Supplementary Materials

Quantitative mass spectrometry to interrogate proteomic heterogeneity in metastatic lung adenocarcinoma and validate a novel somatic mutation CDK12-G879V

Xu Zhang[1]; Khoa Dang Nguyen[1]; Paul Rudnick[2]; Nitin Roper[1]; Emily Kawaler[3]; Tapan K. Maity[1]; Shivangi Asasthi[1]; Shaojian Gao[1]; Romi Biswas[1]; Abhilash Venugopalan[1]; Constance M. Cultraro[1]; David Fenyö[3]; Udayan Guha[1*]

[1]Thoracic and GI Malignancies Branch, Center for Cancer Research, NCI, NIH, Bethesda, MD; [2]Spectragen Informatics LLC, Bainbridge Island, WA; [3]Institute for Systems Genetics and Department of Biochemistry and Molecular Pharmacology, NYU School of Medicine, New York, NY.

*Correspondence: udayan.guha@nih.gov

Table S1: Proteins and peptides identified in super-SILAC experiment (Excel document).

Table S2: Proteins, peptides and phosphopeptides identified in TMT experiment searching against human patient specific database (Excel document).

Table S3: Proteins, peptides and phosphopeptides identified in TMT experiment searching against human patient specific database plus mouse database (Excel document).

Table S4: List of proteins identified from both patient specific database and mouse database.

Table S5: List of validated variant peptides.

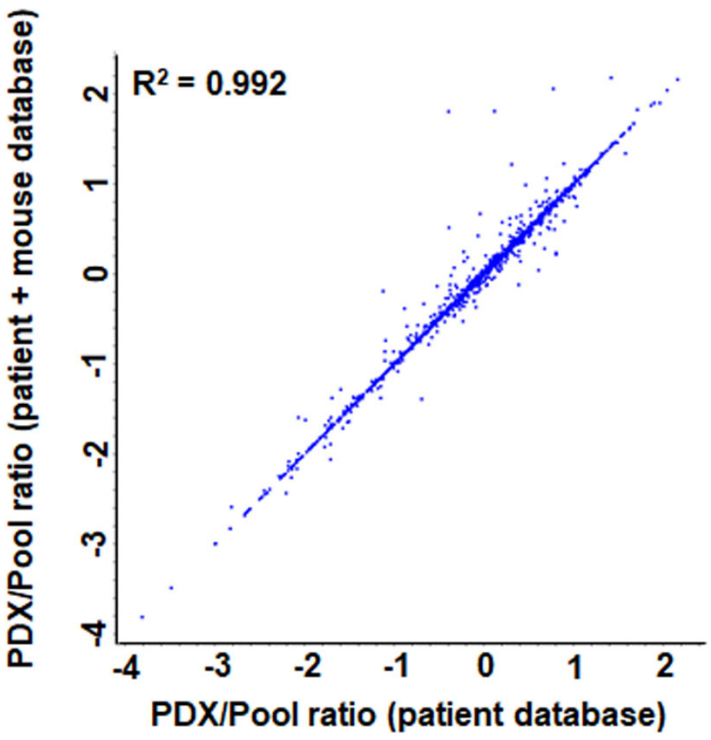

Figure S1: Correlation of the PDX/pool protein ratios from the TMT10plex experiment using patient-specific only and together with mouse database search. X-axis is the ratio from the patient specific database search, Y-axis is the ratio from the combined search of patient specific database and mouse database.

