

SFPQ Directs Histone H3.3 Deposition to R-Loops in DNA Repeats to Protect Genome Stability

Alessandro Ferrando^{1, 6, #}, Michele Giaquinto^{1, #}, Luisa M.R. Napolitano², Giulia Canarutto^{1, 3}, Alessandro Framarini¹, Alice Gambelli¹, Pamela Veneziano Broccia^{1, 4}, Annie Zappone¹, Eleonora Petti^{1, 7}, Chiara Boncristiani¹, Andrea Parlante^{1, 3}, Silvia Onesti², Silvano Piazza^{1, 3}, Roberta Benetti^{5, \$}, Stefan Schoeftner^{1, \$}

1. Dipartimento di Scienze della Vita, Università degli Studi di Trieste, Via E. Weiss 2, 34127 Trieste, Italy.
2. Structural Biology Laboratory, Elettra - Sincrotrone Trieste S.C.p.A, Area Science Park Basovizza, 34149 Trieste, Italy.
3. Computational Biology, International Centre for Genetic Engineering and Biotechnology, 34149 Trieste, Italy.
4. current address, Centro Nacional de Análisis Genómico (CNAG), c/ Baldiri Reixac, 4, Barcelona Science Park - Tower I 08028 Barcelona, Spain.
5. Laboratory of Epigenomics, Department of Medicine, Università degli Studi di Udine, Udine, Italy.
6. current address, Dipartimento di Scienze Cliniche e Biologiche, Università degli Studi di Torino, Regione Gonzole, 10, 10043 Orbassano, Italy
7. current address, IRCCS - Regina Elena National Cancer Institute, Translational Oncology Research Unit, Via Elio Chianesi 53, 00144, Rome, Italy.

These authors contributed equally to the manuscript

\$ Correspondence to: roberta.benetti@uniud.it; sschoeftner@units.it

CONTENTS:

Supplementary figures

Supplementary figure 1: related to main figure 1; Title: Suppression of R-loops and replication stress by SFPQ.

Supplementary figure 2: related to main figure 2; Title: SFPQ *in vitro* binding specificity for nucleic acid substrates.

Supplementary figure 3: related to main figure 3; Title: SFPQ controls DAXX localization and ATRX expression.

Supplementary figure 4: related to main figure 4; Title: SFPQ controls chromatin organization.

Supplementary figure 5: related to main figure 5; Title: Identification of functionally relevant SFPQ domains by deletion analysis.

Supplementary figure 6: related to main figure 6; Title: SFPQ suppresses replication stress and genome instability.

Supplementary figure 7: related to main figure 7; Title: Suppression of R-loops prevents the activation of innate immunity via the cGAS/STING pathway in sarcoma.

Supplementary figure 8: related to main figure 7; Title: Expression of individual signature genes and human sarcoma patient overall survival.

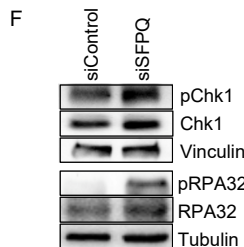
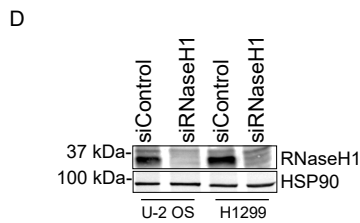
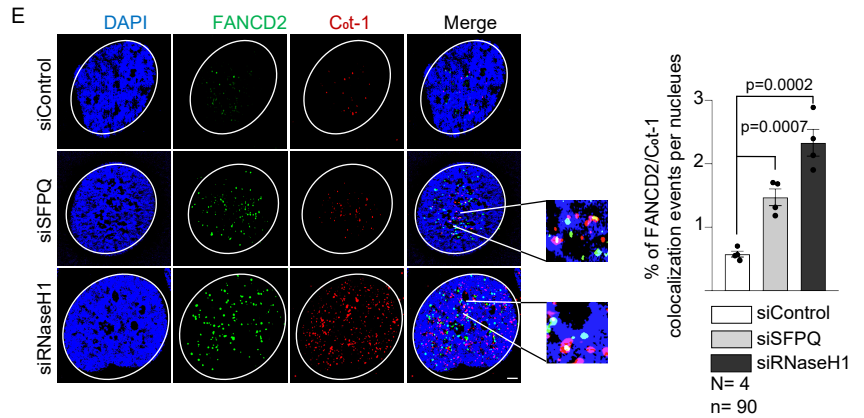
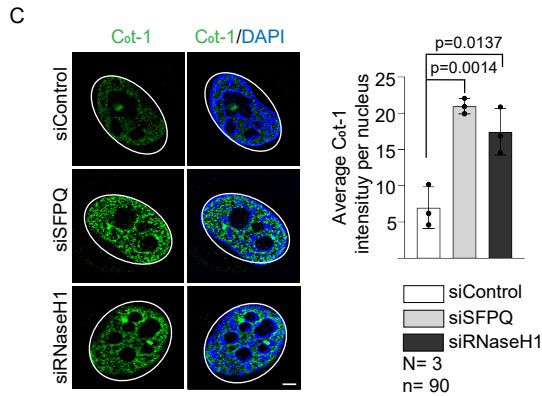
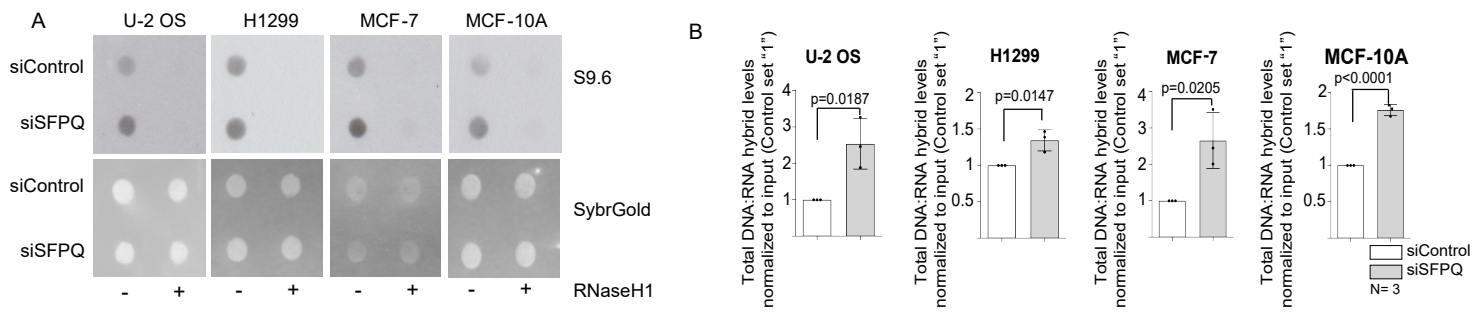
Supplementary figure 9: related to main figure 7; Title: Impact of SFPQ and individual signature genes on human sarcoma patient overall survival.

Supplementary tables:

Supplementary table 1: Title: Signature genes used for Kaplan-Meier analysis in Supplementary Figure 8

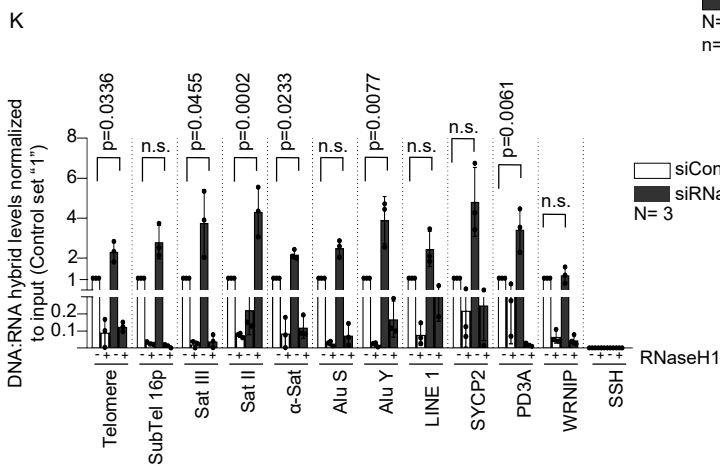
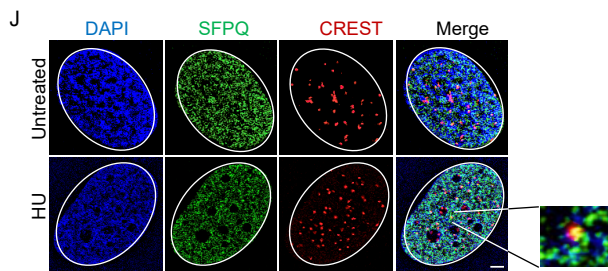
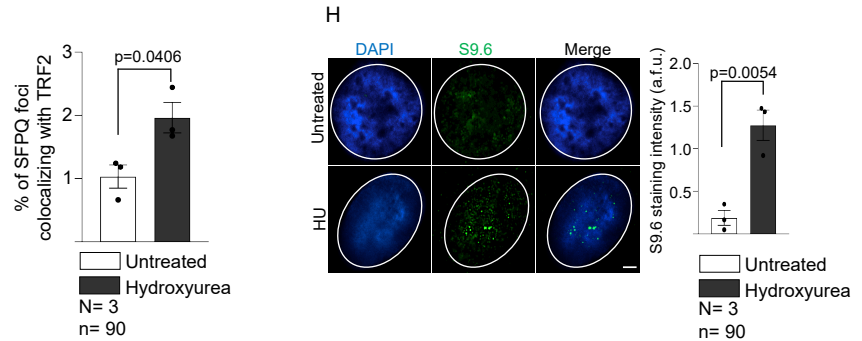
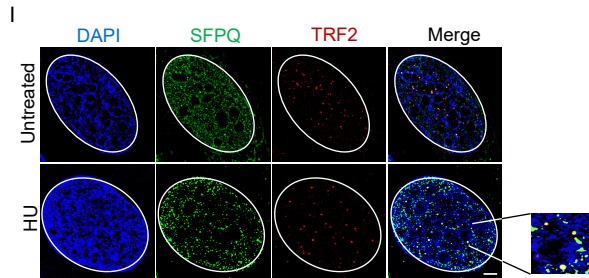
Supplementary table 2: Title: Signature genes used for Kaplan-Meier analysis in Supplementary Figure 9

