

Review

Integrated network analyses for functional genomic studies in cancer

Jennifer L. Wilson^a, Michael T. Hemann^b, Ernest Fraenkel^a, Douglas A. Lauffenburger^{a,*}^a Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, United States^b Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge MA, United States

ARTICLE INFO

Keywords:

Regulatory networks
Signaling pathways
Computational modeling
RNAi

ABSTRACT

RNA-interference (RNAi) studies hold great promise for functional investigation of the significance of genetic variations and mutations, as well as potential synthetic lethalties, for understanding and treatment of cancer, yet technical and conceptual issues currently diminish the potential power of this approach. While numerous research groups are usefully employing this kind of functional genomic methodology to identify molecular mediators of disease severity, response, and resistance to treatment, findings are generally confounded by “off-target” effects. These effects arise from a variety of issues beyond non-specific reagent behavior, such as biological cross-talk and feedback processes so thus can occur even with specific perturbation. Interpreting RNAi results in a network framework instead of merely as individual “hits” or “targets” leverages contributions from all hit/target contributions to pathways *via* their relationships with other network nodes. This interpretation can ameliorate dependence upon individual reagent performance and increase confidence in biological validation. Here we provide background on RNAi studies in cancer applications, review key challenges with functional genomics, and motivate the use of network models grounded in pathway analyses.

© 2013 The Authors. Published by Elsevier Ltd. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

1. Introduction

Discovery of gene products vital for function of a biological system, using gene-interference studies at has become increasing popular because of the capability for RNAi methods for manipulating multiple cellular components in either biased or unbiased manner. These experiments aspire to identify high-confidence “hit” sets as putatively responsible for an experimental phenotype and conceivably imaginable as drug “targets”, although requiring dedicated follow-up tests to buttress confidence in validity. Typically, the findings from the initial “screen” study are compiled as list of individual genes whose knockdown yielded significant alteration of biological system function, and the follow-up validation experiments are considered in isolation. While there are encouraging successes along this avenue, the realization that molecular components executing or governing cell/tissue phenotypic operation work in concert among myriad dynamic partners

– directly and indirectly – motivates appreciation for considering a more integrative perspective on interpretation of RNAi-based functional genomic studies.

‘Concerted’ operation brings to mind an instrumental orchestra as one notional metaphor. Proper generation of a musical program depends on the collective efforts of the players involved, and deviations of any individual in pitch, volume, or timing can produce inappropriate sound and affect the overall orchestral performance as other individuals attempt to adapt – or naturally produce further errors themselves. The sound of any particular individual is rarely decisive, while an instrumental section can either mitigate or amplify aberrations and other instrumental sections may aim to compensate. Accordingly, flawed performance may be viewed as arising from identifiable “drivers” but sustained pathology is more likely manifested by inability of the overall company to find an appropriate new balance *via* diverse modulations. And when aspiring for remediation, as the music proceeds the original deviations no longer remain the most effective points of correction because the propagated adaptations and compensations render a simple “re-set” difficult to achieve dynamically.

We use this integrative, or ‘concerted’ point of view to inform our recommendations about the investigation of cancer systems using RNAi. We offer that a most effective framework uses multi-node pathways for gaining greatest insight about how a system is dysregulated and for how that system might be remediated, and further that this point of view is essential to RNAi analyses.

* Corresponding author at: 77 Massachusetts Ave, 16-343 Cambridge, MA 02139, United States. Tel.: +1 617 252 1629; fax: +1 617 258 0204.

E-mail addresses: jlwilson@mit.edu (J.L. Wilson), hemann@mit.edu (M.T. Hemann), fraenkel@mit.edu (E. Fraenkel), lauffen@mit.edu (D.A. Lauffenburger).

2. RNAi screens as a tool for cancer biology

Because cancer is a mutation-driven disease, many investigators have focused on using genetic characterizations of cancers, yet there are often non-intuitive relationships between gene features and disease phenotypes [1–4]. Much is known about the cancer genome landscape, yet, while hundreds of human genes have been linked to cancer, mutations are not always consistent across patients, and disease severity may not correlate with mutational status alone [2,5–8]. Further, occurrence of drug resistance also does not exhibit direct correlation with mutational status [3,9]. For instance, in pediatric medulloblastoma, systematic measurement of mutation-status and transcriptional profiling revealed that mutation rates are not consistent across pediatric tumors [9,10].

