A

Metric	Raw result	Interpretation
PhyloCSF score	-85.6006	non-coding
CPAT coding probability	3.46%	non-coding
CPC	-1.20592	non-coding

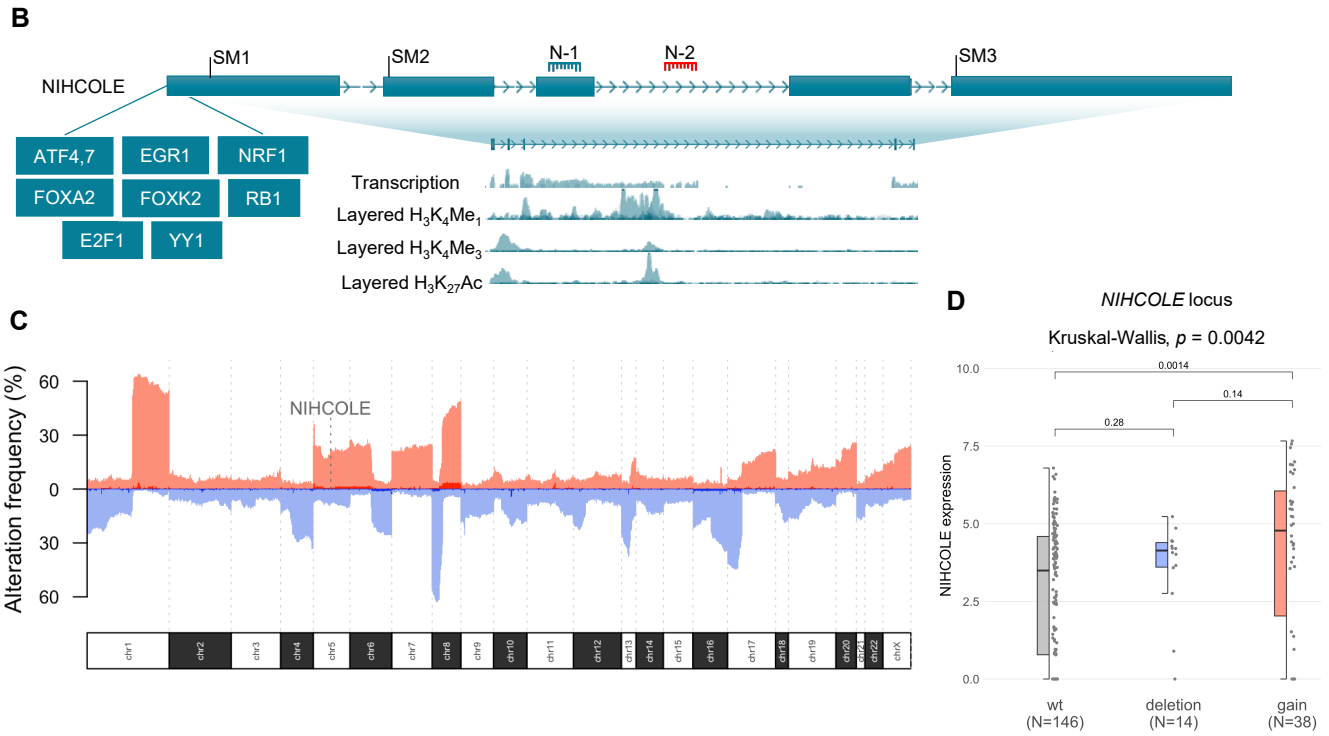


Figure S1. NIHCOLE RNA and *NIHCOLE* locus characterization. A Evaluation of NIHCOLE's coding potential using three different metrics. Raw score and interpretation are shown. **B** Scheme of NIHCOLE's locus. Exons are solid bars; arrowed lines depict introns and transcription direction. ChiP-Seq studies performed by ENCODE have identified transcription factors (depicted in boxes) and H3K4Me1, H3K4Me3 and H3K27Ac tracks (obtained from UCSC). Gapmers are depicted as N-1 and N-2 and their location in exon 3 and intron 3, respectively, is indicated. The relative location of structural motifs SM1, SM2 and SM3 on exons one, two and five respectively, is also shown. **C** Frequency of amplification (red) and deletion (blue) of HCC genomes in the INSERM cohort. NIHCOLE genomic location is indicated. **D** Correlation of NIHCOLE expression with copy number alterations. N = number of samples. p = the significance of the statistical analysis (Kruskal-Wallis ANOVA-test) is indicated.