Supplementary table 3 Title: Synthetic siRNAs used in this study.



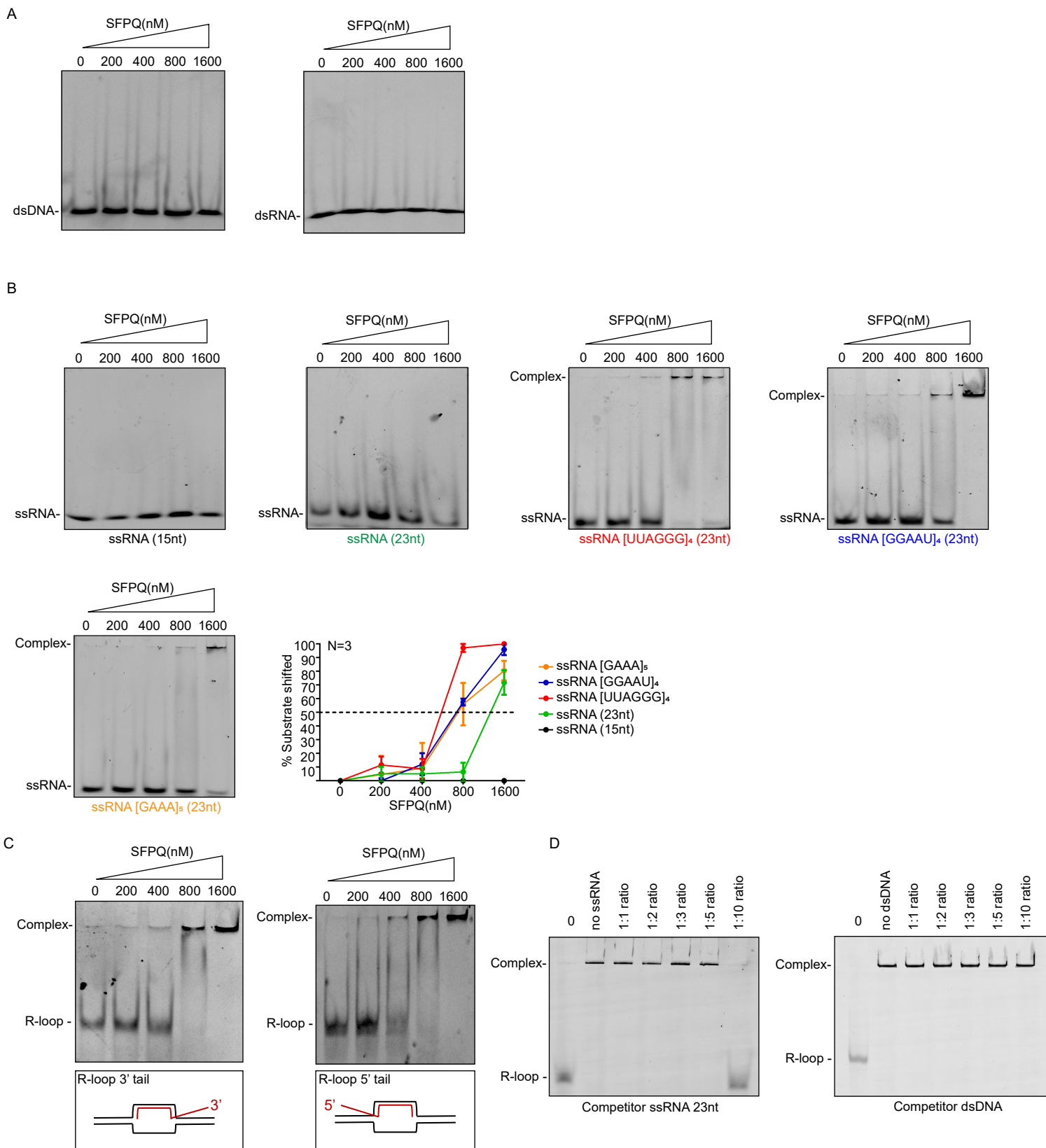
G

U-2 OS			
	Average \pm SD (n=3)	Average \pm SD (n=3)	p-value
Cell cycle phase	siControl (%)	siSFPQ (%)	siCon vs. siSFPQ
subG1 (apoptosis)	1.223 \pm 0.044	2.937 \pm 0.206	0.0002
G0/G1	36.57 \pm 2.857	37.1 \pm 3.200	n.s.
S	19.57 \pm 3.412	16.83 \pm 4.534	n.s.
G2/M	17.43 \pm 0.718	22.87 \pm 1.504	0.0048
>G2	24 \pm 3.377	19.27 \pm 4.565	n.s.



