Protein expression profiles of C3H 10T1/2 murine fibroblasts and of isogenic cells transformed by the H1047R mutant of phosphoinositide 3-kinase (PI3K)

Jonathan Ross Hart,^{1,*} Lujian Liao,^{2,*} Lynn Ueno,¹ John R. Yates III² and Peter K. Vogt¹

¹Department of Molecular and Experimental Medicine ²Department of Chemical Physiology; The Scripps Research Institute; La Jolla, CA USA

Key words: stable isotope labeling, SILAC, mass spectrometry, oncogenic transformation, PI 3-kinase

Abbreviations: Akt, cellular homolog of murine thymoma virus akt8 oncogene; Acacb, acetyl-CoA carboxylase 2; eIF4E, eukaryotic initiation factor 4E; FoxO1, forkhead transcription factors O1; FoxO3a, forkhead transcription factors O3a; Hif1 α , hypoxia-induced factor 1 α ; Ifit1, interferon-induced protein with tetratricopeptide repeats 1; Igtp, interferon gamma-inducible GTP-binding protein; IL-1, interleukin-1; Irs1, insulin receptor substrate 1; Isg15, interferon stimulated gene 15; LC-MS/MS, liquid chromatography-tandem mass spectrometry; Mlck, myosin light chain kinase; Oas2, 2'-5'-oligoadenylate synthetase 2; p70S6K, p70 S6 kinase; Papss, 3'-phosphoadenosine 5'-phosphosulfate synthetase; PI3K, phosphatidylinositol-3-kinase; Pi4k2a, phosphatidylinositol-4-kinase type 2 α ; Pon3, paraoxonase 3; Prss1, protease, serine 1; Pten, phosphatase and tensin homolog deleted on chromosome 10; RCAS, replication competent avian leukosis virus with a splice acceptor; RTKs, receptor tyrosine kinase; SILAC, stable isotope labeling with amino acids in cell culture; Stat, signal transducer and activator of transcription; TOR, target of rapamycin; YB-1, Y-box binding protein 1

We have used stable isotope labeling with amino acids in cell culture (SILAC) in conjunction with tandem mass spectrometry to characterize the proteomes of two isogenic cell lines that differ in the expression of a single oncoprotein, p110 α of PI3K, carrying the H1047R mutation. 51,510 peptides were identified and assigned to 4,201 proteins. Most notable among the proteins that show increased expression in the oncogenically transformed cells are several involved in the interferon response including Isg15, Ifit1, Igtp and Oas2 (interferon stimulated gene 15, interferon-induced protein with tetratricopeptide repeats 1, interferon gamma-inducible GTP-binding protein, 2'-5'-oligoadenylate synthetase 2). Prominent among the downregulated proteins are several involved in cell adhesion as well as proteins that are affected by the negative feedback from PI3K signaling. The differential expressions documented in this analysis suggest novel links between oncogenic PI3K and several signaling pathways. These links will be explored in future studies.

Introduction

Hyperactivation of the phosphatidylinositol kinase (PI3K) pathway is frequently observed in many cancers. It can result from gain-of-function mutations in PI3K, inactivating mutations or loss of PTEN, and activating mutations of upstream signaling partners such as receptor tyrosine kinases or RAS. Active PI3K induces the activation of the downstream kinases AKT and TOR. These in turn regulate the expression of proteins both through transcription factors including FOXO1, FOXO3a (forkhead transcription factors O1, O3a) and HIF1 α (hypoxia-induced factor 1 α) and through translational control mediated by the ribosomal protein S6, YB-1 (Y-box binding protein 1) and release of eIF4E (eukaryotic initiation factor 4E). We have investigated changes induced in the abundance of individual proteins

by the cancer-derived mutant of PI3K H1047R using stable isotope labeling with amino acids in cell culture (SILAC). Tandem mass spectrometry permits the identification and relative quantification of proteins expressed under different conditions. We have applied this method to a pair of isogenic cell lines, the mouse fibroblast line C3H 10T1/2 and the same cell line oncogenically transformed by the H1047R mutant of p110 α , the catalytic subunit of PI3K. We were able to quantify the changes to the global proteome induced by the hyperactive PI3K mutant.

