BC200 63-84 Probe

```
BC200 TAGCTTGAGCCCAGGAGTTCGA
 AluJ TTGCTTGAGCCCAGGAGTTCGA
 AluY TCACCTGAGGTCAGGAGTTCGA
 AluS TCACTTGAGGTCAGGAGTTCGA
  7SL TCGCTTGAGTCCAGGAGTTCTG
```

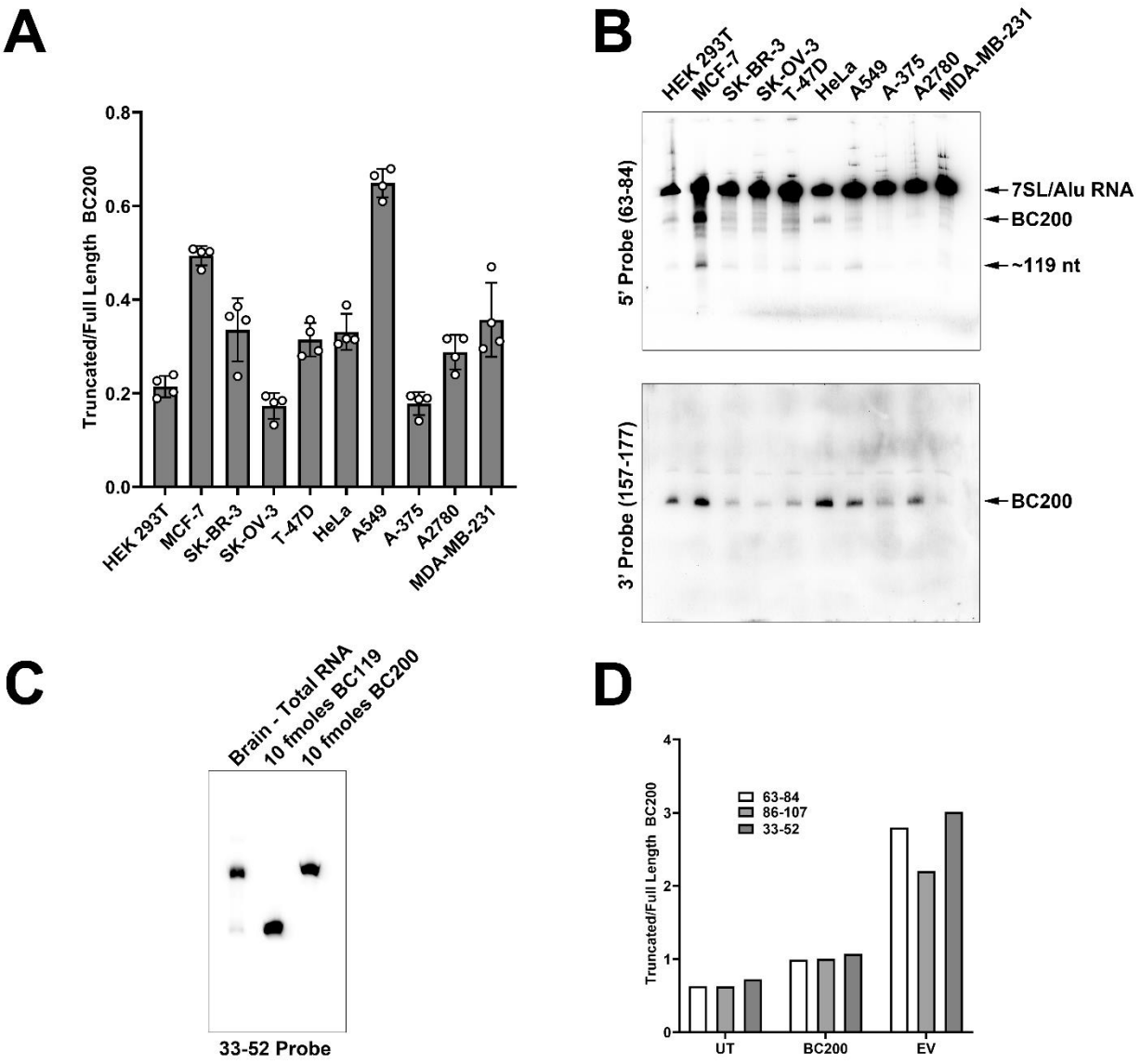
BC200 86-107 Probe

```
BC200 ACCTGCCTGGGCAATATAGCGA
 AluJ ACCAGCCTGGGCAACATAGCGA
 AluY ACCAGCCTGGCCAACATGGTGA
 AluS ACCAGCCTGGCCAACATGGTGA
  7SL TCCAGCCTGGGCAACATAGCGA
```

