Nucleic Acids Research Advance Access published November 21, 2011

Nucleic Acids Research, 2011, 1–8
doi:10.1093/nar/gkr1062

Comparative interactomics with Funcoup 2.0

Andrey Alexeyenko[1,2], Thomas Schmitt[2,3,4], Andreas Tjärnberg[2,3,4,5], Dmitri Guala[2,3,4], Oliver Frings[2,3,4] and Erik L. L. Sonnhammer[2,3,4,5,*]

[1]School of Biotechnology, Royal Institute of Technology, [2]Science for Life Laboratory, Box 1031, SE-17121 Solna, Sweden, [3]Stockholm Bioinformatics Centre, [4]Department of Biochemistry and Biophysics, Stockholm University and [5]Swedish eScience Research Center

Received September 27, 2011; Revised and Accepted October 26, 2011

ABSTRACT

FunCoup (http://FunCoup.sbc.su.se) is a database that maintains and visualizes global gene/protein networks of functional coupling that have been constructed by Bayesian integration of diverse high-throughput data. FunCoup achieves high coverage by orthology-based integration of data sources from different model organisms and from different platforms. We here present release 2.0 in which the data sources have been updated and the methodology has been refined. It contains a new data type Genetic Interaction, and three new species: chicken, dog and zebra fish. As FunCoup extensively transfers functional coupling information between species, the new input datasets have considerably improved both coverage and quality of the networks. The number of high-confidence network links has increased dramatically. For instance, the human network has more than eight times as many links above confidence 0.5 as the previous release. FunCoup provides facilities for analysing the conservation of subnetworks in multiple species. We here explain how to do comparative interactomics on the FunCoup website.

INTRODUCTION

Recent advances in high-throughput biology such as genomics, proteomics and interactomics have led to a massive increase in our knowledge about the functional properties of genes and their encoded proteins. From direct interactions and indirect ones such as correlated functional behaviour, one can infer networks of functional coupling. The FunCoup networks are among the largest reconstructions to date, which can be attributed to the extensive transfer of evidence between species via orthologues and the usage of nine different data source types. By synthesis of multiple data sources, a more comprehensive network can be obtained, with higher quality. One reason for this is that underlying biological networks are indeed composed of different molecular mechanisms of communication between genes and proteins: via protein phosphorylation, complex formation, transcription factor binding, miRNA targeting etc. Secondly, every high-throughput technique has specific advantages and drawbacks. The false-positive rate is often considerable and the false-negative rate is always huge. By combining the signal of functional coupling from heterogeneous sources, true signals will be enforced while false ones will be dampened. The FunCoup (1) framework is a Bayesian approach to turn various raw scores of functional coupling into probabilistic estimates that are then integrated across all types of data and model organisms. The orthologue assignments used by FunCoup for cross-species mapping are obtained from the InParanoid database (2).

Several other databases exist that integrate multiple data sources into networks. Each database has a unique combination of species, data sources, integration methods and user interface. Examples of other multi-species databases are N-Browse (3), ConsensusPathDB (4), I2D (5), GeneMANIA (6), PathwayCommons (7) and APID (8), containing between 3 and 15 species. More extensive species coverage is provided by the VisANT database (9) with 111 species, and STRING (10) with 1100. FunCoup mainly contains species for which there is abundant high-throughput data, i.e. the most popular model organisms. One exception is *Ciona intestinalis* which was included to demonstrate that the framework also works well in the absence of data in the species itself. The requirement for a species to be included is availability of a gold standard set of functional couplings in the same species, so that the input data are evaluated in the proper context. FunCoup has a set of unique scoring functions and an algorithm that creates discretized (binned) mappings between each raw metric score (Pearson linear correlation, PPI score etc.) and the respective likelihood of functional coupling given the raw metric value, dataset, species and type of functional coupling. One consequence

*To whom correspondence should be addressed. Tel: +46852481184; Fax: +46855378214; Email: Erik.Sonnhammer@sbc.su.se

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

of this feature that stands out is that FunCoup assigns both positive and negative evidence scores. As an example, two proteins localized in the same cellular compartment is a positive evidence of being in the same complex, whereas non-overlapping localizations generate an evidence against it. It also helps to avoid overestimation of the total score when summing over a large number of potential evidences.

The FunCoup database is downloadable as flat files (one per species) and can be queried online at the website FunCoup.sbc.su.se. Here a user can simply paste in one or multiple query identifiers and view the local subnetwork. Figure 1 illustrates the results page using the gene DYX1C1 (Dyslexia susceptibility 1 candidate gene 1 protein). At the top, an integrated Java applet jSquid (11) is shown if Java is installed, otherwise a static picture will appear. The size and properties of the subnetwork can be controlled on the query page. For instance, the confidence cut-off can be changed, or the query can be restricted to certain data types or source species. Below the network graph, a table with details on evidences for each link is shown, as well as a table of all the genes. Each query can be saved as a bookmark, and the resulting network can be saved for future use in jSquid.

