

Pain Networks: A web-based resource for the visualisation of pain-related genes in the context of their network associations

James R. Perkins, Jonathan Lees, Ana Antunes-Martins, Ilhem Diboun, Stephen B. McMahon, David L.H. Bennett, Christine Orengo

Institute of Structural and Molecular Biology, University College London, London, UK
The Wolfson Centre for Age-Related Diseases, King's College London, Guy's Campus, London, UK
The Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK
Department of Physiology and Biophysics, Weill Cornell Medical College in Qatar, Education City, Doha, P.O. Box 24144, Qatar

Keywords:
Systems biology
Protein–protein interactions
Protein interaction networks
Web-based resource
Pain genes
Microarrays

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Article Info

Article history:
Received 9 January 2013
Received in revised form 15 August 2013
Accepted 2 September 2013

Abstract

Hundreds of genes are proposed to contribute to nociception and pain perception. Historically, most studies of pain-related genes have examined them in isolation or alongside a handful of other genes. More recently the use of systems biology techniques has enabled us to study genes in the context of the biological pathways and networks in which they operate.

Here we describe a Web-based resource, available at http://www.PainNetworks.org. It integrates interaction data from various public databases with information on known pain genes taken from several sources (eg, The Pain Genes Database) and allows the user to examine a gene (or set of genes) of interest alongside known interaction partners. This information is displayed by the resource in the form of a network.

The Web site can be used to find out more about a gene of interest by looking at the function of its interaction partners. It can also be used to interpret the results of a functional genomics experiment by revealing putative novel pain-related genes that have similar expression patterns to known pain-related genes and by ranking genes according to their network connections with known pain genes.

We expect this resource to grow over time and become a valuable asset to the pain community.

© 2013 The Authors Published by Elsevier B.V. on behalf of International Association for the Study of Pain. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Persistent pain has a complex pathophysiological basis in which multiple mechanisms contribute to its initiation and maintenance [17,23,44,65]. The risk of developing this condition is dependent on the interaction between environmental and genetic factors. Multiple techniques are currently being employed to investigate the molecular basis of pain: transcriptional profiling of experimental models of persistent pain in rodent and human [14,20,33,40], the study of inbred rodent strains and genetically modified animals to understand the genetic basis of acute and chronic pain [1.2,3,4,63], and, more recently, human genetics, describing both high impact variants [31,64] and the use of gene association studies in experimental pain models and patient cohorts [15,66]. Systems biology (defined as “the strategy of pursuing integration of complex data about the interactions in biological systems from diverse experimental sources . . .” [6]), provides a means to organise, integrate, and maximise the utility of these large disparate pain-related data sets [3].

Gene expression profiling experiments using microarray technology represent one means to study pain in an unbiased, high-throughput manner [14,29,33,40]. These functional genomics studies provide a large amount of information on biological systems, including processes involved in pain. However, they generate large gene lists. By considering the relationship between these putative genes and those responsible for pain processes, a biologist can focus on genes interacting with pain-related genes and therefore likely to be pain related.

Protein interactions can be detected using a range of technologies [48], and these data have been collected and made available by several public resources [4,13,18,27,52,57]. By combining gene
expression data with protein networks, built from this interaction/association data, we can identify parts of the network activated under particular conditions and detect differentially expressed genes likely to be involved in pain due to connections with known pain genes.

Network-based approaches have been successfully applied to many human diseases, including cancer [10,47,59], diabetes [38], asthma [25], Parkinson’s [45], Alzheimer’s [30,55], and systemic inflammation [11]. Several resources exist that enable a user to build networks and examine their gene(s) of interest in the context of their network associations [35,58,63].

The PainNetworks resource displays pain-specific gene networks built by combining publicly available protein interaction data with microarray-derived gene expression data from a wide range of pain-related experiments. Users can query with a single gene or list of genes generated by their own transcriptional profiling experiments. The resulting network shows interactions between the query genes and other nonquery genes, including pain-related genes, for example, from the Pain Genes Database [34]. Clustering of the networks can be performed to detect modules of related genes; filtering can be used to select genes from a specific tissue type or anatomical structure of interest. Annotation can also be obtained for network genes and interactions.

Here we describe the Web site and underlying methodology. We also demonstrate the power of the networks to reproduce known associations in some pain-related biological pathways of interest and we give examples of using the site to find out more about genes and experiments of interest.

2. Methods

2.1. Network construction

In order to build the networks presented on the site, protein-protein interaction data were taken from a number of databases, as described below. These interactions were then combined in an unweighted manner to build a network, that is, the network will have an edge between 2 nodes (representing genes/proteins) provided at least one public resource reports an experimental validation for this interaction/association. PainNetworks then allows the user to view a part of this network, using a query gene or set of genes, and to further filter this network using various data sources.

2.2. Interaction/association data

Protein interactions were taken from 6 public databases (MINT, IntAct, BioGRID, DIP, HPRD, and Reactome [4,13,18,27,52,57]) and combined to produce a network.

