Identifying new antiepileptic drugs through genomics-based drug repurposing

Nasir Mirza1,* Greame J. Sills1, Munir Pirmohamed1,† and Anthony G. Marson1,†

1Department of Molecular & Clinical Pharmacology, University of Liverpool, Liverpool L69 3GL, UK

*To whom correspondence should be addressed at: Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, Block A: Waterhouse Building, 1-5 Brownlow Street, Liverpool L69 3GL. Tel: +44 151 794 5549; Fax: +44 151 794 5059; Email: nasir.mirza@liv.ac.uk

Abstract

Currently available antiepileptic drugs (AEDs) fail to control seizures in 30% of patients. Genomics-based drug repurposing (GBR) offers the potential of savings in the time and cost of developing new AEDs. In the current study, we used published data and software to identify the transcriptomic signature of chronic temporal lobe epilepsy and the drugs that reverse it. After filtering out compounds based on exclusion criteria, such as toxicity, 36 drugs were retained. 11 of the 36 drugs identified (>30%) have published evidence of the antiepileptic efficacy (for example, curcumin) or antiepileptogenic affect (for example, atorvastatin) in recognised rodent models or patients. By objectively annotating all ~20,000 compounds in the LINCS database as either having published evidence of antiepileptic efficacy or lacking such evidence, we demonstrated that our set of repurposable drugs is ~6-fold more enriched with drugs having published evidence of antiepileptic efficacy in animal models than expected by chance (P-value <0.006). Further, we showed that another of our GBR-identified drugs, the commonly-used well-tolerated antihyperglycemic sitagliptin, produces a dose-dependent reduction in seizures in a mouse model of pharmacoresistant epilepsy. In conclusion, GBR successfully identifies compounds with antiepileptic efficacy in animal models and, hence, it is an appealing methodology for the discovery of potential AEDs.

Introduction

Epilepsy is amongst the most common neurological disorders: 65 million people have epilepsy worldwide (1). Currently available AEDs (AEDs) have been identified using traditional drug-discovery approaches (2): serendipity, phenotypic screening, hypothesis-driven target-orientated design, and structural modification of existing molecules. These compounds act primarily upon a small number of targets involved mainly in modulating neuronal excitability (2). Currently available AEDs have a number of significant shortcomings. They fail to control seizures in 30% of patients (3), and possess no antiepileptogenic or disease-modifying activity (4,5). In addition, they do not exhibit any meaningful efficacy against significant epilepsy comorbidities (6), such as cognitive impairment and psychiatric conditions. Over 100 years of utilising traditional drug discovery approaches for the development of an increasingly long list of approved AEDs, now approaching 30, has failed to overcome these challenges. A need exists, therefore, for newer more effective AED-discovery approaches.

In recent years, there has been an increasing interest in exploiting genomics-based drug repurposing (GBR) to aid and accelerate the drug discovery process. GBR is based on the following precepts (7). Every biological state and, hence, every disease state can be described by a gene expression signature. Treatments that restore gene expression patterns to their norm are associated with the successful restoration of the disease phenotype (8). The methodology can be summarised as follows.
The first step is to generate a signature of differential gene expression for the disease through a genome-wide gene expression analysis comparing normal tissue with the disease tissue of interest. Then, using the gene expression signature of disease, databases of drug-induced gene expression signatures, such as Connectivity Map (cMap) and Library of Integrated Network-based Cellular Signatures (LINCS) are queried. These databases contain drug-induced signatures of differential gene expression for thousands of compounds. Each signature is generated through genome-wide gene expression analysis comparing cells before and after drug exposure. If disease- and drug-induced signatures are sufficiently negatively correlated (i.e. the genes upregulated in the disease-induced signature are down-regulated in the drug-induced signature and vice versa) then the effect of the drug on transcription is opposite to the effect of the disease. Hence, the drug might revert the disease signature of differential gene expression and the disease phenotype itself. GBR has successfully produced a number of therapeutic leads for different diseases (8–11).

