



The Emerging Potential for Network Analysis to Inform Precision Cancer Medicine

Kivilcim Ozturk^{1,2}, Michelle Dow^{1,2}, Daniel E. Carlin¹,
Rafael Bejar³ and Hannah Carter^{1,2,4,5}

1 - Department of Medicine, Division of Medical Genetics, University of California San Diego, La Jolla, CA 92093, USA

2 - Bioinformatics and Systems Biology Program, University of California San Diego, La Jolla, CA 92093, USA

3 - Moores Cancer Center, Division of Hematology and Oncology, University of California San Diego, La Jolla, CA 92093, USA

4 - Moores Cancer Center and Institute for Genomic Medicine, University of California San Diego, La Jolla, CA 92093, USA

5 - CIFAR, MaRS Centre, West Tower, 661 University Ave., Suite 505, Toronto, ON M5G 1M1, Canada

Correspondence to Hannah Carter: Department of Medicine, Division of Medical Genetics, University of California San Diego, La Jolla, CA 92093, USA hkcarter@ucsd.edu

<https://doi.org/10.1016/j.jmb.2018.06.016>

Edited by Juan Fuxman Bass

Abstract

Precision cancer medicine promises to tailor clinical decisions to patients using genomic information. Indeed, successes of drugs targeting genetic alterations in tumors, such as imatinib that targets BCR–ABL in chronic myelogenous leukemia, have demonstrated the power of this approach. However, biological systems are complex, and patients may differ not only by the specific genetic alterations in their tumor, but also by more subtle interactions among such alterations. Systems biology and more specifically, network analysis, provides a framework for advancing precision medicine beyond clinical actionability of individual mutations. Here we discuss applications of network analysis to study tumor biology, early methods for N-of-1 tumor genome analysis, and the path for such tools to the clinic.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Efforts to catalog the somatic mutations in thousands of tumor genomes have uncovered substantial genetic heterogeneity in cancer. Despite their phenotypic similarities (cancer cells display certain hallmark behaviors [1]), individual tumors rarely share the same mutations. This heterogeneity presents a significant challenge for finding any one treatment that will benefit all patients, necessitating precision approaches that tailor clinical management toward individual patients, or small groups with similar disease.

Studies of tumor genomes have determined that the majority of detected mutations are passengers that do not contribute to the oncogenic process [2–4]. The small number of causal driver mutations in a tumor presents the optimal targets for therapy, as they are specific to tumor cells and interfering with their activity should impair tumor growth. Among

some of the most successful targeted therapies to date are imatinib, which targets the BCR–ABL fusion; gefitinib, which binds and inhibits EGFR; and trastuzumab, which inhibits HER2.

Although targeted therapies have the potential to generate large responses, response rates are often modest. For example, complete responses have been observed in advanced melanomas treated with immune checkpoint inhibitors; however, across tumor types, response rates have yet to exceed ~40% of patients [5–7]. Various factors may contribute to which patients are sensitive to a therapy, including secondary mutations or subclonal heterogeneity. For some therapies, tumor-type specific differences in the underlying molecular networks appear to account for differences in response rate. Targeted inhibition of BRAF V600E is highly effective in melanoma, but not in colorectal cancer where it was found that inhibition caused feedback activation of EGFR which maintained cell

proliferation in the presence of the inhibitor [8]. Thus, the state of individual cancer genes alone may not be enough to inform response. Interactions between multiple alterations in the tumor may need to be taken into consideration.

Furthermore, many tumors do not have a driver mutation that clearly indicates a specific targeted therapy [2,9,10]. Because genes affected by driver mutations (cancer genes) tend to participate in common biological activities such as genome maintenance, cell differentiation or growth signaling [2,11–13], targeting the biological pathways rather than specific cancer genes or driver mutations may provide a strategy that will be effective for a larger number of patients [14,15]. Thus, improvement of precision therapy is likely to require approaches capable of modeling interactions between somatic alterations in tumors and capturing information about the underlying molecular state of the pathways driving tumorigenesis.

Systems biology provides a toolkit for modeling complex biological systems and their constituent interactions. Increasingly, techniques from systems biology such as network analysis are being applied to analyze genomic data in order to better understand disease [16]. Several tasks in precision cancer medicine stand to benefit from network-based approaches including identification of driver mutations, genes and pathways; patient stratification into groups with similar characteristics; identifying therapeutic opportunities; and assessment of cancer risk. Here we review the applications of networks toward these tasks at the cohort and individual patient scale and consider future work needed to bring network analysis of tumor genomes to the clinic.

Network analysis across tumor cohorts

Networks have been employed in the study of tumor cohorts in order to gain insights into tumor biology, to identify putative biomarkers and relevant disease subtypes, and to find possible targets for therapy. In these scenarios, networks are often used to control heterogeneity and to increase statistical power through aggregation (Fig. 1). The structures of the networks can be defined from known molecular interactions (e.g., protein–protein interactions), such that the network itself is a model of the biological system. This confers meaning to interactions within the network such that connected nodes are expected to be more functionally related than distant nodes and traversal of adjacent edges in the network can be loosely interpreted as biological information flow. Networks can also be used to integrate different data types, for example, by defining the state of a node by combining different measurements (mutation, expression, copy number, methylation, etc.), or by mapping events in one data layer to another.

The predominant application of networks to tumor genomic data has been with the intent to define tumor biology by uncovering the driver pathways, genes, and mutations that contribute to tumorigenesis (Fig. 2, Table 1). This task is confounded by mutational heterogeneity; somatic alterations broadly affect a variety of genes, and driver mutations are difficult to distinguish from the more prevalent passengers. Alteration frequency alone is of limited value, leading to a number of methods that attempt to model positive selection of mutations across tumors at the single gene scale [3,17–20]. It is now well established that somatic alterations in tumors aggregate not at the gene level but at the pathway level [1,2], supporting the use of networks to better identify patterns of selection indicating driver events.

Driver pathways

The term pathway is often used to describe a set of molecular interactions that collectively mediate a specific biological activity within a cell [21]. In the context of a molecular network, various biological pathways will be embedded as smaller highly connected subnetworks. Since cancer driving mutations are thought to target genes in a relatively small number of signaling and regulatory pathways [1,2,22,23], biological networks can provide a framework to study the convergence of mutations within connected subnetworks, such that identification of driver pathways can be formulated as a module detection problem (Fig. 2). This is equivalent to a search for positive selection of mutations across groups of functionally related genes. Consequently, module detection approaches are able to implicate infrequently mutated driver genes that would be overlooked based on frequency-based approaches, but fall into frequently altered network neighborhoods [24–27].

Diffusion algorithms have become a popular strategy for identifying cancer network modules (Fig. 3). The first application of network diffusion to identify disease-associated genes was described by Köhler *et al.* [28]. HotNet [24] adapted this method for driver pathways by developing a robust statistical framework for network propagation. HotNet uses mutations as a “heat source” on the network, where heat is allowed to diffuse across network edges, spreading the influence of mutations across the network dependent on the topology. After diffusion, the network can be partitioned to identify “hot” regions that are likely driver pathways enriched for the influence of mutations. The advantage of the diffusion approach is that it naturally penalizes highly connected regions of the network (heat has to be divided across large numbers of edges) where mutation influence will appear more concentrated at random. HotNet2 [25] proposed “insulated” heat diffusion (mathematically equivalent to PageRank

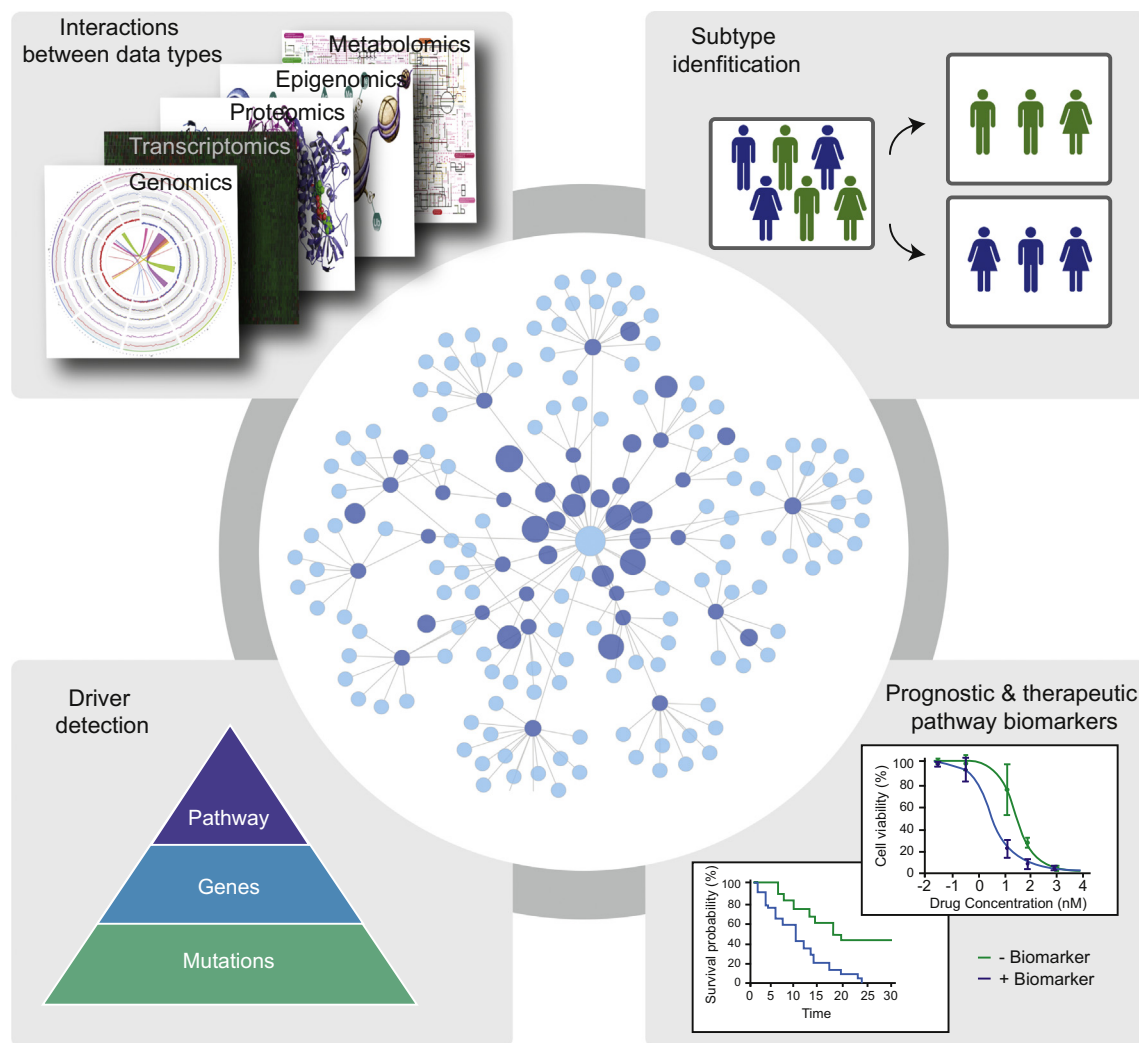


Fig. 1. Overview of network applications in cancer.

[29]) to incorporate edge direction which allows the algorithm to capture a sense of effects being upstream or downstream of causal mutations, and a damping factor that can be tuned to emphasize local topology over distant network regions.

Babaei *et al.* [26] argued that different cancer relevant-pathways may have different sizes, and thus, it may be necessary to diffuse at different scales to detect different pathways. They developed ReMIC [26], which diffuses mutation scores on a hybrid

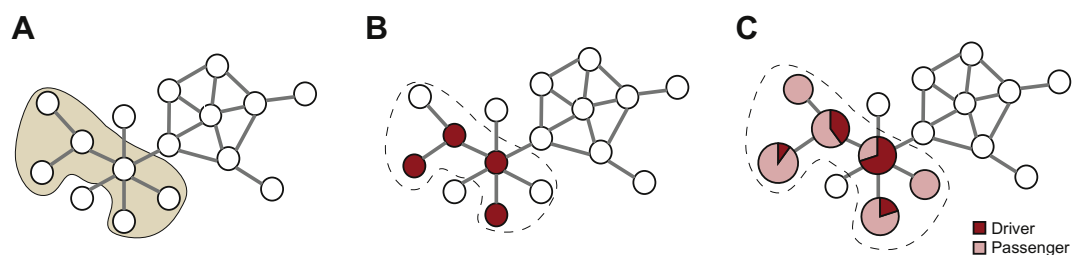


Fig. 2. Schematic diagram of three concepts used for identification of cancer drivers. Nodes represent proteins while edges represent interactions between proteins. (A) Identification of driver pathways, where the nodes within the shaded region belong to the pathway. (B) Identification of driver genes, where red nodes correspond to the proteins encoded by driver genes within the pathway (outlined). (C) Identification of driver mutations targeting the driver genes. Node size represents mutation frequency. Pie charts on the nodes display the percentage of driver (red) and passenger (pink) mutations.

interaction network across a range of parameters and implicates genes that are significantly mutated at any network scale relative to a permuted gene-mutation pairs. Many of the genes identified by this approach were part of the same connected components within the network, and thus considered to participate in the same pathways. VarWalker [27] additionally prioritizes edges where both nodes are recurrently implicated across patients by collapsing patient-specific networks into a single consensus mutation subnetwork that can then be divided into subgraphs to identify pathways. Patient-specific significantly mutated genes are determined using generalized additive models and used as starting nodes for a patient-specific Random Walk with Restart (RWR) on a PPI network. Patient-specific subnetworks are constructed from significant nodes and edges after comparing against a random model. This approach was able to detect both frequently and infrequently mutated cancer genes [27].

NetBox [30] also seeks to identify network modules enriched for altered genes in tumors, but rather than using diffusion, altered network modules are identified based on a higher density of interactions within groups than between groups via the edge betweenness algorithm [31]. NetBox was applied to analyze glioblastoma tumors using a custom interaction network that combines PPIs and signaling pathways (e.g., kinases and phosphorylation targets), using mutation and copy number data to implicate driver modules. In nCOP [32], the problem is framed as identifying small subnetworks in the larger PPI network that cover (e.g., have one or more mutations) in the largest number of patients.

Another class of methods uses variants on the prize collecting Steiner tree algorithm to identify cancer-relevant subnetworks (Fig. 3). This algorithm attempts to find connected components of a network that contain the most “prize” nodes with the minimal number of edges. The prizes in the case of cancer are genes with some evidence for association with the disease. For instance, Sun *et al.* [33] used protein–protein interaction data and known MAPK and PI3K/Akt signaling members to reconstruct their network neighborhoods. In work by Tuncbag *et al.* [34], the prizes are taken to be genes that had phosphorylation changes in cancer cell lines with the goal of reconstructing glioblastoma signaling pathways. Steiner tree algorithms tend to return more parsimonious solutions as compared to diffusion algorithms and are more sensitive to the underlying network topology, since they rely on shorter paths through a network, rather than all possible paths.

Several approaches seek to identify driver networks by combining information from genomic and transcriptomic perturbations. TieDIE [35] infers the structure of cancer associated gene networks by using mutated genes as source nodes on a custom directed PPI network and activated transcription factors as

source nodes on a transcription factor–target network. Nodes implicated by these independent diffusion processes are integrated into a cancer subnetwork connecting the two node sets. TieDIE was applied to patients with luminal *versus* basal subtypes of breast cancer in order to identify subnetworks distinguishing these subtypes [35]. DNA damage repair alterations and downstream activated MYC were characteristic of basal tumors, whereas estrogen receptor and PIK3CA pathway signaling dominated the luminal tumors. iMCMC [36] is also designed based on the idea that variation in gene sequence and expression together might contribute to cancer. It constructs two weighted networks, one from mutations and CNV data and the other from gene expression, where nodes represent a measure of the importance or effect size of a node, and edges represent a measure of similarity or functional relatedness. Then these two networks are combined and mutated core modules are detected via an optimization model followed by statistical tests.