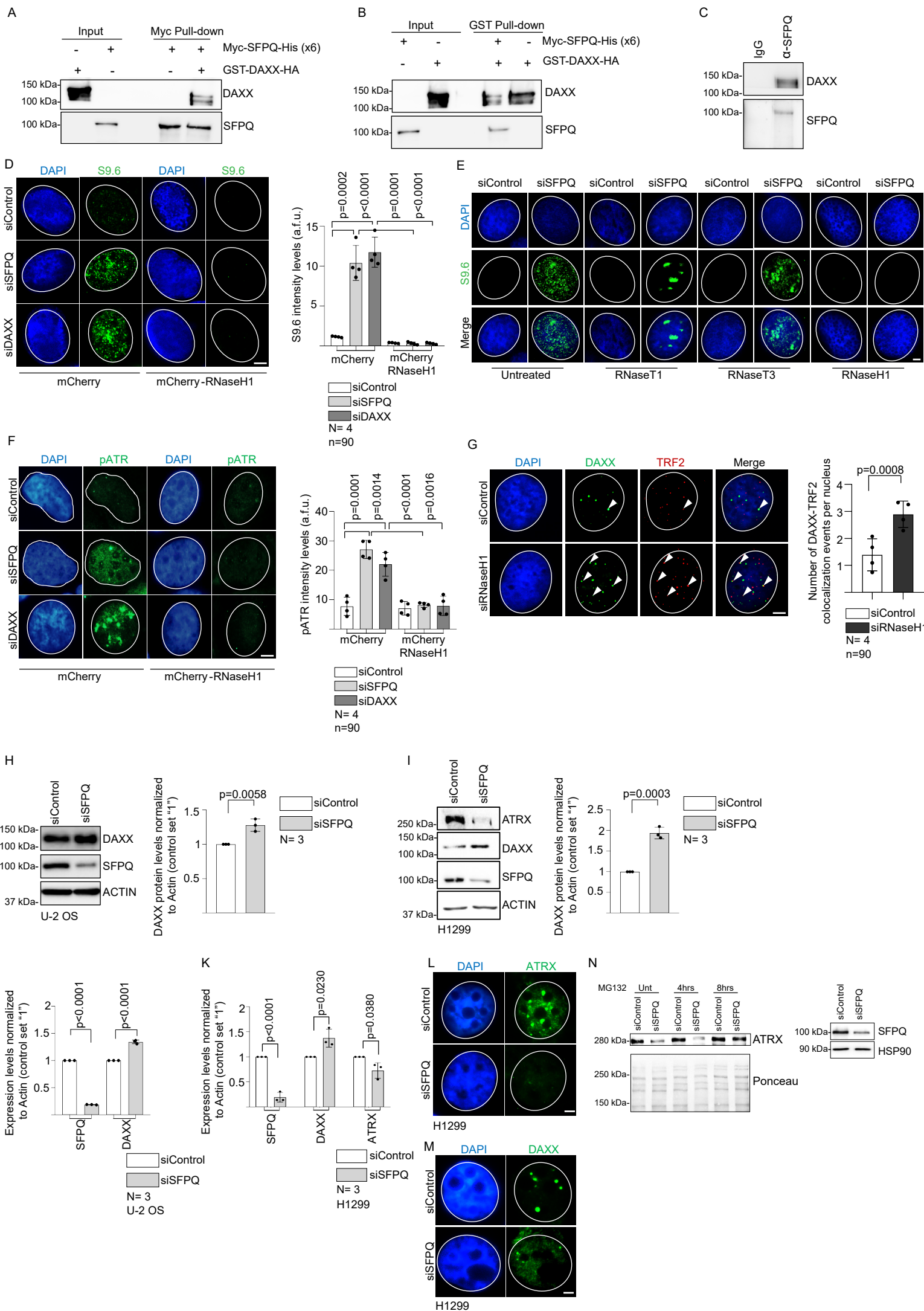
Supplementary Figure 1

- (a) DNA dot-blot experiment using cancer cell lines transfected with indicated siRNAs. Representative images are shown. Prior to loading, control samples were treated with recombinant RNaseH1 to eliminate RNA:DNA hybrids. RNA:DNA hybrids were stained using a monoclonal S9.6 antibodies. As loading control, total DNA was stained using SybrGold.
- (b) Quantifications of dot-blot experiments shown in (a).
- (c) C₀t-1 RNA-FISH performed in H1299 cells transiently transfected with indicated siRNAs (left panel). Quantification of C₀t-1 signal intensity (right panel) as obtained by super-resolution microscopy.
- (d) Western blotting analysis using the indicated antibodies to ensure RNaseH1 knockdown efficiency in both U-2 OS and H1299 cell lines.
- (e) C₀t-1 RNA-FISH combined with anti-FANCD2 immunostaining performed on U-2 OS cells transfected with indicated siRNAs. Images were obtained by super-resolution microscopy. Selected co-localization events are shown as zoomed image. Left panels, representative images. Right panel, quantification of co-localization events per nucleus
- (f) Western blotting analysis using the indicated antibodies. Vinculin and alpha-tubulin specific antibodies were used as loading control.
- (g) Cell cycle flow cytometry analysis of U-2 OS cells transiently transfected with control siRNAs and siRNAs targeting SFPQ.
- (h) Representative image demonstrating increased R-loop levels in hydroxyurea treated U-2 OS cells. In control experiments fixed cells were treated with recombinant RNaseH1 prior to immunofluorescence (not shown). For additional quality controls related to the S9.6 monoclonal antibody, see Supplementary Figure 3e.
- (i) Representative images of combined immunofluorescence with anti-SFPQ and anti-TRF2 antibodies in U-2 OS cells treated with hydroxyurea (HU) or left untreated. Images were obtained by super-resolution microscopy. Right panel, quantification of co-localization events.
- (j) Combined immunofluorescence with anti-SFPQ and anti-CREST (centromere) antibodies in U-2 OS cells treated with hydroxyurea (HU) or left untreated. Right panel, quantification of co-localization events.
- (k) DRIP-qPCR analysis using U-2 OS cells transiently transfected with the indicated siRNAs. Target regions for quantitative PCR analysis are indicated. *SYCP2* and *PD3A* were used as positive control for DAXX and SFPQ, respectively. *WRNIP* and *SSH* were used as negative controls. RNaseH1, treatment of genomic DNA with recombinant RNaseH1 prior to DRIP. Bars represent mean values, error bars indicate standard deviation. N, number of independent experiments. n, number of analysed nuclei. An unpaired, two-sided Student's t-test was used to calculate statistical significance; p-values are shown. Scale bar (1 μ m) applies to all images in respective immunofluorescence panels. DNA was stained using DAPI (4',6-diamidino-2-phenylindole). Source data are provided with this paper.



Supplementary Figure 2

- (a) EMSA using increasing amounts of recombinant SFPQ and dsDNA or dsRNA as nucleic acid substrates.
- (b) EMSA using recombinant SFPQ and the indicated synthetic ssRNAs. Graph shows the percentage of substrates shifted in EMSA experiments.
- (c) EMSA showing binding of recombinant SFPQ to synthetic R-loops containing protruding 3' RNA terminus (left) or 5' RNA terminus (right). Error bars indicate standard deviation; N, number of independent experiments.
- (d) EMSA experiments using labelled R-loop structures and unlabelled competing ssRNA (left panel) or dsDNA (right panel); "0" no recombinant SFPQ added to probe. Molar excess ratio of competitor nucleic acids is indicated. Data in graph are represented as mean values; error bars indicate standard deviation; N= number of independent experiments. Source data are provided with this paper.



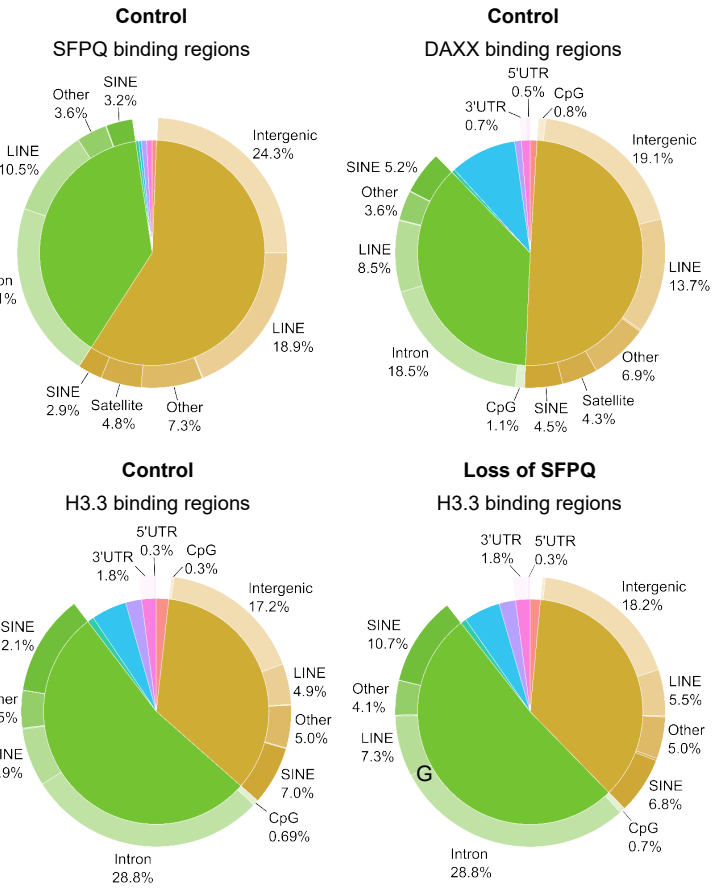
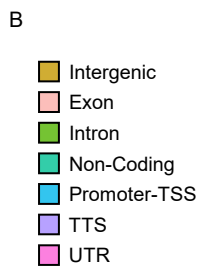
Supplementary Figure 3

- (a) Protein pull-down assay using recombinant His6-myc-SFPQ and GST-DAXX-HA using anti-myc antibody coupled beads. Immunoblots were performed using anti-SFPQ and anti-DAXX antibodies.
- (b) Protein pull-down assay using recombinant His6-Myc-SFPQ and GST-DAXX-HA using glutathione coupled beads. Immunoblots were performed using anti-SFPQ and anti-DAXX antibodies.
- (c) Western blotting of reverse crosslinked proteins obtained from ChIP using antibodies specific for endogenous SFPQ or control IgG. Immunoblots were performed using anti-SFPQ and anti-DAXX antibodies.
- (d) Immunofluorescence staining using monoclonal anti-RNA:DNA hybrid (S9.6) antibodies in U-2 OS cells expressing mCherry or mCherry-RNaseH1 cells after transfection with indicated siRNAs. Left panels, representative images. Right panel, quantification of pan-nuclear immunofluorescence intensity.
- (e) Validation of substrate specificity of S9.6 monoclonal antibodies. Fixed cells were pre-treated with the indicated recombinant enzymes before proceeding with anti-RNA:DNA hybrid immunofluorescence. Only RNaseH1 treatment completely eliminated RNA:DNA hybrid signals.
- (f) Immunofluorescence analysis using anti-p-ATR (Thr1989) antibodies on U-2 OS cells expressing mCherry and mCherry-RNaseH1 after transfection with indicated siRNAs. Left panel, representative images. Right panel, quantification of pan-nuclear immunofluorescence intensity.
- (g) Combined immunofluorescence using anti-DAXX and anti-TRF2 antibodies on U-2 OS cells transfected with indicated siRNAs. Left panel, representative images, arrowheads indicate co-localization events. Right panel, quantification of co-localization events per nucleus.
- (h) Western blotting using whole cells extracts from U-2 OS cells transiently transfected with the indicated siRNAs. Anti-DAXX, anti-SFPQ, and anti-ACTIN specific antibodies were used in western blotting experiments. Right panel, quantification of DAXX protein levels.
- (i) Western blotting using whole cells extracts from H1299 cells transiently transfected with the indicated siRNAs. Anti-ATRX, anti-DAXX, anti-SFPQ, and anti-ACTIN specific antibodies were used in western blotting experiments. Right panel, quantification of DAXX protein levels.
- (j, k) Quantitative real-time PCR on the indicated genes in U-2 OS (j) or H1299 (k) transfected with control and SFPQ specific siRNAs. Actin was used as reference gene.
- (l) Representative images of immunofluorescence analysis using anti-ATRX antibodies. H1299 cells were transiently transfected with indicated siRNAs.
- (m) Representative images of immunofluorescence analysis using anti-DAXX antibodies. H1299 cells were transiently transfected with indicated siRNAs.
- (n) Western blotting of whole cell extracts of H1299 cells transiently transfected with control or SFPQ specific siRNAs. Prior to lysis, cells were treated with MG132 for the indicated time periods (left panel). Right panel, confirmation of SFPQ knock-down using lysates collected prior to MG132 treatment.