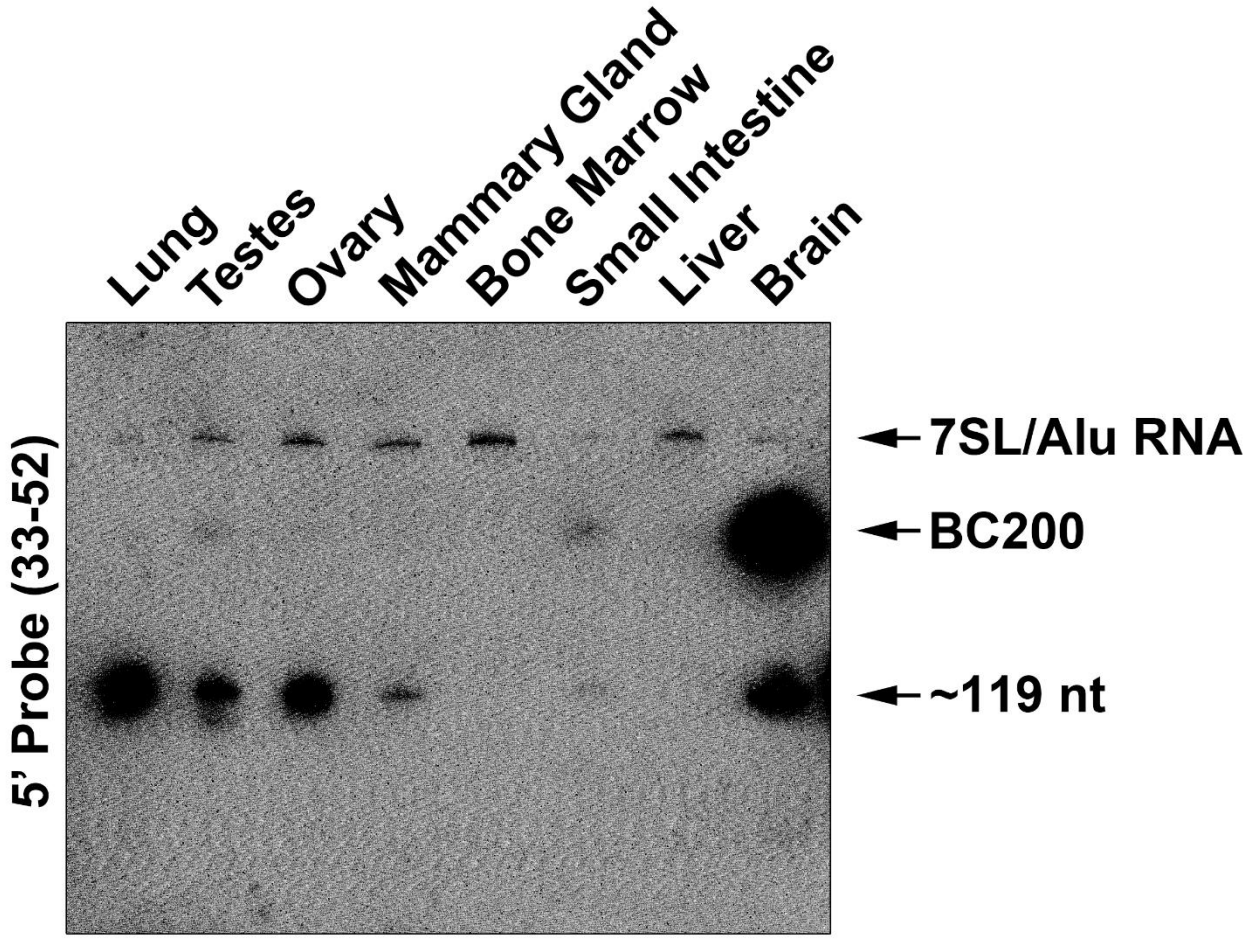
BC200 33-52 Probe

```
BC200 AGCTCTCAGGGAGGCTAAGA
 AluJ AGCACTTTGGGAGGCCGAGG
 AluY AGCACTTTGGGAGGCCGAGG
 AluS AGCACTTTGGGAGGCCGAGG
  7SL AGCTACTCGGGAGGCTGAGG
```

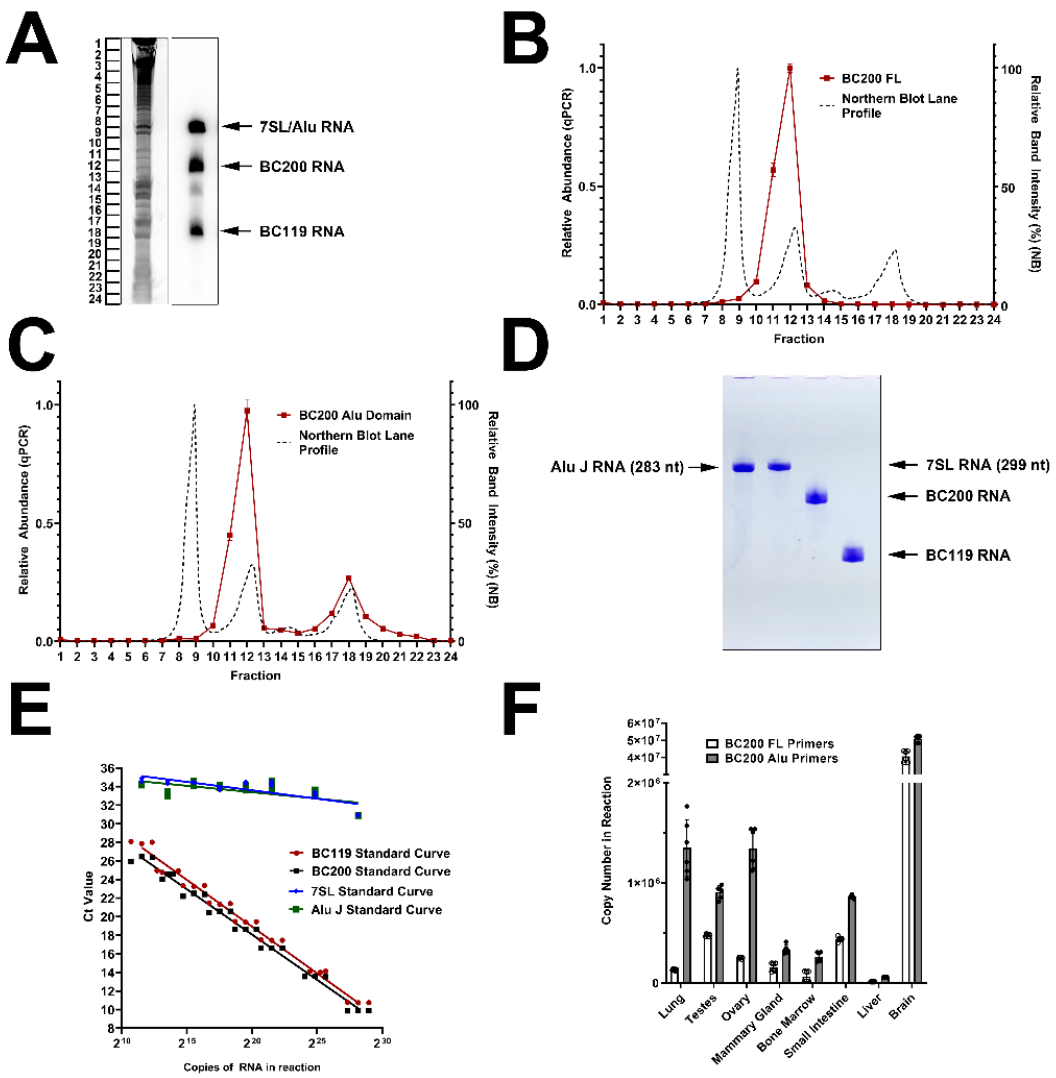
Supplementary Figure 1. Alignment of BC200 targeting probes with Alu consensus and 7SL RNA sequences. Mismatches to the BC200 sequence are highlighted in red.



