OXFORD

Briefings in Bioinformatics, 2015, 1–14

doi: 10.1093/bib/bbv048
Paper

Bioinformatics methods in drug repurposing for Alzheimer's disease

John C. Siavelis, Marilena M. Bourdakou, Emmanouil I. Athanasiadis, George M. Spyrou and Konstantina S. Nikita

Corresponding author. George M. Spyrou, Center of Systems Biology, Biomedical Research Foundation, Academy of Athens, Soranou Ephessiou 4, 115 27 Athens, Greece. Tel.: +30 210 6597151; Fax: +30 210 6597505; E-mail: gspyrou@bioacademy.gr

Abstract

Alarming epidemiological features of Alzheimer's disease impose curative treatment rather than symptomatic relief. Drug repurposing, that is reappraisal of a substance's indications against other diseases, offers time, cost and efficiency benefits in drug development, especially when *in silico* techniques are used. In this study, we have used gene signatures, where up- and down-regulated gene lists summarize a cell's gene expression perturbation from a drug or disease. To cope with the inherent biological and computational noise, we used an integrative approach on five disease-related microarray data sets of hippocampal origin with three different methods of evaluating differential gene expression and four drug repurposing tools. We found a list of 27 potential anti-Alzheimer agents that were additionally processed with regard to molecular similarity, pathway/ontology enrichment and network analysis. Protein kinase C, histone deacetylase, glycogen synthase kinase 3 and arginase inhibitors appear consistently in the resultant drug list and may exert their pharmacologic action in an epidermal growth factor receptor-mediated subpathway of Alzheimer's disease.

Key words: Alzheimer's disease; drug repurposing; gene signatures; Connectivity Map; epidermal growth factor receptor pathway

Introduction

The clinical term of dementia describes chronic progressive disorder of cognitive functions that include language, judgement, orientation and predominantly memory. Alzheimer's disease accounts for 60% of dementias, and by 2050, patients are expected to triple because of life expectancy increase [1, 2]. Unfortunately, effective biomarkers and treatments remain elusive. So far, drugs based on amyloid cascade hypothesis, tau protein hyperphosphorylation, brain vascular pathology, neurotransmitter dysfunction, cell cycle dysregulation, inflammation, oxidative stress and mitochondrial impairment have failed to cure dementia [3]. The mainstream of therapy includes Alzheimer-specific drugs—acetylocholinesterase inhibitors and N-methyl-D-aspartate receptor (NMDA) antagonists—that temporarily restrain disease progression and symptomatic treatment that tackles comorbidities such as depression, psychosis and dysautonomia. In view of the upcoming epidemiological burden, drug repositioning might provide an elegant solution to the need for curative treatments.

Drug repositioning/repurposing/reprofiling identifies new indications for existing drugs. Compared with novel drug

John C. Siavelis is an MSc student at the Department of Medical Physics, Faculty of Medicine, University of Patras, Patras 26500, Greece. His research is focused on bioinformatics.
Marilena M. Bourdakou is a PhD candidate student at the Department of Informatics and Telecommunications, University of Athens, 15784 Ilissia, Greece. Her research is focused on bioinformatics.
Emmanouil I. Athanasiadis is a postdoc researcher at the Center of Systems Biology, Biomedical Research Foundation, Academy of Athens, Soranou Ephessiou 4, 115 27 Athens, Greece. His research interests include bioinformatics and bioimaging.
George M. Spyrou is a staff research scientist–professor level at the Center of Systems Biology, Biomedical Research Foundation, Academy of Athens, Soranou Ephessiou 4, 115 27 Athens, Greece. His research interests include bioinformatics and computational intelligence.
Konstantina S. Nikita is a professor at the School of Electrical and Computer Engineering, National Technical University of Athens, 15573 Athens, Greece. Her research interests lie in the field of bioinformatics and computational intelligence.
Submitted: 19 April 2015; **Received (in revised form):** 12 June 2015

discovery, drug repositioning is faster, safer and cheaper [4]. It is implemented *in vivo*, *in vitro* and *in silico* when drug data originate from living organisms, cell cultures and computational modelling, respectively. *In silico* repurposed drugs deploy bioinformatics to perform chemical structure comparisons, molecular docking, gene analyses, network simulations and chemogenomics data processing [5]; the latter relates drugs to their gene targets. In a systems biology extension of chemogenomics, drugs perturb whole cell gene expression levels rather than single genes. Microarray experiments, where samples are divided by drug treatment or disease state, materialize this claim by simultaneously measuring the expression levels of thousands of genes and generating unique lists of up- and down-regulated genes. These so-called gene signatures enable drug–drug, disease–disease and drug–disease comparisons. Connectivity Map (cMap, http://www.broadinstitute.org/cmap/) is the prototypical chemogenomics tool of such an approach [6]. It consists of 6100 drug signatures that stem from five different cell types treated with 1309 bioactive molecules of various concentrations and experiment duration (perturbagens). Users can provide their own signatures in the form of up- and down-regulated genes and compare them with those of the database. cMap returns a connection score stemming from a two-tailed non-parametric Kolmogorov–Smirnov statistical test. Magnitude and sign quantify the gene signatures' correlation, and biological interpretations build on the signature reversion principle: if a drug treatment anti-correlates with a disease signature, then it opposes to a cell's pathologic gene expression pattern, and by extension to the disease phenotype itself. Applications of cMap range from identifying mode of action of unknown substances to exploring new indications of existing drugs and potential novel therapies for diseases [7]. Since its release in 2006, numerous implementations have refined cMap methodology, providing different enrichment estimations and statistical metrics [8, 9], combining network analyses' features [10, 11] or being embedded in integrative drug repurposing pipelines [12]. More importantly, biological pluralism has been conferred by incorporating different cell lines and compounds under the so-called Library of Integrated Cellular Signatures (LINCS) project [13]. Currently in its second phase, a cloud-based platform (http://www.lincscloud.org/) has been made publicly available, along with in-house tools to process the data.

As a currently incurable disease, Alzheimer gene signatures to query such tools have been constructed: Williams G. analysed the disease signature from conserved genes across animal microarray experiments in a cMap-based drug repurposing tool called SPIED [8]; Hajjo *et al.* queried cMap with two signatures from Alzheimer's disease patients as part of their chemocentric approach [14]; Chen *et al.* processed disease gene signatures from six different brain regions in cMap to functionally describe Alzheimer's pathology [15]. A common feature of the above signatures is that they are derived from either single microarray experiments or summarized consensual gene lists. In this study, we followed a gene signature-based approach on five different gene expression experiments of Alzheimer-affected post-mortem human hippocampus. We designed an original formula to merge together outcomes from different implementations of the cMap methodology, diversifying metrics of signature comparison (SPIEDw, correlation coefficient), statistical processing (sscMap, P-value calculation) and biological factors (LINCS-L1000, cell types and compound enrichment). The main assumption is that a compound's consistent anti-correlation with the disease reinforces its pharmacological significance. We then devised a back-loaded pipeline that emphasizes on structural comparisons, a common alternative in drug repurposing, based on an in-house developed tool. To further elucidate the mode of action, we chose to perform pathway/ontology enrichment as well as network analyses. A common feature among the tools used is that they are either web-based or desktop applications that can easily be embodied in a drug repurposing pipeline and are of integrative nature because of the broad set of their underlying high-quality databases. This multi-tool/application processing provides a more robust drug list with less experiment-oriented and more disease-associated mode-of-action interpretations.

Methods

Data

We sought Alzheimer-related microarray studies in Gene Expression Omnibus (GEO)—a transcriptional data repository [16]. Results were filtered to human hippocampus origin and included five GEO Series (GSE) data sets (Table 1). We focused on hippocampal tissue to analyse a severely affected brain region in Alzheimer, responsible for the early-onset memory deficits and thus treatment desirable, as well as to acquire deconvolved gene signatures from other brain tissue expression profiles. Each data set was log2 transformed and quantile normalized. Subsequent analysis was done in R (v.3.1.2) statistical environment and Microsoft Excel v.2010.

Methods to identify differentially expressed gene

Differential gene identification methods can be classified as univariate, multivariate or hybrid, if a data set's expression levels are calculated for each gene individually, collectively or both. A combination of different class-representative methods can highlight distinct aspects of gene expression, leading to multilevel gene signatures. Here, each of the five data sets were processed with such a combination of methods: Limma, a gene expression linear model that calculates a moderated *t*-statistic [22]; Limma-ChDir, a re-ranking of Limma-significant genes by

Table 1. GEO data sets of Alzheimer's disease-affected human hippocampus

No	Author	GEO accession number	Control number	Patient number	Specimen	Pathologic disease stage
1	Liang W.S. *et al.* [17]	GSE5281	13	10	Laser-captured CA1* hippocampal pyramidal cells	Moderate to severe
2	Blalock E.M. *et al.* [18]	GSE28146	8	22	Laser-captured CA1 hippocampal pyramidal cells	Incipient, moderate, severe
3	Hokama M. *et al.* [19]	GSE36980	10	7	Hippocampal grey matter tissue	Moderate to severe
4	Miller J.A. *et al.* [20]	GSE29378	32	31	CA1 and CA3 hippocampal tissue	Moderate to severe
5	Blair L.J. *et al.* [21]	GSE48350	25	19	Hippocampal tissue	Moderate to severe

*Cornu Ammonis.

