Citation: CPT Pharmacometrics Syst. Pharmacol. (2019) 8, 211–219; doi:10.1002/psp4.12383

ARTICLE

A Network Pharmacology Approach for the Identification of Common Mechanisms of Drug-Induced Peripheral Neuropathy

Guillermo de Anda-Jáuregui[1,2], Brett A. McGregor[1], Kai Guo[1] and Junguk Hur[1,*]

Drug-induced peripheral neuropathy is a side effect of a variety of therapeutic agents that can affect therapeutic adherence and lead to regimen modifications, impacting patient quality of life. The molecular mechanisms involved in the development of this condition have yet to be completely described in the literature. We used a computational network pharmacology approach to explore the Connectivity Map, a large collection of transcriptional profiles from drug perturbation experiments to identify common genes affected by peripheral neuropathy-inducing drugs. Consensus profiles for 98 of these drugs were used to construct a drug–gene perturbation network. We identified 27 genes significantly associated with neuropathy-inducing drugs. These genes may have a potential role in the action of neuropathy-inducing drugs. Our results suggest that molecular mechanisms, including alterations in mitochondrial function, microtubule and cytoskeleton function, ion channels, transcriptional regulation including epigenetic mechanisms, signal transduction, and wound healing, may play a critical role in drug-induced peripheral neuropathy.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Drug-induced peripheral neuropathy (DIPN) is a side effect of many drugs that detrimentally impacts the quality of life of patients and therapeutic adherence. Mechanisms through which drugs can induce neuropathy have been described for a limited number of drugs.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ This study presents a computational approach based on graph theory to explore gene perturbation profiles of neuropathy-inducing drugs and to identify genes with potential functional implications in the development of DIPN.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ We used a network model to integrate high-throughput drug perturbation profiles to identify genes that are commonly affected by neuropathy-inducing drugs. With this model, we explored and identified genes associated with biological functions whose perturbation may be linked to DIPN.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

☑ Our network-based model provides a novel way of generating hypotheses that may drive new experimental efforts to identify mechanisms behind DIPN.

Drug-induced peripheral neuropathy (DIPN), also known as medication-induced or iatrogenic neuropathies, involves damage to the peripheral nervous system as an adverse effect of a therapeutic (or diagnostic) agent.[1–3] DIPN comprises a small subset of neuropathies, accounting for only 2–4% of all neuropathy cases.[4,5] Although this incidence rate may seem small, the condition impacts the patient's quality of life and influences therapeutic adherence.[6] Several drug classes have peripheral neuropathy as a side effect. Typical examples include chemotherapeutics,[7] antibiotics,[8] and HIV treatments.[9]

Currently, it is understood that drugs may induce peripheral neuropathy through different types of damage at the cellular level. These include (i) axonal degeneration through a dying-back mechanism, (ii) segmental demyelination, and (iii) damage to the soma of the neuron.[1,10] However, the molecular entities that can lead to these perturbations are varied. Furthermore, the full spectrum of neuropathy-inducing drugs (NIDs) has not been associated with mechanisms that can explain their link to this condition and an understanding of the molecular entities involved in the development of this adverse reaction is still needed.

The rise of genomic technologies has generated large amounts of biologically relevant data that require novel computational approaches for their analysis.[11] This study of disease from a systems biology perspective has allowed the development of descriptive models that link molecular

[1]Department of Biomedical Sciences, School of Medicine & Health Sciences, University of North Dakota, Grand Forks, North Dakota, USA; [2]Computational Genomics Division, National Institute of Genomic Medicine, Colonia Arenal Tepepan, Delegación Tlalpan, México DF, Mexico. *Correspondence: Junguk Hur (junguk.hur@med.und.edu)

Received: September 10, 2018; accepted: December 27, 2018. doi:10.1002/psp4.12383

alterations to physiological and pathological conditions, leading to new insights. The Connectivity Map (CMap) project,[12,13] which has generated a large collection of transcriptional responses to drug perturbation in cultured human cell lines, is a useful resource for the development of such models.

The use of network theory provides a theoretical framework suitable for the exploration of drug effects at different biological levels[14,15] and for the integration of large-scale drug information.[16] The CMap data have been studied through the construction of network models by others.[17,18] In this work, we systematically identified potential molecular mechanisms that may be implicated in DIPN. We developed a network-based approach to connect peripheral NIDs to targets at the gene expression level based on experimental data from CMap. We identified the most connected genes in this network and evaluated whether this high-degree connectivity was exclusively associated with NIDs. We provide literature-based evidence of the effects that perturbation of these genes has in the neurological setting. We propose this approach as a method to identify genes and associated biological functions that have not been previously used for the study of DIPN.

METHODS

Peripheral NIDs

We used 234 drugs from a previous study of DIPN[10] in which a text-mining approach was used to identify a list of drugs associated with DIPN by using information from drug labels retrieved from the Drugs@FDA database.[19] In addition, this information was complemented using DailyMed[20] and the Side Effect Resource.[21]

CMap

The CMap[12,13] is an important resource for the study of pharmacological effects on gene expression. It contains data from a series of perturbation experiments on a variety of cell lines spanning a variety of experimental conditions. A major asset of this resource is the fact that comparability among samples is achieved by virtue of the experimental design, which aims to reduce batch effects and other artifacts. For our work, we retrieved the CMap transcriptional profiles currently available, which contain expression profiles for 6,100 different experimental conditions. These profiles were analyzed in the Affymetrix Human Genome U133A platform (Affymetrix, Santa Clara, California).

In CMap, each treatment is described by a nonparametric rank-ordered list of all probe sets in the microarray platform. The expression level of each probe after treatment is compared with the expression of the same probe in a vehicle control sample, and the differences between these expression levels are ranked with the highest ranked probe representing the probe exhibiting the maximum up-regulation (or activation) after treatment and the lowest ranked probe exhibiting the maximum down-regulation (or inhibition) after treatment.

