Cell Genomics, Volume *4*

Supplemental information

Diversity of ribosomes at the level of rRNA variation associated with human health and disease

Daphna Rothschild, Teodorus Theo Susanto, Xin Sui, Jeffrey P. Spence, Ramya Rangan, Naomi R. Genuth, Nasa Sinnott-Armstrong, Xiao Wang, Jonathan K. Pritchard, and Maria Barna

Supplementary Figures

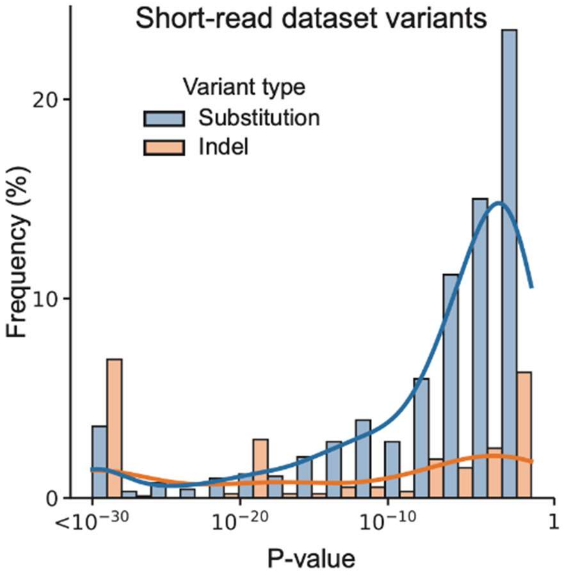


Figure S1. Indels are associated with lower P-values compared to SNVs, related to Figure 1.

Distribution of Mutect2 false discovery rate (FDR)-corrected log 10 likelihood ratio scores of variant existence measured for substitutions and indels.

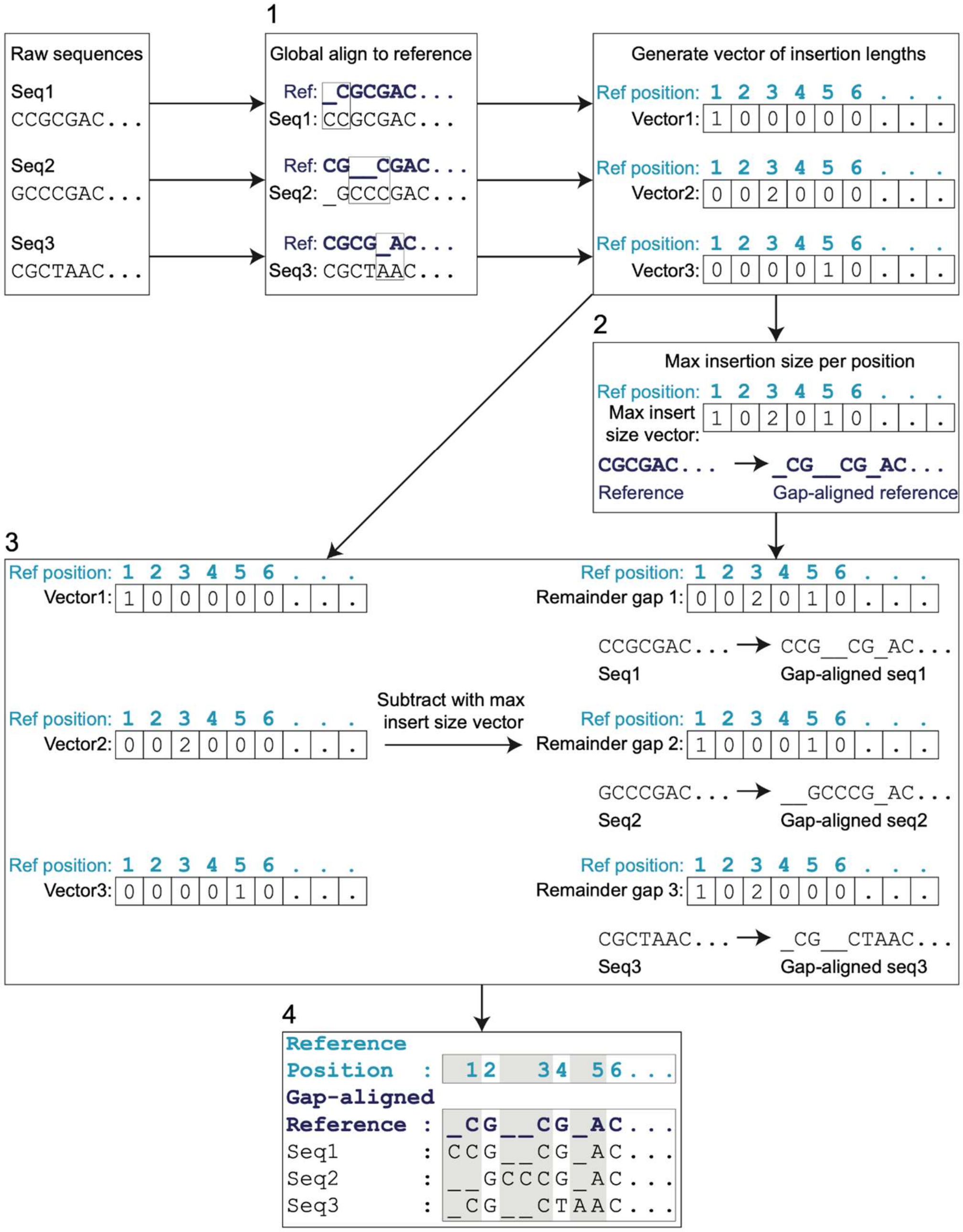

1
Raw sequences
Seq1
CCGCGAC...
Seq2
GCCCGAC...
Seq3
CGCTAAC...
Global align to reference
Ref: _CGCGAC...
Seq1: CCGCGAC...
Ref: CG__CGAC...
Seq2: _GCCCGAC...
Ref: CGCG_AC...
Seq3: CGCTAAC...
Generate vector of insertion lengths
Ref position: 1 2 3 4 5 6 . . .
Vector1: 1 0 0 0 0 0 . . .
Ref position: 1 2 3 4 5 6 . . .
Vector2: 0 0 2 0 0 0 . . .
Ref position: 1 2 3 4 5 6 . . .
Vector3: 0 0 0 0 1 0 . . .
2
Max insertion size per position
Ref position: 1 2 3 4 5 6 . . .
Max insert size vector: 1 0 2 0 1 0 . . .
CGCGAC... → _CG__CG_AC...
Reference
Gap-aligned reference
3
Ref position: 1 2 3 4 5 6 . . .
Vector1: 1 0 0 0 0 0 . . .
Ref position: 1 2 3 4 5 6 . . .
Remainder gap 1: 0 0 2 0 1 0 . . .
CCGCGAC... → CCG__CG_AC...
Seq1
Gap-aligned seq1
Ref position: 1 2 3 4 5 6 . . .
Vector2: 0 0 2 0 0 0 . . .
Subtract with max insert size vector
Ref position: 1 2 3 4 5 6 . . .
Remainder gap 2: 1 0 0 0 1 0 . . .
GCCCGAC... → __GCCCG_AC...
Seq2
Gap-aligned seq2
Ref position: 1 2 3 4 5 6 . . .
Vector3: 0 0 0 0 1 0 . . .
Ref position: 1 2 3 4 5 6 . . .
Remainder gap 3: 1 0 2 0 0 0 . . .
CGCTAAC... → _CG__CTAAC...
Seq3
Gap-aligned seq3
4
Reference
Position : 12 34 56...
Gap-aligned
Reference : _CG__CG_AC...
Seq1 : CCG__CG_AC...
Seq2 : __GCCCG_AC...
Seq3 : _CG__CTAAC...

Figure S2. Reference Gap Alignment (RGA) steps, related to Figure 1.

The four steps in the RGA method are illustrated with matching numbers to the steps in the main text.

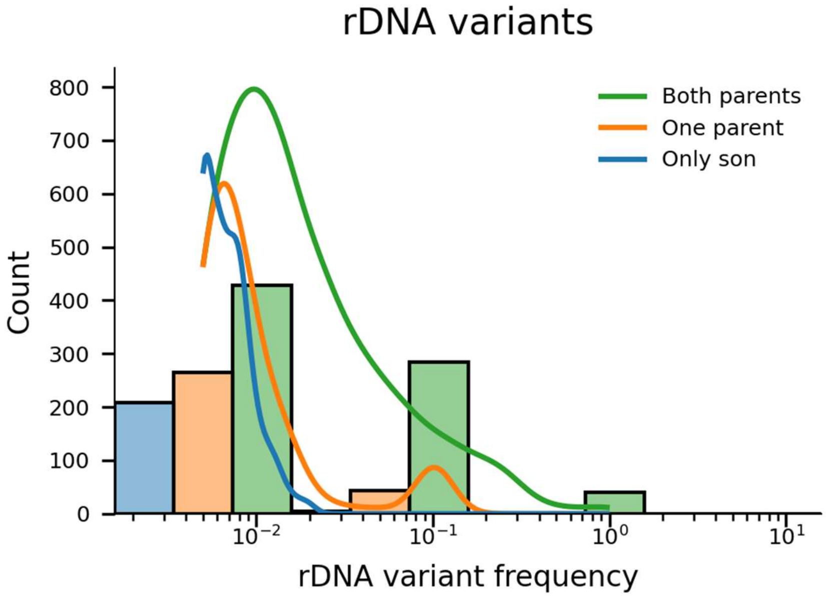


Figure S3. rDNA variants frequecies from a progeny sample support variant heritability, related to Figure 1.

rDNA variant frequencies of the GIAB Ashkenazy son are shown in a histogram and are color coded in blue if the are not found in neither parents, orange if found in only one of the parents and in green if found in both.

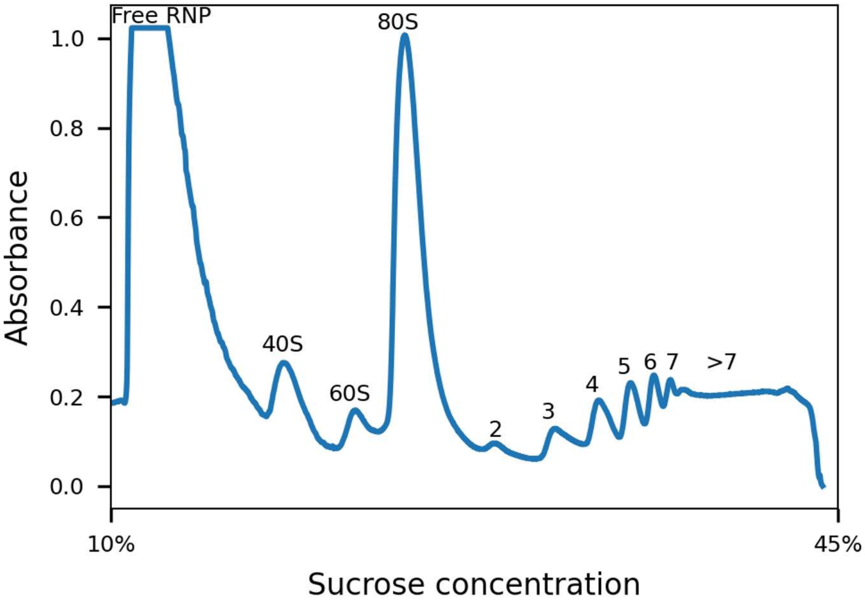


Figure S4. polysome profile from H7-hESC 10-45% sucrose gradient fractionation, related to Figure 1.

H7-hESC A260 trace showing the free ribonucleoproteins (RNP), free 40S and 60S subunits, 80S monosomes, and polysomes (marked with 2-7 and >7).

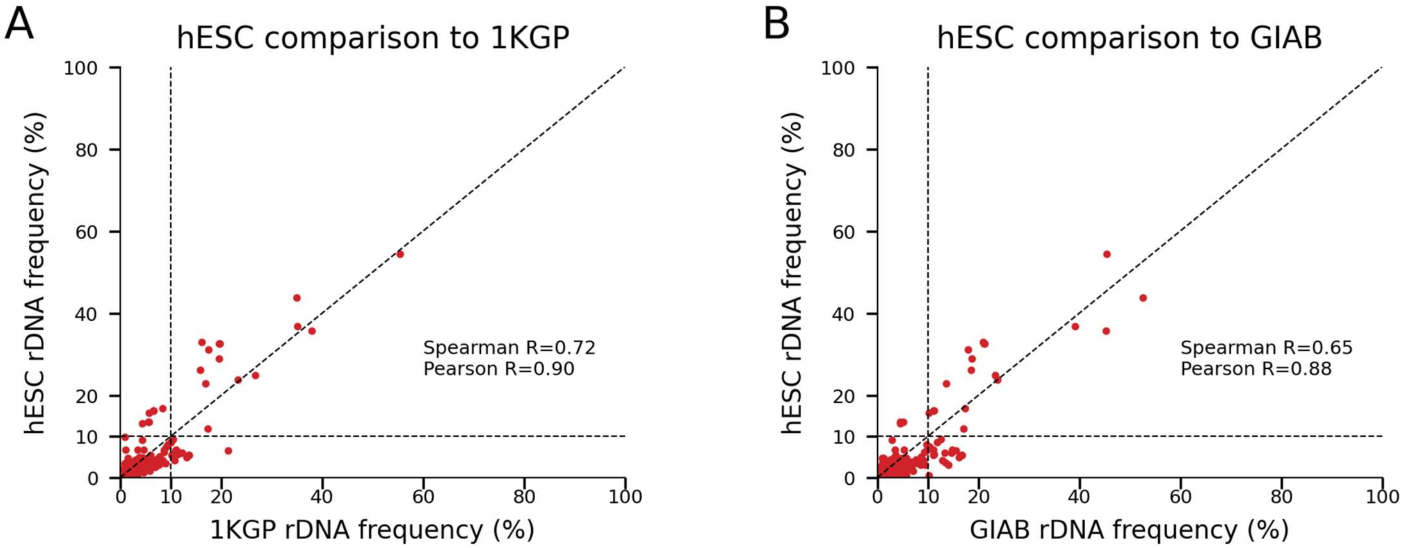


Figure S5: 28S rDNA variant frequencies comparing 1KGP and GIAB to the H7-hESC, related to Figure 1.

A. Scatter plot comparing 28S rDNA variant frequencies found in long-read of the 1KGP (x-axis) and in the H7-hESC (y-axis). Pearson and Spearman correlations are indicated.

