

Seeing chordate evolution through the *Ciona* genome sequence

Cristian Cañestro, Susan Bassham and John H Postlethwait

Address: Institute of Neuroscience, 1254 University of Oregon, 1425 E. 13th Avenue, Eugene, OR 97403, USA.

Correspondence: John H Postlethwait. E-mail: jpostle@uoneuro.uoregon.edu

Published: 3 March 2003

Genome **Biology** 2003, **4**:208

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2003/4/3/208>

© 2003 BioMed Central Ltd

Abstract

A draft sequence of the compact genome of the sea squirt *Ciona intestinalis*, a non-vertebrate chordate that diverged very early from other chordates, including vertebrates, illuminates how chordates originated and how vertebrate developmental innovations evolved.

“Mr. Kovalevsky has lately observed that the larvae of ascidians are related to the Vertebrata, in their manner of development, in the relative position of the nervous system, and in possessing a structure closely like the chorda dorsalis of vertebrate animals;... Thus, if we may rely on embryology, ever the safest guide in classification, it seems that we have at last gained a clew to the source whence the Vertebrata were derived.”

Charles Darwin [1]

The organism

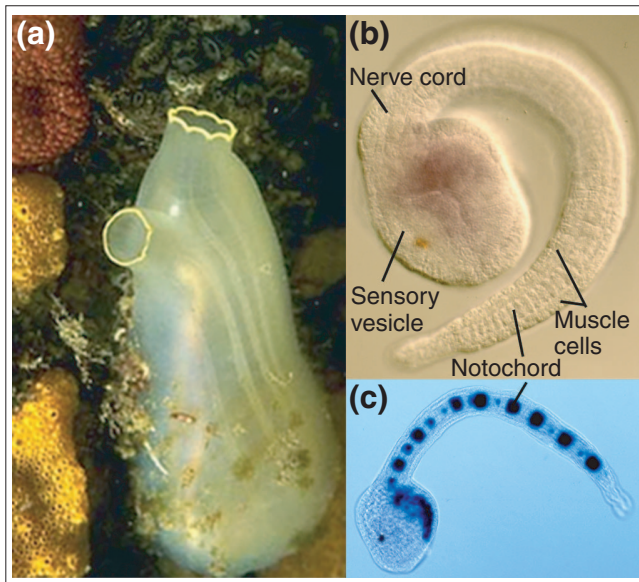
Early chordates were gentle, filter-feeding marine organisms, peaceful grazers of photosynthetic prokaryotes. Their descendants include rapacious vertebrates, with sense organs honed to detect prey, cunning nervous systems that out-wit hapless victims, and bony armor that protects them from counter-attack. What new genetic mechanisms led to the evolution of the innovative developmental programs that produced these novel vertebrate features? A recent article [2] announcing a draft genome sequence of the sea squirt *Ciona intestinalis* provides insights into not only the evolution of vertebrate novelties but also the entire phylum Chordata.

The adult *C. intestinalis* is a diaphanous sack that clings to a rock or to the piling of a pier, sucking in water through its mouth and filtering plankton through perforations - gill slits - in its pharynx. This simple adult ‘sea squirt’ disguised its true

evolutionary heritage so well that for a long time it languished in classifications among the mollusks, or even the “Vermes”. In 1866, Kovalevsky recognized that this ‘gilt-rimmed vase’ (see Figure 1a) and its relatives each developed from a beautiful larva that was obviously a chordate and a cousin of the vertebrates. The ‘tadpole’ larva has a notochord supporting a propulsive tail and a dorsal nervous system (Figure 1b). This microscopic larva settles from the plankton, glues its ‘face’ to a rock, and metamorphoses by resorbing its tail, restructuring its nervous system, and deploying a pair of siphons through which it will filter algae from the sea for the rest of its days.