In quantification blots, bars represent mean values, error bars indicate standard deviation. N, number of independent experiments. n, number of analysed nuclei. An unpaired, two-sided Student's t-test was used to calculate statistical significance; p-values are shown. Scale bar (1 μ m) applies to all images in respective immunofluorescence panels. DNA was stained using DAPI (4',6-diamidino-2-phenylindole). siRNAs are listed in Supplementary Table 3. Source data are provided with this paper.

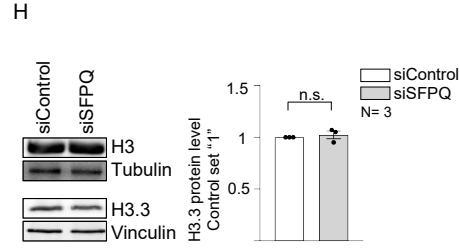
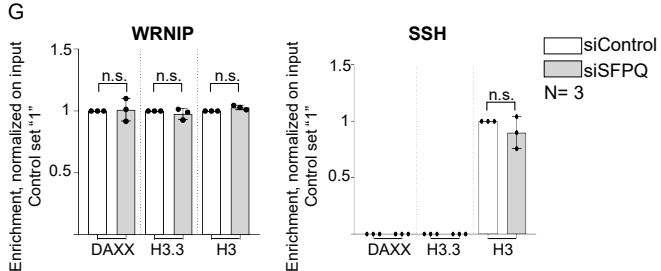
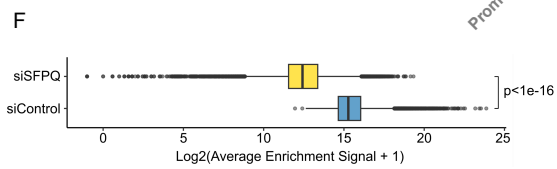
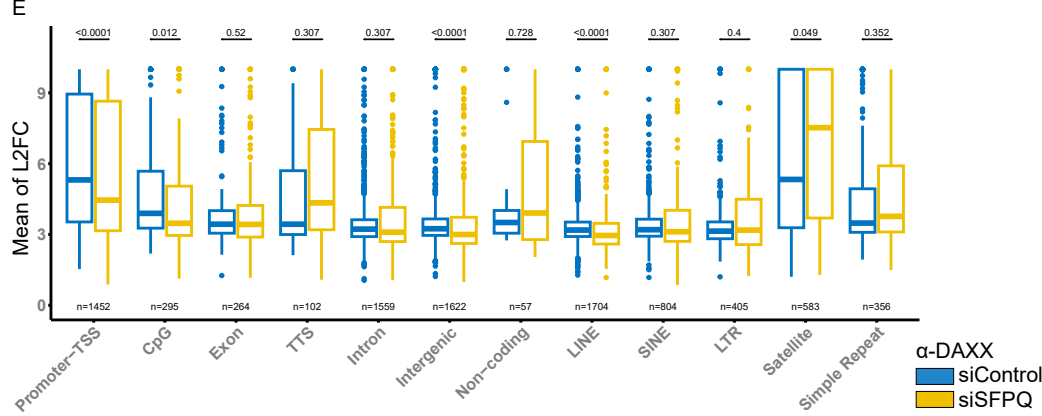
A

	SFPQ	DAXX	H3.3
siControl	6375	6643	29869
siSFPQ	//	2816	55388



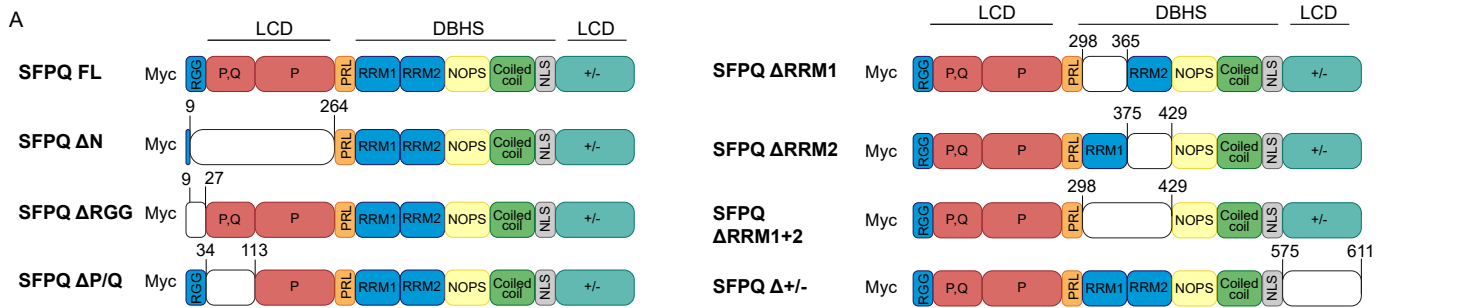
C

SFPQ binding regions (%)			DAXX binding regions (%)			H3.3 binding regions (%)		
Region	siCon		Region	siCon	siSFPQ	Region	siCon	siSFPQ
3'UTR	0.80		3'UTR	0.72	0.50	3'UTR	1.77	1.77
5'UTR	0.02		5'UTR	0.54	1.59	5'UTR	0.32	0.32
CpG	0.09		CpG	1.99	5.34	CpG	1.02	1.02
Exon	0.61		Exon	0.96	2.45	Exon	1.46	1.80
Intergenic	24.30		Intergenic	19.10	15.07	Intergenic	18.14	17.20
Intron	21.16		Intron	18.53	14.85	Intron	28.77	28.79
LINE	29.35		LINE	22.29	11.61	LINE	12.86	11.80
Non-Coding	0.41		Non-Coding	0.50	0.94	Non-Coding	0.82	0.93
Other	8.82		Other	7.77	6.71	Other	8.33	8.57
Promoter-TSS	0.39		Promoter-TSS	9.54	21.12	Promoter-TSS	5.24	5.07
Satellite	4.96		Satellite	4.34	7.35	Satellite	0.29	0.13
Simple Repeat	2.26		Simple Repeat	3.04	3.68	Simple Repeat	1.08	1.17
SINE	6.10		SINE	9.71	7.64	SINE	17.52	19.14
TTS	0.74		TTS	0.96	1.15	TTS	2.38	2.29



Supplementary Figure 4

- (a) Table reporting the number of peaks (e.g. binding sites) of indicated proteins obtained by ChIP-Seq of indicated experimental conditions. Control, U-2 OS cells transiently transfected with control siRNAs; loss of SFPQ, U-2 OS cells transiently transfected with SFPQ specific siRNAs.
- (b) Pie charts representing the distribution of SFPQ (top left panel), DAXX (top right panel), or H3.3 (bottom left panel) peaks across the indicated sequence categories after ChIP-Seq. Chromatin from control siRNA transfected U-2 OS cells was used for the analysis. Bottom right panel, ChIP using chromatin from U-2 OS cells transiently transfected with SFPQ specific siRNAs; distribution of H3.3 peaks across different sequence categories is shown.
- (c) Table reporting the percentage of binding sites of indicated proteins at listed genome regions. Chromatin from control siRNA (siCon) and SFPQ specific siRNAs (siSFPQ) transfected U-2 OS cells was used.
- (d) Venn diagram representing the number of unique or common SFPQ and DAXX peaks obtained by ChIP seq using chromatin from U-2 OS cells transiently transfected with control and SFPQ specific siRNAs.
- (e) Box-plot representing DAXX mean Log₂ Fold Change at reported genomic regions in control (blue) and SFPQ depleted (yellow) U-2 OS cells. n, number of peaks for reported genomic region for both conditions. P-values were calculated with Wilcox test and adjusted for False Discovery Rate. Center, median (50th percentile); box bounds, Q1 (25th percentile) and Q3 (75th percentile); whiskers, minimum and maximum non-outlier values (within 1.5×IQR of Q1/Q3); dots, outliers. Promoter-TSS = Promoter-Transcription Starting Site; CpG = CpG islands; TTS = Transcription Termination Site; LINE = Long Interspersed Nuclear Element; SINE = Short Interspersed Nuclear Elements; LTR = Long Terminal Repeat.
- (f) Box plot representing area under the curve (AUC) calculation for individual H3.3 peaks in control and SFPQ depleted U-2 OS cells. p-value was calculated using the Wilcoxon rank sum test with continuity correction. Center, median (50th percentile); box bounds, Q1 (25th percentile) and Q3 (75th percentile); whiskers, minimum and maximum non-outlier values (within 1.5×IQR of Q1/Q3); dots, outliers.
- (g) Chromatin immunoprecipitation (ChIP) assay performed on U-2 OS cells transiently transfected with indicated siRNAs. The graph reports the binding of DAXX, H3.3, and H3 to negative control regions in *WRNIP* (left panel) and *SSH* genes (right panel). Alterations are shown as fold-change between control and SFPQ knock-down U-2 OS cells.
- (h) Western blotting of U-2 OS cells transiently transfected with the indicated siRNAs. Loss of SFPQ does not impact on H3.3 or H3 protein levels.
- In quantification blots g and h, bars represent the mean values, error bars indicate standard deviation. N= number of independent experiments. An unpaired, two-sided Student's t-test was used to calculate statistical significance; p-values are shown. Source data are provided with this paper.