These resources contain information on interactions that have been experimentally detected using a variety of experimental techniques (for example, yeast 2-hybrid, tandem affinity purification). Further details on these techniques are provided in [48]. Only interactions that have a PubMed ID reference are included in this resource; the relevant PubMed ID for an interaction and its source can be viewed by clicking on an edge in a network.

Proteins were mapped to their genes, allowing the data from different databases to be combined using a common identifier. This allowed us to resolve ambiguous mapping between databases. Direct protein-protein interactions (such as interactions curated from the literature or found using yeast 2-hybrid screens) are displayed in the network as solid lines. Indirect associations (such as associations inferred through co-complex membership) are displayed as broken lines.

The focus of PainNetworks is on experimentally defined interactions. However, we also provide the user with the ability to add predicted interactions from the STRING resource [58], v 9.0. Only high-confidence predictions (ie, with a STRING score >800) are included. More details on the STRING score can be found at the resource itself, http://www.string-db.org, and in [43]. We also provide predicted protein interaction data derived from homology-inferred inheritance of interactions between orthologues (HIP-PO-C, HIPPO-DB; see below for details). As with many features of the Web site, this option is not enabled by default, but can be selected by the user and may be valuable for poorly annotated, less-well-studied genes for which few interactions are known, or for genes that cannot be assayed easily using experimental methods.

Interaction data are currently available for humans, rats, and mice. Therefore, the site has 3 potential species-centric views. Each view constructs the network using the interaction data available for that species. In addition, direct physical interactions detected in another species are transferred to that network, as described below.

2.3. Transferring annotation between species

The known pain genes deposited in the Pain Genes Database are from mice [34]. The protein-protein interaction (PPI) and Online Mendelian Inheritance in Man (OMIM) data are largely from humans; though some PPI data come from other species. The pain-related gene expression microarray data are produced using rat models, and the tissue-specific expression data are from humans. To integrate the data so they can be viewed together on the Web site, annotations must be transferred to one of the species, and therefore it is necessary to map genes between the species. PainNetworks currently allow the data to be transferred to mice, rats, or humans. Thus, there are 3 different possible views of the site: human-centric, mouse-centric, or rat-centric.

The Ensembl Compara method for identifying orthologues is used to map genes between species [62]. In order to transfer PPIs between species, a method similar to the interlogs procedure [41] is used; we name this procedure, “Homologous Inheritance of Protein-Protein Interactions by Orthology – Conservative” (HIP-PO-C). Given an interaction between a pair of genes in one species, if both of these genes have unambiguous orthologues in another species, this interaction can be inherited. This method has been shown to work well in previous studies [51]. Inherited interactions will be visualised as a green edge in the network, as opposed to blue for interactions not inherited from other species. For example, when looking at the mouse interaction network (ie, using the mouse-centric view), inherited interactions found using human-based methodology would be coloured green; interactions discovered using mouse-based methodology would be coloured blue.

A second method for protein interaction inheritance is also included, named HIPPO-DB (Homologous Inheritance of Protein-Protein interactions by Orthology – Domain Based). This method uses a combination of CATH and Pfam domain classifications [19,37,50] to identify the likely subset of domain interactions underpinning a given protein interaction. This method provides increased confidence in protein interactions and indicates the domain combination that may be underpinning the interaction.

Predicted gene associations taken from the STRING database [58] are available for the human-centric view, and are not transferred between species. Table 1 shows the total numbers of genes and interactions in the respective networks.

2.4. Pain-related gene lists

The user can input a set of query genes in order to produce a network. However, the Web site also contains lists of pain-related genes obtained from public resources: 1) The OMIM database, which is an online database of diseases and related genes (containing traditional “Mendelian” diseases and more complex polygenic diseases). 2) The Pain Genes Database (PGD), maintained by Jeffrey
Mogil’s Pain Genetics Lab. This resource contains genes that, according to a published study, lead to a change in pain-related behaviour when knocked out in mice [34]. 3) The Algynomics Pain Research Panel. This is a custom microarray that allows the user to assay single nucleotide polymorphisms in a number of genes that may have some involvement in pain-related processes: http://www.algynomics.com/pain-research-panel.html. Table 1 shows the number of pain-related genes present in the various interaction networks for the different organisms, and the number of interactions between them.

PainNetworks will be updated regularly to capture any new microarray, RNA-seq experiments, additional genes placed in the PGD, OMIM, or other published pain-based resources, and for future versions of the Algynomics Pain Research Panel. The user can use genes from these resources to annotate their network, or to identify clusters in the network with an over-representation of pain-related genes.

Note that there is not an exact match between the number of pain-related genes in PainNetworks and PGD or OMIM. This is due to ambiguities between databases with regard to gene identifiers: a given gene in one resource might not have an equivalent entry in another, or might have more than one.

Currently, PainNetworks contains, for humans, 10,297 genes with Ensembl IDs, with 40,832 direct interactions between them. The respective numbers for mice are 2104 and 2716, and for rats, 545 and 438. Most efforts to map interactions and associations between genes/proteins focussed on humans; this explains the discrepancy in numbers between the different organisms.