Since Kovalevsky’s revelation, ascidians have generated considerable excitement in the discussion of vertebrate origins. Classical anatomists entertained the possibility that an ancient ascidian evolved a capacity to mature its gonads while remaining in the motile larval form, skipping a sessile adult stage, and that its descendants gave rise to the first vertebrate. Contemporary biologists acknowledge the evolutionary importance of ascidians differently - their genome and embryos offer a periscopic view on early chordate biology, but the ascidians themselves are not our ancestors.

Ascidians and their two sister classes, the larvaceans (or appendicularians) and thaliaceans, define the urochordates that, together with vertebrates and cephalochordates (such as amphioxus) make up the phylum Chordata (Figure 2). Urochordates split off earliest from other chordates, and the characteristics we share with them may therefore help to

**Figure 1**

Ciona intestinalis. **(a)** The *Ciona* adult develops from a tadpole larva by a dramatic metamorphosis, during which it resorbs its tail, modifies its central nervous system, and transforms its digestive system into incurrent and excurrent siphons that filter plankton through perforations in the pharynx (gill slits). Image courtesy of Andrew Martinez. **(b)** An early tadpole hybridized with a probe for the alcohol dehydrogenase gene *CiAdh3* to show the anterior endoderm [27]. Chordate features such as the notochord and the muscular tail are visible, and the positions of the sensory vesicle (brain) and dorsal nerve cord are indicated. **(c)** Electroporation is an efficient procedure by which to introduce DNA into *Ciona* for functional experiments, as is shown in this larva expressing a reporter construct made by Corbo *et al.* [4], a *lacZ* gene driven by the *brachyury* promoter in the notochord.

shape a portrait of the echordate, the ancestor of modern chordates. Chordates, like echinoderms and hemichordates, are deuterostomes, a division of animals that are distinguished from other animals (protostomes) by the development of the embryonic blastopore into the anus rather than the mouth.

Because of their phylogenetic position and their beautiful embryos, ascidians are popular among embryologists, and there has been impressive progress in the application of molecular tools to their study. These tools include loss-of-function studies using antisense morpholino oligonucleotides [3]; electroporation, which facilitates both the introduction of transgenes (Figure 1c) [4] and identification of *cis*-regulatory elements [5]; and the induction of mutations that block development in *C. savignyi* [6], a species whose genome is also being sequenced. The sequence of the *C. intestinalis* genome will greatly facilitate these functional analyses.

The sequence

A consortium of biologists and genomicists from sea coasts around the world (Australia, Canada, France, Italy, Japan,

Scotland, and the USA), led by the US DOE Joint Genome Institute, used a whole-genome shotgun approach to determine the sequence of much of the 159 megabase (Mb) genome of a single *C. intestinalis* individual from California (Figure 2b). The size of the *C. intestinalis* genome is in the same range as genomes of protosomes; it is around 50 Mb larger than that of the nematode *Caenorhabditis elegans* and 100 Mb smaller than most insects [7]. The *Ciona* genome is only about 5% of the size of the human genome, and 43% of the dense genome of the pufferfish *Fugu rubripes* (365 Mb).

The *Ciona* sequence database [8] reports 116.7 Mb of unique sequence in 2,501 scaffolds longer than 3 kilobases (kb). Identification of genes and their annotation were facilitated by the Kyoto full-length cDNA project [9]. This has produced a database [10] that adds significant value to the sequence by providing expression data for thousands of sequences over several life-cycle stages. The team [2] identified 15,852 gene models in *Ciona*, a number comparable to other protosomes (there are 13,639 and 19,518 currently identified in *D. melanogaster* and *C. elegans*, respectively, according to the Ensembl project [11]) (Figure 2). On the other hand, this number is only about half the estimated total number of gene models in the human and mouse genomes (30,000, although at the time of this writing Ensembl predicted just 22,980 and 22,444 genes, respectively) [12]. *F. rubripes* appears to have an upper limit of about 38,000 genes [13] (or 35,180 according to Ensembl); this larger number is due to the genome duplication event that preceded the radiation of teleost fish [14].

