

Supplementary Materials

Immune-Signatures for Lung Cancer Diagnostics: Evaluation of Protein Microarray Data Normalization Strategies

Stefanie Brezina, Regina Soldo, Roman Kreuzhuber, Philipp Hofer, Andrea Gsur and Andreas Weinhaeusel

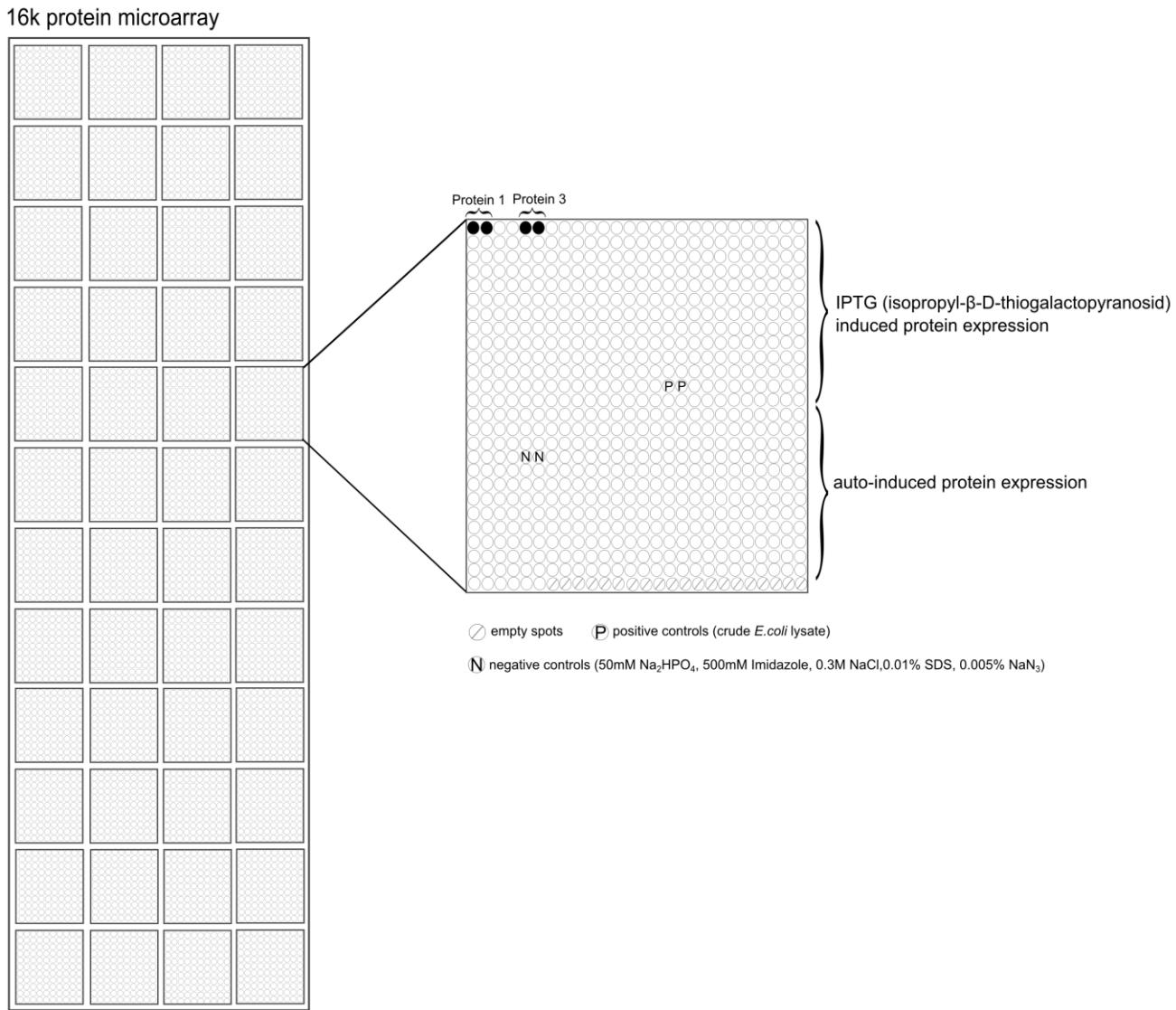


Figure S1. 16k protein microarray design. The array is sub-divided into 48 blocks, each containing 676 spots. Protein spots are printed in duplicate. As positive control crude *E. coli* lysate (0.3, 0.4, and 