In our orchestral analogy, these investigations are akin to rating the quality of the company using each players' individual audition. This perspective lacks context and an understanding of the player's contribution to the orchestra's performance. To account for this context, investigators have turned to RNA-mediated interference (RNAi) technologies to fine tune a genetic player's ability. These tools manipulate genetic features at a functional level and may be a complementary approach for studying the non-intuitive relationship between mutation, expression, and disease phenotype [6,9,11] just as a conductor may better appreciate a musician's performance while playing within their section.

From an engineering perspective, gene-interference experiments are attractive experiments for understanding cancer because of the opportunity to modulate gene function under diverse, potentially relevant conditions. Investigators have targeted single genes, or multiple genes together, in large scale screens, as well as pathway specific studies [6,9]. When investigating genetic amplifications in liver cancer, one group simultaneously explored the role of these amplification events and the relative contribution of the *in vivo* environment with a genome-scale RNAi screen [12,13]. In this instance, and many others, RNAi screens afford the opportunity to explore numerous targets simultaneously. The Achilles Project from the Broad Institute added another dimension to genome-wide screens by drastically increasing the scale of their investigation and challenging the reproducibility of shRNA libraries. They introduced a library of shRNAs into more than 100 established cancer cell lines and identified functional phenotypes that were common and unique to each cell line [14,15]. Researchers can take advantage of varying RNAi reagent targeting efficacy to create titrations of gene interference, known as epi-allelic series [14,16]. This technique manipulates variation in mRNA expression to create a gradient of disease phenotype. As expected, this approach created varying lymphoma phenotypes which increased in disease severity as shRNA targeting efficiency against p53 increased [16]. While we note here only a few investigations, RNAi experiments lend themselves to the perturbation of many more parameters: multiple cues, multiple dosing schemes, multiple environments, and multiple time points.

RNAi reagents hold significant advantages over other interference methods, such as small molecule inhibitors. More specifically, siRNA offers the advantage of isoform specificity and enables fine-tuning of individual isoform expression and activity. For an investigation of T-cell Erk regulation, researchers used epi-allelic series with siRNAs against ERK1 and ERK2 to identify the role of these kinases on downstream IL-2 production [17]. The epi-allelic series again showed a correlation between siRNA targeting efficiency and phenotype. In addition, the researchers identified that IL-2 production scaled with total ERK activation and was not isoform specific. When comparing the siRNA-mediated effects on IL-2 to those of a MEK inhibitor's effect, they also found that the gene-interference methods reduced IL-2 production to a greater extent than chemical inhibitor dosing at an equivalent level of ERK activation [2,8,17]. From this finding they inferred that ERK may also have a role as a

scaffold in downstream IL2 production; such a phenomenon may have not been indicated using only either approach alone.

3. RNAi screening challenges

Gene interference screens are quickly becoming high-throughput, but they are poorly suited to the well-accepted data analysis tools from other 'omics biology experiments. Birmingham et al. provide a thorough review of statistical adaptations for target discovery from RNAi experiments [1,3]. Generally, these adaptations consist of normalization, and some means of 'top-hit' identification based on outstanding performance relative to the remaining population. However, inconsistent reagent performance limits statistical power and subsequent validation of these candidates often fails.

Variability in RNAi screening data can derive from a variety of factors, both off-target and crosstalk events, and cause varying rates of false positives and false negatives in RNAi screens, reducing confidence in final hit selection [6,7,10]. Off-target events are a non-specific result of the experimental reagents, and may include the inadvertent knockdown of additional transcripts through microRNA-like effects and the incomplete knockdown of a protein target due to a protein half-life greater than the experimental timeline. Crosstalk events, on the other hand, are a result of the biological response to RNAi perturbation as opposed to the experimental reagents used. These events may include increased expression of transcripts normally repressed by microRNAs that have to compete for use of the internal degradation machinery, and increased expression or activity of proteins which are compensatory for the RNAi target [6,9,11].

