

Supplementary information – table of content

A) 8 Supplementary figures

Figure S1. Details for the RNA-seq data.

Figure S2. Summary of detected transcript and proteins in the cell line A-431.

Figure S3. Summary of detected transcript and proteins in the cell line U-251MG.

Figure S4. Overlap of detected gene products by the three methods in the cell lines A-431 and U-2OS.

Figure S5. Comparison of the gene products detected by the three methods RNA-Seq, MS and confocal IF in the different categories of the Venn diagram in Figure 3B.

Figure S6. Correlation of transcript and protein level changes between cell lines.

Figure S7. Randomized correlation of transcript and protein ratios between the three cell lines.

Figure S8. Correlation between transcript and protein levels.

21 Supplementary tables

Supplementary Table 1. The number of analyzed and expressed genes for each platform in the three cell lines.

Supplementary Table 2. The number of genes from various protein classes present in each of the cell lines based on MS intensities (log2).

Supplementary Table 3. The number of genes from various protein classes present in each of the cell lines based on RPKM values (log2).

Supplementary Table 4. Gene Ontology (GO)-based enrichment analysis of the proteins in the “Similar”-group in Figure 1F.

Supplementary Table 5. GO-based enrichment analysis of the proteins in the “Slightly changed”-group in Figure 1F.

Supplementary Table 6. GO-based enrichment analysis of the proteins in the “Substantially changed”-group in Figure 1F.

Supplementary Table 7. GO-based enrichment analysis of the genes from the “Similar” category of Figure 2E.

Supplementary Table 8. GO-based enrichment analysis of the genes from the “Slightly changed” category of Figure 2E.

Supplementary Table 9. GO-based enrichment analysis of the genes from the “Substantially changed” category of Figure 2E.

Supplementary Table 10. GO-based enrichment analysis of the genes from the “Cell-type specific” category of Figure 2E.

Supplementary Table 11. GO-based enrichment analysis of the genes from the “Not expressed” category of Figure 2E.

Supplementary Table 12. GO-based enrichment analysis of the genes expressed exclusively in U-2 OS on the transcript level.

Supplementary Table 13. GO-based enrichment analysis of the genes expressed exclusively in U-251 MG on the transcript level.

Supplementary Table 14. GO-based enrichment analysis of the genes expressed exclusively in A-431 on the transcript level.

Supplementary Table 15. The mRNA abundance in cell line A-431 for the categories defined in Figure 2E.

Supplementary Table 16. The mRNA abundance in cell line U-251 MG for the categories defined in Figure 2E.

Supplementary Table 17. GO-based enrichment analysis of the “grey” category of Figure 3C.

Supplementary Table 18. GO-based enrichment analysis of the “blue” category of Figure 3C.

Supplementary Table 19. GO-based enrichment analysis of the “yellow” category of Figure 3C.

Supplementary Table 20. GO-based enrichment analysis of the “green” category of Figure 3C.

Supplementary Table 21. GO-based enrichment analysis of the “red” category of Figure 3C.

Supplementary references

2 supplementary references

Supplementary datasets

Supplementary dataset 1. Protein expression for the proteins analyzed by IF in the three cell lines.

Supplementary dataset 2. MS intensity scores and SILAC ratios for the proteins detected by MS in the three cell lines.

Supplementary dataset 3. RPKM values and RPKM ratios for the proteins detected by RNA-seq in the three cell lines.

Supplementary Figures and Legends

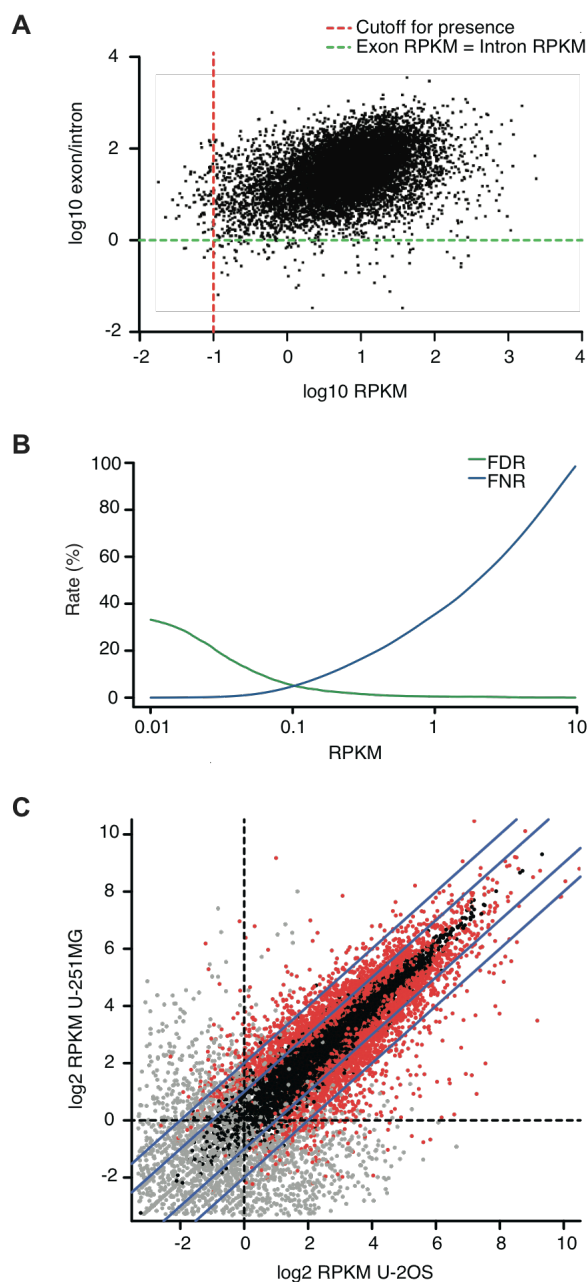


Figure S1. Details for the RNA-seq data. The raw RNA-Seq data was mapped with BWA³¹ allowing a maximum of five mismatches, and only unique reads were retained for further analysis. RPKM values were calculated for each Ensembl gene as previously described⁵. **A.** Scatter plot of exon/intron RPKM ratio versus exon RPKM for 9,630 protein coding genes where both exons and introns were detected. Most genes have a ratio of >10 indicating that the RNA-seq method interrogates mature transcripts. **B.** In order to identify a relevant threshold value for detecting present genes, the approach described in Ramsköld et al¹⁰ was used to estimate false positive (FDR) and false negative rates (FNR) for different threshold values. A RPKM threshold value of 0.1 was set to detect presence, which corresponds to a false

discovery rate and false negative rate of 5%. C. Scatter plot of log₂ RPKM values in U-2 OS versus U-251 MG. For each gene, a Z-test for fractions was used to compare the read count proportions between the two cell lines. Genes with a significant difference ($p < 0.01$) after using FDR for multiple testing correction, are shown in red and those considered not significant are shown in black. The remaining genes, for which less than 10 reads were detected and hence the test could not be performed, are shown in grey. The blue lines display a fold change of two and four between the two cell lines, which are the underlying limits used for categorizing the genes in Figure 2E.

