ELSEVIER

*Systems-level drug response phenotypes combined with network models offer an exciting means for elucidating the mechanisms of action of polypharmacological agents, including multitargeted natural products.*



# Transcriptional response networks for elucidating mechanisms of action of multitargeted agents

**Milla Kibble[1], Suleiman A. Khan[1], Niina Saarinen[2], Francesco Iorio[3], Julio Saez-Rodriguez[3,4], Sari Mäkelä[2] and Tero Aittokallio[1,5]**

[1] Institute for Molecular Medicine Finland (FIMM), Biomedicum Helsinki 2U, Tukholmankatu 8, University of Helsinki, Helsinki 00014, Finland
[2] Institute of Biomedicine, Turku Center for Disease Modeling & Functional Foods Forum, University of Turku, Turku 20014, Finland
[3] European Molecular Biology Laboratory – European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SD, UK
[4] Joint Research Centre for Computational Biomedicine (JRC-COMBINE) – RWTH Aachen University, Faculty of Medicine, D-52074 Aachen, Germany
[5] Department of Mathematics and Statistics, Quantum, University of Turku, Turku 20014, Finland

**Drug discovery is moving away from the single target-based approach towards harnessing the potential of polypharmacological agents that modulate the activity of multiple nodes in the complex networks of deregulations underlying disease phenotypes. Computational network pharmacology methods that use systems-level drug–response phenotypes, such as those originating from genome-wide transcriptomic profiles, have proved particularly effective for elucidating the mechanisms of action of multitargeted compounds. Here, we show, via the case study of the natural product pinosylvin, how the combination of two complementary network-based methods can provide novel, unexpected mechanistic insights. This case study also illustrates that elucidating the mechanism of action of multitargeted natural products through transcriptional response-based approaches is a challenging endeavor, often requiring multiple computational–experimental iterations.**

## Introduction

Over the past few decades, drug discovery has adopted a target-based approach, searching for compounds that target a specific disease-causing gene product [1]. Perhaps the most successful example of the target-based approach was the discovery of imatinib, a BCR-ABL tyrosine kinase inhibitor (TKI), for the treatment of several cancers, including Philadelphia chromosome-positive chronic myeloid leukemia (CML) [2]. However, recently, the paradigm of searching for highly

*Corresponding author:* Kibble, M. (milla.kibble@helsinki.fi)

**Milla Kibble** received her PhD in pure mathematics from University College London in 2000. She subsequently worked at Numbercraft Ltd., using analytics techniques to gain information from large data sets. Since then, she has combined teaching of mathematics and statistics with research into probability and graph theory applications to biological data in the Biomathematics Group at Turku University and the Stochastics Group at Aalto University. Milla joined the Institute for Molecular Medicine Finland (FIMM) as a senior researcher in 2014 and her research focuses on network pharmacology approaches to elucidate mechanisms of action and efficacy of drugs and their combinations, with a particular interest in natural products.

**Suleiman Ali Khan** received his PhD in computer science from Aalto University in 2015, with a primary focus on statistical machine learning and bioinformatics, in particular the application and development of novel Bayesian multiview methods to model structural responses of drugs. He subsequently joined FIMM as a postdoctoral researcher and has been working on the identification of drug response factors using integrative and predictive machine-learning methods, with a specific emphasis on combining information from multiple genomic data sets. His current research interests include computational biology, medicine, and statistical machine learning.

**Tero Aittokallio** received his PhD in applied mathematics from the University of Turku in 2001. He did his postdoctoral training in the Systems Biology Lab at the Institut Pasteur (2006–2007), where he focused on network biology applications using high-throughput experimental assays. Dr Aittokallio then launched his independent career as a principal investigator in the Turku Biomathematics Research Group in 2007, and received a 5-year appointment as an Academy of Finland Research Fellow (2007–2012). In 2011, he started as EMBL Group Leader at FIMM, where his research group is developing systems medicine solutions for biomedical problems, with a particular interest in personalized medicine.


http://dx.doi.org/10.1016/j.drudis.2016.03.001

**GLOSSARY**

**Batch effect** the similarity of gene expression or other molecular profiles observed for unrelated stimuli in cells grown or processed simultaneously.
**Drug network** a network in which drugs are represented by nodes and nodes are connected if two drugs are similar in a predefined manner (here, in terms of their transcriptional response).
**Drug repositioning** the discovery of new uses for approved drugs, which have well-known safety and pharmacokinetic profiles [23].
**Gene expression profile** a list of probe set identifiers, ranked according to differential gene expression in a particular cellular context, with the most upregulated probe sets at the top of the list.
**Mechanism of action (MoA)** the set of cellular targets or other mechanisms of a compound that produce its pharmacological effect in a given cellular context.
**Network biology** systems-level modeling of biological processes, pathways, and molecular interactions, rather than focusing on individual players only.
**Network pharmacology** conceptualization of the fact that, in certain cases, to have an effect, drugs must target multiple pathways in a disease network and/or work synergistically with other drugs.
**Synergy** the joint action of drugs that when used together increase each other's effectiveness. See [80] for a discussion of the challenges of modeling synergistic interactions.

selective ligands has been questioned, because of high late-stage clinical attrition rates, which can be largely attributed to lack of clinical efficacy and safety [3–6]. With the recognition that many approved drugs are in fact multitargeted, an alternative systems biology approach to drug discovery is increasing in popularity. The so-called 'network pharmacology paradigm' [7] (see Glossary) makes use of systems-level disease models, aiming at finding compounds (or combinations thereof) that can modulate the activity of multiple nodes in the complex network of impaired mechanisms and deregulated interactions underlying a disease phenotype (i.e., the 'disease network'), either by targeting several pathways or by working synergistically, potentially with fewer adverse effects [8,9].

Therefore, an essential part of the modern drug discovery process is the elucidation of the multiple target mechanisms of compounds to better understand their phenotypic effects, both therapeutic and adverse. Here, computational methods offer great potential to provide initial hypotheses for *in vitro* and *in vivo* target validation studies. In particular, a class of computational drug-discovery methods that use comprehensive drug response phenotypes, such as those originating from genome-wide transcriptomic profiles, has proved particularly effective in elucidating compound mechanisms of action (MoA). These methods were initiated about a decade ago, with the advent of the Connectivity Map (CMap) data resource [10]. New types of drug response profile are continuously being produced, for example, from community efforts such as the Library of Integrated Network-based Cellular Signatures project (LINCS, http://www.lincsproject.org/), which aims to provide a comprehensive reference data resource of cellular response signatures to a range of small-molecule and genetic perturbations.

Here, we describe how the combination of two complementary computational network pharmacology methods, namely Mode of Action by Network Analysis (MANTRA) [11] and Group Factor Analysis (GFA) [12], can provide novel, unexpected MoA predictions. As a specific case, we illustrate the operation of these methods in the context of multitargeted natural products. Although less popular during the era of target-based drug discovery, natural products still make a substantial contribution to current new drugs [13–16]. Recent reviews have highlighted the largely untapped potential of network pharmacology approaches in natural product research [17–20]; we chose as our case study a relatively little-researched natural polyphenol, pinosylvin (3,5-dihydroxy-*trans*-stilbene). Pinosylvin has been shown to have antiproliferative and proapoptotic efficacy on prostate cancer cells *in vitro* and *in vivo* [21], but its MoA has so far been poorly understood. In this study, we show that the true value of the data-driven network-based methods is their potential to lead to unbiased hypotheses that might not otherwise have been conceived and, hence, to truly novel and even surprising findings.

## Transcriptional response networks to elucidate drug MoA

We illustrate the operation of the computational network pharmacology methods using genome-wide gene expression response profiles as their input. The main public resource for such data is the CMap resource, which comprises genome-wide transcriptional responses of various human cell lines to more than 1000 bioactive small molecules (Table 1). Given that our example compound pinosylvin is not included in this resource, we produced a response profile for pinosylvin in house (Box 1). Since the advent of the CMap resource, researchers have successfully discovered cellular targets of compounds and identified candidates for drug repositioning by searching for commonalities in the phenotypic responses using the simple idea that, if two drugs elicit similar transcriptional responses, then they could share a common MoA (hence, potentially a therapeutic application) even if directly acting on different intracellular targets. The original CMap tool compares transcriptional response profiles using a method similar to Gene Set Enrichment Analysis (GSEA) [22], based on the Kolmogorov-Smirnov statistic. Furthermore, computational researchers have provided an abundance of methods for measuring the similarity between genome-wide transcriptional responses.

