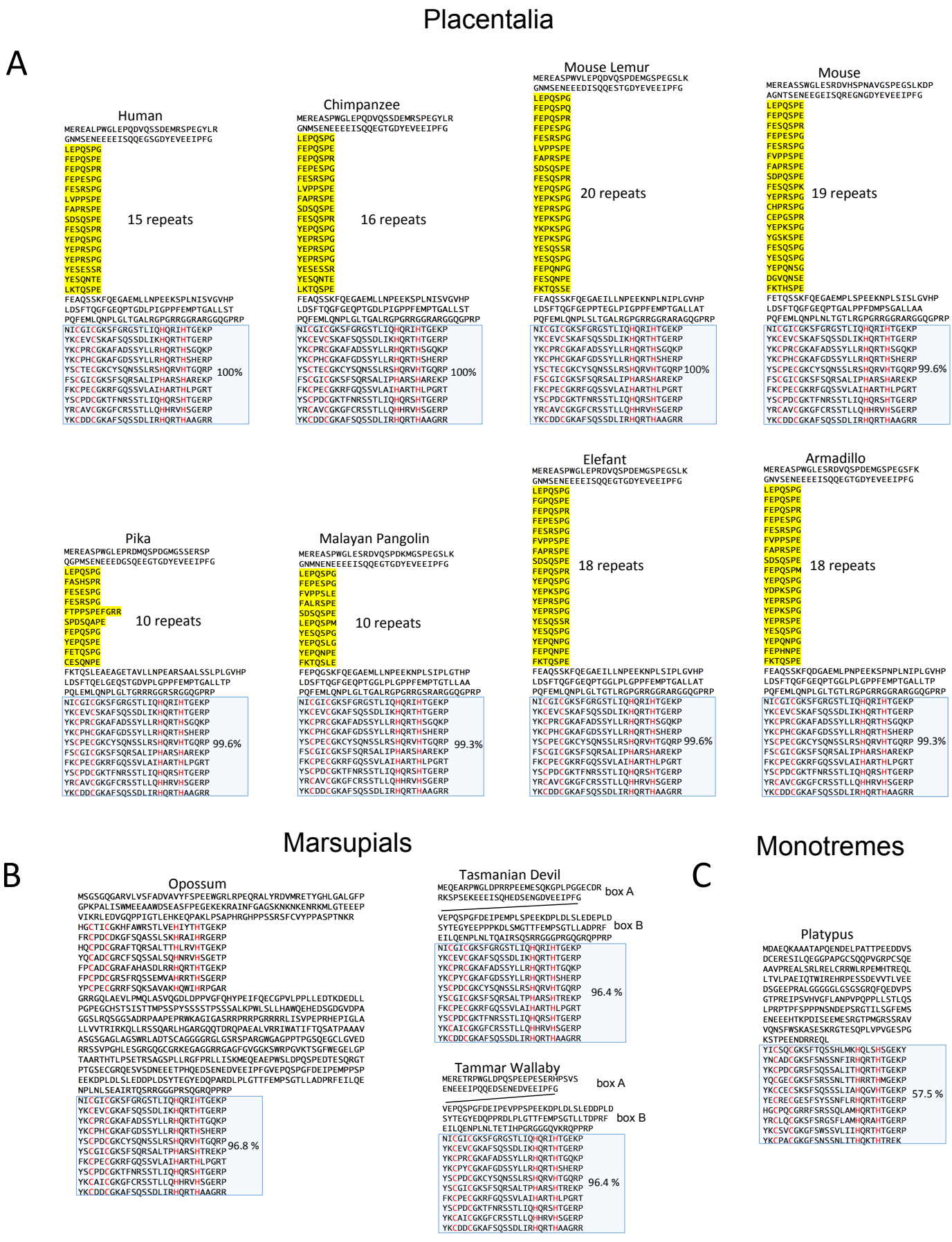

Placentalia
A
Human
15 repeats
Chimpanzee
16 repeats
Mouse Lemur
20 repeats
Mouse
19 repeats
Pika
10 repeats
Malayan Pangolin
10 repeats
Elefant
18 repeats
Armadillo
18 repeats
100%
99.6%
99.3%
Marsupials
B
Opossum
Tasmanian Devil
box A
box B
96.4 %
Tammar Wallaby
box A
box B
96.4 %
96.8 %
Monotremes
C
Platypus
57.5 %

Figure S1

Figure S1

Protein sequence of ZNF768 in various species of mammals. A domain of 10 zinc fingers is highly conserved in all mammals (blue box).
(**A**) An array containing 10 - 20 heptad repeats (yellow box) is a characteristic of placental animals. (**B**) Marsupials Opossum, Tammar Wallaby and Tasmanian devil lack the array of heptad-repeats. ZNF768 Tammar Wallaby and Tasmanian devil contain box A and box B. (**C**) Monotreme Platypus lacks the array of heptad repeats, box A and box B, and shows further a reduced conservation of the zinc finger domain (blue box). Sorting was performed according the phylogenetic relationship.