2.5. Expression data and analysis

The current datasets used by the Web site consist of gene expression data taken from microarray analyses of 4 animal pain models, performed by researchers in the London Pain Consortium and their collaborators, using rat as the model organism. Three of these are published [16,40].

Details on experimental design have either been published or are provided on the Web site and are summarized in Table 2. Further details of all currently available experiments can be viewed by clicking on the experiments tab. Lists of genes and fold changes can be downloaded for each experiment in a variety of formats, or browsed online. More datasets will be added in time, and we invite members of the pain community to contact us to upload datasets to the site. Users can also view their own expression datasets through the site locally (client side), without having to send any data via the Web. Instructions on how to do this are given on the site, along with an example file.

All microarray-derived gene expression datasets were produced using Affymetrix GeneChip technology. The resultant data were preprocessed using the RMA package [9,26]. Differential gene expression was calculated using limma (Linear Models for Microarray Data) [54], which produced a P-value for each gene probed by the array. False discovery rate was estimated from these P-values using the Benjamini–Hochberg method [7]. Fold changes in gene expression were computed between case samples (ie, animals in which pain-related behaviour was induced) and control samples (ie, the corresponding control for the model). A false discovery rate of 0.1 was used to decide if a gene was differentially expressed between case and control samples.

2.6. Filtering to obtain context specific networks

Networks can be filtered using various methods to make the networks more context-specific. For example, a list of Gene Ontology (GO) category names can be entered, and only genes annotated with one or more of these terms would be retained in the network. In this way, networks can be filtered to only display genes involved in specific biological processes (eg, GO:0006954, inflammatory response), molecular functions (eg, GO:0005244, voltage-gated ion channel activity), or expressed in given cellular components (eg, GO:0045202, synapse).

The user can also filter the network to retain genes differentially expressed in one or more of the gene expression microarray experiments. They may also filter the network so that it retains interactions where both genes show significant differential expression (in the same direction) in user-specified experiments. Such filtering can greatly reduce the size of the interaction network, making it easier to examine and interpret.

| Name | Experimental details | Species | Platform | Contributor | Reference |
|------|----------------------|---------|----------|-------------|-----------|
| HIV  | Rats received concomitant delivery of dDC (Zalcitabine) and HIV-gp120 to the sciatic nerve. Gene expression in L5-DRG at 15 days was compared to vehicle-receiving controls | Rat | Affymetrix Rat Genom 230 2.0 | Klio Maratou/ Andrew Rice | [40] |
| SNT-L5 | Spinal nerve transection (SNT) was performed on the rat L5 spinal nerve. Gene expression in the L5-DRG (dorsal root ganglion) at 15 days was compared to that of the L5-DRG for rats that received sham surgery | Rat | Affymetrix Rat Genom 230 2.0 array | Klio Maratou/ Andrew Rice | [40] |
| NGF  | Rats received recombinant nerve growth factor (NGF) for seven days through a pump placed subcutaneously beneath the scapulae. Gene expression in the L4 and L5-DRG was compared to vehicle receiving controls | Rat | Affymetrix Rat Genom 230 2.0 array | Stephen McMahon | Unpublished |
| SNI  | Spared nerve injury (SNI) was performed on tibial and peroneal branches of the sciatic nerve of rats. Gene expression in the L4 and L5-DRG was compared to sham surgery receiving control animals | Rat | Affymetrix Rat Genom 230 2.0 array | David Vega-Avelaira/Maria Fitzgerald | [61] |
Networks can also be filtered using tissue-specific expression data, in order to produce networks that only contain genes expressed in a given tissue/anatomical structure. These data are taken from the Human Protein Atlas Project [60]. This project combines Antibody-Based Proteomic and tissue-microarray based techniques in order to profile protein expression across various cells and tissues. For example, the user is able to filter the network so that it only contains genes expressed in a given cell or tissue, such as peripheral nerve tissue.

### 2.7. Clustering the network

Since there are large amounts of experimental data available for many genes, it is very easy to generate large networks (ie, >100 nodes), which are difficult to interpret. Therefore, PainNetworks allows the user to cluster the network in order to break it up into smaller groups of related genes. The advantages of clustering are 2-fold: to help visualisation and to find modules enriched in specific functions. We use established graph-clustering algorithms to find clusters in the network. The aim of these algorithms is to find groups in the network with a high number of interactions between group members, but a small number of interactions between members of the group and the other genes in the network outside the group.

Two graph-clustering algorithms are implemented: Louvain clustering [8] and Links clustering [2]. Louvain looks for distinct clusters of genes in the network, not allowing a given gene to belong to more than one network. Links clustering groups the edges of the network, that is, the interactions. This second approach has the potential advantage that it allows communities to be found with overlapping nodes; the same gene can belong to different clusters in the network. This fits in well with the functional pleiotropy of many genes.

