Phosphotyrosine Signaling Proteins that Drive Oncogenesis Tend to be Highly Interconnected

Grigoriy Koytiger‡§, Alexis Kaushansky§¶, Andrew Gordus§||, John Rush**,
Peter K. Sorger‡ ‡‡, and Gavin MacBeath‡‡‡

Mutation and overexpression of receptor tyrosine kinases or the proteins they regulate serve as oncogenic drivers in diverse cancers. To better understand receptor tyrosine kinase signaling and its link to oncogenesis, we used protein microarrays to systematically and quantitatively measure interactions between virtually every SH2 or PTB domain encoded in the human genome and all known sites of tyrosine phosphorylation on 40 receptor tyrosine kinases and on most of the SH2 and PTB domain-containing adaptor proteins. We found that adaptor proteins, like RTKs, have many high affinity bindings sites for other adaptor proteins. In addition, proteins that drive cancer, including both receptors and adaptor proteins, tend to be much more highly interconnected via networks of SH2 and PTB domain-mediated interactions than nononcogenic proteins. Our results suggest that network topological properties such as connectivity can be used to prioritize new drug targets in this well-studied family of signaling proteins. Molecular & Cellular Proteomics 12: 10.1074/mcp.M112.025858, 1204–1213, 2013.

Receptor Tyrosine Kinase (RTK) signaling networks evolved in Metazoans to process extracellular cues and elicit cellular responses such as differentiation, proliferation or migration. Canonical RTK signaling is initiated when a ligand binds to the extracellular domain of its cognate receptor, inducing receptor dimerization, activation of the intracellular tyrosine kinase domain, and auto-phosphorylation of intracellular tyrosine residues (1). These phosphotyrosine (pTyr) residues then serve as recruitment sites for one or more of the ∼120 SH2 domains (2) and 44 PTB domains (3) encoded in the human genome. Upon recruitment, many adaptor proteins bearing SH2 and PTB domains themselves become phosphorylated on tyrosine residues by active receptor or cytosolic tyrosine kinases. This second set of phosphorylation events leads to the recruitment of additional SH2 and PTB domain-containing proteins (1). For example, the binding of GRB2 to pTyr427 of SHC1 (or pTyr317 in the p52 SHC1 isoform) induces activation of the RAS/MAPK kinase cascade (4). Similarly, binding of phosphatidylinositol 3-kinase (PI3KR1, p85) to pTyr612 of IRS1 initiates signaling through the PI3K/ AKT cascade in cells exposed to insulin or insulin-like growth factors (5).

RTKs share a similar set of interactors and activate many of the same response pathways, including RAS/MAPK and PI3K/AKT cascades (1). They do not, however, necessarily elicit the same phenotypes. For example, both EGFR and NTRK1 induce MAPK signaling in PC12 cells, but EGFR triggers proliferation whereas NTRK1 promotes differentiation. Both of these phenotypes are dependent on ERK activity (6). This phenomenon is also observed clinically, where only a subset of RTKs have been shown to drive cancer despite sharing many downstream pathways (7). Current qualitative representations of signaling networks as linear cascades are inadequate to explain the diverse phenotypes that arise downstream of different RTKs.

Aberrant signal processing by RTK networks has been causally linked to cancer development, maintenance, and progression in many human tissues. Well-studied examples include overexpression of ERBB2 in breast cancer (8), KIT in testicular germ cell tumors (9), and MET in gastric cancer (10). Constitutive activating mutations of RTKs, such as those observed in the RET kinase (11, 12) in multiple endocrine neoplasia type 2 or KIT in gastrointestinal stromal tumors (13) are also capable of driving oncogenesis. Similarly, mutation or overexpression of SH2 domain-containing cytosolic proteins such as ABL kinase (14) or the PI3K p85 regulatory subunit (15) can also drive cancer, in these cases by inducing constitutive enzymatic activity that is decoupled from upstream signaling events.

Recently, Barabási and colleagues advanced a mathematical argument that network driver nodes, the nodes that control information flow in a network, should not be highly interconnected (16). If this notion is extended to cancer, in which
signaling networks are substantially altered or rewired, we would expect that proteins driving oncogenesis would not be highly interconnected. We sought to determine experimentally if there is indeed a link between network connectivity and the propensity of a protein to drive cancer. In making this determination, we cannot rely solely on literature-derived interaction networks (17, 18) as they are confounded by study bias (19, 20). Specifically, oncogenic proteins are more intensively studied than nononcogenic proteins, potentially resulting in a bias in terms of number of binding partners. As a means to collect systematic pTyr-mediated interaction data, in vivo methods like the yeast two-hybrid system are not suitable as they do not allow for control over post-translational modification events (21). Systematic co-immunoprecipitation coupled with mass spectrometry is also problematic, as many interactions mediated by tyrosine phosphorylation are transient, with half lives on the order of seconds, and any particular cell type expresses only a subset of the proteome (22). Previous systematic research on the binding specificity of SH2/PTB domains has therefore been performed using phosphorylated peptides and in vitro binding assays. For example, pioneering studies of this type involving peptide libraries uncovered consensus binding motifs for a variety of SH2 and PTB domains (23–25). It is now apparent, however, that these motifs are simplified views of in vivo selectivity and that specificity is also defined by “anti-motifs” representing excluded contacts for SH2 domains (26), SH3 domains (27), and probably other modular interaction domains as well. As a consequence, it is very difficult to predict accurately whether an SH2 or PTB domain will bind to a particular sequence known to be phosphorylated in vivo.