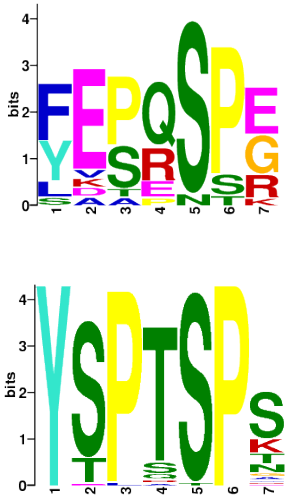
Reference sequence of mammalian ZNF768 proteins:

Human, *Homo sapiens*, NP_078947.3
Chimpanzee, *Pan troglodytes*, XP_016785186.1
Mouse lemur, *Microcebus murinus*, XP_012619672.1
Mouse, *Mus musculus*, NP_666314.1
Pika, *Ochotona princeps*, XP_012782987
Malayan pangolin, *Manis javanica*, XP_017519428.1
Elefant, *Loxodonta africana*, XP_010596820.1
Armadillo, *Dasypus novemcinctus*, XP_012375444.1
Opossum, *Monodelphis domestica*, XP_007498557.2
Tasmanian devil, *Sarcophilus harrisii*, XP_012398319.1
Tammar wallaby, *Macropus eugenii*, ENSMEUP00000000462
Platypus, *Ornithorhynchus anatinus,* ENSOANP00000018579

A

MEREALPWGLEPQDVQSSDEMRSPEGYLR
GNMSENEEEEISQQEGSGDYEVEEIPFG
LEPQSPG
FEPQSPE
FEPQSPR
FEPESPG
FESRSPG
LVPPSPE
FAPRSPE
SDSQSPE
FESQSPR
YEPQSPG
YEPRSPG
YEPRSPG
YESESSR
YESQNTE
LKTQSPE
FEAQSSKFQEGAEMLLNPEEKSPLNISVGVHP
LDSFTQGFGEQPTGDLPIGPPFEMPTGALLST
PQFEMLQNPLGLTGALRGPGRRGGRARGGQGPRP
NICGICGKSFGRGSTLIQHQRIHTGEKP
YKCEVCSKAFSQSSDLIKHQRTHTGERP
YKCPRCGKAFADSSYLLRHQRTHSGQKP
YKCPHCGKAFGDSSYLLRHQRTHSHERP
YSCTECGKCYSQNSSLRSHQRVHTGQRP
FSCGICGKSFSQRSALIPHARSHAREKP
FKCPECGKRFGQSSVLAIHARTHLPGRT
YSCPDCGKTFNRSSTLIQHQRSHTGERP
YRCAVCGKGFCRSSTLLQHHRVHSGERP
YKCDDCGKAFSQSSDLIRHQRTHAAGRR

B



C

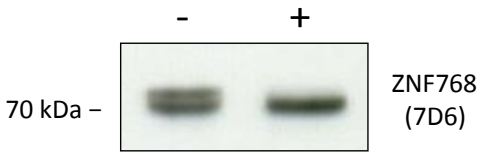


Figure S2

Phosphorylation of ZNF768. (**A**) Confirmed phosphorylation sites in human ZNF768 protein (www.cellsignal.com). (**B**) Heptad repeat consensus motif for ZNF768 determined from the 15 heptad repeats in human ZNF768 and Pol II CTD (Fig. 1B) using MEME. (**C**) Western blot of cellular extracts of U2OS cells before (-) and after (+) treatment with alkaline phosphatase.