Characteristic Direction that defines a control-disease separating hyperplane and estimates gene directional cosines [23]; and mAP-KL, a multiple hypothesis tested gene summarization through affinity propagation algorithm based on Krzanowski and Lai clustering index [24]. After data set processing, probe set ids were matched to gene symbols according to the platforms' annotation files. We discarded probe sets that did not have a match and maintained the most differentially expressed one in cases of gene symbol correspondence to multiple probe sets. In Limma, we kept top 500 up-regulated and 500 down-regulated genes based on logFC (fold change) from the gene list with unadjusted P-value ≤ 0.05. Number of genes corresponds to the input length limit of cMap. In Limma-ChDir, we reordered the above gene list according to the directional cosines to acquire gene sets of similar size. Parameter gamma was set to zero (default), and cases of sign discord between Limma and ChDir were rejected. In mAP-KL, we obtained all of the returned genes after setting maxT feature selection size (genes to be considered for clustering) to 1500 and maximum number of clusters to 50. Parameters were arbitrarily chosen after rerun experimentations to ensure both up- and down-regulated genes in the function output.

Drug repurposing tools

Up- and down-regulated genes–collectively termed as disease signatures–queried four drug repurposing tools: cMap (build 2), the prototype of drug gene expression comparisons; SPIEDw, a platform-independent implementation of cMap with novel correlation coefficient-based similarity score [25]; sscMap, a Java application of more robust statistical significance on grounds of a connection strength estimation [26]; and LINCS-L1000, the advanced version of cMap with significantly increased drug treatments, cell types and gene signatures based on L1000 high-throughput technology [12].

For cMap and sscMap, a preceding step was undertaken to convert gene symbols to unique databases' required format of HG-U133A probe sets via NetAffx database [27]. Six common probe sets between up- and down-regulated gene lists were found and removed. Instead of explicit gene expression levels, notation of $+1$ and -1 for up- and down-regulation was used for appropriate SPIEDw and sscMap input format.

Negatively signed drugs consisted of cMap permuted results with P-value < 0.05, SPIEDw entire outcome (results with significance value > 2), sscMap treatment set score normalized to unity with a tolerance of one false connection among all possible drugs ($P < 1/1309$, strictest parameterization) and LINCS-L1000 summarized results with mean connectivity score across the four cell lines in which the perturbagen connected most strongly to the query (best score 4) < -0.9 (proposed threshold for significant results). For each tool, we calculated the mean (μ) and SD (σ) of common sign drug scores that stemmed from at least two different methods.

To preserve the richness of LINCS-L1000 database, further analysis of the results was dichotomized. Drugs that appeared in two or more tools from cMap, sscMap and SPIEDw were scored according to the relation:

$$Score1 = \frac{\mu_{cMap} * \frac{1}{CV_{cMap}} + \mu_{SPIEDw} * \frac{1}{CV_{SPIEDw}} + \mu_{sscMap} * \frac{1}{CV_{sscMap}}}{\frac{1}{CV_{cMap}} + \frac{1}{CV_{SPIEDw}} + \frac{1}{CV_{sscMap}}},$$

where μ and CV are the mean value and the coefficient of variation (σ/μ) of each tool, respectively (we predefined a value of 0.001 for SD $= 0$ cases). Scores of common drugs were summed across data sets, and eight drugs were acquired. In contrast, we collected LINCS-L1000 compounds that appeared in at least two data sets and calculated their mean value (Score 2) to acquire 20 drugs.

The final outcome comprises 27 unique drug names (one drug overlap), of which 11 were further grouped into four pharmacological classes. Significance of its each class over-representation (P-value) was estimated as the probability of getting the same number of drugs in the class if 27 compounds were randomly picked from the LINCS-L1000 database. Class members in the total database were manually retrieved from compound sellers catalogues.

Analysis of the drug list

The final drug list was further processed with:

- ChemBioServer: a web application for searching, filtering and comparing drug structures [28]. We used 2D SDF files of both the resultant drugs and Food and Drug Administration-approved Alzheimer's therapeutics in the tool's hierarchical clustering function. Threshold value of Soergel distance ≤ 0.4 was selected on the grounds of pharmacological action-based guilt-by-association principle; trichostatin-a and vorinostat are both well-studied histone deacetylase (HDAC) inhibitors and were grouped together at this level.
- Enrichr: an integrative gene set enrichment analysis tool, to evaluate drugs' pathways and ontology [29]. For each gene list input, the tool calculates an enrichment score based on a modified Fisher exact test, where binomial probability distribution merges with a z-score that corrects the randomness of the result. Enrichr allows wide access to biological information by analysing transcription factors, pathways, ontology, disease/drugs, cell types and others. Results were ranked according to the combined score > 5 [c $=$ log (adjusted P-value)*|z-score|], which approximately corresponds to Benjamini–Hochberg adjusted P-values < 0.1.
- NetworkAnalyst: a web-based tool for network construction, analysis and visual inspection [30]. NetworkAnalyst has an underlying database of 14 755 proteins and 145 955 human protein–protein interactions, along with in-tool functions to provide extensive network exploration. Top 3 ranked nodes were obtained in Hub Explorer, and modules with P-value < 0.05 were highlighted in Module Explorer applications.
- Mode of Action by NeTwoRk Analysis (Mantra 2.0): a publicly available drug network clustered according to network theory and based on drugs' gene signature similarity metrics [31]. Given a compound's gene signature input, Mantra automatically incorporates it into the network in which the topology can reveal shared intra-cluster features (mode of action, chemical structure, etc).
- Compound Digest App: a lincscloud.org application that estimates drug–drug or drug–gene similarity based on transcriptional responses from treatment, gene knockdown and gene overexpression experiments. The same threshold of LINCS-L1000 drug filtering (best score $4 \geq 90$) was considered significant.

Our approach is summarized in the diagram below (Figure 1)

Results

Drug targets

Five data sets of Alzheimer-affected human hippocampus were collected and further processed via three class-distinct differential expression identification methods, Limma, Limma-ChDir

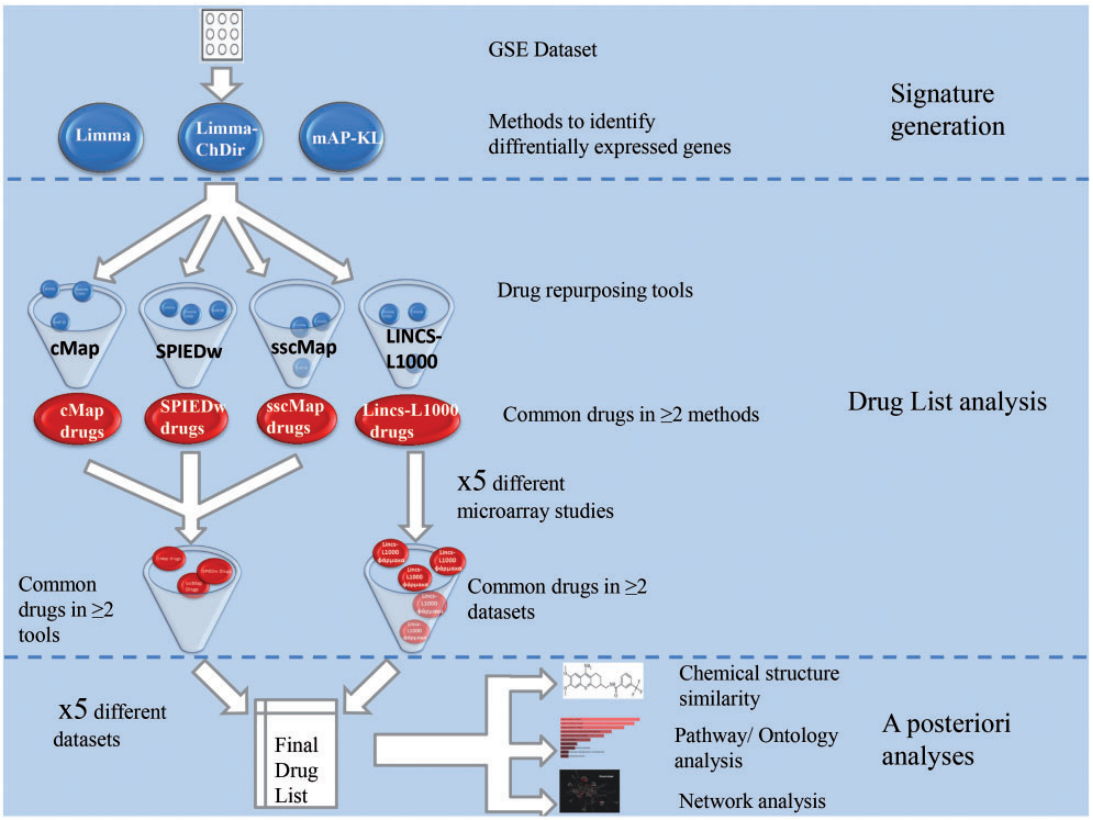


Figure 1. Analysis workflow: initially, each GSE data set is processed with Limma, Limma-ChDir and mAP-KL methods to find differentially expressed genes. The generated lists of up- and down-regulated genes are used as input in four drug repurposing tools, cMap, SPIEDw, sccMap and LINCS-L1000. For each tool, common drugs extracted from gene lists of two or more methods are qualified to the next stage. LINCS-L1000 subsequent analysis is separated from cMap, SPIEDw and sscMap results: the latter are combined to acquire common drugs in more than two tools, and the whole procedure is repeated for all original five data sets, whilst LINCS-L1000 results comprise common drugs across the five data sets. The final drug list—a summation of the above drugs—is further processed with regard to structure similarity, pathway/ontology enrichment and network analysis.

and mAP-KL. The generated gene signatures were used as input to four cMap-based implementations (cMap, SPIEDw, sscMap, LINCS-L1000) to diversify results on account of the novelties each tool conveyed. Derived drugs were combined and scored according to an original formula. The pipeline has resulted in 27 drugs (Table 2). We observe that common anti-Alzheimer agents—tacrine, donepezil, galantamine, memantine and rivastigmine—are absent. Except for rivastigmine that eludes the analysis because it only participates in LINCS-L1000 as treatment instance and not as summarized result, the rest display mean values close to zero and relatively high SDs. Therefore, cMap implementations cannot detect their established therapeutic value.

