

RESEARCH

Open Access

Predicting targeted drug combinations based on Pareto optimal patterns of coexpression network connectivity

Nadia M Penrod¹, Casey S Greene^{2,3} and Jason H Moore^{2,3*}

Abstract

Background: Molecularly targeted drugs promise a safer and more effective treatment modality than conventional chemotherapy for cancer patients. However, tumors are dynamic systems that readily adapt to these agents activating alternative survival pathways as they evolve resistant phenotypes. Combination therapies can overcome resistance but finding the optimal combinations efficiently presents a formidable challenge. Here we introduce a new paradigm for the design of combination therapy treatment strategies that exploits the tumor adaptive process to identify context-dependent essential genes as druggable targets.

Methods: We have developed a framework to mine high-throughput transcriptomic data, based on differential coexpression and Pareto optimization, to investigate drug-induced tumor adaptation. We use this approach to identify tumor-essential genes as druggable candidates. We apply our method to a set of ER⁺ breast tumor samples, collected before ($n = 58$) and after ($n = 60$) neoadjuvant treatment with the aromatase inhibitor letrozole, to prioritize genes as targets for combination therapy with letrozole treatment. We validate letrozole-induced tumor adaptation through coexpression and pathway analyses in an independent data set ($n = 18$).

Results: We find pervasive differential coexpression between the untreated and letrozole-treated tumor samples as evidence of letrozole-induced tumor adaptation. Based on patterns of coexpression, we identify ten genes as potential candidates for combination therapy with letrozole including EPCAM, a letrozole-induced essential gene and a target to which drugs have already been developed as cancer therapeutics. Through replication, we validate six letrozole-induced coexpression relationships and confirm the epithelial-to-mesenchymal transition as a process that is upregulated in the residual tumor samples following letrozole treatment.

Conclusions: To derive the greatest benefit from molecularly targeted drugs it is critical to design combination treatment strategies rationally. Incorporating knowledge of the tumor adaptation process into the design provides an opportunity to match targeted drugs to the evolving tumor phenotype and surmount resistance.

Background

A great deal of effort has been directed toward the identification of molecular targets that drive oncogenesis and the development of novel therapeutics that interact with these targets [1-6]. However, tumor cells have a remarkable ability to adapt to such treatments through functional redundancies and activation of compensatory signaling

pathways that enable them to tolerate the presence of targeted drugs. Thus, despite making important contributions to the treatment of cancer, the success of targeted therapies has been limited by resistance.

The predominant strategy for overcoming resistance is to combine drugs that act through ancillary mechanisms to block the functional redundancies and compensatory signaling pathways that serve as escape routes for cell survival. This strategy is supported by studies showing that complex networks, including the networks of molecular interactions that underlie biological function, are vulnerable to coordinated attacks at multiple targets [7,8], and

*Correspondence: jason.h.moore@dartmouth.edu

²Department of Genetics, Geisel School of Medicine at Dartmouth College, HB7937 One Medical Center Dr, Lebanon NH 03766, USA

³Institute for Quantitative Biomedical Sciences, Geisel School of Medicine at Dartmouth College, HB7937 One Medical Center Dr, Lebanon NH 03766, USA
Full list of author information is available at the end of the article

by functional genomics screens with RNA-mediated interference showing that cells can be increasingly sensitized to a molecularly targeted drug by inhibiting a second complementary target concurrently [9]. While this strategy is intuitive and may appear straightforward, selecting the best combination of targets to maximize tumor cell death while minimizing collateral damage and toxicity presents a tremendous challenge. Furthermore, it does not take into account the evolving tumor phenotype that emerges through the adaptation process in response to drug perturbation.

To address this challenge we have developed a framework to identify tumor-essential genes as potential drug targets by mining high-throughput transcriptomic data based on coexpression patterns where coexpression serves as a proxy for coregulation or participation in the same biological processes [10,11]. We apply this method to tumor samples taken from breast cancer patients undergoing preoperative letrozole treatment. This allows us to identify essential genes in the primary and residual tumors capturing changes in essentiality as the tumors adapt to the drug.

Letrozole is a non-steroidal aromatase inhibitor that binds competitively and reversibly to the aromatase enzyme and, in effect, inhibits the production of estrogen by blocking the conversion of androgens into estrogens. Estrogen regulates cell growth and differentiation influencing the development and progression of breast cancer by binding to and activating estrogen receptors (ERs). ERs participate in cell signaling and regulate gene expression through the activation or repression of gene transcription [12].

Letrozole is used neoadjuvantly to reduce the volume of large operable, locally advanced, and inoperable ER⁺ breast cancers in postmenopausal patients [13,14]. Efforts have been made to enhance the effects of letrozole by combining it with other drugs to reduce further tumor burden in responders and to develop effective treatment strategies for nonresponders [15-17]. To date, these combinations have led to only modest increases in clinical response. For example, combining letrozole with the mTOR inhibitor everolimus increases response rates, determined by clinical palpitation, at a moderately statistically significant level ($P = 0.062$) relative to letrozole treatment alone [15]. This indicates that the effects of letrozole can be enhanced by combining it with other molecularly targeted drugs, but it also suggests that there is room for improvement in choosing the most effective combinations for letrozole in this setting.

Here we assess patterns of differential coexpression among patient tumors sampled before and after letrozole treatment. Based on these coexpression patterns we identified tumor-essential genes and letrozole-induced tumor-essential genes as potential candidates for combination

therapy with neoadjuvant letrozole treatment. We show that coexpression is a suitable measure of tumor adaptation to drug perturbation by validating letrozole-induced coexpression relationships in an independent data set.

Methods

Data description

The initial analysis was performed with transcriptomic data generated from core biopsies of ER⁺ breast tumors at diagnosis ($n = 58$) and again following a 90-day course of neoadjuvant treatment with the drug letrozole ($n = 60$) [18,19]. Inclusion criteria required the samples to contain at least 20% malignant tissue. RNA was extracted, amplified, and hybridized to Affymetrix HG-U133A GeneChip arrays. The data are publicly available through the Gene Expression Omnibus (GEO) database [GEO:GSE20181].

An independent data set was used for replication. The replication data are also transcriptomic profiles generated from core biopsies of ER⁺ breast tumors at diagnosis ($n = 18$) and again following a 90-day course of neoadjuvant treatment with the drug letrozole ($n = 18$) [20]. Inclusion criteria required the samples to contain at least 50% malignant tissue. RNA was extracted, amplified, and hybridized to Affymetrix HG-U133 Plus 2.0 GeneChip arrays. The data are publicly available through GEO [GEO:GSE10281].

Data processing

We downloaded and processed the raw probe intensity (CEL) files for each data set independently. We used a custom chip definition file (CDF) to ensure we were using the most recent probe annotations and to filter the Affymetrix probe sets to include only those probes that uniquely map to genes [21]. Data were background corrected, normalized, and summarized using the robust multi-array average algorithm [22] as implemented in the R statistical language [23].

Differential expression analysis

Differential expression analysis between untreated and letrozole-treated tumor samples was conducted using the linear models for microarray data (limma) method [24] implemented in the limma package in R. We chose this method based on its robust performance across a variety of sample sizes and noise levels [25]. To correct for multiple hypothesis testing, genes at a false discovery rate (FDR) below 5% were considered differentially expressed at a statistically significant level. We performed coexpression analysis on the set of differentially expressed genes.

Differential coexpression analysis

Using the subset of genes found to be differentially expressed by letrozole treatment, we generated two sets of

coexpressed gene pairs, those that occur in the untreated tumor samples and those that occur in the letrozole-treated tumor samples. To identify coexpressed gene pairs, we calculated the first-order Spearman's partial correlation coefficients and associated P values [26] between the expression levels for each pairwise combination of genes. Spearman's correlation allows us to identify both linear and non-linear coexpression relationships and has been shown to outperform the more commonly applied Pearson's correlation coefficient at identifying coexpression relationships among genes within the same pathways and among functionally related transcription factors [27].

Gene pairs with a coexpression P value that met an FDR-based significance threshold of $\alpha = 0.01$ were retained. This significance threshold was chosen based on simulations carried out by de la Fuente *et al.* [26]. To validate this threshold for selecting coexpressed gene pairs in our data, we used permutation testing to model the null hypothesis that there are no coexpression relationships among genes in these data sets (Additional file 1). Permutation tests were designed to randomize the expression values for each gene, across samples, within each time point. Following randomization, we calculated coexpression as described above and counted the number of partial correlation coefficients that met our significance threshold. This process was repeated 1,000 times to generate a null distribution. The observed numbers of significant coexpression relationships, for untreated and treated tumors in both data sets, fall to the right of the upper bound in the matched null distribution ($P < 0.001$) (Additional file 1) allowing us to reject the null hypothesis by showing that more gene pairs were coexpressed than would be expected by random chance when a significance threshold of $\alpha = 0.01$ is applied. The complete results of the coexpression analysis are presented in Additional files 2, 3, 4, 5, 6 and 7.

Annotating coexpressed gene pairs

We annotated each gene to the Gene Ontology (GO) biological process [28], Kyoto encyclopedia of genes and genomes (KEGG) [29], and Reactome [30] databases through Bioconductor. We found common processes and pathways by intersecting the annotations for each pair of genes.

