The BioGRID Interaction Database

Michael S. Livstone[1], Bobby-Joe Breitkreutz[2], Chris Stark[2], Lorrie Boucher[2], Andrew Chatr-Aryamontri[3], Rose Oughtred[1], Julie Nixon[3], Teresa Reguly[2], Jennifer Rust[1], Andrew Winter[3], Kara Dolinski[1], Mike Tyers[2,3].

[1]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA; [2]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada; [3]Wellcome Trust Centre for Cell Biology, The University of Edinburgh, Edinburgh, UK.

The goal of the Biological General Repository for Interaction Datasets (BioGRID) is to archive and freely disseminate collections of genetic and protein interactions from major model organisms. BioGRID currently contains over 355,000 interactions curated from high-throughput datasets and individual focused studies from over 23,000 publications in the primary literature. Complete coverage of the entire literature for both the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe* has been achieved and efforts to expand curation across multiple species are underway.

Displaying Interactions

The new search results display provides a summary of the total number of curated interactions for the gene or protein of interest, as well as the number of physical or genetic interactions for each interactor. The interactions may be filtered according to whether they are low or high throughput. Details are provided for different experiment types including the role of the interactor and experimental throughput. Mousing over icons in the notes column reveals further details including phenotypes, post-translational modifications, quantitative scores, and notes, if applicable. The summary view shown below may be switched to a table view that can be sorted by interactor, experiment type, publication, or throughput, among other options.

PhosphoGRID Database

Direct links are provided from BioGRID's Interaction Summary page to various resources including PhosphoGRID, a database of experimentally verified phosphorylation sites curated from the *S. cerevisiae* literature. In this example, phosphorylation sites in the protein sequence of CDC28 are shown in red and protein kinase motifs are highlighted in blue. Details for each phosphorylation site are provided in the table below, or in a pop-up window viewable by mousing over a phosphorylated residue of interest. PhosphoGRID documents over 5000 phosphorylated residues found in 1495 gene products. All curated data in PhosphoGRID are freely available and may be downloaded at www.phosphogrid.org.

All data in BioGRID are freely available and may also be downloaded via:

http://www.thebiogrid.org/

Searching BioGRID

BioGRID may be searched using various protein or gene identifiers, including UniProt and NCBI IDs, as well as keywords and PubMed IDs. Searches may be limited to an organism of interest, or else performed on all 50 organisms including yeast, human, and mouse. In this example, CDC28 is entered as the search term and *S. cerevisiae* is selected as the organism.

New project in conjunction with other databases

BioGRID is initiating a new project to provide high-quality coverage of biological processes of relevance to human disease. Our initial focus is on genes in the Wnt signaling pathway, recently co-curated as part of the Gene Ontology (GO) Consortium Reference Genome project. BioGRID and WormBase curators are curating *C. elegans* physical and genetic interactions using papers annotated with the GO evidence codes IPI (Inferred from Physical Interaction) or IGI (Inferred from Genetic Interaction). BioGRID captures specific information not collected during GO curation, such as specific experimental systems and (for genetic interactions) phenotypes from the Worm Phenotype Ontology, and displays additional notes and details (see figure below). We will soon expand this focused curation into additional organisms. By coordinating our efforts with the GO Consortium and the Linking Animal Models to Human Disease Initiative (LAMHDI) Consortium, we hope to provide a more complete picture of important processes.


BioGRID Database Statistics

The number of interactions ("edges" in the table below) and the number of papers curated in BioGRID has nearly tripled since July 2006. 99% of interactions come from the six model organisms shown, with 68% coming from *S. cerevisiae*.