Generation of unique drug sample profiles

The CMap data include transcriptomic profiles for 98 of the 234 NIDs listed in our previous work.[10] Two of these NIDs (exemestane and topiramate) had only one profile corresponding to one experimental condition tested, whereas the remaining 96 were profiled in more than one experimental condition (see **Table S1**). To have a representative ranked expression profile for each drug, we followed the Kruskal–Borda strategy described by Iorio *et al.*[22] For all the rank-ordered lists originating from the same drug, a distance metric was computed (in this implementation, Spearman's Footrule is used). The two closest samples were merged through a majority voting system and reranked until a single consensus profile for the drug was obtained. The result was a ranked expression matrix in which each column contains a drug, and each row contains the rank of the transcript's expression level. Transcripts belonging to the same genes were aggregated to the gene level using the maximum expression level among them.

Drug–gene perturbation network

To model the relationships between NIDs and genes experimentally found in the CMap, we employed a network-based approach. A network is a mathematical object composed of a set of nodes and a set of edges or links representing relationships between these nodes. In the case of a biological network, for instance, nodes can represent molecules, whereas edges can represent the physical and chemical interactions between them. In this work, we modeled the effects of the drug on gene expression as a bipartite network, which is composed of the following two classes of nodes: drugs and genes. Edges connect drugs and genes, representing an action of drug treatment on gene expression.

The network space was populated with the nodes of two types, representing the 98 NIDs whose perturbation profiles were included in CMap and the 12,438 genes measured in the microarray platform. An edge in the network was drawn between a drug and a gene if the gene was ranked in the top 100 positions (up-regulation) or the bottom 100 positions (down-regulation). This criterion was decided because the available data consist of ranked gene lists for each drug, which do not allow, for instance, to select any number of genes above a certain threshold (e.g., a significance threshold for differential expression values). The changes in gene expression induced by drugs can be generally thought of as markers of gene susceptibility to drug perturbation. The network model is then a representation of the potential susceptibility of a given gene to be affected by different drugs.

Construction of null models

To assess the significance of network properties, the generation of comparable networks through a null model is necessary. This should reflect the nature of the phenomenon being modeled through the biological network to be useful.[23] Two main questions arise when analyzing the of NID perturbation network generated in this work: (i) whether the NID perturbation network structure is different from a randomly generated network and (ii) whether the network properties, particularly the degree, of a particular gene in the NID perturbation network can be associated exclusively to NIDs and not to other non-neuropathy–inducing drugs. To answer these questions, we generated networks using two different null models.

The first null model consists of generating 98 artificial drug profiles with rank values randomly assigned to all 12,438 genes. The top 100 and the bottom 100 ranked genes were selected to draw a link between the drugs and genes. This is equivalent to a random rewiring, preserving the degree distribution of the drug layer nodes. We refer to this model as the *randomly generated null model*. We use this model to assess whether the network topology of the pharmacologically relevant network is different from that expected by randomly connecting drugs and genes.

The second null model evaluates whether the topological properties, such as degree, of a given node are exclusive to the neuropathy-related network. In this model, 98 drugs were randomly selected from the complete list of drugs included in the CMap data set (n = 1,211), excluding the 98 NIDs. Then, the top 100 and bottom 100 ranked genes for each profile are connected to the drug nodes. We refer to this model as the *randomly selected drug model.* With each model, an ensemble of 5,000 networks was generated.

Degree as a measure of relative importance of genes in the neuropathy context

Genes that are affected by a larger number of NIDs may be more related to the underlying mechanisms driving the pathological phenotype associated with these drugs. A fundamental network-based metric that can be used to identify such genes is their *degree*—the number of neighbors in the graph, which in our model represents the number of drugs that can affect that gene.

The distribution of degree values is one of the most defining properties of any given network and can be used to categorize the nature of a particular network as well as for comparison purposes.[24] To test the pharmacological significance of the drug perturbation network, the degree distribution of our experimental network was compared with those of the randomly generated null model networks using the Kolmogorov–Smirnov and χ^2 tests. In addition, Hellinger distance and Jensen Shannon divergence were computed.

The genes were selected based on their higher degree in the NID network, considering that genes that were connected to 10 or more drugs (10% of the total evaluated drugs) were more likely to be related to the neuropathic condition. The rationale behind this selection cut-off was that randomly generated networks such as the ones from the first null model did not have nodes with a degree higher than 10.

To further refine this selection, a secondary filtering was used, specifically whether a gene, identified as susceptible to NIDs by having a high degree, is more susceptible to these drugs than to other drugs that are not known to induce peripheral neuropathy. For each gene, an empirical distribution of degree frequency was constructed from the randomly selected drug networks. The degrees of the highly connected genes (degree $\geq$ 10) in the NID perturbation network were compared against these empirical distributions by calculating a Z-score. Those genes with a Z-score equal to or above 1.96 (equivalent to P value < 0.05) were considered to have a high degree in the neuropathy context.

Function identification and literature-based validation

We explored the National Center for Biotechnology Information (NCBI) gene database[25] to identify the functions in which the identified genes are generally involved. We used this information to generate a list of functions that, through associations to these genes, may be involved in the development of DIPN. In addition, we looked for the identified genes in a list of known housekeeping genes.[26] We queried PubMed abstracts using SciMiner,[27] a web-based literature-mining tool, to find papers where these functions are reported in a neurological context. The queries took the form of *function* AND *nerve*, *function* AND *neuron*, or *function* AND *pain*, for instance, "transcription AND neuron."

RESULTS

Network parameters

As illustrated in **Figure 1**, a network of 98 NIDs, of the original 234 drugs reported in our previous work,[10] and their perturbed genes was generated based on gene expression perturbation profiled in CMap. The CMap drug perturbation profile data contained a set of 12,438 genes; however, 5,300 of these genes were not connected (either through up- or down-regulation in the context of the CMap perturbation experiments) to any of the NIDs. The remaining 7,138 genes were perturbed by at least one drug and therefore were included in the network, having a degree of one or higher. Of these connected genes, 2,556 (35.81%) genes were connected exclusively to a single NID. The parameters of this network are summarized in **Table 1**, and **Supplementary Material S1** contains the complete network in gml format, which can be visualized with Cytoscape 3.[28]

One of the fundamental network measures is the cumulative degree distribution, which describes the cumulative frequency of degree values in the network. **Figure 2** illustrates the degree distribution (in terms of 1 – the cumulative distribution function) of gene nodes in the network as well as the distributions for comparable networks generated through the null model. The experimental distribution, represented by the solid line, was different (average Benjamini–Hochberg BH-corrected P values for Kolmogorov–Smirnov = 3.4×10^{-6}, $\chi^2 = 1.2 \times 10^{-6}$; average Hellinger distance = 0.1026, average Jensen Shannon divergence = 0.0105) from those of the randomly generated null model networks, whose average is represented by the dotted line.