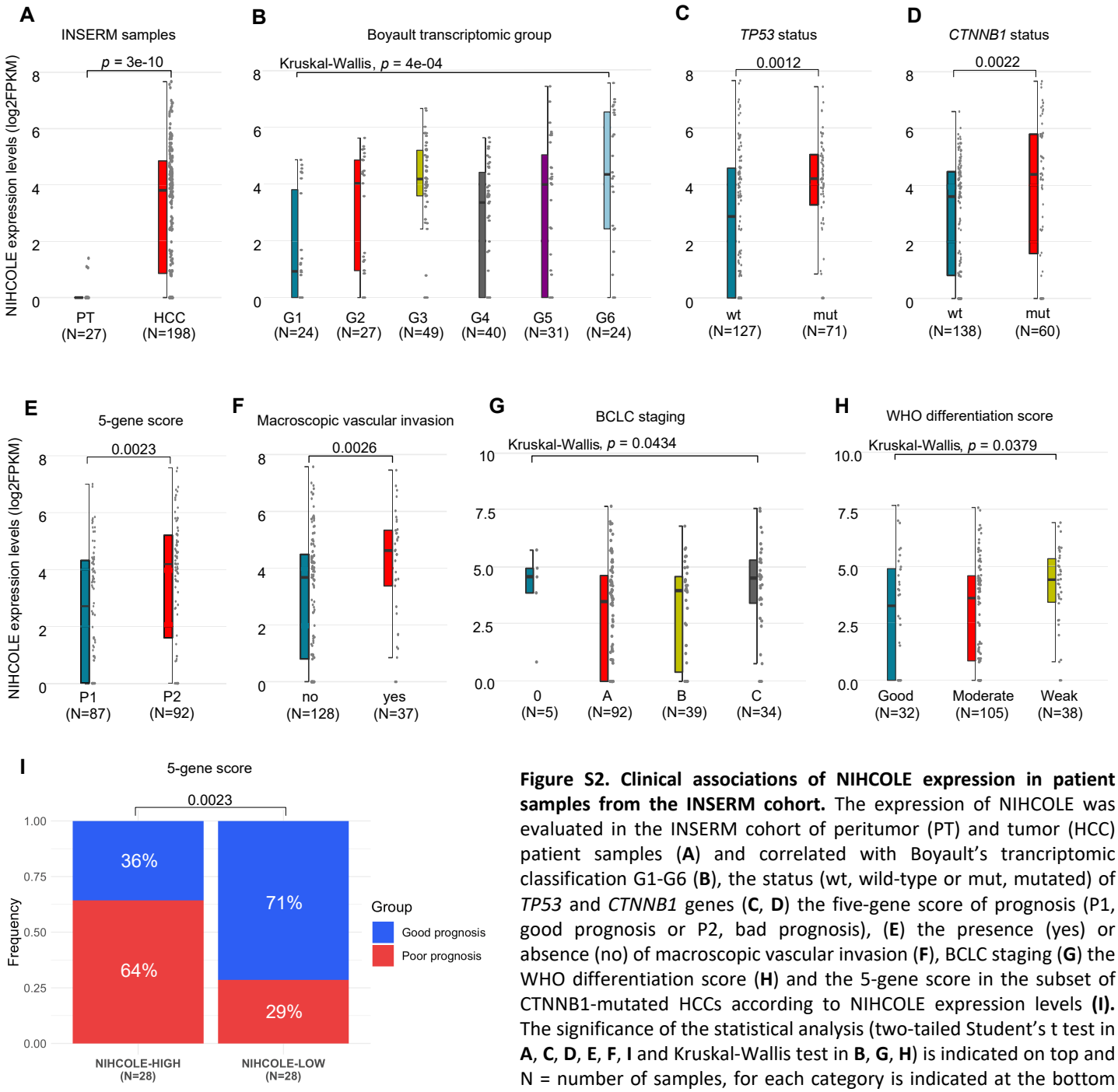


Figure S2. Clinical associations of NIHCOLE expression in patient samples from the INSERM cohort. The expression of NIHCOLE was evaluated in the INSERM cohort of peritumor (PT) and tumor (HCC) patient samples (**A**) and correlated with Boyault's trancriptomic classification G1-G6 (**B**), the status (wt, wild-type or mut, mutated) of *TP53* and *CTNNB1* genes (**C**, **D**) the five-gene score of prognosis (P1, good prognosis or P2, bad prognosis), (**E**) the presence (yes) or absence (no) of macroscopic vascular invasion (**F**), BCLC staging (**G**) the WHO differentiation score (**H**) and the 5-gene score in the subset of CTNNB1-mutated HCCs according to NIHCOLE expression levels **(I).** The significance of the statistical analysis (two-tailed Student's t test in **A**, **C**, **D**, **E**, **F**, **I** and Kruskal-Wallis test in **B**, **G**, **H**) is indicated on top and N = number of samples, for each category is indicated at the bottom of each graph.