GBR holds much promise for drug discovery in epilepsy. It is known that hundreds of genes influence susceptibility to epilepsy (12) and are altered in the epileptic focus (13,14). By targeting these dysregulated genes, and restoring global gene expression patterns to their healthy norm, repurposed drugs could ameliorate the diverse pathological changes that create and promote the hyperexcitable neuronal network which generates seizures. Although this hypothesis is yet to be tested in epilepsy, it has been shown in other diseases (8) that treatments which restore gene expression patterns to their norm are more effective in restoring the normal clinical phenotype. This approach offers the potential of efficacy against epilepsy, epileptogenesis, and the comorbidities that share genes, biological pathways and pathophysiological mechanisms with epilepsy.

**Results**

The transcriptomic signature of epilepsy and the drugs that reverse it

The transcriptomic signature of disease was extracted from a published RNA sequencing analysis of a mouse model of chronic temporal lobe epilepsy (TLE) (16). The genes constituting the transcriptomic signature of TLE used in the current study are listed in [Supplementary Material, Table S1](http://www.broadinstitute.org/gsea/). Drugs that reverse the epileptic transcriptomic signature were identified using the Library of Integrated Network-based Cellular Signatures (LINCS) data and software. 123 compounds ([Supplementary Material, Table S2](http://www.broadinstitute.org/gsea/)) met the inclusion criterion (LINCS mean connectivity score threshold of -85). After filtering out compounds ([Supplementary Material, Table S2](http://www.broadinstitute.org/gsea/)) based on the exclusion criteria of toxicity, parental route of administration, lack of animal or human dosage data, or BBB-impermeability, 36 compounds remained ([Supplementary Material, Table S3](http://www.broadinstitute.org/gsea/)). We found no explicit published evidence of BBB-impermeability for any of the drugs; most have never undergone an in vitro or in vivo assessment of BBB-permeability. The known primary mode of action (17) of each of the 36 compounds is listed in [Supplementary Material, Table S3](http://www.broadinstitute.org/gsea/).

**Dissimilar drugs modulate the same biological pathways**

We used the Drug-Set Enrichment Analysis (DSEA) tool to demonstrate that the identified compounds, although pharmacologically diverse, are alike in modulating biological pathways. DSEA showed that the unselected sample of five GBR-identified drugs with varying modes of action shared 134 (out of 669) Reactome pathways at a significance threshold of \( P < 0.05 \) ([Supplementary Material, Table S4](http://www.broadinstitute.org/gsea/)). The median number of significant pathways shared by 1000 randomly-selected five-compound drug-sets was 31, the mean was 33.2, and the range was 10-98 ([Supplementary Materials, Table S5 and Fig. S1](http://www.broadinstitute.org/gsea/)). Hence, none of the 1000 random drug-sets shared \( \geq 134 \) pathways, demonstrating that the number of pathways shared by GBR-identified drugs is statistically significant (permutation-based \( p \)-value <0.001).

**Repurposable drugs are better than conventional AEDs at reversing the pathway dysregulation underlying epilepsy**

The Reactome pathways most significantly dysregulated in TLE were identified by performing Gene-Set Enrichment Analysis ([http://www.broadinstitute.org/gsea/](http://www.broadinstitute.org/gsea/)) separately on the up- and down-regulated genes constituting the transcriptomic signature of TLE identified above. The hundred most up-regulated and the hundred most down-regulated pathways were noted; the degree of enrichment was statistically significant (FDR <0.05) for all of these 200 pathways. 29 pathways were both up- and down-regulated; these pathways were excluded. It should be noted that in the Reactome database, pathways are assembled into a hierarchy of biological processes. The top hierarchy of biological processes represents very broad categories, such as ‘Developmental Biology’ and ‘Neuronal System’, within which subcategories can be differently dysregulated; this was the case with the 29 pathways both up- and down-regulated in the current analysis. Of the 142 (out of 200) pathways remaining, one was not present in the DSEA database (as the database excludes the very smallest and largest pathways) leaving 141 pathways to be considered in the present analysis.

We hypothesized that repurposable drugs are better than conventional AEDs at rectifying the dysregulated pathways identified above. In turn, each repurposable drug present in the in the DSEA database was compared with the set of conventional AEDs present in the DSEA database. [Supplementary Material, Table S6 and Fig. 1](http://www.broadinstitute.org/gsea/) show the percentage of dysregulated pathways corrected most effectively by each drug. In each comparison, the repurposable drug was the best performing compound for the highest proportion of pathways. The fraction of dysregulated pathways for which the repurposable drug surpassed conventional AEDs was statistically significant (binomial test FDR < 0.05) for all comparisons.