Mutual exclusivity is a pathway-level signature of selection, in that once a driver mutation has activated or inactivated a pathway, it is rare to observe additional driver mutations in that pathway in the same tumor. Finding cancer pathways can then be formulated as a problem of identifying sets of functionally related genes that demonstrate mutual exclusivity of somatic alteration, where biological networks can be used to approximate functionally related genes (Fig. 3). In addition, driver pathways should ideally explain the maximum number of tumors. Several methods seek sets of genes with high exclusivity that cover a high proportion of tumor samples, *de novo* from somatic mutation data, without prior interaction or pathway information, including Miller *et al.* [37], Dendrix [38], and Multi-Dendrix [39]. While Miller *et al.* [37] identifies driver modules from gene networks constructed based on pairwise exclusivity scores, Dendrix [38] assigns a single weight score to a group of genes based on mutual exclusivity of the group and searches for the highest scoring group with a stochastic search method. MDPFinder [40] and Multi-Dendrix [39] use the same weight score. While MDPFinder evaluates the performance of different search techniques (an exact method and a stochastic method that allows for integration of expression data), Multi-Dendrix [39] focuses on identifying multiple mutated pathways based on the observation that simultaneous perturbation of several pathways can be required for cancer [1]. Szczurek and Beerenwinkel [41] developed a probabilistic, generative model to detect cancer pathways while accounting for error rates. In contrast to these approaches, MEMo [42] relies on prior interaction information to identify driver oncogenic network modules by analyzing maximal cliques that belong to the same pathway for patterns of mutual exclusivity across multiple patients. Other methods

integrate mutual exclusivity with other information. iMCMC [36] uses mutual exclusivity to assign edge weights when building its somatic mutation/CNV network, thereby using it as a measure of functional similarity. In addition to being useful for driver pathway identification, mutual exclusivity can provide independent evidence to validate pathways predicted by alternative approaches [26].

Driver genes

While detection of driver network modules naturally implicates the constituent genes as cancer associated, several methods have been developed to implicate genes directly (Fig. 2). Direct implication of genes may reduce false positive driver gene prediction in cases where not all genes in a network module have equal oncogenic potential. Although many gene-level methods rely on patterns of mutation, networks have also been applied to implicate driver genes.

Many network-based disease gene prediction algorithms rely on the guilt-by-association concept, which is based on the observation that phenotypically similar genes tend to be co-located in biological networks [43,44]. There are different ways of measuring topological proximity of genes in a network [45]. Local approaches focus on genes that are either in direct physical interaction, or are located close to known disease genes. For example, Wu *et al.* [46] used the shortest path algorithm to predict gastric cancer-related genes based on gene–gene interaction networks. However, the noisiness and incompleteness of the background network can limit the accuracy of predictions made based on local approaches. Global approaches define proximity based on the overall network topology, for example, using algorithms such as random walk with restart (RWR) [28], kernel diffusion [28], or network propagation [47] or by transforming the network into a probabilistic graphical model [48]. These approaches may help overcome local incompleteness of the network as they use information beyond the local subnetwork in which a gene is located. On the other hand, MUFFINN [49], a pathway-centric method that identifies cancer genes based on mutational information of both individual genes and their neighbors in functional networks, reported that analysis of mutations in indirect neighbors via diffusion algorithms did not improve predictive performance compared to analysis of only direct neighbors. NetSig [50] also implicates novel cancer genes based on mutations affecting a gene's direct network neighbors. These methods that combine mutational data with molecular network information can complement gene-based statistical tests [17,19,20,51] by uncovering drivers from the long-tail of infrequently mutated genes that have significantly mutated neighbors.

Employing only topological information while overlooking other functional information might limit driver

gene detection methods. The integration of phenotype similarity information into biological networks has been shown to improve cancer driver gene prediction [45,52,53]. Incorporation of phenotype information can also help avoid implicating genes based on false-positive interactions present in networks [52] or missing network edges [53]. As cancer is a heterogeneous disease and different classes of somatic perturbation (mutation, expression, methylation) can result in similar consequences for biological function, incorporation of multi-omics data into molecular networks can further improve cancer driver gene predictions [54,55].

Identifying genes functionally targeted by CNVs can be particularly difficult, since CNVs usually overlap many genes. Indeed, large CNVs may promote tumorigenesis through simultaneous effects on multiple genes that are positive or negative regulators of cell proliferation [56]. CONEXIC approaches this problem using the assumption that the expression of a driver gene should correlate with the expression of genes in its downstream module. CONEXIC [57] learns a network of modulator genes and associated gene expression modules from matched tumor CNV and expression data, then uses the model to identify small numbers of modulator genes within CNV regions that best explain variation in module gene expression. Matched CNV and gene expression data are not always available, and several methods have been developed to implicate driver genes within CNVs in the absence of transcriptional measurements. MAXDRIVER [58] relies on a hybrid network, combining a custom gene functional similarity network, gene–disease associations, and a disease phenotype similarity network and then identifies driver genes within CNV regions based on maximal information flow between phenotype and genotype nodes in the hybrid network, where phenotypes are specific cancers.

Just as not all genes in a pathway have equal oncogenic potential, not all driver mutations within a cancer gene will alter function in the same way. Proteins often have multiple functions, and mutations can perturb specific functions while preserving others. Approaches have been developed to identify unexpected clustering of mutations in protein domains [59], in 3D structure [60,61], or in general [62,63]. Driver genes can then be implicated on the basis of mutation clustering that suggests a particular function is statistically perturbed in cancer.

Protein interactions with distinct partners can serve as a proxy for the different functional activities of a protein [64,65]. Integrating protein structure with protein interaction networks can allow reassignment of amino acids to specific interactions as interface residues, thereby mapping mutations to specific protein activities [66–68]. Several groups have developed scoring strategies that prioritize network edges enriched for somatic mutations in cancer

under the assumption that recurrently perturbed interactions (edges) in tumors implicates the associated genes (nodes) as potential cancer drivers (Fig. 4). Some methods evaluate the ratio of observed to expected mutations in interface regions, controlling for the size of the region relative to the size of the protein [69] or the size and the amino acid composition [70]. An alternative approach uses the non-synonymous to synonymous (dN/dS) ratio at interfaces, a signature of selective pressure that has been used to identify cancer genes [71], to evaluate whether interfaces are unexpectedly biased toward functional mutations [72]. Mechismo quantifies the consequences of amino acid substitutions at interfaces based on the expected pairing of specific residues across interfaces [73], enabling an assessment of the likely effect of the mutation based on the shift toward a more or less frequently observed amino acid pair. Some analyses additionally estimate the impact of the amino acid substitution on the stability of the protein complex [68,72,74] with methods that use force fields to evaluate binding affinity changes (e.g., FoldX [75]).

Alternative splicing can also play an important role in cancer by perturbing molecular interaction networks [76]. Differential inclusion of exons that form interaction interfaces in cancer can point to potential mechanisms of tumorigenesis. Climente-González *et al.* [76] demonstrated that protein domain families that are frequently mutated in tumors are also significantly perturbed by alternative splicing. Interestingly, the set of tissue-specific exons that form binding motifs was found to alter interactions in signaling networks known to be enriched in cancer genes, suggesting the potential for dysregulation of normal splicing differences to promote cancer [77]. Moreover, cancer mutations can also more broadly perturb splicing. DrAS-Net [78] uses a network-based strategy to study how somatic mutations impact the alternative splicing of neighboring genes in a functional interaction network. This study uncovered extensive effects of somatic mutations on alternative splicing in multiple tumor types.

Networks can be adapted to identify various types of cancer-associated elements. Hypothesizing that cancer-associated miRNAs would consistently dysregulate their target genes, Xu *et al.* [79] prioritized novel prostate cancer miRNAs from known miRNA-targets using a network of dysregulated miRNA target-pairs, based on joint analysis of miRNA and mRNA expression profiles. Li *et al.* [80] uncovered prognostic miRNA signatures in glioma malignant progression by *de novo* inferring the functional targets of miRNAs from an miRNA–mRNA regulatory network constructed from paired miRNA–mRNA expression profiles. Le [81] applied a variety of module finding approaches to a miRNA similarity network based on miRNA-target similarity, finding

that these approaches worked well to identify known disease modules and implicating six new miRNAs in breast cancer pathogenesis. Novel miRNAs were also implicated in breast cancer by Hamed *et al.* [82], who analyzed the proximity of somatic mutations to regulatory interactions between miRNAs and mRNAs as well as transcription factors and their target genes. Creixell *et al.* [83] analyzed mutations in the context of kinase signaling to identify altered patterns of post-translational modification that likely contribute to cancer. In general, non-coding RNAs and post-translational modifications have received less attention than somatic alterations and DNA methylation changes affecting protein-coding genes; however, there is ample evidence that they play an important role in tumorigenesis [84,85].

Driver mutations

While identifying driver pathways and genes can provide insight into tumor biology and identify new candidate therapeutic targets, the challenge of discriminating driver mutations from passengers persists. In a clinical setting, it is important to identify actionable mutations in cancer that indicate a particular treatment course. Even well-established cancer genes can carry passenger mutations (Fig. 2), and the driver or passenger status of a mutation may determine whether response to a gene-targeted therapy is possible. While many methods exist to predict whether a mutation is a driver [86,87], networks provide a means by which multiple omic data layers can be used as evidence of functional effect.

Transcriptional data can be used to discriminate drivers under the assumption that driver mutations impact the gene expression levels of the proteins they target, and hence their interacting partners and/or the proteins within the same pathway, while passenger mutations should not display a strong effect [8,88]. DriverNet [88] uses this concept to prioritize mutations by their effect on transcriptional networks by relating the mutations to disrupted transcriptional patterns via a bipartite graph and assigns statistical significance to candidate predictions. With this approach, DriverNet identified multiple rare drivers in breast and ovarian cancer. ParadigmShift [89] predicts the functional effect of a mutation on the pathway neighborhood of the targeted gene by employing a belief-propagation algorithm to deduce mutation activity based on gene expression and copy number data upstream and downstream of particular mutations. It has been shown to be effective in predicting novel mutations in glioblastoma, ovarian and lung squamous cancers. These methods are complementary to frequency-based methods as they have been shown to be successful in detecting rare drivers.

Table 1. Network-based methods for identifying driver events in cancer across multiple scales (pathway, gene, miRNA, or mutation)

Method	Application	Approach	URL	Reference
HotNet	Driver pathway identification	Diffusion	http://compbio.cs.brown.edu/projects/hotnet/	[24]
HotNet2	Driver pathway identification	Diffusion	http://compbio.cs.brown.edu/projects/hotnet2/	[25]
ReMIC	Driver pathway identification	Diffusion	http://bioinformatics.tudelft.nl/users/sepideh-babaei	[26]
VarWalker	Driver pathway identification	Diffusion	https://bioinfo.uth.edu/VarWalker.html	[27]
NetBox	Driver pathway identification	Edge density	http://sanderlab.org/tools/netbox.html	[30]
nCOP	Driver pathway identification	Network-informed set coverage	https://github.com/Singh-Lab/nCOP	[32]
Sun <i>et al.</i>	Driver pathway identification	Prize collecting Steiner tree	–	[33]
Tuncbag <i>et al.</i>	Driver pathway identification	Prize collecting Steiner tree	–	[34]
TieDIE	Driver pathway identification	Diffusion	https://sysbiowiki.soe.ucsc.edu/tiedie	[35]
iMCMC	Driver pathway identification	Network fusion and optimization-based module finding	–	[36]
Miller <i>et al.</i>	Driver pathway identification	Mutual exclusivity	http://brl.bcm.tmc.edu/rme/index.rhtml	[37]
Dendrix	Driver pathway identification	Mutual exclusivity	http://compbio.cs.brown.edu/projects/dendrix/	[38]
MDPFinder	Driver pathway identification	Mutual exclusivity	http://zhangroup.aporc.org/ShiHuaZhang	[40]
Multi-Dendrix	Driver pathway identification	Mutual exclusivity	http://compbio.cs.brown.edu/projects/multi-dendrix/	[39]
Szczurek and Beerenwinkel	Driver pathway identification	Mutual exclusivity	–	[41]
MEMo	Driver pathway identification	Mutual exclusivity	http://sanderlab.org/tools/memo.html	[42]
Wu <i>et al.</i>	Driver gene identification	Shortest path algorithm	–	[46]
Conflux	Driver gene identification	Probabilistic graphical model	–	[48]
MUFFINN	Driver gene identification	Functional impact on network neighbors	http://www.inetbio.org/muffinn/	[49]
NetSig	Driver gene identification	Functional impact on network neighbors	http://www.lagelab.org/resources	[50]
Luo <i>et al.</i>	Driver gene identification	Data integration; mutation and phenotype similarity	–	[52]
Liu <i>et al.</i>	Driver gene identification	Data integration; mutation and phenotype similarity	–	[53]
Lin <i>et al.</i>	Driver gene identification	Data integration; mutation and phenotype similarity	–	[45]
Cantini <i>et al.</i>	Driver gene identification	Data integration; multi-omics	https://github.com/lcan88/Gene4x	[54]
Yuan <i>et al.</i>	Driver gene identification	Data integration; multi-omics	–	[55]
CONEXIC	Driver gene identification	Drivers within CNV regions	http://www.c2b2.columbia.edu/danapeerlab/html/conexic.html	[57]
MAXDRIVER	Driver gene identification	Drivers within CNV regions	–	[58]
Porta-Pardo <i>et al.</i>	Driver gene identification	Ratio of observed to expected mutations in interface regions	https://github.com/eduardporta/e-Driver	[69]
Raimondi <i>et al.</i>	Driver gene identification	Ratio of observed to expected mutations in interface regions	–	[70]
Engin <i>et al.</i>	Driver gene identification	dN/dS ratio at interfaces	–	[72]
Mechismo	Driver gene identification	Pair potentials of amino acid substitutions at interfaces	http://mechismo.russelllab.org	[73]
Climente-González <i>et al.</i>	Driver gene identification	Network perturbation by alternative splicing	–	[76]
Buljan <i>et al.</i>	Driver gene identification	Network perturbation by alternative splicing	–	[77]
DrAS-Net	Driver gene identification	Alternative splicing perturbation by mutations	http://www.bio-bigdata.com/dras_net/index.jsp	[78]
Xu <i>et al.</i>	Driver miRNA identification	Data integration; regulatory interactions (miRNA–mRNA)	–	[79]
Li <i>et al.</i>	Driver miRNA identification	Data integration; regulatory interactions (miRNA–mRNA)	–	[80]
Le	Driver miRNA identification	Diffusion	–	[81]
Hamed <i>et al.</i>	Driver miRNA identification	Data integration; regulatory interactions (TF–gene and miRNA–mRNA)	–	[82]
Creixell <i>et al.</i>	Network-attacking mutation (NAM) identification	Data integration; post-translational modifications	http://lindinglab.science/	[83]
DriverNet	Driver mutation identification	Effects on transcriptional networks	http://shahlab.ca/projects/drivernet/	[88]
PARADIGM-SHIFT	Driver mutation identification	Local network effects	https://github.com/paradigmshift/paradigmshift.github.com	[89]