For an overview of all developments over the past 10 years, many of which represent different variations of the original GSEA method, we refer the reader to the previous reviews by Iorio *et al.* [23] and by Qu *et al.* [24] of *in silico* approaches using *in vivo* and *in vitro* transcriptomic drug response profiles, and we comment on the new methods that have been published subsequent to the aforementioned reviews. A key recent development is the addition of new types of information source to the network-based methods. For example, Jahchan *et al.* [25] used additional information on the enriched pathways of the molecules that are most similar to the compound of interest in terms of transcriptional response, Laenen *et al.* [26] incorporated a functional protein association network into their computational model and Wu *et al.* [27] incorporated an adverse effect profile based on differential expression of essential genes. Quan *et al.* [28] used an approach to transcriptional response similarity assessment that utilizes the

**TABLE 1**

**Resources used during the case study**

| Resource | URL/availability/developers | Description | Application in the present study | Refs |
|---|---|---|---|---|
| **CMap** | http://www.broadinstitute.org/cmap/ Publicly available; produced by the Broad Institute of MIT and Harvard | Database of genome-wide gene expression profiles produced upon treatment of 1309 compounds on cancer cell lines | Transcriptional profiles of compounds for comparison with the pinosylvin profile | [10] |
| **MANTRA** | http://mantra.tigem.it/; analysis done using a web interface; publicly available | Computational tool for MoA analysis of drugs and identification of known and approved candidates for drug repositioning | Used to derive hypotheses on MoA of pinosylvin | [11,85] |
| **GFA** | http://research.cs.aalto.fi/pml/software/GFAsparse/; analysis done using publicly available R code of GFA | Computational approach for identifying multiple mechanisms of a drug, while extracting representative set of genes | Used to derive hypotheses on MoA of pinosylvin | [12] |
| **KIBA** | Analysis done using a simple R-script based on ChEMBL compound–target database | KIBA score summarizes potency of drug–target interactions using multiple bioactivity types | To obtain target profiles for compounds used in the analysis | [38] |
| **ChEMBL** | http://www.ebi.ac.uk/chembl; publicly available; run by the European Bioinformatics Institute, Cambridge, UK | Database with bioactivity measurements for almost 1.5 million distinct compounds and over 10 000 protein targets | To obtain target profiles for compounds used in the analysis | [86] |
| **Khan *et al.*** | Implementation is available on request from the authors | Drug–target data in this article were obtained from ChEMBL, DrugBank, DUD, and ZINC; additionally targets were extracted from supplementary material in [11]. In total, 716 CMap compounds had target information | To obtain target profiles for compounds used in the analysis | [84] |
| **DrugBank** | http://www.drugbank.ca/; publicly available | Database containing >4100 drug entries and >14 000 protein or drug target sequences linked to these drug entries | To obtain target profiles for compounds used in the analysis | [87] |
| **DUD** | http://dud.docking.org/; publicly available | Database of 2950 active compounds against a total of 40 targets, originally designed to provide decoys for benchmarking virtual screening | To obtain target profiles for compounds used in the analysis | [88] |
| **ZINC** | http://zinc.docking.org/; publicly available | Database containing about 90 million commercially available compounds for structure-based virtual screening, as well as their target information | To obtain target profiles for compounds used in the analysis | [89] |
| **Pathway Commons** | http://www.pathwaycommons.org/; publicly available | Collection of publicly available pathway data from multiple organisms, containing data on 31 698 pathways and 1 151 476 interactions from 18 data sources | To link targets of closest neighbors of pinosylvin to their underlying pathways | [90] |

idea of drug-induced transcriptional modules [29]. Also, some novel probabilistic methodologies are beginning to emerge [30,31] as well as a new combined unsupervised and supervised approach [32]. Finally, Napolitano *et al.* [33] conceived Drug-Set Enrichment Analysis to investigate commonalities in the MoA of a set of pharmacologically diverse compounds that are nevertheless able to induce a common phenotype. An application of a similar method to cystic-fibrosis phenotype correctors is presented in [34].

Most recently, Woo *et al.* [35] introduced a method, called DeMAND, for predicting the MoA of a compound using tissue-specific regulatory networks and Cmap-type transcriptional profiles in response to *in vitro* or *in vivo* compound perturbations. Specifically, they combined genome-wide gene expression profiles from human lymphoma cells with a lymphoma-specific gene interaction network, and used the Kullback–Leibler divergence [36] to establish compound-mediated dysregulation of given gene nodes in the network. Although the authors were able to experimentally prove their MoA predictions, one limitation of this approach is the requirement for high-quality, context-specific gene regulatory networks. However, as the authors comment, these might well be temporary limitations, because many groups are working on reliable methods to produce such networks for various cancer types. Once available on a large scale, such tissue-specific models have the potential to lead to the identification of more effective and safe therapeutic targets [37].

Here, we focus in more detail on two recent complementary network-based methods, namely MANTRA [11], and GFA [12], for elucidation of compound MoA. In brief, the MANTRA method is based on the GSEA method to measure the similarity between transcriptional response profiles; it uses solely the transcriptional response data, which are aggregated across the different cancer cell lines to dilute cell line-specific effects on transcription as well as batch effects and other confounding factors, such as multiple drug

**BOX 1**

## Production of pinosylvin gene expression response profiles

We aimed to discover the unknown MoA of pinosylvin via the comparison of its transcriptional response profile to those of compounds in CMap of known MoA. The CMap data resource comprises genome-wide transcriptional response profiles of 1309 small molecules in human cell lines, mainly the prostate cancer cell line PC3 and the breast cancer cell line MCF7. Given that pinosylvin is not one of the small molecules present in CMap, in this analysis, gene expression response profiles of pinosylvin at different concentrations on PC3 and MCF7 cell lines were produced as input into the MANTRA and GFA methods to gain further insights into the molecular mechanisms behind the effects of pinosylvin. The MCF7 cell line is the main cell line used in the CMap data, but we also wanted to include the lesser-used PC3 cell line to enable assessment of prostate cancer-specific action and to facilitate analysis of noncontext-specific action by combining the results over the two cell lines (Fig. I).

Given that the optimal concentration is not known for many compounds, most of the CMap compounds were applied to cell lines at a concentration of 10 μM. A subset of compounds was also applied at concentrations reported to be effective in cell culture or to approximate the maximum attainable plasma concentrations after therapeutic dosing, while another subset was profiled at a range of concentrations to explore the sensitivity of results to dose. To retain similarity in the methodology, we used a concentration of 10 μM for pinosylvin. We also produced expression data for the dose of 40 μM because previous studies have shown this to be a physiologically relevant, therapeutically effective, and nontoxic dose. The choice of dose for the other 1309 CMap compounds could affect the gene expression profiles and, thus, also the results produced from their comparison.

Four batches of the three doses (0 μM, 10 μM, and 40 μM) were produced for each cell line, yielding a total of 24 samples. Unlike in the CMap data, which was produced over a period of 1 year, here the batches for each cell line were produced and RNA extracted simultaneously under the same conditions and, therefore, batch effect was not deemed to be an issue. Following the application of pinosylvin to the cell lines and RNA extraction for each dose–cell line pair, the Limma statistical package in R was used to assess differential expression with respect to the untreated hybridization. The probe-set identifiers were then ranked by the modified *t*-statistic. The selected panel of differentially expressed probe set identifiers in each dose–cell line pair are called 'gene expression signatures'. The two gene expression profiles at dose 10 μM for MCF7 and PC3 had few probe sets with significant differential expression and, hence, were disregarded. Thus, in the analysis, we proceeded with the two gene expression profiles at a 40 μM dose, one for each cell line (Fig. I).

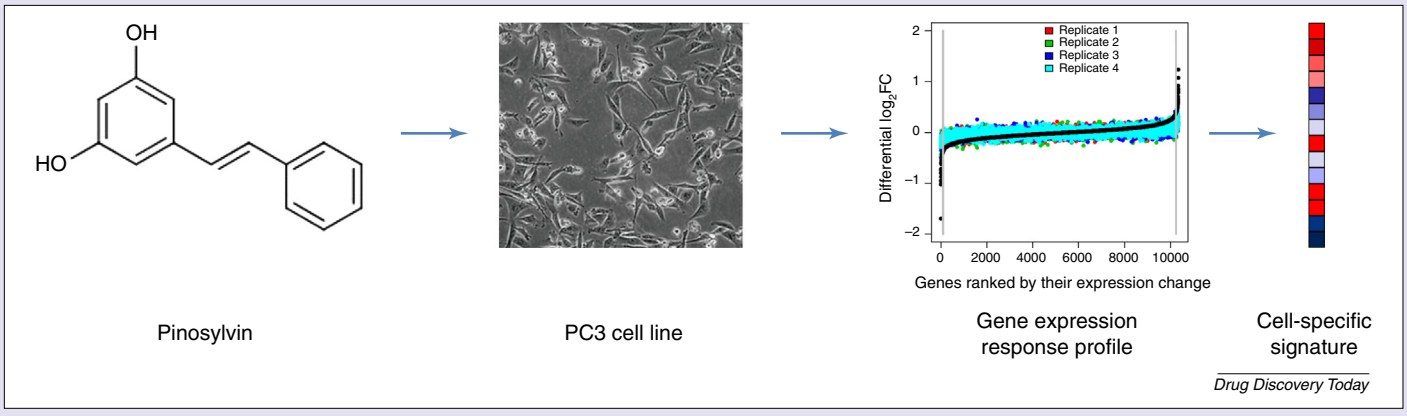


**FIGURE I**

Pinosylvin was applied to a specific cancer cell line, here PC3, to elicit a transcriptional response profile that forms the input into the computational methods.

dosages. By contrast, the GFA method is a probabilistic model that considers cell line-specific transcriptional responses and can also incorporate structural information about the compounds. We hypothesized that the use of such complementary methods would lead to a more comprehensive elucidation of MoA of multitargeted compounds. We begin by describing the two methods and their outputs using as an example case the natural product pinosylvin. Finally, we demonstrate the added value of using more than one method in providing valuable insights that single methods alone could not predict.