The *Ciona* genome is tightly packed; it has 50% of the number of genes found in the human genome but only 5% of the quantity of DNA. A major factor influencing genome compaction is the abundance of interspersed repeats and transposable elements (TEs). In the mosquito *Anopheles gambiae*, for instance, TEs account for about 16% of the euchromatic genome and 60% of the heterochromatic genome; this contrasts with the compact genome of *D. melanogaster* with 2% and 8%, respectively [7]; and just 3% of the compact pufferfish genome is interspersed repeats, far below the 35-40% in mammals [13]. In *Ciona*, Dehal *et al.* [2] identify several high-copy-number repeat classes, which account for about 11% of the genome. Although this analysis does not describe any TEs, a previous systematic study identified several types in *Ciona* [15]. The *Ciona* genome draft now provides the raw data for an exhaustive analysis of TEs, which will illuminate their roles in evolutionary history of genome architecture.

The innovations

What does the *Ciona* sequence [2] tell us about the innovations of vertebrates, chordates, and deuterostomes? Almost 62% of *Ciona* genes (9,883) have a detectable protosome

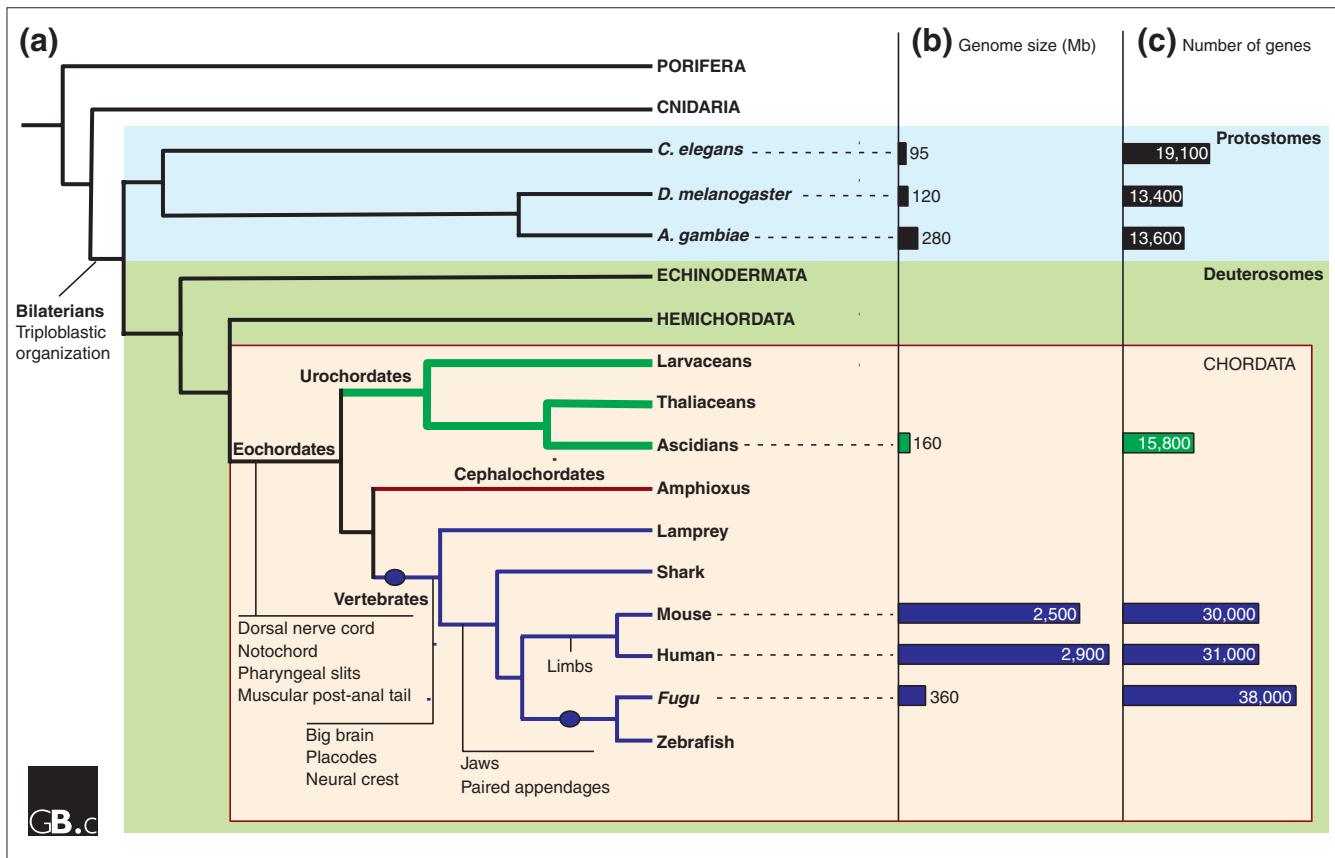


Figure 2

A phylogeny of chordate evolution. (a) A phylogenetic tree depicting the likely relationships between chordates and other animals, timing of genome amplifications (solid circles), and origin of novel characteristics. Phyla are in capitals; Porifera includes sponges; Cnidaria includes jellyfish and hydra; Echinodermata includes sea stars and sea urchins and Hemichordata includes marine acorn worms. (b) Genome size (Mb) and (c) estimated gene number of sequenced animal genomes are represented by bars.

homolog, and these presumably constitute an ancient core of genes common to bilaterian animals. A few hundred *Ciona* genes, including phytochelatin synthase and hemocyanin, have stronger similarity to genes of protostomes than to any vertebrate gene. Either these are ancient bilaterian genes whose vertebrate orthologs have been lost or changed beyond recognition, or they were perhaps acquired by horizontal transfer from protostomes. Conversely, around 15% of *Ciona* genes (2,570) lack a clear protostome homolog yet have a vertebrate counterpart. These could have arisen in the deuterostome lineage after the split from protostomes, or alternatively they could be homologs of ancient bilaterian genes that have diverged beyond detection or been lost from modern protostomes. The genome sequence of an echinoderm, an outgroup to the chordates, will help determine if these genes are ancient within the deuterostomes [16].