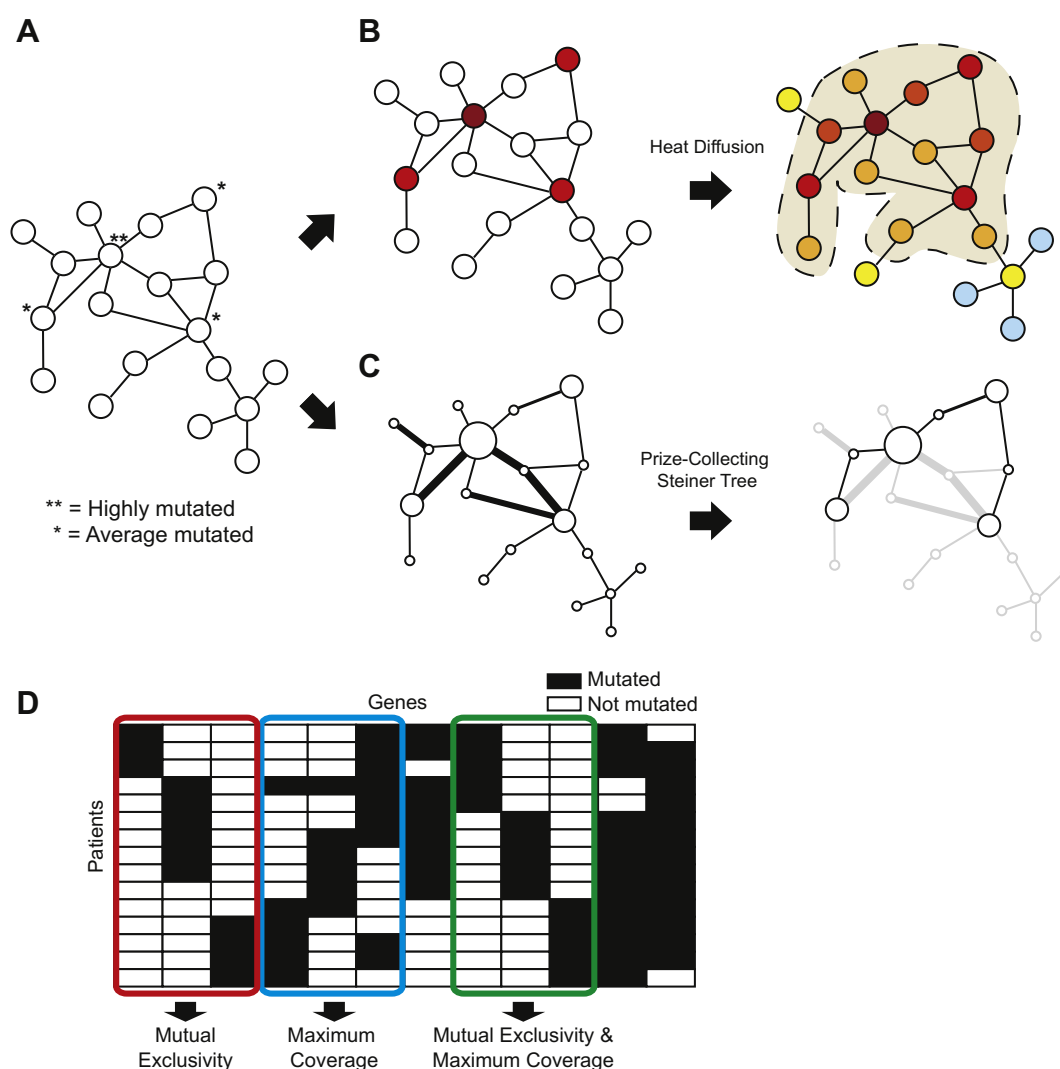


Fig. 3. Summary of several commonly used algorithmic approaches in the identification of driver pathways/subnetworks. (A) An example network with nodes representing proteins and edges representing interactions between proteins. Nodes marked with an asterisk are mutated proteins. (B) A heat diffusion process where each node is initially assigned a color based on the mutation score of the corresponding gene. Heat diffuses across the edges of the network where intensity of the node colors denotes the mutational influence on the protein. In the end, significantly mutated subnetworks (outlined) are reported. (C) A prize-collecting Steiner tree approach where node size represents mutation score and edge width represents the confidence in the interaction (other concepts can be used for determining the node size and edge width). Algorithm reports the connected components with the most "prize" nodes and the least number of edges. (D) A mutual exclusivity approach that identifies sets of genes that simultaneously maximize mutual exclusivity and coverage of patient samples.

Clinically relevant patient stratification with networks

Understanding tumor biology is central to identifying novel drug targets and building a body of knowledge that can help with the interpretation of tumor genomes (Table 2). A directly relevant aspect of this is defining clinically informative cancer subtypes that can indicate relevant characteristics of a tumor, including long-term prognosis or likely drug sensitivities. Subtypes are well established in some tumors (e.g., molecular subtypes

in breast cancer), but less obvious in others, and are often defined based on the presence or absence of specific molecular markers, or based on similar patterns of gene expression. Different approaches seem to work for different cancer types; gene expression can be used to stratify breast and ovarian cancer [90–94]; however, could not stratify tumors into clinically informative subtypes in some cancers [93,95].

Because tumor genomes are so heterogeneous, it is rare that a single molecular event (e.g., HER2 upregulation in breast cancer) defines an entire

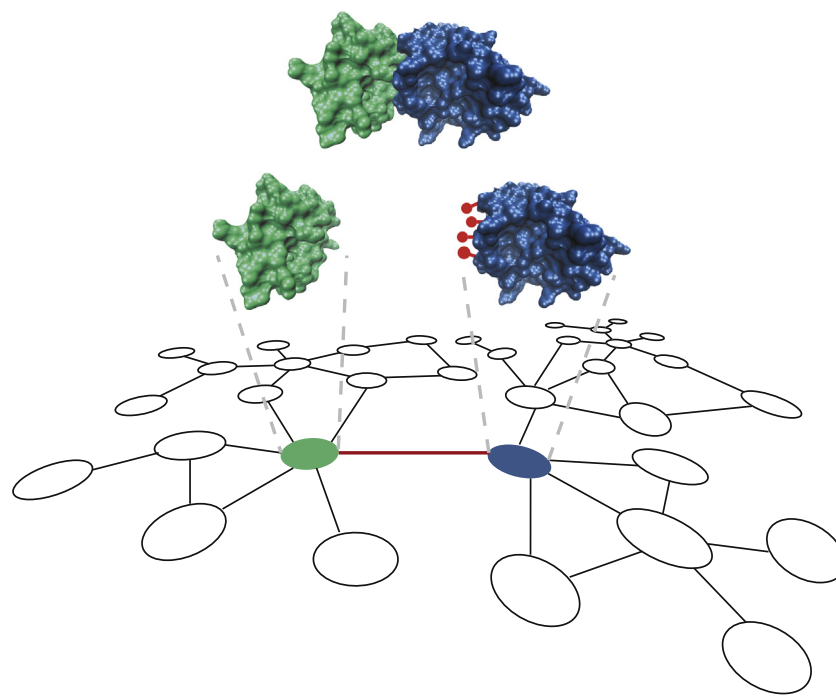


Fig. 4. A network-based approach for integration of protein–protein interaction networks with protein 3D structures and mutation data to identify cancer driver genes. Nodes represent proteins and edges represent interactions between proteins in the network. Unexpected mutational enrichment (red lollipops) on the interaction interface region of a protein (blue) implicates the encoding gene as potential cancer driver.

subtype. Networks can be used to manage this heterogeneity to group samples with similar molecular profiles that might indicate tumor similarity even though different genes play synonymous roles [96]. Indeed, in some tumor types, molecular subtypes are naturally defined by the activity of a particular biological pathway (e.g., Wnt and SHH subtypes of medulloblastoma), but the pathway is activated by different mutations across patients.

Using mutation profiles to infer similar patient clusters requires that the data be processed to make it more amenable to statistical learning. Tumor mutation profiles are sparse and binary, with little overlap. Network-based stratification (NBS) [97] integrates molecular networks with mutation profiles to stratify tumors into clinically meaningful subtypes by clustering patients with mutations in similar network regions. This is done by first projecting the binary mutation profiles of each patient onto a gene interaction network, followed by the application of network propagation to spread the influence of mutations over the network. Then the “network-smoothed” patient profiles, which are neither binary nor sparse, are clustered into subtypes using network regularized non-negative matrix factorization combined with consensus clustering to ensure stability of the final clustering [97]. NetNorM normalizes patient mutation profiles based on gene network topology, adding or removing mutations for particular patients conditional on the distribution of mutations on the network [98] instead of network smoothing to overcome the sparse and binary nature of mutation data. NBS or similar approaches have successfully

stratified cancer cohorts into subtypes that are related to clinical outcomes such as patient survival, response to therapy, or tumor histology in ovarian, lung [97], colorectal, head and neck, kidney [99], endometrial [97,99,100], and prostate cancers [101].

In its original application, NBS used a fixed gene network for all cancer types. It has recently been suggested that interpatient heterogeneity could be better overcome by using cancer-type-specific networks. To this end, He *et al.* [100] employed expression data to create cancer-type-specific significant co-expression networks that were then used with somatic mutation data in an NBS approach. By focusing on the disease-specific network, it was possible to identify survival-associated subtypes in uterine corpus endometrial carcinoma cancer that were not detected by the original NBS method. There are also similar approaches [102,103] that integrate network architecture information with gene expression profiles (as opposed to somatic mutation data like in NBS) to assign weights to genes in a gene by patient matrix which is then clustered to stratify patients into groups and discover cancer subtypes. Smoothing expression across the network emphasizes groups of related genes with similar expression patterns across samples, thereby identifying more stable signals within the expression data.

Tumor stratification may be further improved by integrating multiple omic data layers. PARADIGM [104] integrates copy number variation, expression, and pathway-level data to infer patient-specific genetic activities. PARADIGM demonstrates that

clustering patients based on their significant pathway perturbations divides them into clinically relevant subgroups in glioblastoma multiforme more successfully than using gene-level data (gene expression or copy number variation) in isolation. A method called SNF [105] constructs separate patient-by-patient similarity networks from DNA methylation, mRNA expression, and miRNA expression, then fuses them into one network that represents the full spectrum of underlying data and can be used for cancer subtype discovery via spectral clustering. This approach not only avoids noise and bias from different data types but also takes advantage of the intersecting information of the overall molecular characteristics of the cohort. ndmaSNF [106] also takes advantage of the SNF framework, extending it by adding somatic mutation data to find cancer subtypes and demonstrating that these subtypes are characterized by distinct survival profiles. Overall, incorporating multi-omic data to a fused network is reported to enhance the power for discovering subtypes that correlate with survival or other clinical features in cancer.

Establishing the effectiveness of stratification using cost-effective gene panels is key to bring tumor stratification approaches into the clinic. In a recent study, Zhong *et al.* [99] demonstrated that small panels are effective in clustering tumors across 13 major cancer subtypes and even outperform full exome data for most cancer types. This is likely due to the enrichment of clinically important genes and cancer drivers in such panels compared to the full exome data which is dominated by passenger mutations [2]. Interestingly, the original NBS reported performance loss when silent or non-functional mutations were filtered [97], but the decrease in performance was not attributable to an effect on overall mutation burden.

Network analysis to identify therapeutic opportunities

Networks can be a useful tool in providing insights and guidance for designing cancer therapies, or even identifying possible synergistic drug combinations. Synergistic drug combinations are particularly attractive because they tend to be effective at lower doses, which can reduce toxicity and make it more difficult for tumors to develop resistance [107]. Li *et al.* [108] provided a report of recent tools and databases available for predicting synergistic drug combinations with omics data. To predict drug responses or design drug combination, models have integrated disease signaling network data and transcriptomic expression data [109,110], protein interactions, protein–DNA interactions [111,112], and pathway–pathway interactions [113]. Using network modeling, researchers can view the interactions between drugs, cancer-related genes, therapeutic targets, and signaling pathways as a system, and thereby enhance the ability to suggest clinically applicable novel therapeutic approaches.

As an example, Lu *et al.* [114] performed Boolean network analysis of epithelial cell transformation under normal and inflammatory conditions, implicating key regulatory modules in malignant transformation. By perturbing specific nodes and running simulations on the network, they were able to predict effective drug combinations, validating that ceramide and PIK3CA/AKT/MTOR pathway inhibitors had synergistic anti-cancer effects *in vitro*. A similar analysis of TP53 signaling mapped perturbations of the TP53 pathway to a state space of proliferation, senescence, or cell death, and predicted that nutlin-3, an MDM2 inhibitor, would have limited effectiveness in isolation, but would synergize with WIP1 inhibition to increase cell death [115].

Proteins involved in the same biological pathway or protein complex should also be transcriptionally co-regulated in order to maintain the necessary stoichiometry of constituent molecules. Deviations in co-regulation of proteins from a normal setting, evaluated using networks of gene or protein co-expression, has been shown to create exploitable drug sensitivities [116,117]. Somatic mutations were found to generate effects that were transmitted from the affected proteins to distant tightly co-regulated gene products via the PPI network [117], suggesting the necessity to look beyond the mutated genes themselves to understand oncogenic mechanism. Diffusion-based approaches have been adapted to evaluate possible downstream consequences of individual mutations [118], and could potentially be helpful for defining a search space for mutation-associated differential protein co-regulation.

Of note, Lee *et al.* [119] reported that while simultaneous application of EGFR inhibitors and DNA-damaging chemotherapy did not result in synergy, time-staggered application significantly sensitized triple-negative breast cancer cells to chemotherapy. The authors chose EGFR signaling because of its known cross-talk with the DNA damage repair pathway. This suggests that dynamic network models could provide a strategy for systematically evaluating opportunities for drugs targeting one pathway to expose therapeutic vulnerabilities in another. This avenue remains largely unexplored, however, because high-density time course measurements are required and are difficult to obtain in an *in vivo* setting.

Networks as prognostic biomarkers

Another important aspect of clinical decision making is understanding prognosis. Some tumors may be slow to progress to malignancy and patients may be over-treated (as is common for ductal carcinoma *in situ* and prostate tumors), whereas others may be very aggressive and merit aggressive clinical intervention. Another important aspect of prognosis is monitoring for tumor progression post-treatment. Due to inter-tumor and inter-patient

heterogeneity, individual somatic alterations or over-expressed proteins have limited value as biomarkers. One alternative is to use a panel of biomarkers; however, this comes with a risk of potential overfitting as there is a large number of possible combinations to explore [120]. Networks have been applied to optimize selection of relevant biomarkers [121] and have even been used directly as biomarkers themselves [122].

When clinically relevant tumor subtypes have been established, biomarkers are needed to confidently place patients within those subtypes. mCGfinder [123] uses a matrix decomposition framework over a gene interaction network to find genes where alterations are more concentrated in subsets of patients. This approach prioritizes recurrently altered genes that are biased toward subtypes and can thus be used to assign new patients, provided that a new patient has mutations in the relevant genes. Weighted gene co-expression network analysis (WGCNA) [124] provides a strategy to identify modules of co-expressed genes that distinguish groups of patients, and was able to implicate modules associated with tumor stage or grade in serous ovarian cancer. Stage and grade are markers of disease aggressiveness that are typically determined from histological analysis of tumor biopsies. CAERUS [125] was able to predict cancer-free *versus* recurrent disease status in both breast and ovarian cancers by analyzing the correlation of expression levels of selected marker genes and their network neighbors.

Survival time or time to recurrence post-treatment can serve as a proxy for aggressiveness of tumor subtype. Biomarkers associated with short *versus* long survival can be helpful in determining whether a patient's disease should be monitored or treated aggressively. Mutations and gene expression levels can be assessed as prognostic biomarkers by analyzing their correlation with time to recurrence or death. Cox proportional hazards model [126] is a popular method performed for this task, but due to high dimensionality of mutation and expression data, it can suffer from overfitting. Use of networks has been suggested to overcome this issue by reducing dimensionality in a way that accounts for the relationship between genes. Network-based Cox regression models have been reported to improve accuracy of survival prediction in ovarian cancer (Net-Cox [127] and DegreeCox [128]). These methods incorporate a matrix representation of network information into the model such that the model is constrained to focus on genes that are particularly relevant. Net-Cox focuses on genes that are functionally related and co-expressed, whereas DegreeCox focuses on genes that are central in the network. Both methods have been shown to outperform Lasso and Ridge regression, approaches to dimensionality reduction that ignore the relationship between genes.

As it has been shown that non-redundant features are more effective in a machine learning setting, several approaches use networks to reduce redundancy in the space of features that are used for survival analysis. For example, under the assumption that miRNA–mRNA target pairs will have correlated expression, Gade *et al.* [121] built a bipartite network of mRNA and miRNA target information in order to fuse miRNA and mRNA expression profiles and guide selection of non-redundant mRNA and miRNA features. A method called GIREN uses an interaction network to group features, assuming that proximity in the network indicates functional redundancy [129].

Other methods have used measures of pathway activity as features for Cox regression instead of individual genes. For example, Huang *et al.* [130] used pathway deregulation scores; Eng *et al.* [131] created a pathway-specific score by first using Lasso to select genes within pathways, and then combining the direction of gene activity into a patient-specific pathway index score; and Zhang *et al.* [132] first used Cox modeling to determine a prognostic score for each pathway based on the predictive value of their constituent genes, and then used the pathway prognostic scores as features for survival analysis.

Biomarkers of metastasis provide an alternative measure of progression that can be assessed from biopsy of the primary tumor and can be useful for identifying patients that may benefit from more aggressive therapeutic approaches. Chuang *et al.* [122] analyzed gene expression on a PPI network to identify subnetworks differentially expressed between metastatic and non-metastatic breast tumors, enabling classification of tumors as metastatic based on gene expression profiles. These subnetwork biomarkers were more reproducible across cohorts than biomarkers based on individual genes. Because transitions in biological systems are often abrupt, time series analysis has been used to search for tipping points or dynamic network biomarkers that indicate an imminent state change. Yang *et al.* [133] applied this approach in the setting of hepatocellular carcinoma and implicated a subnetwork centered around CALM3 as a biomarker of pulmonary metastasis.

