

Functional evaluation of domain-domain interactions and human protein interaction networks

Andreas Schlicker¹, Carola Huthmacher, Fidel Ramírez,
Thomas Lengauer, and Mario Albrecht²

Department of Computational Biology and Applied Algorithmics
Max Planck Institute for Informatics
Stuhlsatzenhausweg 85
66123 Saarbrücken
Germany
andreas.schlicker@mpi-inf.mpg.de
mario.albrecht@mpi-inf.mpg.de

Abstract: Large amounts of protein and domain interaction data are being produced by experimental high-throughput techniques and computational approaches. To gain insight into the value of the provided data, we used our new similarity measure based on the Gene Ontology to evaluate the molecular functions and biological processes of interacting proteins or domains. The applied measure particularly addresses the frequent annotation of proteins or domains with multiple Gene Ontology terms. Using our similarity measure, we compare predicted domain-domain and human protein-protein interactions with experimentally derived interactions. The results show that our similarity measure is of significant benefit in quality assessment and confidence ranking of domain and protein networks. We also derive useful confidence score thresholds for dividing domain interaction predictions into subsets of low and high confidence.

1 Introduction

Experimental high-throughput techniques have produced enormous amounts of protein-protein interaction (PPI) data for different species [1]. These data can now be mined for new information on the functions and interrelationships of proteins [2]. In particular, different bioinformatics methods, mainly based on the homology of protein sequences, have supported the large-scale prediction of human protein networks [3-8]. Additionally, manually curated literature data and four large-scale yeast-2-hybrid maps have recently become available for the human interactome [9-13]. However, in contrast to predicted data, the experimental coverage of the human interactome is still low. To predict protein interaction networks, domain-domain interactions (DDIs) are often taken into account [8, 14-16]. For this purpose, different sets of DDIs have been predicted using bioinformatics methods [16-18] and supplement experimental DDI sets derived from 3D structure data [19, 20].

¹Presenting author at GCB

²Corresponding author

The Gene Ontology (GO) consortium provides a standardized vocabulary that is commonly used to annotate genes and their products with biological processes and molecular functions [21]. This annotation particularly allows for assessing the functional similarity of genes or their products. Resnik [22] and Lin [23] introduced semantic similarity measures for the comparison of single terms in “is-a” ontologies. Both measures are based on the information content of ontology terms. Based on these semantic similarity measures, several methods for the functional comparison of gene products have been introduced. Lu *et al.* [24] and Lin *et al.* [25] evaluated the usefulness of different features, ranging from expression profiles to functional relationships between genes, for the prediction of PPIs. They concluded that functional similarity based on GO annotation leads to high accuracy in predicting PPIs. Wu *et al.* also introduced new similarity measures between GO terms and proteins [26]. Their measures were used to create a predicted network of PPIs and to evaluate genome-scale datasets. Very recently, Guo *et al.* assessed the applicability of GO-based similarity measures to human regulatory pathways [27]. They showed that the functional similarity between two proteins decreases as their distance within the same regulatory pathway increases.

One problem with existing GO-based similarity measures is that they do not account for the frequent annotations of gene products or protein domains with multiple GO terms or that they simply average over all annotations. To address this problem, we use our novel GO similarity measure that explicitly deals with this functional multiplicity [28]. The measure is applied to ranking the interaction networks and the corresponding prediction methods based on the overall functional similarity of the interacting proteins or domains. The comparison of experimentally derived sets with predicted sets of DDIs using our GO similarity measure results in confidence score thresholds separating low- and high-confidence subsets of predicted DDIs. In addition, we utilize our measure to analyze experimental and predicted networks of human protein interactions.³

2 Materials and Methods

2.1 Experimental and predicted datasets

Two experimental sets of DDIs were taken from iPfam [19] and the database of 3D interacting domains (3did) [20] and compared to three sets of predicted interactions between Pfam-A domains [29]. The first predicted set is InterDom, a database of putatively interacting domains compiled from different data sources [17]. The other two sets are taken from two recent publications by Liu, Liu, and Zhao (LLZ) [16] and by Riley *et al.* (domain pair exclusion analysis, DPEA). Their bioinformatics approaches are methodological extensions of an expectation-maximization algorithm first applied to the prediction of domain interactions by Deng *et al.* in 2002 [15]. The DDI prediction methods assign a confidence score (CS) to each DDI and rank the predicted DDIs