0.5 mg mL⁻¹) and as negative control elution buffer was spotted onto the slides. Control spots are randomly distributed over the whole array. In total, 281 positive and 82 negative control spots are printed onto the 16k microarray.

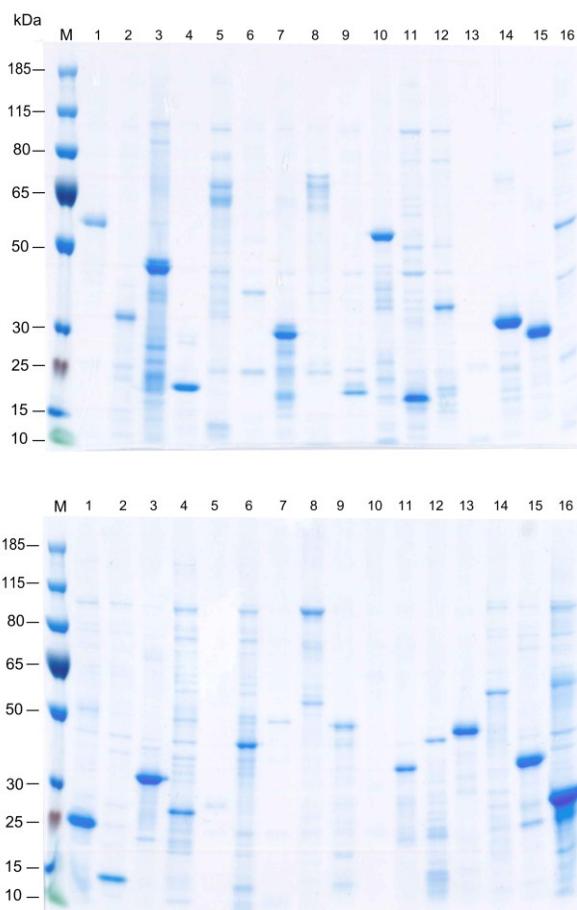


Figure S2. SDS-PAGE (NuPAGE Novex, 4%–12% Bis-tris) of 2 µg UniPEx protein eluates (lanes 1–16) after Ni-NTA purification. PageRuler™ Plus Prestained Protein Ladder (Life Technologies, Carlsbad, CA, USA) was used.