**The set of identified compounds is enriched with AEDs**

At the time of conducting this analysis, there were 19,767 unique compounds within the LINCS database. Using the strategy detailed in Methods of objective computational data retrieval and parsing, followed by careful manual filtering based on pre-defined exclusion criteria, 203 drugs (1%) were identified as possessing published evidence of antiepileptic efficacy in animal models.

Of the 123 drugs included in our transcriptionally-repurposed drug-set, seven (5.7%) possessed published evidence of antiepileptic efficacy, which represents an approximately six-fold enrichment.
LINCS database compounds were randomly divided into 159 sets of 123 drugs each. The median number of AEDs within the random drug-sets was 1, the mean was 1.2, and the range was 0-4. None of the random drug-sets included seven or more drugs with published evidence of antiepileptic efficacy, confirming that the enrichment of AEDs within the transcriptionally-repurposed drug-set is statistically significant, with a permutation-based p-value of \( < 0.006 \) (\( =1/159 \) random drug-sets). This enrichment was also statistically significant according to the hypergeometric distribution (\( P < 0.0003 \)).

Further details about the data used for completing this enrichment analysis, including evidence used for considering or discounting drugs as antiepileptic, can be found in the Supplementary Material.

In addition to above unsupervised computational approach for parsing information from Medline abstracts, which was applied to all 123 drugs identified by the LINCS computational algorithm, a more detailed manual literature-review was performed in order to identify additional evidence of antiepileptic efficacy for the 36 compounds that were selected after application of exclusion criteria; results for this have been summarised in Table 1.

**In vivo validation**

Sitagliptin, one of the drugs identified using our GBR pipeline, was chosen for in vivo validation as it is a commonly-used well-tolerated licensed medication. Levetiracetam, an AED which is well characterised and effective in this model (18,19), was included in the drug screening to verify adequate execution of the screening process. Sitagliptin produced a dose-dependent reduction in seizure scores (Supplementary Material, Fig. S2) in the 6 Hz psychomotor seizure mouse model (18,21). The mean \( \pm \) standard error seizure scores for the vehicle, sitagliptin 125 mg/kg, sitagliptin 250 mg/kg and sitagliptin 500 mg/kg groups were 0.93 \( \pm \) 0.12, 0.73 \( \pm \) 0.18, 0.60 \( \pm \) 0.14, and 0.47 \( \pm \) 0.13, respectively. In keeping with published reports, the AED levetiracetam (18,19) also produced a reduction in seizures scores. There was a statistically significant difference in seizure scores across the groups (Kruskal–Wallis p-value < 0.05). The difference between control and treatment group seizure scores was statistically significant (two-sided Mann-Whitney U test FDR < 0.05) for the highest dose of sitagliptin evaluated (500 mg/kg) and for levetiracetam. Individual seizure scores are shown in Supplementary Material, Table S9.

**Discussion**

In this study, we have applied GBR to epilepsy, demonstrated that this identifies a set of compounds significantly enriched with drugs possessing efficacy in animal models of epilepsy, and presented in vivo validation of a GBR-identified drug.

Our GBR strategy—identifying compounds using the LINCS algorithm and excluding those unsuited for oral administration—identified 36 drugs. Using DSEA, we showed that our sample of GBR-identified compounds exert a similar modulatory effect on 134 Reactome pathways; significantly greater than expected by chance (permutation-based p-value < 0.001). This observation is useful for two reasons. Firstly, the GBR-identified compounds appear to have little in common with each other: they are pharmacologically diverse with different molecular targets. Dissimilar drugs with disparate modes of action would seem unlikely to induce the same antiepileptic phenotype. Our analysis shows that, in spite of their different modes of action, GBR-identified compounds modulate the same biological pathways, enabling them to
induce the same therapeutic effect. Secondly, the DSEA algorithm is based on gene expression data from cMap, not LINCS. Hence, it is encouraging to see that an analysis using a paradigm and a database of drug-induced gene expression signatures different from that used to perform GBR corroborates the pathway-level functional similarity of our GBR-identified compounds.