³Abbreviations: ATX, ataxin; BP, biological process; CS, confidence score; DDI, domain-domain interaction; GO, Gene Ontology; HTT, huntingtin; MF, molecular function; PPI, protein-protein interaction; Y2H, yeast two-hybrid.

according to the score. InterDom uses different data sources to infer DDIs and calculates the CS based on the support from each source [17]. LLZ and DPEA compute maximum-likelihood estimates to derive a CS, and we use the probability λ and the log-odds score E as CS from LLZ and DPEA, respectively [16, 18]. The pfam2go file from the GO web site (<http://www.geneontology.org/external2go/pfam2go>) contains a mapping of Pfam-A domains to GO terms. This file (downloaded on July 7, 2005) was used to annotate the Pfam-A domains with GO terms. Table 1 summarizes the number of DDIs in each dataset.

Table 1: Total number of Pfam-A domains in the different datasets of DDIs (column 'Total'). The columns for biological process ('BP') and molecular function ('MF') contain the fraction of interactions whose interacting domains are both annotated with GO.

Dataset	Total	BP (%)	MF (%)
iPfam	3,046	52.07	56.30
3did	3,034	49.51	54.19
InterDom	29,957	27.07	37.64
LLZ	9,160	17.75	19.64
DPEA	3,005	22.40	24.19

We also analyzed six predicted sets of human PPIs named Bioverse [6], HiMAP [8], HomoMINT [7], Sanger [4], OPHID [5], and POINT [3]. Additionally, subsets of core interactions with high confidence were derived from Bioverse, HiMAP and Sanger. The Bioverse-core set contains very reliable interactions based on a sequence similarity threshold of at least 80% between human and the homolog of the source species [30], HiMAP-core interactions have a large likelihood ratio [8], and Sanger-core comprises only predictions with the greatest experimental support [4]. Additionally, we assembled five consensus sets named ConSet n that consist of protein interactions contained in at least n predicted interactomes, with n ranging from 2 to 6.

As experimental datasets, we downloaded the manually curated human protein reference database (HPRD) [13], release of 13 September 2005, and two yeast two-hybrid (Y2H) maps that we named 'Vidal' [10] and 'Wanker' [11] after the senior authors. We also merged the two Y2H maps into the combined dataset Vidal & Wanker. Both Y2H maps and the HPRD data became available after the six predicted human networks were published. Further experimental PPIs were extracted from the published networks of direct and indirect interaction partners for ataxins (ATX) [12] and huntingtin (HTT) [9]. These networks include Y2H and literature-derived datasets, which we call ATX-/HTT-Y2H and ATX-/HTT-literature, respectively. The ATX-interologs set comprises interactions from the ATX network that have been derived by mapping interologs [12], and thus we regard it as another predicted set of PPIs. Generally, the diverse gene and protein accession numbers of the PPI sets were mapped to NCBI Entrez gene identifiers [31]. The mapping of Entrez gene identifiers to GO annotations was obtained from NCBI (<ftp://ftp.ncbi.nih.gov/gene/DATA/gene2go.gz>). Furthermore, we compiled another set of PPIs using the interacting proteins that underlie iPfam DDIs with both domains belonging to different proteins. This set was annotated from two different sources, that is, with the GO annotation from the UniProt release 5.4 (IUP-set) and with GO terms from the pfam2go file (IPG-set).

2.2 Functional similarity measure

The GO controlled vocabulary consists of three different ontologies: biological process (BP), molecular function (MF), and cellular component. The ontologies are organized as directed acyclic graphs with terms being represented as nodes and parent-child relationships as edges. There are two types of edges: “is-a” links, indicating that the child is an instance of its parent, and “part-of”, used if the child is a component of its parent. Each node may have several parents and children.

Our semantic similarity measure is an extension of previous measures by Resnik and Lin [22, 23]. As suggested by Resnik, we defined the probability of a term as its relative frequency of occurrence in a set of annotated gene products. The root node of each ontology has the probability 1. We used the GO annotation of all proteins in the UniProt release 5.4 for the calculation of term frequencies. The semantic similarity of two terms is defined as follows:

$$sim(t_1, t_2) = \max_{a \in CA} \left(\frac{2 * \log p(a)}{\log p(t_1) + \log p(t_2)} * (1 - p(a)) \right),$$

where t_1 and t_2 are GO terms, $p(t_1)$ and $p(t_2)$ their probabilities, and CA is the set of their common ancestors in the graph. This similarity measure takes into account how similar and detailed both terms t_1 and t_2 are, and it ranges from 0 (for terms that only have the root node in common) to 1.