Results

Transformation of C3H 10T1/2 murine fibroblasts with the H1047R mutant of p110 α , the catalytic subunit of PI3K. We generated 10T1/2 cells that express tva, the cell surface receptor

*Correspondence to: Jonathan Ross Hart and Lujian Liao; Email: jhart@scripps.edu and lliao@scripps.edu

Submitted: 02/04/11; Accepted: 02/07/11

DOI: 10.4161/cc.10.6.15077

of avian sarcoma and leukosis viruses, envelope subgroup A.^{1,2} 10T1/2 cells were transfected with the pCB6-tva vector expression vector (a gift from Paul Bates) followed by selection with 500 µg/ml of G418. Several G418-resistant colonies were picked and expanded. One of these clones (T1) was highly sensitive to oncogenic transformation by several oncogenes including *src* and *ras* as mediated by the avian retroviral expression vector RCAS(A).³ This clone was infected with the RCAS(A) vector expressing the H1047R mutant of p110α.⁴ Transformed cell foci were picked with a capillary pipette, expanded and then tested for anchorage-independent growth following published techniques.⁵ Several colonies growing in soft nutrient agar were isolated. They consisted of rounded, poorly adhering cells that grew significantly faster than the parental 10T1/2 tva cells, which were flat, fibroblastic and adhered tightly to the plastic substrate. One of the transformed clones was used in the present study. The cultivation of the transformed and the parental cell lines was conducted in the presence of 200 µg/mL G418 except during labeling for SILAC when the selection agent was omitted from the culture medium (see below).

Western blots of the selected colony revealed elevated expression of p110α and constitutive phosphorylation of Akt and ribosomal protein S6, independent of serum stimulation. In contrast, the parental 10T1/2 cell line showed phosphorylation of Akt and S6 only after serum stimulation. These results document the constitutive activity of the H1047R mutant of p110α in the transformed C3H 10T1/2 cells. The H1047R-transformed cells were also dependent upon PI3K activity for growth and proliferation. The addition of PI3K inhibitors to these cells significantly decreased proliferation. The growth of wild type C3H 10T1/2 cells was less sensitive to inhibitors of PI3K.

Stable isotope labeling for protein quantitation. The H1047R-transformed cells were labeled with U-¹⁵N₄-arginine (U-¹³C, U-¹⁵N, uniformly labeled with ¹³C and ¹⁵N respectively) and U-¹³C₆,U-¹⁵N₂-lysine, and the untransformed cells were labeled with U-¹²C₆,U-¹⁴N₄-arginine and U-¹²C₆,U-¹⁴N₂-lysine for use in SILAC analysis. For this procedure, all unlabeled arginine and lysine must be removed from both the culture medium and the serum. Initially, a custom DMEM medium lacking arginine and lysine was prepared and supplemented with dialyzed fetal bovine serum lacking all amino acids. Labeled arginine and lysine were added for incorporation into proteins. However, this medium failed to support the growth of both the parental and the transformed cell line. The problem was overcome by switching to MCDB 170 medium supplemented with dialyzed serum and bovine pituitary extract.⁶ This medium supported robust growth and the addition of bovine pituitary extract did not interfere with isotopic labeling. Bovine pituitary extract may prove beneficial in stable isotope labeling experiments with other cell lines that cannot be grown with dialyzed serum. Cells were labeled in triplicate during five passages in the modified media. H1047R-transformed cells were cultivated in the heavy medium, and the parental cell line was grown in the medium containing the light amino acids. Cells were harvested by trypsinization and counted. H1047R-transformed and parental cells were then mixed at a 1:1 ratio. The cells were lysed and fractionated into cytoplasmic and

nuclear fractions. The resulting proteins were analyzed by multidimensional LC-MS/MS (liquid chromatography-tandem mass spectrometry).

