BMC Bioinformatics

RESEARCH ARTICLE　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Genome and network visualization facilitates the analyses of the effects of drugs and mutations on protein-protein and drug-protein networks

Arnaud Céol[*], Lisette G. G. C. Verhoef, Mark Wade and Heiko Muller

Abstract

Background: Biologists generally interrogate genomics data using web-based genome browsers that have limited analytical potential. New generation genome browsers such as the Integrated Genome Browser (IGB) have largely overcome this limitation and permit customized analyses to be implemented using plugins. We illustrate the use of a plugin for IGB that exploits advanced visualization techniques to integrate the analysis of genomics data with network and structural approaches.

Results: We show how visualization technologies that combine both genomics and network biology can facilitate the selection of the key amino acid contacts from protein-protein and protein-drug interactions. Starting from the MDM2-P53 interaction, which is a high-value target for cancer therapy, and Nutlin, the parent small molecule of an MDM2 antagonist that is currently in clinical trials, we show that this method can be generalized to analyze how drugs and mutations can interfere with both protein-protein and drug-protein networks. We illustrate this point by two additional use-cases exploring the molecular basis of tamoxifen side effects and of drug resistance in chronic myeloid leukemia patients.

Conclusions: Combined network and structure biology approaches provide key insights into both the genetic and the edgetic roles of variants in diseases. 3D interactomes facilitate the identification of disease-relevant interactions that can then be specifically targeted by drugs. Recent advances in molecular interaction and structure visualization tools have greatly simplified the mapping of mutated residues to molecular interaction interfaces. Such approaches can now also be integrated with genome visualization tools to enable comparative analyses of interaction contacts.

Keywords: Protein networks, Structures, Genome variants, Genome browsing, Visualization, Protein-protein interaction, Protein-ligand interactions

---

* Correspondence: arnaud.ceol@iit.it
Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di
Tecnologia (IIT), Via Adamello 16, Milan I-20139, Italy

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 216 of 269

Background

The wealth of available genomics data has created an analysis bottleneck; efficient means of candidate gene prioritization that permit the integration, comparison and interpretation of genome-scale molecular information are therefore urgently needed. In particular, network and structural biology criteria must be applied to candidate genes in order to pinpoint their contribution to the development of cancer-specific phenotypic traits and to evaluate their druggability. In general terms, the former can be evaluated by studying the connectivity of candidate genes with bona-fide cancer genes in protein-protein interaction (PPI) networks, while the latter can be deduced from structural information on the protein of interest.

Disease networks (networks of disorders and disease proteins) facilitate the identification of connections between disease-causing gene defects [1, 2] and mapping genomic regions to the networks of interactions with structures (3D networks) allows data-driven hypothesis generation about mutations that are likely to have edgetic effects [3], i.e. a loss or gain of interaction. Protein-ligand interactions (PLIs) can reveal valuable hints about the druggability of candidate genes as well as the contact regions of their binding partners. The development of drugs interfering with PPIs is challenging due to the nature of the PPI interface. The first successful example was the 1986 report of peptide-mediated inhibition of Herpes simplex type I ribonucleotide reductase [4, 5]. Since then, important progress has been made. It has been recognized that druggable PPIs largely fall into four different classes [6]. The first class employs short peptide epitopes that are bound by another protein. The second class is based on secondary structure epitopes where a single peptide from one protein binds to a groove presented by the binding partner. The major part of the binding energy is provided by a small number of amino acids called hotspots. The third class of PPIs utilizes large and shallow globular regions called tertiary structure epitopes. The fourth class is based on allosteric interactions. The second class of PPIs has sparked particular interest due to the presence of binding pockets and hotspot amino acids that can be identified by experimental and computational methods, and whose function can be disrupted by relatively small compounds developed using traditional medicinal chemistry [6]. It is reasonable to assume that hotspot amino acids are preferred targets of mutations in cancer and other diseases. Therefore, computational tools facilitating the mapping of mutations identified in genomic screens to protein structures are expected to be of considerable help in identifying hotspots.

Knowledge of the structures of proteins and PPIs is required to understand the mechanism of action of drugs and how they interfere with protein networks. The Protein Data Bank (PDB) [7] contains almost 30,000 structures that include human molecules. Those structures describe more than 4500 human direct interactions (estimation based on the number of experimental structures available in Interactome3D version 2015_05). Although the coverage of the protein network is far from being complete, high quality models such as those proposed by Interactome3D [8] have extended the human PPI network with models for 4294 additional human PPIs. It is important to note that all interactions in Interactome3D, including those based on a model only, are supported by experimental evidence (e.g., yeast two hybrid or co-immunoprecipitation), and that the modeling framework has been validated to ensure that only high quality models are provided.

Combining data on protein structures and genomic screens for integrated analysis is a non-trivial task. We have recently developed a plugin for IGB [9] that uses advanced visualization techniques to integrate the analysis of genomics data with network and structural biology approaches [10]. The plugin automatically maps genomic regions to protein sequence and interaction structures and identifies residues in contact with proteins, nucleic acids or small molecules. Here we show how such visualization technologies that combine both genomics and network biology can be used to map genomic variations to molecular networks, and to identify hotspots based on protein-protein and protein-drug interactions. This allows the end user to generate hypotheses regarding drug- and ligand-dependent perturbations of PPI networks, and provides predictions as to how specific mutations might have an impact on drug resistance.

Results and discussion

Protein-protein and protein-ligand interactions in structure databases

The PDB contains structures for 5387 human proteins from UniProt (02/2015) [11]. Those structures describe the coordinates of either a single molecule or of a complex that can include one or more binding partners (proteins, nucleic acids, drugs or other compounds). Based on PDB, Interactome3D has built a network of 8880 human interactions with structure (including experimental structures and models). Many of the proteins in this 3D PPI network also interact with drugs and other compounds. We compared the interfaces in PPI structures to the interfaces of protein-ligand interactions (PLIs). Our premise was that compounds that share one or more contact residues with a protein are potentially capable of interfering with the PPI.

