Use of a Label-Free Quantitative Platform Based on MS/MS Average TIC to Calculate Dynamics of Protein Complexes in Insulin Signaling

Xuemei Yang,[1] Adam Friedman,[2] Shailender Nagpal,[3] Norbert Perrimon,[2] and John M. Asara[1,2]

[1]Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215; [2]Harvard Medical School, Boston, Massachusetts 02115; and [3]Innovagene Informatics, Framingham, Massachusetts 02143

A label-free quantification strategy including the development of in-house software (NakedQuant) to calculate the average TIC across all spectral counts in tandem affinity purification (TAP)-tagging liquid chromatography-mass spectrometry MS/MS (LC/MS/MS) experiments was applied to a large-scale study of protein complexes in the MAPK portion of the insulin signaling pathway from *Drosophila* cells. Dynamics were calculated under basal and stimulating conditions as fold changes. These experiments were performed in the context of a core service model with the user performing the TAP immunoprecipitation and the MS core performing the MS and informatics stops. The MS strategy showed excellent coverage of known components in addition to potentially novel interactions.

KEY WORDS: LC/MS/MS, quantification, proteomics, signal transduction, average TIC, spectral counting, protein–protein interaction, networks

Large-scale proteomics experiments similar to those introduced by Gavin et al.[1] are needed to interrogate protein–protein interactions in cellular signaling pathways to uncover new targets for disease and understand biological functions. Tandem affinity purification (TAP)-mass spectrometry (MS) experiments on fifteen bait proteins in the MAPK pathway in *Drosophila* S2R+ cells, with and without insulin and stimulation. Protein differences were quantified by calculating average total ion current (TIC) values from all identified peptide MS/MS spectra (spectral counts)/protein from data-dependent liquid chromatography (LC)/MS/MS runs.[2]

We developed a software suite called NakedQuant v1.0 that uses several features, including protein grouping across biological samples by BLAST, normalization of individual proteins, or entire biological samples, and fold-change calculations. From the output, protein–protein interaction networks were assembled based on the protein signal changes between the basal and stimulated conditions. The network revealed excellent coverage of known bait protein interactions and many novel interactionsin the MAPK signaling pathway. The TAP procedure helps to reduce nonspecific interactions. These data show that novel interactions in signaling pathways through protein–protein interaction studies from immunoprecipitations (IP) of intact protein complexes are effective using simple label-free MS approaches. The project also shows that large-scale experiments are possible within the "core" service model.

ADDRESS CORRESPONDENCE TO: John M. Asara, Beth Israel Deaconess Medical Center, 3 Blackfin Circle, Boston, MA 02115, USA; Phone: 617-735-2651; E-mail: jasara@bidmc.harvard.edu

MATERIALS AND METHODS

TAP Tag and MAPK Bait Protein Preparation (Figure 1)

TAP tags were incorporated into 15 bait proteins of the MAPK branch of the insulin signaling pathway in S2R+ Drosophila cells. For each bait, cell pools were untreated or treated with insulin for 10 min. Cells were then lysed and immunoprecipitated using a IgG-Sepharose column and TEV protease elution, followed by a calmodulin-Sepharose (calcium) column and EGTA elution. Protein complex elutions were cleaned by TCA precipitation, reduced/alkylated with iodoacetamide, and digested with trypsin overnight. The digests were cleaned using C18 ZipTip, SpeedVac-concentrated, and injected into the LC-MS/MS system.

Data Acquisition/Validation (Figure 2)

Data-dependent LC/MS/MS experiments were run using a Proxeon EasynLC at 300 nL/min coupled to a Thermo LTQ-Orbitrap XL for two replicates of each bait condition (basal and insulin-stimulated) for a total of 64 LC/MS/MS experiments including TAP tag controls over a 6-month period. The data were searched against the reversed Fly Base protein database as a result of its completeness using the Sequest algorithm within Proteomics Browser software. Protein sequence lists were validated by setting a 1.5% false

discovery rate threshold for protein identifications based on the number of reversed database hits and requiring at least two unique peptides/protein.