We define the unique neighborhood of a node as the set of neighbors of a node that are not shared by any other node in the network. It is a measure of the overlap between genes targeted by each drug. **Figure 3** illustrates the different distributions of unique neighborhood sizes for drug nodes in the NID network (represented with the black, shaded distribution) in the randomly generated null model networks (represented with multiple colored distributions to the right) and for the randomly selected drug networks (multiple colored distributions to the left). The displacement of the distribution to the left indicates that, on average, the drugs are affecting genes unaffected by any other drug in the network.

The largest connected component of a network is the subgraph that contains the most connected nodes in a network. The size of this component can be used as a simple measure of network cohesiveness. **Figure 4** shows the size of the largest connected component of the network for the

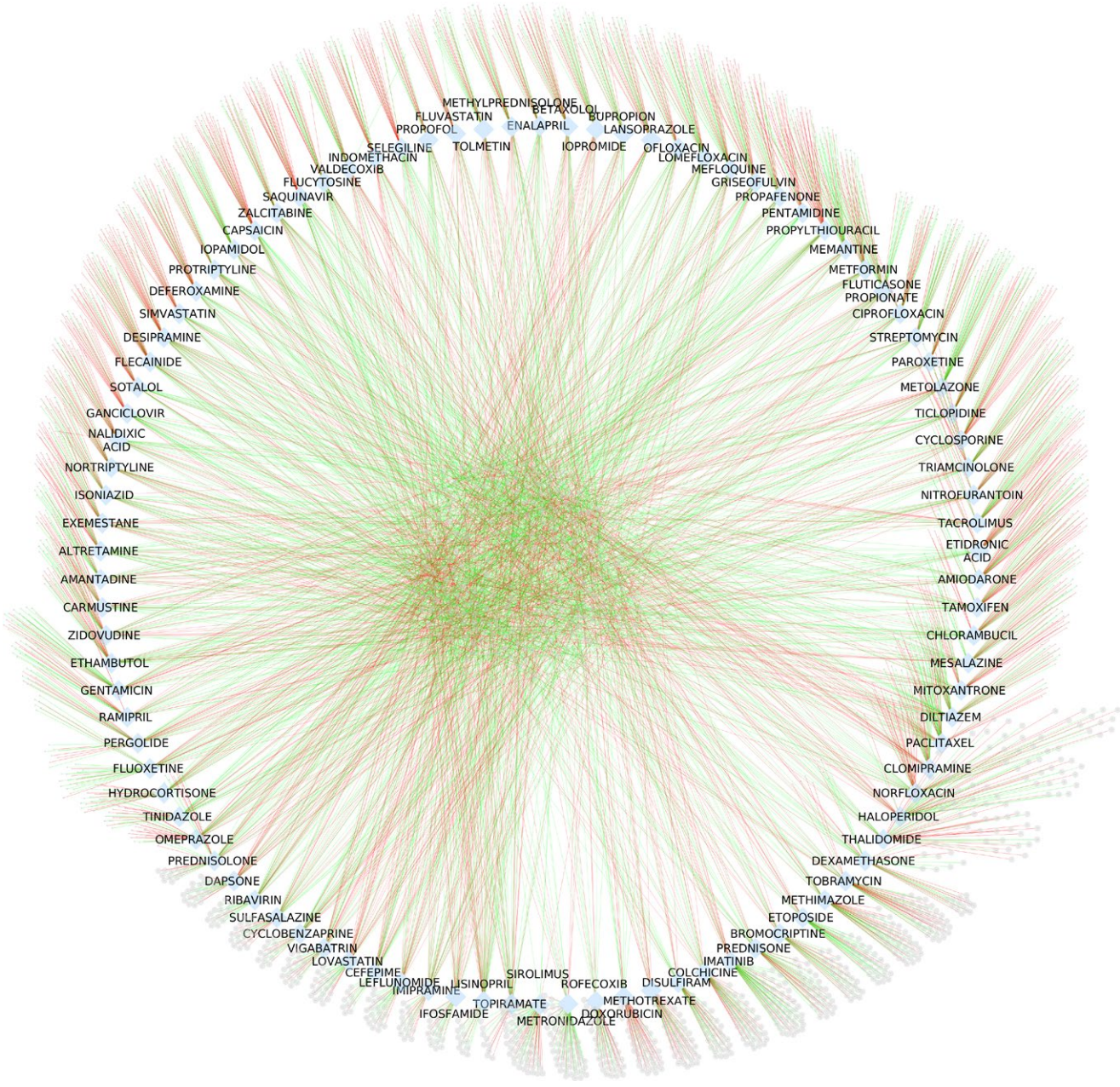

Figure 1 Drug–gene perturbation network visualization. Drugs are arranged in a circle (blue nodes), linked to genes through either upregulation (red links) or downregulation (green links). Genes perturbed by a single drug are oriented outside, whereas genes perturbed by multiple drugs are oriented inside the drug node circle. Transparency and sizes of nodes and edges were adjusted (based on degree and edge betweenness) for visualization purposes.

NID network compared with the size distribution of the randomly generated networks and the randomly selected drug networks. It is shown that the size of the largest connected component of the NID network is significantly smaller than that of the randomly generated networks and larger than that of the randomly selected drug networks.

Genes with high degrees and their functional relevance

In this work, highly connected genes are considered more likely to be involved in the neuropathic condition. We defined these highly connected genes as those with a degree of 10 or higher; 64 such genes were identified, which can be found in **Table S2**. Of these genes, 27 were found with a degree value significantly higher (Z-score ≥ 1.96) than that expected from the ensemble of networks derived from randomly selected drug networks. These genes and their biological functions can be found in **Table 2** .