B. Same as (A) for GIAB (x-axis).

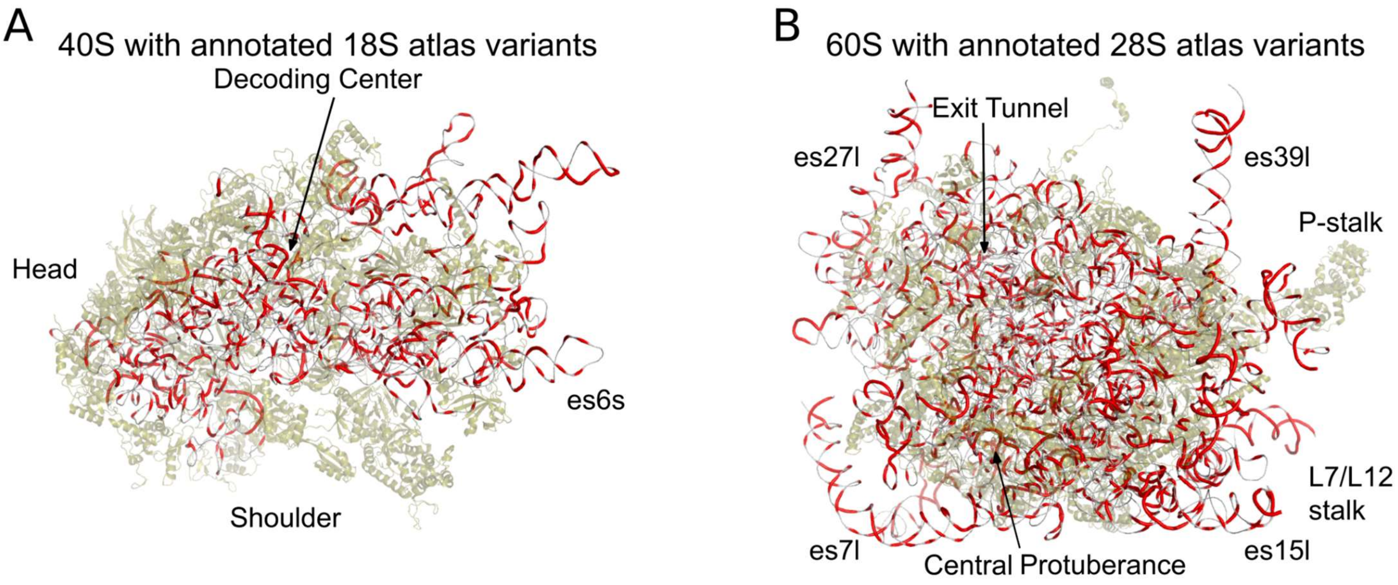


Figure S6. 3D structure of the 40S and 60S with variants, related to Figure 1.

A. 40S 3D structure with 18S atlas variants. Ribosomal proteins are presented in semitransparent green, 18S rRNA in gray. Positions with variants are highlighted in red.

B. Same as (A) for the 60S and the 28S atlas variants.

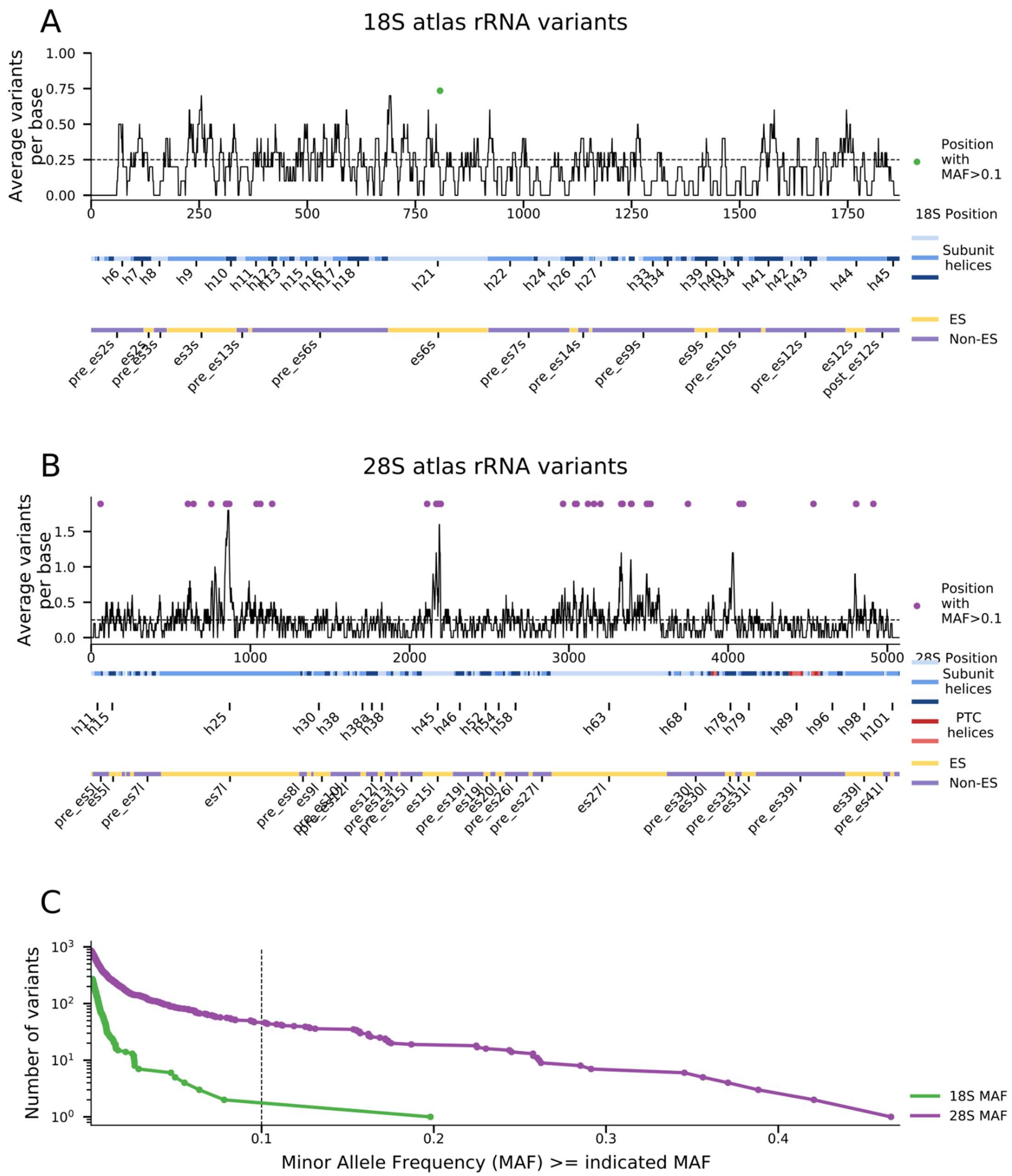


Figure S7. 18S and 28S variant distributions, related to Figure 1.

A. Number of variants at 18S positions. X axis is the nucleotide position along the 18S. Below the X axis are annotations for 18S helix regions and ES regions. Y axis is the average number of atlas variants at a given position for a window size of 20 bases. Green dot at x = 807 (at helix h21 or es6s) is a position with a rRNA variant with minor allele frequency (MAF) >0.1. This dot is also plotted in panel (C).
B. Same as (A) but for the 28S. Purple dots annotate all positions with a variant with MAF>0.1. All purple dots are presented in panel (C).
C. rRNA allele frequency spectrum plot. Values in the X-axis indicate a MAF, and the Y axis are the number of variants with at least the X-axis matched MAF. For example, there are two variants with MAF >0.4 and 6 variants with MAF>0.3. The green plot matches 18S variants and purple plot matches the 28S variants. Individual variants are marked with a dot on the line plot. Dashed line indicates MAF equal to 0.1. The green and purple dots which are in panels (A) and (B) are on the right side of the dashed black line marking MAF=0.1.

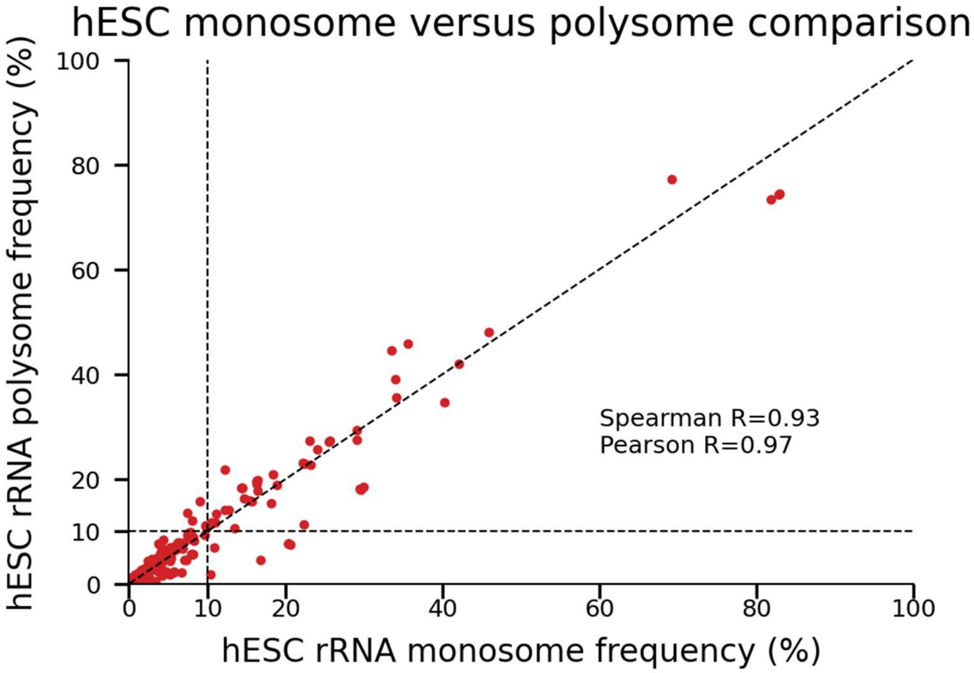


Figure S8. H7-hESC 28S rRNA variant frequencies in monosomes and polysomes, related to Figure 1.

Scatter plot of 28S rRNA variant frequencies from the H7-hESC comparing polysome fractions. In the x-axis rRNA variant frequencies are calculated from monosomes and in the y-axis from polysomes. Pearson and Spearman correlations are indicated.

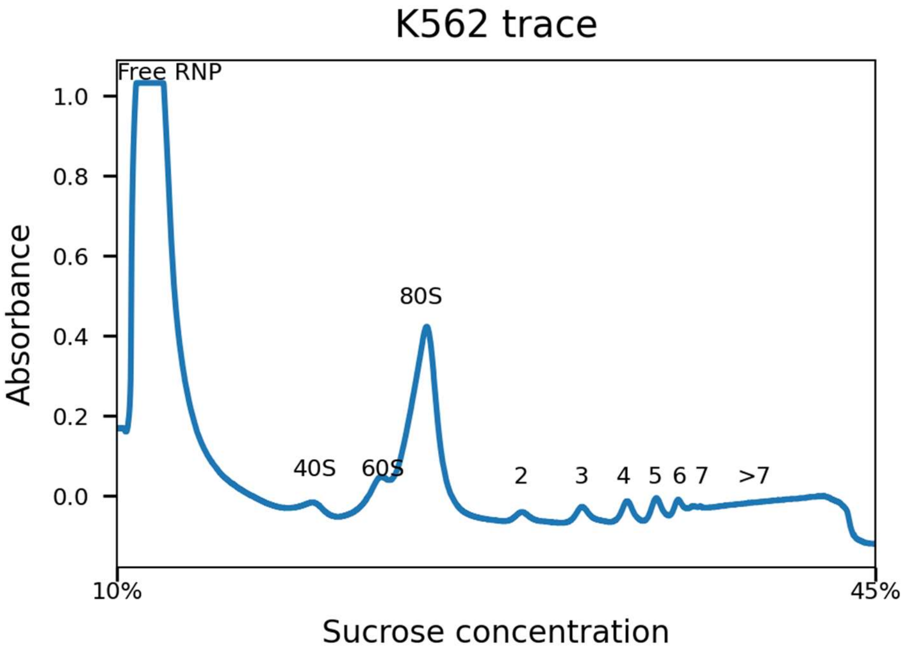


Figure S9. polysome profile from K562 cell-line in 10-45% sucrose gradient fractionation, related to Figure 1.

K562 cancer cell-line A260 trace showing the free ribonucleoproteins (RNP), free 40S and 60S subunits, 80S monosomes, and polysomes (marked with 2-7 and >7).

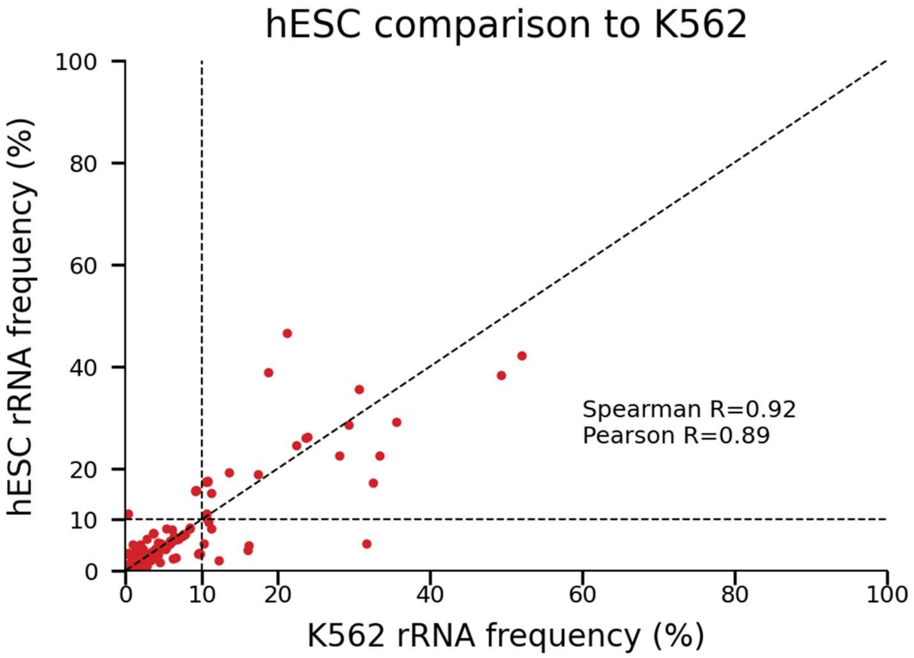


Figure S10. 28S rRNA variant frequencies comparing K562 to the H7-hESC, related to Figure 1.

Scatter plot comparing 28S rRNA variant frequencies found in the K562 cancer cell-line (x-axis) and in the H7-hESC (y-axis). Pearson and Spearman correlations are indicated.