This semantic similarity measure for single GO terms can be expanded to a functional similarity measure of gene products. Let g_1 and g_2 be two gene products annotated with the GO term sets GO^1 and GO^2 of size N and M , respectively. The similarity matrix S containing all pair-wise similarity values is computed as

$$s_{ij} = sim(GO_i^1, GO_j^2), \forall i \in \{0, \dots, N\}, \forall j \in \{0, \dots, M\}.$$

The row vectors and column vectors of matrix S represent the two possible directions of comparing g_1 and g_2 . While the similarity computed from g_1 to g_2 (*rowScore*) is defined as the average over the row maxima, the similarity from g_2 to g_1 (*columnScore*) is defined as the average over the column maxima:

$$rowScore = \frac{1}{N} \sum_{i=1}^N \max_{1 \leq j \leq M} s_{ij}; \quad columnScore = \frac{1}{M} \sum_{j=1}^M \max_{1 \leq i \leq N} s_{ij}.$$

The *rowScore* and the *columnScore* are always between 0 and 1. Furthermore, we define the functional similarity of two gene products with respect to one ontology as

$$GOscore(g_1, g_2) = \max\{rowScore(g_1, g_2), columnScore(g_1, g_2)\}.$$

We refer to this *GOscore* as *MFscore* for MF and *BPscore* for BP. One important aspect of this score is that it allows for comparing gene products with multiple functions. This

property is especially important when comparing GO annotations of domains because they occur in diverse proteins involved in different processes. For more details on our GO similarity measure, see Schlicker *et al.* [28].

3 Results and Discussion

3.1 Comparing confidence scores for domain interactions

The predictions of DDIs by InterDom, LLZ and DPEA are compiled from diverse data sources using different bioinformatics methods. To gain insight into the similarity and the quality of the predictions, we compared the predicted sets of DDIs with each other and to the experimentally derived sets iPfam and 3did. The overlap of the datasets InterDom, LLZ and DPEA regarding Pfam-A domains as well as regarding their predicted interactions are given in Table 2. LLZ and DPEA share many Pfam-A domains and predicted DDIs with InterDom, while the overlap between LLZ and DPEA is much smaller.

Table 2: Overlap of the InterDom, LLZ and DPEA datasets with regard to Pfam-A domains and predicted domain interactions. Each number refers to the percentage of domains or interactions in the row datasets that are also contained in the respective column dataset. Percentages in parentheses give the number of DDIs shared between two datasets in ratio to the overall number of DDIs with interacting domains contained in both datasets.

	Pfam-A domains (%)			Domain-domain interactions (%)		
Dataset	InterDom	LLZ	DPEA	InterDom	LLZ	DPEA
InterDom	100.0	44.4	25.1	100.0 (100.0)	11.4 (19.3)	4.8 (23.2)
LLZ	79.3	100.0	26.9	58.8 (72.7)	100.0 (100.0)	10.6 (60.8)
DPEA	86.5	51.9	100.0	78.9 (89.3)	32.9 (62.2)	100.0 (100.0)

Figure 1 and Table S1 give an overview of the overlap of the experimental interactions contained in iPfam and 3did and the three sets of predicted interactions InterDom, LLZ and DPEA. 11.8% of the DDIs predicted by DPEA are confirmed by iPfam or 3did, whereas only 7.4% and 3.0% of the DDIs predicted by InterDom and LLZ, respectively, are in common with iPfam or 3did. Thus, DPEA appears to be the best of the three prediction methods.

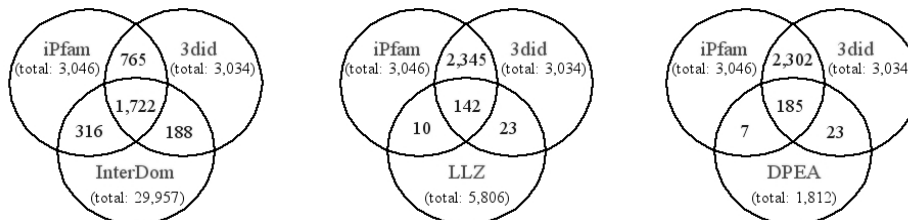


Fig. 1: Overlap of the datasets containing predicted or experimental Pfam-A domain interactions.