Based on the functions identified for this list of genes, a systematic literature revision was performed. **Figure 5** shows the results of the SciMiner queries to identify the biological functions whose perturbations we propose to be

Table 1 Drug–gene perturbation network parameters

Parameter	Values
Drugs	98
Genes	12,438
Connected genes	7,138
Edges	19,600
Maximum degree (connected genes)	15
Number of connected components	1

associated with DIPN in neurological or pain-related contexts. **Figure 5a** shows the number of papers identified by these queries. **Figure 5b** shows the number of genes whose perturbations are associated with these queries. Finally, **Figure 5c** shows the number of papers in which genes identified in our NID network are found in the context of these queries (see **Table S3** for the list of PubMed identifiers for each gene–query pair).

DISCUSSION

In this work, data from high-throughput perturbation experiments of NIDs were integrated into a bipartite network model of drugs and perturbed genes. We observed that this network has a unique topology, with a degree distribution distinct from comparable random bipartite networks. The neighborhood of each drug node was unique and in each case was defined by a set of genes that are only perturbed by this drug and no other. None of the highly perturbed genes were connected to more than 15% of the NIDs, which suggests the involvement of diverse mechanisms leading to DIPN. These mechanisms may be associated with the alteration of biological functions in which the most connected genes in this network are involved. Therefore, we identified these highly connected genes in the NID network and demonstrated that their degree was high exclusively in the context of NIDs. Through systematic queries of the current biomedical literature using a text-mining approach, we identified instances in which perturbations of these genes have been reported in neurological or pain-related settings, some of which are discussed below.

The topology of the NID network was completely different from those of the randomly generated null model networks with respect to the degree distribution of gene nodes, overlap in gene neighbors of drug nodes, and size of the largest connected components. The degree distribution of drug nodes was constant because the network was constructed using the same number (100) of genes for each drug; however, the degree distribution of gene nodes was variable from 1 to 15. The degree distribution of the NID network reflects a higher quantity of disconnected gene nodes (that is, with degree 0) compared with the null model (see **Figure 2**). It also shows that in the NID network, there are genes with a degree value higher than the highest found in the null model.

The overlap in the gene neighbors of drug nodes was also different between the NID network and the randomly generated networks. This finding suggests that the genes affected by NIDs are not randomly distributed throughout the genome but, rather, certain genes are preferentially affected by these drugs. However, our results indicate that this preference is not exclusive to NIDs, as the networks generated from randomly selected non-NIDs also had smaller unique neighborhoods when compared with the randomly generated networks.

All drug nodes in the NID network, as well as the networks generated from the null models, are part of a single connected component, that is, a path can be traced from any drug to any other drug. The sizes of the largest connected component were significantly different between the NID network and those from both null models. The randomly generated networks had larger connected components than either the NID network or randomly selected drug networks.

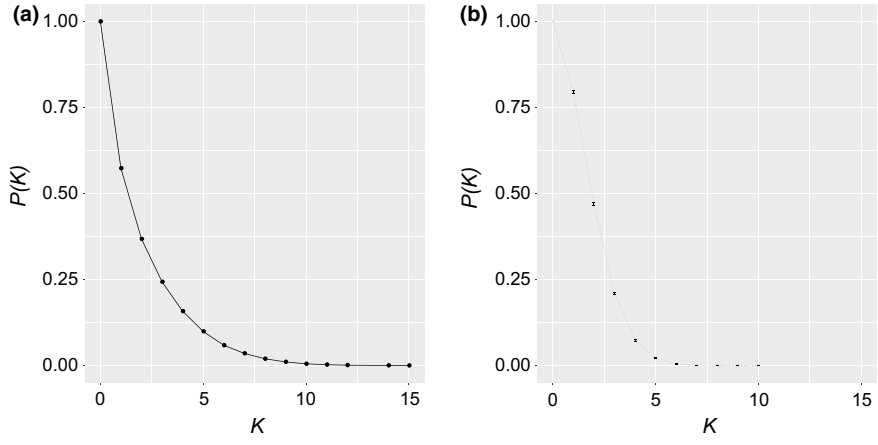

Figure 2 Degree frequency distribution. (**a**) The values for 1–(degree cumulative distribution function) ($P(K)$) vs. degree (K) for the neuropathy-inducing drug–gene perturbation network as a thick black line is illustrated. (**b**) The average value of $P(K)$ vs. K is shown as a dotted line for the networks in the randomly generated null model ensemble (5,000 networks). For each value of K, error bars are shown indicating the 5th and 95th percentile regions of values observed in the randomly generated null model ensemble networks. The major difference between the two panels is that the distributions for the random networks (**b**) have a maximum degree value of 10, whereas the drug-induced peripheral neuropathy network (**a**) has a maximum degree value of 15.