We also evaluated each gene pair for functional relationships based on empirical data with networks from the Integrated Multi-species Prediction (IMP) web server [31]. These gene networks were generated as described in Park *et al.* [32] and integrate data sources that include wet biochemical evidence including the IntAct, MINT, MIPS, and BioGRID databases. In these gene networks, edges represent the posterior probability of a functional relationship between two genes. Therefore, each edge is interpretable as the posterior probability, given a large compendium of empirical data collected from

human-derived samples, that two genes work together to carry out a specific biological process. We overlaid our coexpressed gene pairs onto these networks to determine the likelihood that a functional relationship exists between the pairs of genes we identify.

For novel gene pairs that replicate, we used IMP to predict functional relationships directly and to identify bridging genes that connect coexpressed gene pairs. For this purpose we considered edges above a probability threshold of 0.70. This cutoff is stringent: only 0.042% of edges in the network (141,214 / 333,452,400) have sufficient evidence to place them above this threshold. Functional descriptions of the genes in the results were taken from GeneCards [33].

Pareto identification of tumor-essential genes

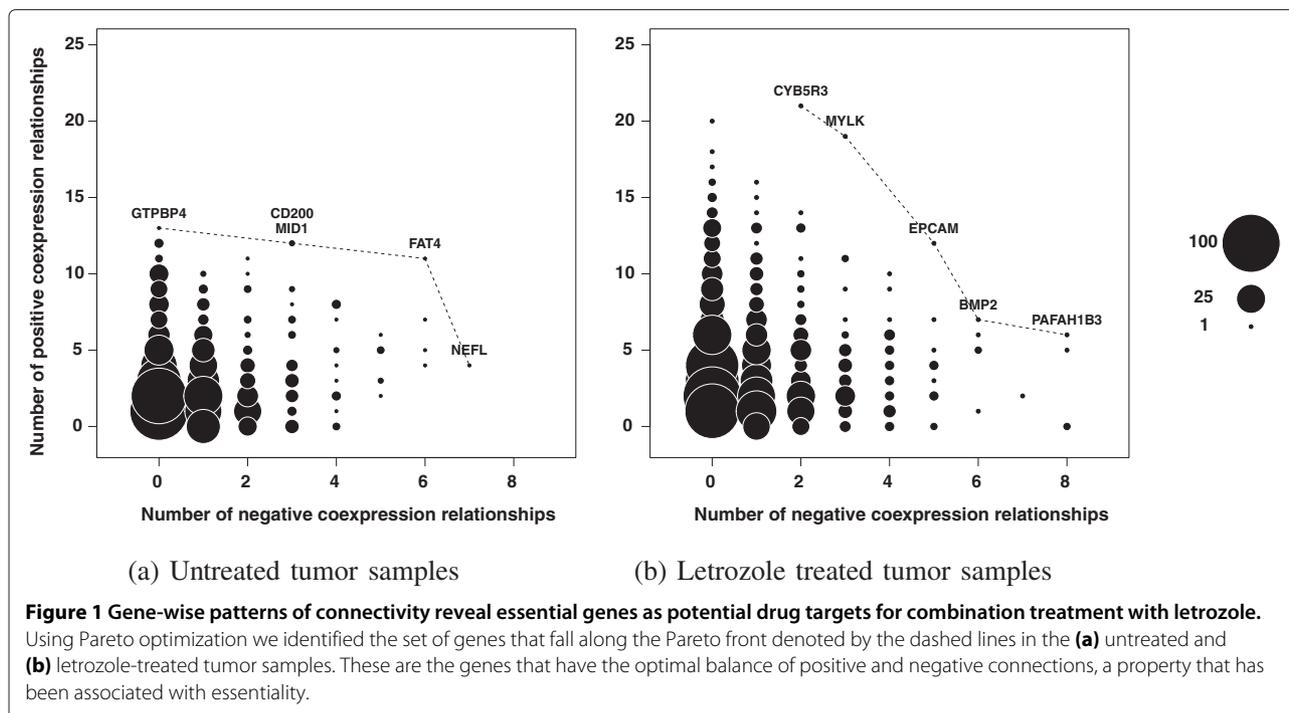
Studies in model organisms demonstrate that essential genes tend to have a combination of many positive and many negative genetic interactions [34]. Based on these findings we used coexpression as a proxy for coregulation and we identified essential genes as those that have many positively and many negatively coexpressed gene partners. This presents a multi-objective optimization problem because we were trying to maximize two variables, the number of positive partners and the number of negative partners, simultaneously. It is unlikely that a single gene will maximize both of these objectives, so instead of looking for a single solution, we used Pareto optimization, a multi-objective optimization algorithm, to identify the set of genes that most closely maximize both objectives. To illustrate this, we plotted the number of positive partners by the number of negative partners for each gene (Figure 1) and identified the genes that fall along the leading edge of the data, termed the Pareto front. The genes that lie along the Pareto front have more positively and negatively coexpressed gene partners than any gene falling to the left of this curve. We consider each of these genes to be essential and thus a potential drug target.

Results

Letrozole induces differential coexpression in ER⁺ breast tumors

Treatment with the aromatase inhibitor letrozole changes gene expression globally, resulting in a marked downregulation of genes involved in cell-cycle processes including mitosis and DNA metabolism and an upregulation of genes involved in wounding and immune responses, skin and vasculature development, and cell adhesion [19].

Building on this knowledge, using transcriptomic profiles for ER⁺ breast tumor biopsy samples collected before ($n = 58$) and after ($n = 60$) a course of neoadjuvant treatment with letrozole, we selected the subset of genes that are differentially expressed for coexpression analysis. These data allowed us to generate two snapshots of



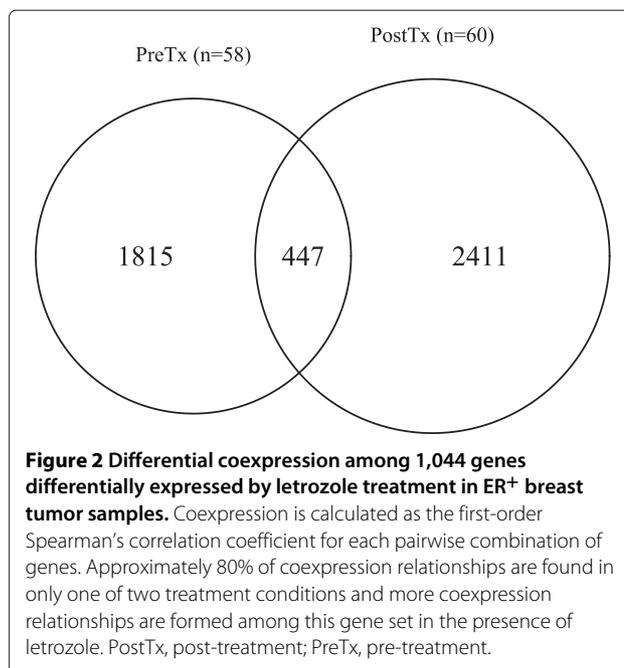
gene–gene relationships: those that occurred among these genes in the untreated tumors and those that occurred among these genes in the residual tumors, which have adapted to tolerate the presence of the drug. We defined coexpression as a statistically significant Spearman’s correlation coefficient in a partial correlation model. These specifications allowed us to find both linear and non-linear relationships and to focus on direct gene–gene relationships by excluding gene pairs that are coexpressed due to a common regulator.

We found considerable differential coexpression among genes between the untreated and letrozole-treated tumor samples. Approximately 80% of pairwise relationships occurred in only one of the two treatment conditions (Figure 2). Furthermore, we identified 1.26 times as many pairwise relationships in the letrozole-treated tumor samples as in the untreated tumor samples among the same set of genes. These dynamic coexpression relationships provide evidence of tumor adaptation emphasizing the context-dependent nature of gene–gene relationships and suggesting that the functional relationships among genes change as the tumors adapt to perturbation by the drug.

Each coexpressed gene pair has either a positive connectivity or a negative connectivity based on the sign of the correlation coefficient that connects the two genes. Gene pairs with positive connectivity have expression levels that are directly correlated. Gene pairs with negative connectivity have expression levels that are inversely correlated. In agreement with previous work on coexpression analysis among human genes [35], we identified more

coexpression relationships of positive connectivity than negative connectivity in both the untreated and letrozole-treated tumor samples.

Among genes that have sustained coexpression relationships in both the untreated and letrozole-treated tumor samples, the connectivity patterns were conserved, indicating that the nature of the relationships between



these genes does not change in the presence of the drug. We removed these common connections leaving only those gene pairs that represent differential coexpression between the two treatment conditions for further analysis.

Pairwise coexpression relationships are supported by known biological evidence

To confirm that we had identified pairwise relationships with biological relevance we mapped each pair of genes to the Gene Ontology (GO), KEGG, and Reactome databases. We also looked for evidence of functional relationships by querying IMP, a web server that mines empirical data to provide a predictive probability that a pair of genes work together within a biological process. We found that 42% of the coexpressed gene pairs in untreated tumors and 45% of the coexpressed gene pairs in letrozole-treated tumors are supported by at least one of these sources of biological evidence.

Furthermore, we looked for evidence of the biological effects of drug treatment. Among the pathway and process databases, GO has the highest coverage for our gene set (88%) compared to KEGG (38%) and Reactome (35%). So we isolated GO biological process terms that are exclusively represented by gene pairs in the letrozole-treated tumor samples. We found that these processes correspond to both the intended effects and side effects of the drug (Additional file 8). Examples include decreased mitosis, bone density loss [36], hypercholesterolemia [36], arthralgia and myalgia [37,38].