It is interesting to note the lack of conventional licensed AEDs amongst our GBR-identified compounds, which suggests that the GBR-identified compounds are more effective than some conventional AEDs at reversing the transcriptomic signature of TLE. In keeping with this, we showed, using DSEA-based analysis, that each repurposable drug in our sample is better than all tested conventional AEDs in rectifying the highest number of dysregulated biological pathways underlying TLE. Again, it is encouraging to see that an analysis using a paradigm and a database of drug-induced gene expression signatures different from that used to perform GBR corroborates the superiority of the tested GBR-identified compounds over the tested conventional AEDs at correcting the TLE transcriptomic signature. At the same time, these findings suggest that the tested conventional AEDs are not reliant upon correcting the transcriptomic signature of disease in order to exert their therapeutic effect; traditional modes of action, such as blocking voltage-gated sodium or calcium channels, may be sufficient for them to produce an anti-seizure effect in most (but not all) individuals. Whether the signature-reversion effect of GBR-identified compounds translates into a more effective, broader or disease-modifying clinical effect should be investigated in future functional studies. However, our bioinformatic, literature-based and in vivo analysis hints at this possibility, as discussed below.

A number of the GBR-identified drugs have published evidence of antiepileptic or antiepileptogenic efficacy in animal models. Almost a third of the final 36 compounds identified have functional evidence of efficacy in animal models of epilepsy. We demonstrated that the GBR-identified drugs are significantly enriched (permutation-based p-value < 0.006) with drugs having published evidence of antiepileptic efficacy in animal models by using a novel objective methodology, based upon the auto-parsing of relevant information from Medline abstracts. We demonstrated this enrichment within all 123 compounds identified by the LINCS algorithm, rather than to just the 36 compounds that remained after filtering, in order to avoid the possibility of perceived bias in the manual selection of the 36 compounds.

As mentioned above, one motivation for seeking transcriptome-rectifying therapies is the anticipation of antiepileptogenic efficacy. This expectation is supported by our findings: almost half of the compounds with published data from animal models have evidence of an antiepileptogenic effect (Table 1). An example is the drug lestaartinib. Using the rat hypoxia-induced

### Table 1. Drugs, from amongst the 36 GBR-identified compounds, that have evidence of antiepileptic or antiepileptogenic efficacy in animal models or humans.

<table>
<thead>
<tr>
<th>Drug</th>
<th>Mechanism of action</th>
<th>Summarised evidence of antiepileptic and antiepileptogenic efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-iodotubercidin</td>
<td>Adenosine kinase inhibitor</td>
<td>Suppression of epileptiform activity evoked by brief, high-frequency stimulation in rat hippocampal slices (57). Antiepileptic effect in rodent bicuculline (58), kainic acid (59), maximal electric shock (60,61) models.</td>
</tr>
<tr>
<td>Androstenol</td>
<td>Neurosteroid, GABA&lt;sub&gt;A&lt;/sub&gt; receptor modulator</td>
<td>Antiepileptic effect in the rodent 6 Hz electroshock and pentylenetetrazol (62) models</td>
</tr>
<tr>
<td>Atorvastatin</td>
<td>HMG-CoA reductase inhibitor</td>
<td>Antiepileptic effect in the rodent DBA/2 audiogenic seizures (63), increasing current electroshock (24,64), kainic acid (65), pentylenetetrazol (32,64,66–68), and quinolinic acid (69) models. Antiepileptogenic effect in the rodent pentylenetetrazol (24) and WAG/Rij (25) models. Association with reduced risk of poststroke epilepsy in patients (26).</td>
</tr>
<tr>
<td>Camptothecin</td>
<td>Inhibitor of DNA topoisomerase I</td>
<td>Antiepileptic effect in a Drosophila model of epilepsy (70)</td>
</tr>
<tr>
<td>Curcumin</td>
<td>Antioxidant</td>
<td>Antiepileptic effect in the rodent isoniaizid (87,88), metrazol (87) and pentylenetetrazol (89) model.</td>
</tr>
<tr>
<td>FGIN-1-27</td>
<td>Mitochondrial benzodiazepine receptor agonist</td>
<td>Antiepileptic effect in the rodent isoniaizid (87,88), metrazol (87) and pentylenetetrazol (89) model.</td>
</tr>
<tr>
<td>Lestaartinib</td>
<td>Tyrosine kinase inhibitor</td>
<td>Antiepileptogenic effect in rats: administration of lestaartinib within 12 hours of the first neonatal seizure attenuates increased susceptibility to seizures in later life (22).</td>
</tr>
<tr>
<td>Mepacrine</td>
<td>Multiple actions, including inhibition of PL&lt;sub&gt;A&lt;/sub&gt;</td>
<td>Antiepileptic effect in the rodent lithium-pilocarpine (34) model.</td>
</tr>
<tr>
<td>Quinidine</td>
<td>Sodium channel blocker</td>
<td>Antiepileptic effect in electroshock (90) and penicillin-induced (91) seizures in cats.</td>
</tr>
<tr>
<td>Resveratrol</td>
<td>Produced naturally by several plants; principal mode of action unknown</td>
<td>Antiepileptic effect in the rodent FeCl&lt;sub&gt;3&lt;/sub&gt; (92), increasing current electroshock (73,74), iron-induced epileptogenesis (75), kainic acid (76), penicillin-induced epileptiform activity (77), pentylenetetrazol (33,73,78–81), and pilocarpine (82,83) models. Anticonvulants in pilocarpine- (84) and pentylenetetrazol-induced status epilepticus (73). Protects against pentylenetetrazol-induced kindling (85), and exerts a favourable disease-modifying effect in the kainic acid model (86).</td>
</tr>
<tr>
<td>Sitagliptin</td>
<td>Dipeptidyl peptidase-4 inhibitor</td>
<td>Antiepileptic effect: sitagliptin produces a dose-dependent reduction in seizures in a mouse-model of drug-resistant epilepsy (current study)</td>
</tr>
</tbody>
</table>
epilepsy model, researchers showed that administration of lese-taurinib at the time of the epileptogenic insult attenuates susceptibility to seizures later in life (22). Within our list also is the first drug shown to be potentially antiepileptogenic in man: atorvastatin (23). After studies demonstrating the antiepileptogenic effects of atorvastatin in rodent models (24,25), Guo et al. have recently shown that administration of statins to patients with early-onset poststroke seizures is associated with reduced risk of poststroke epilepsy (26). In the study by Guo and colleagues, most patients received atorvastatin; there were not enough data to conduct appropriate subgroup analyses because too few patients were given simvastatin or rosvastatin.