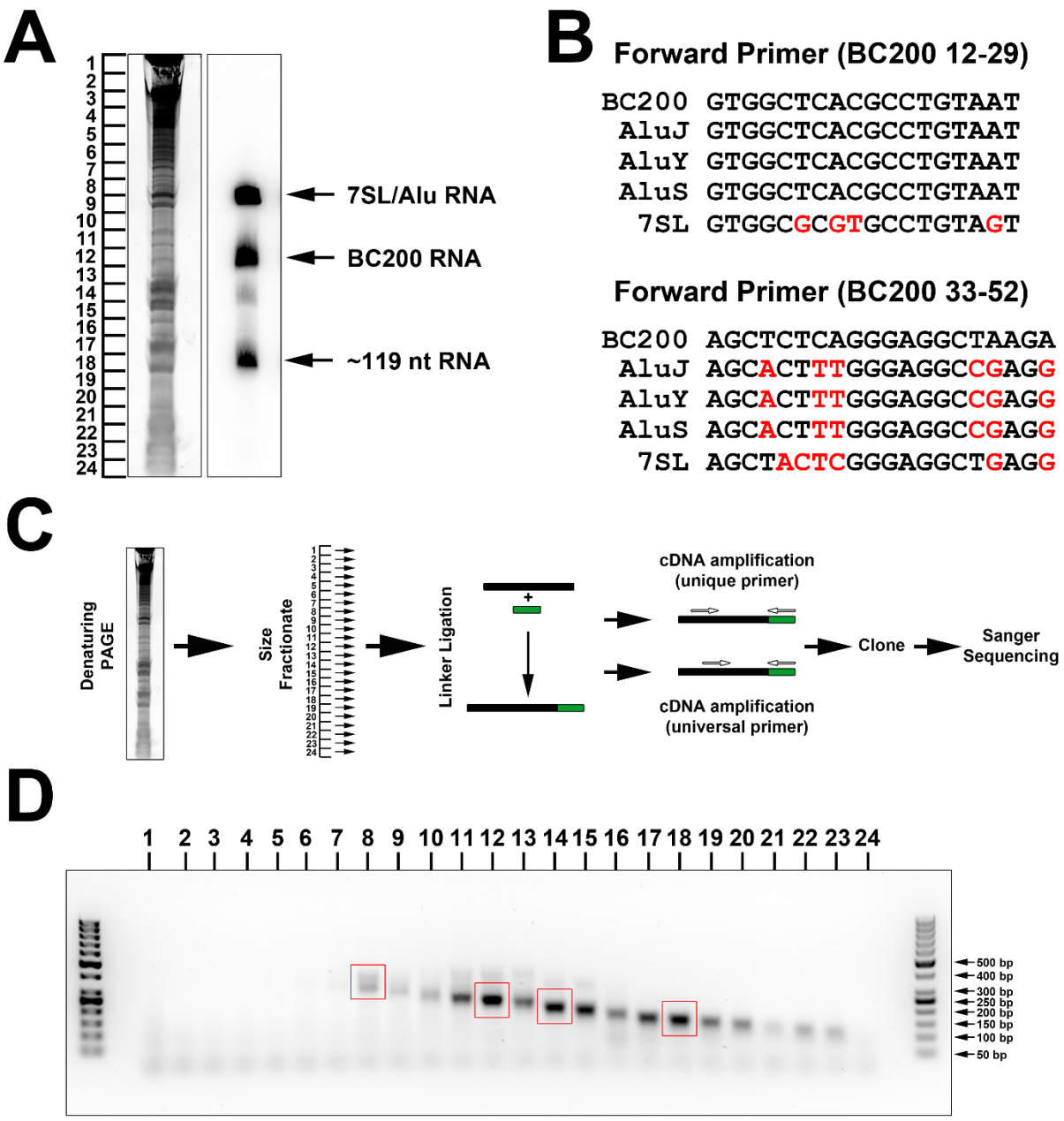
Supplementary Figure 2. Expression of BC200 and BC120 in a panel of human cell lines. (A) Intensity of the ~119 band in Figure 2A was assessed relative to the corresponding BC200 band by densitometry. Integrated band intensity (sum of background subtracted pixel intensity) for BC119 was divided by the same value obtained from the BC200 band. Data represents the mean of four replicates +/- standard deviation. (B) Endogenous BC200 was assessed by northern blot using 5 µg total RNA harvested from the indicated cell lines 48 hours after plating with the indicated probes. (C) 2 ug of RNA from normal human brain was separated by denaturing urea-TBE polyacrylamide gel electrophoresis alongside 10 fmoles of *in-vitro* transcribed BC119 and BC200 and detected by northern blot with a probe targeting BC200 nucleotides 33-52. (D) Quantification of the bands in Figure 2B by densitometry reveals similar observed ratios between the ~119 nt fragment and BC200 regardless of the detection probe used.