Rather surprisingly, 21% of *Ciona* genes (3,399) have no clear homolog in the fly, worm, pufferfish, or human genomes (under high-stringency Smith-Waterman alignment along 60% of the target protein). Although this group

might include poorly modeled genes or genes that have been broken by the ends of contigs, some may also have evolved so rapidly that they have lost significant resemblance to their orthologs, or they may be genes specific to urochordates or to ascidians. Resolution of this last ambiguity would require the sequencing of a distant urochordate genome, such as that of a larvacean [17].

What can the *Ciona* sequence tell us about vertebrate innovations? It can expedite the hunt for genes that are important in the development of the embryonic tissues that are widely believed to have initiated the evolution of the vertebrates [18]. These include neural crest and ectodermal placodes, which contribute to the paired sensory organs and head skeleton that are so apparent in vertebrates but are seemingly lacking from other chordates. If these tissues evolved after vertebrate origins, what functions do crest and placode genes have in an ascidian? Might such genes reveal the evolutionary precursors of these tissues? Many genes central to these questions have already been investigated in ascidians by targeted sequencing [19], but the

Ciona sequencing project has detected more peripheral players. For instance, three *Ciona* genes (ci0100131069, ci0100130876 and ci0100135383) appear to be homologous to vertebrate genes encoding the olfactomedin family; one vertebrate member of this family, *Noelin-1*, encodes a protein that makes the neural tube competent to generate neural crest cells [20]. Two ascidian genes (ci0100149361 and ci0100140298) are homologous to the vertebrate *HAND1* and *HAND2* genes, which are also important in crest development [21]. And *Ciona* has two orthologs (ci0100136347 and ci0100130219) of the vertebrate *Prox1* gene, a marker of lens, otic, olfactory, and ganglia placodes [22]. The discovery of genes with only weak homology to olfactory receptors, however, undermines the prospects of discovering an olfactory placode precursor. Nonetheless, developmental biologists will be kept busy for some time studying the expression and function of these newly revealed genes.