It is clear that epilepsy and its comorbidities, such as anxiety, depression and cognitive impairment, share genes, biological pathways and pathophysiological mechanisms (27–31). Hence, another motivation for seeking transcriptome-rectifying antiepileptic therapies is the anticipation of efficacy against epilepsy comorbidities. Again, this expectation is supported by our findings: a number of compounds with published data from animal models of epilepsy have evidence of an effect against comorbidities in the models. For example, atorvastatin ameliorates anxiety, depression and cognitive impairment (25,32); resveratrol is effective against cognitive impairment and depression (33), and mepacrine ameliorates cognitive impairment (34) in rodent models of epilepsy.

We wished to show that, in addition to the compounds with published evidence of antiepileptic and antiepileptogenic efficacy found to be enriched within our drug-set, other GBR-identified compounds yet unstudied in animal models of epilepsy also have therapeutic efficacy. We demonstrated that the commonly-used well-tolerated licenced antihyperglycaemic agent sitagliptin produces a dose-dependent reduction in seizure scores in a mouse model of pharmacoresistant epilepsy—the 6 Hz psychomotor seizures model. Seizures in this model are resistant to benzodiazepines and other AEDs (29,30), and the effects of sitagliptin are dose-dependent and reproducible (31). Sitagliptin at twice the equivalent of the maximum dose used in the current study has been administered to rodents for repeated doses without adverse effect (36,37), and sitagliptin is effective in several different animal models of epilepsy (38). The success of sitagliptin in animal models of epilepsy allows us to use sitagliptin as a positive control to investigate whether other GBR-identified compounds show similar effects.

Identifying the transcriptomic signature of epilepsy and the drugs that reverse it

A transcriptomic signature of chronic temporal lobe epilepsy (TLE) was generated as follows. Differentially expressed genes
were extracted from the results of a recently published (16) high-throughput RNA sequencing analysis comparing hippocampal tissue from normal mice and from mice with pilocarpine-induced chronic epilepsy (mice exhibiting spontaneous recurrent seizures 6 weeks after pilocarpine-induced status epilepticus). The mouse genes were mapped to human homologues listed in the Mouse Genome Database (44). Human homologues were mapped to Affymetrix U133A microarray probe identifiers listed in the NetAffx database (45). Probes corresponding to the most differentially expressed mouse genes (FDR <0.05 & fold change ≥2) were retained.