Once clusters have been built, they can be ranked according to their size or the number of pain-related genes they contain. Enriched GO categories for each cluster are also displayed. Enrichment is performed using a method similar to that employed by the DAVID functional enrichment tool [21]; for each GO category, their size or the number of pain-related genes they contain. Enriched GO categories for each cluster are also displayed. Enrichment is performed using a method similar to that employed by the DAVID functional enrichment tool [21]; for each GO category, the proportion of category members in the cluster is compared to the proportion of category members outside of the cluster using Fisher’s exact test. Categories are deemed significant if they have a P-value of <0.05 following Bonferroni correction for multiple testing, and if the proportion of category members inside the cluster is at least 3 times higher than the proportion outside the cluster.

### 2.8. Additional gene annotation data

Additional functional annotation can be obtained for the genes in the network by clicking on the relevant node. This will bring up a box containing details of the GO categories to which the gene belongs, drugs that can target the protein encoded by the gene (taken from DrugBank [28]), and any potential domains the protein contains (InterPro [24], Pfam [50]). Clicking on the network selection tab allows the user to highlight the genes in the network that have drug-binding annotations, predicted coiled-coil regions (ncoils), predicted transmembrane regions (TMHMM [32]), and signal peptide sequence features (SignalP [49]).

### 2.9. Pain gene enrichment score

The ratio of interactions that a gene has with known pain genes, compared to all interactors, multiplied by 100, is deemed the Pain Gene Enrichment (PGE) score. It is viewable in the table of network genes found in the “Network Genes” tab of the Web site. In order to contextualise this score, a “percentile” value is given in the same table. This value represents the percentage of genes in the resource that have a higher PGE score than a given gene. For example, for a gene with a PGE score of 90: if there were a total of 10,000 genes in the resource and 100 of these genes had a PGE-score higher than 90, the percentile for this gene would be 10.

### 2.10. Ranking a set of query genes by their probability to be pain associated

As well as providing a PGE score, we provide a ranking for query lists of genes, for example, lists of differentially expressed genes identified by a microarray experiment. This ranking is generated by the GeneMANIA algorithm, a publicly available method that is well established and shown to be effective in analysing protein interactions.

| Use | Gene(s)/experiment involved | Description | Link |
|-----|-----------------------------|-------------|------|
| Find out more about a gene of interest | NRG1 (Neuregulin 1) | Enter NRG1 into the search bar, examine the resultant network | http://www.youtube.com/watch?v=5VnU72V-YLa6 |
| Use the results of a gene expression microarray experiment to highlight differentially expressed genes in a network | ATF3 (Cyclic AMP-dependent transcription factor) | Enter ATF3 and filter results by direct physical interactions. The results of a pain-related microarray gene expression experiment can then be used to highlight nodes in the network | http://www.youtube.com/watch?v=A6e7OyK58CM |
| Look at predicted associations for a given gene for which no known physical interactions exist | VGF (VGF nerve growth factor inducible) | VGF is used as a query gene, but no interactions are found using default search parameters. However, the user can include predicted association in their query – this will return a network of VGF and its predicted associations | http://www.youtube.com/watch?v=Z0ix7uahJdg |
| Make a large network from a group of related genes and cluster in order to look for subsets of highly connected genes | OPRD1, OPRK1, OPRM1 and ORL1 (opioid receptors) | These genes are used to query the site and generate the network. This resulting network is then clustered, these clusters are then sorted by the number of pain-related genes they contain | http://www.youtube.com/watch?v=4RdFMBojego |
| Search for experimental interactions between a large list of genes identified as differentially expressed from a microarray gene expression experiment | SNT (spinal nerve transaction model): LS vs Sham | In this search the top six genes from the SNT experiment (largest fold changes amongst the significantly differentially expressed genes) were used as query genes. The resultant network shows the interactions between these genes. Genes are: VIP, AR1, CSR1, POSTN, TSLP, CLEC7A | http://www.youtube.com/watch?v=1Y9a8RdJKEs |
| Look for interaction partners for a given gene that are druggable using approved drugs | SCN10A (sodium channel, voltage-gated, type X, alpha subunit) | Network produced using SCN10A as a query genes. Annotation used to show that many of the interactors are druggable, and that many are predicted to be membrane bound proteins | http://www.youtube.com/watch?v=1Mta9GqL014 |
network topologies to determine the likelihood that a query gene is functionally associated with a set of genes in the network known to be involved in a particular biological process. We use the default GO biological process-weighted default kernel combination, and apply label affinity propagation as described in [63] to derive scores for ranking. The ranking can be viewed as an extra column in the Network Genes tab of the site.

2.11. Implementation

The Web site was implemented using Python and CSS. Cytoscape Web [39], a Flash-based plug-in, available from http://cytoscapeweb.cytoscape.org/, was used to visualise the networks.

In addition to customising the content of their network, the user can move the protein components around on the screen interactively, allowing them to focus on whatever part of the network they wish to examine further.

A detailed reference manual, video tutorials, and several interactive search tutorials and are also available from the site.