Organism	Type	July 2006 (2.0.17)			August 2010 (3.0.67)		
		nodes	edges	papers	nodes	edges	papers
A. thaliana (thale cress)	PI	0	0	0	1,735	4,719	747
	GI	0	0	0	88	174	55
C. elegans (worm)	PI	2,790	4,433	1	2,813	4,663	12
	GI	0	0	0	1,030	2,112	5
D. melanogaster (fly)	PI	6,997	22,133	2	7,396	24,480	167
	GI*	1,189	10,314	1,493	982	9,994	1,466
H. sapiens (human)	PI	3,380	7,238	178	9,467	48,368	10,203
	GI	0	0	0	479	463	178
S. cerevisiae (budding yeast)	PI	5,144	49,297	3,267	5,783	90,769	5,444
	GI	3,352	24,636	3,796	5,357	146,081	5,606
S. pombe (fission yeast)	PI	0	0	0	1,441	4,019	769
	GI	0	0	0	1,340	11,527	953
All other organisms	PI	284	620	73	2,288	2,985	830
Total	ALL	19,176	118,671	7818	30,665	347,966	23,451

Increase in BioGRID data content, 2006-2010. * from FlyBase

Distribution of BioGRID users

BioGRID traffic sources
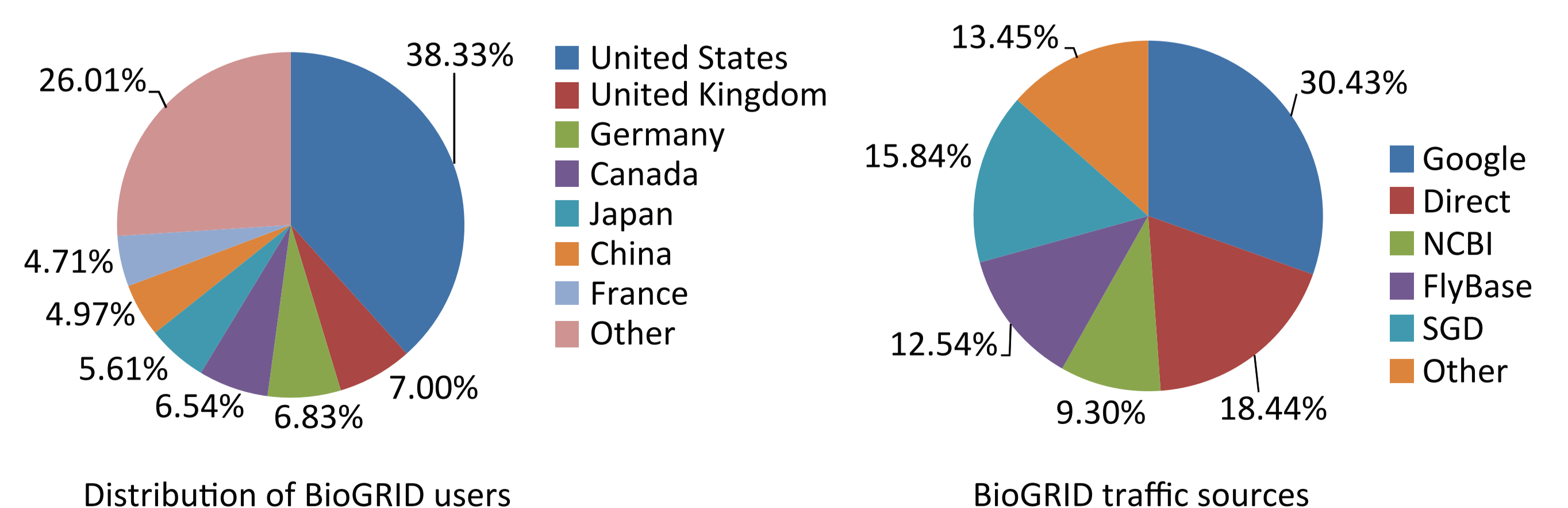
References

(1) Breitkreutz BJ, Stark C, Reguly T, Boucher L, Breitkreutz A, Livstone M, Oughtred R, Lackner DH, Bähler J, Wood V, Dolinski K, Tyers M. "The BioGRID Interaction Database: 2008 update." *Nucleic Acids Res.* 2008 Jan;*36*(Database issue):D637-40.