Many approaches attempt to compensate for off-target effects. One method utilizes multiple RNAi reagents against the same gene, and only considers the gene a hit if multiple reagents yield a similar phenotype [6,9]. However, the ability to identify true positives from redundant reagents is complicated by the targeted gene product's context within the cell [9,13]. For example, unintended effects are less likely for gene targets with highly specific, non-redundant roles or those that exist in linear pathways. However, for highly connected genes or those involved in multiple pathways, there is a greater chance of biological crosstalk, and thus varied results between redundant siRNAs [9,15].

A genome-wide screen for homologous recombination (HR) mediators highlights the role of unintended effects and how redundant RNAi reagents may mislead results [12,16]. For instance, 5 out of 10 RNAi reagents against the HIRIP3 gene decreased capacity for homologous recombination. While all reagents successfully reduced mRNA expression, rescue experiments with RNAi-resistant mRNA failed to recover homologous recombination activity. Further, relative mRNA expression changes did not correlate with changes in homologous recombination.

Computational analyses of sequence similarity between siRNA reagents and non-targeted, mRNA transcripts can predict off-target effects but is imperfect in all situations. Genome-wide enrichment of seed sequences (GESS) analysis looks for enrichment of non-targeted 3' UTR regions in siRNA sense and antisense sequences [14,16]. In theory, these 3' UTR matches identify unintended target genes and subsequent modulation of these genes should recapitulate the phenotype erroneously assigned to the original siRNA. The method successfully identifies genes enriched in active siRNAs for multiple screens, and can filter primary screening hits to decrease the false positive rate [14,17].

In the previously mentioned screen for homologous recombination mediators, GESS analysis identified a significant enrichment for RAD51 3' UTR in the high-scoring, non-RAD51 siRNAs [12]. As expected, RAD51 mRNA was depleted in the presence of 4 of the 7 siRNAs against HIRIP3 and RAD51 mRNA levels better correlated

with changes in the homologous recombination phenotype than HIRIP3 mRNA levels. Yet, only 1 of the 7 HIRIP3 siRNAs actually contained the seed match for the RAD51 UTR demonstrating that additional cross-talk events may occur in the presence of the HIRIP3 siRNAs. While GESS successfully identified RAD51 mRNA levels as the true predictor for homologous recombination, it was unable to fully explain the observed changes in this gene's transcription, as all HIRIP3 siRNAs did not reduce RAD51.

4. Biological networks add contextual value

A network framework enables researchers to consider contextual influences on how pathway components assimilate, integrate and propagate knowledge in a manner that is distinct from the list model [18,19]. More specifically, a network motif, consisting of a coherent group of functionally related genetic regulators, may better explain an observed phenotype where statistically-ranked lists are insufficient [4,20]. Already, these network motifs for target discovery have led to better understanding of the non-intuitive relationships between genotype and disease phenotype and identification of better therapeutic targets [4,8].

Networks can be useful for predicting drug targets and also for selecting drug combinations [19]. Their functional context provides rational selection of single targets as well as combinatorial targets that could synergistically affect a desired phenotype because they consider pathway membership [19]. Where toxicity had previously constrained the selection of combination therapies, researchers may now instead prioritize combinations based on specificity to controlling a particular phenotype. Understanding macromolecular pathways led to conclusions about synergies between doxorubicin and TNF-alpha therapies [21]. The authors showed that administering TNF-alpha as an adjuvant to doxorubicin treatment increased apoptotic cell death in the presence of low-levels of DNA damage by using an integrated network approach. Without pathway and network-level information, this non-intuitive relationship may have been missed.

Network interpretation has already added depth to non-intuitive instances of drug resistance. Recently, Wilson et al. showed that growth-factors within the tumor microenvironment may increase resistance to kinase inhibitor therapy [22]. While this might seem counterintuitive in a linear-process formalism, considering the cell's underlying signaling network make these results less surprising. Wagner et al. used network inference methods to create interaction networks by combining systematic RNAi-perturbation data with phosphorylation information at multiple time points for six receptor-tyrosine kinases (RTKs) (EGFR, FGFR1,c-Met,IGF-1R,NTRK2, and PDGFR β) [23]. From the resulting networks, they clustered each RTK network, identifying core signaling components shared between all RTKs as well as cluster-specific modules. They postulated that modules shared between RTKs within the same cluster could explain resistance to targeted RTK therapy. More specifically, if RTKs of a particular class shared signaling components and affected the same downstream phenotypes, then these within-cluster RTKs could compensate for chemical inhibition by actuating the original downstream phenotype [23]. They demonstrated this compensation within the EGFR/c-Met/FGFR1 cluster by showing correlation of receptor expression with resistance to therapies targeted to other within-cluster RTKs.

