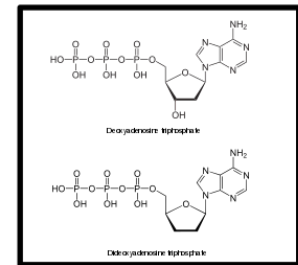
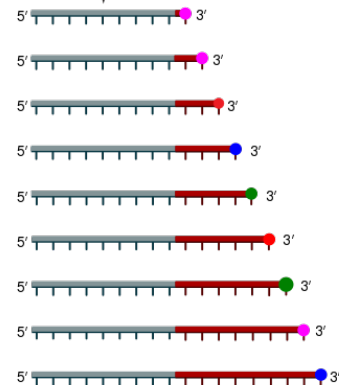
sequence Sanger “chain termination method”

① Reaction mixture

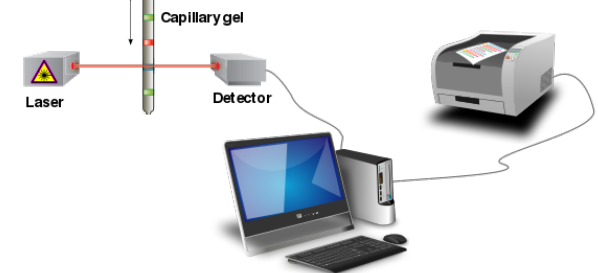
- ▶ Primer and DNA template
- ▶ DNA polymerase
- ▶ ddNTPs with flouochromes
- ▶ dNTPs (dATP, dCTP, dGTP, and dTTP)



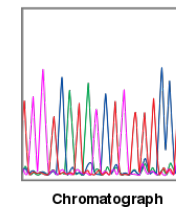
② Primer elongation and chain termination



③ Capillary gel electrophoresis separation of DNA fragments



④ Laser detection of flouochromes and computational sequence analysis



Genome assembly

Genome size : 520MB

N50 contigs 9,177 bp

Longest N50 contigs 25,666 bp

Number of scaffolds : 3,032

Genome coverage : 11.5

Genome outcome

The amphioxus sequence reveals key features of the genome of the last common ancestor of all chordates