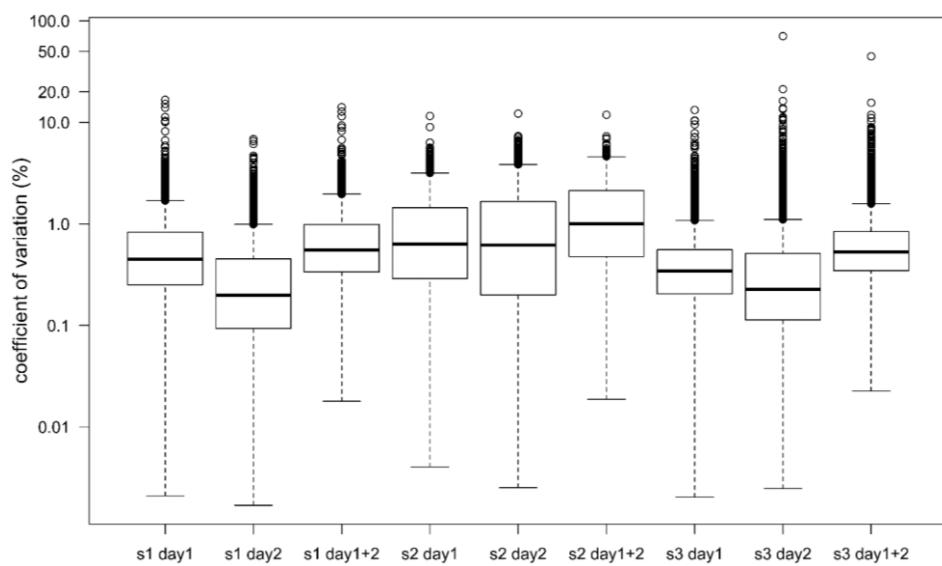


Figure S3. Boxplot representing intra-experiment and inter-experiment coefficient of variation (CV) values of all proteins on the 16k microarray. Each box represents a replicate of a specimen (s1, s2, s3) which was processed on the microarray in triplicates on the same day (day1 and day2 = intra-experiment comparison) and on consecutive days (day1+2=inter-experiment comparison).

Figure S4. Overlaps of significant antigens derived from class comparison analysis with different data pre-processing methods (DWD, ComBat, quantile normalization, unnormalized) including the whole data set with all lung cancer cases *versus* controls (“All (r1-6)”), and “single-run” as well as “cross-run” analyses of the four histologic entities. Right upper half shows relative overlaps (%) and left lower half shows absolute overlaps. QNORM = quantile normalization, SCLC = small cell lung cancer, SqLC = squamous cell lung cancer, LCLC = large cell lung cancer, AdCa = adenocarcinoma of the lung, r = experimental run.