5. Functional pathways explain target variability

A meta-analysis of nine RNAi screens for HIV-replication factors used functional enrichment to explain discrepancies across high-scoring targets from each screen [24]. When they investigated

the percentage of scoring targets across three screens, this overlap only included a modest 3–6% of gene targets. They show that variability between screens, variability between experimental timing and toxicity thresholds all contributed to the minimal overlap among these screens. However, when they looked at gene membership in GO ontology categories, they found much greater overlap in the enrichment of GO categories across screens than in the individual gene targets. This finding indicates that a more global, functional filter is useful for identifying true positives from highly variable RNAi screens.

Additionally, using functional pathway membership increased experimental validation rates in an RNAi screen for DNA-damage mediators [25]. The authors screened all protein-coding genes in *Drosophila melanogaster* and compared top hits to an analogous screen in *Saccharomyces cerevisiae*, but did not see a statistically significant overlap between screening targets [25]. They expanded their target list by identifying functional pathways that contained gene targets from their screen in *Drosophila melanogaster* and were able to find enrichment of targets from the *Saccharomyces cerevisiae* screen within this set. Further, using this expanded set for validation experiments identified false negatives from the original screen. These results reaffirm the utility of filtering data by pathway membership to identify true positives and also using pathway membership as a search space for false negatives.

In a pioneering study, Jones et al. demonstrated the significance of using pathway context in a patient setting [26]. They performed a global analysis of mutations in pancreatic cancers, but found little overlap in the specific mutations across patients. However, they instead found a core set of signaling pathways that consistently enriched for patient-specific mutations. They postulate that targeting the physiological consequences of these pathways instead of the individual mutations would improve therapeutic development [26]. If we consider the discrepancy between RNAi reagent performance across replicates as similar to the mutational differences between patients, these findings present more motivation for using a pathway-centered approach for functional genomic studies.

6. Network integration is a viable tool for hypothesis generation

Given the importance of understanding the functional context of a genetic alteration, network methods are a useful computational tool. Additionally, these tools enable the incorporation of multiple data sets and experiments to create more holistic interpretations of biological systems. Because of the availability of many experimental datasets through various databases, data integration will be influential in future investigations [27]. Here, we review a few integrated network approaches and highlight how networks have improved the interpretation of biological investigations and affected further hypothesis generation.

In metastatic breast cancer, integrating copy-number variation (CNV) and gene expression data across multiple samples accurately predicted novel drivers of disease [28]. The authors used a refined method for first identifying recurrent CNVs from gene expression data and then used a Bayesian methodology to create a network of mutated genes. From this network, they found master regulators by selecting genes that had a high authority score. Mathematically, the authority score identified genes with a statistically significant number of outgoing connections as compared to the mean number of connections. To test their hypotheses about mediators for breast cancer, they performed an siRNA screen testing the effect of gene interference on cell viability. Of the gene targets that had the greatest effect on cell viability, they found a significant enrichment of their high-authority regulators [28]. This finding demonstrates that networks can synchronize disparate datasets and that network properties are viable characterizations for finding novel regulators.

Utilizing a data-integration approach, Huang et al. constructed an EGFRVIII-specific signaling network for glioblastoma multiforme (GBM) by incorporating proteomic and transcriptional data [29]. More interestingly, they used this network to identify, test and validate novel therapies. Their final networks consisted of a high-confidence set of experimental data points as well as gene targets not included in the original set, but rather added through known protein-protein interactions. From this network set, they systematically expanded targets for therapeutic intervention by identifying targets with known chemical inhibitors and ranking them based on their proximity to the core functional network. From this target set, they identified compounds in clinical trial with known effects on cancer systems and chemical inhibitors not yet tested for GBM [29].