Table 2 High-degree genes, neuropathy-specificity Z-score, and biological role

Entrez gene ID	Symbol	Description	Degree	*Z*-score	Involved in[a]
54820	*NDE1*	nudE neurodevelopment protein 1	15	1.93	Microtubule organization
5971	*RELB*	RELB proto-oncogene, NF-kB subunit	14	2.16	Transcription factor
25819	*NOCT*	Nocturnin	14	4.06	Circadian regulation
64319	*FBRS*	Fibrosin	14	2.61	Fibroblast proliferation
2152	*F3*	Coagulation factor III, tissue factor	12	2.18	Coagulation
5296	*PIK3R2*	Phosphoinositide-3-kinase regulatory subunit 2	12	2.54	Signal transduction
28990	*ASTE1*	Asteroid homolog 1	12	3.9	Uncharacterized
51564	*HDAC7*	Histone deacetylase 7	12	2.51	Histone modification
57827	*C6orf47*	Chromosome 6 open reading frame 47[b]	12	2.28	Uncharacterized
1179	*CLCA1*	Chloride channel accessory 1	11	3.21	Ion channel
1271	*CNTFR*	Ciliary neurotrophic factor receptor	11	3.15	Neurite outgrowth
2356	*FPGS*	Folylpolyglutamate synthase[b]	11	2.16	Folate metabolism
9816	*URB2*	URB2 ribosome biogenesis 2 homolog (S. cerevisiae)	11	2.08	Uncharacterized
27156	*RSPH14*	Radial spoke head 14 homolog	11	2.28	Microtubule organization
51224	*TCEB3B*	Elongin A2	11	3.29	Transcription elongation
54332	*GDAP1*	Ganglioside induced differentiation associated protein 1	11	2.1	Mitochondrial metabolism
56672	*AKIP1*	A-kinase interacting protein 1	11	3.58	Signal transduction
2161	*F12*	Coagulation factor XII	10	3.53	Coagulation
3783	*KCNN4*	Potassium calcium-activated channel subfamily N member 4	10	2.52	Ion channel
4998	*ORC1*	Origin recognition complex subunit 1	10	2.99	Cell cycle control
5393	*EXOSC9*	Exosome component 9	10	2.2	RNA degradation
5565	*PRKAB2*	Protein kinase AMP-activated noncatalytic subunit beta 2	10	2.15	Signal transduction
9827	*RGP1*	RGP1 homolog, RAB6A GEF complex partner 1[b]	10	3.37	Signal transduction
22994	*CEP131*	Centrosomal protein 131	10	2.71	Microtubule organization
26468	*LHX6*	LIM homeobox 6	10	3.9	Transcriptional regulation
54714	*CNGB3*	Cyclic nucleotide gated channel beta 3	10	2.74	Ion channel
79157	*MFSD11*	Major facilitator superfamily domain containing 11[b]	10	1.95	Solute carrier ([49]Perland *et al*., 2016)

[a]Based on annotation from the NCBI gene database[25] unless otherwise noted. [b]Housekeeping gene according to Eisenberg and Levanon.[26]

This finding is consistent with what has been previously discussed regarding the preferential targeting of certain genes by drugs; experimentally, the drugs were found to act in a subset of the genome, leaving untargeted genes disconnected in the network. Between the NID and the randomly selected drug networks, the NID network's largest component was significantly larger. This result shows a difference in the topology of the NID network when compared with the networks of non-NIDs.

Although the NID network included more genes with a higher degree than the randomly generated networks (64 genes with a degree of 10 or higher), no single gene was connected to all NIDs. Therefore, we propose that DIPN may not be explained by the perturbation of a single gene (and associated biological function) but, rather, that each NID may have effects on different biological functions (thus altering the expression of the associated genes), which collectively lead to the neuropathic condition. We focused on the group of genes that, based on the higher degree in the NID network, were susceptible to the effects of more neuropathic drugs. Those genes with a high degree exclusively associated to NIDs (**Table 2**) serve as proxies to the biological functions that are perturbed by NIDs. The evidence of these genes' perturbation in neurological contexts was

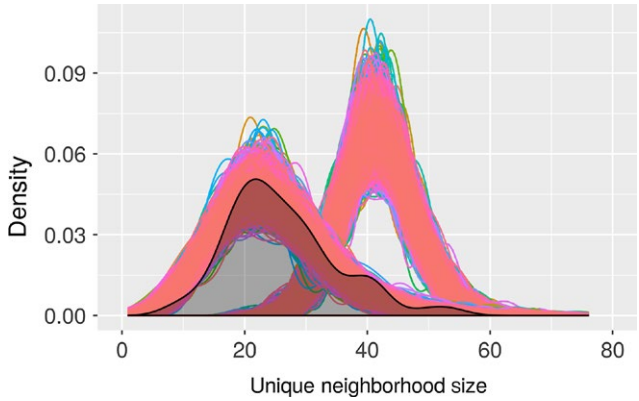

Figure 3 Unique neighborhood size distribution. The unique neighborhood size distribution for drug nodes in the drug-induced peripheral neuropathy network and the null models are shown. Distributions for the randomly generated networks are seen to the right, centered in a size 40. Distributions for the randomly selected drug networks are seen to the left, with a mode near size 20. The distribution for the drug-induced peripheral neuropathy network is shown in black shading; it should be noted that because this distribution lies completely in the region of randomly selected drug networks, it is indicated that the drug-induced peripheral neuropathy network differs in terms of neighborhood size distribution from the randomly generated networks but not from the randomly selected drug networks.

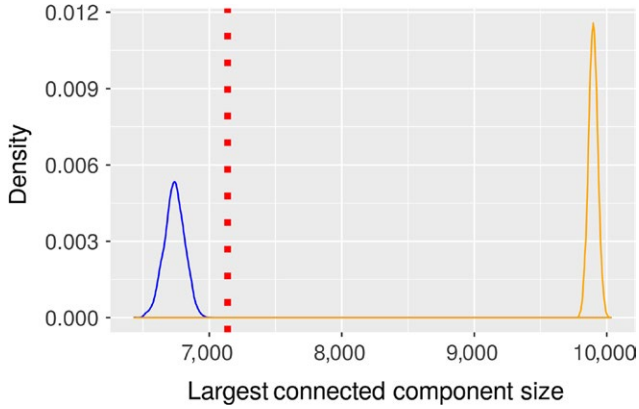

Figure 4 Largest connected component size distribution. The largest connected component size distribution is shown for the null models. The distribution for the randomly generated network model is shown to the right in yellow (mean value = 9,898). The distribution for the randomly selected drug network model is shown to the left in blue (mean value = 67,367). A dotted line indicates the largest connected component size for the drug-induced peripheral neuropathy network, consisting of 7,236 nodes.

collected using literature mining (**Figure 5**), which gives credence to the idea of these genes being potentially involved in the neuropathy.

Some of the genes identified using the NID network have strong associations to a neuropathic mechanism. For instance, impaired function of neuronal mitochondria has been proposed as having an important role in the development of neuropathies.[29] One of the genes identified in this work is ganglioside-induced differentiation-associated protein 1 (*GDAP1*), a gene that encodes an outer mitochondrial membrane protein. It acts as a regulator of mitochondrial fission,[30] and there is abundant evidence associating alterations of this gene to the hereditary Charcot-Marie-Tooth neuropathy (CMT). Different mutations of this gene have been observed and linked to different clinical manifestations of CMT.[31,32] Some mechanisms involving changes in mitochondrial movement, abnormal distribution,[33] and perturbations in mitochondrial fission[34] in CMT have been associated with mutations of *GDAP1*. Our results may provide a starting point for further exploration of therapies targeting mitochondrial processes,[35] which might have beneficial effects in the treatment of DIPN.