Because different mutated and differentially expressed genes characterize different tumor types and subtypes, different biomarkers of prognosis and progression are likely to be necessary for each disease. Wu *et al.* [46] identified new prognostic genes for gastric cancer using proximity to known gastric cancer genes in a protein interaction network. Tissue-specific networks may provide a strategy to better identify disease-specific biomarkers [134]. Yuan *et al.* [135] constructed a leukemia-specific protein–protein interaction network by filtering a global network against curated annotations in order to search for biomarkers for leukemia. Tissue-

Table 2. Network-based methods for patient stratification and biomarker identification (prognostic, therapeutic, progression)

Method	Application	Approach	URL	Reference
NBS	Patient stratification	Clustering patients based on network-smoothed mutation profiles	http://chianti.ucsd.edu/~mhofree/NBS/	[97]
NetNorM	Patient stratification	NBS with mutation profile normalization instead of network smoothing	https://github.com/marineLM/NetNorM	[98]
He <i>et al.</i>	Patient stratification	NBS using cancer-type-specific co-expression networks	–	[100]
NCIS	Patient stratification	Network-informed co-clustering using expression data	http://bioen-compbio.bioen.illinois.edu/NCIS/ ^a	[102]
NetBC	Patient stratification	Network-informed bi-clustering using expression data	http://mlda.swu.edu.cn/codes.php?name=NetBC ^a	[103]
PARADIGM	Patient stratification	Data integration; mutation and pathway-level activities	http://sbenz.github.io/Paradigm/	[104]
SNF	Patient stratification	Data integration; multi-omic patient similarity networks	http://compbio.cs.toronto.edu/SNF/SNF/Software.html	[105]
ndmaSNF	Patient stratification	SNF framework with addition of somatic mutation data	–	[106]
Zhong <i>et al.</i>	Patient stratification	NBS using small panels of exome-level data	–	[99]
DrugComboRanker	Therapeutic biomarkers	Integrating disease signaling network data and transcriptomic expression data	Available upon request	[109]
EPN	Therapeutic biomarkers	Integrating disease signaling network data and transcriptomic expression data	Available upon request	[110]
Wu <i>et al.</i>	Therapeutic biomarkers	Data integration; protein interactions, protein–DNA interactions, and signaling pathways	–	[111]
Choi <i>et al.</i>	Therapeutic biomarkers	Network perturbation analysis	Available upon request	[112]
Chen <i>et al.</i>	Therapeutic biomarkers	Pathway–pathway interaction network-based synergy evaluation	–	[113]
SimpleBool	Therapeutic biomarkers	Boolean network analysis on a knowledge-based network	https://github.com/SimpleBool/SimpleBool	[114]
Choi <i>et al.</i>	Therapeutic biomarkers	Boolean network modeling and attractor landscape analysis	Available upon request	[115]
IMAHP	Therapeutic biomarkers	Protein co-regulation analysis to explore drug sensitivities	–	[116]

Lee <i>et al.</i>	Therapeutic biomarkers	Effect of time-staggered application of drugs on therapeutic response	–	[119]
Roumeliotis <i>et al.</i>	Variant Interpretation	Impact of genomic alterations on protein networks	–	[117]
Engin <i>et al.</i>	Variant Interpretation	Diffusion to evaluate downstream consequences of network perturbations	–	[118]
mCGfinder	Subtype biomarkers	Matrix decomposition to prioritize alterations concentrated in subsets of patients	https://github.com/USTC-Hllab/mCGfinder	[123]
Sun <i>et al.</i>	Subtype biomarkers	Weighted gene co-expression network analysis (WGCNA)	https://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/Rpackages/WGCNA/	[124]
CAERUS	Subtype biomarkers	Co-expression analysis of marker genes and network neighbors	http://www.oicr.on.ca/research/ouellette/caerus ^a	[125]
Net-Cox	Survival analysis	Network-informed Cox regression	http://compbio.cs.umn.edu/Net-Cox/	[127]
DegreeCox	Survival analysis	Network-informed Cox regression	http://sels.tecnico.ulisboa.pt/gitlab/averissimo/degree-cox	[128]
Gade <i>et al.</i>	Survival analysis	Data integration; mRNA and miRNA expression profiles	Available as supplementary material	[121]
GIREN	Survival analysis	Data integration; gene measurements and interaction information	https://github.com/lzcyzm/GIREN	[129]
Huang <i>et al.</i>	Survival analysis	Pathway-based Cox regression	–	[130]
Eng <i>et al.</i>	Survival analysis	Pathway-based Cox regression	–	[131]
Zhang <i>et al.</i>	Survival analysis	Pathway-based Cox regression	http://www.ssg.uab.edu/bhglm/	[132]
Yang <i>et al.</i>	Prognostic biomarkers	Time series analysis for dynamic network biomarkers indicating imminent state change	–	[133]
Chuang <i>et al.</i>	Prognostic biomarkers	Detection of subnetworks associated with metastatic breast cancer	–	[122]
Wu <i>et al.</i>	Prognostic biomarkers	Proximity to known cancer genes in interaction network	–	[46]
Yuan <i>et al.</i>	Prognostic biomarkers	Filtering a global interaction network against curated annotations	–	[135]
Zhou <i>et al.</i>	Prognostic biomarkers	Data integration; mRNA, miRNA and long non-coding RNA interactions	–	[136]

^a Link listed in the original publication of the method is not valid.

specific interaction networks for RNAs are less well established. Zhou *et al.* [136] constructed a network from mRNA, miRNA, and long non-coding RNA interactions from measurements derived from ovarian cancers and uncovered ten potential prognostic lncRNA biomarkers that classified patients into high- and low-risk subgroups with significantly different survival outcomes.

Patient-specific methods

In the clinical setting, decisions must be made about the treatment of individual patients. Even clinical trials are increasingly incorporating individual genomic information by placing patients into “baskets” according to the genes in their tumor that harbor mutations [137]. However, a growing body of evidence suggests that patients with different mutations affecting the same gene can have very different responses to the same treatment [138]. For instance, different responses to HER2/HER3 inhibition by neratinib were recently reported for different mutations and tumor types [139]. Response to cetuximab, a drug used for treatment of colorectal cancer, also differs for specific mutations in the KRAS protein [140]. There are also examples of warfarin, carbamazepine, and clopidogrel being shown to be ineffective in the presence of certain genetic perturbations [141,142]. This underscores the need for tools to accurately interpret patient-specific ‘omics data to advance precision cancer medicine (Table 3).

Although affordable omics-based technologies have enabled great leaps in understanding of the mechanisms underlying the biology and clinical characteristics of cancer, few of these techniques work at the single-sample level. There is increasing interest in the development of bioinformatics tools that are able to interpret ‘omics data for individual subjects in disease related studies (e.g., personalome) [143]. In the setting of cancer, early methods have taken advantage of the availability of tumor ‘omics data to develop computational tools for single-patient tasks including identifying activated pathways, implicating personal driver mutations and predicting prognosis or drug sensitivities. Such N-of-1 methodologies could eventually inform patient assignments in clinical trials or support decision making in the clinic.

N-of-1 pathway identification

Studies that identify activated pathways from transcriptome profiles have thus far relied on the following three principles: (i) the statistical universe is a single patient; (ii) significance is derived from aggregating data on gene sets or biological modules from single or paired samples from the same patient; and (iii) pathway-level information is able to answer questions of clinical importance [143–147].

Transcriptome-based analyses can be formulated to allow interpretation [148,149] in terms of detection of deregulated gene sets related to mechanisms at various scales of biological organization (e.g., DNA repair, signaling, immune response). Using expression and pathway information, several N-of-1 studies have developed unidirectional pathway scores relying on a single [144,147] or paired samples from the same individual [145]. These scores are designed to find enrichment for extreme values within a pathway, for example, by comparing the normalized sum of rank-weighted gene expression for genes in the pathway to all genes not in that pathway, after which pathways can be ranked by their scores [144]. Paired sample analysis takes advantage of having two or more samples from the same patient, either as a time-series, or as samples from a diseased *versus* non-diseased site, and then by using a statistical test (e.g., the Wilcoxon signed-rank test) to identify pathways that are overall up- or down-regulated between samples [145]. Variants have also been developed to detect pathway enrichment from bi-directional dysregulation which is ubiquitous in biological systems. N-of-1-pathways MixEnrich uses mixture modeling to cluster genes with normal *versus* aberrant (up- or down-regulated) expression, enabling detection of enrichment based on bi-directional effects [147]. kMen also groups genes based on normal *versus* aberrant expression by clustering transcripts based on the absolute value of the log2 fold change between paired samples [146]. N-of-1 transcriptome-based methods have been applied to multiple cancer types, including head and neck cancer [144,147], lung adenocarcinoma [145], ovarian, and breast cancer [146,149,150], and have been shown to distinguish tumor samples from matched normal and healthy control samples, and to correlate with survival.

For methods aimed at informing clinical decisions, an additional requirement is effective extraction and reporting of clinically relevant and actionable targets from complex data for single samples [151]. Schissler *et al.* [143] refined the N-of-1-pathway framework wherein the Wilcoxon signed-rank test and Mahalanobis distance are combined to create a “clinical relevance metric” which provides information about the magnitude of deregulation of pathways with biological or clinical significance. The clinical utility of the score was evaluated using bootstrapping, simulation, and evaluation in biological replicates as well as by evaluating the ability of pathway scores to predict long term survival in breast cancer. OncoRep derives multiple layers of information from single-sample RNA-seq data in the setting of breast cancer, including molecular subtype classification, altered genes and pathways, gene fusions, and clinically actionable mutations [151]. The tool provides an HTML report of findings, as well as drug recommendations based on DrugBank [152] and PharmGkb [153].

Table 3. Network-based methods for N-of-1 tumor analyses to identify altered pathways or driver mutations

Method	Application	Approach	URL	Reference
Ingenuity Pathway Analysis	N-of-1 pathway analysis	Regulatory network inference and scoring from expression data using a curated causal network from the Ingenuity Knowledge Base	http://www.ingenuity.com	[148]
Pathifier	N-of-1 pathway analysis	Inferring pathway dysregulation scores for each tumor sample on the basis of expression data	www.weizmann.ac.il/pathifier/	[149]
FAIME	N-of-1 transcriptomic analysis	Pathway and molecular functional profiles from gene expression	http://www.lussiergroup.org/publications/FAIME/	[144]
N-of-1-pathway	N-of-1 transcriptomic analysis	Pathway and molecular functional profiles from gene expression (paired samples)	http://lussierlab.org/publications/N-of-1-pathways	[145]
MixEnrich	N-of-1 transcriptomic analysis	Detects bidirectionally responsive pathways using mixture models to group genes	http://lussiergroup.org/publications/MixEnrich	[147]
kMEn	N-of-1 transcriptomic analysis	Detects bidirectionally responsive pathways using K-means to group genes (paired samples)	http://www.lussierlab.org/publications/kMEn/	[146]
N-OF-1-PATHWAYS-MD	N-of-1 transcriptomic analysis	Approximates the magnitude of dysregulation of pathways with biological or clinical significance (paired samples)	http://www.lussierlab.net/publications/N-of-1-pathways	[143]
OncoRep	N-of-1 transcriptomic analysis	Derives multiple layers of clinically relevant information from single sample RNA-Seq data	https://bitbucket.org/sulab/oncorep/	[151]
Merid <i>et al.</i>	N-of-1 driver prediction	Co-occurrence of mutations and CNVs in a functional network	http://research.scilifelab.se/andrej_alexeyenko/downloads.html	[154]
DawnRank	N-of-1 driver prediction	Ranks mutated genes for overall impact on differential expression of downstream genes in the molecular interaction network	http://bioen-compbio.bioen.illinois.edu/DawnRank/ ^a	[155]
OncoIMPACT	N-of-1 driver prediction	Formulates the minimum set cover problem on a bipartite network linking mutated genes to downstream targets identified from network path analysis	https://sourceforge.net/projects/oncoimpact/	[156]
SSN	N-of-1 driver prediction	Sample-specific networks based on expression differences relative to a panel of reference samples	http://sysbio.sibcb.ac.cn/cb/chenlab/software.htm	[157]
SCS	N-of-1 driver prediction	Formulates the minimum set cover problem on a bipartite patient-specific network linking mutated genes to downstream targets identified using control theory	http://sysbio.sibcb.ac.cn/cb/chenlab/software.htm	[158]

^a Link listed in the original publication of the method is not valid.

N-of-1 driver prediction

Robust personalized driver prediction is required for discovering rare causal events in cancer that are often obscured by tumor heterogeneity and may provide important information for selecting effective therapies. Whereas Merid *et al.* [154] developed a strategy to identify drivers based on co-occurrence of mutations and CNVs from individual tumors in a functional network in the absence of expression data, most driver prediction methods integrate mutation and expression. In general, most mutations in tumors are expected to be passengers with little impact on gene expression, therefore this problem can be framed as identifying the minimum set of mutations that explain the maximum variation in gene expression conditioned on connectivity in a network. DawnRank [155] and OncoIMPACT [156] predict personalized drivers based on the impact of mutated genes on the overall differential expression of downstream genes in the molecular interaction network. While DawnRank uses network propagation to rank genes that broadly influence gene expression while controlling for the effects of network topology, OncoIMPACT finds paths between mutated and differentially expressed, defining parameters related to the path length and degree of nodes along the path to constrain predictions. OncoIMPACT uses the resulting paths to construct a bipartite network connecting mutated genes with differentially expressed genes and applies the minimum set cover criteria to select drivers.

The previously described methods rely on a single global network for prioritizing driver genes; however, it is also possible to create patient-specific networks from transcriptional measurements. This is the approach taken by sample-specific networks (SSN) [157], which constructs networks that emphasize expression differences relative to a panel of reference samples. This approach was applied to the TCGA to characterize cancer at a network level and to identify individual-specific disease modules and driver genes without using mutation information. The Single-sample Controller Strategy creates patient-specific networks by focusing on the regions of a molecular network most relevant to mutations and differentially expressed genes [158], where differential expression is determined from paired tumor-normal samples. Personalized drivers are again identified through mapping to a bipartite network, in this case by relying on control theory, and then by applying the minimum set cover. Early results from these analyses suggest that patient-specific networks may be more effective for identifying the drivers in individual tumors [158], and aggregating single-sample drivers across tumors could lead to new insights for some cancer types [157].

While some N-of-1 methods truly require only data from a single sample, others may rely on cohort analysis for parameter fitting [156], or require com-

parison against a control or reference population [149,157]. Although N-of-1 network analyses show promise for enabling single-patient inference, it is rare for tumors to undergo whole transcriptome and exome sequencing and even less frequent that matched tumor-normal transcriptomic and genomic data are available. Thus, there are currently practical limitations for the broad adoption of such approaches.

Road to the clinic

Mutations and copy number changes of individual genes have long been used as clinical biomarkers for the selection of therapies in patients with various forms of cancer. In most cases, these genetic abnormalities are directly related to the targeted therapy in question, like BCR-ABL and imatinib, activating EGFR mutations and erlotinib, and HER2 amplification and trastuzumab. These relationships are not cryptic and are easy to demonstrate in pre-clinical studies, making it easier to develop them in clinical trials. Historically, genetic associations with response to these types of targeted agents have been robust, dramatic, and extremely meaningful clinically. This has made the adoption of molecularly targeted therapies and their associated genetic biomarkers routine in the area of oncology.

There have also been examples of more complex genomic information used to tailor therapy in cancer patients, although the road to clinical implementation of these approaches has been more challenging. For example, patterns of gene expression in patients with early-stage, estrogen receptor-positive breast cancer can predict the likelihood of disease recurrence and the expected risk reduction from adjuvant chemotherapy. The prototypic clinical implementation of this approach was with the Oncotype DX test that examines the expression of 21 genes in a breast tumor sample [159]. This panel was reduced from a larger analysis of 250 candidate genes validated in three independent cohorts and can be performed on fresh-frozen, paraffin-embedded tissue [160]. The results of the test are given as a single score that estimates the risk of recurrence and can be roughly divided into a binary decision with only a small range of indeterminate values. Subsequent studies and years of clinical use have repeatedly supported the analytic validity and predictive value of this test which is now considered a standard of care [161–164]. Several similar approaches have since been implemented in breast and other cancers including colorectal and prostate.