A unique feature of the FunCoup website is the possibility to perform 'comparative interactomics' such that subnetworks of different species are aligned with each other using orthologues. Network alignment is an emerging field that has received attention not only because it can predict protein function but also because on the proteome scale it is an algorithmically and computationally very challenging problem. Several tools exist, for instance NetworkBlast (12), IsoRankN (13), Graemlin (14) and GraphCrunch (15). These use different methods and heuristics to align networks on the basis of features

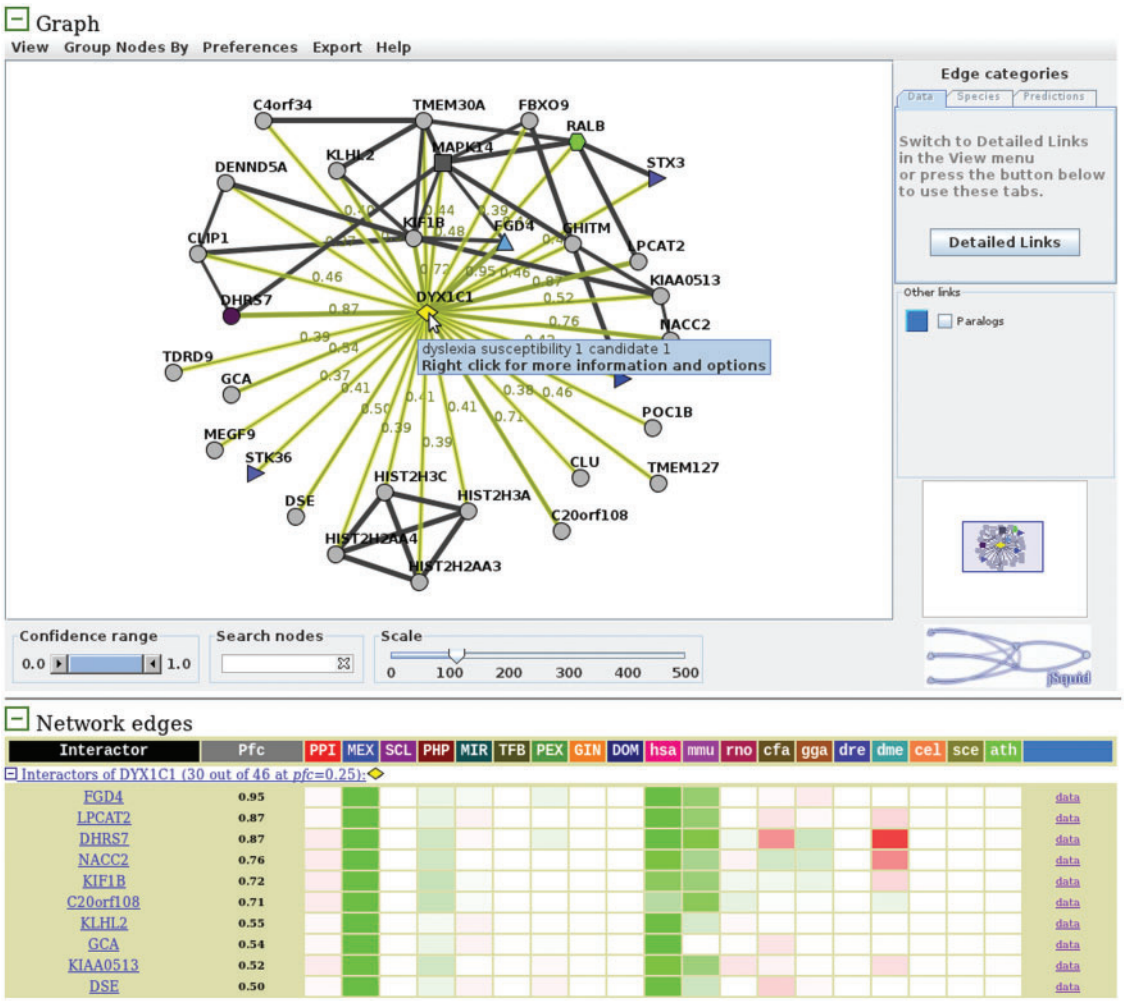

Figure 1. The main results page of FunCoup for the query DYX1C1 (human) and a cut-off of $pfc > 0.25$. The upper panel shows the subnetwork graph in the jSquid java applet. The query is shown as a yellow diamond and its neighbours in the FunCoup network either as grey balls or with shape/colour if it is assigned to a KEGG pathway. The confidence values of the edges to a node are only shown in the graph upon mouseover. Edges can also show relative support from different data types or species, or show predicted type by activating Detailed Links. The nodes are movable and can be assigned a new shape or colour. Groups of nodes can be selected with mouse rubberband and be collapsed. Below the graph, subnetwork details are shown for each link to indicate the level of support from each data type and species. The link 'data' at the right shows the raw scores of the underlying evidences.

