BMC Bioinformatics

CrossMark

DrugGenEx-Net: a novel computational platform for systems pharmacology and gene expression-based drug repurposing

Naiem T. Issa[1], Jordan Kruger[2†], Henri Wathieu[3†], Rajarajan Raja[4], Stephen W. Byers[1,2] and Sivanesan Dakshanamurthy[1,2*]

Abstract

Background: The targeting of disease-related proteins is important for drug discovery, and yet target-based discovery has not been fruitful. Contextualizing overall biological processes is critical to formulating successful drug-disease hypotheses. Network pharmacology helps to overcome target-based bottlenecks through systems biology analytics, such as protein-protein interaction (PPI) networks and pathway regulation.

Results: We present a systems polypharmacology platform entitled DrugGenEx-Net (DGE-NET). DGE-NET predicts empirical drug-target (DT) interactions, integrates interaction pairs into a multi-tiered network analysis, and ultimately predicts disease-specific drug polypharmacology through systems-based gene expression analysis. Incorporation of established biological network annotations for protein target-disease, −signaling pathway, −molecular function, and protein-protein interactions enhances predicted DT effects on disease pathophysiology. Over 50 drug-disease and 100 drug-pathway predictions are validated. For example, the predicted systems pharmacology of the cholesterol-lowering agent ezetimibe corroborates its potential carcinogenicity. When disease-specific gene expression analysis is integrated, DGE-NET prioritizes known therapeutics/experimental drugs as well as their contra-indications. Proof-of-concept is established for immune-related rheumatoid arthritis and inflammatory bowel disease, as well as neuro-degenerative Alzheimer's and Parkinson's diseases.

Conclusions: DGE-NET is a novel computational method that predicting drug therapeutic and counter-therapeutic indications by uniquely integrating systems pharmacology with gene expression analysis. DGE-NET correctly predicts various drug-disease indications by linking the biological activity of drugs and diseases at multiple tiers of biological action, and is therefore a useful approach to identifying drug candidates for re-purposing.

Keywords: DrugGenEx-NET, TMFS, Polypharmacology, Gene expression analysis, Rheumatoid arthritis, Inflammatory bowel disease, Parkinson's disease, Alzheimer's disease

Background

Modern drug discovery endeavors are only rarely translated into acceptable clinical success rates [1]. Preclinical drug discovery initiatives have been gene-centric with a focus on finding drugs for targets of interest with high binding affinity and selectivity [2]. It is increasingly accepted, however, that disease states exhibit biological complexity, and that the gene-centric view neglects physiologic context by isolating the target in an artificial environment [3]. Furthermore, drugs arising from de novo design are likely to have many unknown targets given the limited scope of biochemical assays, thus leading to both clinical toxicity and unanticipated novel disease indications [4]. Systems pharmacology, the integration of systems biology with network pharmacology, is a mechanism-centric solution that considers the global physiological environment of disease states and allows for the discovery of drugs or combinations of

* Correspondence: sd233@georgetown.edu
†Equal contributors
[1]Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington DC 20057, USA
[2]Department of Biochemistry & Molecular Biology, Georgetown University, Washington DC 20057, USA
Full list of author information is available at the end of the article

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 2 of 18

drugs that may simultaneously target multiple nodes of the disease-associated network [5]. Initiatives utilizing network analysis have led to successful drug discovery efforts [6–11].

As most FDA-approved drugs are considered safe and simultaneously exhibit multi-target effects, drug repurposing is an optimal strategy for harnessing the strength of polypharmacology [12]. Current methods do not utilize high-throughput approaches to empirically determine drug-target associations and subsequently contextualize them using systems biology. Here, we have created a novel computational systems pharmacology platform, entitled DGE-NET, that: (1) accurately predicts drug-protein target interactions, (2) assesses drug effects through systems analysis of cumulative predicted targets for each drug, and (3) formulates drug-disease associations through gene expression analysis and polypharmacology.

DGE-NET was first applied to a set of 3,671 FDA approved and experimental drugs across 2,335 human protein target crystal structures for potential drug repurposing. Drugs were then associated with biological effects, which include molecular functions, signaling pathways, protein-protein interactions (PPIs) and diseases, through association with their predicted targets. Drug-biological effect predictions were validated at multiple tiers using findings in the literature and experimentally determined associations from annotated databases. Over 50 drug-disease and 100 drug-pathway associations were validated. DGE-NET also provided further evidence for unexpected toxicities, such as the potential carcinogenic properties of the cholesterol absorption blocker ezetimibe. Drug-target and drug-biological effect signatures were also statistically associated with clinical disease-relevant protein targets, PPIs, pathways, and functions obtained from differential gene expression analysis. DGE-NET incorporated a novel drug prioritization scheme that ranks drugs matched to a disease based on its polypharmacology at each tier of biological action.

For proof-of-concept, DGE-NET was applied to human-derived gene expression datasets obtained for rheumatoid arthritis (RA), inflammatory bowel disease (IBD), Alzheimer's disease (AD), and Parkinson's disease (PD). DGE-NET was validated by prioritizing approved drugs and biologics as well as those currently being examined repurposing, and also revealed drugs contraindicated in those conditions, such as tetracyclines in IBD. DGE-NET is first computational platform we know of that predicts novel protein binding signatures of FDA-approved drugs and subsequently matches drug action at multiple levels of biological activity to gene expression-based characterization of disease perturbation. It stands as an effort to address the pressing need for models that account for the complexity of multi-tiered interactions for better simulations of disease states and predictive therapeutics. In summary, DGE-NET is a novel computational method for gene expression- and systems polypharmacology-driven drug repurposing.

Methods

Collection of FDA-approved drugs, experimental molecules, and protein target curation

Spatial Data Files (SDF) of drugs and experimental molecules containing spatial atom connectivity information were obtained from DrugBank [13], the NCGC Pharmaceutical Collection [14], FDA (www.FDA.gov), and BindingDB [15]. Energy-minimized 3D structures were prepared using Schrodinger's LigPrep [16] algorithm at pH 7.0. Human protein crystal structures were obtained from RCSB (www.rcsb.org). Only X-ray structures with <2.5 angstrom resolution and a reference co-crystallized ligand were chosen. Protein structures were further processed to remove non-biologically relevant chains (i.e. those that do interact with the ligand), metal ions, and all heteroatoms (i.e. non-cofactors, solvent molecules). Structures were then prepared using ProteinPrep in Schrodinger to relax the structures and optimize hydrogen bonds at pH 7.0. After processing, the dataset included 3,671 drugs and 2,335 protein target crystal structures.

Predicting Drug-Target (DT) signatures

DGE-NET utilizes a modified version of our "Train, Match, Fit and Streamline" (TMFS) method [17] for generating reliable binding signature predictions. Briefly, TMFS is a proteochemometric method that predicts the binding potential of a protein-ligand complex by integrating docking, three-dimensional shape, and ligand physicochemical descriptors (Fig. 1). GLIDE [18] was used to dock molecules into protein pockets identified by the reference ligand, and QikProp [19] was used to generate the following ligand-specific physicochemical descriptors: (1) solvent-accessible surface area, (2) volume, (3) dipole, (4) # H-bond acceptors, (5) # H-bond donors, (6) globularity, (7) ionization potential, and (8) electron affinity. Strike [20] was used to generate Tanimoto similarity coefficients to quantify the similarity of ligand physicochemical descriptors to that of the bioactive reference molecules found in the protein complex crystal structures. Ligand and pocket 3D shapes were quantified using a spherical harmonics expansion approach [21] and ligand-reference molecule/ligand-protein pocket shape similarities were quantified using a Euclidean distance metric. After docking scores, shape similarity, Euclidian distance scores, and ligand-based descriptor similarity scores were derived by the tools described above, a common scheme was used to normalize these scores, wherein each is transformed into a 0–1 range, 1 being the most favorable score present. These metrics were combined into a comprehensive Z-score
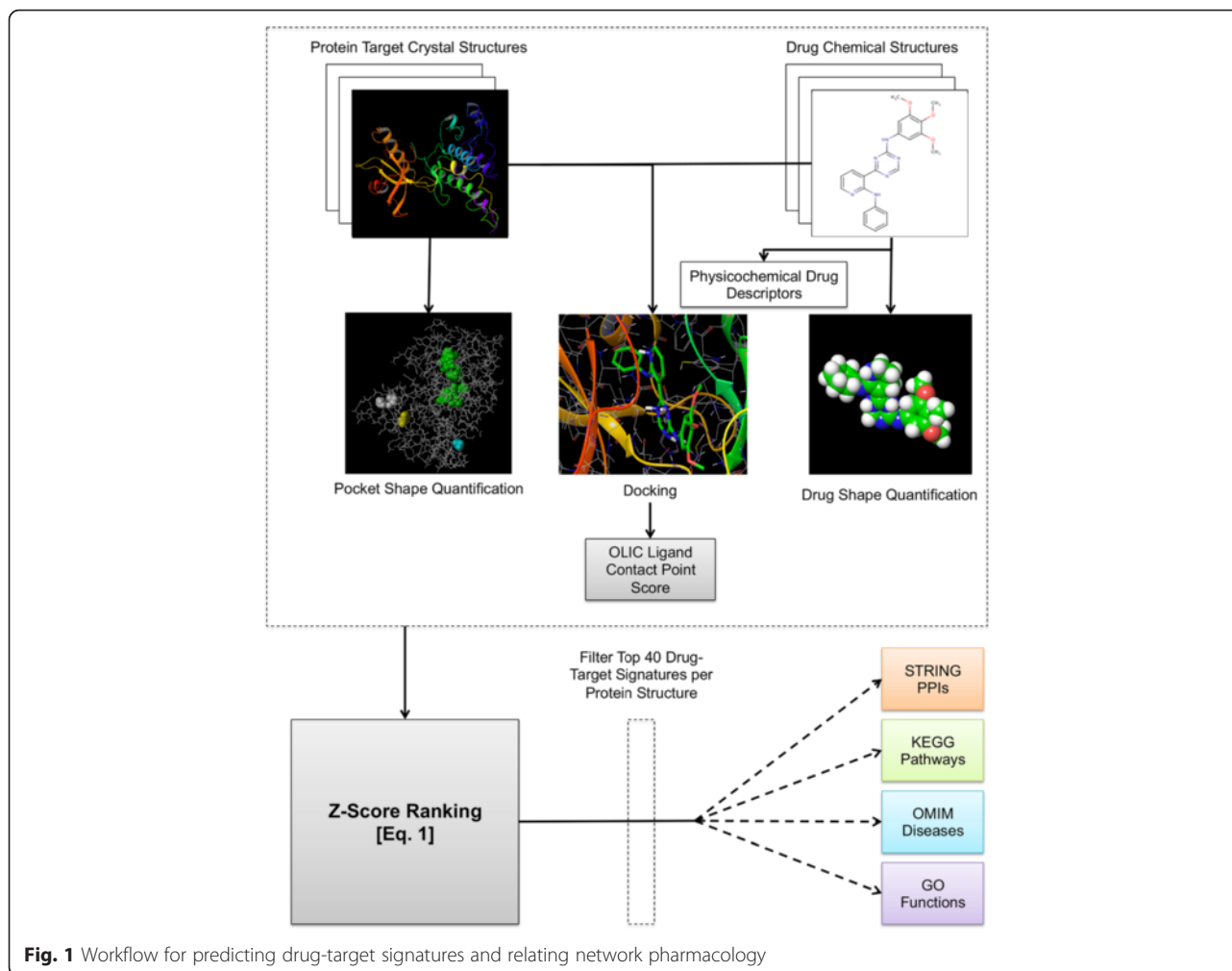
Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 3 of 18

Fig. 1 Workflow for predicting drug-target signatures and relating network pharmacology

that was used to rank ligands such that the top-ranking molecules are considered most likely to bind. The Z-score for a unique ligand (l) –protein (p) co-crystallized with reference ligand r is as follows:

$$Z(l,r,p) = w_k Y(l,p)$$
$$+ \sum_{m=1}^{M} \left[w_m f_m(l,p) + w'_m f'_m(l,r) \right]$$
$$+ \sum_{n=1}^{N} [X_n(l,r) + CS(OLIC)] \quad (1)$$

Y is the normalized docking score with weight $w_{k=4}$. The first summation term is the normalized shape similarity score for ligand-to-protein pocket $f_m(l,p)$ and ligand-to-reference $f'_m(l,r)$ with weights $w_{m=1}$ and $w'_{m=2}$, and the second summation term corresponds to the sum of the Tanimoto similarity coefficients between the ligand and reference for physicochemical descriptors. Aforementioned weights for docking, protein shape

similarity, and ligand shape similarity, respectively, were found to maximize the accuracy of TMFS in predicting top protein targets from publically available experimental data. Lastly, $CS(OLIC)$ is a correction term based on the similarity of contact points created between the ligand and reference to the protein target. It was assumed that drugs have similar experimental activity if their interaction involves similar binding site residues and interaction patterns to that of the reference. The top 40-scoring drugs were considered as "hits" for a given protein target for subsequent network analysis. The top 40 drugs were chosen as they represent the top 1 % of all the drugs in our dataset, a fraction that is typically employed in virtual screening protocols [17].