Overview of differentially expressed proteins. We identified 51,510 peptides (Sup. Table 1) which could be assigned to 4,201 proteins (Sup. Table 2). 99 proteins were upregulated in the H1047R-transformed cells ($p < 0.05$), and 49 proteins were downregulated ($p < 0.05$) by a factor of ≥ 2 (Sup. Tables 3 and 4). In the H1047R cell line, there was also a systematic overexpression of all proteins by an average of 10%. This difference was not corrected, because H1047R-expressing cells contain more protein on a per cell basis as a result of the hyperactivation of the TOR kinase.⁷ A selection of upregulated proteins are shown in Table 1, downregulated proteins in Table 2.

Among the highly upregulated proteins are several involved in interferon response including Isg15, Ifit1, Igtp and Oas2 (Interferon stimulated gene 15, interferon-induced protein with tetratricopeptide repeats 1, interferon gamma-inducible GTP-binding protein, 2'-5'-oligoadenylate synthetase 2). Also upregulated are caspase 1, Pon3 (paraoxonase 3), Acacb (acetyl-CoA carboxylase 2) and Mlck (myosin light chain kinase). Other differentially expressed proteins are encoded by the vector for the H1047R mutant of p110α, including sequences of the retroviral Gag, Pol and Env proteins, which are not expressed in the parental cell line.

There are many downregulated proteins as well. Irs1 (insulin receptor substrate 1) is strongly downregulated. This is expected, as strong signaling through the PI3K pathway leads to the activation of p70S6K (p70 S6 kinase) and consequent feedback phosphorylation of IRS1. This phosphorylation of Irs1 induces the ubiquitination and destruction of the protein.⁸⁻¹⁰ The abundance of Pi4k2a, (phosphatidylinositol-4-kinase type 2α), another member of the phosphoinositide metabolic pathway, is also significantly reduced. This protein has been recently identified as an essential part of autophagy¹¹ and also regulates angiogenesis and HIF1α (hypoxia-inducible factor 1α) activity. Additional downregulated proteins are cadherins, catenins and integrins involved in cellular morphology and attachment. The proteins with the lowest abundance ratio are derived from the culture medium or the processing of the cells including dialyzed serum peptides, PRSS1 (protease, serine 1), keratin, hemoglobin as well as the trypsin used for the digestion of the proteins.

Statistical and gene ontology analysis. The proteomics data produced from this experiment are fundamentally different from nucleic acid microarray data. Whereas a microarray can detect changes in essentially all mRNA transcripts, only a fraction of the proteins in the cell can be quantified by mass spectrometry. The detection of a peptide in LC-MS/MS is not determined by abundance alone; rather, solubility, chromatographic behavior, ionization and possible interaction of the peptide ion with other compounds can also play critical roles. As a result, not all peptides are detected in repeat analyses of the same digest, and this loss of data makes the multivariate statistical methods used in microarray analysis inappropriate. In order to circumvent this problem, two methods are in common use, either reject all proteins that have a value missing in one sample or set an arbitrary threshold and analyze overrepresentation.

Table 1. Proteins with enhanced expression in PI3K H1047R-transformed cells

Ratio transformed to control	Protein	Protein name
13.48	Isg15	Isg15 ubiquitin-like modifier
8.95	Oas2	2'-5' oligoadenylate synthetase 2, isoform CRA_a
6.05	Cd80	CD80 molecule
5.18	Ifit1	Interferon-induced protein with tetratricopeptide repeats 1
4.33	Errfi1	ERBB receptor feedback inhibitor 1
3.79	Akr1c13	Aldo-keto reductase family 1 member C13
3.72	Igtp	interferon gamma induced GTPase
2.72	Adfp	Adipophilin
2.34	Nudt11	Nudt10 Diphosphoinositol polyphosphate phosphohydrolase 3-alpha
2.23	Hsp90b1	heat shock protein 90 kDa beta (Grp94), member 1
1.59	Eif2b2	eukaryotic translation initiation factor 2B, subunit 2beta
1.51	Stk10	serine/threonine kinase 10