Label-Free Quantification by Average TIC

Validated protein reports were imported into in-house-developed software (NakedQuant v1.0). The software was developed for label-free MS/MS-based quantification. Proteins were grouped across 31 biological samples (15 bait proteins, basal and stimulated, and one TAP control) by BLAST (93% protein identity/78% peptide identity). A baseline of one spectral count and 3.50 × 10E5 TIC was set for proteins not detected in some biological conditions. Proteins were then normalized across the entire biological sample set by setting the total number of spectral counts and/or average TIC to equal values or based on single bait proteins between conditions (Figure 3). Average TIC was calculated using total ion current (TIC) values from all identified MS/MS spectra (spectral counts) per identified protein. Fold changes between the basal and stimulated states for each bait protein were then calculated and displayed in a matrix format. Quantitative interactions were displayed graphically using Cytoscape.

RESULTS

See Figures 4–8.

CONCLUSIONS

Protein–protein interactions (814) were identified from 15 bait TAP-MS experiments under basal and stimulated conditions, representing 526 proteins.

The canonical network and sub-networks were dynamically established with excellent coverage/bait protein in the MAPK insulin signaling pathway.

A dynamic network was established successfully using average TIC from data-dependent LC-MS/MS/MS experiments.

Software was developed to handle large-scale, label-free quantitative projects.