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

such as sequence similarity, network topology similarity, functional similarity or structural similarity. Performing network alignment globally is however not very practical and runtimes are very long. For a given gene or gene set of interest, it is often more useful to consider the local subnetwork and search for its optimal alignment against another organism's network. FunCoup performs an orthology-based subnetwork alignment around query gene(s). This was already possible in version 1, but only in a mode that mostly aligns nodes sharing edges with evidence transferred from the other species. Version 2.0 employs a much stricter method, where the network alignment is based only on evidence from the species itself. This way conserved functional associations with independent support in each species are found. Such alignments are however considerably less frequent. The new stricter method is now the default mode, and a large part of this paper is devoted to showing how to carry out such analyses online on the website.

NEW FEATURES IN RELEASE 2

Beyond adding the new species dog, chicken and zebra fish, the data sources for functional coupling in FunCoup 2.0 have been updated for the already included species. A new data type GIN (genetic interactions) has been added for yeast, based on the correlation between genetic interaction profiles of two genes (16). Several data types have been substantially improved by using more comprehensive sources, e.g. the UniDomInt database (17) for domain interactions, while others have been improved by better score functions, e.g. the PPI score. In particular, we were in a position to consider microarray expression sets from a much broader choice than when building version 1. For each species we selected the most comprehensive (number of distinct conditions and probed transcripts) and informative (higher likelihood of functional coupling given co-expression) datasets.

Confidence values *pfc* were calculated for each predicted link from the final Bayesian scores (FBS, sum of log likelihood ratios from individual input sets) according to:

$$pfc = \frac{1}{1+e^{-FBS-\ln P(FC)}}$$

where $P(FC)$, the prior probability that 'two randomly picked proteins are functionally coupled' is set to 0.001. A *pfc* value for each gene-gene link is now incorporated into all the flat files, in addition to the FBS and its components classified by contributing evidence classes. Users downloading a whole network can thus study versions of it based on e.g. solely protein–protein interactions, a union of co-expression and sub-cellular co-localization, or data from a certain species, just like users of the web query interface.

The inclusion of more comprehensive data and data of higher quality has greatly increased the total evidence and yields more accurate predictions. We raised the minimum *pfc* cut-off from 0.02 to 0.1, yet predict more functional couplings for most of the species. Table 1 shows the network sizes in FunCoup 2.0. Considering only links with *pfc* > 0.1, the number of links has grown 2–10 times. The vertebrate networks have grown the most, which is not surprising as the newly introduced species are also vertebrates. Also, the network of *Arabidopsis thaliana* has grown 8-fold which can be explained, apart from a significant increase in input data from this species, also by the fact that it contains multiple inparalogs (co-orthologues) in clusters with vertebrates. Each inparalog thus receives functional coupling evidence from the orthologue(s).

Table 1. Total network sizes in FunCoup 2.0

Species	Nr of links	Nr of genes in network
A. thaliana	1 943 407	15 278
C. elegans	1 664 577	13 459
C. familiaris	1 749 034	17 550
C. intestinalis	3 97 038	4540
D. melanogaster	1 276 343	11 679
D. rerio	1 999 528	13 033
G. gallus	1 134 553	12 458
H. sapiens	4 675 444	21 087
M. musculus	4 315 860	20 147
R. norvegicus	3 066 419	16 425
S. cerevisiae	449 522	5354

The networks were pruned to only contain links with confidence > 0.1.

For all species, on average about 70% of the links with a *pfc* of 0.5 or higher in FunCoup 1.0 are conserved in FunCoup 2.0. For the most confident links, *pfc* of 0.99 or higher, we even see a conservation of 90%. The observed loss can be explained by changes in the underlying datasets or changes in orthology assignments provided by InParanoid.

Figure 2 shows the relative evidence contribution stratified by data type or species. Compared to version 1.0, the relative data-type contributions are similar, but mRNA co-expression is now even more dominating, accounting for 50–65% of the support. The fractions of support from the species' own data have also increased, although it is still true for all species that more than 50% of the evidence is contributed by other species.

The FunCoup 2.0 networks are scale-free and highly interconnected. Fitting a power law function to the degree frequency distribution gives $P(k) = 0.1\,k^{-0.8}$, where k is the node degree, for the human network. These are the same regression coefficients as for FunCoup 1.0 links with $pfc > 0.1$.

GENE SET ANALYSIS

The FunCoup website features many options and parameter choices under 'More options'. The default values of these were set to suitable settings for single gene queries. However, the website can also be used to analyse large gene sets, up to a few hundred genes. Such gene sets may have been obtained from a functional genomics experiment, for instance all genes that were significantly differentially expressed between two conditions.