We subsequently sought mechanistic similarities between the resultant compounds and anti-Alzheimer drugs, namely inhibition of acetylocholisterase or NMDA antagonism. Skimmianine—an alkaloid furoquinoline—inhibits acetylcholinesterase and may act to replenish the cholinergic deficit [32]. Paradoxically, biperidin is also negatively signed, despite being a cholinergic antagonist [33]. Significant drugs of the final list are defined as overlaps of: (i) the separated analyses (6-bromoindirubin-3'-oxime), (ii) the majority (at least three) of the data sets (SB-216763, inhibitor BEC) and (iii) drug classes (HDAC inhibitors; vorinostat, trichostatin-a, HC-toxin, panobinostat, dacinostat and protein kinase C (PKC) inhibitors; bisindolylmaleimide, enzastaurin, PKCbeta-inhibitor, rottlerin). These are further grouped into four over-represented drug classes (Table 3).

Arginase (ARG) occurs in two isoforms (ARG1, ARG2) and converts arginine to ornithine and urea. Ornithine is further metabolized to polyamines (putrescine, spermidine, spermine) that promote cell proliferation. ARG competes with the nitric oxide synthase for their common substrate, arginine; the latter enzyme metabolizes arginine to the vasodilator nitric acid (NO) [34].

Glycogen synthase kinase 3 (GSK3) exists in two isoforms (GSK3α, GSK3β) and is mainly located in the cytoplasm. It participates in the Wnt and insulin pathways, regulating cell proliferation, migration, inflammation, glucose metabolism and apoptosis. In Alzheimer's disease, it is implicated in the hyperphosphorylation of tau protein [35].

HDACs are classified into four classes (I, IIA/IIB, III, IV) and are primarily found in the nucleus. They are associated with epigenetic modification—inherited genetic changes without nucleotide sequence alterations—by preventing the acetylation of histones—DNA structural proteins—and the subsequent binding of transcription factors for gene expression. They are involved in Notch signalling pathways, regulating cell cycle and apoptosis [36].

PKC belongs to the superfamily of protein kinases AGC (PKA, PKG and PKC) and occurs in 15 isoforms. Typically, it is activated by diacylglycerol—after the hydrolysis of the cytoplasmic membrane phospholipids—or calcium ions and phosphorylates other proteins on serine and threonine residues. PKC is a multifunctional protein that mediates numerous cellular functions such as desensitazation of receptors, transcriptional

Table 2. Final drug list

	Data sets	Drug name	Pharmacologic class/action	Score 1	Score 2
cMap, SPIEDw,sscMap	1	–	–	–	
	2	Rottlerin	Experimental/inhibitor of PKCδ	−0.62	
	3	Dapsone	Antimycobacterial/anti-inflammatory	−0.53	
		Biperiden	Antiparkinson/anticholinergic	−0.58	
	4	Skimmianine	Experimental/acetylcholinesterase inhibitor	−0.37	
		Bupropion	Antidepressant/noradrenaline and dopamine reuptake inhibitor	−0.62	
		Milrinone	Cardiac stimulant/inhibitor of phospho-diesterase-3 (PDE3)	−0.66	
	5	Pioglitazone	Antidiabetic/activator of PPAR-γ (peroxisome proliferator-activated receptor gamma)	−0.32	
LINCS-L1000	1, 4	Rucaparib	Experimental/inhibitor of poly-ADP ribose polymerase (PARP)		−0.96
		Nifedipine	Antihypertensive/calcium channel blocker		−0.96
	1, 5	BRD-K05331696	–		−0.94
		SN-38	Irinotecan metabolite/inhibitor of topoisom-erase I		−0.96
		Isoliquiritigenin	Experimental/sirtuin activator		−0.96
	3, 4	Forskolin	Experimental/activator of adenylyl cyclase		−0.91
	3, 5	Vorinostat	Antineoplastic/inhibitor of HDAC		−0.94
		Trichostatin-a	Antifungal/inhibitor of HDAC		−0.95
		HC-toxin	Experimental/inhibitor of HDAC		−0.95
		THM-I-94	–		−0.96
		Panobinostat	Experimental/inhibitor of HDAC		−0.96
		Dacinostat	Experimental/inhibitor of HDAC		−0.98
		KM-00927	–		−0.98
	4, 5	Bisindolylmaleimide	Experimental/inhibitor of PKC		−0.95
		THZ-2-98-01	–		−0.95
		Enzastaurin	Experimental/inhibitor of PKCβ		−0.97
		PKCbeta-inhibitor	Experimental/inhibitor of PKCβ		−0.97
	1, 4, 5	Inhibitor BEC	Experimental/inhibitor of ARG		−0.96
	3, 4, 5	SB-216763	Experimental/inhibitor of GSK3		−0.95
CMap, SPIEDw, sscMap LINCS-L1000	4, 5	6-bromoindirubin-3'-oxime	Experimental/inhibitor of GSK3	−0.88	−0.92

Data set origin, drug name, pharmacologic class/mechanism and the two scores related to cMap, SPIEDw, sscMap and LINCS-L1000 results are displayed.

Table 3. Over-represented drug classes in the final list, drug members and probability of obtaining the same classes by chance (*P*-value)

Drug class	Drug member	*P*-value
PKC inhibitor	Bisindolylmaleimide, enzastaurin, PKCbeta-inhibitor, rottlerin	1.59E-05
ARG inhibitors	Inhibitor BEC	0.00824931
HDAC inhibitors	Vorinostat, trichostatin-a, HC-toxin, panobinostat, dacinostat	9.67E-07
GSK3 inhibitors	GSK-3-inhibitor-IX, SB 216763	0.00223632

regulation, immune response and cell proliferation, as well as memory and learning processes in the brain [37].

Chemical structure similarity

To elucidate potential anti-Alzheimer properties through mechanisms other than the primary drug action, we exploited ChemBioServer to compare the molecular structure with drugs of clinical use. We carried out a hierarchical clustering based on Soergel distance and Ward linkage (Figure 2). The following seven drug groups were produced: (i) dapsone, KM-00927, (ii) donepezil, galantamine, (iii) forskolin, rottlerin, (iv) THM-I-94, dacinostat, panobinostat, (v) HC-toxin, SN-38, (vi) enzastaurin, PKCbeta-inhibitor and (vii) bisindolylmaleimide, SB-216763. We notice that no drug clusters with typical anti-Alzheimer drugs, including skimmianine that shares common anticholinesterase activity. With respect to the other drug similarities, we reason that KM-00927 may possess anti-inflammatory properties,

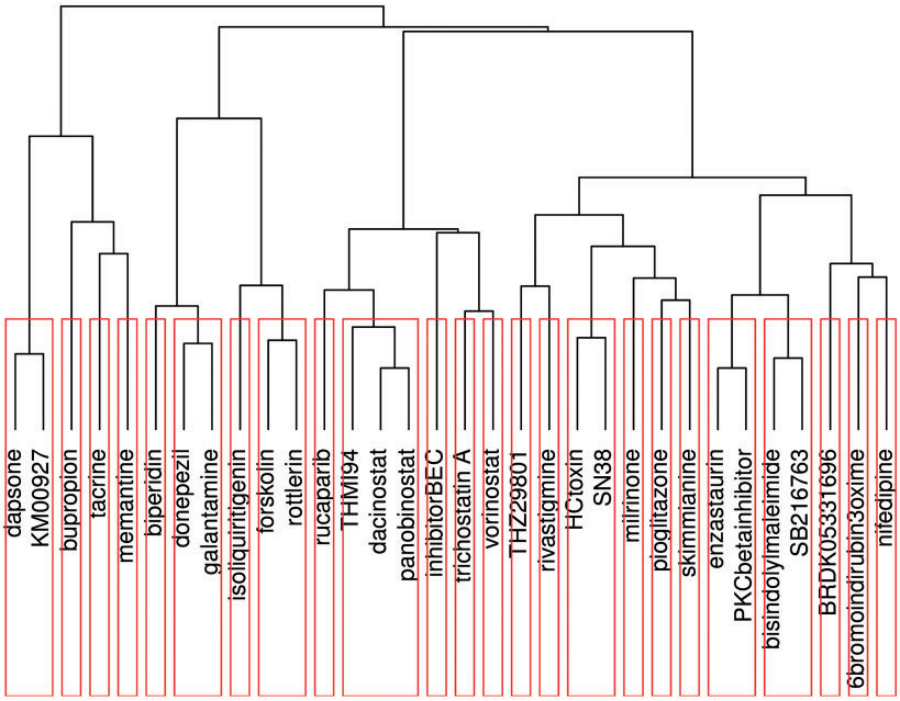


Figure 2. ChemBioServer hierarchical clustering of the final 27 compounds along with marketed anti-Alzheimer drugs (tacrine, rivastigmine, donepezil, galantamine and memantine) based on structure similarity. The different groups in each box are thresholded at Soergel distance value 0.4.

rottlerin could activate adenylyl cyclase, THM-I-94 might act as HDAC inhibitor and bisindolylmaleimide could inhibit GSK3 in addition to PKC, although inverse correlations cannot be excluded.