## Operation of the MANTRA method

As detailed in Box 2, the MANTRA method [11] produces a ranked list of CMap compounds based on the similarity of their gene expression profiles to that of the input compound, here pinosylvin (Table 2). The underlying idea of this network-based approach is to discover the MoA of the compound of interest via neighboring compounds that produce a similar transcriptional response and that have a known MoA. Such a 'guilt-by-association' approach requires accurate MoA information for the most similar CMap small molecules, something that is not necessarily easy to find. In our case study, the target profiles for the neighboring compounds used in the MANTRA analysis were obtained using multiple means: we applied the KIBA method [38] to curated ChEMBL bioactivity data; we also performed manual literature searches and combined bioactivity data from several drug-target databases, including DrugBank, DUD, and ZINC (Tables 1 and 2). This drives the search for a pinosylvin MoA hypothesis towards fundamental MoA. We begin by looking individually at the closest neighbors of pinosylvin, and then continue to look at the closest neighbors as a group. We end this section by describing a novel insight into the information that can be gained about groups of compounds.

## Closest neighbor approach

The simple idea behind looking at the drugs with the closest transcriptional response to pinosylvin is that two drugs with a similar MoA [by this we mean having cellular target(s) in common that contribute to the treatment response] will produce similar transcriptional changes. Out of all 1309 small molecules in the CMap data set, Chembridge 5707885 had the most similar transcriptional response to pinosylvin, as measured by the MANTRA method (Table 2). However, little is known about this compound apart from it having a transcriptional response similar to the natural phenol gossypol, which also occurs in the pinosylvin list, at position ten. Gossypol has been shown to have anticancer effects via various mechanisms [39,40], but it is difficult to pinpoint potential targets for pinosylvin from this association because gossypol is known to act on a widerange of targets, as discussed later. Hence, we move on, and consider the second closest neighbor, trifluoperazine.

Trifluoperazine, a dopamine receptor $D_2$ (DRD2) antagonist, is an antipsychotic drug of the phenothiazine group. It is also an antagonist of serotonin type 2 receptors 5-$HT_{2A}$ and 5-$HT_{2C}$, and a calmodulin (CaM) inhibitor (Table 2). There is also evidence of trifluoperazine inhibiting tumor growth [41,42]. This led us to question whether pinosylvin is a G-protein-coupled receptor (GPCR) ligand or a CaM inhibitor and, in particular, whether some of its anticancer effects might come via these mechanisms. There have been numerous links between GPCR ligands and anticancer activity [43] and CaM has an established role in cancer, both through apoptosis and invasion [44]. However, to our knowledge, the question of whether natural stilbenoids, such as pinosylvin, have an effect on prostate cancer via a GPCR-related mechanism has not been posed previously.

First, we tested pinosylvin against a set of 46 standard GPCRs. Although exceptionally high binding values were not obtained for any single target, ligand competition did seem selective to 5-$HT_2$ receptors. Building on the previous finding that 5-$HT_{2B}$ receptors are expressed in PC3 cells, and that a 5-HT2B antagonist, SB-21550, inhibits PC3 cell proliferation [45], we further tested pinosylvin together with the known 5-$HT_{2B}$ agonist BW723C86 in cell viability and cell toxicity assays. However, BW723C86 did not abolish the antiproliferative activity of pinosylvin in PC3 cells, which suggests that the effect of pinosylvin is not mediated by 5-$HT_{2B}$ receptors. So, for now, we put a GPCR-related mechanism out of the picture for pinosylvin, but we return to this line of enquiry later.

Second, we tested CaM inhibition using an invasion assay. Given that pinosylvin and our model compound trifluoperazine (a CaM inhibitor) had no effect on PC3 cell invasion, it is inconclusive whether pinosylvin inhibits CaM. However, our model-based investigation did lead to the interesting side result that CaM inhibition is not key to invasion in PC3 cells.

Looking further down the list of compounds having a similar transcriptional response to that produced by pinosylvin, we find compounds that produce a range of transcriptional and post-transcriptional modifications (Table 2). For example, Cytochalasin B affects the actin cytoskeleton [46] and so might induce broad transcriptional changes. Gossypol induces changes in signal transduction mediators involved in multiple protein expression and activity, and interacts with cellular and mitochondrial membranes [47]. Taken together with the fact that pinosylvin is low in the lists of closest neighbors of trifluoperazine and celastrol (the natural product fifth in the list of closest neighbors to pinosylvin), we postulated that pinosylvin acts on multiple targets, specifically on a subset of targets of trifluoperazine and celastrol.

**BOX 2**

### MANTRA method of Iorio *et al.*

The MANTRA method first produces a so-called single 'Prototype Ranked List' (PRL) of genes for each of the 1309 compounds in the CMap data set. In other words, the cell line–dose level gene expression profiles for a given compound are merged into a single profile for the compound using a rank-aggregation algorithm [81]. This ranked list aims to capture the consensus transcriptional response of a compound across doses and cell lines, and also to overcome batch-effect issues in the CMap data. To enable comparison, we also produced a PRL for pinosylvin; however, we were in a balanced situation whereby we had an equal number of compound treatments over the two cell lines and, hence, we could produce the merged ranked list simply by summing the rank positions for a given probe set and reranking, rather than using the more complicated merging algorithm. Given that we used the HG-U133 + PM array plate in our experiment, we reduced the 54 715 probe sets in our data to match the 22 283 probe sets in the CMap data produced on the Affymetrix GeneChip® HG-U133A array to enable further computational analysis.

Once we had a PRL for pinosylvin to compare with the PRLs of all 1309 CMap compounds (Fig. II), the comparison was done via a distance measure based on Gene Set Enrichment Analysis (GSEA) to output a list of closest neighbors to pinosylvin (i.e., a ranked list of those CMap drugs with the most similar merged gene expression profiles to that of pinosylvin). The idea was that compounds with similar transcriptional profiles might share MoA. Hence, the next step was to look at the MoA of the nearest neighbors to pinosylvin. The analysis was done using the MANTRA 2.0 tool, which is an easy-to-use web-based tool for applying the MANTRA method (Table 1, main text). As expected, when the individual non-merged MCF7 and PC3 expression profiles for pinosylvin were input into the MANTRA tool, these lists were the most similar to the merged pinosylvin list out of all CMap PRLs.

An additional output of MANTRA 2.0 is a visualization of a drug network, where drugs are nodes and are joined by an edge if the distance (a measure of the similarity in transcriptional response) between them is below a certain threshold (here, a distance score of less than 0.8). A 'community' is defined as a group of nodes densely interconnected with each other and with fewer connections to nodes outside the group. In total, 106 communities were identified using an affinity propagation algorithm [82]; each community was coded with a numerical identifier and one of its nodes was identified as the 'exemplar' of the community (i.e., the drug whose effect best represents the effects of the other drugs in the community).

The top 17 nearest neighbors of the polyphenols pinosylvin, resveratrol, and genistein were included in the network diagrams for visualization purposes and the network diagram was modified using Cytoscape [83] (Fig. II).