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

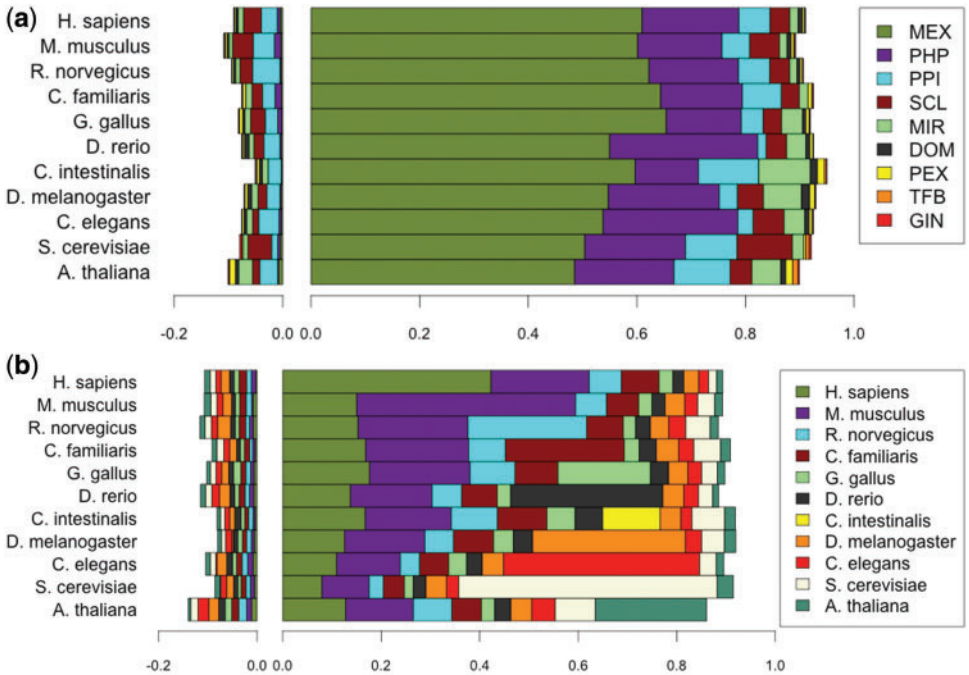

Figure 2. The relative contribution of evidence in FunCoup 2.0 categorized by (**A**) data type and (**B**) species of origin. Positive contributions are shown to the right and negative to the left. The total amount of evidence (LLRs) was normalized within each species so that the negative and positive contributions sum to 1. Evidence data types are: MEX: mRNA co-expression; PHP: phylogenetic profile similarity; PPI: protein–protein interaction; SCL: sub-cellular co-localization; MIR: co-miRNA regulation by shared miRNA targeting; DOM: domain interactions; PEX: protein co-expression; TFB: shared transcription factor binding; GIN: genetic interaction profile similarity.

For gene set analysis, the query settings should be changed. The most important parameter is the Network Distance, i.e. the number of steps to take from the query gene(s). This is by default set to 1, and although it can be increased to 3 this often gives prohibitively large subnetworks for even single queries because FunCoup's networks are rich in hubs. Moreover, as it is a small-world network (average path between two nodes is about 4.5 edges), larger network distances are not always biologically meaningful. Hence, for a large set of query genes, it is recommended to set it to 0, which means that only links between the query genes are searched for (setting it to 1 will often generate many thousands of links). Such large networks are impossible to analyse graphically in jSquid. On the other hand, a cut-off is usually applied to limit the number of links (default 30 most confident), but this would then represents a tiny fraction of all the links.

We thus recommend the following procedure:

(i) Enter gene set identifiers (many types are supported) into the query box.
(ii) Set network distance to 0 and confidence cut-off to 0.5.
(iii) Run query. If the subnetwork appears as a single module rather than as a set of disjoint clusters, consider raising the confidence cut-off. Not that the confidence cut-off can also be raised in jSquid with a slider.
(iv) Identify clusters and select genes with mouse rubberband (drag with left button), select 'copy' from the drop-down menu (right button), and paste cluster member's IDs into a new query box. This is easiest with the option 'Label network nodes with ENSEMBL IDs' as the gene IDs then do not get species prefixes.
(v) Set network distance to 1 and confidence cut-off to 0.5
(vi) Run query. Consider lowering the confidence cut-off and/or increasing the number of links cut-off to get a larger subnetwork.

This analysis can also be done with multiple gene sets, to investigate whether the sets belong to separate network clusters or not. A common application is when two gene sets are obtained by complementary approaches, and one wants to test the hypothesis that they are significantly related. This can currently not be done statistically on the website, but a new separate tool CrossTalkZ can perform such tests.

COMPARATIVE INTERACTOMICS

In comparative genomics, a common strategy is to first map orthologues between species and then carry out a range of different analyses on these to understand their independent evolution since the split from a single gene in the last common ancestor. At a higher level, one can ask the question how conserved entire pathways are between species. This requires a method to identify relevant sub-networks and map them between species. FunCoup provides this for its entire networks, not limited to

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

known pathways. Orthologous genes enable alignment of subnetworks between different species. As FunCoup's networks are incomplete, this can only provide the picture given the current knowledge. Nonetheless, this functionality still gives useful insights into degree of conservation of pathways and other functional modules.

