Data in Brief 6 (2016) 715–721

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data article

Proteome-wide dataset supporting the study of ancient metazoan macromolecular complexes

Sadhna Phanse [a,1], Cuihong Wan [a,b,1], Blake Borgeson [b,1], Fan Tu [b], Kevin Drew [b], Greg Clark [c], Xuejian Xiong [d,e], Olga Kagan [a], Julian Kwan [a,d], Alexandr Bezginov [c], Kyle Chessman [d,e], Swati Pal [d,e], Graham Cromar [d,e], Ophelia Papoulas [b], Zuyao Ni [a], Daniel R. Boutz [b], Snejana Stoilova [a], Pierre C. Havugimana [a], Xinghua Guo [a], Ramy H. Malty [g], Mihail Sarov [h], Jack Greenblatt [a,d], Mohan Babu [g], W. Brent Derry [d,e], Elisabeth R. Tillier [c], John B. Wallingford [b,f], John Parkinson [d,e], Edward M. Marcotte [b,f], Andrew Emili [a,d,*]

[a] *Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario, Canada*
[b] *Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX, USA*
[c] *Department of Medical Biophysics, Toronto, Ontario, Canada*
[d] *Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada*
[e] *Hospital for Sick Children, Toronto, Ontario, Canada*
[f] *Department of Molecular Biosciences, University of Texas at Austin, Austin, TX, USA*
[g] *Department of Biochemistry, University of Regina, Regina, Saskatchewan, Canada*
[h] *Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany*

ARTICLE INFO

Article history:
Received 15 September 2015
Received in revised form
17 November 2015
Accepted 23 November 2015
Available online 12 December 2015

Keywords:
Proteomics
Metazoa
Protein complexes

ABSTRACT

Our analysis examines the conservation of multiprotein complexes among metazoa through use of high resolution biochemical fractionation and precision mass spectrometry applied to soluble cell extracts from 5 representative model organisms *Caenorhabditis elegans, Drosophila melanogaster*, *Mus musculus*, *Strongylocentrotus purpuratus,* and *Homo sapiens*. The interaction network obtained from the data was validated globally in 4 distant species (*Xenopus laevis*, *Nematostella vectensis*, *Dictyostelium discoideum*, *Saccharomyces cerevisiae*) and locally by targeted affinity-purification

* Corresponding author at: Banting and Best Department of Medical Research, Terrence Donnelly Center for Cellular and Biomolecular Research, University of Toronto, CCBR, Rm 914, 160 College Street, Toronto, Ontario, Canada M5S 3E1. Tel.: +1 617 610 4042; fax: +1 416 978 8528.
E-mail address: andrew.emili@utoronto.ca (A. Emili).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.dib.2015.11.062

Biochemical Fractionation

experiments. Here we provide details of our massive set of supporting biochemical fractionation data available *via* ProteomeXchange (PXD002319-PXD002328), PPIs *via* BioGRID (185267); and interaction network projections via (http://metazoa.med.utoronto.ca) made fully accessible to allow further exploration. The datasets here are related to the research article on metazoan macromolecular complexes in Nature [1].

Specifications Table

Subject area	Biology
More specific subject area	Metazoan proteomics
Type of data	Set of tables
How data was acquired	Biochemical fractionation combined with quantitative mass spectrometry using LTQ XL; LTQ Orbitrap Velos
Data format	Raw and processed data
Experimental factors	• Whole body lysate from worm (*Caenorhabditis elegans*) • AX4 cells from amoeba (*Dictyostelium discoideum*) • 2 cell types in fly (*Drosophila melanogaster*), • 5 cell lines in human (*Homo sapiens*), • Embryonic stem cells from mice (*Mus musculus*) • Unfertilized sea anemone eggs (*Nematostella vectensis*) • Log-phase culture of wild type yeast W303 strain (*Saccharomyces cerevisiae*) • 5 different development stages in sea urchin (*Strongylocentrotus purpuratus*), • Stage 15–19 embryos, adult male heart and liver from frog *(Xenopus laevis)*
Experimental features	Combination of biochemical fractionation with quantitative mass spectrometry for 6387 fractions obtained from 69 different experiments, to examine the composition of soluble multiprotein complexes among diverse animal models.
Data source location	*Toronto, Canada*
Data accessibility	• Biochemical fractionations – ProteomeXchange (PXD002319-PXD002328) • PPIs – BioGRID (185267) • Complexes and interaction network projections – http://metazoa.med.utoronto.ca • MS1 and MS2 elution profiles, correlation scores and ortholog mappings – http://metazoa.med.utoronto.ca • Supplementary data with research article.

Value of the data

- Macromolecular complexes drive essential biological processes, yet their ubiquity across phyla is unclear. By applying a human-centric approach on the merged data for 5 species obtained through fractionation and mass spectrometry, and subsequent computational analysis we identified 16,655 high confidence protein–protein interactions and 981 putative functional modules encompassing 2153 broadly-conserved proteins found in virtually all multicellular eukaryotes.