Pathway–ontology analysis

To assemble a representative gene set of Alzheimer's disease, we obtained the most common differentially expressed genes from the generated gene signatures across the majority of the studies (three or more). Common statistically significant genes may best summarize the disease-related transcriptional perturbation. A list of 154 genes was used as input in Enrichr, a publicly available, multi-database enrichment analysis tool. We limited our search to pathway and ontology analysis: a pathway is a directional graph with distinct relationships between nodes that cooperatively exert a biological effect; gene ontology is constructed on the basis of a controlled vocabulary to describe cellular compartment, molecular function and biological process. We used KEGG [38], WikiPathways [39], Reactome [40] and Biocarta [41] databases for pathway analysis, and all terms of Gene Ontology enrichment, namely Biological Process, Cellular Component and Molecular Function [42]. The following terms were deemed as significant (Table 4)

- Physiological and pathological hypertrophy of the heart (*Homo sapiens*): the term refers to normal heart growth after systematic exercise and pathological hypertrophy in heart failure and other diseases. Growth factors such as EGF and IGF1, cytokines, hormones and mechanic stress activate the pathway, which induces the transcription of hypertrophy genes and synthesis of cell growth proteins.
- Axon guidance: nerve cells communicate via cell membrane protrusions called axons. The mechanism by which an axon is directed to the receptive neuronal cell involves interaction with molecules (netrins, slits, ephrins, semaphorins), cell adhesion proteins (cell adhesion molecules), growth factors, neurotransmitters, etc.
- Depolarization of the presynaptic terminal triggers the opening of calcium channels: in membranes of presynaptic nerve cells, electrical signals (action potentials) stimulate the opening of Ca^{2+} channels to initiate neurotransmitter release in the synaptic cleft.
- Sarcolemma: the cell membrane of a striated muscle fibre cell.
- Cell–substrate adherens junction, cell–substrate junction, adherens junction, anchoring junction, focal adhesion: a collection of terms sharing common genes that mediate cell–cell and cell–extracellular matrix communication.
- Collagen binding: collagen, as the dominant protein of the extracellular matrix, binds to the cell via collagen receptors (CBP) and regulates cell proliferation, migration, adhesion, etc.
- Growth factor binding and transmembrane protein kinase activity: the first term refers to interactions of cells with proteins or polypeptides that promote cell growth or proliferation, while the latter is a general descriptor of signal transmission via catalysis of the reaction: protein + ATP = phosphoprotein + ADP.

Among the enriched pathways, most frequently appeared genes consist of interleukin 6 signal transducer (IL6ST), collagen, type VI, alpha 1 (COL6A1) and epidermal growth factor receptor (EGFR).

Network analysis

To reduce the bias of preprocessed pathway databases, we built our own network using NetworkAnalyst web application. The network was constructed based on the previous 154 genes that acted as seeds to find their direct neighbours (first-order interactors). A trim function was interposed to extract the minimally connected subgraph of the original seed proteins and reduce the bias towards well-studied proteins. The resultant network consists of 357 nodes and 857 edges. For our analysis, we used

Table 4. Enrichr statistically significant pathway/ontology terms

Database	Term	*P*-value	Adjusted *P*-value	z-score	Combined score	Genes
KEGG	–					
WikiPathways	Physiological and pathological hypertrophy of the heart (*Homo sapiens*)	0.000187	0.0305	−2.006	7.000	MAPK11, JUN, LIFR, **IL6ST**
Reactome	Axon guidance	0.000143	0.0512	−2.211	6.571	ENAH, CACNB2, NRP2, CACNB4, **COL6A1**, ITSN1, ARPC4, **IL6ST**, MET, SRGAP1, **EGFR**
	Depolarization of the presynaptic terminal triggers the opening of calcium channels	0.000272	0.0512	−1.877	5.579	CACNB2, CACNB4, CACNG2
BioCarta	–					
GO biological process	–					
GO cellular component	Sarcolemma (GO:0042383)	0.00023	0.0413	−2.150	6.850	DTNA, ANXA1, AHNAK, **COL6A1**, SSPN
	Cell–substrate adherens junction (GO:0005924)	0.00085	0.0554	−2.367	6.849	ENAH, ANXA1, AHNAK, PALLD, TGFB1I1, LUC7L3, SORBS1, **EGFR**, PDLIM7, DDR2
	Cell–substrate junction (GO:0030055)	0.000923	0.0554	−2.364	6.840	ENAH, ANXA1, AHNAK, PALLD, TGFB1I1, LUC7L3, SORBS1, EGFR, PDLIM7, DDR2
	Adherens junction (GO:0005912)	0.002087	0.0834	−2.309	5.737	ENAH, ANXA1, AHNAK, PALLD, TGFB1I1, LUC7L3, SORBS1, EGFR, PDLIM7, DDR2
	Anchoring junction (GO:0070161)	0.002655	0.0834	−2.296	5.703	ENAH, ANXA1, AHNAK, PALLD, TGFB1I1, LUC7L3, SORBS1, EGFR, PDLIM7, DDR2
	Focal adhesion (GO:0005925)	0.00278	0.0834	−2.271	5.642	ENAH, ANXA1, AHNAK, PALLD, TGFB1I1, LUC7L3, EGFR, PDLIM7, DDR2
GO biological process	Collagen binding (GO:0005518)	0.000186	0.0555	−2.267	6.552	ACHE, ECM2, ABI3BP, ANTXR1, DDR2
	Growth factor binding (GO:0019838)	0.000557	0.0625	−2.363	6.551	TGFBR3, NRP2, **COL6A1**, LIFR, **IL6ST**, **EGFR**
	Transmembrane receptor protein kinase activity (GO:0019199)	0.000629	0.0625	−2.242	6.216	TGFBR3, NRP2, MET, EGFR, DDR2

Most frequently found genes across databases are in bold.

the Hub Explorer application to estimate the topological measures of degree and betweenness centrality: number of node connections with other proteins in the network and number of shortest paths passing through the node, respectively. We note c-Jun (JUN), Yes-associated protein 1 (YAP1) and EGFR are ranked at the top of the list if sorted according to betweenness centrality—a measure that highlights a network's 'bottleneck' nodes. Their significance is retained even if another network analysis tool is used based on the same protein–protein interactions called CytoNCA [43] (new JUN, YAP1 and EGFR positions are third, fifth and second, respectively).

- EGFR, a transmembrane glycoprotein with tyrosine kinase activity, belongs to the superfamily of HER/ErbB receptors, which consists of four members: (i) EGFR or HER1/ErbB1, (ii) HER2/ErbB2, (iii) HER3/ErbB3 and (iv) HER4/ErbB4. Structurally, it is divided into three sections: the extracellular portion which is composed of about 620 amino acids, a small hydrophobic transmembrane segment and a cytoplasm with tyrosine kinase activity. When ligands (EGF, TGF, amphiregulin, EPGN, BTC, heparin-binding EGF and others) bind to the EGFR, they induce receptor's homo- or heterodimerization, namely the connection with another EGFR or similar receptor. The new conformation activates the cytoplasmic enzyme activity and initiates a downstream signalling cascade [44].
- Proteins Jun (c-Jun, JunD, JunB) constitute a family of transcription factors that either omodimerize or heterodimerize with ATF (ATF-2, ATF-a) and Fos (c-Fos, FosB, Fra-1, Fra-2) factors to synthesize the Activator protein 1 (AP-1). The activation of AP-1 complex is mediated by the MAPK kinases, namely the c-Jun-NH2-terminal (JNK) kinases. Phosphorylation of c-Jun subunit releases the AP-1 from a wider repressor complex and allows its transition to the nucleus to initiate transcription. Overactivation of the Jun protein promotes cell proliferation, whereas down-regulation mediates apoptosis and cell differentiation [45].
- YAP1 is a transcription co-activator, inducing gene expression via activating transcription factors. Functionally, it is located downstream of the Hippo pathway; the latter includes conserved proteins that inhibit cell proliferation and promote apoptosis during organogenesis. When phosphorylated by specific kinases (Mst, WW45, Lats, Mob), YAP1 moves to the cytoplasm where it is deactivated. In contrast, non-phosphorylated form localizes at the cell nucleus where it is associated with transcription factors (PEBP2a, TEAD) to promote cell growth [46].

We have distinguished EGFR by both pathway/ontology enrichment and network analysis, and we set forth to investigate its role in the network. To elaborate more in the network topology, we also identified subgraphs with more inter-connectivity than randomly expected estimated by random walk-based Walktrap Algorithm of in-tool Module Explorer application. Significant modules displayed ubiquitin C (UBC) and YAP1 as their respective hub proteins (by degree centrality). We also identified Amyloid Precursor protein (APP), the considered culprit of Alzheimer pathology, among the resultant network nodes, despite its absence in the original seed proteins. Path Explorer section of NetworkAnalyst was used to determine shortest paths between APP and EGFR. Signalling

