Computational functional genomics-based approaches in analgesic drug discovery and repurposing

Catharina Lippmann^{‡,1}, Dario Kringel^{‡,2}, Alfred Ultsch³ & Jörn Lötsch^{*,1,2}

¹Fraunhofer Institute of Molecular Biology & Applied Ecology – Project Group Translational Medicine & Pharmacology (IME-TMP), Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

²Institute of Clinical Pharmacology, Goethe-University, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

³DataBionics Research Group, University of Marburg, Hans-Meerwein-Straße 6, 35032 Marburg, Germany

*Author for correspondence: Tel.: +49 69 6301 4589; Fax: +49 69 6301 7636; j.loetsch@em.uni-frankfurt.de

‡Authors contributed equally

Persistent pain is a major healthcare problem affecting a fifth of adults worldwide with still limited treatment options. The search for new analgesics increasingly includes the novel research area of functional genomics, which combines data derived from various processes related to DNA sequence, gene expression or protein function and uses advanced methods of data mining and knowledge discovery with the goal of understanding the relationship between the genome and the phenotype. Its use in drug discovery and repurposing for analgesic indications has so far been performed using knowledge discovery in gene function and drug target-related databases; next-generation sequencing; and functional proteomics-based approaches. Here, we discuss recent efforts in functional genomics-based approaches to analgesic drug discovery and repurposing and highlight the potential of computational functional genomics in this field including a demonstration of the workflow using a novel R library 'dbtORA'.

First draft submitted: 14 March 2018; Accepted for publication: 30 April 2018; Published online: 24 May 2018

Keywords: computational methods • data-driven research • data science • dbtORA R library • drug discovery • knowledge discovery • machine learning

The modern pharmacological treatment of pain started 150 years ago when Sertürner isolated morphine from opium [1]. 80 years later, Knorr and Filene synthesized the first synthetic drug, called antipyrine [2], and at the end of the 19th century, acetylsalicylic acid was patented by Hoffmann [3]. Since then, several further analgesics were developed; however, mainly belonging to the two major classes of opioids and anti-inflammatory nonopioids. In the late 20th century, the spectrum of pharmaceuticals against pain has been broadened by substances repurposed from antiepileptics or antidepressants such as gabapentin, pregabalin or amitriptyline, respectively. Within the current century, novel primary analgesic drugs are increasingly developed [4], for example, ziconotide, a peptide derived from the toxin of the fish-hunting snail *Conus magus* [5]. Despite this long history of pharmacological drug development, chronic pain has still a high prevalence [6,7] affecting up to 30% of the world's population [8]. With half a billion individuals in the EU, this projects to 80 million Europeans [9] and implies high socioeconomic costs [10]. The importance of pain treatment as a medical need is emphasized by considerably public funding activities such as by the EU [11].

Genetic approaches have been included in pain research since the late 60s of the last century [12], addressing genetic variants that modulate the perception of pain or the response to analgesics. Currently, >500 genes have been implicated in the modulation of pain-related phenotypes [13]. A particular research focus is set on genes that confer insensitivity to pain as a promising source of druggable targets in an analgesic context [14]. However, the complexity of pain is increasingly addressed in genome-wide approaches exceeding the study of single variants of genes. This has become possible due to recent technological advances in biomedical science enabling the access to the full human genome [15,16]. The novel research area of functional genomics investigates the biochemical, cellular and/or physiological properties of each and every gene product [17] with the goal of understanding the

relationship between the genome and the phenotype on a global genome-wide scale. These developments have made the research on pain and analgesic drugs a data-rich discipline that can be mastered only by making use of parallel developments in computer science including the establishment of knowledge bases, mining methods for big data, machine-learning methods and artificial intelligence [18]. These methods have been applied to drug discovery and repurposing for analgesic indications as summarized in the following.

Current evidence of functional genomics approaches at analgesic drug research

Functional genomics combines data derived from various processes related to DNA sequence, gene expression and protein function, such as coding and noncoding transcription, protein translation, protein-DNA, protein-RNA and protein-protein interactions, and together, these data are used to model interactive and dynamic networks that regulate gene expression, cell differentiation and cell-cycle progression [19]. A literature search for these items was conducted in PubMed at www.ncbi.nlm.nih.gov/pubmed on 16 November 2017. Specifically, the search string 'functional AND (genomic OR genomics) AND (pain OR *nocicepti* OR hyperalgesi* OR allodyn*) AND pharmacol* AND (genome OR transcriptome OR proteome OR metabolome OR interactome) NOT review' obtained 115 hits, of which most referred to genome-wide expression analyses or genetic association studies of small groups of genes with pain-related phenotypes. In this field, the use of functional genomics for the explicit purpose of drug discovery and repurposing, and in particular for analgesic indications without merely discussing general findings as offering such a possibility, has been pursued in a limited selection of major approaches (Figure 1), comprising knowledge discovery in gene function and drug target-related databases for drug discovery and repurposing; the use of next-generation sequencing (NGS) for analgesic drug research purposes; and functional proteomics-based approaches to analgesic target identification. These major scenarios of the use of functional genomics-based approaches for analgesic drug research make use of several freely available data bases and mining tools (Table 1) and will be summarized and exemplified in the following.

Knowledge discovery in gene function & drug target-related databases

Functional genomics of pain in analgesic drug development makes use of rapid advancements in the understanding of the role of the involved molecular pathways and genetic control mechanisms [18]. The biological roles of proteins can now be assigned to their coding genes thanks to concomitant developments in bioinformatics that facilitate the analysis of fundamental biological functions by providing databases or knowledge bases [47] and the software to exploit the stored content (Table 1). Hierarchical structured knowledge in the form of ontologies, has gained a scientific value for knowledge discovery in biomedical including pharmacological research.

Several different methods have been introduced for the implementation of *in silico* drug development and repurposing tasks [48]. All methods have in common that they analyze information about the interaction of drugs with molecular processes and relate this interaction with knowledge about the molecular pathophysiology of the disease of interest, which in the present context is a clinical setting of pain. A further common feature is the analysis of large and complex datasets for similarity between drug and disease-related features. This implies the use of so-called big data, which exceed by size and complexity the means of many classical data analytical methods, entering the environment of machine learning and artificial intelligence that have developed from computer science rather than from statistics. The approaches are data driven and, as machine learning for instance, focus at optimization and performance of algorithms without requiring prior assumptions about relationships between the variables [49,50]. Nevertheless, statistics is an integral component of at several steps of these analyses.

Methods of knowledge discovery enabling the use of computational functional genomics for drug repurposing are rapidly evolving and therefore, classification attempts [48] may be transitory. For the present overview, we observed four main concepts applied in knowledge discovery-based drug discovery and repurposing that differ with respect to the kind of drug-related genomic information that is associated to the disease-related genomic information. Specifically, different approaches focus on the analysis of the interactions of ligands (drugs) with their targets; the transcriptional effects of drugs; interactions of drugs with molecular network pathways; or employ the knowledge about the regulation of disease-relevant biological processes to draw conclusions about a possible involvement of drugs. However, while the majority of computational approaches have been increasingly applied to several different biomedical topics, only a subset has been applied so far within the limited field of analgesic drug discovery and repurposing, which will be outlined as follows.