We started by retrieving the list of structures from PDB that include a compound and extracted all contacts between proteins and ligands (Fig. 1a). Then, we
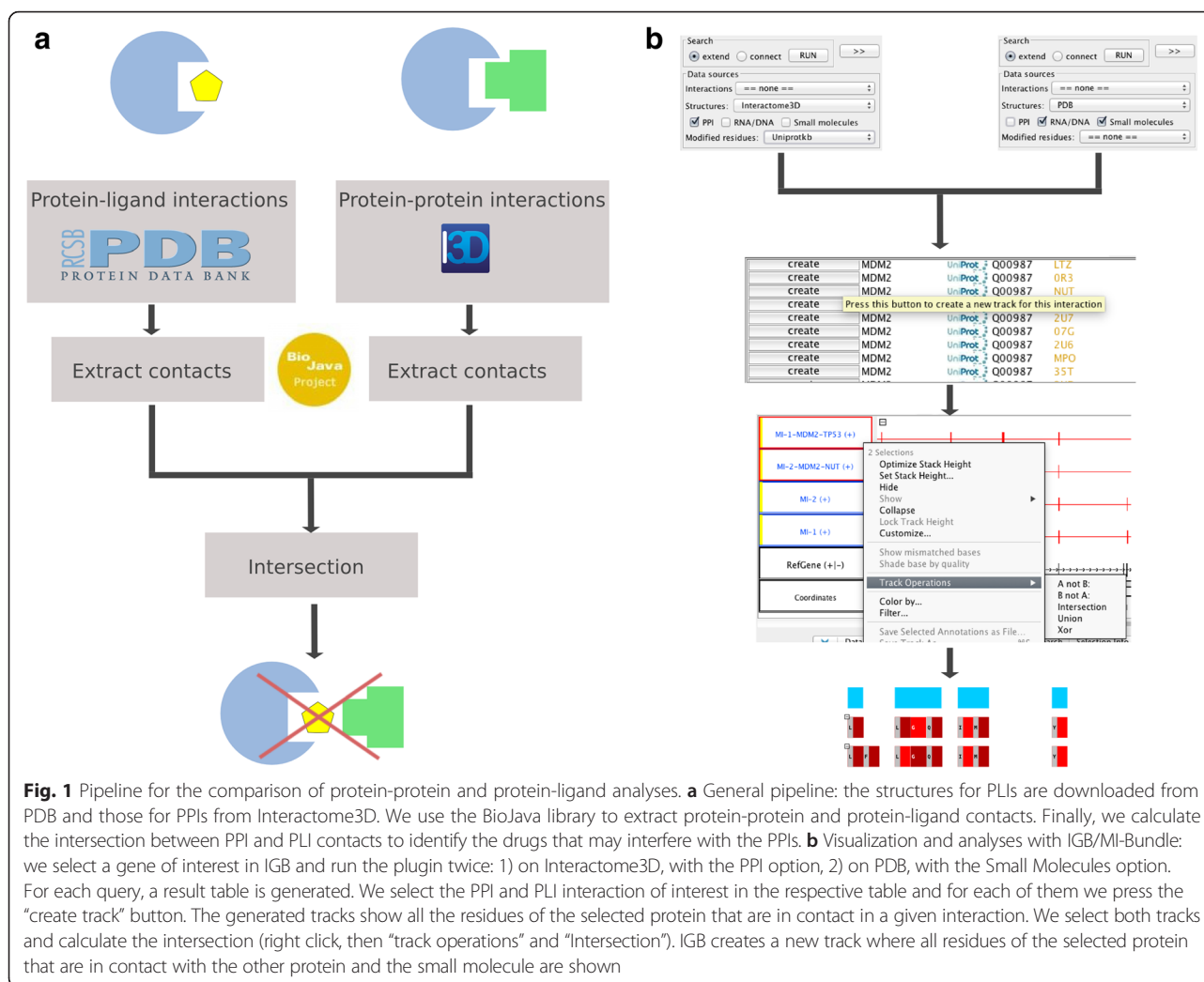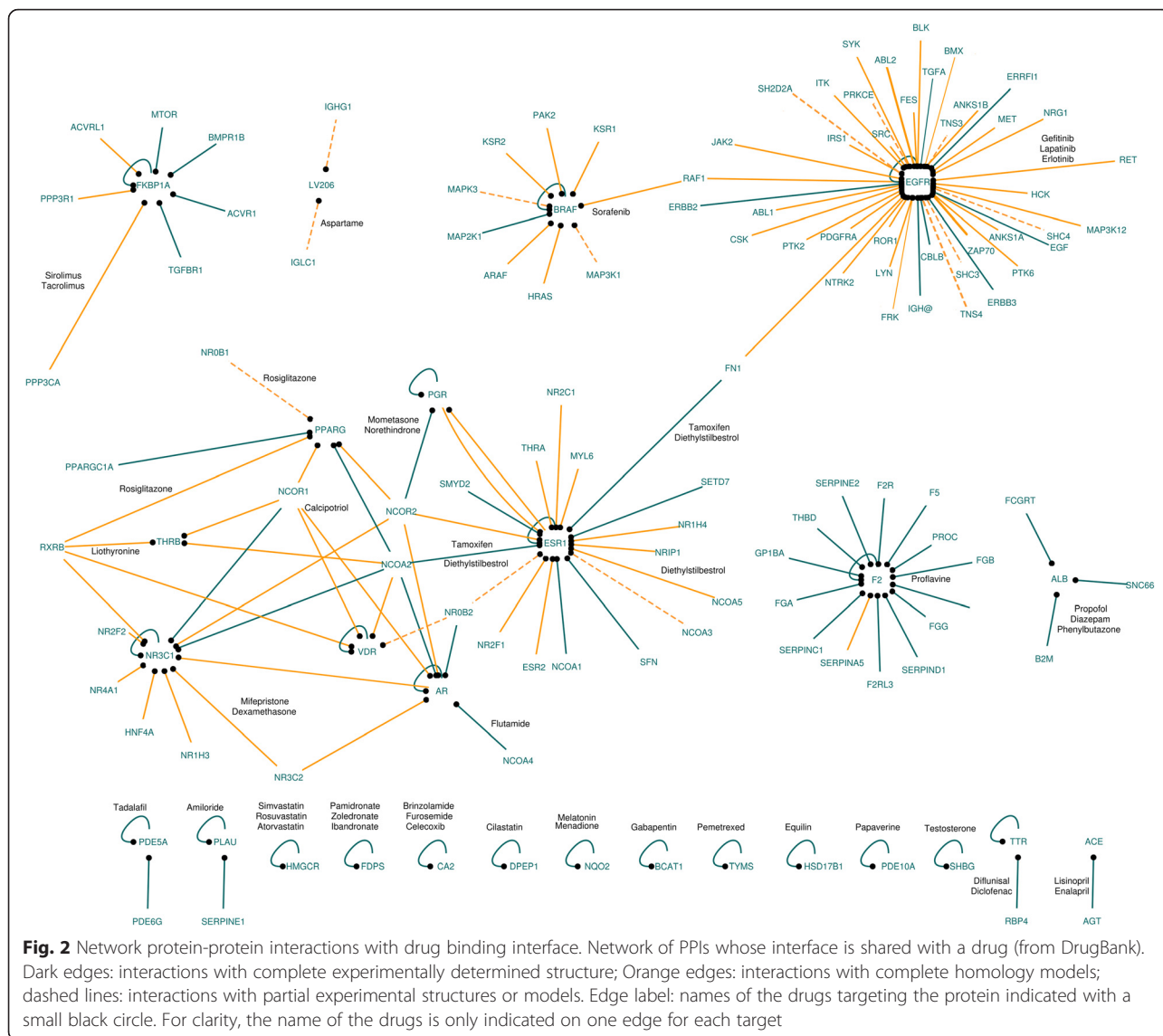
Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 217 of 269