This comparative interactomics feature was already present in FunCoup 1.0, but has been modified to enable more specific studies. A particular caveat to be aware of when running FunCoup in multi-species mode is the fact that FunCoup uses orthology to transfer evidence of functional coupling between species. Therefore, links between orthologues often share the same evidence, and a network alignment of genes whose subnetwork is based on all available evidence does not say much about the actual network conservation given evidence from the species itself. Hence, by default, FunCoup in multi-species mode now displays networks based only on the species' own data. The drawback is that the evidence base becomes highly reduced and few links have high confidence, which can give a very reduced network in some species. To return to the mode when all orthology-transferred evidences are allowed, check the option 'Use evidence from all species'. Such alignments should be interpreted with caution however, as many of the edges that appear conserved are actually based on the same evidence. In this mode, a user should always inspect the species source of the couplings to make sure that they are different. Note that the multi-species mode supports displaying conservation in more than two species simultaneously (up to all the eleven). Examples of such universally conserved sub-networks include e.g. RNA-polymerase sub-units, see Figure 3.

The multi-species mode is activated by checking 'Show sub-network(s) in several organisms' under 'More options'. Here one can choose which species to show the subnetwork in by holding Ctrl and clicking with the mouse. Figure 4 shows an example with subnetworks in human and *Caenorhabditis elegans*. Note that in multi-species mode, genes are coloured according to species and gene names are prefixed by a three-letter species code (not with the option to display ENSEMBL IDs). In this example, we used the human gene RAD50, a DNA repair protein, as a query, and asked for the human and *C. elegans* subnetworks. Several of the neighbours of human RAD50 are orthologues to the neighbours of *C. elegans* rad-50, for instance SMC3, SMC1A, HDAC1/2 and TRRAP. Other neighbours such as SMC6 have orthologues that are linked indirectly to rad-50 in *C. elegans*. Overall, the conservation of this network module is striking given the high evolutionary distance between human and worm, and that the evidences for functional coupling come independently from either species.

PUBLISHED FunCoup USES

FunCoup is linked to by many on-line gene annotation databases. A form of tight integration is realized in the

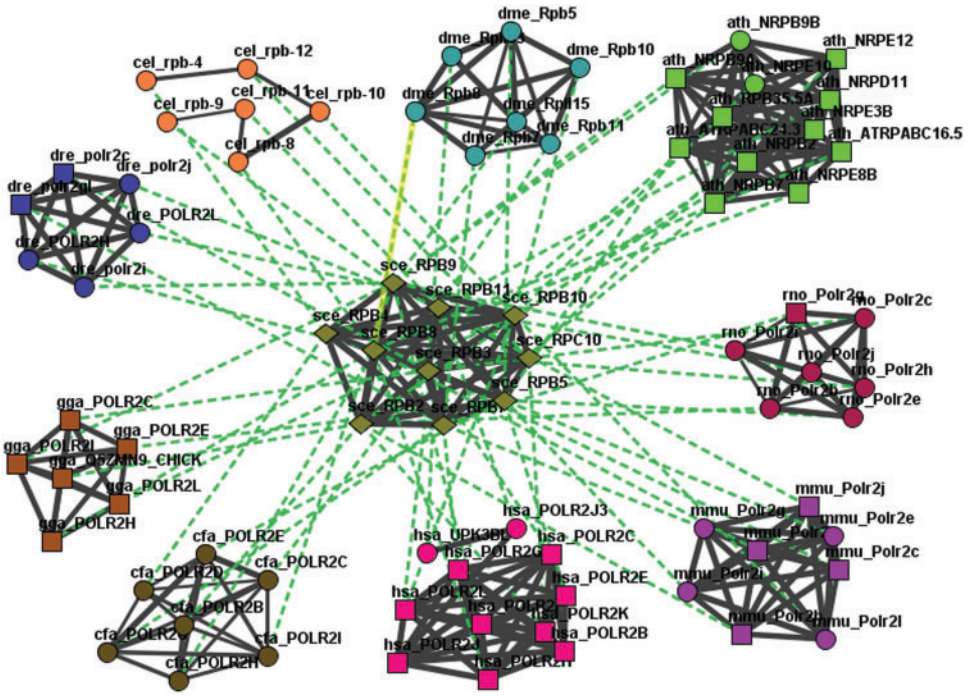

Figure 3. Example of comparative interactomics with FunCoup. Subunits of RNA-polymerase II in *S. cerevisiae* were used as query genes (diamonds in the centre). These were retrieved as genes with ENSEMBL descriptions that contain 'DNA-directed RNA polymerase II * kDa polypeptide': RPB6, RPB11, RPC10, RPB10, RPB7, RPB5, RPB2, RPB9, RPB3, RPB8, RPB4. The subnetwork in all FunCoup species was asked for at network distance 0 (only links between query genes and their orthologs). Green dotted lines connect orthologues, while black solid lines indicate functional coupling. A significant amount of evidence ($pfc > 0.5$) comes from each individual species itself, but for clarity only black summary lines are shown, representing all species' evidences. The nodes are coloured according to species, and labelled with a species prefix (cfa = *Canis familiaris*, etc.).