To overcome these limitations, it is necessary to test interactions between binding domains and peptides bearing physiological sequences one at a time and in a noncompetitive setting. To perform such assays in an unbiased, high-throughput, and quantitative manner, we used protein domain microarrays (28). In brief, 134 purified recombinant SH2 and PTB domains were printed as microarrays in individual wells of 96-well microtiter plates. The arrays were then probed with fluorescently labeled phosphopeptides derived from known sites of tyrosine phosphorylation on human proteins. By probing the arrays with eight different concentrations of each peptide, full saturation binding curves were obtained, providing an estimate of the equilibrium dissociation constant for each biochemical interaction. Our previous work using this assay focused on nine RTKs: FGFR1, IGF1R, MET, NTRK2, PDGFR β (29) and the four ErbB receptors (3, 30). In these studies, we found that the arrays correctly identify most previously reported interactions (3). In addition, they invariably highlight new interactions, many of which we validated biologically: they occur in nonengineered cells and play important roles in signal transduction (3, 30), (31–33). We now use this approach to systematically quantify, on a nearly proteome-wide level, interactions between SH2 or PTB domains and known sites of tyrosine phosphorylation on both human RTKs and the adaptor proteins themselves. We find a very high degree of connectivity that challenges conventional, linear views of receptor-adaptor interaction. Moreover, RTKs and adaptor proteins that have been shown to play a causal role in cancer tend to mediate substantially more interactions than those that do not. This suggests that these connectivity profiles may provide insight into how networks are rewired and may help prioritize new targets for anti-cancer drug discovery.

EXPERIMENTAL PROCEDURES

A detailed protocol for preparing protein domain microarrays has been described previously (28).

Determining Interaction Affinity—Following peptide binding, arrays were scanned at multiple PMT voltages on a Tecan LS400 microarray scanner on both Cy5 and Cy3 channels. Spots with saturated pixels were eliminated. The remaining spots were fit to a line that allows for the conversion of all the Cy3 values measured at different PMT voltages to the same scale. The fold-over-background value of a titration was determined by taking the trimmed mean of the Cy3 values for each domain-peptide titration divided by the Cy3 values of the Thioredoxin control spots. The mean Cy3 value of the Thioredoxin control spots was then subtracted from the Cy3 value of the domain spot and the spots were normalized by the Cy5 value.

Each domain was printed in quadruplicate in each well and the arrays were probed using eight different concentrations of each peptide, ranging from 5 μM down to 10 nM. These 32 data points were then fit to the following equation using MATLAB®s robust fit function with bisquare weights:

\[ F_{\text{obs}} = F_{\text{max}}[\text{Peptide}]/(K_D + [\text{Peptide}]) \quad \text{(Eq. 1)} \]

where \( F_{\text{obs}} \) is the mean fluorescence of replicate spots, \( F_{\text{max}} \) is the fluorescence at saturation, \([\text{Peptide}]\) is the total concentration of phosphopeptide, and \( K_D \) is the equilibrium dissociation constant. Robust fitting procedures are more resistant to the presence of outlier data resulting from missed spots, fluorescent debris or other aberrations (34). Only titrations with fold-over-background values in the top 10% of the data and \( R^2 \) values over 0.9 were kept. Replicate affinities were then averaged in log space. The MATLAB® code used to perform the analysis is supplied in supplementary information.

Determining Oncogene Status—Oncogene status for RTKs was determined using the Sanger Institute Cancer Gene Census (7, 35). The Sanger list contains only those genes for which a strong causal link to cancer has been established. Because many SH2/PTB domain-containing proteins have no catalytic activity of their own, they serve principally to amplify and propagate signals from upstream proteins. For this reason, adaptors were assigned as having a role in cancer by their presence in a union of the Sanger list and other cancer gene lists (36–38) compiled by the Bushman lab (http://micrобр230.med.upenn.edu/links/genelist).

Statistical Tests for Enrichment of Binding—To test against the null hypothesis that the distribution of connectivity is equivalent for both Oncogene and non-Oncogene classes, we performed the nonparametric Mann-Whitney U test in MATLAB® R2011a (The MathWorks, Inc., Natick, MA). Only those interactions with dissociation constants below 1 μM were considered.