Adaptive coexpression propounds druggable targets for combination therapy

Gene-wise analysis shows that, regardless of letrozole treatment status, most genes have only a few coexpression partners and a propensity toward relationships of positive connectivity while a few genes have many coexpression partners usually incorporating both positive and negative connectivities (Figure 1). In general, genes tend to form more coexpression relationships in the presence of the drug with a noticeable increase in the number of relationships of negative connectivity.

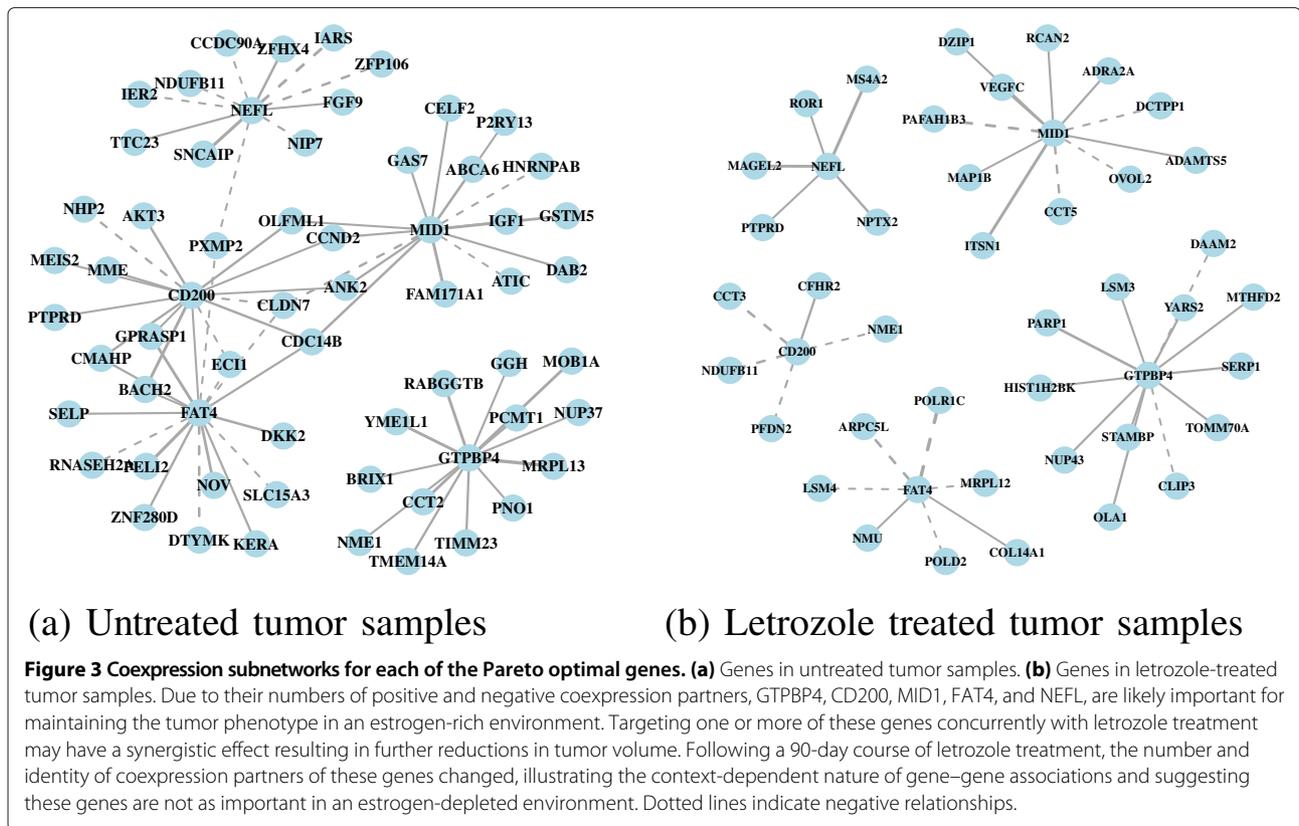
Our goal was to identify druggable targets that will synergize with neoadjuvant letrozole treatment. Our strategy was to identify the genes that have connectivity patterns consistent with those of essential genes because these are the points at which the tumors are likely to be vulnerable to a targeted attack. Based on empirical data showing a tendency for essential genes to form many relationships of both positive and negative connectivities, termed double connectivity [34,39], we used Pareto optimization (see Methods) to identify essential genes as those that maximize the numbers of positive and negative coexpression relationships, simultaneously.

We first identified genes with high double connectivity in the untreated tumors as genes that are likely important for maintaining the tumor phenotype in an estrogen-rich environment. This gene set includes the GTPase GTPBP4, the glycoprotein CD200, the microtubule-associated MID1, the cadherin FAT4, and the neurofilament NEFL (Figure 1a). We see context-dependent associations among these genes and their coexpression partners illustrated by the tendency to form more coexpression relationships prior to letrozole treatment and to associate with a different set of genes under each treatment condition (Figure 3).

Each gene in this set has the potential to be a druggable target. Targeting one or more of these genes concurrently with the inhibition of estrogen signaling, through letrozole treatment, has the potential to enhance letrozole's ability to reduce tumor volume. There is limited literature regarding the functional role of GTPBP4 in the context of cancer. One report suggests that inhibition of this gene could be effective by showing an inverse relationship between the expression level of GTPBP4 in breast tumors carrying wild-type p53 and patient survival [40].

The other four genes in this set share coexpression relationships to form a connected subnetwork, which suggests that targeting just one of these genes could effectively modulate the expression of the others. Based on their biological roles in the context of cancer, it appears that the inhibition of CD200 and FAT4 would be effective in reducing tumor volume. A series of studies demonstrated that overexpression of CD200 promotes tumor growth and metastasis of breast cancer in immunocompetent mice through suppression of the immune response, a process that can be reversed by treatment with an anti-CD200 monoclonal antibody [41,42]. FAT4 is a member of the Hippo signaling pathway and has been classified as a putative tumor suppressor in breast cancer [43], although, this is a context-dependent designation as it has also been shown to play roles in tumorigenesis and planar cell polarity [44], a process linked to metastasis. FAT4 has been recently implicated as a druggable target [45] but to date there are no drugs that specifically target this gene.

In contrast, the downregulation or loss of MID1 and NEFL has been associated with more aggressive disease. MID1 was recently shown to mediate the ubiquitin-dependent degradation of $\alpha 4$ [46], a regulator of mTOR and cell-cycle progression, which is highly expressed in breast cancer [47]. NEFL is an independent prognostic indicator of disease-free survival in early stage breast cancer where low expression correlates with worse outcome [48]. Our coexpression networks show that expression of FAT4 is positively correlated with CD200 and negatively correlated with both MID1 and NEFL suggesting that inhibition of FAT4 may be the optimal target for neoadjuvant co-treatment with letrozole because its modulation



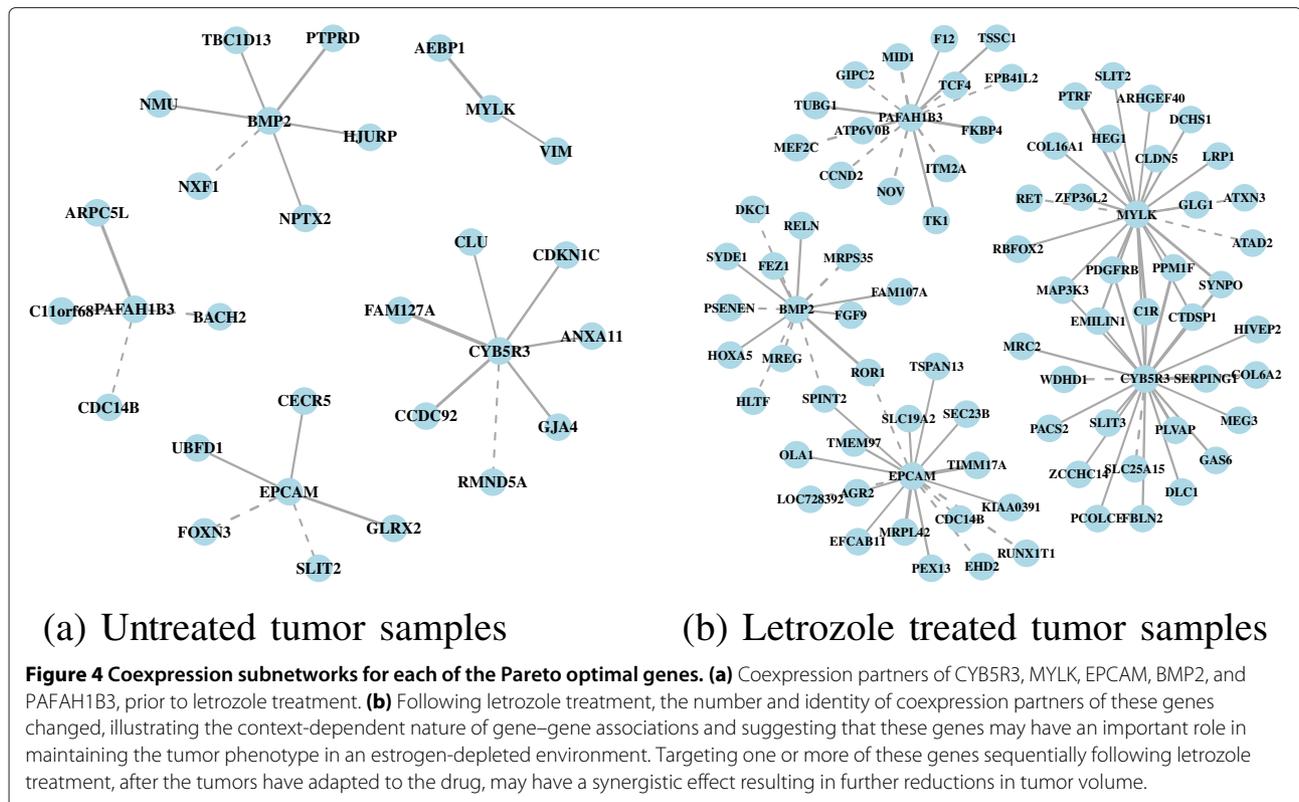
may downregulate CD200 while upregulating MID1 and NEFL.