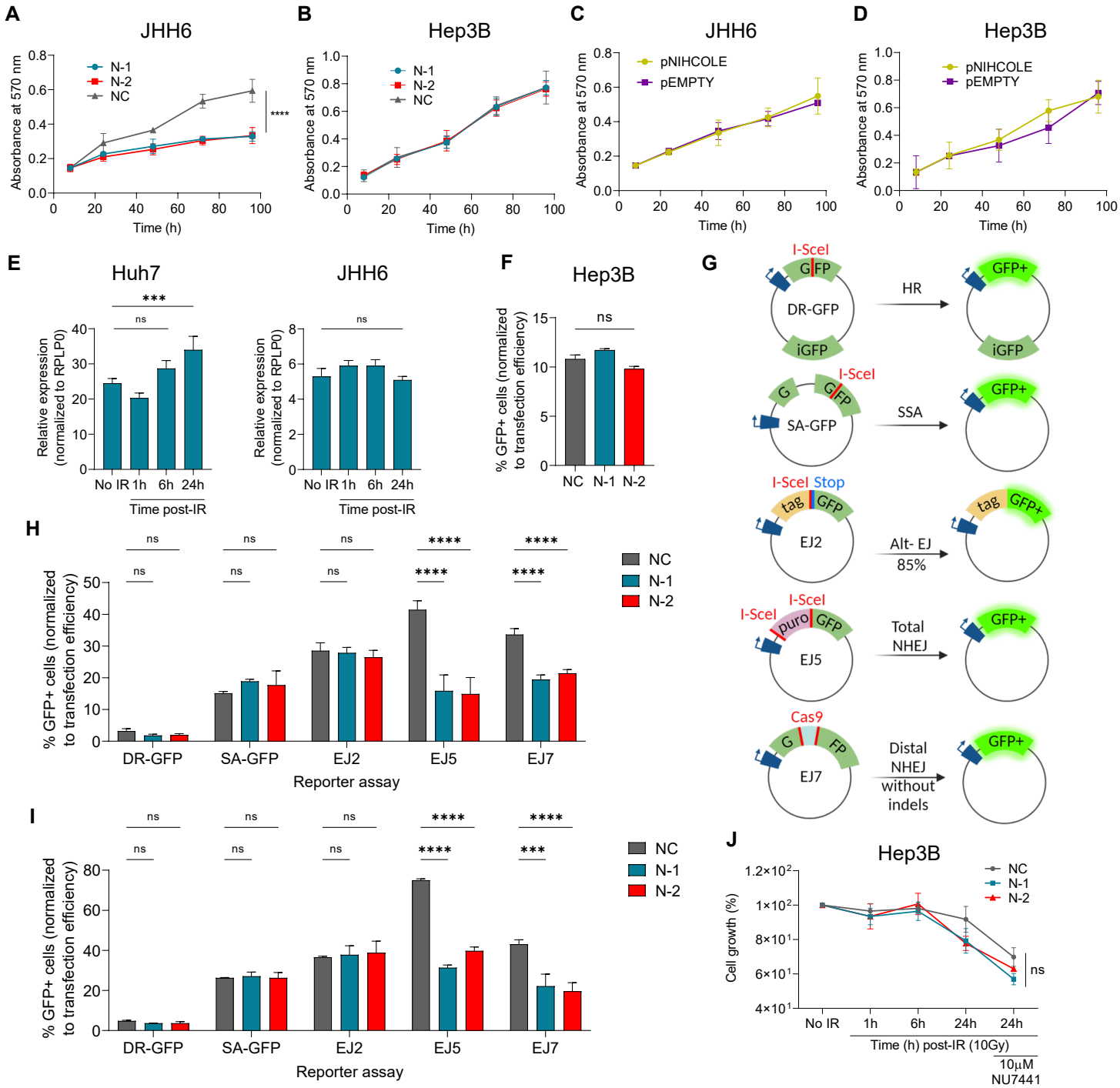