RTK and Adaptor Expression in Tumor Samples—cBio Cancer Genomics Data Server Matlab Toolbox v1.04 was used to retrieve
Oncogenes Are Highly Interconnected

mRNA sequencing data corresponding to 1437 tumors from breast cancer, colon, and rectal adenocarcinoma (39) and clear cell kidney cancer collected by The Cancer Genome Atlas Network. We used a 1 read per kilobase of exon model per million mapped reads (RPKM) threshold to separate genes which are expressed. This threshold has been shown to correspond to 1 transcript per C2C12 cell (40) and optimizes overlap with mass spectrometry results in HeLa cells (41). We consider this to be an upper bound estimate to the total number of potential interaction partners.

RESULTS AND DISCUSSION

Data Collection—We started by compiling a list of known sites of tyrosine phosphorylation on human RTKs and on all SH2 and PTB domain-containing proteins listed in the PhosphoSitePlus database (42). We restricted our studies to experimentally verified sites of tyrosine phosphorylation as nonphysiologically relevant sites, when artificially phosphorylated, may also bind SH2/PTB domains (43) and thereby confound any systems level conclusions drawn from the resulting data. We focused on 40 of the 53 Uniprot-annotated RTKs that had more than three phosphorylation sites to include only those receptors whose biology is sufficiently understood to enable systematic comparison. This resulted in a total of 729 unique phosphopeptides that we successfully synthesized, fluorescently labeled, and purified using high performance liquid chromatography (supplemental Table S1). We then used each of the peptides to probe microarrays comprising virtually every human SH2 and PTB domain (supplemental Table S2). The arrays were probed at eight concentrations of each peptide, ranging from 10 nM to 5 μM (see, for example, Fig. 1A). By repeating this process for each phosphopeptide we were able to generate a quantitative interaction map for the receptor (Fig. 1B). This process was repeated for all of the remaining RTKs to generate, for the first time, a global, systematic, and unbiased view of RTK recruitment (Fig. 1C).

Larger versions of the connectivity diagrams for each RTK are available in supplemental Fig. S1. Only interactions with $K_D < 1 \mu M$ are included in these diagrams. SH2 domain-mediated interactions of lower affinity have been shown to be physiologically relevant intramolecularly, but it is generally accepted that most bona fide intermolecular interactions exhibit dissociation constants below 1 μM (44–46).

As with any assay, one invariably encounters false positives and false negatives. With protein domain microarrays, the most frequent source of false positives is nonspecific binding between peptide and domain, whereas false negatives likely arise from low surface activity of the domains, presumably because some domains denature on the slide surface or are preferentially immobilized in a way that blocks access to the binding site. To estimate the stochastic false positive rate, we conservatively assumed that domains having the lowest frequency of binding were actually inactive in our assay and exhibited no true-positive interactions (any interactions identified for these domains are by definition false positives). We observed that domains in the bottom 20% of the binding frequency spectrum accounted for 20 interactions out of the 37,490 titrations performed on the domains, corresponding to a stochastic false positive rate of $5.33 \times 10^{-4}$. If we extend this rate to all the titrations performed (195,546), we estimate that 104 of the 2808 interactions we identified are liable to be stochastic false positives (false discovery rate of 3.72%). As there is no reason to believe that false positive errors are observed preferentially with phosphopeptides derived from cancer-causing proteins, the primary conclusions of this study are not affected.

Another potential limitation of our approach is that the current list of physiological sites of tyrosine phosphorylation in PhosphoSitePlus may be incomplete. This would be particularly problematic if nononcogenic RTKs were less well annotated than oncogenic ones. To investigate this possibility, we used phosphotyrosine-directed mass spectrometry (30) to study six nononcogenic members of the Ephrin class of RTKs - EphA2, EphA3, EphA4, EphB2, EphB3, and EphB4. We overexpressed each receptor in HEK293T cells, a procedure that induced receptor auto-activation and phosphorylation in all six cases. Receptors were then immunoprecipitated using an anti-pTyr antibody and subjected to targeted and untargeted μLC-MS/MS. Using this approach, we were able to identify 32 out of the 38 known sites of intracellular tyrosine phosphorylation (84% sensitivity). Remarkably, we did not identify any sites of tyrosine phosphorylation beyond those already reported in the PhosphoSitePlus database (supplemental Table S4). This suggests that the many high throughput, pTyr-directed mass spectrometric studies that have been used to populate PhosphoSitePlus are not biased against nononcogenic receptors and that the existing list of tyrosine phosphorylation events, at least on RTKs, is nearly complete.