Fig. 1 Pipeline for the comparison of protein-protein and protein-ligand analyses. **a** General pipeline: the structures for PLIs are downloaded from PDB and those for PPIs from Interactome3D. We use the BioJava library to extract protein-protein and protein-ligand contacts. Finally, we calculate the intersection between PPI and PLI contacts to identify the drugs that may interfere with the PPIs. **b** Visualization and analyses with IGB/MI-Bundle: we select a gene of interest in IGB and run the plugin twice: 1) on Interactome3D, with the PPI option, 2) on PDB, with the Small Molecules option. For each query, a result table is generated. We select the PPI and PLI interaction of interest in the respective table and for each of them we press the "create track" button. The generated tracks show all the residues of the selected protein that are in contact in a given interaction. We select both tracks and calculate the intersection (right click, then "track operations" and "Intersection"). IGB creates a new track where all residues of the selected protein that are in contact with the other protein and the small molecule are shown

inspected all human PPIs in the pre-calculated structural human network from Interactome3D, including models, to identify contacts in common with the ligands (residues with Buried Surface Area > 1 Å^2, see methods). We identified 1253 ligands that bind at the interface of 1123 human binary PPIs (meaning that at least one contact is shared with PLI). The interactions are based on 473 complete experimental structures and 650 complete homology models (>80 % coverage) or partial experimental structures or models.

Many of those ligands (for instance solvents, salts, molecules used for experimental purposes, and cofactors) may be irrelevant when looking for molecules that may alter the PPI network. Therefore, we restricted the list of ligands in PDB to include only small molecules approved by the U.S. Food and Drug Administration (FDA) listed in DrugBank [12] (see Methods). Of the 1253 interface compounds we identified above, 44 are approved small molecules that are cross-referenced in DrugBank. Those 44 drugs are at the interface of 151

PPIs (Fig. 2). None of those compounds were originally designed to target a PPI. This network contains several well-known drug targets, including cancer targets such as EGFR and BRAF. The amino acids in contact with the drugs are also in contact with many proteins involved in signal transduction (46 %, Gene Ontology enrichment *p*-value: 7e−19), in response to stimulus (85 %, *p*-value 3e−16), and they are mainly membrane proteins (72 %, *p*-value: 8e-7, see Additional file 1 for the complete list of enriched terms). Interestingly, many of those contacts have been identified in complete homology models (68 interactions) or partial experimental structures or models (14 interactions), rather than in complete experimental structures (68). Some interactions have been modeled from the homologs of the interacting proteins in other closely related species. For example, the NCOA2-VDR interaction (uniprotkb:Q15596 and uniprotkb:P11473) was modeled based on a structure that used the Zebrafish homolog of human VRD (PDB:3o1d). An experimental structure can also be used to model other