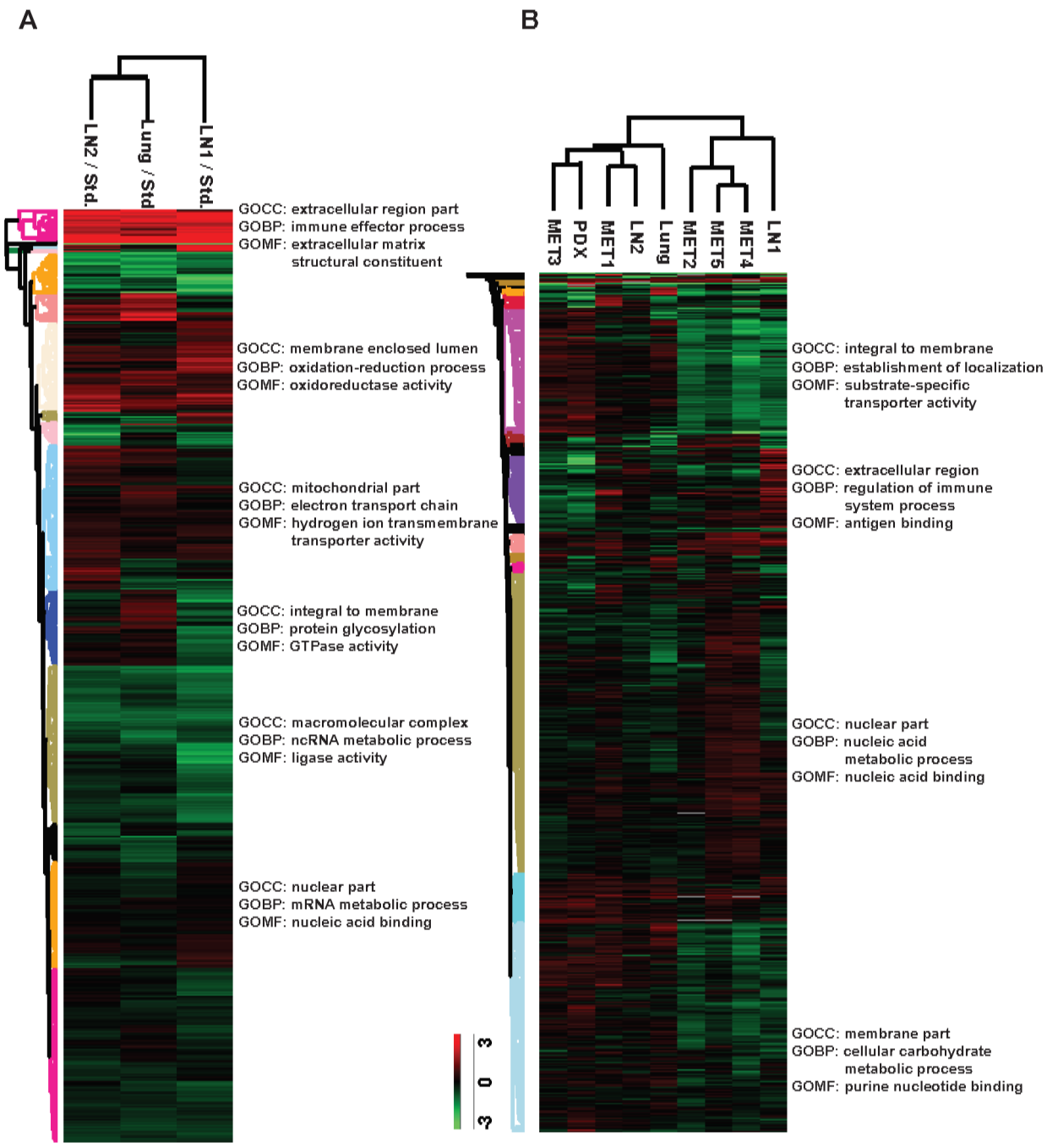

Figure S2. Hierarchical clustering of patient tumor tissue proteome.

(A) Hierarchical clustering of proteins and three biopsy tumor tissues from super-SILAC experiment shows that LN1 is separated from lung and LN2 tissues. (B) Hierarchical clustering of proteins and nine tissues from TMT experiment. Two clusters of tissues are grouped based on the protein expression.