Another example is potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4 (*KCNN4*), encoding an intracellular calcium-activated potassium channel. *KCNN4* is expressed in the nervous system with distinctive expression patterns across cellular and subcellular regions.[36] This channel has just recently been targeted and shown to reverse tactile allodynia in a model of peripheral nerve injury in rodents.[37]

The exploration of the NID network allows the identification of such genes, which have been identified in other neuropathic settings, but it also points to those less explored. For instance, nudE nuclear distribution gene E homolog 1 (*NDE1*) is a gene involved in microtubule organization and encoding a protein that is part of the dynein complex. Although there is no evidence directly linking *NDE1* to neuropathies, other elements of the dynein complex have been found to malfunction in cases where defects in axonal transport lead to CMT.[38,39] Ciliary neurotrophic factor receptor (*CNTFR*) is another gene identified in the NID network that has not been associated with neuropathies. However, it is involved in neuronal survival,[40] and the signaling pathways activated through these receptors lead to protective effects against neurotoxicity in dopaminergic cells.[41]

Some of the genes identified through the network analysis have been barely explored. We would like to call special attention to nocturnin (*NOCT*), a gene involved in circadian regulation. The network model identifies it as one of the most connected (affected by 14 drugs) and, most important, one of the most exclusively associated only to NIDs. Although we found no information in the literature related to a role of these genes in a neurological (let alone neuropathic) setting, we believe, based on the network model, that further experimental exploration of these genes may provide insight on neuropathies.

When interpreting the significance of these highly connected genes in networks derived from transcriptional information such as ours, one important consideration is that housekeeping genes, playing functionally central roles, may also be topologically central.[42] We compared our highly connected genes with a curated list of housekeeping genes[26] and identified 12 such housekeeping genes among the 64 highly connected genes, which suggest no significant enrichment (P value = 0.2 by hypergeometric test). Among the 27 significant highly connected genes, there were only four housekeeping genes (**Table 2**). It is plausible to think that in the case of these genes, their

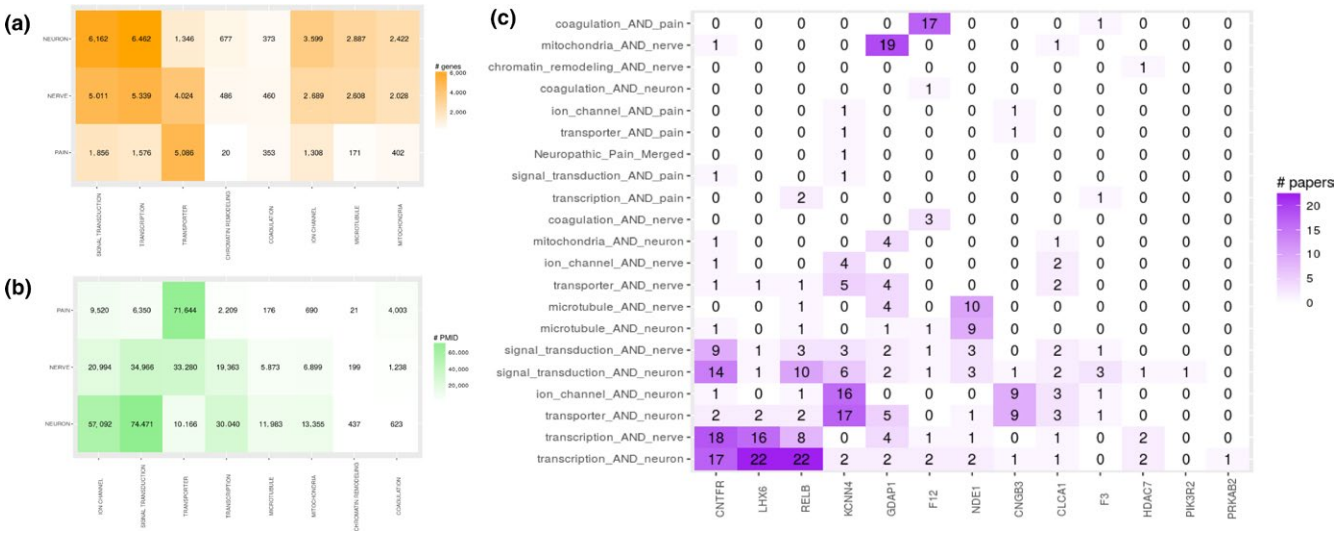

Figure 5 Heatmaps of literature-based validation: (**a**) the number of genes associated to a particular query (combination of keywords), (**b**) the number of papers indexed in PubMed, identified using PubMed IDs (PMID), retrieved by a combination of keywords, (**c**) the number of papers in which a given gene is associated to a particular query.

perturbation by NIDs affects the functionality of the cell; the mechanism through which this perturbation may induce the neuropathic condition remains to be described.

There are certain limitations to our model, mostly imposed by the availability of high-throughput experimental data. The coverage of NIDs in the CMap is limited and only 98 of 234 NIDs were included in the present study. The high-throughput profiles available in CMap were generated using gene perturbation experiments in cancer cell lines (*MCF7* Michigan Cancer Foundation-7, *ssMCF7* ss = serum starved, *HL60* Human Leukemia, *PC3* Prostate Cancer, and *SKMEL5* Skin Melanoma). The responses observed in these cell lines may not completely reflect the activity of a drug *in vivo* and in the context of drug-induced neuropathy. These cell lines were not derived from nerve or brain tissues. Although the CMap data may not reflect the specificity in neuronal tissues, they are still useful in understanding overall drug-perturbed transcriptional responses and have been used for other nervous system conditions[43–45] as well as for adverse drug reaction studies.[16,46]

The available perturbation profiles are provided as ranked gene lists, which forbids the use of individual selection criteria to identify the neighborhood of genes that are significantly perturbed by each drug. This consideration, along with the representativity of CMap data of the neurological context, was behind the decision of modeling only gene perturbation in the current study as opposed to using gene up/down regulation. If these limitations could be resolved, it would be possible to further discuss the contributions of these gene perturbations in terms of loss and gain of functions, which would further lead to the potential use of drugs inducing opposite perturbations in a drug-repurposing setting.