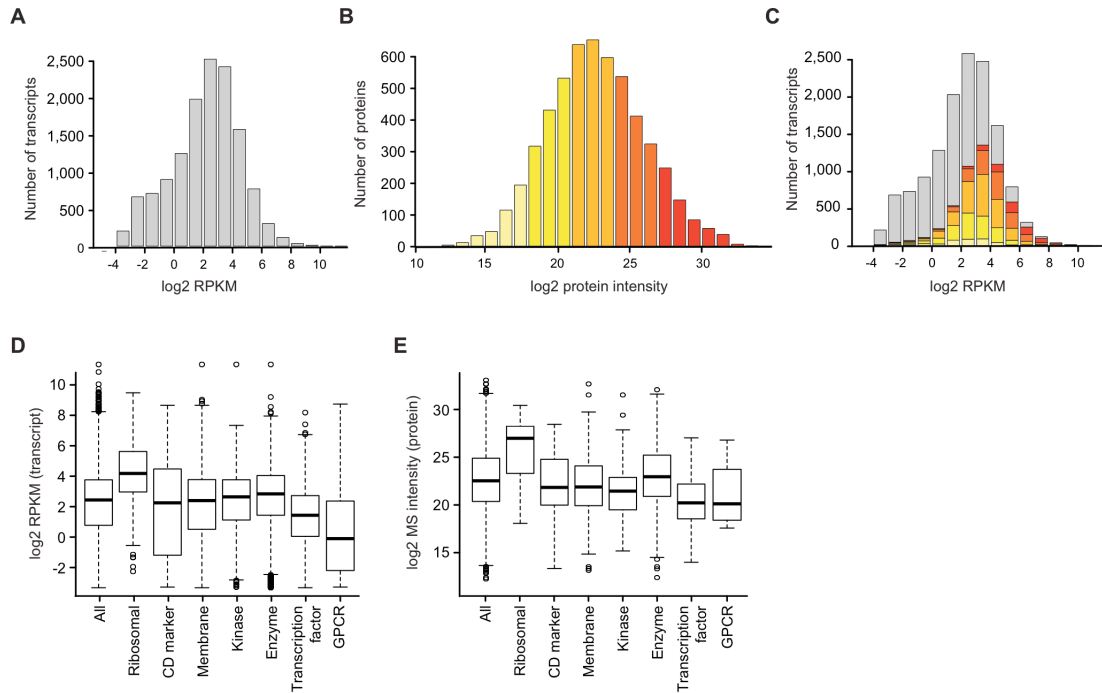


Figure S2. Summary of detected transcript and proteins in the cell line A-431. **A**, Distribution of log₂ RPKM values for all detected transcripts (n=13,075). **B**, Distribution of log₂ MS intensity values for all detected proteins (n=5,408). Bars are colored according to MS intensity ranging from light yellow (low MS intensity) to dark red (high MS intensity). **C**, Overlay of detected proteins onto the histogram of detected transcripts. Transcripts also detected on a protein level are colored according to the MS intensity in **B**. **D**, Distribution of log₂ RPKM values for all transcripts detected in eight protein categories (Supplementary Table 2). **E**, Distribution of log₂ MS Intensity values for all proteins detected in the same protein categories (Supplementary Table 4). For correlation between RNA-seq and MS intensity data see Supplementary Fig. 4.

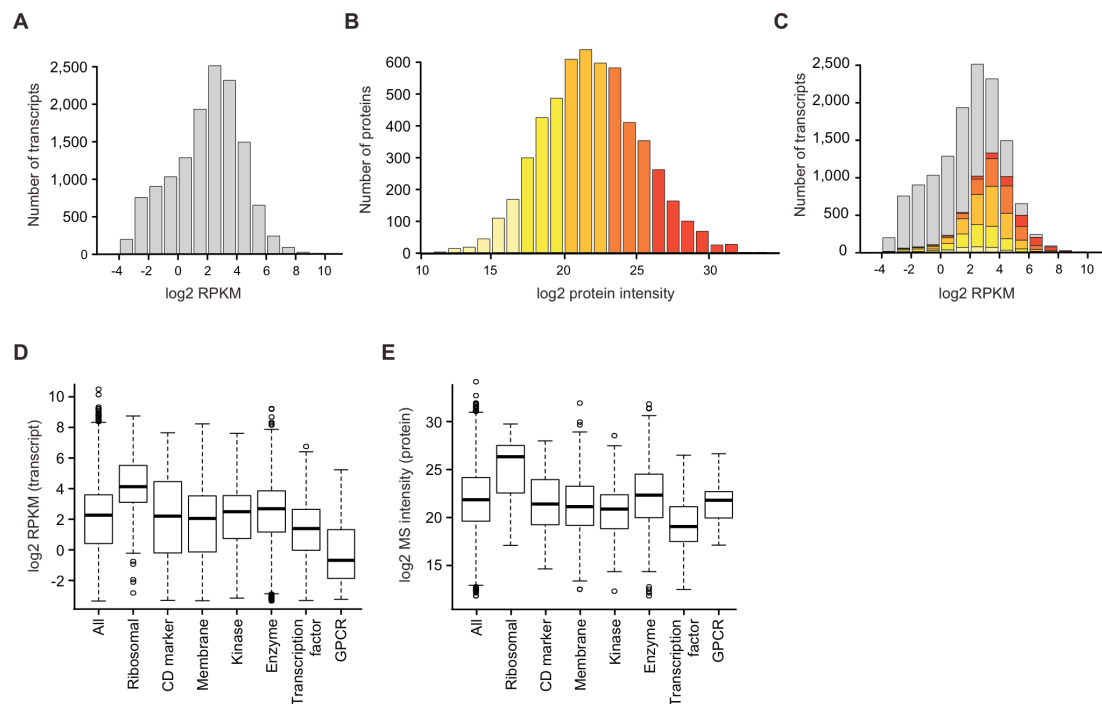


Figure S3. Summary of detected transcript and proteins in the cell line U-251MG. **A**, Distribution of log₂ RPKM values for all detected transcripts (n=13,459). **B**, Distribution of log₂ MS intensity values for all detected proteins (n=5,368). Bars are colored according to MS intensity ranging from light yellow (low MS intensity) to dark red (high MS intensity). **C**, Overlay of detected proteins onto the histogram of detected transcripts. Transcripts also detected on a protein level are colored according to the MS intensity in b. **D**, Distribution of log₂ RPKM values for all transcripts detected in eight protein categories (Supplementary Table 2). **E**, Distribution of log₂ MS Intensity values for all proteins detected in the same protein categories (Supplementary Table 4). For correlation between RNA-seq and MS intensity data see Supplementary Fig. 4.

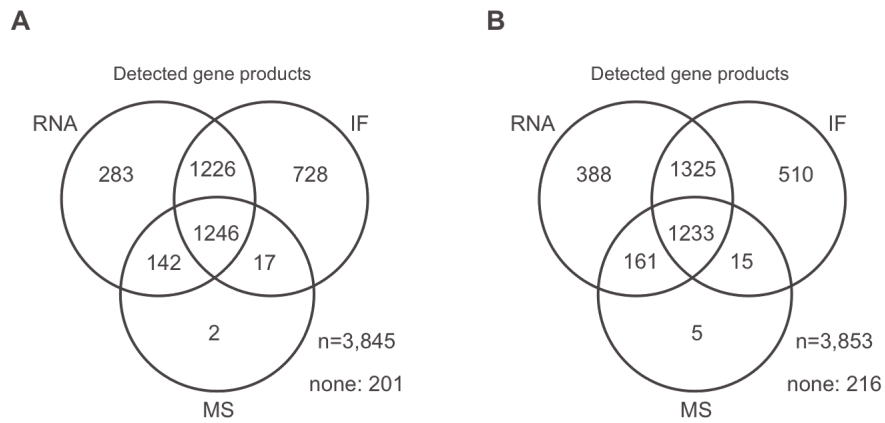


Figure S4. Overlap of detected gene products by the three methods in the cell lines A-431 and U-2OS. **A**, Overlap between the three methods RNA-Seq, MS and confocal IF for all genes studied by IF (n=3,845) in A-431 cells. **B**, Overlap between the three methods for all genes studied by IF (n=3,853) in U-2OS cells.

