Science Signaling
AAAS

` **Integrative Analysis of Genome-Wide RNA Interference Screens**

Jason D. Berndt, Travis L. Biechele, Randall T. Moon and Michael B. Major (12 May 2009)

Science Signaling **2** (70), pt4. [DOI: 10.1126/scisignal.270pt4]

Science Signaling (ISSN 1937-9145) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005.

BIOINFORMATICS

Integrative Analysis of Genome-Wide RNA Interference Screens

Jason D. Berndt,* Travis L. Biechele, Randall T. Moon, Michael B. Major

Published 12 May 2009; Volume 2 Issue 70 pt4

High-throughput genetic screens have exponentially increased the functional annotation of the genome over the past 10 years. Likewise, genome-scale efforts to map DNA methylation, chromatin state and occupancy, messenger RNA expression patterns, and disease-associated genetic polymorphisms, and proteome-wide efforts to map protein-protein interactions, have also created vast resources of data. An emerging trend involves combining multiple types of data, referred to as integrative screening. Examples include papers that report integrated data generated from large-scale RNA interference screens on the Wnt/β-catenin pathway with either genotypic or proteomic data in colorectal cancer. These studies demonstrate the power of data integration to generate focused, validated data sets and to identify high-confidence candidate genes for follow-up experiments. We present the ongoing evolution and new strategies for the integrative screening approach with respect to understanding and treating human disease.

Presentation Notes

Slide 1: Science Signaling *logo*

The slideshow and notes for this presentation are provided by *Science Signaling* (www.sciencesignaling.org).

Slide 2: Integrative analysis of genome-wide RNA interference screens

Genomic and proteomic screens generate enormous, complex, and often intractable data sets. We present the concept that the integration of multiple types of screening data [for example, short interfering RNA (siRNA) data and proteomics data] can generate testable models of signal transduction and help to define the cellular context in which these models are applicable to human disease.

Slide 3: Development of integrated genome-wide screening approaches

Technological advancements made over the past 20 years allow researchers to address global questions about signal transduction and cell behavior. Two advancements in particular, genome sequencing and the discovery of RNA interference (RNAi), have enabled genome-wide reverse genetic screens in various model systems. As a result of continued improvements in technology, the cost associated with large-scale screening is decreasing. Consequently, the number of papers published that use large-scale screening technologies has increased exponentially (*1*). Although these experiments provide invaluable functional annotation for every gene in the genome, they also yield immense data sets that require calculated reduction.

Slide 4: Benefits of genome-wide RNAi screens

There are a number of reasons why an investigator would perform a genome-wide functional screen. When properly constructed, these screens are unbiased, comprehensive, and internally controlled—meaning that they are designed to test every gene in a genome, including all the known members of a pathway, under the same conditions and at the same time. This eliminates artifacts that occur when comparing data from different experiments, labs, or tissues, or combinations thereof. Perhaps most important, the resulting screen data provide an invaluable hypothesis-generating tool, one that benefits both the investigating laboratory and the scientific community.

Slide 5: Defining new regulators of Wnt/β-catenin signal transduction

We performed a genome-wide, near-saturation RNAi screen for genes involved in Wnt/β-catenin signal transduction in human colorectal carcinoma cells (*2*). Here we describe a method for arriving at a highly validated list of hits from a siRNA screen, and we justify the purpose of integrating siRNA hits with information about protein-protein interactions. We applied this approach to the Wnt/β-catenin signal transduction pathway to create an integrated physical and functional map that illustrates the relationships of both known and previously unknown mediators of Wnt signaling that could lead to novel therapeutic insights for treating colorectal cancer.

Slide 6: Wnt/β-catenin signal transduction

According to the National Cancer Institute, there were ~150,000 new cases and 50,000 deaths from colorectal cancer in 2008, accounting for ~10% of new cancer cases and cancer-related deaths. The lifetime risk of developing colorectal cancer is 1 in 19 (*3*). In 90% of sporadic colorectal cancers, the Wnt/β-catenin pathway is rendered constitutively active by somatic mutations. Most common are mutations that truncate the adenomatosis polyposis coli (APC) protein or those in β-catenin that prevent its phosphorylation. Either of these types of mutations leads to the accumulation of β-catenin and increased β-catenin–dependent transcription. Ultimately, activation of β-catenin–dependent transcription drives neoplastic transformation (*4*). Our goal was to define the suite of proteins that regulate Wnt/β-catenin signaling in colorectal carcinoma cells with siRNA screening. This slide shows the core of the Wnt/β-catenin pathway. In the absence of Wnt, β-catenin is phosphorylated and targeted for proteasome-mediated degradation. In the presence of the ligand Wnt, the APC complex is inhibited, and the β-catenin protein accumulates, translocates to the nucleus, and stimulates β-catenin–dependent gene expression. Mutations that inactivate the APC complex have the same effect, except that the pathway is now active in the absence of the Wnt ligand.