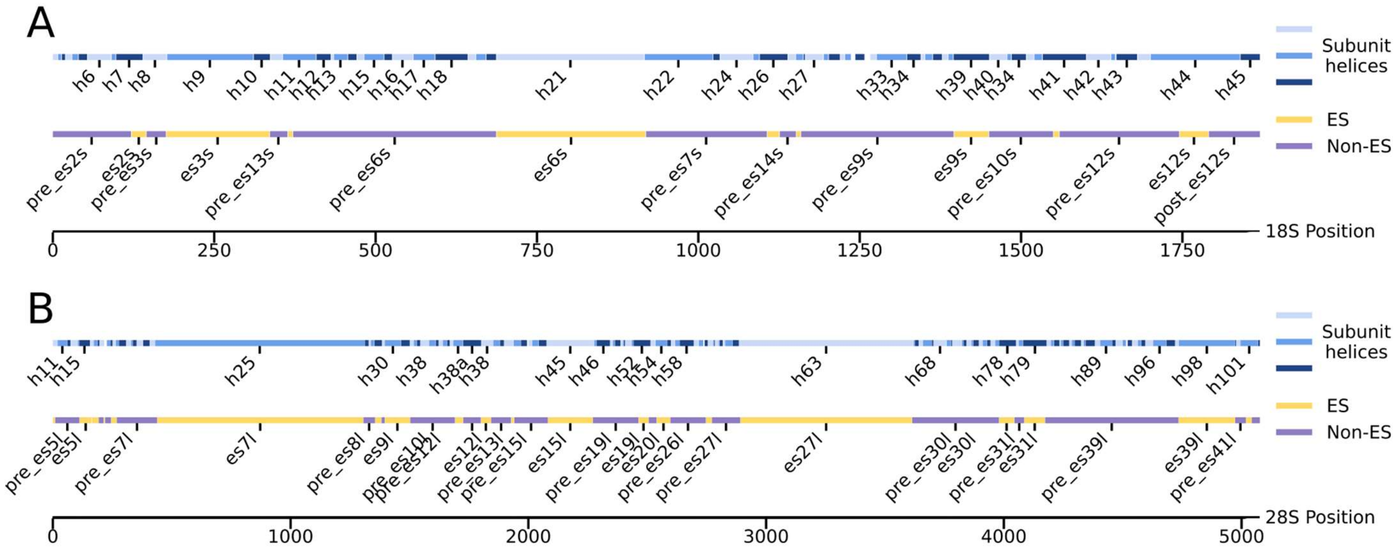


Figure S11. 18S and 28S region annotation, related to Figure 1.

A. 18S helix and ES region annotation. Only helices and ES regions with at least 20 bases are labeled

B. 28S helix and ES region annotation. Only helices and ES regions with at least 40 bases are labeled

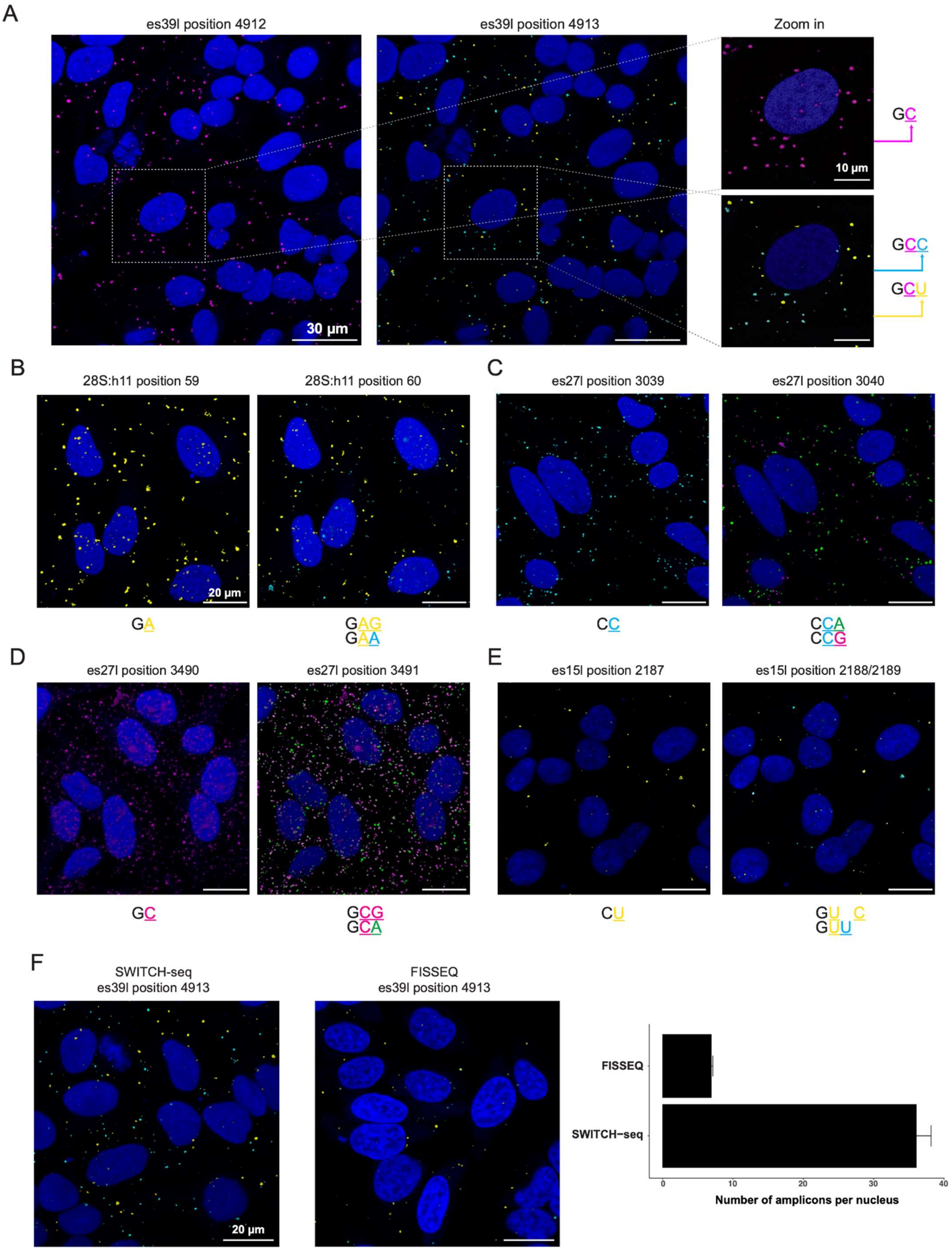

A
es39l position 4912
es39l position 4913
Zoom in
30 µm
10 µm
GC
GCC
GCU
B
28S:h11 position 59
28S:h11 position 60
20 µm
GA
GAG
GAA
C
es27l position 3039
es27l position 3040
CC
CCA
CCG
D
es27l position 3490
es27l position 3491
GC
GCG
GCA
E
es15l position 2187
es15l position 2188/2189
CU
GU C
GUU
F
SWITCH-seq
es39l position 4913
FISSEQ
es39l position 4913
20 µm
FISSEQ
SWITCH−seq
0
10
20
30
40
Number of amplicons per nucleus

Figure S12. *In-situ* sequencing of rRNA variants, related to Figure 2.

A. Two rounds of representative fluorescent *in situ* sequencing images of HeLa cells (DAPI staining in blue) are presented for the es39l-probed region. We identified a non-variable base C (magenta) at position 4912. At position 4913, two alternative sequences were revealed: the known reference sequence C (cyan) and the alternative variant U (yellow).
B. Similar to (A) for 28S:h11 where G and A are detected at position 60
C. Similar to (A) for es27l where A and G are detected at position 3040
D. Similar to (A) for es27l where G and A are detected at position 3491
E. Similar to (A) for es15 where a U insertion is detected at position 2188.
F. Representative fluorescence images comparing SWITCH-seq and FISSEQ in the detection of the rRNA variant at es39l position 4913.

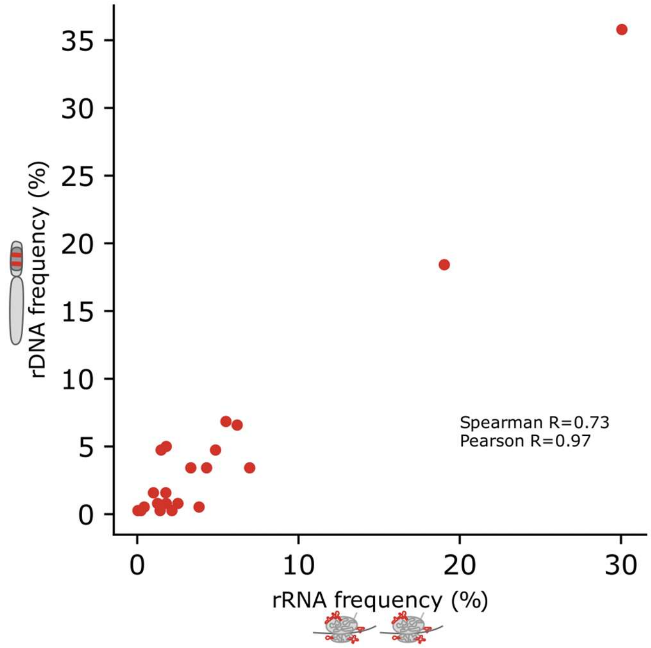


Figure S13. 28S rRNA subtype frequencies in h7-hESC, related to Figure 3.

Scatter plot showing the frequencies of 28S rRNA subtypes in rRNA (x-axis) and rDNA (y-axis) in H7-hESC

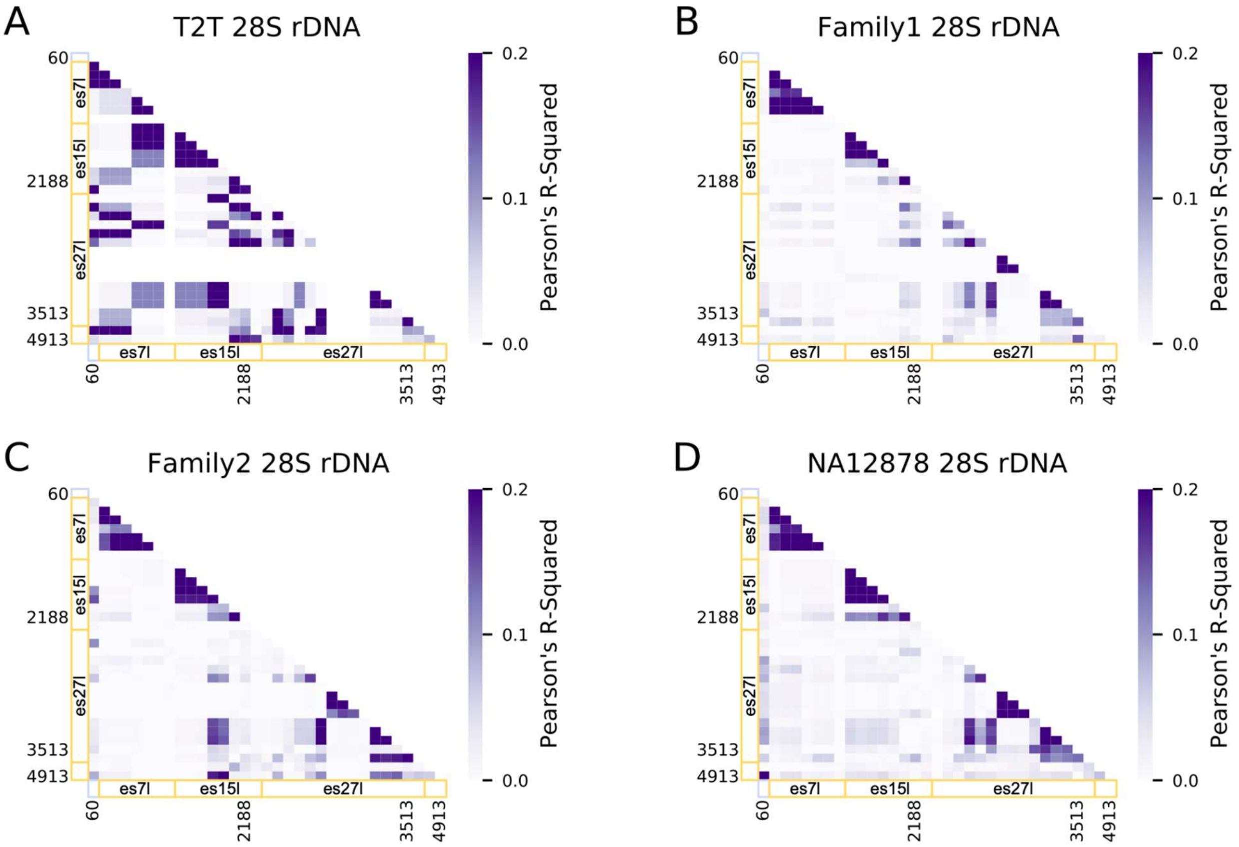
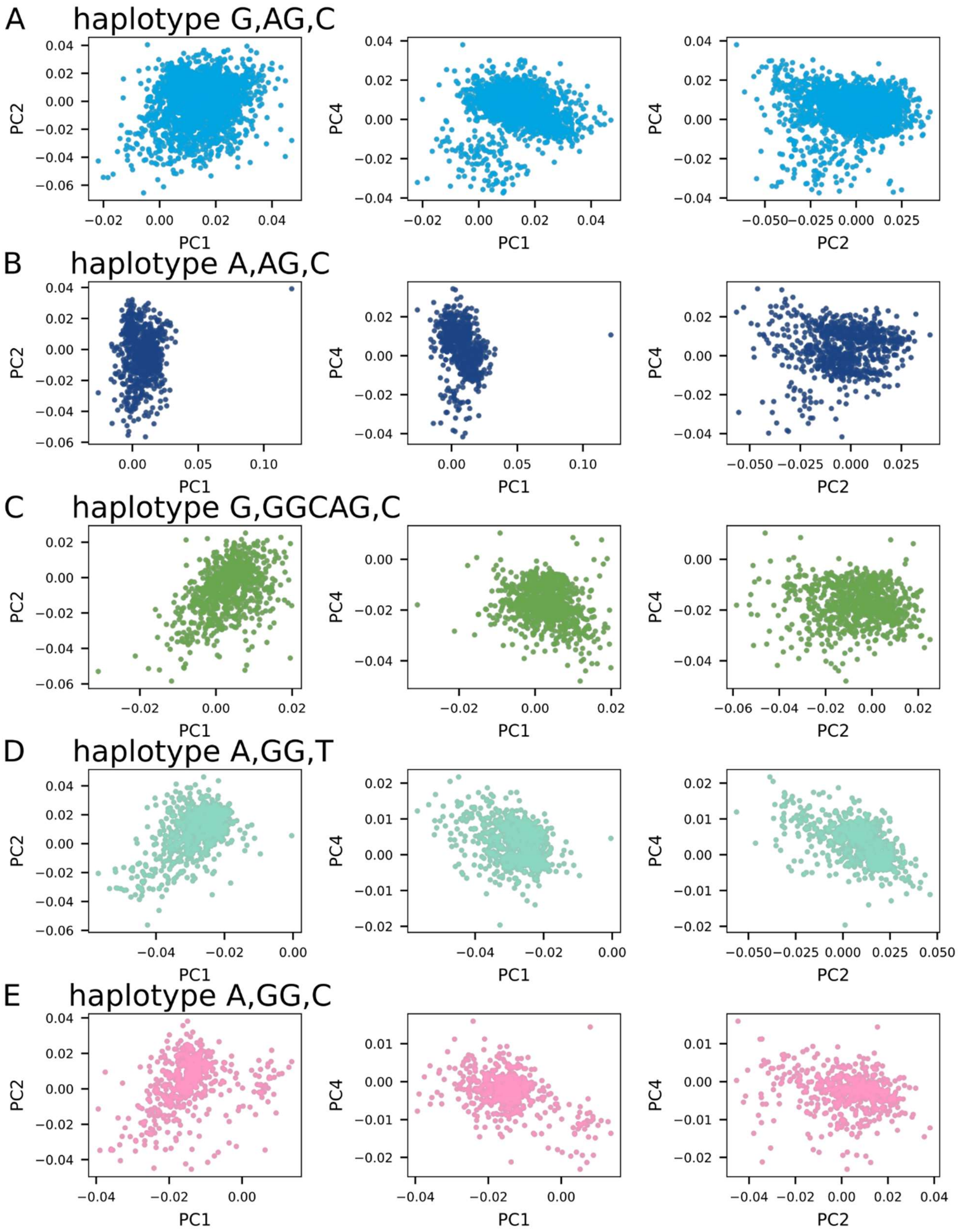


Figure S14. 28S subtypes in T2T and GIAB, related to Figure 3.

