

CIPRO 2.5: *Ciona intestinalis* Protein Database - a unique integrated repository of large-scale omics data, bioinformatic analyses, and curated annotation, with ability for user rating and comments

Toshinori Endo^{1*}, Keisuke Ueno¹, Kouki Yonezawa¹, Katsuhiko Mineta¹, Kohji Hotta², Yutaka Satou³, Liyx Yamada⁴, Michio Ogasawara⁵, Hiroki Takahashi⁶, Ayako Nakajima⁷, Mia Nakachi⁷, Mamoru Nomura⁷, Junko Yaguchi⁷, Alu Konno⁷, Yasunori Sasakura⁷, Akiyasu C Yoshizawa⁸, Hisaaki Taniguchi⁸, Chisato Yamasaki⁹, Miho Sera⁹, Tadashi Imanishi⁹, Kazuo Inaba⁷

From Beyond the Genome: The true gene count, human evolution and disease genomics
Boston, MA, USA. 11-13 October 2010

CIPRO database (<http://cipro.ibio.jp/2.5>) is an integrated protein database for a tunicate species *Ciona intestinalis* that is part of the Urochordata. Although CIPRO provides proteomic and transcriptomic data on a single species, the animal is considered unique in the evolutionary tree, representing a possible origin of the vertebrates and is a good model for understanding chordate evolution, including human evolution. Furthermore, *C. intestinalis* has been one of the favorites of developmental biologists; therefore, a lot of amount of accumulated knowledge on its development, morphology, in addition to the recent genome sequence and gene expression data exists. The CIPRO database aims to collect published data and to present unique information, including the unpublished transcriptomic and proteomic data and human curated annotation, for the use of researchers in biology and bioinformatics. The current database contains 89,673 unique sequences covering all the proteins from all the gene models on this species; the number was reduced to 70,493 by similarity clustering. Of these sequences, more than 5,000 proteins are manually annotated based on the large-scale transcriptomic, proteomic and bioinformatic data. Those annotations can be subjected to be qualification by rating, curation, and

comments by named and anonymous users through the web site of CIPRO database. Unique features of CIPRO database include:

(i) **Original experimental data** Unpublished experimental data, including 2D-PAGE with the identified protein spots by protein mass fingerprint (PMF) MS analysis, expressions or localizations of protein and RNA across developmental stages and tissues, altogether summarized in a single chart for the comparison among status and methods. RNA expressions are observed by microarray and EST. Each protein is linked to an independent Ascidian Proteome Database summarizing large-scale MS-based proteomic analyses.

(ii) **Whole *Ciona intestinalis* proteome database** Proteins across gene models are presented: all protein models derived from published gene models are incorporated, including Kyoto model (KG), KH (successor of KG model), PROCITS, JGI's versions 1 and 2, and Ensembl (version 58.2) are incorporated. Identical sequences across gene models are shown.

(iii) **Original comprehensive user-friendly interfaces** Bioinformatic analyses and prediction results are summarized in pictures for grasp at a glance: homology search, cytolocalization, secondary structure prediction combined with modification sites, such include phosphorylation and three-dimensional structures.

¹Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan
Full list of author information is available at the end of the article

(iv) Comparative analysis data for disease association Comparison with human genome: map location of human homologues is graphically shown with associated disease information. Comparative data for other model organisms are also included.

(v) Community-wide curation capability opened to users To facilitate progressive improvement of annotation by visited users, users can place additional annotation for the protein name and/or comments, which will be subjected to rating by the followed data viewers. To aid curation by wide community, information for literature and essence of matched motif patterns and other related protein information are shown with the links.

(vi) Useful search facilities Various search methods are provided including blast homology, free text, partial sequence, protein mass fragment, and cross item searches.

Author details

¹Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan. ²Department of Biosciences and Informatics, Keio University, Yokohama 223-8522, Japan. ³Department of Zoology, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606 8502, Japan. ⁴Sugashima Marine Biological Laboratory, Nagoya University, 429-63 Sugashima, Toba 517-0004, Japan. ⁵Department of Biology, Faculty of Science, Chiba University, Inage-Ku, Chiba 263-8522, Japan. ⁶National Institute for Basic Biology, Nishigonaka 38, Myodaiji, Okazaki 444-8585, Japan. ⁷Shimoda Marine Research Center, University of Tsukuba, 5-10-1 Shimoda, Shizuoka, 415-0025, Japan. ⁸Institute for Enzyme Research, the University of Tokushima, 3-18-15, Kuramoto-cho, Tokushima 770-8503, Japan. ⁹Biomedical Information Research Center, Advanced Industrial Science and Technology, Waterfront Bio-IT Research Building, 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan.

Published: 11 October 2010

doi:10.1186/gb-2010-11-S1-P11

Cite this article as: Endo *et al.*: CIPRO 2.5: *Ciona intestinalis* Protein Database - a unique integrated repository of large-scale omics data, bioinformatic analyses, and curated annotation, with ability for user rating and comments. *Genome Biology* 2010 11(Suppl 1):P11.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