Slide 7: Results of a siRNA-based screen of Wnt/β-catenin signaling in colorectal carcinoma

The workflow and results of our siRNA screen are depicted. For the primary and secondary screens, we used colorectal carcinoma cell lines stably transduced with a β-catenin–responsive luciferase-based transcriptional reporter. First, we screened pools of siRNAs, each pool targeting 1 of approximately 20,000 transcripts. 740 siRNA pools passed our criteria in the primary screen for a positive identification, a "hit." Given that not all siRNAs targeting a specific transcript will result in protein knockdown, screening with pooled siRNAs increases the chance of achieving knockdown without increasing the size of the screen. However, screening with pooled

Howard Hughes Medical Institute, Department of Pharmacology, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA 98109, USA.

***Corresponding author. E-mail, jdberndt@u.washington.edu**

siRNA increases the number of off-target effects, leading to more false-positive screen hits. To eliminate these false positives, the pools representing hits were deconvoluted and individual siRNAs were rescreened. In addition, to eliminate cell type–specific hits, deconvoluted pools were also screened in a second cell line. In some cases, additional individual siRNAs targeting selected transcripts were screened. 119 genes passed the secondary screen in two cell lines.

Slide 8: What is the best approach to data reduction?

After the secondary screen, we were satisfied that we had a rigorously validated set of siRNAs that targeted mRNAs encoding proteins that regulated the β-catenin–responsive reporter. The remaining challenge is one that every investigator faces when dealing with genome-scale screens: Which of the hits should be pursued in depth? For the remainder of this presentation, we present methods of data reduction to generate testable models of signal transduction using the data and strategies of Major *et al*. (*2*) to illustrate the approaches.

Slide 9: Use endogenous target genes to validate the siRNA screen and to prioritize the hits

In order to validate the reporter-based screen hits and to provide a measure of specificity to the Wnt/β-catenin pathway, we performed a tertiary screen. Specifically, we quantitated the abundance of endogenous mRNAs that require β-catenin–mediated transcription for expression after the introduction of siRNAs passing the secondary screen. For the primary screen, we, like many investigators, used an artificial reporter as the screen readout to maximize the signal-to-noise ratio and thereby increase the sensitivity of the high-throughput screen. We used an optimized β-catenin–responsive luciferase reporter that has a very large dynamic range of responsiveness (*5*). Because of the artificial nature of a transcriptional reporter, screens based on these assays include potential false-positive identifications. In the tertiary screen, quantifying the transcribed mRNAs of endogenous β-catenin target genes allowed us to identify β-catenin–regulated genes that are also regulated by a given siRNA screen hit. Simultaneously, by examining the expression of genes that are not regulated by Wnt/β-catenin signaling, such as housekeeping genes, we demonstrated the specificity of the siRNA screen hit to the Wnt/β-catenin signal transduction pathway. In general, when running large-scale screens, our experience supports the inclusion of a follow-up validation screen, one based on an endogenous, rather than an engineered, metric of signaling activity.

Slide 10: Defining the endogenous β-catenin target genes in human colorectal cancer

In order to validate hits by examining their activity on endogenous target genes, one first needs to know the endogenous target genes that are relevant to the system. Because we were interested in the Wnt/β-catenin pathway in colorectal cancer, we needed to define the genes that are regulated by β-catenin (*CTNNB1* in the slide) in colorectal carcinoma. Because DLD1 and SW480 cell lines, which are human colorectal cancer cell lines, contain mutant forms of APC, the Wnt pathway is constitutively active. Thus, we used siRNA-mediated silencing of β-catenin and cDNA microarray expression profiling to define the β-catenin gene signature. Of the 43,675 transcripts surveyed, we found 329 that were regulated by all five β-catenin siRNAs tested in DLD1 cells. Of these, 40 genes were also regulated in SW480 cells. We used a time-course microarray analysis after β-catenin siRNA transfection to further narrow this list to genes that are directly regulated by β-catenin. Next, we performed microarray analysis on cancerous and adjacent normal tissues from patients with colon adenocarcinomas. Using unsupervised hierarchical clustering, we found that the β-catenin gene signature successfully segregated disease from nondisease tissue in 96% of cases.