Oncogenic RTKs Are Highly Connected—By examining the connectivity profile of the 40 RTKs at various affinity thresholds (Fig. 2A), we sought to identify whether a link exists between connectivity and oncogenicity. We determined whether or not an RTK is an oncogene based on its inclusion in the Sanger Institute’s Cancer Gene Census (7) (http://www.sanger.ac.uk/genetics/CGP/Census/), which seeks to determine a strict causal (and not merely correlative) relationship between cancer development and mutation and/or overexpression of a given gene. Based on these assignments, we found that oncogenic RTKs have a significantly higher median connectivity in our interaction data set than nononcogenic RTKs. At an affinity threshold of 1 μM, for example, the median number of binding partners is 21 for nononcogenic RTKs and 56 for oncogenic receptors, corresponding to a ~2.5-fold difference in the number of interactions (Mann-Whitney $U$ test $p = 2.77 \times 10^{-5}$). Nononcogenic RTKs also have a median of six phosphorylation sites, whereas oncogenic RTKs have nine phosphorylation sites (Mann-Whitney $U$ test $p = 0.007$), corresponding to a 50% increase. This indicates that the primary reason for the increased connectivity of oncogenic RTKs is...
the presence of more promiscuous pTyr docking sites and only secondarily an increase in the number of docking sites. For instance, the nononcogenic receptor CSF1R, although it has twice as many phosphorylation sites as NTRK3 (8 versus 4), has far fewer high-affinity \(K_D/H^{11021}\) binding partners (6 versus 58). In addition, the RTK with the greatest number of phosphorylation sites in our study is the nononcogenic ERBB4 with 16. ERBB4 has only 18 binding partners, however, which places it in the bottom 25% of the RTKs we studied.

Many interactions annotated here may be of too low affinity to bind at appreciable levels to a receptor that is present at low surface density or at a low level of activation. When a cell becomes malignant, however, gene amplification or overexpression may make phosphorylated receptors (or adaptors) sufficiently abundant that low affinity interactions are enabled and downstream signaling activated. This notion is supported by the observation, in PC12 cells, that EGFR mediates proliferation when present at normal levels but differentiation when overexpressed (6). These phenomena cannot be captured in qualitative diagrams of signaling and highlight the need to think of these interactions as being contextually conditional and existing on a quantitative spectrum without a single fixed threshold.

**Fig. 1.** Determining binding affinity for SH2/PTB domains to phosphopeptides. A, Example titration of 5-(6)-TAMRA-labeled phosphopeptide corresponding to the sequence encompassing KIT-pTyr900. Titrations highlighted in green correspond to previously identified binding partners (63). Curves are colored according to affinity (see legend). B, Total connectivity for phosphorylation sites of the oncogenic RTK, KIT. Red circles corresponding to phosphorylation sites, blue circles to SH2 domains, and yellow circles to PTB domains. Lines are colored according to affinity. C, Visualization of the binding profile of the 40 RTKs analyzed in the study. RTKs are sorted in descending order by the total number of links with affinities lower than 1 \(\mu M\).
Adaptor-Adaptor Binding—Focusing only on the initial recruitment of adaptors to RTKs provides an incomplete view of the full complexity of early RTK signaling. Adaptors can themselves become tyrosine phosphorylated and interact with each other, forming multi-protein complexes at the receptor. By systematically probing our SH2/PTB domain microarrays with phosphopeptides derived from 334 known pTyr sites on adaptor proteins, we identified a dense network of interactions among the adaptor proteins themselves (Fig. 3A). For example, the well-studied tyrosine kinase ABL1, which is responsible for driving the development of chronic myeloid leukemia when fused with the protein BCR, is currently annotated to mediate 15 interactions through its SH2 domain (47). In our study, we identified 53 biochemical binding partners, including known interactions with DAB1, ERBB2, PTK2, and SHC1 (47). Our data may also resolve some outstanding questions in the field. For example, it has previously been reported that the Vav1 SH2 domain binds to a phosphorylation site on BCR/ABL and that this interaction is critical for activation of Rac-1 and BCR/ABL-mediated leukemogenesis.

Fig. 2. Oncogenic RTKs are enriched for binding partners. A, Cumulative density distribution of RTK interactions at different affinity thresholds. Oncogenic RTKs (7, 35) in red have significantly more binding partners than nononcogenic RTKs in gray; median of 56 versus 21 interactions with affinity less than 1 μM (Mann-Whitney U-Test p = 2.8 × 10⁻⁵). B, Example of connectivity between sites of tyrosine phosphorylation on the nononcogenic RTK CSF1R and SH2/PTB domains.

Molecular & Cellular Proteomics 12.5
Previous attempts to identify the site of binding by constructing a series of Tyr to Phe mutants, focusing in particular on the Vav1 consensus binding motif, were unsuccessful (48). Our data reveal an interaction between Vav1 and pTyr917 of BCR/ABL (Uniprot id: A9UF07), which was not included in the earlier mutagenesis experiments.

RTK-mediated activation of the MAPK pathway is well studied (49, 50), but we nevertheless find many new interactions that impinge on this pathway. RTKs vary in their ability to recruit early signaling proteins in this pathway, such as SHC1, GRB2, and PTPN11. Here, we find that these adaptors have the potential to interact with many other signaling proteins as

---

**Fig. 3.** Adaptors are densely interconnected. A, Adaptor proteins displayed in a circle with the size of the band corresponding to protein length (64). Lines begin at sites of tyrosine phosphorylation and terminate in either a yellow triangle corresponding to a PTB domain or a blue triangle corresponding to an SH2 domain and are color coded according to affinity. B, Visualization of the binding partners of ABL1 SH2 domain (in-links) and phosphorylation sites of ABL1 (out-links). C, RTK activation of the MAPK cascade is a varied process that integrates binding of GRB2 to SHC1 and SHP2 as well as to the RTK. D, The combinatorial complexity of RTK activation of MAPK is liable to be even greater as GRB2, SHC1 and PTPN11 (SHP2) are also capable of binding other adaptors, as can be seen by the diagrams, as well as other RTKs (not visualized).
well, increasing the complexity of MAPK pathway activation (Fig. 3D). This view aligns well with the idea that RTKs function not as discrete molecular machines, as typically depicted, but as pleomorphic ensembles that are highly complex and dynamic with many rapidly interchanging states (22, 51).