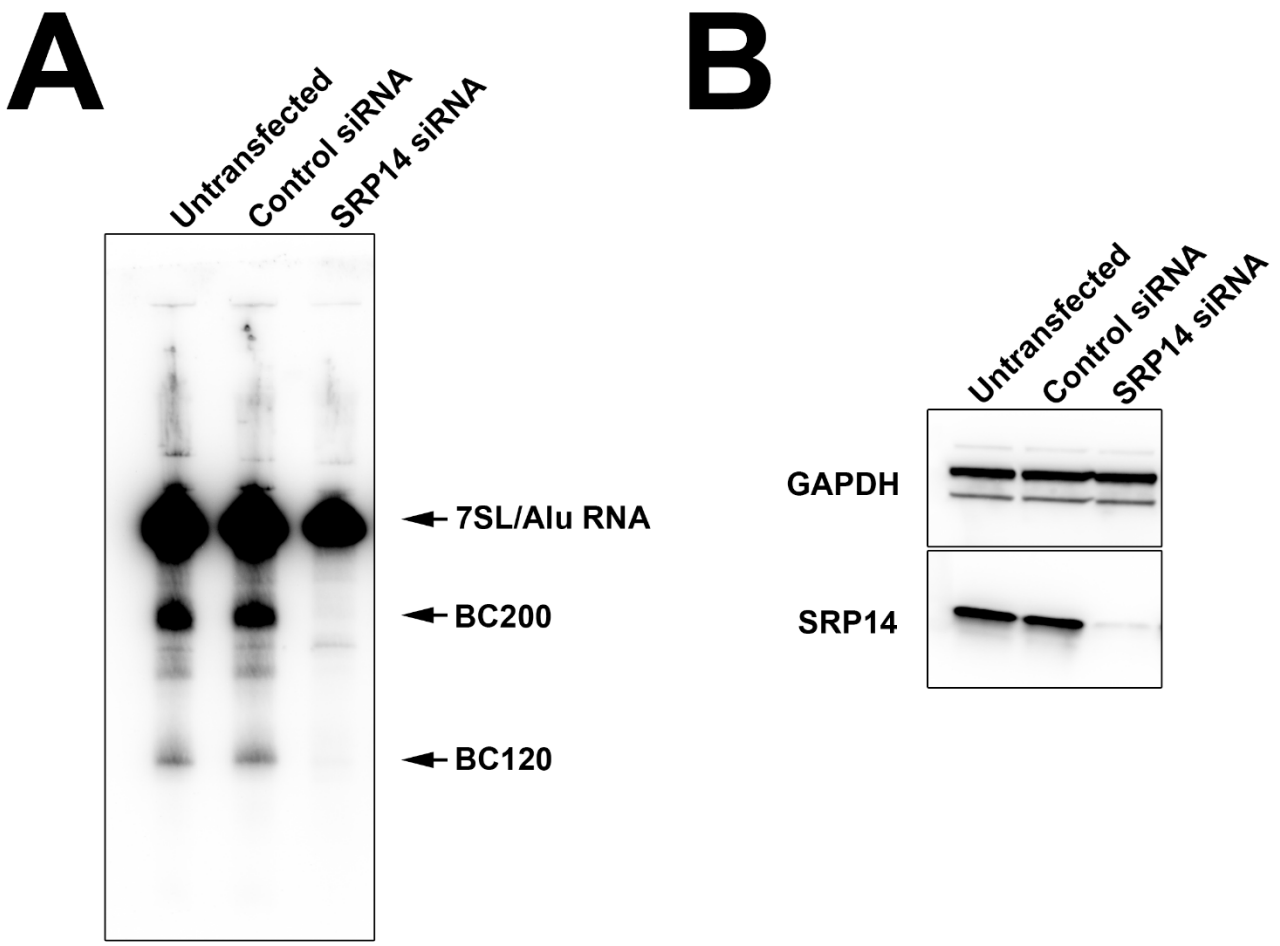
Supplementary Figure 3. The BC200 truncation is expressed in normal human tissues (A) High exposure of northern blot from Figure 2D. 2 µg of total RNA from the indicated normal human tissues was northern blotted with a probe targeting BC200 nucleotides 33-52. In the case of Brain, 500 ng total RNA was loaded to balance the signal intensity.



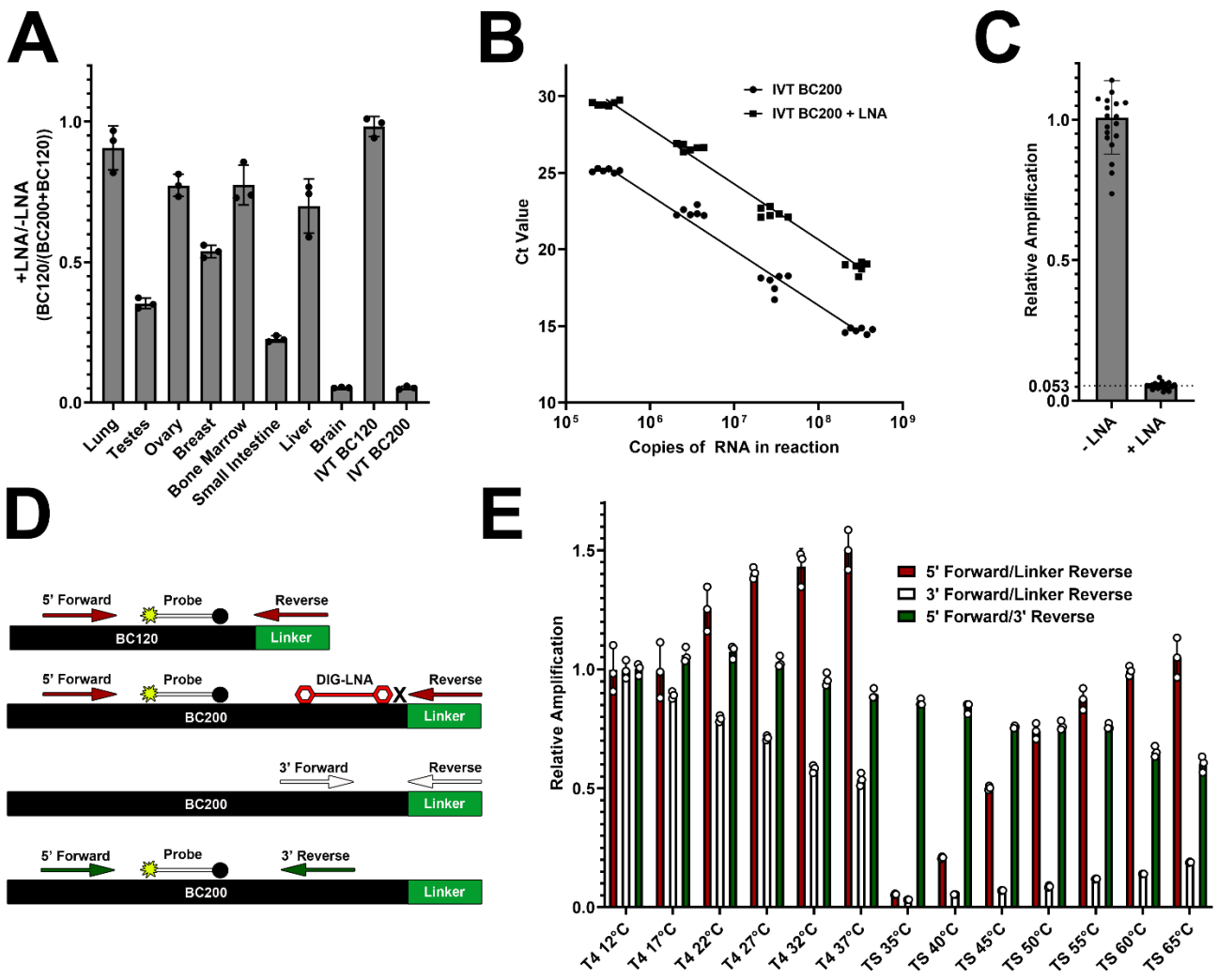
Supplementary Figure 4. A primer/probe set in the BC200 Alu domain can differentiate BC120 from Alu and 7SL RNA. (A) Total RNA from MCF-7 cells separated by denaturing TBE-Urea PAGE. Identical lanes were separated and used for total RNA staining with SYBR gold (panel 1) or northern blot with a probe targeting BC200 nucleotides 63-84 (panel 2). Bands were excised from a third lane for extraction and purification of the size fractionated RNA according to the template indicated on the left of the total RNA stain panel. (B) Size fractionated and purified RNA was used as a template for one-step RT-qPCR using a primer/probe set that spanned the length of BC200 (BC200 FL). qPCR data (left y-axis) is normalized to the fraction with the highest measured concentration of BC200 (lowest Ct value) and represents the mean of three replicates +/- standard deviation. Dashed line represents the densitometry lane profile of the northern blot in (A) plotted on the right y-axis aligned with the size separated fractions as indicated in the x-axis. (C) as in (B) using a primer/probe set complementary to the BC200 Alu domain. (D) Indicated RNAs were *in-vitro* transcribed, purified, and analyzed by denaturing TBE-urea PAGE. Gels were stained with the total RNA stain Toluidine Blue-O. (E) Standard curves generated by RT-qPCR using a primer/probe set complementary to the BC200 Alu domain with the indicated RNAs as template. Individual data points are shown for three replicates. Linear regression was performed with GraphPad Prism software to obtain a line of best fit. (F) (B) 25 ng of the RNA from a panel of human tissues was used as a template for RT-qPCR using a primer/probe set that spans the full-length BC200 (BC200 FL) or the BC200 Alu domain (BC200 Alu). Data represents the mean of 6 replicates +/- standard deviation.