Adaptive immunity involving lymphocytes appears to have arisen within the vertebrates; Dehal *et al.* [2] could not find *Ciona* homologs of genes involved in this system, including immunoglobulins, T-cell receptors, or major histocompatibility complex (MHC) genes. Furthermore, they found that, although *Ciona* has orthologs for each of the 14 proteasome genes, it lacks immunoproteasome-specific genes, suggesting that *Ciona* has no specific system for presenting antigens. Nevertheless, ascidians have a potent innate immune system [23], including possible complement genes and several lectins.

Genome amplifications

Gene counts show that mice and humans have about twice as many genes as ascidians. When did genome-amplification events happen with respect to the origin of vertebrate innovations? Did amplification occur by whole-genome duplication or by the independent amplification of many small regions of the genome? And if it occurred by genome duplication, was there one round of duplication or more? A sequenced non-vertebrate chordate genome fully aligned along complete chromosomes would provide data with which to critically examine these issues. Although about 85% of the available *Ciona* sequence is in 905 scaffolds longer than 20 kb, these are not yet aligned along chromosomes. Substantial value would be added to the sequence if a meiotic or radiation-hybrid mapping panel were constructed, as a way to order markers from each scaffold. Then, long-range comparative synteny analysis could be brought to bear on the issue of the mechanism of vertebrate genome amplification.

Unfortunately, few ancient branches survive on the chordate tree (Figure 2). But the extant lineages diverged at key nodes for resolving important questions about chordate evolution. The sequence of a larvacean, a new model for molecular

development [24] with a genome half the size of the ascidian [25], would allow generalizations about ancestral urochordates. Cephalochordates are the sister group of vertebrates, and the genome sequence of an amphioxus would, in conjunction with urochordate sequences, suggest the content of the genome before the genome amplification events that occurred with, and probably facilitated, the origin of vertebrate developmental novelties [26]. Finally, we need genomic evidence from the earliest diverging vertebrate whose embryos can be studied, namely a lamprey. Sifting through these genomes to design functional experiments on developmental regulatory mechanisms will help us learn how meek, unassuming grazers evolved into vicious predators like you and me.

Acknowledgements

We thank an NSF IGERT program grant DGE 9972830 in Evolution of Development and Genomics (S.B.), a grant EX2002-0059 from the Ministry of Education, Culture and Sports from the Spanish Government (C.C.), and NIH grant R01RR10715 (J.H.P.) for support.

