

ARTICLE

Open Access

# Drug repurposing for Alzheimer's disease based on transcriptional profiling of human iPSC-derived cortical neurons

Gareth Williams<sup>1</sup>, Ariana Gatt<sup>1</sup>, Earl Clarke<sup>1</sup>, Jonathan Corcoran<sup>1</sup>, Patrick Doherty<sup>1</sup>, David Chambers<sup>1</sup> and Clive Ballard<sup>2</sup>

## Abstract

Alzheimer's disease is a complex disorder encompassing multiple pathological features with associated genetic and molecular culprits. However, target-based therapeutic strategies have so far proved ineffective. The aim of this study is to develop a methodology harnessing the transcriptional changes associated with Alzheimer's disease to develop a high content quantitative disease phenotype that can be used to repurpose existing drugs. Firstly, the Alzheimer's disease gene expression landscape covering severe disease stage, early pathology progression, cognitive decline and animal models of the disease has been defined and used to select a set of 153 drugs tending to oppose disease-associated changes in the context of immortalised human cancer cell lines. The selected compounds have then been assayed in the more biologically relevant setting of iPSC-derived cortical neuron cultures. It is shown that 51 of the drugs drive expression changes consistently opposite to those seen in Alzheimer's disease. It is hoped that the iPSC profiles will serve as a useful resource for drug repositioning within the context of neurodegenerative disease and potentially aid in generating novel multi-targeted therapeutic strategies.

## Introduction

Global gene expression profiling can be thought of as a high content quantitative phenotypic measure characterising tissue<sup>1</sup>, cell type in, for example, the heterogeneous context of the brain<sup>2–4</sup> and revealing diversity within a previously thought homogeneous population<sup>5</sup>. Further, biological state dynamics can be modelled through temporal patterns of expression<sup>6</sup>. In the therapeutic context, it has been established that disease-associated expression changes can distinguish between disease states and are consistent across independent data sets, thus facilitating the identification of robust biomarkers<sup>7</sup>. Disease-associated gene changes point to modulated pathways and affected cell types, thus providing valuable insights into mechanisms<sup>8</sup>. Interestingly, the quantitative nature of the transcriptional phenotype has allowed for a direct

mapping of disease to potential therapeutic<sup>9–12</sup>. Here the obvious hypothesis is that drugs tending to reverse the expression changes seen in the disease state may act to reverse the biological changes associated with the disease itself. An important caveat here is that some expression changes associated with Alzheimer's disease (AD) may in fact be compensatory and beneficial. Drug repurposing or repositioning has resulted in successful initiatives across several maladies<sup>13–19</sup>. Further, and of more specific interest to the present project, drugs with profiles showing significant anti-correlation to AD gene changes have been shown to be conspicuous for their reported neuroprotective activities<sup>12</sup>. In a recent development, disease-associated gene expression changes have begun to be inferred from genomic risk variant data with the Genotype-Tissue Expression repository<sup>20</sup> and harnessed to predict repurposing candidates for major psychiatric conditions<sup>21</sup>. Although there is some intriguing psychotherapeutic association of the candidate drugs in this approach, the predicted transcriptional perturbation does

Correspondence: Gareth Williams ([gareth.2.williams@kcl.ac.uk](mailto:gareth.2.williams@kcl.ac.uk))

<sup>1</sup>Wolfson Centre for Age-Related Diseases, King's College London, London Bridge, London SE1 1UL, UK

<sup>2</sup>College of Medicine and Health, University of Exeter, Exeter EX1 2LU, UK

© The Author(s) 2019



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

not have an overlap with that seen in diseased brain tissue [G. Williams, unpublished observation]. In the absence of further validation of the predicted gene changes, one must fall back on data from patient samples.

There are no established disease-modifying drugs for the treatment of AD, there have been no new symptomatic treatments licensed for AD for >20 years and the pipeline of emerging therapies is very limited. Target-based drug research in AD has led to many insights into the disease and provided the research community with useful tool compounds. However, the promising results seen in the laboratory have so far failed to be carried over to the clinic and this has led to researchers casting around for novel, non-target-based approaches<sup>22</sup>. The main aim of transcription-based drug discovery is not target discovery, but rather the discovery of drugs that have a disease-modulating effect based on their global transcriptional activity. A particularly attractive aspect of the approach is that it naturally lends itself to repositioning existing drugs thereby bypassing the hurdles that novel entities must overcome on the road to the clinic. AD has been extensively studied in relation to the expression changes following pathological and cognitive decline<sup>23–26</sup>. The wealth of data points to consistent and characteristic changes associated with the disease and thereby makes a repositioning strategy particularly attractive.

The application of gene expression profiling to drug repositioning is limited at present by the fact that full drug profiles are available only on a restricted set of immortalised human cell lines. This data is provided by the Broad Institute connectivity map project (CMAP)<sup>11</sup>. A more extensive drug set has been profiled on a variety of induced pluripotent stem cell (iPSC)-derived cells, including neural stem cells and differentiated cortical neurons. However, this data constituting the LINCS project<sup>27</sup> is based on profiling a set of 1000 landmark genes and then using an optimised linear mapping to generate full profiles. This motivated the present initiative to define the full expression profiles of the CMAP candidate drugs in the more AD relevant cell type of iPSC-derived cortical neurons. The new phenotypes can then be compared to the CMAP profiles and more pertinently scored against the disease profiles to see whether they preserve or enhance their anti-correlation with AD. In this context, iPSC-derived cortical neurons have now been established as a model system for the study of neurological diseases especially the tracing of the effects of disease-related genetic variants<sup>28–31</sup>. This model provides for an efficient moderate throughput platform to assess the transcriptional effects of the candidate drugs in a more neurological context. It must be remembered, however, that AD is a complex pathology also involving multiple cell types, such as microglia and astrocytes. In this context, assaying drug perturbations within isolated

iPSC cultures facilitates an important but limited insight into the disease.

The motivation for the work presented here is to generate a neuronal-specific transcriptional database of compounds with a view to drug repositioning in AD and other neurodegenerative conditions. The initial compound set was assembled based on CMAP profiles that showed a tendency to reverse AD-associated expression changes observed across a variety of independent studies. The drug candidates were then profiled for their transcriptional effects on iPSC-derived human cortical neurons. The results indicate that at the global level there is a degree of correspondence between the CMAP and iPSC profiles. Furthermore, 51 of the drugs have profiles that drive transcription changes counter to those in AD. The consistently regulated genes correspond to those implicated in AD. It is hoped that the transcriptional data for these drugs will be of use to the wider community of researchers interested in neurodegenerative conditions and facilitate further repositioning efforts.