Figure 1. Overview of computational functional genomics approaches selected for their use in a pain and analgesic drug research context. Following a PubMed search for ‘functional AND (genomic OR genomics) AND (pain OR *nocicepti* OR hyperalgesi* OR allodyni*) AND pharmacol* AND (genome OR transcriptome OR proteome OR metabolome OR interactome) NOT review’, the use of functional genomics for the explicit purpose of drug discovery and repurposing, and in particular for analgesic indications, indicated a limited selection of major approaches pursued in this field, comprising knowledge discovery in gene function and drug target-related databases for drug discovery and repurposing; the use of next-generation sequencing for analgesic drug research purposes; and functional proteomics-based approaches to analgesic target identification.

Approaches focusing on drug targets

Ligand-based approaches at drug discovery and repurposing search for suitable compounds are based on the concept that similar compounds with respect to the interaction with biological structures tend to have similar biological properties [48]. For example, the design of a multitarget inhibitor [51] of the heat shock protein 90 [52] interactom was facilitated through ligand-based analysis of known inhibitors available in the ChEMBL database (www.ebi.ac.uk/chembl/ [35]), using site binding similarity analyses. Following the identification of promising targets within the heat shock protein complex with key importance in cancer, compounds with shared activity against more than a single target provided suitable candidates that for selected compounds could be shown to agree with the results of analysis based on chemical structure similarity between drugs and targets [51]. The implication of multitarget inhibition in analgesic therapy has been proposed as a combination of low-dose inhibitors of inflammatory responses along the prostaglandin synthesis pathway, providing a possible approach to maximize anti-inflammatory effects while keeping side effects low [53]. However, published uses of ligand-based computational screening methods are not yet available.

Ligand-based approaches also include large-scale screening of chemical compounds for interaction with molecular targets, which in addition to new laboratory techniques such as reprogrammed nociceptor neurons from fibroblasts [54], make use of computational methods such as *in silico* chemogenomics, high-throughput docking, 3D bioinformatics target binding comparison and target and ligand annotation and ontologies [55]. For example, in a broad chemogenomic mining approach at target profiling of 767 chemicals against 684 targets of therapeutic relevance, drugs targeting G protein-coupled receptors were found as displaying most heterogeneous pharmaco-

Table 1. Examples of data sources and computational tools used for the data science approach to drug repurposing from functional genomics knowledge.

Site name	URL	Ref.
Gene names and functions		
AmiGO (search utility for GO)	http://amigo.geneontology.org/	[20]
Gene Ontology (GO)	www.geneontology.org/	[21]
HUGO Gene Nomenclature Committee	www.genenames.org/	[22]
Database for Annotation, Visualization and Integrated Discovery (DAVID)	https://david.ncifcrf.gov	[23]
NCBI gene index database	www.ncbi.nlm.nih.gov/gene/	
GeneCards	www.genecards.org	[24]
Human diseases		
Online Mendelian Inheritance in Man (OMIM®) database	www.ncbi.nlm.nih.gov/omim	
MalaCards: The human disease database	www.malacards.org	[25]
PainGenes database	www.jbldesign.com/jmogil/enter.html	[26]
The Human Pain Genetics Database (HPGDB)	https://humanpaingenetics.org/hpgdb/	[27]
Comparative Toxicogenomics Database (CTD)	http://ctd.mdibl.org	[28]
Biological pathways		
Pathway Commons	www.pathwaycommons.org	[29]
Kyoto Encyclopedia of Genes and Genomes (KEGG)	www.genome.jp/kegg/	[30]
Drugs, small molecules and/or (potential) targets		
DrugBank database	www.drugbank.ca	[31,32]
Connectivity map (CMap)	https://clue.io/cmap	[33]
DrugSig	http://biotechlab.fundan.edu.cn/database/drugsig/	[34]
Thomson Reuters Integrity database (nonfree)	https://integrity.thomson-pharma.com	
ChEMBL	www.ebi.ac.uk/chembl/	[35]
UniProtKB/Swiss-Prot	www.uniprot.org	[36–38]
Gene Expression Omnibus (GEO)	www.ncbi.nlm.nih.gov/geo	[39]
Software		
R software	http://CRAN.R-project.org/	[40]
Drug versus Disease (DvD)	www.ebi.ac.uk/saezrodriguez/DVD/	[41]
GeneTrail	http://genetrail.bioinf.uni-sb.de/	[42]
Bioconductor	http://bioconductor.org/	[43]
R package 'org.Hs.eg.db'	https://bioconductor.org/packages/org.Hs.eg.db/	[44]
R package 'GO.db'	http://bioconductor.org/packages/GO.db/	[45]
R package 'dbtORA'	https://github.com/IME-TMP-FFM/dbtORA	
Graphviz	www.graphviz.org	[46]
Literature		
PubMed	www.ncbi.nlm.nih.gov/pubmed/	
All resources are publicly available, most of them free of charge.		

logical profiles, which makes pleiotropic uses of drugs including drug repurposing an option, in particular, when addressing opioid, *N*-methyl-D-aspartate and 5-hydroxytryptamine type 3 receptors that qualify as among other indications as analgesics [56].

Furthermore, so-called computational molecular docking approaches aim at scoring the potential complementarity of small molecules to binding sites by analyzing the mutual agreement between the chemical structures [57]. This is used for both, drug discovery and repurposing, depending on the novelty of the compound in the pharmacological context. Virtual screenings are therefore performed either by docking a known drug into a large set of different target structures, or by docking a database-listed approved drug into the intended specific target. Molecular docking is therefore restricted to studies in which a 3D structure of the target is available through crystallography, nuclear magnetic resonance or comparative models [48]. Computational molecular docking has been successfully

used in an analgesic context in a study where the drugs listed in the DrugBank database [31,32] were docked into 35 crystal structures of MAPK14 [58]. This identified the *Bcr Abl* tyrosin kinase inhibitor nilotinib as a repurposing candidate for the treatment of inflammatory diseases. As in other studies using computational functional genomics methods (e.g., [13,59]), the authors provided support for their computational finding based on published evidence from experimental studies, which supported 31 of the top predicted interactions [58]. Furthermore, computational docking studies were successful in identifying the membrane protein cadherin-11, which has been implied in rheumatoid arthritis [60], as a target of the *COX-2* inhibitor celecoxib [61].

Approaches focusing on transcriptional effects

Transcriptional responses are used to create perturbation profiles of gene expressions treated by drugs versus untreated genes [62]. These drug discovery and repurposing approaches are based on the comparisons of differential gene expression profiles, for example, from GEO (Gene Expression Omnibus) [39] either induced by the drug or associated with disease. This requires the knowledge about the changes in gene expression induced by a drug, which can be accessed via comparing the expression profile of a cell before and after exposure to a drug; and the knowledge about changes in gene expression associated with the disease of interest. Two slightly differing approaches have been proposed comprising the ‘signature reversion’ and the ‘guilt by association’ methods.