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 218 of 269

Fig. 2 Network protein-protein interactions with drug binding interface. Network of PPIs whose interface is shared with a drug (from DrugBank). Dark edges: interactions with complete experimentally determined structure; Orange edges: interactions with complete homology models; dashed lines: interactions with partial experimental structures or models. Edge label: names of the drugs targeting the protein indicated with a small black circle. For clarity, the name of the drugs is only indicated on one edge for each target
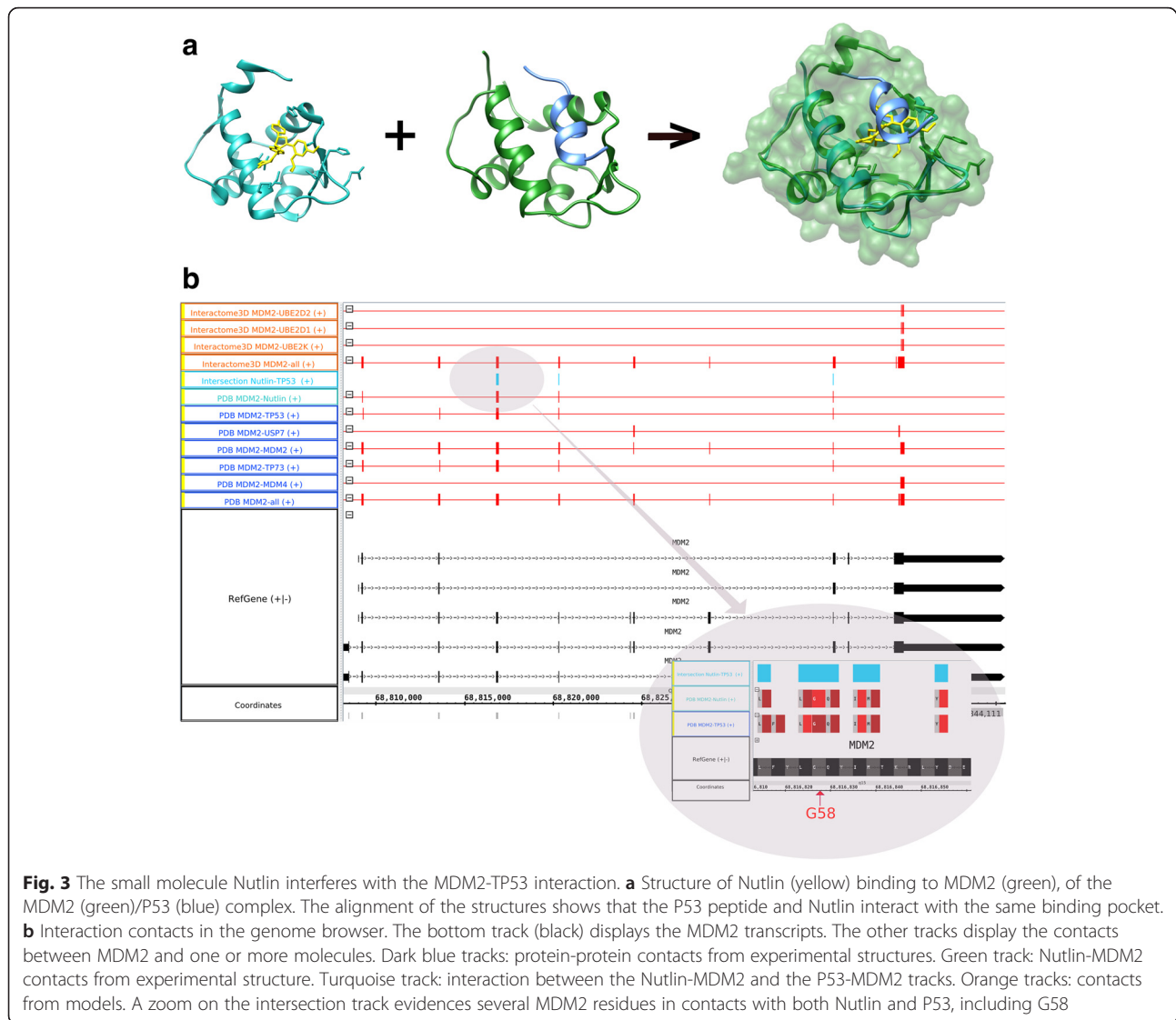
interactions in the network. For instance, the model for the BRAF-RAF1 (uniprotkb:P15056 and uniprotkb:P04049) interaction is based on the experimental structure of the BRAF homodimer (PDB: 4ehe,4mbj). Since these results were able to pinpoint cancer-relevant pathways, and also predict the influence of small molecules on PPIs, we next used genome and network visualization to analyze some examples of drug interactions.

Drugs in protein-protein networks: a visualization perspective

By intersecting the structural protein-protein and protein ligand networks above, we observed that many small molecules, including several approved drugs, could potentially compete with other proteins for binding at interaction sites. However, these analyses required programming skills, the download of datasets with diverse

formats from external repositories, and basic knowledge of network and structural biology. To overcome these issues, and make the tool more accessible to a broader scientific user base, we have integrated the analysis pipeline into a single visualization tool as an IGB plugin called Molecular Interaction (MI) bundle [10]. Several improvements have been implemented and integrated (version 2.9 of the plugin), including the option of querying PDB for ligands only (in previous versions, PPIs were always included), and the integration of information from DrugBank (including drug name and ID) if a cross-reference is available.

Here we used this plugin to visually investigate several examples of drug-PPI interactions, and to integrate mutations and drug-ligand data in order to identify gatekeeper mutations (mutations that abrogate binding of drugs to their targets).

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 219 of 269

Fig. 3 The small molecule Nutlin interferes with the MDM2-TP53 interaction. **a** Structure of Nutlin (yellow) binding to MDM2 (green), of the MDM2 (green)/P53 (blue) complex. The alignment of the structures shows that the P53 peptide and Nutlin interact with the same binding pocket. **b** Interaction contacts in the genome browser. The bottom track (black) displays the MDM2 transcripts. The other tracks display the contacts between MDM2 and one or more molecules. Dark blue tracks: protein-protein contacts from experimental structures. Green track: Nutlin-MDM2 contacts from experimental structure. Turquoise track: interaction between the Nutlin-MDM2 and the P53-MDM2 tracks. Orange tracks: contacts from models. A zoom on the intersection track evidences several MDM2 residues in contacts with both Nutlin and P53, including G58

The analyses can be done in 5 steps (Fig. 1b):

1) Choose drug target and select the corresponding gene in IGB,
2) Select PDB and small molecules in the MI-Bundle to obtain a list of structures with ligands,
3) Select Interactome3D to obtain a list of PPIs with either a structure or a model,
4) Create a new track for each single interaction of interest: the PLI and the PPI by clicking on the "create track" button on the selected row in the result table,
5) Select tracks of interest and calculate the intersection using basic IGB functionalities.