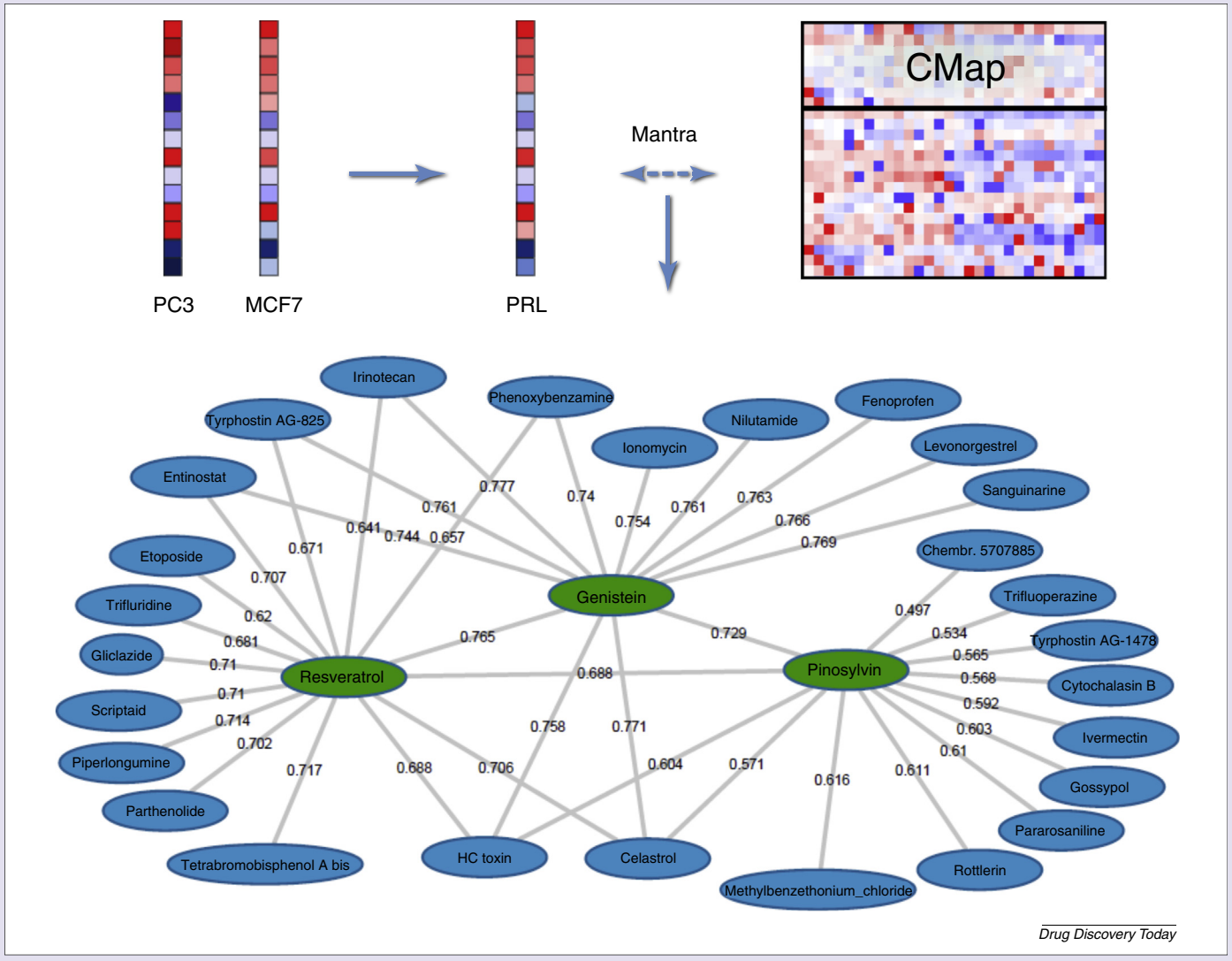


**FIGURE II**

MANTRA 2.0 mapped the closest compounds in terms of transcriptional response to pinosylvin, genistein, and resveratrol onto a drug-network.

## Using pathway information about closest neighbors

The transcriptional response profiles for a pair of drugs might be similar, even if their targets are not the same but if the drugs target the same pathway upstream of the translational effects. With this in mind, we mapped the known targets of the top five neighbors of pinosylvin (Table 2) onto pathways using Pathway Commons (Table 1), and found that all except trifluoperazine point to an inhibition of nuclear factor kappa B (NF-κB) downstream of the epidermal growth factor receptor (EGFR) (Fig. 1). This led to the hypothesis that the transcriptional profiles of these four drugs are similar to that of pinosylvin because of the ultimate inhibition of NF-κB. Subsequently, we found evidence to support our data-derived hypothesis: Lee *et al.* [48] showed that pinosylvin suppressed the production of proinflammatory mediators through the inhibition of the NF-κB pathway, a finding corroborated by Laavola *et al.* [49]. It is possible that this inhibition of NF-κB by pinosylvin is responsible for the proapoptotic effect seen in PC-3M-luc2 cells [21], as is the case for the polyphenol curcumin through blocking phosphorylation of IκBα [50].

Although NF-κB is involved in many processes and several signaling pathways, the example serves to highlight the idea of using pathway information from all of the closely related compounds to obtain novel MoA hypotheses. Extra support for the NF-κB hypothesis comes from the observation that celastrol (the fifth most similar drug to pinosylvin in terms of transcriptional response) is a potent antioxidant and anti-inflammatory drug [51] that has also been reported to have anticancer properties [52,53] and to inhibit NF-κB activation through IKKβ inhibition [54].

## Insights for groups of compounds

So far, we have used the MANTRA method to predict the MoA for pinosylvin. However, we end this section with a description of how the method also has potential to add insights into a group of compounds of the researcher's choice. Resveratrol (*trans*-3,5,4′-trihydroxystilbene) is a much-researched metabolite of pinosylvin, with a similar structure and many reported health benefits and anticancer properties [55–57]. Although it is not necessarily the case that two structurally similar compounds work in a similar way, as can be exemplified by the two structurally similar flavones

**TABLE 2**

**Ranked list of the top five drugs based on the similarity of their gene expression profiles to that of pinosylvin, produced by the MANTRA method[a]**

| Compound | MANTRA network distance | Information and/or targets | Source of target information |
|---|---|---|---|
| **Chembridge 5707885 2-Methoxyethyl 7-(4-chlorophenyl)-4-(2-fluorophenyl)-2-methyl-5-oxo-4,6,7,8-tetrahydro-1H-quinoline-3-carboxylate** | 0.497 | No target information is available for this Chembridge drug. The CMap drug with the closest transcriptional profile to Chembridge 5707885 is gossypol, which has the targets below: | |
| | | Protein kinase C (PKC) inhibitor | [91] |
| | | Aldose reductase (ALDR1) inhibitor | KIBA method |
| | | Induced myeloid cell leukemia **1** (MCL1) | |
| | | Apoptosis regulator B cell lymphoma 2 (BCL2) | |
| **Trifluoperazine** | 0.534 | DRD2 antagonist | ChEMBL report card |
| | | $5\text{-HT}_{2C}$ receptor antagonist | |
| | | $5\text{-HT}_{2A}$ receptor antagonist | |
| | | Alpha-1a adrenergic receptor (ADRA1A) | Data in [84] |
| | | Potassium-transporting ATPase (ATP4A) | |
| | | Mitotic checkpoint serine/threonine-protein kinase (BUB1) | |
| | | Neuron-specific vesicular protein calcyon (DRD1IP) | MANTRA (Drugbank) |
| | | CaM inhibition | [92] |
| **Tyrphostin AG-1478** | 0.565 | EGFR inhibition | KIBA method |
| | | MAP kinase p38 alpha (MAPK14) | |
| | | MAPK-interacting serine/threonine-protein kinase MNK1 (MKNK1) | |
| **Cytochalasin B** | 0.568 | Src inhibition | [93] |
| **Celastrol** | 0.571 | Inhibition of NF-κB activation through IKKβ inhibition | [54] |
| | | More information on the many targets of celastrol | [94] |

[a] Here we list the known targets of the drugs and the sources of this target information. The full list of closest compounds to pinosylvin in terms of transcriptional response, as produced by the MANTRA method, can be found in Table S1 in the supplemental information online.

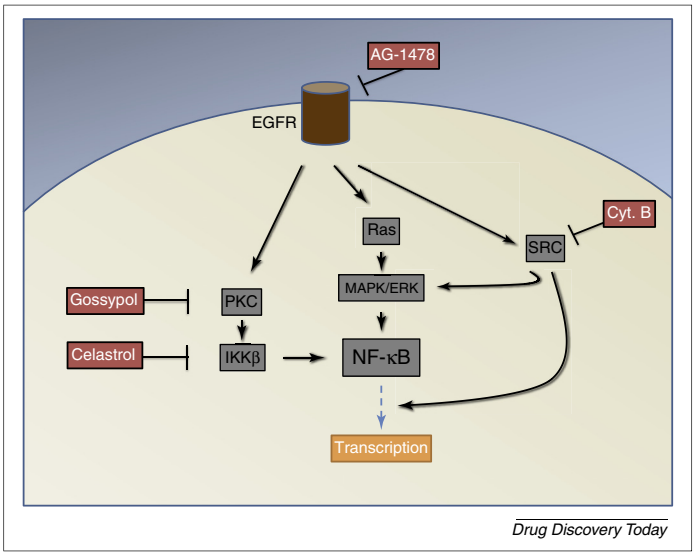


**FIGURE 1**

Pathway diagram showing relations between the cellular targets of four of the closest neighbors of pinosylvin. We mapped the known targets of the five most similar compounds in terms of their transcriptional response onto pathway diagrams using Pathway Commons (with Chembridge 5707885 replaced by the most similar drug having known targets, namely gossypol). This mapping led to the hypothesis that the transcriptional profiles of four of these five drugs, namely gossypol, tyrphostin AG-1478, cytochalasin B, and celastrol, are similar to that of pinosylvin because of the ultimate inhibition of nuclear factor kappa B (NF-κB). The target abbreviations are explained in Table 2 (main text).

apigenin and naringenin [58], it would be interesting to investigate whether this is true for pinosylvin and resveratrol. By contrast, genistein (4′,5,7-trihydroxyisoflavone) is a much-researched natural polyphenol that is structurally different to pinosylvin, and is not normally thought of as having similar mechanisms to the stilbenes. However, genistein is in Phase II clinical trials for prostate cancer (ClinicalTrials.gov Identifier: NCT01126879). Therefore, it is interesting to investigate how similar or different the lesser-researched polyphenol pinosylvin is to these much-researched polyphenols, resveratrol and genistein.

To investigate similarities between pinosylvin, resveratrol, and genistein, we mapped their closest compounds onto a network diagram, where links between compounds correspond to significant similarities in transcriptional response according to MANTRA (Box 2). First, the network approach used here highlighted unexpected similarities between pinosylvin and genistein. In particular, the transcriptional response that was most similar to that of genistein, was that of pinosylvin out of all 1309 CMap compounds, suggesting that a certain mode of action of pinosylvin is also common to genistein. Second, celastrol has links to pinosylvin and also to resveratrol and genistein. However, pinosylvin, resveratrol, and genistein are not among the top neighbors of celastrol. This suggests that a certain aspect of celastrol is common to each of these compounds. This aspect could be inhibition of NF-κB, because celastrol, resveratrol, and genistein are all known to inhibit NF-κB [54,59]. It is not unexpected that pinosylvin, genistein, and resveratrol should have links to celastrol, given that

natural polyphenols are known to have anti-inflammatory properties.