Drug discovery via similarity of drug & disease-related differential expression patterns

The basis of the ‘signature reversion’ method is given by so-called signatures, which are defined by a set of genes combined with a pattern of expressional changes. As elaborated elsewhere [63], the leading idea is the comparison of these signatures associated with either the drug or the disease. The former can be detected experimentally by comparing the expression profile of a cell before and after exposure to a drug, which characterizes the effects of the active compounds on the transcriptional program. The latter is a result of commonly established biomedical research methods addressing the differential expression of sets of genes associated with pathophysiological processes. The repurposing candidate can be found by similarity analyses among the available signatures of drugs versus diseases, that is, if the signature of a drug corresponds to the disease signature, the drug qualifies as a therapeutic agent for this disease as it is expected to be able to reverse the disease-induced transcriptional changes. Signatures are available from publicly accessible databases such as the connectivity map (<https://clue.io/cmap> [33,64]) (Table 1), which provides a reference collection of gene-expression profiles from cultured human cells treated with bioactive small molecules, and additionally enables the comparison of signatures of small molecules used as drugs or other chemicals with signatures of pathophysiological processes or diseases [33]. This is complemented by the DrugSig database (<http://biotechlab.fundan.edu.cn/database/drugsig/> [34]) providing drug response-related microarray data annotated with drug and target-related information from connectivity map, further public databases and scientific literature. Recently another approach to expand the connectivity map has been proposed. A novel high-throughput gene expression profiling method, ‘L1000’, was described that allowed expanding the connectivity map in three dimensions comprising small molecule perturbations; genetic perturbations; and cell lines [65]. In the above approaches, disease-relevant information is usually needed, which can be queried from databases such as MalaCards (www.malacards.org [25]) that provide empirically found sets of genes regulated in diseases.

Drug repurposing via similarity of differential expression patterns to drugs known to work

In the ‘guilt-by-association’ approach, a repurposing candidate is sought via similarity analysis of a compound’s signature of differential gene expression with that of a drug known to work in the particular disease. The implementation of this idea is pursued via constructing a pathway–drug network representing known associations between drugs and pathways, in which network propagation is initialized from nodes selected on the basis of known disease pathways and known drugs associated with the phenotype of interest [66].

Approaches focusing on pathway-based network interactions

A network can be modeled as a connected graph with nodes and edges. Each node can represent either an individual molecular entity (e.g., a drug), its biological target, a modifier molecule within a biological process, or a target pathway, while an edge represents either a direct or indirect interaction between two connected nodes [48]. The idea of network-based drug repurposing is to take synergetic interactions of drugs, genes and biological processes into account. Each pathway can mention information on a series of biological events such as biochemical reactions, physical interactions, transcriptional responses and phosphorylation and enzyme catalysis [67]. This was applied, for

example, to identify drug repurposing candidates for Crohn's disease [67]. Specifically, a multilayer causal network was constructed that consisted of chains drawn from drugs to the disease, integrating heterogeneous expert-curated biological resources that unlike traditional 'one drug – one target – one disease' causal models systematically considered all possible chains connecting drugs to the disease via target- and gene-involved pathways [67]. Approved drugs and their targets were queried from the DrugBank database [31], target-involved pathways were obtained from the Pathway Commons database [29] and from the Kyoto Encyclopedia of Genes and Genomes (KEGG [30]). Downstream genes were queried from the Pathway Commons database, and disease-related genes were obtained from the Comparative Toxicogenomics database (CTD [28]), the latter also providing pairs of known therapeutic drug–disease relationships.

Based on the CauseNet method, the authors identified five drug repurposing candidates for Crohn's disease comprising anakinra, which is an IL-1 receptor antagonist-approved drug for the treatment of rheumatoid arthritis. Further network analysis showed that via binding at the IL-1 receptor, multiple pathways may be affected such as osteoclast differentiation and amoebiasis, which in turn may involve genes involved in the risk of Crohn's disease such as *NCF4* and *FN1*. Among the repurposing candidates figured also the mast cell stabilizing drug nedocromil, which the authors show to potentially also target *HSP90AA1* and *FPRI*, affecting multiple nucleotide-binding oligomerization domain-like receptor signaling pathways, and further affecting multiple genes with a role in Crohn's disease including *IL6*, *TNF*, *NLRP3*, *NOD2*, *FPR2* and *IL10*. Finally, the authors report the TNF- α antibodies adalimumab and prednisolone as repurposing candidates in the context of Crohn's disease [68]. The authors concluded that the computational method enables the discovery of new uses of existing drugs based on casual inference in a layered drug–target–pathway–gene–disease network. However, in the present context of analgesic drug research, no similar successes with pathway-based computational methods have been reported so far.

Approaches focusing on alterations in biological functions

Phenotypic drug discovery approaches try to address incompletely understood complexities of diseases [69] without relying on the knowledge about specific drug targets or hypotheses about their particular roles in the trait of interest. A focus on the relevant similarities between drugs and diseases differing from target-based strategies was set in the so-called process pharmacology [70], in which diseases are regarded as a result of alterations of the activity in one or several biological processes involved in the pathophysiology of the disease while the molecular drug targets serve merely as a link between drug and disease-relevant biological processes. Biological processes are defined according to the GeneOntology knowledge base (GO) as “a series of events or molecular functions with a defined beginning and end” [20]. The concept directly accommodates the therapeutic setting of treating a disease rather than a molecular structure, and provides a suitable basis for analgesic drug classification and repurposing.

R implementation of extended over-representation analysis: dbtORA

The 'process pharmacology' approach requires the identification of GO knowledge base-derived biological processes characterizing the biological roles of a particular set of genes. It will be used as an example of computational functional genomics research in analgesic drug discovery. As in contrast to other approaches the necessary software has not yet been described, the here used dbtORA R package, freely available at <https://github.com/IME-TMP-FFM/dbtORA>, is described in more detail in the Supplementary Materials.

In the present context of pain and analgesics research, a suitable set of genes was queried from the PainGenes database (www.jbldesign.com/jmogil/enter.html [26]). This mainly includes genes found in at least three independent studies in transgenic mice to contribute to the modulation of pain. Further genes were added because they were recognized as causing human hereditary diseases associated with extreme pain phenotypes or being targets of novel analgesics [18].