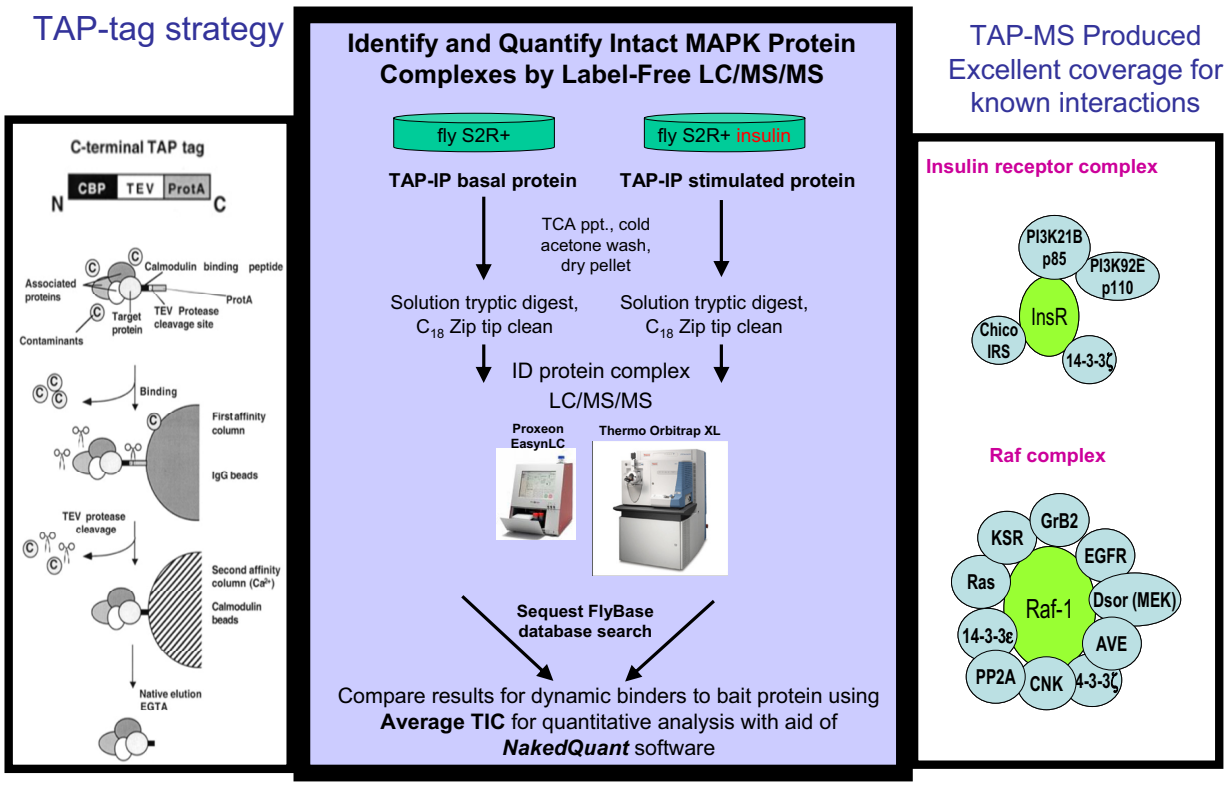
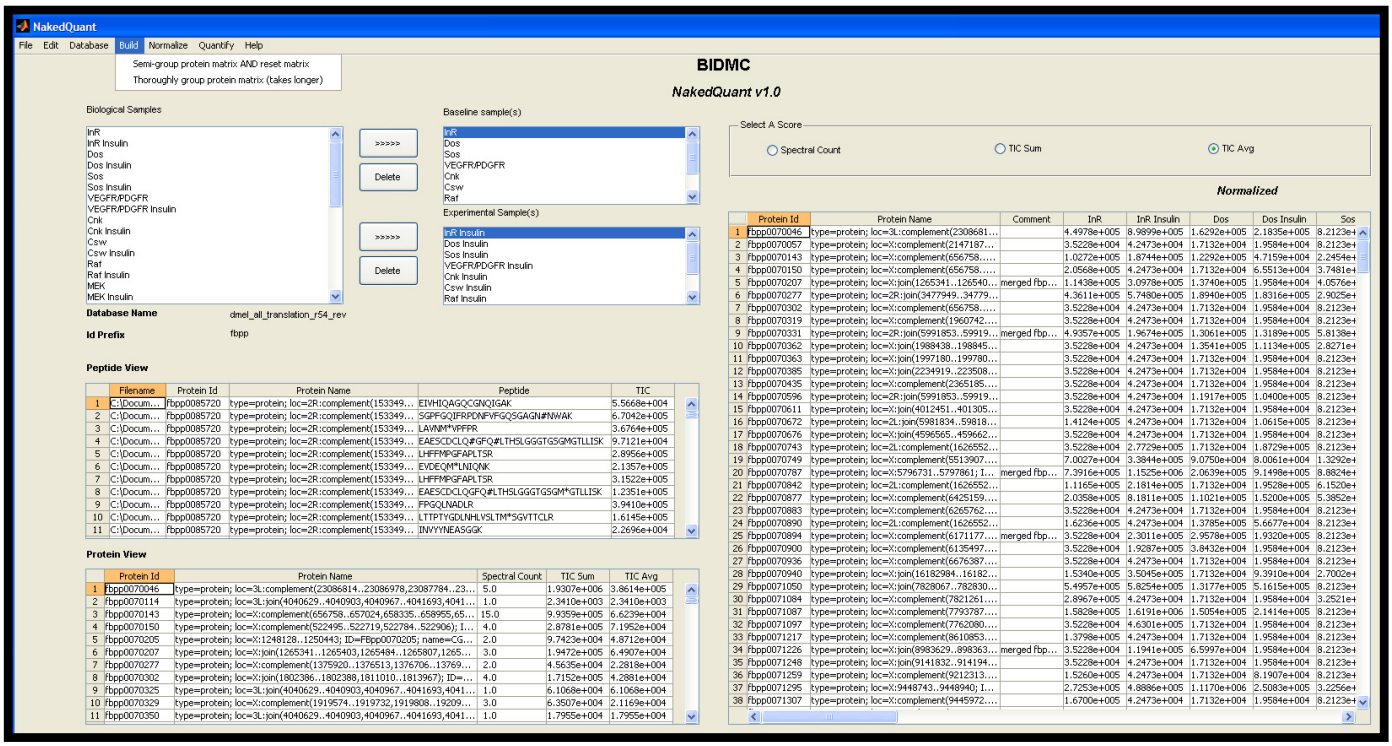


FIGURE 1

Identify and quantify intact MAPK protein complexes by label-free LC/MS/MS. CBP, Calmodulin-binding protein; TEV, Tobacco Etch virus; ProtA, protein A; InsR, insulin receptor; IRS, insulin receptor substrate; ppt, precipitate; ID, identification.