Oncogenic Adaptors are Highly Connected—To probe the relationship between connectivity and oncogenicity in the adaptor-adaptor layer of binding, we organized domains by sequence similarity (Fig. 4A) (52). The blue bars indicate how many peptides bind to each domain (in-links) and the green bars show how many binding partners interact with phosphorylation sites on the corresponding protein (out-links). Simple inspection reveals a wide range of binding promiscuity within the SH2 and PTB domain families. For example, proteins in the SRC, HCK, and PI3K N-terminal families of domains have relatively high numbers of binding partners and are frequently annotated as oncogenes. When we sort the domains by the number of in-links and out-links, from least to most connected, it is clear that oncogenic proteins (red) are preferentially enriched among proteins with high in-connectivity (median 20 versus 8, Mann-Whitney U test \( p = 3.4 \times 10^{-3} \)) and out-links (9 versus 1, Mann-Whitney U test \( p = 1.0 \times 10^{-4} \)) than nononcogenic adaptors. SH2 domains bind a higher median number of peptides (15.5 versus 2, Mann-Whitney U test \( p = 1.7 \times 10^{-3} \)) but share a similar number of out-links (Mann-Whitney U Test \( p > 0.01 \)).

Oncogenes are Highly Interconnected

The quantitative interaction data generated in this study (supplemental Table S3) should prove generally useful in the fields of network modeling (53) and interaction prediction (54). With the caveat that our data necessarily include false positives and false negatives, we also hope that the novel interactions identified in these studies will aid efforts to uncover signaling pathways downstream of RTKs. To facilitate the use of our data, the supplementary information for this paper contains the annotated interaction list (supplemental Table S3) and raw microarray image intensities as well as the MATLAB® scripts used to fit the titration curves, perform statistical analyses, and generate figures (Supplementary Code).

Our work shows that the genes whose mutation or overexpression drives cancer tend to be highly connected hubs. Moreover, the most highly connected proteins in our networks are the primary targets of new anticancer drugs (e.g. erlotonib for EGFR (55), lapatanib for ERBB2 (56), vandetanib for RET (57), sunitinib for KIT (58), and imatanib for ABL1 (14)). Analogously, viral proteins preferentially target hub proteins during infection (59, 60). This combined evidence suggests that disease states arise preferentially from the perturbation of network hubs and drugs should target these same hub proteins.

Fig. 4. Oncogenic adaptors are more highly promiscuous. A, Phylogeny of adaptors shows that genetically related domains share similar binding profiles. For example the SRC class (SRC, YES1, FGR) of tyrosine kinases all show a large number of in-links. This can be because of sequence determinants of selectivity as well as the bias in the phosphoproteome to sequences that interact with the domain family. B, Adaptors sorted according to the number of peptides with which they interact (in-links; the x axis) and the number of adaptors that interact with them (out-links; y axis). Oncogenic adaptors have both higher median in-links (20 versus 8, Mann-Whitney U test \( p = 3.4 \times 10^{-3} \)) and out-links (9 versus 1, Mann-Whitney U test \( p = 1.0 \times 10^{-4} \)) than nononcogenic adaptors. SH2 domains bind a higher median number of peptides (15.5 versus 2, Mann-Whitney U test \( p = 1.7 \times 10^{-3} \)) but share a similar number of out-links (Mann-Whitney U Test \( p > 0.01 \)).
In conclusion, we have generated a systematic map covering a substantial fraction of the potential interactions between SH2 or PTB domains and sites of tyrosine phosphorylation on RTKs and adaptor proteins. These interactions are very poorly represented in existing unbiased human interactomes (61, 62) despite extensive evidence they play essential roles in signal transduction. We observe a high degree of connectivity among RTKs and adaptor proteins, and among adaptor proteins themselves. This is in contrast to the usual depiction of receptor-proximal signaling as a series of linear pathways connecting sites of tyrosine phosphorylation on RTKs to a few adaptor proteins and then to a few core signaling molecules such as ERK and AKT. Of course, the actual complexity of the resulting network in any particular cell type depends not only on the affinities of the interactions (as determined here) but also on the relative abundance of RTKs and SH2/PTB-containing proteins. mRNA sequencing data of 1430 diverse tumor types collected by the TCGA consortium shows that they express a median of 78% of the phosphotyrosine signaling proteins analyzed in this study. In fact, over half of these proteins are expressed ubiquitously (in >95% of the tumor samples; supplemental Fig. S2). Thus, it is highly likely that receptor and adaptor proteins combine to form a highly interconnected mesh with the potential to perform complex signal processing functions.