There are three principle lessons that can be learned from the Oncotype DX example that apply to newer, and more complex, network-derived biomarkers. First, the experimentally generated model must be reduced to practice in a way that is economically feasible, reproducible, and analytically valid across a range of conditions (e.g., sample

quantity, age, quality, fixation, etc.). Second, the result of the assay must be unambiguous and easy to interpret. A propensity score associated with a clinical outcome like the Oncotype DX RS score or a binary result like “positive” or “negative” are more interpretable than a list of likely cancer driver pathways untethered to a specific drug or therapy. Finally, the assay will require extensive validation in multiple cohorts and must outperform the existing standard of care. Even then, clinical adoption and inclusion in consensus treatment guidelines (often required for insurance reimbursement) can take time. Partnering with cooperative groups, study sponsors, and clinical tissue banks to examine already collected samples would greatly accelerate the validation process.

In some cases, genomic analysis could lead to a discovery which then enables the identification of a non-genetic biomarker that is easier to detect.

Unsupervised gene expression analysis of diffuse large B-cell lymphomas, for example, found two highly reproducible gene profiles dubbed germinal B-cell like and activated B-cell like for their resemblance to normal germinal and activated B-cell signatures [165]. The distinction between activated B-cell like and germinal B-cell like subtypes of diffuse large B-cell lymphoma was clinically significant as these groups had very different outcomes after standard treatment. While the expression signature could be predicted by analyzing a small number of genes, it was noted that cell surface markers identifiable with immunohistochemistry could also distinguish these groups [166–168]. At the time, this assay was much more accessible and familiar to pathologists, enabling its rapid adoption in practice. This example reinforces the point that network analyses that identify vulnerabilities in cancer

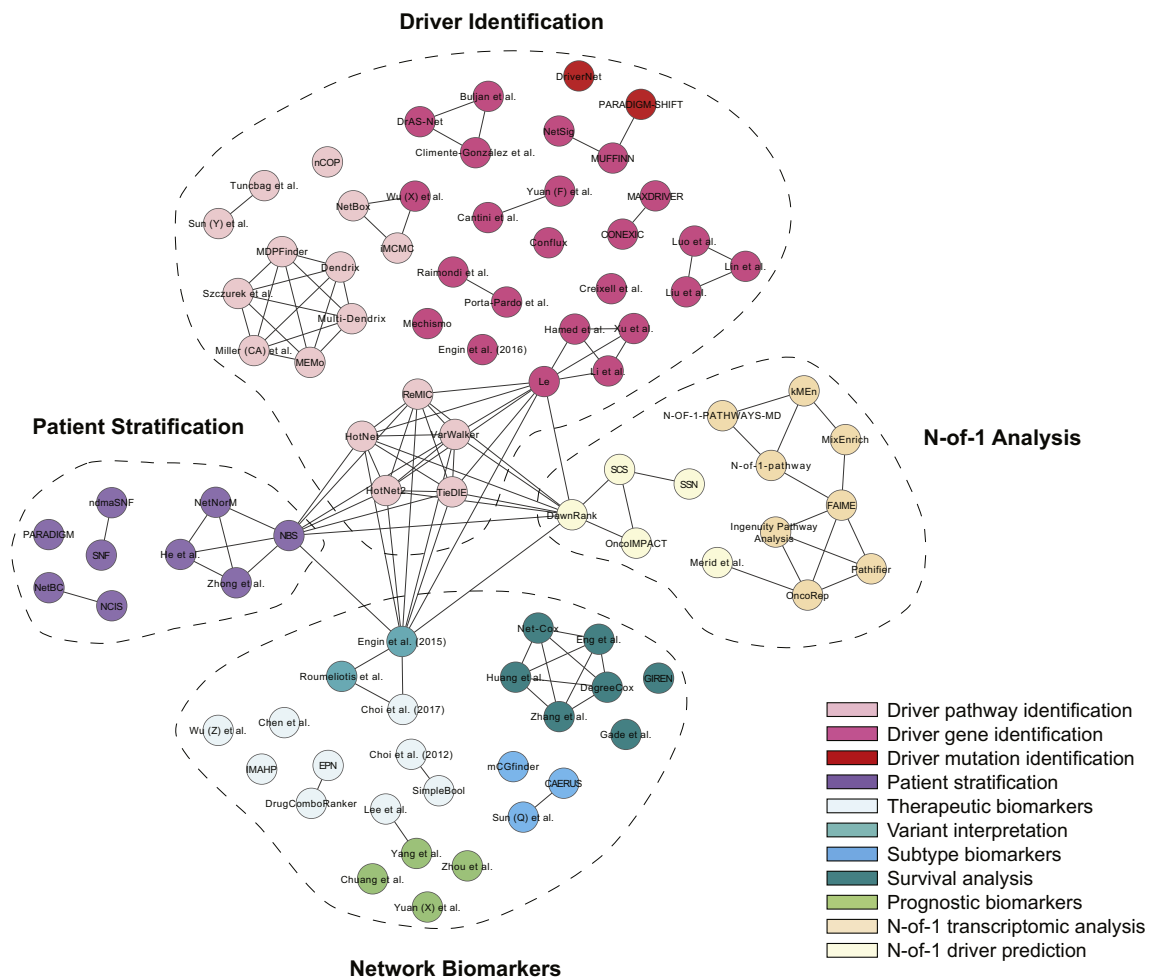


Fig. 5. Network of network-based methods used in cancer. Nodes represent the methods and edges represent similarity in the utilized algorithmic approaches. Node color represents the application of the method. Similar applications are grouped together (dashed line) under four main categories: driver identification for pathways, genes (or miRNAs or alternative splicing) or mutations, patient stratification, network biomarkers (therapeutic biomarkers, variant interpretation, subtype biomarkers, survival analysis, prognostic biomarkers), and N-of-1 analysis including transcriptomic analysis and driver prediction.

patients must be reduced to practice in a way that is accessible to both those doing the testing and those receiving the results. Being able to generate the requisite data with established methods will lower the barrier to clinical implementation.

The types of network and machine learning algorithms described here will enable more precise and individualized classification of many cancer types. For rarer tumor types, large-scale validation studies of recommended therapies may not be possible. Instead, we will need to use approaches like CancerLinQ, developed by the American Society of Clinical Oncology, to register and follow outcomes of every patient evaluated with these tools and treated based on the predictions they make. This will enable the evaluation of competing approaches in a variety of clinical contexts, ultimately providing confidence about the value of these techniques.

Discussion and conclusion

Networks have proved repeatedly to be a powerful tool for studying biological systems in health and disease. In this review, we examined algorithmic and analytic innovations in the application of network analysis to cancer, a setting uniquely characterized by an abundance of molecular data including multiple 'omic data layers and paired disease-normal samples. In this setting, networks have been used to overcome inter-tumoral heterogeneity, facilitate data integration and capture interactions among genomic, epigenetic, and transcriptomic alterations. Across samples, networks have helped implicate driver pathways, genes, and mutations; identify cancer subtypes; stratify patients with similar disease; and uncover novel biomarkers of prognosis and therapeutic response (Fig. 5). Networks can also support single-sample inference, helping to identify activated pathways and personal driver mutations. Across all of these settings, networks are an important departure point for systems biology; they provide a fundamental structure for moving beyond single-gene explanations for disease phenomena. However, there are many practical considerations to be addressed before network methods can be translated for applications in precision medicine.

Acknowledgments

This work was supported by the National Institutes of Health grant DP5-OD017937 to H.C., SDCSB/CCMI Systems Biology training grant (GM085764 and CA209891) to K.O., National Library of Medicine training grant T15LM011271 to M.D., and CIFAR fellowship to H.C.

Received 12 March 2018;

Received in revised form 30 May 2018;

Accepted 6 June 2018

Available online 15 June 2018

Keywords:

cancer systems biology;
precision cancer medicine;
network analysis

References

- [1] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, *Cell* 144 (2011) 646–674, <https://doi.org/10.1016/j.cell.2011.02.013>.
- [2] B. Vogelstein, N. Papadopoulos, V.E. Velculescu, S. Zhou, L.A. Diaz, K.W. Kinzler, Cancer genome landscapes, *Science* 339 (80) (2013) 1546–1558, <https://doi.org/10.1126/science.1235122>.
- [3] H. Carter, S. Chen, L. Isik, S. Tyekucheva, V.E. Velculescu, K.W. Kinzler, B. Vogelstein, R. Karchin, Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations, *Cancer Res.* 69 (2009) 6660–6667, <https://doi.org/10.1158/0008-5472.CAN-09-1133>.
- [4] E.D. Pleasance, R.K. Cheetham, P.J. Stephens, D.J. McBride, S.J. Humphray, C.D. Greenman, I. Varela, M.-L. Lin, G.R. Ordóñez, G.R. Bignell, K. Ye, J. Alipaz, M.J. Bauer, D. Beare, A. Butler, R.J. Carter, L. Chen, A.J. Cox, S. Edkins, P.I. Kokko-Gonzales, N.A. Gormley, R.J. Grocock, C.D. Haudenschild, M.M. Hims, T. James, M. Jia, Z. Kingsbury, C. Leroy, J. Marshall, A. Menzies, L.J. Mudie, Z. Ning, T. Royce, O.B. Schulz-Trieglaff, A. Spiridou, L.A. Stebbings, L. Szajkowski, J. Teague, D. Williamson, L. Chin, M.T. Ross, P.J. Campbell, D.R. Bentley, P.A. Futreal, M.R. Stratton, A comprehensive catalogue of somatic mutations from a human cancer genome, *Nature* 463 (2010) 191–196, <https://doi.org/10.1038/nature08658>.
- [5] B.D. Shields, F. Mahmoud, E.M. Taylor, S.D. Byrum, D. Sengupta, B. Koss, G. Baldini, S. Ransom, K. Cline, S.G. Mackintosh, R.D. Edmondson, S. Shalin, A.J. Tackett, Indicators of responsiveness to immune checkpoint inhibitors, *Sci. Rep.* 7 (2017) 807, <https://doi.org/10.1038/s41598-017-01000-2>.
- [6] D.J. Hermel, P.A. Ott, Combining forces: the promise and peril of synergistic immune checkpoint blockade and targeted therapy in metastatic melanoma, *Cancer Metastasis Rev.* 36 (2017) 43–50, <https://doi.org/10.1007/s10555-017-9656-2>.
- [7] J.M. Pitt, M. Vétizou, R. Daillère, M.P. Roberti, T. Yamazaki, B. Routy, P. Lepage, I.G. Boneca, M. Chamillard, G. Kroemer, L. Zitvogel, Resistance mechanisms to immune-checkpoint blockade in cancer: tumor-intrinsic and -extrinsic factors, *Immunity* 44 (2016) 1255–1269, <https://doi.org/10.1016/j.immuni.2016.06.001>.
- [8] A. Prahallad, C. Sun, S. Huang, F. Di Nicolantonio, R. Salazar, D. Zecchin, R.L. Beijersbergen, A. Bardelli, R. Bernards, Unresponsiveness of colon cancer to BRAF (V600E) inhibition through feedback activation of EGFR, *Nature* 483 (2012) 100–103, <https://doi.org/10.1038/nature10868>.

- [9] C.L. Sawyers, The cancer biomarker problem, *Nature* 452 (2008) 548–552, <https://doi.org/10.1038/nature06913>.
- [10] R.L. Schilsky, Implementing personalized cancer care, *Nat. Rev. Clin. Oncol.* 11 (2014) 432–438, <https://doi.org/10.1038/nrclinonc.2014.54>.
- [11] L.D. Wood, D.W. Parsons, S. Jones, J. Lin, T. Sjöblom, R.J. Leary, D. Shen, S.M. Boca, T. Barber, J. Ptak, N. Silliman, S. Szabo, Z. Dezzo, V. Ustyanksky, T. Nikolskaya, Y. Nikolsky, R. Karchin, P.A. Wilson, J.S. Kaminker, Z. Zhang, R. Croshaw, J. Willis, D. Dawson, M. Shipitsin, J.K.V. Willson, S. Sukumar, K. Polyak, B.H. Park, C.L. Pethiyagoda, P.V.K. Pant, D.G. Ballinger, A.B. Sparks, J. Hartigan, D.R. Smith, E. Suh, N. Papadopoulos, P. Buckhaults, S.D. Markowitz, G. Parmigiani, K.W. Kinzler, V.E. Velculescu, B. Vogelstein, The genomic landscapes of human breast and colorectal cancers, *Science* 318 (2007) 1108–1113, <https://doi.org/10.1126/science.1145720>.
- [12] M.A. Ali, T. Sjöblom, Molecular pathways in tumor progression: from discovery to functional understanding, *Mol. Biosyst.* 5 (2009) 902–908, <https://doi.org/10.1039/b903502h>.
- [13] L.A. Garraway, E.S. Lander, Lessons from the cancer genome, *Cell* 153 (2013) 17–37, <https://doi.org/10.1016/j.cell.2013.03.002>.
- [14] E.S. Knudsen, J.Y.J. Wang, Targeting the RB-pathway in cancer therapy, *Clin. Cancer Res.* 16 (2010) 1094–1099, <https://doi.org/10.1158/1078-0432.CCR-09-0787>.
- [15] L. Spring, A. Bardia, S. Modi, Targeting the cyclin D-cyclin-dependent kinase (CDK) 4/6-retinoblastoma pathway with selective CDK 4/6 inhibitors in hormone receptor-positive breast cancer: rationale, current status, and future directions, *Discov. Med.* 21 (2016) 65–74 (<http://www.ncbi.nlm.nih.gov/pubmed/26896604>, accessed March 9, 2018).
- [16] A.-L. Barabási, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, *Nat. Rev. Genet.* 12 (2011) 56–68, <https://doi.org/10.1038/nrg2918>.
- [17] M.S. Lawrence, P. Stojanov, P. Polak, G.V. Kryukov, K. Cibulskis, A. Sivachenko, S.L. Carter, C. Stewart, C.H. Mermel, S.A. Roberts, A. Kiezun, P.S. Hammerman, A. McKenna, Y. Drier, L. Zou, A.H. Ramos, T.J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M.L. Cortés, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D.I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A.M. Dulak, J. Lohr, D.-A. Landau, C.J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S.A. McCarroll, J. Mora, R.S. Lee, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S.B. Gabriel, C.W.M. Roberts, J.A. Biegel, K. Stegmaier, A.J. Bass, L.A. Garraway, M. Meyerson, T.R. Golub, D.A. Gordenin, S. Sunyaev, E.S. Lander, G. Getz, Mutational heterogeneity in cancer and the search for new cancer-associated genes, *Nature* 499 (2013) 214–218, <https://doi.org/10.1038/nature12213>.
- [18] N.D. Dees, Q. Zhang, C. Kandoth, M.C. Wendl, W. Schierding, D.C. Koboldt, T.B. Mooney, M.B. Callaway, D. Dooling, E.R. Mardis, R.K. Wilson, L. Ding, MuSiC: identifying mutational significance in cancer genomes, *Genome Res.* 22 (2012) 1589–1598, <https://doi.org/10.1101/gr.134635.111>.
- [19] C.J. Tokheim, N. Papadopoulos, K.W. Kinzler, B. Vogelstein, R. Karchin, Evaluating the evaluation of cancer driver genes, *Proc. Natl. Acad. Sci.* 113 (2016) 14330–14335, <https://doi.org/10.1073/pnas.1616440113>.
- [20] A. Gonzalez-Perez, N. Lopez-Bigas, Functional impact bias reveals cancer drivers, *Nucleic Acids Res.* 40 (2012), e169, <https://doi.org/10.1093/nar/gks743>.
- [21] M.P. Cary, G.D. Bader, C. Sander, Pathway information for systems biology, *FEBS Lett.* 579 (2005) 1815–1820, <https://doi.org/10.1016/J.FEBSLET.2005.02.005>.
- [22] W.C. Hahn, R.A. Weinberg, Modelling the molecular circuitry of cancer, *Nat. Rev. Cancer* 2 (2002) 331–341, <https://doi.org/10.1038/nrc795>.
- [23] B. Vogelstein, K.W. Kinzler, Cancer genes and the pathways they control, *Nat. Med.* 10 (2004) 789–799, <https://doi.org/10.1038/nm1087>.
- [24] F. Vandin, E. Upfal, B.J. Raphael, Algorithms for detecting significantly mutated pathways in cancer, *J. Comput. Biol.* 18 (2011) 507–522, <https://doi.org/10.1089/cmb.2010.0265>.
- [25] M.D.M. Leiserson, F. Vandin, H.-T. Wu, J.R. Dobson, J.V. Eldridge, J.L. Thomas, A. Papoutsaki, Y. Kim, B. Niu, M. McLellan, M.S. Lawrence, A. Gonzalez-Perez, D. Tamborero, Y. Cheng, G.A. Ryslik, N. Lopez-Bigas, G. Getz, L. Ding, B.J. Raphael, Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes, *Nat. Genet.* 47 (2015) 106–114, <https://doi.org/10.1038/ng.3168>.
- [26] S. Babaei, M. Hulsman, M. Reinders, J. de Ridder, Detecting recurrent gene mutation in interaction network context using multi-scale graph diffusion, *BMC Bioinformatics* 14 (2013) 29, <https://doi.org/10.1186/1471-2105-14-29>.
- [27] P. Jia, Z. Zhao, VarWalker: personalized mutation network analysis of putative cancer genes from next-generation sequencing data, *PLoS Comput. Biol.* 10 (2014), e1003460, <https://doi.org/10.1371/journal.pcbi.1003460>.
- [28] S. Köhler, S. Bauer, D. Horn, P.N. Robinson, Walking the interactome for prioritization of candidate disease genes, *Am. J. Hum. Genet.* 82 (2008) 949–958, <https://doi.org/10.1016/j.ajhg.2008.02.013>.
- [29] S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine, *Comput. Netw. ISDN Syst.* 30 (1998) 107–117, [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X).
- [30] E. Cerami, E. Demir, N. Schultz, B.S. Taylor, C. Sander, Automated network analysis identifies core pathways in glioblastoma, *PLoS One* 5 (2010), e8918, <https://doi.org/10.1371/journal.pone.0008918>.
- [31] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002) 7821–7826, <https://doi.org/10.1073/pnas.122653799>.
- [32] B.H. Hristov, M. Singh, Network-based coverage of mutational profiles reveals cancer genes, *Cell Syst.* 5 (2017) 221–229.e4, <https://doi.org/10.1016/j.cels.2017.09.003>.
- [33] Y. Sun, C. Ma, S. Halgamuge, The node-weighted Steiner tree approach to identify elements of cancer-related signaling pathways, *BMC Bioinformatics* 18 (2017), 551, <https://doi.org/10.1186/s12859-017-1958-4>.
- [34] N. Tuncbag, P. Milani, J.L. Pokorny, H. Johnson, T.T. Sio, S. Dalin, D.O. Iyegbe, F.M. White, J.N. Sarkaria, E. Fraenkel, Network modeling identifies patient-specific pathways in glioblastoma, *Sci. Rep.* 6 (2016), 28668, <https://doi.org/10.1038/srep28668>.
- [35] E.O. Paull, D.E. Carlin, M. Niepel, P.K. Sorger, D. Haussler, J.M. Stuart, Discovering causal pathways linking genomic events to transcriptional states using Tied Diffusion Through Interacting Events (TieDIE), *Bioinformatics* 29 (2013) 2757–2764, <https://doi.org/10.1093/bioinformatics/btt471>.