Relating drug-target predictions to diseases, pathways, functions, and protein-protein interactions

Predicted drug-target associations were associated with diseases, signaling pathways and molecular functions for network analysis (Fig. 1). Protein targets were cross-referenced using the unique PDB entry with UniProt

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 4 of 18

[22]. Because many crystal structures may correspond to the same protein, collapsing them using UniProt reduces the total number of protein target nodes. A list of genes associated with the protein were obtained from each UniProt entry and mapped to Online Mendelian Inheritance in Man (OMIM) Morbidity Map [23] gene-disease associations, a procedure modeled after Yildirim et al. [24]. Drugs are connected to a disease via mapping of their target genes to their associated disease. Thus, a drug is connected to a disease if its predicted targets have disease genes associated with the disorder. In the DT-disease network, all disorders associated with a predicted protein target will be associated with the drug.

Disease-associated targets were also annotated with KEGG pathway [25, 26] and Gene Ontology (GO) molecular function [27, 28] information using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) Functional Annotation Tool (FAT) [29, 30]. FAT was also used to annotate pathways and functions for a given drug via its predicted direct and indirect targets through protein-protein interactions using FDR <0.25. Protein-protein interactions (PPIs) were extracted from the ExPASy STRING database [31] using a confidence score cutoff of 0.95. Any PPI pairs where one of the partners did not exist in our protein target dataset were excluded. A gene list comprised of a drug's predict direct targets as well as those targets' interacting partners was subjected to DAVID annotation. For example, if Drug A was predicted to interact with Target A and Target B, and Target A also interacted with Protein C while Target B interacted with Protein D and Protein E, then the gene list for Drug A would consist of the following: Target A, Target B, Protein C, Protein D, and Protein E.

Annotating disease and pathway categories

The disease categories from Medical Subject Headings (MeSH) were used for annotation of disease names corresponding to OMIM disorder entries. Approximately 93 % of the diseases were mapped to a disease category. The Comparative Toxicogenomics Database (CTD) [32] was used to map 75 % of the diseases; the remaining diseases were manually curated, with 71 % of these providing a partial or close match. Diseases that mapped to multiple disease categories were manually evaluated to determine a primary disease category. This was done by determining what the primary clinically treated category is for a disease. For example, the disease systemic lupus erythematosus is primarily an autoimmune disorder but can be considered as "skin and connective tissue" if the disease process involves the facial malar rash. Diseases in which a primary category could not be determined were categorized as multiple. Pathways were manually organized into categories based on metabolic/cellular processes and diseases as annotated by KEGG.

Incorporation of disease gene expression data with systems pharmacology

A schematic of DGE-NET is illustrated in Fig. 2. Differential gene expression analysis on Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) microarray data was performed for RA (GSE55235 and GSE55457), IBD (GSE52746 and GSE11223), AD (GSE29378), and PD (GSE7621). Differentially expressed genes between normal and diseased patient biopsies with adjusted P values < 0.05 (using GEO2R [33]) were obtained. GEO2R is a R-based publicly accessible web tool for analyzing GEO-deposited gene expression data (http://www.ncbi.nlm.nih.gov/geo/
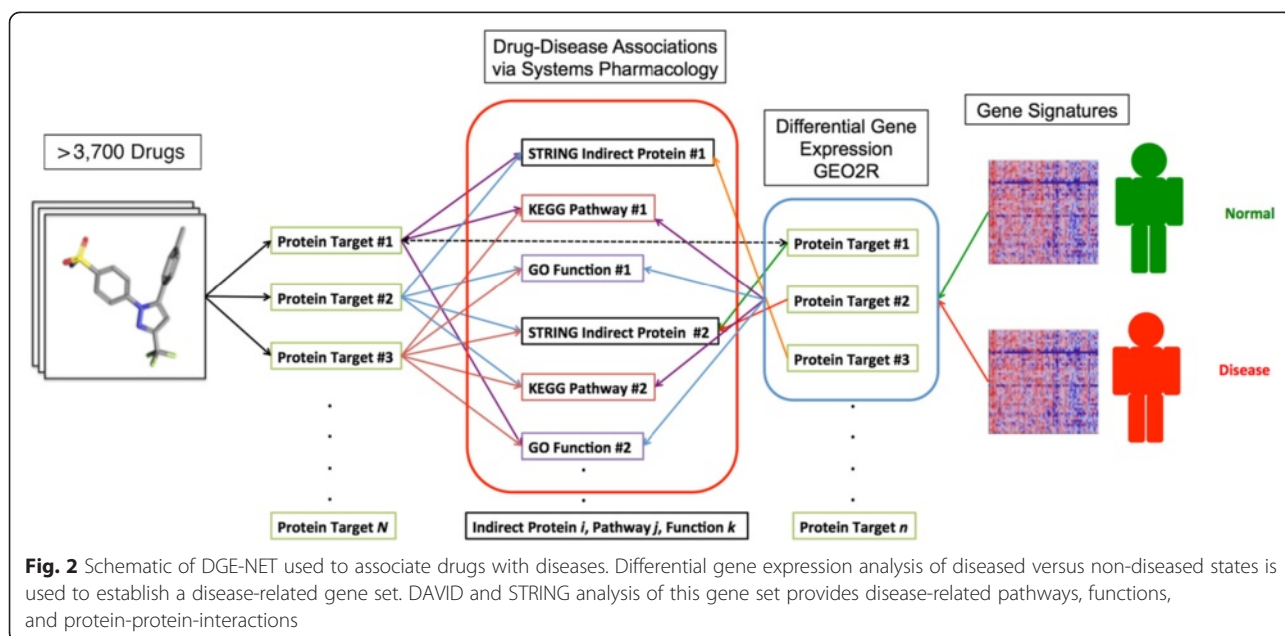

Fig. 2 Schematic of DGE-NET used to associate drugs with diseases. Differential gene expression analysis of diseased versus non-diseased states is used to establish a disease-related gene set. DAVID and STRING analysis of this gene set provides disease-related pathways, functions, and protein-protein-interactions

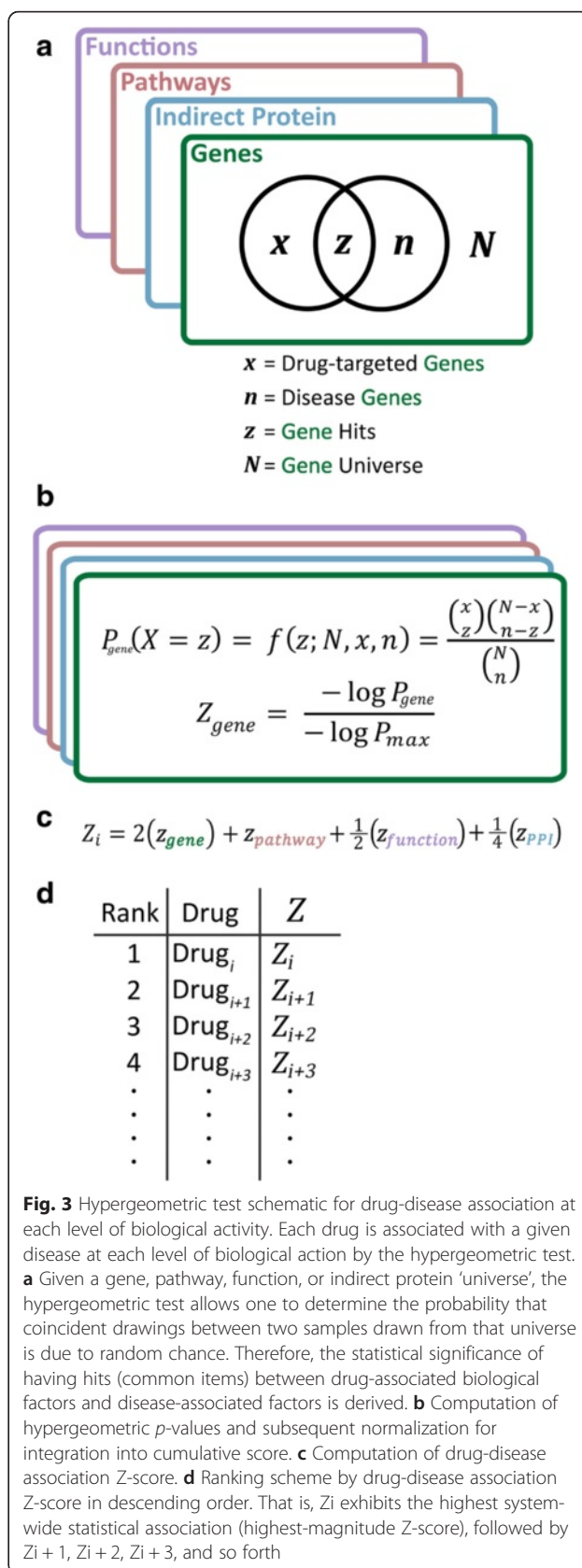Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 5 of 18