## Materials and methods

### The AD-associated transcriptional landscape

The NCBI GEO database<sup>32</sup> was queried for series containing samples derived from postmortem AD patient brains for various stages of the disease. Similarly, murine AD model brain samples were also collected based on relevant query key words: 5xFAD, 3xTG, Alzheimer's disease+mouse. Profiles were generated based on relative levels of non-disease and disease state sample averages, with the scaled fold level defined as  $f = \frac{(d)-(c)}{(d)+(c)}$ , where the brackets indicate averages of the control (*c*) and disease (*d*) samples. The statistical significance is measured by Student's *t* test and those folds falling below the 95% confidence interval were dropped as were those with folds of <20%. The human disease versus control AD set comprises 21 profiles derived from 13 series (NCBI GEO accession: GSE84422<sup>24</sup>, GSE37263<sup>33</sup>, GSE36980<sup>34</sup>, GSE39420<sup>35</sup>, GSE1297<sup>23</sup>, GSE29378<sup>36</sup>, GSE48350<sup>37</sup>, GSE15222<sup>25</sup>, GSE26972<sup>38</sup>, GSE37264<sup>39</sup>, GSE28146<sup>40</sup>, GSE5281<sup>41</sup>, GSE13214<sup>42</sup>) showing intra-profile consistency based on the regression scores for significant (Student's *t* test  $p < 0.05$ ) correlations, see Supplementary Table 1. To capture brain region variability, the number of profiles is greater than the number of series. In Supplementary Table 2, the extent of intra- versus inter-series AD profile correlation scores are given showing that in many cases the variability in brain region profiles is greater than that between independent series. Cognitive decline was based on decline in Mini-Mental State Examination (MMSE)<sup>43</sup> represented by two profiles from two independent series and Clinical Dementia Rating (CDR)<sup>44</sup> profiles from one series. Similarly, series

corresponding to murine models of AD were gathered from 5xFAD and 3xTG mice resulting in seven profiles from three series (NCBI GEO accession: GSE50521<sup>45</sup>, GSE119756<sup>46</sup>, GSE101144<sup>47</sup>, GSE77574<sup>48</sup>) for the 5xFAD set and nine profiles from eight series (NCBI GEO accession: GSE31624, GSE15128<sup>49</sup>, GSE36237, GSE92926<sup>50</sup>, GSE60460, GSE60911<sup>51</sup>, GSE36981<sup>34</sup>, GSE35210) for the 3xTG set. Series corresponding to BRAAK stage progression (NCBI GEO accession: GSE1297, GSE84422, GSE48350, GSE106241<sup>52</sup>) were generated with a linear mixed model analysis, by fitting the gene expression level across the samples in the series to a linear function of the BRAAK stage with categorical calls on cell type and gender as covariates. The resulting residual correlation *Z* score for gene expression against BRAAK stage constituted the BRAAK profile. Profiles corresponding to full BRAAK progression were not considered to be sufficiently different to the overt disease profiles derived from the same series, where disease assignment is also based on BRAAK staging. However, gene expression changes driving mild BRAAK pathology should capture early disease biology invisible in the overt profiles. In total, six profiles corresponding to mild BRAAK pathology, level 0 to level 3, formed the mild BRAAK AD set. Similar profiles were generated for psychiatric measures MMSE and CDR (NCBI GEO accession: GSE48350, GSE1297, GSE84422). In the case of the MMSE profile, the regression signs were reversed as MMSE scores decrease with disease progression, see Table 1 for an overall comparison of the profile sets.

Representative profiles for each set were based on genes showing consistent changes across the constituent profiles. In particular, the sense changes (upregulation and downregulation calls) for significantly regulated genes were summed over the profiles and only those genes retained that had an absolute regulation fraction of >20% and with a significant regulation statistic measured by Student's *t* test of  $p < 0.05$ . Owing to the categorical nature of the representative profiles, correlation with the iPSC profiles was based on an enrichment analysis. The

enrichment score was generated based on a binomial probability sum with gene probabilities scaled according to their frequencies in SPIED<sup>53</sup>.

#### CMAP profiles

CMAP data were downloaded from the Broad connectivity map site ([www.broadinstitute.org/connectivity-map-cmap](http://www.broadinstitute.org/connectivity-map-cmap))<sup>11</sup>. This consisted of probe sets for each sample ranked according to expression level relative to batch control. The data consist of 6100 samples covering 1260 drugs and 4 cell types. The relative probe expression ranks, defined as  $1 - 2 \frac{R - R_{\min}}{R_{\max} - R_{\min}}$ , where *R* is the rank of a given gene's expression change ( $R_{\max}$  being the highest and  $R_{\min}$  being the lowest ranks), were averaged over replicates ignoring cell type and filtered based on significance using a one-sample Student's *t* test. For genes with multiple probes, the probe with the largest significant change was mapped to the gene. This resulted in a unique profile for each drug in CMAP. The compound data can be queried through SPIED<sup>53</sup>.

#### iPSC profiles

Following the dominant CMAP treatment protocol, cell cultures were treated for 6 h and at compound concentrations of 10  $\mu$ M. The iPSC expression samples were generated on the Affymetrix Human Genome U133 Plus 2.0 Array platform from ThermoFisher Scientific.

Human iPSC-derived cerebral cortical neurons (HyCCNs; Ax0026) were cultured as per the manufacturer's guidelines ([www.axolbio.com/page/neural-stem-cells-cerebral-cortex](http://www.axolbio.com/page/neural-stem-cells-cerebral-cortex)). Each drug treatment at a concentration of 10  $\mu$ M for 6 h was performed on 3 independent HyCCN cultures (average density 300 K/cm<sup>2</sup>) and RNA from each treated well extracted by direct cell lysis and recovery using the Absolutely RNA Microprep Kit (Agilent, as per the manufacturer's guidelines). Each drug-treated plate also consisted of a vehicle-only control set of triplicate cultures. Integrity of total RNAs was determined using an Agilent Bioanalyser as per the manufacturer's instructions and only samples with RNA

**Table 1** The AD sets show varying degrees of overlap

	AD	BRAAKmild	COGI	5xFAD	3xTG
AD	11.26 $\pm$ 0.45	-1.81 $\pm$ 0.38	13.34 $\pm$ 1.08	3.38 $\pm$ 0.33	0.05 $\pm$ 0.11
BRAAKmild		4.43 $\pm$ 1.00	-1.03 $\pm$ 0.83	-0.35 $\pm$ 0.34	-0.04 $\pm$ 0.20
COGI			15.10 $\pm$ 3.47	3.09 $\pm$ 0.70	0.26 $\pm$ 0.23
5xFAD				13.23 $\pm$ 1.67	0.46 $\pm$ 0.21
3xTG					-0.06 $\pm$ 0.25

The overt AD profile set is highly correlated with the cognitive decline profiles. There is a degree of overlap with the 5xFAD profiles but poor agreement with the mild BRAAK and 3xTG animal profiles. The 3xTG profile set is conspicuous for not being internally consistent or having significant overlap with the other AD sets. The numbers in the table correspond to the average *Z* score across pairs in the sets, excluding correlations of profiles with themselves

integrity number >7 were progressed to transcriptome analysis. Transcriptome changes driven by exposure to the candidate drugs were determined using the Nugen Ovation V2 labelling system (<https://www.nugen.com/products>) followed by Human U133 Plus 2 GeneChips as per the manufacturer's instructions ([www.thermofisher.com/order/catalog/product/900466](http://www.thermofisher.com/order/catalog/product/900466)).