Other criteria for prediction quality are the CS and the rank assigned to domain interactions observed experimentally. The distributions of CSs show that many interactions in iPfam and 3did receive a high CS by LLZ and a low CS by InterDom and DPEA (Figure S1). However, DDIs contained in iPfam and 3did are assigned top ranks by all three prediction methods (Figure S2). Surprisingly, further analyses indicate only weak correlations between CSs and ranks of different prediction methods (Figures S3-S5). However, DDIs from iPfam that are predicted by two different computational methods are assigned a good rank by at least one method. This suggests that all methods are able to detect correct domain interactions. Further details on the results are described in the online supplement.

3.2 Background distribution and randomized domain networks

In order to obtain a background distribution, all available Pfam-A domains (release 17.0) were mapped to BP and MF terms of GO, and the distributions of the *MFscore* and *BPscore* for all pairs of Pfam-A domains were calculated (Figure S6). Apparently, most domain pairs have very low *MFscore*, which indicates that the molecular functions of the domains are generally quite distinct. The mean is about 0.1 and the median is 0. The *BPscore* is distributed similarly, but there are fewer domain pairs with *BPscore* below 0.1. This finding is also reflected by the higher mean and median of 0.23 and 0.17, respectively. These results indicate that the *BPscore* should generally be higher than the *MFscore*.

Subsequently, we randomized all DDI networks in our analysis to determine a possible bias towards specific functions or processes. This was accomplished by keeping one of the two nodes of the interaction edges fixed while randomly shuffling the other nodes of the edges. The obtained distributions are all very similar and closely resemble the background distribution for BP and MF (Figures S7 and S8). The distributions of the randomized experimental iPfam and 3did networks for BP contain more DDIs with *BPscore* below 0.1, but fewer with *BPscore* between 0.1 and 0.2 in contrast to the predicted datasets. The means and medians of all randomized experimental and predicted networks are similar, suggesting that neither of the networks is biased towards specific processes or functions.

3.3 Computing and analyzing *GOscore* distributions

The *BPscore* distributions for iPfam and 3did (Figure 2) show that most DDIs have a very high similarity score exceeding 0.8, which means that the corresponding interacting domains are part of the same process or closely related processes. This is supported by high means of about 0.9 and medians of almost 1. The distributions for the predicted sets InterDom or DPEA look alike. Interestingly, only one third of the predicted interactions have a *BPscore* above 0.8. Furthermore, both datasets include a large fraction of interactions with *BPscore* below 0.4, indicating almost no functional similarity between the domains. The mean is 0.51 for both datasets and the medians 0.39 and 0.41 for InterDom and DPEA, respectively. The LLZ predictions contain substantially fewer interactions with high *BPscore*, and many more interactions with very low *BPscore*. This is reflected by the relatively low mean of 0.35 and the median of 0.2. In summary,

DPEA performs slightly better than InterDom, and both show a better performance than LLZ.

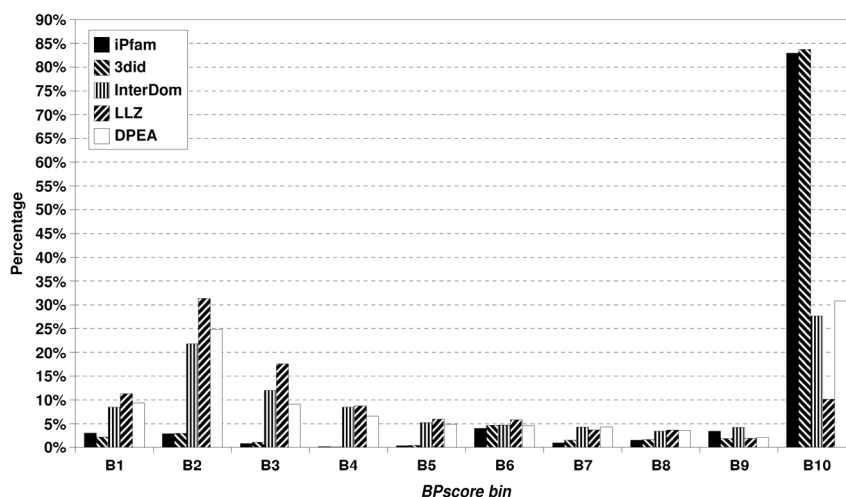


Fig. 2: BPscore distribution for the different datasets of experimental DDIs (iPfam and 3did) and predicted DDIs (InterDom, LLZ and DPEA). The BPscore bins correspond to the following intervals: B1: [0.0, 0.1[; B2: [0.1, 0.2[; B3: [0.2, 0.3[; B4: [0.3, 0.4[; B5: [0.4, 0.5[; B6: [0.5, 0.6[; B7: [0.6, 0.7[; B8: [0.7, 0.8[; B9: [0.8, 0.9[; B10: [0.9, 1.0].