geo2r/). The differential gene list was subjected to functional systems biology annotation as noted above. For disease sets, multiple testing correction yielded few genes having significantly differential expression. Nominal *P*-values < 0.05 were therefore used to allow for robust overrepresentation analyses. For the IBD set (normal colonic tissue control versus active IBD without anti-TNF therapy), the top 1,500 up-regulated and top 1,500 down-regulated genes were taken to create a list of 3,000 genes – the maximum number that DAVID accepts. All other datasets resulted in differential gene lists of fewer than 3,000 genes.
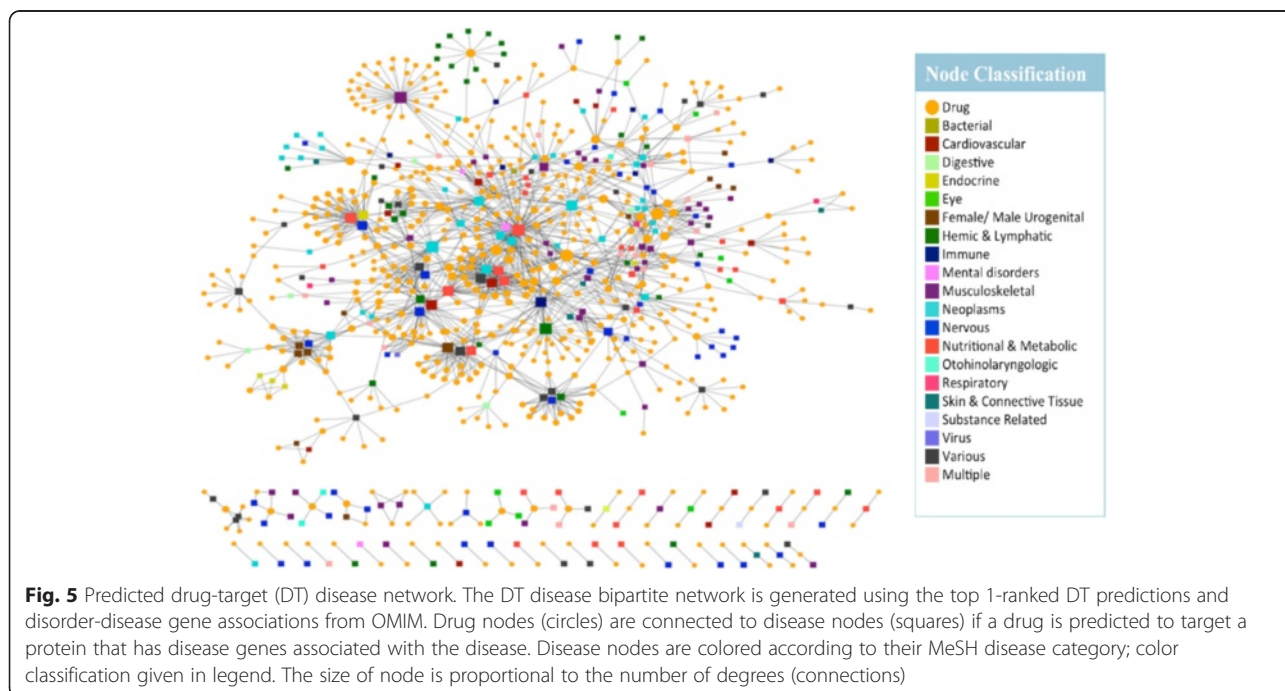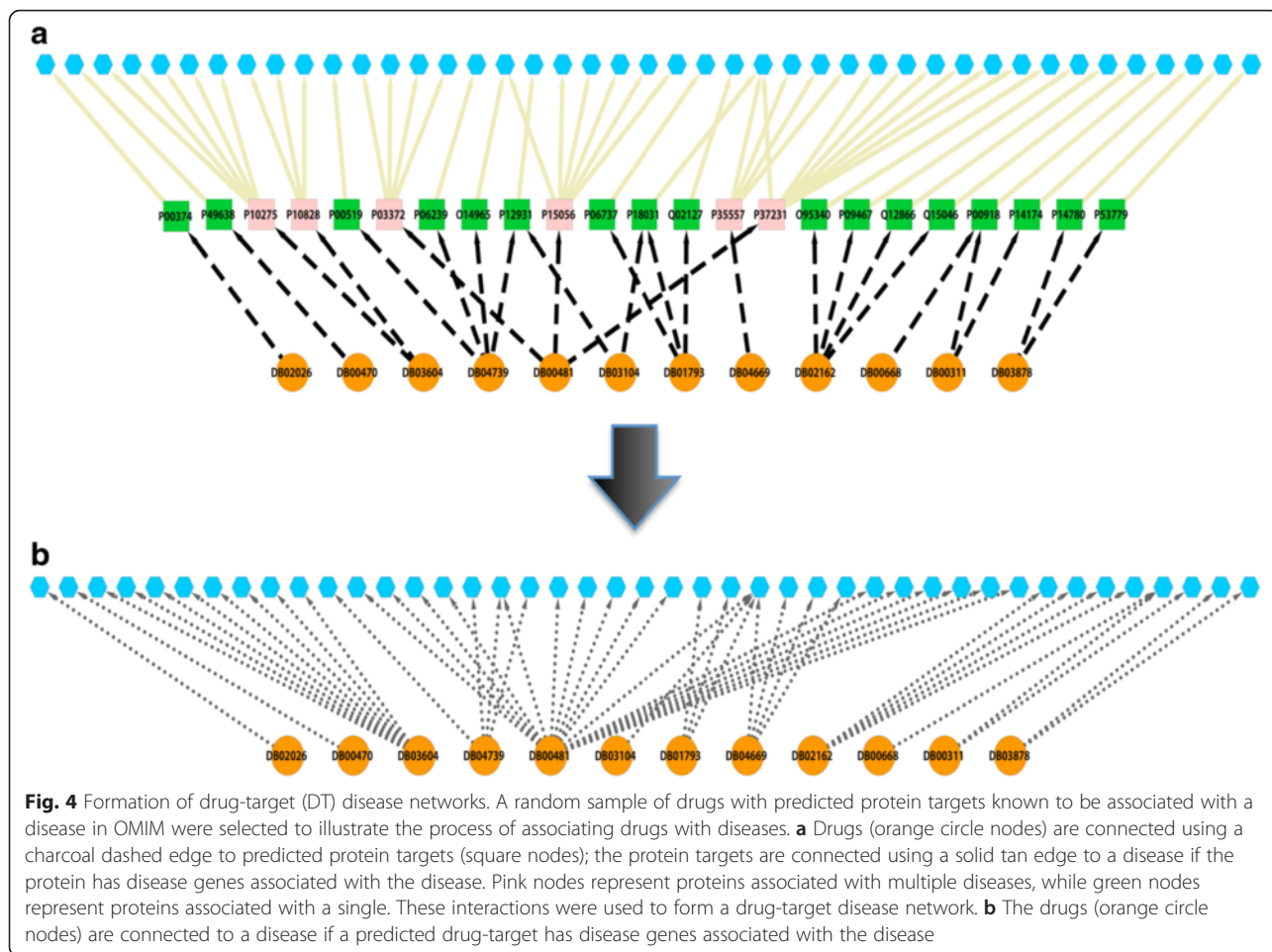
Using drug-target signatures from TMFS and the DGIdb [34], a comprehensive resource of experimentally determined drug-target associations curated from multiple large publically available databases, drugs were associated with diseases using the hypergeometric test (Fig. 3a) in R [35] at each of the following biological levels: direct protein targets, cell signaling pathways, molecular functions and PPIs. Drugs with $P < 0.05$ had their *P*-values log-transformed and normalized to the value of the most significantly-associated drug, resulting in values on the 0–1 unit range as illustrated in Fig. 3b. All non-significant *P*-values were automatically normalized to a value of 0. Normalization minimizes discrepancies found in the *P*-value ranges between different biological effect categories.

For each drug i, normalized values corresponding to each biological effect tier were used to calculate a drug-disease association Z-score used for ranking:

$$Z_i = aA + bB + cC + dD \tag{2}$$

where A, B, C, and D correspond to the normalized values for drug-direct target, –pathway, –function, and –PPI associations, respectively. In illustrative Fig. 3c, A, B, C, and D correspond to zgene, zpathway, zfunction, and zPPI, respectively. Associated weights a, b, c, and d were set to the values of 2, 1, 0.5, and 0.25, respectively, as to prioritize direct binding of disease-regulated gene products with each subsequent level of activity receiving lower weights (Fig. 3c). This configuration was determined to best prioritize experimentally validated drugs for the given indication, and allowed for drugs highly associated with disease mechanisms at pathway, function, and indirect proteins levels to be recognized as candidates even when gene-level significance of association was poor. PPIs were given the least weight as many interactions tend to occur simultaneously within the diseased cell and prioritizing relevant interactions is difficult due to the simultaneous expression of thousands of proteins. Drugs are ranked in descending order by Z-score (Fig. 3d). High Z-scores indicate a drug's potential to most significantly and simultaneously target the greatest amount of direct proteins, pathways, functions and PPIs associated with the disease.

Fig. 3 Hypergeometric test schematic for drug-disease association at each level of biological activity. Each drug is associated with a given disease at each level of biological action by the hypergeometric test. **a** Given a gene, pathway, function, or indirect protein 'universe', the hypergeometric test allows one to determine the probability that coincident drawings between two samples drawn from that universe is due to random chance. Therefore, the statistical significance of having hits (common items) between drug-associated biological factors and disease-associated factors is derived. **b** Computation of hypergeometric *p*-values and subsequent normalization for integration into cumulative score. **c** Computation of drug-disease association Z-score. **d** Ranking scheme by drug-disease association Z-score in descending order. That is, Zi exhibits the highest system-wide statistical association (highest-magnitude Z-score), followed by Zi + 1, Zi + 2, Zi + 3, and so forth

Fig. 4 Formation of drug-target (DT) disease networks. A random sample of drugs with predicted protein targets known to be associated with a disease in OMIM were selected to illustrate the process of associating drugs with diseases. **a** Drugs (orange circle nodes) are connected using a charcoal dashed edge to predicted protein targets (square nodes); the protein targets are connected using a solid tan edge to a disease if the protein has disease genes associated with the disease. Pink nodes represent proteins associated with multiple diseases, while green nodes represent proteins associated with a single. These interactions were used to form a drug-target disease network. **b** The drugs (orange circle nodes) are connected to a disease if a predicted drug-target has disease genes associated with the disease

Fig. 5 Predicted drug-target (DT) disease network. The DT disease bipartite network is generated using the top 1-ranked DT predictions and disorder-disease gene associations from OMIM. Drug nodes (circles) are connected to disease nodes (squares) if a drug is predicted to target a protein that has disease genes associated with the disease. Disease nodes are colored according to their MeSH disease category; color classification given in legend. The size of node is proportional to the number of degrees (connections)

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 7 of 18

Thus, drugs with the highest Z-scores are prioritized for repurposing due to their systems-wide effects.

Results & discussions

Prediction of empirical drug-disease associations

DGE-NET predicted drug associations to diseases with known etiologies by way of direct gene aberrations, as annotated in OMIM (Fig. 4). The DT-disease network contains 562 drugs (only those appearing as the top 1-ranked for their respective protein target) and 296 diseases, with the largest component containing 498 drugs (Fig. 5; Additional file 1: Table S1). The neoplasm and "nutritional and metabolic" disease classes are found centrally, reflecting the large number of drugs already approved for them and a notable potential for repurposing.

Given their topology in the network, associated drugs have potential polypharmacology to other disease classes. More specialized diseases tend to occupy peripheral areas of the DT-disease network, exhibiting a smaller degree of node connectivity and suggesting increasingly unique pathogenic factors. Such diseases include digestive, urogenital, "hemic and lymphatic", and respiratory disorders. By contrast, the DT-cancer network exhibits high connectivity, with the average degree of drug nodes being 1.7 and 57 of 159 having a degree higher than 1 (Fig. 6). 26 drugs are predicted to target colorectal cancer, several of which are also predicted to target breast cancer. This is reflected in clinical practice, where several drugs are utilized across multiple cancers. The biologically sensible topology of the network provides further validation: biologically-related cancers are
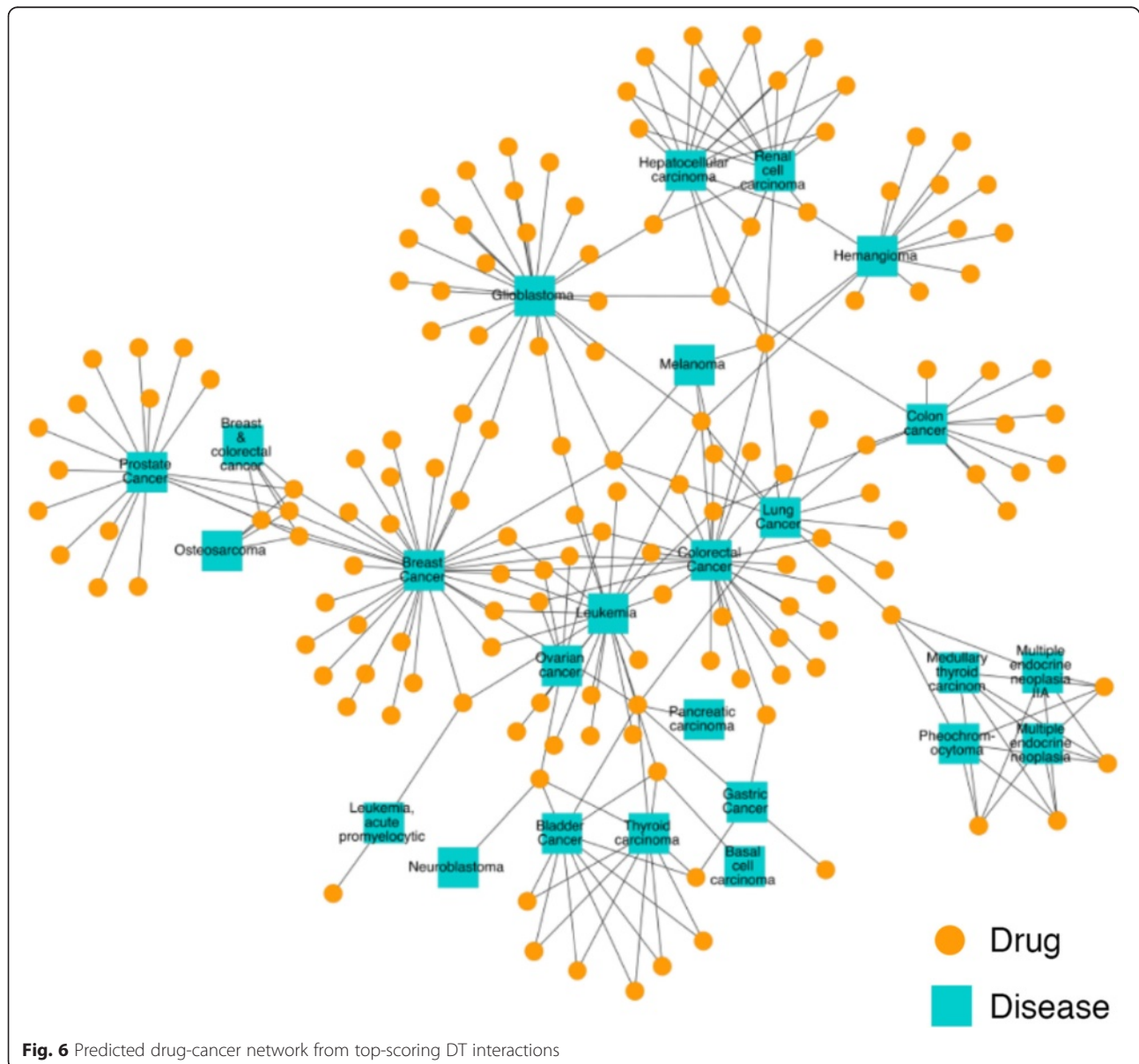

Fig. 6 Predicted drug-cancer network from top-scoring DT interactions

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 8 of 18

clustered together through their predicted drugs. For instance, the bottom right cluster contains the endocrine gland tumors medullary thyroid carcinoma, multiple endocrine neoplasia (MEN), and pheochromocytoma, whereas the unique endothelial-originating hemangioma is found isolated in the top right.