We also identified genes with high double connectivity in the letrozole-treated tumors as genes that are essential in an estrogen-depleted environment. Targeting these genes sequentially after estrogen signaling has been inhibited by letrozole has the potential to reduce further tumor volume by blocking escape pathways as they emerge while the tumors try to adapt to the drug. In the letrozole-treated tumors, the essential gene set includes the enzyme CYB5R3, the kinase MYLK, the antigen EPCAM, the growth factor BMP2, and the acetylhydrolase PFAH1B3 (Figure 1b). These genes tend to have more coexpression partners following letrozole treatment relative to the untreated tumor samples (Figure 4). And again, the set of genes acting as coexpression partners differs following letrozole treatment, showing the context-dependent nature of gene–gene associations.

The expression levels of CYB5R3 and MYLK are positively correlated. The CYB5R3 gene plays a functional role in redox homeostasis by maintaining the balance of NAD⁺ /NADH within cells. It has been linked to cancer through its association with mitochondrial dysfunction [49]. Mitochondrial dysfunction promotes tumor growth in a condition-dependent manner [50]. MYLK is associated with breast tumor metastasis through *in vitro* studies showing its role in mediating migration and invasion of

the MDA-MB-231 cell-line [51] and the intravasation of breast cancer cells through an endothelial cell layer [52]. Inhibition of either of CYB5R3 or MYLK could be effective at halting tumor progression because the inhibition of one of these genes should modulate the expression of the other.

The expression levels of EPCAM and BMP2 are negatively correlated. EPCAM is a cell-adhesion molecule that has been associated with cell signaling, proliferation, differentiation, migration and metastasis [53] and used as a marker of both primary tumors and circulating tumor cells for patients with breast cancers and other endothelium-derived tumors [54,55]. Silencing EPCAM gene expression in breast tumor cell lines *in vitro* results in a dramatic decrease in metabolic activity, cell migration and invasion [56]. BMP2, a member of the TGF β superfamily, is a target gene of ER signaling, which is down-regulated in the presence of estrogen [57]. By forming a heterodimer with BMP7, BMP2 acts as a TGF β antagonist and prevents bone metastases in a mouse model of breast cancer [58]. Due to the negative coexpression relationship that exists between EPCAM and BMP2, inhibition of EPCAM may upregulate BMP2 to contribute to the prevention of metastasis. Notably, several drugs have been developed to target EPCAM as cancer therapeutics [59,60]. To our knowledge, the association between PFAH1B3 and breast cancer is novel.



Replication highlights biologically relevant and novel coexpression relationships

To determine if coexpression relationships induced by letrozole treatment are generalizable, we did a replication analysis with an independent data set. This data includes transcriptomic profiles for 18 ER⁺ breast tumor biopsy samples collected before and after a course of neoadjuvant treatment with letrozole. For consistency, we used only the subset of genes that were differentially expressed by letrozole treatment in both data sets resulting in a set of 263 genes for coexpression analysis. Confirming our earlier finding, there was patent differential coexpression among this set of genes for both data sets, with an increase in the number of pairwise relationships among genes in the letrozole-treated samples (Figure 5). With fewer samples in the replication data, we had limited statistical power to detect patterns of coexpression; however, those relationships that do replicate provide validation for letrozole-induced tumor adaptation.

We validated six gene–gene relationships induced by letrozole treatment (Table 1). One gene pair is supported by strong biological evidence and the other five gene pairs validate novel relationships. To attach functional meaning to these novel findings we generated functional subnetworks in IMP that incorporate additional genes to make connections between the coexpressed gene pairs. The first validated relationship is a positive connection between the

ribonucleotide reductase RRM2 and the DNA topoisomerase TOP2A, two genes that map to the DNA replication pathway. They have a high probability of a functional interaction in IMP (0.88) and are downregulated by letrozole treatment in agreement with the effects of blocking ER signaling [61].

The next two validated gene pairs involve a long non-coding RNA, LINC00341, of unknown function. LINC00341 is coexpressed with RUNX1T1, a proto-oncogene and transcriptional repressor, which interacts with DNA-bound transcription factors, and MEF2C, a transcription factor involved in myogenesis and muscle cell differentiation maintenance. Functionally, RUNX1T1 and MEF2C are linked through two intermediates, SIN3A and both HDAC4 and HDAC9, all of which are transcriptional repressors (Figure 6a). This suggests that LINC00341 is part of a complex that regulates transcription of MEF2C, a gene that has previously been shown to be regulated by long non-coding RNAs [62].

The remaining three validated gene pairs constitute a subnetwork representative of the epithelial-to-mesenchymal transition (EMT), a process associated with wound healing and metastasis (Figure 6b). The two key genes that complete the functional subnetwork among these gene pairs, the glycoproteins FN1 and SPARC, are bona fide markers of EMT [63,64]. FN1 functionally connects a pair of coexpressed glycoproteins, FBLN1 and

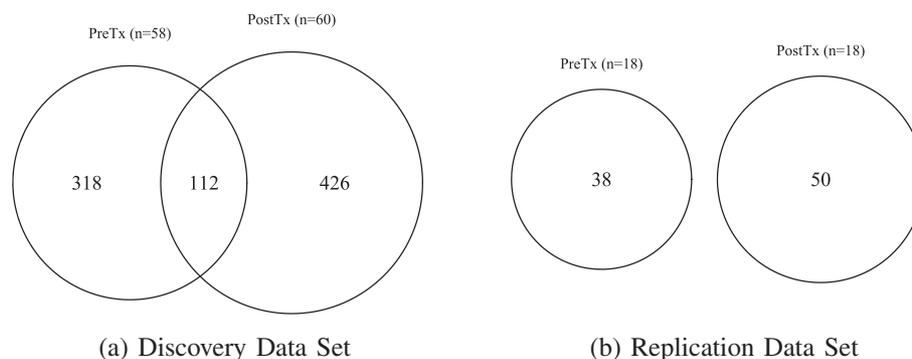


Figure 5 Differential coexpression among 263 genes differentially expressed by letrozole treatment. Two independent data sets of ER⁺ breast tumor samples were used. Coexpression was calculated as the first-order Spearman's correlation coefficient for each pairwise combination of genes. In both the (a) discovery and (b) validation data sets, most coexpression relationships are unique to one of the two treatment conditions with an increase in the number of letrozole-induced coexpressed gene pairs. PostTx, post-treatment; PreTx, pre-treatment.

FSTL1, and SPARC is connected to IGFBP7, a tumor suppressor, which creates a functional link between a candidate tumor suppressor, PDGFRL, and a mesenchymal factor, FSTL1. In addition to FN1 and SPARC, other well-established EMT-associated genes are also upregulated in these tumor samples following letrozole treatment including TWIST1, SNAI2, ZEB1, and ZEB2 [65]. We did not identify any of these replicated coexpression relationships in the untreated tumor samples as evidence that the residual tumor cells have undergone a functional reorganization during adaptation to tolerate the presence of the drug.

Discussion

Here, we introduced a method to prioritize genes that have coexpression and connectivity patterns consistent with those of essential genes [34] as potential drug targets in the design of rational combination therapies for the treatment of cancer. We applied this method to predict combination therapy targets based on the adaptive response of ER⁺ breast tumors to neoadjuvant treatment with the aromatase inhibitor letrozole. We used coexpression as a proxy for functional relationships and found that adaptation to drug perturbation is evident in the differential coexpression patterns we observed between the untreated and letrozole-treated tumor samples. This is consistent with previous work showing that functional relationships among genes are dependent on the cellular state and local environment and reflected in patterns of coexpression [10].

We confirmed that many of the coexpressed gene pairs we identified have known biological relevance, but we also found pairs that are not yet annotated to the same processes or pathways and do not yet have empirical evidence that predicts a functional relationship. Perhaps the most obvious reason for this is annotation bias, which occurs

because well-studied genes are assigned many annotations while the understudied genes may not be annotated at all [66,67]. In our analysis, 26% of the genes have one or fewer annotations. Presumably, many of these genes are multifunctional, serving to connect related biological pathways that will not be revealed through annotation analysis alone. This is one of the reasons we incorporated IMP as a discovery tool, to move beyond curated annotations to find functional relationships supported by empirical data.