The NCBI GEO hosts 145,000 samples on this platform, making it the most popular array chip. The relative expression levels of probes were collected for the GEO data and the iPSC control data. The ranks were scaled to lie between zero for the highest expression probe and unity for the lowest. The relative rank of each probe was defined as  $\frac{r_0-r}{r_0}$  for  $r < r_0$  and  $\frac{r_0-r}{1-r_0}$  for  $r > r_0$ , where  $r$  and  $r_0$  are the average probe ranks over the iPSC samples and the set of samples deposited on GEO, respectively. Probes were then mapped to genes and, in the case of degeneracy, the probe with the largest relative rank mapping to the gene. The gene rank profile was taken to be related to the relative gene expression characterising iPSCs.

Drug treatment profiles were based on statistically filtered ratios of drug-treated and control groups. These were generated based on a combined set of 554 samples, which were robust multiarray averaging normalised. The samples were distributed over 23 plates with the corresponding dimethyl sulfoxide controls. Transcriptional profiles for the 153 drugs were generated based on normalising to the plate control and multiple plate drug replicates kept as separate profiles. The drug set is enriched for CMAP based anti-AD potential (153). Rapamycin, which has a well-defined transcriptional signature, served as a positive control. The expression changes were either measured as scaled folds filtered for significance with Student's  $t$  test or as  $Z$  scores, with significance based on the magnitude of  $Z$ . Degenerate probes were mapped to genes based on the dominant probe responses.

## Results

### AD-associated expression changes

To capture as much as possible of the transcriptional landscape of AD, different categories were defined based on overt disease versus healthy profiles, profiles following early pathological and cognitive measures, together with those from animal models, as described in 'Materials and methods'. There is a good degree of overlap between the overt AD profiles and those following cognitive decline, see Table 1, but it was reasoned that there is sufficient variability to give rise to unique drug candidates, see section on 'CMAP candidates'. The early BRAAK stage profiles show little overlap with overt or cognitive decline profiles, see Table 1, and thus it is anticipated that these profiles may shed light on distinct early stage pathology and early therapeutic intervention. The animal model data naturally separates into those based on the 5xFAD, which

is consistent with AD as can be seen in Supplementary Table 3, and those based on 3xTG, showing little overlap with AD profiles or internal consistency. A similar analysis also including rat models of AD has been carried out by Hargis and Blalock<sup>54</sup>. Animal model data were included in this study because the expression changes seen in the model systems have established causes, i.e. the inserted mutations, 5xFAD or 3xTG in our case. Consequently, candidate drugs reversing these changes may have more focused mechanisms of action. Furthermore, the evidence for neuroprotection is to a large extent derived from experiments in animal models.

### CMAP candidates

In general, transcription-based repositioning results in tens of candidates out of a total of just over a thousand drugs constituting CMAP<sup>13–19</sup>. The relatively small number of compounds that are put forward for rigorous bio-assaying to establish firmer evidence for a disease-modulating potential of course reflects the experimental resource required. The basis of the present project was to select candidates to populate a database of iPSC profiles for drugs biased towards their predicted anti-AD and wider neuroprotective activities. It was therefore reasoned that the thresholds for deeming a drug a repositioning candidate had to be relaxed to allow for over a hundred candidates to be taken forward. To this end, five AD-based profile sets that capture distinct aspects of the disease were separately queried against CMAP and three selection criteria were applied. In the first instance, data were gathered on the anti-correlation rank of each compound, with compounds showing a high rank in either of the profiles considered as candidates, see Supplementary Table 4 for the complete candidate list. A second selection was based on consistency of the anti-correlation across profiles in each set, and finally some compounds with conspicuously high anti-correlations with individual profiles were added to the set. The full list of compounds is given in Supplementary Table 4 and consists of 153 compounds. Interestingly, among these drugs are established neuroprotective entities and AD therapeutics, see below.

### iPSC profiles

As a first step in establishing the phenotype of the model cell system, the overall iPSC transcriptional profile was queried against a database of publicly deposited gene expression profiles via SPIED<sup>12,53</sup>, see 'Materials and methods'. The top 1000 genes in the iPSC rank profile consists of 959 upregulated and 41 downregulated genes and this served as a query in the SPIED search. It is perhaps worth pointing out here that the level of gene expression unique to a given cell type will tend to be elevated relative to a background consisting of a variety of

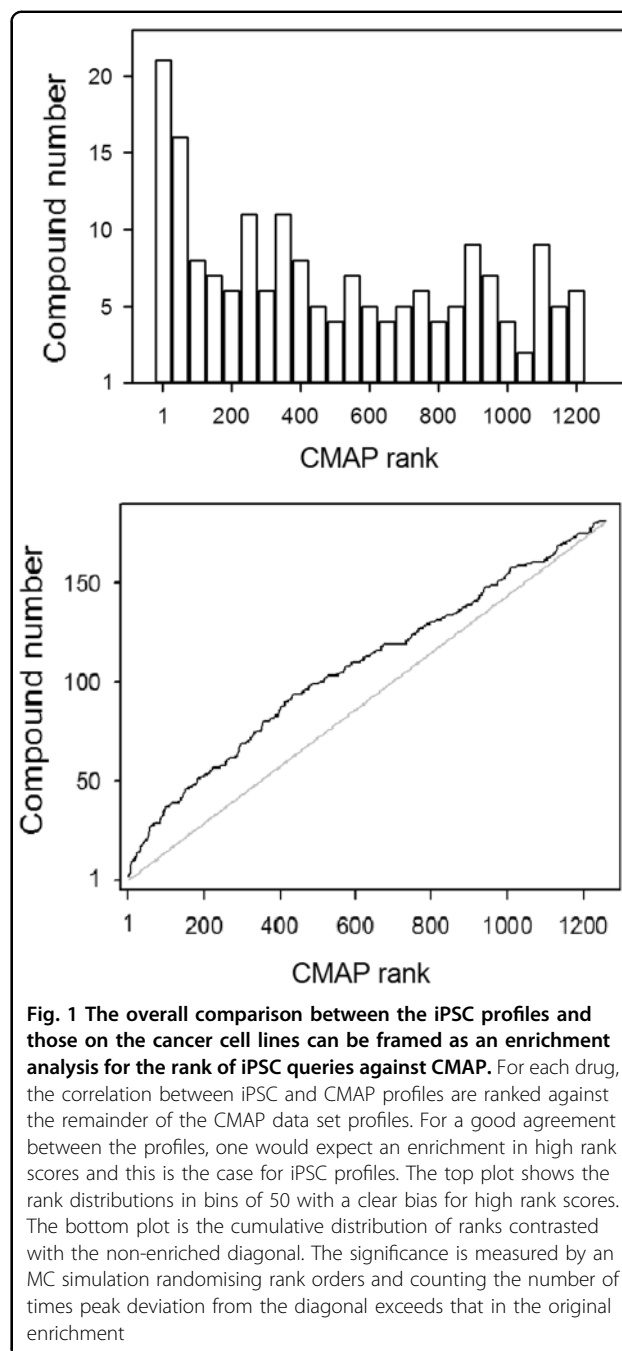
tissue types. An analogy would be in the context of division of labour one is characterised by what one does not by what one does not do. The top SPIED hits show a high correlation with human brain-derived samples, validating the cell's lineage, see Supplementary Table 5.