Drugs that reverse the epileptic transcriptomic signature were identified using the Library of Integrated Network-based Cellular Signatures (LINCS) data and software (http://apps.lincscloud.org/query; date last accessed June 06, 2016). LINCS is successor to the highly successful Connectivity Map (46), with a vastly expanded number of drugs profiled in multiple human cell lines, producing more than a million transcriptomic drug signatures.

The LINCS software quantitates the (dis)similarity between the disease and each of the drug signatures. For each drug signature, the software generates a 'connectivity score' ranging from +100 to -100. A high positive connectivity score indicates that the corresponding compound's transcriptomic signature is like that of the disease signature, while a high negative connectivity score indicates that the compound's transcriptomic signature is opposite that of the disease. As each drug has been profiled in multiple human cell lines, there are multiple transcriptomic signatures for each drug, and multiple connectivity scores for each drug-disease combination. For each drug, the mean of the four highest connectivity scores across different cell lines is the default statistic for summarising the relationship of the drug-disease pair, in the LINCS algorithm. The cell lines utilised for calculating the mean connectivity score are not pre-selected by tissue of origin; any tissue-specific noise is thought to be drowned out by the powerful directionality of the induced signature (47). The mean connectivity score threshold for drug selection may be set at different points: a threshold score of ~90 has been proposed and used in published literature (48). In the current study, a mean connectivity score threshold of ~85 was chosen in order to generate a longer list of drugs for evaluation.

Drugs identified using the above process were then filtered based on the following exclusion criteria:

1) Drugs included in the National Institute for Occupational Safety and Health List of Antineoplastic and Other Hazardous Drugs (49), or known to be toxic, according to published literature
2) Drugs requiring parenteral administration
3) Drugs with no published data to indicate effective in vivo doses in humans or animals
4) Drugs reported to BBB-impermeable in published literature

For the drugs that remained after the above filtering process, the primary known mode of action was identified through literature review.

Do dissimilar drugs modulate the same biological pathways?

The drugs identified above appeared to have little in common with each other: they were pharmacologically diverse with different molecular targets. Dissimilar drugs with disparate modes of action would appear unlikely to induce the same antiepileptic phenotype. We hypothesized that the identified drugs, in spite of their different modes of action, modulate the same biological pathways, enabling them to induce the same therapeutic effect.

To test this hypothesis, Drug-Set Enrichment Analysis (DSEA), which is based upon concepts similar to those first introduced by Hegde and colleagues (50), was utilized. DSEA (51) is designed to search the transcriptional responses of different drugs for shared pathways whose genes are upregulated (or downregulated) by the drugs in the set. If drugs in a drug-set tend to modulate the same pathway more than the other drugs in the database, this pathway will be detected by DSEA. A set of drugs of interest is tested against a database of pathways. Each pathway in the database is stored as a ranked list of drugs, sorted from the one most up-regulating the pathway to the one most down-regulating it. Given a query-set of drugs, DSEA checks for each pathway whether the drugs tend to be significantly ranked at the top (or the bottom) of the list, by applying a Kolmogorov-Smirnov (KS) test. A KS p-value can thus be computed for each pathway. F-values assess how much the ranks of the chosen drugs are consistently small (up-regulation) or large (down-regulation) for each pathway. The final output of DSEA is a list of pathways ranked by the p-value.

The current DSEA database is based upon the cMap 2.0 data-set containing transcriptional profiles for 1309 drugs. Of the compounds identified through our drug repurposing pipeline, five (camptothecin, mepacrine, quinidine, resveratrol and scriptaid) are found in the cMap 2.0 and, hence, DSEA database. This subset is an unselected sample of five different drugs with widely varying modes of action and, hence, a suitable collection with which to test our hypothesis that GBR-identified drugs, though seemingly dissimilar, modulate the same pathways.

The five drugs were analysed using the DSEA webservice (http://dsea.tigem.it; date last accessed June 06, 2016). The number of significant Reactome pathways (p < 0.05) shared by these five drugs was counted. In order to determine if the number of pathways shared by the GBR-identified drugs is statistically significant, 1000 randomly-selected five-compound drug-sets were also analysed through the webservice, and the number of shared pathways was counted for each random set.