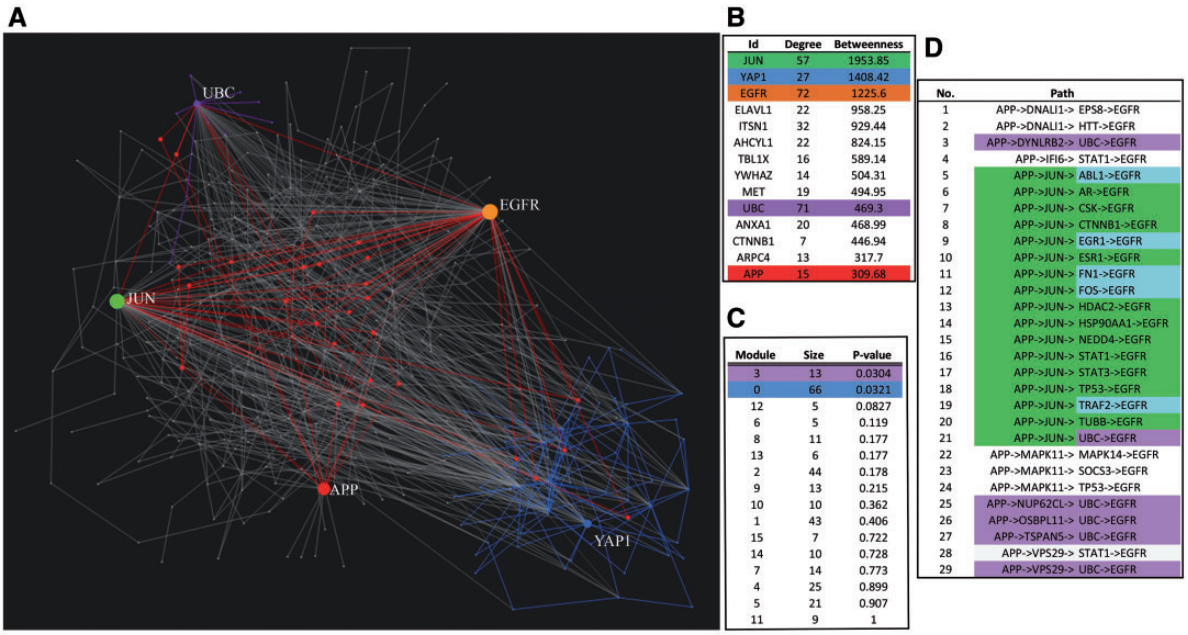


Figure 3. Network Analyst-derived protein–protein interactions network and subsequent in-tool analyses' results. (**A**) Network constructed with the 154 genes shared across the majority of data sets (i.e. at least three) as seed proteins. 'Trim' function was implemented to the first-order neighbours' network to extract a minimally connected subgraph of the original seed proteins. Final network has 357 nodes and 857 edges. (**B**) Sorted nodes according to betweenness centrality measure. Top nodes consist of JUN, YAP1 and EGFR and are highlighted according to the network's colouring. (**C**) Subgraphs with relatively more inter-connectivity than randomly expected (modules) estimated by random walk-based Walktrap Algorithm. Top results according to *P*-value are provided. Hub proteins (by degree centrality) of the modules include UBC and YAP1 as highlighted respectively. (**D**) APP (Amyloid beta A4 protein), a renowned pathogenic protein of Alzheimer's disease, is generated by the network construction despite its absence from the original seed proteins. Occupation of the 15th position in the betweenness centrality ranked list is illustrated. Shortest paths between APP and EGFR (red coloured in the network) indicate that potential signals propagate principally through the JUN node and UBC and YAP1 modules.

flow is mainly interceded by JUN node, as well as the two modules previously defined, proposing an indirect EGFR–APP interplay (Figure 3).

Gene signature-based approach to mode of action identification

While previous analyses have underlined EGFR's significance in Alzheimer gene expression perturbation, we have found no specific inhibitor of the receptor in the final drug list, although such agents are included in the repurposing tools used. To investigate the drugs' mode of action, we exploited Mantra tool's capability to interpolate compounds to its underlying drug network clusters according to gene signature similarity. Drugs that group together are supposed to share enriched features. We have retrieved cluster (community) ids for each of the drugs in the final derived list that are also included in Mantra's database (12 of the total 27 compounds). Results consist of enriched communities no. 16 and no. 32, namely HDAC inhibitors (vorinostat, trichostatin-a, HC toxin) and Topoisomerase (TopoI) inhibitors (skimmianine, nifedipine). Both drug classes modify DNA to confer epigenetic cell changes.

To overcome the limited drug overlap between Mantra and the derived final list, we exploited Compound Digest of the Lincscloud database (http://apps.lincscloud.org/compound_digest), where compound–compound enrichment scores are provided based on the gene signature similarity for 2920 entries (included are 24 of the total 27 final drugs). We accumulated the most significant drug neighbours of each compound and constructed the drug–drug unweighted network in Cytoscape platform [47]. By applying Newman's modularity clustering algorithm, we ended up with five clusters, among which one coincides with a Mantra's cluster (vorinostat, trichostatin-a, HC-toxin). To further elucidate the mode of action of each cluster, we devised a similar approach for drug–gene signature similarities derived from knockdown and mutated overexpression experiments. The underlying assumption is that if a drug correlates with a knockdown/overexpression gene study, then the drug could inhibit/stimulate this gene's expression. Top regulated genes were considered those that appear in the majority of the intra-cluster population (Figure 4).

Injection of prior knowledge to mode of action identification

In the previous analyses, we initially processed Alzheimer data sets to derive 27 drugs that potentially reverse transcriptional changes. Based on the same data sets, we correlated EGFR with Alzheimer's disease in terms of pathway/ontology enrichment and network analysis. We also hypothesized that because we have found no EGFR inhibitor, mode of action is mediated by shared mechanisms in gene signature-based clusters of predefined and originally constructed networks.

We assume that drugs might act in EGFR-mediated downstream signalling cascades. These include Ras/Raf/MAPK, PI3K/Akt, PLCγ, STAT and src kinase activation. The receptor's activity is regulated by Cbl-mediated endocytosis and subsequent degradation in lysosomes [48]. Drug targets of two of the over-represented classes relate to the above path: PKC and GSK3 appear on PLCγ and PI3K/Akt, respectively. For the remaining classes, we turn to Compound Digest App results and document EGFR-related genes from each cluster's top results. A *P*-value

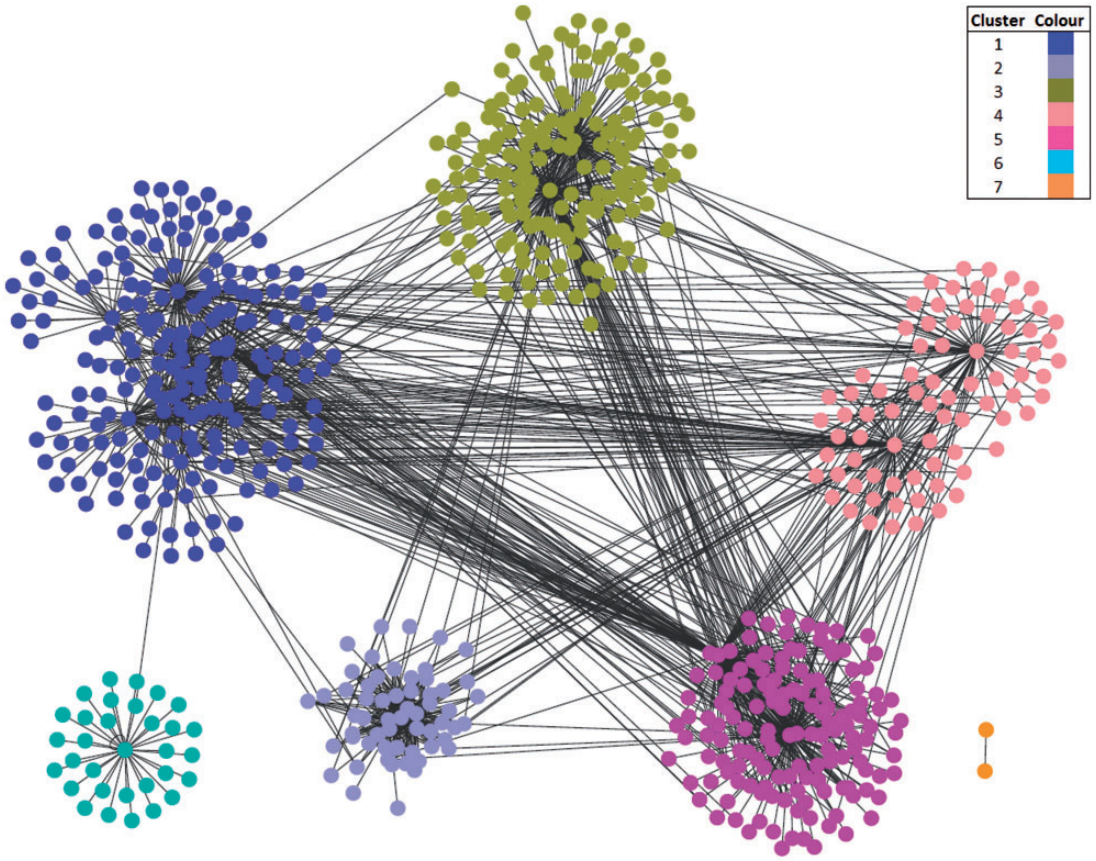


Figure 4. Drug–drug LINCS signature similarity-based network. Nearest neighbour drugs have been accumulated for the 24 (seed drugs) of the 27 final compounds through lincscloud web app, 'Compound Digest'. Best score $4 \geq 90$ has been placed as threshold (mean connectivity score across the four cell lines in which the perturbagen connected most strongly to the query). Resultant network after implementation of Girvan–Newman clustering algorithm is displayed.

calculating the probability of randomly picking EGFR-related genes—as retrieved from KEGG database—from the total gene population of the application is calculated (Table 5). Inhibitor BEC clusters with drugs that down-regulate an isomorph of GSK3B gene (GSK3A, *P*-value = 0.01736), while HDAC inhibitors modulate KRAS, AKT3, JUN, MYC and ERBB3 (*P*-value = 0.00012) (Figure 5). Their action in EGFR-signalling cascade is depicted pictorially in Figure 5.