Software was developed called NakedQuant v1.0 that calculates fold changes based on label-free MS/MS data (Spectral counts, Average TIC and Sum TIC)

FIGURE 2

NakedQuant v1.0 label-free quantitative software suite.

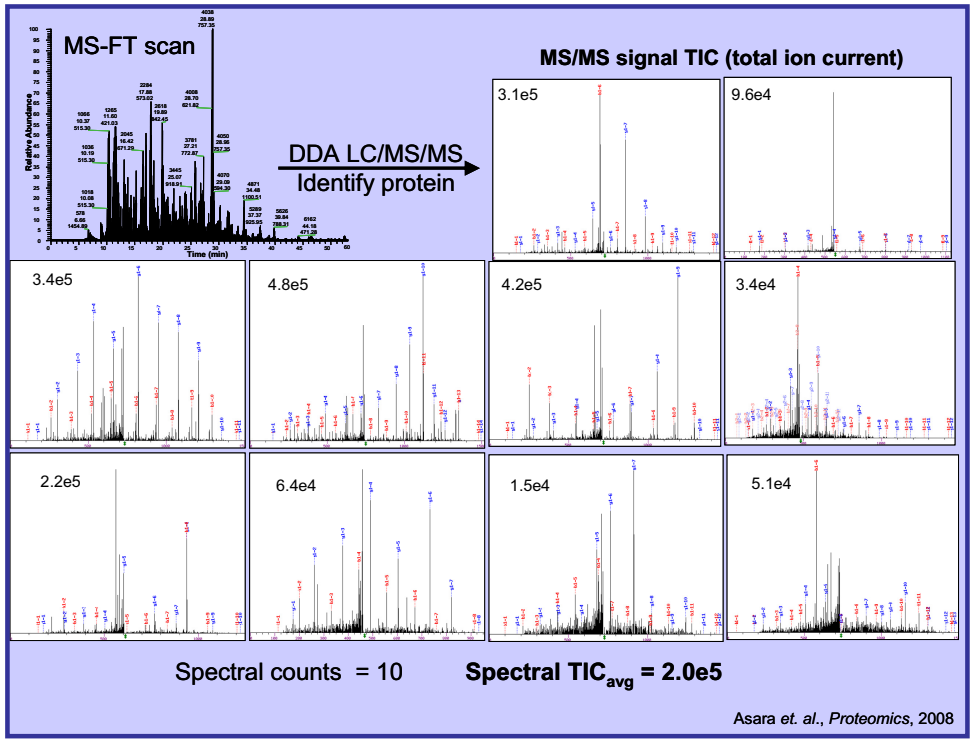


FIGURE 3

Example of average TIC calculation for label-free quantification.[2] FT, Fourier transform; DDA, data-dependent acquisition.