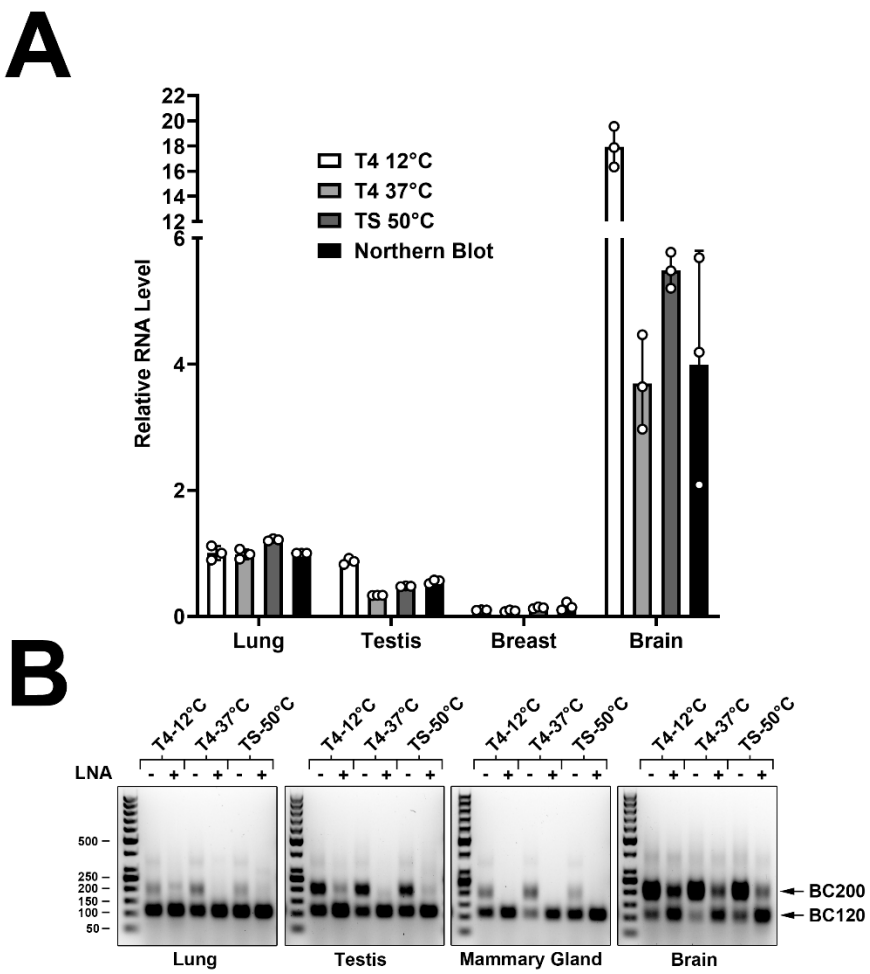
Supplementary Figure 5. Sequencing of size-fractionated RNA. (A) Total RNA from MCF-7 cells was separated by denaturing TBE-Urea PAGE. Identical lanes were separated and used for total RNA staining with SYBR gold (panel 1) or northern blot with a probe targeting BC200 nucleotides 63-84 (panel 2). Bands were excised from a third lane for extraction and purification of the size fractionated RNA according to the template indicated on the left of the total RNA stain panel. Size fractionated RNA was 3' ligated to an oligonucleotide linker. (B) Alignment of the forward primers used to amplify cDNA with the indicated Alu consensus sequences and the 7SL RNA. Mismatches with the BC200 sequence are highlighted in red. (C) Flow chart of the method used to sequence amplified RNA. (D) Agarose gel electrophoresis of the amplification products using the 12-29 forward primer shown in (B) and a reverse primer complementary to the ligated oligonucleotide linker. Bands indicated by red box were excised, purified and blunt-end cloned for sequencing analysis.