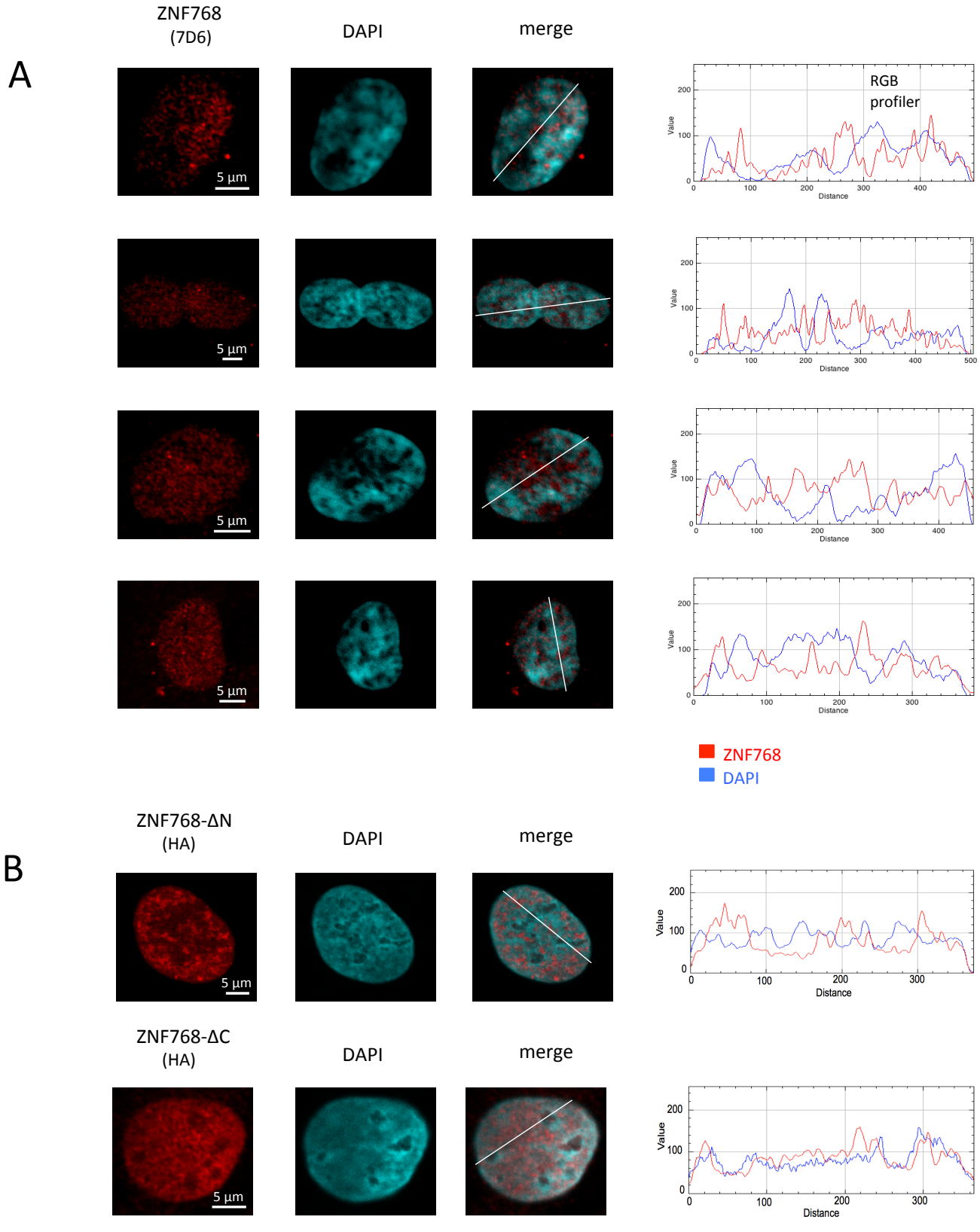


Figure S3

Confocal images of ZNF768 in U2OS cells. (**A**) Endogenous ZNF768 was stained with mAb 7D6 (red), chromatin with DAPI, and merged. RGB profiler is shown on the right hand site, white line marks the scanned area. (**B**) Confocal images of ZNF768 mutants (Structure of mutants is described in Figure 2E).

A

	# merged regions	unique (1/4)	diff cell (2/4)	consistent cell (2/4)	consistent cell+ (3/4)	all (4/4)
Raji (rep. 1)	15854 (56.1%)	7324 (25.2%)	631 (80.82%)	3048 (62.11%)	336 (89.88%)	2747 (98.33%)
Raji (rep. 2)	9242 (80.08%)	1169 (55.77%)	305 (96.72%)	3048 (62.11%)	216 (99.07%)	2747 (98.33%)
U2OS (rep. 1)	8983 (84.31%)	2421 (60.43%)	876 (86.53%)	716 (79.75%)	1654 (95.16%)	2747 (98.33%)
U2OS (rep. 2)	4527 (88.67%)	358 (29.33%)	60 (78.33%)	716 (79.75%)	87 (90.8%)	2747 (98.33%)
unique regions	21012 (58.09%)	11272 (36.07%)	936 (86%)	3764 (65.46%)	2293 (94.59%)	2747 (98.33%)

B

	MIRs	MIRs with motif	MIRs with stringent motif	motifs genome-wide	stringent motifs genome-wide
total	579294	70760	13812	423846	14852
bound by ZNF768	13168	11187	7481	12875	7706
Percentage bound	2.3	15.8	54.2	3.0	51.9

Figure S4

ZNF768 peaks in Raji and U2OS cells. (**A**) Number of peaks and percentage of peaks (in brackets) containing the binding motif in replicate experiment 1 and 2 in Raji and U2OS cells. Explanation of column headings: # merged regions = total number of peaks in each sample, overlapping peaks (peak centers ±100 bp) were merged within samples and across all 4 samples to obtain unique peak regions; unique (1/4) = number of peaks identified only in one sample; diff cell (2/4) = number of peaks identified in one replicate for each cell type; consistent cell (2/4) = number of peaks identified in both replicates for one cell type but not in any replicate for the other cell type; consistent cell+ (3/4) = number of peaks identified in both replicates for one cell type and one replicate for the other cell type; all (4/4) = number of peaks identified in all four samples. (**B**) ZNF768 peaks in MIR sequences and in the whole genome.