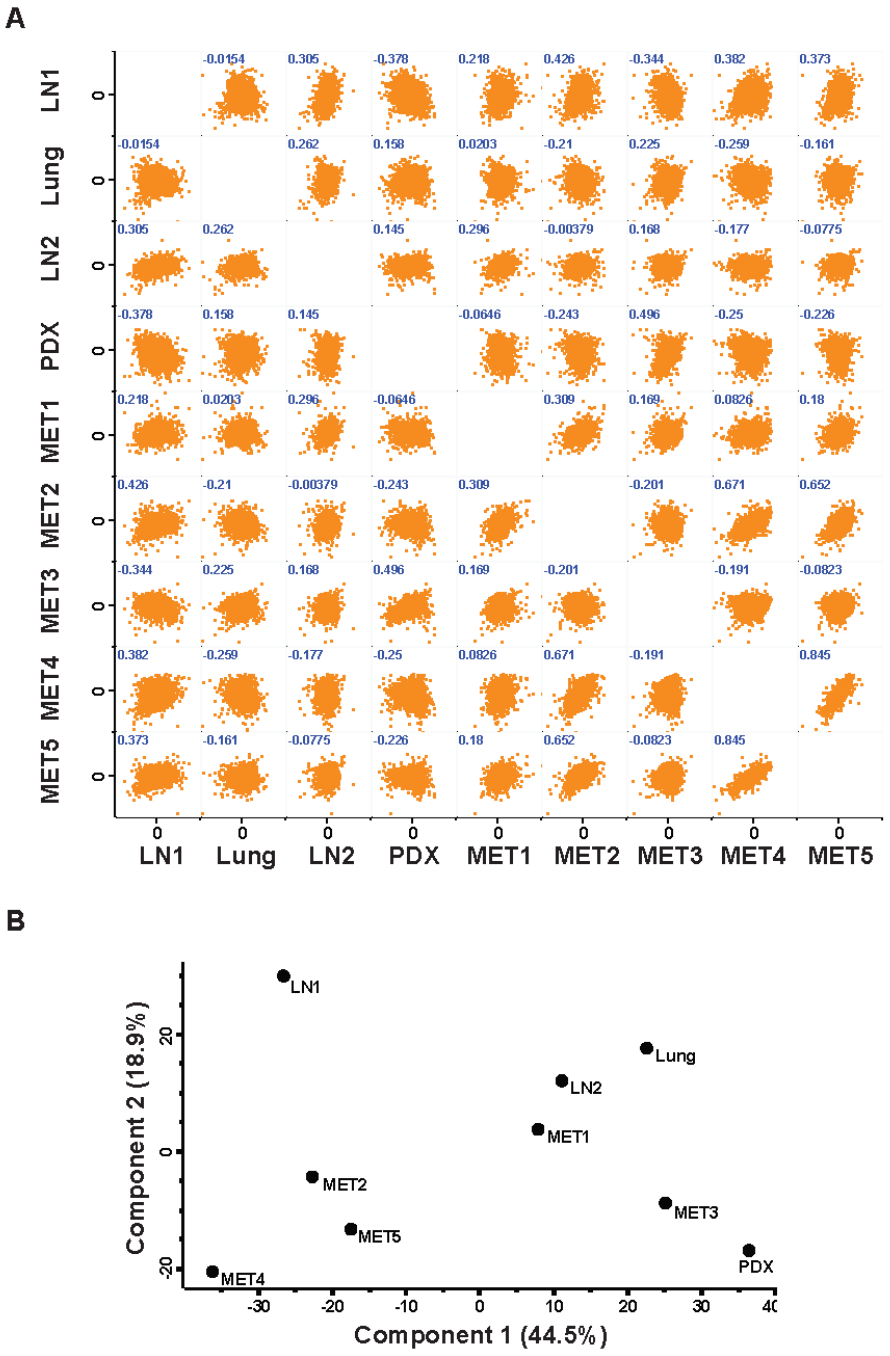

Figure S3. Comparison of lung and lymph node tissue proteome.

(A) Correlation of protein expression between the three tumor biopsies, one PDX, and five autopsy samples from the TMT experiment. (B) Principal component analysis of the tissues based on the protein expression shows two clusters, the same as the hierarchical clusters.