This resulted in a set of $n = 540$ genes identified empirically as relevant to pain. Their biological roles were queried from the GO (www.geneontology.org/) [21]. In the GO, the knowledge about the biological roles of genes is formulated using a controlled and clearly defined vocabulary of GO terms [71,72]. Particular biological roles of the $n = 540$ 'pain genes', among all human genes, were analyzed by means of overrepresentation analysis (ORA), which compared the occurrence of the specific set of genes annotated at particular GO terms with the expected occurrence of all human genes at these terms. The significance of a GO term associated with a particular gene was calculated by means of Fisher's exact test [73] with α correction for multiple testing according to Bonferroni [74]. A minimum number of five genes annotated to a term and a p-value threshold of 8.74×10^{-5} were set as limits; the latter was computed based on a preliminary ORA with a p-value threshold of 0.01, from which all p-values of terms

were submitted to a computed ABC analysis [75], which is an item categorization technique suitable to identify ‘the important few’ from a set of positive items.

The result of the ORA was a representation of the complete knowledge about the biological processes influenced by the set of pain genes as a polyhierarchy organized in a directed acyclic graph (DAG [76]). The complete highly complex DAG is shown in the [Supplementary Figure 1](#); however, to make the result comprehensible, functional abstraction [77], which partitions the graph into distinct representations of coherent parts of the complete knowledge (so called ‘functional areas’) that present the complete DAG at maximum coverage, certainty, information value and conciseness [77], was applied. The resulting 18 biological processes described the knowledge about the functional genomics of pain ([Figure 2](#)) in agreement with earlier analyses [13,18] as composed of intuitively pain-relevant biological processes such as ‘signaling’ (GO:0023052), ‘response to stimulus’ (GO:0050896), ‘behavior’ (GO:0007610), which included processes such as ‘learning or memory’ or ‘developmental process’ (GO:0032502), which pointed at neuronal plasticity via processes related to neuron or axon development, immune system process (GO:0002376) pointing at the known role of the immune system in persistent pain [78–81]. In addition, more general processes resulted such as ‘multicellular organismal process’ (GO:0032501) or ‘localization’ (GO:0051179), which mainly referred to ion and neurotransmitter transport, and ‘rhythmic process’ (GO:0048511), which has been recognized in various pain conditions [82–84].

Biological process-based analgesics drug classification

Combining several lines of functional biomedical genomics research with machine learning methods, ‘process pharmacology’ was demonstrated to be suitable framework for the classification of analgesic drugs [70]. Specifically, knowledge discovery in the DrugBank [31,32] and Gene Ontology [21] databases allowed the construction of a ‘drug target versus biological process’ matrix as a combination of drug versus genes and genes versus processes matrices (for details, see [Supplementary Materials](#)). The phenotypic functional genomics-based drug research approach was applied to $n = 79$ analgesics drugs queried from the DrugBank database, which were associated with $n = 102$ genes respectively molecular targets that in turn were annotated to $d = 928$ biological processes as reported previously [85]. Unsupervised machine learning was used to identify structures within the data space [50]. Specifically, each drug was represented as a vector in a $d = 928$ dimensional feature space of impacts strengths on processes. To explore this feature space, a projection of the high-dimensional information onto a 2D grid was used. For this purpose, a grid of artificial neurons on a self-organizing map was used previously [70,85].

Biological process-based analgesics drug repurposing

In the framework of ‘process pharmacology’, computational functional genomics approaches for analgesics drug repurposing may employ the association of drugs with biological processes. The search for repurposing candidates bases on machine-learned identification of similarities in the high-dimensional data space. While using a set of > 500 pain-related genes [13], appeared a too heterogeneous basis, a functionally more focused subset of pain genes was successfully used [14]. Specifically, genes causally involved in human insensitivity to pain are in a particular focus of analgesic drug discovery and repurposing. Substances targeting, for example, *SCN9A*, in which loss-of-function variants cause congenital inability to perceive pain [86], have reached clinical development [4]. Currently, $n = 20$ of such genes have been reported [14]. A functional genomics analysis proposed that among the functions of this set of genes, the processes related to nervous system development and to ceramide and sphingosine signaling pathways are particularly important [14]. This is in line with earlier suggestions to use these pathways as therapeutic target in pain. Following establishment of the functional genomics of hereditary insensitivity to pain, the biological processes were used for a similarity analysis with the functional genomics of $n = 4834$ database-queried drugs. Using unsupervised machine learning implemented as emergent self-organizing maps, a cluster of $n = 22$ drugs emerged sharing important functional genomic features with hereditary insensitivity to pain. As discussed previously [87], for more than the half of the members of this cluster, evidence about an implication in pain could be found in the literature. While it appears unlikely that this will be true for any random set of drugs, a statistically significant difference of the findings with positive hits expected by chance was not tested in that proof-of-concept assessment. By contrast, when using the present method to identify pain-relevant genes [13], there, of a set of 34 hits, 33 hits could be supported by empirical evidence whereas for a random set of 34 genes, only three hits were obtained. This suggests that the method may be suitable for analgesics drug repurposing.

Figure 2. Information graphics about the computational functional genomics of pain. Graphical representation of the main biological functions exerted by the set of $n = 540$ genes that have been implicated in the modulation of pain, grouped into 18 primary 'functional areas' arranged circularly in the figure with names given horizontally at the periphery of the circle. Each of these functional areas comprises groups of biological processes in which the genes annotated to the functional area are involved at a more detailed and specific level. All processes are based on the statistically significant results of a GeneOntology knowledge base over-representation analysis with subsequent algorithmic identification of functional areas, according to their informational importance. The areas qualify as headlines representing particular aspects (taxonomies) of the complete polyhierarchy at maximum coverage, certainty, information value and conciseness [77]. The figure has been created using the R software package (version 3.4.3 for Linux; <http://CRAN.R-project.org/>) [40].

Next-generation sequencing-based approaches to analgesic drug discovery

The development of NGS [88] has facilitated unrestricted access to the whole genetic information in agreement with the wider acceptance of whole genomic information as a valuable method in clinical research [89]. NGS techniques have been broadly applied in functional genomics research [90–93]. Due to their better cost-effectiveness than that of Sanger sequencing, they probably will emerge as the dominant genomics technology [94] extending to gene-expression profiling [95], genome annotation [96], small non-coding RNA discovery [97] and detection of

aberrant transcription [98] previously dominated by microarrays [99,100], up to analyzing the epigenome for analgesic drug discovery including histone modifications [100] and DNA methylation [101,102]. NGS has shifted the research focus from population genome-wide association studies aimed at identifying common genetic variation toward exome sequencing targeted at rare gene variants of high effect or at patterns of noncoding occurred in noncoding portions of the genome increasingly recognized as functionally relevant [103].

So far, NGS has been applied in a pharmacogenetic context analyzing the genetic background of the effects of lidocaine combined with epinephrine, which had been administered following oral surgery in male and female Caucasians [104]. Similarly, with respect to opioid receptors, NGS has been applied to identify a genetic background of extremely high opioid dosing demands for analgesia [105]. A machine-learned classifier was able to identify such patients at a moderately high sensitivity and specificity based on a pattern of variants in opioid receptor genes. At a broader level, NGS was applied onto the DNA of subjects exhibiting extreme pain phenotypes such as insensitivity to pain or episode of paroxysmally exaggerated pain [106] with the aim to identify novel drug targets from identified variants in genes. However, while supporting concepts of pharmacogenomics-based analgesic therapy, the clinical utility is not yet apparent [107] and personalized opioid therapies are still obtained mainly by dose titration [108]. Thus, while progress in identifying the genetics of pain and analgesia has made, results are just at the beginning to appear.