Supplementary Figure 5

(a) Schematic representation of SFPQ domains and mutants. Full-length (FL) SFPQ is composed of the listed domains: arginine-glycine-glycine (RGG) box (blue); N- and C-terminal low-complexity domains (LCDs); proline/glutamine-rich subdomain (P,Q, red); proline-rich domain (P, red); PR linker (PRL, orange); the DBHS conserved region, containing RNA-recognition motifs (RRM1 and RRM2, blue), NonA/paraspeckles (NOPS) domain (yellow), the coiled-coil domain (green), a nuclear-localization sequence (NLS) (in grey) and the +/- domain (bright green). Domains that were deleted in the individual expression constructs are shown as white boxes. Numbers indicate amino acid positions in full-length SFPQ.

(b) Immunoprecipitation experiments using U-2 OS cells transiently co-transfected with an expression vector for full-length, HA-tagged DAXX and a control vector (pcDNA) or expression vectors encoding the indicated, myc-tagged SFPQ versions (FL, full length). Immunoblots were performed using anti-myc and anti-HA antibodies.

(c) Schematic representation of expression vectors encoding for the myc-tagged SFPQ P-domain or nFS-domain, fused with a nuclear localization signal (NLS) and a GFP open reading frame.

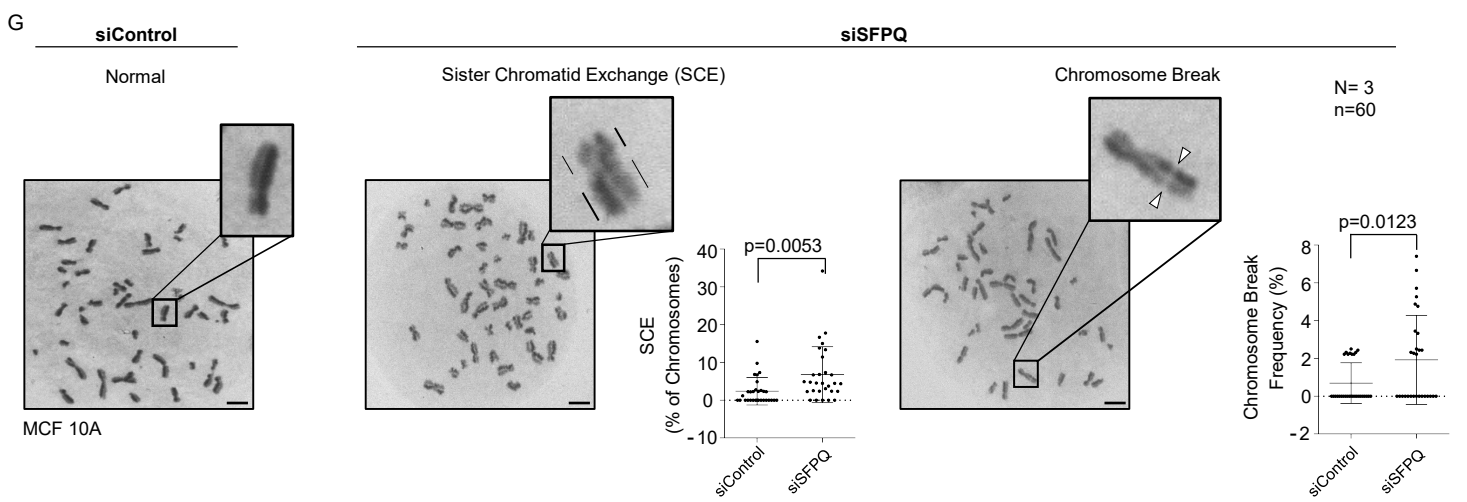
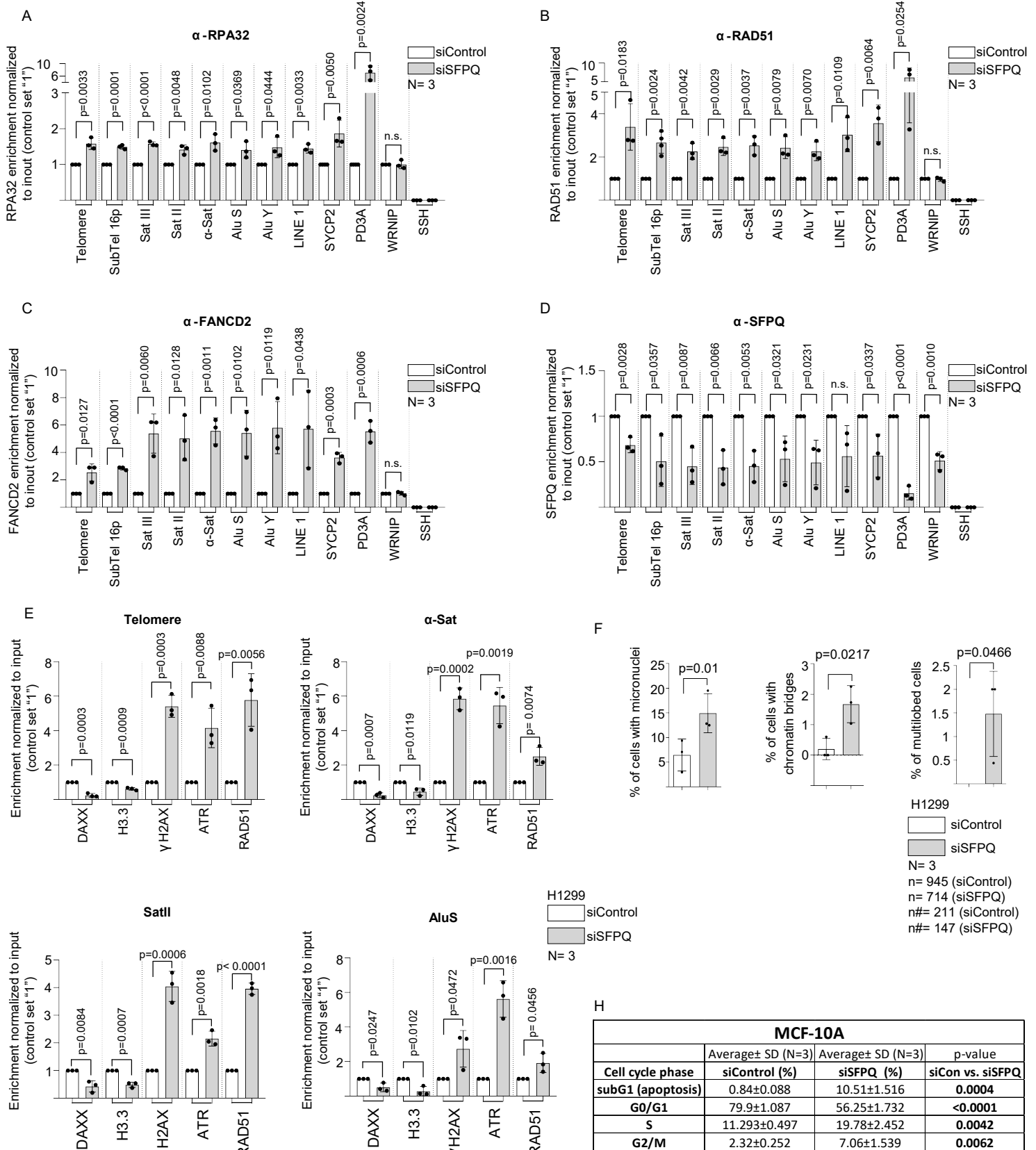
(d) Immunoblot performed using whole cell extracts from U-2 OS cells stably expressing myc-tagged full length (FL) or listed SFPQ mutants that were transiently transfected with control and SFPQ specific siRNAs. Western blotting was performed using the indicated antibodies.