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

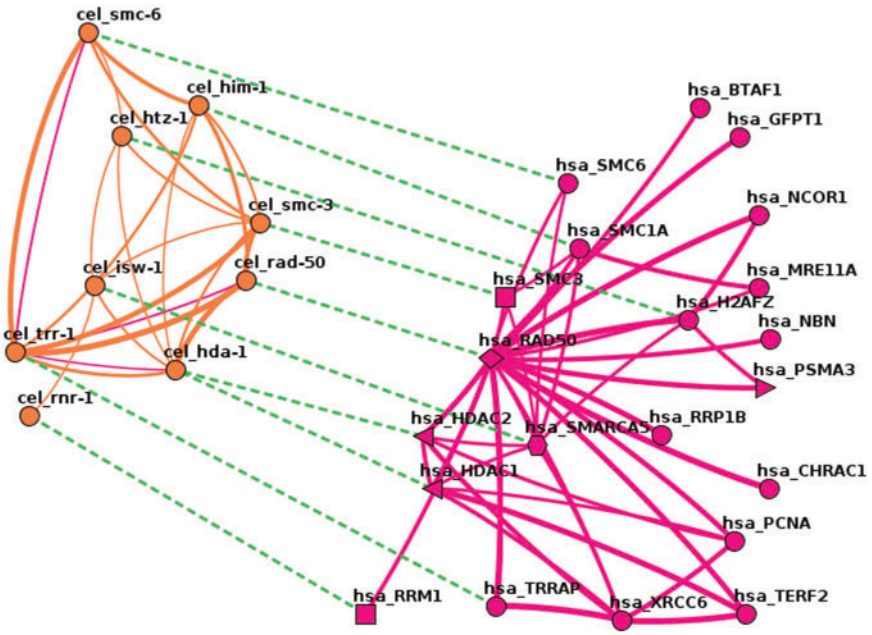

Figure 4. Example of comparative interactomics with FunCoup. The human gene RAD50 (shown as a diamond, the major hub) was used as a query, and subnetworks in human and *C. elegans* with links more confident than 0.5 were asked for. The human subnetwork is shown to the right and the *C. elegans* network to the left. Gene names in respective species are prefixed hsa_ and cel_. *C. elegans* genes are coloured orange, as are supporting functional coupling evidence links from *C. elegans*. Likewise, human genes and evidence are coloured cerise. Note that most of the evidence, but not all, comes from the species itself. Evidence support from any other species was hidden in this graph. Functional coupling links are drawn as solid lines with the width proportional to the confidence, while orthologue links are shown as dashed green lines.

Gerontome database (18) of ageing-related genes. Here, the graphical network viewer jSquid is launched to show the nearest interaction partners predicted by FunCoup.

A common situation in molecular biology is when experiments lead to multiple separated gene clusters. The question is then whether those clusters are significantly associated with each other. For example, ref. 19 looked for biological processes enriched when disabling an oxidative stress response gene and found two distinct processes, proteolysis and ageing. Network analysis with FunCoup revealed a close interconnection between these two clusters, supporting their functional coupling.

Skjølberg *et al*. (20) used FunCoup to investigate and characterize the functional interactions of genes that are differentially expressed after irradiation with ultraviolet light in fission yeast *Schizosaccharomyces pombe*. Since *S. pombe* is currently not part of the FunCoup database the corresponding orthologues in *Saccharomyces cerevisiae* were used for the network analysis. The authors showed that the genes induced by irradiation form a strongly interconnected cluster in FunCoup that involves mainly genes related to translation and transcription.

In both experimental and statistics-based (e.g. genome-wide association studies) biological research, it is important to secure additional evidence that might support or invalidate a certain hypothesis. Reynolds *et al*. (21) used linkage disequilibrium mapping to obtain a list of genes potentially implicated in Alzheimer-related dementia. Using the FunCoup network, the authors analysed the genes' functional relatedness to Alzheimer's disease by the enrichment of common interactors. They found evidence for involvement of previously known Alzheimer genes and one of the novel candidates, TOM1L2. For the rest of the list, no support from the network analysis was found. Thus, the genetic research was successfully complemented with an independent line of evidence.

METHODS

We here list changes in methods compared to version 1.0 and major changes in input data. For a complete list of all 53 input datasets, we refer to the on-line table provided on the FunCoup website under 'Input data'.