Proteomics in analgesic drug discovery & repurposing

Proteomics addresses the entire content of proteins in a cell, a tissue or an organism, complementing genomic approaches by delivering information about protein isoforms, post-translational modifications such as glycosylation and phosphorylation, protein–protein interaction or protein stability and degradation [109]. Proteomic methods have been widely used for identification of potential biomarkers in several disease conditions, [110] and their application to pain mediators or regulating transmitters is obvious [111].

By extending single molecule analyses to whole biological networks and pathways, a molecular overview of the biological processes involved in chronic pain can be obtained [112]. Indeed, an analysis of spinal protein expression in rats exposed to repeated intrathecal injection of morphine to investigate morphine tolerance and addiction identified eight regulated proteins possibly qualifying as targets for the prevention of morphine tolerance [113]. Moreover, combination of whole-genome sequencing with functional genomic knowledge allowed the association of proteins of unknown function with a physiological role in pain [114]. Specifically, using homology cloning and sequencing of whole genomes in cell lines with the framework of so-called reverse pharmacology [115] orphan receptors, previously identified as natural ligands of nociceptin, which is a neuropeptide-modulating nociception by inhibiting neuronal activity [116], could be recognized as ORL1 [117]. Ligands consisting of brain-penetrating agonists or antagonists have recently been screened in animal models of acute and inflammatory pain to evaluate the ORL1 receptor as target for novel nonopioid analgesics [117].

Further uses of proteomics in analgesic drug discovery and repurposing are structure-based methods. The idea is that proteins with similar structures are likely to have similar functions and recognize similar ligands. Therefore, protein comparison is used as a method to identify secondary targets of an approved drug [48] such as the search for protein pairs sharing ligands on the basis of a known common ligand to both members of the pair, which can be employed to identify off-target effects of known drugs [118].

Discussion & conclusion

Functional genomics combines data from several sources, directly related to genetic research or indirectly via the study of gene products or of interactions between gene products and their targets up to complex networks that combine gene expression, proteomic targets and drug-related features. This complexity, which accommodates the complexity of pain and of the effects of analgesic drugs, requires an adequate reflection in the data analysis strategies. Suitable approaches are provided by contemporary methods of data science [119] including methods of artificial intelligence with their currently most popular implementation of machine learning in pain research [120]. This facilitates the extraction of information from data to derive new knowledge [121] adopting so-called DIKW (data – information – knowledge – wisdom) research approaches. Machine learning is currently increasingly used in pain and analgesia [120] and, for example, was able to produce genetic biomarkers of pain sensitivity [122] or opioid effects [105] that outperformed classical approaches.

Several approaches have been shown to provide suitable paths to new analgesic uses for known drugs. Most approaches so far generated short lists of repurposing candidates preselected among many more available compounds,

hence, shortening the search for repurposing candidates substantially. Published results so far could be demonstrated to be biologically plausible and were supported by evidence for an implication in pain or analgesia. However, none of the presented computational projects has so far succeeded in the development of a clinically approved repurposed drug. This may be due to the novelty of computational approaches that are under fast development. One of their limitations is the dependency of the correctness and completeness of database information. However, it should be mentioned that insufficient evidence or information is a general problem in science affecting both, data-driven and hypothesis-driven research. A more specific problem related to drug research is that certain classes of drugs are likely to be differently suited to the approaches summarized above under drug discovery in gene function and drug target-related databases. For example, drugs that induce changes very close to transcriptional machinery such as topoisomerase or histone deacetylase inhibitors may have stronger and more unspecific effects on gene expression than drugs that function in a more indirect mechanism such as antibodies that block cytokine binding to receptors. A further problem, in particular with machine learning, is overfitting, that is, performing rote learning ending up in describing noise or irrelevant relationships rather than the true relationship between data. However, measures against the latter are implemented as a standard of the analyses, such as splitting the data in training and tests sets and performing cross-validation virtual experiments, not substantially differing from classic approaches that, for example, in gene variant association studies, require an independent validation cohort. Hence, weaknesses of computational methods are not unique while mainly being specific for these particular techniques. Therefore, computational approaches in functional genomics of pain and analgesic drugs complement rather than replace classical approaches that are based on laboratory experiments of clinical investigations. The examples presented in this review demonstrate the ability of computational approaches to substantially accelerate the discovery of drug-repurposing candidates for analgesic indications.

Future perspective

The accelerating combination of biomedical laboratory and clinical research with novel methods of data science bears the potential to reshape our approaches to drug discovery and development. Novel approaches are being proposed at decreasing intervals and their application has already led to working results. These approaches require a new level of interdisciplinary collaboration where experts in biomedical research, biostatistics, mathematics and computer science not only contribute their particular skills to a collaborative project, but increasingly exceed their own fields of expertise to adopt the understanding of the contributing disciplines in order to achieve a comprehensive understanding of the acquisition, processing and interpretation of the data. This will combine human and artificial intelligence in the field of analgesic drug development and possibly accelerate the development of novel analgesic drugs to provide the necessary treatments for pain, in particular persistent pain, which is currently still a major healthcare problem affecting millions of patients worldwide.

Executive summary

- Functional genomics combines data derived from various processes related to DNA sequence, gene expression, and protein function, such as coding and noncoding transcription, protein translation, and protein–DNA, protein–RNA, and protein–protein interactions.
- Several computational functional genomics based approaches to drug development have been developed in the last few years, comprising (i) knowledge discovery in gene function and drug target related databases for drug discovery and repurposing, (ii) the use of next generation sequencing for analgesic drug research purposes and (iii), functional proteomics based approaches to analgesic target identification.
- The search for new analgesics increasingly includes computational functional genomics with the goal of understanding the relationship between the genome and the pain-related phenotype and the interaction of drugs with this system.
- A subset of respective methods has already been applied to pain-related research topics and results are promising in the sense of proof-of-concept; however, without having led to the development of novel analgesics yet.
- To exemplify the concept of computational functional genomics based research in pain and analgesia, the creation of a GeneOntology knowledgebase derived picture of the biological functions important to pain from a genetics perspective is demonstrated in this summary.

Supplementary data

Supplementary information includes a detailed description of the use of the R library 'dbtORA' (file 'dbtORA.docx') and the complete set of files related to the generation of the computational functional genomics analysis of pain (file: dbtORAfiles.zip), which among output files generated by the R library incusing the list of the $n = 540$ 'pain genes' used for the analysis (Schmerz-Gene2016Mar540Genes540.names). The single files are described in the 'dbtORA.docx' document.