### Comparison of iPSC and CMAP profiles

The extent to which an iPSC profile correlates with its CMAP equivalent can be assessed by querying the CMAP database with the iPSC profile and ranking the CMAP equivalent. The extensively studied perturbagen rapamycin served as a positive control and eight independent profiles were generated to assess the degree to which these profiles are consistent with each other and with the rapamycin profile in CMAP. The rapamycin profiles had consistently high overlaps among themselves, but less so with the CMAP profile, with only one returning rapamycin as a top hit, rank seven, in a CMAP query, see Supplementary Fig. 1. In Supplementary Fig. 2, iPSC and CMAP profile pairs with the four highest CMAP query ranks are shown. Overall, there are 30 significantly correlating and 8 anti-correlating pairs. The overall comparison of the iPSC and CMAP profiles can be framed in terms of an enrichment analysis for the rank of the equivalent compound hit and the significance can be assessed with Kolmogorov–Smirnov (KS) statistic on the maximal deviation from the zero-enrichment diagonal line. The KS statistic furnishes an objective measure of the robustness of the iPSC profiles and suggest that iPSC profiles based on a  $Z$  score threshold of  $|Z| > 3$ , see ‘Materials and methods’ for details, capture most of the compound-associated changes. The enrichment is that of the rank of a given iPSC compound score with itself in CMAP. The enrichment plot is shown in Fig. 1. The KS statistic is highly significant with the chance of a random compound association beating the enrichment maximum of  $p = 5.1E-6$ .

### Relation of iPSC profiles to AD

Further to assessing the extent to which compounds orchestrate similar expression changes in the cancer cell lines and differentiated cortical neurons, it is critical to test whether the drugs also act in an anti-AD manner in the neuronal context. To this end, the drug profiles were scored against five representative AD reprofiles derived from the AD sets defined above, see ‘Materials and methods’ for details. Table 2 lists the compounds with at least two significant anti-correlations with the representative AD profiles, which will be referred to as AD hit compounds (ADC). The ADC set show a relatively high degree of intra-profile correlation as compared to other iPSC profile pairs, see Fig. 2. The average correlations in terms of regression  $Z$  scores are: 2.43 for ADC pairs and 0.77 for all other pairs. It is therefore of interest to see to

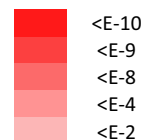


**Fig. 1** The overall comparison between the iPSC profiles and those on the cancer cell lines can be framed as an enrichment analysis for the rank of iPSC queries against CMAP. For each drug, the correlation between iPSC and CMAP profiles are ranked against the remainder of the CMAP data set profiles. For a good agreement between the profiles, one would expect an enrichment in high rank scores and this is the case for iPSC profiles. The top plot shows the rank distributions in bins of 50 with a clear bias for high rank scores. The bottom plot is the cumulative distribution of ranks contrasted with the non-enriched diagonal. The significance is measured by an MC simulation randomising rank orders and counting the number of times peak deviation from the diagonal exceeds that in the original enrichment

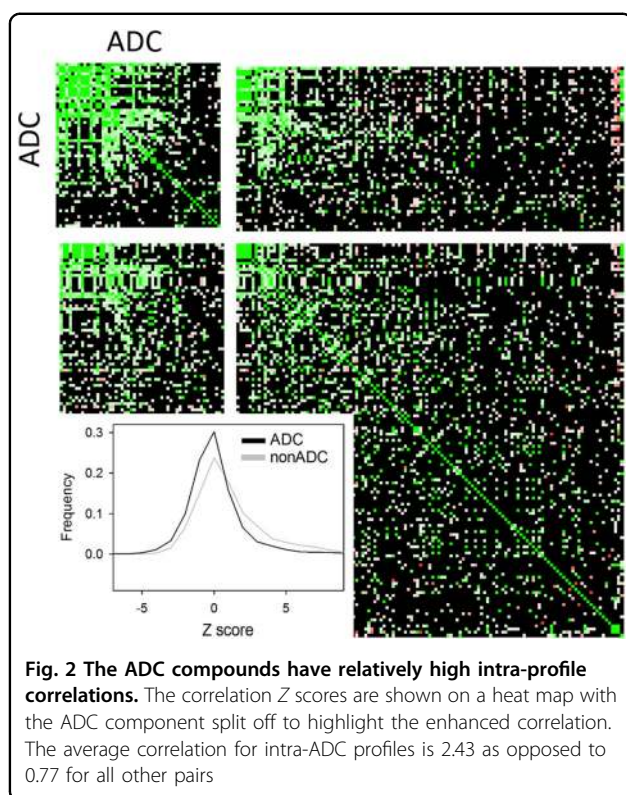
what extent the ADC set regulate a common set of transcripts. In Fig. 3, the common ADC target genes are shown demonstrating a high degree of consistency with a clearly defined set of upregulated and downregulated gene cohort. To get an idea of the underlying biological networks that are being perturbed by the ADC, a pathway enrichment analysis was performed on each of the profiles in the ADC set. The consistently positively and negatively regulated pathways defined by an enrichment in the upregulated and downregulated gene sets, respectively,

**Table 2 Compounds with iPSC profiles showing anti-correlation with at least two representative AD profiles, referred to as the ADC set**