Table 2. Proteins with reduced expression in PI3K H1047R-transformed cells

Ratio control to transformed	Protein	Protein name
6.82	IRS1	Insulin receptor substrate 1
2.88	Cdh2	cadherin 2, type 1, N-cadherin (neuronal)
2.62	Rrm1	ribonucleotide reductase M1
1.99	Ctnnb1	catenin (cadherin associated protein), beta1
1.89	Cat	catalase
1.87	Nedd4	neural precursor cell expressed, developmentally down-regulated 4
1.74	Cd44	CD44 molecule (Indian blood group)
1.65	Ctnnd1	catenin (cadherin-associated protein), delta
1.57	Gnai2	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2

These methods reduce the amount of information produced in a protein expression, and therefore we have developed a different approach. Many peptides are measured numerous times in LC-MS/MS. We used each measurement of a peptide, including molecular weight and ratio, as evidence for that particular protein. We pooled all the evidence for a protein to make an inference about the mean abundance ratio and variance of that protein. Using that information, we then compared this distribution to the distribution of all evidence using Student's *t*-test. With this approach, we identified proteins that are up- or downregulated by the presence of the H1047R mutant of PI3K.

We extended this analysis to protein functions. The Gene Ontology database associates particular functions, subcellular localizations and activities with groups of specific proteins. For each of these classes in the Gene Ontology database, we pooled the evidence from the abundance measurements and compared the resulting distribution with the distribution of all evidence. This analysis identifies specific functional classes that have been altered by the expression of H1047R. Of the 5,439 assignable functions, 2,788 differ significantly ($p < 0.05$) in the comparison of normal versus H1047R-transformed cells. Because of this multiplicity of altered functions, we ordered them by effect size, applying Cohen's *d* estimator to the data. Cohen's *d* estimator

provides a value for the number of standard deviations that separate a particular distribution from the mean. Of those functions that differ significantly from the mean ($p < 0.05$) and have a large effect size ($d > 0.8$), 517 are downregulated and 330 are upregulated (Sup. Tables 5 and 6).

The most upregulated functions are associated with the retroviral vector, including the Gag, Pol or Env proteins. Ignoring these obvious results, we find large effect sizes associated with the upregulation of Isg15 functionalization, 2'-5'-oligoadenylate synthetase activity, aspartic-type endopeptidase activity (Cathepsin D) and positive regulation of interleukin-1 secretion. Both Isg15 and 2'-5'-oligoadenylate synthetase activity are strongly associated with interferon response, which in turn is driven by the Stat family of transcription factors.

Among the downregulated processes, the largest effect sizes are seen with the trivial contaminants from the dialyzed serum, the cell culture medium and the trypsin used in digestion. The next-ranking downregulated function is the positive regulation of PI3K signaling resulting from the negative feedback that is initiated by the p70S6K-mediated phosphorylation of Irs1.⁸⁻¹⁰ Also downregulated are sulfation of proteins via the Papss (3'-phosphoadenosine-5'-phosphosulfate synthetase), sialylation and fucosylation of proteins, cell adhesion and ribonucleotide reductase.

Discussion

Transformation of C3H 10T1/2 cells by p110 α H1047R. The C3H 10T1/2 cells used in this study were transformed by infection with an RCAS retrovirus expressing the p110 α H1047R mutant. H1047R induces constitutive signaling through Akt and TOR and is needed for the replication of these cells. The protocol for transforming and selecting the cells may have led to secondary mutations, and such secondary mutations may be required for transformation. However, secondary mutations do not abrogate the role of H1047R in these cells but rather support it. H1047R expression is driven by a viral promoter and its level is much higher than that of endogenous wild type p110 α . The high levels of H1047R can cause changes in the expression of the endogenous p110 α and of p110 β as a result of competition for p85 regulatory subunits.