However, as more high-throughput data sets are released, the model may be refined and updated. The CMap itself has been expanded and integrated to the larger Library of Integrated Network-Based Cellular Signatures,[47] which will increase the coverage of drugs evaluated and provide perturbation profiles in cell lines more representative of the tissues of interest. Furthermore, the integration of other data sources beyond gene perturbation profiles into network models may allow for deeper and more meaningful interpretation of the role of genes in pathological conditions. The sources of this information vary, as they may come from general drug resources or information related to the specific subject such as the different forms of peripheral neuropathy used in our study. However, integrating the information of these resources to generate network models may not be trivial, even if the available information may be already represented as a network, and the differences in construction methods would make the networks incomparable unless adaptations are made.

The study of pharmacology from a systems perspective and the use of networks for this purpose are quickly becoming the norm. Network models are needed to understand the on and off-target effects of drugs that are involved in their therapeutic and adverse effects.[48] The work presented in this article is an example of an application of network models that may provide new insights into the nature of DIPN.

Supporting Information. Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (www.psp-journal.com).

Table S1. A summary of experimental conditions available in the Connectivity Map for the 98 neuropathy-inducing drugs.
Table S2. Genes with a degree of 10 or higher in the drug–gene perturbation network and their associated *Z*-scores.
Table S3. Lists of PubMed identifiers for each pair of gene–query corresponding to **Figure 5c.**
Supplementary Material S1. Neuropathy-inducing drug–gene network in gml format.

Acknowledgments. The authors would like to thank Dr Jane P.F. Bai from the US Food and Drug Administration for critical comments on this manuscript.

Funding. This research was partially supported by the National Institute of Health (R24 DK082841 to J.H.), the University of North Dakota Post-Doc Pilot Grant (to K.G.), and the North Dakota EPSCoR through National Science Foundation grant #OIA-1355466 (to B.A.M.).

Conflict of Interests. The authors declared no competing interests for this work.

Author Contributions. G.D.J., B.M., and J.H. wrote the manuscript; G.A.J. and J.H. designed the research; G.A.J., K.G., B.M., and J.H. performed the research; G.A.J., K.G., and B.M. analyzed the data.