3. Results

3.1. Overview of the site

Here we present a publicly available resource, PainNetworks. The resource combines data from various sources in order to allow the user to view pain-related biological networks. These networks can be used for various purposes: to learn more about a gene of interest, to further analyse a group of genes that have a putative role in pain, and to look for novel drug targets. A further benefit
is the ability to submit gene lists obtained by high-throughput studies in order to use the networks to identify a more focused subset of genes likely to be associated with pain-related processes.

Therefore, the purpose of the site ranges from educational through hypothesis generation to the interpretation of large data-sets obtained in high-throughput functional genomics studies.

In this section, we will demonstrate the power of the resource to recognize known pathways implicated in pain and we will describe some example usage scenarios for the site. We have tried to make the site as intuitive as possible, however, there is a large amount of functionality and a number of ways to view the networks and export the data. Although we will describe some key features of the site here, more extensive details are provided in the reference manual: http://www.PainNetworks.org/tutorials/RefMan.pdf. This manual describes how to use all of the features of the site, illustrated with screenshots. There are also video tutorials, available at http://www.youtube.com/channel/UCfl06Zr51-Sy3Bq9Neca_w0A/videos?view=0, which present potential usage scenarios. It is also possible to search the site using example queries presented on the site itself, easily accessible from the panel to the left of the network panel on the site’s homepage.

A nonexhaustive summary of potential uses for the site is also presented in Table 3. Below we demonstrate the performance of the resource in reproducing information on known biological pathways reported as pathway figures in the literature. We do this by reporting the proportion of interactions between the pathway genes, shown in published pathway figures that are also highlighted on the PainNetworks Web site.

Given that the resource uses experimental data on protein interactions (both direct and indirect), the PainNetworks Web site is unlikely to match these diagrams perfectly. Pathway diagrams taken from the literature are necessarily reductive models of the real processing that occurs within the cell, biased towards the system being considered. It is therefore possible that some of the proteins in the pathway are also involved in other processes [36,53]. Furthermore, with regards to any missing connections in the networks constructed by PainNetworks, the site uses protein interaction data stored in public databases, and these resources do not currently capture all the protein interactions reported in the literature. However, they are the most comprehensive public source of such data.

Despite these caveats, we explored how well PainNetworks could reproduce two such canonical pathway diagrams, shown in Figs. 1 and 2, taken from http://pathwaymaps.com/maps/652/ and [46]. Fig. 1 shows the TrkA-NGF signalling pathway; PainNetworks built the network shown in this figure using the pathway members as query genes, filtering the network to include only physical interactions and only considering interactions between the query genes. As we see in Fig. 1, PainNetworks was able to recover 76% of the interactions in the original pathway. Moreover, PainNetworks displays some additional interactions between the query genes that are not shown in the canonical pathway figure (6 interactions, 5 of which involve the same gene, PIK3R1). Since these interactions are associated with experiments reported in one of the public databases integrated by PainNetworks, these links may suggest novel associations between genes in the network that had not previously been considered.

Fig. 2 shows another example in which PainNetworks was used to reproduce the NRG1-ErbB signalling pathways, which has been...
implicated in pain [12]. A figure taken from a recent review [46] was used to make the comparison. PainNetworks was queried using all the members of the pathway (NRG1, ErbB2 and ErbB3): the resource is able to find all of the interactions within the

Fig. 4. Looking for druggable proteins associated with opioid receptors. Building the network using all 4 opioid receptors and their interactors, as shown in panel (A), results in a large network. Different ways of reducing the network to make it more interpretable are presented and explained further in the text: (B) and (C) show the results of clustering the original network to find groups of genes that interact with each other more frequently than they interact with genes outside of the group. (C) Genes highlighted in red represent genes with known drug targets, according to DrugBank [27]. This annotation is obtained from the “Network selection” tab. (D) Shows the network with indirect interactions removed, leaving only direct physical interactions.
canonical pathway diagram (Fig. 2C), plus many more interactions with other genes from outside the pathway (Fig. 2A). These may be good targets for putative novel pain genes implicated in this pathway.

3.3. Using the information in PainNetworks to suggest putative novel pain genes

Fig. 3 shows the number of interactions for each gene in the resource with at least one known interaction (i.e., network degree, x-axis), plotted against the number of these interactions that are pain related (y-axis).

The red points on the figure represent pain-related genes, taken from the PGD; the white nodes show the genes not reported in PGD. Three further analyses that refine the network are shown in the bottom half of the figure. Fig. 4B shows one of the clusters found by clustering the original network by the Louvain clustering, which partitions the genes into separate, nonoverlapping clusters. We see that three of the opioid receptors are placed into a cluster with beta-2-adrenergic receptor (ADRB2), pro-opiomelanocortin

have been transferred to humans, using the strategies described in Methods.

3.4. Using PainNetworks to find out more about genes of interest and their interactors

3.4.1. Opioid receptors

In Fig. 4, we show some network analyses of the 4 different human opioid receptors – the δ opioid receptor, κ opioid receptor, μ opioid receptor, and nociceptin receptor – known to be implicated in pain. Fig. 4A shows the results of a simple search of the Web site using the genes that encode these receptors: OPRD1, OPRK1, OPRM1, and ORL1, in the human-centric view. This query results in a very large network, too large to interpret visually. Therefore, only interactions that involve the query genes are shown in the network; interactions between the interactors of the query genes are not shown.