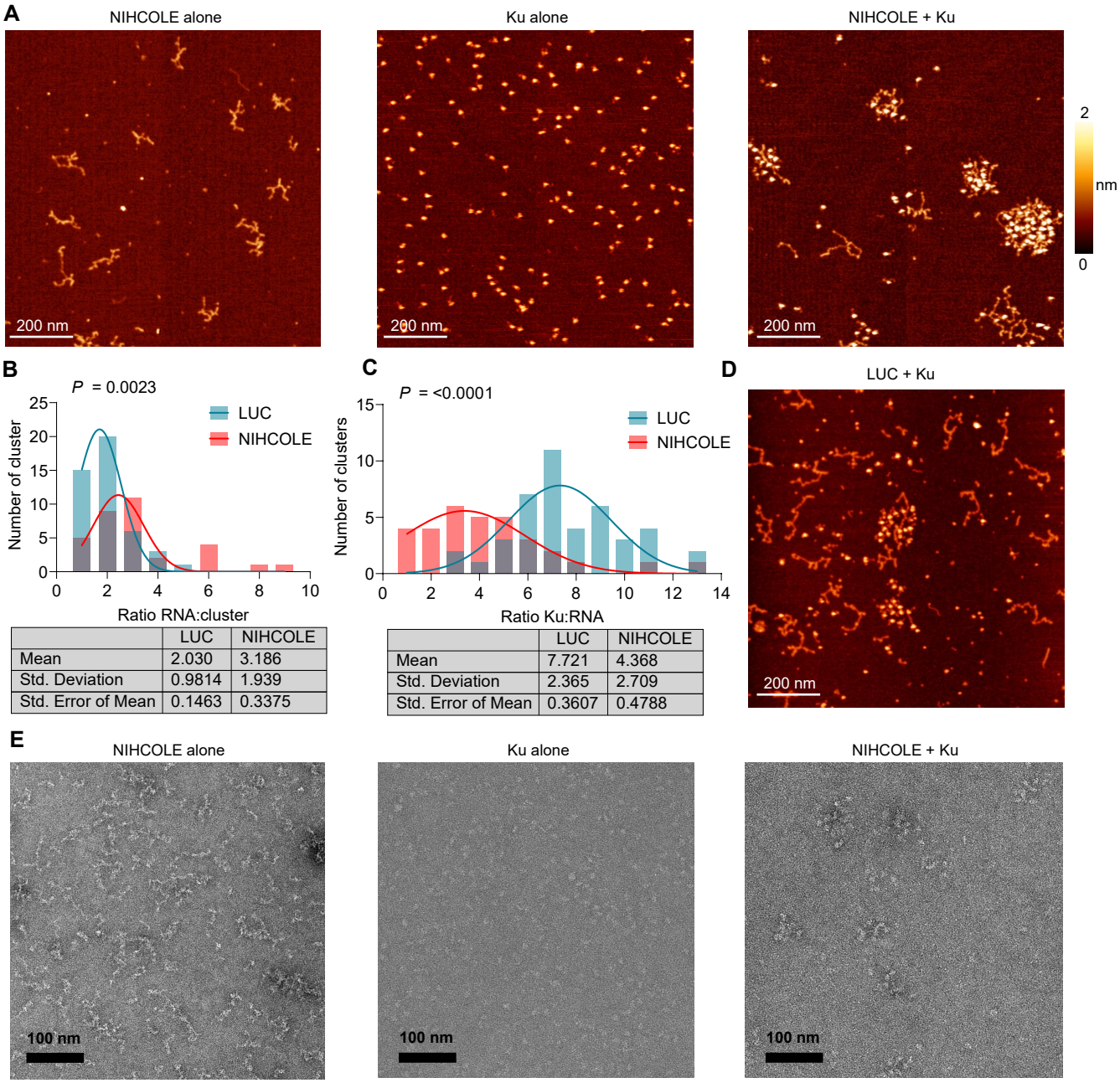