To view the supplementary data that accompany this paper please visit the journal website at:

<https://www.futuremedicine.com/doi/suppl/10.2217/pgs-2018-0036>

Financial & competing interests disclosure

This work has been funded by the Landesoffensive zur Entwicklung wissenschaftlich – ökonomischer Exzellenz (LOEWE), LOEWE-Zentrum für Translationale Medizin und Pharmakologie (J Lötsch), which provided the necessary means to make the R programming of the dbtORA library possible, and by the European Union Seventh Framework Programme (FP7/2007 - 2013) under grant agreement no. 602919 (J Lötsch, GLORIA). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. Sertürner F. Darstellung der reinen mohnsäure (opiumsäure) nebst einer chemischen untersuchung des opiums. *J. Pharmacie* 14, 47–93 (1806).
2. Knorr L. Darstellung von oxypyrazolen durch einwirkung von acetessigestern, ihrer substitutionsprodukten und homologen auf hydrazine (D.P. 264298 vom 22. Juli 1883). *Berichte der Deutschen Chemischen Gesellschaft* 17, 149 (1884).
3. Dreser H. Pharmakologisches über aspirin (acetylsalicylsäure). *Pflügers Arch.* 76, 306–318 (1899).
4. Lötsch J, Geisslinger G. Pharmacogenetics of new analgesics. *Br. J. Pharmacol.* 163(3), 447–460 (2011).
5. Schmidt A, Lötsch J, Freynhagen R, Geisslinger G. Ziconotide for treatment of severe chronic pain. *Lancet* 375(9725), 1569–1577 (2010).
6. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur. J. Pain* 10(4), 287–333 (2006).
7. Elliott AM, Smith BH, Penny KI, Smith WC, Chambers WA. The epidemiology of chronic pain in the community. *Lancet* 354(9186), 1248–1252 (1999).
8. Elzahaf RA, Tashani OA, Unsworth BA, Johnson MI. The prevalence of chronic pain with an analysis of countries with a Human Development Index less than 0.9: a systematic review without meta-analysis. *Curr. Med. Res. Opin.* 28(7), 1221–1229 (2012).
9. Reid KJ, Harker J, Bala MM *et al.* Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. *Curr. Med. Res. Opin.* 27(2), 449–462 (2011).
10. Hill L, Schug SA. Recent advances in the pharmaceutical management of pain. *Expert Rev. Clin. Pharmacol.* 2(5), 543–557 (2009).
11. Kringel D, Lötsch J. Pain research funding by the European Union Seventh Framework Programme. *Eur. J. Pain* 19(5), 595–600 (2015).
12. Godinova AM. [Genetic analysis of migraine]. *Zh. Neuropatol. Psikhiatr. Im. S. S. Korsakova* 65(8), 1132–1138 (1965).
13. Ultsch A, Kringel D, Kalso E, Mogil JS, Lötsch J. A data science approach to candidate gene selection of pain regarded as a process of learning and neural plasticity. *Pain* 157(12), 2747–2757 (2016).
14. Lötsch J, Lippmann C, Kringel D, Ultsch A. Integrated computational analysis of genes associated with human hereditary insensitivity to pain. A drug repurposing perspective. *Front. Neurosci.* 10, 252 (2017).
15. Venter JC, Adams MD, Myers EW *et al.* The sequence of the human genome. *Science* 291(5507), 1304–1351 (2001).
16. Lander ES, Linton LM, Birren B *et al.* Initial sequencing and analysis of the human genome. *Nature* 409(6822), 860–921 (2001).
17. Gibson G, Muse SV. *A Primer Of Genome Science*. Sinauer Associates, MA, USA (2009).
18. Lötsch J, Doebering A, Mogil JS, Arndt T, Geisslinger G, Ultsch A. Functional genomics of pain in analgesic drug development and therapy. *Pharmacol. Ther.* 139(1), 60–70 (2013).
19. Bunnik EM, Le Roch KG. An introduction to functional genomics and systems biology. *Adv. Wound Care* 2(9), 490–498 (2013).
20. Carbon S, Ireland A, Mungall CJ *et al.* AmiGO: online access to ontology and annotation data. *Bioinformatics* 25(2), 288–289 (2009).
21. Ashburner M, Ball CA, Blake JA *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25(1), 25–29 (2000).
22. Seal RL, Gordon SM, Lush MJ, Wright MW, Bruford EA. Genenames.org: the HGNC resources in 2011. *Nucleic Acids Res.* 39, D514–D519 (2011).

23. Dennis G, Sherman BT, Hosack DA *et al*. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* 4(9), R60 (2003).
24. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. *Trends Genet.* 13(4), 163 (1997).
25. Rappaport N, Nativ N, Stelzer G *et al*. MalaCards: an integrated compendium for diseases and their annotation. *Database (Oxford)* 2013, bat018 (2013).
26. Lacroix-Fralish ML, Ledoux JB, Mogil JS. The Pain Genes Database: an interactive web browser of pain-related transgenic knockout studies. *Pain* 131(1–2), 3e1–4e1 (2007).
27. Beraldo Meloto C, Benavides R, Lichtenwalter RN *et al*. The Human Pain Genetics Database (HPGDB): a resource dedicated to human pain genetics research. *Pain* 159(4), 749–763 (2018).
28. Davis AP, Murphy CG, Saraceni-Richards CA, Rosenstein MC, Wiegers TC, Mattingly CJ. Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical–gene–disease networks. *Nucleic Acids Res.* 37, D786–D792 (2009).
29. Cerami EG, Gross BE, Demir E *et al*. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.* 39, D685–D690 (2011).
30. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* 38, D355–D360 (2010).
31. Wishart DS, Knox C, Guo AC *et al*. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* 36, D901–D906 (2008).
32. Wishart DS, Knox C, Guo AC *et al*. DrugBank: a comprehensive resource for *in silico* drug discovery and exploration. *Nucleic Acids Res.* 34, D668–D672 (2006).
33. Lamb J, Crawford ED, Peck D *et al*. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313(5795), 1929–1935 (2006).
34. Wu H, Huang J, Zhong Y, Huang Q. DrugSig: a resource for computational drug repositioning utilizing gene expression signatures. *PLoS ONE* 12(5), e0177743 (2017).
35. Gaulton A, Bellis LJ, Bento AP *et al*. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 40, D1100–D1107 (2012).
36. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45(D1), D158–D169 (2017).
37. Boeckmann B, Bairoch A, Apweiler R *et al*. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* 31(1), 365–370 (2003).
38. Bairoch A, Apweiler R. The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucleic Acids Res.* 24(1), 21–25 (1996).
39. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30(1), 207–210 (2002).
40. R Development Core Team. R: a language and environment for statistical computing. the R Foundation for Statistical Computing, Vienna, Austria (2011). <http://www.R-project.org/>
41. Pacini C, Iorio F, Goncalves E *et al*. DvD: an R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data. *Bioinformatics* 29(1), 132–134 (2013).
42. Backes C, Keller A, Kuentzer J *et al*. GeneTrail – advanced gene set enrichment analysis. *Nucleic Acids Res.* 35, W186–W192 (2007).
43. Huber W, Carey VJ, Gentleman R *et al*. Orchestrating high-throughput genomic analysis with bioconductor. *Nat. Methods* 12(2), 115–121 (2015).
44. Carlson M. org.Hs.eg.db: genome wide annotation for human (2017). <https://bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html>
45. Carlson M. GO.db: a set of annotation maps describing the entire Gene Ontology (2017). <https://bioconductor.org/packages/release/data/annotation/html/GO.db.html>
46. Gansner ER, North SC. An open graph visualization system and its applications to software engineering. *Software – Practice and Experience* 30(11), 1203–1233 (2000).
47. Brodie ML, Mylopoulos J. Knowledge bases and databases: semantic vs computational theories of information. *N. Directions Database Syst.* Ablex, 186–218 (1986).
48. March-Vila E, Pinzi L, Sturm N *et al*. On the integration of *in silico* drug design methods for drug repurposing. *Front. Pharmacol.* 8, 298 (2017).
49. Shalev-Shwartz S, Ben-David S. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, Cambridge, UK (2014).
50. Murphy KP. *Machine Learning: A Probabilistic Perspective*. The MIT Press, MA, USA (2012).