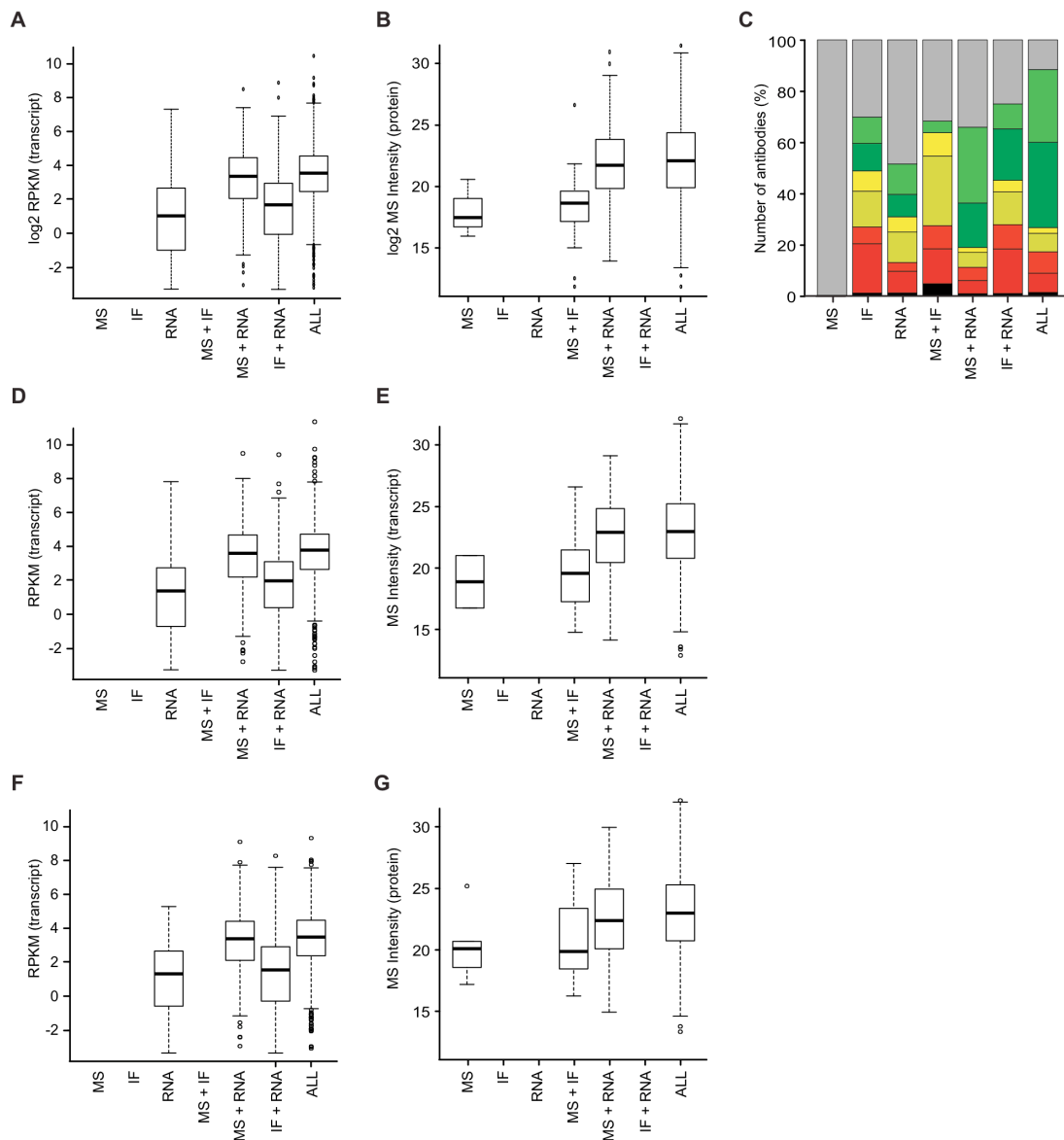


Figure S5. Comparison of the gene products detected by the three methods RNA-Seq, MS and confocal IF in the different categories of the Venn diagram in Figure 3B. **A**, Distribution of log₂ RPKM values for transcripts detected by RNA-Seq in U-251 MG in the different categories. Number of transcripts per category from left to right; 0/0/605/0/23/1,302/3,149. **B**, Distribution of log₂ MS intensity values in the different categories for U-251 MG. Number of proteins per category from left to right; 3/0/0/158/23/0/1,403. **C**, Validation results using Western Blot (WB) on U-251 MG cell lysate, for the used antibodies in the different categories. The WB results are classified as: no bands detected (grey), supportive – single band corresponding to the predicted size in kDa (+/-20%) (bright green), supportive – Band of predicted size in kDa (+/-20%) with additional bands present (green), uncertain – single band larger than predicted size in kDa (+20%) but partly supported by predicted transmembrane region, signal peptide or by other available data (bright yellow) uncertain – single band differing more than +/-20% from predicted size in kDa and not supported by predicted transmembrane region, signal peptide or by other available data (yellow), not supportive – Weak band of predicted size but with additional bands of higher

intensity also present (bright red), not supportive – only bands not corresponding to the predicted size. **D**, Distribution of log₂ RPKM values for transcripts detected by RNA-Seq in A-431 in the different categories. Number of transcripts per category from left to right; 0/0/728/0/15/1,325/3,217. **E**, Distribution of log₂ MS intensity values in the different categories for A-431. Number of proteins per category from left to right; 0/0/728/0/15/1,325/1,407. **F**, Distribution of log₂ RPKM values for transcripts detected by RNA-Seq in U-2 OS in the different categories. Number of transcripts per category from left to right; 0/0/510/0/17/1,226/3,083. **G**, Distribution of log₂ MS intensity values in the different categories for U-2 OS. Number of proteins per category from left to right; 0/0/728/0/15/1,325/1,414.