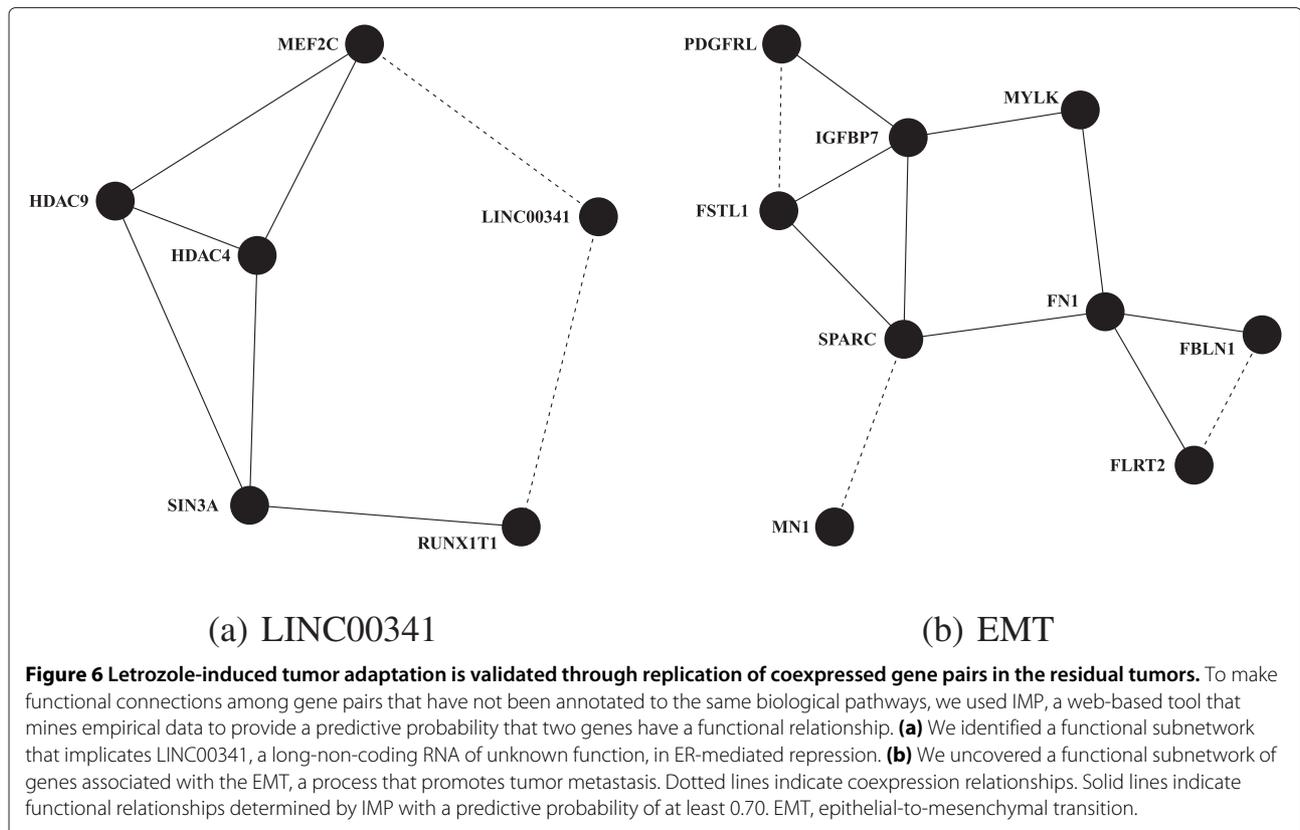
Repeated sampling of tumors before and after letrozole treatment allowed us to capture dynamic changes in gene expression and coexpression, illustrating changes in the functional relationships among genes that are induced by the drug. In this way, the adaptive response becomes a process that can be exploited to identify context-dependent targets. In total, we have identified ten Pareto optimal genes as potential targets for use in combination with letrozole. Of these genes, EPCAM stands out because opportunely, several monoclonal antibodies have already been developed against EPCAM as cancer therapeutics, including the well-tolerated, fully humanized version, adecatumumab [59]. Inhibition of EPCAM with adecatumumab has only been tested in patients with advanced disease. As a single agent, adecatumumab shows activity in metastatic breast cancer, but does not lead to tumor regression [68]. The combination of docetaxel and adecatumumab in a Phase IB trial achieved a clinical benefit, defined as a complete or partial response or stable disease, in 44% of patients with relapsed or refractory advanced-stage breast cancer [69].

Based on our findings, the addition of adecatumumab following an initial period of letrozole treatment should enhance the anti-tumor effects of letrozole alone. Suitably, recent trials have demonstrated that patients continue to derive a clinical benefit from neoadjuvant letrozole for

Table 1 Replication of letrozole-induced coexpression relationships

Gene 1	Gene 2	Gene 1 expression change	Gene 2 expression change	PCOR coefficient discovery data	PCOR coefficient replication data	GO term	KEGG term	REACTOME term	IMP
RRM2	TOP2A	↓	↓	0.44	0.90	DNA replication	-	-	0.88
LINC00341	RUNX1T1	↑	↑	0.34	0.78	-	-	-	-
MN1	SPARC	↑	↑	0.45	0.76	-	-	-	-
FBLN1	FLRT2	↑	↑	0.43	0.70	-	-	-	-
MEF2C	LINC00341	↑	↑	0.37	0.66	-	-	-	-
FSTL1	PDGFRL	↑	↑	0.47	0.63	-	-	-	0.10

We identified coexpression relationships among a set of 263 genes that were differentially expressed in two independent data sets of breast tumor biopsies following 90 days of letrozole treatment. The coexpressed gene pairs that validated are shown here with their pathway annotations and IMP score, which can be interpreted as the predictive probability that the two genes have a functional interaction based on a large compendium of empirical data. GO, Gene Ontology; IMP, Integrated Multi-species Prediction; KEGG, Kyoto encyclopedia of genes and genomes; PCOR, Partial correlation coefficient.



up to one year of treatment [70-73], making sequential therapy a fitting option. Moreover, metastasis is virtually prevented in mice when treated with a murine-specific version of adecatumumab [74], which suggests that this combination has the potential to be a long-term treatment strategy for the management of ER⁺ breast cancer as a chronic condition in elderly patients [75].

Despite differences in inclusion criteria and the limited sample size of the replication data, we were able to replicate six letrozole-induced coexpression relationships as validation of letrozole-induced adaptation. Two of the novel relationships that replicate provide clues about the function of the uncharacterized long non-coding RNA LINC00341. We have shown that LINC00341 is coexpressed with both RUNX1T1 and MEF2C (Figure 6a). RUNX1T1 is part of a corepressor complex that interacts with SIN3A *in vivo* [76]. SIN3A interacts with HDACs 4 and 9, specifically binding the catalytic domain of HDAC 9 in cells derived from B-cell tumors [77]. HDAC4 and HDAC9 also physically interact with MEF2C repressing MEF2C-dependent transcription [78,79]. Inhibition of SIN3 activity in breast cancer cells leads to the derepression of silenced genes, such as ESR1 α , restoring sensitivity to tamoxifen treatment [80]. Through the same mechanism, inhibiting HDACs in combination with letrozole is more effective at suppressing tumor growth in a

xenograft model than either treatment alone [81]. In this context, through guilt-by-association [11], it appears that LINC00341 may play a role in ER-mediated transcriptional repression.

We also showed that three validated gene pairs constitute a subnetwork representative of the EMT (Figure 6b). This is in agreement with a previous study showing that breast tumors contain cells with both epithelial and mesenchymal markers, the latter being associated with residual tumor following either chemotherapy or letrozole treatment in breast cancer [20]. EMT-derived cells can differentiate into mature osteoblasts, adipocytes or chondrocytes, and they have the ability to invade and migrate, homing toward wound sites [82] and participating in the invasion-metastasis cascade [83]. The SPARC and FN1 genes have an established association with EMT [63,64]. IGFBP7, a secreted tumor suppressor, can discriminate circulating endothelial cells of cancer patients from those of healthy donors [84] and, in this context, it functionally connects SPARC, PDGFRL, a gene that is highly expressed as primary melanomas transition into metastatic melanomas [85], FSTL1, a diffusible mesenchymal factor that can independently determine the cell fate of the endothelium [86], and MYLK, a multifunctional kinase that is involved in epithelial cell survival, is required for epithelial wound healing, and is included in

our list of Pareto optimal genes for the letrozole-treated tumors.

Notably, although EPCAM was not in the subset of 263 genes, it is also a marker of EMT and circulating endothelial cells [54,87]. Suppression of EPCAM attenuates tumor progression and downregulates transcription factors that are involved in EMT reprogramming [88]. We have validated the EMT pathway as a biological process involved in tumor adaptation to letrozole treatment and identified two potential targets within this pathway, MYLK and EPCAM, in the discovery data set as letrozole-induced essential genes, whose targeting should have a synergistic effect with neoadjuvant letrozole treatment.

We have focused on using the adaptive process at a single treatment time point to identify a letrozole-induced essential gene as a second target for sequential therapy. Because tumors comprise heterogeneous cell populations, it is likely that letrozole acts as a selective pressure, changing the proportions of clonal populations within the tumor, in addition to modulating gene expression within individual cells. This combination of tumor evolution and adaptation provides the tumor with a plethora of ways to resist the effects of the drug. In light of this, we believe this approach will reach its full potential when applied serially throughout the course of treatment with the sequential addition of drugs until the tumor has regressed enough to be completely resected or until there is no evidence of disease. If we can understand how relationships between genes change in response to a given treatment, we can plan interventions that will interfere with the adaptation process, preventing the development of resistance.