	AD	BRAAKmild	COGI	5xFAD	3xTG	DESCRIPTION
TRIOXSALEN	-0.22	-0.43	-0.37		-0.33	furanocoumarin and a psoralen derivative
ALLANTOIN		-0.64	-0.35		-1	urea hydantoin used in dermatological preparations
FLUMETHASONE		-0.33	-0.33	-0.33		corticosteroid for topical use
GALANTHAMINE		-0.5	-0.37		-0.33	cholinesterase inhibitor
OXAPROZIN	-0.34		-0.2		-0.2	NSAID
RISPERIDONE	-0.29	-0.38	-0.54			antipsychotic
SULFAMONOMETHOXINE	-0.36	-0.6	-0.33			anti-myocardial
(cis-) NANOPHINE	-0.21	-0.27				angiogenesis inhibitor
ACETYLSALICYLSALICYLIC ACID		-0.33	-0.38			Aspirin impurity
ALBENDAZOLE		-0.32			-0.6	anthelmintic
AMINOPURINE		-0.39	-0.24			purine analog of guanine and adenine
ATRACTYLOSIDE		-0.29	-0.21			toxic glycoside
BENZTHIAZIDE	-0.28	-0.71				anti-hypertensive
BEPHENIUM		-0.24	-0.29			anthelmintic
BROMOPRIDE		-0.29	-0.25			dopamine antagonist
CEPHALOTHIN		-0.57	-0.23			cephalosporin antibiotic
CHLORPROMAZINE		-0.46			-0.33	antipsychotic
CYCLOPENTHIAZIDE		-0.25	-0.43			anti-hypertensive
DILAZEP		-0.21	-0.41			adenosine reuptake inhibitor
DIPYRIDAMOLE			-0.33	-0.45		inhibits blood clot formation
DOXORUBICIN			-0.36		-0.5	chemotherapeutic
DROPROPIZINE		-0.59	-0.2			cough suppressant
ERGOCALCIFEROL		-0.25	-0.33			vitamin D
ESTRIOL	-0.5		-0.38			steroid, weak estrogen
ETOMIDATE	-0.35		-0.33			anaesthetic
FENBUFEN		-0.38	-0.22			NSAID
FLUOCINONIDE	-0.51		-0.53			Glucocorticoid
GLAFENINE		-0.38	-0.26			NSAID
HYDROFLUMETHIAZIDE	-0.39		-0.44			oral thiazide used to treat hypertension and edema
IPRONIAZIDE		-0.51	-0.33			Monoamine oxidase inhibitor
KAWAIN	-0.22	-0.25				sedative, anxiolytic, psychotropic
LEVAMISOLE		-0.54	-0.47			used to treat parasitic worm infections
MESTRANOL		-0.24	-0.21			estrogen receptor agonist
MONOBENZONE		-0.29	-0.36			depigmentation drug
N6-METHYLADENOSINE		-0.39	-0.3			methylated RNA
NEOSTIGMINE	-0.21		-0.31			cholinesterase inhibitor
NETILMICIN		-0.78	-0.5			aminoglycoside antibiotic
OFLOXACIN		-0.33	-0.41			antibiotic
OXANTEL		-0.57		-0.41		anthelmintic
OXOLINIC ACID		-0.5	-0.33			antibiotic
PHENACETIN			-0.41	-0.57		pain-relieving and fever-reducing drug
PHENINDIONE		-0.28	-0.79			Vitamin K antagonist
PHENOXYBENZAMINE		-0.31	-0.22			anti-hypertensive NP
PROBENECID		-0.25	-0.29			increases uric acid excretion in the urine
PROPARACAINE	-0.3				-0.5	topical anesthetic
PYRANTEL		-0.52			-0.2	anthelmintic
RONIDAZOLE		-0.5	-0.29			nitroimidazole antibiotic
TELENZEPINE		-0.35	-0.26			antimuscarinic
THIOGUANOSINE		-0.3	-0.25			chemotherapeutic
TRICHLORMETHIAZIDE		-0.36	-0.27			diuretic
XAMOTEROL		-0.4		-0.33		$\beta$ 1 adrenergic receptor agonist



The numbers are the correlation  $\frac{n|+n|-n|+n|}{n|+n|+n|+n|}$  and the associated binomial enrichment score is reflected in the red intensity. The compound descriptions are given and those with reported neuroprotective activity are highlighted in grey



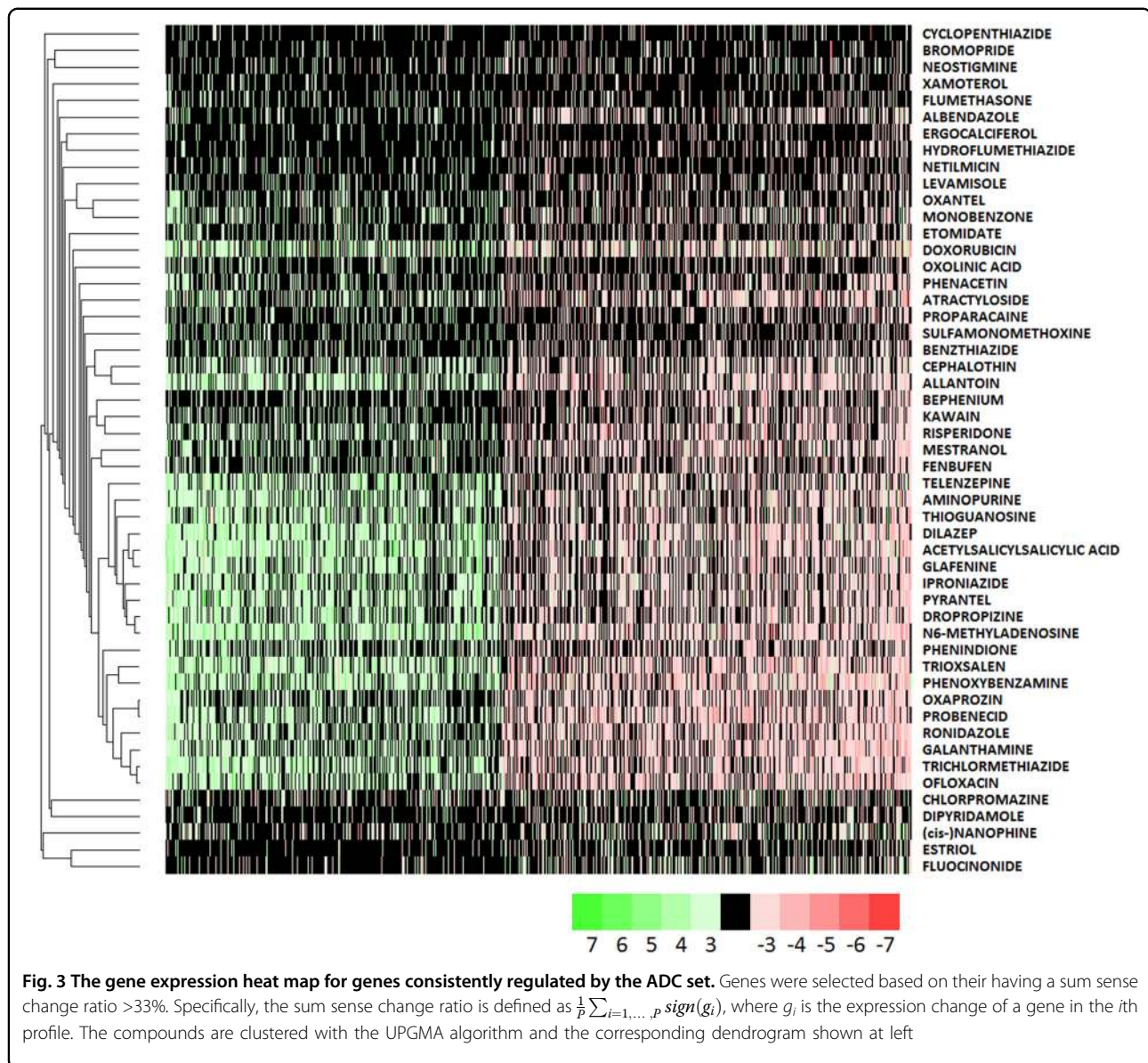
are given in Supplementary Table 6, and these point to key processes associated with AD that underpin the potential therapeutic action of the drugs. The enrichment for the AD, Parkinson's disease and mitochondrial pathways in the positively regulated gene sets is driven by the upregulation of cytochrome c oxidases, ubiquinone oxidoreductases and ATP synthases. These are all key players in mitochondrial function, which is known to be compromised in AD<sup>55,56</sup>, with growing evidence that gene variation affecting mitochondrial function may play a role in AD<sup>57,58</sup>. The downregulated set appears to be less consistent. Nonetheless, the enrichment of immune-associated pathways points to a possible anti-inflammatory activity of the candidate drugs. Interestingly, the following drugs have been reported to have neuroprotective activity: fluocinonide<sup>59</sup>, kawain<sup>60–63</sup>, allantoin<sup>64</sup>, dipyrindamole<sup>65–67</sup>, estriol<sup>68</sup>, levamisole<sup>69</sup>, mycophenolic acid<sup>70</sup>, neostigmine<sup>71</sup>, probenecid<sup>72,73</sup>, chlorpromazine<sup>74</sup>, and phenoxybenzamine<sup>75</sup>, and xamoterol has been reported to ameliorate neuroinflammation and pathology in 5xFAD mice<sup>76</sup> and shown to enhance cognition in a Down syndrome mouse model<sup>77</sup>. The atypical antipsychotic risperidone prescribed to manage psychosis in AD has demonstrated neuroprotection in animal models of ischemia<sup>78</sup>. Furthermore, cholinesterase inhibition is a therapeutic strategy for AD<sup>79</sup> and there are two such inhibitors in the candidate list with galantamine