Acknowledgments—We thank Jeffrey Knott, Susan Rogers, and Colleen Hunter (Cell Signaling Technology Inc.) for peptide synthesis, Bogdan A. Budnick and William S. Lane (FAS Center for Systems Biology) for mass spectrometry, Vijay Warrier for Ephrin receptor proteins, Bogdan A. Budnick and William S. Lane (FAS Center for Systems Biology) for mass spectrometry, Vijay Warrier for Ephrin receptor proteins, Colleen Hunter (Cell Signaling Technology Inc.) for peptide synthesis, and Taran S. Gujral (Harvard Medical School) for helpful discussions and Taran S. Gujral (Harvard Medical School) for helpful discussions and Taran S. Gujral (Harvard Medical School) for helpful discussions and Taran S. Gujral (Harvard Medical School) for helpful discussions. We thank Jeffrey Knott, Susan Rogers, and Colleen Hunter (Cell Signaling Technology Inc.) for peptide synthesis, Bogdan A. Budnick and William S. Lane (FAS Center for Systems Biology) for mass spectrometry, Vijay Warrier for Ephrin receptor proteins, Bogdan A. Budnick and William S. Lane (FAS Center for Systems Biology) for mass spectrometry, Vijay Warrier for Ephrin receptor proteins, Colleen Hunter (Cell Signaling Technology Inc.) for peptide synthesis, and Taran S. Gujral (Harvard Medical School) for helpful discussions and Taran S. Gujral (Harvard Medical School) for helpful discussions.

This work was supported by an award from the W. M. Keck Foundation and by grants from the National Institutes of Health (R33 CA128726 and RO1 GM072872). G.K. was the recipient of the AstraZeneca Graduate Fellowship and the Max Tisher Graduate Fellowship. A.K. was supported in part by the NIH Molecular, Cellular, and Chemical Biology Training Grant (5 T32 GM07598).

This article contains supplemental Figs. S1 and S2 and Tables S1 to S4.

‡‡ To whom correspondence should be addressed: Department of Systems Biology, Harvard Medical School, 200 Longwood Ave., Boston, MA 02115. Tel.: 617-432-6902; Fax: 617-432-6990; E-mail: gavin_macbeath@harvard.edu.

REFERENCES

1. Lemmon, M. A., and Schlessinger, J. (2010) Cell signaling by receptor tyrosine kinases. Cell 141, 1117–1134
2. Liu, B. A., Jablonowski, K., Raina, M., Arcé, M., Pawson, T., and Nash, P. D. (2006) The Human and Mouse Complement of SH2 Domain Proteins–Establishing the Boundaries of Phosphotyrosine Signaling. Mol. Cell 22, 651–663
3. Jones, R. B., Gordus, A., Kralj, J. A., and MacBeath, G. (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature 439, 168–174
4. Salcini, A. E., McGlade, J., Pelicci, G., Nicoletti, I., Pawson, T., and Pelicci, P. G. (1994) Formation of Shc-Grb2 complexes is necessary to induce neoplastic transformation by overexpression of Shc proteins. Oncogene 9, 2827–2836
5. Esposito, D. L., Li, Y., Cama, A., and Quon, M. J. (2001) Tyr(612) and Tyr(632) in human insulin receptor substrate-1 are important for full activation of insulin-stimulated phosphatidylinositol 3-kinase activity and translocation of GLUT4 in adipose cells. Endocriology 142, 2833–2840
6. Marshall, C. J. (1995) Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. Cell 80, 179–185
7. Santarius, T., Shipley, J., Brewer, D., Stratton, M. R., and Cooper, C. S. (2010) A census of amplified and overexpressed human cancer genes. Nat. Rev. Cancer 10, 59–64
8. Perren, T. J. (1991) c-erbB-2 oncogene as a prognostic marker in breast cancer. Br. J. Cancer 63, 328–332
9. McIntyre, A., Summersgill, B., Gryglewicz, B., Gillis, A. J., Stoop, J., van Gorp, R. J., Dennis, N., Fisher, C., Huddart, R., Cooper, C., Clark, J., Oosterhuis, J. W., Looijenga, L. H., and Shipley, J. (2005) Amplification and overexpression of the KIT gene is associated with progression in the seminoma subtype of testicular germ cell tumors of adolescents and adults. Cancer Res. 65, 8085–8089
10. Comoglio, P. M., Giordano, S., and Trusolino, L. (2008) Drug development of MET inhibitors: targeting oncogene addiction and epide. Nat. Rev. Drug Discovery 7, 504–516
11. Mulligan, L. M., Kwok, J. B., Healey, C. S., Elsdon, M. J., Eng, C., Gardner, E., Love, D. R., Mole, S. E., Moore, J. K., and Papi, L. (1993) Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. Nature 363, 458–460
12. Hofstra, R. M., Landsvater, R. M., Ceccherini, I., Stulp, R. P., Stelwagen, T., Luo, Y., Pasini, B., Höppener, J. W., van Amstel, H. K., and Romeo, G. (1994) A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid cancer. Nature 367, 375–376
13. Hirotia, S., Isozaki, K., Moriyama, Y., Hashimoto, K., Nishida, T., Ishiguro, S., Kawano, K., Hanada, M., Kurata, A., Takeda, M., Tunio, G., Matsu- zawa, Y., Kanakura, Y., Shimomura, Y., and Kitamura, Y. (1998) Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. Science 279, 577–580
14. Drucker, B. J., Sawyer, C. L., Kantarjian, H., Resta, D. J., Reese, S. F., Ford, J. M., Capdeville, R., and Talpaz, M. (2001) Activity of a specific inhibitor of the BCR-ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. N. Engl. J. Med. 344, 1038–1042
15. Cheung, L. W., Hennessy, B. T., Li, J., Yu, S., Myers, A. P., Djordjevic, B., Lu, Y., Stemke-Hale, K., Dyer, M. D., Zhang, F., Ju, Z., Cantley, L. C., Scherer, S. E., Liang, H., Liu, K. H., Broadus, R. R., and Mills, G. B. (2011) High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. Cancer Discovery 1, 170–185
16. Liu, Y. Y., Slotine, J. J., and Barabási, A. L. (2011) Controllability of complex networks. Nature 473, 167–173
17. Jonason, P. F., and Bates, P. A. (2006) Global topological features of cancer proteins in the human interactome. Bioinformatics 22, 2291–2297
18. Wachi, S., Yoneda, K., and Wu, R. (2005) Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer elucidates a novel mechanism for regulation of PTEN protein stability. Cancer Discovery 1, 170–185
19. Hakes, L., Robertson, D. L., and Oliver, S. G. (2005) Effect of dataset selection on the topological interpretation of protein interaction networks. BMC Genomics 6, 131
20. Hakse, E., Pinney, J. W., Robertson, D. L., and Lovell, S. C. (2008) Protein–protein interaction networks and biology—what’s the connection? Nat. Biotechnol. 26, 69–72
21. Phizicky, E., Bastiaens, P. I., Zhu, H., Snyder, M., and Fields, S. (2003) Protein analysis on a proteomic scale. Nature 422, 208–215
22. Mayer, B. J., Blinov, M. L., and Loew, L. M. (2009) Molecular machines or molecular motors? An oriented peptide array library (OPAL) strategy to study protein-protein interactions. J. Biol. Chem. 279, 8802–8807
Oncogenes are Highly Interconnected