Conclusions

The advantage of molecularly targeted drugs is that they selectively act on cancerous cells leading to fewer side effects and better patient outcomes. However, tumors are dynamic living systems that modulate gene expression and coexpression relationships as part of an adaptive response that facilitates robustness in the face of these targeted perturbations. By focusing on patterns of coexpression in breast tumors, before and after letrozole treatment, we were able to capture this adaptive response and identify tumor-essential genes and letrozole-induced tumor-essential genes as potential candidates for combination therapy with neoadjuvant letrozole treatment. Given complete data sets of serially sampled tumors throughout a course of treatment, this approach could be an effective means of designing adaptive treatment strategies that respect the context-dependent functions of genes and the resilience of tumor cells, providing an opportunity to refine further the process of personalized medicine by pairing targeted drugs with evolving tumor phenotypes.

Additional files

Additional file 1: Permutation testing for coexpression.

Additional file 2: Coexpressed gene pairs in the discovery data set: untreated.

Additional file 3: Coexpressed gene pairs in the discovery data set: letrozole treated.

Additional file 4: Coexpressed gene pairs in the discovery data set (replication study): untreated.

Additional file 5: Coexpressed gene pairs in the discovery data set (replication study): letrozole treated.

Additional file 6: Coexpressed gene pairs in the replication data set: untreated.

Additional file 7: Coexpressed gene pairs in the replication data set: letrozole treated.

Additional file 8: Table of GO terms that are exclusively shared by coexpressed gene pairs in the letrozole-treated tumors and their associated biological effects.

Abbreviations

EMT: epithelial-to-mesenchymal transition; ER: estrogen receptor; FDR: false discovery rate; Gene Expression Omnibus; GEO: GO, Gene Ontology; IMP: Integrated Multi-species Prediction; KEGG: Kyoto encyclopedia of genes and genomes; limma: linear models for microarray data.

Competing interests

The authors declare that there are no competing interests.

Authors' contributions

NP conceived the study and performed the analyses. NP, CG and JM designed the study and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This work was supported by National Institutes of Health grants LM009012 and LM010098. The publication costs for this article were funded by JM.

Author details

¹Department of Pharmacology and Toxicology, Geisel School of Medicine at Dartmouth College, HB7937 One Medical Center Dr, Lebanon NH 03766, USA.

²Department of Genetics, Geisel School of Medicine at Dartmouth College, HB7937 One Medical Center Dr, Lebanon NH 03766, USA. ³Institute for Quantitative Biomedical Sciences, Geisel School of Medicine at Dartmouth College, HB7937 One Medical Center Dr, Lebanon NH 03766, USA.

Received: 30 September 2013 Accepted: 22 April 2014

Published: 30 April 2014

References

1. Fisher B, Redmond C, Brown A, Wolmark N, Wittliff J, Fisher ER, Plotkin D, Bowman D, Sachs S, Wolter J, Frelick R, Desser R, LiCalzi N, Geggie P, Campbell T, Elias G, Prager D, Koontz P, Volk H, Dimitrov N, Gardner B, Lerner H, Shibata H: **Treatment of primary breast cancer with chemotherapy and tamoxifen.** *New Engl J Med* 1981, **305**:1–6.
2. Bisagni G, Cocconi G, Scaglione F, Fraschini F, Pfister C, Trunet P: **Letrozole, a new oral non-steroidal aromatase inhibitor in treating postmenopausal patients with advanced breast cancer. A pilot study.** *Ann Oncol* 1996, **7**:99–102.
3. Buzdar AU, Jones SE, Vogel CL, Wolter J, Plourde P, Webster A: **A phase III trial comparing anastrozole (1 and 10 milligrams), a potent and selective aromatase inhibitor, with megestrol acetate in postmenopausal women with advanced breast carcinoma.** *Cancer* 1997, **79**:730–739.
4. Kaufmann M, Bajetta E, Dirix LY, Fein LE, Jones SE, Zilembo N, Dugardyn JL, Nasurdi C, Mennel RG, Cervek J, Fowst C, Polli A, di Salle E, Arkhipov A, Piscitelli G, Miller LL, Massimini G: **Exemestane is superior to megestrol acetate after tamoxifen failure in postmenopausal women with**

- advanced breast cancer: results of a phase III randomized double-blind trial.** *J Clin Oncol* 2000, **18**:1399–1411.
5. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, Gianni L, Baselga J, Bell R, Jackisch C, Cameron D, Dowsett M, Barrios CH, Steger G, Huang CS, Andersson M, Inbar M, Lichinitser M, Lang I, Nitz U, Iwata H, Thomssen C, Lohrisch C, Suter TM, Ruschoff J, Suto T, Greatorex V, Ward C, Straehle C, McFadden E, et al.: **Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer.** *New Engl J Med* 2005, **353**:1659–1672.
 6. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Davidson NE, Tan-Chiu E, Martino S, Paik S, Kaufman PA, Swain SM, Pisansky TM, Fehrenbacher L, Kutteh LA, Vogel VG, Visscher DW, Yothers G, Jenkins RB, Brown AM, Dakhil SR, Mamounas EP, Lingle WL, Klein PM, Ingle JN, Wolmark N: **Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer.** *New Engl J Med* 2005, **353**:1673–1684.
 7. Albert R, Jeong H, Barabási AL: **Error and attack tolerance of complex networks.** *Nature* 2000, **406**:378–382.
 8. Ágoston V, Csermely P, Pongor S: **Multiple weak hits confuse complex systems: a transcriptional regulatory network as an example.** *Phys Rev E* 2005, **71**:051909.
 9. Whitehurst AW, Bodemann BO, Cardenas J, Ferguson D, Girard L, Peyton M, Minna JD, Michnoff C, Hao W, Roth MG, Xie XJ, White MA: **Synthetic lethal screen identification of chemosensitizer loci in cancer cells.** *Nature* 2007, **446**:815–819.
 10. Li KC: **Genome-wide coexpression dynamics: theory and application.** *PNAS* 2002, **99**:16875–16880.
 11. Wolfe C, Kohane I, Butte A: **Systematic survey reveals general applicability of ‘guilt-by-association’ within gene coexpression networks.** *BMC Bioinformatics* 2005, **6**:227.
 12. Björnström L, Sjöberg M: **Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes.** *Mol Endocrinol* 2005, **19**:833–842.
 13. Eiermann W, Paepke S, Appfelstaedt J, Llombart-Cussac A, Eremin J, Vinholes J, Mauriac L, Ellis M, Lassus M, Chaudri-Ross A, Dugan M, Borgs M, Semiglazov V: **Preoperative treatment of postmenopausal breast cancer patients with letrozole: a randomized double-blind multicenter study.** *Ann Oncol* 2001, **12**:1527–1532.
 14. Dixon JM, Renshaw L, Dixon J, Thomas J: **Invasive lobular carcinoma: response to neoadjuvant letrozole therapy.** *Breast Cancer Res Treat* 2011, **130**:871–877.
 15. Baselga J, Semiglazov V, van Dam P, Manikhas A, Bellet M, Mayordomo J, Campone M, Kubista E, Greil R, Bianchi G, Steinseifer J, Molloy B, Tokaji E, Gardner H, Phillips P, Stumm M, Lane HA, Dixon JM, Jonat W, Rugo HS: **Phase II randomized study of neoadjuvant everolimus plus letrozole compared with placebo plus letrozole in patients with estrogen receptor-positive breast cancer.** *J Clin Oncol* 2009, **27**:2630–2637.
 16. Forero-Torres A, Saleh MN, Galleshaw JA, Jones CF, Shah JJ, Percent JJ, Nabell LM, Carpenter JT, Falkson CI, Krontiras H, Urist MM, Bland KI, De Los Santos JF, Meredith RF, Caterinicchia V, Bernreuter WK, O'Malley JP, Li Y, LoBuglio AF: **Pilot trial of preoperative (neoadjuvant) letrozole in combination with bevacizumab in postmenopausal women with newly diagnosed estrogen receptor- or progesterone receptor-positive breast cancer.** *Clin Breast Cancer* 2010, **10**:275–280.
 17. Fasching PA, Jud SM, Hauschild M, Kümmel S, Schütte M, Warm M, Hanf V, Muth M, Baier M, Schulz-Wendtland R, Beckmann MW, Lux MP: **Anticancer activity of letrozole plus zoledronic acid as neoadjuvant therapy for postmenopausal patients with breast cancer: FEMZONE trial results.** *Cancer Res* 2012, **72**(24 Supplement):PD07–02.
 18. Miller W, Larionov A, Renshaw L, Anderson T, White S, Murray J, Murray E, Hampton G, Walker J, Ho S, Krause A, Evans DB, Dixon JM: **Changes in breast cancer transcriptional profiles after treatment with the aromatase inhibitor, letrozole.** *Pharmacogenet Genom* 2007, **17**:813–826.
 19. Miller W, Larionov A, Anderson T, Evans D, Dixon J: **Sequential changes in gene expression profiles in breast cancers during treatment with the aromatase inhibitor, letrozole.** *Pharmacogenomics J* 2012, **12**:10–21.
 20. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, Rimm DL, Wong H, Rodriguez A, Herschkowitz JI, Fand C, Zhang X, He X, Pavlick A, Gutierrez MC, Renshaw L, Larionov AA, Faratian D, Hilsenbeck SG, Perou CM, Lewis MT, Rosen JM, Chang JC: **Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features.** *PNAS* 2009, **106**:13820–13825.
 21. Dai M, Wang P, Boyd A, Kostov G, Athey B, Jones E, Bunney W, Myers R, Speed T, Akil H, Watson SJ, Meng F: **Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data.** *Nucleic Acids Res* 2005, **33**:e175.
 22. Irizarry R, Hobbs B, Collin F, Beazer-Barclay Y, Antonellis K, Scherf U, Speed T: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**:249–264.
 23. R Core Team: **R: A language and environment for statistical computing.** 2013. [http://www.r-project.org/]
 24. Smyth G: *Limma: Linear Models for Microarray Data.* New York: Springer; 2005.
 25. Jeffery I, Higgins D, Culhane A: **Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data.** *BMC Bioinformatics* 2006, **7**:359.
 26. de la Fuente A, Bing N, Hoeschele I, Mendes P: **Discovery of meaningful associations in genomic data using partial correlation coefficients.** *Bioinformatics* 2004, **20**:3565–3574.
 27. Kumari S, Nie J, Chen HS, Ma H, Stewart R, Li X, Lu MZ, Taylor WM, Wei H: **Evaluation of gene association methods for coexpression network construction and biological knowledge discovery.** *PLoS One* 2012, **7**:e50411.
 28. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene Ontology: tool for the unification of biology.** *Nat Genet* 2000, **25**:25–29.
 29. Kanehisa M, Goto S: **KEGG: Kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**:27–30.
 30. Vastrik I, D'Eustachio P, Schmidt E, Joshi-Tope G, Gopinath G, Croft D, de Bono B, Gillespie M, Jassal B, Lewis S, Matthews L, Wu G, Birney E, Stein L: **Reactome: a knowledge base of biologic pathways and processes.** *Genome Biol* 2007, **8**:R39.
 31. Wong AK, Park CY, Greene CS, Bongo LA, Guan Y, Troyanskaya OG: **IMP: a multi-species functional genomics portal for integration, visualization and prediction of protein functions and networks.** *Nucleic Acids Res* 2012, **40**:W484–W490.
 32. Park CY, Wong AK, Greene CS, Rowland J, Guan Y, Bongo LA, Burdine RD, Troyanskaya OG: **Functional knowledge transfer for high-accuracy prediction of under-studied biological processes.** *PLoS Comput Biol* 2013, **9**:e1002957.
 33. Safran M, Solomon I, Shmueli O, Lapidot M, Shen-Orr S, Adato A, Ben-Dor U, Esterman N, Rosen N, Peter I, Olender T, Chalifa-Caspi V, Lancet D: **GeneCards 2002: towards a complete, object-oriented, human gene compendium.** *Bioinformatics* 2002, **18**:1542–1543.
 34. Costanzo M, Baryshnikov A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, Prinz J, St Onge RP, VanderSluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M, Deshpande R, Li Z, Lin ZY, Liang W, Marback M, Paw J, San Luis BJ, Shuteriqi E, Tong AHY, van Dyk N, et al.: **The genetic landscape of a cell.** *Science* 2010, **327**:425.
 35. Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P: **Coexpression analysis of human genes across many microarray data sets.** *Genome Res* 2004, **14**:1085–1094.
 36. Breast International Group (BIG) 1-98 Collaborative Group, Thürlimann B, Keshaviah A, Coates AS, Mouridsen H, Mauriac L, Forbes JF, Paridaens R, Castiglione-Gertsch M, Gelber RD, Rabaglio M, Smith I, Wardley A, Wardly A, Price KN, Goldhirsch A: **A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer.** *New Engl J Med* 2005, **353**:2747.
 37. Castel LD, Hartmann KE, Mayer IA, Saville BR, Alvarez J, Boomershine CS, Abramson VG, Chakravarthy AB, Friedman DL, Cella DF: **Time course of arthralgia among women initiating aromatase inhibitor therapy and a postmenopausal comparison group in a prospective cohort.** *Cancer* 2013, **119**:2375–2382.
 38. Presant CA, Bosserman L, Young T, Vakli M, Horns R, Upadhyaya G, Ebrahimi B, Yeon C, Howard F: **Aromatase inhibitor-associated arthralgia and/or bone pain: frequency and characterization in non-clinical trial patients.** *Clin Breast Cancer* 2007, **7**:775–778.