Discussion

Drug relations with Alzheimer's disease

Our analysis of Alzheimer's disease gene signatures resulted in 27 drugs that are structurally different from drugs in current clinical use, namely acetylcholinesterase inhibitors and NMDA channel blockers. Rather, they seem to target specific EGFR-related proteins. Among the proposed drugs, we notice common substances with other gene signature endeavours based on direct or indirect pharmacologic similarities (milrinone, H-7 and camptothecin [8], troglitazone [13] and vorinostat and trischostatin [14]). Traditional drug repurposing, arising from research in cell lines, animal models and clinical trials, favours dapsone and pioglitazone as well [49]. Dapsone has been tested with contradictory results, while pioglitazone has demonstrated clinical benefit in Alzheimer's disease patients with diabetes comorbidity.

The above findings justify the search for a relationship between EGFR and Alzheimer's disease. In Alzheimer, EGFR is overexpressed in amyloid plaques' neurons, astrocytes and endothelial cells. Aβ oligomers activate the EGFR, and cleavage products of presenilines regulate transcription factors that control the receptor's gene expression [50]. Additionally, inhibition of EGFR by gefitinib and erlotinib prevents memory disorders in Aβ42-overexpression species of Drosophila, although apoptosis is not obviated [51].

To find out the significance of the results, we manually curated available literature in Google Scholar search engine, based on a reversed hypothesis: were these drugs to proceed in preclinical or clinical trials, would their mode of action suppress or promote Alzheimer-disease features? While a number of drugs may act therapeutically, others seem to induce pathology (Table 6). Gene perturbation alone is unable to accurately predict treatment options, although discrepancies pertaining to disease genetic variability and experimental settings (dose, duration, age, stage of disease, etc) should be considered as well. The case of ARG inhibitor recapitulates these concerns: M. J. Kan *et al.* have shown that ARG and ornithine decarboxylase inhibitor, although expected to increase neurotoxic nitric oxide levels, can protect mice from Alzheimer-like pathology [52].

To test the brain region dependence of the results, we applied the same workflow to three data sets from entorhinal, temporal and frontal brain regions, respectively (GSE5281, GSE36980). No significant commonalities among them are observed and, interestingly, these seem to attenuate the further we move from hippocampus (Figure 6). Because Alzheimer's

Table 5. Enriched modules of the clustered drug–drug network

Cluster	Mantra community (neighbour drug)	Seed members	Down-regulated genes (no. times found)	Up-regulated genes (no. times found)	Highlighted genes overrepresentation (P-value)
1	32 (Skimmianine)*	Rucaparib, **nifedipine**, inhibitor BEC, forskolin, enzastaurine, THZ-2-98-01, SB-216763, PKCbeta-inhibitor	**GSK3A** (5)	GPR139 (2)	0.01736842
2	16	**Vorinostat**, **trischostatin-a**, panobinostat, dacinostat, THM-I-94, KM-00927, **HC-toxin**	**KRAS**, PDHX, PSMB2, ERGIC2, TMED10, COPS2, GMPS, TERF1, **AKT3**, PAF1, PSMA1, LRSAM1, RYK, PPAP2B, **JUN**, PPIA, INPP5D, **MYC**, UBC, FAS, EIF2AK3, **ERBB3**, ABAT, BIRC6, GNE, ARF4, PAPOLA, OGG1, ATR, KNDC1 (7)	CDX2, HNF4A, EBF1, SOX2 (7)	0.000119848
3	13	BRD-K05331696, isoliquiritigenin, **bupropion**	PSMB2, PSMD1, PSMA1, KEAP1, PSMA3, TMED10, PSMD2, ZNF785, SULT1A3, **GSK3A**, **KRAS**, SNCA, UBQLN2 (2)	NFE2L2, CTBP1 (2)	0.01922224
4	–	Bisindolylmaleimide, SN-38	RRM1, RPA2, PPIE, SUPT5H, PAF1, ABAT, TYMS, NR2F2, **JUN**, YWHAH, RPA1, ZFX, RAD9A, NOTCH2NL, ATXN3, WT1, TMEM2, PPFIBP2, MYB (2)	–	0.2417124
5	0, 62	**GSK-3-inhibitor-IX**, **rottlerin**	COPA, COPB2, VCP, EIF2S2, EIF2B3, AFG3L2, RXRA, PHB2, HSPA5, TMED10, YME1L1, EIF2B2, CDK6, NR2F2, IARS2, COG4, MTA1, TTK, USF1, ATP5L, BIRC5, SAP18, MB, EIF2B5, XPNPEP1, ENTPD6, CHMP2A, HSPE1, CASC3 (2)	PHB, IFNG, SLC37A4 (2)	–
6	27	**Dapsone**	–	–	–
7	81	**Biperidin**	–	–	–

*Skimmianine is absent from LINCS Compound Digest drugs, but colocalizes with nifedipine in Mantra's Community 32.
The Mantra community, seed members, signature correlated up- and down-regulated genes, and *P*-values of EGFR genes' over-representation out of the total down-regulated gene lists are tabulated. Mantra communities refer to pre-composed drug clusters based on gene signature similarities of cMap (build 2) compounds after various network clustering algorithms were used. Highlighted* in bold* in seed members' column are each LINCS cluster's drugs also discovered in Mantra's network. Down-/up-regulated genes derived from drug–gene knockdown/overexpression experiments' signature similarity, where the same threshold (best score $4 \geq 90$) has been used as before. Top genes according to the number of times found are provided. Highlighted* in bold* are genes that appear in KEGG-curated EGFR pathway. Last column shows *P*-values for these EGFR pathway-related genes' over-representation estimated by a hypergeometric distribution.

disease causes progressive neurodegeneration, eventually affecting almost the whole brain, we can argue that rucaparib, inhibitor BEC and HDAC inhibitors (trichostatin-a, panobinostat, dacinostat) pertain to global disease-modifying capability. On the contrary, GSK3 inhibitors are more hippocampal-specific drugs. It is clear that a greater number of data sets are needed to firmly support the arguments made.

Pipeline limitations

Limitations of our study relate to the Alzheimer's disease features, the bioinformatics tools used and the study's methodology. Alzheimer's pathology implicates almost all cell functions so that significance of highlighted pathways is diluted. Moreover, Alzheimer's disease share common pathologic features with other neurodegenerative diseases and listed drugs may not be disease-specific. As far as gene signatures are concerned, they ignore epigenetic and post-transcription modifications. In the Alzheimer's context, signatures derive from post-mortem specimens and may capture the underlying hypoxia rather than dementia-related perturbations *per se*.

Because we have considered the summarized drug signatures in cMap results, we have, in essence, tested similarities between genes ubiquitously expressed across different cell lines. These genes cannot represent direct effects of neurotransmission modulation—the mode of action of current anti-Alzheimer drugs. Rather, the gene signatures may capture downstream events shared by the majority of cells, particularly cell cycle regulation and differentiation, for they relate to cancer cell types. It may seem that current treatment options are unable to regulate such downstream signals, thus their inability to cure the disease and only provide symptomatic relief.

Drug repurposing tools amplify the biological gap; they use different cells from neurons, batches are both concentration- and time-limited and antineoplastic drugs overpopulate their database. Methodologically, the limited number of data sets and their separate processing may have allowed disease heterogeneity to decorrelate final results. Furthermore, disease signatures of 1000 genes provided small absolute enrichment scores, close to noise values. Regarding the *a posteriori* analyses, we proposed EGFR downstream protein modulation as the sole mode of action of the enriched results. This assumption ignores a