Figure S3. Additional characterization of NIHCOLE expression and perturbations in Huh7 and JHH6 cells. A-D Cells were transfected with either a control gapmer (NC) or NIHCOLE targetting gapmers (N-1 and N-2) or plasmids expressing NIHCOLE (pNIHCOLE) or an empty vector (pEMPTY). The proliferation of control and NIHCOLE-depleted JHH6 (**A**) or HEP3B (**B**) cells was followed over time using MTT assays. Similarly, the proliferation of NIHCOLE-overexpressing or control JHH6 (**C**) or HEP3B (**D**) cells was followed over time using MTT assays. Graphs shows mean ± SEM (*n* = 2). **E** Expression of NIHCOLE after irradiation. JHH6 and Huh7 cells were collected prior to irradiation or at different time points after 10Gy irradiation and analyzed for their expression of NIHCOLE. Graph shows mean ± SD (n = 4). **F-I** Control and NIHCOLE gapmers were co-transfected with GFP DSB repair reporters in Huh7 (**H**) and JHH6 (**I**) cells. HEP3B cell were only transfected with the EJ5 reporter (**F**).The percentage of GFP positive cells was normalized to transfection efficiency. Graphs shows mean ± SEM (*n* = 2 for **F** and **I**; and *n* = 3 for **H**). **G** Diagram of GFP DSB reporter assays to measure homologous recombination (HR), single-strand annealing (SSA), alternative end-joining (Alt-EJ), total NHEJ and distal NHEJ without indels. Red lines depict the sites of DSBs induced by I-SceI endonuclease or Cas9. **J** Cell growth of HEP3B transfected with control or NIHCOLE gapmers, unchallenged or treated with radiation for the indicated times and after radiation plus NU7441 inhibitor measured on a Cellometer cell counter. Graphs shows mean ± SEM (*n* = 2). The significance of the statistical analysis (two-tailed Student's t in **A**, **E**, **F** and **J** and two-way ANOVA in **H** and **I**) is indicated in each graph and summarized as: not significant (ns); * <0.05; **<0.01; ***<0.001; ****<0.0001.