We illustrate this pipeline with the examples of the p53-activating drug Nutlin, the estrogen receptor modulator Tamoxifen, and the development of resistance to Imatinib in chronic myeloid leukemia patients (see also Additional file 2: Additional example of drugs targeting PPIs).

The MDM2-P53 druggable interaction, a case study

P53 is a transcription factor that mediates cell cycle arrest, senescence, or apoptosis in response to DNA damage or oncogenic stress. Consistent with its bona fide tumor suppressor status, the function of p53 or its regulators is altered in most cancers. The MDM2 oncoprotein (uniprotkb: Q00987) is overexpressed in a large fraction of human tumors, and exerts its activity primarily via inhibition of P53 (uniprotkb: P04637). The N terminus of P53 is required for its activity as a transcription factor; it is also the critical domain required for binding to MDM2 [13]. Once bound by MDM2, P53 transactivation function is inhibited. Furthermore, the intrinsic E3 ubiquitin ligase activity of MDM2 leads to

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 220 of 269

proteasome-dependent P53 degradation. Thus, targeting of this interaction has received great attention, as it is expected to reactivate P53 in the ~50 % of tumors that retain a wild type allele [14].
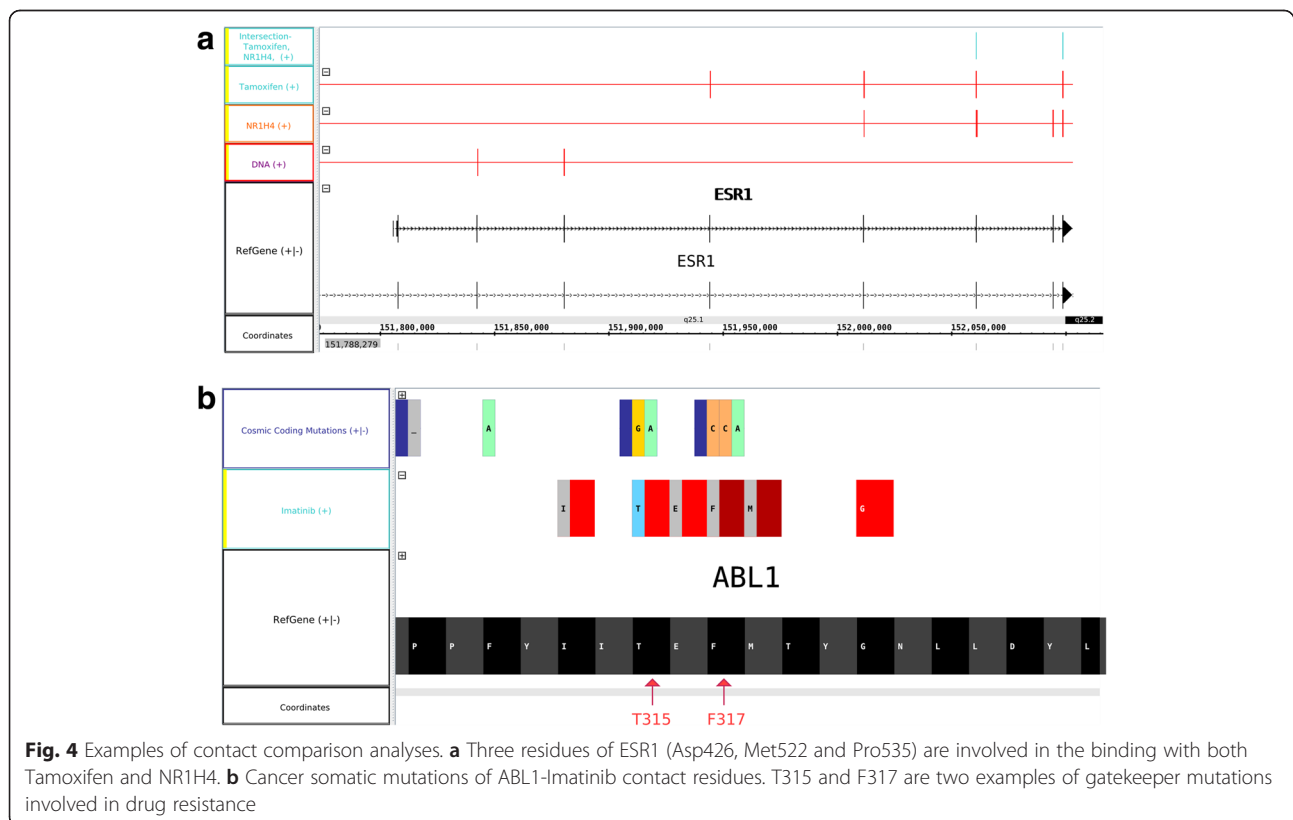
Starting our pipeline from the *MDM2* gene (Fig. 3), we pinpointed in the MDM2/P53 track several residues important for the binding between those two molecules; A cell-based bimolecular luciferase complementation (BiLC) assay indicates an effect of the mutation of the G58 residue on the PPI (see Additional file 2: Selection of the key amino acid contacts from protein-protein and protein-drug interactions for experimental manipulation). Other residues (e.g. V75) had already been reported in literature [15].

In the list of partners of MDM2, we found the cis-imidazoline, Nutlin (NUT, CHEBI:46,742), which was identified during screening of a compound diversity set performed by Roche [16]. The residues on the intersection track (MDM2/P53 and MDM2/Nutlin) are the MDM2 residues that Nutlin shields from p53. These residues are Phe19, Trp23, and Leu26, which are located in an alpha-helical region of the P53 N terminus that binds to the N-terminal MDM2 hydrophobic pocket [17]. The imidazoline scaffold of Nutlin essentially mimics these three critical P53 residues; the compound therefore competes with endogenous P53 for binding to MDM2. In the absence of a structure between MDM2 and P53 and knowing that Nutlin disrupts this interaction, it would have been possible to exploit our strategy to infer some of the contact residues between MDM2 and P53.