In summary, we used four strategies to define the colorectal β-catenin gene signature: (i) We eliminated off-target effects by assaying with multiple siRNAs against each gene; (ii) we eliminated cell type differences by assaying the siRNAs in at least two different cell lines; (iii) we identified probable direct targets by performing a time-course analysis with the siRNAs; and (iv) we validated our β-catenin gene signature in vivo by testing its ability to distinguish normal colon tissue from colon cancer tissue.

Slide 11: Do siRNA screen hits regulate endogenous β-catenin target genes?

Once the list of endogenous target genes is established, the next step is to validate the hits from the siRNA secondary screen, using the expression of these genes as the endogenous readout. In our study, we used multiplex quantitative reverse transcription polymerase chain reaction [Fluidigm platform (*6*)] and microarray expression analyses to quantify the expression of β-catenin target genes after siRNA transfection. Of the 78 secondary screen hits that were profiled, we found 38 genes encoding proteins that regulated at least five endogenous β-catenin target genes. Thus, by validating our reporter-based siRNA screen hits on a set of endogenous β-catenin target genes, we generated a list of high-confidence previously unknown regulators of the Wnt/β-catenin pathway in colorectal cancer cell lines.

Slide 12: Results of our siRNA screen validation strategy

By filtering the results of the primary siRNA screen through two additional validation screens, the number of potential positive hits can be reduced from the tens of thousands to a more reasonable number that are much more likely to be real hits and not false positives. For our system, of the initial ~28,000 siRNAs screened, we identified 38 triply validated regulators of β-catenin–dependent transcription. We have now repeated the primary and secondary screens in two additional cell lines that are not derived from colon tissue, HEK293T and HT1080 (unpublished data). By comparing the hits of the colorectal siRNA screen and these new screens, we hope to determine cell type–specific requirements for Wnt/β-catenin signal transduction. Also, by comparing microarray data after β-catenin knockdown among these cell lines, we have defined both common and cell type–specific β-catenin target genes (unpublished data).

Thus, we have a high-confidence list of genes whose loss of function affects Wnt/β-catenin signaling. From here, one can imagine several ways to identify the genes that are best for follow-up analyses.

Slide 13: Sources of data for integrative analysis

There are at least eight ways to limit the candidate gene set to select the ones for further analysis, in addition to the simplest approach of selecting the siRNA with the most significant deviation from the mean. However, because of the idiosyncrasies of siRNA-mediated knockdown (for example, differences due to the relative stability of proteins), this approach may not lead to the most interesting proteins. Assigning quasi-arbitrary significance thresholds to define a siRNA as a hit can result in false-negative and false-positive identifications. Moreover, a list of *z* scores gives no indica-

tion of the mechanism by which the products of these genes affect the cellular process of interest (here, the Wnt/β-catenin pathway); and finally, the magnitude of impact of a siRNA on the pathway does not provide any information about the importance of that protein for the disease etiology or about the therapeutic potential of manipulating that protein.

To address these limitations, one can further annotate the hit list from a functional genetic screen using various different types of data. Much of these data are available in publically accessible databases or can be acquired in parallel with the genetic screen. However, gene annotation was traditionally done on a low-throughput gene-by-gene basis using literature and database searches and follow-up experiments. More recently, semiautomated, nonbiased approaches to annotate entire gene lists have proven not only useful but also powerful in identifying relevant hits for follow-up assays. For the remainder of this presentation, we focus on three types of data integration: First, integration of siRNA data and protein interaction networks (PINs); second, integration of siRNA data and protein targets identified by small-molecule screens; and third, integration of siRNA data and disease-risk alleles identified in genome-wide association studies.