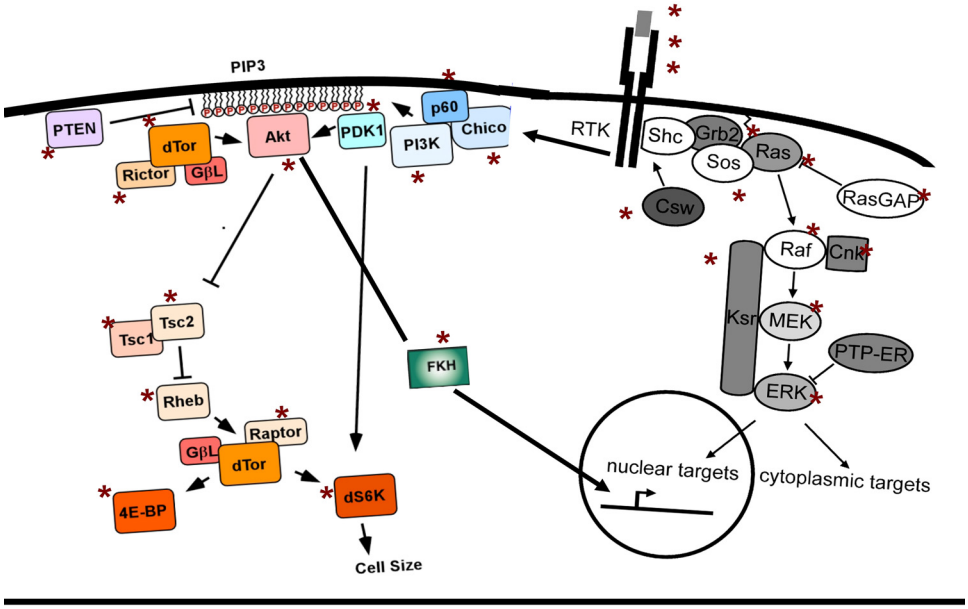


FIGURE 4

The Drosophila insulin signaling pathway. PIP3, Phosphatidylinositol 3,4,5-triphosphate; PTEN, phosphatase and tensin homologue deleted on chromosome 10; dTor, target of rapamycin Drosophila; PDK1, phosphoinositide-dependent kinase-1; FKH, forkhead; RTK, receptor tyryosine kinase; PTP-ER, protein tyrosine phosphatase-ERK/enhancer.

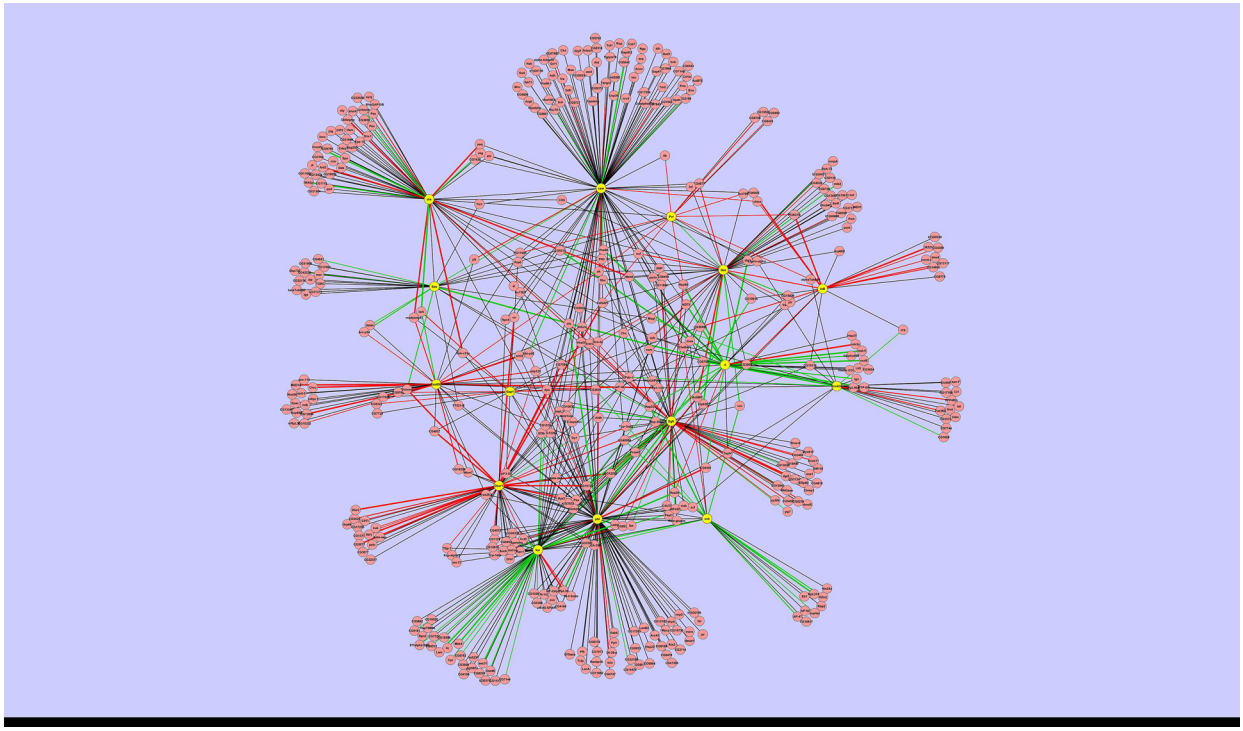


FIGURE 5

LC/MS/MS modeling of the Dynamic MAPK pathway protein–protein interaction network.