A. Correlation coefficient (Pearson's r^2) heatmap of T2T 28S rDNA for the same positions analyzed in the h7-hESC (**Figure 1F, 3A**). X-axis and Y-axis are annotated by regions. Helix regions are annotated by light blue and ES regions are annotated by yellow. H7-hESC positions with higher r^2 are annotated.
B. Same as (A) for the Chinese Han family trio from the GIAB dataset.
C. Same as (A) for the Ashkenazi family trio from the GIAB dataset.
D. Same as (A) for the HA12878 cell line from the GIAB dataset.



A haplotype G,AG,C

B haplotype A,AG,C

C haplotype G,GGCAG,C

D haplotype A,GG,T

E haplotype A,GG,C

Figure S15. 28S haplotype groups as shown by Bray-Curtis Principal Coordinate Analysis (PCoA), related to Figure 3.

PCoA of 386 28S rDNA sequences from each GIAB sample. The first, second and fourth PCs are presented for 28S sequences that belong to different haplotypes where each haplotype is presented in a separte panel A-E. Each dot is a complete 28S rDNA sequence with similarity between sequences measured on 6-mers. The colors correspond to coloring an rDNA sequence by its 3 position haplotype shown in the main Figure 3D.

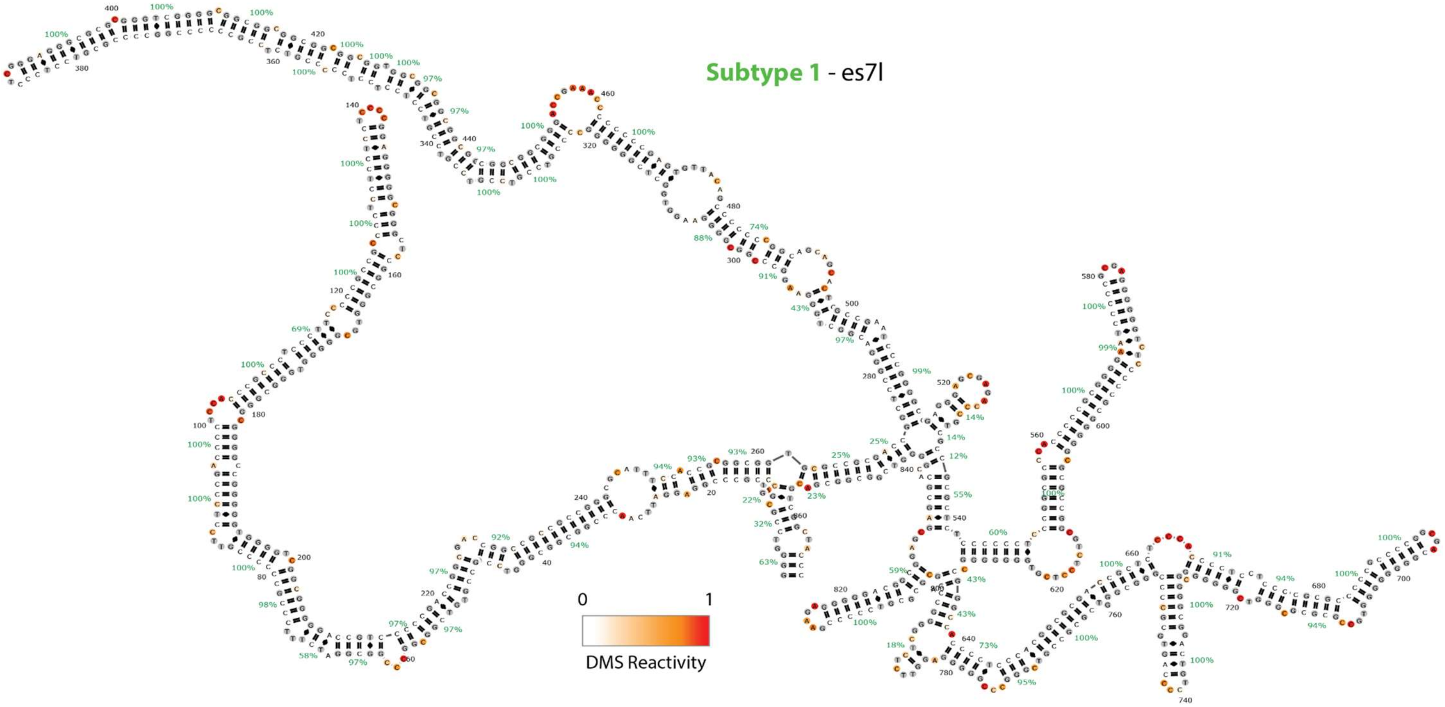


Figure S16. In-cell DMS with long-read sequencing shows accessibility differences in es7l of different 28S subtypes, related to Figure 4.

RNA secondary structure of es7l predicted secondary structure for subtype 1 (A, GGCAG, T).

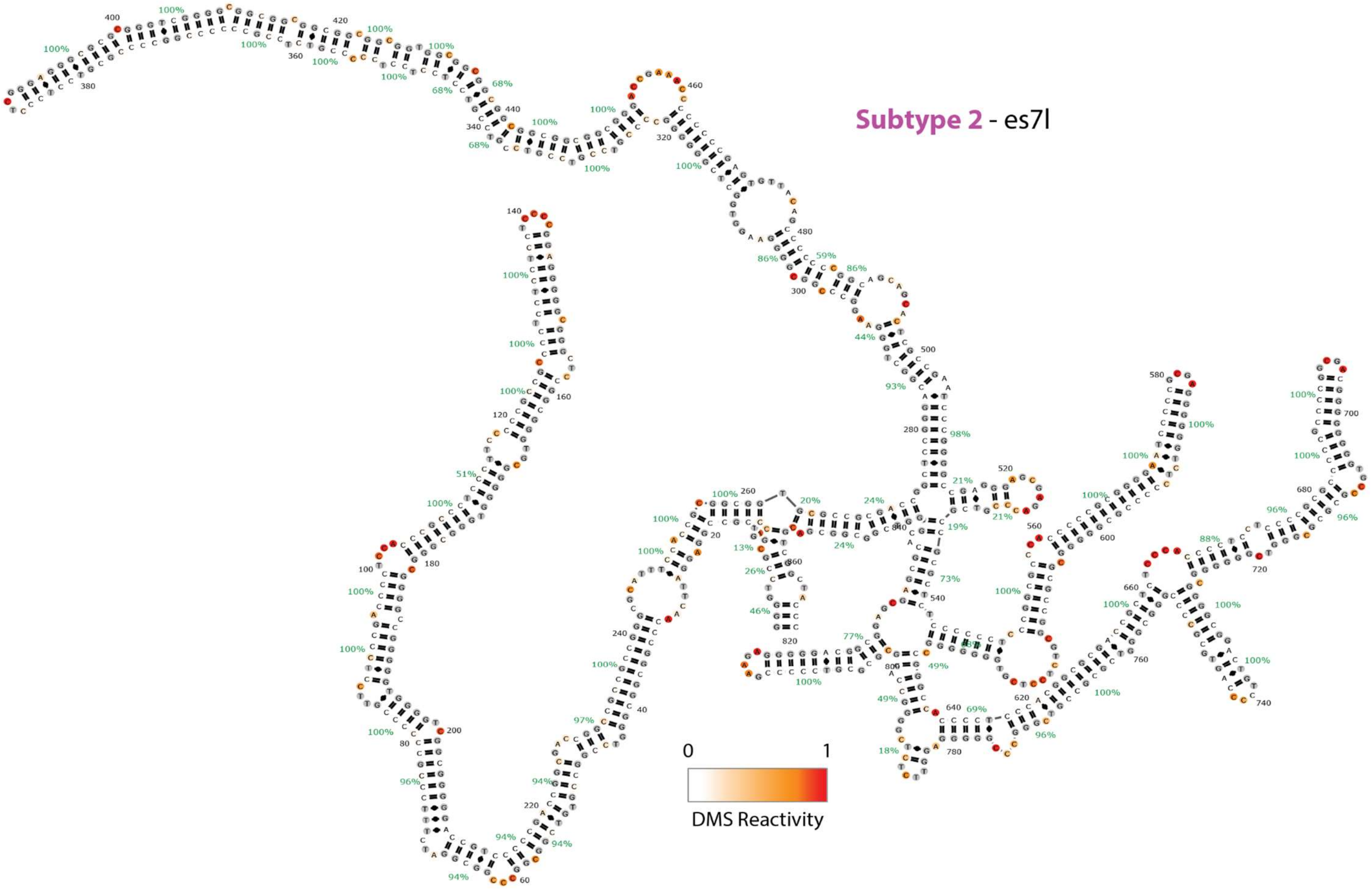


Figure S17. In-cell DMS with long-read sequencing shows accessibility differences in es7l of different 28S subtypes, related to Figure 4.

RNA secondary structure of es7l predicted secondary structure for subtype 2 (G, AG, C).

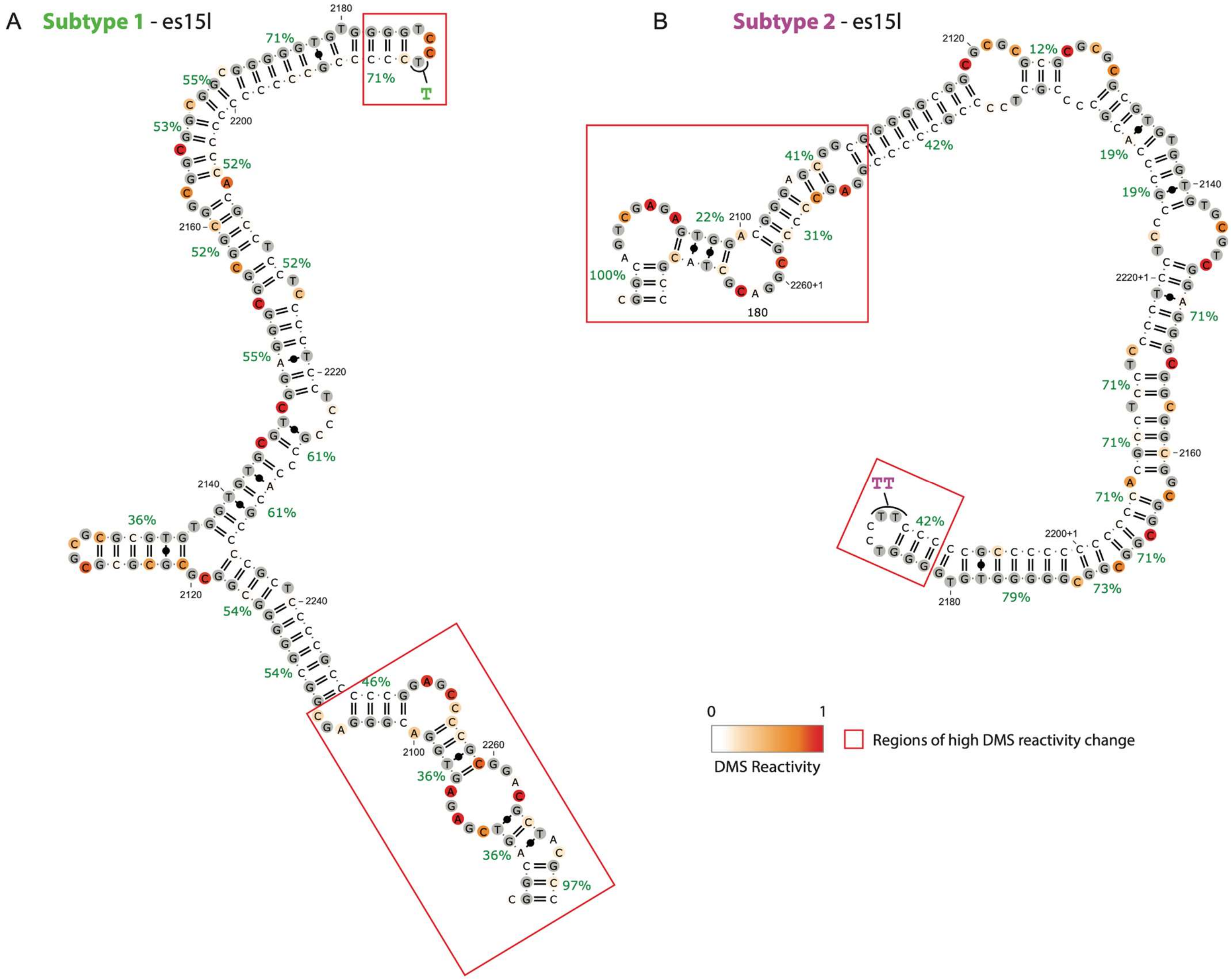


Figure S18. In-cell DMS with long-read sequencing shows that es15l of different 28S subtypes have different RNA 2D structure, related to Figure 4.

A. RNA secondary structure of es15l predicted secondary structure for subtype 1 (A, GGCAG, T).
B. RNA secondary structure of es15l predicted secondary structure for subtype 2 (G, AG, C).