To highlight differences between compounds in a group, one can compare the lists of closest neighbors for each compound. For example, despite their structural similarity, resveratrol is not the most similar compound to pinosylvin in terms of transcriptional response; resveratrol is number 46 in the ordered list of compounds that are similar to pinosylvin. Therefore, we hypothesized that these two compounds work differently in some fundamental way. By looking at the drugs that are solely in the list of closest compounds to resveratrol (Table S2 in the supplemental information online), but not in the list for pinosylvin (Table S1 in the supplemental information online), we can search for clues as to potential differences between these two structurally similar compounds. For example, monobenzone occurs in the resveratrol list at position 21, but does not produce a transcriptional response similar to that of pinosylvin. Monobenzone, which is structurally similar to resveratrol, is a drug used to treat the loss of skin color (vitiligo). This hints that one difference between resveratrol and pinosylvin is connected to melanin synthesis or excretion [60].

## Operation of the GFA method

In contrast to MANTRA, the GFA method of Khan *et al.* [12] segregates the drugs and their responses into so-called 'components', which are cell line specific and group together drugs with similar transcriptional responses over a particular subset of genes that are learned in a data-driven fashion (Box 3). A compound can be linked to several components, making it possible to identify

**BOX 3**

### GFA method of Khan *et al.*

GFA [12,95] is a recent computational approach, based on the well-established factor analysis scheme, that decomposes transcriptional data sets into 'components' (also known as factors), which in turn can be used to create hypotheses for underlying biological processes. The method uses the gene expression response of each cell line separately to identify both the cross-cancer consensus effects as well as cancer-specific responses. Each component represents a distinct response in the data, and can be interpreted as a set of drugs, activating or repressing a set of genes in one or more cell lines. Therefore, the method can identify multiple mechanisms of each drug, while extracting a representative set of genes for each.

The GFA method uses the treatment versus control differential expression data directly into the modeling process. To this end, the raw gene expression data of CMap were preprocessed to obtain a $\log_2$FC differential response for both of the cell lines separately. Using the expression data of the most abundant platform in CMap, HG-U133A, a total of 1154 drugs were found in both MCF7 and PC3 cell lines. Multiple concentrations of the drugs were combined to obtain a single reliable profile for each drug [84]. The gene expression of pinosylvin was processed analogously to obtain the treatment versus control differential gene expression response. With a threshold of $|\log_2 FC| > 0.5$, 164 genes were identified as expressed in either MCF7 or PC3 by pinosylvin treatment. GFA was then run on the two matrices of responses, MCF7 and PC3; each comprising 1155 drugs (including pinosylvin) and 164 genes.

The analysis was done using the publically available R code of GFA (Table 1, main text). The model identified 30 components, 17 shared between MCF7 and PC3, with 13 responses specific to either of the cancer types. One shared component captured a pinosylvin response common on both MCF7 and PC3 (labeled #1), while one specific component (labeled #2) captured the MCF7-specific response of pinosylvin and related drugs. Each of these components is characterized by a set of drugs that regulate the expression of genes in a similar fashion, potentially forming hypotheses on the common action mechanisms of pinosylvin and correspondingly identified drugs. The remaining components captured responses of other drugs. Fig. III shows an example of a component in which pinosylvin is active, namely the MCF7-specific component 2, named the 'steroidal component' because it is highly enriched for steroidal compounds.

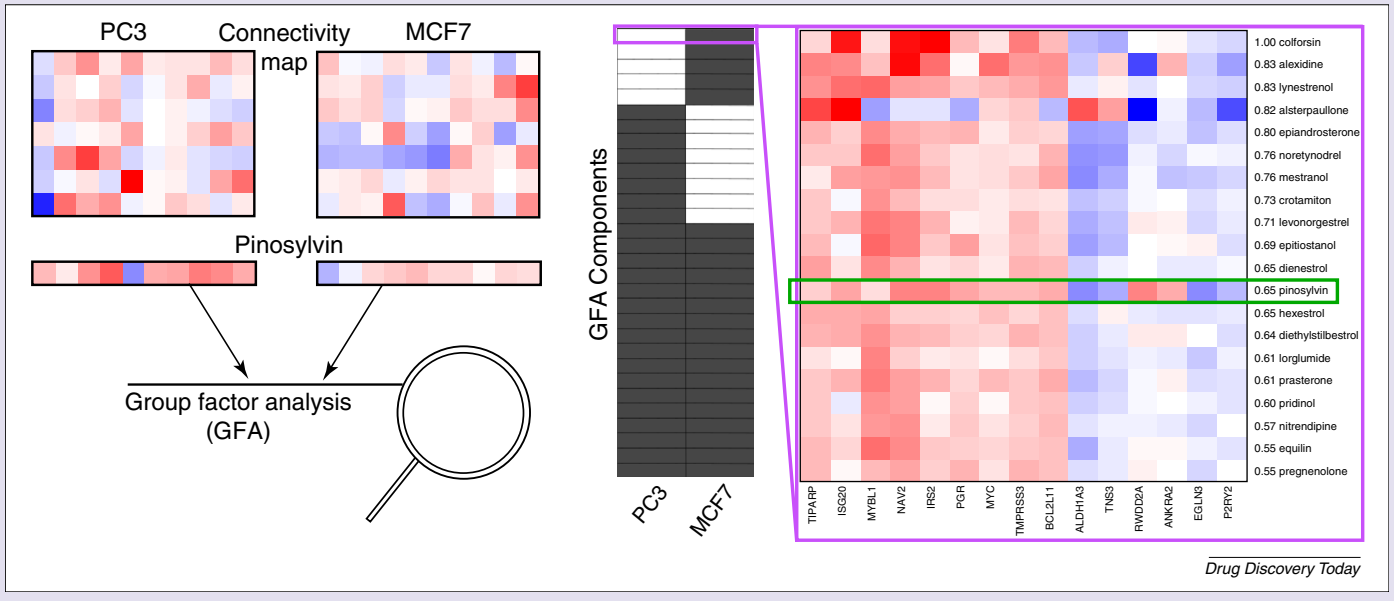


**FIGURE III**

Group Factor Analysis (GFA) decomposed the two gene expression response data sets spanning the common 1155 drugs into separate components.

multiple context-dependent MoAs for a compound. Moreover, because responses of drugs can be unique in various cancer types, GFA makes it possible to automatically identify which responses are common between cancers and which are specific to any one of the subtypes. The cell-specific responses might remain elusive when cell lines are merged, as is the standard choice in MANTRA. GFA also has the potential to use structural descriptors of compounds, such as 3D Pentacle descriptors, to identify the chemical properties linked to the responses. However, because there were structurally similar compounds to pinosylvin in the CMap data (e.g., resveratrol), we did not want the structural effects to dominate the findings and so did not include any structural descriptors in our analysis.

## Key GFA components

The GFA model identified 30 components from the CMap data combined with the pinosylvin expression profiles. Seventeen of these components were shared between MCF7 and PC3 cell lines, and 13 captured responses specific to either of the cancer types. Each of these components was characterized by a set of drugs that

