Table S1. Contents of mappingout.txt

Column	Name	Comment
1	Peptide sequence	
2	Peptide numbering	Ordered by peptide's genomic coordinates
3	Peptide length	
4	Gene symbol	HGNC symbol
5	Protein accession	
6	PSM count	
7	Peptide relative intensity	Mean of PSMs relative intensity for this peptide.
8	Standard deviation	Standard deviation of PSMs relative intensity for this peptide.
9	Cluster	Formed by peptide level quantitative analysis
10	Detected clusters NO.	The number of clusters formed for this gene
11	Variants NO.	Number of splice variants of this gene in EVDB/ECgene Database
12	NOMV	Number of variants the peptide mapped to
13	Accession id of variants	Accession id of known variants the peptide mapped to
14	Chromosome	
15	Chr_start	Genomic start coordinates of this peptide
16	Chr_end	Genomic end coordinates of this peptide
17	Trans_start	Transcriptional start position of this peptide
18	Trans_end	Transcriptional end position of this peptide
19	Strand	"+" or "-"
20	Exon1	The exon which the first amino acid of this peptide mapped to
21	Exon2	The exon which the last amino acid of this peptide mapped to

Note: column 9,10 will be 0 if PQPQ is not used. For novel splice junction peptides, Trans_start and Trans_end will be length of coding sequence for this peptide localized on the exon1 and exon2.

Table S2. Contents of genestatistic.txt

Column	Name	Description
1	Gene symbol	
2	Identified clusters NO.	Number of clusters formed in peptide level analysis
3	Known splice variants NO.	
4	PSM count	Number of PSMs
5	Unique peptide	Number of unique peptides identified for this gene
6	Identified variants	Number of identified splice variants
7	Variants ID	Accession number of identified splice variants

Note: column 2 will be 0 if PQPQ is not used. Column 6 and 7 will only report splice variants identified from genes that have two or more known splice variants.

Table S3. Comparison of splice variant and SVSPs identification between public domain dataset (4Skin iPS cells mixed with hESCs and fibroblast cells) and our A431 cell line dataset

	P_Mix1	P_Mix2	A_whole	A_heavy	A_medium	A_light
Genes input to SpliceVista	5441	5543	6945	7260	7244	5900
Unique peptides input to SpliceVist	27374	28948	32969	36077	36417	28726
Average sequence coverage	14.80%	15.24%	14.24%	14.06%	14.78%	14.87%
splice variant identified by SpliceVista	390	397	607	672	681	512
Splice variant specific peptides	1287	1392	1680	1963	2027	1309

Note: public domain dataset starts with prefix: P; A431 cell line dataset starts with prefix :A

iTRAQ	113	114	115	116	117	118	119	121
	0h	0h	2h	2h	6h	6h	24h	24h
Whole cell	x	x	x	x	x	x	x	x
Light	x	x	x	x	x	x	x	x
Medium	x	x	x	x	x	x	x	x
Heavy	x	x	x	x	x	x	x	x

Fig.S1. Experimental setup for A431 cell line time course data. A431 human cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% Fetal Bovine Serum, 1% Penicillin/Streptomycin, non-essential amino acids, and sodium pyruvate, all from Invitrogen. Twenty-four hours after seeding, A431 cell line cultures (in duplicates) were treated with Gefitinib and harvested at 2h, 6h and 24h after treatment by trypsination and washed in PBS. Controls were left untreated (duplicates at 0h). Protein samples were further extracted from three subcellular fractions (light, medium, heavy) and whole cell and then digested by trypsin. Trypsin digested peptides from the four different time points and cellular fractions were labeled by 8-plex iTRAQ. The following is subcellular fractionation procedures. For whole cell protein extraction, cells were lysed using buffer A (10mM Hepes pH 7.5, 10mM KCl, 1.5mM MgCl, HALT protease/phosphatase inhibitors, from Pierce/Thermo Scientific) supplemented with 2% SDS. After vortexing, samples were heated to 95°C for 5min and sonicated with a probe sonicator (Bandelin Sonopuls, Buch and Holm) using 2 x 15s, 50% duty cycle, 50% power. For subcellular fractionation cells were re-suspended in buffer A and incubated on ice for 20min. Cells were then disrupted by 30 strokes with a Kontes douncer (2ml, pestle B small clearance 0.0005-0.0025 inch, Kimbel Chase/Sigma-Aldrich). After centrifugation at 600 g (5min, 4°C), the pellet (Heavy fraction, corresponding roughly to plasma membrane and nuclear fraction) was washed once in buffer A, and proteins from the pellet were extracted by the method used for whole cell extraction. The supernatant was centrifuged for pelleting of cytoplasmic organelles (Medium fraction, 100,000 g, 1h, 4°C). The supernatant from the ultra-centrifugation was supplemented with 1% SDS and saved as the soluble fraction (Light fraction). Proteins from the pellet were extracted by the method used for whole cell extraction. A431 protein extracts from whole cell and from the three subcellular fractions were processed following a slightly modified FASP protocol. The resulting peptide mixtures were arranged into four sets (whole, light, medium and heavy) and peptides at different time points in each set were labeled with iTRAQ8plex (ABsciex). After pooling, all iTRAQ8-labeled peptide sets were cleaned by strong cation exchange solid phase extraction (SCX-SPE, strata-X-C, Phenomenex), and aliquoted for IEF.