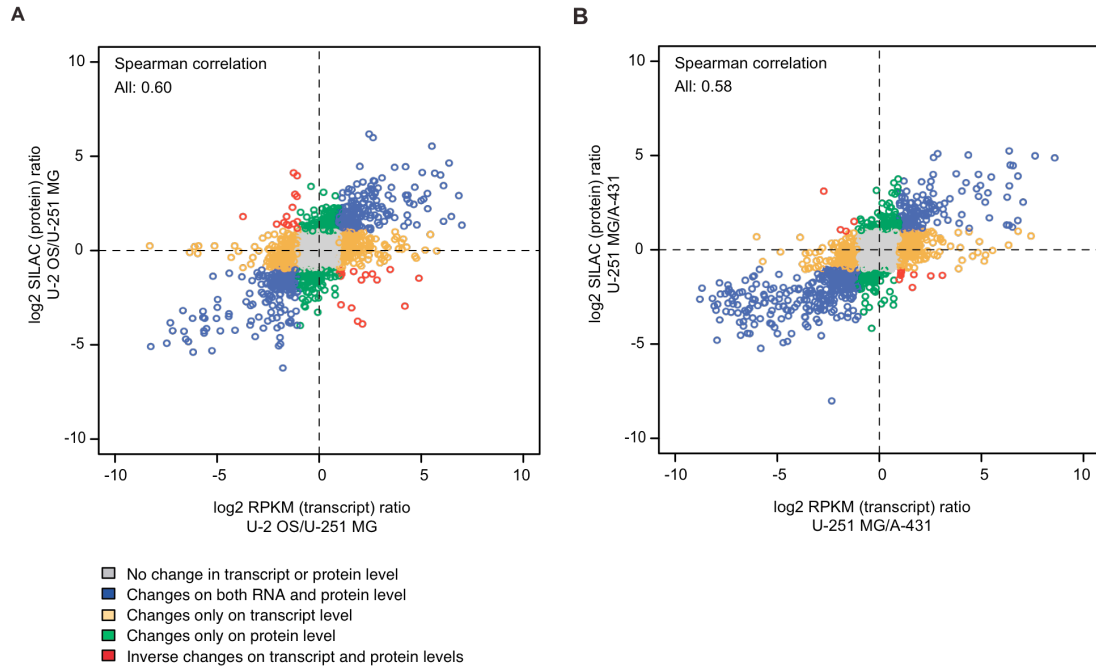


Figure S6. Correlation of transcript and protein level changes between cell lines. Correlation between changes on protein (log₂ SILAC ratio) and transcript levels (log₂ RPKM ratio). **A**, U-2OS over U-251MG cells. **B**, U-251MG over A-431 cells.

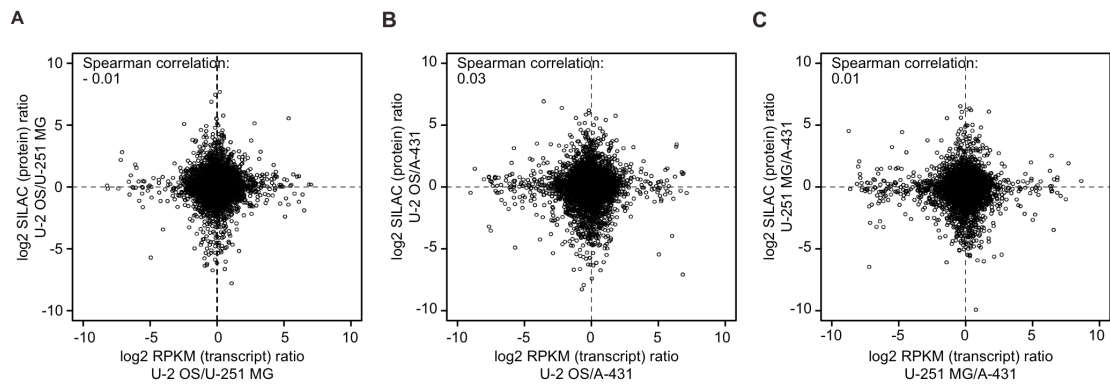


Figure S7. Randomized correlation of transcript and protein ratios between the three cell lines. Transcript / protein pairs of log₂-ratios were randomly selected from the total set of genes expressed on both protein and transcript level. **A**, U-2 OS over U-251 MG cells. **B**, U-2 OS over A-431 cells. **c**, U-251 MG over A-431 cells.

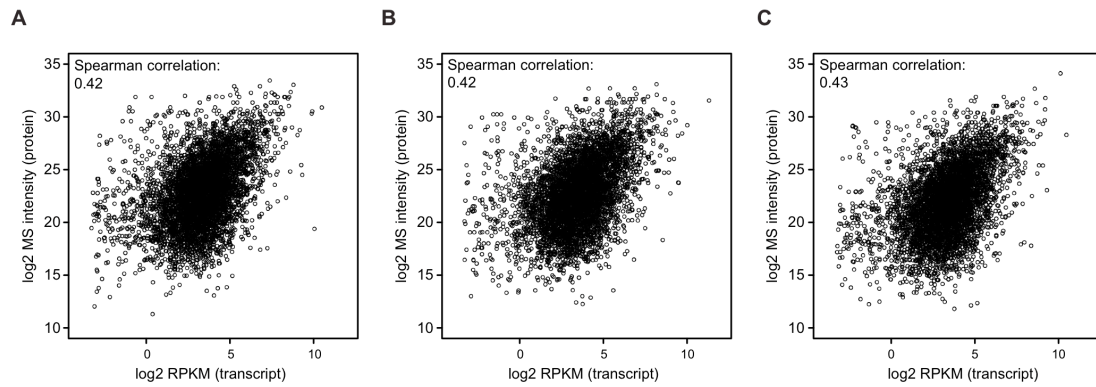


Figure S8. Correlation between transcript and protein levels. Scatter plot of log₂ RPKM and log₂ MS intensity values of gene products detected at both transcript and protein levels. **A**, U-2OS cells (n=5,210). **B**, A-431 cells (n=5,158). **C**, U-251MG cells (n=5,197).

Supplementary Tables

Supplementary Table 1. The number of analyzed and expressed genes for each platform in the three cell lines.

Platform	Cell line	Total analyzed genes	Expressed genes	Fraction expressed genes (%)
MS	U-2 OS	21,248	5,405	25
MS	U-251	21,248	5,368	25
MS	A-431	21,248	5,408	25
RNA	U-2 OS	20,963	14,064	67
RNA	U-251	20,963	13,459	64
RNA	A-431	20,963	13,075	62
IF	U-2 OS	3,874	3,101	80
IF	U-251	3,872	3,168	82
IF	A-431	3,866	3,234	84

Supplementary Table 2. The number of genes from various protein classes present in each of the cell lines based on MS intensities (log2). The collection of protein classes has been previously described¹⁴, but here a class of human cell differentiation molecules ("CD markers") (<http://www.uniprot.org/docs/cdlist>), and a class of membrane proteins based on a majority-decision method ("Membrane") have been added³².

Class	Total number of genes	U-2 OS expressed genes	U-251 MG expressed genes	A-431 expressed genes
All	5,480	5,405	5,368	5,408
Ribosomal	161	134	134	134
CD marker	360	66	65	67
Membrane	5,590	904	897	908
Kinase	504	179	180	180
Enzyme	2,594	1,035	1,031	1,037
Transcription factor	1,416	128	126	126
GPCR	773	7	7	8

Supplementary Table 3. The number of genes from various protein classes present in each of the cell lines based on RPKM values (log2). The collection of protein classes has been previously described¹⁴, but here a class of human cell differentiation molecules ("CD markers") (<http://www.uniprot.org/docs/cdlist>), and a class of membrane proteins based on a majority-decision method ("Membrane") have been added³².