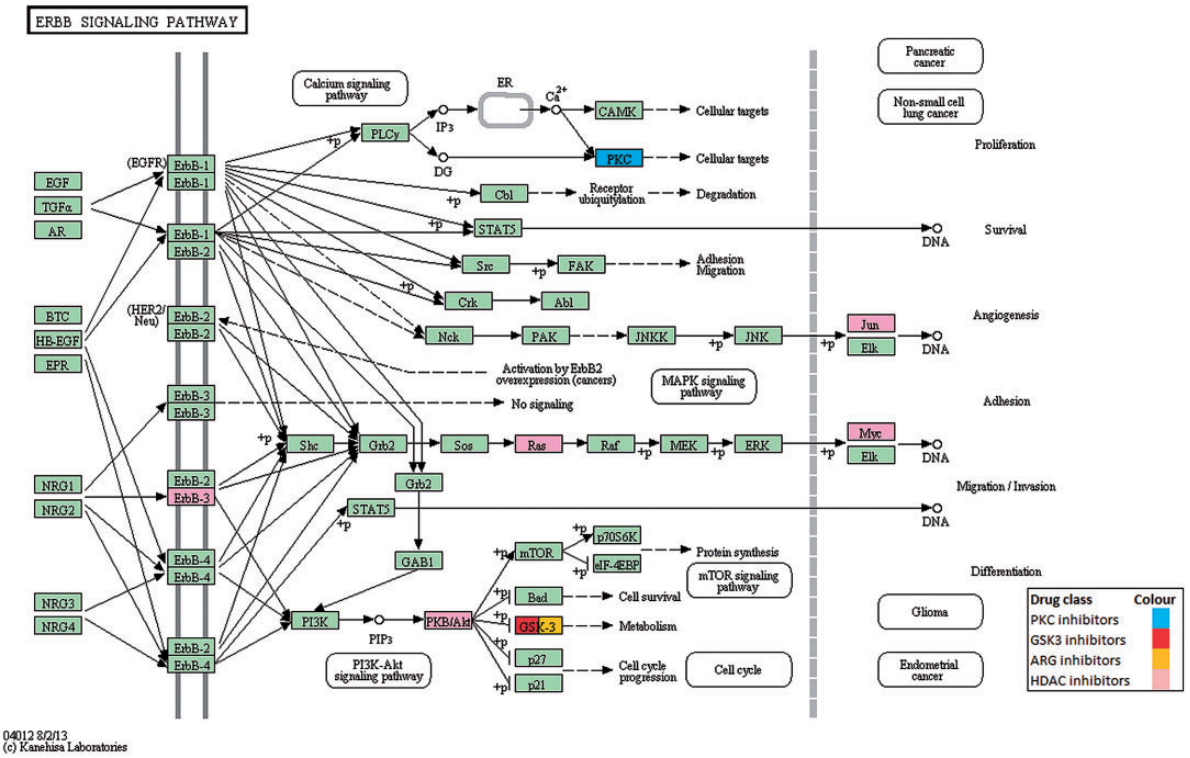
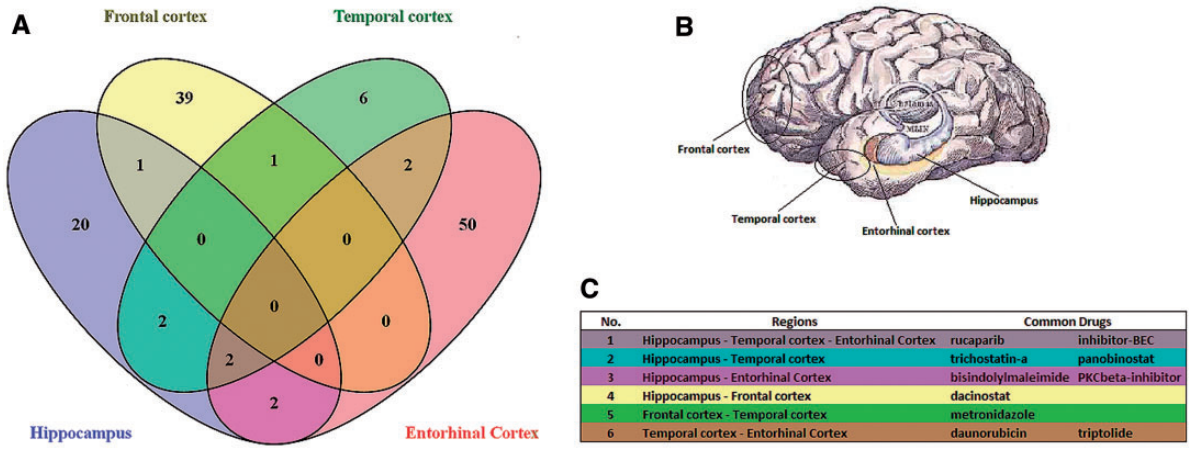


Figure 5. KEGG pathway of Erbb signalling cascade. Highlighted are PKC, HDAC, ARG and GSK3 inhibitors hypothetical targets as proposed by drug-knockdown experiments' gene signature similarities.



No.	Regions	Common Drugs	
1	Hippocampus - Temporal cortex - Entorhinal Cortex	rucaparib	inhibitor-BEC
2	Hippocampus - Temporal cortex	trichostatin-a	panobinostat
3	Hippocampus - Entorhinal Cortex	bisindolylmaleimide	PKCbeta-inhibitor
4	Hippocampus - Frontal cortex	dacinostat	
5	Frontal cortex - Temporal cortex	metronidazole	
6	Temporal cortex - Entorhinal Cortex	daunorubicin	triptolide

Figure 6. Venn diagram of common drugs from four different brain region data sets' analyses. (**A**) Entorhinal, temporal and frontal cortex related data sets (GSE5281 and GSE36980) were processed similarly to the data sets of hippocampal origin. (**B**) Topographic relationship among the brain regions. Entorhinal cortex and hippocampus are medial structures of temporal cortex. (**C**) Region-specific common drugs.

systems biology perspective of pharmacology, a substance's ability to intervene in a variety of pathways. In addition, pathway/ontology enrichment and network analysis are biased in favour of well-studied cellular functions and proteins. Finally, discordances between the computational-predicted and theoretical drug actions against Alzheimer warrant careful reception of the results.

Conclusion

There is a need to discover treatment options for Alzheimer in view of the disease incidence rise in the upcoming years. *In silico* drug repurposing tackles speed, cost and safety issues in drug research. Our method of gene signature comparison tools proposes that PKC, HDAC, ARG and GSK3 inhibitors could reverse Alzheimer-induced gene expression pattern, thus tackling the disease phenotype. The majority of them are linked to EGFR-downstream signalling regulation, which in turn participates in Alzheimer's pathology. We propose that a network simulation of the drug actions could clarify the underlying modes of action. Other drug properties should be evaluated as well, regarding pharmacokinetics, pharmacodynamics, blood–brain barrier permeability and toxicity. These are *sine qua non* factors of drug development in

Table 6. Literature search for drug–Alzheimer relationship

No.	Drug name	Relation to Alzheimer's disease
Potential anti-Alzheimer agents		
1	Dapsone	Reduces amyloid plaques [53]
2	Rottlerin	Protects from Pb-induced accumulation of b-amyloid [54]
3	Skimmianine	Inhibits acetylcholinesterase [32]
4	Bupropion	Increases psychomotor activities in APP23 mice [55]
5	Milrinone	Increases intracellular cGMP and cAMP that enhance synaptic plasticity [56]
6	Pioglitazone	Ameliorates cognitive deficits in 3xTg-AD mice [57]
7	Rucaparib	Inhibits b-amyloid induction of PARP-mediated apoptosis in SH-SY5Y neuroblastoma cells and TgCRND8 transgenic mice [58]
8	Forskolin	In high concentration, it activates PP-2A that reduces PKA-induced tau phosphorylation [59]
9	Isoliquiritigenin	Inhibits amyloid aggregation and 5-lipoxygenase-induced inflammation [60]
10	HDAC inhibitors (vorinostat, trichostatin-a, HC-toxin, panobinostat, dacinostat)	Restore cognitive impairment by enhancing synaptic plasticity [61]
11	GSK3 inhibitors (6-bromoindirubin-3'-oxime, SB-216763)	Facilitate long-term potentiation and neurogenesis via inhibition of apoptosis and tau hyperphosphorylation [62]
Potential Alzheimer inducers		
1	Nifedipine	Long term use impairs cognition [63]
2	Biperiden	Causes cognitive deficits because of cholinergic antagonism [33]
3	SN-38	Drug of same class, camptothecin, induces neuronal apoptosis [64]
4	Inhibitor BEC	Low ARG levels increase neurotoxic NO free radicals [65]
5	PKC inhibitors (bisindolylmaleimide, enzastaurin, PKCbeta-inhibitor)	Reduced PKC levels induce memory impairment, amyloid accumulation, tau hyperphosphorylation and inflammation [66]

Results are divided to drugs that either tackle Alzheimer mechanisms or induce disease features.

Alzheimer, especially when the disease complexity imposes a polypharmacological approach as shown in other neurodegenerative diseases [67].

Key Points

- *In silico* drug repurposing may facilitate the discovery of new agents for so far incurable diseases such as Alzheimer's disease.
- Accumulation of Alzheimer microarray studies in public repositories has enabled a multi-data set processing via diverse methods and tools based on the gene signature concept.
- We defined a drug set with enriched drug targets of PKC, HDAC, ARG and GSK3.
- Molecular similarity, pathway/ontology enrichment, network analysis and gene signature-based drug clustering proposed a relation between EGFR and Alzheimer's disease.
- PKC, HDAC, ARG and GSK3 inhibitors may exert anti-Alzheimer properties in an EGFR-mediated way, although further studies are needed to confirm the hypothesis.

References

1. Weiner MF, Lipton AM. *Clinical Manual of Alzheimer Disease and other Dementias*. Arlington, VA: American Psychiatric Publishing, 2012.
2. Hebert LE, Weuve J, Scherr PA, *et al*. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. *Neurology* 2013;**80**(19):1778–83.
3. Anand R, Gill KD, Mahdi AA. Therapeutics of Alzheimer's disease: past, present and future. *Neuropharmacology* 2014;**76**:27–50.
4. Barratt MJ, Frail DE. *Drug Repositioning: Bringing New Life to Shelved Assets and Existing Drugs*. Hoboken, NJ: John Wiley & Sons, 2012.
5. Liu Z, Fang H, Reagan K, *et al*. *In silico* drug repositioning-what we need to know. *Drug Discov Today* 2013;**18**(3–4):110–15.
6. Lamb J, Crawford ED, Peck D, *et al*. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;**313**:1929–35.
7. Qu XA, Rajpal DK. Applications of Connectivity Map in drug discovery and development. *Drug Discov Today* 2012;**17**(23–24):1289–98.
8. Williams G. A searchable cross-platform gene expression database reveals connections between drug treatments and disease. *BMC Genomics* 2012;**13**(1):12.
9. Zhang SD1, Gant TW. A simple and robust method for connecting small-molecule drugs using gene-expression signatures. *BMC Bioinformatics* 2008;**9**:258.
10. Iorio F, Rittman T, Ge H *et al*. Transcriptional data: a new gateway to drug repositioning? *Drug Discov Today* 2013;**18**(7–8):350–7.
11. Iorio F, Bosotti R, Scacheri E, *et al*. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci USA* 2010;**107**(33):14621–6.
12. Pacini C, Iorio F, Gonçalves E, *et al*. DvD: An R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data. *Bioinformatics* 2013;**29**(1):132–4.
13. Broad Institute. Library of Integrated Cellular Signatures (LINCS). http://www.lincscloud.org/(23 January 2015, date last accessed).
14. Hajjo R, Setola V, Roth BL, *et al*. A chemocentric informatics approach to drug discovery: identification and experimental validation of selective estrogen receptor modulators as ligands of 5-hydroxytryptamine-6 receptors and as potential cognition enhancers. *J Med Chem* 2012;**55**(12):5704–19.