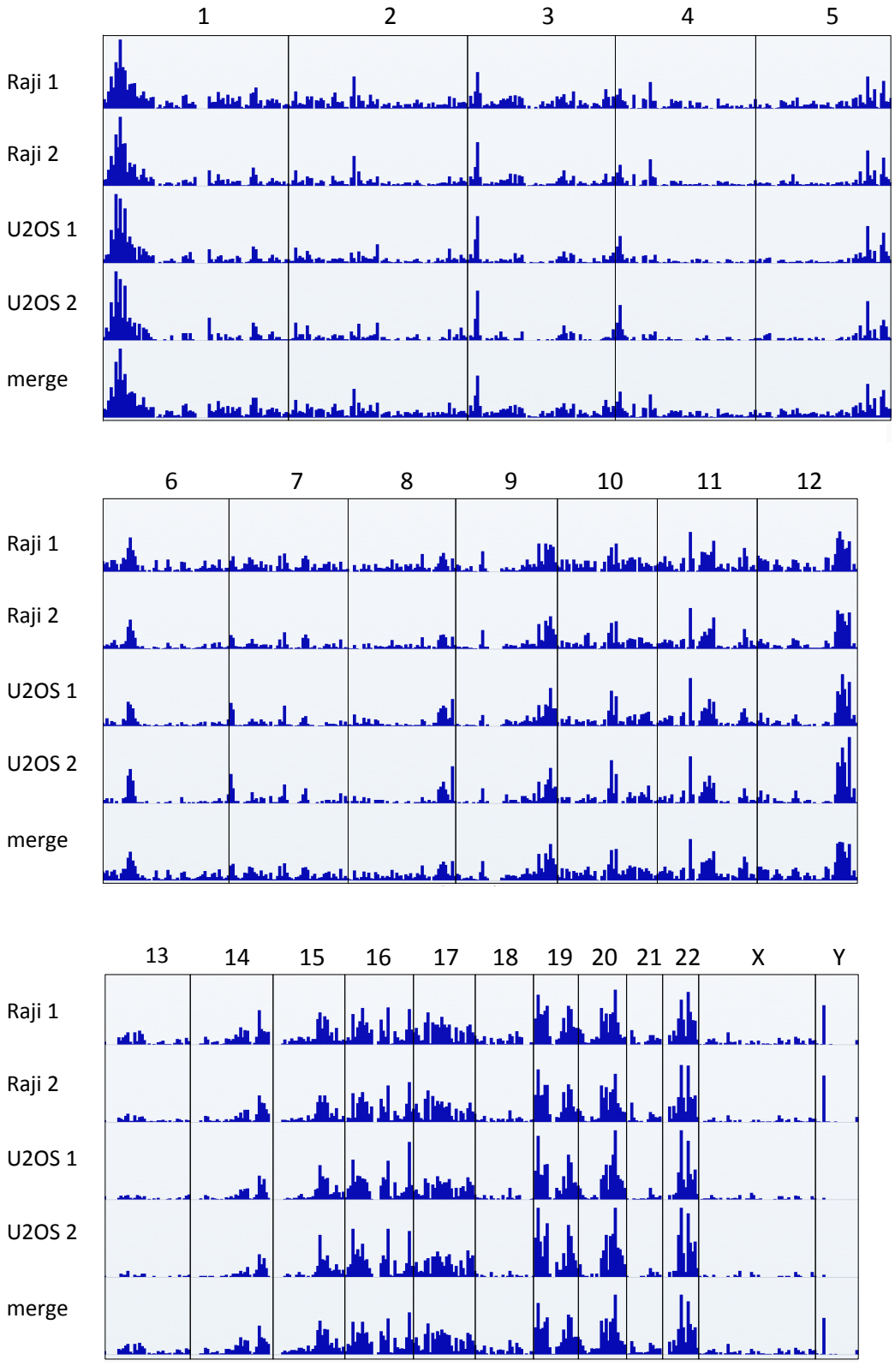
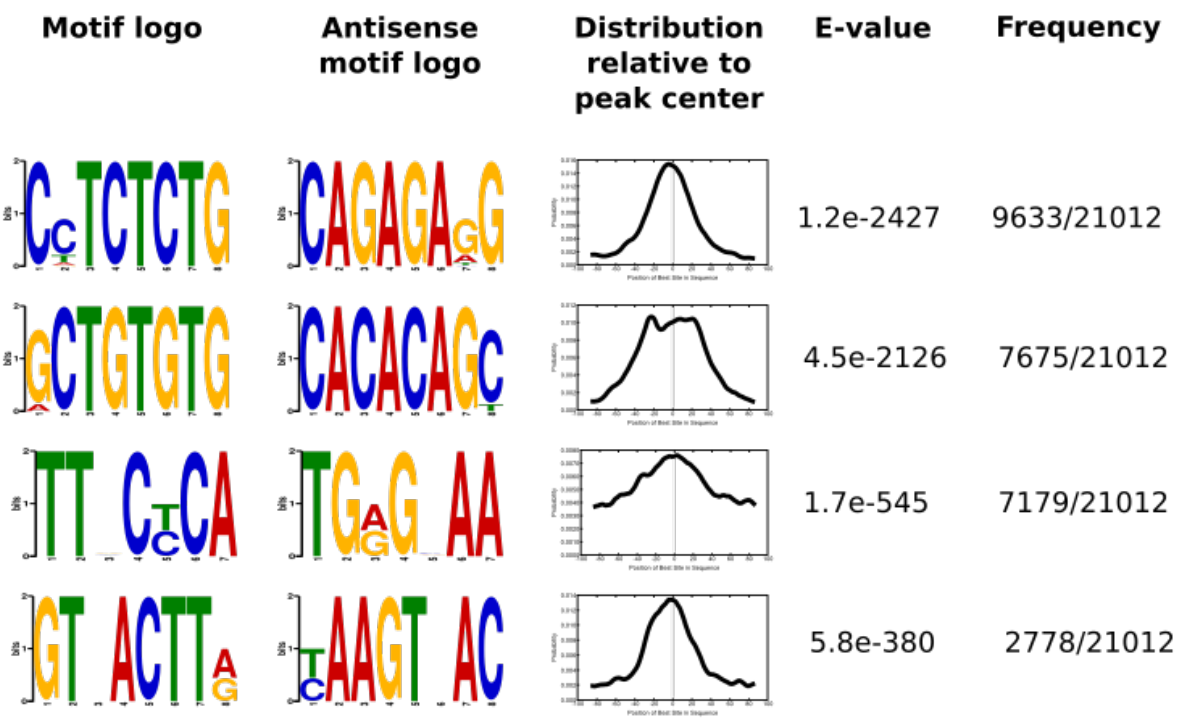


Figure S5
Distribution of ZNF768 peaks in ChIP experiments in Raji and U2OS cells over chromosomes 1 – 22, and X and Y chromosome.

A



B

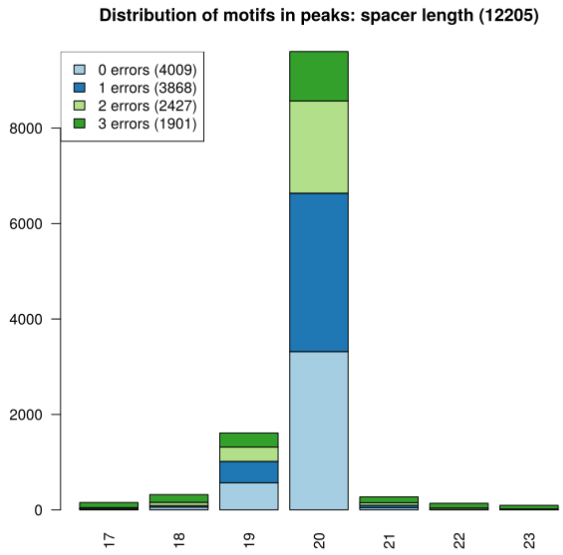


Figure S6

(**A**) Top four motif logos identified using MEME-ChIP in merged unique ZNF768 peaks in Raji and U2OS cells. Frequency indicates the number of merged unique peaks containing each motif. (**B**) Number of peaks matching the ZNF768 binding motif in Fig. 2A with a spacer length of 20±3bp. For each spacer length, the number of peaks containing the anchor regions with a certain number of mismatches is indicated.