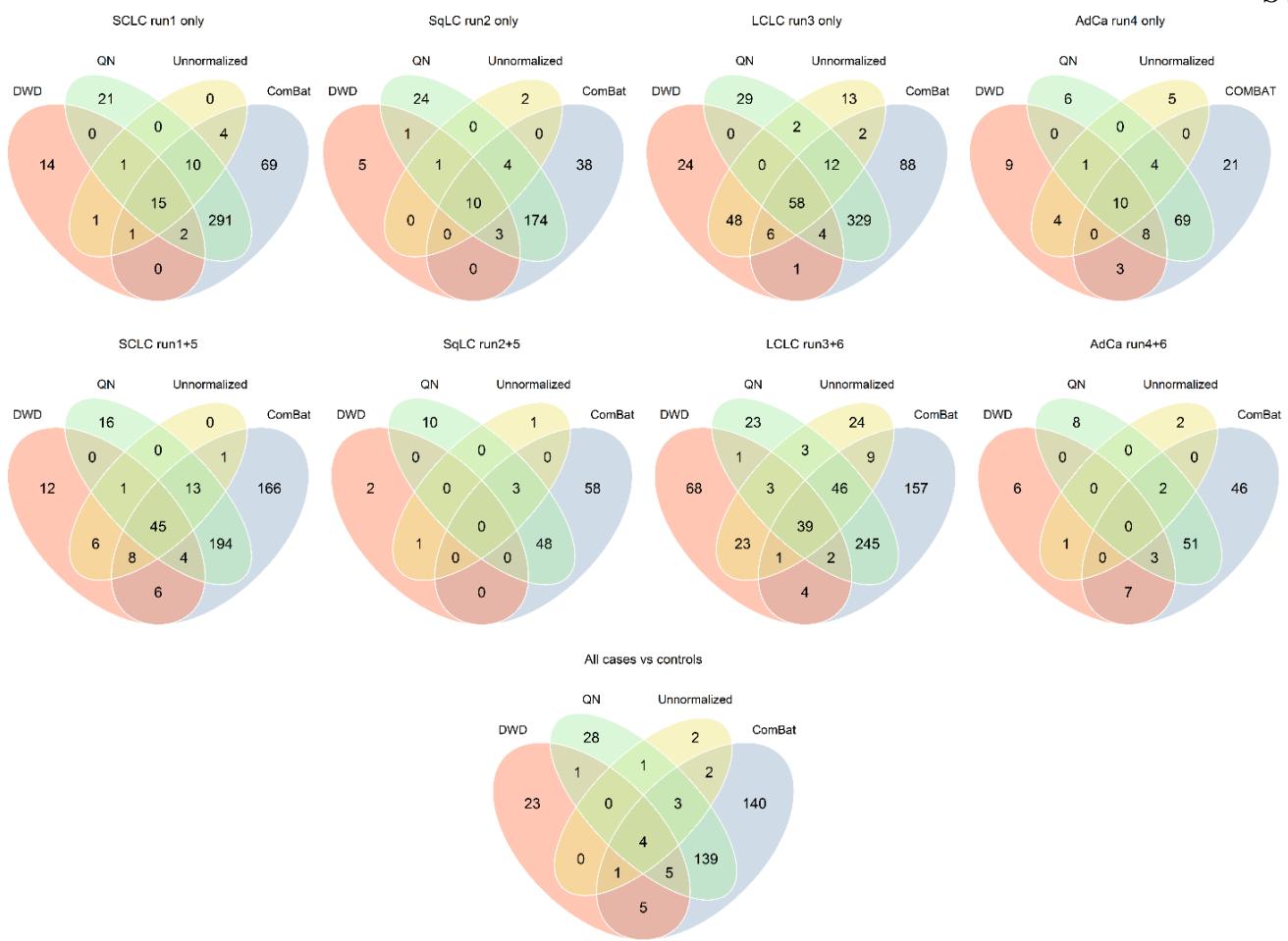


Figure S5. Overlaps of significant antigens derived from class comparison analysis ($p < 0.001$) with different data pre-processing methods (DWD, QN = quantile normalization, unnormalized, ComBat) for “single-run” analysis (run1 only to run4 only) and “cross-run” analysis (run1+5 to run4+6) for distinct histologic entities of lung cancer (SCLC = small cell lung cancer, SqLC = squamous cell lung cancer, LCLC = large cell lung cancer, AdCa = adenocarcinoma of the lung), as well as for all cases *versus* all controls including all samples.

Table S1. Classifier list established with class prediction using quantile normalized data and 100 recursive feature elimination using all cases (n = 100) versus all controls (n = 100). For class prediction results compare Table 5.