as an established AD therapeutic<sup>80</sup>, while neostigmine exhibits poor blood–brain barrier penetrance and is therefore not in clinical use for AD. There does not appear to be any gene expression signature distinguishing compounds with reported neuroprotective activities from the other ADC compounds. This is to be expected as not all compounds have been assayed for neuroprotection and biological activity is not expected to be solely encoded in the transcriptome.

## Discussion

Neurodegenerative diseases present a therapeutic challenge due to the difficulty in establishing a clear protein or mechanistic culprit for classic target-based intervention. Another hurdle is a consequence of the temporal extent of disease progression and the probable need to treat before overt symptom onset. This is a particular problem in designing clinical trials. With this in mind, alternatives to target-based approaches are increasingly being pursued. One recent report compared Parkinson's disease (PD) incidence and chronic therapeutic use data from the Norwegian Prescription Database ([www.norpd.no](http://www.norpd.no)), showing that salbutamol use reduced PD risk<sup>81</sup>. A middle ground between target-based and epidemiological approaches is a methodology based on the disease phenotype gleaned from gene expression changes observed in pathological states. Underlying this approach is the observation that disease states can effectively be represented by characteristic expression changes, in the sense that these changes are consistent and can function as high content quantitative biomarkers. One avenue available to drug repositioning is to use these transcriptional phenotypes together with the hypothesis that an anti-correlation in phenotypes is indicative of the therapeutic potential of the compound. Whereas the transcriptional landscape of neurodegeneration and AD in particular has been well characterised, the corresponding data for compounds are either limited to full profiles defined on non-neuronal proliferating cells or partial profiles on iPSC-derived neuronal cells. The basis of the present study is to go some way to fill this gap in the compound-associated transcriptome with an emphasis on drugs with an anti-AD potential.

In the context of defining the neurotherapeutic potential of candidate drugs, a further development would be to treat wild-type or mutant AD mice with the compounds and measure expression changes in the brain, along the lines of the DrugMatrix project<sup>82</sup>. This approach would have the advantage of including non-neuronal factors contributing to AD pathology such as inflammation. However, practical considerations limit whole-animal approaches to smaller drug sets and will therefore form part of a subsequent endeavour based on a more limited set of drug candidates selected based on the iPSC data.



In the present work, we have established an AD transcriptional profile landscape and shown this to have a high degree of internal consistency. This disease-associated transcriptional landscape served as the basis for selecting a series of candidate drugs from the CMAP database of cancer cell line profiles, which were then assayed for their transcriptional effect on iPSC-derived cortical neurons. The iPSC profiles show a degree of overlap with the corresponding CMAP profiles, with a highly significant overall comparison in terms of the ranks observed for iPSC queries of CMAP. Out of the 153 iPSC drug profiles, 51, termed the ADC set, showed a high degree of anti-correlation with transcriptional changes seen in AD. A pathway enrichment analysis performed on each of the ADC set showed that pathways related to mitochondrial

function were commonly upregulated while commonly downregulated pathways represented immune-associated pathways. Interestingly, these pathological features are found in multiple neurodegenerative disorders, such as PD and Huntington's disease, and it would be of interest to investigate whether these compounds may have wider therapeutic potential. Notably, 18 of the ADC drugs already have established neuroprotective ability in published studies. Whereas we expect that initial CMAP filtering against AD profiles has led to increased likelihood of discovering compounds that tend to reverse AD-associated expression changes in the context of iPSC cultures, this can only be rigorously assessed by generating iPSC profiles for a series of compounds randomly selected from the CMAP database, which is outside the



scope of the present study. In conclusion, approaches to identifying a broader range of candidate therapies for AD are urgently needed. It is therefore expected that the iPSC database will serve as a useful platform for drug repositioning across multiple neuropathological disorders as well as AD.

#### Acknowledgements

This work was funded by the Wellcome foundation: A systematic programme to develop and evaluate the best candidate treatments for repositioning as therapies for Alzheimer's disease (SMART-AD) reference 102001/Z/13/Z.

#### Conflict of interest

C.B. reports grants and personal fees from Lundbeck and Acadia and personal fees from Roche, Orion, GSK, Otusaka, Heptares and Lilly outside the submitted work. The other authors declare that they have no conflict of interest.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41398-019-0555-x>).