A

M1: 5´- *CAGT***GCTGTGTGACCTTGGGCAAGTCACTTAACCTCTCTG***CAGT* -3´

M2: 5´- *CAGT***GCTGTGTG**CAGTCAGTCAGTCAGTCAGT**CCTCTCTG***CAGT* -3´

M3: 5´- *CAGT*CAGT**TGTGACCTTGGGCAAGTCACTTAACCTC**CAGT*CAGT* -3´

B

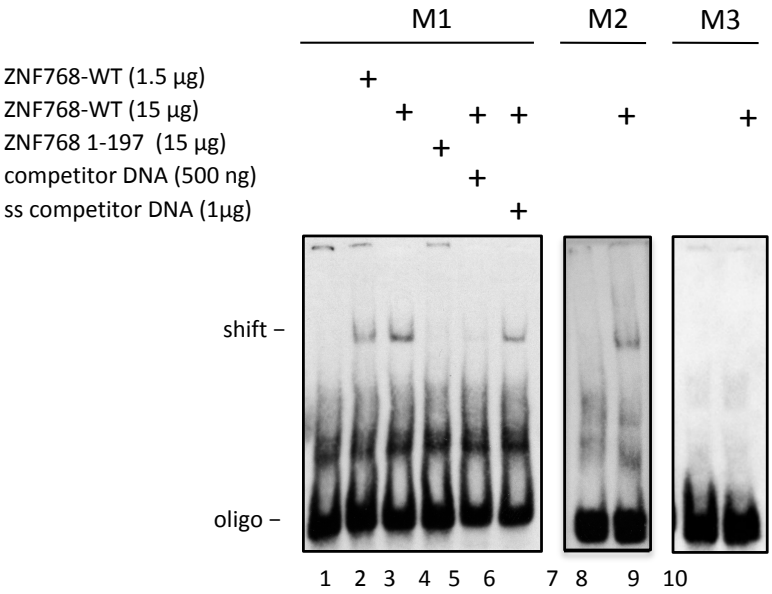
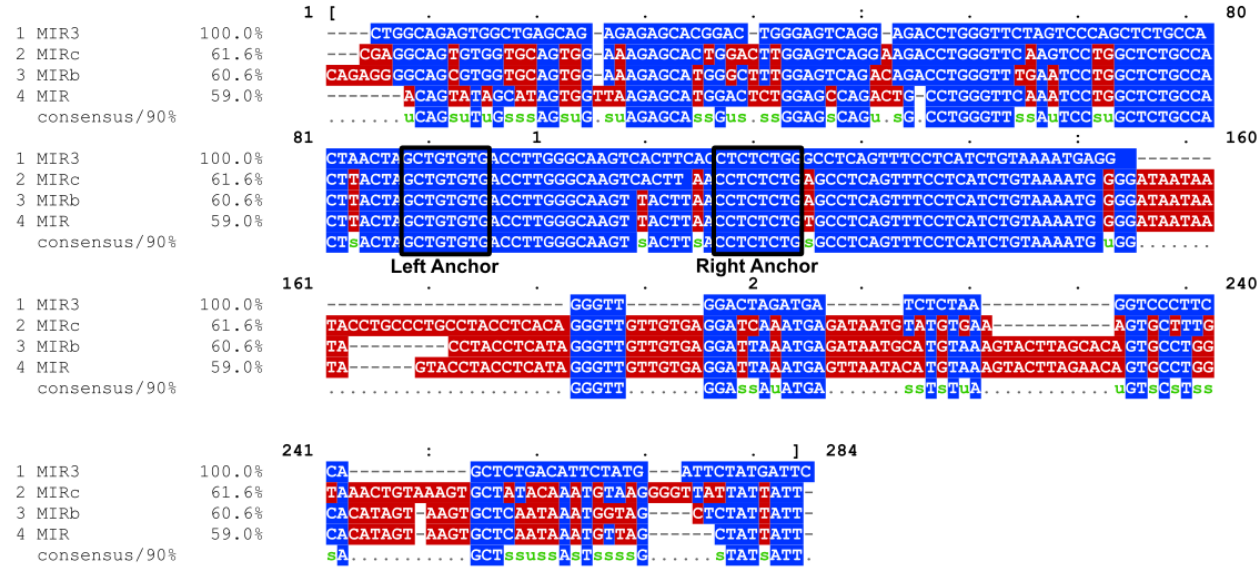


Figure S7

Electrophoretic mobility shift assay (EMSA) with ZNF768 protein. (**A**) Oligonucleotide with the sequence of the ZNF768 binding motif (M1, black sequence), with replacement of the spacer sequence (M2, red sequences), and with partial replacement of the anchor sequences (M3, red sequences). (**B**) Double-stranded fragments M1 – M3 were end-labelled with DIG-11-dUTP and analyzed in extracts with recombinant ZNF768 protein in EMSA.