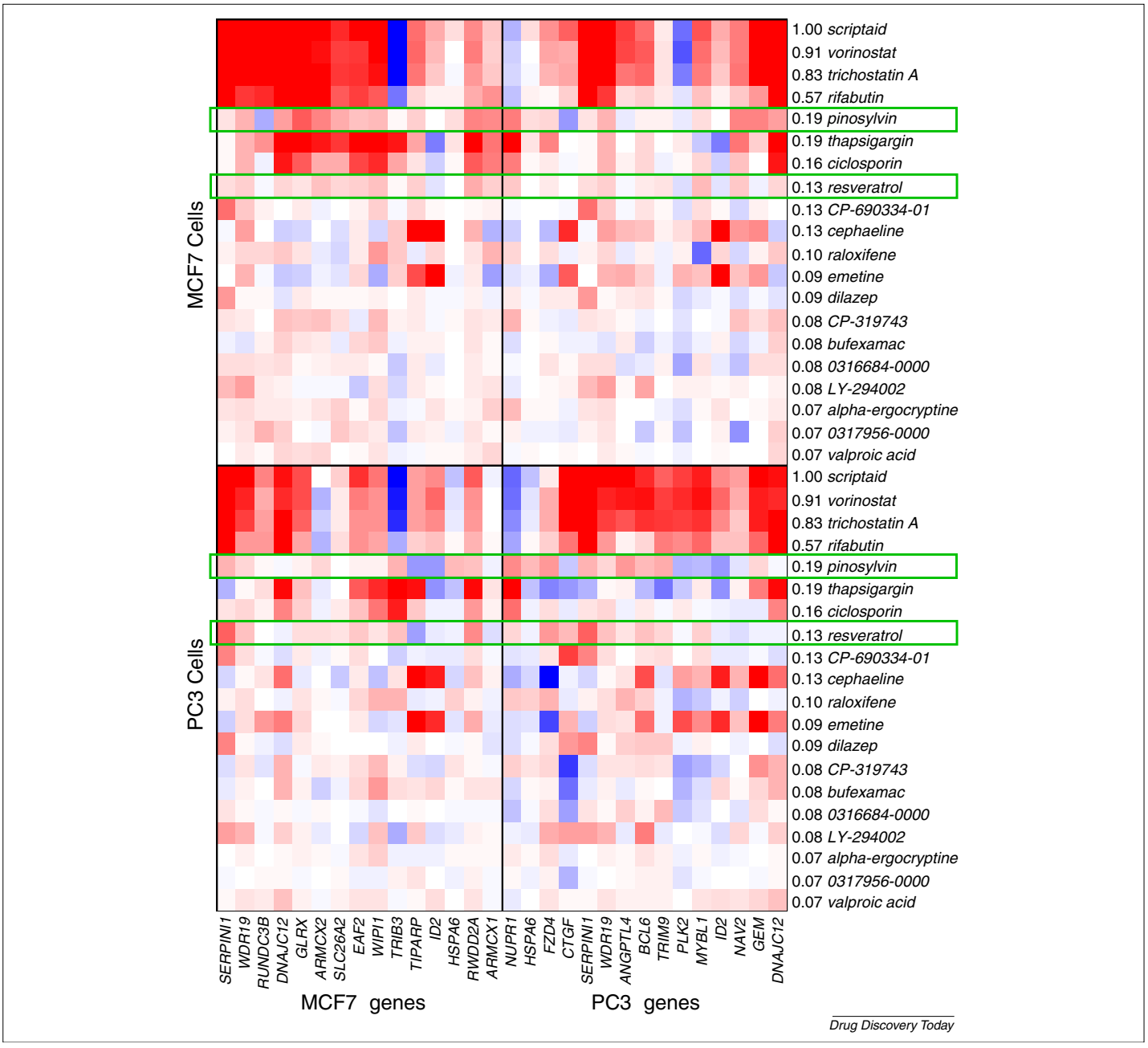


**FIGURE 2**

Component 1 'histone deactylase (HDAC) component' produced by the Group Factor Analysis (GFA) method. The GFA method segregates drugs and their transcriptional responses into so-called 'components', where each component is characterized by a set of drugs that regulate the expression of genes in a similar fashion, forming hypotheses on the common action mechanisms. Component 1 captured a response of pinosylvin common to both MCF7 and PC3 cell lines. This component is characterized by the three HDAC inhibitors, scriptaid, vorinostat, and trichostatin A, thereby predicting a HDAC response for pinosylvin. Red indicates genes that are upregulated and blue genes that are downregulated.

regulate the expression of genes in a similar fashion, potentially forming hypotheses on the common action mechanisms. Two components were relevant to pinosylvin, labeled as components 1 and 2, while the remaining components captured responses of other drugs.

Component 1 (Fig. 2) captured a response common on both MCF7 and PC3. This is a histone deacetylase (HDAC) inhibitor component characterized by the three HDAC inhibitors, scriptaid, vorinostat, and trichostatin A. Both pinosylvin and resveratrol are present in this component and, hence, as well as predicting a HDAC response for pinosylvin, such a role is also suggested for resveratrol (for which there is some prior evidence [61]). There is some evidence that HDAC inhibitors can reprogram the NF-κB response in cancer cells [62], linking this prediction with the NF-κB prediction above.

Component 2 (Box 3, Fig. III) captured the MCF7-specific response of pinosylvin and related drugs. The top drug in this component is colforsin, a water-soluble derivative of forskolin, which increases levels of cAMP. Cyclic AMP is an important signaling molecule in the regulation of prostate cancer [63] and breast cancer [64] cell proliferation. This led us to hypothesize that pinosylvin increases levels of cAMP. Interestingly, previous studies have shown that the structurally similar resveratrol increases the amount of cAMP by directly inhibiting cAMP-specific phosphodiesterases [65] (including PDE4, which is one of the most abundant cAMP-selective PDEs in PC3 cells [66] and also present in MCF7 [67]). Furthermore, Liu *et al.* [68] showed that genistein increases cAMP in insulin-secreting INS-1 cells, possibly primarily via enhanced adenylate cyclase activity.

Given that there were also numerous steroidal compounds in component 2, we hypothesized that pinosylvin also works in a manner similar to steroids. As regards steroid receptors, nuclear estrogen receptors (ERα and ERβ) are of particular interest, because they are known to interact with a range of steroidal and nonsteroidal compounds. However, if interaction with estrogen receptors was the common factor in component 2, then one would expect genistein to also be present in this component, because it is known to be an ERα agonist in MCF7 cells [69]. We continue this line of enquiry in the next section by taking into account the output also from the MANTRA method.

Taken together, the putative targets identified for pinosylvin (HDACs, GPCRs, and cAMP) are well in line with earlier data on the biological activities of pinosylvin, such as inhibition of cancer cell proliferation, induction of apoptosis, sensitization to TRAIL, and anti-inflammatory activities [21,49,70]. Further mechanistic studies are needed to confirm the relevance of these signaling molecules in mediating the anticarcinogenic and/or anti-inflammatory effects of pinosylvin. However, as regards to the putative steroid-like properties, our recent data indicate that pinosylvin, unlike genistein, does not target the classical nuclear steroid receptors ERα and the androgen receptor (AR) (L. Polari, PhD thesis, University of Turku, 2015, http://urn.fi/URN:ISBN:978-951-29-6101-6). Therefore, future studies focusing on the other steroid-related processes, such as steroid biosynthesis, or nongenomic steroid signaling, are warranted.

## Combining information from both methods leads to novel insights

Combining the GPCR hypothesis for pinosylvin obtained from the MANTRA method with the above observations on steroidal compounds in component 2 obtained from the GFA method, led us to question whether the common factor in component 2 could be the G-protein-coupled estrogen receptor (GPER), also known as GPR30, and to ask whether pinosylvin binds to GPER. Furthermore, might the antiproliferative and proapoptotic effects seen in prostate cancer occur via this mechanism? This would fit with the prediction, discussed in the previous section, that pinosylvin increases cAMP, because GPER agonists have been shown to increase cAMP production [71]. In addition, GPER belongs to the same GPCR family as 5HT-2, which was highlighted as a potential target by the MANTRA method. It is also known that GPER is expressed both in MCF7 [72] and PC3 cells [73].

GPER is a challenging target to validate, because there are no specific markers for GPER-mediated signaling. Furthermore, GPER ligand-binding studies are far from straightforward [71]. However, there is evidence that GPER is a relevant target in cancers, in particular for breast cancer and castration-resistant prostate cancer. GPER is expressed in clinical breast and prostate cancer specimens, and the GPER-selective agonist G1 modulates the growth of breast and prostate cancer cells *in vitro* and *in vivo* xenografts [71,73,74]. The polyphenols genistein and resveratrol have both been shown to have a GPER-related mechanism, genistein in certain cancer cell lines, [71] and, therefore, one can also hypothesize such a role for pinosylvin. However, a multitude of MoAs has been proposed for these two well-researched polyphenols [75,76], and it is not clear *a priori* which of them apply to pinosylvin.

The current data-driven exploration highlighted GPER as a potential target for pinosylvin, and the existing studies provided supporting published evidence for this novel hypothesis in the contexts of both prostate and breast cancer. In conclusion, by combining the hypotheses from the two computational models, we developed a hypothesis that would not have been obtained from either method alone.

## Concluding remarks and future directions

The elucidation of compound MoA is a key part of the drug discovery process. We have provided an overview of the common concepts in the computational network pharmacology methods, which use transcriptional response profiles to provide hypotheses for the MoA of compounds. The specific case study described here highlights how the data-driven approaches can provide researchers with completely unintuitive hypotheses for MoA, thus potentiating novel findings, once carefully validated using *in vitro* or *in vivo* assays. In addition, our case study adds to evidence that the network-based methods are particularly useful for natural products, which typically are multitargeted [20]. However, as demonstrated over the course of the application case study, experimental validation of the model predictions can be far from trivial, and establishing the MoA of a compound might require several iterations of the computational and experimental phases.

Using the MANTRA and GFA methods, we developed hypotheses for the MoA of the less-researched natural polyphenol pinosylvin. In particular, the MANTRA method predicted a GPCR-related mechanism and CaM inhibition, although this remained inconclusive following our initial experiments, and inhibition of NF-κB, for which we found literature evidence. The MANTRA method also hinted at commonalities between pinosylvin and genistein, which are stronger than those between

pinosylvin and a structurally more similar polyphenol, resveratrol. Using the GFA method, a HDAC response was predicted for pinosylvin, as well as an increase in levels of cAMP and steroid-like activity. By combining the information from these two computational methods, we were able to predict a GPER-mediated role for pinosylvin in prostate cancer. The question of whether natural stilbenoids have an effect on prostate cancer via a GPCR-related mechanism has not been posed previously, underscoring the potential of such unbiased approaches to suggest novel hypotheses.