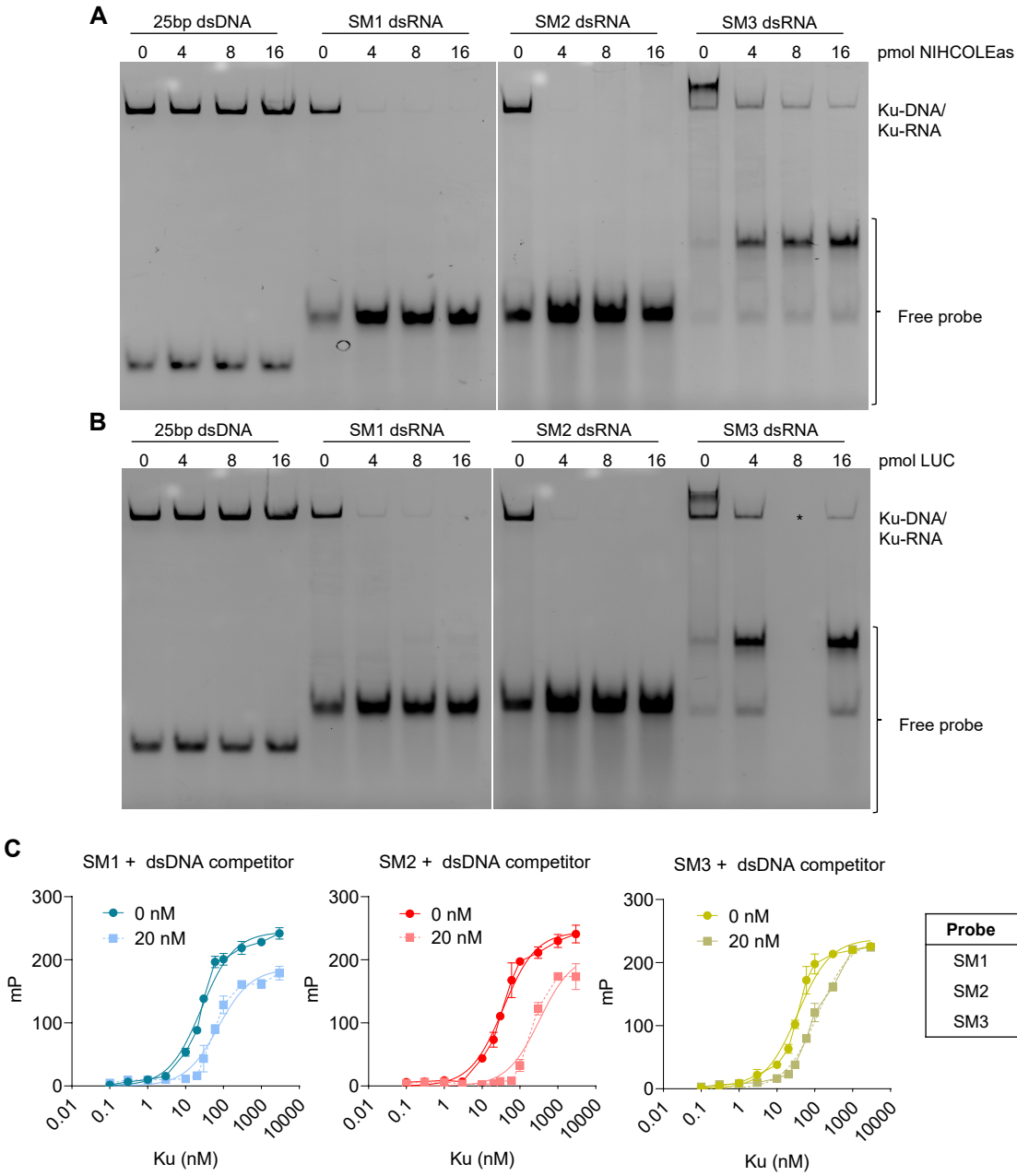


	LUC	NIHCOLE
Mean	2.030	3.186
Std. Deviation	0.9814	1.939
Std. Error of Mean	0.1463	0.3375

	LUC	NIHCOLE
Mean	7.721	4.368
Std. Deviation	2.365	2.709
Std. Error of Mean	0.3607	0.4788

Mean	8.195
Std. Deviation	2.859
Std. Error of Mean	0.1110

Figure S4. NIHCOLE interacts directly with Ku as evaluated by AFM and negative stain EM. A Large field of view (1000x1000 nm) AFM image of NIHCOLE alone, Ku alone, NIHCOLE and Ku. Z scale from dark to bright is 0-2 nm. Scale in shown at the bottom-left corner. **B-C** Estimations of the ratios of RNA molecules per cluster (**B**) and Ku molecules per RNA (**C**) Graph shows a histogram of frequency distribution and non-linear Gaussian regression (n = 45 for LUC, and n = 33 for NIHCOLE). The significance of the statistical analysis (P = two-tailed Mann-Whitney test) and a table of descriptive statistics is shown. **D** Large field of view (1000x1000 nm) AFM image of luciferase messenger (LUC) and Ku. Z scale from dark to bright is 0-2 nm. Scale in shown at the bottom-left corner. **E** Additional images observed in the electron microscope after negative staining of NIHCOLE alone, Ku alone or NIHCOLE and Ku. Ku is observed as globular densities. **F** Estimation of the number of Ku globular densities per NIHCOLE-Ku cluster. Graph shows a histogram of frequency distribution and non-linear Gaussian regression (n = 663). A table of descriptive statistics is shown.