(e) Colocalization of TRF2 and phosphorylated ATR (Thr1989) in U-2 OS cells transiently overexpressing dominant negative SFPQ (SFPQ dP and SFPQ nFS, left panel). Right panel, quantification of immunofluorescence data.

(f) Immunofluorescence analysis of U-2 OS cells stably expressing myc-tagged SFPQ mutants after transient transfection with control siRNAs or siRNAs specifically targeting the 5'UTR of endogenous SFPQ. Anti-myc and anti-p-ATR(Thr1989) specific antibodies were used for immunodetection. Representative images are shown. Experiments were carried out in parallel with data shown in Fig. 5e. Quantification is shown in Fig. 5f and is limited to myc-positive cells.

(g) Immunofluorescence analysis of U-2 OS cells stably expressing myc-tagged SFPQ mutants after transient transfection with control siRNAs or siRNAs specifically targeting the 5'UTR of endogenous SFPQ. Anti-myc and anti-DAXX specific antibodies were used for immunodetection. Representative images are shown. Experiments were carried out in parallel with data shown in Fig. 5g.

(e-g) For quantifications, data represents the mean, error bars indicate standard deviation. N= number of independent experiments. n= number of analysed nuclei. An unpaired, two-sided Student's t-test was used to calculate statistical significance. Scale bar (1 μ m) applies to all images in respective immunofluorescence panels. DNA was stained using DAPI (4',6-diamidino-2-phenylindole). Source data are provided with this paper.



Supplementary Figure 6

(a - d) Chromatin immunoprecipitation (ChIP) assay experiments performed on U-2 OS cells transiently transfected with control siRNAs and siRNAs targeting SFPQ. Quantitative real-time PCR was carried out using specific primer pairs for the indicated genome regions. *PD3A* and *WRNIP* were used as positive control for SFPQ and DAXX binding, respectively. *SSH* was used as negative control. Graphs reports the binding of RPA32 (a), RAD51 (b), FANCD2 (c), and SFPQ (d) to the indicated regions. Fold change values were calculated using control siRNA transfected cells as reference. Amplification values from ChIP using IgG from pre-immune serum was subtracted during normalization.

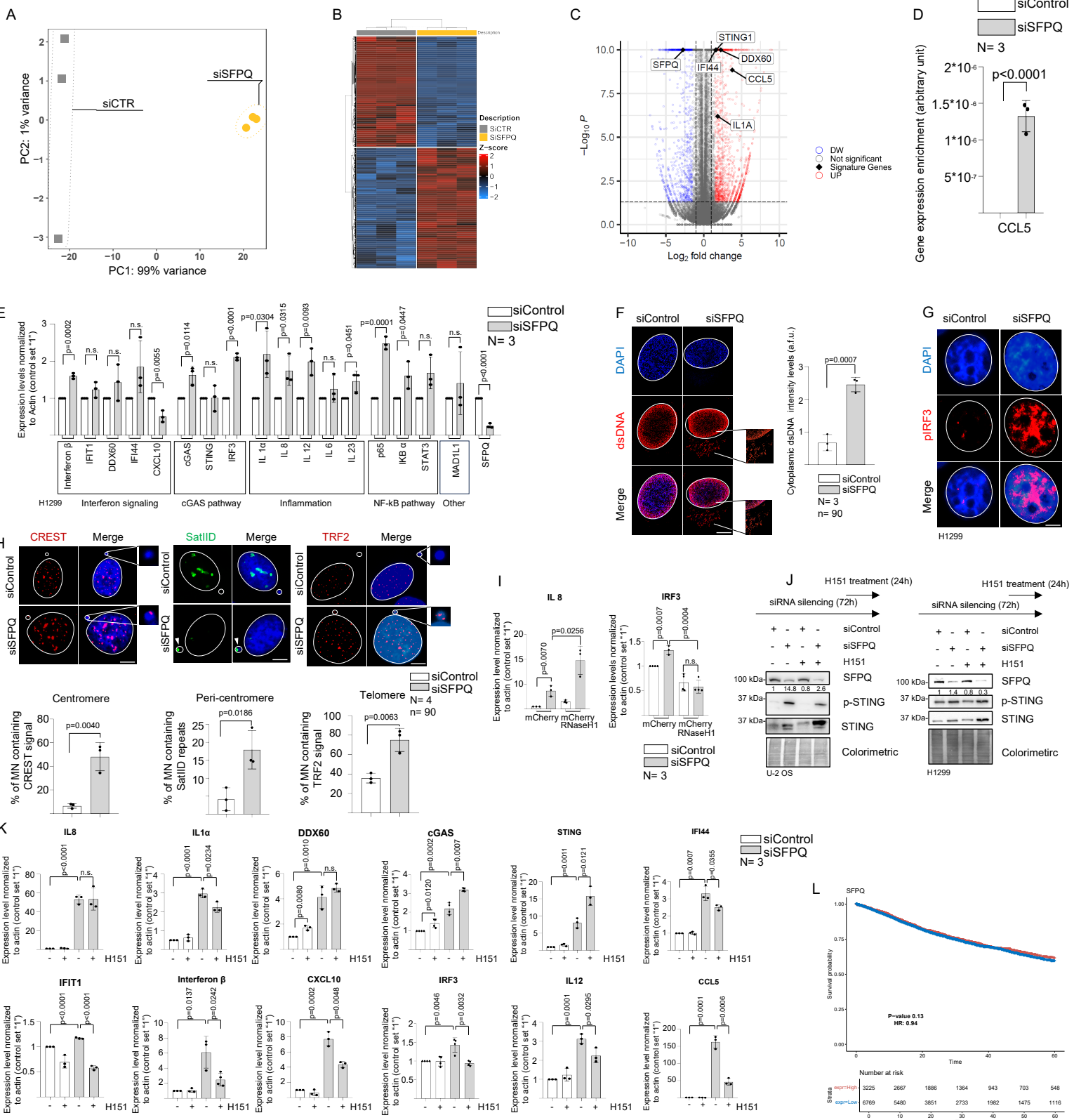
(e) Chromatin immunoprecipitation (ChIP) assay using chromatin from H1299 cells transiently transfected using control siRNAs and siRNAs targeting SFPQ. Graphs report the binding of DAXX, H3.3, γ H2AX, ATR, and RAD51 to the indicated genome region of interest, as detected by quantitative real-time PCR. Enrichment is expressed as fold-change values normalize to control siRNA transfected H1299 cells. Amplification values from ChIP using IgG from pre-immune serum was subtracted during normalization.

(f) Quantification of time-lapse microscopy experiments using H1299 cells stably expressing GFP-tagged H2B and transiently transfected with control siRNAs and siRNAs targeting SFPQ. Frequency of micronuclei, chromatin bridges and multilobal cells are shown. n, number of analysed nuclei for the micronuclei experiment; n#=# number of mitotic cells analysed for the chromatin bridges and multilobal cells.

(g) Metaphase chromosome aberrations observed in MCF 10A cells transiently transfected with control siRNAs and siRNAs targeting SFPQ. Representative Giemsa-stained, normal metaphase chromosomes (siControl) and chromosomes from SFPQ knock-down cells with increased sister chromatid exchange (SCE) rate and chromosomal breaks are shown. Fat and thin lines along chromosome indicate a representative event of sister chromatid exchange. Empty arrowheads indicate position of chromosome break. Graphs show the frequency of chromosome aberration events; n= number of analysed metaphases. Scale bar (1 μ m) applies to all images. DNA was stained using Giemsa.