The grey nodes show pain-associated genes, taken from the PGD; the white nodes show the genes not reported in PGD. Three further analyses that refine the network are shown in the bottom half of the figure. Fig. 4B shows one of the clusters found by clustering the original network by the Louvain clustering, which partitions the genes into separate, nonoverlapping clusters. We see that three of the opioid receptors are placed into a cluster with beta-2-adrenergic receptor (ADRB2), pro-opiomelanocortin
(POMC), and a number of other genes from the pain genes database. Fig. 4B shows this same network cluster, but with druggable targets highlighted in red, generated using the “Network selection” tab. Fig. 4D shows the results of filtering the original network so that only direct physical interactions (not indirect interactions based on co-complex membership) are shown. Because the network produced by this method is small, the interactions between the nonopioid genes are displayed. Investigation of these networks can further elucidate pathways in which opioid compounds operate by revealing novel genes highly connected to these receptors. Some of these novel genes may be druggable proteins that might influence opioid receptor activity.

3.4.2. Angiotensin II receptor, type 2 gene analysis

Fig. 5 shows how PainNetworks can be used to find out more about the angiotensin II receptor, type 2 gene (AGTR2). A recent phase II clinical trial showed that an inhibitor of this receptor
(EMA 401) is effective in reducing pain reported by patients with postherpetic neuralgia [42]. Fig. 5A shows AGTR2 and its network interactions. A large number of the interactors of AGTR2 are pain-associated genes: 76 of its 284 interactors are in the pain-related gene list; this would give it a PGE score of 27. A PGE score of 27 is higher than the score associated with 97% of the genes in Pain-Networks and within the range of PGE-scores obtained by other known pain genes.

Fig. 5B shows the network formed by considering direct physical interactions involving AGTR2, along with related proteins AGTR1 (angiotensin II receptor, type 1) and AGTRAP (angiotensin II receptor-associated protein).

Genes that change in expression following spinal nerve transection (quantified using gene expression microarrays) are highlighted in Fig. 6. We notice that many of the genes in this network are differentially expressed, suggesting a potential role of this network in neuropathic pain. Thus, PainNetworks could be used as a starting point to further elucidate the mechanisms by which AGTR2 leads to pain relief.

3.4.3. Cyclic AMP-dependent transcription factor

Cyclic AMP-dependent transcription factor (ATF3) has been shown to change in expression in a number of microarray experiments profiling gene expression in animal models of pain [33]. Fig. 6 shows ATF3 together with its network associations/interactions (A). Although ATF3 shows differential expression in a number of different neuropathic pain models, relatively few of its interaction partners are known to be pain associated. We see only 2 genes from the pain-associated gene list interacting with ATF3, NKRB1, and HDAC4. However, if we click on the “Fold Change” link for the spinal nerve transection, transected L5 nerve vs sham experiment, we see that many of ATF3’s network interaction partners are differentially expressed in this experiment (Fig. 6B). As we would expect, many of these differentially expressed interaction partners are also transcription factors. This suggests that spinal nerve transection may be leading to changes in gene expression through the actions of a network of transcription factors during the induction phase of neuropathic pain. This supports the hypothesis that ATF3 is indeed associated with the response to nerve injury, and these differentially expressed genes could represent further targets for experimental investigation.

3.5. Using PainNetworks to filter putative pain genes identified by a microarray experiment

PainNetworks can also be used to analyse the results of a functional genomics experiment. Fig. 7 shows the network returned by querying the site using the differentially expressed genes with the largest fold changes from the spinal nerve transection experiment described above (Fig. 7A). We see that the network returned is quite large; however, by zooming in and inspecting different parts of the network, we see that there is a highly connected module that contains a large number of pain genes (Fig. 7D).

3.6. Conclusions

To conclude, we have developed a resource for the pain community that allows a pain scientist to view protein networks based on their gene(s) of interest. PainNetworks combines publicly available data on protein interactions, gene associations, gene expression data, and protein annotations. The value of the resource is increased considerably by including information on known pain genes given in the PGD of Mogil et al. [34], and by presenting information on putative gene functions reported in GO [5]. Networks provided by our resource allow a pain researcher to consider the function of their genes based on the functions of their interactors.
Our hope is that this novel view of the data will lead to hypothesis generation and a fuller understanding of the biological processes involved in pain. A particularly valuable aspect of the resource is the ability to interpret large gene datasets returned from high throughput functional genomics experiments. Users can obtain information on which genes in their list of putative novel targets are more likely to be pain associated because they have more interactions with known pain genes. The resource will expand over time: we plan to add human genome-wide association study data, data from more transcriptional profiling experiments, and to increase the number of species covered. We expect that PainNetworks will become a valuable asset to the pain community.

Conflict of interest statement

None of the authors have financial or other relationships that might lead to a conflict of interest in the study.