51. Anighoro A, Stumpfe D, Heikamp K *et al.* Computational polypharmacology analysis of the heat shock protein 90 interactome. *J. Chem. Inf. Model.* 55(3), 676–686 (2015).
52. Tissieres A, Mitchell HK, Tracy UM. Protein synthesis in salivary glands of *Drosophila melanogaster*: relation to chromosome puffs. *J. Mol. Biol.* 84(3), 389–398 (1974).
53. Lötsch J, Geisslinger G. Low-dose drug combinations along molecular pathways could maximize therapeutic effectiveness while minimizing collateral adverse effects. *Drug Discov. Today* 16(23–24), 1001–1006 (2011).
54. Wainger BJ, Buttermore ED, Oliveira JT *et al.* Modeling pain *in vitro* using nociceptor neurons reprogrammed from fibroblasts. *Nat. Neurosci.* 18(1), 17–24 (2015).
55. Bredel M, Jacoby E. Chemogenomics: an emerging strategy for rapid target and drug discovery. *Nat. Rev. Genet.* 5, 262 (2004).
56. Gregori-Puigjane E, Mestres J. A ligand-based approach to mining the chemogenomic space of drugs. *Comb. Chem. High Throughput Screen.* 11(8), 669–676 (2008).
57. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* 3(11), 935–949 (2004).
58. Li YY, An J, Jones SJ. A computational approach to finding novel targets for existing drugs. *PLoS Comput. Biol.* 7(9), e1002139 (2011).
59. Kharkar PS, Borhade S, Dangi A, Warriar S. In search of novel anti-inflammatory agents: computational repositioning of approved drugs. *J. Comp. Sci.* 10, 217–224 (2015).
60. Kiener HP, Lee DM, Agarwal SK, Brenner MB. Cadherin-11 induces rheumatoid arthritis fibroblast-like synoviocytes to form lining layers *in vitro*. *Am. J. Pathol.* 168(5), 1486–1499 (2006).
61. Dakshanamurthy S, Issa NT, Assefnia S *et al.* Predicting new indications for approved drugs using a proteochemometric method. *J. Med. Chem.* 55(15), 6832–6848 (2012).
62. Filzen TM, Kutchukian PS, Hermes JD, Li J, Tudor M. Representing high throughput expression profiles via perturbation barcodes reveals compound targets. *PLoS Comp. Biol.* 13(2), e1005335 (2017).
63. Iorio F, Rittman T, Ge H, Menden M, Saez-Rodriguez J. Transcriptional data: a new gateway to drug repositioning? *Drug Discov. Today* 18(7–8), 350–357 (2013).
64. Lamb J. The connectivity map: a new tool for biomedical research. *Nat. Rev. Cancer* 7(1), 54–60 (2007).
65. Subramanian A, Narayan R, Corsello SM *et al.* A Next Generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 171(6), 1437.e1417–1452.e1417 (2017).
66. Jadamba E, Shin M. A systematic framework for drug repositioning from integrated omics and drug phenotype profiles using pathway-drug network. *BioMed Res. Int.* 2016, 7147039 (2016).
67. Li J, Lu Z. Pathway-based drug repositioning using causal inference. *BMC Bioinformatics* 14(Suppl. 16), S3 (2013).
68. Dudley JT, Sirota M, Shenoy M *et al.* Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. *Sci. Transl. Med.* 3(96), 96ra76 (2011).
69. Moffat JG, Vincent F, Lee JA, Eder J, Prunotto M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* 16(8), 531–543 (2017).
70. Lötsch J, Ultsch A. Process pharmacology: a pharmacological data science approach to drug development and therapy. *CPT Pharmacometrics Syst. Pharmacol.* 5(4), 192–200 (2016).
71. Camon E, Magrane M, Barrell D *et al.* The Gene Ontology Annotation (GOA) project: implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Res.* 13(4), 662–672 (2003).
72. Camon E, Magrane M, Barrell D *et al.* The Gene Ontology Annotation (GOA) database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res.* 32, D262–D266 (2004).
73. Fisher RA. On the interpretation of chi square from contingency tables, and the calculation of p. *J. R. Stat. Soc.* 85(1), 87–94 (1922).
74. Bonferroni CE. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8, 3–62 (1936).
75. Ultsch A, Lötsch J. Computed ABC analysis for rational selection of most informative variables in multivariate data. *PLoS ONE* 10(6), e0129767 (2015).
76. Thulasiraman K, Swamy MNS. *Graphs: Theory And Algorithms*. Wiley, New York, NY, USA (1992).
77. Ultsch A, Lötsch J. Functional abstraction as a method to discover knowledge in gene ontologies. *PLoS ONE* 9(2), e90191 (2014).
78. Tian L, Ma L, Kaarela T, Li Z. Neuroimmune crosstalk in the central nervous system and its significance for neurological diseases. *J. Neuroinflammation* 9, 155 (2012).
79. Von Hehn CA, Baron R, Woolf CJ. Deconstructing the neuropathic pain phenotype to reveal neural mechanisms. *Neuron* 73(4), 638–652 (2012).
80. Calvo M, Dawes JM, Bennett DL. The role of the immune system in the generation of neuropathic pain. *Lancet Neurol.* 11(7), 629–642 (2012).