25. Zhou, S., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J. (1993) SH2 domains recognize specific phosphopeptide sequences. Cell 72, 767–778

26. Liu, B. A., Jablonowski, K., Shah, E. E., Engelmann, B. W., Jones, R. B., and Nash, P. D. (2010) SH2 domains recognize contextual peptide sequence information to determine selectivity. Mol. Cell Proteomics 9, 2381–2404

27. Gorelik, M., Stanger, K., and Davidson, A. R. (2011) A Conserved residue in the yeast Bem1p SH3 domain maintains the high level of binding specificity required for function. J. Biol. Chem. 286, 19470–19477

28. Kaushansky, A., Allen, J. E., Gordus, A., Stifter, M. A., Karp, E. S., Chang, B. H., and MacBeath, G. (2010) Quantifying protein–protein interactions in high throughput using protein domain microarrays. Nat. Protoc. 5, 773–790

29. Lefort, A. M., and Schlessinger, J. (2010) Cell signaling by receptor tyrosine kinases. Cell 141, 1117–1134

30. Kaushansky, A., Gordus, A., Budnik, B. A., Lane, W. S., Rush, J., and MacBeath, G. (2008) System-wide investigation of ErbB4 reveals 19 sites of Tyr phosphorylation that are uniquely selective in their recruitment properties. Chem. Biol. 15, 808–817

31. Jung, A. S., Kaushansky, A., Macbeath, G., and Kaushansky, K. (2011) Tyro3 is a receptor in thrombopoietin-induced pluripotent cell proliferation by promoting Akt signaling. Cell Cycle 10, 1838–1844

32. Mehlitz, A., Banhart, S., Mäurer, A. P., Kaushansky, A., Gordus, A. G., Zielecki, J., Macbeath, G., and Meyer, T. F. (2010) Tarp regulates early Chlamydia-induced host cell survival through interactions with the human adaptor protein SHC1. J. Cell Biol. 190, 143–157

33. Boettcher, J. P., Kirchner, M., Churin, Y., Kaushansky, A., Pompiaia, M., Thompson, A., Brinkley, B., Macbeath, G., and Meyer, T. F. (2010) Tyrosine-phosphorylated caveolin-1 blocks bacterial uptake by inducing Vav2-RhoA-mediated cytoskeletal rearrangements. PLoS Biol. 8

34. Huber, P. J. (2003) Robust Statistics, John Wiley & Sons

35. Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Flys, P., Ogilvie, I., Clegg, S., Packer, B., A., and Pritchard, T. (2004) RTCGD: retroviral tagged cancer gene database.