Our experience with growing these cells with dialyzed serum shows that they are not growth factor-independent. Replication of these cells requires the addition of dialyzed fetal bovine serum and of bovine pituitary extract to the medium. The dependence on these media supplements raises the possibility that the growth factors provided by these supplements fundamentally change protein expression. Bovine pituitary extract stimulates numerous different signaling pathways. It could obscure some of the changes induced by H1047R, because bovine pituitary extract also stimulates the PI3K pathway through endogenous mechanisms. However, both the transformed and the control cell lines were treated with equal amounts of bovine pituitary extract. Therefore changes to the proteome are fundamentally due to H1047R expression and should be regarded as changes observed under serum stimulation rather than changes observed under serum starvation.

The proteomic profile. The differentially expressed proteins in p110 α -transformed cells form a profile that shows characteristic features of a cancer cell. These include upregulation of proliferation-related proteins, secretion of IL-1, drug resistance and unfolded protein response. Among the downregulated functions there are several that cause changes of the cell surface and reduce adhesiveness as well as autophagy. Additionally, the data confirm several features known to be associated with the hyperactivated PI3K pathway, including enhanced metabolism of phosphatidylinositols, negative regulation of several RTKs and negative regulation of Irs1. The proteomic profile of the p110 α -transformed cells also shows an unexpected feature: enhanced expression of interferon-stimulated genes that are driven by Stat-mediated transcription. Although Stat proteins are often activated in cancer¹²⁻¹⁸ and in some PI3K-driven tumors, such activation has not been observed previously in p110 α -transformed cells. We are currently investigating the relevance of this outstanding feature for the process of PI3K-induced oncogenic transformation.

Statistical analysis. Evidence and degrees of freedom. In shotgun proteomics, it is common that a single protein is measured numerous times in a single experiment. The statistical evaluation of such complex data sets is not fully developed. Data produced from two independent preparations of the same cells are clearly

independent experiments. In the analysis of the present data, measurements of multiple peptides of the same protein were also treated as independent. Similarly, if a specific protein peptide was measured multiple times, all measurements were considered as independent.

Large expression changes. Peptides representing large changes in expression are observed as “singletons” where a tandem mass spectrum is assignable to a particular protein based on sequence data but lacks either the heavy or light isotope correlate. Such data usually represent large changes where the other state (heavy or light isotope labeled peptide) falls below the limit of detection and is not detected. Data from such singletons are included in the analysis and the ratios given represent the signal intensity divided by the noise background. This ratio corresponds to the minimum up- or downregulation of the singleton protein. For statistical tests, all ratios obtained are treated as equalities and not inequalities. This treatment results in more stringent testing and in an underestimation of the true ratios for these peptides.

Pooled data. Peptide measurements can be linked to a specific protein, and by pooling the measurements, a relative expression of the protein is obtained. Likewise, specific proteins are linked to functions, cellular localizations and processes through the gene ontology database. Each term in the gene ontology database has specific proteins associated with it, and we have pooled all the measurements for the peptides belonging to these proteins and compared this pooled distribution to the distribution of the complete data set. In some cases there are disproportionate numbers of measurements originating from certain members of a specific gene ontology pool. The pool distribution then closely resembles that of the member with the largest number of measurements. The differences in the number of measurements are a result of expression level, ease of ionization, solubility and numerous other factors that could bias our statistical tests. We have tested this possibility by pooling proteins instead of peptides for the gene ontology analysis and found no evidence for significant bias.

Some peptides cannot be unambiguously assigned to a single protein, usually because that particular protein has several closely related isoforms. For these cases, the individual peptide measurements have only been added to the pool once and not once for every possible protein assignment.

Materials and Methods

Cell lines. The mouse embryonic fibroblast cell line C3H 10T1/2 expressing the avian retroviral subgroup A receptor, *tva*, were infected by the RCAS(A)-p110 α -H1047R virus.¹⁹⁻²¹ This virus transformed the fibroblasts and allowed them to form colonies in soft agar. The colonies were picked and grown to establish the 10T1/2-H1047R cell line. Cells were grown in DMEM (Invitrogen, Carlsbad, CA) supplemented with 10% fetal calf serum (Omega, Tarzana, CA), penicillin, streptomycin, L-glutamine (Sigma, St. Louis, MO) and 200 ug/mL G418 (Invitrogen, Carlsbad, CA).