Acknowledgements

We acknowledge the Wellcome Trust for a strategic award to the London Pain Consortium. D.L.H.B., J.R.P., S.B.M. and C.O. are members of the London Pain Consortium. A.A.-M., S.B.M., D.L.H.B., J.L., and C.O. are part of the Europain Collaboration, which has received support from the Innovative Medicines Initiative Joint Undertaking, under grant agreement no 115007, resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (FP7/2007–2013) and EFPIA companies’ in-kind contribution. D.L.H.B. is a senior Wellcome clinical scientist (095698/z/11/z).

References

[1] Abrahamsen B, Zhao J, Asante CO, Cendan CM, Marsh S, Martinez-Barbera JP, Nassar MA, Dickson AH, Wood JN. The cell and molecular origin of experimental, cold, and inflammatory pain. Science 2008;321:702–5.

[2] Ahn Y-J, Bagrow JP, Lehmann S. Link communities reveal multiscale complexity in networks. Nature 2010;466:761–4.

[3] Antunes-Martins A, Perkins JR, Lees J, Hildebrandt T, Orenco G, Bennett DLJ. Systems biology approaches to finding novel pain mediators. Wiley Interdiscip Rev Syst Biol Med 2012.

[4] Aragón R, Achuthan P, Alam-Farague Y, Armean I, Bridge A, Derow C, Feuermann M, Ghanbarian AT, Lambiotte R, Khadake J, Kasarskis A, Lewis S, Mateje JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25–9.

[5] Baitaluk M. System biology of gene regulation. Methods Mol Biol 2009;569:55–87.

[6] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B (Methodological) 1995;57:289–300.

[7] Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of complex in networks. Nature 2010;466:761–4.

[8] Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech: Theor Exp 2008;10:P10008.

[9] Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19:521–34.

[10] Bonnici N, Berenguer A, Díez I, Reina O, Medina I, Dapoz J, Moreno V, Pujana MA. Biological properties, properties and molecular wiring diagrams of candidate low-penetration breast cancer susceptibility genes. BMC Med Genomics 2008;1:62.

[11] Calvano SE, Xiao W, Richards DR, Feliciano RM, Baker HV, Cho RJ, Chen RO, Brownstein BH, Cobb JP, Tschoeke SK, Miller-Graziano C, Moldawer LL, Mindrinos MN, Davis RW, Tempkin RS, Lowry SF. Inflamm and Host Response to Injury Large Scale Collab. Res. Program. A network-based analysis of systemic inflammation in humans. Nature 2005;437:1032–7.

[12] Calvo M, Zhu N, Tsantoulas C, Ma Z, Grist J, Loeb JA, Bennett DLJ. Neuregulin-erbB signaling promotes microglial proliferation and chemotaxis contributing to microgliosis and pain after peripheral nerve injury. J Neurosci 2010;30:5437–40.

[13] Ceol A, Chatti Aryamonti A, Licata L, Peluso D, Briganti L, Perfetto L, Castagnoli L, Cesareni G. MINT, the molecular interaction database: 2009 update. Nucleic Acids Res 2010;38:D553–9.

[14] Costigan M, Belfor K, Karchewski L, Griffin RS, D’Souza D, Allchorne A, Sitarski J, Mannion JW, Pratt RE, Wolff CJ. Replicate high-density rat genome oligonucleotide microarrays reveal hundreds of regulated genes in the dorsal root ganglion after peripheral nerve injury. BMC Neurosci 2002;13:16:1–4.

[15] Costigan M, Belfor K, Griffin RS, D’Souza D, Allchorne A, Sitarski J, Mannion JW, Pratt RE, Wolff CJ. Replicate high-density rat genome oligonucleotide microarrays reveal hundreds of regulated genes in the dorsal root ganglion after peripheral nerve injury. BMC Neurosci 2002;13:16:1–4.

[16] Croft D, O’Kelly G, Wu C, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalavslakaya I, Mahajan S, May B, Nedge W, Schmidt E, Shamsosky Y, Yung C, Birney E, Hermjakob H, D’Esteachio P, Stein L. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 2011;39:D961–7.

[17] Cuff AL, Sillitoe I, Lewis T, Clegg J, Rentzsch R, Furniss C, Pellegrini-Calace M, Jorg D, Thornton J, Orenco CA. Extending CATH: extending coverage of the protein structure universe and linking structure with function. Nucleic Acids Res 2011;39:D420–6.

[18] Davies JM, Calvo M, Perkins JR, Paterson JJ, Kewenter W, Hobbs C, Kaan TKY, Orenco C, Bennett DLJ, McManhom SB. CLUMES microarray irradiation-induced pain. Sci Transl Med 2011;3:90ra60.

[19] Dennis Jr G, Sherman BT, Hgosac DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Analysis, Visualization, and Integrated Discovery. Genome Biol 2003;4:4–P.