Figure S9 contains the *MFscore* distributions of all datasets. Interestingly, the distributions for iPfam and 3did are quite distinct from the other distributions. Almost 80% of the domain interactions in iPfam or 3did have an *MFscore* above 0.8, which indicates related molecular functions annotated to the interacting domains. In addition, both sets contain very few interactions with very low *MFscore*. The means of over 0.8 and the medians of almost 1 corroborate this interpretation. The predictions made by InterDom and DPEA show similar distributions, but rather low means and medians. Similar to the findings for the *BPscore* distribution, predictions made by LLZ show a lower *MFscore*. As in the case of the *BPscore* distribution, InterDom and DPEA have similar performance and both perform significantly better than LLZ.

3.4 Deriving confidence score thresholds

The methods InterDom, LLZ and DPEA all provide CSs for the prediction of DDIs. However, in order to utilize sets of predicted interactions in practice, it is important to derive reasonable thresholds for low- and high-confidence sets of DDIs. It is to be expected that the functional similarity of domains predicted to interact increases as the confidence in these predictions rises. To verify this expectation, we used different CS thresholds to calculate the *GOscore* means and medians of all interactions with a CS larger than the respective threshold. We also calculated the overlap of these interactions with iPfam and 3did.

Figure 3 shows the change in *BPscore* mean and median, and the change in dataset size with varying CS threshold for the DPEA dataset. When raising the DPEA CS threshold

from 3 to 6, the *BPscore* median increases from slightly over 0.4 to almost 1, and the mean raises from 0.51 to approximately 0.7. The *MFscore* median and the overlap with iPfam and 3did show a steep increase in this CS range (Figures S10 and S11). Consequently, we suggest assigning predictions with a CS between 3 and 6, and above 6 to DPEA subsets of low- and high-confidence DDIs, respectively.

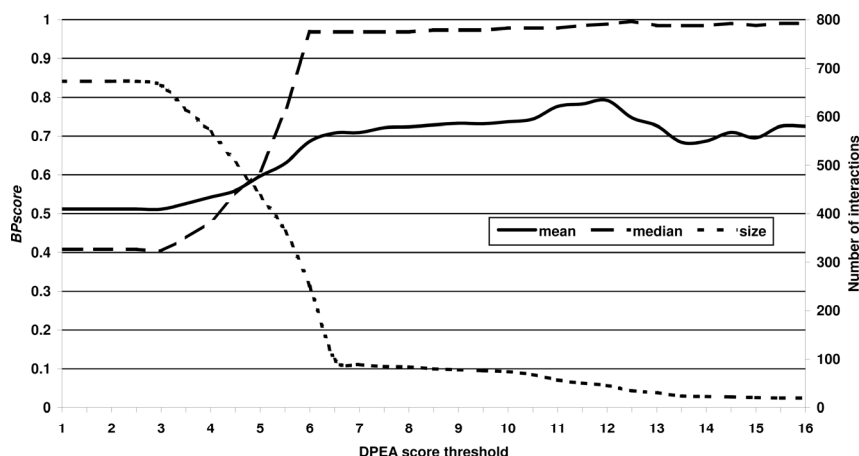


Fig. 3: Change in *BPscore* mean and median, and in dataset size with varying confidence score threshold for DPEA. Size refers to the number of DDIs with confidence score above the threshold.

The analysis of the InterDom set reveals that the *BPscore* median reaches 0.98 with a CS threshold of 30 (Figure S12). The *BPscore* mean is 0.68 at this point and increases with higher thresholds. The same score development holds true for *MFscore*, but it is shifted slightly towards higher thresholds (Figure S13). At a threshold of 60, the dataset consists of 1,888 interactions and the median increase diminishes. The overlap with iPfam and 3did increases with rising InterDom score and is about 27% for a threshold of 60 (Figure S14). Altogether, these results suggest a threshold of 60 for InterDom predictions of high confidence.

The analysis of LLZ predictions reveals that the *BPscore* mean and median, and the overlap with iPfam and 3did are very low over the whole CS range (Figures S15-S17). These results do not allow deriving any reasonable CS threshold for some LLZ subset of DDIs.