A



B

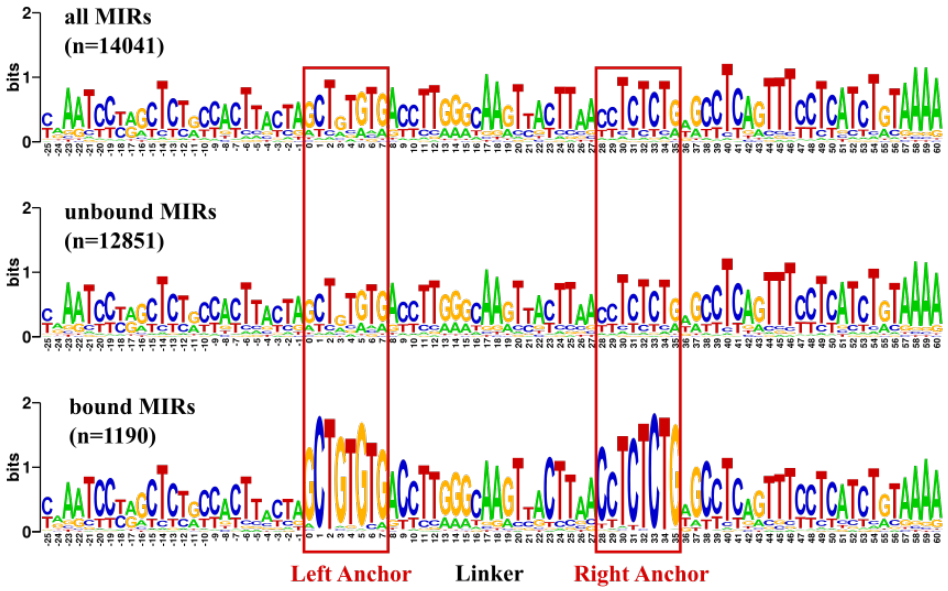
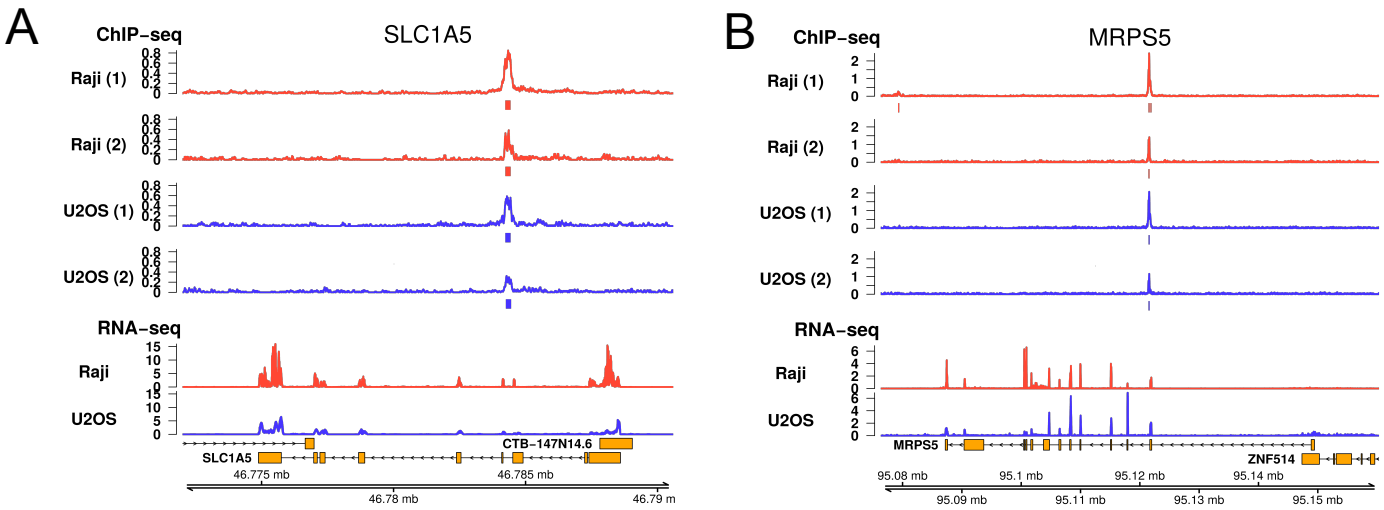


Figure S8

(**A**) Alignment of human MIR3, MIR3c, MIR3b, and MIR consensus sequences. The left and right anchor sequences of the ZNF768 binding motif from Fig. 3A are highlighted by black boxes. (**B**) Motif logos for human MIR sequences that align without gaps to the MIR consensus sequence in the region of the ZNF768 binding motif +25bp on either side (=14,041 MIR sequences). Logos are shown separately for all of these MIRs (top row), MIRs not bound by ZNF768 (middle row) and MIRs bound by ZNF768 (bottom row). MIRs not bound by ZNF768 show no particular conservation in human for the binding motif. The MIR core sequence covers the sequence from 93 nt to 159 nt (Smit and Riggs, Nucleic Acids Res. 23, 98-102).