UniqueID	Gene Symbol	p-value	UniqueID	Gene Symbol	p-value
RZPDp9027L0419Q	KIF5A	1.00 × 10 ⁻⁶	RZPDp9027G0320Q	TXN2	1.35 × 10 ⁻¹
RZPDp9028L0418Q	EIF3M	1.05 × 10 ⁻⁴	RZPDp9028M0217Q	TOMM20	1.30 × 10 ⁻¹
RZPDp9028L1618Q	NFYA	1.32 × 10 ⁻⁴	RZPDp9027G197Q	NOL11	1.30 × 10 ⁻¹
RZPDp9028L1514Q	FPGS	1.96 × 10 ⁻⁴	RZPDp9027D1316Q	NECAP1	1.13 × 10 ⁻¹
RZPDp9028G0420Q	NFKBIA	2.91 × 10 ⁻⁴	RZPDp9028C229Q	SNRNP48	9.64 × 10 ⁻²
RZPDp9027G0711Q	TRIOBP	3.61 × 10 ⁻⁴	RZPDp9028K2012Q	ARHGEF1	9.59 × 10 ⁻²
RZPDp9027H168Q	BCAS2	7.07 × 10 ⁻⁴	RZPDp9028L201Q	LRP1	6.11 × 10 ⁻²
RZPDp9028O1314Q	HLA-E	1.93 × 10 ⁻³	RZPDp9027E1817Q	EXOSC10	5.74 × 10 ⁻²
RZPDp9028N0516Q	ADI1	2.13 × 10 ⁻³	RZPDp9028B0517Q	PPP6R1	5.62 × 10 ⁻²
RZPDp9028K1819Q	TBCB	2.34 × 10 ⁻³	RZPDp9027I1618Q	GABBR1	5.61 × 10 ⁻²
RZPDp9028J0319Q	FYN	2.92 × 10 ⁻³	RZPDp9028N1413Q	TP53BP2	5.48 × 10 ⁻²
RZPDp9027K1912Q	TRAK1	3.07 × 10 ⁻³	RZPDp9028G059Q	AKAP13	5.06 × 10 ⁻²
RZPDp9028L1519Q	ARHGEF18	4.83 × 10 ⁻³	RZPDp9028I1316Q	EDARADD	4.99 × 10 ⁻²
RZPDp9028A0112Q	TRIM28	9.32 × 10 ⁻³	RZPDp9028H1220Q	RRP1B	4.68 × 10 ⁻²
RZPDp9027K1911Q	SNCB	1.02 × 10 ⁻²	RZPDp9028B0915Q	DUSP2	4.14 × 10 ⁻²
RZPDp9028K169Q	PRMT1	1.05 × 10 ⁻²	RZPDp9028F218Q	ATXN2L	4.08 × 10 ⁻²
RZPDp9027C1912Q	TMEM222	1.35 × 10 ⁻²	RZPDp9028A1219Q	NLRP1	2.27 × 10 ⁻²
RZPDp9028N136Q	MUC2	1.55 × 10 ⁻²	RZPDp9028F019Q	HERC2	2.26 × 10 ⁻²
RZPDp9028K0714Q	NFKB1	2.04 × 10 ⁻²	RZPDp9028P168Q	CEP250	2.17 × 10 ⁻²
RZPDp9027I062Q	G3BP2	2.15 × 10 ⁻²	RZPDp9027G1218Q	EPS8	1.98 × 10 ⁻²
RZPDp9028O1617Q	PLCG1	2.98 × 10 ⁻²	RZPDp9028C2019Q	GGA2	1.92 × 10 ⁻²
RZPDp9028H1517Q	ACO2	3.14 × 10 ⁻²	RZPDp9027A196Q	CUL7	1.77 × 10 ⁻²
RZPDp9028D191Q	COMP	3.66 × 10 ⁻²	RZPDp9028A1014Q	NLRC5	1.64 × 10 ⁻²
RZPDp9028G0616Q	APBB1	3.76 × 10 ⁻²	RZPDp9028D058Q	UTP14A	1.54 × 10 ⁻²
RZPDp9027G0710Q	FAM192A	3.99 × 10 ⁻²	RZPDp9028P024Q	ARHGDIA	1.26 × 10 ⁻²