Drug-disease predictions were validated via data found in the primary literature (Additional file 2: Table S2). Out of 526 predicted drug-disease associations, 51 were validated. Full coverage is not attainable, as many drug-disease associations have not yet been examined. Nonetheless, some predicted drug-disease combinations have been well studied, such as lisinopril for diabetes-associated microvascular complications [36]. Other associations include the anti-hookworm mebendazole for hepatocellular carcinoma and the antibiotic ceftriaxone for bladder cancer. Thus, for diseases with strong single-gene known associations, DGE-NET is able to reliably predict clinically relevant drug-disease associations by forming accurate drug-target associations. These data collectively demonstrate the ability of DGE-NET to establish known and novel drug-disease associations.

Expansion of the drug-target prediction space to systems pharmacology

Many diseases exhibit complexity in implicating multiple perturbations rather than single deciding gene associations, and this necessitates a complex systems pharmacology perspective for clinical treatment. Drugs were therefore associated with pathways using KEGG annotations of their predicted targets. Mazindol (DB00368) and sulfadiazine (DB00359) had the least number of predicted pathways (Fig. 7). Mazindol is a tricyclic anorexigenic known to affect the noradrenergic, dopaminergic and serotonergic pathways (KEGG Drug D00367). Sulfadiazine is a sulfonamide used to treat bacterial infections by specifically inhibiting the folate biosynthesis pathway (KEGG Drug D00587). DGE-NET was able to recapitulate their specificity for those pathways. Alternatively, kinase inhibitors and nucleoside analogs such as nelarabine (DB01280) disrupt multiple pathways (Fig. 7). The KEGG Drug corpus was also used to validate 103 drug-pathway associations across 59 drugs (Table 1). Thus, DGE-NET is able to reliably associate drugs with biological pathways important in disease processes.

DGE-NET also related predicted DT signatures to molecular functions (Fig. 7). Deferasirox, an iron chelator, was predicted to affect the greatest number of molecular functions. According to the Institute for Safe Practices, deferasirox was the second most suspected drug in reported patient deaths [37]. This may be due to its potential to disrupt many molecular functions as predicted by DGE-NET. Anti-neoplastic drugs were also predicted to alter a large number of functions (Fig. 8). This reflects
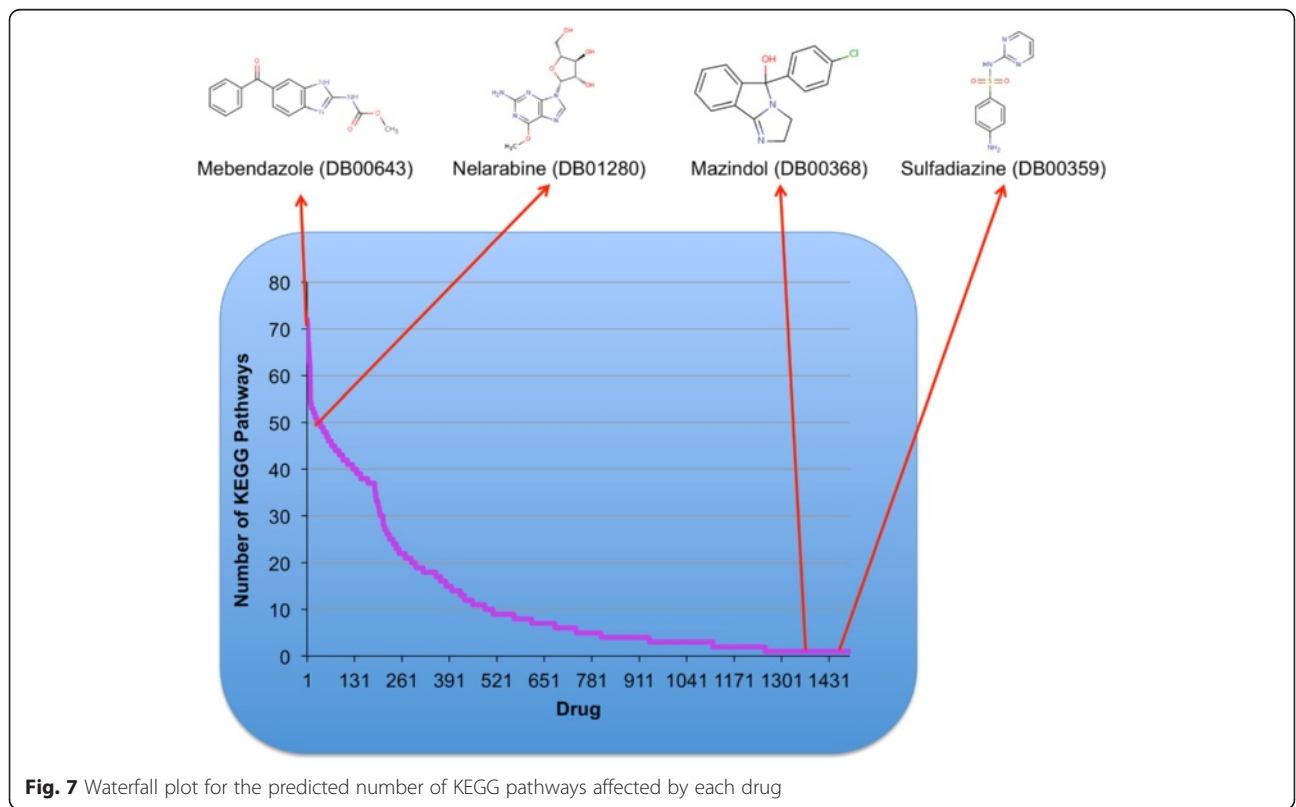

Fig. 7 Waterfall plot for the predicted number of KEGG pathways affected by each drug

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 9 of 18

Table 1 Validations of predicted drug-pathway associations via the KEGG Drug database

Drug	KEGG Drug ID	KEGG Pathway
Acetohexamide	D00219	Type II diabetes mellitus
Aripiprazole	D01164	Gap junction
Bezafibrate	D01366	Adipocytokine signaling pathway
Bicalutamide	D00961	Pathways in cancer, Prostate cancer
Candesartan	D00626	Vascular smooth muscle contraction
Carvedilol	D00255	Vascular smooth muscle contraction
Celecoxib	D00567	VEGF signaling pathway
Cilostazol	D01896	Insulin signaling pathway
Clozapine	D00283	Gap junction
Conivaptan	D01236	Vascular smooth muscle contraction
Danazol	D00289	Oocyet meiosis, Progesterone-mediated oocyte maturation, Pathways in cancer
Dasatinib	D03658	MAPK signaling pathway, ErbB signaling pathway, Cytokine-cytokine receptor interaction, VEGF signaling pathway, Pathways in cancer, Chronic myeloid leukemia
Diflunisal	D00130	VEGF signaling pathway
Domperidone	D01745	Gap junction
Droperidol	D00308	Gap junction
Drospirenone	D03917	Aldosterone-regulated sodium transport
Dydrogesterone	D01217	Oocyte meiosis, Progesterone-mediated oocyte meiosis
Eltrombopag	D03978	Cytokine-cytokine receptor interaction, Jak-STAT signaling pathway
Epoprostenol	D00106	Vascular smooth muscle contraction
Eprosartan	D04040	Vascular smooth muscle contraction
Erlotinib	D07907	MAPK signaling pathway, ErbB signaling pathway, Cytokine-cytokine receptor interaction, Pathways in cancer, Pancreatic cancer, Non-small cell lung cancer
Fenofibrate	D00565	Adipocytokine signaling pathway
Floxuridine	D04197	Pyrimidine metabolism
Flupenthixol	D01044	Gap
Flurbiprofen	D00330	VEGF signaling pathway
Flutamide	D00586	Pathways in cancer, Prostate cancer
Gemcitabine	D02368	Purine metabolism, Pyrimidine metabolism
Gliclazide	D01599	Type II diabetes mellitus
Glipizide	D00335	Type II diabetes mellitus
Haloperidol	D00136	Gap junction
Imatinib	D01441	MAPK signaling pathway, Cytokine-cytokine receptor interaction, Hematopoietic cell lineage, Pathways in cancer, Chronic myeloid leukemia
Indacaterol	D09318	Endocytosis
Indomethacin	D00141	VEGF signaling pathway
Ketoprofen	D00132	VEGF signaling pathway

Table 1 Validations of predicted drug-pathway associations via the KEGG Drug database *(Continued)*

Lapatinib	D04024	MAPK signaling pathway, ErbB signaling pathway, Cytokine-cytokine receptor pathway, Pathways in cancer
Levonorgestrel	D00950	Oocyte meiosis, Progesterone-mediated oocyte maturation
Losartan	D08146	Vascular smooth muscle contraction
Methysergide	D02357	Gap junction
Milrinone	D00417	Progesterone-mediated oocyte maturation
Mitiglinide	D01854	Type II diabetes mellitus
Naproxen	D00118	VEGF signaling pathway
Nilutamide	D00965	Pathways in cancer, Prostate cancer
Norethindrone	D00182	Oocyte meiosis, Progesterone-mediated oocyte maturation
Olmesartan	D01204	Vascular smooth muscle contraction
Oxaprozin	D00463	VEGF signaling pathway
Piroxicam	D00127	VEGF signaling pathway
Progesterone	D00066	Oocyte meiosis, Progesterone-mediated oocyte maturation
Propericiazine	D01485	Gap junction
Regadenoson	D05711	Vascular smooth muscle contraction
Risperidone	D00426	Vascular smooth muscle contraction, Gap junction
Salsalate	D00428	VEGF signaling pathway
Silodosin	D01965	Vascular smooth muscle contraction
Sorafenib	D08524	MAPK signaling pathway, ErbB signaling pathway, Cytokine-cytokine receptor interaction, Chemokine signaling pathway, mTOR signaling pathway, VEGF signaling pathway, Natural killer cell mediated cytotoxicity, Pathways in cancer, Renal cell carcinoma
Sulindac	D00120	VEGF signaling pathway
Sunitinib	D06402	MAPK signaling pathway, Cytokine-cytokine receptor interaction, VEGF signaling pathway, Pathways in cancer
Telmisartan	D00627	Vascular smooth muscle contraction
Testosterone	D00075	Pathways in cancer, Prostate cancer
Vandetanib	D06407	MAPK signaling pathway, ErbB signaling pathway, Cytokine-cytokine receptor interaction, VEGF signaling pathway, Pathways in cancer

their polypharmacology as a class of drugs, as they are designed to affect cell signaling and growth through multiple mechanisms. As a result, these drugs also exhibit high toxicity. Such analysis of molecular function can have the advantage of identifying broad- or specific-acting drugs for enriched clinical efficacy or minimized toxicity.

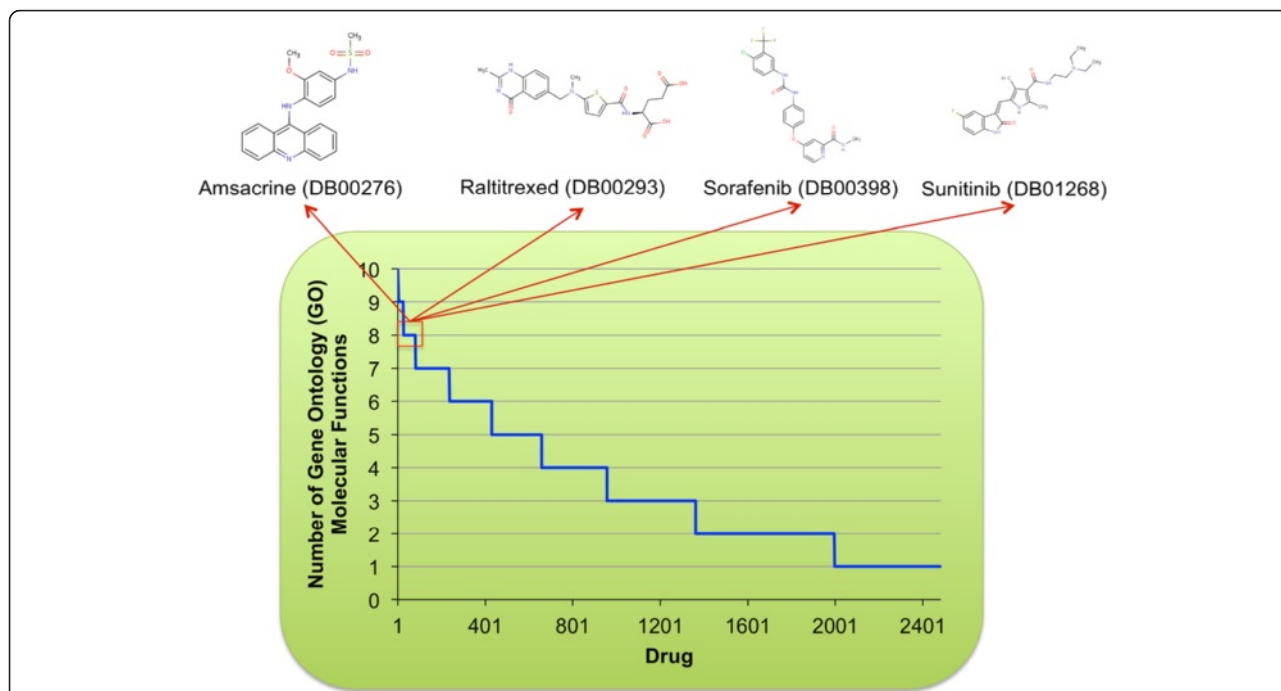The incorporation of protein-protein interactions (PPIs) further increased the robustness of DGE-NET,