References

1. Darwin C: *The Descent of Man, and Selection in Relation to Sex*. New York: Hurst; 1874.
2. Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomoso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, et al.: **The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins**. *Science* 2002, **298**:2157-2167.
3. Satou Y, Imai KS, Satoh N: **Action of morpholinos in *Ciona* embryos**. *Genesis* 2001, **30**:103-106.
4. Corbo JC, Levine M, Zeller RW: **Characterization of a noto-chord-specific enhancer from the Brachyury promoter region of the ascidian, *Ciona intestinalis***. *Development* 1997, **124**:589-602.
5. Harafuji N, Keys DN, Levine M: **Genome-wide identification of tissue-specific enhancers in the *Ciona* tadpole**. *Proc Natl Acad Sci USA* 2002, **99**:6802-6805.
6. Moody R, Davis SW, Cubas F, Smith WC: **Isolation of developmental mutants of the ascidian *Ciona savignyi***. *Mol Gen Genet* 1999, **262**:199-206.
7. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, et al.: **The genome sequence of the malaria mosquito *Anopheles gambiae***. *Science* 2002, **298**:129-149.
8. **JGI: *Ciona intestinalis* v1.0** [<http://genome.jgi-psf.org/ciona4/ciona4.home.html>]
9. Satou Y, Takatori N, Fujiwara S, Nishikata T, Saiga H, Kusakabe T, Shin-i T, Kohara Y, Satoh N: ***Ciona intestinalis* cDNA projects: expressed sequence tag analyses and gene expression profiles during embryogenesis**. *Gene* 2002, **287**:83-96.
10. **Ghost Database: a *Ciona intestinalis* cDNA resource** [<http://ghost.zool.kyoto-u.ac.jp/indexr1.html>]
11. **Ensembl Genome Browser** [<http://www.ensembl.org/>]
12. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexander M, An P, et al.: **Initial sequencing and comparative analysis of the mouse genome**. *Nature* 2002, **420**:520-562.
13. Aparicio S, Chapman J, Stupk E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S, Smit A, et al.: **Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes***. *Science* 2002, **297**:1301-1310.
14. Amores A, Force A, Yan Y-L, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang Y-L, et al.: **Zebrafish *hox* clusters and vertebrate genome evolution**. *Science* 1998, **282**:1711-1714.
15. Simmen MW, Bird A: **Sequence analysis of transposable elements in the sea squirt, *Ciona intestinalis***. *Mol Biol Evol* 2000, **17**:1685-1694.

16. Cameron RA, Mahairas G, Rast JP, Martinez P, Biondi TR, Swartzell S, Wallace JC, Poustka AJ, Livingston BT, Wray GA, et al.: **A sea urchin genome project: sequence scan, virtual map, and additional resources.** *Proc Natl Acad Sci USA* 2000, **97**:9514-9518.
17. Wada H, Satoh N: **Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA.** *Proc Natl Acad Sci USA* 1994, **91**:1801-1804.
18. Gans C, Northcutt RG: **Neural crest and the origin of vertebrates: a new head.** *Science* 1983, **220**:268-274.
19. Holland LZ, Holland ND: **Evolution of neural crest and placodes: amphioxus as a model for the ancestral vertebrate?** *J Anat* 2001, **199**:85-98.
20. Bronner-Fraser M: **Molecular analysis of neural crest formation.** *J Physiol Paris* 2002, **96**:3-8.
21. Howard M, Foster DN, Cserjesi P: **Expression of HAND gene products may be sufficient for the differentiation of avian neural crest-derived cells into catecholaminergic neurons in culture.** *Dev Biol* 1999, **215**:62-77.
22. Rodriguez-Niedenfuhr M, Papoutsi M, Christ B, Nicolaides KH, von Kaisenberg CS, Tomarev SI, Wilting J: **Prox1 is a marker of ectodermal placodes, endodermal compartments, lymphatic endothelium and lymphangioblasts.** *Anat Embryol (Berl)* 2001, **204**:399-406.
23. Davidson B, Swalla BJ: **A molecular analysis of ascidian metamorphosis reveals activation of an innate immune response.** *Development* 2002, **129**:4739-4751.
24. Bassham S, Postlethwait J: **Brachyury (T) expression in embryos of a larvacean urochordate, Oikopleura dioica, and the ancestral role of brachyury.** *Dev Biol* 2000, **220**:322-333.
25. Seo H-C, Kube M, Edvardsen RB, Jensen MF, Beck A, Spriet E, Gorsky G, Thompson EM, Lehrach H, Reinhardt R, et al.: **Miniature genome in the marine chordate Oikopleura dioica.** *Science* 2001, **294**:2506.
26. Holland PWH, Garcia-Fernández J, Williams NA, Sidow A: **Gene duplications and the origins of vertebrate development.** *Dev* 1994, (suppl):125-133.
27. Cañestro C, Godoy L, González-Duarte R, Albalat R: **Comparative expression analysis of Adh3 during arthropod, urochordate, cephalochordate and vertebrate development challenges its predicted housekeeping role.** *Evol Dev*, in press.