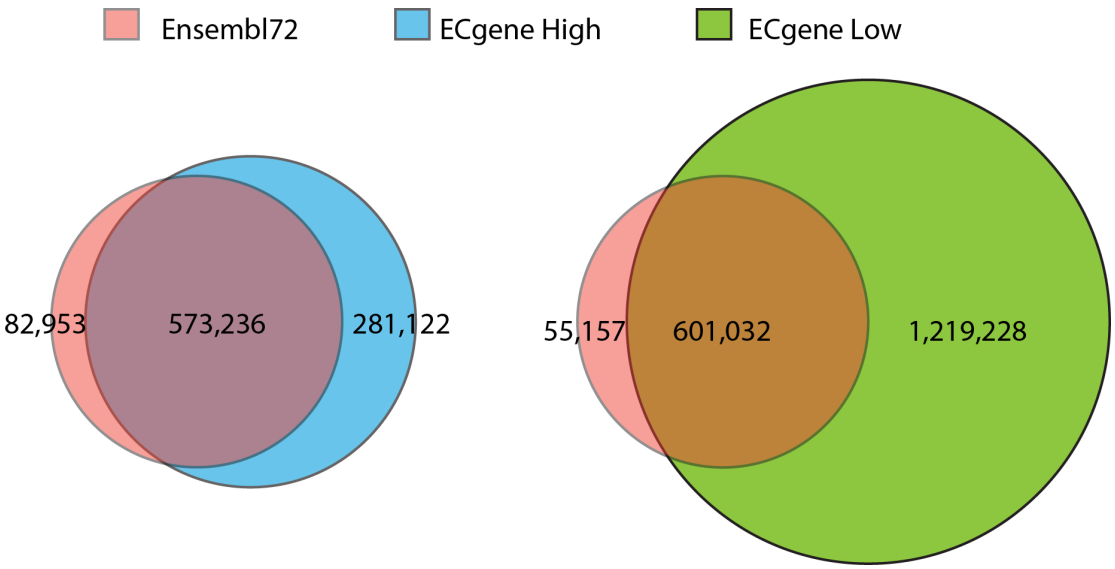

Fig.S2. Overlap of trypsin digested peptides (6aa-40aa long) from Ensembl 72 human protein database, and ECgene human splice variant database at high and low evidence level.

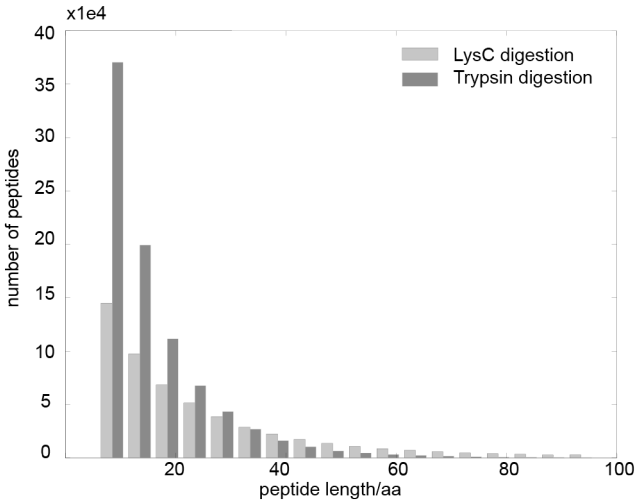

Fig.S3. peptide length distribution using LysC or Trypsin digestion. Only peptides >= 6aa were counted. By trypsin digestion, 96% peptides are between 6aa and 40 aa long, while by lysC digestion, 82% peptides are between 6aa and 40 aa long.