Issa et al. BMC Bioinformatics (2016) 17:202

Page 10 of 18

Fig. 8 Waterfall plot for the predicted number of GO molecular functions affected by each drug. Inset highlights four anti-neoplastic drugs predicted to disrupt the greatest number of functions from the anti-neoplastic drug class

providing insight into unexpected biological similarities among drugs. For example, fluoxymesterone (DB01185) and amscarine (DB00276) are chemically and structurally unrelated. However, our method predicted that they would bind androgen receptor and B-Raf, respectively, both of which interact with MAPK1. It is through the PPI with MAPK1 that these drugs link to pathways in cancer (KEGG hsa:05200). Other drug-PPI validations are listed in Table 2 [38–48]. To highlight the importance of PPIs in attaining a mechanistic understanding of drug effects, we specifically assessed the predicted effects of ezetimibe (Fig. 9; Additional file 3: Table S3). Ezetimibe (DB00983) is a cholesterol-lowering drug

used for improving cardiovascular health and has also been associated with increased incidence of cancer [49, 50]. PPIs derived from predicted targets for ezetimibe are highly clustered, indicating that the affected biological space is tightly coordinated through those targets and greatly perturbed by the actions of ezetimibe (Fig. 9). These clustered interacting targets are mainly involved in cell growth, differentiation and signal transduction. Functional annotation using both direct and indirect ezetimibe targets implicates pathways and functions involved in carcinogenesis (Additional file 3: Table S3). Thus, the present DGE-NET prediction of ezetimibe's pro-tumorigenic effects warrants further investigation.

Table 2 Validations of predicted drug-PPI interactions

Drug Name	Protein #1 (direct binding partner)	Protein #2 (PPI)	Reference
Bicalutamide	ABL1	CASP9	Danquah et al. Pharm Res. 26(9):2081–92. (2009) [38]
	ABL1	CCNA2	Katayama et al. Int J Oncol. 36(3):553–62. (2010) [39]
	ABL1	MAPK11	Malinowska et al. Endocr Relat Cancer. 16:155–169. (2009) [40]
Cladribine	ADA	DCK	Sasvári-Székely et al. Biochem Pharmacol. 56(9):1175–1179. (1998) [41]
Chlordiazepoxide	AKT1	NR3C1	Curtin et al. Brain Behav, Immun. 23(4): 535–547. (2009) [42]
Progeterone	AR	F2	Oger et al. Arterioscler Thromb Vasc Biol. 23:1671–1676. (2003) [43]
Cyproterone	AR	CASP3	Eckle et al. Toxicol Pathol. 32:9–15. (2004) [44]
	AR	NR3C1	Honer et al. Mol Pharmacol. 63(5):1012–1020. (2003) [45]
Telmisasrtan	BCL2	IL2	Syrbe et al. Hypertens Res. 30(6):521–527. (2007) [46]
Sorafenib	BRAF	PRKCQ	Jane et al. J Pharmacol Exp Ther. 319(3):1070–1080. (2006) [47]
Methotrexate	DHFR	CDK2	Maddika et al. J Cell Sci. 121:979–988. (2008) [48]

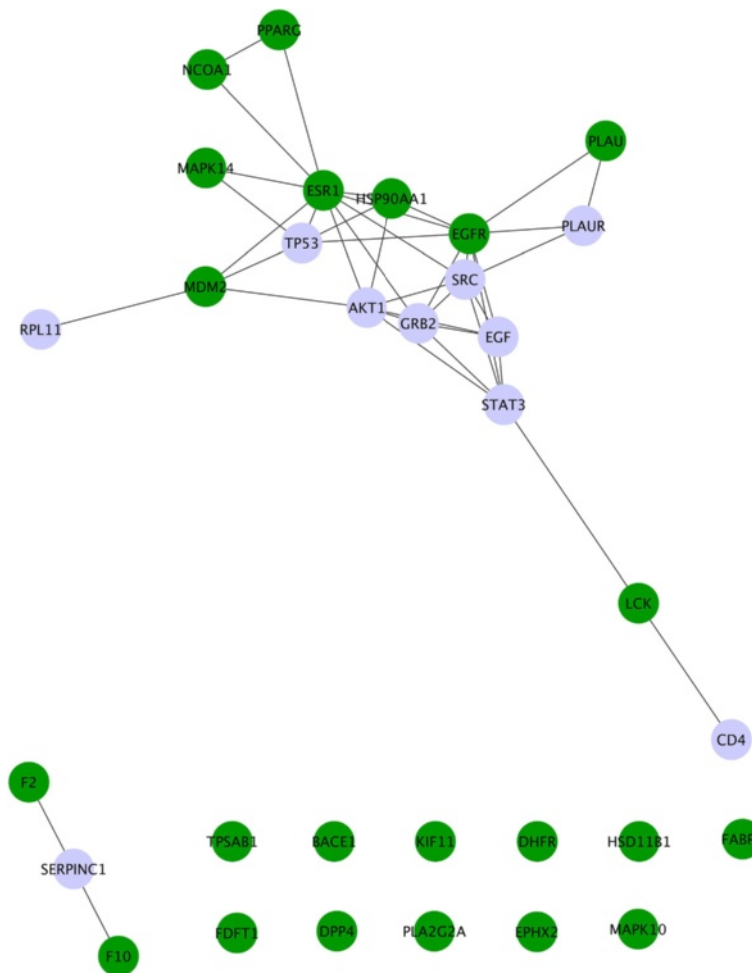Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 11 of 18

Fig. 9 Ezetimibe protein-protein interaction (PPI) network. Direct targets (green nodes) predicted for ezetimibe from TMFS were used to establish interactions between direct targets as well as indirect targets (light purple nodes) using the ExPASy STRING database with a confidence score cutoff greater than 0.95

Incorporation of autoimmune disease-related gene expression data for polypharmacology-driven drug repurposing

Autoimmune diseases are systemic or local pro-inflammatory pathologies with multiple etiologies. Current therapeutics such as corticosteroids, methotrexate and anti-TNF biologics focus on regulating inflammation, and immunosuppression. In addition to acting non-specifically these medications do not address the full extent of effector tissue pathobiology. A treatment approach rooted in polypharmacology may be more efficacious and offers the potential for limiting side effects. For proof-of-concept, we apply DGE-NET as a gene expression-based polypharmacology prediction method (Fig. 2) for rheumatoid arthritis and inflammatory bowel disease.

Rheumatoid arthritis (RA) is a painful multi-joint destructive disease. Joint synovium, usually 1–2 cells thick, becomes inflamed and reaches multicellular thickness due to infiltration of immune effector cells and activation

and subsequent proliferation of fibroblast-like synoviocytes (FLS). Cellular molecular cross-talk, infiltration and proliferation lead to pannus formation, which acts analogously to an invasive tumor and causes joint destruction. As FLS cells are critical mediators of RA, we applied our method using differentially expressed genes when comparing activated FLS cells from RA patients and quiescent FLS cells from non-RA patients (GSE55235 and GSE55457). A consensus drug list was constructed by combining the top 100 (~ Top 10 % of total drug database) predicted drugs for each study and extracting those that are present in one or both lists, ranked by mean association Z-score (Additional file 4: Table S4). Shown in Table 3 are those drugs from the consensus drug list that are currently used for RA, or have been found to be potentially useful in the clinic [51–54, 97–103]. Drugs currently used in the clinic were recapitulated in our list, such as anti-TNF biologics adalimumab and etanercept, as well as the NSAID sulindac. Non-approved

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 12 of 18

Table 3 Validations of predicted drug indications for RA and IBD from consensus drug lists, ordered by drug list ranking

Rheumatoid Arthritis (RA)	Reference for Validation	Inflammatory Bowel Disease (IBD)	Reference for Validation
Alvocidib	Sekine et al. J Immunol. 180(3):1954–1961 (2008) [51]	Sulfasalazine	Klotz et al. N Engl J Med. 303(26):1499–1502 (1980) [105]
Karenitecin	Liu et al. Med Res Rev. 35(4):753-89 (2015) [53]	Olsalazine	Baumgart et al. Lancet. 369(9573):1641–1657 (2007) [106]
Sulindac	Brogden et al. Drugs. 16(2):97–114 (1978) [97]	Tetomilast	Keshavarzian et al. Expert Opin Investig Drugs. 16(9):1489–1506 (2007) [107]
Sunitinib	Fuyura et al. Mod Rheumatol. 24(3):487–491 (2013) [52]	Inosine	Mabley et al. Am J Physiol Gastrointest Liver Physiol. 284(1):G138-G144 (2003) [108]
INCB28050	Taylor et al. Ann Rheum Dis. 73:A31 (2014) [99]	Thioproperazine	Lechin et al. J Clin Gastroenterol. 4(5):445–450 (1982) [56]
Amodiaquine	Kersley et al. Lancet. 2(7108):886–888 (1959) [54]	Etoricoxib	El Miedany et al. Am J Gastroenterol. 101(2):311–317 (2006) [109]
Raltitrexed	van der Heijden et al. Scand J Rheumatol. 43(1):9–16 (2014) [100]	Balsalazide	Carter et al. Gut. 53(Suppl 5):V1-V16 (2004) [110]
BIRB 796	Page et al. Arthritis Rheum. 62(11):3221–3231 (2010) [101]	Thalidomide	Gerich et al. Ailment Pharmacol Ther. 41(5):429–437 (2015) [59]
Adalimumab	Weinblatt et al. Arthritis Rheum. 48(1):35–45 (2003) [102]	Rosiglitazone	Ramakers et al. J Clin Immunol. 27(3):275–283 (2007) [57]
Etanercept	Moreland et al. Ann Intern Med. 130(6):478–486 (1999) [103]	Irbesartan	Ray et al. Gut. 62(S1):A525-A525 (2013) [60]
Minocycline	O'Dell et al. Arthritis Rheum. 40(5):842–848 (1997) [104]	Chloroquine	Nagar et al. Int Immunopharmacol. 21(2):328–335 (2014) [111]

drugs currently being studied for RA also appeared. These include kinase inhibitors such as alvocidib [51] and sunitinib [52], the topoisomerase inhibitor karenitecin [53], and the chloroquine-related compound amodiaquine [54]. Predicted RA indication for these drugs, which are generally anti-cancer agents, illustrates an important mechanistic underpinning of RA with respect to FLS cells in that activated FLS mimic cancer cell progression [55]. Regardless of the activating stimulus (e.g. TNF-α), our polypharmacological method focuses on downstream gene expression, signaling, and functional effects in activated FLS cells. This highlights the cancer-like mechanisms of pathogenesis and prioritizes those drugs that are able to simultaneously disrupt the greatest number of those mechanisms. In addition, because antibodies have single-target effects, we were surprised by their predicted indications for RA. However, if that target has many pathology-related pleiotropic downstream effects, such as TNF-α, then such drugs would be prioritized due to the pathway and function terms in our equation. Thus, DGE-NET is capable of making important polypharmacological associations beyond immediate gene targets.

DGE-NET also predicted drugs for inflammatory bowel disease (IBD), also a multi-etiological immune-related collection of disorders. Differential gene expression analysis was performed by, comparing normal and inflamed bowel tissues (GSE52746 and GSE11223). Like

the RA dataset, a consensus list of the Top 100 drugs obtained from each IBD study was constructed (Additional file 5: Table S5), and therapeutic validations from this list are recapitulated in Table 3 [56–60, 104–110]. Our method predicted the known IBD drug sulfasalazine, serving as an important litmus. Other predicted drugs that are promising in experimental settings and from diverse chemical classes include the anti-psychotic thioproperazine [56], the anti-diabetic thiazlidinedione rosiglitazone [57], the leukotriene receptor antagonist tetomilast [58], and thalidomide [59]. Interestingly, DGE-NET predicted the angiotensin receptor blocker (ARB) irbesartan as potential therapy. A recent preliminary study implicates the role of angiotensin receptors in intestinal fibrosis in Crohn's disease [60], a type of IBD, but greater investigation is needed.

