Figure S1.
Probe signal (y-axis) as a function of distance to enzyme cut PmeI (x-axis). The green lines indicate +/- 600 bases from the enzyme cut.

Figure S2.
These figures show the results of bisulfite validation for 3 loci.
A) a gene, two de-methylated probes in the same experiment.
B) AluSq element, microarray methylation profiling performed twice, each time demonstrating the methylation of the region.
C) AluY element, methylation profiling performed 5 times, each time demonstrating the de-methylation of the region.

Each figure is composed of a picture generated using UCSC genome browser to plot the location of the probe in the genomic context. Below, a textual summary that shows the location of the probe (**o**), CpG island (*****) repetitive elements (**T**-LTR , **A**-Alu, **1**- Line1 **2**- Line2, **=**-other, **S**-Sine). Vertical bars indicate the location of an enzyme recognition site for AciI, HHaI and McrBC enzymes. The region in the text is analogous to the one shown above in the UCSC genome browser graphic. The region surrounding the probe (highlighted in YELLOW on top of the McrBC endonuclease cleavage track) has been bisulfite treated, amplified and sequenced. The outcome of sequencing is shown in the bottom portion of the figure. The sequenced reads were aligned to the genomic sequence using clustalW. Subsequently the alignments were collapsed to only show locations of Cs in CpGs. An unmethylated CpG is indicated as "**O**" and methylated as "**@**".

Figure S3A
Ordering within plots. Per-experiment average methylation levels of the most informative subset of L1P and the least informative probes near DNA transposons and AluSq regions. Experiments are not ordered. The faint blue line indicates the average values of significant L1P probes in normal, non-tumor adjacent, tumor and sperm experiments (from top to bottom).

Figure S3B
Ordering within plots. Per-experiment average methylation levels of the most informative subset of L1P and the least informative probes near DNA transposons and AluSq regions. Experiments ordered within their groups based on L1P - mean(AluSq + DNA) probe values. The faint blue line indicates the average values of significant L1P probes in normal, non-tumor adjacent, tumor and sperm experiments (from top to bottom)

Figure S4A.
Example of a per-experiment plot showing average methylation levels of 4 categories of genomic compartments. See Supplementary Section 3.

Figure S4B.

Example of per-category plot showing distributions of average methylation levels of 4 categories of genomic compartments. See Supplementary Section 3

Figure S5.
Normal, Non-Tumor Adjacent, Tumor, Sperm variance in methylation levels ordered by repetitive element.

Figure S6.
Normal, Non-Tumor Adjacent,, Tumor, Sperm variance in methylation levels ordered by repetitive element, MaLR.

Figure S7.
Normal, Non-Tumor Adjacent,, Tumor, Sperm variance in methylation levels ordered by repetitive element, Alu.

Figure S8.
Normal, Non-Tumor Adjacent, Tumor, Sperm variance in methylation levels ordered by repetitive element, ERV.

Figure S9.
Normal, Non-Tumor Adjacent, Tumor, Sperm variance in methylation levels ordered by repetitive element, SVA.

Figure S10.
Normal, Non-Tumor Adjacent, Tumor, Sperm variance in methylation levels ordered by repetitive element, L1P.

Figure S11.
Composite figure explaining the repeat CpG content using alignment plots and bin plots. The top portion, enveloped using the gray lines describes the specific repetitive element (alignment).
A) All sequences annotated by repeat masker as belonging to AluYb were filtered to output only those for which there is a probe on the microarray. From all these sequences, the ones that are shorter than the median of all sequences were further excluded to focus the subsequent analysis on longer (possibly full length) elements. The alignment of 60 of those sequences selected at random was then created using clustalw2 (1.0.11) program and standard parameters. The output of clustalw2 was subsequently parsed using a series of custom python and R scripts to create the figure. The nucleotides are colored to highlight the GC rich regions using the following scheme: G - yellow, C - green, A - light gray, T - pink
B)
Summary of the alignment in A), histogram indicates relative conservation of CpG at a given location. Per alignment locus, the number of CpGs was counted and normalized by the number of sequences aligned (i.e. 60)
C)

Summary of the alignment in A), (inverted) stacked histogram that summarizes a distribution of nucleotides at a given location. The bars are stacked from the most abundant nucleotide on the top, to the least abundant on the bottom. The colors are the same as in the part A). The most abundant nucleotide is marked in the middle of its bar.