3.5 Comparing human protein interaction networks

We calculated the *BPscore* for all datasets of PPIs. Table 3 summarizes the results ranked by the average *BPscore*. The *BPscore* means range from 0.82 for Bioverse-core to 0.37 for Wanker PPI set. While the average *BPscores* for the predicted datasets vary significantly, the experimental Y2H datasets have rather low mean *BPscore*. In contrast, predicted datasets such as both HiMAP datasets and Bioverse-core as well as the manually curated sets HPRD and HTT-literature receive high mean scores. The different results for the HTT and ATX networks also indicate that literature-curated, carefully

validated, PPIs reach a higher *BPscore* than PPIs derived by high-throughput experiments.

Table 3: Ranking of predicted and experimental protein networks based on *BPscore*. The column 'Scored' contains the fraction of PPIs with an assigned *BPscore*. The two rightmost columns give the percentages of PPIs contained in HPRD or the combined Y2H set Vidal & Wanker.

Dataset	Interactions	Scored (%)	mean <i>BPscore</i>	HPRD (%)	Vidal & Wanker (%)
Bioverse-core	3,266	83.2	0.823	28.9	1.1
IPG-set	1,931	45.9	0.815	15.9	0.7
HiMAP-core	8,832	84.6	0.813	9.1	0.6
HiMAP	38,378	89.4	0.799	3.8	0.2
IUP-set	1,931	22.8	0.764	15.9	0.7
ConSet6	484	77.5	0.709	21.3	1.2
HPRD	20,121	86.1	0.662	100.0	0.6
HTT-literature	428	97.4	0.643	90.2	0.2
ConSet5	1,565	73.2	0.642	16.1	1.3
Bioverse	233,941	81.4	0.572	1.5	0.1
ConSet3	10,844	66.5	0.561	9.2	0.8
ConSet4	4,747	67.1	0.559	10.2	0.9
ConSet2	38,258	69.3	0.556	6.0	0.4
Sanger-core	11,131	65.3	0.551	4.5	0.6
ATX-literature	4,796	67.5	0.537	46.9	39.1
HomoMINT	10,870	57.5	0.510	5.6	0.7
OPHID	28,255	62.6	0.499	4.4	0.2
Vidal	2,754	40.2	0.471	3.5	100.0
HTT-Y2H	164	62.2	0.456	3.8	5.1
POINT	98,528	56.9	0.451	2.6	0.2
Sanger	67,518	62.3	0.427	1.3	0.1
ATX-interologs	1,527	62.0	0.418	6.8	1.2
ATX-Y2H	770	39.9	0.394	1.4	1.0
Wanker	2,033	54.8	0.370	1.2	100.0

The *BPscore* means of the iPfam-derived IUP- and IPG-sets with the same PPIs, but distinct GO annotations, are 0.76 and 0.81, respectively. These values are lower than the mean of the corresponding DDIs in iPfam, which may be partly due to the fact that we excluded self-interactions in the two PPI sets. The score distributions for the IUP- and IPG-sets show that using the GO annotation of proteins or Pfam domains leads to different results (Figure S18). In contrast to the small increase in mean *BPscore*, the distributions of the IUP- and IPG-sets differ significantly. In comparison, the manually curated HPRD set has a mean similarity measure of 0.66. The distribution of this set shows that over 50% of the interactions have a *BPscore* above 0.7 (Figure S20). However, 10% of the interactions have a score between 0.1 and 0.2. The consensus PPI

sets ConSet1-4 show a similar mean *BPscore*, and ConSet5 and ConSet6 score higher, but they constitute small interaction sets only.

Especially on the lower ranks, the *BPscore* ranking of the datasets is similar to rankings resulting from the computed HPRD or Y2H verification rate (Table 3), that is, the percentage of interactions contained in HPRD or the combined Y2H dataset Vidal & Wanker. The predicted Bioverse-core set and the consensus sets have the best verification rates with respect to HPRD. The fact that the Vidal and Wanker sets have published validation rates of 78% and 62-66%, respectively, agrees well with the slightly higher mean *BPscore* 0.47 of Vidal in contrast to the mean 0.36 of Wanker [10, 11]. The lower mean *BPscore* of Wanker may also be due to the use of many protein fragments in contrast to full-length proteins employed by Vidal [10, 11].

4 Conclusions

Following the idea that interacting domains or proteins should have highly similar biological process (BP) annotation and, to a smaller degree, similar molecular function (MF) annotation, we evaluated the functional similarity of three predicted and two experimental domain-domain interaction (DDI) networks as well as several predicted and experimental human protein-protein interaction (PPI) networks. Furthermore, we analyzed to which extent predicted DDIs or PPIs overlap with experimentally derived interactions.