Class	Total number of genes	U-2 OS expressed genes	U-251 MG expressed genes	A-431 expressed genes
All	14,616	14,064	13,459	13,075
Ribosomal	161	145	147	146
CD marker	360	182	155	168
Membrane	5,590	3,356	3,191	2,992
Kinase	504	430	416	404
Enzyme	2,594	2,038	1,973	1,931
Transcription factor	1,416	1,064	1,033	974
GPCR	773	159	149	124

Supplementary Table 4. Gene Ontology (GO)-based enrichment analysis of the proteins in the “Similar”-group in Figure 1F. 3,171 out of 3,228 Ensembl gene identifiers mapped to DAVID. All genes in the plot were used as background (n=4,938). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
macromolecular complex	Cellular Component	1192	2.8E-41	2.1E-38
gene expression	Biological Process	840	2.5E-37	1.1E-33
ribonucleoprotein complex	Cellular Component	362	5.5E-35	2.1E-32
cellular macromolecule metabolic process	Biological Process	1376	5.6E-34	1.2E-30
nuclear part	Cellular Component	793	1.9E-31	4.8E-29
macromolecule metabolic process	Biological Process	1424	1.3E-29	1.8E-26
nucleic acid binding	Molecular Function	795	1.6E-28	2.4E-25
RNA metabolic process	Biological Process	473	4.4E-28	4.6E-25
intracellular organelle	Cellular Component	2290	6.9E-26	1.3E-23
organelle	Cellular Component	2291	8.2E-26	1.2E-23
intracellular organelle part	Cellular Component	1478	1.6E-25	2.0E-23
nucleus	Cellular Component	1361	6.7E-25	7.3E-23
organelle part	Cellular Component	1481	7.5E-25	7.1E-23
membrane-bounded organelle	Cellular Component	2094	8.4E-25	7.1E-23
intracellular membrane-bounded organelle	Cellular Component	2091	1.2E-24	8.9E-23
nuclear lumen	Cellular Component	617	5.4E-23	3.7E-21
RNA processing	Biological Process	341	8.7E-23	7.3E-20
RNA binding	Molecular Function	400	1.3E-22	1.0E-19
cellular macromolecule biosynthetic process	Biological Process	652	1.2E-21	8.2E-19
intracellular organelle lumen	Cellular Component	761	2.6E-21	1.7E-19

Supplementary Table 5. GO-based enrichment analysis of the proteins in the “Slightly changed”-group in Figure 1F. 1,477 out of 1,481 Ensembl gene identifiers mapped to DAVID. All genes in the plot were used as background (n=4,938). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
plasma membrane	Cellular Component	294	6.2E-12	3.6E-9
catalytic activity	Molecular Function	669	1.6E-9	2.2E-6
alcohol metabolic process	Biological Process	82	5.1E-9	1.8E-5
plasma membrane part	Cellular Component	195	2.4E-8	7.0E-6
cell junction	Cellular Component	75	5.0E-8	9.7E-6
membrane	Cellular Component	540	1.0E-7	1.5E-5
monosaccharide binding	Molecular Function	18	1.2E-7	8.1E-5
membrane part	Cellular Component	450	5.4E-7	6.2E-5
intrinsic to plasma membrane	Cellular Component	69	7.0E-7	6.7E-5
regulation of biological quality	Biological Process	168	8.5E-7	1.5E-3
integral to plasma membrane	Cellular Component	66	1.8E-6	1.5E-4
lipid metabolic process	Biological Process	104	2.2E-6	2.7E-3
steroid biosynthetic process	Biological Process	23	5.1E-6	4.6E-3
cell adhesion	Biological Process	69	8.0E-6	5.8E-3
sterol metabolic process	Biological Process	24	1.0E-5	6.0E-3
biological adhesion	Biological Process	69	1.1E-5	5.6E-3
sterol biosynthetic process	Biological Process	18	1.4E-5	6.3E-3
hexose metabolic process	Biological Process	46	1.6E-5	6.3E-3
actin binding	Molecular Function	68	1.6E-5	7.1E-3
homeostatic process	Biological Process	89	2.3E-5	8.3E-3

Supplementary Table 6. GO-based enrichment analysis of the proteins in the “Substantially changed”-group in Figure 1F. 227 out of 229 Ensembl gene identifiers mapped to DAVID. All genes in the plot were used as background (n=4,938). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
plasma membrane	Cellular Component	74	6.5E-13	1.7E-10
biological adhesion	Biological Process	26	5.0E-9	8.2E-6
plasma membrane part	Cellular Component	50	7.6E-9	9.8E-7
cell adhesion	Biological Process	25	2.1E-8	1.7E-5
response to stimulus	Biological Process	65	4.1E-7	2.3E-4
extracellular region	Cellular Component	25	1.7E-6	1.5E-4
extracellular space	Cellular Component	14	4.2E-6	2.7E-4
extracellular region part	Cellular Component	17	4.7E-6	2.4E-4
cell-cell junction	Cellular Component	14	6.0E-6	2.6E-4
intrinsic to plasma membrane	Cellular Component	20	9.4E-6	3.5E-4
multicellular organismal process	Biological Process	62	1.6E-5	6.7E-3
integral to plasma membrane	Cellular Component	19	2.1E-5	6.7E-4
immune response	Biological Process	17	2.2E-5	7.2E-3
response to wounding	Biological Process	16	3.9E-5	1.1E-2
membrane part	Cellular Component	84	5.1E-5	1.4E-3
calcium ion binding	Molecular Function	22	5.7E-5	2.2E-2
protease binding	Molecular Function	5	6.4E-5	1.2E-2
transmembrane receptor activity	Molecular Function	11	9.2E-5	1.2E-2
response to external stimulus	Biological Process	21	9.2E-5	2.1E-2
membrane	Cellular Component	96	1.2E-4	3.0E-3

Supplementary Table 7. GO-based enrichment analysis of the genes from the “Similar” category of Figure 2E. 5,097 genes were identified out of 5,187 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Gene Count	P-value	Benjamini
intracellular	Cellular Component	3745	9.10E-250	7.20E-247
intracellular part	Cellular Component	3656	3.70E-240	1.50E-237
intracellular organelle	Cellular Component	3198	3.80E-193	9.90E-191
organelle	Cellular Component	3201	3.80E-193	7.60E-191
membrane-bounded organelle	Cellular Component	2952	3.90E-190	6.20E-188
intracellular membrane-bounded organelle	Cellular Component	2949	1.40E-189	1.80E-187
nucleus	Cellular Component	2007	1.70E-133	1.90E-131
cellular macromolecule metabolic process	Biological Process	2054	1.00E-125	4.80E-122
intracellular organelle part	Cellular Component	1738	2.30E-124	2.30E-122
organelle part	Cellular Component	1743	4.60E-123	4.00E-121
cellular metabolic process	Biological Process	2447	3.80E-120	9.10E-117
nuclear part	Cellular Component	913	4.80E-109	3.80E-107
macromolecule metabolic process	Biological Process	2130	2.60E-101	4.10E-98
intracellular organelle lumen	Cellular Component	880	5.70E-101	4.10E-99
membrane-enclosed lumen	Cellular Component	901	1.00E-97	6.80E-96
organelle lumen	Cellular Component	885	3.50E-96	2.10E-94
nuclear lumen	Cellular Component	737	3.30E-90	1.90E-88
cellular process	Biological Process	3349	7.00E-88	8.30E-85
primary metabolic process	Biological Process	2421	5.20E-85	5.00E-82
metabolic process	Biological Process	2614	2.30E-84	1.80E-81