MDM2 is involved in three additional interactions for which a structure is available. We created a new track to display the contacts with each of those: MDM2, USP7 (uniprotkb:Q93009) and MDM4 (uniprotkb:O15151). Interestingly, the MDM2 homo-dimerization site intersects with the MDM2-Nutlin interface, suggesting that Nutlin may also interfere with MDM2 homodimerization. Conversely, the contacts that MDM2 makes with USP7 and MDM4 are distinct from the ones with Nutlin: The MDM2/USP7 and MDM2/MDM4 interactions may not be affected by this ligand, suggesting an edgetic effect of this compound. Our prediction that Nutlin does not interfere with the MDM2/MDM4 interaction is supported by data showing that MDM2 and MDM4 co-immunoprecipitate following Nutlin treatment, which is consistent with Nutlin-stimulated, MDM2-dependant degradation of MDM4 [18, 19].

In order to determine whether Nutlin might perturb other interactions in which MDM2 participates, we repeated the previous query with Interactome3D resulting in the extension of the 3D coverage of the MDM2 network with models for interactions with three E2 enzymes (Ubiquitin-conjugating enzyme E2), UBE2D1 (uniprotkb:P51668), UBE2D2 (uniprotkb:P62837), and UBE2K (uniprotkb:P61086). Based on the models, none of

Fig. 4 Examples of contact comparison analyses. **a** Three residues of ESR1 (Asp426, Met522 and Pro535) are involved in the binding with both Tamoxifen and NR1H4. **b** Cancer somatic mutations of ABL1-Imatinib contact residues. T315 and F317 are two examples of gatekeeper mutations involved in drug resistance

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 221 of 269

those interactions share contact regions with Nutlin; those models nevertheless confirm the importance of the C-terminal MDM2 RING domain in mediating its E3 ubiquitin ligase activity [20]. It is worth noting that the contacts are based on structures and models that are not complete, and that additional interfaces may exist.

Another mechanism for tamoxifen action

Intriguingly, our modeling suggests a novel PPI target for Tamoxifen, a small molecule that is used in the clinic to treat estrogen-dependent breast cancer [21]. Tamoxifen's main mechanism-of-action is to compete with estrogen for binding to the estrogen receptor (ESR1), thereby preventing the pro-survival and proliferative effects of this hormone [22]. In addition to steroid hormones, ESR1 has several protein-binding partners, including NR1H4 (Fig. 2), a nuclear receptor that regulates the metabolism of cholesterol, lipid, and glucose. We observed that both Tamoxifen and NR1H4 share a common binding interface with ESR1 (Fig. 4a). Following the steps outlined above, an intersection track confirmed that Tamoxifen and NR1H4 have common contact residues with ESR1. Thus, Tamoxifen might antagonize the interaction between ESR1 and NR1H4. This could have functional consequences, since NR1H4-dependent growth of tumor cell lines in vitro is dependent on the ESR1/NR1H4 interaction [23].

Mutations inducing drug resistance: the case of imatinib

Although many effective anticancer drugs have been developed, advances are often hampered by the development of drug resistance. There are several mechanisms that cause drug resistance, including target cell extrinsic mechanisms, tumor microenvironment, and gene polymorphisms that affect gene expression, for instance [24]. Gatekeeper mutations are defined as those that abrogate binding of drugs to their targets, dramatically reducing the clinical efficacy of treatment. A classical gatekeeper mutation causes resistance to Imatinib, a successful ATP-competitive BCR-ABL kinase inhibitor.

A search for ligands of ABL1 in PDB indicates that the structure of this protein in contact with Imatinib has been resolved (PDB: 2hyy and 3pyy). A new track can be created to display the residues that are actually in contact. We also downloaded from COSMIC [25] a list of missense mutations and loaded them into the browser. The comparison of mutated and contact residues identified several common mutations, which indeed engender resistance to Imatinib (T315I and F317L) [26]. Importing mutation data in IGB is an easy task. In the future, in combination with the MI-Bundle, it will be possible to submit any newly discovered mutation to PLI analyses to identify new gatekeeper mutations.

Conclusions

Genomics and molecular analysis tools are now widely available and sufficiently user friendly to allow cross-disciplinary investigation of the impact of genomic variants on biological systems. In this manuscript, we have used an extension for the IGB to illustrate the identification of key residues for the binding of MDM2 to P53, of Tamoxifen to NR1H4, and of gatekeeper mutations inducing resistance to Imatinib. Our IGB plugin greatly simplifies the integration of genomics data with network and structural analyses of both protein-protein and PLIs.

The particular example of the MDM2-P53 interaction is derived from the intersection of 3D protein-protein and protein-ligand networks. Nevertheless, the structural coverage of the protein network is largely incomplete. The compound-binding proteins we have identified in this work are involved in many interactions for which no structure is available. Thus, we are likely missing additional interactions that may be targeted by small molecules. However, it is possible to increase the structural coverage of the protein network with Interactome3D, which includes PPI models for interaction supported by experimental evidence. Although indirect interactions (e.g., in large complexes detected by co-immunoprecipitation) are excluded from the underlying experimental PPI network in Interactome3D, false positives may still limit the accuracy of the pipeline. The score implemented by the MI bundle may help to identify the most reliable interactions.

By looking at the intersection between protein-ligand contacts and protein-protein contacts, we were able to identify small molecules that may interfere with PPIs. To be clear: our aim with this modeling approach is not to diminish the importance of the conventional approaches to discover the mode of action of those drugs. However, we suggest that our integrated tool can be exploited to generate new hypotheses. It may also be useful for the identification of drugs or molecular scaffolds that can be repurposed in order to block PPIs.

Here, we focused on approved small molecules. Nevertheless, many other ligands have been crystalized together with target proteins and deposited in PDB, including experimental molecules and biotech drugs (peptide, protein or nucleic acid drugs). By extending the analyses to these additional compounds, we believe that it will be possible to generate predictions regarding their effects on disease-relevant PPI networks.