Common peaks for ZNF768



U2OS specific peaks for ZNF768

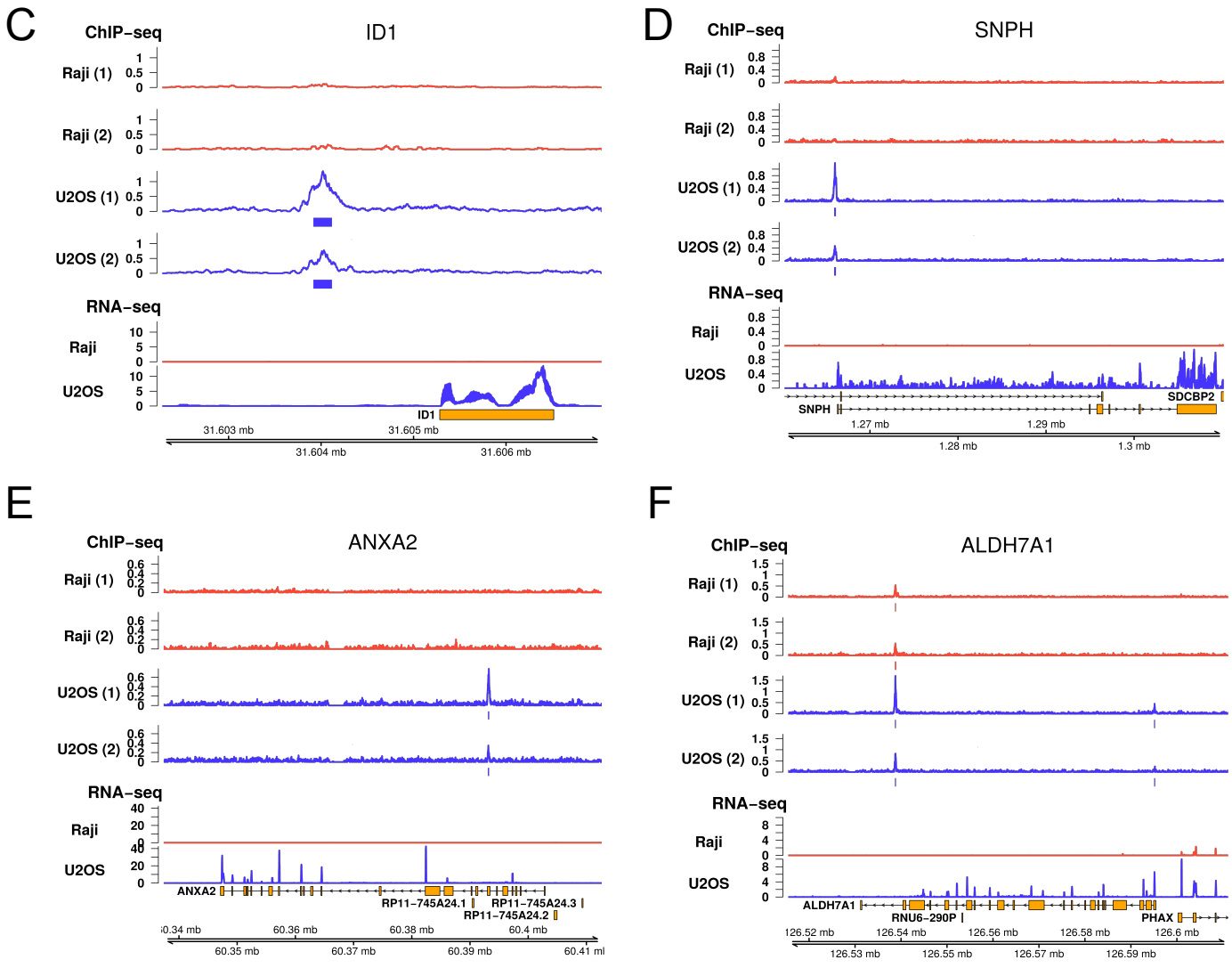


Figure S9

Raji specific peaks for ZNF768

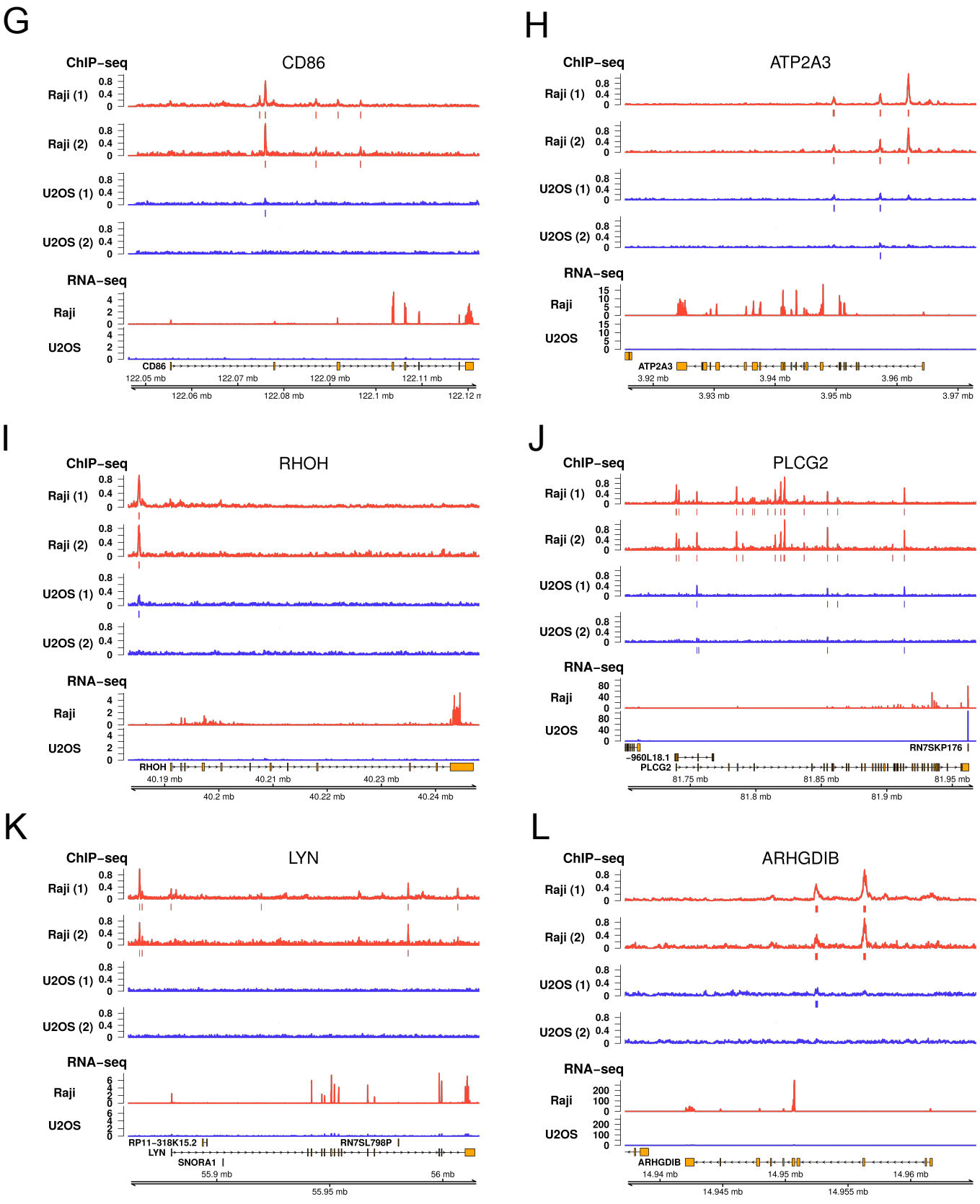


Figure S9

Figure S9

ChIP-seq (replicates shown separately) and RNA-seq (mean of 4 replicates) read coverage (in counts per million) for example genes. Identified peaks are shown as rectangles below the corresponding ChIP-seq sample. Genomic coordinates and gene annotation (boxes=exons, lines=introns, strand indicated by arrowheads) are shown in the bottom row. (**A,B**) Consistent ZNF768 binding in both cell types for the genes Solute Carrier Family 1 Member 5 (SLC1A5) and Mitochondrial Ribosomal Protein S5 (MRPS5). (**C**-**F**) Strong peaks for ZNF768 were associated with the