Probe	*Ki* (nM)	R^2
SM1	7.60	0.9792
SM2	2.05	0.9778
SM3	8.19	0.9792

Figure S5. Ku is an RNA binding protein and DNA behaves as a competitor for Ku-RNA binding. A, B Recombinant Ku was assessed for binding to RNA in the presence of *in vitro* transcribed full length antisense NIHCOLE (NIHCOLEas) (**A**) or the messenger of luciferase (LUC) (**B**). Four pmol of each probe were incubated with 4 pmol of Ku with increasing molar ratios of full-length RNAs under binding conditions. Reaction complexes were separated by native electrophoresis. A representative image of three independent experiments is shown. **C** Binding curves for Ku-RNA binding were measured by FPA before (0 nM, darker color) and after addition of the competitor 25bp dsDNA (20nM, lighter shade). Non-linear regression with a mixed inhibition model was used to calculate Ki values for all three SMs and results are shown in the box to the right. Graphs show mean ± SD (*n* = 2).

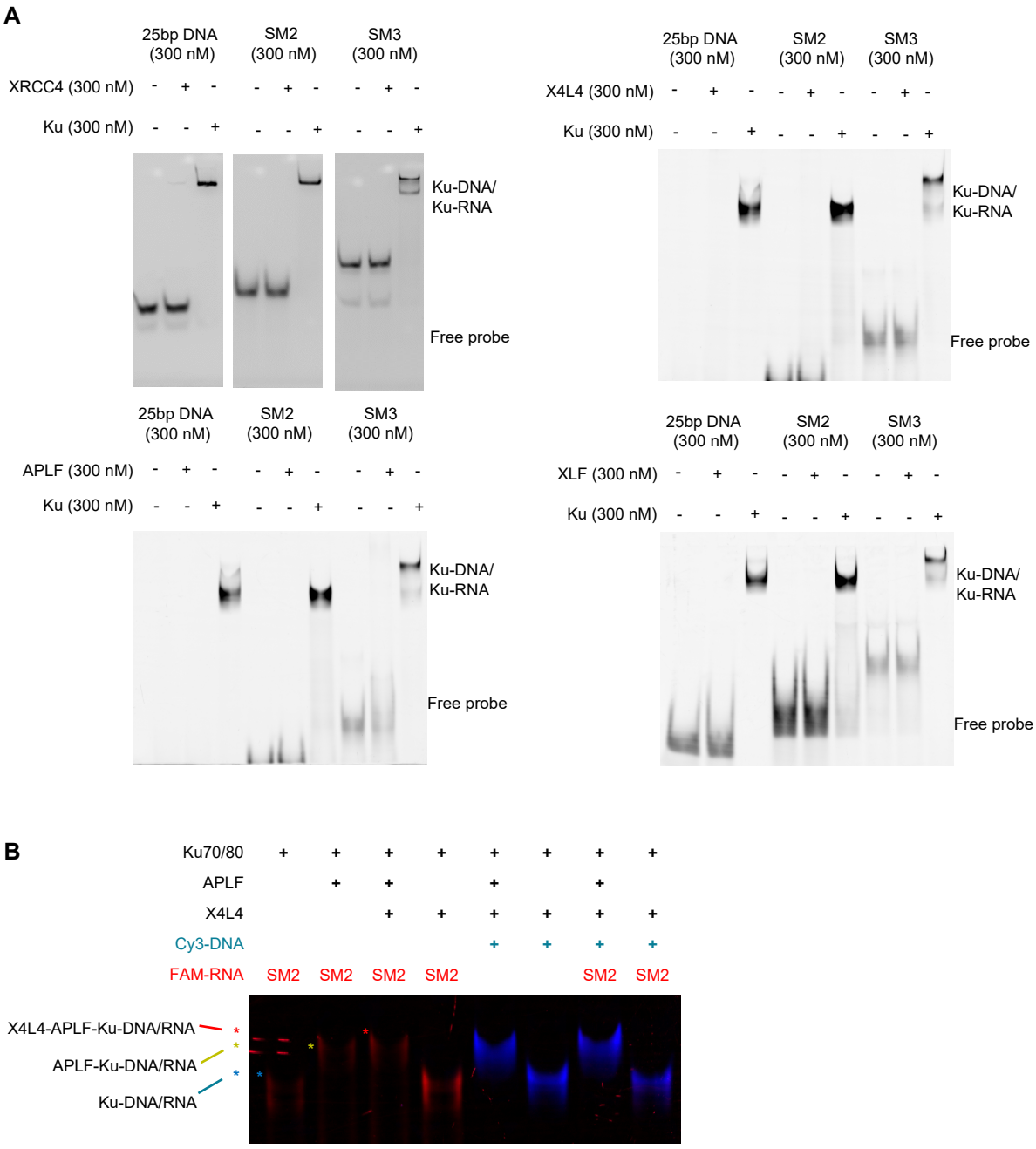