Slide 14: Use proteomics to validate siRNA screens and find functional protein complexes

For our analysis of the Wnt/β-catenin pathway in colorectal cancer cells, we chose to perform a proteomic analysis of siRNA hits in order to address some of the limitations of siRNA screening alone. This integrated approach allowed us to test the predictions that siRNA screen hits encode known components of the Wnt/β-catenin pathway and their first-order interactors. Further, because proteins do not function in isolation, we could identify physical interactions between siRNA screen hits. Because the stringent three-screen validation methods that we used may have resulted in the classification of some of the siRNA hits as false negatives, this approach of integrating the siRNA screen hits with protein interaction data can also be used to rescue some of these false negatives.

Slide 15: Integrating proteomic and siRNA screens

Various tools are available for integrating expression data or RNAi screen data with protein interaction data. We used Cytoscape (*7*) software, and here we demonstrate a seven-step approach for generating PINs.

- Step 1 is defining a core set of proteins of interest.
- Step 2 is selecting one of these proteins to use as "bait" from which to build the network.
- Step 3 is finding all of the published interactions with the bait protein and constructing the base network of proteins (nodes) and their relations (edges; white lines).
- Step 4 is performing a protein interaction screen and adding these new interacting proteins and relations (blue lines) to the PIN.
- Step 5 is overlaying the siRNA screen data onto the PIN.
- Step 6 is creating additional PINs for other bait proteins.
- Step 7 is combining all of the PINs with the hits from the siRNA screen identified into a single extended PIN.

To put this into the context of the Wnt/β-catenin screen, we first identified a core set of proteins in the Wnt/β-catenin pathway and then selected 20 of these as baits for the construction of a PIN. We curated all known protein interactions with the bait proteins from the STRING database (*8*). To identify additional, previously unknown interactions (blue edges), we performed affinity purification tandem mass spectrometry (AP-MS/MS) with a subset of the bait proteins. Finally, we colored all the nodes according to their activity in the secondary siRNA screen, with gray representing pathway components that were not identified in the siRNA screen and red and green representing the hits that either inhibited Wnt/β-catenin–dependent gene expression or were required for Wnt/β-catenin–dependent gene expression, respectively. After performing this analysis for multiple bait proteins, we combined the PINs to create an extended PIN of Wnt signaling.

Although the extended Wnt PIN allows us to map the most highly validated hits from our screen, we may have lost valuable data by not including some of the medium-confidence siRNA hits. Thus, we also created a PIN from AP-MS/MS and bioinformatics data (not shown) that included the proteins encoded by a subset of the medium-confidence siRNA hits (those that passed the secondary screen in the DLD1 cell line) We then integrated this PIN with the extended Wnt PIN.

Slide 16: Expectations and results of data integration

Integrating siRNA screen data with protein interaction data permits statistical testing of at least two predictions and provides mechanistic insight into the potential functions of the hits.

Perhaps the most obvious prediction that this integration tests is that the siRNA screen hits encode proteins that physically associate with the Wnt pathway. Specifically, we found a statistically significant enrichment of the siRNA screen hits in a PIN comprising the core Wnt pathway and its first-order interacting proteins. By comparison, we did not observe our siRNA screen hits within a PIN for 300 random proteins and their first-order interactors. The integrative analysis also allows the prediction that proteins encoded by siRNA screen hits form physical complexes, as one would expect given that proteins do not act in isolation. Again, using 300 randomly selected proteins as a comparative baseline, we found a statistically significant enrichment of protein-protein interactions within our siRNA screen hits.

Finally, for many of the siRNA screen hits, we identified interacting proteins of known biochemical function. This insight then instructs hypothesis generation as to the mechanism by which the siRNA hit controls Wnt signal transduction. For example, we further characterized AGGF1, which had been previously reported as a secreted angiogenic factor (*9*). With the integrated screens, we found that AGGF1 is required for Wnt/β-catenin signaling and that it associates with a member of the SWI/SNF chromatin remodeling complex. AGGF1 localized to the nucleus, where it associated with the promoters of β-catenin target genes.

Slide 17: Caveats of proteomic-siRNA integrative screening

Although this is a powerful approach for validating RNAi screen hits and for predicting functionality of the screen hits, it does have limits. Here we list several of the current limitations. First, many of the siRNA hits will not map onto a physical map of the pathway because of the indirect nature by which they regulate signal transduction. Second, because there are false-positive and false-negative discovery rates associated with both screening technologies, it remains a certainty that integration

of the data sets will not eliminate all false leads. Finally, a lack of comprehensiveness in the databases and overrepresented proteins in biochemical assays may bias the results.