New *PPI* score

In FunCoup 1, we did not include prey–prey interactions from large studies. In FunCoup 2.0, we use all prey–prey interactions by introducing a penalty term for them in the PPI score that combines the probabilistic scores S_+ (for being coupled) and S_- (for 'not' being coupled):

$$S_{PPI} = \frac{S_+}{S_+ + S_-}, \text{ where}$$

$$S_- = (1 - P(PPI)) \prod_{P=1}^{|Papers|} \prod_{a=1}^{Assays_p|} pc_-$$

$$S_+ = P(PPI) \prod_{p=1}^{|Papers|} \prod_{a=1}^{|Assays_p|} \frac{pc_+}{\sqrt{|Assays_p(A,B)| \cdot \log_2(|IP_a(A,B,..)|)\pi_{A,B}}}$$

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

S_+ has acquired a new term $\pi_{A,B}$ which penalizes for the number of prey–prey relationships in the assay a. If both A and B appear as preys in a and there are at least one other prey in a then $\pi_{A,B} = \ln(|PP(A,B,..)|)$, where $|PP(A,B,..)|$ represents the number of prey-prey relationships in a. If not, $\pi_{A,B} = 1$.

Thus, the score increases with

(i) the number of individual published reports on the interaction between proteins A and B and
(ii) the number of separate experiments that validated interaction between A and B within the same report

and decreases with

(i) number of partners $|IP_a|$ other than A and B reported in the same interaction in the same experiment, i.e. for multi-protein experiments,
(ii) number of prey–prey interactions in the experiment (if A and B were both preys).

The probabilities

(i) $P(PPI)$, 'an interaction exists between a pair of proteins', 0.001;
(ii) pc_+, 'a single positive report is published given the interaction is true', 0.1; and
(iii) pc_-, 'a single positive report is published given the interaction is false' 0.001

were assigned arbitrarily (to the same values as in FunCoup 1.0).

As a result, we can employ much more information on pairwise relations between proteins than a strict bait–prey approach could. In total, there were 1 446 285 prey–prey relations for the seven organisms for which we could get enough data from IntAct (same list as in FunCoup 1). The increase was very significant for human, *Mus musculus*, *Rattus norvegicus* and *S. cerevisiae,* and not so strong in *A. thaliana* and *C. elegans* (number of available relations less than doubled). The impact of prey–prey relations was relatively weak but significant. Alone they were not sufficient for predicting functional coupling, but they can serve as additional evidence.

In FunCoup 2.0 we switched to only use the IntAct database (22) for PPI data as we reasoned that all reliable interactions previously collected from other PPI sources are already in IntAct.

Domain interactions

We switched to using the UniDomInt database (17) for domain interactions, as it is an amalgamation of nine predicted domain interaction databases. The UniDomInt score, which reflects the level of support among the source databases, was used directly during Bayesian training. In each species, the domain interactions were first mapped to protein pairs using Pfam 25 (23) and then to gene pairs using Ensembl 63 BioMart (24). Interactions with a UniDomInt score of 0 were not used.

Sub-cellular localization

We switched to using the 'filtered annotations' of each species from the Gene Ontology (25). GO terms were autocompleted up to the highest level of the Cellular Component Ontology. Gene identifiers were mapped to ENSEMBL gene identifiers using Ensembl 63 BioMart.

Discretization

Each continuous score was discretized into bins during Bayesian training. In FunCoup 1.0 we used a maximum of 10 bins, but after further testing we found it to be more optimal to set the maximum to seven bins.

FUNDING

Swedish Research Council, Swedish eScience Research Center, and Stockholm University. Funding for open access charge: Swedish Research Council.

Conflict of interest statement. None declared.