We demonstrated that the application of functional similarity measures is not restricted to the validation of PPIs [27], but also useful for DDIs. Our analysis of DDIs revealed that the BP similarity of interacting domains is generally higher than the corresponding MF similarity. This observed difference between BP and MF similarity agrees well with findings by Guo *et al.* for PPIs using other GO similarity measures [27]. The difference may be partly due to the fact that interacting domains or proteins may perform different functions though they act in similar processes. Another reason may be that GO terms are more densely connected in the top levels of the BP ontology than of the MF ontology.

The iPfam-derived IUP- and IPG-sets encompass the same PPIs, but the IUP-set is annotated with the GO terms of the proteins in UniProt and the IPG-set with the GO terms of the Pfam domains. The comparison of these two sets revealed that the *BPscore* results depend on the annotation used. This indicates that the choice of the annotation source contributes to the differing findings for DDIs and PPIs. Moreover, a higher number of proteins annotated with diverse BPs may decrease the mean *BPscore* of protein networks in contrast to sets of DDIs annotated with more generic GO terms.

In agreement with our results on human protein interaction networks, Reguly *et al.* observed for yeast interaction datasets that the GO annotation of literature-curated PPI sets is more coherent than the GO annotation of high-throughput PPI sets [32]. Since manually curated datasets of PPIs taken from scientific literature have a higher mean *BPscore* than most predicted and high-throughput sets, the latter sets may contain a significant number of false interactions or a large amount of proteins involved in novel processes. This can lead to a considerable decrease in *BPscore*. Furthermore, proteins described in the literature may be annotated particularly well with GO. Therefore, a

more thorough analysis of the PPI results using alternative measures will be required to explain differences between predicted and experimental datasets.

Our functional similarity analysis in conjunction with an evaluation of the overlap between experimentally derived and predicted DDIs allowed the definition of confidence score thresholds for DDI prediction results. These thresholds are useful for improving PPI predictions based on DDIs as well as for assessing the confidence of PPIs derived by high-throughput experiments. In the future, incorporating other similarity criteria besides GO may improve the confidence assessment of predicted interactions further. As the coverage and quality of GO annotations improves, the importance of approaches that use functional similarity for the validation and prediction of PPIs and DDIs will increase.

Acknowledgements

We are grateful to Francisco S. Domingues and the anonymous reviewers for useful comments on the manuscript. Part of this study was financially supported by the German National Genome Research Network (NGFN) and by the German Research Foundation (DFG), contract number KFO 129/1-1. This work was conducted in the context of the BioSapiens Network of Excellence funded by the European Commission under grant number LSHG-CT-2003-503265.

References

1. Sharan, R., Ideker, T.: Modeling cellular machinery through biological network comparison. *Nat Biotechnol* **24** (2006) 427-433
2. Bork, P., Jensen, L.J., von Mering, C., Ramani, A.K., Lee, I., Marcotte, E.M.: Protein interaction networks from yeast to human. *Curr Opin Struct Biol* **14** (2004) 292-299
3. Huang, T.W., Tien, A.C., Huang, W.S., Lee, Y.C., Peng, C.L., Tseng, H.H., Kao, C.Y., Huang, C.Y.: POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome. *Bioinformatics* **20** (2004) 3273-3276
4. Lehner, B., Fraser, A.G.: A first-draft human protein-interaction map. *Genome Biol* **5** (2004) R63
5. Brown, K.R., Jurisica, I.: Online predicted human interaction database. *Bioinformatics* **21** (2005) 2076-2082
6. McDermott, J., Bumgarner, R., Samudrala, R.: Functional annotation from predicted protein interaction networks. *Bioinformatics* **21** (2005) 3217-3226
7. Persico, M., Ceol, A., Gavrila, C., Hoffmann, R., Florio, A., Cesareni, G.: HomoMINT: an inferred human network based on orthology mapping of protein interactions discovered in model organisms. *BMC Bioinformatics* **6 Suppl 4** (2005) S21
8. Rhodes, D.R., Tomlins, S.A., Varambally, S., Mahavisno, V., Barrette, T., Kalyana-Sundaram, S., Ghosh, D., Pandey, A., Chinnaiyan, A.M.: Probabilistic model of the human protein-protein interaction network. *Nat Biotechnol* **23** (2005) 951-959
9. Goehler, H., Lalowski, M., Stelzl, U., Waelter, S., Stroedicke, M., Worm, U., Droege, A., Lindenberg, K.S., Knoblich, M., Haenig, C., et al.: A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. *Mol Cell* **15** (2004) 853-865
10. Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., et al.: Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437** (2005) 1173-1178

11. Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., et al.: A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122** (2005) 957-968
12. Lim, J., Hao, T., Shaw, C., Patel, A.J., Szabo, G., Rual, J.F., Fisk, C.J., Li, N., Smolyar, A., Hill, D.E., et al.: A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* **125** (2006) 801-814
13. Mishra, G.R., Suresh, M., Kumaran, K., Kannabiran, N., Suresh, S., Bala, P., Shivakumar, K., Anuradha, N., Reddy, R., Raghavan, T.M., et al.: Human protein reference database - 2006 update. *Nucleic Acids Res* **34** (2006) D411-414
14. Wojcik, J., Schachter, V.: Protein-protein interaction map inference using interacting domain profile pairs. *Bioinformatics* **17 Suppl 1** (2001) S296-305
15. Deng, M., Mehta, S., Sun, F., Chen, T.: Inferring domain-domain interactions from protein-protein interactions. *Genome Res* **12** (2002) 1540-1548
16. Liu, Y., Liu, N., Zhao, H.: Inferring protein-protein interactions through high-throughput interaction data from diverse organisms. *Bioinformatics* **21** (2005) 3279-3285
17. Ng, S.K., Zhang, Z., Tan, S.H., Lin, K.: InterDom: a database of putative interacting protein domains for validating predicted protein interactions and complexes. *Nucleic Acids Res* **31** (2003) 251-254
18. Riley, R., Lee, C., Sabatti, C., Eisenberg, D.: Inferring protein domain interactions from databases of interacting proteins. *Genome Biol* **6** (2005) R89
19. Finn, R.D., Marshall, M., Bateman, A.: iPfam: visualization of protein-protein interactions in PDB at domain and amino acid resolutions. *Bioinformatics* **21** (2005) 410-412
20. Stein, A., Russell, R.B., Aloy, P.: 3did: interacting protein domains of known three-dimensional structure. *Nucleic Acids Res* **33** (2005) D413-417
21. Gene Ontology Consortium: The Gene Ontology (GO) project in 2006. *Nucleic Acids Res* **34** (2006) D322-326
22. Resnik, P.: Using information content to evaluate semantic similarity in a taxonomy. *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (1995) 448-453
23. Lin, D.: An information-theoretic definition of similarity. *Proceedings of the Fifteenth International Conference on Machine Learning* (1998) 296-304
24. Lu, L.J., Xia, Y., Paccanaro, A., Yu, H., Gerstein, M.: Assessing the limits of genomic data integration for predicting protein networks. *Genome Res* **15** (2005) 945-953
25. Lin, N., Wu, B., Jansen, R., Gerstein, M., Zhao, H.: Information assessment on predicting protein-protein interactions. *BMC Bioinformatics* **5** (2004) 154
26. Wu, X., Zhu, L., Guo, J., Zhang, D.Y., Lin, K.: Prediction of yeast protein-protein interaction network: insights from the Gene Ontology and annotations. *Nucleic Acids Res* **34** (2006) 2137-2150
27. Guo, X., Liu, R., Shriver, C.D., Hu, H., Liebman, M.N.: Assessing semantic similarity measures for the characterization of human regulatory pathways. *Bioinformatics* **22** (2006) 967-973
28. Schlicker, A., Domingues, F.S., Rahnenführer, J., Lengauer, T.: A new measure for functional similarity of gene products based on Gene Ontology. *BMC Bioinformatics* **7** (2006) 302
29. Finn, R.D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., et al.: Pfam: clans, web tools and services. *Nucleic Acids Res* **34** (2006) D247-251
30. Yu, H., Luscombe, N.M., Lu, H.X., Zhu, X., Xia, Y., Han, J.D., Bertin, N., Chung, S., Vidal, M., Gerstein, M.: Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res* **14** (2004) 1107-1118
31. Maglott, D., Ostell, J., Pruitt, K.D., Tatusova, T.: Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* **33** (2005) D54-58
32. Reguly, T., Breitkreutz, A., Boucher, L., Breitkreutz, B.J., Hon, G.C., Myers, C.L., Parsons, A., Friesen, H., Oughtred, R., Tong, A., et al.: Comprehensive curation and analysis of global interaction networks in *Saccharomyces cerevisiae*. *J Biol* **5** (2006) 11