In the interferon-stimulatory DNA (ISD) sensing pathway, an integrated network approach proved successful for identifying novel regulators of this process and for testing new therapeutics [30]. In this analysis, the authors created an interaction network of potential ISD regulators by combining direct interacting partners of known ISD pathway components with interacting pairs from their own quantitative mass-spectrometry experiments. Perturbation of this compendium network with RNAi reagents identified Abcf1, Cdc37, and Ptpn1 as effectors of the ISD-sensing response to dsDNA. In this situation, curating and expanding interaction information around known pathway components successfully identified novel genes for the ISD response. The authors also measured ISD-pathway induction after treatment with chemical inhibitors against their novel genes and demonstrated a reduction in deleterious interferon production. These results show that integration is useful for developing new hypotheses for therapeutic development and supports the Jones et al. perspective concerning efficacy of designing therapeutic options around downstream pathway physiology [26].

Data integration within a network framework also added depth to understanding metabolic disorders using SNP and genetic linkage data [31]. In this investigation, researchers created a network where interactions depended upon significant co-expression and linkage data between genes. Using optimization, they selected highly connected gene sub-modules and then used these modules for further hypothesis generation. Many sub-modules were enriched for genetic features that were significantly associated with disease traits (fat mass, weight, plasma insulin levels, etc.) and one sub-module was significantly enriched for genetic features with significant correlation to all disease traits. They expanded this module, and created a macrophage-driven superior module from which they selected and further perturbed genetic loci. From these perturbations, they were able to demonstrate the sub-network's contribution to the observed disease traits and classify genetic features previously not associated with metabolic traits. This collective understanding of genetic interactions ultimately created a more comprehensive view of their dataset and novel hypotheses distinct from those that would be identified from looking at gene-trait associations independently.

Recently, our own efforts investigating *in vivo* mediators of Acute Lymphoblastic Leukemia (ALL) have employed a data integration approach to ascertain GO biological function enrichment rather than to looking at screening targets independently (unpublished). A B-cell model of ALL was infected with a genome-scale shRNA library and after infection, cells were plated *in vitro* or tail-vein injected into syngeneic recipient mice. After disease developed, cells were harvested and sequenced for final shRNA representation. To analyze this data we used Simultaneous Analysis of Multiple Networks (SAMNet), which is a flow-based formalism which relates screening hits to downstream expression data using the interactome as a guide for possible connections among the data [32]. The method generated a network enriched for functional pathways, such as developmental processes, that are known to play a

role in ALL – whereas these were not identified when analyzing experimental data independently. This enrichment increases confidence that RNAi hits identified within the network are true positives. Further, SAMNet adds targets, or nodes, to the network that were not present in the original high-scoring target set, making it possible to hypothesize about potential false negatives in the data.

In these examples, data analysis in isolation was insufficient for discovering novel regulators and targets for therapeutic intervention. Instead, a concerted network approach, integrating multiple data sets or experimental results, improved target identification and created testable hypotheses for therapeutic development.

7. Reciprocal engineering for future gene-interference investigations

Understanding and modulating cancer requires a concerted understanding of gene function and appreciation for each gene's pathway membership. Much like an orchestra, the performance of the group depends on the collective group effort rather than the ability of any one player. Auditioning players individually is important for assessing skills and musicality, yet their full potential depends on their ability to contribute to the sound of the group. Gene-interference studies are the experimental parallel of 'auditioning', yet their interpretation is limited if each player is considered in isolation. Instead, the conductor must observe the player within his section to see if deficiencies affect the overall sound or if the sound of his peers compensate for his weaknesses. In the same way, building biological networks using RNAi experimental data analyzes the player in his section, and uses his pathway membership to assess his effect on the sound of the orchestra.

'Network Filtering' techniques will increasingly become a secondary post-processing step to statistical analyses for gene-interference studies. We have conceptualized how network motifs may complement existing statistical approaches in Fig. 1. The