[20] Díaz JS, Rincón RJ, Barceló J, Puértolas J-S, Zhong FY, Levitin DJ, Mogil JS. Patterns of pain: triangulation: bridging linkage and molecular-network information for identifying candidate genes in Alzheimer’s disease. Proc Natl Acad Sci USA 2009;106:15148–53.

[21] Foulkes T, Wood JN. Pain genes. PLoS Genet 2008;4:e1000006.

[22] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[23] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[24] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[25] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[26] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[27] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[28] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[29] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[30] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[31] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[32] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[33] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[34] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.

[35] Frome HN, Wood JN. The Pain Genes Database: an update. PLoS Genet 2008;4:e1000086.
Lees J, Yeats C, Perkins J, Sillito I, Rentzas R, Dessailly BH, Orenco C. Gene3D: a domain-based resource for comparative genomics, functional annotation and protein network analysis. Nucleic Acids Res 2012;40:D465–71.

Liu M, Liberton A, Kong SW, Lai WR, Park PJ, Kohane IS, Kasif S. Network-based analysis of affected biological processes in type 2 diabetes models. PLoS Genet 2007;3:e96.

Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD. Cytoscape Web: an interactive web-based network browser. Bioinformatics 2010;26:2347–8.

Maratou K, Wallace VJC, Hasnie FS, Okuse K, Hosseini R, Jina N, Blackjord B, Phey T, Oreno G, Dickenson AH, McMahon SB, Rice ASC. Comparison of dorsal root ganglion gene expression in rat models of traumatic and HIV-associated neuropathic pain. Eur J Pain 2009;13:387–98.

Matthews LR, Vaglio P, Reboul J, Ge H, Davis BP, Garrels J, Vincent S, Vidal M. Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs”. Genome Res 2001;11:2120–6.

McCarthy TD, Desem N, Kitson G, Rice ASC, Dworkin BH, Bountra C, Anand P, Raff M. Clinical development of EM401: an angiotensin II type 2 receptor antagonist as a potential new treatment for neuropathic pain. In: 4th World Congress on Pain, International Association for the Study of Pain (IASP) 2012. Available: http://www.iasp-pain.org/Milan/Program/Milan2012_Program.pdf.

Moring C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglerini M, Joufrré N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res 2005;33:D433–7.

Mogil JS. Pain genetics: past, present and future. Trends Genet 2012;28:258–66.

Moran LB, Graeber MB. Towards a pathway definition of Parkinson’s disease: a complex disorder with links to cancer, diabetes and inflammation. Neurogenetics 2008;9:1–13.

Newburn J, Birchmeier C. Nrg1/ErbB signaling networks in Schwann cell and myelination. Semin Cell Dev Biol 2010;21:922–8.

Nibbe RK, Markowitz S, Myeroff L, Ewing R, Chance MR. Discovery and scoring of protein interaction subnetworks discriminative of late stage human colon cancer. Mol Cell Proteomics 2009;8:827–45.

Perkins JR, Diboun I, Dessailly BH, Lees JG, Orengo C. Transient protein-protein interactions: structural, functional, and network properties. Structure 2010;18:1233–43.

Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Genome Res 2009;19:327–35.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD. The Pfam protein families database. Nucleic Acids Res 2010;38:D290–301.

Petroski KA, Otsuka K, Sato Y, Sato M, Bork P, Eddy SR, Bateman A, Finn RD. The Pfam database in 2013: protein families and domains. Nucleic Acids Res 2013;41:D296–305.

Petersen TN, Nielsen H, von Heijne G, Brunak S. The PHD web server: interactive web-based prediction of protein’s histidinyl and cysteinyl phosphorylation. Nucleic Acids Res 2009;37:W514–8.

Pfeuffer T, Oughtred R, Jansen VD, van Helden J, von Mering C. STRING: known and predicted protein-protein interactions from whole-genome phylogenetic analysis. Nucleic Acids Res 2012;40:D698–704.

Petersen TN, Nielsen H, Brunak S. Identifying potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs”. Genome Res 2001;11:2120–6.

Petroski KA, Otsuka K, Sato Y, Sato M, Bork P, Eddy SR, Bateman A, Finn RD. The Pfam database in 2013: protein families and domains. Nucleic Acids Res 2013;41:D296–305.

Pfeuffer T, Oughtred R, Jansen VD, van Helden J, von Mering C. STRING: known and predicted protein-protein interactions from whole-genome phylogenetic analysis. Nucleic Acids Res 2012;40:D698–704.

Petersen TN, Nielsen H, Brunak S. Identifying potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs”. Genome Res 2001;11:2120–6.

Petroski KA, Otsuka K, Sato Y, Sato M, Bork P, Eddy SR, Bateman A, Finn RD. The Pfam database in 2013: protein families and domains. Nucleic Acids Res 2013;41:D296–704.

Pfeuffer T, Oughtred R, Jansen VD, van Helden J, von Mering C. STRING: known and predicted protein-protein interactions from whole-genome phylogenetic analysis. Nucleic Acids Res 2012;40:D698–704.

Petersen TN, Nielsen H, Brunak S. Identifying potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs”. Genome Res 2001;11:2120–6.