In addition to recapitulating known drug associations, we predicted the drugs topotecan and mebendazole for repurposing to rheumatoid arthritis. Topotecan is a DNA topoisomerase 1 (Top1) inhibitor used for NSCLC cancer and has been given both orally and intravenously. Topoisomerases have been implicated in rheumatoid arthritis etiology [61], and the established Top1 inhibitor camptothecin (CPT) has been shown to be effective in a murine collagen-induced RA model [62]. Koo et al. developed a novel nanocarrier for CPT called CPT-SSM-VIP, which denotes micelles to overcome solubility issues and vasoactive intestinal peptide (VIP) for active targeting. As CPT provides evidence for Top1 inhibition

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 13 of 18

in RA, we also pursued topotecan. Although it can be inferred that topotecan could be an effective anti-arthritic via topoisomerase, many other unreported targets were predicted for topotecan that could mediate potential efficacy. These include multiple tyrosine-protein kinases (BTK, CSK, LCK, TTK, ITK, LYN), non-tyrosine kinases (AURK1, PIK3CG), as well as cyclin A2. Mebendazole is an anti-hookworm tubulin inhibitor with anti-cancer potential through mammalian crossover tubulin [63] and kinase inhibition [64]. We previously predicted many novel protein kinase targets for mebendazole [17]. Kinase inhibition is a sought after therapeutic strategy for rheumatoid arthritis, especially as non-biologic treatment alternatives and for methotrexate-resistant cases [65–67]. Inhibitors of spleen tyrosine kinase (Syk) and Janus kinases (Jaks) have shown short-term efficacy, but other kinases inhibitors with good long-term effect profiles may also exist. Other kinases implicated in RA pathogenesis include aurora kinases [68] and cyclin-dependent kinases (CDKs) [69]. Mebendazole may serve as a good non-biologic disease-modifying antirheumatic drug (DMARD) given its historic use, low toxicity profile, and its effect on multiple kinases.

In another proof-of-concept, we applied DGE-NET to two neurodegenerative disorders, Alzheimer's disease (AD) and Parkinson's disease (PD). Table 4 summarizes those drugs predicted to be in the top 50 for AD and PD by DGE-NET that are currently validated for standard or potential therapeutic use [70–86]. The complete top 50 predicted drugs for these diseases and their validations are found in Additional file 6: Table S6. Others listed are currently undergoing pre-clinical or clinical investigation. Of note is that memantine, an approved drug for AD, appears beyond the top 50 but within the top 500. This drug exhibits less polypharmacology but is still effective given the importance of its direct targets and pathways for AD disease processes (i.e. NMDA receptor antagonism reducing glutamate excitotoxicity of neurons [87]). Thus, it can be hypothesized that drugs found higher up in the rank list may be more effective than the current clinical standards of care as those drugs theoretically alter a greater proportion of disease-associated protein targets and biological effects simultaneously.

Sunitinib has been identified as a lead candidate having the potential to mitigate the development of oxidant injury to endothelial cells associated with AD [79]. Sunitinib could affect the vascular activation mechanisms of pathogenesis in AD by reducing the expression of amyloid beta, thrombin, tumor necrosis factor alpha, interleukin-1 beta, interleukin-6, and matrix metalloproteinase 9, and other factors associated with neurodegenerative disorders [79, 88, 89]. This anti-angiogenic property has been

Table 4 Validations of top 50 predicted drug indications for AD and PD, ordered by ranking

Alzheimer's Disease (AD)	Reference for Validation	Parkinson's Disease (PD)	Reference for Validation
Rasagiline	Weinreb et al. Neurotherapeutics. (6)1:163–74. (2009) [70]	Dextroamphetamine	Parkes et al. J Neurol Neurosurg Psychiatry. 38(3):232–7 (1975) [83]
Interferons	Grimaldi et al. J Neuroinflammation. 11:30 (2014) [71]	Orphenadrine	Bersani et al. Clin Neuropharmacol. 13(6):500–6 (1990) [84]
Calcium	Woods et al. Adv Exp Med Biol. 740:1193–217 (2012) [72]	Quinacrine	Tariq et al. Brain Res Bull. 54(1):77–82 (2001) [85]
Dovitinib	Li et al. Medical Hypotheses. (80)4:341–44. (2013) [73]	Atomoxetine	Weintraub et al. Neurology. 75(5):448–55 (2010) [86]
Somatropin Recombinant	Ling et al. Growth Horm IGF Res. (17)4:336–41 (2007) [74]		
Aripiprazole	De Deyn et al. Expert Opin. Pharmacother. (14)4:459–74 (2013) [75]		
Clozapine	Tariot et al. Clin Geriatr Med. (17)2:359–76 (2001) [76]		
Quercetin	Ansari et al. J Nutr Biochem. 20(4):269–75 (2009) [77]		
Flavopiridol	Pallàs et al. Med Hypotheses. 64(1):120–3 (2005) [78]		
Sunitinib	Grammas et al. J Alzheimers Dis. 40(3):619–30 (2014) [79]		
Risperidone	Katz et al. Int J Geriatr Psychiatry. (60)2:107–15 (2007) [80]		
Genistein	Valles et al. Brain Res. 1312:138–44 (2010) [81]		
Dasatinib	Dhawan et al. J Neuroinflammation. 9:117 (2012) [82]		

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 14 of 18

previously shown to be a major component of the anti-cancer acitivity of sunitinib [90]. Figure 10 illustrates the polypharmacology of sunitinib, at each level of biological activity, predicted by DGE-NET to coincide with significantly AD-associated factors. Single-agent or combination therapies that exploit multiple aspects of disease process are assumed to be efficacious, requiring lower dosages than current therapies and reducing the likelihood of resistance.

In addition to therapeutic drug repurposing candidates, DGE-NET reported drugs that are known to be contra-indicated for their respective diseases. Minocycline and tretinoin, both of which are used to treat acne, may have IBD toxicity. Minocycline is a tetracycline antimicrobial with a potential association with IBD (Additional file 5: Table S5) [91]. Tretinoin is a topical retinoid that is structurally related to isotretinoin, an oral medication used for more severe acne. While tretinoin itself is safe, isotretinoin has been implicated in causing IBD (Additional file 5: Table S5) [92], though this finding is controversial. It could be extrapolated that if tretinoin was given orally and at higher doses that IBD may be a consequence. Others include methysergide, a prophylactic drug that is contra-indicated for RA and other collagen diseases (Additional file 4: Table S4) [93], indomethacin, a non-

selective non-steroidal anti-inflammatory drug known to exacerbate IBD (Additional file 5: Table S5) [94, 95], quetiapine, an atypical antipsychotic associated with increased cognitive decline in AD (Additional file 6: Table S6) [96], and methamphetamine, which has been linked with an increased risk of PD (Additional file 6: Table S6), [97]. The appearance of these drugs is likely due to DGE-NET not discriminating between agonistic and antagonistic effects of drugs but rather forming non-directional drug-target-effect associations. Counter-therapeutic drug actions are therefore incorporated, so long as they correspond with disease-associated biological activity.

Conclusions

DGE-NET is able to predict drug-target interactions and contextualize their biological effects at the levels of protein-protein interactions, biological pathways, and molecular functions. It further integrates gene expression signatures for identification of systems-based disease-relevant targets and prioritization of drugs that exhibit a desired polypharmacology. DGE-NET recapitulated known therapeutic and contraindicated drugs for rheumatoid arthritis and inflammatory bowel disease and led to the identification of mebendazole
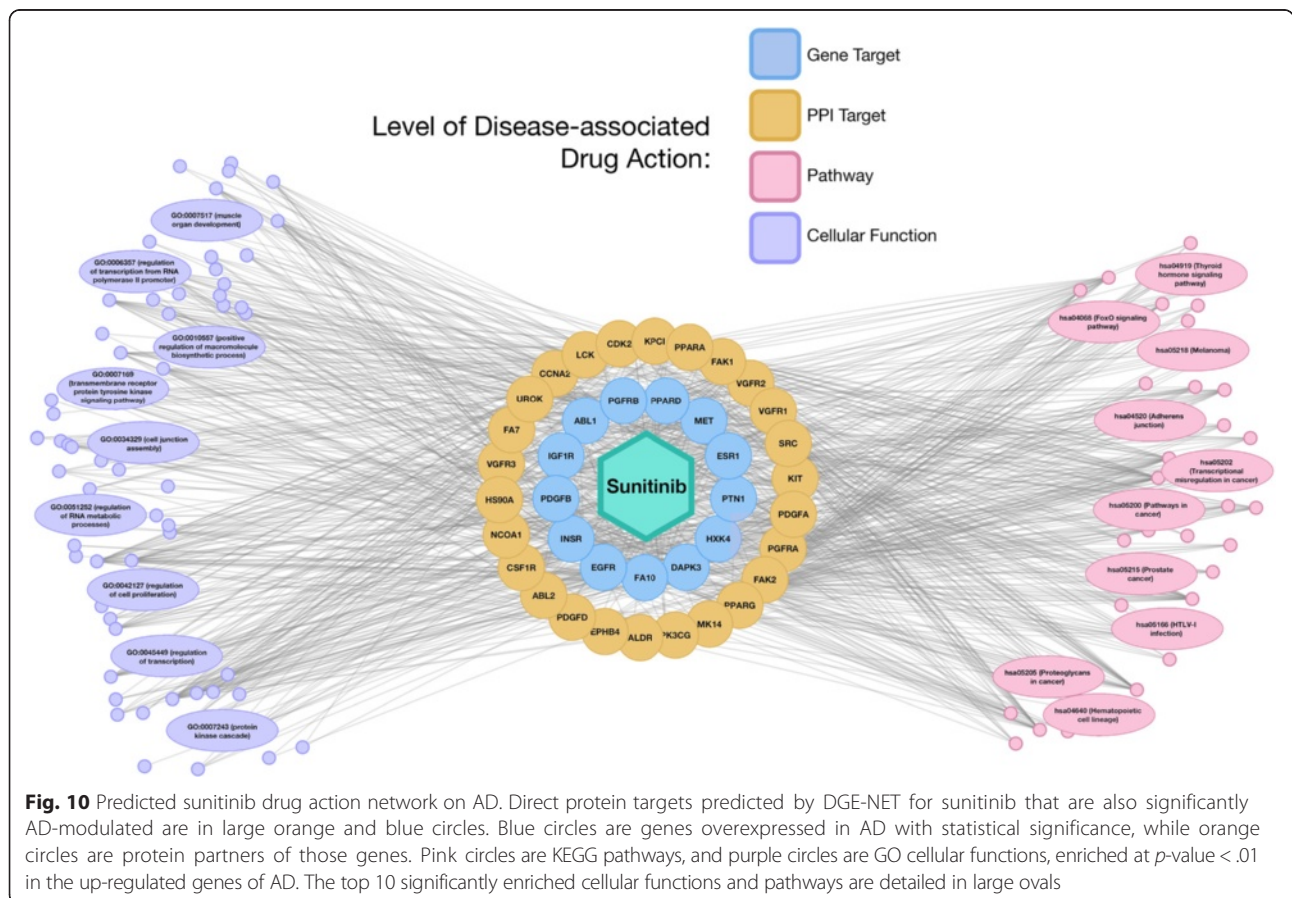

Fig. 10 Predicted sunitinib drug action network on AD. Direct protein targets predicted by DGE-NET for sunitinib that are also significantly AD-modulated are in large orange and blue circles. Blue circles are genes overexpressed in AD with statistical significance, while orange circles are protein partners of those genes. Pink circles are KEGG pathways, and purple circles are GO cellular functions, enriched at *p*-value < .01 in the up-regulated genes of AD. The top 10 significantly enriched cellular functions and pathways are detailed in large ovals

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 15 of 18

as drug repurposing candidate for rheumatoid arthritis. Its ability to do so can also be extended to other small molecules with the potential to act as endogenous drugs to alter physiology, such as metabolites. We are currently pursuing the application of DGE-NET to cancer-associated metabolites to potentially explain the mechanisms behind metabolite-disease phenotypic associations. DGE-NET ultimately assists in the formulation of drug-disease hypotheses poised for clinical success.

Differential gene expression analysis is one way of assessing disease pathogenesis to find therapeutic targets. DGE-NET is the first computational tool that associates drugs with diseases through multiple tiers of systems biology obtained via gene expression analysis. This not only aids in finding effective drugs but helps bypass issues that arise from traditional gene sequencing approaches such as un-actionable mutations in single nucleotide polymorphisms, which is currently an important limitation in oncology. Importantly, DGE-NET in its current form does not differentiate agonist or antagonist effects of drugs. The next iteration will include this improvement so that DGE-NET can better discriminate between therapeutic agents and drugs that are contraindicated.