Slide 18: Sources of data for integrative analysis

Although integrating siRNA and proteomic screens provides a foundation for understanding the mechanistic roles of several proteins in a particular pathway, it does not help to address other limitations of siRNA screening, such as the contribution of hits to disease etiology or potential therapies associated with the inhibition of these genes. These questions may be better addressed by other types of integrative screens—for example, by integrating the siRNA screen data with data from small-molecule screens or with data from disease-associated genetic polymorphism screens. For the Wnt/β-catenin pathway, we are beginning to perform these other types of integrative screens.

Slide 19: Use drug screens to validate siRNA screens and find potential therapeutics

Small-molecule screens have greatly expanded the repertoire of chemical tools for the investigation of signaling pathways and the development of drugs for use in patients. Small-molecule and siRNA screens are complementary approaches in that each is founded on chemically distinct approaches of modulating protein activity. Thus, the integration of small-molecule and siRNA screens can eliminate some false-negative and false-positive identifications resulting from either technique alone. Of equal importance, drug-protein relationships that are important for the phenotype of interest (for example, Wnt signaling) might be revealed by such integration. As such, this approach provides mechanistic insight into novel chemical tools and lead compounds for therapeutic intervention.

Slide 20: Integrating drug and siRNA screens

Although the targets of most small molecules are unknown, there is an increasing knowledge of drug-protein interactions available in public databases and in the literature. Very simply, there are three steps to integrating drug-protein interaction data with siRNA screen data. The first step is to identify the relevant bioactive compounds. This can be achieved in several ways, depending on the biological question at hand. We, for example, performed a small-molecule screen using the β-catenin–responsive luciferase reporter that we previously used for the siRNA screens. The second step is to identify all of the drug-protein interactions for the selected bioactive compounds and create a drug-protein network. Finally, the hits from the siRNA screen are merged onto the network and color-coded on the basis of their activity in the assay. In this example, we curated all known protein-drug interactions from the STITCH (*10*) database (where an interaction can be either physical or functional) for 11 compounds identified in our small-molecule screen on the β-catenin–responsive transcriptional reporter. We then colored these proteins according to whether they were required for β-catenin–dependent gene expression or inhibited β-catenin–dependent gene expression in the secondary siRNA screen. This limited integrated analysis revealed two common nodes: FOS (a transcription factor) and CDK5 (cyclin-dependent kinase 5). Both proteins were previously reported to bind β-catenin (*11*, *12*). Though this integration is narrow in scope, it serves as a proof of concept that the nonbiased integration of siRNA screens and small molecule-screens can yield mechanistic insight into potentially therapeutic regulators of the Wnt pathway.

Slide 21: Caveats of drug-siRNA integrative screening

Drug-siRNA integration suffers from many of the same limitations as siRNA-proteomics integration. For instance, the previous example was potentially hindered by the fact that the screens were conducted in cell lines of different tissue origin. In addition, the specific limitations of small-molecule and siRNA screens may result in nonoverlapping sets of false-negative identifications.

Slide 22: Integration of disease-associated mutation maps with siRNA data

A third and complementary approach to integrating different types of screens is the use of data from genome-wide association studies. The causative loci linked with the majority of disease-risk haplotypes have yet to be mapped. However, because of increasing numbers of genome-wide association studies, advancing technologies, such as deep sequencing and comparative genomic hybridization on microarrays, high-density single-nucleotide polymorphism maps, and advanced computational methods, there is unprecedented spatial resolution in linkage maps and sensitivity in the allelic detection. Thus, causative genetic polymorphisms and copy number changes are being identified for an increasing number of diseases. This information identifies the relative contributions of alleles to disease severity, penetrance, and incidence for the population but has limited utility in the identification of molecular mechanisms and therapies. Moreover, historically, follow-up studies on candidate genes and signaling pathways were performed on a gene-by-gene basis resulting in a long lag time between identification and therapy. The integration of these data with RNAi screen data provides phenotypic annotation of the disease-associated genotype, which may help to guide the selection of proteins for follow-up assays.