Stable isotope labeling. Cells were labeled in a defined media based upon MCDB 170.⁶ The media was prepared as described except arginine, lysine and DL- α -lipoic acid were withheld and

the media was supplemented with 140 nM hydrocortisone (Sigma, St. Louis, MO), 5 ng/mL insulin (bovine) (Sigma, St. Louis, MO), 10 ng/mL recombinant EGF (Sigma, St. Louis, MO), 100 mM ethanolamine (Sigma, St. Louis, MO), 100 mM phosphoethanolamine (Sigma, St. Louis, MO), 10% dialyzed fetal calf serum (Omega, Tarzana, CA), penicillin, streptomycin (Sigma, St. Louis, MO), L-glutamine (Invitrogen, Carlsbad, CA) and bovine pituitary extract (Hammond Cell Tech, Windsor, CA). U-¹²C₆, U-¹⁴N₄-arginine and U-¹²C₆, U-¹⁴N₂-lysine were added to this media to form the light labeling media or U-¹³C₆, U-¹⁵N₂-lysine and U-¹³C₆, U-¹⁵N₄-arginine (Cambridge Isotope Labs, Andover, MA) to form heavy labeling media. Cells were cultured in isotopically labeled media for five passages before lysis.

Cells were lysed using the nuclear fractionation procedure. Cells were released from plates using trypsin to form cell suspensions. The cell suspensions were counted by coulter counter and mixed 1:1 H1047R:10T1/2. These cell suspensions were centrifuged and the supernatants removed. The cell pellet was resuspended in 5 mL of 10 mM HEPES, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM dithiothreitol, 0.05% Igepal CA-630 (Sigma, St. Louis, MO), 1 mM PMSF, 2 mM NaVO₄, complete protease inhibitor cocktail (Roche). This suspension was left on ice for 5 minutes followed by centrifugation at 800x g at 4°C. The supernatant was saved as the cytoplasmic fraction. The pellet was resuspended in 1,122 µL of 5 mM HEPES (Sigma, St. Louis, MO), 1.5 mM MgCl₂ (Sigma, St. Louis, MO), 0.2 mM EDTA (Fisher, Pittsburgh, PA), 0.5 mM dithiothreitol (Sigma, St. Louis, MO), 26% glycerol (Fisher, Pittsburgh, PA), 1 mM PMSF (Sigma, St. Louis, MO), 2 mM NaVO₄ (Sigma, St. Louis, MO), complete protease inhibitor cocktail and 78 µL 4.6 M NaCl. The resulting suspension was homogenized in a dounce homogenizer for 20 strokes. The lysate was centrifuged for 20 minutes at 15,000x g at 4°C and the supernatant saved as nuclear fraction. The cytoplasmic and nuclear fractions were acetone precipitated prior to digestion and analysis.

Methods for mass spectrometry and data analysis. The protein pellets were solubilized with 8 M urea/Invitrosol (Invitrogen, Carlsbad, CA), reduced and alkylated, diluted with 4x volumes of 100 mM Tris-HCl and digested with trypsin overnight. After digestion, the pH was adjusted to 2.5 using 90% formic acid. Peptides from the protein digest were analyzed by a 12-step MudPIT experiment as previously described in reference 22. In each duty cycle of mass analysis, one high-resolution (60,000) MS spectrum was acquired using the orbitrap analyzer, followed by six data-dependent MS/MS scans using the linear ion trap analyzer. For MS/MS analysis, normalized collision energy of 35% was used throughout the CID phase.