Our analyses are restricted to human proteins. It will be interesting to extend it to other species, both model organisms (many structures are generated based on the homologs of human proteins), and in host-pathogen interactions. To this end it may be useful to model not only PPIs, but also PLIs.

For our initial validation and testing, we have focused our analyses solely on contact residues, considering that one

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 222 of 269

contact may be enough to imply a steric clash when the two partners are superimposed. We didn't assess the extent of cases in which the protein structure of a PLI includes a protein interface that is significantly different from the interface between the same protein and another protein (PPI). We would increase the accuracy of the method by checking the number of clashing atoms. Indeed, in the examples based on the MI-bundle, we were able to visually verify that the respective proteins and drugs share several residues. This current limitation will also be restrained by extending the intersection analysis to the entire interaction interface. Employment of other tools and databases may be useful in this regard [27, 28]. In addition, it will be interesting to discriminate biological interfaces from crystal artifacts. This can be done for instance with the EPPIC server [28] (see Additional file 2: Selection of the key amino acid contacts from protein-protein and protein-drug interactions for experimental manipulation).

We have shown that all our analyses can be done without prior bioinformatics skills by using IGB together with the MI bundle. The software allows loading, visualization, comparison, and analysis of protein-protein and PLIs. Importantly, such analyses are not limited to the current knowledge deposited in public databases. The MI-bundle allows a local copy of an Interactome3D directory to be used. Any user that generates new protein networks (e.g. by yeast two hybrid) can submit them to the Interactome3D server and download the resulting structural network. Consequently, the user will be able to apply the plugin to their own data to discover new putative drug-PPI interactions.

As a genome browser, IGB accepts a large number of genomic formats, such as BED, VCF or GTF; it can also communicate directly with DAS servers [29], such as those implemented by the UCSC Genome Browser [30, 31] or Ensembl [32]. Therefore, it is easy to interrogate newly discovered mutations for their potential to induce drug resistance (gatekeeper mutations).

Finally, we did not consider the features of PPIs that may influence the success of an inhibitor in this manuscript. In a recent review, Smith and Gestwicki [6] have analyzed how amenable each of the four PPI classes are to inhibition, according to whether the interface is classified as "Loose" or "Tight", and "Wide" or "Narrow". Integrating this information in our frameworks will facilitate the identification of the most promising PPI targets.

Methods
Drugs and compounds
We used the mapping file available at [33] to associate the compounds in the structures to DrugBank IDs (DrugBank version 4.2) [12]. This file only contains approved drugs. We downloaded the file [34] to obtain the list of known targets from DrugBank.

We used the mapping between Uniprot and PDB provided by the SIFTS initiative [35].

Structures
We worked on a local mirror of the PDB database updated on May 2015. For PPIs, we downloaded the representative human dataset from Interactome3D, version 2015_05. The structures in Fig. 3 have been aligned and displayed with the UCSF Chimera package [36].

Contacts
In order to identify the contacts between the pairs of proteins or the pairs of protein-compound identified in the previous step, we measure the available buried surface area as described in [10]. We consider as contact all residues with Buried Surface Area > 1 Å^2. Parsing of PDB files and calculation of the available surface area was done with the BioJava libraries version 4.1 [37].

Mutation data
We have downloaded the VCF file of all coding mutations in Cosmic [25] version 4.1 for the human reference genome version GRCh38 from [38].

Network visualization
The network figures have been generated with Cytoscape version 3.2.0 [39].

Gene ontology enrichment
We used the BINGO plugin [40] for Cytoscape to obtain GO enrichments on the list of partners of the drug targets. To avoid a bias toward proteins with structures, we used as reference set all the human proteins included in the human 3D network of Interactome3D. We downloaded from GOA [41] the GO annotations for human (generated: 2015-07-20). We used the GOA slim ontology in order to obtain a smaller list of GO terms that covers all GO categories (Biological Processes, Molecular Function, Cellular Compartment).

Analyses with IGB and the MI-bundle
We downloaded and installed IGB and the MI-Bundle as described at [42]. All analyses in IGB are done on the human reference genome version GRCh38.

Additional files

Additional file 1: Additional table: Gene Ontology terms enriched in the partners of drug targets. (PDF 58 kb)

Additional file 2: Additional results: Selection of the key amino acid contacts from protein-protein and protein-drug interactions for experimental manipulation; Additional example of drugs targeting PPIs. (PDF 499 kb)

Céol *et al. BMC Bioinformatics* 2016, **17**(Suppl 4):54

Page 223 of 269

Abbreviations
FDA: the Food and Drug Administration; IGB: integrated genome browser; MI: molecular interaction; PDB: the Protein Data Bank; PLI: protein-ligand interaction; PPI: protein-protein interaction.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
AC carried out the analyses, developed the software and wrote the manuscript. LGGCV performed the G58A mutagenesis and produced BiLC data (see Additional file 2) and wrote the manuscript. MW conceived/analyzed experiments and wrote the manuscript. HM proposed and coordinated the project and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements
We thank Geoff Wahl and Yao-Cheng Li (Salk Institute, La Jolla) for reagents related to the split luciferase experiments. We would also like to thank and Jose Manuel Duarte and Guido Capitani for their help with the EPPIC data.