- [36] J. Zhang, S. Zhang, Y. Wang, X.-S. Zhang, Identification of mutated core cancer modules by integrating somatic mutation, copy number variation, and gene expression data, *BMC Syst. Biol.* 7 (Suppl. 2) (2013) S4, <https://doi.org/10.1186/1752-0509-7-S2-S4>.
- [37] C.A. Miller, S.H. Settle, E.P. Sulman, K.D. Aldape, A. Milosavljevic, Discovering functional modules by identifying recurrent and mutually exclusive mutational patterns in tumors, *BMC Med. Genet.* 4 (2011), 34, <https://doi.org/10.1186/1755-8794-4-34>.
- [38] F. Vandin, E. Upfal, B.J. Raphael, De novo discovery of mutated driver pathways in cancer, *Genome Res.* 22 (2012) 375–385, <https://doi.org/10.1101/gr.120477.111>.
- [39] M.D.M. Leiserson, D. Blokh, R. Sharan, B.J. Raphael, Simultaneous identification of multiple driver pathways in cancer, *PLoS Comput. Biol.* 9 (2013), e1003054, <https://doi.org/10.1371/journal.pcbi.1003054>.
- [40] J. Zhao, S. Zhang, L.-Y. Wu, X.-S. Zhang, Efficient methods for identifying mutated driver pathways in cancer, *Bioinformatics* 28 (2012) 2940–2947, <https://doi.org/10.1093/bioinformatics/bts564>.
- [41] E. Szczurek, N. Beerenwinkel, Modeling mutual exclusivity of cancer mutations, *PLoS Comput. Biol.* 10 (2014), e1003503, <https://doi.org/10.1371/journal.pcbi.1003503>.
- [42] G. Ciriello, E. Cerami, C. Sander, N. Schultz, Mutual exclusivity analysis identifies oncogenic network modules, *Genome Res.* 22 (2012) 398–406, <https://doi.org/10.1101/gr.125567.111>.
- [43] M. Oti, H.G. Brunner, The modular nature of genetic diseases, *Clin. Genet.* 71 (2007) 1–11, <https://doi.org/10.1111/j.1399-0004.2006.00708.x>.
- [44] S. Oliver, Guilt-by-association goes global, *Nature* 403 (2000) 601–603, <https://doi.org/10.1038/35001165>.
- [45] L. Lin, T. Yang, L. Fang, J. Yang, F. Yang, J. Zhao, Gene gravity-like algorithm for disease gene prediction based on phenotype-specific network, *BMC Syst. Biol.* 11 (2017), 121, <https://doi.org/10.1186/s12918-017-0519-9>.
- [46] X. Wu, H. Tang, A. Guan, F. Sun, H. Wang, J. Shu, Finding gastric cancer related genes and clinical biomarkers for detection based on gene–gene interaction network, *Math. Biosci.* 276 (2016) 1–7, <https://doi.org/10.1016/j.mbs.2015.12.001>.
- [47] O. Vanunu, O. Magger, E. Ruppin, T. Shlomi, R. Sharan, Associating genes and protein complexes with disease via network propagation, *PLoS Comput. Biol.* 6 (2010), e1000641, <https://doi.org/10.1371/journal.pcbi.1000641>.
- [48] A.M. Mezlini, A. Goldenberg, Incorporating networks in a probabilistic graphical model to find drivers for complex human diseases, *PLoS Comput. Biol.* 13 (2017), e1005580, <https://doi.org/10.1371/journal.pcbi.1005580>.
- [49] A. Cho, J.E. Shim, E. Kim, F. Supek, B. Lehner, I. Lee, MUFFINN: cancer gene discovery via network analysis of somatic mutation data, *Genome Biol.* 17 (2016), 129, <https://doi.org/10.1186/s13059-016-0989-x>.
- [50] H. Horn, M.S. Lawrence, C.R. Chouinard, Y. Shrestha, J.X. Hu, E. Worstell, E. Shea, N. Illic, E. Kim, A. Kamburov, A. Kashani, W.C. Hahn, J.D. Campbell, J.S. Boehm, G. Getz, K. Lage, NetSig: network-based discovery from cancer genomes, *Nat. Methods* 15 (2017) 61–66, <https://doi.org/10.1038/nmeth.4514>.
- [51] C.H. Mermel, S.E. Schumacher, B. Hill, M.L. Meyerson, R. Beroukhim, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, *Genome Biol.* 12 (2011) R41, <https://doi.org/10.1186/gb-2011-12-4-r41>.
- [52] J. Luo, S. Liang, Prioritization of potential candidate disease genes by topological similarity of protein–protein interaction network and phenotype data, *J. Biomed. Inform.* 53 (2015) 229–236, <https://doi.org/10.1016/j.jbi.2014.11.004>.
- [53] B. Liu, M. Jin, P. Zeng, Prioritization of candidate disease genes by combining topological similarity and semantic similarity, *J. Biomed. Inform.* 57 (2015) 1–5, <https://doi.org/10.1016/j.jbi.2015.07.005>.
- [54] L. Cantini, E. Medico, S. Fortunato, M. Caselle, Detection of gene communities in multi-networks reveals cancer drivers, *Sci. Rep.* 5 (2015), 17386, <https://doi.org/10.1038/srep17386>.
- [55] F. Yuan, W. Lu, Prediction of potential drivers connecting different dysfunctional levels in lung adenocarcinoma via a protein–protein interaction network, *Biochim. Biophys. Acta* (2017) <https://doi.org/10.1016/j.bbdis.2017.11.018>.
- [56] T. Davoli, A.W. Xu, K.E. Mengwasser, L.M. Sack, J.C. Yoon, P.J. Park, S.J. Elledge, Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome, *Cell* 155 (2013) 948–962, <https://doi.org/10.1016/j.cell.2013.10.011>.
- [57] U.D. Akavia, O. Litvin, J. Kim, F. Sanchez-Garcia, D. Kotliar, H.C. Causton, P. Pochanard, E. Mozes, L.A. Garraway, D. Pe'er, An integrated approach to uncover drivers of cancer, *Cell* 143 (2010) 1005–1017, <https://doi.org/10.1016/j.cell.2010.11.013>.
- [58] Y. Chen, J. Hao, W. Jiang, T. He, X. Zhang, T. Jiang, R. Jiang, Identifying potential cancer driver genes by genomic data integration, *Sci. Rep.* 3 (2013), 3538, <https://doi.org/10.1038/srep03538>.
- [59] M.L. Miller, E. Reznik, N.P. Gauthier, B.A. Aksoy, A. Korkut, J. Gao, G. Ciriello, N. Schultz, C. Sander, Pan-cancer analysis of mutation hotspots in protein domains, *Cell Syst.* 1 (2015) 197–209, <https://doi.org/10.1016/j.cels.2015.08.014>.
- [60] H. Stehr, S.-H.J. Jang, J.M. Duarte, C. Wierling, H. Lehrach, M. Lappe, B.M.H. Lange, The structural impact of cancer-associated missense mutations in oncogenes and tumor suppressors, *Mol. Cancer* 10 (2011) 54, <https://doi.org/10.1186/1476-4598-10-54>.
- [61] A. Kamburov, M.S. Lawrence, P. Polak, I. Leshchiner, K. Lage, T.R. Golub, E.S. Lander, G. Getz, Comprehensive assessment of cancer missense mutation clustering in protein structures, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) E5486–E5495, <https://doi.org/10.1073/pnas.1516373112>.
- [62] P. Yue, W.F. Forrest, J.S. Kaminker, S. Lohr, Z. Zhang, G. Cavet, Inferring the functional effects of mutation through clusters of mutations in homologous proteins, *Hum. Mutat.* 31 (2010) 264–271, <https://doi.org/10.1002/humu.21194>.
- [63] D. Tamborero, A. Gonzalez-Perez, N. Lopez-Bigas, Onco-driveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes, *Bioinformatics* 29 (2013) 2238–2244, <https://doi.org/10.1093/bioinformatics/btt395>.
- [64] Q. Zhong, N. Simonis, Q.-R. Li, B. Charlotiaux, F. Heuze, N. Klitgord, S. Tam, H. Yu, K. Venkatesan, D. Mou, V. Swearingen, M.A. Yildirim, H. Yan, A. Dricot, D. Szeto, C. Lin, T. Hao, C. Fan, S. Milstein, D. Dupuy, R. Brasseur, D.E. Hill, M.E. Cusick, M. Vidal, Edgetic perturbation models of human inherited disorders, *Mol. Syst. Biol.* 5 (2009), 321, <https://doi.org/10.1038/msb.2009.80>.
- [65] S. Yi, S. Lin, Y. Li, W. Zhao, G.B. Mills, N. Sahni, Functional variomics and network perturbation: connecting genotype to phenotype in cancer, *Nat. Rev. Genet.* 18 (2017) 395–410, <https://doi.org/10.1038/nrg.2017.8>.