RZPDp9028P028Q	NEDD9	4.35 × 10 ⁻²	RZPDp9028J074Q	SLC9A3R2	1.14 × 10 ⁻²
RZPDp9028D093Q	COL4A1	6.78 × 10 ⁻²	RZPDp9027D2118Q	COPE	1.04 × 10 ⁻²
RZPDp9028P048Q	ERCC5	7.07 × 10 ⁻²	RZPDp9028I186Q	AGRN	1.01 × 10 ⁻²
RZPDp9027G078Q	C10orf35	7.16 × 10 ⁻²	RZPDp9028G0917Q	SETD2	5.60 × 10 ⁻³
RZPDp9028I0219Q	D2HGDH	8.84 × 10 ⁻²	RZPDp9028L2017Q	RPS25	4.88 × 10 ⁻³
RZPDp9028N152Q	CCDC88C	9.99 × 10 ⁻²	RZPDp9028N0410Q	FBF1	4.74 × 10 ⁻³
RZPDp9027J039Q	SBK1	1.13 × 10 ⁻¹	RZPDp9027E197Q	AKR1C4	1.52 × 10 ⁻³
RZPDp9028G1711Q	RSBN1	1.46 × 10 ⁻¹	RZPDp9028A0915Q	SMYD5	1.48 × 10 ⁻³
RZPDp9028I021Q	IGHG1	1.52 × 10 ⁻¹	RZPDp9028H1210Q	NFKB1	1.29 × 10 ⁻³
RZPDp9027K1920Q	HMGB2	1.82 × 10 ⁻¹	RZPDp9028I2212Q	PSAP	9.28 × 10 ⁻⁴
RZPDp9027B1912Q	VIMP	2.74 × 10 ⁻¹	RZPDp9028M1019Q	TRIM78P	7.68 × 10 ⁻⁴
RZPDp9028A0521Q	GOLGA7	7.65 × 10 ⁻¹	RZPDp9028M219Q	ZEB1	6.39 × 10 ⁻⁴
RZPDp9028J131Q	SFN	9.72 × 10 ⁻¹	RZPDp9028I2112Q	RCSD1	5.71 × 10 ⁻⁴
RZPDp9028K1319Q	SRPR	9.07 × 10 ⁻¹	RZPDp9028M099Q	SRRM2	3.32 × 10 ⁻⁴
RZPDp9028F0420Q	FAM21A/FAM21C	8.46 × 10 ⁻¹	RZPDp9028O1020Q	UQCRC1	2.93 × 10 ⁻⁴
RZPDp9027G075Q	GLOD4	5.81 × 10 ⁻¹	RZPDp9028D249Q	MDFIC	2.03 × 10 ⁻⁴
RZPDp9028A133Q	WNK2	3.42 × 10 ⁻¹	RZPDp9028M0811Q	ADH5	1.94 × 10 ⁻⁴
RZPDp9027K199Q	KCTD15	2.76 × 10 ⁻¹	RZPDp9028D1211Q	SREBF2	1.06 × 10 ⁻⁴
RZPDp9027K1910Q	PPP1CA	2.50 × 10 ⁻¹	RZPDp9028D213Q	MEGF6	6.79 × 10 ⁻⁵
RZPDp9027P138Q	MC1R	2.46 × 10 ⁻¹	RZPDp9028F0114Q	TP53	6.65 × 10 ⁻⁵
RZPDp9028N196Q	ALDOA	2.00 × 10 ⁻¹	RZPDp9028A2418Q	PSMC4	3.29 × 10 ⁻⁵
RZPDp9028L013Q	CLDN5	1.58 × 10 ⁻¹	RZPDp9028E218Q	PCBP1	2.58 × 10 ⁻⁵
RZPDp9028A1721Q	BAZ1A	1.41 × 10 ⁻¹	RZPDp9028I2111Q	ZC3H13	7.50 × 10 ⁻⁶
RZPDp9027P046Q	FGFR3	1.40 × 10 ⁻¹	RZPDp9028E1012Q	SUMO1P3	5.30 × 10 ⁻⁶
RZPDp9028C0813Q	MED20	1.37 × 10 ⁻¹	RZPDp9028M236Q	CD81	5.00 × 10 ⁻⁶