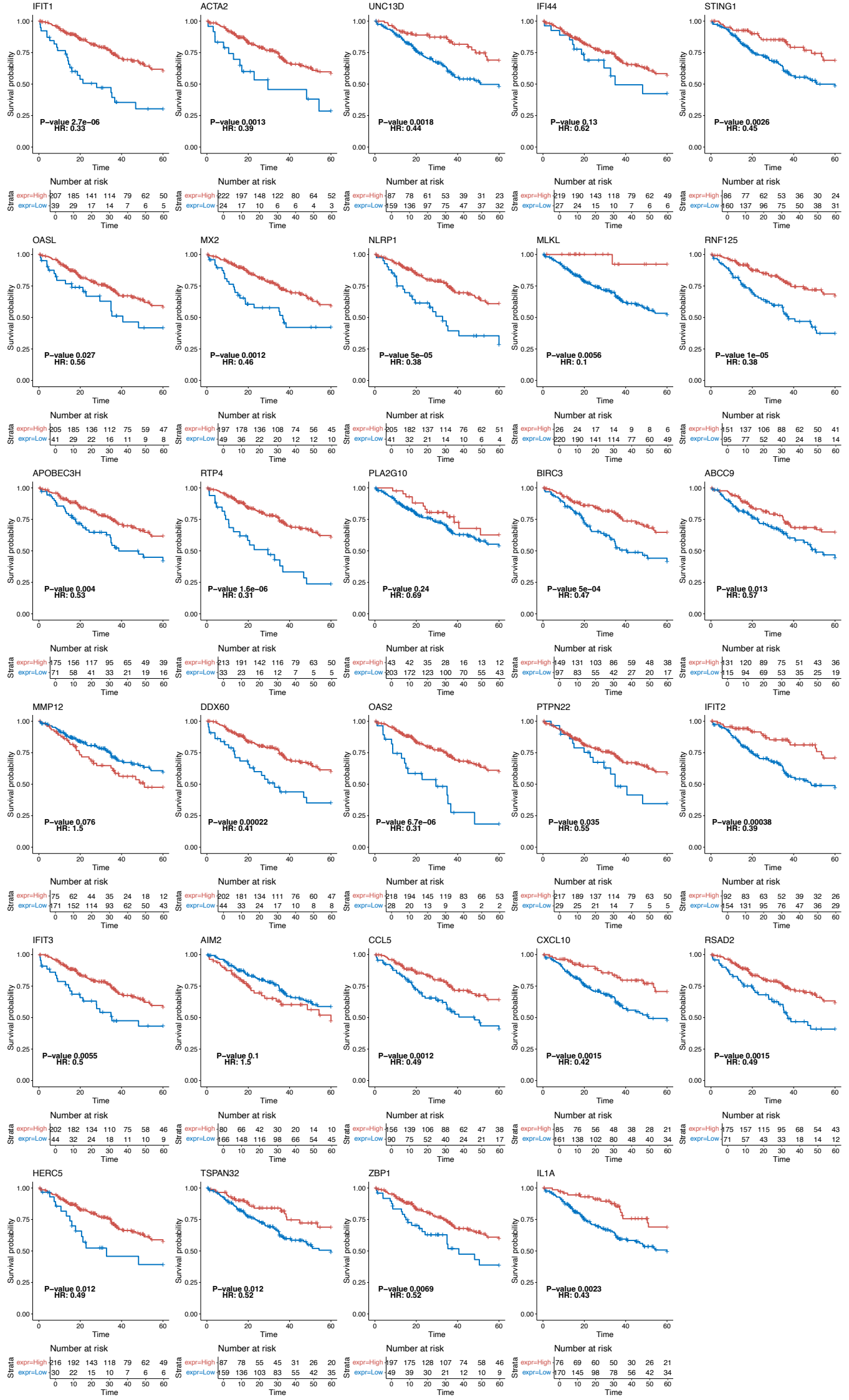
(h) Cell cycle flow cytometry analysis of MCF 10A cells transiently transfected with control siRNAs and siRNAs targeting SFPQ

For quantifications, bars represent mean values, error bars indicate standard deviation. N= number of independent experiments. An unpaired, two-sided Student's t-test was used to calculate statistical significance; p-values are shown. Source data are provided with this paper.



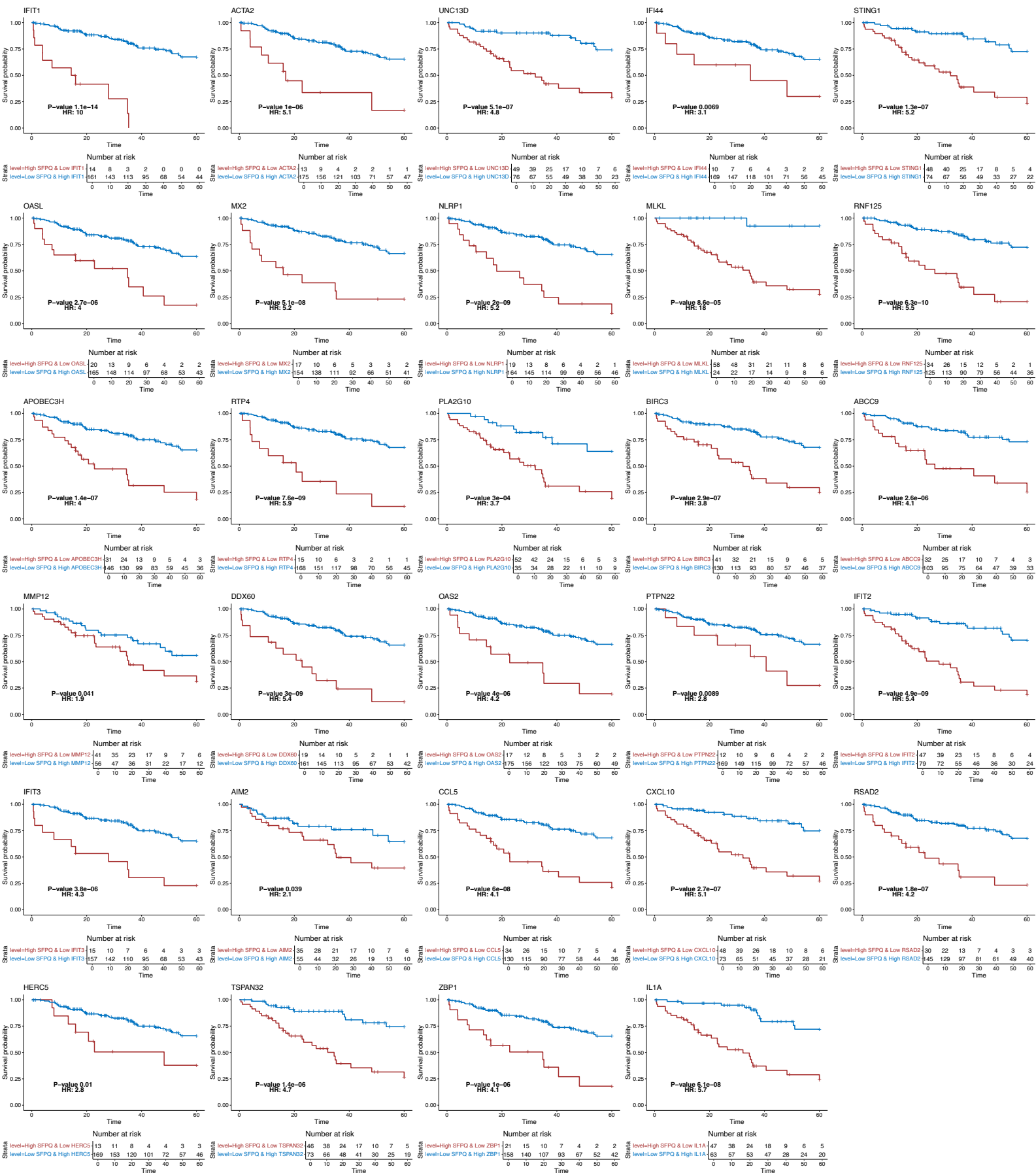
Supplementary Figure 7

- (a) Principal Component Analysis (PCA) performed on RNA-Seq experiment using three biological replicates. U-2 OS cells were transfected with control siRNA and siRNA targeting SFPQ. PC1 shows 99% variance and PC2 1%. Yellow dots correspond to SFPQ knock-down U-2 OS cells, grey squares correspond to U-2 OS cells transfected with control siRNAs.
- (b) Heatmap of Z-score of differentially expressed genes (DEGs) after siRNA mediated depletion of SFPQ from U-2 OS cells silenced with indicated siRNA. Clustering was performed using Euclidean distance and “average” as clustering method. Red, upregulated genes; blue, downregulated genes.
- (c) Volcano plot representing genes that are differentially expressed in U-2 OS cells transfected with control and SFPQ specific siRNAs. Selected genes from the innate signature genes (Fig. 7b) and SFPQ are highlighted. Red, upregulated genes; blue, downregulated genes; gray, not significantly altered, black diamond, innate immunity signature gene.
- (d) Quantitative RT-PCR of CCL5 in U-2 OS cells after transient siRNA mediated depletion of SFPQ
- (e) Quantitative RT-PCR of indicated genes in H1299 cells transfected with control or SFPQ specific siRNAs. N, number of independent experiments. An unpaired, two-sided Student's t-test was used to calculate statistical significance; p-values are shown.
- (f) Immunofluorescence analysis using specific antibodies detecting dsDNA in the cytosol of experimental U-2 OS cells. Left panels, representative images; right panel, quantification of cytoplasmic dsDNA signal intensity.
- (g) Immunofluorescence analysis using specific antibodies for the phosphorylated version of IRF3 (pIRF3) on H1299 cells transiently transfected with control siRNAs and siRNAs targeting the SFPQ mRNA. Representative images are shown.
- (h) Immunofluorescence and DNA-FISH analyses on U-2 OS cells transiently transfected with control and SFPQ specific siRNAs; representative images are shown in panels on the left. Left panels, anti-CREST; central panel, SatIID DNA FISH; right panels, anti-TRF2 antibodies. Arrowheads indicate micronuclei in representative images. Quantification of micronuclei positive for CREST antigen (bottom left), SatIID DNA (bottom central) or TRF2 (bottom right) are shown.
- (i) U-2 OS cells with doxycycline inducible expression constructs for mCherry or mCherry-RNaseH1 were transiently transfected with the indicated siRNAs. IL-8 and IRF3 expression was monitored by quantitative RT-PCR.
- (j) Representative western-blot of U-2 OS (left) or H1299 (right) cells transfected with indicated siRNAs. Experimental cells were treated with the STING inhibitor H151 or left untreated. Western blotting was performed using the indicated antibodies.
- (k) Quantitative real-time PCR using cDNA from U-2 OS cells transfected with indicated siRNAs and treated with the STING inhibitor H151, or left untreated.
- (l) Kaplan-Meier plot analysis showing no correlation between SFPQ expression and patient overall survival using data from TCGA All Pancancer data. Patients were categorized based on the expression levels of SFPQ and their survival was censored at 60 months. HR is the hazard ratio calculated by the Cox regression model. HR p-value is 0.127. Patient numbers are indicated.
- For quantifications, bars represent mean values, error bars indicate standard deviation. N= number of independent experiments, n= number of analyzed nucleus. An unpaired, two-sided Student's t-test was used to calculate statistical significance. Scale bar (1 μ m) applies to all images in respective immunofluorescence panels. DNA was stained using DAPI (4',6-diamidino-2-phenylindole). Source data are provided with this paper.