- [66] M. Vázquez, A. Valencia, T. Pons, Structure-PPI: a module for the annotation of cancer-related single-nucleotide variants at protein–protein interfaces, *Bioinformatics* 31 (2015) 2397–2399, <https://doi.org/10.1093/bioinformatics/btv142>.
- [67] R. Mosca, J. Tenorio-Laranga, R. Olivella, V. Alcalde, A. Céol, M. Soler-López, P. Aloy, dSysMap: exploring the edgetic role of disease mutations, *Nat. Methods* 12 (2015) 167–168, <https://doi.org/10.1038/nmeth.3289>.
- [68] M. Li, F.L. Simonetti, A. Goncarenco, A.R. Panchenko, MutaBind estimates and interprets the effects of sequence variants on protein–protein interactions, *Nucleic Acids Res.* 44 (2016) W494–W501, <https://doi.org/10.1093/nar/gkw374>.
- [69] E. Porta-Pardo, L. Garcia-Alonso, T. Hrabe, J. Dopazo, A. Godzik, A pan-cancer catalogue of cancer driver protein interaction interfaces, *PLoS Comput. Biol.* 11 (2015), e1004518. <https://doi.org/10.1371/journal.pcbi.1004518>.
- [70] F. Raimondi, G. Singh, M.J. Betts, G. Apic, R. Vukotic, P. Andreone, L. Stein, R.B. Russell, Insights into cancer severity from biomolecular interaction mechanisms, *Sci. Rep.* 6 (2016), 34490. <https://doi.org/10.1038/srep34490>.
- [71] C. Greenman, P. Stephens, R. Smith, G.L. Dalglish, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, S. Edkins, S. O'Meara, I. Vastrik, E.E. Schmidt, T. Avis, S. Barthorpe, G. Bhamra, G. Buck, B. Choudhury, J. Clements, J. Cole, E. Dicks, S. Forbes, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, A. Jenkinson, D. Jones, A. Menzies, T. Mironenko, J. Perry, K. Raine, D. Richardson, R. Shepherd, A. Small, C. Tofts, J. Varian, T. Webb, S. West, S. Widaa, A. Yates, D.P. Cahill, D.N. Louis, P. Goldstraw, A.G. Nicholson, F. Brasseur, L. Looijenga, B.L. Weber, Y.-E. Chiew, A. de Fazio, M.F. Greaves, A.R. Green, P. Campbell, E. Birney, D.F. Easton, G. Chenevix-Trench, M.-H. Tan, S.K. Khoo, B.T. Teh, S.T. Yuen, S.Y. Leung, R. Wooster, P.A. Futreal, M.R. Stratton, Patterns of somatic mutation in human cancer genomes, *Nature* 446 (2007) 153–158, <https://doi.org/10.1038/nature05610>.
- [72] H.B. Engin, J.F. Kreisberg, H. Carter, Structure-based analysis reveals cancer missense mutations target protein interaction interfaces, *PLoS One* 11 (2016), e0152929. <https://doi.org/10.1371/journal.pone.0152929>.
- [73] M.J. Betts, Q. Lu, Y. Jiang, A. Drusko, O. Wichmann, M. Utz, I.A. Valtierra-Gutiérrez, M. Schlesner, N. Jaeger, D.T. Jones, S. Pfister, P. Lichter, R. Eils, R. Siebert, P. Bork, G. Apic, A.-C. Gavin, R.B. Russell, Mechismo: predicting the mechanistic impact of mutations and modifications on molecular interactions, *Nucleic Acids Res.* 43 (2015), e10. <https://doi.org/10.1093/nar/gku1094>.
- [74] Y. Dehouck, J.M. Kwasigroch, M. Rooman, D. Gilis, BeAtMuSiC: prediction of changes in protein–protein binding affinity on mutations, *Nucleic Acids Res.* 41 (2013) W333–W339, <https://doi.org/10.1093/nar/gkt450>.
- [75] R. Guerois, J.E. Nielsen, L. Serrano, Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations, *J. Mol. Biol.* 320 (2002) 369–387, [https://doi.org/10.1016/S0022-2836\(02\)00442-4](https://doi.org/10.1016/S0022-2836(02)00442-4).
- [76] H. Climente-González, E. Porta-Pardo, A. Godzik, E. Eyra, The functional impact of alternative splicing in cancer, *Cell Rep.* 20 (2017) 2215–2226, <https://doi.org/10.1016/j.celrep.2017.08.012>.
- [77] M. Buljan, G. Chalancon, S. Eustermann, G.P. Wagner, M. Fuxreiter, A. Bateman, M.M. Babu, Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks, *Mol. Cell* 46 (2012) 871–883, <https://doi.org/10.1016/j.molcel.2012.05.039>.
- [78] Y. Li, N. Sahni, R. Pancsa, D.J. McGrail, J. Xu, X. Hua, J. Coulombe-Huntington, M. Ryan, B. Tychon, D. Sudhakar, L. Hu, M. Tyers, X. Jiang, S.-Y. Lin, M.M. Babu, S. Yi, Revealing the determinants of widespread alternative splicing perturbation in cancer, *Cell Rep.* 21 (2017) 798–812, <https://doi.org/10.1016/j.celrep.2017.09.071>.
- [79] J. Xu, C.-X. Li, J.-Y. Lv, Y.-S. Li, Y. Xiao, T.-T. Shao, X. Huo, X. Li, Y. Zou, Q.-L. Han, X. Li, L.-H. Wang, H. Ren, Prioritizing candidate disease miRNAs by topological features in the miRNA target-dysregulated network: case study of prostate cancer, *Mol. Cancer Ther.* 10 (2011) 1857–1866, <https://doi.org/10.1158/1535-7163.MCT-11-0055>.
- [80] Y. Li, J. Xu, H. Chen, J. Bai, S. Li, Z. Zhao, T. Shao, T. Jiang, H. Ren, C. Kang, X. Li, Comprehensive analysis of the functional microRNA–mRNA regulatory network identifies miRNA signatures associated with glioma malignant progression, *Nucleic Acids Res.* 41 (2013), e203. <https://doi.org/10.1093/nar/gkt1054>.
- [81] D.-H. Le, Network-based ranking methods for prediction of novel disease associated microRNAs, *Comput. Biol. Chem.* 58 (2015) 139–148, <https://doi.org/10.1016/j.compbiolchem.2015.07.003>.
- [82] M. Hamed, C. Spaniol, A. Zapp, V. Helms, Integrative network-based approach identifies key genetic elements in breast invasive carcinoma, *BMC Genomics* 16 (Suppl. 5) (2015) S2, <https://doi.org/10.1186/1471-2164-16-S5-S2>.
- [83] P. Creixell, E.M. Schoof, C.D. Simpson, J. Longden, C.J. Miller, H.J. Lou, L. Perryman, T.R. Cox, N. Zivanovic, A. Palmeri, A. Wesolowska-Andersen, M. Helmer-Citterich, J. Ferkinghoff-Borg, H. Itamochi, B. Bodenmiller, J.T. Erler, B. E. Turk, R. Linding, Kinome-wide decoding of network-attacking mutations rewiring cancer signaling, *Cell* 163 (2015) 202–217, <https://doi.org/10.1016/j.cell.2015.08.056>.
- [84] R.I. Joh, C.M. Palmieri, I.T. Hill, M. Motamedi, Regulation of histone methylation by noncoding RNAs, *Biochim. Biophys. Acta* 1839 (2014) 1385–1394, <https://doi.org/10.1016/j.bbagr.2014.06.006>.
- [85] J. Cao, The functional role of long non-coding RNAs and epigenetics, *Biol. Proced. Online* 16 (2014) 11, <https://doi.org/10.1186/1480-9222-16-11>.
- [86] B.J. Raphael, J.R. Dobson, L. Oesper, F. Vandin, Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine, *Genome Med.* 6 (2014) 5, <https://doi.org/10.1186/gm524>.
- [87] F. Cheng, J. Zhao, Z. Zhao, Advances in computational approaches for prioritizing driver mutations and significantly mutated genes in cancer genomes, *Brief. Bioinform.* 17 (2016) 642–656, <https://doi.org/10.1093/bib/bbv068>.
- [88] A. Bashashati, G. Haffari, J. Ding, G. Ha, K. Lui, J. Rosner, D.G. Huntsman, C. Caldas, S.A. Aparicio, S.P. Shah, DriverNet: uncovering the impact of somatic driver mutations on transcriptional networks in cancer, *Genome Biol.* 13 (2012) R124, <https://doi.org/10.1186/gb-2012-13-12-r124>.
- [89] S. Ng, E.A. Collisson, A. Sokolov, T. Goldstein, A. Gonzalez-Perez, N. Lopez-Bigas, C. Benz, D. Haussler, J.M. Stuart, PARADIGM-SHIFT predicts the function of mutations in multiple cancers using pathway impact analysis, *Bioinformatics* 28 (2012) i640–i646, <https://doi.org/10.1093/bioinformatics/bts402>.
- [90] R.W. Tothill, A.V. Tinker, J. George, R. Brown, S.B. Fox, S. Lade, D.S. Johnson, M.K. Trivett, D. Etemadmoghadam, B.

- Locandro, N. Traficante, S. Fereday, J.A. Hung, Y.-E. Chiew, I. Haviv, D. Australian Ovarian, Cancer Study Group, D. Gertig, A. DeFazio, D.D.L. Bowtell, Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome, *Clin. Cancer Res.* 14 (2008) 5198–5208, <https://doi.org/10.1158/1078-0432.CCR-08-0196>.
- [91] J.S. Reis-Filho, L. Pusztai, Gene expression profiling in breast cancer: classification, prognostication, and prediction, *Lancet (Lond. Engl.)* 378 (2011) 1812–1823, [https://doi.org/10.1016/S0140-6736\(11\)61539-0](https://doi.org/10.1016/S0140-6736(11)61539-0).
- [92] P.A. Konstantinopoulos, D. Spentzos, B.Y. Karlan, T. Taniguchi, E. Fountzilas, N. Francoeur, D.A. Levine, S.A. Cannistra, Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer, *J. Clin. Oncol.* 28 (2010) 3555–3561, <https://doi.org/10.1200/JCO.2009.27.5719>.
- [93] T.C.G.A.R. Cancer Genome Atlas Research Network, Integrated genomic analyses of ovarian carcinoma, *Nature* 474 (2011) 609–615, <https://doi.org/10.1038/nature10166>.
- [94] P.A. Konstantinopoulos, D. Spentzos, S.A. Cannistra, Gene-expression profiling in epithelial ovarian cancer, *Nat. Clin. Pract. Oncol.* 5 (2008) 577–587, <https://doi.org/10.1038/ncponc1178>.
- [95] T.C.G.A. Cancer Genome Atlas Network, Comprehensive molecular characterization of human colon and rectal cancer, *Nature* 487 (2012) 330–337, <https://doi.org/10.1038/nature11252>.
- [96] F. Vandin, P. Clay, E. Upfal, B.J. Raphael, Discovery of mutated subnetworks associated with clinical data in cancer, *Pac. Symp. Biocomput.* (2012) 55–66 (<http://www.ncbi.nlm.nih.gov/pubmed/22174262>, accessed March 8, 2018).
- [97] M. Hofree, J.P. Shen, H. Carter, A. Gross, T. Ideker, Network-based stratification of tumor mutations, *Nat. Methods* 10 (2013) 1108–1115, <https://doi.org/10.1038/nmeth.2651>.
- [98] M. Le Morvan, A. Zinovyev, J.-P. Vert, NetNorM: capturing cancer-relevant information in somatic exome mutation data with gene networks for cancer stratification and prognosis, *PLoS Comput. Biol.* 13 (2017), e1005573, <https://doi.org/10.1371/journal.pcbi.1005573>.
- [99] X. Zhong, H. Yang, S. Zhao, Y. Shyr, B. Li, Network-based stratification analysis of 13 major cancer types using mutations in panels of cancer genes, *BMC Genomics* 16 (2015) S7, <https://doi.org/10.1186/1471-2164-16-S7-S7>.
- [100] Z. He, J. Zhang, X. Yuan, Z. Liu, B. Liu, S. Tuo, Y. Liu, Network based stratification of major cancers by integrating somatic mutation and gene expression data, *PLoS One* 12 (2017), e0177662, <https://doi.org/10.1371/journal.pone.0177662>.
- [101] L. Yang, S. Wang, M. Zhou, X. Chen, W. Jiang, Y. Zuo, Y. Lv, Molecular classification of prostate adenocarcinoma by the integrated somatic mutation profiles and molecular network, *Sci. Rep.* 7 (2017), 738, <https://doi.org/10.1038/s41598-017-00872-8>.
- [102] Y. Liu, Q. Gu, J.P. Hou, J. Han, J. Ma, A network-assisted co-clustering algorithm to discover cancer subtypes based on gene expression, *BMC Bioinformatics* 15 (2014) 37, <https://doi.org/10.1186/1471-2105-15-37>.
- [103] G. Yu, X. Yu, J. Wang, Network-aided bi-clustering for discovering cancer subtypes, *Sci. Rep.* 7 (2017), 1046, <https://doi.org/10.1038/s41598-017-01064-0>.
- [104] C.J. Vaske, S.C. Benz, J.Z. Sanborn, D. Earl, C. Szeto, J. Zhu, D. Haussler, J.M. Stuart, Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM, *Bioinformatics* 26 (2010) i237–i245, <https://doi.org/10.1093/bioinformatics/btq182>.
- [105] B. Wang, A.M. Mezlini, F. Demir, M. Fiume, Z. Tu, M. Brudno, B. Haibe-Kains, A. Goldenberg, Similarity network fusion for aggregating data types on a genomic scale, *Nat. Methods* 11 (2014) 333–337, <https://doi.org/10.1038/nmeth.2810>.
- [106] C. Yang, S.-G. Ge, C.-H. Zheng, ndmaSNF: cancer subtype discovery based on integrative framework assisted by network diffusion model, *Oncotarget* 8 (2017) 89021–89032, <https://doi.org/10.18632/oncotarget.21643>.
- [107] T.-C. Chou, Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies, *Pharmacol. Rev.* 58 (2006) 621–681, <https://doi.org/10.1124/pr.58.3.10>.
- [108] X. Li, G. Qin, Q. Yang, L. Chen, L. Xie, Biomolecular network-based synergistic drug combination discovery, *Biomed. Res. Int.* 2016 (2016) 8518945, <https://doi.org/10.1155/2016/8518945>.
- [109] L. Huang, F. Li, J. Sheng, X. Xia, J. Ma, M. Zhan, S.T.C. Wong, DrugComboRanker: drug combination discovery based on target network analysis, *Bioinformatics* 30 (2014) i228–i236, <https://doi.org/10.1093/bioinformatics/btu278>.
- [110] G. Jin, H. Zhao, X. Zhou, S.T.C. Wong, An enhanced Petri-net model to predict synergistic effects of pairwise drug combinations from gene microarray data, *Bioinformatics* 27 (2011) i310–i316, <https://doi.org/10.1093/bioinformatics/btr202>.
- [111] Z. Wu, X.-M. Zhao, L. Chen, A systems biology approach to identify effective cocktail drugs, *BMC Syst. Biol.* 4 (Suppl. 2) (2010) S7, <https://doi.org/10.1186/1752-0509-4-S2-S7>.
- [112] M. Choi, J. Shi, Y. Zhu, R. Yang, K.-H. Cho, Network dynamics-based cancer panel stratification for systemic prediction of anticancer drug response, *Nat. Commun.* 8 (2017), 1940, <https://doi.org/10.1038/s41467-017-02160-5>.
- [113] D. Chen, H. Zhang, P. Lu, X. Liu, H. Cao, Synergy evaluation by a pathway–pathway interaction network: a new way to predict drug combination, *Mol. Biosyst.* 12 (2016) 614–623, <https://doi.org/10.1039/c5mb00599j>.
- [114] J. Lu, H. Zeng, Z. Liang, L. Chen, L. Zhang, H. Zhang, H. Liu, H. Jiang, B. Shen, M. Huang, M. Geng, S. Spiegel, C. Luo, Network modelling reveals the mechanism underlying colitis-associated colon cancer and identifies novel combinatorial anti-cancer targets, *Sci. Rep.* 5 (2015), 14739, <https://doi.org/10.1038/srep14739>.
- [115] M. Choi, J. Shi, S.H. Jung, X. Chen, K.-H. Cho, Attractor landscape analysis reveals feedback loops in the p53 network that control the cellular response to DNA damage, *Sci. Signal.* 5 (2012) ra83, <https://doi.org/10.1126/scisignal.2003363>.
- [116] J.D. Lapek, P. Greninger, R. Morris, A. Amzallag, I. Pruteanu-Malinici, C.H. Benes, W. Haas, Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities, *Nat. Biotechnol.* 35 (2017) 983–989, <https://doi.org/10.1038/nbt.3955>.
- [117] T.I. Roumeliotis, S.P. Williams, E. Gonçalves, C. Alsinet, M. Del Castillo Velasco-Herrera, N. Aben, F.Z. Ghavidel, M. Michaut, M. Schubert, S. Price, J.C. Wright, L. Yu, M. Yang, R. Dienstmann, J. Guinney, P. Beltrao, A. Brazma, M. Pardo, O. Stegle, D.J. Adams, L. Wessels, J. Saez-Rodriguez, U.