Table S2. Classifier list established with class prediction using ComBat-adjusted data and 25 greedy pairs using all cases (n = 100) *versus* all controls (n = 100). For class prediction results compare Table 5.

UniqueID	Gene Symbol	p-value
RZPDp9027C1510Q	CNPPD1	< 1 × 10 ⁻⁷
RZPDp9027H1613Q	NONO	< 1 × 10 ⁻⁷
RZPDp9027L0419Q	KIF5A	< 1 × 10 ⁻⁷
RZPDp9027J0410Q	U2AF1	1.00 × 10 ⁻⁷
RZPDp9027A0816Q	SPTBN4	1.00 × 10 ⁻⁷
RZPDp9027H1611Q	WBP11	2.00 × 10 ⁻⁷
RZPDp9027M0416Q	PRKAG1	4.00 × 10 ⁻⁷
RZPDp9027N199Q	MRPL10	5.00 × 10 ⁻⁷
RZPDp9027N1810Q	BRD2	1.10 × 10 ⁻⁶
RZPDp9027N047Q	TANK	1.30 × 10 ⁻⁶
RZPDp9028F183Q	SIRT7	1.40 × 10 ⁻⁶
RZPDp9027F0610Q	WHSC2	1.40 × 10 ⁻⁶
RZPDp9027J1913Q	RPL18	1.60 × 10 ⁻⁶
RZPDp9027G0410Q	PIN1	2.40 × 10 ⁻⁶
RZPDp9027K065Q	CAP1	2.70 × 10 ⁻⁶
RZPDp9027F079Q	ZNF638	3.60 × 10 ⁻⁶
RZPDp9028P1819Q	XPO4	4.40 × 10 ⁻⁶
RZPDp9027J0512Q	SGK2	5.00 × 10 ⁻⁶
RZPDp9027L1612Q	EDC4	5.20 × 10 ⁻⁶
RZPDp9027H168Q	BCAS2	7.80 × 10 ⁻⁶
RZPDp9028L0418Q	EIF3M	8.10 × 10 ⁻⁶
RZPDp9027D1610Q	MARCH2	1.00 × 10 ⁻⁵
RZPDp9028L1519Q	ARHGEF18	7.02 × 10 ⁻⁵
RZPDp9027L045Q	FAM189B	1.00 × 10 ⁻⁴
RZPDp9028L1514Q	FPGS	2.04 × 10 ⁻⁴
RZPDp9027F194Q	R3HCC1	8.80 × 10 ⁻³
RZPDp9027A1112Q	GSTM4	1.30 × 10 ⁻³
RZPDp9027G1218Q	EPS8	1.28 × 10 ⁻³
RZPDp9027L236Q	RPL9	8.37 × 10 ⁻⁴
RZPDp9027O0115Q	OTUD1	5.74 × 10 ⁻⁴
RZPDp9027A2412Q	POLR2B	3.16 × 10 ⁻⁴
RZPDp9028I095Q	C12orf32	7.67 × 10 ⁻⁵
RZPDp9027M114Q	AGT	2.04 × 10 ⁻⁵
RZPDp9028D213Q	MEGF6	1.77 × 10 ⁻⁵
RZPDp9028A2418Q	PSMC4	9.40 × 10 ⁻⁶
RZPDp9028I2111Q	ZC3H13	8.40 × 10 ⁻⁶
RZPDp9028G0813Q	IMPDH2	5.70 × 10 ⁻⁶
RZPDp9027A1812Q	RRP9	5.70 × 10 ⁻⁶
RZPDp9028E1012Q	SUMO1P3	5.60 × 10 ⁻⁶
RZPDp9028M126Q	ZNF629	4.80 × 10 ⁻⁶
RZPDp9028H1213Q	RABGGTB	4.20 × 10 ⁻⁶
RZPDp9028M2311Q	CFDP1	4.10 × 10 ⁻⁶
RZPDp9027A093Q	SRSF3	3.60 × 10 ⁻⁶
RZPDp9028E098Q	CLIP1	3.40 × 10 ⁻⁶
RZPDp9028I2310Q	ZNFX1	3.20 × 10 ⁻⁶
RZPDp9028M236Q	CD81	2.50 × 10 ⁻⁶
RZPDp9027M1013Q	WDR73	1.10 × 10 ⁻⁶
RZPDp9027E239Q	OCIAD2	8.00 × 10 ⁻⁷
RZPDp9027P239Q	U2SURP	4.00 × 10 ⁻⁷
RZPDp9028H1210Q	NFKB1	3.00 × 10 ⁻⁷