15. Chen F, Guan Q, Nie ZY, *et al*. Gene expression profile and functional analysis of Alzheimer's disease. *Am J Alzheimers Dis Other Demen* 2013;**28**(7):693–701.
16. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;**30**(1):207–10.
17. Liang WS, Reiman EM, Valla J, *et al*. Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. *PNAS* 2008;**105**(11):4441–6.
18. Blalock EM, Buechel HM, Popovic J, *et al*. Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. *J Chem Neuroanat* 2011;**42**(2):118–26.
19. Hokama M, Oka S, Leon J, *et al*. Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama Study. *Cereb Cortex* 2014;**24**(9):2476–88.
20. Miller JA, Woltjer RL, Goodenbour JM, *et al*. Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med* 2013;**5**(5):48.
21. Blair LJ, Nordhues BA, Hill SE, *et al*. Accelerated neurodegeneration through chaperone-mediated oligomerization of tau. *J Clin Invest* 2013;**123**(10):4158–69.
22. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;**3**(1):1–25.
23. Clark NR, Hu KS, Feldmann AS, *et al*. The characteristic direction: a geometrical approach to identify differentially expressed genes. *BMC Bioinformatics* 2014;**15**:79.
24. Sakellariou A, Sanoudou D, Spyrou G. Combining multiple hypothesis testing and affinity propagation clustering leads to accurate, robust and sample size independent classification on gene expression data. *BMC Bioinformatics* 2012;**13**:270.
25. Williams G. SPIEDw: a searchable platform-independent expression database web tool. *BMC Genomics* 2013;**14**(1):765.
26. Zhang S-D, Gant TW. sscMap: an extensible Java application for connecting small-molecule drugs using gene-expression signatures. *BMC Bioinformatics* 2009;**10**:236.
27. Liu G, Loraine AE, Shigeta R, *et al*. NetAffx: affymetrix probesets and annotations. *Nucleic Acids Res* 2003;**31**(1):82–6.
28. Athanasiadis E, Cournia Z, Spyrou G. ChemBioServer: a web-based pipeline for filtering, clustering and visualization of chemical compounds used in drug discovery. *Bioinformatics* 2012;**28**(22):3002–3.
29. Chen EY, Tan CM, Kou Y, *et al*. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013;**14**:128.
30. Xia J, Benner MJ, Hancock REW. NetworkAnalyst—integrative approaches for protein–protein interaction network analysis and visual exploration. *Nucleic Acids Res* 2014;**42**:W167–74.
31. Carrella D, Napolitano F, Rispoli R, *et al*. Mantra 2.0: an online collaborative resource for drug mode of action and repurposing by network analysis. *Bioinformatics* 2014;**30**(12): 1787–8.
32. Yang ZD, Zhang DB, Ren J, *et al*. Skimmianine, a furoquinoline alkaloid from Zanthoxylum nitidum as a potential acetylcholinesterase inhibitor. *Med Chem Res* 2012;**21**:722–5.
33. Szczodry O, van der Staay FJ, Arndt SS. Modelling Alzheimer-like cognitive deficits in rats using biperiden as putative cognition impairer. *Behav Brain Res* 2014;**274**:307–11.
34. Durante W, Johnson FK, Johnson RA. Arginase: a critical regulator of nitric oxide synthesis and vascular function. *Clin Exp Pharmacol Physiol* 2007;**34**(9):906–11.
35. Forde JE, Dale TC. Glycogen synthase kinase 3: A key regulator of cellular fate. *Cell Mol Life Sci* 2007;**64**:1930–44.
36. Smith CL. A shifting paradigm: histone deacetylases and transcriptional activation. *BioEssays* 2008;**30**:15–24.
37. Mellor H, Parker PJ. The extended protein kinase C superfamily. *Biochem J* 1998;**332**(Pt 2):281–92.
38. Kanehisa M, Goto S, Sato Y, *et al*. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 2014;**42**:D199–205.
39. Kelder T, van Iersel MP, Hanspers K, *et al*. WikiPathways: building research communities on biological pathways. *Nucleic Acids Res* 2012;**40**:D1301–7.
40. Milacic M, Haw R, Rothfels K, *et al*. Annotating cancer variants and anti-cancer therapeutics in reactome. *Cancers* 2012;**4**(4):1180–211.
41. Nishimura D. BioCarta. *Biotech Softw Internet Report* 2001;**2**(3):117–20.
42. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet* 2000;**25**(1):25–9.
43. Tang Y, Li M, Wang J, *et al*. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of biological networks. *Biosystems* 2015;**127**:67–72.
44. Jorissen RN, Walker F, Pouliot N, *et al*. Epidermal growth factor receptor: mechanisms of activation and signalling. *Exp Cell Res* 2003;**284**:31–53.
45. Schreck I, Al-Rawi M, Mingot JM, *et al*. C-Jun localizes to the nucleus independent of its phosphorylation by and interaction with JNK and vice versa promotes nuclear accumulation of JNK. *Biochem Biophys Res Commun* 2011;**407**(4):735–40.
46. Zhao B, Lei Q-Y, Guan K-L. The Hippo-YAP pathway: new connections between regulation of organ size and cancer. *Curr Opin Cell Biol* 2008;**20**(6):638–46.
47. Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**(11):2498–504.
48. Chung BM. Endocytic regulation of EGFR signaling. *Interdiscip Bio Cent* 2012;**4**:1–7.
49. Appleby BS, Nacopoulos D, Milano N, *et al*. A review: treatment of Alzheimer's disease discovered in repurposed agents. *Dement Geriatr Cogn Disord* 2013;**35**:1–22.
50. Siddiqui S, Fang M, Ni B, *et al*. Central role of the EGF receptor in neurometabolic aging. *Int J Endocrinol* 2012;**2012**:739428.
51. Wang L, Chiang H-C, Wu W, *et al*. Epidermal growth factor receptor is a preferred target for treating Amyloid-β–induced memory loss. *Proc Natl Acad Sci USA* 2012;**109**(41):16743–48.
52. Kan MJ, Lee JE, Wilson JG *et al*. Arginine deprivation and immune suppression in a mouse model of Alzheimer's disease. *J Neurosci* 2015;**35**(15):5969–82.
53. Walker D, Lue L-F. Anti-inflammatory and immune therapy for Alzheimer's disease: current status and future directions. *Curr Neuropharmacol* 2007;**5**(4):232–43.
54. Behl M, Zhang Y, Shi Y, *et al*. Lead-induced accumulation of β-amyloid in the choroid plexus: role of low density lipoprotein receptor protein-1 and protein kinase C. *Neurotoxicology* 2010;**31**(5):524–32.
55. Neumeister KL, Riepe MW. Bupropion and citalopram in the APP23 mouse model of Alzheimer's disease: a study in a dryland maze. *Int J Alzheimer's Dis* 2012;**2012**:673584.
56. Reneerkens OAH, Rutten K, Steinbusch HWM, *et al*. Selective phosphodiesterase inhibitors: a promising target for cognition enhancement. *Psychopharmacology* 2009;**202**(1–3):419–43.
57. Searcy JL, Phelps JT, Pancani T, *et al*. Long-term pioglitazone treatment improves learning and attenuates pathological

markers in a mouse model of Alzheimer's disease. *J Alzheimers Dis* 2012;**30**(4):943–61.
58. Martire S, Fuso A, Rotili D, *et al*. PARP-1 modulates amyloid beta peptide-induced neuronal damage. Lee MK, ed. *PLoS One* 2013;**8**(9):e72169.
59. Tian Q, Zhang JX, Zhang Y, *et al*. Biphasic effects of forskolin on tau phosphorylation and spatial memory in rats. *J Alzheimers Dis* 2009;**17**(3):631–42.
60. Chen YP, Zhang ZY, Li YP, *et al*. Syntheses and evaluation of novel isoliquiritigenin derivatives as potential dual inhibitors for amyloid-beta aggregation and 5-lipoxygenase. *Eur J Med Chem* 2013;**66**:22–31.
61. Fischer A. Targeting histone-modifications in Alzheimer's disease. *What is the evidence that this is a promising therapeutic avenue?* Neuropharmacology 2014;**80**:95–102.
62. King MK, Pardo M, Cheng Y, *et al*. Glycogen synthase kinase-3 inhibitors: Rescuers of cognitive impairments. *Pharmacol Ther* 2014;**141**(1):1–12.
63. Maxwell CJ, Hogan DB, Ebly EM. Calcium-channel blockers and cognitive function in elderly people: results from the Canadian Study of Health and Aging. *Can Med Assoc J* 1999;**161**(5):501–6.
64. Morris EJ, Geller HM. Induction of neuronal apoptosis by camptothecin, an inhibitor of DNA topoisomerase-I: evidence for cell cycle-independent toxicity. *J Cell Biol* 1996;**134**(3):757–70.
65. Vural H, Sirin B, Yilmaz N, *et al*. The role of arginine-nitric oxide pathway in patients with Alzheimer disease. *Biol Trace Elem Res* 2009;**129**:58–64.
66. Alkon DL, Sun MK, Nelson TJ. PKC signaling deficits: a mechanistic hypothesis for the origins of Alzheimer's disease. *Trends Pharmacol Sci* 2007;**28**(2): 51–60.
67. Boll M-C, Bayliss L, Vargas-Cañas S, *et al*. Clinical and biological changes under treatment with lithium carbonate and valproic acid in sporadic amyotrophic lateral sclerosis. *J Neurol Sci* 2014;**340**(1–2):103–8.