SUPPLEMENTARY MATERIALS AND METHODS

iTRAQ quantitative proteomics

Reduction, alkylation and trypsin digestion

To identify downstream signals of ZNF322A, proteins were extracted from A549 cells treated with negative control siRNA as control sample, or with ZNF322A-specific siRNA as knockdown ZNF322A sample. We used the method as described in our previous paper.[1]

iTRAQ labeling and LC-MS/MS analysis

We performed iTRAQ labeling and LC-MS/MS analyses which were modified from our previous paper.[1] Total of 300 μg peptides for each sample were used for iTRAQ labeling (AB SCIEX). Peptides were mixed and dried with a centrifugal evaporator (CVE-2000). iTRAQ-labeling peptides were dried and then fractionated in a Waters e2695 Alliance HPLC system (Waters Corporation) using a 2.1 × 200-mm polysulfoethyl A column containing 5 μm particles with 200-m pore size (PolyLC). Each fraction was dried and then dissolved in 0.1% TFA for desalting using ZipTips™ (Millipore).

For mass spectrometry of iTRAQ-labeled peptides, each sample was reconstituted in 24 μl buffer A [0.1% (v/v) formic acid (FA) in H_2O] and analyzed by LC-ESI-Q-TOF mass spectrometry (Waters SYNAPT® G2 HDMS; Waters Corp.). Samples were injected into a 180 μm × 2 cm capillary trap column and separated by a 75 μm × 25 cm nanoACQUITY UPLC™ 1.7 μm Ethylene Bridged Hybrid (BEH) C18 column using a nanoACQUITY Ultra Performance LC™ System (Waters Corp.).

NanoLockSpray™ source (Waters Corp.) is used for accurate mass measurements, and the lock mass channel was sampled every 30 seconds. The mass spectrometer was calibrated with a synthetic human [Glu^{1}]-Fibrinopeptide B solution (1 pmol/μl; Sigma-Aldrich) delivered through the NanoLockSpray source. Data acquisition was operated in the data directed analysis (DDA). The DDA method included one full MS scan (*m/z* 350-1700, 1 second) and three MS/MS scans (*m/z* 100-1990. 1.5 seconds for each scan) sequentially on the three most intense ions present in the full scan mass spectrum. Each sample was analyzed in duplicate.

<u>Data processing and analysis</u>

The peak lists of MS/MS spectra were generated by Mascot Distiller v2.3.2 (Matrix Science) and searched against the sequence database (containing 533,657 sequence entries) of the Swiss-Prot human database using Mascot search engine v2.3.02 (Matrix Science). The tolerance of precursor peptide was ±0.3. The tolerance of fragment ion was ±0.2. Search parameters were as follows: two missed cleavages made from the trypsin digestion, deamidated (NQ), oxidation (M), iTRAQ4plex (K), iTRAQ4-plex (N-term), and methylthio (C). Peptide charge was set to the relative molecular mass (Mr), instrument was set to ESI-QUAD-TOF and decoy database was searched. Mascot search results and ions score were filtered using an expectation value ($P<0.05$). We performed a decoy database search against a randomized decoy database created to evaluate the false discovery rate (FDR) by Mascot using identical search parameters and validation criteria. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium[2] via the PRIDE partner repository with the data set identifier PXD001123 and DOI 10.6019/PXD001123.

Signature ions (m/z = 114, 115, 116 and 117) and peptides were analyzed using Multi-Q software (v1.6.5.4).[3] Before quantitation of the expression of each protein, the peak intensity of the iTRAQ signature ion was normalized as Method 1 of our previous study.[1] For determining expression ratio of identified proteins from A549 cells with control siRNA and si-*ZNF322A*, control samples were equally divided as C1 (iTRAQ 114-labeled) and C2 (iTRAQ 115-labeled), whereas si-*ZNF322A* samples were equally divided as T1 (iTRAQ 116-labeled) and T2 (iTRAQ 117-labeled). Calculations of *S*-value, *R*-value and $\pm 2\sigma_S$ (±2SD of *S*-values) were used to identify significantly regulated proteins as previously described.[1]

Calculation of *S*-value: $S = \frac{1}{2}\log_2\left(\frac{T_1 \cdot C_2}{T_2 \cdot C_1}\right) = \frac{1}{2}\left[\log_2\left(\frac{T_1}{C_1}\right) - \log_2\left(\frac{T_2}{C_2}\right)\right]$