Are repurposable drugs better than conventional AEDs at reversing the pathway dysregulation underlying epilepsy?

We wished to determine if repurposable drugs were more effective than conventional AEDs at correcting the pathway dysregulation underlying TLE. The DSEA dataset was again used for this analysis. As stated above, each pathway in the DSEA database is stored as a ranked list of drugs, sorted from the one most up-regulating the pathway to the one most down-regulating it. Hence, for any pathway up- or down-regulated in disease, the DSEA dataset can be utilized for identifying the compound most effective at reversing the dysregulation.

As mentioned above, of the compounds identified through our drug repurposing pipeline, five (camptothecin, mepacrine, quinidine, resveratrol and scriptaid) are found in the DSEA database. The following eight conventional AEDs are in the DSEA database: acetazolamide, carbamazepine, ethosuximide, gabapentin, primidone, topiramate, valproic acid and vigabatrin. This set of drugs is an unselected sample of different compounds with widely varying modes of action and, hence, a suitable collection with which to test our hypothesis that
repurposable drugs are better than conventional AEDs at reversing the pathway dysregulation underlying TLE.

The Reactome pathways most significantly dysregulated in TLE were identified by performing Gene-Set Enrichment Analysis (http://www.broadinstitute.org/gsea/; date last accessed June 06, 2016) separately on the up- and down-regulated genes constituting the transcriptomic signature of TLE identified above. The hundred most up-regulated and the hundred most down-regulated pathways were noted; the degree of enrichment was statistically significant (FDR < 0.05) for all of these 200 pathways.

The following strategy was adopted:

1) in turn, compare each repurposable drug with the set of conventional AEDs,
2) for each dysregulated pathway, determine if the repurposable drug is better than conventional AEDs at rectifying the dysregulation,
3) determine if the repurposable drug outperforms conventional AEDs for more pathways than expected by chance alone.

The exact binomial test (R 3.2.4) was used to determine if the fraction of dysregulated pathways for which the repurposable drug surpasses conventional AEDs is statistically significant.

Is the set of identified compounds enriched with AEDs?

We hypothesized that the set of compounds identified through GBR is enriched with drugs of known antiepileptic efficacy. In order to test this hypothesis, the following strategy was devised:

1) Each of the compounds in the LINCS database was annotated as either having published evidence of antiepileptic efficacy or lacking published evidence of antiepileptic efficacy.
2) Based on the above, we determined the number of compounds with evidence of the antiepileptic efficacy amongst our set of repurposable drugs.
3) To determine if compounds with antiepileptic efficacy are overrepresented in our set of repurposable drugs, the LINCS compounds were divided into random sets of the same size as the repurposed set, and the number of drugs with antiepileptic efficacy was determined for each random set.

For added robustness, we also used the hypergeometric distribution (R 3.2.4) to determine the statistical significance of the enrichment of compounds with antiepileptic efficacy within our set of repurposable drugs.

For the above analysis, we included all 123 drugs identified by the LINCS computational algorithm, rather than the 36 that remained after application of exclusion criteria, as the latter manual filtering steps might inadvertently advantage our chosen drug set. Annotation of the LINCS compounds according to the evidence of antiepileptic efficacy was not limited to licensed AEDs, but included all drugs with evidence of efficacy in recognised animal models of epilepsy and seizures. In order to ensure that this process is objective, an unsupervised computational approach was adopted for parsing information from Medline abstracts about the antiepileptic efficacy of the LINCS compounds. As there are ~20,000 drugs in the LINCS database, a solely manual literature search would, in any case, be impracticable. A clear predefined search strategy and vocabulary of search terms was used for this data collection. Abstracts were downloaded from Medline and then automatically parsed (auto-parsed) using Unix command line text manipulation tools. Methodological details are provided in the Supplementary Material, including search terms and synonyms used for the Medline search and auto-parsing, and the Unix commands utilised. In order to maximise the relevance of the auto-parsed data, we sought sentence-level concurrence of search terms, which increases specificity and precision of data retrieval (52).

A Medline abstract sentence was extracted if the following search terms concurred within it: the name of a compound from the LINCS database, suppression (or a synonym thereof), seizures (or a synonym thereof), and the name of a recognised rodent model of epilepsy or seizures.