- McDermott, J.S. Choudhary, Genomic determinants of protein abundance variation in colorectal cancer cells, *Cell Rep.* 20 (2017) 2201–2214, <https://doi.org/10.1016/j.celrep.2017.08.010>.
- [118] H.B. Engin, M. Hofree, H. Carter, Identifying mutation specific cancer pathways using a structurally resolved protein interaction network, *Pac. Symp. Biocomput.* (2015) 84–95 (<http://www.ncbi.nlm.nih.gov/pubmed/25592571>, accessed August 21, 2017).
- [119] M.J. Lee, A.S. Ye, A.K. Gardino, A.M. Heijink, P.K. Sorger, G. MacBeath, M.B. Yaffe, Sequential application of anti-cancer drugs enhances cell death by rewiring apoptotic signaling networks, *Cell* 149 (2012) 780–794, <https://doi.org/10.1016/j.cell.2012.03.031>.
- [120] A.J. Vargas, C.C. Harris, Biomarker development in the precision medicine era: lung cancer as a case study, *Nat. Rev. Cancer* 16 (2016) 525–537, <https://doi.org/10.1038/nrc.2016.56>.
- [121] S. Gade, C. Porzelius, M. Fälth, J.C. Brase, D. Wuttig, R. Kuner, H. Binder, H. Sülthmann, T. Beissbarth, Graph based fusion of miRNA and mRNA expression data improves clinical outcome prediction in prostate cancer, *BMC Bioinformatics* 12 (2011) 488, <https://doi.org/10.1186/1471-2105-12-488>.
- [122] H.-Y. Chuang, E. Lee, Y.-T. Liu, D. Lee, T. Ideker, Network-based classification of breast cancer metastasis, *Mol. Syst. Biol.* 3 (2007), 140, <https://doi.org/10.1038/msb4100180>.
- [123] J. Xi, A. Li, M. Wang, A novel network regularized matrix decomposition method to detect mutated cancer genes in tumour samples with inter-patient heterogeneity, *Sci. Rep.* 7 (2017), 2855, <https://doi.org/10.1038/s41598-017-03141-w>.
- [124] Q. Sun, H. Zhao, C. Zhang, T. Hu, J. Wu, X. Lin, D. Luo, C. Wang, L. Meng, L. Xi, K. Li, J. Hu, D. Ma, T. Zhu, Gene co-expression network reveals shared modules predictive of stage and grade in serous ovarian cancers, *Oncotarget* 8 (2017) 42983–42996, <https://doi.org/10.18632/oncotarget.17785>.
- [125] K.X. Zhang, B.F.F. Ouellette, CAERUS: predicting CAnCER oUtcomeS using relationship between protein structural information, protein networks, gene expression data, and mutation data, *PLoS Comput. Biol.* 7 (2011), e1001114, <https://doi.org/10.1371/journal.pcbi.1001114>.
- [126] D.R. Cox, Models and life-tables regression, *J. R. Stat. Soc. Ser. B* 34 (1972) 187–220.
- [127] W. Zhang, T. Ota, V. Shridhar, J. Chien, B. Wu, R. Kuang, Network-based survival analysis reveals subnetwork signatures for predicting outcomes of ovarian cancer treatment, *PLoS Comput. Biol.* 9 (2013), e1002975, <https://doi.org/10.1371/journal.pcbi.1002975>.
- [128] A. Verissimo, A.L. Oliveira, M.-F. Sagot, S. Vinga, DegreeCox—a network-based regularization method for survival analysis, *BMC Bioinformatics* 17 (2016), 449, <https://doi.org/10.1186/s12859-016-1310-4>.
- [129] L. Lin Zhang, H. Hui Liu, Y. Yufei Huang, X. Xuesong Wang, Y. Yidong Chen, J. Jia Meng, Cancer progression prediction using gene interaction regularized elastic net, *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 14 (2017) 145–154, <https://doi.org/10.1109/TCBB.2015.2511758>.
- [130] S. Huang, C. Yee, T. Ching, H. Yu, L.X. Garmire, A novel model to combine clinical and pathway-based transcriptomic information for the prognosis prediction of breast cancer, *PLoS Comput. Biol.* 10 (2014), e1003851, <https://doi.org/10.1371/journal.pcbi.1003851>.
- [131] K.H. Eng, S. Wang, W.H. Bradley, J.S. Rader, C. Kendzierski, Pathway index models for construction of patient-specific risk profiles, *Stat. Med.* 32 (2013) 1524–1535, <https://doi.org/10.1002/sim.5641>.
- [132] X. Zhang, Y. Li, T. Akinyemiju, A.I. Ojesina, P. Buckhaults, N. Liu, B. Xu, N. Yi, Pathway-structured predictive model for cancer survival prediction: a two-stage approach, *Genetics* 205 (2017) 89–100, <https://doi.org/10.1534/genetics.116.189191>.
- [133] B. Yang, M. Li, W. Tang, W. Liu, S. Zhang, L. Chen, J. Xia, Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma, *Nat. Commun.* 9 (2018), 678, <https://doi.org/10.1038/s41467-018-03024-2>.
- [134] H. Chen, Z. Zhu, Y. Zhu, J. Wang, Y. Mei, Y. Cheng, Pathway mapping and development of disease-specific biomarkers: protein-based network biomarkers, *J. Cell. Mol. Med.* 19 (2015) 297–314, <https://doi.org/10.1111/jcmm.12447>.
- [135] X. Yuan, J. Chen, Y. Lin, Y. Li, L. Xu, L. Chen, H. Hua, B. Shen, Network biomarkers constructed from gene expression and protein–protein interaction data for accurate prediction of leukemia, *J. Cancer* 8 (2017) 278–286, <https://doi.org/10.7150/jca.17302>.
- [136] M. Zhou, X. Wang, H. Shi, L. Cheng, Z. Wang, H. Zhao, L. Yang, J. Sun, Characterization of long non-coding RNA-associated ceRNA network to reveal potential prognostic lncRNA biomarkers in human ovarian cancer, *Oncotarget* 7 (2016) 12598–12611, <https://doi.org/10.18632/oncotarget.7181>.
- [137] D. Cescon, L.L. Siu, Cancer clinical trials: the rear-view mirror and the crystal ball, *Cell* 168 (2017) 575–578, <https://doi.org/10.1016/j.cell.2017.01.027>.
- [138] E. Mardis, Many mutations in one clinical-trial basket, *Nature* 554 (2018) 173–175, <https://doi.org/10.1038/d41586-018-01312-x>.
- [139] D.M. Hyman, S.A. Piha-Paul, H. Won, J. Rodon, C. Saura, G.I. Shapiro, D. Juric, D.I. Quinn, V. Moreno, B. Doger, I.A. Mayer, V. Boni, E. Calvo, S. Loi, A.C. Lockhart, J.P. Erinjeri, M. Scaltriti, G.A. Ulaner, J. Patel, J. Tang, H. Beer, S.D. Selcuklu, A.J. Hanrahan, N. Bouvier, M. Melcer, R. Murali, A.M. Schram, L.M. Smyth, K. Jhaveri, B.T. Li, A. Drilon, J.J. Harding, G. Iyer, B.S. Taylor, M.F. Berger, R.E. Cutler, F. Xu, A. Butturini, L.D. Eli, G. Mann, C. Farrell, A.S. Lalani, R. P. Bryce, C.L. Arteaga, F. Meric-Bernstam, J. Baselga, D.B. Solit, HER kinase inhibition in patients with HER2- and HER3-mutant cancers, *Nature* 554 (2018) 189–194, <https://doi.org/10.1038/nature25475>.
- [140] E. Van Cutsem, C.-H. Köhne, E. Hitre, J. Zaluski, C.-R. Chang Chien, A. Makhson, G. D’Haens, T. Pintér, R. Lim, G. Bodoky, J.K. Roh, G. Folprecht, P. Ruff, C. Stroh, S. Tejpar, M. Schlichting, J. Nippgen, P. Rougier, Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer, *N. Engl. J. Med.* 360 (2009) 1408–1417, <https://doi.org/10.1056/NEJMoa0805019>.
- [141] D.A. Flockhart, T. Skaar, D.S. Berlin, T.E. Klein, A.T. Nguyen, Clinically available pharmacogenomics tests, *Clin. Pharmacol. Ther.* 86 (2009) 109–113, <https://doi.org/10.1038/clpt.2009.39>.
- [142] E.J. Topol, Pharmacy benefit managers, pharmacies, and pharmacogenomic testing: prescription for progress? *Sci. Transl. Med.* 2 (2010), 44cm22, <https://doi.org/10.1126/scitranslmed.3001067>.
- [143] A.G. Schissler, V. Gardeux, Q. Li, I. Achour, H. Li, W.W. Piegorsch, Y.A. Lussier, Dynamic changes of RNA-sequencing expression for precision medicine: N-of-1-pathways Mahalanobis distance within pathways of single

- subjects predicts breast cancer survival, *Bioinformatics* 31 (2015) i293–i302, <https://doi.org/10.1093/bioinformatics/btv253>.
- [144] X. Yang, K. Regan, Y. Huang, Q. Zhang, J. Li, T.Y. Seiwert, E.E.W. Cohen, H.R. Xing, Y.A. Lussier, Single sample expression-anchored mechanisms predict survival in head and neck cancer, *PLoS Comput. Biol.* 8 (2012), e1002350. <https://doi.org/10.1371/journal.pcbi.1002350>.
- [145] V. Gardeux, I. Achour, J. Li, M. Maienschein-Cline, H. Li, L. Pesce, G. Parinandi, N. Bahroos, R. Winn, I. Foster, J.G.N. Garcia, Y.A. Lussier, “N-of-1-pathways” unveils personal deregulated mechanisms from a single pair of RNA-Seq samples: towards precision medicine, *J. Am. Med. Inform. Assoc.* 21 (2014) 1015–1025, <https://doi.org/10.1136/amiajnl-2013-002519>.
- [146] Q. Li, A.G. Schissler, V. Gardeux, J. Berghout, I. Achour, C. Kenost, H. Li, H.H. Zhang, Y.A. Lussier, kMEn: analyzing noisy and bidirectional transcriptional pathway responses in single subjects, *J. Biomed. Inform.* 66 (2017) 32–41, <https://doi.org/10.1016/j.jbi.2016.12.009>.
- [147] Q. Li, A.G. Schissler, V. Gardeux, I. Achour, C. Kenost, J. Berghout, H. Li, H.H. Zhang, Y.A. Lussier, N-of-1-pathways MixEnrich: advancing precision medicine via single-subject analysis in discovering dynamic changes of transcriptomes, *BMC Med. Genet.* 10 (2017), 27. <https://doi.org/10.1186/s12920-017-0263-4>.
- [148] A. Krämer, J. Green, J. Pollard, S. Tugendreich, Causal analysis approaches in Ingenuity Pathway Analysis, *Bioinformatics* 30 (2014) 523–530, <https://doi.org/10.1093/bioinformatics/btt703>.
- [149] A. Livshits, A. Git, G. Fuks, C. Caldas, E. Domany, Pathway-based personalized analysis of breast cancer expression data, *Mol. Oncol.* 9 (2015) 1471–1483, <https://doi.org/10.1016/j.molonc.2015.04.006>.
- [150] V. Gardeux, A.D. Arslan, I. Achour, T.-T. Ho, W.T. Beck, Y.A. Lussier, Concordance of deregulated mechanisms unveiled in underpowered experiments: PTBP1 knockdown case study, *BMC Med. Genet.* 7 (Suppl. 1) (2014) S1, <https://doi.org/10.1186/1755-8794-7-S1-S1>.
- [151] T. Meißner, K.M. Fisch, L. Gioia, A.I. Su, OncoRep: an n-of-1 reporting tool to support genome-guided treatment for breast cancer patients using RNA-sequencing, *BMC Med. Genet.* 8 (2015) 24, <https://doi.org/10.1186/s12920-015-0095-z>.
- [152] D.S. Wishart, C. Knox, A.C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, J. Woolsey, DrugBank: a comprehensive resource for in silico drug discovery and exploration, *Nucleic Acids Res.* 34 (2006) D668–D672, <https://doi.org/10.1093/nar/gkj067>.
- [153] M. Whirl-Carrillo, E.M. McDonagh, J.M. Hebert, L. Gong, K. Sangkuhl, C.F. Thorn, R.B. Altman, T.E. Klein, Pharmacogenomics knowledge for personalized medicine, *Clin. Pharmacol. Ther.* 92 (2012) 414–417, <https://doi.org/10.1038/clpt.2012.96>.
- [154] S.K. Merid, D. Goranskaya, A. Alexeyenko, Distinguishing between driver and passenger mutations in individual cancer genomes by network enrichment analysis, *BMC Bioinformatics* 15 (2014) 308, <https://doi.org/10.1186/1471-2105-15-308>.
- [155] J.P. Hou, J. Ma, DawnRank: discovering personalized driver genes in cancer, *Genome Med.* 6 (2014), 56. <https://doi.org/10.1186/s13073-014-0056-8>.
- [156] D. Bertrand, K.R. Chng, F.G. Sherbaf, A. Kiesel, B.K.H. Chia, Y.Y. Sia, S.K. Huang, D.S.B. Hoon, E.T. Liu, A. Hillmer, N. Nagarajan, Patient-specific driver gene prediction and risk assessment through integrated network analysis of cancer omics profiles, *Nucleic Acids Res.* 43 (2015), e44. <https://doi.org/10.1093/nar/gku1393>.
- [157] X. Liu, Y. Wang, H. Ji, K. Aihara, L. Chen, Personalized characterization of diseases using sample-specific networks, *Nucleic Acids Res.* 44 (2016), e164. <https://doi.org/10.1093/nar/gkw772>.
- [158] W.-F. Guo, S.-W. Zhang, L.-L. Liu, F. Liu, Q.-Q. Shi, L. Zhang, Y. Tang, T. Zeng, L. Chen, Discovering personalized driver mutation profiles of single samples in cancer by network control strategy, *Bioinformatics* (2018) <https://doi.org/10.1093/bioinformatics/bty006>.
- [159] M. Cronin, M. Pho, D. Dutta, J.C. Stephans, S. Shak, M.C. Kiefer, J.M. Esteban, J.B. Baker, Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay, *Am. J. Pathol.* 164 (2004) 35–42, [https://doi.org/10.1016/S0002-9440\(10\)63093-3](https://doi.org/10.1016/S0002-9440(10)63093-3).
- [160] S. Paik, S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, F.L. Baehner, M.G. Walker, D. Watson, T. Park, W. Hiller, E.R. Fisher, D.L. Wickerham, J. Bryant, N. Wolmark, A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer, *N. Engl. J. Med.* 351 (2004) 2817–2826, <https://doi.org/10.1056/NEJMoa041588>.
- [161] M. Cronin, C. Sangli, M.-L. Liu, M. Pho, D. Dutta, A. Nguyen, J. Jeong, J. Wu, K.C. Langone, D. Watson, Analytical validation of the Oncotype DX genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer, *Clin. Chem.* 53 (2007) 1084–1091, <https://doi.org/10.1373/clinchem.2006.076497>.
- [162] K.S. Albain, W.E. Barlow, S. Shak, G.N. Hortobagyi, R.B. Livingston, I.-T. Yeh, P. Ravdin, R. Bugarini, F.L. Baehner, N.E. Davidson, G.W. Sledge, E.P. Winer, C. Hudis, J.N. Ingle, E.A. Perez, K.I. Pritchard, L. Shepherd, J.R. Gralow, C. Yoshizawa, D.C. Allred, C.K. Osborne, D.F. Hayes, Breast Cancer Intergroup of North America, Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial, *Lancet Oncol.* 11 (2010) 55–65, [https://doi.org/10.1016/S1470-2045\(09\)70314-6](https://doi.org/10.1016/S1470-2045(09)70314-6).
- [163] M. Dowsett, J. Cuzick, C. Wale, J. Forbes, E.A. Mallon, J. Salter, E. Quinn, A. Dunbier, M. Baum, A. Buzdar, A. Howell, R. Bugarini, F.L. Baehner, S. Shak, Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study, *J. Clin. Oncol.* 28 (2010) 1829–1834, <https://doi.org/10.1200/JCO.2009.24.4798>.
- [164] W.J. Gradishar, B.O. Anderson, R. Balassanian, S.L. Blair, H.J. Burstein, A. Cyr, A.D. Elias, W.B. Farrar, A. Forero, S. H. Giordano, M.P. Goetz, L.J. Goldstein, S.J. Isakoff, J. Lyons, P.K. Marcom, I.A. Mayer, B. McCormick, M.S. Moran, R.M. O'Regan, S.A. Patel, L.J. Pierce, E.C. Reed, K.E. Salerno, L.S. Schwartzberg, A. Sitapati, K.L. Smith, M. Lou Smith, H. Soliman, G. Somlo, M. Telli, J.H. Ward, D.A. Shead, R. Kumar, NCCN Guidelines Insights: Breast Cancer, Version 1.2017, *J. Natl. Compr. Cancer Netw.* 15 (2017) 433–451 (<http://www.ncbi.nlm.nih.gov/pubmed/28404755>, accessed March 8, 2018).
- [165] A.A. Alizadeh, M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, J.I.

- Powell, L. Yang, G.E. Marti, T. Moore, J. Hudson, L. Lu, D.B. Lewis, R. Tibshirani, G. Sherlock, W.C. Chan, T.C. Greiner, D.D. Weisenburger, J.O. Armitage, R. Warnke, R. Levy, W. Wilson, M.R. Grever, J.C. Byrd, D. Botstein, P.O. Brown, L.M. Staudt, Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, *Nature* 403 (2000) 503–511, <https://doi.org/10.1038/35000501>.
- [166] D.W. Scott, G.W. Wright, P.M. Williams, C.-J. Lih, W. Walsh, E.S. Jaffe, A. Rosenwald, E. Campo, W.C. Chan, J.M. Connors, E.B. Smeland, A. Mottok, R.M. Braziel, G. Ott, J. Delabie, R.R. Tubbs, J.R. Cook, D.D. Weisenburger, T.C. Greiner, B.J. Glinsmann-Gibson, K. Fu, L.M. Staudt, R.D. Gascoyne, L.M. Rimsza, Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue, *Blood* 123 (2014) 1214–1217, <https://doi.org/10.1182/blood-2013-11-536433>.
- [167] P.N. Meyer, K. Fu, T.C. Greiner, L.M. Smith, J. Delabie, R.D. Gascoyne, G. Ott, A. Rosenwald, R.M. Braziel, E. Campo, J.M. Vose, G. Lenz, L.M. Staudt, W.C. Chan, D.D. Weisenburger, Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab, *J. Clin. Oncol.* 29 (2011) 200–207, <https://doi.org/10.1200/JCO.2010.30.0368>.
- [168] W.W.L. Choi, D.D. Weisenburger, T.C. Greiner, M.A. Piris, A.H. Banham, J. Delabie, R.M. Braziel, H. Geng, J. Iqbal, G. Lenz, J.M. Vose, C.P. Hans, K. Fu, L.M. Smith, M. Li, Z. Liu, R.D. Gascoyne, A. Rosenwald, G. Ott, L.M. Rimsza, E. Campo, E.S. Jaffe, D.L. Jaye, L.M. Staudt, W.C. Chan, A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy, *Clin. Cancer Res.* 15 (2009) 5494–5502, <https://doi.org/10.1158/1078-0432.CCR-09-0113>.