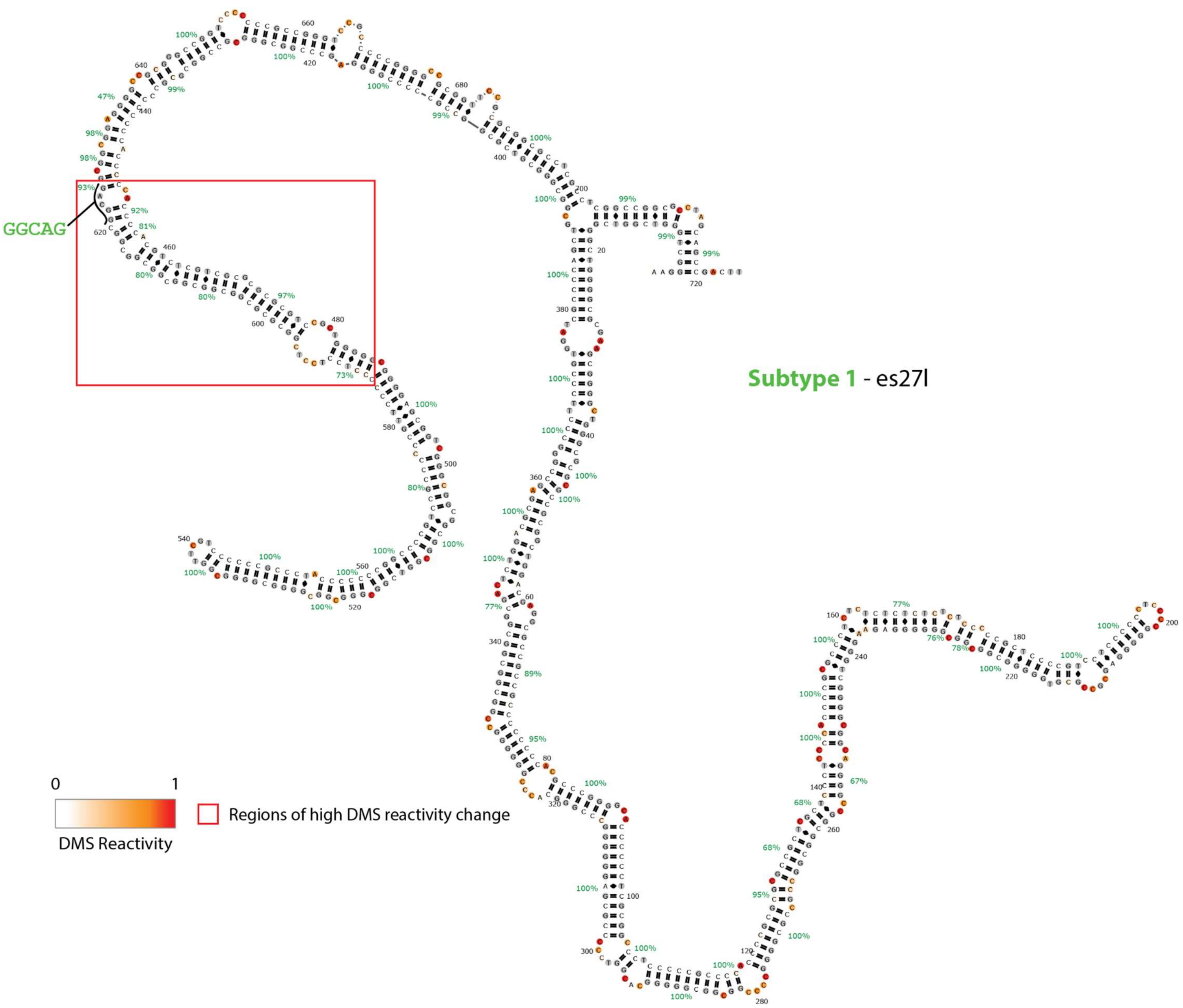


Figure S19. In-cell DMS with long-read sequencing shows that es27l of different 28S subtypes have different RNA 2D structure, related to Figure 4.

RNA secondary structure of es15l predicted secondary structure for subtype 1 (A, GGCAG, T).

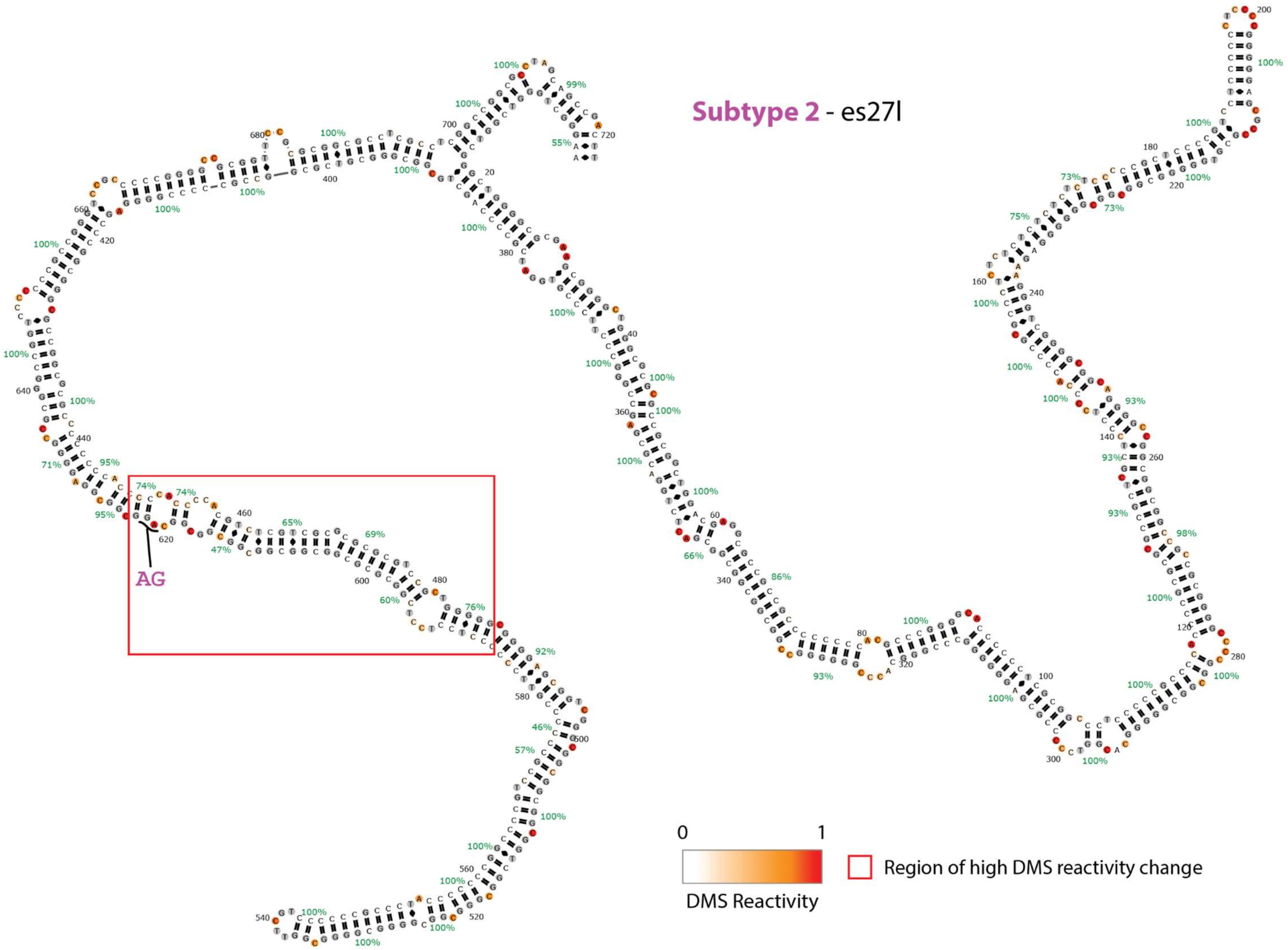


Figure S20. In-cell DMS with long-read sequencing shows that es27l of different 28S subtypes have different RNA 2D structure, related to Figure 4.

RNA secondary structure of es15l predicted secondary structure for subtype 2 (G, AG, C).

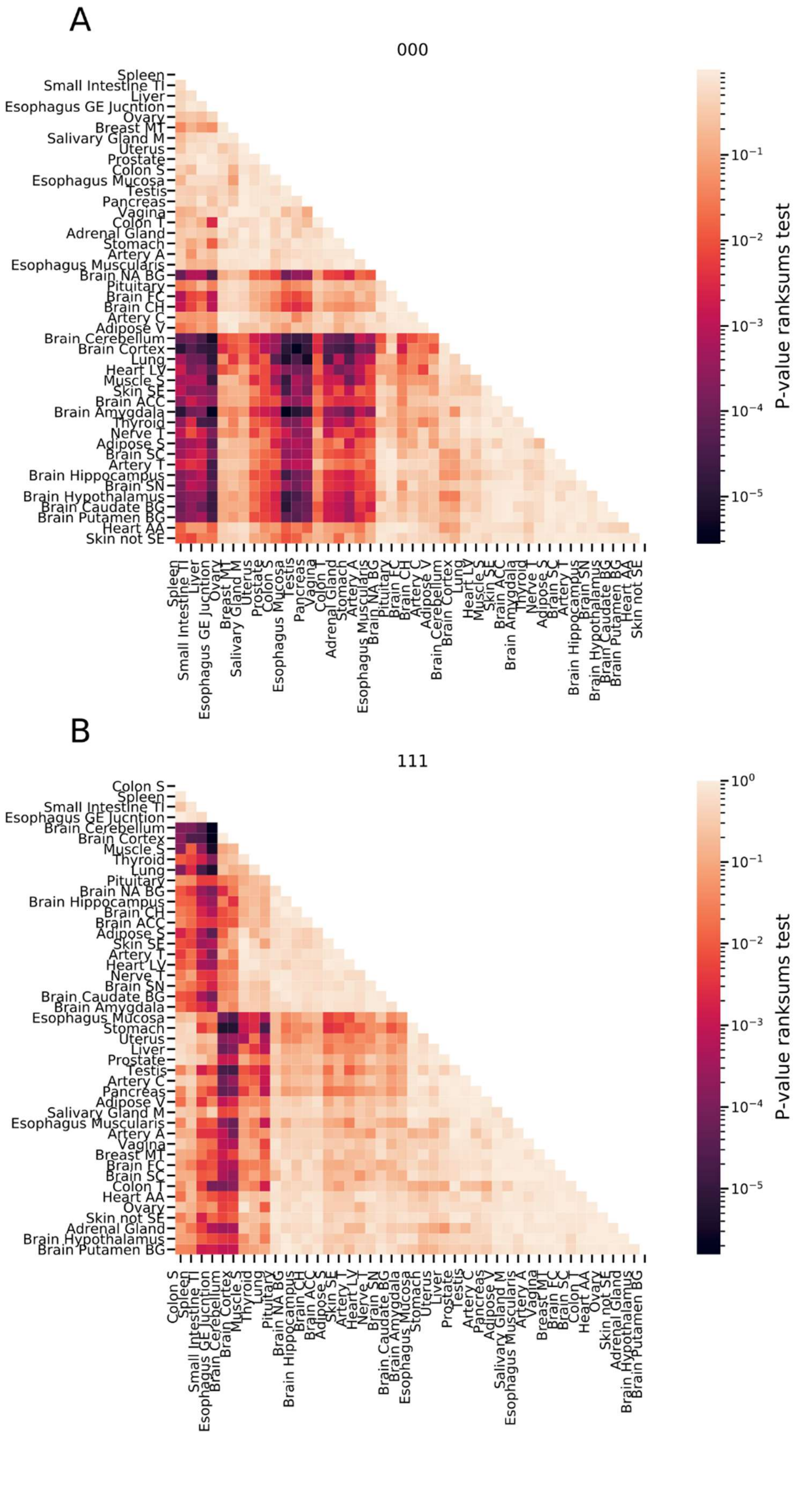

A
000
Spleen
Small Intestine TI
Liver
Esophagus GE Jucntion
Ovary
Breast MT
Salivary Gland M
Uterus
Prostate
Colon S
Esophagus Mucosa
Testis
Pancreas
Vagina
Colon T
Adrenal Gland
Stomach
Artery A
Esophagus Muscularis
Brain NA BG
Pituitary
Brain FC
Brain CH
Artery C
Adipose V
Brain Cerebellum
Brain Cortex
Lung
Heart LV
Muscle S
Skin SE
Brain ACC
Brain Amygdala
Thyroid
Nerve T
Adipose S
Brain SC
Artery T
Brain Hippocampus
Brain SN
Brain Hypothalamus
Brain Caudate BG
Brain Putamen BG
Heart AA
Skin not SE
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
P-value ranksums test
B
111
Colon S
Spleen
Small Intestine TI
Esophagus GE Jucntion
Brain Cerebellum
Brain Cortex
Muscle S
Thyroid
Lung
Pituitary
Brain NA BG
Brain Hippocampus
Brain CH
Brain ACC
Adipose S
Skin SE
Artery T
Heart LV
Nerve T
Brain SN
Brain Caudate BG
Brain Amygdala
Esophagus Mucosa
Stomach
Uterus
Liver
Prostate
Testis
Artery C
Pancreas
Adipose V
Salivary Gland M
Esophagus Muscularis
Artery A
Vagina
Breast MT
Brain FC
Brain SC
Colon T
Heart AA
Ovary
Skin not SE
Adrenal Gland
Brain Hypothalamus
Brain Putamen BG
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
P-value ranksums test

Figure S21 (1 out of 5 similar plots). rRNA subtype expression levels are tissue specific, related to Figure 5.

A. A heatmap showing FDR-corrected rank sum test P-values comparing the expression levels of the rRNA subtype with the haplotype sequence of G,AG,C (titled 000) across tissues. The tissues are ordered by average hierarchical clustering of the rank sum corrected P-values.
B. Same as (A) for haplotype A,GGCAG,T (titled 111) across tissues.