MS2 and MS3 spectra were analyzed using the following software analysis tools. Both the light and heavy spectra were searched with the ProLucid algorithm²³ against the IPI mouse database (<ftp://ftp.ebi.ac.uk/pub/databases/IPI/>, version 3.52, releasing date November 21, 2008), that was concatenated to a decoy database in which the sequence for each entry in the original database was reversed. The search parameters include a static

cysteine modification of 57.02146. No trypsin specificity was required for all peptides. To search the heavy isotope enriched peptides, the static modification of arginine and lysine was set for a mass addition of 10.00827 and 8.014199, respectively. The database search results were assembled and filtered using the DTASelect program with a spectra level false positive rate of less than 0.5%, mass accuracy of 20 ppm. Only full tryptic peptides were allowed. Under such filtering conditions, the estimated false positive rate was below 2% at the peptide level and below 5% at the protein level. Peptides that passed the filter were quantified using the in-house developed software Census.²⁴

Data from LC-MS/MS contains peptide molecular weight, sequence information and heavy and light ion intensities. Intensity ratios are calculated as $r = I_{\text{light}}/I_{\text{heavy}}$. The sequence information is used to identify the protein or proteins that match the particular sequence. In the case of a peptide matching several potential proteins, the peptide is assigned a unique identifier, which maps to all potential proteins.

Statistical analysis. All peptide intensity ratios were combined into a single dataset. The logarithm of the ratios was computed and was used to construct the total population distribution and from this a mean and variance were calculated. The degrees of freedom were set to the total number of independent peptide measurements. This excludes counting the same peptide for multiple protein assignments. All t-tests and Cohen *d* estimates used this data set as a reference.

For each individual protein, the peptide intensities were pooled and used to determine mean and variance for that particular protein. This mean and variance were tested using the unequal variance t-test versus the total population to determine the p-value. Effect sizes for individual proteins were estimated using Cohen's *d* estimator.

For functional analysis, each protein in our data set was associated with gene ontology terms including inherited terms using release 73 of GOA (Gene Ontology Annotation) mouse from UniProtKB/EMBL-EBI.²⁵ For every term, all the protein evidence associated with a term was pooled into a new data set. From this data set mean and variance were calculated. These were compared to the total data set using two-tail t-test assuming unequal variance to determine the p-value. Effect sizes for the individual terms were estimated using Cohen's *d*.

For significance we have chosen $p < 0.05$ and $f > 0.8$ (Cohen's definition of a large effect size) as our threshold values.²⁶

Acknowledgements

The C3H 10T1/2 tva cells were generated in this lab by Masa Aoki who kindly shared this resource with the authors of this paper. This work was supported by NIH grants R01 CA078230 (J.R.H., L.U., P.K.V.) and P30 NS057096 (L.L., J.R.Y.). This is manuscript 21014 of The Scripps Research Institute.

Note

Supplemental materials can be found at: <http://www.landesbioscience.com/journals/cc/article/15077>