Availability of data and materials

Because DGE-NET is applied to publicly available data, the authors have provided a tutorial which describes the step-wise implementation of DGE-NET, in Additional file 7.

Additional files

Additional file 1: Table S1. Predicted drug-target-disease associations using OMIM. For each human protein target crystal structure, the top 40-ranked drugs were associated with a disease through their predicted target. (XLSX 464 kb)

Additional file 2: Table S2. Validations of predicted drug-disease associations from the literature. (XLS 38 kb)

Additional file 3: Table S3. Predicted systems pharmacology of ezetimibe. Targets predicted to directly associate with ezetimibe and their interacting protein partners (confidence score cutoff greater than 0.95) were used to infer pathways and molecular functions (FDR < 0.25) that could be perturbed by ezetimibe. (XLS 42 kb)

Additional file 4: Table S4. Consensus drug rank list for Rheumatoid Arthritis. (XLS 47 kb)

Additional file 5: Table S5. Consensus drug rank list for Inflammatory Bowel Disease. (XLS 51 kb)

Additional file 6: Table S6. Top 50 predicted drugs and validations for Alzheimer's Disease and Parkinson's Disease. (XLS 39 kb)

Additional file 7: Tutorial outlining the manual implementation of the DrugGenEx-NET methodology. (DOCX 14 kb)

Abbreviations

AD: Alzheimer's disease; CDK: Cyclin-dependent kinase; CPT: Camptothecin; CTD: Comparative toxicogenomics database; DAVID: Database for annotation, visualization, and integrated discovery; DGE-NET: DrugGenEx-Net; DMARD: Disease-modifying antirheumatic drug; DT: Drug-target; FAT: Functional annotation tool; FLS: Fibroblast-like synoviocytes; IBD: Inflammatory bowel disease; MEN: Multiple endocrine neoplasia; MeSH: Medical subject headings; OMIM: Online mendelian inheritance in man; PD: Parkinson's disease; PPI: Protein-protein interactions; RA: Rheumatoid arthritis; SDF: Spatial data file; TMFS: Train, match, fit, streamline; VIP: Vasoactive intestinal peptide.

Author details

[1]Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington DC 20057, USA. [2]Department of Biochemistry & Molecular Biology, Georgetown University, Washington DC 20057, USA. [3]Georgetown University Medical Center, Washington DC 20057, USA. [4]George Mason University, 4400 University Dr, Fairfax, VA 22030, USA.

References

1. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov. 2010;9:203–14.
2. Bajorath J. Rational drug discovery revisited: interfacing experimental programs with bio- and chemo-informatics. Drug Discov Today. 2001;6: 989–95.
3. Chanda SK, Caldwell JS. Fulfilling the promise: drug discovery in the post-genomic era. Drug Discov Today. 2003;8:168–74.
4. Pujol A et al. Unveiling the role of network and systems biology in drug discovery. Trends Pharmacol Sci. 2010;31(3):115–23.
5. Anighoro A, Bajorath J, Rastelli G. Polypharmacology: challenges and opportunities in drug discovery. J Med Chem. 2014;57:7874–87.
6. Cheng F, Liu C, Jiang J, Lu W, Li W, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. PLoS Comput Biol. 2012;8:e1002503.
7. Campillos M, Kuhn M, Gavin AC, Jensen LJ, Bork P. Drug target identification using side-effect similarity. Science. 2008;321:263–6.
8. Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, et al. Large-scale prediction and testing of drug activity on side-effect targets. Nature. 2012;486:361–7.
9. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313:1929–35.
10. Chen B, Ding Y, Wild DJ. Assessing drug target association using semantic linked data. PLoS Comput Biol. 2012;8:e1002574.
11. Hu G, Agarwal P. Human disease-drug network based on genomic expression profiles. PLoS One. 2009;4:e6536.
12. Medina-Franco JL, Giulianotti MA, Welmaker GS, Houghten RA. Shifting from the single to the multitarget paradigm in drug discovery. Drug Discov Today. 2013;18(9–10):495–501.
13. Knox C, Law V, Jewison T, Liu P, Ly S, et al. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic Acids Res. 2011;39:D1035–41.
14. Huang R, Southall N, Wang Y, Yasgar A, Shinn P, et al. The NCGC Pharmaceutical Collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. Sci Transl Med. 2011; 3(80):80ps16.
15. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. Nucleic Acids Res. 2007;35:D198–201.
16. Schrödinger, LLC. Schrödinger Release 2013–3: LigPrep, version 2.8. New York: Schrödinger, LLC; 2013.

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 16 of 18

17. Dakshanamurthy S, Issa NT, Assefina S, Seshasayee A, Peters OJ, et al. Predicting new indications for approved drugs using a proteochemometric method. J Med Chem. 2012;55:6832–48.

18. Schrödinger, LLC. Small-Molecule Drug Discovery Suite 2013–3: Glide, version 6.1. New York: Schrödinger, LLC; 2013.

19. Schrödinger, LLC. Small-Molecule Drug Discovery Suite 2013–3: QikProp, version 3.8. New York: Schrödinger, LLC; 2013.

20. Schrödinger, LLC. Small-Molecule Drug Discovery Suite 2013–3: Strike, version 2.4. New York: Schrödinger, LLC; 2013.

21. Kahraman A, Morris R, Laskowski R, Thornton J. Shape variation in protein binding pockets and their ligands. J Mol Biol. 2007;368:283–301.

22. The UniProt Consortium. Activities at the Universal Protein Resource (UniProt). Nucleic Acids Res. 2014;D1:D191–8.

23. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, et al. The human disease network. Proc Natl Acad Sci U S A. 2007;104:8685–90.

24. Yildirim MA, Goh KI, Cusic ME, Barabasi AL, Vidal M. Drug-target network. Nat Biotechnol. 2007;25:1119–26.

25. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28:27–30.

26. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, et al. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014;42:D199–205.

27. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. Nat Genet. 2000;25:25–9.

28. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, et al. AmiGO: online access to ontology and annotation data. Bioinformatics. 2009;25:288–9.

29. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009; 4:44–57.

30. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37:1–13.

31. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013;D1:D808–15.

32. Davis AP, Murphy CG, Johnson R, Lay JM, Lennon-Hopkins K, et al. The comparative toxicogenomics database: update 2013. Nucleic Acids Res. 2013;D1:D1104–14.

33. Barrett T, Whilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets- update. Nucleic Acids Res. 2013; D41:D991–5.

34. Griffith M, Griffith OL, Coffman AC, Weible JV, McMichael JF, et al. DGIdb – Mining the druggable genome for personalized medicine. Nat Methods. 2013;10:1209–10.

35. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

36. Zuanetti G, Latini R, Maggioni AP, Franzosi M, Santoro L, et al. Effect of the ACE inhibitor lisinopril on mortality in diabetic patients with acute myocardial infarction: data from the GISSI-3 study. Circulation. 1997;96: 4239–41.

37. ISMP (2010). "ISMP QuarterWatch(TM)" 15 (12). ISMP Medication Safety Alert. pp. 1–3

38. Danquah M, Li F, Duke III CB, Miller DD, Mahato RI. Micellar delivery of bicalutamide and embelin for treating prostate cancer. Pharm Res. 2009;26: 2081–92.

39. Katayama H, Murashima T, Saeki Y, Nishizawa Y. The pure anti-androgen bicalutamide inhibits cyclin A expression both in androgen-dependent and-independent cell lines. Int J Oncol. 2010;36:553–62.

40. Malinowska K, Neuwirt H, Cavarretta IT, Bektic J, Steiner H, Dietrich H, Moser PL, Fuchs D, Hobisch A, Culig Z. Interleukin-6 stimulation of growth of prostate cancer in vitro and in vivo through activation of the androgen receptor. Endocr Relat Cancer. 2009;16(1):155–69. doi:10.1677/ERC-08-0174. Epub2008Nov14.

41. Sasvári-Székely M, Spasoukotskaja T, Szóke M, Csapó Z, Turi Á, Szántó I, et al. Activation of deoxycytidine kinase during inhibition of DNA synthesis by 2- chloro-2′-deoxyadenosine (Cladribine) in human lymphocytes. Biochem Pharmacol. 1998;56:1175–9.

42. Curtin NM, Boyle NT, Mills KH, Connor TJ. Psychological stress suppresses innate IFN-γ production via glucocorticoid receptor activation: Reversal by the anxiolytic chlordiazepoxide. Brain Behav Immun. 2009;23:535–47.

43. Oger E, Alhenc-Gelas M, Lacut K, Blouch MT, Roudaut N, Kerlan V, et al. Differential effects of oral and transdermal estrogen/progesterone regimens on sensitivity to activated protein C among postmenopausal women a randomized trial. Arterioscler Thromb Vasc Biol. 2003;23:1671–6.

44. Eckle VS, Buchmann A, Bursch W, Schulte-Hermann R, Schwarz M. Immunohistochemical detection of activated caspases in apoptotic hepatocytes in rat liver. Toxicol Pathol. 2004;32:9–15.

45. Honer C, Nam K, Fink C, Marshall P, Ksander G, Chatelain RE, et al. Glucocorticoid receptor antagonism by cyproterone acetate and RU486. Mol Pharmacol. 2003;63:1012–20.

46. Syrbe U, Moebes A, Scholze J, Swidsinski A, Dorffel Y. Effects of the angiotension II type 1 receptor antagonist Telmisartan on monocyte adhesion and activation in patients with essential hypertension. Hypertens Res. 2007;30:521.

47. Jane EP, Premkumar DR, Pollack IF. Coadministration of sorafenib with rottlerin potently inhibits cell proliferation and migration in human malignant glioma cells. J Pharmacol Exp Ther. 2006;19:1070–80.

48. Maddika S, Ande SR, Wiechec E, Hansen LL, Wesselborg S, Los M. Akt-mediated phosphorylation of CDK2 regulates its dual role in cell cycle progression and apoptosis. J Cell Sci. 2008;121:979–88.

49. Drazen JM, D'Agostino RB, Ware JH, Morrissey S, Curfman GD. Ezetimibe and cancer – an uncertain association. N Engl J Med. 2008;359:1398–9.

50. Rossebo AB, Pedersen TR, Boman K, et al. Intensive lipid lowering with simvastatin and ezetimibe in aortic stenosis. N Engl J Med. 2008;359:1343–56.

51. Sekine C et al. Successful treatment of animal models of rheumatoid arthritis with small-molecule cyclin-dependent kinase inhibitors. J Immunol. 2008;180(3):1954–61.

52. Furuya K et al. Therapeutic effects of sunitinib, one of the anti-angiogenetic drugs, in a murine arthritis. Mod Rheumatol. 2014;24(3):487–91.

53. Liu YQ, Li WQ, Morris-Natschke SL, et al. Perspectives on biologically active camptothecin derivatives. Med Res Rev. 2015;35(4):753–89.

54. Kersley GD, Palin AG. Amodiaquine and hydroxychloroquine in rheumatoid arthritis. Lancet. 1959;274(7108):886–8.

55. Ainola MM et al. Pannus invasion and cartilage degradation in rheumatoid arthritis: involvement of MMP-3 and interleukin-1b. Clin Exp Rheumatol. 2005;23:644–50.

56. Lechin F et al. Treatment of ulcerative colitis with thioproperazine. J Clin Gastroenterol. 1982;4(5):445–50.

57. Ramakers JD et al. The PPARγ agonist rosiglitazone impairs colonic inflammation in mice with experimental colitis. J Clin Immunol. 2007;27(3): 275–83.

58. Schreiber S, et al. A randomized, placebo-controlled, phase II study of tetomilast in active ulcerative colitis. Gastroenterology. 2007;132(1):76–86.

59. Gerich ME, et al. Long-term outcomes of thalidomide in refractory Crohn's disease. Aliment Pharmacol Ther. 2015;41(5):429–37.

60. Ray S et al. PTH-102 preliminary evidence for a role of the renin angiotensin system in intestinal fibrosis in crohn's disease using angiotensin receptor immunohistochemistry. Gut. 2013;62 Suppl 1:A252–2.

61. Jackson JK et al. Topoisomerase inhibitors as anti-arthritic agents. Inflamm Res. 2008;57(3):126–34.

62. Koo OM, Rubinstein I, Önyüksel H. Actively targeted low-dose camptothecin as a safe, long-acting, disease-modifying nanomedicine for rheumatoid arthritis. Pharm Res. 2011;28(4):776–87.

63. Sasaki J, Ramesh R, Chada S, Gomyo Y, Roth JA, Mukhopadhyay T. The anthelmintic drug mebendazole induces mitotic arrest and apoptosis by depolymerizing tubulin in non-small cell lung cancer cells. Mol Cancer Ther. 2002;1:1201–9.