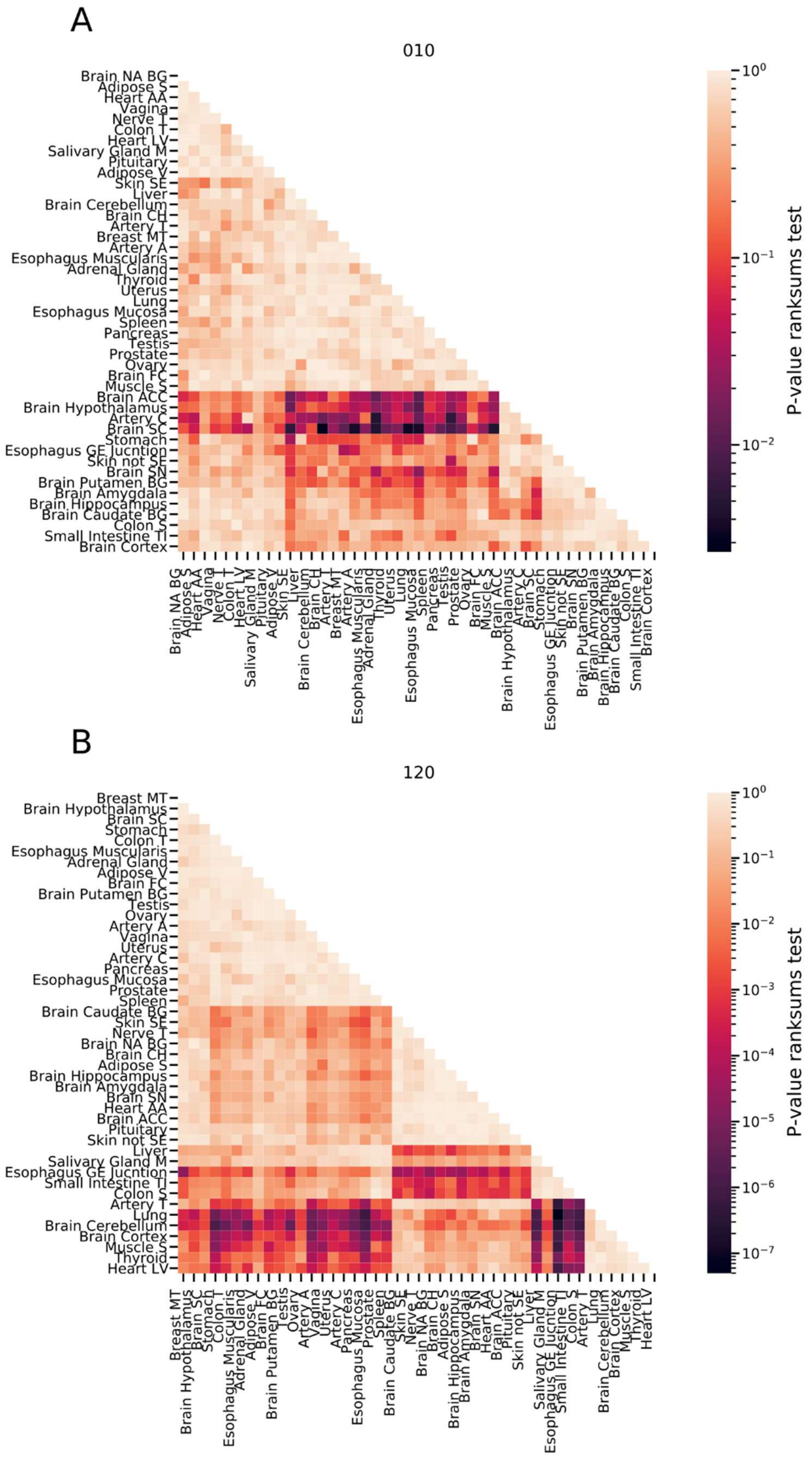

A
010
Brain NA BG
Adipose S
Heart AA
Vagina
Nerve T
Colon T
Heart LV
Salivary Gland M
Pituitary
Adipose V
Skin SE
Liver
Brain Cerebellum
Brain CH
Artery T
Breast MT
Artery A
Esophagus Muscularis
Adrenal Gland
Thyroid
Uterus
Lung
Esophagus Mucosa
Spleen
Pancreas
Testis
Prostate
Ovary
Brain FC
Muscle S
Brain ACC
Brain Hypothalamus
Artery C
Brain SC
Stomach
Esophagus GE Jucntion
Skin not SE
Brain SN
Brain Putamen BG
Brain Amygdala
Brain Hippocampus
Brain Caudate BG
Colon S
Small Intestine TI
Brain Cortex
10^0
10^{-1}
10^{-2}
P-value ranksums test
B
120
Breast MT
Brain Hypothalamus
Brain SC
Stomach
Colon T
Esophagus Muscularis
Adrenal Gland
Adipose V
Brain FC
Brain Putamen BG
Testis
Ovary
Artery A
Vagina
Uterus
Artery C
Pancreas
Esophagus Mucosa
Prostate
Spleen
Brain Caudate BG
Skin SE
Nerve T
Brain NA BG
Brain CH
Adipose S
Brain Hippocampus
Brain Amygdala
Brain SN
Heart AA
Brain ACC
Pituitary
Skin not SE
Liver
Salivary Gland M
Esophagus GE Jucntion
Small Intestine TI
Colon S
Artery T
Lung
Brain Cerebellum
Brain Cortex
Muscle S
Thyroid
Heart LV
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
P-value ranksums test

Figure S22 (2 out of 5 similar plots). rRNA subtype expression levels are tissue specific, related to Figure 5.

A. A heatmap showing FDR-corrected rank sum test P-values comparing the expression levels of the rRNA subtype with the haplotype sequence of G,GGCAG,C (titled 010) across tissues**.** The tissues are ordered by average hierarchical clustering of the rank sum corrected P-values.
B. Same as (A) for haplotype A,GG,C (titled 120) across tissues**.**

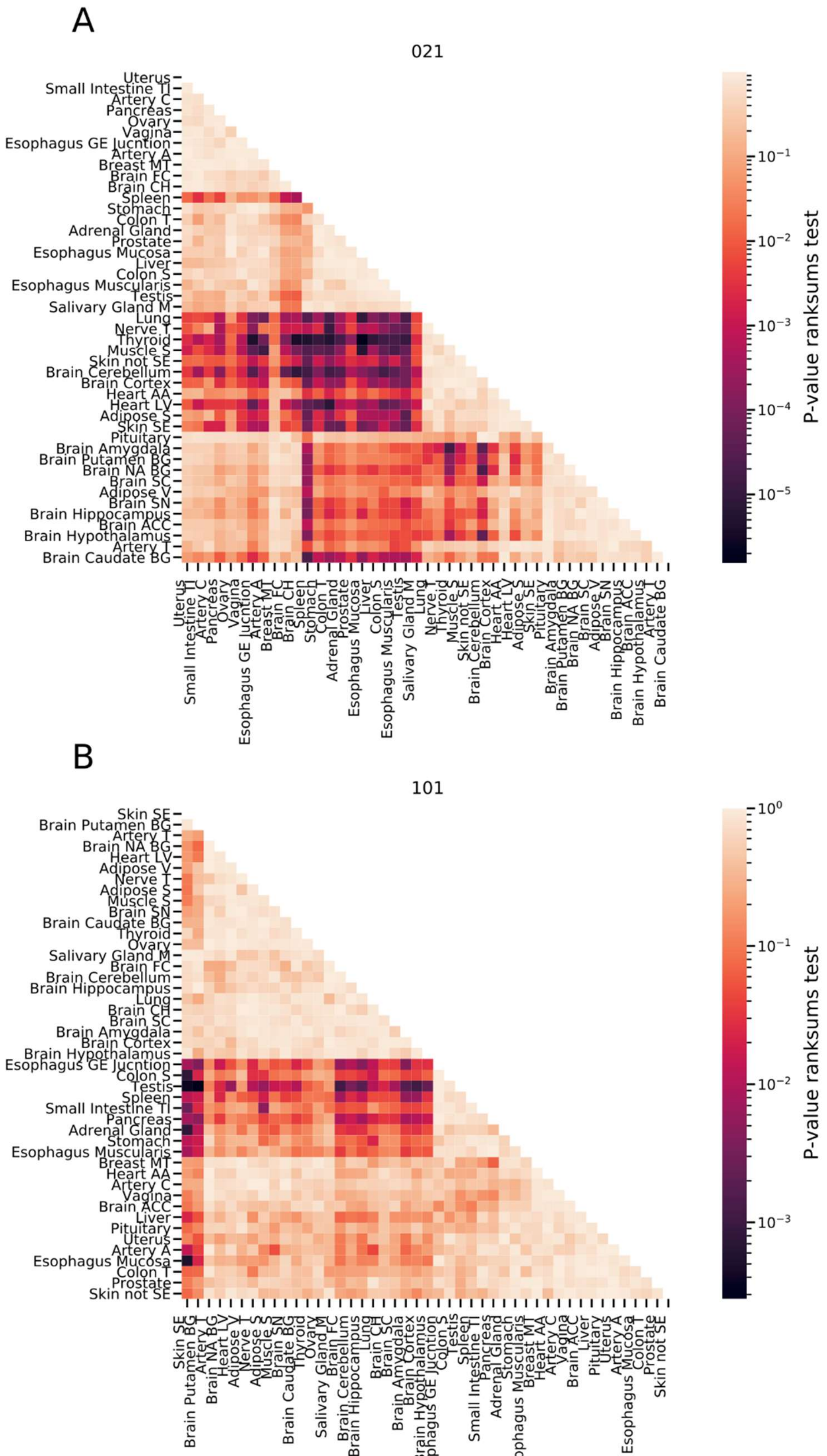

A
021
Uterus
Small Intestine TI
Artery C
Pancreas
Ovary
Vagina
Esophagus GE Jucntion
Artery A
Breast MT
Brain FC
Brain CH
Spleen
Stomach
Colon T
Adrenal Gland
Prostate
Esophagus Mucosa
Liver
Colon S
Esophagus Muscularis
Testis
Salivary Gland M
Lung
Nerve T
Thyroid
Muscle S
Skin not SE
Brain Cerebellum
Brain Cortex
Heart AA
Heart LV
Adipose S
Skin SE
Pituitary
Brain Amygdala
Brain Putamen BG
Brain NA BG
Brain SC
Adipose V
Brain SN
Brain Hippocampus
Brain ACC
Brain Hypothalamus
Artery T
Brain Caudate BG
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
P-value ranksums test
B
101
Skin SE
Brain Putamen BG
Artery T
Brain NA BG
Heart LV
Adipose V
Nerve T
Adipose S
Muscle S
Brain SN
Brain Caudate BG
Thyroid
Ovary
Salivary Gland M
Brain FC
Brain Cerebellum
Brain Hippocampus
Lung
Brain CH
Brain SC
Brain Amygdala
Brain Cortex
Brain Hypothalamus
Esophagus GE Jucntion
Colon S
Testis
Spleen
Small Intestine TI
Pancreas
Adrenal Gland
Stomach
Esophagus Muscularis
Breast MT
Heart AA
Artery C
Vagina
Brain ACC
Liver
Pituitary
Uterus
Artery A
Esophagus Mucosa
Colon T
Prostate
Skin not SE
10^{0}
10^{-1}
10^{-2}
10^{-3}
P-value ranksums test

Figure S23 (3 out of 5 similar plots). rRNA subtype expression levels are tissue specific, related to Figure 5.

A. A heatmap showing FDR-corrected rank sum test P-values comparing the expression levels of the rRNA subtype with the haplotype sequence of G,GG,C (titled 021) across tissues. The tissues are ordered by average hierarchical clustering of the rank sum corrected P-values.
B. Same as (A) for haplotype A,AG,T (titled 101) across tissues.

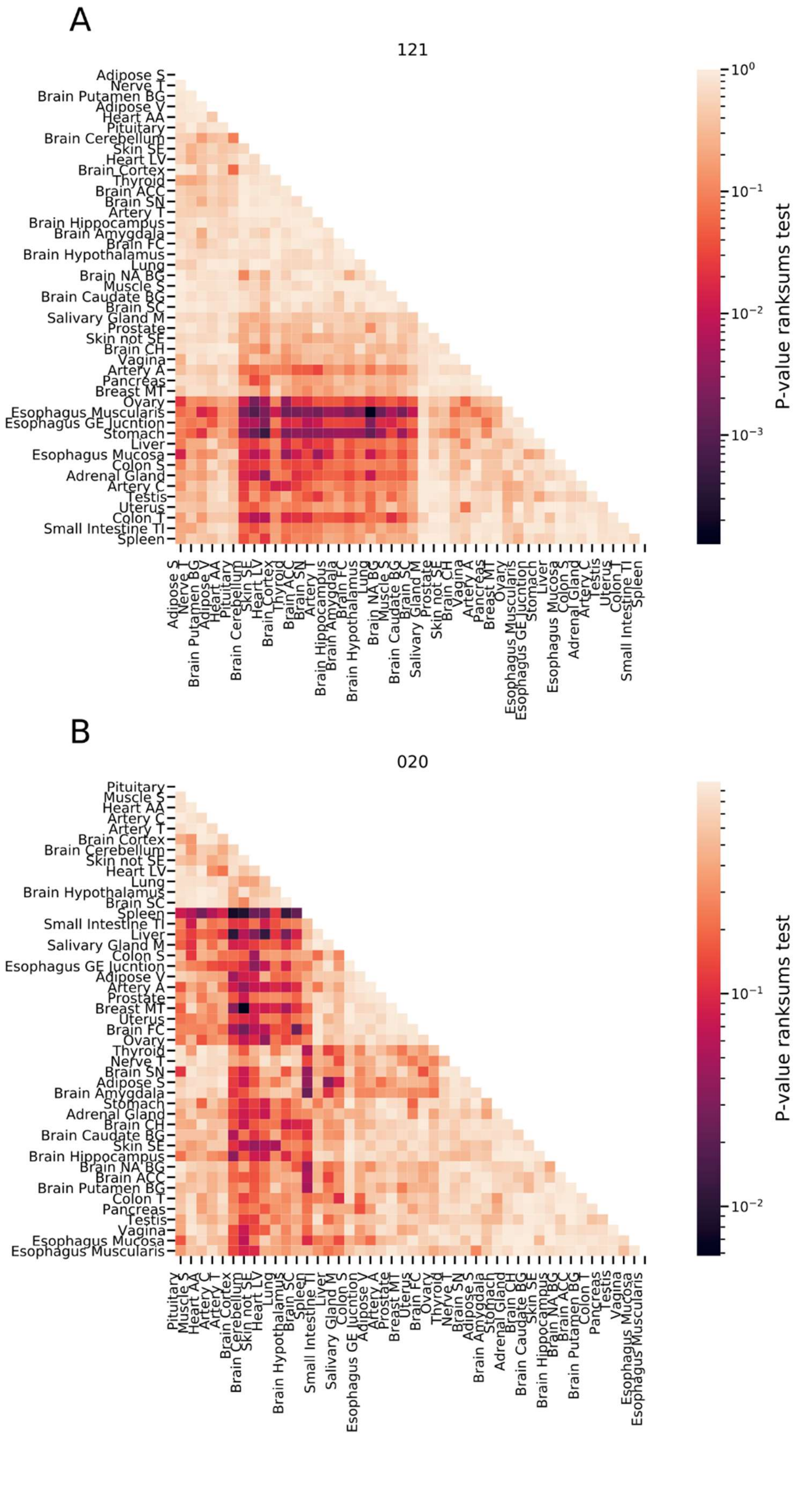

A
121
Adipose S
Nerve T
Brain Putamen BG
Adipose V
Heart AA
Pituitary
Brain Cerebellum
Skin SE
Heart LV
Brain Cortex
Thyroid
Brain ACC
Brain SN
Artery T
Brain Hippocampus
Brain Amygdala
Brain FC
Brain Hypothalamus
Lung
Brain NA BG
Muscle S
Brain Caudate BG
Brain SC
Salivary Gland M
Prostate
Skin not SE
Brain CH
Vagina
Artery A
Pancreas
Breast MT
Ovary
Esophagus Muscularis
Esophagus GE Jucntion
Stomach
Liver
Esophagus Mucosa
Colon S
Adrenal Gland
Artery C
Testis
Uterus
Colon T
Small Intestine TI
Spleen
10^0
10^{-1}
10^{-2}
10^{-3}
P-value ranksums test
B
020
Pituitary
Muscle S
Heart AA
Artery C
Artery T
Brain Cortex
Brain Cerebellum
Skin not SE
Heart LV
Lung
Brain Hypothalamus
Brain SC
Spleen
Small Intestine TI
Liver
Salivary Gland M
Colon S
Esophagus GE Jucntion
Adipose V
Artery A
Prostate
Breast MT
Uterus
Brain FC
Ovary
Thyroid
Nerve T
Brain SN
Adipose S
Brain Amygdala
Stomach
Adrenal Gland
Brain CH
Brain Caudate BG
Skin SE
Brain Hippocampus
Brain NA BG
Brain ACC
Brain Putamen BG
Colon T
Pancreas
Testis
Vagina
Esophagus Mucosa
Esophagus Muscularis
10^{-1}
10^{-2}
P-value ranksums test

Figure S24 (4 out of 5 similar plots). rRNA subtype expression levels are tissue specific, related to Figure 5.