Figure S6. SM2 and SM3 binding to NHEJ factors. A Recombinant Ku, XRCC4, X4L4, APLF and XLF were assessed for binding to dsDNA, SM2 and SM3. 300nM of labeled probe was incubated with 300nM of each protein under binding conditions. Complexes were separated by native electrophoresis. A representative image of three independent experiments is shown. **B** Recombinant Ku, APLF, and X4L4 were assessed for binding to FAM-labeled SM2 and Cy3-labeled DNA. 200nM of each labeled probe was incubated with 200nM Ku, 600nM APLF and 600nM X4L4 under binding conditions. Resulting complexes were separated by native electrophoresis and visualized with a FAM filter. A representative image out of three independent experiments is shown.

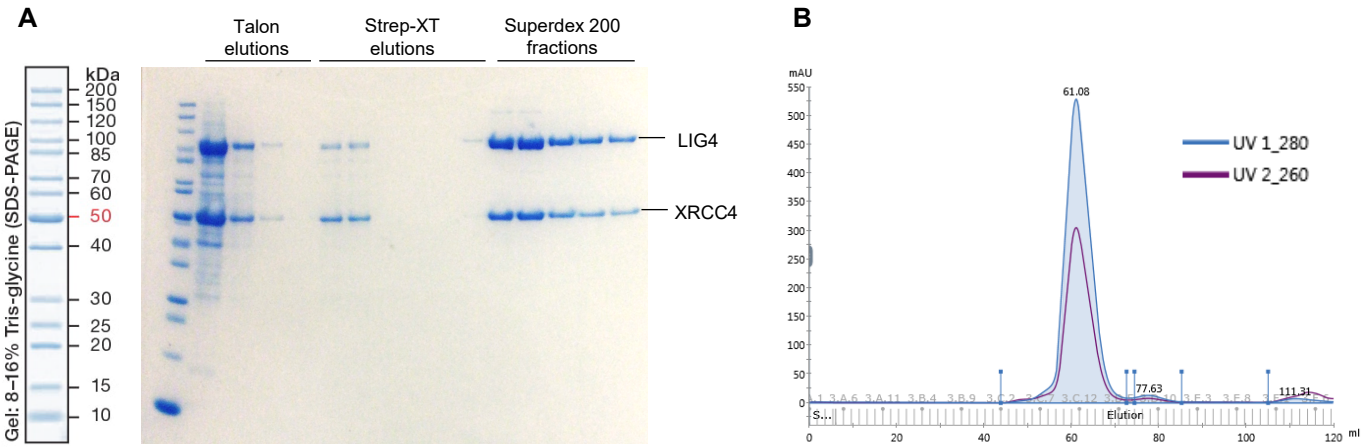


Figure S7. Purification of X4L4 complex. A Denaturing SDS-PAGE of protein fractions obtained during the purification of XRCC4/LIG4 complex. Molecular weight marker is shown to the left. Bands corresponding to LIG4 and XRCC4 are indicated to the right. **B** Chromatogram of Superdex size exclusion XRCC4/LIG4 complex purification.