Supplementary Table 8. GO-based enrichment analysis of the genes from the “Slightly changed” category of Figure 2E. 6,436 out of 6,607 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
intracellular	Cellular Component	3952	1.70E-40	1.50E-37
intracellular part	Cellular Component	3822	1.50E-36	6.40E-34
cell part	Cellular Component	4950	1.10E-29	3.00E-27
cell	Cellular Component	4950	1.30E-29	2.90E-27
intracellular membrane-bounded organelle	Cellular Component	2933	2.30E-26	4.00E-24
membrane-bounded organelle	Cellular Component	2935	2.60E-26	3.80E-24
cytoplasm	Cellular Component	2719	2.70E-25	3.30E-23
intracellular organelle	Cellular Component	3231	4.70E-24	5.00E-22
organelle	Cellular Component	3232	1.30E-23	1.30E-21
binding	Molecular Function	4349	5.90E-19	1.30E-15
cellular process	Biological Process	3699	7.70E-18	4.40E-14
cellular metabolic process	Biological Process	2421	2.10E-15	6.00E-12
cytoplasmic part	Cellular Component	1838	3.70E-15	3.20E-13
metabolic process	Biological Process	2729	3.10E-13	5.90E-10
nucleus	Cellular Component	1855	2.90E-12	2.30E-10
cellular macromolecule metabolic process	Biological Process	1911	7.30E-12	1.00E-08
primary metabolic process	Biological Process	2477	9.20E-12	1.00E-08
catalytic activity	Molecular Function	1898	4.40E-09	4.90E-06
transferase activity, transferring phosphorus-containing groups	Molecular Function	408	1.20E-08	9.20E-06
metal ion binding	Molecular Function	1514	1.50E-08	8.10E-06

Supplementary Table 9. GO-based enrichment analysis of the genes from the “Substantially changed” category of Figure 2E. 1,649 out of 1,691 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
system development	Biological Process	317	5.80E-16	2.10E-12
extracellular region part	Cellular Component	161	1.70E-15	8.50E-13
multicellular organismal development	Biological Process	366	1.30E-14	2.50E-11
anatomical structure development	Biological Process	332	1.40E-14	1.80E-11
developmental process	Biological Process	387	5.70E-13	5.50E-10
extracellular region	Cellular Component	261	2.70E-12	7.00E-10
nervous system development	Biological Process	167	4.10E-12	3.10E-09
plasma membrane	Cellular Component	440	7.90E-12	1.30E-09
cell adhesion	Biological Process	116	1.20E-11	7.50E-09
biological adhesion	Biological Process	116	1.30E-11	6.90E-09
extracellular matrix	Cellular Component	73	1.30E-11	1.60E-09
plasma membrane part	Cellular Component	286	1.40E-11	1.40E-09
membrane	Cellular Component	751	1.50E-10	1.30E-08
proteinaceous extracellular matrix	Cellular Component	67	1.70E-10	1.30E-08
cell junction	Cellular Component	92	6.00E-10	3.80E-08
regulation of developmental process	Biological Process	107	6.80E-09	3.30E-06
multicellular organismal process	Biological Process	469	7.10E-09	3.00E-06
extracellular matrix part	Cellular Component	33	8.00E-09	4.50E-07
calcium ion binding	Molecular Function	132	1.50E-08	1.80E-05
organ development	Biological Process	222	1.60E-08	6.10E-06

Supplementary Table 10. GO-based enrichment analysis of the genes from the “Cell-type specific” category of Figure 2E. 1,905 out of 2,053 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
substrate specific channel activity	Molecular Function	97	6.30E-20	7.10E-17
channel activity	Molecular Function	99	7.30E-20	4.10E-17
passive transmembrane transporter activity	Molecular Function	99	8.80E-20	3.30E-17
ion channel activity	Molecular Function	93	8.90E-19	2.50E-16
multicellular organismal process	Biological Process	524	3.30E-18	1.20E-14
intrinsic to membrane	Cellular Component	655	1.30E-17	5.70E-15
extracellular region	Cellular Component	294	2.20E-17	5.10E-15
transmembrane transporter activity	Molecular Function	158	2.40E-16	5.00E-14
gated channel activity	Molecular Function	76	8.60E-16	1.70E-13
substrate-specific transmembrane transporter activity	Molecular Function	145	3.00E-15	4.80E-13
ion transmembrane transporter activity	Molecular Function	132	3.40E-15	4.70E-13
integral to membrane	Cellular Component	626	4.40E-15	6.60E-13
plasma membrane part	Cellular Component	307	7.00E-15	8.00E-13
transporter activity	Molecular Function	185	1.40E-14	1.70E-12
ion transport	Biological Process	134	5.30E-14	9.70E-11
integral to plasma membrane	Cellular Component	188	5.80E-14	5.30E-12
cell-cell signaling	Biological Process	112	9.00E-14	1.10E-10
intrinsic to plasma membrane	Cellular Component	189	2.50E-13	1.90E-11
membrane part	Cellular Component	737	6.20E-13	4.10E-11
plasma membrane	Cellular Component	463	7.70E-13	4.40E-11

Supplementary Table 11. GO-based enrichment analysis of the genes from the “Not expressed” category of Figure 2E. 4,307 out of 5,425 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

Term	GO Category	Count	P-Value	Benjamini
olfactory receptor activity	Molecular Function	344	2.30E-229	3.10E-226
sensory perception of smell	Biological Process	347	1.30E-213	5.30E-210
G-protein coupled receptor activity	Molecular Function	529	7.40E-210	5.00E-207
sensory perception of chemical stimulus	Biological Process	362	1.40E-202	2.90E-199
G-protein coupled receptor protein signaling pathway	Biological Process	608	1.70E-196	2.40E-193
transmembrane receptor activity	Molecular Function	637	4.80E-175	2.20E-172
receptor activity	Molecular Function	771	4.30E-159	1.40E-156
sensory perception	Biological Process	463	6.90E-157	7.20E-154
cognition	Biological Process	477	2.90E-138	2.40E-135
cell surface receptor linked signal transduction	Biological Process	736	7.00E-127	4.80E-124
molecular transducer activity	Molecular Function	824	4.90E-122	1.30E-119
signal transducer activity	Molecular Function	824	4.90E-122	1.30E-119
system process	Biological Process	634	5.80E-119	3.50E-116
neurological system process	Biological Process	540	9.40E-114	4.90E-111
extracellular region	Cellular Component	734	3.50E-94	1.80E-91
response to stimulus	Biological Process	1046	6.50E-89	3.00E-86
multicellular organismal process	Biological Process	1209	1.00E-80	4.20E-78
intrinsic to membrane	Cellular Component	1448	1.50E-70	3.80E-68
integral to membrane	Cellular Component	1409	1.60E-70	2.70E-68
plasma membrane	Cellular Component	1094	1.10E-67	1.30E-65

Supplementary Table 12. GO-based enrichment analysis of the genes expressed exclusively in U-2 OS on the transcript level. 859 out of 922 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
extracellular region	Cellular Component	137	5.30E-10	1.80E-07
neurotransmitter binding	Molecular Function	21	6.90E-09	5.20E-06
multicellular organismal process	Biological Process	237	7.80E-09	2.00E-05
neurotransmitter receptor activity	Molecular Function	20	9.60E-09	3.60E-06
intrinsic to membrane	Cellular Component	289	4.90E-08	8.50E-06
integral to plasma membrane	Cellular Component	90	6.90E-08	8.10E-06
substrate specific channel activity	Molecular Function	42	7.20E-08	1.80E-05
ion channel activity	Molecular Function	41	9.40E-08	1.70E-05
intrinsic to plasma membrane	Cellular Component	91	9.90E-08	8.60E-06
cell-cell signaling	Biological Process	54	1.20E-07	1.60E-04
plasma membrane	Cellular Component	214	1.30E-07	9.10E-06
channel activity	Molecular Function	42	1.90E-07	2.80E-05
passive transmembrane transporter activity	Molecular Function	42	2.00E-07	2.50E-05
gated channel activity	Molecular Function	35	2.20E-07	2.30E-05
transmembrane transporter activity	Molecular Function	69	1.20E-06	1.20E-04
integral to membrane	Cellular Component	273	1.30E-06	7.80E-05