Slide 23: Integrated genotype and RNAi screening

Increasing numbers of genome-wide sequencing and mapping studies provide public data on candidate mutations associated with disease. In this example, we used a set of probable "driver mutations" (that is, mutations that cause the cancer) identified in a genome-wide screen of colorectal cancers (*12*). We asked whether any of the mutations identified in this screen were also proteins identified in our siRNA screen of the Wnt/β-catenin pathway in colorectal carcinoma cells. A similar study was recently published (*13*). For illustration purposes, we created a PIN for 320 siRNA hits that passed our secondary screen in DLD1 cells. Many of these did not interact with any other protein (orphan nodes), several were in small networks of two to four proteins, and a few were part of a larger network. We then colored those hits if they were also probable driver mutations in colorectal cancer (blue). Mutational inactivation of the APC protein activates Wnt/β-catenin signaling in ~80% of colorectal cancers (*4*). As expected, we identified APC in the integration of these screens. We also identified BCL9, another known member of the Wnt pathway that is required for β-catenin–dependent transcription (*14*). In addition, we found two other proteins common to both screens. One of these is an orphan node; another has physical interactions with other siRNA hits in the Wnt/β-catenin pathway. The facts that these proteins are both modulators of Wnt/β-catenin signaling and are mutant in colorectal cancer suggest that their contribution to the progression of colorectal disease may depend on their role in the Wnt/β-catenin pathway.

Slide 24: Colorectal cancer genetic polymorphisms in the Wnt/β-catenin pathway

In an approach complementary to that described in the previous slide, one could integrate genotypic data with a PIN for the pathway of interest irrespective of whether the proteins in that pathway were hits in a siRNA screen. For example, we used 20 core Wnt/β-catenin pathway members as baits to develop a PIN from the STRING database and selected primary literature references (*2*). We then asked whether any of the proteins in or associated with the Wnt/β-catenin pathway were also identified in the genetic polymorphism screen in colorectal cancer (*12*). In addition to APC and BCL9, we identified five additional proteins: TCF7L2 (a member of the TCF family of transcription factors); TP53 (the tumor suppressor p53); and SMADs 2, 3, and 4 [signaling mediators of transforming growth factor–β (TGF-β) superfamily members], using this method. TCF7L2 is a transcription factor required for Wnt signaling in some contexts and was a probable false negative in our siRNA screen (*15*). The identification of TP53 and SMADs suggests that pathway crosstalk between the Wnt and p53 pathways, Wnt and TGF-β pathways, or both pathways may be important for colorectal cancer etiology.

Slide 25: Caveats of genotype-RNAi integrative screening

As with the other integration techniques, integrating genotypic data with data generated in a reverse genetic screen has limitations.

Slide 26: Conclusions

The practice of combining two data sets created from diverse methodologies allows one to hone large intractable data sets into easily testable models. In the examples we presented here, we integrated siRNA screen hits that met arbitrarily determined statistical criteria. However, every siRNA tested had an associated score that could be represented on a continuum. Similarly, candidate small molecules and their targets are given similar scores, and polymorphic alleles have scores representing degrees of association with a particular disease (risk). By quantitatively combining these scores and determining a probability of co-occurrence, one could reduce the number of false negatives while simultaneously providing a confidence measure for positive identifications. Moreover, quantitative analysis allows the integration of multiple forms of data in the creation of models, an approach that has recently been successfully applied to predict functional networks in yeast and mice (*16*, *17*). As technologies improve and costs fall, we will be increasingly exposed to large data sets describing complex biological systems. Whether these data translate to knowledge and new therapeutic avenues depends on our ability to comprehend and distill the growing wealth of data.

Slide 27: People who did the work

Editor's Note: This contribution is not intended to be equivalent to an original research paper. Note, in particular, that the text and associated slides have not been peer-reviewed.