References
1. Goh K, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L. The human disease network. Proc Natl Acad Sci U S A. 2007;104:8685–90.
2. Ideker T, Sharan R. Protein networks in disease. Genome Res. 2008;18:644–52.
3. Sahni N, Yi S, Zhong Q, Jailkhani N, Charloteaux B, Cusick ME, et al. Edgotype: a fundamental link between genotype and phenotype. Curr Opin Genet Dev. 2013;23(6):649–57.
4. Dutia BM, Frame MC, Subak-Sharpe JH, Clark WN, Marsden HS. Specific inhibition of herpesvirus ribonucleotide reductase by synthetic peptides. Nature. 1986;321:439–41.
5. Cohen EA, Gaudreau P, Brazeau P, Langelier Y. Neutralization of herpes simplex virus ribonucleotide reductase activity by an oligopeptide-induced antiserum directed against subunit H2. J Virol. 1986;60:1130–3.
6. Smith MC, Gestwicki JE. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. Expert Rev Mol Med. 2012;14:997–1003.
7. Velankar S, Alhroub Y, Alili A, Best C, Boutselakis HC, Caboche S, et al. PDBe: protein data bank in Europe. Nucleic Acids Res. 2011;39:D402–10.
8. Mosca R, Céol A, Aloy P. Interactome3D: adding structural details to protein networks. Nat Methods. 2013;10:47–53.
9. Nicol JW, Helt GA, Blanchard SG, Raja A, Loraine AE. The integrated genome browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics. 2009;25:2730–1.
10. Céol A, Müller H. The MI bundle: enabling network and structural biology in genome visualization tools. Bioinformatics. 2015;31(22):3679–81. doi:10.1093/bioinformatics/btv431.
11. Magrane M, Consortium U. UniProt Knowledgebase: a hub of integrated protein data. Database (Oxford). 2011;2011:bar009. doi:10.1093/database/bar009.
12. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, et al. DrugBank 4.0: Shedding new light on drug metabolism. Nucleic Acids Res. 2014;42.
13. Momand J, Zambetti GP, Olson DC, George D, Levine AJ. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. Cell. 1992;69:1237–45.
14. Wade M, Li Y-C, Wahl GM. MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nat Rev Cancer. 2013;13:83–96.
15. Freedman D a, Epstein CB, Roth JC, Levine a J. A genetic approach to mapping the p53 binding site in the MDM2 protein. Mol Med. 1997;3:248–59.
16. Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science. 2004;303:844–8.
17. Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ, et al. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. Science. 1996;274:948–53.
18. Wade M, Wong ET, Tang M, Stommel JM, Wahl GM. Hdmx modulates the outcome of p53 activation in human tumor cells. J Biol Chem. 2006;281:33036–44.
19. Xia M, Knezevic D, Tovar C, Huang B, Heimbrook DC, Vassilev LT. Elevated MDM2 boosts the apoptotic activity of p53-MDM2 binding inhibitors by facilitating MDMX degradation. Cell Cycle. 2008;7:1604–12.
20. Fang S, Jensen JP, Ludwig RL, Vousden KH, Weissman AM. Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53. J Biol Chem. 2000;275:8945–51.
21. Huang B, Warner M, Gustafsson J-Å. Estrogen receptors in breast carcinogenesis and endocrine therapy. Mol Cell Endocrinol. 2014;147:381–7.
22. Jordan VC. Tamoxifen (ICI46,474) as a targeted therapy to treat and prevent breast cancer. Br J Pharmacol. 2006;147:S269–76.
23. Journe F, Laurent G, Chaboteaux C, Nonclercq D, Durbecq V, Larsimont D, et al. Farnesol, a mevalonate pathway intermediate, stimulates MCF-7 breast cancer cell growth through farnesoid-X-receptor-mediated estrogen receptor activation. Breast Cancer Res Treat. 2008;107:49–61.
24. Barouch-Bentov R, Sauer K. Mechanisms of drug resistance in kinases. Expert Opin Investig Drugs. 2011;20:153–208.
25. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer. Nucleic Acids Res. 2011;39(Database issue):D945–50.
26. Reddy EP, Aggarwal a K. The Ins and outs of Bcr-Abl inhibition. Genes. Cancer. 2012;3:447–54.
27. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774–97.
28. Duarte JM, Srebniak A, Schärer M a, Capitani G. Protein interface classification by evolutionary analysis. BMC Bioinformatics. 2012;13:334.
29. Jenkinson AM, Albrecht M, Birney E, Blankenburg H, Down T, Finn RD, et al. Integrating biological data—the distributed annotation system. BMC Bioinformatics. 2008;9 Suppl 8:S3.
30. Raney BJ, Cline MS, Rosenbloom KR, Dreszer TR, Learned K, Barber GP, et al. ENCODE whole-genome data in the UCSC genome browser (2011 update). Nucleic Acids Res. 2011;2011(39):D871–5.
31. Hinrichs AS. The UCSC genome browser database: update 2006. Nucleic Acids Res. 2006;34:D590–8.
32. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S. Ensembl. Nucleic Acids Res. 2015;2015:43.
33. DrugBank mapping file. http://www.drugbank.ca/system/downloads/current/drug_links.csv.zip. Accessed July 2015.
34. List of targets in DrugBank. http://www.drugbank.ca/system/downloads/current/drug_target_uniprot_links.csv.zip. Accessed July 2015.
35. Velankar S, Dana JM, Jacobsen J, van Ginkel G, Gane PJ, Luo J, et al. SIFTS: structure integration with function, taxonomy and sequences resource. Nucleic Acids Res. 2013;41(Database issue):D483–9.
36. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–12.
37. Prlic A, Yates A, Bliven SE, Rose PW, Jacobsen J, Troshin PV, et al. BioJava: an open-source framework for bioinformatics in 2012. Bioinformatics. 2012;28:2693–5.
38. VCF file for all coding mutations in Cosmic. https://cancer.sanger.ac.uk/cosmic/files?data=/files/grch38/cosmic/v73/VCF/CosmicCodingMuts.vcf.gz. Accessed July 2015.
39. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, et al. A travel guide to cytoscape plugins. Nat Methods. 2012;9:1069–76.
40. Maere S, Heymans K, Kuiper M. BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. 2005;21:3448–9.
41. Dimmer EC, Huntley RP, Alam-Faruque Y, Sawford T, O'Donovan C, Martin MJ, et al. The UniProt-GO annotation database in 2011. Nucleic Acids Res. 2012;40(Database issue):D565–70.
42. MI Bundle home page: http://cru.genomics.iit.it/igbmibundle. Accessed July 2015.