Supplementary Table 13. GO-based enrichment analysis of the genes expressed exclusively in U-251 MG on the transcript level. 524 out of 559 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
substrate specific channel activity	Molecular Function	41	6.10E-13	3.70E-10
nervous system development	Biological Process	74	6.80E-13	1.60E-09
channel activity	Molecular Function	41	1.90E-12	5.90E-10
passive transmembrane transporter activity	Molecular Function	41	2.00E-12	4.20E-10
ion channel activity	Molecular Function	39	4.80E-12	7.30E-10
substrate-specific transmembrane transporter activity	Molecular Function	56	4.20E-10	5.10E-08
transporter activity	Molecular Function	70	4.20E-10	4.30E-08
transmembrane transporter activity	Molecular Function	59	5.90E-10	5.20E-08
multicellular organismal process	Biological Process	171	1.50E-09	1.80E-06
transmission of nerve impulse	Biological Process	33	2.60E-09	2.00E-06
ion transmembrane transporter activity	Molecular Function	50	2.70E-09	2.00E-07
substrate-specific transporter activity	Molecular Function	59	5.50E-09	3.80E-07
gated channel activity	Molecular Function	30	5.50E-09	3.40E-07
glutamate receptor activity	Molecular Function	11	6.00E-09	3.30E-07
cation channel activity	Molecular Function	28	6.80E-09	3.50E-07
chemical homeostasis	Biological Process	40	7.80E-09	4.60E-06
ion transport	Biological Process	51	8.50E-09	4.00E-06
synaptic transmission	Biological Process	29	1.40E-08	5.60E-06
voltage-gated cation channel activity	Molecular Function	20	1.60E-08	7.40E-07
system process	Biological Process	77	1.80E-08	5.90E-06

Supplementary Table 14. GO-based enrichment analysis of the genes expressed exclusively in A-431 on the transcript level. 525 out of 572 Ensembl gene identifiers mapped to DAVID. All genes analyzed with RNA-Seq were used as background (n=19,063). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
immune response	Biological Process	39	1.40E-07	2.80E-04
immune system process	Biological Process	46	6.70E-06	6.40E-03
intrinsic to membrane	Cellular Component	174	1.40E-05	3.80E-03
defense response	Biological Process	32	4.10E-05	2.60E-02
membrane part	Cellular Component	200	4.30E-05	5.90E-03
integral to membrane	Cellular Component	166	5.50E-05	5.00E-03
extracellular region	Cellular Component	76	7.00E-05	4.80E-03
inflammatory response	Biological Process	20	2.50E-04	1.10E-01
fibroblast growth factor receptor antagonist activity	Molecular Function	4	2.70E-04	1.50E-01
interleukin-1 receptor antagonist activity	Molecular Function	4	2.70E-04	1.50E-01
FAD binding	Molecular Function	9	3.10E-04	8.70E-02
heme binding	Molecular Function	11	4.20E-04	7.90E-02
interleukin-1 receptor activity	Molecular Function	4	4.60E-04	6.50E-02
interleukin-1 receptor binding	Molecular Function	5	4.90E-04	5.60E-02
membrane	Cellular Component	211	5.40E-04	2.90E-02
cytokine activity	Molecular Function	14	6.10E-04	5.80E-02
tetrapyrrole binding	Molecular Function	11	7.20E-04	5.90E-02
icosanoid metabolic process	Biological Process	7	8.60E-04	2.80E-01
iron ion binding	Molecular Function	18	8.90E-04	6.30E-02
unsaturated fatty acid metabolic process	Biological Process	7	1.30E-03	3.50E-01

Supplementary Table 15. The mRNA abundance in cell line A-431 for the categories defined in Figure 3. The mRNA abundance has been divided in three similar sized fractions (low, medium and high) and the number of genes and percent of genes within each category is shown.

Category	Number of genes	Low	Medium	High
Similar levels	5,187	560 (11%)	1,997 (38%)	2,630 (51%)
Slightly changed levels	5,921	1,916 (32%)	1,957 (33%)	2,048 (35%)
Significantly changed levels	1,395	663 (48%)	352 (25%)	380 (27%)
Cell-type specific	572	479 (84%)	66 (12%)	27 (5%)

Supplementary Table 16. The mRNA abundance in cell line U-251 MG for the categories defined in Figure 2E. The mRNA abundance has been divided in three similar sized fractions (low, medium and high) and the number of genes and percent of genes within each category is shown.

Category	Number of genes	Low	Medium	High
Similar levels	5,187	532 (10%)	1,982 (38%)	2,673 (52%)
Slightly changed levels	6,246	2,258 (36%)	2,084 (33%)	1,904 (30%)
Significantly changed levels	1,467	8160 (59%)	362 (25%)	245 (17%)
Cell-type specific	559	522 (93%)	33 (6%)	4 (1%)

Supplementary Table 17. GO-based enrichment analysis of the “grey” category of Figure 3C. 2,992 out of 3,003 Ensembl gene identifiers mapped to DAVID. The genes detected on the transcript and protein level in both U-2 OS and A-431 were used as background (n=4,734). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
gene expression	Biological Process	760	1.30E-18	5.70E-15
cellular macromolecule metabolic process	Biological Process	1269	2.50E-17	5.40E-14
macromolecular complex	Cellular Component	1056	1.00E-15	7.40E-13
nuclear part	Cellular Component	722	2.70E-15	9.90E-13
macromolecule metabolic process	Biological Process	1312	5.10E-14	7.40E-11
RNA metabolic process	Biological Process	427	3.40E-13	3.70E-10
membrane-bounded organelle	Cellular Component	1958	3.90E-13	9.70E-11
intracellular membrane-bounded organelle	Cellular Component	1955	5.20E-13	9.60E-11
intracellular organelle lumen	Cellular Component	710	1.60E-12	2.30E-10
membrane-enclosed lumen	Cellular Component	725	4.30E-12	5.30E-10
intracellular organelle	Cellular Component	2129	4.50E-12	4.70E-10
organelle	Cellular Component	2130	5.00E-12	4.60E-10
organelle lumen	Cellular Component	711	8.30E-12	6.90E-10
intracellular	Cellular Component	2515	1.50E-11	1.10E-09
nucleus	Cellular Component	1252	1.60E-11	1.10E-09
intracellular organelle part	Cellular Component	1353	1.80E-11	1.10E-09
nuclear lumen	Cellular Component	564	2.10E-11	1.20E-09
cellular macromolecule biosynthetic process	Biological Process	594	2.60E-11	2.30E-08
organelle part	Cellular Component	1357	3.10E-11	1.70E-09
nucleic acid binding	Molecular Function	710	5.20E-11	8.10E-08