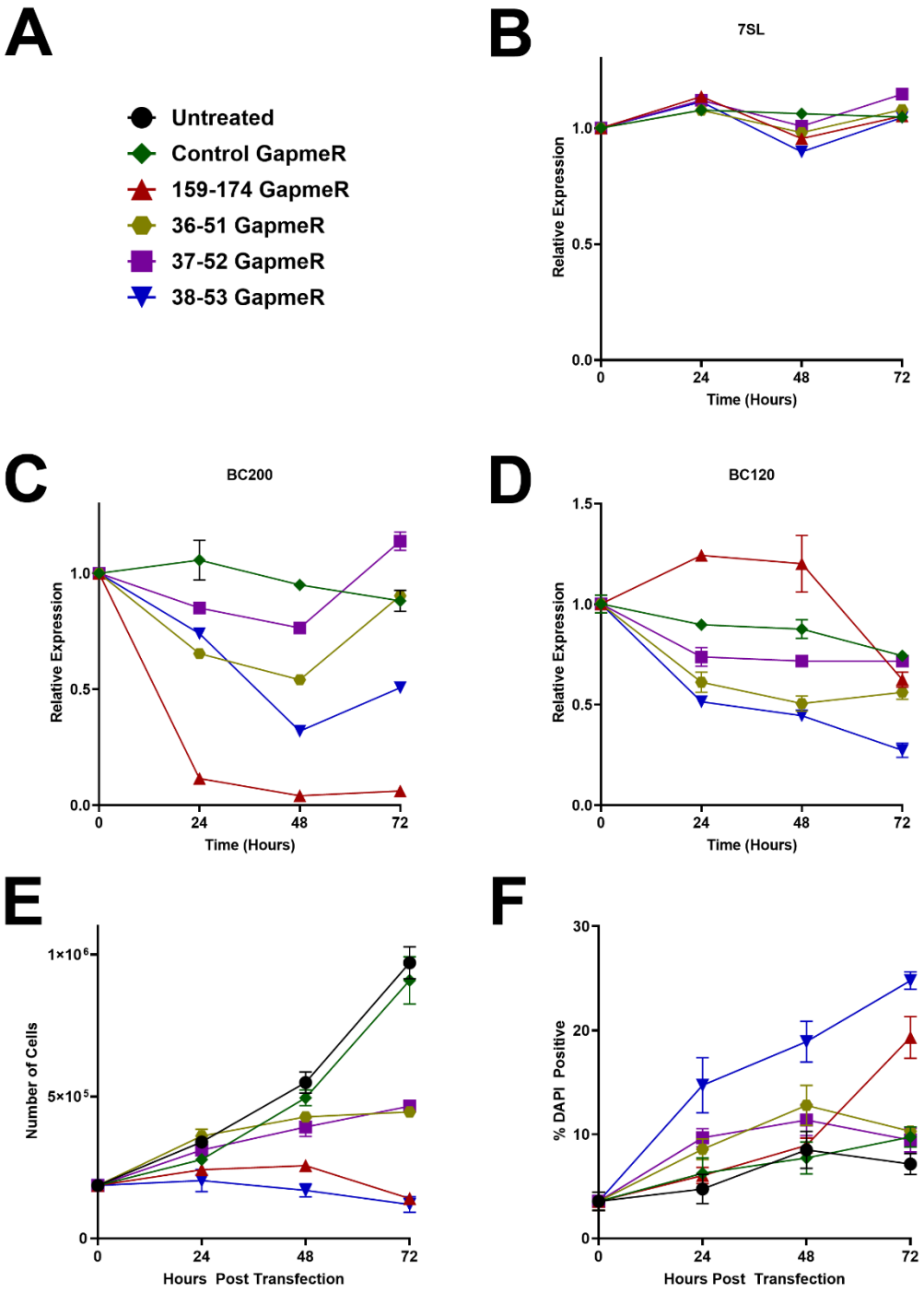
Supplementary Figure 6. SRP14 Knock-down abrogates BC200 and BC120 expression. (A) Total RNA from MCF-7 cells transfected with the indicated siRNAs was separated by denaturing TBE-Urea PAGE. Northern blot was performed with a probe targeting BC200 nucleotides 63-84. (B) In parallel, half of the cells used in (A) were lysed in RIPA buffer and total protein was separated by SDS-PAGE. Western blot was performed to monitor SRP14 knock-down with GADPH as a loading control. Similar results were observed upon SRP9 knock-down.



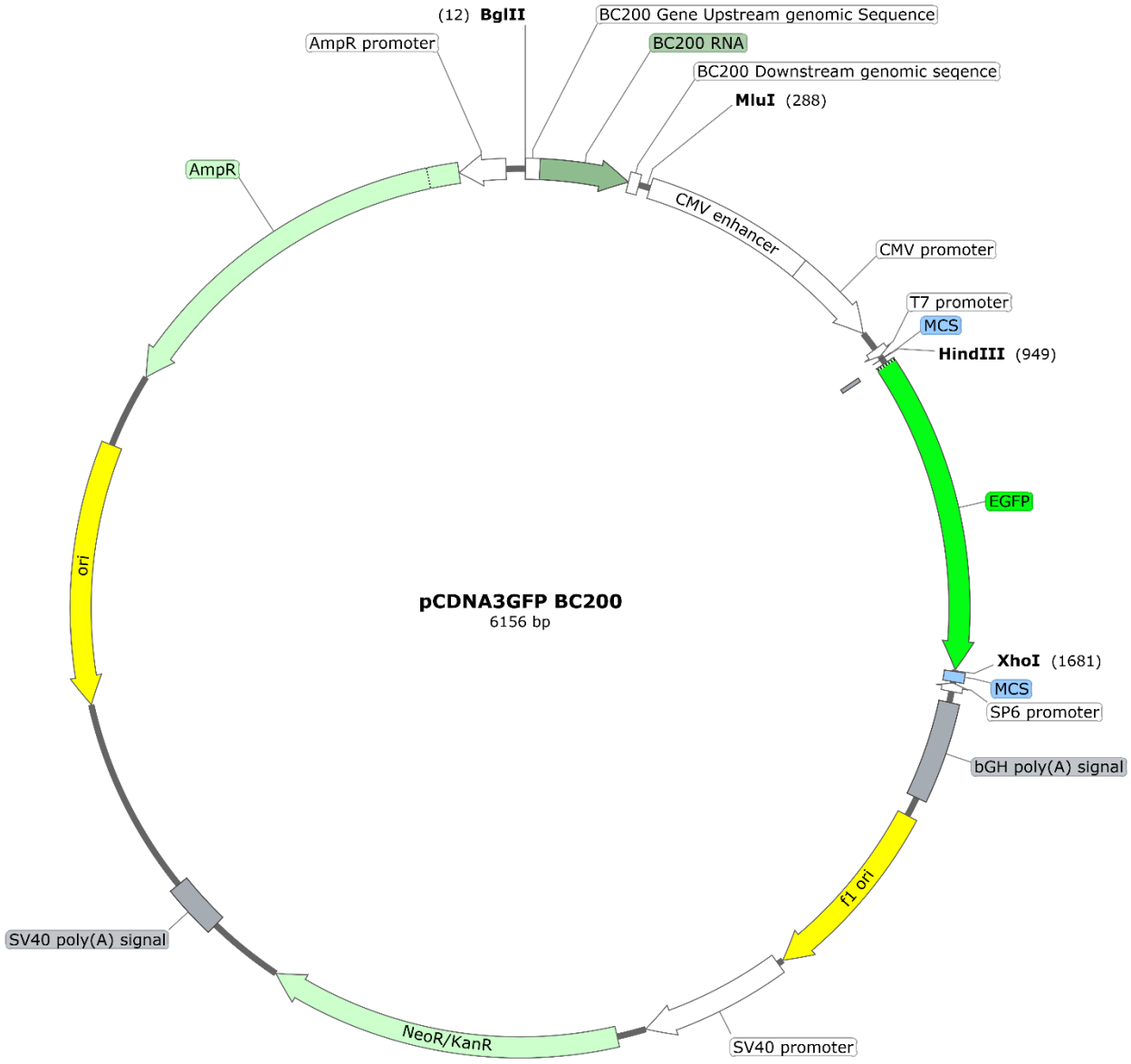
Supplementary Figure 7. Improvement of BC120 specific amplification by optimizing linker ligation bias. (A) 25 ng of ligated RNA from the indicated tissues was used as a template for RT-qPCR in the presence and absence of the blocking DIG-labelled LNA using the primer/probe sets shown in Figure 7A. For the indicated *in-vitro* transcribed, linker-ligated RNAs, 5 µl of RNA at a concentration of 10 pM was used as a template. Data is relative to the untreated (no LNA) reactions and represents the mean of 3 replicates +/- standard deviation. (B) A dilution series of *in-vitro* transcribed, linker ligated BC200 RNA was used as a template for RT-qPCR in the presence and absence of the blocking LNA probe. Data was fit to a straight line using non-linear regression (GraphPad Prism). (C) Bar graph comparing the relative amplification at the four concentrations assessed in (B) +/- blocking LNA. Dashed line represents the mean reduction in the presence of DIG-labelled LNA. While blocking was ~95% effective it was not complete. (D) Schematic demonstrating the location of primers and probes in the context of linker ligated BC120 and BC200. Primer set indicated