References

1. The number of publications illustrated was determined using the following PubMed query: ("HIGH-THROUGHPUT"[Title/Abstract] OR "GENOME-WIDE"[Title/Abstract] OR "GENOMEWIDE"[Title/Abstract] OR "LARGE-SCALE"[Title/Abstract]) AND ("SCREEN*"[Title/Abstract]).
2. M. B. Major, B. S. Roberts, J. D. Berndt, S. Marine, J. Anastas, N. Chung, M. Ferrer, X. Yi, C. L. Stoick-Cooper, P. D. von Haller, L. Kategaya, A. Chien, S. Angers, M. MacCoss, M. A. Cleary, W. T. Arthur, R. T. Moon, New regulators of Wnt/?-catenin signaling revealed by integrative molecular screening. *Sci. Signal.* **1**, ra12 (2008).
3. http://seer.cancer.gov/statfacts/html/colorect.html.
4. R. T. Moon, A. D. Kohn, G. V. De Ferrari, A. Kaykas, WNT and β-catenin signalling: Diseasesand therapies. *Nat. Rev. Genet.* **5**, 691–701 (2004).
5. T. L. Biechele, R. T. Moon, Assaying β-catenin/TCF transcription with β-catenin/TCF transcription-based reporter constructs. *Methods Mol. Biol.* **468**, 99–110 (2009).
6. S. L. Spurgeon, R. C. Jones, R. Ramakrishnan, High-throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS ONE* **3**, e1662 (2008).
7. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
8. B. Snel, G. Lehmann, P. Bork, M. A. Huynen, STRING: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **28**, 3442–3444 (2000).
9. X. L. Tian, R. Kadaba, S. A. You, M. Liu, A. A. Timur, L. Yang, Q. Chen, P. Szafranski, S. Rao, L. Wu, D. E. Housman, P. E. DiCorleto, D. J. Driscoll, J. Borrow, Q. Wang, Identification of an angiogenic factor that when mutated causes susceptibility to Klippel-Trenaunay syndrome. *Nature* **427**, 640–645 (2004).
10. M. Kuhn, C. von Mering, M. Campillos, L. J. Jensen, P. Bork, STITCH: Interaction networks of chemicals and proteins. *Nucleic Acids Res.* **36**, D684–D688 (2007).
11. K. Toualbi, M. C. Guller, J. L. Mauriz, C. Labalette, M. A. Buendia, A. Mauviel, D. Bernuau, Physical and functional cooperation between AP-1 and β-catenin for the regulation of TCF-dependent genes. *Oncogene* **26**, 3492–3502 (2007).
12. L. D. Wood, D. W. Parsons, S. Jones, J. Lin, T. Sjoblom, R. J. Leary, D. Shen, S. M. Boca, T. Barber, J. Ptak, N. Silliman, S. Szabo, Z. Dezso, 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. Willson, S. Sukumar, K. Polyak, B. H. Park, C. L. Pethiyagoda, P. V. 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**, 1108–1113 (2007).
13. R. Firestein, A. J. Bass, S. Y. Kim, I. F. Dunn, S. J. Silver, I. Guney, E. Freed, A. H. Ligon, N. Vena, S. Ogino, M. G. Chheda, P. Tamayo, S. Finn, Y. Shrestha, J. S. Boehm, S. Jain, E. Bojarski, C. Mermel, J. Barretina, J. A. Chan, J. Baselga, J. Tabernero, D. E. Root, C. S. Fuchs, M. Loda, R. A. Shivdasani, M. Meyerson, W. C. Hahn, CDK8 is a colorectal cancer oncogene that regulates β-catenin activity. *Nature* **455**, 547–551 (2008).
14. T. Kramps, O. Peter, E. Brunner, D. Nellen, B. Froesch, S. Chatterjee, M. Murone, S. Zullig, K. Basler, Wnt/Wingless signaling requires BCL9/Legless-mediated recruitment of pygopus to the nuclear β-catenin-TCF complex. *Cell* **109**, 47–60 (2002).
15. V. Korinek, N. Barker, P. J. Morin, D. van Wichen, R. de Weger, K. W. Kinzler, B. Vogelstein, H. Clevers, Constitutive transcriptional activation by a β-catenin-Tcf complex in $APC^{-/-}$ colon carcinoma. *Science* **275**, 1784–1787 (1997).
16. Y. Guan, C. L. Myers, R. Lu, I. R. Lemischka, C. J. Bult, O. G. Troyanskaya, A genomewide functional network for the laboratory mouse. *PLOS Comput. Biol.* **4**, e1000165 (2008).
17. J. Zhu, B. Zhang, E. N. Smith, B. Drees, R. B. Brem, L. Kruglyak, R. E. Bumgarner, E. E. Schadt, Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat. Genet.* **40**, 854–861 (2008).

10.1126/scisignal.270pt4

Citation: J. D. Berndt, T. L. Biechele, R. T. Moon, M. B. Major, Integrative analysis of genome-wide RNA interference screens. *Sci. Signal.* **2**, pt4 (2009).