39. Gustin MP, Paultre CZ, Randon J, Bricca G, Cerutti C: **Functional meta-analysis of double connectivity in gene coexpression networks in mammals.** *Physiol Genomics* 2008, **34**:34–41.
40. Lunardi A, Di Minin G, Provero P, Dal Ferro M, Carotti M, Del Sal G, Collavin L: **A genome-scale protein interaction profile of *Drosophila* p53 uncovers additional nodes of the human p53 network.** *PNAS* 2010, **107**:6322–6327.
41. Gorczynski RM, Chen Z, Diao J, Khatri I, Wong K, Yu K, Behnke J: **Breast cancer cell CD200 expression regulates immune response to EMT6 tumor cells in mice.** *Breast Cancer Res Treat* 2010, **123**:405–415.
42. Podnos A, Clark DA, Erin N, Yu K, Gorczynski RM: **Further evidence for a role of tumor CD200 expression in breast cancer metastasis: decreased metastasis in CD200R1KO mice or using CD200-silenced EMT6.** *Breast Cancer Res Treat* 2012, **136**:117–127.
43. Qi C, Zhu Y, Hu L, Zhu Y: **Identification of Fat4 as a candidate tumor suppressor gene in breast cancers.** *Int J Cancer* 2009, **124**:793–798.
44. Saburi S, Hester I, Goodrich L, McNeill H: **Functional interactions between Fat family cadherins in tissue morphogenesis and planar polarity.** *Development* 2012, **139**:1806–1820.
45. Sadeqzadeh E, Bock CE, Thorne RF: **Sleeping giants: emerging roles for the fat cadherins in health and disease.** *Med Res Rev* 2013. doi:10.1002/med.21286.
46. Du H, Huang Y, Zaghulula M, Walters E, Cox TC, Massiah MA: **The MID1 E3 ligase catalyzes the polyubiquitination of Alpha4 ($\alpha 4$), a regulatory subunit of protein phosphatase 2A (PP2A) novel insights into MID1-mediated regulation of PP2A.** *J Biol Chem* 2013, **288**:21341–21350.
47. Chen L, Lai Y, Li D, Zhu X, Yang P, Li W, Zhu W, Zhao J, Li X, Xiao Y, Zhang Y, Xing XM, Wang Q, Zhang B, Lin YC, Zeng JL, Zhang SX, Liu CX, Li ZF, Zeng XW, Lin ZN, Zhuang ZX, Chen W: **$\alpha 4$ is highly expressed in carcinogen-transformed human cells and primary human cancers.** *Oncogene* 2011, **30**:2943–2953.
48. Li XQ, Li L, Xiao CH, Feng YM: **NEFL mRNA expression level is a prognostic factor for early-stage breast cancer patients.** *PLoS One* 2012, **7**:e31146.
49. de Cabo R, Siendones E, Minor R, Navas P: **CYB5R3: a key player in aerobic metabolism and aging?** *Aging (Albany NY)* 2010, **2**:63–68.
50. Sanchez-Alvarez R, Martinez-Outschoorn UE, Lamb R, Hult J, Howell A, Gandara R, Sartini M, Rubin E, Lisanti MP, Sotgia F: **Mitochondrial dysfunction in breast cancer cells prevents tumor growth: understanding chemoprevention with metformin.** *Cell Cycle* 2013, **12**:172–182.
51. Harrison SM, Knifley T, Chen M, O'Connor KL: **LPA, HGF, and EGF utilize distinct combinations of signaling pathways to promote migration and invasion of MDA-MB-231 breast carcinoma cells.** *BMC Cancer* 2013, **13**:501.
52. Khuon S, Liang L, Dettman RW, Sporn PH, Wysolmerski RB, Chew TL: **Myosin light chain kinase mediates transcellular intravasation of breast cancer cells through the underlying endothelial cells: a three-dimensional FRET study.** *J Cell Sci* 2010, **123**:3.
53. Trzpis M, McLaughlin PM, de Leij LM, Harmsen MC: **Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule.** *Am J Pathol* 2007, **171**:386–395.
54. Coumans FA, Ligthart ST, Uhr JW, Terstappen LW: **Challenges in the enumeration and phenotyping of CTC.** *Clin Cancer Res* 2012, **18**:5711–5718.
55. Went P, Vasei M, Bubendorf L, Terracciano L, Tornillo L, Riede U, Kononen J, Simon R, Sauter G, Baeuerle P: **Frequent high-level expression of the immunotherapeutic target Ep-CAM in colon, stomach, prostate and lung cancers.** *BJC* 2006, **94**:128–135.
56. Osta WA, Chen Y, Mikhitarian K, Mitas M, Salem M, Hannun YA, Cole DJ, Gillanders WE: **EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy.** *Cancer Res* 2004, **64**:5818–5824.
57. Yamamoto T, Saatcioglu F, Matsuda T: **Cross-talk between bone morphogenic proteins and estrogen receptor signaling.** *Endocrinology* 2002, **143**:2635–2642.
58. Buijs J, van der Horst G, van den Hoogen C, Cheung H, de Rooij B, Kroon J, Petersen M, van Overveld P, Pelger R, van der Pluijm G: **The BMP2/7 heterodimer inhibits the human breast cancer stem cell subpopulation and bone metastases formation.** *Oncogene* 2011, **31**:2164–2174.
59. Kurtz JE, Dufour P: **Adecatumumab: an anti-EpCAM monoclonal antibody, from the bench to the bedside.** *Expert Opin Biol Ther* 2010, **10**:951–958.
60. Goere D, Flament C, Rusakiewicz S, Poirier-Colame V, Kepp O, Martins I, Pesquet J, Eggermont AM, Elias D, Chaput N, Zitvogel L: **Potent immunomodulatory effects of the trifunctional antibody catumaxomab.** *Cancer Res* 2013, **73**:4663–4673.
61. Ellis MJ, Rosen E, Dressman H, Marks J: **Neoadjuvant comparisons of aromatase inhibitors and tamoxifen: pretreatment determinants of response and on-treatment effect.** *J Steroid Biochem Mol Biol* 2003, **86**:301–307.
62. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I: **A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA.** *Cell* 2011, **147**:358–369.
63. Tiezzi DG, Fernandez SV, Russo J: **Epithelial mesenchymal transition during the neoplastic transformation of human breast epithelial cells by estrogen.** *Int J Oncol* 2007, **31**:823.
64. Lien H, Hsiao Y, Lin Y, Yao Y, Juan H, Kuo W, Hung MC, Chang K, Hsieh F: **Molecular signatures of metaplastic carcinoma of the breast by large-scale transcriptional profiling: identification of genes potentially related to epithelial-mesenchymal transition.** *Oncogene* 2007, **26**:7859–7871.
65. Polyak K, Weinberg RA: **Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits.** *Nat Rev Cancer* 2009, **9**:265–273.
66. Greene CS, Troyanskaya OG: **Accurate evaluation and analysis of functional genomics data and methods.** *Ann NY Acad Sci* 2012, **1260**:95–100.
67. Gillis J, Pavlidis P: **Assessing identity, redundancy and confounds in Gene Ontology annotations over time.** *Bioinformatics* 2013, **29**:476–482.
68. Schmidt M, Scheulen M, Dittrich C, Obrist P, Marschner N, Dirix L, Rüttinger D, Schuler M, Reinhardt C, Awada A: **An open-label, randomized phase II study of adecatumumab, a fully human anti-EpCAM antibody, as monotherapy in patients with metastatic breast cancer.** *Ann Oncol* 2010, **21**:275–282.
69. Schmidt M, Rüttinger D, Sebastian M, Hanusch C, Marschner N, Baeuerle P, Wolf A, Göppel G, Oruzio D, Schlimok G, Steger GG, Wolf C, Eiermann W, Lang A, Schuler M: **Phase IB study of the EpCAM antibody adecatumumab combined with docetaxel in patients with EpCAM-positive relapsed or refractory advanced-stage breast cancer.** *Ann Oncol* 2012, **23**:2306–2313.
70. Krainick-Strobel UE, Lichtenegger W, Wallwiener D, Tulusan AH, Jänicke F, Bastert G, Kiesel L, Wackwitz B, Paepke S: **Neoadjuvant letrozole in postmenopausal estrogen and/or progesterone receptor positive breast cancer: a phase IIb/III trial to investigate optimal duration of preoperative endocrine therapy.** *BMC Cancer* 2008, **8**:62.
71. Llombart-Cussac A, Guerrero A, Galán A, Carañana V, Buch E, Rodríguez-Lescure Á, Ruiz A, Fuster Diana C, Guillem Porta V: **Phase II trial with letrozole to maximum response as primary systemic therapy in postmenopausal patients with ER/PgR (+) operable breast cancer.** *Clin Transl Oncol* 2012, **14**:125–131.
72. Allevi G, Strina C, Andreis D, Zanon V, Bazzola L, Bonardi S, Foroni C, Milani M, Cappelletti M, Gussago F, Agugini S, Giardini R, Martinotti M, Fox SB, Harris AL, Bottini A, Berruti A, Generali D: **Increased pathological complete response rate after a long-term neoadjuvant letrozole treatment in postmenopausal oestrogen and/or progesterone receptor-positive breast cancer.** *Br J Cancer* 2013, **108**:1587–1592.
73. Dixon J, Renshaw L, Macaskill EJ, Young O, Murray J, Cameron D, Kerr GR, Evans DB, Miller WR: **Increase in response rate by prolonged treatment with neoadjuvant letrozole.** *Breast Cancer Res Treat* 2009, **113**:145–151.
74. Lutterbuese P, Brischwein K, Hofmeister R, Crommer S, Lorenczewski G, Petersen L, Lippold S, da Silva A, Locher M, Baeuerle PA, Schlereth B: **Exchanging human Fc γ 1 with murine Fc γ 2a highly potentiates anti-tumor activity of anti-EpCAM antibody adecatumumab in a syngeneic mouse lung metastasis model.** *Cancer Immunol Immunother* 2007, **56**:459–468.