in red will amplify BC120 but will be largely blocked from amplifying BC200. Primer set indicated in white will only amplify linker ligated BC200 and primer set indicated in green will amplify total BC200 RNA. The white primer set did not allow for the use of a Taqman probe and as such detection was based on SYBR green based RT-qPCR. (E) RT-qPCR analysis of linker-ligated total MCF-7 RNA using the primer sets outlined in (D). Data represents the mean of 3 replicates +/- standard deviation and are expressed relative to the 12°C ligation samples with T4 RNA ligase (T4 12°C). TS indicates thermostable 5' App DNA/RNA ligase. Elevated ligation temperatures preferentially improve ligation efficiency to the BC120 truncation.



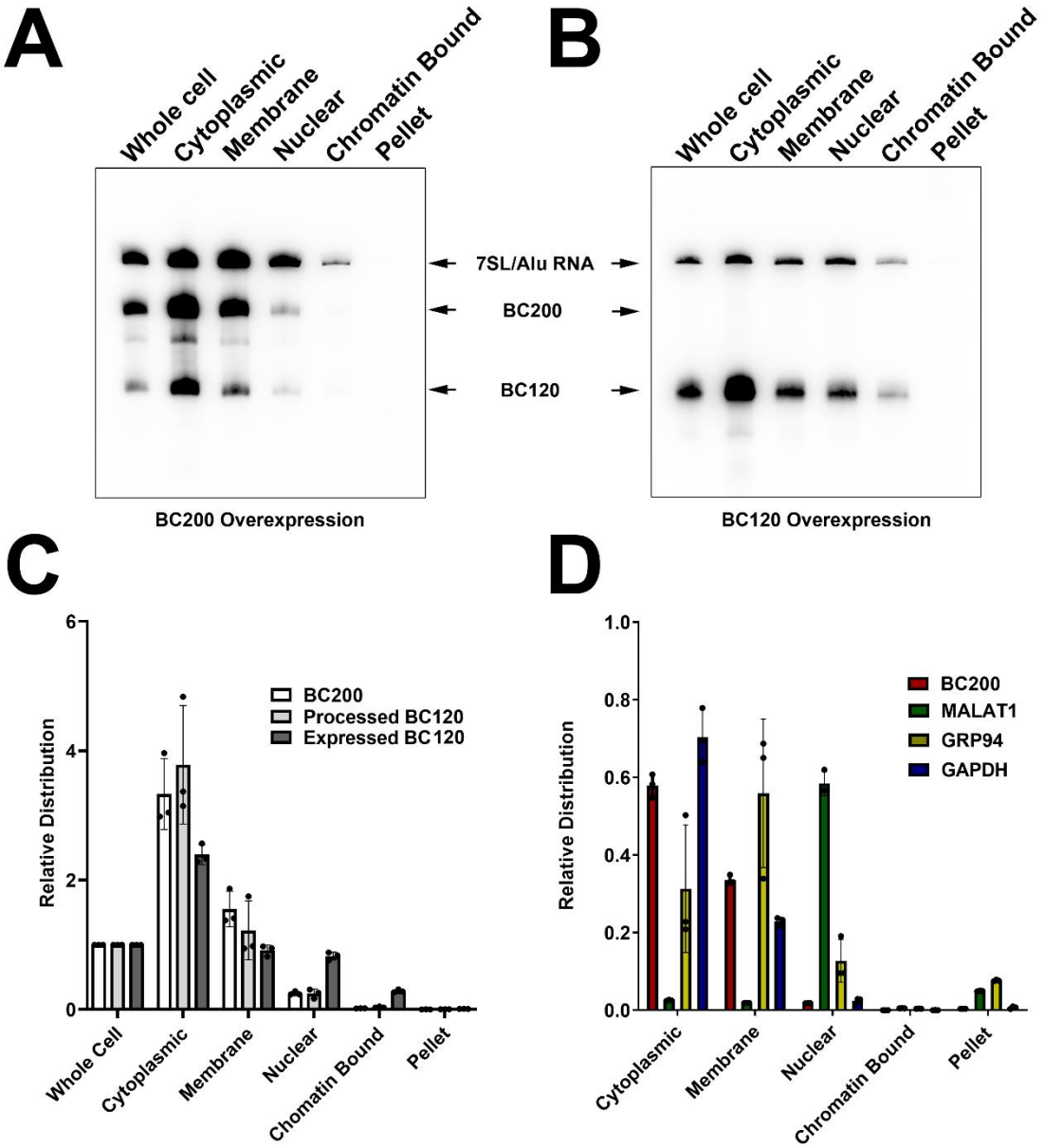
Supplementary Figure 8. Analysis of BC120 expression in normal human tissues using optimized ligation conditions. (A) 25 ng of ligated RNA from the indicated tissues was used as a template for RT-qPCR in the presence the blocking DIG-labelled LNA using the primer/probe sets shown in Figure 7A. Ligation was performed with T4 RNA ligase at 12° and 37°C (T4 12°C, T4 37°C) as well as with thermostable 5' App DNA/RNA ligase at 50°C (TS 50°C). Data is relative to the T4 12°C ligated normal lung total RNA. Black bars represent relative quantification of northern blots (representative blot in Figure 6A) corrected for underloading of brain total RNA and represents the mean of 3 replicates +/- standard deviation. Elevated ligation temperatures give quantification results in line with the northern blot analysis by ameliorating the low temperature ligation bias towards B200. (B) Amplification products from (A) were separated by agarose gel electrophoresis and stained with SYBR Safe nucleic acid dye.