36. Huret, J. L., Minor, S. L., Dorkeld, F., Dessen, P., and Bernheim, A. (2000) Atlas of genetics and cytogenetics in oncology and haematology, an comprehensive resource for investigating the structure and function of human genes. Biochim. Biophys. Acta 1516, 808–817

37. Sjoblom, T., Jones, S., Wood, L. D., Parsons, D. W., Lin, J., Barber, T. D., Gill, J. H., Weir, B. A., McPherson, J., Tu, C.-H., Liu, C., Gabriel, S., Wendl, M. C., Wexler, S., Kim, S. A., Marra, M. A., Delaney, A. L., Mardis, E. R., Ding, L., and Ley, T. J. (2007) The Cancer Genome Atlas (TCGA): an integrated genomic resource for oncology and basic biomedical research. Nature 447, 719–724

38. Akagi, K., Suzuki, T., Stephens, R. M., Jenkins, N. A., and Copeland, N. G. (2002) Mapping the consistent core of common genes in breast and colorectal cancers. Cancer Cell 1, 727–737

39. Lemmon, M. A., and Schlessinger, J. (2010) Cell signaling by receptor tyrosine kinases. Nature Revs Cancer 10, 177–183

40. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Biotechnol. 26, 551–555

41. Hurett, J. L., Minor, S. L., Dorkeld, F., Dessen, P., and Bernheim, A. (2000) Atlas of genetics and cytogenetics in oncology and haematology, an comprehensive resource for investigating the structure and function of human genes. Biochim. Biophys. Acta 1516, 808–817

42. Schlessinger, J. (1995) Measurement of the binding of tyrosyl phosphopeptides to SH2 domains recognize specific phosphopeptide sequences. Cell 72, 767–778

43. Kaushansky, A., Gordus, A., Chang, B. H., MacBeath, G. (2008) System-wide investigation of ErbB4 reveals 19 sites of Tyr phosphorylation that are uniquely selective in their recruitment properties. Chem. Biol. 15, 808–817

44. Huber, P. J. (2003) Robust Statistics, John Wiley & Sons

45. Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Flys, P., Ogilvie, I., Clegg, S., Packer, B., A., and Pritchard, T. (2004) RTCGD: retroviral tagged cancer gene database.

46. Huret, J. L., Minor, S. L., Dorkeld, F., Dessen, P., and Bernheim, A. (2000) Atlas of genetics and cytogenetics in oncology and haematology, an comprehensive resource for investigating the structure and function of human genes. Biochim. Biophys. Acta 1516, 808–817

47. Sjoblom, T., Jones, S., Wood, L. D., Parsons, D. W., Lin, J., Barber, T. D., T., Marshall, M., Down, T., Hubbard, T., Wooster, R., Flys, P., Ogilvie, I., Clegg, S., Packer, B., A., and Pritchard, T. (2004) RTCGD: retroviral tagged cancer gene database. Nucleic Acids Res. 32, D523–D527

48. Network, T. C. G. A. (2012) Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330–337

49. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 621–628

50. Nagaraj, N., Wilsie, J. R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pålbo, S., and Mann, M. (2011) Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. 7, 548

51. Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Stitzenberg, E. E., Murr, W., Latham, V., and Sullivan, M. (2011) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Research 40, D261–D270

52. Kaushansky, A., Gordus, A., Chang, B., Rush, J., and MacBeath, G. (2008) A quantitative study of the recruitment potential of all intracellular tyrosine residues on EGFR, FGFR1 and IGF1R. Mol. Biol. Cell 9, 643–653

53. Leduc, J., Lemmon, M. A., Zhou, M., Green, J., Bottled, M. C., and Schlessinger, J. (1995) Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: a reanalysis. Proc. Natl. Acad. Sci. U.S.A. 92, 3199–3203

54. Piccione, E., Case, R. D., Domchek, S. M., Hu, P., Chaudhuri, M., Backer, J. M., Schlessinger, J., and Shoelson, S. E. (1993) Phosphatidylinositol 3-kinase p85 SH2 domain specificity defined by direct phosphopeptide/
Oncogenes are Highly Interconnected

Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J. M., Cevik, S., Simon, C., Smet, A.-S. de, Dann, E., Smolyar, A., Vinayagam, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R. R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M. E., Roth, F. P., Hill, D. E., Tavernier, J., Wanker, E. E., Barabási, A. L., and Vidal, M. (2009) An empirical framework for binary interactome mapping. Nature Methods 6, 83–90

63. Lennartsson, J., Wernstedt, C., Engström, U., Hellman, U., and Rönnstrand, L. (2003) Identification of Tyr900 in the kinase domain of c-Kit as a Src-dependent phosphorylation site mediating interaction with c-Crk. Experimental Cell Research 288, 110–118

64. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., and Marra, M. A. (2009) Circos: An information aesthetic for comparative genomics. Genome Research 19, 1639–1645