75. Wink C, Woensdregt K, Nieuwenhuijzen G, van der Sangen M, Hutschemaekers S, Roukema J, Tjan-Heijnen V, Voogd A: **Hormone treatment without surgery for patients aged 75 years or older with operable breast cancer.** *Ann Surg Oncol* 2012, **19**:1185.
76. Wang J, Hoshino T, Redner RL, Kajigaya S, Liu JM: **ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex.** *PNAS* 1998, **95**:10860–10865.
77. Petrie K, Guidez F, Howell L, Healy L, Waxman S, Greaves M, Zelent A: **The histone deacetylase 9 gene encodes multiple protein isoforms.** *J Biol Chem* 2003, **278**:16059–16072.
78. Haberland M, Arnold MA, McAnally J, Phan D, Kim Y, Olson EN: **Regulation of HDAC9 gene expression by MEF2 establishes a negative-feedback loop in the transcriptional circuitry of muscle differentiation.** *Mol Cell Biol* 2007, **27**:518–525.
79. Wang AH, Bertos NR, Vezmar M, Pelletier N, Crosato M, Heng HH, Thng J, Han J, Yang XJ: **HDAC4, a human histone deacetylase related to yeast HDA1, is a transcriptional corepressor.** *Mol Cell Biol* 1999, **19**:7816–7827.
80. Farias EF, Petrie K, Leibovitch B, Murtagh J, Chornet MB, Schenk T, Zelent A, Waxman S: **Interference with Sin3 function induces epigenetic reprogramming and differentiation in breast cancer cells.** *PNAS* 2010, **107**:11811–11816.
81. Sabnis GJ, Golubeva O, Chumsri S, Nguyen N, Sukumar S, Brodie AM: **Functional activation of the estrogen receptor- α and aromatase by the HDAC inhibitor entinostat sensitizes ER-negative tumors to letrozole.** *Cancer Res* 2011, **71**:1893–1903.
82. Battula VL, Evans KW, Hollier BG, Shi Y, Marini FC, Ayyanan A, Wang Ry, Brisken C, Guerra R, Andreeff M, Mani SA: **Epithelial-mesenchymal transition-derived cells exhibit multilineage differentiation potential similar to mesenchymal stem cells.** *Stem Cells* 2010, **28**:1435–1445.
83. Kalluri R, Weinberg RA: **The basics of epithelial-mesenchymal transition.** *J Clin Invest* 2009, **119**:1420.
84. Smirnov DA, Foulk BW, Doyle GV, Connelly MC, Terstappen LW, O'Hara SM: **Global gene expression profiling of circulating endothelial cells in patients with metastatic carcinomas.** *Cancer Res* 2006, **66**:2918–2922.
85. Riker AI, Enkemann SA, Fodstad O, Liu S, Ren S, Morris C, Xi Y, Howell P, Metge B, Samant RS, Shevde LA, Li W, Eschrich S, Daud A, Ju J, Matta J: **The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis.** *BMC Med Genomics* 2008, **1**:13.
86. Umez T, Yamanouchi H, Iida Y, Miura M, Tomooka Y: **Follistatin-like-1, a diffusible mesenchymal factor determines the fate of epithelium.** *PNAS* 2010, **107**:4601–4606.
87. Tveito S, Andersen K, Kåresen R, Fodstad Ø: **Analysis of EpCAM positive cells isolated from sentinel lymph nodes of breast cancer patients identifies subpopulations of cells with distinct transcription profiles.** *Breast Cancer Res* 2011, **13**:R75.
88. Lin CW, Liao MY, Lin WW, Wang YP, Lu TY, Wu HC: **Epithelial cell adhesion molecule regulates tumor initiation and tumorigenesis via activating reprogramming factors and epithelial-mesenchymal transition gene expression in colon cancer.** *J Biol Chem* 2012, **287**:39449–39459.

doi:10.1186/gm550

Cite this article as: Penrod et al.: Predicting targeted drug combinations based on Pareto optimal patterns of coexpression network connectivity. *Genome Medicine* 2014 **6**:33.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