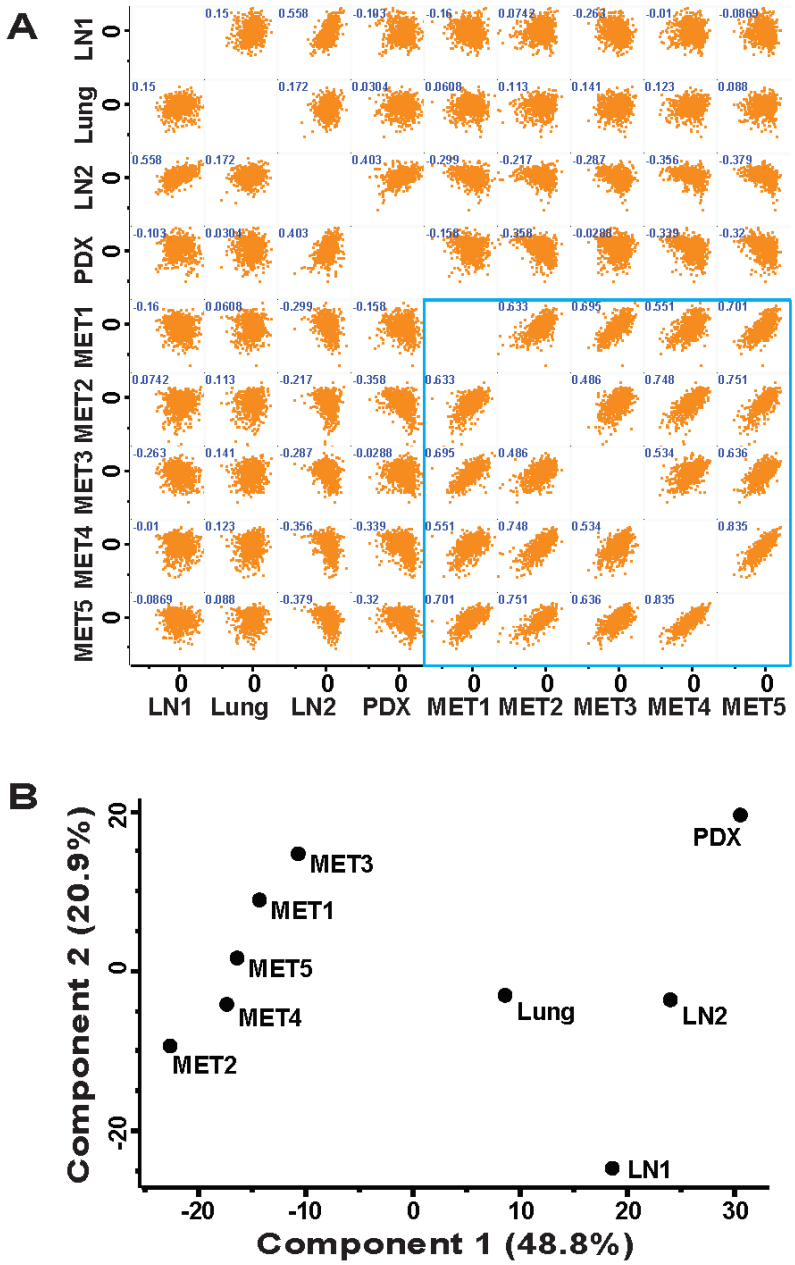

Figure S4. Comparison of phosphoproteome of lung and lymph node tissues.

(A) Correlation of protein phosphorylation between the three tumor biopsies, one PDX, and five autopsy samples from the TMT experiment. The five autopsy samples show higher correlation than the other samples (the blue box). (B) Principal component analysis of the tissues based on the protein phosphorylation differentiates the autopsy tumor tissues from the other tumor biopsies and PDX sample.

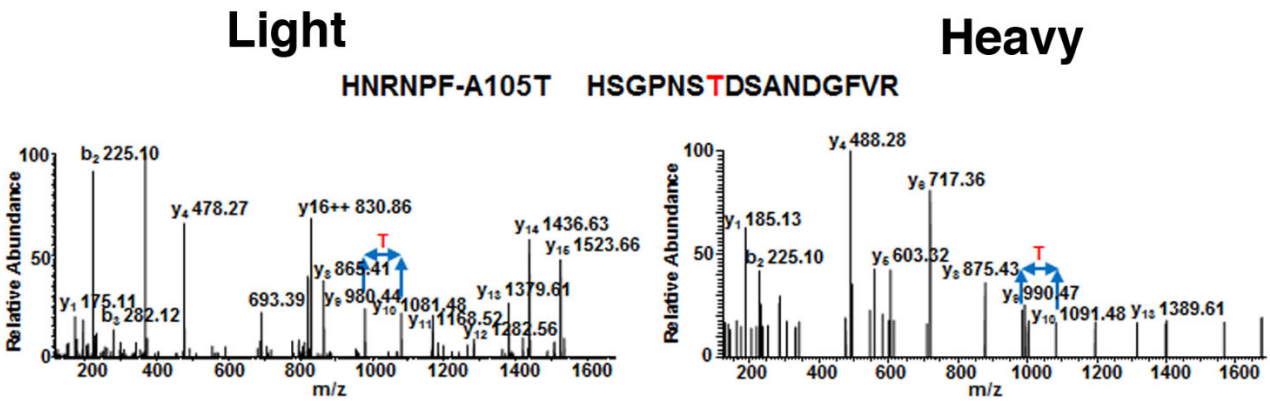

Figure S5. Confirmation of the mutant peptide, HNRNPF-A105T identified from DDA data searched by the patient-specific database. MS/MS spectrum of the light endogenous mutant peptide (left panel) and the heavy amino acid labelled synthesized peptide (right panel) show the presence of the mutation in HNRNPF.

Proteins	Peptides	Transitions
CDK12-G879V	LADFVLAR	791.4410 (y7), 605.3770 (y5), 458.3085 (y4), 720.4039 (y6), 730.4134 (b7)
	LADFVLAR^	730.4122 (y6), 730.4134 (b7), 801.4493 (y7), 468.3168 (y4), 615.3852 (y5),
FASN-R1439Q	GILADEDSSQPVWLK	642.3974 (y5), 1303.6165 (y11), 1374.6536 (y12), 1059.5469 (y9), 944.5200 (y8)
	GILADEDSSQPVWLK^	650.4116 (y5), 1311.6307 9y11), 1067.5611 (y9), 1382.6678 (y12), 952.5342 (y8)
HNRNPF-A105T	HSGPNSTDSANDGFVR	282.1197 (b3), 593.3042 (y5), 707.3471 (y6), 865.4163 (y8), 580.2474 (b6)
	HSGPNSTDSANDGFVR^	717.3554 (y6), 282.1197 (b3), 875.4245 (y8), 580.2474 (b6), 603.3125 (y5)

Figure S6: List of transitions monitored for each mutant peptide.