Supplementary Table 18. GO-based enrichment analysis of the “blue” category of Figure 3C. All 623 Ensembl gene identifiers mapped to DAVID. The genes detected on the transcript and protein level in both U-2 OS and A-431 were used as background (n=4,734). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
plasma membrane	Cellular Component	161	2.50E-16	9.10E-14
cell adhesion	Biological Process	46	2.00E-10	5.30E-07
biological adhesion	Biological Process	46	2.70E-10	3.50E-07
plasma membrane part	Cellular Component	103	8.00E-10	1.60E-07
intrinsic to plasma membrane	Cellular Component	45	1.00E-09	1.40E-07
membrane	Cellular Component	261	1.40E-09	1.40E-07
integral to plasma membrane	Cellular Component	43	4.40E-09	3.60E-07
anchoring junction	Cellular Component	30	4.50E-08	3.00E-06
calcium ion binding	Molecular Function	48	8.20E-08	6.70E-05
membrane part	Cellular Component	211	8.20E-07	4.80E-05
cell junction	Cellular Component	38	1.40E-06	7.10E-05
extracellular region	Cellular Component	45	2.60E-06	1.20E-04
integral to membrane	Cellular Component	136	2.90E-06	1.20E-04
intrinsic to membrane	Cellular Component	141	3.70E-06	1.40E-04
adherens junction	Cellular Component	25	4.80E-06	1.60E-04
actin filament-based process	Biological Process	32	5.40E-06	4.70E-03
cytoskeletal protein binding	Molecular Function	51	8.00E-06	3.20E-03
actin binding	Molecular Function	37	1.70E-05	4.50E-03
apical junction complex	Cellular Component	15	1.70E-05	5.40E-04
apicolateral plasma membrane	Cellular Component	15	2.60E-05	7.70E-04

Supplementary Table 19. GO-based enrichment analysis of the “yellow” category of Figure 3C. 692 out of 706 Ensembl gene identifiers mapped to DAVID. The genes detected on the transcript and protein level in both U-2 OS and A-431 were used as background (n=4,734). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
nucleosome	Cellular Component	12	1.50E-03	5.10E-01
protein-DNA complex	Cellular Component	14	1.00E-02	9.10E-01
MHC class I peptide loading complex	Cellular Component	4	1.20E-02	8.50E-01
response to DNA damage stimulus	Biological Process	37	1.30E-02	1.00E+00
nuclease activity	Molecular Function	18	1.30E-02	1.00E+00
mitochondrial transport	Biological Process	12	1.50E-02	1.00E+00
intronless viral mRNA export from host nucleus	Biological Process	5	2.00E-02	1.00E+00
structural constituent of muscle	Molecular Function	5	2.10E-02	1.00E+00
transcription export complex	Cellular Component	5	2.20E-02	9.20E-01
mitotic cell cycle checkpoint	Biological Process	9	2.30E-02	1.00E+00
response to stress	Biological Process	86	2.40E-02	1.00E+00
response to bacterium	Biological Process	10	2.50E-02	1.00E+00
regulation of immune system process	Biological Process	21	2.60E-02	1.00E+00
coenzyme binding	Molecular Function	22	2.80E-02	1.00E+00
monocarboxylic acid metabolic process	Biological Process	24	3.10E-02	1.00E+00
nucleosome assembly	Biological Process	12	3.20E-02	1.00E+00
DNA damage response, signal transduction resulting in induction of apoptosis	Biological Process	6	3.20E-02	1.00E+00
cellular response to stress	Biological Process	46	3.30E-02	1.00E+00
response to molecule of bacterial origin	Biological Process	8	3.40E-02	1.00E+00
sarcolemma	Cellular Component	6	3.50E-02	9.60E-01

Supplementary Table 20. GO-based enrichment analysis of the “green” category of Figure 3C. 375 out of 376 Ensembl gene identifiers mapped to DAVID. The genes detected on the transcript and protein level in both U-2 OS and A-431 were used as background (n=4,734). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
cellular homeostasis	Biological Process	23	1.90E-04	3.30E-01
catalytic activity	Molecular Function	179	4.60E-04	2.50E-01
di-, tri-valent inorganic cation homeostasis	Biological Process	11	1.20E-03	7.30E-01
isomerase activity	Molecular Function	16	1.40E-03	3.50E-01
cation homeostasis	Biological Process	12	1.40E-03	6.30E-01
multicellular organismal process	Biological Process	82	1.50E-03	5.40E-01
cellular chemical homeostasis	Biological Process	15	1.90E-03	5.50E-01
cellular ion homeostasis	Biological Process	15	1.90E-03	5.50E-01
cellular cation homeostasis	Biological Process	11	2.10E-03	5.30E-01
cellular di-, tri-valent inorganic cation homeostasis	Biological Process	10	2.30E-03	5.00E-01
ion homeostasis	Biological Process	15	2.70E-03	5.20E-01
chemical homeostasis	Biological Process	17	3.30E-03	5.40E-01
transmembrane receptor activity	Molecular Function	11	3.90E-03	5.60E-01
multicellular organismal development	Biological Process	65	5.80E-03	7.20E-01
biological regulation	Biological Process	159	6.50E-03	7.20E-01
developmental process	Biological Process	72	8.50E-03	7.80E-01
homeostatic process	Biological Process	26	8.60E-03	7.60E-01
cellular response to oxidative stress	Biological Process	6	9.70E-03	7.80E-01
potassium ion transport	Biological Process	5	1.00E-02	7.80E-01
lymphocyte proliferation	Biological Process	5	1.00E-02	7.80E-01

Supplementary Table 21. GO-based enrichment analysis of the “red” category of Figure 3C. All 17 Ensembl gene identifiers mapped to DAVID. Note, relatively high p-values due to the small gene set in this category. The genes detected on the transcript and protein level in both U-2 OS and A-431 were used as background (n=4,734). The GO categories and terms for the top 20 hits sorted by enrichment p-value are reported here along with gene count and Benjamini-Hochberg multiple testing correction.

GO Term	GO Category	Count	P-Value	Benjamini
anatomical structure development	Biological Process	7	1.40E-02	9.90E-01
muscle organ development	Biological Process	3	2.00E-02	9.70E-01
system development	Biological Process	6	3.30E-02	9.80E-01
cell differentiation	Biological Process	5	4.10E-02	9.80E-01
developmental process	Biological Process	7	4.20E-02	9.60E-01
organ development	Biological Process	5	4.70E-02	9.50E-01
cellular developmental process	Biological Process	5	5.30E-02	9.40E-01
hydrolase activity	Molecular Function	7	5.40E-02	9.90E-01
negative regulation of specific transcription from RNA polymerase II promoter	Biological Process	2	6.00E-02	9.40E-01
multicellular organismal process	Biological Process	7	6.20E-02	9.30E-01
negative regulation of gene-specific transcription	Biological Process	2	6.40E-02	9.10E-01
multicellular organismal development	Biological Process	6	7.70E-02	9.30E-01
biological regulation	Biological Process	11	8.60E-02	9.40E-01
regulation of specific transcription from RNA polymerase II promoter	Biological Process	2	9.60E-02	9.40E-01
cytoskeletal part	Cellular Component	4	9.90E-02	1.00E+00
negative regulation of cell cycle	Biological Process	2	1.00E-01	9.40E-01

Supplementary references

References to supplementary information

- ³¹ Li, H. and Durbin, R., Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25** (14), 1754 (2009).
- ³² Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M., and Berglund, L., Prediction of the human membrane proteome. *Proteomics* **10** (6), 1141 (2010).