1. Weimer, L.H. Medication-induced peripheral neuropathy. *Curr. Neurol. Neurosci. Rep.* **3**, 86–92 (2003).
2. Toyooka, K. & Fujimura, H. Iatrogenic neuropathies. *Curr. Opin. Neurol.* **22**, 475–479 (2009).
3. Vilholm, O.J., Christensen, A.A., Zedan, A.H. & Itani, M. Drug-induced peripheral neuropathy. *Basic Clin. Pharmacol. Toxicol.* **115**, 185–192 (2014).
4. Jain, K.K. Drug-induced peripheral neuropathies. *Drug Neurol. Disord.* **2**, 263–294 (2001).
5. Pratt, R.W. & Weimer, L.H. Medication and toxin-induced peripheral neuropathy. *Semin. Neurol.* **25**, 204–216 (2005).
6. Makanjuola, T., Taddese, H.B. & Booth, A. Factors associated with adherence to treatment with isoniazid for the prevention of tuberculosis amongst people living with HIV/AIDS: a systematic review of qualitative data. *PLoS ONE* **9**, e87166 (2014).
7. Ocean, A.J. & Vahdat, L.T. Chemotherapy-induced peripheral neuropathy: pathogenesis and emerging therapies. *Support. Care Cancer* **12**, 619–625 (2004).
8. Mafukidze, A.T., Calnan, M. & Furin, J. Peripheral neuropathy in persons with tuberculosis. *J. Clin. Tuberc. Other Mycobact. Dis.* **2**, 5–11 (2016).
9. Dalakas, M.C. Peripheral neuropathy and antiretroviral drugs. *J. Peripher. Nerv. Syst.* **6**, 14–20 (2001).
10. Hur, J., Guo, A.Y., Loh, W.Y., Feldman, E.L. & Bai, J.P.F. Integrated systems pharmacology analysis of clinical drug-induced peripheral neuropathy. *CPT Pharmacomet. Syst. Pharmacol.* **3**, e114 (2014).
11. Ultsch, A., Kringel, D., Kalso, E., Mogil, J.S. & Lötsch, J. A data science approach to candidate gene selection of pain regarded as a process of learning and neural plasticity. *Pain* **157**, 2747–2757 (2016).
12. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
13. Lamb, J. The Connectivity Map: a new tool for biomedical research. *Nat. Rev. Cancer* **7**, 54–60 (2007).
14. Harrold, J.M., Ramanathan, M. & Mager, D.E. Network-based approaches in drug discovery and early development. *Clin. Pharmacol. Ther.* **94**, 651–658 (2013).
15. Zhao, S. & Li, S. Network-based relating pharmacological and genomic spaces for drug target identification. *PLoS ONE* **5**, e11764 (2010).
16. Lee, S., Lee, K.H., Song, M. & Lee, D. Building the process-drug-side effect network to discover the relationship between biological processes and side effects. *BMC Bioinform.* **12** (suppl. 2), S2 (2011).
17. Cheng, F., Zhao, J., Fooksa, M. & Zhao, Z. A network-based drug repositioning infrastructure for precision cancer medicine through targeting significantly mutated genes in the human cancer genomes. *J. Am. Med. Inform. Assoc.* **23**, 681–691 (2016).
18. Cheng, F. *et al.* Systems biology-based investigation of cellular antiviral drug targets identified by gene-trap insertional mutagenesis. *PLoS Comput. Biol.* **12**, e1005074 (2016).
19. US Food and Drug Administration. https://www.accessdata.fda.gov/scripts/cder/daf/ accessed november 2017 <U.S. Drugs@ FDA> (2014).
20. National Institutes of Health. Health & Human Services daily med. <http://dailymed.nlm.nih.gov/dailymed/index.cfm> (2005).
21. Kuhn, M., Letunic, I., Jensen, L.J. & Bork, P. The SIDER database of drugs and side effects. *Nucleic Acids Res.* **44**, D1075–D1079 (2016).
22. Iorio, F. *et al.* Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl. Acad. Sci. U S A* **107**, 14621–14626 (2010).
23. Iorio, F. *et al.* Efficient randomization of biological networks while preserving functional characterization of individual nodes. *BMC Bioinform.* **17**, 542 (2016).
24. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47 (2002).
25. Brown, G.R. *et al.* Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res.* **43**, D36–D42 (2015).
26. Eisenberg, E. & Levanon, E.Y. Human housekeeping genes, revisited. *Trends Genet.* **29**, 569–574 (2013).
27. Hur, J., Schuyler, A.D., States, D.J. & Feldman, E.L. SciMiner: web-based literature mining tool for target identification and functional enrichment analysis. *Bioinformatics* **25**, 838–840 (2009).
28. Su, G., Morris, J.H., Demchak, B. & Bader, G.D. Biological network exploration with cytoscape 3. *Curr. Protoc. Bioinform.* **2014**, 8.13.1–8.13.24 (2014).
29. Chandrasekaran, K. *et al.* Mitochondrial transcription factor A regulation of mitochondrial degeneration in experimental diabetic neuropathy. *Am. J. Physiol. - Endocrinol. Metab.* **309**, E132–E141 (2015).
30. Milone, M. & Benarroch, E.E. Mitochondrial dynamics: general concepts and clinical implications. *Neurology* **78**, 1612–1619 (2012).
31. Pezzini, I. *et al.* GDAP1 mutations in Italian axonal Charcot-Marie-Tooth patients: phenotypic features and clinical course. *Neuromuscul. Disord.* **26**, 26–32 (2016).
32. Kabzińska, D., Kotruchow, K., Cegielska, J., Hausmanowa-Petrusewicz, I. & Kochański, A. A severe recessive and a mild dominant form of Charcot-Marie-Tooth disease associated with a newly identified Glu222Lys GDAP1 gene mutation. *Acta Biochim. Pol.* **61**, 739–744 (2014).
33. Pla-Martín, D. *et al.* Silencing of the Charcot-Marie-Tooth disease-associated gene GDAP1 induces abnormal mitochondrial distribution and affects Ca^{2+} homeostasis by reducing store-operated Ca^{2+} entry. *Neurobiol. Dis.* **55**, 140–151 (2013).
34. Huber, N., Guimaraes, S., Schrader, M., Suter, U. & Niemann, A. Charcot-Marie-Tooth disease-associated mutants of GDAP1 dissociate its roles in peroxisomal and mitochondrial fission. *EMBO Rep.* **14**, 545–552 (2013).
35. Areti, A., Yerra, V.G., Komirishetty, P. & Kumar, A. Potential therapeutic benefits of maintaining mitochondrial health in peripheral neuropathies. *Curr. Neuropharmacol.* **14**, 593–609 (2016).
36. Turner, R.W. *et al.* Neuronal expression of the intermediate conductance calcium-activated potassium channel KCa3. 1 in the mammalian central nervous system. *Pflügers Arch. J. Physiol.* **467**, 311–328 (2015).
37. Staal, R.G.W. *et al.* Inhibition of the potassium channel KCa3. 1 by senicapoc reverses tactile allodynia in rats with peripheral nerve injury. *Eur. J. Pharmacol.* **795**, 1–7 (2017).
38. d'Ydewalle, C., Benoy, V. & Bosch, L. Van Den Charcot-Marie-Tooth disease: emerging mechanisms and therapies. *Int. J. Biochem. Cell Biol.* **44**, 1299–1304 (2012).
39. Weedon, M.N. *et al.* Exome sequencing identifies a DYNC1H1 mutation in a large pedigree with dominant axonal Charcot-Marie-Tooth disease. *Am. J. Hum. Genet.* **89**, 308–312 (2011).
40. Davis, S. *et al.* The receptor for ciliary neurotrophic factor. *Science* **253**, 59–63 (1991).
41. Jeong, K.H., Nam, J.H., Jin, B.K. & Kim, S.R. Activation of CNTF/CNTFRα signaling pathway by hRheb (S16H) transduction of dopaminergic neurons *in vivo*. *PLoS ONE* **10**, e0121803 (2015).
42. Goh, K.-I. *et al.* The human disease network. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8685–8690 (2007).
43. Kidnapillai, S. *et al.* The use of a gene expression signature and connectivity map to repurpose drugs for bipolar disorder. *World J. Biol. Psychiatry* https://doi.org/10.1080/15622975.2018.1492734.[e-pub ahead of print]. .
44. Siavelis, J.C., Bourdakou, M.M., Athanasiadis, E.I., Spyrou, G.M. & Nikita, K.S. Bioinformatics methods in drug repurposing for Alzheimer's disease. *Brief. Bioinform.* **17**, 322–335 (2016).
45. Xiao, S.J. *et al.* Gene expression profiling coupled with Connectivity Map database mining reveals potential therapeutic drugs for Hirschsprung disease. *J. Pediatr. Surg.* **53**, 1716–1721 (2018).
46. Hur, J., Liu, Z., Tong, W., Laaksonen, R. & Bai, J.P.F. Drug-induced rhabdomyolysis: from systems pharmacology analysis to biochemical flux. *Chem. Res. Toxicol.* **27**, 421–432 (2014).
47. Keenan, A.B. *et al.* The library of integrated network-based cellular signatures NIH program: system-level cataloging of human cells response to perturbations. *Cell Syst.* **6**, 13–24 (2018).
48. Cheng, F. *et al.* Network-based approach to prediction and population-based validation of *in silico* drug repurposing. *Nat. Commun.* **9**, 2691 (2018).
49. Perland, E., Lekholm, E., Eriksson, M.M., Bagchi, S., Arapi, V. & Fredriksson, R. The Putative SLC Transporters Mfsd5 and Mfsd11 Are Abundantly Expressed in the Mouse Brain and Have a Potential Role in Energy Homeostasis. *PLoS ONE* **11**(6), e0156912 (2016). https://doi.org/10.1371/journal.pone.0156912. ISSN 1932-6203. PMC 4896477. PMID 27272503.