Received: 22 March 2019 Revised: 21 May 2019 Accepted: 17 July 2019

Published online: 06 September 2019

#### References

- Su, A. I. et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA* **101**, 6062–7 (2004).
- Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R. & O'Keefe, S. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–47 (2014).
- Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–76 (2007).
- Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–42 (2015).
- Chen, R., Wu, X., Jiang, L. & Zhang, Y. Single-cell RNA-Seq reveals hypothalamic cell diversity. *Cell Rep.* **18**, 3227–41 (2017).
- Busch, H. et al. Gene network dynamics controlling keratinocyte migration. *Mol. Syst. Biol.* **4**, 199 (2008).
- Golub, T. R. et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–7 (1999).
- Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–51 (2013).
- Marton, M. J. et al. Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat. Med.* **4**, 1293–301 (1998).
- Hughes, T. R. et al. Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–26 (2000).
- Lamb, J. et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–35 (2006).
- Williams, G. A searchable cross-platform gene expression database reveals connections between drug treatments and disease. *BMC Genomics* **13**, 12 (2012).
- Wei, G. et al. Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell* **10**, 331–42 (2006).
- Zhang, D. et al. Ouabain mimics low temperature rescue of F508del-CFTR in cystic fibrosis epithelial cells. *Front. Pharmacol.* **3**, 176 (2012).
- Sirota, M. et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci. Transl. Med.* **3**, 96ra77 (2011).
- Dudley, J. T. et al. Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci. Transl. Med.* **3**, 96ra76 (2011).
- Chang, M., Smith, S., Thorpe, A., Barratt, M. J. & Karim, F. Evaluation of phenoxylbenzamine in the CFA model of pain following gene expression studies and connectivity mapping. *Mol. Pain* **6**, 56 (2010).
- Kunkel, S. D. et al. mRNA expression signatures of human skeletal muscle atrophy identify a natural compound that increases muscle mass. *Cell Metab.* **13**, 627–638 (2011).
- Walf-Vorderwulbecke, V. et al. Targeting acute myeloid leukemia by drug-induced c-MYB degradation. *Leukemia* **32**, 882–9 (2018).
- Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–5 (2013).
- So, H. C. et al. Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry. *Nat. Neurosci.* **20**, 1342–9 (2017).
- Corbett, A., Williams, G. & Ballard, C. Drug repositioning in Alzheimer's disease. *Front. Biosci.* **7**, 184–8 (2015).
- Blalock, E. M. et al. Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc. Natl Acad. Sci. USA* **101**, 2173–8 (2004).
- Wang, M. et al. Integrative network analysis of nineteen brain regions identifies molecular signatures and networks underlying selective regional vulnerability to Alzheimer's disease. *Genome Med.* **8**, 104 (2016).
- Webster, J. A. et al. Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445–58 (2009).
- Readhead, B. et al. Multiscale analysis of independent Alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. *Neuron* **99**, 64–82 e7 (2018).
- Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437–52 e17 (2017).
- Dolmetsch, R. & Geschwind, D. H. The human brain in a dish: the promise of iPSC-derived neurons. *Cell* **145**, 831–4 (2011).
- Egawa, N. et al. Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci. Transl. Med.* **4**, 145ra04 (2012).
- Ochalek, A. et al. Neurons derived from sporadic Alzheimer's disease iPSCs reveal elevated TAU hyperphosphorylation, increased amyloid levels, and GSK3B activation. *Alzheimers Res. Ther.* **9**, 90 (2017).
- Kondo, T. et al. Modeling Alzheimer's disease with iPSCs reveals stress phenotypes associated with intracellular Aβeta and differential drug responsiveness. *Cell Stem Cell* **12**, 487–96 (2013).
- Barrett, T. et al. NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res.* **35**, D760–5 (2007).
- Tan, M. G. et al. Genome wide profiling of altered gene expression in the neocortex of Alzheimer's disease. *J. Neurosci. Res.* **88**, 1157–69 (2010).
- Hokama, M. et al. Altered expression of diabetes-related genes in Alzheimer's disease brains: the Hisayama study. *Cereb. Cortex* **24**, 2476–88 (2014).
- Antonell, A. et al. A preliminary study of the whole-genome expression profile of sporadic and monogenic early-onset Alzheimer's disease. *Neurobiol. Aging* **34**, 1772–8 (2013).
- Miller, J. A., Woltjer, R. L., Goodenbour, J. M., Horvath, S. & Geschwind, D. H. Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med.* **5**, 48 (2013).
- Berchtold, N. C. et al. Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proc. Natl Acad. Sci. USA* **105**, 15605–10 (2008).
- Berson, A. et al. Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. *EMBO Mol. Med.* **4**, 730–42 (2012).
- Lai, M. K., Esiri, M. M. & Tan, M. G. Genome-wide profiling of alternative splicing in Alzheimer's disease. *Genom. Data* **2**, 290–2 (2014).
- Blalock, E. M., Buechel, H. M., Popovic, J., Geddes, J. W. & Landfield, P. W. Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. *J. Chem. Neuroanat.* **42**, 118–26 (2011).
- Liang, W. S. et al. Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol. Genomics* **28**, 311–22 (2007).
- Silva, A. R. et al. Transcriptional alterations related to neuropathology and clinical manifestation of Alzheimer's disease. *PLoS ONE* **7**, e48751 (2012).
- Pangman, V. C., Sloan, J. & Guse, L. An examination of psychometric properties of the mini-mental state examination and the standardized mini-mental state examination: implications for clinical practice. *Appl. Nurs. Res.* **13**, 209–13 (2000).
- Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. & Martin, R. L. A new clinical scale for the staging of dementia. *Br. J. Psychiatry.* **140**, 566–72 (1982).