Calculation of *R*-value: $R = \frac{1}{2}\log_2\left(\frac{T_1 \cdot T_2}{C_1 \cdot C_2}\right) = \frac{1}{2}\left[\log_2\left(\frac{T_1}{C_1}\right) + \log_2\left(\frac{T_2}{C_2}\right)\right]$

Functional enrichment analysis

Gene set analysis[4] was used to enrich the functions associated with the iTRAQ-identified protein list (see the complete list of total 1,108 proteins in Supplementary Table S2). The gene sets were prepared from biological process category of GO database. If the size of set is less than 10 or greater than 500, this gene set was discarded. The R value derived from iTRAQ experiment was normalized into a range of -1 and 1 and used as the abundance metric to weight proteins in running sum statistics. The enrichment score (ES) of a gene set *S* was defined as:

$$ES = \sum_{i=1}^{|S|} r_i^2$$

where r_i is the normalized R value of a proteins in gene set S. The statistical significance for association of S and protein expression is obtained by comparing the

observed value of ES and its permutation distribution which run 10,000 times gene set permutation. A relation map among these gene sets with p-value < 0.05 was generated by EnrichmentMap[5] with similarity score > 0.5 using combination of Jaccard and overlap coefficients and visualized by Cytoscape[6].

We also performed the Ingenuity Pathway Analysis (IPA, Ingenuity Systems) database (application build version 171496, and content version 14197757) to analyze the associated networks or diseases and disorders of the 64 significantly differential expressed proteins. A score greater than or equal to 3 was considered statistically significant ($P < 0.01$).

Flow cytometry

Approximately 4×10^5 cells were grown in 6 cm dishes. Cells were collected, washed once with PBS, fixed with ice-cold 90% ethanol then collected by centrifugation. Pellets were resuspended in 1 ml of PBS mixture [20 μg/ml propidium iodide (Sigma-Aldrich), 200 μg/ml RNase A, and 1 μl Triton X-100] and then incubated at 37°C in the dark for 15 minutes. Determination of cell cycle distribution was performed by FACScan Flow cytometer (BD, Franklin Lakes, NJ, USA) and calculated using ModFIT LT 2.0 version software (BD).

Immunoprecipitation assay

For immunoprecipitation, 600 μg cell lysates were incubated with 20 μl Protein G/Protein A agarose beads (IP05; Calbiochem Co.) and 4 μg of the appropriate antibody, including ZNF322A, HA, c-Jun, Rabbit-IgG, and Mouse-IgG at 4°C for 2 hours. Complexes were then washed 3 times with 1× immonoprecepitation buffer (50

mM Tris-HCl, pH 7.5, 150 mM NaCl, 20 mM α-glycerol-phosphate, 1% NP-40, 5 mM EDTA). Proteins were eluted by boiling in 2× SDS loading buffer, separated by 8% SDS-PAGE, then blotted with HA or c-Jun antibody.

Mutation analysis of *p53* gene

Tumors were analyzed for DNA sequence alterations in exons 5 to 11 of *p53* gene as described previously[7]. PCR fragments were sequenced using ABI 377 automatic sequencer (PE Applied Biosystems). Sequencing analysis of the opposite strand for all of the mutations were repeated at least once using independent PCR products and primers listed in Supplementary Table S3.

References

1. Wu YH, Hu CW, Chien CW, Chen YJ, Huang HC, Juan HF. Quantitative proteomic analysis of human lung tumor xenografts treated with the ectopic ATP synthase inhibitor citreoviridin. *PLoS One* 2013; **8**: e70642.
2. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D *et al*. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* 2014; **32**: 223-226.
3. Lin WT, Hung WN, Yian YH, Wu KP, Han CL, Chen YR *et al*. Multi-Q: a fully automated tool for multiplexed protein quantitation. *J Proteome Res* 2006; **5**: 2328-2338.
4. Dinu I, Potter JD, Mueller T, Liu Q, Adewale AJ, Jhangri GS *et al*. Improving gene set analysis of microarray data by SAM-GS. *BMC Bioinformatics* 2007; **8**: 242.

5. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 2010; **5**: e13984.
6. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011; **27**: 431-432.
7. Wang YC, Chen CY, Chen SK, Cherng SH, Ho WL, Lee H. High frequency of deletion mutations in p53 gene from squamous cell lung cancer patients in Taiwan. *Cancer Res* 1998; **58**: 328-333.