Protein Id	Protein Name	Comment	TIC Avg InR	TIC Avg InR Insulin	TIC Avg Dos	TIC Avg Dos Insulin	TIC Avg Sos	TIC Avg Sos Insulin	TIC Avg VEGFR/PDGFR	TIC Avg VEGFR/PDGFR Insulin	TIC Avg Cnk	TIC Avg Cnk Insulin	TIC Avg Fold Change InR Insulin/InR	TIC Avg Fold Change Dos Insulin/Dos	TIC Avg Fold Change Sos Insulin/Sos	TIC Avg Fold Change VEGFR/PDGFR Insulin/VEGFR/PDGFR	TIC Avg Fold Change Cnk Insulin/Cnk
fbpp0082838	type=protein; loc=3R:12968419..12910131; ID=FBpp0082838; name=Huk-PA; parent=FBgn0038476,FBtr0083397; dbxref=GB_protein:AAN13733.1,FlyBase:FBpp0082838,FlyBase_Annotation_IDs:CG5175-PA,GB_protein:AAN13733; MD5=3604b35c28d2e4e24b66258ae5a78685; length=570;		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.00E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	-1.39	-1.251	-2.251
fbpp0082928	type=protein; loc=3R:complement(13362911..13363016,13363078..13365028,13365084..13365586); ID=FBpp0082928; name=l(3)07682-PA; parent=FBgn0010926,FBtr0083504; dbxref=FlyBase:FBpp0082928,FlyBase_Annotation_IDs:CG5824-		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	-1.39	-1.679	-2.251
fbpp0082987	type=protein; loc=3R:join(14068568..14068631,14073155..14073624,14073693..14073873,14073937..14074010); ID=FBpp0082987; name=14-3-3epsilon-PA; parent=FBgn0020238,FBtr0083565; dbxref=FlyBase:FBpp0082987,FlyBase_Annotation_IDs:CG31196-PA,GB_protein:AAF55519	merged fbpp0082988 merged fbpp0082990 merged fbpp0082989	3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	4.99E+05	3.29E+05	1.206	1.143	-1.39	-1.679	-1.518
fbpp0083030	type=protein; loc=3R:complement(14044239..14045945); ID=FBpp0083030; name=CG7183-PA; parent=FBgn0038583,FBtr0083610; dbxref=FlyBase:FBpp0083030,FlyBase_Annotation_IDs:CG7183-PA,GB_protein:AAF55511.1,GB_protein:AAF55511; MD5=2cb5d7e99ef550b0736ac1d00b008c8		3.52E+04	1.26E+05	1.71E+04	7.16E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	3.578	4.181	-1.39	-1.679	-2.251
fbpp0083178	type=protein; loc=3R:join(15290547..15290568,15302845..15302999,15303174..15306638,15306702..15307016,15307079..15307384,15307536..15308038,15308286..15308373,15309626..15310666,15310762..15310994,15311331..15311492,15311554..15311713,15313818..15313913,1		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	-1.39	-1.679	-2.251
fbpp0083227	type=protein; loc=3R:join(15703600..15703633,15703721..15703930,15703999..15704972); ID=FBpp0083227; name=Arc42-PA; parent=FBgn0038742,FBtr0083817; dbxref=FlyBase:FBpp0083227,FlyBase_Annotation_IDs:CG4703-		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	-1.39	-1.679	-2.251
fbpp0083286	type=protein; loc=3R:join(16143992..16144007,16144699..16144752,16145290..16145432,16145491..16145722,16145937..16146173,16147198..16147358,16147424..16147602,16147671..16148864,16148937..16149171); ID=FBpp0083286; name=CG5077-PA; parent=FBgn0038786,FBtr0	merged fbpp0083287	3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	-1.39	-1.679	-2.251
fbpp0083348	type=protein; loc=3R:complement(16455469..16455597,16455661..16455780,16455849..16456019,16456079..16456385,16456442..16457029,16457231..16457437,16457500..16459244); ID=FBpp0083348; name=PI3K92E-PA; parent=FBgn0015279,FBtr0083940; dbxref=FlyBase:FBpp0083		3.52E+04	2.46E+05	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	6.981	1.143	-1.39	-1.679	-2.251
fbpp0083365	type=protein; loc=3R:join(16582959..16582971,16598518..16599604,16617467..16617585,16617659..16617971,16618036..16618277,16619438..16619573,16622866..16622999,16623067..16623195,16623258..16623398,16623469..16623575,16624037..16624131,16631669..16631844);		2.22E+05	4.25E+04	8.16E+04	1.96E+04	4.16E+05	1.82E+05	2.69E+04	1.60E+04	6.61E+04	2.04E+04	-5.233	-4.165	-2.288	-1.679	-3.238
fbpp0083366	type=protein; loc=3R:join(16617469..16617585,16617659..16617971,16618036..16618277,16619438..16619573,16622866..16622999,16623067..16623195,16623258..16623398,16623469..16623575,16624037..16624131,16625739..16626022,16628012..16628273,16631669..16631844);		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.01E+05	2.69E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	6.103	-1.679	-2.251
fbpp0083371	type=protein; loc=3L:join(8067575..8067925,8086159..8086437,8091446..8091825,8092246..8092636,8095993..8097918,8097980..8098159,8098227..8098396,8098540..8098757,8099116..8100049,8101739..8101823); ID=FBpp0111510; name=Ect4-PD; parent=FBgn0085402,FBtr0112		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	3.95E+05	2.58E+05	9.99E+04	2.33E+04	1.04E+04	1.206	1.143	4.805	-2.586	-2.251
fbpp0083415	type=protein; loc=3R:join(16942026..16942355,16942418..16942782,16942839..16943266,16943318..16943577); ID=FBpp0083415; name=ppan-PA; parent=FBgn0010770,FBtr0084012; dbxref=FlyBase:FBpp0083415,FlyBase_Annotation_IDs:CG5786-PA,GB_protein:AAF55851.1,GB_prot		3.52E+04	2.48E+05	1.71E+04	1.10E+05	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	7.046	6.403	-1.39	-1.679	-2.251
fbpp0083441	type=protein; loc=3R:complement(16887948..16889056,16889148..16889551,16889620..16889786); ID=FBpp0083441; name=Cortactin-PA; parent=FBgn0025865,FBtr0084039; dbxref=FlyBase:FBpp0083441,FlyBase_Annotation_IDs:CG3637-		3.52E+04	4.25E+04	1.71E+04	1.96E+04	8.21E+04	5.91E+04	2.69E+04	1.60E+04	2.33E+04	1.04E+04	1.206	1.143	-1.39	-1.679	-2.251