A. A heatmap showing FDR-corrected rank sum test P-values comparing the expression levels of the rRNA subtype with the haplotype sequence of A,GG,T (titled 121) across tissues**.** The tissues are ordered by average hierarchical clustering of the rank sum corrected P-values.
B. Same as (A) for haplotype G,GG,C (titled 020) across tissues**.**

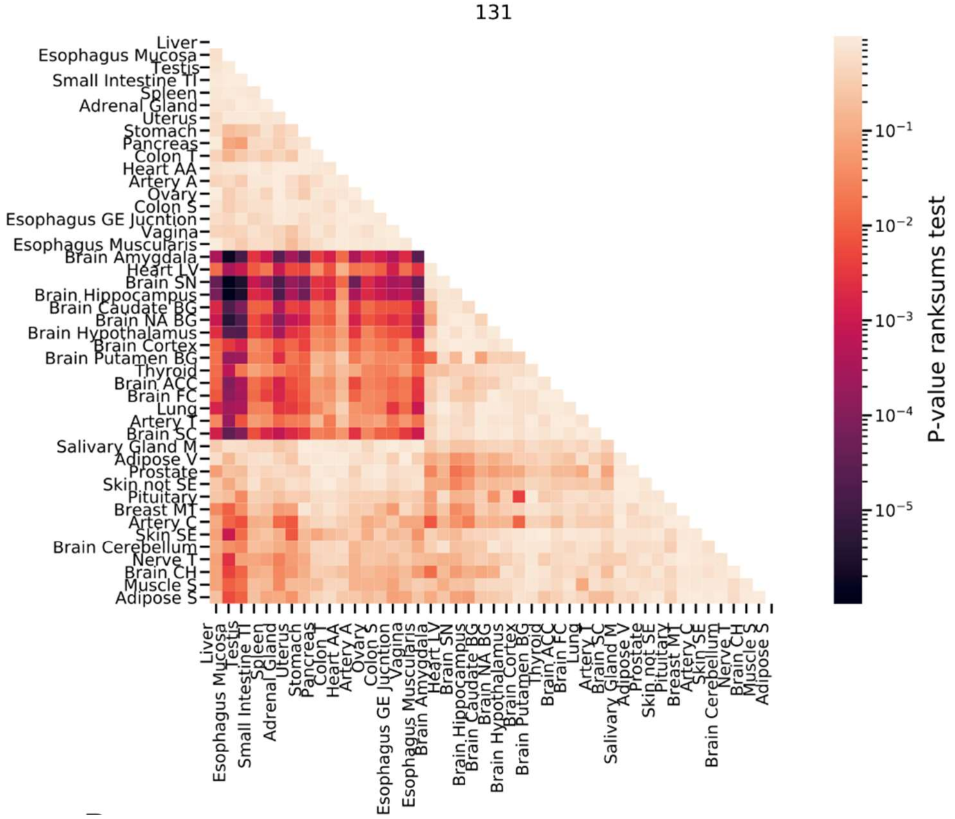


Figure S25 (5 out of 5 similar plots). rRNA subtype expression levels are tissue specific, related to Figure 5.

A heatmap showing FDR-corrected rank sum test P-values comparing the expression levels of the rRNA subtype with the haplotype sequence of A,GGCGGCAG,T (titled 131) across tissues**.** The tissues are ordered by average hierarchical clustering of the rank sum corrected P-values.

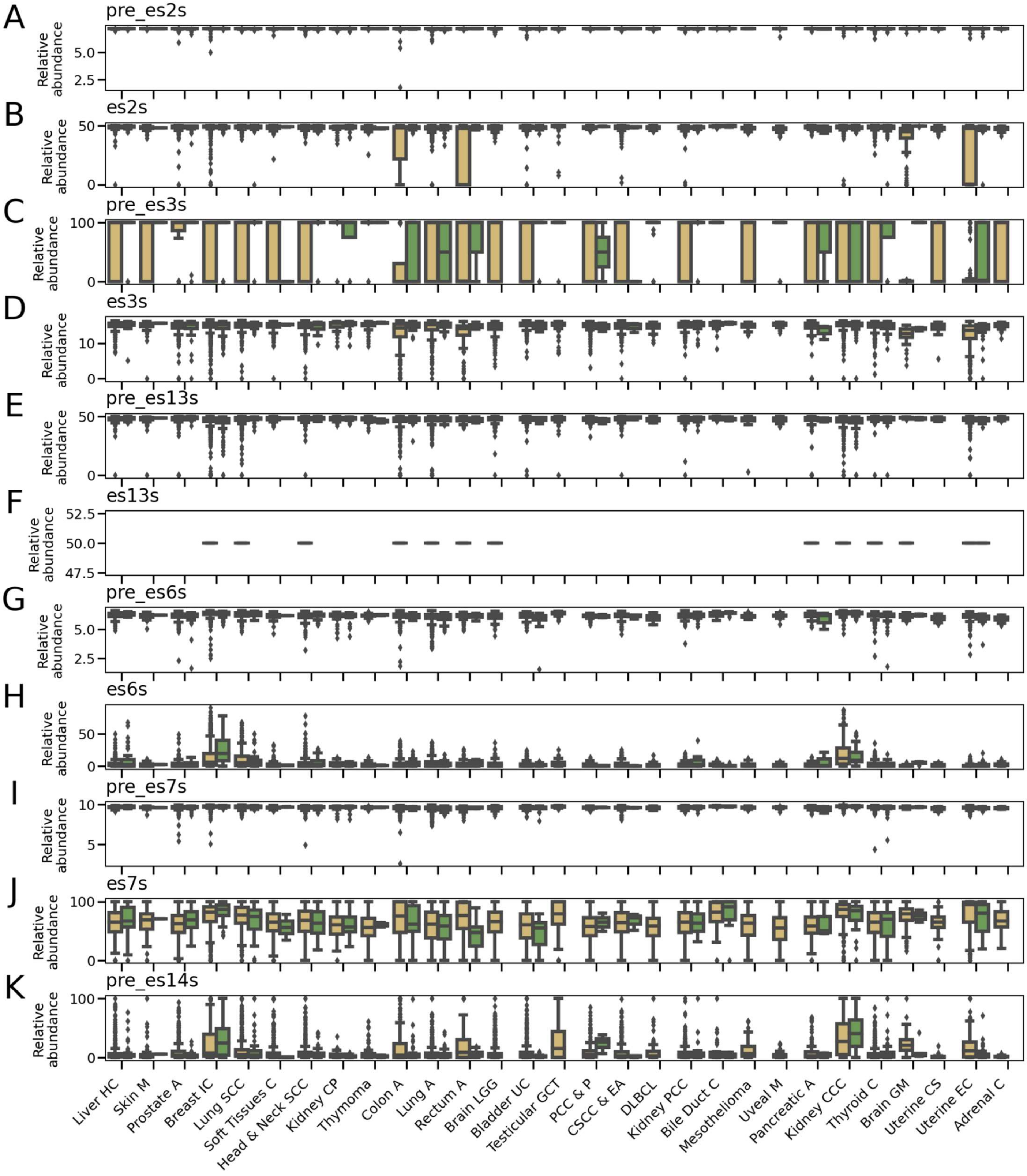


Figure S26. Cancer-specific rRNA variants relative-abundances (1 of 6 figures), related to Figure 6. (A-K) Scatter plot of top abundant regional rRNA variants relative abundances for TCGA cancer and control biopsy samples (cancer and control samples are in yellow and green boxes respectively **Table**

S23 for region to regional variant conversion and the P-value for comparing case/control). The top most abundant rRNA regional variant is presented per ES/non-ES region across tissues. The x-axis is the same for all panels and is displayed in (K).
Abbreviations:

Adrenal C = Adrenocortical Carcinoma
Bile Duct C = Cholangiocarcinoma
Bladder UC = Bladder Urothelial Carcinoma
Brain GM = Brain Glioblastoma Multiforme
Brain LGG = Brain Lower Grade Glioma
Breast IC = Breast Invasive Carcinoma
Colon A = Colon Adenocarcinoma
CSCC & EA = Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma
DLBCL = Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
Head & Neck SCC = Head and Neck Squamous Cell Carcinoma
Kidney CCC = Kidney Renal Clear Cell Carcinoma
Kidney CP = Kidney Chromophobe
Kidney PCC = Kidney Renal Papillary Cell Carcinoma
Liver HC = Liver Hepatocellular Carcinoma
Lung A = Lung Adenocarcinoma
Lung SCC = Lung Squamous Cell Carcinoma
Pancreatic A = Pancreatic Adenocarcinoma
PCC & P = Pheochromocytoma and Paraganglioma
Prostate A = Prostate Adenocarcinoma
Rectum A = Rectum Adenocarcinoma
Skin M = Skin Cutaneous Melanoma
Soft Tissues C = Soft Tissues Carcinoma
Testicular GCT = Testicular Germ Cell Tumors
Thyroid C = Thyroid Carcinoma
Uterine CS = Uterine Carcinosarcoma
Uterine EC = Uterine Corpus Endometrial Carcinoma
Uveal M = Uveal Melanoma

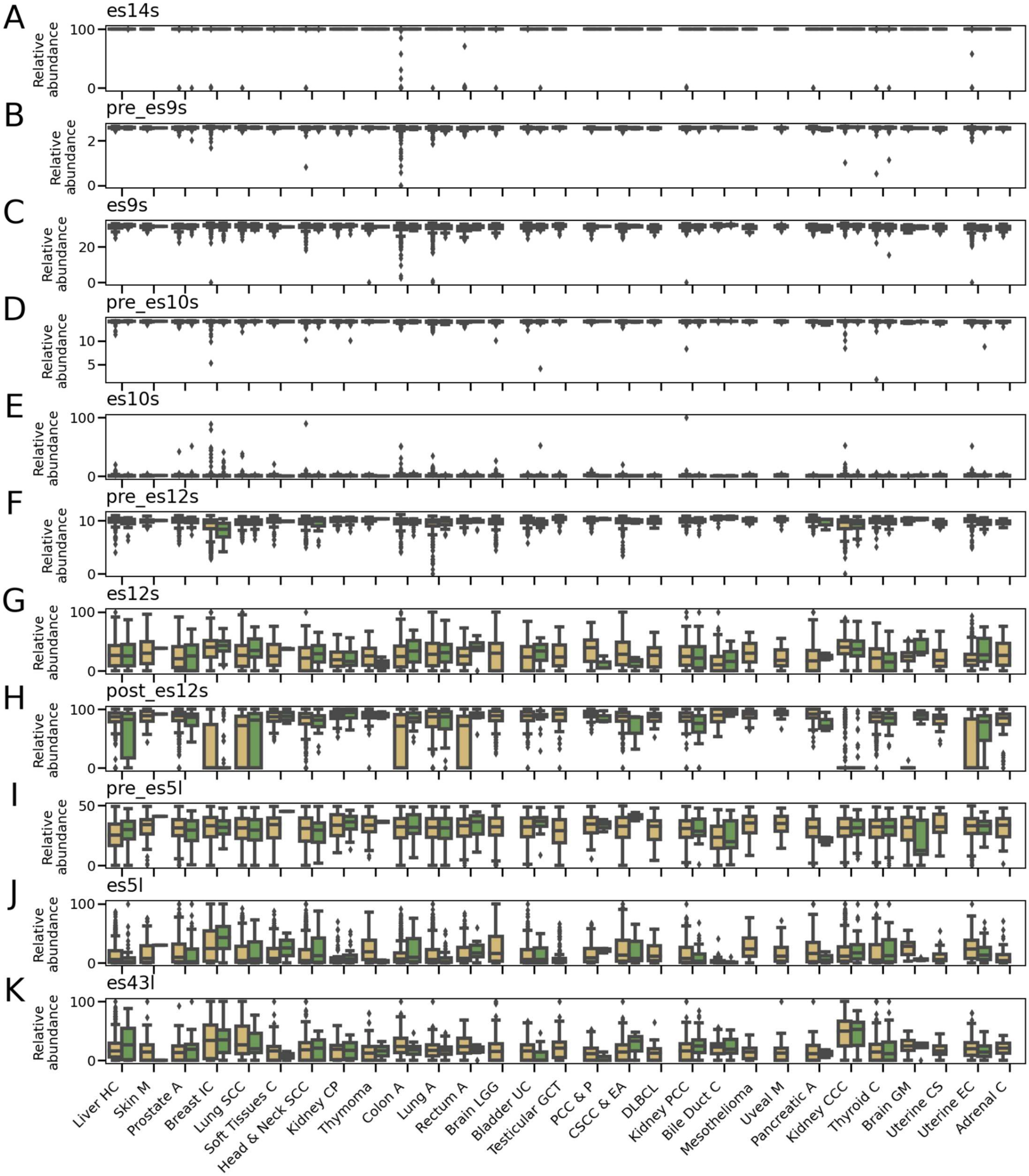


Figure S27. Cancer-specific rRNA variants relative-abundances (2 of 6 figures), related to Figure 6. (A-K) Scatter plot of different regional rRNA variants, relative abundances for TCGA cancer biopsy samples (cancer and control samples are in yellow and green boxes respectively). The top most abundant rRNA regional variant is presented per ES/non-ES region across tissues (**Table S23** for

region to regional variant conversion and the P-value for comparing case/control). The x-axis is the same for all panels and is displayed in (K). X-axis cancer type full name for the abbreviations are listed at the bottom of **Figure S26**.