promoter region of the ID1 and SNPH genes and the gene body of the ANXA2 and ALDH7A1 genes in U2OS cells, but are or only faintly visible in Raji cells. (**G**-**L**) Peaks were detected in Raji cells for the genes CD86, ATP2A3, RHOH, PLCG2, LYN, and ARHGDIB. No or weak peaks were identified in U2OS.

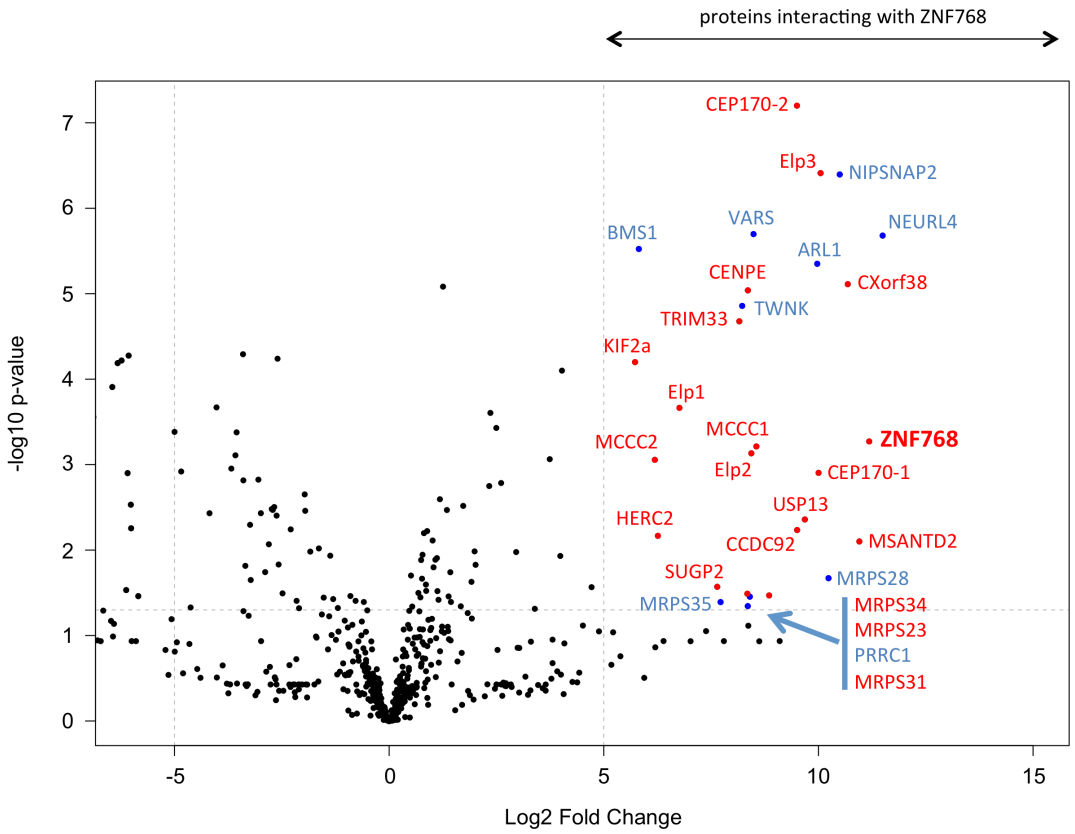


Figure S10

Volcano plot of the ZNF768-specific interactome of Raji cells. Interactions of proteins with log fold change higher than 5 and p-value >0.05 are indicated. Factors depicted in red were detected also among the 30 best interactors in U2OS cells. Dots on the left hand site indicate interactors of a control antibody directed against the nucleolar protein Pes1. Complete list of interactors is shown in Supplementary Table 4.