FIGURE 6

Example of grouped proteins TIC average fold-change output from NakedQuant v1.0.

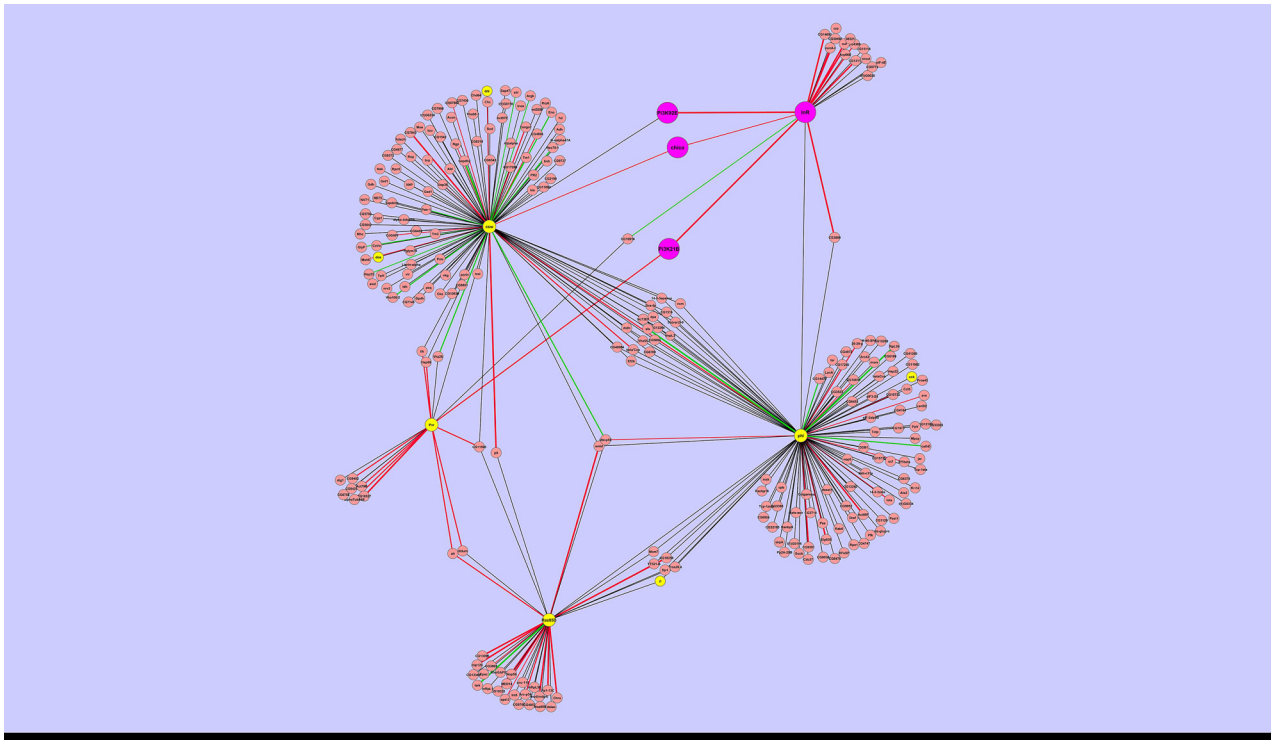


FIGURE 7

The Dynamic PI3K sub-network in MAPK pathway by LC/MS/MS.

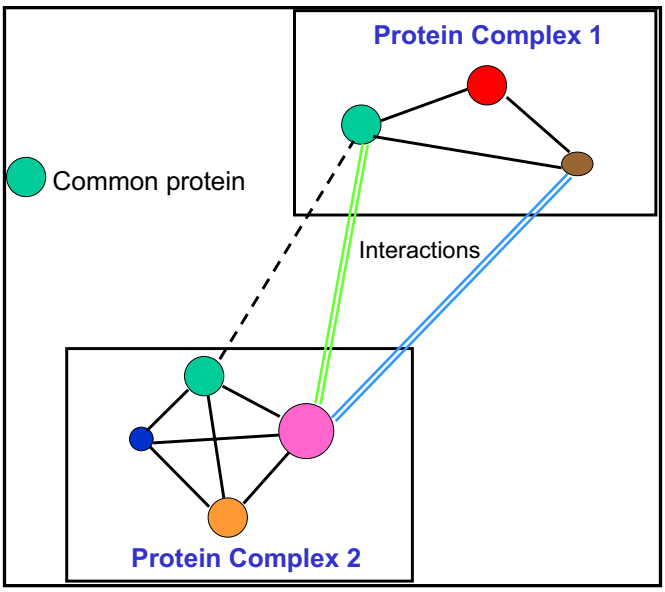


FIGURE 8

IP-MS-based approach for finding protein–protein interactions.

REFERENCES

1. Gavin AC, Bösche M, Krause R, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 2002;10:141–147.
2. Asara JM, Christofk HR, Freimark LM, Cantley LC. A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics* 2008;8:994–999.
3. Friedman A, Perrimon N. A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signaling. *Nature* 2006; 444:230–234.