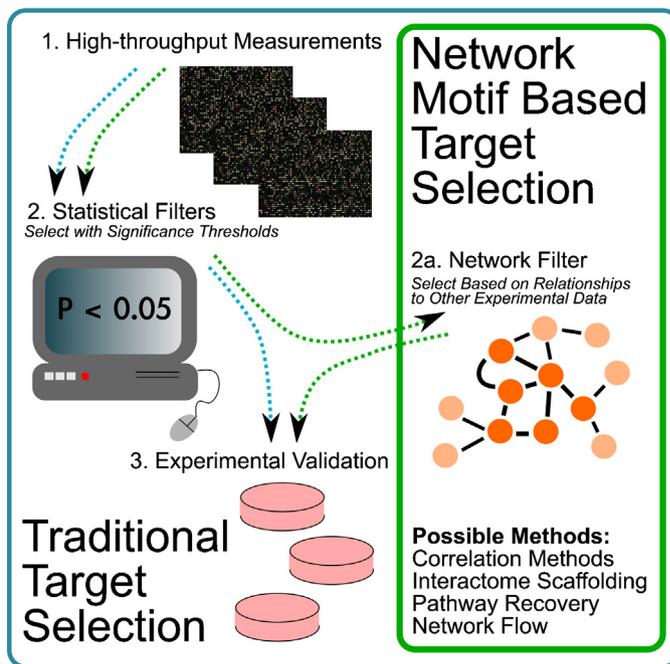


Fig. 1. Network filtering uses functional relations to identify candidate targets. Traditional methods for target selection collect high-throughput data, then use statistical methods to filter data for top candidates and further experimental validation. A network filtering approach would collect high-throughput data, perform statistical analyses, but would add an additional network-construction step to find functional consensus among top targets before identifying candidate regulators or predicted pharmacological targets and moving to experimental validation.

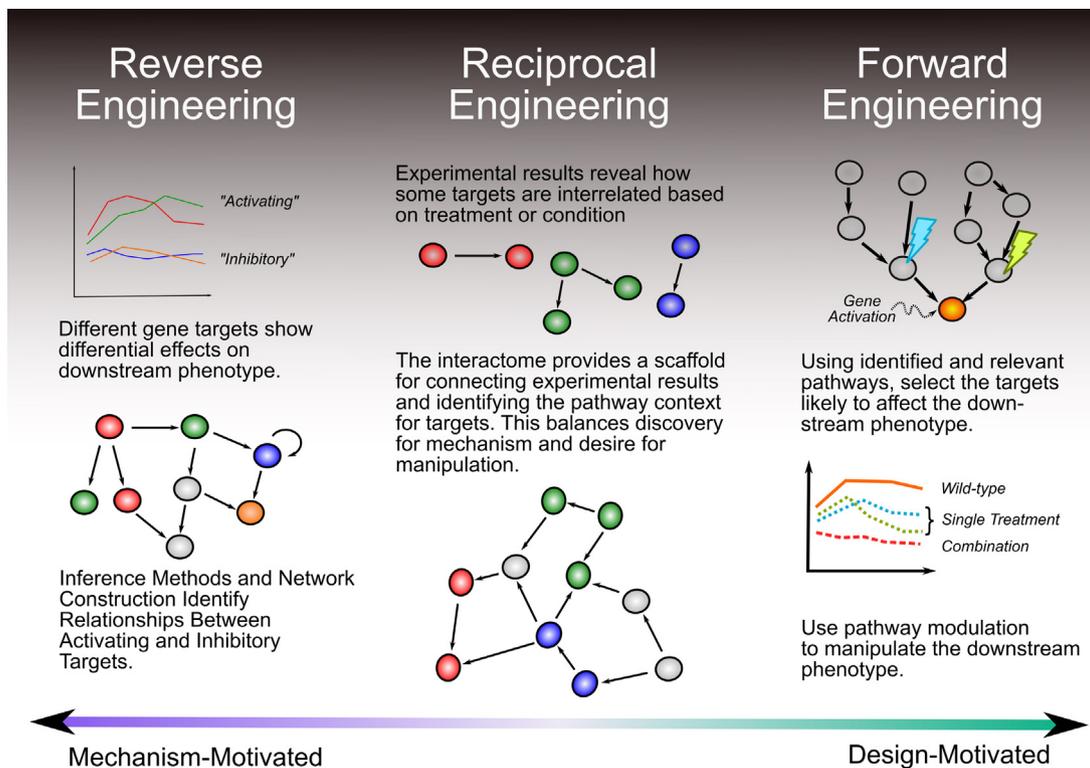


Fig. 2. Reciprocal engineering balances forward and reverse engineering principles. Reciprocal Engineering balances the mechanism-driven paradigm of reverse engineering with the design-motivated paradigm of forward engineering. In the graphical representation, reverse engineering is exemplified by the desire to uncover mechanisms explaining relationships among the data – for example using correlation to infer relationships among expression. On the other hand, forward engineering is using a known network to manipulate the system – in this example, inhibiting node functions to alter downstream gene activation. Blending both paradigms, Reciprocal Engineering balances relationship discovery with design motives to create networks that infer biological function as well as offer points for manipulation.