In the future, computational network pharmacology methods will benefit from new phenotypic response data, for example that being produced by the Library of Integrated Network-based Cellular Signatures project (LINCS, http://www.lincsproject.org/), as well as high-quality, context-specific gene regulatory networks and more reliable drug–target and pathway information. Although novel computational methods are simultaneously being created, it is useful to also consider novel applications and combinations of existing methods and their extensions, as highlighted through the current work. For example, Iorio *et al.* [77] recently extended the MANTRA computational pipeline to filter out the effect of transcriptional changes resulting from nonspecific secondary MoA. In particular, they used an iterative, network-guided, semisupervised approach to refine the gene expression signature of the compound of interest to produce a transcriptional signature representative of the primary MoA. They showed how the approach can be used to disentangle the indirect mitotic arrest and general microtubule disruption effects of paclitaxel from its primary microtubule stabilization effect, and further used the method to find novel microtubule stabilizers [77].

MANTRA and GFA are both designed to identify similarities between drugs based on their large-scale phenotypic response profiles. Further advances in computational approaches might benefit cases where the goal is to illuminate differences between two drugs. For instance, a possible future direction would be to explore how these methods can be extended for optimally identifying differences between response mechanisms of two candidate drugs. Furthermore, future efforts to consider multiple levels of MoA for the neighboring compounds in the MANTRA method, beyond direct target effects to also include indirect and downstream signaling effects, might yield a wider understanding of the various drug response pathways. Finally, the GFA method learns the response similarities in an entirely data-driven fashion. Supplementing the data-driven nature with prior information of known pathway interactions and biochemical processes might prove useful for prediction of action mechanisms.

The concept of network pharmacology not only considers single multitargeted drugs to be of potential use to treat complex diseases, but also combinations of compounds, the rationale being that a combination of drugs might target nodes on compensatory pathways, countering problems such as emerging drug resistance. We did not touch on approaches for predicting effective drug combinations here, but computational methods that can prioritize the most therapeutically effective combinations of compounds for experimental validation are important because of the impracticality of testing all possible drug combinations, given that the number of combinations increases exponentially with the number of drugs to be screened. These methods are reviewed elsewhere (e.g., [78,79]).

We hope that, through this detective story into the MoA of pinosylvin, we have furnished researchers new to computational network pharmacology methods with the background to carry out similar studies for their own compounds of interest, while providing more experienced researchers in the field with novel insights into how the methods can be applied and their results experimentally validated.

## Acknowledgments

The authors thank Krister Wennerberg and Jing Tang for many useful discussions on this work, and especially Jing Tang for his help in finding the targets of the MANTRA neighboring compounds of pinosylvin. We would also like to thank Diego Carrella for his help in implementing the MANTRA tool.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.drudis.2016.03.001.

## References

1 Swinney, D.C. (2013) Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin. Pharmacol. Ther.* 93, 299–301
2 Capdeville, R. *et al.* (2002) Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat. Rev. Drug Discov.* 1, 493–502
3 Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* 3, 711–716
4 Hutchinson, L. and Kirk, R. (2011) High drug attrition rates: where are we going wrong? *Nat. Rev. Clin. Oncol.* 8, 189–190
5 Arrowsmith, J. (2012) A decade of change. *Nat. Rev. Drug Discov.* 11, 17–18
6 Waring, M.J. *et al.* (2015) An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discov.* 14, 475–486
7 Hopkins, A.L. (2007) Network pharmacology. *Nat. Biotechnol.* 25, 1110–1111
8 Kitano, H. (2007) A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* 6, 202–210
9 Lehár, J. *et al.* (2009) Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.* 27, 659–666
10 Lamb, J. *et al.* (2006) The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929–1935
11 Iorio, F. *et al.* (2010) Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14621–14626
12 Khan, S.A. *et al.* (2014) Identification of structural features in chemicals associated with cancer drug response: a systematic data-driven analysis. *Bioinformatics* 30, i497–i504
13 Li, J.W-H. and Vederas, J.C. (2009) Drug discovery and natural products: end of an era or an endless frontier? *Science* 325, 161–165
14 Tao, L. *et al.* (2014) Nature's contribution to today's pharmacopeia. *Nat. Biotechnol.* 32, 979–980
15 Harvey, A.L. *et al.* (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* 14, 111–129
16 Shen, B. (2015) A new golden age of natural products drug discovery. *Cell* 163, 1297–1300
17 Buriani, A. *et al.* (2012) Omic techniques in systems biology approaches to traditional Chinese medicine research: present and future. *J. Ethnopharmacol.* 140, 535–544
18 Azmi, A.S. *et al.* (2012) Can network pharmacology rescue neutraceutical cancer research? *Drug Discov. Today* 17, 807–809
19 Pelkonen, O. *et al.* (2014) Why is research on herbal medicinal products important and how can we improve its quality? *J. Tradit. Complement. Med.* 4, 1–7
20 Kibble, M. *et al.* (2015) Network pharmacology applications to map the unexplored target space and therapeutic potential of natural products. *Nat. Prod. Rep.* 32, 1249–1266