45. Paesler, K. et al. Limited effects of an eF2alphaS51A allele on neurological impairments in the 5xFAD mouse model of Alzheimer's disease. *Neural Plast.* **2015**, 825157 (2015).
46. Boeddrich, A. et al. The anti-amyloid compound DO1 decreases plaque pathology and neuroinflammation-related expression changes in 5xFAD transgenic mice. *Cell Chem. Biol.* **26**, 109–20 e7 (2019).
47. Neuner, S. M., Heuer, S. E., Huentelman, M. J., O'Connell, K. M. S. & Kaczorowski, C. C. Harnessing genetic complexity to enhance translatability of Alzheimer's disease mouse models: a path toward precision medicine. *Neuron* **101**, 399–411 e5 (2019).
48. Marsh, S. E. et al. The adaptive immune system restrains Alzheimer's disease pathogenesis by modulating microglial function. *Proc. Natl Acad. Sci. USA* **113**, E1316–25 (2016).
49. Pereson, S. et al. Progranulin expression correlates with dense-core amyloid plaque burden in Alzheimer disease mouse models. *J. Pathol.* **219**, 173–81 (2009).
50. Castillo, E. et al. Comparative profiling of cortical gene expression in Alzheimer's disease patients and mouse models demonstrates a link between amyloidosis and neuroinflammation. *Sci. Rep.* **7**, 17762 (2017).
51. Sykora, P. et al. DNA polymerase beta deficiency leads to neurodegeneration and exacerbates Alzheimer disease phenotypes. *Nucleic Acids Res.* **43**, 943–59 (2015).
52. Marttinen, M. et al. A multiomic approach to characterize the temporal sequence in Alzheimer's disease-related pathology. *Neurobiol. Dis.* **124**, 454–68 (2019).
53. Williams, G. SPIEDw: a searchable platform-independent expression database web tool. *BMC Genomics* **14**, 765 (2013).
54. Hargis, K. E. & Blalock, E. M. Transcriptional signatures of brain aging and Alzheimer's disease: what are our rodent models telling us? *Behav. Brain Res.* **322**, 311–28 (2017).
55. Swerdlow, R. H. & Khan, S. M. A "mitochondrial cascade hypothesis" for sporadic Alzheimer's disease. *Med. Hypotheses* **63**, 8–20 (2004).
56. Moreira, P. I., Carvalho, C., Zhu, X., Smith, M. A. & Perry, G. Mitochondrial dysfunction is a trigger of Alzheimer's disease pathophysiology. *Biochim. Biophys. Acta* **1802**, 2–10 (2010).
57. Lakatos, A. et al. Association between mitochondrial DNA variations and Alzheimer's disease in the ADNI cohort. *Neurobiol. Aging* **31**, 1355–63 (2010).
58. Alvarez, V. et al. Mitochondrial transcription factor A (TFAM) gene variation and risk of late-onset Alzheimer's disease. *J. Alzheimers Dis.* **13**, 275–80 (2008).
59. Wang, J. et al. Identification of select glucocorticoids as Smoothened agonists: potential utility for regenerative medicine. *Proc. Natl Acad. Sci. USA* **107**, 9323–8 (2010).
60. Assemi, M. Herbs affecting the central nervous system: ginkgo, kava, St. John's wort, and valerian. *Clin. Obstet. Gynecol.* **44**, 824–35 (2001).
61. Backhaus, C. & Krieglstein, J. Extract of kava (*Piper methysticum*) and its methysticin constituents protect brain tissue against ischemic damage in rodents. *Eur. J. Pharmacol.* **215**, 265–9 (1992).
62. Schmidt, N. & Ferger, B. Neuroprotective effects of (+/-)-kavain in the MPTP mouse model of Parkinson's disease. *Synapse* **40**, 47–54 (2001).
63. Wruck, C. J. et al. Kavalactones protect neural cells against amyloid beta peptide-induced neurotoxicity via extracellular signal-regulated kinase 1/2-dependent nuclear factor erythroid 2-related factor 2 activation. *Mol. Pharmacol.* **73**, 1785–95 (2008).
64. Ahn, Y. J. et al. Effects of allantoin on cognitive function and hippocampal neurogenesis. *Food Chem. Toxicol.* **64**, 210–6 (2014).
65. Blake, A. D. Dipyridamole is neuroprotective for cultured rat embryonic cortical neurons. *Biochem. Biophys. Res. Commun.* **314**, 501–4 (2004).
66. Farinelli, S. E., Greene, L. A. & Friedman, W. J. Neuroprotective actions of dipyridamole on cultured CNS neurons. *J. Neurosci.* **18**, 5112–23 (1998).
67. Guo, S., Stins, M., Ning, M. & Lo, E. H. Amelioration of inflammation and cytotoxicity by dipyridamole in brain endothelial cells. *Cereb. Dis.* **30**, 290–6 (2010).
68. MacKenzie-Graham, A. et al. Estriol-mediated neuroprotection in multiple sclerosis localized by voxel-based morphometry. *Brain Behav.* **8**, e01086 (2018).
69. Shukry, M. et al. Pinacidil and levamisole prevent glutamate-induced death of hippocampal neuronal cells through reducing ROS production. *Neurol. Res.* **37**, 916–23 (2015).
70. Ebrahimi, F. et al. Time dependent neuroprotection of mycophenolate mofetil: effects on temporal dynamics in glial proliferation, apoptosis, and scar formation. *J. Neuroinflamm.* **9**, 89 (2012).
71. Qian, J. et al. A combination of neostigmine and anisodamine protects against ischemic stroke by activating alpha7nAChR. *Int. J. Stroke* **10**, 737–44 (2015).
72. Colin-Gonzalez, A. L. & Santamaria, A. Probenecid: an emerging tool for neuroprotection. *CNS Neurol. Disord. Drug Targets* **12**, 1050–65 (2013).
73. Vamos, E., Voros, K., Zadori, D., Vecsei, L. & Klivenyi, P. Neuroprotective effects of probenecid in a transgenic animal model of Huntington's disease. *J. Neural Transm.* **116**, 1079–86 (2009).
74. Geng, X. et al. Neuroprotection by chlorpromazine and promethazine in severe transient and permanent ischemic stroke. *Mol. Neurobiol.* **54**, 8140–50 (2017).
75. Rau, T. F., Kothival, A., Rova, A., Rhoderick, J. F. & Poulsen, D. J. Phenox-ybenzamine is neuroprotective in a rat model of severe traumatic brain injury. *Int. J. Mol. Sci.* **15**, 1402–17 (2014).
76. Ardestani, P. M. et al. Modulation of neuroinflammation and pathology in the 5XFAD mouse model of Alzheimer's disease using a biased and selective beta-1 adrenergic receptor partial agonist. *Neuropharmacology* **116**, 371–86 (2017).
77. Faizi, M. et al. Comprehensive behavioral phenotyping of Ts65Dn mouse model of Down syndrome: activation of beta1-adrenergic receptor by xamoterol as a potential cognitive enhancer. *Neurobiol. Dis.* **43**, 397–413 (2011).
78. Yan, B. C. et al. Neuroprotection of posttreatment with risperidone, an atypical antipsychotic drug, in rat and gerbil models of ischemic stroke and the maintenance of antioxidants in a gerbil model of ischemic stroke. *J. Neurosci. Res.* **92**, 795–807 (2014).
79. Birks, J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst. Rev.* CD005593 (2006).
80. Scott, L. J. & Goa, K. L. Galantamine: a review of its use in Alzheimer's disease. *Drugs* **60**, 1095–122 (2000).
81. Mittal, S. et al. beta2-Adrenoreceptor is a regulator of the alpha-synuclein gene driving risk of Parkinson's disease. *Science* **357**, 891–8 (2017).
82. Hardt, C. et al. ToxDB: pathway-level interpretation of drug-treatment data. *Database (Oxford)* **2016**, baw052 (2016).