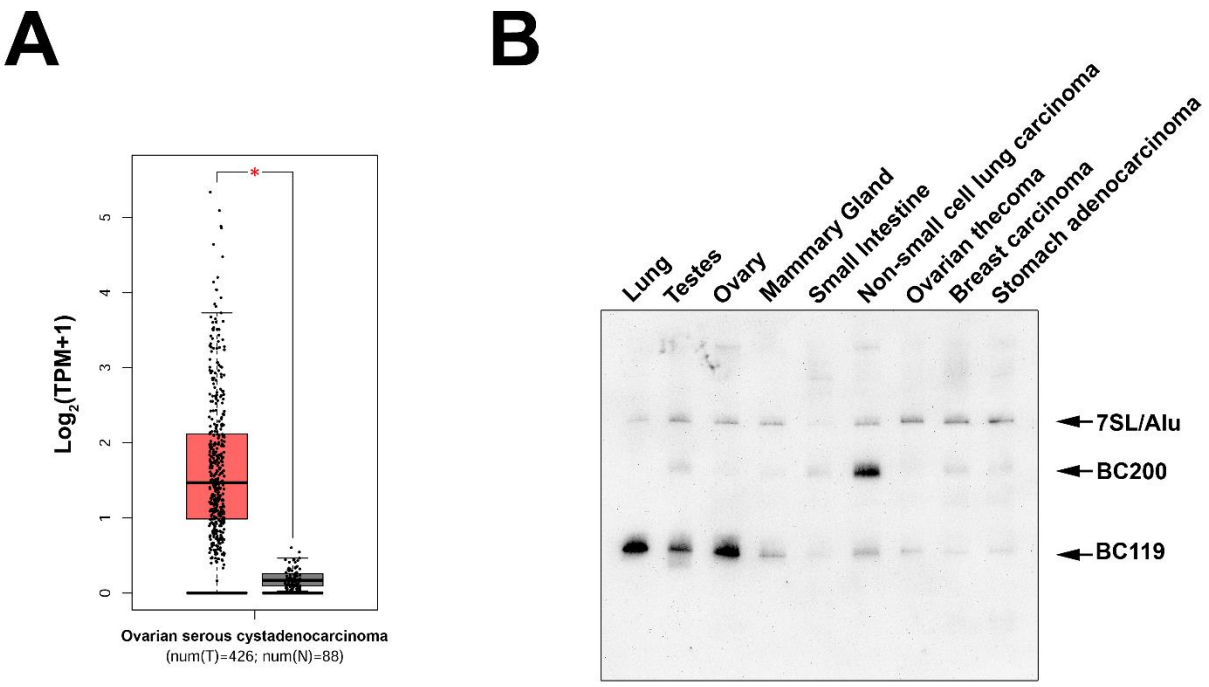
Supplementary Figure 9. Knockdown of BC120 and BC200 with Alu domain targeting LNA GapmeRs. (A) Legend for the line graphs presented in B-F. (B,C,D) MCF-7 cells were transfected with the indicated LNA GapmeRs. RNA was used as a template for RT-qPCR with primer/probe sets to quantify expression of 7SL (B), BC200 (C) and BC120 (D). Data is relative to the untreated cells and represents the mean of 3 replicates +/- standard deviation. (E) MCF-7 cells were transfected and counted at the indicated time points. Data represents the calculated total number of cells in each well and is the mean of 4 wells per condition +/- standard deviation. (F) As in (E), to the same samples a cell impermeable viability dye (DAPI) was added to each sample and the proportion of DAPI positive cells was measured at each timepoint for each condition. Data represents the mean of 4 wells per condition +/- standard deviation.



Supplementary Figure 10. Design of BC200 GFP reporter constructs. Schematic of GFP reporter plasmids with BC200 expression cassette inserted into the BglII/MluI restriction endonuclease cutting sites upstream of the CMV enhancer.



Supplementary Figure 11. Subcellular localization of processed and overexpressed BC120. (A) MCF-7 cells were transfected with a plasmid to overexpress full-length BC200. 24 hours later subcellular fractionation was performed, and RNA was purified from each of the indicated fractions. 2 µg of RNA from the indicated fractions was separated by denaturing TBE Urea-PAGE and used for northern blotting with a probe complementary to BC200 nucleotides 63-84. (B) As in (A); however, cells were transfected with a plasmid to overexpress the BC120 truncation alone. (C) Densitometry quantification of the data in (A) and (B). Data represents the mean quantification from 3 separate northern blots +/- standard deviation relative to the "whole cell" sample. (D) 25 ng of RNA from each of the fractions used in (A) and (B) was used as a template for RT-qPCR to demonstrate accurate separation of sub-cellular fractions. MALAT1 is a marker for nuclear RNA, GRP94 is associated with endoplasmic reticulum and GAPDH is primarily cytosolic. Data represents the mean of 3 replicates +/- standard deviation relative to the "whole cell" sample.



Supplementary Figure 12. Assessment of BC120 expression in additional cancer types. (A) Analysis of BC200 expression in ovarian cancer and normal ovary using the GEPIA webserver. Data represents sequencing analysis of BC200 in 426 patient samples (red box) from The Cancer Genome Atlas (TCGA) and 88 normal samples (gray box) from TCGA as well as the GTEx Project (Genotype-Tissue Expression). * indicates p value less than 0.01 by one-way ANOVA. (B) 2 µg of RNA from the indicated normal and tumour tissue types was separated by denaturing TBE Urea-PAGE and used for northern blotting with a probe complementary to BC200 nucleotides 33-52.