Seminars in Cancer Biology 23 (2013) 213–218

Contents lists available at ScienceDirect

Seminars in Cancer Biology

journal homepage: www.elsevier.com/locate/semcancer

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 lethalities, 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.



## 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).

1044-579X/$ – see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.semcancer.2013.06.004

## 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