network provides context for any RNAi reagent's performance and leverages a gene target's pathway membership to identify true and false positives. This approach assumes that if multiple RNAi target genes enrich within a pathway, they are more likely contributing to an experimental phenotype. Similarly, if an RNAi-targeted gene stands out experimentally, but does not have interactions with other targets, this target may be a false positive. In this mindset, networks offer a filter for removing false positives from entering the final 'hit-list'. Where RNAi study results are affected by unintended effects, noise-reduction techniques will have a large impact on data interpretation.

However, there is still much to be learned about gene pathways and their modulation for cancer systems. While pathways provide context for gene function, pathways themselves are context specific and require systematic perturbation [5]. For instance, melanoma patients with BRAF mutations have drastically different sensitivity to therapy than colorectal cancer patients with similar mutations [5]. As such, many considerations remain about interpreting pathways and using them to refine experimental evidence.

Conceptualizing pathways instead of individual molecules changes the hypotheses generated as well as the experimental validations that follow [18]. For instance, because of the increased level of interconnectedness, disease modules are computationally identifiable by graph theory parameters such as clustering coefficients, and shorter path lengths. Further, designing validation experiments around these modules may provide novel insight into understanding disease and also improve correlation between predicted perturbation and experimental phenotype [8,18]. This conceptualization further requires consideration of network properties instead of only experimental phenotype and may instead prioritize candidates based on number of connections (network degree) or enrichment against randomized networks. Yet, it is

unclear which graphical parameters are most predictive of false positives or false negatives and whether these parameters are consistently predictive across multiple systems.

As network motif discovery becomes more common, we envision an accompanying shift in approach to these methods. This shift incorporates reverse engineering principles through a desire to find models that explain system behaviors as well as forward engineering principles in which the investigator designs the system to control a particular phenotype. We propose that future efforts to construct and manipulate cancer networks will use an 'Reciprocal Engineering' approach (Fig. 2). In this 'Reciprocal Engineering' mindset, researchers balance motivation to explain a system with motivation to design a controllable system. Already researchers have alluded to the value of 'Integrated Interactomics' and how future data integration which balances motivations for understanding and prediction, changes our investigation of cancer systems [4].

Recent commentaries in the area underscore the potential impact of this paradigm shift. These articles concur with the notion that signaling pathways drive cancer progression, and are a rich source of targets for therapeutic development [5,20]. Both biological network models and gene-interference studies are cutting edge techniques that have greatly added to our understanding of cancer systems. As such, future endeavors merging these growing fields will enhance understanding of cancer systems and improve ability to manipulate a complicated disease.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgements

The authors would like to thank our funding sources: NSF Graduate Research Fellowship Program (JLW), NCI Integrative Cancer Biology Program grant U54-CA112967 (DAL/EF/MH), and NCI grant U01-CA155758 (DAL/MH).