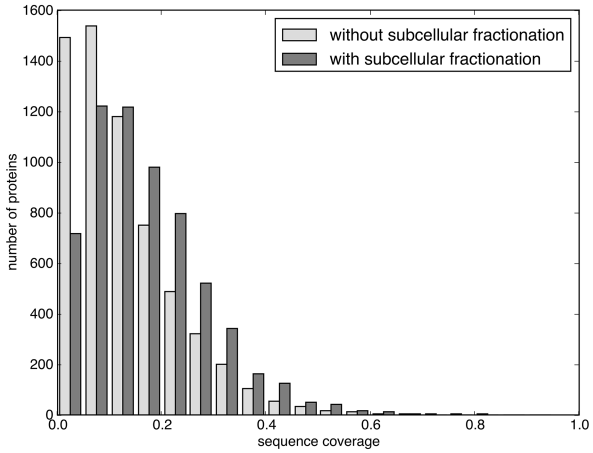

Fig.S4. Distribution of protein sequence coverage for proteins identified with and without subcellular fractionation. 6211 proteins that are both identified with and without subcellular fractionation are included in this plot.

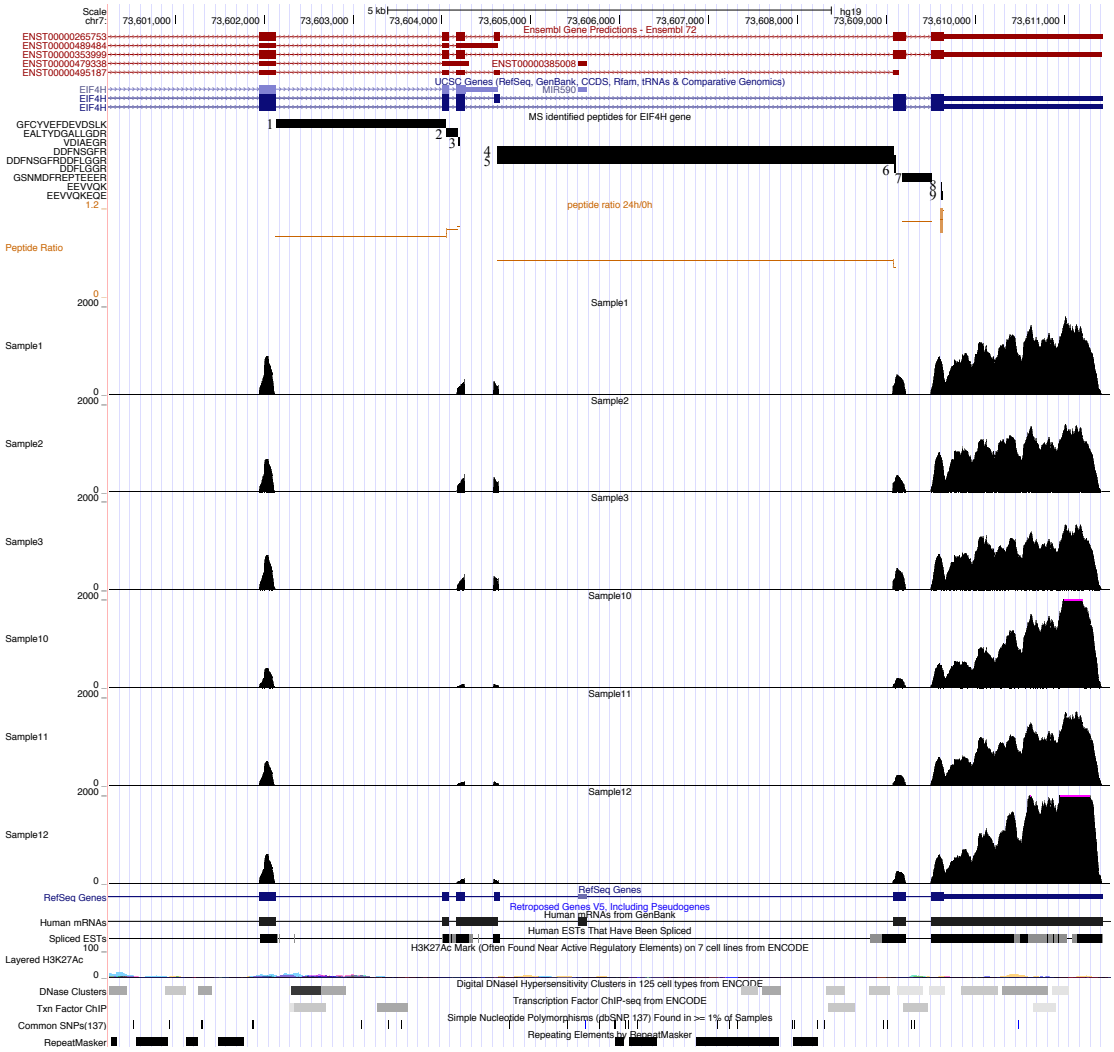

Fig.S5. Comparison of the same samples between protein level data (MS based proteomics) and RNA level data (RNA-seq) by UCSC genome browser. Black boxes are peptides, in total nine peptides were identified for EIF4H by MS. The start and end position of a black box indicates the genomic position of the first and last amino acid of this peptide, the box length is not scaled to peptide length. There are five transcripts in the Ensembl gene track (red), two of them (on the first and third row with thicker exon boxes) are protein-coding transcripts. The orange lines indicate the peptide ratio (24h/0h) between time points 24h and 0h after Gefitinib treatment. If two peptides (peptide 4 and 5) overlap in genomic region, the mean of 24h/0h ratio of the two peptides is shown as peptide ratio of that region. Splice variant specific peptide 4 and 5 (uniquely map to ENST00000265753) showed significant down-regulation at 24 h after Gefitinib treatment (see Fig. 7). Sample1, 2 and 3 are read counts from RNA sequencing of biological triplicates of A431 cell line at 0h (no drug treatment); sample10, 11 and 12 are read counts from RNA sequencing of biological triplicates of A431 cell line at 24h after Gefitinib treatment. All samples were plotted in the same scale (read count: 