64. Li L, Liu Y, Zhang Q, Zhou H, Zhang Y, Yan B. Comparison of cancer cell survival triggered by microtubule damage after turning Dyrk1B kinase on and off. ACS Chem Biol. 2014;9:731–42.

65. Opar A. Kinase inhibitors attract attention as oral rheumatoid arthritis drugs. Nat Rev Drug Discov. 2010;9(4):257–8.

66. Gomez-Puerta JA, Mócsai A. Tyrosine kinase inhibitors for the treatment of rheumatoid arthritis. Curr Top Med Chem. 2013;13:760.

67. Cohen S et al. Co-administration of the JAK inhibitor CP-690,550 and methotrexate is well tolerated in patients with rheumatoid arthritis without need for dose adjustment. Br J Clin Pharmacol. 2010;69(2):143–51.

68. Glant TT, et al. Differentially expressed epigenome modifiers, including aurora kinases A and B, in immune cells in rheumatoid arthritis in humans and mouse models. Arth Rheum. 2013;65(7):1725–35.

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 17 of 18

69. Raychaudhuri S, Remmers EF, Lee AT, et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet. 2008;40(10):1216–23.

70. Weinreb O, Mandel S, Bar-Am O, Yogev-Falach M, Avramovich-Tirosh Y, Amit T, Youdim MB. Multifunctional neuroprotective derivatives of rasagiline as anti-Alzheimer's disease drugs. Neurotherapeutics. 2009;6(1):163–74.

71. Grimaldi LM, Zappalà G, Iemolo F, Castellano AE, Ruggieri S, Bruno G, Paolillo A. A pilot study on the use of interferon beta-1a in early Alzheimer's disease subjects. J Neuroinflammation. 2014;11:30.

72. Woods NK, Padmanabhan J. Neuronal calcium signaling and Alzheimer's disease. Adv Exp Med Biol. 2012;740:1193–217.

73. Li JS, Yao ZX. Modulation of FGF receptor signaling as an intervention and potential therapy for myelin breakdown in Alzheimer's disease. Med Hypotheses. 2013;80(4):341–4.

74. Ling FA, Hui DZ, Ji SM. Protective effect of recombinant human somatotropin on amyloid beta-peptide induced learning and memory deficits in mice. Growth Horm IGF Res. 2007;17(4):336–41.

75. De Deyn PP, Drenth AF, Kremer BP, Oude Voshaar RC, Van Dam D. Aripiprazole in the treatment of Alzheimer's disease. Expert Opin Pharmacother. 2013;14(4):459–74.

76. Tariot PN, Ryan JM, Porsteinsson AP, Loy R, Schneider LS. Pharmacologic therapy for behavioral symptoms of Alzheimer's disease. Clin Geriatr Med. 2001;17(2):359–76.

77. Ansari MA, Abdul HM, Joshi G, Opii WO, Butterfield DA. Protective effect of quercetin in primary neurons against Abeta(1–42): relevance to Alzheimer's disease. J Nutr Biochem. 2009;20(4):269–75.

78. Pallàs M, Verdaguer E, Jordà EG, Jiménez A, Canudas AM, Camins A. Flavopiridol: an antitumor drug with potential application in the treatment of neurodegenerative diseases. Med Hypotheses. 2005;64(1):120–3.

79. Grammas P, Martinez J, Sanchez A, Yin X, Riley J, Gay D, Desobry K, Tripathy D, Luo J, Evola M, Young A. A new paradigm for the treatment of Alzheimer's disease: targeting vascular activation. J Alzheimers Dis. 2014;40(3):619–30.

80. Katz IR, Jeste DV, Mintzer JE, Clyde C, Napolitano J, Brecher M. Comparison of risperidone and placebo for psychosis and behavioral disturbances associated with dementia: a randomized, double-blind trial. Risperidone Stud Group J Clin Psychiatry. 1999;60(2):107–15.

81. Valles SL, Dolz-Gaiton P, Gambini J, Borras C, Lloret A, Pallardo FV, Viña J. Estradiol or genistein prevent Alzheimer's disease-associated inflammation correlating with an increase PPAR gamma expression in cultured astrocytes. Brain Res. 2010;1312:138–44.

82. Dhawan G, Combs CK. Inhibition of Src kinase activity attenuates amyloid associated microgliosis in a murine model of Alzheimer's disease. J Neuroinflammation. 2012;2(9):117.

83. Parkes JD, Tarsy D, Marsden CD, Bovill KT, Phipps JA, Rose P, Asselman P. Amphetamines in the treatment of Parkinson's disease. J Neurol Neurosurg Psychiatry. 1975;38(3):232–7.

84. Bersani G, Grispini A, Marini S, Pasini A, Valducci M, Ciani N. 5-HT2 antagonist ritanserin in neuroleptic-induced parkinsonism: a double-blind comparison with orphenadrine and placebo. Clin Neuropharmacol. 1990; 13(6):500–6.

85. Tariq M, Khan HA, Al Moutaery K, Al DS. Protective effect of quinacrine on striatal dopamine levels in 6-OHDA and MPTP models of Parkinsonism in rodents. Brain Res Bull. 2001;54(1):77–82.

86. Weintraub D, Mavandadi S, Mamikonyan E, Siderowf AD, Duda JE, Hurtig HI, Colcher A, Horn SS, Nazem S, Ten Have TR, Stern MB. Atomoxetine for depression and other neuropsychiatric symptoms in Parkinson disease. Neurology. 2010;75(5):448–55.

87. Hashimoto R, Hough C, Nakazawa T, Yamamoto T, Chuang DM. Lithium protection against glutamate excitotoxicity in rat cerebral cortical neurons: involvement of NMDA receptor inhibition possibly by decreasing NR2B tyrosine phosphorylation. J Neurochem. 2002;80(4):589–97.

88. Perez-Gracia JL, Prior C, Guillén-Grima F, et al. Identification of TNF-alpha and MMP-9 as potential baseline predictive serum markers of sunitinib activity in patients with renal cell carcinoma using a human cytokine array. Br J Cancer. 2009;101(11):1876–83. doi:10.1038/sj.bjc.6605409.

89. Wrasidlo W, Crews LA, Tsigelny IF, Stocking E, Kouznetsova VL, Price D, Paulino A, Gonzales T, Overk CR, Patrick C, Rockenstein E, Masliah E. Neuroprotective effects of the anti-cancer drug sunitinib in models of HIV neurotoxicity suggests potential for the treatment of neurodegenerative disorders. Br J Pharmacol. 2014;171(24):5757–73. doi:10.1111/bph.12875.

90. Sanchez A, Tripathy D, Yin X, Luo J, Martinez JM, and Grammas P. Sunitinib enhances neuronal survival in vitro via NF-κB-mediated signaling and expression of cyclooxygenase-2 and inducible nitric oxide synthase. J Neuroinflammation 2013. 10(93). doi: 10.1186/1742-2094-10-93.

91. Margolis DJ et al. Potential association between the oral tetracycline class of antimicrobials used to treat acne and inflammatory bowel disease. Am J Gastroenterol. 2010;105(12):2610–6.

92. On SC, Zeichner J. Isotretinoin updates. Dermatol Ther. 2013;26(5):377–89.

93. US Natl Inst Health; DailyMed. Current Medication Information for sansert (methysergide maleate) tablet, coated (February 2006). Available from, as of March 6, 2012: http://dailymed.nlm.nih.gov/dailymed/lookup. cfm?setid=3fae28ee-700e-4d4f-a040-02ef01a2aeb4

94. Felder JB, Korelitz BI, Rajapakse R, Schwarz S, Horatagis AP, Gleim G. Effects of nonsteroidal antiinflammatory drugs on inflammatory bowel disease: a case–control study. Am J Gastroenterol. 2000;95(8):1949–54.

95. Takeuchi K, Smale S, Premchand P, Maiden L, Sherwood R, Thjodleifsson B, Bjornsson E, Bjarnason I. Prevalence and mechanism of nonsteroidal anti-inflammatory drug-induced clinical relapse in patients with inflammatory bowel disease. Clin Gastroenterol Hepatol. 2006;4(2):196–202.

96. Ballard C, Margallo-Lana M, Juszczak E, Douglas S, Swann A, Thomas A, O'Brien J, Everratt A, Sadler S, Maddison C, Lee L, Bannister C, Elvish R, Jacoby R. Quetiapine and rivastigmine and cognitive decline in Alzheimer's disease: randomised double blind placebo controlled trial. BMJ. 2005; 330(7496):874.

97. Callaghan RC, Cunningham JK, Sykes J, Kish SJ. Increased risk of Parkinson's disease in individuals hospitalized with conditions related to the use of methamphetamine or other amphetamine-type drugs. Drug Alcohol Depend. 2012;120(1–3):35–40.

98. Brogden RN, Heel RC, Speight TM, Avery GS. Sulindac: a review of its pharmacological properties and therapeutic efficacy in rheumatic diseases. Drugs. 1978;16(2):97–114.

99. Taylor P, Genovese M, Keystone E, Schlichting D, Beattie S, Macias W. Baricitinib, an oral janus kinase inhibitor, in the treatment of rheumatoid arthritis: safety and efficacy in an open-label, long-term extension study1. Ann Rheum Dis. 2014;73:A31. doi:10.1136/annrheumdis-2013-205124.71.

100. van der Heijden JW, Assaraf YG, Gerards AH, Oerlemans R, Lems WF, Scheper RJ, Dijkmans BA, Jansen G. Methotrexate analogues display enhanced inhibition of TNF-α production in whole blood from RA patients. Scand J Rheumatol. 2014;43(1):9–16. doi:10.3109/03009742.2013.797490. Epub 30 Aug 2013.

101. Page TH, Brown A, Timms EM, Foxwell BM, Ray KP. Inhibitors of p38 suppress cytokine production in rheumatoid arthritis synovial membranes: does variable inhibition of interleukin-6 production limit effectiveness in vivo? Arthritis Rheum. 2010;62(11):3221–31. doi:10.1002/art.27631.

102. Weinblatt ME, Keystone EC, Furst DE, Moreland LW, Weisman MH, Birbara CA, Teoh LA, Fischkoff SA, Chartash EK. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. Arthritis Rheum. 2003;48(1):35–45.

103. Moreland LW, Schiff MH, Baumgartner SW, Tindall EA, Fleischmann RM, Bulpitt KJ, Weaver AL, Keystone EC, Furst DE, Mease PJ, Ruderman EM, Horwitz DA,Arkfeld DG, Garrison L, Burge DJ, Blosch CM, Lange ML, McDonnell ND, Weinblatt ME. Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. Ann Intern Med. 1999;130(6):478–86.

104. O'Dell JR, Haire CE, Palmer W, Drymalski W, Wees S, Blakely K, Churchill M, Eckhoff PJ, Weaver A, Doud D, Erikson N, Dietz F, Olson R, Maloley P, Klassen LW, Moore GF. Treatment of early rheumatoid arthritis with minocycline or placebo: results of a randomized, double-blind, placebo-controlled trial. Arthritis Rheum. 1999;40(5):842–8.

105. Klotz U, Maier K, Fischer C, Heinkel K. Therapeutic efficacy of sulfasalazine and its metabolites in patients with ulcerative colitis and Crohn's disease. N Engl J Med. 1980;303(26):1499–502.

106. Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. Lancet. 2007;369(9573):1641–57.

107. Keshavarzian A, Mutlu E, Guzman JP, Forsyth C, Banan A. Phosphodiesterase 4 inhibitors and inflammatory bowel disease: emerging therapies in inflammatory bowel disease. Expert Opin Investig Drugs. 2007;16(9): 1489–506.

108. Mabley JG, Pacher P, Liaudet L, Soriano FG, Haskó G, Marton A, Szabo C, Salzman AL. Inosine reduces inflammation and improves survival in a murine model of colitis. Am J Physiol Gastrointest Liver Physiol. 2003;284(1): G138–44. Epub 28 Aug 2002.

Issa *et al. BMC Bioinformatics* (2016) 17:202

Page 18 of 18

109. El Miedany Y, Youssef S, Ahmed I, El Gaafary M. The gastrointestinal safety and effect on disease activity of etoricoxib, a selective cox-2 inhibitor in inflammatory bowel diseases. Am J Gastroenterol. 2006;101(2):311–7.

110. Carter MJ, Lobo AJ, Travis SP, IBD Section, British Society of Gastroenterology. Guidelines for the management of inflammatory bowel disease in adults. Gut. 2004;53 Suppl 5:V1–V16.

111. Nagar J, Ranade S, Kamath V, Singh S, Karunanithi P, Subramani S, Venkatesh K, Srivastava R, Dudhgaonkar S, Vikramadithyan RK. Therapeutic potential of chloroquine in a murine model of inflammatory bowel disease. Int Immunopharmacol. 2014;21(2):328–35. doi:10.1016/j.intimp.2014.05.005. Epub 21 May 2014.