REFERENCES

1. Alexeyenko,A. and Sonnhammer,E.L.L. (2009) Global networks of functional coupling in eukaryotes from comprehensive data integration. *Genome Res.*, **19**, 1107–1116.
2. Ostlund,G., Schmitt,T., Forslund,K., Köstler,T., Messina,D.N., Roopra,S., Frings,O. and Sonnhammer,E.L.L. (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.*, **38**, D196–D203.
3. Kao,H.-L. and Gunsalus,K.C. (2008) Browsing multidimensional molecular networks with the generic network browser (N-Browse). *Current Protoc. Bioinform.*, Chapter 9, Unit 9.11.
4. Kamburov,A., Wierling,C., Lehrach,H. and Herwig,R. (2009) ConsensusPathDB–a database for integrating human functional interaction networks. *Nucleic Acids Res.*, **37**, D623–D628.
5. Niu,Y., Otasek,D. and Jurisica,I. (2010) Evaluation of linguistic features useful in extraction of interactions from PubMed; application to annotating known, high-throughput and predicted interactions in I2D. *Bioinformatics*, **26**, 111–119.
6. Warde-Farley,D., Donaldson,S.L., Comes,O., Zuberi,K., Badrawi,R., Chao,P., Franz,M., Grouios,C., Kazi,F., Lopes,C.T. *et al.* (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.*, **38**, W214–W220.
7. Cerami,E.G., Gross,B.E., Demir,E., Rodchenkov,I., Babur,O., Anwar,N., Schultz,N., Bader,G.D. and Sander,C. (2011) Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.*, **39**, D685–D690.
8. Prieto,C. and De Las Rivas,J. (2006) APID: Agile Protein Interaction DataAnalyzer. *Nucleic Acids Res.*, **34**, W298–W302.
9. Hu,Z., Hung,J.-H., Wang,Y., Chang,Y.-C., Huang,C.-L., Huyck,M. and DeLisi,C. (2009) VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. *Nucleic Acids Res.*, **37**, W115–121.
10. Szklarczyk,D., Franceschini,A., Kuhn,M., Simonovic,M., Roth,A., Minguez,P., Doerks,T., Stark,M., Muller,J., Bork,P. *et al.* (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.*, **39**, D561–D568.
11. Klammer,M., Roopra,S. and Sonnhammer,E.L.L. (2008) jSquid: a Java applet for graphical on-line network exploration. *Bioinformatics*, **24**, 1467–1468.
12. Kalaev,M., Smoot,M., Ideker,T. and Sharan,R. (2008) NetworkBLAST: comparative analysis of protein networks. *Bioinformatics*, **24**, 594–596.

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011

13. Liao,C.-S., Lu,K., Baym,M., Singh,R. and Berger,B. (2009) IsoRankN: spectral methods for global alignment of multiple protein networks. *Bioinformatics*, **25**, i253–i258.
14. Flannick,J., Novak,A., Srinivasan,B.S., McAdams,H.H. and Batzoglou,S. (2006) Graemlin: general and robust alignment of multiple large interaction networks. *Genome Res.*, **16**, 1169–1181.
15. Kuchaiev,O., Stevanović,A., Hayes,W. and Pržulj,N. (2011) GraphCrunch 2: software tool for network modeling, alignment and clustering. *BMC Bioinform.*, **12**, 24.
16. Costanzo,M., Baryshnikova,A., Bellay,J., Kim,Y., Spear,E.D., Sevier,C.S., Ding,H., Koh,J.L.Y., Toufighi,K., Mostafavi,S. *et al.* (2010) The genetic landscape of a cell. *Science*, **327**, 425–431.
17. Björkholm,P. and Sonnhammer,E.L.L. (2009) Comparative analysis and unification of domain–domain interaction networks. *Bioinformatics*, **25**, 3020–3025.
18. Kwon,J., Lee,B. and Chung,H. (2010) Gerontome: a web-based database server for aging-related genes and analysis pipelines. *BMC Genomics*, **11(Suppl. 4)**, S20.
19. Fensgård,Ø., Kassahun,H., Bombik,I., Rognes,T., Lindvall,J.M. and Nilsen,H. (2010) A two-tiered compensatory response to loss of DNA repair modulates aging and stress response pathways. *Aging*, **2**, 133–159.
20. Skjølberg,H.C., Fensgård,O., Nilsen,H., Grallert,B. and Boye,E. (2009) Global transcriptional response after exposure of fission yeast cells to ultraviolet light. *BMC Cell Biol.*, **10**, 87.
21. Reynolds,C.A., Hong,M.-G., Eriksson,U.K., Blennow,K., Wiklund,F., Johansson,B., Malmberg,B., Berg,S., Alexeyenko,A., Grönberg,H. *et al.* (2010) Analysis of lipid pathway genes indicates association of sequence variation near SREBF1/TOM1L2/ATPAF2 with dementia risk. *Hum. Mol. Genet.*, **19**, 2068–2078.
22. Aranda,B., Achuthan,P., Alam-Faruque,Y., Armean,I., Bridge,A., Derow,C., Feuermann,M., Ghanbarian,A.T., Kerrien,S., Khadake,J. *et al.* (2010) The IntAct molecular interaction database in 2010. *Nucleic Acids Res.*, **38**, D525–531.
23. Finn,R.D., Mistry,J., Tate,J., Coggill,P., Heger,A., Pollington,J.E., Gavin,O.L., Gunasekaran,P., Ceric,G., Forslund,K. *et al.* (2010) The Pfam protein families database. *Nucleic Acids Res.*, **38**, D211–D22.
24. Flicek,P., Amode,M.R., Barrell,D., Beal,K., Brent,S., Chen,Y., Clapham,P., Coates,G., Fairley,S., Fitzgerald,S. *et al.* (2011) Ensembl 2011. *Nucleic Acids Res.*, **39**, D800–D806.
25. The Gene Ontology Consortium, Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S. *et al.* (2000) Gene ontology: tool for the unification of biology. *Nat. Genet.*, **25**, 25–29.

Downloaded from http://nar.oxfordjournals.org/ at Spris bibliotek och utredningsbank on December 2, 2011