0-2000). The down-regulation of peptide 4, 5 and 6 identified at protein level by MS was also observed at RNA level, shown as a decrease of read count in corresponding genomic region. UCSC custom track files used to generate this figure are pasted in the end of this file.

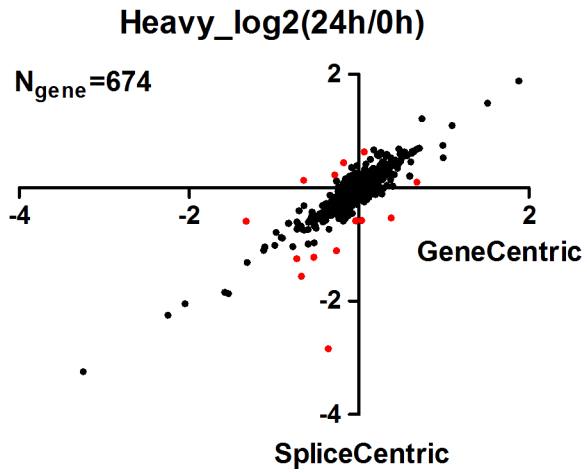

Fig.S6. Comparison of 24h/0h log2 fold change in heavy cell fraction between spliceCentric and geneCentric analysis. 674 genes with SVSPs identified were included. Red dots are genes which showed log2(24h/0h) differences between gene centric and splice variant centric analysis larger than 0.5.

UCSC custom track files used to generate Fig.S3: EIF4H_bed.txt, EIF4H_bedgraph.txt.

EIF4H_bed.txt

browser position chr7:73602128-73609635

track name=Peptide description="MS identified peptides for EIF4H gene"

chr7	73609609	73609626	EEVVQK
chr7	73609079	73609099	DDFLGGR
chr7	73604621	73609078	DDFNSGFR
chr7	73609175	73609506	GSNMDFREPTEEER
chr7	73602128	73604043	GFCYVEFDEVDSLK
chr7	73604044	73604166	EALTYDGALLGDR
chr7	73609609	73609635	EEVVQKEQE
chr7	73604176	73604196	VDIAEGR
chr7	73604621	73609099	DDFNSGFRDDFLGGR

EIF4H_bedgraph.txt

browser position chr7:73602128-73609635

browser full altGraph

track type=bedGraph name="Peptide Ratio" description="peptide ratio 24h/0h" visibility=full color=200,100,0 altColor=0,100,200 priority=20

chr7	73609609	73609626	0.9125
chr7	73609079	73609099	0.4063
chr7	73604621	73609078	0.4845
chr7	73609175	73609506	1.008
chr7	73602128	73604043	0.8105
chr7	73604044	73604166	0.9045
chr7	73609609	73609635	1.1695
chr7	73604176	73604196	0.9535