The results of the auto-parsing were manually reviewed to exclude any irrelevant results. Exclusion was based on clear objective pre-defined criteria: drugs were excluded if the information parsed (1) did not clearly indicate evidence of independent clinical antiepileptic efficacy, or (2) indicated promotion, rather than inhibition, of seizures by the drug.

It is important to note that these data are not presented as a comprehensive list of all drugs with published evidence of antiepileptic efficacy, but rather as a tool for objectively testing the overrepresentation of such drugs within the repurposable drug set.

In addition to above unsupervised computational approach for parsing information from Medline abstracts, which was applied to all 123 drugs identified by the LINCS computational algorithm, a more detailed manual literature-review was performed in order to identify additional evidence of antiepileptic efficacy for the 36 compounds that were selected after application of exclusion criteria.

In vivo validation

For in vivo validation, we wished to choose a drug that: (1) was previously unstudied in animal models of epilepsy, and (2) had Medicines and Healthcare products Regulatory Agency (MHRA) or European Medicines Agency (EMA) approval and extensive data of routine safe clinical use in humans, as such a drug would encounter the fewest research and regulatory barriers on the repurposing pathway and will be more readily adopted by the clinical community. After excluding drugs that already have independent published evidence of antiepileptic efficacy (Table 1), sitagliptin (an antihyperglycaemic) was the highest ranked MHRA/EMA-approved drug in routine clinical use in this country. Hence, sitagliptin was chosen for in vivo validation. The 6 Hz psychomotor seizure mouse model (20), a mouse model of pharmacoresistant epilepsy (18,21), was used for validating the efficacy of this drug. Sitagliptin was evaluated at 3 doses—125, 250 and 500 mg/kg—chosen because they exert a therapeutic effect without producing toxicity in rodent models of brain diseases (36,53–55). Levetiracetam (100 mg/kg), an AED which is well characterised and effective in this model (18,19), was included in the drug screening to verify adequate execution of the screening process. Physiological saline was used as the vehicle control substance. Hence, five groups of mice were studied: sitagliptin 125 mg/kg, sitagliptin 250 mg/kg, sitagliptin 500 mg/kg, levetiracetam 100 mg/kg, and saline.

There were fifteen male NMRI mice (Janvier Labs, France), 22 – 31 g body weight range, in each group. The animal experiments were performed by the Porsolt Research Laboratory (France).
Compounds were administered intraperitoneally (i.p.) 60 minutes before application of the current. Mice were administered a rectangular current (44 mA, rectangular pulse: 0.2 ms pulse width, 3 s duration, 6 Hz) via corneal electrodes connected to a constant current shock generator (Ugo Basile: type 7801). Seizure scores were recorded for each animal. Seizure scores were based on the occurrence and severity of forelimb clonus: absent (0), mild (1) or strong (2). Presence or absence of Straub tail was also recorded, but it was noted that Straub tail was observed in some animals in the absence of forelimb seizures and, hence, this feature was not used in the evaluation. Each test was performed blind.

Statistical analysis

For the bioinformatic analyses, permutation-based approaches used to determine statistical significance have been described in the relevant sections above. A two-sided exact binomial test (R 3.2.4) was used to determine if the fraction of dysregulated pathways for which the repurposed drug surpasses conventional AEDs is statistically significant. For the in vivo analysis, seizure scores across all groups were compared using the Kruskal–Wallis test; threshold of statistical significance was FDR $\alpha = 0.05$. For post hoc analysis, seizure scores for each of the drug treated groups were compared with the vehicle control group using a two-sided Mann-Whitney U test, with Benjamini and Hochberg false discovery rate (FDR) correction for multiple testing; threshold of statistical significance was FDR $\alpha = 0.05$. All statistical calculations were performed using R 3.2.4. G*Power version 3.1 (56) was used to calculate the power of the two-tailed Mann-Whitney U test ($\alpha = 0.05$).

Supplementary Material

Supplementary Material is available at HMG online.

Acknowledgements

MP acknowledges support from the NHS Chair of Pharmacogenetics and the MRC Centre for Drug Safety Science.

Conflict of Interest statement. None declared.

Funding

This work was supported by the Medical Research Council (Medical Research Council Research Training Fellowship awarded to NM) and the National Institute for Health Research (National Institute for Health Research Senior Investigator award to MP).

References