81. Rosen S, Ham B, Mogil JS. Sex differences in neuroimmunity and pain. *J. Neurosci. Res.* 95(1–2), 500–508 (2017).
82. Haus E, Sackett-Lundeen L, Smolensky MH. Rheumatoid arthritis: circadian rhythms in disease activity, signs and symptoms, and rationale for chronotherapy with corticosteroids and other medications. *Bull. NYU Hosp. Jt Dis.* 70(Suppl. 1), 3–10 (2012).
83. Gibbs JE, Ray DW. The role of the circadian clock in rheumatoid arthritis. *Arthritis Res. Ther.* 15(1), 205 (2013).
84. Gilron I, Ghasemlou N. Chronobiology of chronic pain: focus on diurnal rhythmicity of neuropathic pain. *Curr. Opin. Support. Palliat. Care* 8(4), 429–436 (2014).
85. Lötsch J, Ultsch A. A machine-learned computational functional genomics-based approach to drug classification. *Eur. J. Clin. Pharmacol.* 72(12), 1449–1461 (2016).
86. Cox JJ, Reimann F, Nicholas AK *et al.* An SCN9A channelopathy causes congenital inability to experience pain. *Nature* 444(7121), 894–898 (2006).
87. Lötsch J, Kringel D. Use of computational functional genomics in drug discovery and repurposing for analgesic indications. *Clin. Pharmacol. Ther.* doi:10.1002/cpt.960 (2017) (Epub ahead of print).
88. Metzker ML. Sequencing technologies – the next generation. *Nat. Rev. Genet.* 11(1), 31–46 (2010).
89. Frank M, Prenzler A, Eils R, Graf Von Der Schulenburg JM. Genome sequencing: a systematic review of health economic evidence. *Health Econ. Rev.* 3(1), 29 (2013).
90. Bennett ST, Barnes C, Cox A, Davies L, Brown C. Toward the 1000 dollars human genome. *Pharmacogenomics* 6(4), 373–382 (2005).
91. Mikkelsen TS, Ku M, Jaffe DB *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153), 553–560 (2007).
92. Jones-Rhoades MW, Borevitz JO, Preuss D. Genome-wide expression profiling of the Arabidopsis female gametophyte identifies families of small, secreted proteins. *PLoS Genet.* 3(10), 1848–1861 (2007).
93. Cheung F, Haas BJ, Goldberg SM, May GD, Xiao Y, Town CD. Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology. *BMC Genomics* 7, 272 (2006).
94. Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92(5), 255–264 (2008).
95. Torres TT, Metta M, Ottenwalder B, Schlotterer C. Gene expression profiling by massively parallel sequencing. *Genome Res.* 18(1), 172–177 (2008).
96. Emrich SJ, Barbazuk WB, Li L, Schnable PS. Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Res.* 17(1), 69–73 (2007).
97. Morin RD, O'Connor MD, Griffith M *et al.* Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 18(4), 610–621 (2008).
98. Morozova O, Marra MA. From cytogenetics to next-generation sequencing technologies: advances in the detection of genome rearrangements in tumors. *Biochem. Cell Biol.* 86(2), 81–91 (2008).
99. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235), 467–470 (1995).
100. Barski A, Cuddapah S, Cui K *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* 129(4), 823–837 (2007).
101. Esteller M. The necessity of a human epigenome project. *Carcinogenesis* 27(6), 1121–1125 (2006).
102. Callinan PA, Feinberg AP. The emerging science of epigenomics. *Hum. Mol. Genet.* 15(1), R95–R101 (2006).
103. Khurana E, Fu Y, Chakravarty D, Demichelis F, Rubin MA, Gerstein M. Role of non-coding sequence variants in cancer. *Nat. Rev. Genet.* 17(2), 93–108 (2016).
104. Kim H, Ramsay E, Lee H, Wahl S, Dionne RA. Genome-wide association study of acute post-surgical pain in humans. *Pharmacogenomics* 10(2), 171–179 (2009).
105. Kringel D, Ultsch A, Zimmermann M *et al.* Emergent biomarker derived from next-generation sequencing to identify pain patients requiring uncommonly high opioid doses. *Pharmacogenomics J.* 17(5), 419–426 (2017).
106. Williams FM, Scollen S, Cao D *et al.* Genes contributing to pain sensitivity in the normal population: an exome sequencing study. *PLoS Genet.* 8(12), e1003095 (2012).
107. Lane-Ladd SB, Pineda J, Boundy VA *et al.* CREB (cAMP response element-binding protein) in the locus coeruleus: biochemical, physiological, and behavioral evidence for a role in opiate dependence. *J. Neurosci.* 17(20), 7890–7901 (1997).
108. Hajj A, Khabbaz L, Laplanche JL, Peoc'h K. Pharmacogenetics of opiates in clinical practice: the visible tip of the iceberg. *Pharmacogenomics* 14(5), 575–585 (2013).
109. Niederberger E, Geisslinger G. Proteomics in neuropathic pain research. *Anesthesiology* 108(2), 314–323 (2008).
110. Banks RE, Dunn MJ, Hochstrasser DF *et al.* Proteomics: new perspectives, new biomedical opportunities. *Lancet* 356(9243), 1749–1756 (2000).

111. Huang HL, Cendan CM, Roza C *et al*. Proteomic profiling of neuromas reveals alterations in protein composition and local protein synthesis in hyper-excitabile nerves. *Mol. Pain* 4, 33 (2008).
112. Moore DM, McCrory C. The proteomics of intrathecal analgesic agents for chronic pain. *Curr. Neuropharmacol.* 15(2), 198–205 (2017).
113. Liaw WJ, Tsao CM, Huang GS *et al*. Phosphoproteomics and bioinformatics analyses of spinal cord proteins in rats with morphine tolerance. *PLoS ONE* 9(1), e83817 (2014).
114. Meunier JC. Nociceptin/orphanin FQ and the opioid receptor-like ORL1 receptor. *Eur. J. Pharmacol.* 340(1), 1–15 (1997).
115. Mills A, Duggan MJ. Orphan seven transmembrane domain receptors: reversing pharmacology. *Trends Biotechnol.* 12(2), 47–49 (1994).
116. Chiou LC, Liao YY, Fan PC *et al*. Nociceptin/orphanin FQ peptide receptors: pharmacology and clinical implications. *Curr. Drug Targets* 8(1), 117–135 (2007).
117. Meunier JC. Utilizing functional genomics to identify new pain treatments: the example of nociceptin. *Am. J. Pharmacogenomics* 3(2), 117–130 (2003).
118. Chen YC, Tolbert R, Aronov AM, McGaughey G, Walters WP, Meireles L. Prediction of protein pairs sharing common active ligands using protein sequence, structure, and ligand similarity. *J. Chem. Inf. Model.* 56(9), 1734–1745 (2016).
119. President's Information Technology Advisory C. Report to the President: computational science: ensuring America's competitiveness (2005). www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf
120. Lotsch J, Ultsch A. Machine learning in pain research. *Pain* 159(4), 623–630 (2017).
121. Dhar V. Data science and prediction. *Communications of the ACM* 56(12), 64–73 (2013).
122. Lötsch J, Ultsch A. A machine-learned knowledge discovery method for associating complex phenotypes with complex genotypes. Application to pain. *J Biomed. Inform.* 46(5), 921–928 (2013).