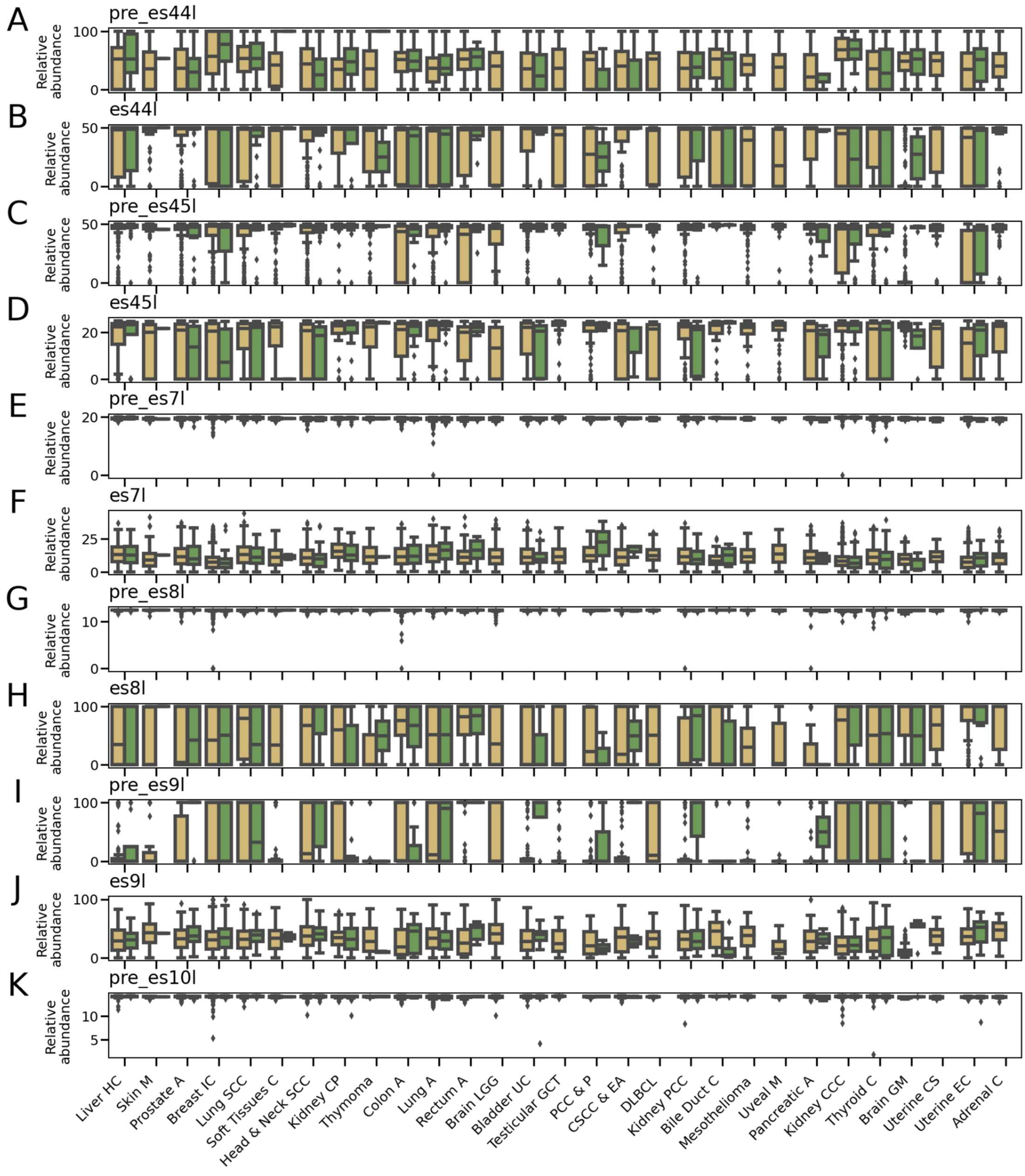


Figure S28. Cancer-specific rRNA variants relative-abundances (3 of 6 figures), related to Figure 6. (A-K) Scatter plot of different regional rRNA variants, relative abundances for TCGA cancer biopsy samples (cancer and control samples are in yellow and green boxes respectively). The top most

abundant rRNA regional variant is presented per ES/non-ES region across tissues (**Table S23** for region to regional variant conversion and the P-value for comparing case/control). The x-axis is the same for all panels and is displayed in (K). X-axis cancer type full name for the abbreviations are listed at the bottom of **Figure S26**.

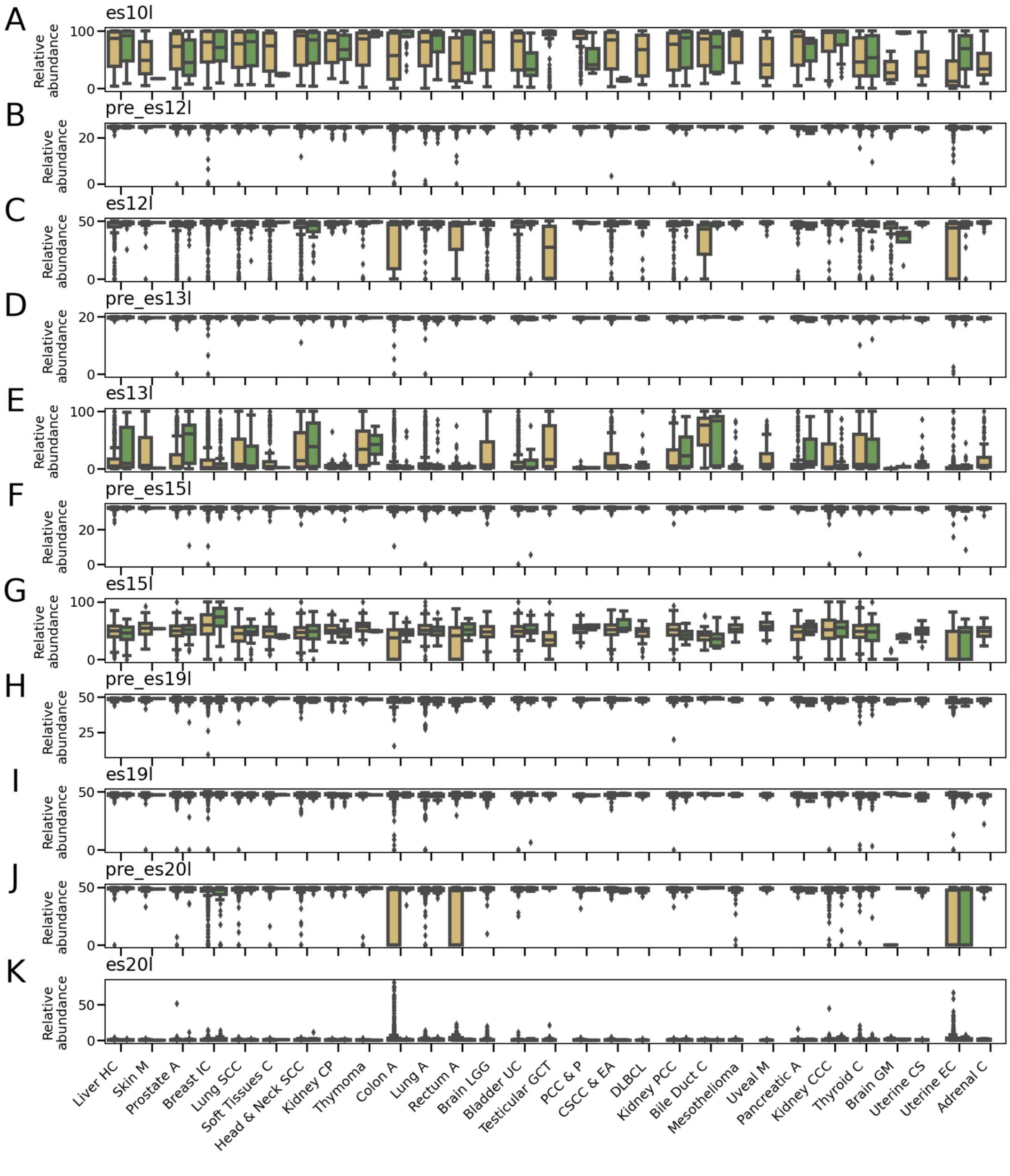


Figure S29. Cancer-specific rRNA variants relative-abundances (4 of 6 figures), related to Figure 6. (A-K) Scatter plot of different regional rRNA variants, relative abundances for TCGA cancer biopsy samples (cancer and control samples are in yellow and green boxes respectively). The top most

abundant rRNA regional variant is presented per ES/non-ES region across tissues (**Table S23** for region to regional variant conversion and the P-value for comparing case/control). The x-axis is the same for all panels and is displayed in (K). X-axis cancer type full name for the abbreviations are listed at the bottom of **Figure S26**.

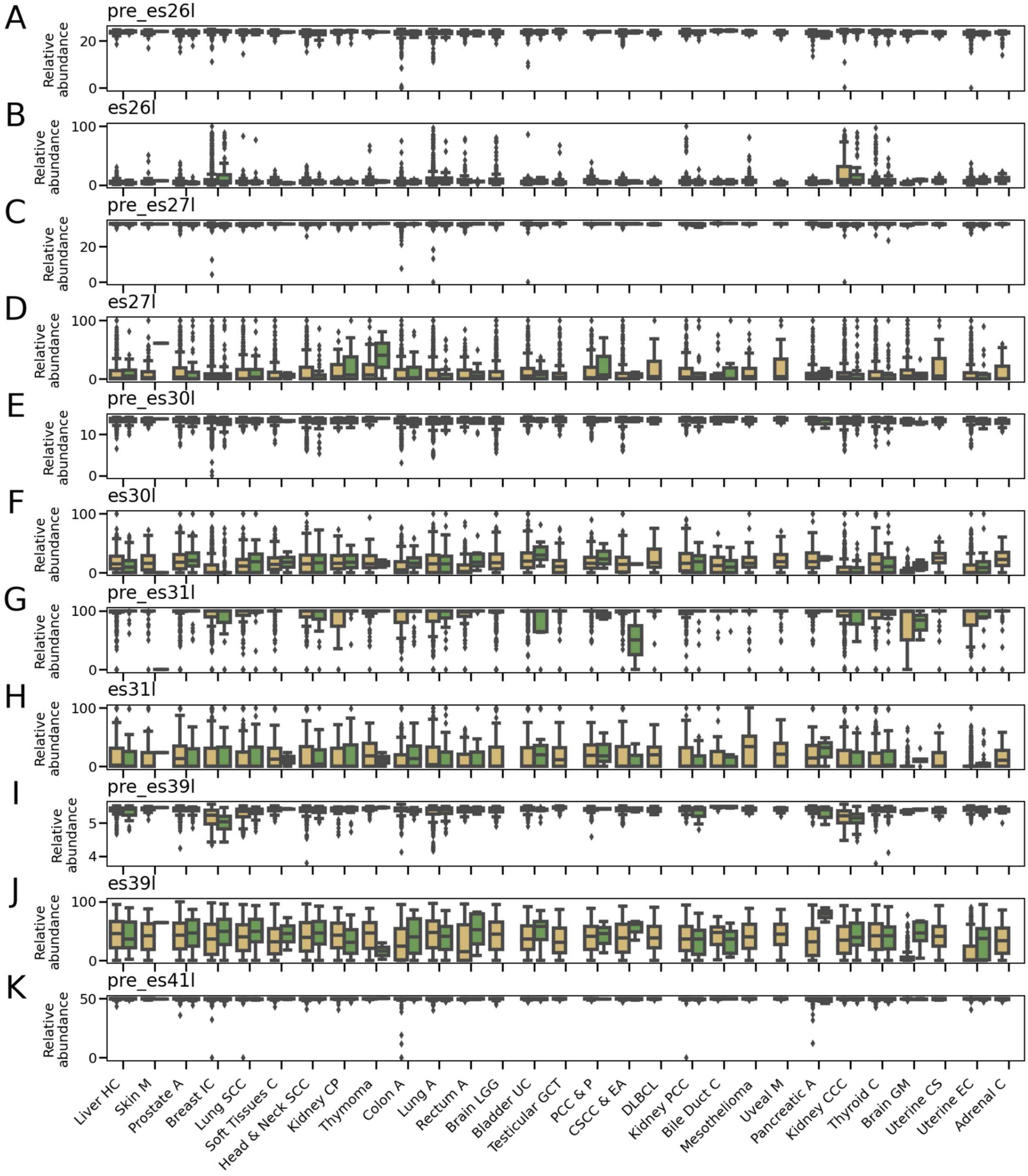


Figure S30. Cancer-specific rRNA variants relative-abundances (5 of 6 figures), related to Figure 6. (A-K) Scatter plot of different regional rRNA variants, relative abundances for TCGA cancer biopsy samples (cancer and control samples are in yellow and green boxes respectively). The top most

abundant rRNA regional variant is presented per ES/non-ES region across tissues (**Table S23** for region to regional variant conversion and the P-value for comparing case/control). The x-axis is the same for all panels and is displayed in (K). X-axis cancer type full name for the abbreviations are listed at the bottom of **Figure S26**.

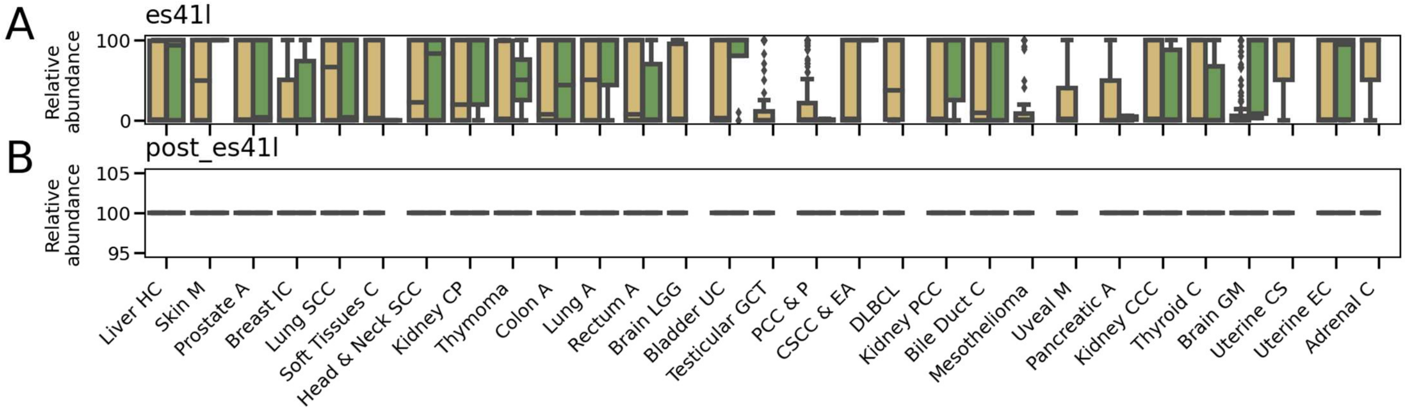


Figure S31. Cancer-specific rRNA variants relative-abundances (6 of 6 figures), related to Figure 6. (A-B) Scatter plot of different regional rRNA variants, relative abundances for TCGA cancer biopsy samples (cancer and control samples are in yellow and green boxes respectively). The top most abundant rRNA regional variant is presented per ES/non-ES region across tissues (**Table S23** for region to regional variant conversion and the P-value for comparing case/control). The x-axis is the same for all panels and is displayed in (B). X-axis cancer type full name for the abbreviations are listed at the bottom of **Figure S26**.