References

- [1] Birmingham A, Selfors LM, Forster T, Wrobel D, Kennedy CJ, Shanks E, et al. Statistical methods for analysis of high-throughput RNA interference screens. *Nat Methods* 2009;6(8):569–75.
- [2] Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009;458(7239):719–24.
- [3] McLeod HL. Cancer pharmacogenomics: early promise, but concerted effort needed. *Science* 2013;339(6127):1563–6.
- [4] Berger B, Peng J, Singh M. Computational solutions for omics data. *Nat Rev Genet* 2013;14(5):333–46.
- [5] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science* 2013;339(6127):1546–58.
- [6] Mohr S, Bakal C, Perrimon N. Genomic screening with RNAi: results and challenges. *Annu Rev Biochem* 2010;79(1):37–64.
- [7] Kaelin WG. Molecular biology. Use and abuse of RNAi to study mammalian gene function. *Science* 2012;337(6093):421–2.
- [8] Vidal M, Cusick ME, Barabási A-L. Interactome networks and human disease. *Cell* 2011;144(6):986–98.
- [9] Sigoillot FD, King RW. Vigilance and validation: keys to success in RNAi screening. *ACS Chem Biol* 2011;6(1):47–60.
- [10] Pugh TJ, Weeraratne SD, Archer TC, Pomeranz Krummel DA, Auclair D, Bochicchio J, et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* 2012;488(7409):106–10.
- [11] Jackson AL, Linsley PS. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rev Drug Discov* 2010;9(57):1–11.
- [12] Adamson B, Smogorzewska A, Sigoillot FD, King RW, Elledge SJ. A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat Cell Biol* 2012;14(3):318–28.
- [13] Zender L, Xue W, Zuber J, Semighini CP, Krasnitz A, Ma B, et al. An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. *Cell* 2008;135(5):852–64.
- [14] Sigoillot FD, Lyman S, Huckins JF, Adamson B, Chung E, Quattrocchi B, et al. A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. *Nature Methods* 2012;9(4):363–6.
- [15] Mesirov J, Garraway L. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *PNAS* 2011;108:12372–7.
- [16] Hemann MT, Fridman JS, Zilfou JT, Hernando E, Paddison PJ, Cordon-Cardo C, et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet* 2003;33(3):396–400.
- [17] Wille L, Kemp ML, Sandy P, Lewis CL, Lauffenburger DA. Epi-allelic Erk1 and Erk2 knockdown series for quantitative analysis of T cell Erk regulation and IL-2 production. *Mol Immunol* 2007;44(12):3085–91.
- [18] Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* 2011;12(1):56–68.
- [19] Kholodenko B, Yaffe MB, Kolch W. Computational approaches for analyzing information flow in biological networks. *Sci Signal* 2012;5(220):re1.
- [20] Yaffe MB. The scientific drunk and the lamppost: massive sequencing efforts in cancer discovery and treatment. *Sci Signal* 2013;6(269):pe13–23.
- [21] Tentner AR, Lee MJ, Ostheimer GJ, Samson LD, Lauffenburger DA, Yaffe MB. Combined experimental and computational analysis of DNA damage signaling reveals context-dependent roles for Erk in apoptosis and G1/S arrest after genotoxic stress. *Mol Syst Biol* 2012;8:1–18.
- [22] Wilson TR, Fridlyand J, Yan Y, Penuel E, Burton L, Chan E, et al. Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature* 2012;487(7408):505–9.
- [23] Wagner J, Wolf-Yadlin A, Sevecka M, Grenier J, Root D, Lauffenburger DA, et al. Receptor tyrosine kinases fall into distinct classes based on their inferred signaling networks. *Sci Signal* 2013 [in press].
- [24] Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, et al. Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog* 2009;5(5): e1000437.
- [25] Ravi D, Wiles AM, Bhavani S, Ruan J, Leder P, Bishop AJR. A network of conserved damage survival pathways revealed by a genomic RNAi screen. *PLoS Genet* 2009;5(6):e1000527.
- [26] Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* 2008;321(5897):1801–6.
- [27] Kholodenko B, Yaffe MB, Kolch W. Computational approaches for analyzing information flow in biological networks. *Sci Signal* 2012;5(220).
- [28] Tran LM, Bin Zhang, Zhang Z, Zhang C, Xie T, Lamb J, et al. Inferring causal genomic alterations in breast cancer using gene expression data. *BMC Syst Biol* 2011;5(1):121.
- [29] Huang S-SC, Clarke DC, Gosline SJC, Labadorf A, Chouinard C, Gordon W, et al. Linking proteomic and transcriptional data through the interactome and epigenome reveals a map of oncogene-induced signaling. *PLoS Comput Biol* 2013;9(2):e1002887.
- [30] Lee MN, Roy M, Ong S-E, Mertins P, Villani A, Li W, et al. Identification of regulators of the innate immune response to cytosolic DNA and retroviral infection by an integrative approach. *Nat Immunol* 2012;14(2):179–85.
- [31] Chen Y, Zhu J, Lum PY, Pinto S, MacNeil DJ, Zhang C, et al. Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008;452(7186):429–35.
- [32] Gosline SJC, Spencer S, Orsu O, Fraenkel E. SAMNET: a network-based approach to integrate multi-dimensional high throughput datasets. *Integr Biol* 2012;4(11):1415–27.