Supplementary Figure 8

Kaplan-Meier plot analysis showing correlation between higher signature genes expression and improved patient overall survival using data from TCGA pan-cancer sarcoma data set. Patients were categorized based on the expression levels of SFPQ and their survival was censored at 60 months. Hazard ratios, hazard ratios p-values and curves p-values are shown in table. SFPQ dependent, innate immunity gene expression signature genes were obtained by RNA-Seq (Figure 7b; *IFIT1*, *ACTA2*, *UNC13D*, *IFI44*, *STING1*, *OASL*, *MX2*, *NLRP1*, *MLKL*, *RNF125*, *APOBEC3H*, *RTP4*, *PLAG2G10*, *BIRC3*, *ABCC9*, *MMP12*, *DDX60*, *OAS2*, *PTPN22*, *IFIT2*, *IFIT3*, *AIM2*, *CCL5*, *CXCL10*, *RSAD2*, *HERC5*, *TSPAN32*, *ZBP1* and *IL1A*).



Supplementary Figure 9

Kaplan-Meier curves showing human sarcoma patient overall survival. Patients were stratified for high SFPQ expression combined with low expression of individual innate immunity signature genes (High SFPQ & Low Signature gene) or low SFPQ expression paired with concomitant high expression of individual innate immunity signature gene (Low SFPQ & High Signature gene). Signature genes are shown in Fig. 7b (*IFIT1*, *ACTA2*, *UNC13D*, *IFI44*, *STING1*, *OASL*, *MX2*, *NLRP1*, *MLKL*, *RNF125*, *APOBEC3H*, *RTP4*, *PLAG2G10*, *BIRC3*, *ABCC9*, *MMP12*, *DDX60*, *OAS2*, *PTPN22*, *IFIT2*, *IFIT3*, *AIM2*, *CCL5*, *CXCL10*, *RSAD2*, *HERC5*, *TSPAN32*, *ZBP1* and *IL1A*). Data was obtained from the TCGA human sarcoma dataset. HR, hazard ratio; p-values are shown.

Genes	Hazard Ratio	Cox p-value	Curve p-value
IFIT1	0,33	7,96714E-06	2,67369E-06
ACTA2	0,39	0,002000781	0,001348573
UNC13D	0,44	0,002440335	0,001841574
IFI44	0,62	0,133092185	0,129863035
STING1	0,45	0,00327299	0,00255384
OASL	0,56	0,028964038	0,026911402
MX2	0,46	0,001588507	0,001210675
NLRP1	0,38	9,47749E-05	4,98899E-05
MLKL	0,1	0,024248661	0,005641339
RNF125	0,38	2,13785E-05	1,02687E-05
APOBEC3H	0,53	0,004594883	0,003959534
RTP4	0,31	5,48798E-06	1,55978E-06
PLA2G10	0,69	0,245223767	0,241800774
BIRC3	0,47	0,000674768	0,000497186
ABCC9	0,57	0,014039285	0,012818809
MMP12	1,5	0,077323832	0,075508698
DDX60	0,41	0,000343743	0,000216458
OAS2	0,31	2,05849E-05	6,71604E-06
PTPN22	0,55	0,037546475	0,035058001
IFIT2	0,39	0,000613407	0,000383493
IFIT3	0,5	0,006489905	0,005501578
AIM2	1,5	0,106869681	0,104716246
CCL5	0,49	0,001499657	0,001194913
CXCL10	0,42	0,002039146	0,00147084
RSAD2	0,49	0,001891899	0,001515908
HERC5	0,49	0,014489297	0,012457109
TSPAN32	0,52	0,01327683	0,011676956
ZBP1	0,52	0,007947196	0,00693515
IL1A	0,43	0,003033951	0,002252034

Supplementary table 1:**Signature genes used for Kaplan-Meier analysis in Supplementary Figure 8**

For each gene in the signature, survival association was assessed using a Cox proportional hazards regression model (coxph), with gene expression dichotomized into High vs Low groups (cutpoints automatically determined using surv_cutpoint). Hazard ratios (HRs) and corresponding two-sided Wald test P values are reported. In parallel, Kaplan-Meier survival curves were compared using a two-sided log-rank (Mantel-Cox) test, and the resulting p-values are shown. Survival times were administratively censored at 60 months. Unless otherwise stated, nominal p-values are reported and no multiple-comparison adjustment was applied.

Genes	Hazard Ratio	Cox p-value	Curve p-value
IFIT1	10	3,21161E-10	1,08625E-14
ACTA2	5,1	1,12685E-05	1,03729E-06
UNC13D	4,8	4,71427E-06	5,09283E-07
IFI44	3,1	0,01012175	0,006871674
STING1	5,2	1,78702E-06	1,33046E-07
OASL	4	1,41251E-05	2,7003E-06
MX2	5,2	9,88878E-07	5,08101E-08
NLRP1	5,2	7,30442E-08	2,01602E-09
MLKL	18	0,004471184	8,56785E-05
RNF125	5,5	3,07331E-08	6,26056E-10
APOBEC3H	4	1,03567E-06	1,36556E-07
RTP4	5,9	3,43473E-07	7,55301E-09
PLA2G10	3,7	0,000704387	0,000298946
BIRC3	3,8	1,71543E-06	2,88659E-07
ABCC9	4,1	1,37089E-05	2,55617E-06
MMP12	1,9	0,04455394	0,041021578
DDX60	5,4	1,16745E-07	2,99843E-09
OAS2	4,2	2,2216E-05	3,99467E-06
PTPN22	2,8	0,012278704	0,008909227
IFIT2	5,4	1,41978E-07	4,93325E-09
IFIT3	4,3	2,20086E-05	3,80746E-06
AIM2	2,1	0,044136418	0,039442714
CCL5	4,1	5,61442E-07	6,02353E-08
CXCL10	5,1	3,315E-06	2,68284E-07
RSAD2	4,2	1,48505E-06	1,83241E-07
HERC5	2,8	0,013556484	0,010019077
TSPAN32	4,7	1,11456E-05	1,40928E-06
ZBP1	4,1	6,5599E-06	1,0193E-06
IL1A	5,7	1,35056E-06	6,12375E-08

Supplementary table 2:**Signature genes used for Kaplan-Meier analysis in Supplementary Figure 9**

For each gene in the signature, survival association was assessed using a Cox proportional hazards regression model (coxph), with gene expression dichotomized into High vs Low groups (cutpoints automatically determined using surv_cutpoint). Hazard ratios (HRs) and corresponding two-sided Wald test P values are reported. In parallel, Kaplan-Meier survival curves were compared using a two-sided log-rank (Mantel-Cox) test, and the resulting p-values are shown. Survival times were administratively censored at 60 months. Unless otherwise stated, nominal p-values are reported and no multiple-comparison adjustment was applied.

Target	Company	Cod/Sequence
Control-non target	Dharmacon	ON-TARGETplus Non-targeting Control siRNAs (D-001810-01-05)
Human SFPQ	Dharmacon	ON-TARGETplus Human SFPQ siRNA (L-006455-00-0005)
Human SFPQ 5'-UTR	Eurofins	5'- CCACGUUCCUGAGCGUCU(dTdT)-3' This paper
Human SFPQ 5'UTR scramble	Eurofins	5' - GTTCTGCTGCCGTCCTA (dTdT) – 3
Human DAXX	Eurofins	5'- GGAGUUGGAUCUCUCAGAA(dTdT)-3' Chen et al., 2003 [10]
Human RNaseH1	Eurofins	5'-ACAAACCAAAGAGCGGAAAUUCAUG(dTdT) -3' (Arora R. et al. 2014 [11])

Supplementary table 3: Synthetic siRNAs used in this study