21 Yatkin, E. *et al.* (2014) Novel lignan and stilbenoid mixture shows anticarcinogenic efficacy in preclinical pC-3M-luc2 prostate cancer model. *PLoS ONE* 9, e93764
22 Subramanian, A. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550
23 Iorio, F. *et al.* (2013) Transcriptional data: a new gateway to drug repositioning? *Drug Discov. Today* 18, 350–357
24 Qu, X.A. and Rajpal, D.K. (2012) Applications of Connectivity Map in drug discovery and development. *Drug Discov. Today* 17, 1289–1298
25 Jahchan, N.S. *et al.* (2013) A drug repositioning approach identifies tricyclic antidepressants as inhibitors of small cell lung cancer and other neuroendocrine tumors. *Cancer Discov.* 3, 1364–1377
26 Laenen, G. *et al.* (2013) Finding the targets of a drug by integration of gene expression data with a protein interaction network. *Mol. Biosyst.* 9, 1676–1685
27 Wu, Z. *et al.* (2013) Drug repositioning framework by incorporating functional information. *IET Syst. Biol.* 7, 188–194
28 Quan, Y. *et al.* (2014) Elucidating pharmacological mechanisms of natural medicines by biclustering analysis of the gene expression profile: a case study on curcumin and Si-Wu-Tang. *Int. J. Mol. Sci.* 16, 510–520
29 Iskar, M. *et al.* (2013) Characterization of drug-induced transcriptional modules: towards drug repositioning and functional understanding. *Mol. Syst. Biol.* 9, 662
30 Jin, G. *et al.* (2012) A novel method of transcriptional response analysis to facilitate drug repositioning for cancer therapy. *Cancer Res.* 72, 33–44
31 Parkkinen, J.A. and Kaski, S. (2014) Probabilistic drug connectivity mapping. *BMC Bioinformatics* 15, 113
32 Pritchard, J.R. *et al.* (2013) Predicting cancer drug mechanisms of action using molecular network signatures. *Mol. Biosyst.* 9, 1604–1619
33 Napolitano, F. *et al.* (2016) Drug-set enrichment analysis: a novel tool to investigate drug mode of action. *Bioinformatics* 32, 235–241
34 Hegde, R.N. *et al.* (2016) Unravelling druggable signalling networks that control F508del-CFTR proteostasis. *eLife* 4, e10365
35 Woo, J.H. *et al.* (2015) Elucidating Compound mechanism of action by network perturbation analysis. *Cell* 162, 441–451
36 Kullback, S. and Leibler, R.A. (1951) On information and sufficiency. *Ann. Math. Stat.* 22, 79–86
37 Cichonska, A. *et al.* (2015) Identification of drug candidates and repurposing opportunities through compound-target interaction networks. *Expert Opin. Drug Discov.* 10, 1333–1345
38 Tang, J. *et al.* (2014) Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. *J. Chem. Inf. Model.* 54, 735–743
39 Huang, Y.-W. *et al.* (2006) Molecular mechanisms of (−)-gossypol-induced apoptosis in human prostate cancer cells. *Anticancer Res.* 26, 1925–1933
40 Volate, S.R. *et al.* (2010) Gossypol induces apoptosis by activating p53 in prostate cancer cells and prostate tumor-initiating cells. *Mol. Cancer Ther.* 9, 461–470
41 Chen, M.-H. *et al.* (2011) Gene expression-based chemical genomics identifies potential therapeutic drugs in hepatocellular carcinoma. *PLoS ONE* 6, e27186
42 Gil-Ad, I. *et al.* (2006) Phenothiazines induce apoptosis in a B16 mouse melanoma cell line and attenuate *in vivo* melanoma tumor growth. *Oncol. Rep.* 15, 107–112
43 Dorsam, R.T. and Gutkind, J.S. (2007) G-protein-coupled receptors and cancer. *Nat. Rev. Cancer* 7, 79–94
44 Berchtold, M.W. and Villalobo, A. (2014) The many faces of calmodulin in cell proliferation, programmed cell death, autophagy, and cancer. *Biochim. Biophys. Acta* 1843, 398–435
45 Dizeyi, N. *et al.* (2005) Expression of serotonin receptors 2B and 4 in human prostate cancer tissue and effects of their antagonists on prostate cancer cell lines. *Eur. Urol.* 47, 895–900
46 MacLean-Fletcher, S. and Pollard, T.D. (1980) Mechanism of action of cytochalasin B on actin. *Cell* 20, 329–341
47 Dodou, K. *et al.* (2005) Investigations on gossypol: past and present developments. *Expert Opin. Investig. Drugs* 14, 1419–1434
48 Lee, J. *et al.* (2006) Involvement of nuclear factor-kappaB in the inhibition of pro-inflammatory mediators by pinosylvin. *Planta Med.* 72, 801–806
49 Laavola, M. *et al.* (2015) Pinosylvin and monomethylpinosylvin, constituents of an extract from the knot of Pinus sylvestris, reduce inflammatory gene expression and inflammatory responses *in vivo*. *J. Agric. Food Chem.* 63, 3445–3453
50 Deeb, D. *et al.* (2004) Curcumin sensitizes prostate cancer cells to tumor necrosis factor-related apoptosis-inducing ligand/Apo2L by inhibiting nuclear factor-kappaB through suppression of IkappaBalpha phosphorylation. *Mol. Cancer Ther.* 3, 803–812
51 Allison, A.C. *et al.* (2001) Celastrol, a potent antioxidant and anti-inflammatory drug, as a possible treatment for Alzheimer's disease. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 25, 1341–1357
52 Boridy, S. *et al.* (2014) Celastrol targets proteostasis and acts synergistically with a heat-shock protein 90 inhibitor to kill human glioblastoma cells. *Cell Death Dis.* 5, e1216
53 Sha, M. *et al.* (2013) Celastrol induces apoptosis of gastric cancer cells by miR-146a inhibition of NF-κB activity. *Cancer Cell Int.* 13, 50
54 Lee, J.-H. *et al.* (2006) Inhibition of NF-kappa B activation through targeting I kappa B kinase by celastrol, a quinone methide triterpenoid. *Biochem. Pharmacol.* 72, 1311–1321
55 Tomé-Carneiro, J. *et al.* (2013) Resveratrol and clinical trials: the crossroad from *in vitro* studies to human evidence. *Curr. Pharm. Des.* 19, 6064–6093
56 Chowdhury, S.A. *et al.* (2005) Tumor-specificity and apoptosis-inducing activity of stilbenes and flavonoids. *Anticancer Res.* 25, 2055–2063
57 Carter, L.G. *et al.* (2014) Resveratrol and cancer: focus on *in vivo* evidence. *Endocr. Relat. Cancer* 21, R209–R225
58 Arango, D. *et al.* (2013) Molecular basis for the action of a dietary flavonoid revealed by the comprehensive identification of apigenin human targets. *Proc. Natl. Acad. Sci. U. S. A.* 110, E2153–E2162
59 Nam, N.-H. (2006) Naturally occurring NF-kappaB inhibitors. *Mini Rev. Med. Chem.* 6, 945–951
60 Lee, T.H. *et al.* (2014) Inhibitory effects of resveratrol on melanin synthesis in ultraviolet B-induced pigmentation in Guinea pig skin. *Biomol. Ther.* 22, 35–40
61 Vahid, F. *et al.* (2015) The role dietary of bioactive compounds on the regulation of histone acetylases and deacetylases: a review. *Gene* 562, 8–15
62 Place, R.F. *et al.* (2005) HDAC inhibition prevents NF-kappa B activation by suppressing proteasome activity: down-regulation of proteasome subunit expression stabilizes I kappa B alpha. *Biochem. Pharmacol.* 70, 394–406
63 Merkle, D. and Hoffmann, R. (2011) Roles of cAMP and cAMP-dependent protein kinase in the progression of prostate cancer: cross-talk with the androgen receptor. *Cell. Signal.* 23, 507–515
64 Zivadinovic, D. *et al.* (2005) Membrane estrogen receptor-alpha levels in MCF-7 breast cancer cells predict cAMP and proliferation responses. *Breast Cancer Res.* 7, R101–R112
65 Park, S.-J. *et al.* (2012) Resveratrol ameliorates aging-related metabolic phenotypes by inhibiting cAMP phosphodiesterases. *Cell* 148, 421–433
66 Hamilton, T.K. *et al.* (2013) Potential therapeutic applications of phosphodiesterase inhibition in prostate cancer. *World J. Urol.* 31, 325–330
67 Drees, M. *et al.* (1993) 3′,5′-Cyclic nucleotide phosphodiesterase in tumor cells as potential target for tumor growth inhibition. *Cancer Res.* 53, 3058–3061
68 Liu, D. *et al.* (2006) Genistein acutely stimulates insulin secretion in pancreatic beta-cells through a cAMP-dependent protein kinase pathway. *Diabetes* 55, 1043–1050
69 Hsieh, C.-Y. *et al.* (1998) Estrogenic effects of genistein on the growth of estrogen receptor-positive human breast cancer (MCF-7) cells *in vitro* and *in vivo*. *Cancer Res.* 58, 3833–3838
70 Park, E.-J. *et al.* (2013) Suppression of Src/ERK and GSK-3/β-catenin signaling by pinosylvin inhibits the growth of human colorectal cancer cells. *Food Chem. Toxicol.* 55, 424–433
71 Prossnitz, E.R. and Arterburn, J.B. (2015) International Union of Basic and Clinical Pharmacology. XCVII. G protein-coupled estrogen receptor and its pharmacologic modulators. *Pharmacol. Rev.* 67, 505–540
72 Ahola, T.M. *et al.* (2002) Progestin upregulates G-protein-coupled receptor 30 in breast cancer cells. *Eur. J. Biochem.* 269, 2485–2490
73 Chan, Q.K.Y. *et al.* (2010) Activation of GPR30 inhibits the growth of prostate cancer cells through sustained activation of Erk1/2, c-jun/c-fos-dependent upregulation of p21, and induction of G(2) cell-cycle arrest. *Cell Death Differ.* 17, 1511–1523
74 Lam, H.-M. *et al.* (2014) Targeting GPR30 with G-1: a new therapeutic target for castration-resistant prostate cancer. *Endocr. Relat. Cancer* 21, 903–914
75 Mahmoud, A.M. *et al.* (2014) Soy isoflavones and prostate cancer: a review of molecular mechanisms. *J. Steroid Biochem. Mol. Biol.* 140, 116–132
76 Han, G. *et al.* (2015) Anti-tumor effects and cellular mechanisms of resveratrol. *Drug Discov. Ther.* 9, 1–12
77 Iorio, F. *et al.* (2015) A semi-supervised approach for refining transcriptional signatures of drug response and repositioning predictions. *PLoS ONE* 10, e0139446
78 Ryall, K.A. and Tan, A.C. (2015) Systems biology approaches for advancing the discovery of effective drug combinations. *J. Cheminformatics* 7, 7
79 Bulusu, K.C. *et al.* (2015) Modelling of compound combination effects and applications to efficacy and toxicity: state-of-the-art, challenges and perspectives. *Drug Discov. Today* 21, 225–238
80 Tang, J. *et al.* (2015) What is synergy? The Saariselkä agreement revisited. *Exp. Pharmacol. Drug Discov.* 6, 181
81 Iorio, F. *et al.* (2009) Identifying network of drug mode of action by gene expression profiling. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* 16, 241–251
82 Frey, B.J. and Dueck, D. (2007) Clustering by passing messages between data Points. *Science* 315, 972–976

83 Shannon, P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504
84 Khan, S.A. *et al.* (2012) Comprehensive data-driven analysis of the impact of chemoinformatic structure on the genome-wide biological response profiles of cancer cells to 1159 drugs. *BMC Bioinformatics* 13, 112
85 Carrella, D. *et al.* (2014) Mantra 2.0: an online collaborative resource for drug mode of action and repurposing by network analysis. *Bioinformatics* 30, 1787–1788
86 Gaulton, A. *et al.* (2012) ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 40, D1100–D1107
87 Wishart, D.S. *et al.* (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 34, D668–D672
88 Huang, N. *et al.* (2006) Benchmarking sets for molecular docking. *J. Med. Chem.* 49, 6789–6801
89 Irwin, J.J. and Shoichet, B.K. (2005) ZINC: a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* 45, 177–182
90 Cerami, E.G. *et al.* (2011) Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.* 39, D685–D690
91 Jarvis, W.D. *et al.* (1994) Induction of apoptotic DNA fragmentation and cell death in HL-60 human promyelocytic leukemia cells by pharmacological inhibitors of protein kinase C. *Cancer Res.* 54, 1707–1714
92 Finlayson, A.E. and Freeman, K.W. (2009) A cell motility screen reveals role for MARCKS-related protein in adherens junction formation and tumorigenesis. *PLoS ONE* 4, e7833
93 Kim, M.-Y. *et al.* (2014) Cytochalasin B modulates macrophage-mediated inflammatory responses. *Biomol. Ther.* 22, 295–300
94 Kannaiyan, R. *et al.* (2011) Molecular targets of celastrol derived from Thunder of God Vine: potential role in the treatment of inflammatory disorders and cancer. *Cancer Lett.* 303, 9–20
95 Virtanen, S. *et al.* (2012) Bayesian group factor analysis. *Proc. AISTATS. J. Mach. Learn. Res. W&CP* 22, 1269–1277