References

1. Rong L, Bates P. Analysis of the subgroup A avian sarcoma and leukosis virus receptor: The 40-residue, cysteine-rich, low-density lipoprotein receptor repeat motif of Tva is sufficient to mediate viral entry. *J Virol* 1995; 69:4847-53.
2. Gilbert JM, Bates P, Varmus HE, White JM. The receptor for the subgroup A avian leukosis-sarcoma viruses binds to subgroup A but not to subgroup C envelope glycoprotein. *J Virol* 1994; 68:5623-8.
3. Oh J, Julius JG, Ferris AL, Hughes SH. Construction and characterization of a replication-competent retroviral shuttle vector plasmid. *J Virol* 2002; 76:1762-8.
4. Kang S, Bader AG, Vogt PK. Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic. *Proc Natl Acad Sci USA* 2005; 102:802-7.
5. Jiang H, Vogt PK. Constitutively active Rheb induces oncogenic transformation. *Oncogene* 2008; 27:5729-40.
6. Hammond SL, Ham RG, Stampfer MR. Serum-free growth of human mammary epithelial cells: Rapid clonal growth in defined medium and extended serial passage with pituitary extract. *Proc Natl Acad Sci USA* 1984; 81:5435-9.
7. Kim DH, Sarbassov DD, Ali SM, King JE, Latek RR, Erdjument-Bromage H, et al. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* 2002; 110:163-75.
8. Shah OJ, Wang Z, Hunter T. Inappropriate activation of the TSC/Rheb/mTOR/S6K cassette induces IRS1/2 depletion, insulin resistance and cell survival deficiencies. *Curr Biol* 2004; 14:1650-6.
9. Harrington LS, Findlay GM, Gray A, Tolacheva T, Wigfield S, Rebholz H, et al. The TSC1-2 tumor suppressor controls insulin-PI3K signaling via regulation of IRS proteins. *J Cell Biol* 2004; 166:213-23.
10. Zhande R, Mitchell JJ, Wu J, Sun XJ. Molecular mechanism of insulin-induced degradation of insulin receptor substrate 1. *Mol Cell Biol* 2002; 22:1016-26.
11. Behrends C, Sowa ME, Gygi SP, Harper JW. Network organization of the human autophagy system. *Nature* 2010; 466:68-76.
12. Chiarle R, Simmons WJ, Cai H, Dhall G, Zamo A, Raz R, et al. Stat3 is required for ALK-mediated lymphomagenesis and provides a possible therapeutic target. *Nat Med* 2005; 11:623-9.
13. Yu H, Jove R. The STATs of cancer—new molecular targets come of age. *Nat Rev Cancer* 2004; 4:97-105.
14. Bromberg JF, Wrzeszczynska MH, Devgan G, Zhao Y, Pestell RG, Albanese C, et al. Stat3 as an oncogene. *Cell* 1999; 98:295-303.
15. Bromberg JF, Horvath CM, Besser D, Lathem WW, Darnell JE Jr. Stat3 activation is required for cellular transformation by v-src. *Mol Cell Biol* 1998; 18:2553-8.
16. Turkson J, Bowman T, Garcia R, Caldenhoven E, De Groot RP, Jove R. Stat3 activation by Src induces specific gene regulation and is required for cell transformation. *Mol Cell Biol* 1998; 18:2545-52.
17. Garcia R, Yu CL, Hudnall A, Catlett R, Nelson KL, Smithgall T, et al. Constitutive activation of Stat3 in fibroblasts transformed by diverse oncoproteins and in breast carcinoma cells. *Cell Growth Differ* 1997; 8:1267-76.
18. Yu CL, Meyer DJ, Campbell GS, Lerner AC, Carter-Su C, Schwartz J, et al. Enhanced DNA-binding activity of a Stat3-related protein in cells transformed by the Src oncoprotein. *Science* 1995; 269:81-3.
19. Pinney DF, Emerson CP Jr. 10T1/2 cells: An in vitro model for molecular genetic analysis of mesodermal determination and differentiation. *Environ Health Perspect* 1989; 80:221-7.
20. Miura Y, Kainuma M, Jiang H, Velasco H, Vogt PK, Hakomori S. Reversion of the Jun-induced oncogenic phenotype by enhanced synthesis of sialosylactosylceramide (GM3 ganglioside). *Proc Natl Acad Sci USA* 2004; 101:16204-9.
21. Bader AG, Kang S, Vogt PK. Cancer-specific mutations in PIK3CA are oncogenic in vivo. *Proc Natl Acad Sci USA* 2006; 103:1475-9.
22. Liao L, McClatchy DB, Park SK, Xu T, Lu B, Yates JR, 3rd. Quantitative analysis of brain nuclear phosphoproteins identifies developmentally regulated phosphorylation events. *J Proteome Res* 2008; 7:4743-55.
23. Xu T, Venable JD, Park SK, Cociorva D, Lu B, Liao L, et al. ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. *Mol Cell Proteomics* 2006; 5:174.
24. Park SK, Venable JD, Xu T, Yates JR, 3rd. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat Methods* 2008; 5:319-22.
25. Barrell D, Dimmer E, Huntley RP, Binns D, O'Donovan C, Apweiler R. The GOA database in 2009—an integrated Gene Ontology Annotation resource. *Nucleic Acids Res* 2009; 37:396-403.
26. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Routledge Academic 1988.