The bottom portion characterizes the genomic context of the repetitive element family (bin plot).
The first 4 sub-plots, characterize all repetitive elements of a particular class in the human genome. The bins of plot D summarize the distribution of CpG counts in all sequences of all repetitive elements from a given lineage (the central bin marked in red) and 1,500 bases up~ and downstream from the repeat in 100 base increments per bin. The distribution of CpG in the repeat bin and external bins are presented in the form of a standard box and whisker plot, where the thick line inside the box indicates a median, the box is drawn around 25^{th} and 75^{th} percentiles, and the outliers are indicated as dots.

Plots E and F keep the binning structure of the sequence as in plot D, and show the average number of potential enzyme cuts among all the sequences per bin normalized to the size of the bin. Gray lines indicate the standard deviation.

Plot G is pertinent to the red bin of Plot D, it shows the distribution of sizes of all genomic repeats of a given family which were included in the central bin of plot D.

The plot H-K are analogous to D-G, with the only difference being that now the subset of repeats with a probe on the microarray are included in this analysis.

The plot L shows the repeat associated probe's location with respect to the repetitive element.

The supplementary figures S12 through S21 contains the bin plots for all families and subfamilies of probed repetitive elements discussed in the main text of the article. 54 alignment plots for these families are available in a zipped archive in the Supplementary Materials.

Figure S12.
CpG content summary for lineages of probed MaLR repetitive elements and 1,500 nucleotides up and down stream from the repeats. The flanking sequence is displayed in bins of 100 bases long.

Figure S13.
CpG content summary for lineages of probed Alu repetitive elements and 1,500 nucleotides up and down stream from the repeats. The flanking sequence is displayed in bins of 100 bases long.

Figure S14.

CpG content summary for lineages of probed ERV repetitive elements and 1,500 nucleotides up and down stream from the repeats. The flanking sequence is displayed in bins of 100 bases long.

Figure S15.
CpG content summary for lineages of probed SVA repetitive elements and 1,500 nucleotides up and down stream from the repeats. The flanking sequence is displayed in bins of 100 bases long.

Figure S16.
CpG content summary for lineages of probed L1P repetitive elements and 1,500 nucleotides up and down stream from the repeats. The flanking sequence is displayed in bins of 100 bases long.

Figure S17.
CpG content side-by-side comparison of an older versus younger lineage within repetitive family, MaLR: MLT1C - MST1A

Figure S18.
CpG content side-by-side comparison of an older versus younger lineage within repetitive family, SVA: SVA_F - SVA_B

Figure S19.
CpG content side-by-side comparison of an older versus younger lineage within repetitive family, L1P: L1PA17 - L1PA4

Figure S20.
CpG content side-by-side comparison of an older versus younger lineage within repetitive family, L1P: L1PA3 - L1HS

Figure S21.
CpG content side-by-side comparison of an older versus younger lineage within repetitive family, Alu: AluYd8 - AluYb9

Figure S22.
ASCIIMAPs showing the location of probes interrogating MaLR elements.
The text provides a summary of a region using ASCII characters (generated using a tool ASCIIMap). The ASCIIMap tracks show the location of the probe (**o**). The location of a CpG island is marked underneath (*****) as are the locations of repetitive elements in the area (**1**-Line1, **2**-Line2, **T**-LTR(MaLR), **S**- SINE, **=**-other, **A**-Alu, etc.). The vertical bars (**I**) indicate the presence of an enzyme recognition site for AciI, HHaI and McrBC enzymes respectively. The resolution of 1 character is about 15 nucleotides. The regions shown are approximately 1kb in length. The genomic coordinates of each regions are given using HG16 build of the human genome. Note that in some cases a probe appears to map inside a region designated as a repetitive element. The probe design algorithm always ensures that the sequence is unique before designing a probe.

The apparent paradox that a repetitive element may have parts that are unique sequences can be explained by considering the age of the repetitive elements for which the probe is designed. For example, an element of the family MLT1C, 85 MYO: over a span of millions of years since it appeared in its original form in the genome, its sequence have deteriorated from its consensus so much that although the element can still be classified as MLT1C now (based on the overall structure and certain sequence patterns), its sequences acquired enough random mutations that the probe algorithm can recognize certain parts within this MLT1C as unique in the genome. This can be best illustrated in the Supplementary Figures S12 through S16, in the “repeat associated probes per bin” panels. For repetitive element families that are younger, i.e. the elements that haven’t had evolutionary time to acquire mutations differentiating them from their respective consensus, the probe designer most likely designs the probe within the 100 bases flanking region of the repetitive element. Conversely, for the older repetitive elements (20, 30, 40+ MYO), the probe designer is able to find regions that have uniquely diverged from the global consensus of the repeat family. The archive of alignments of repetitive elements within their families included in the Supplementary Materials further illustrates how similar or divergent the sequences are within a lineage of repetitive elements.

Figure S1.

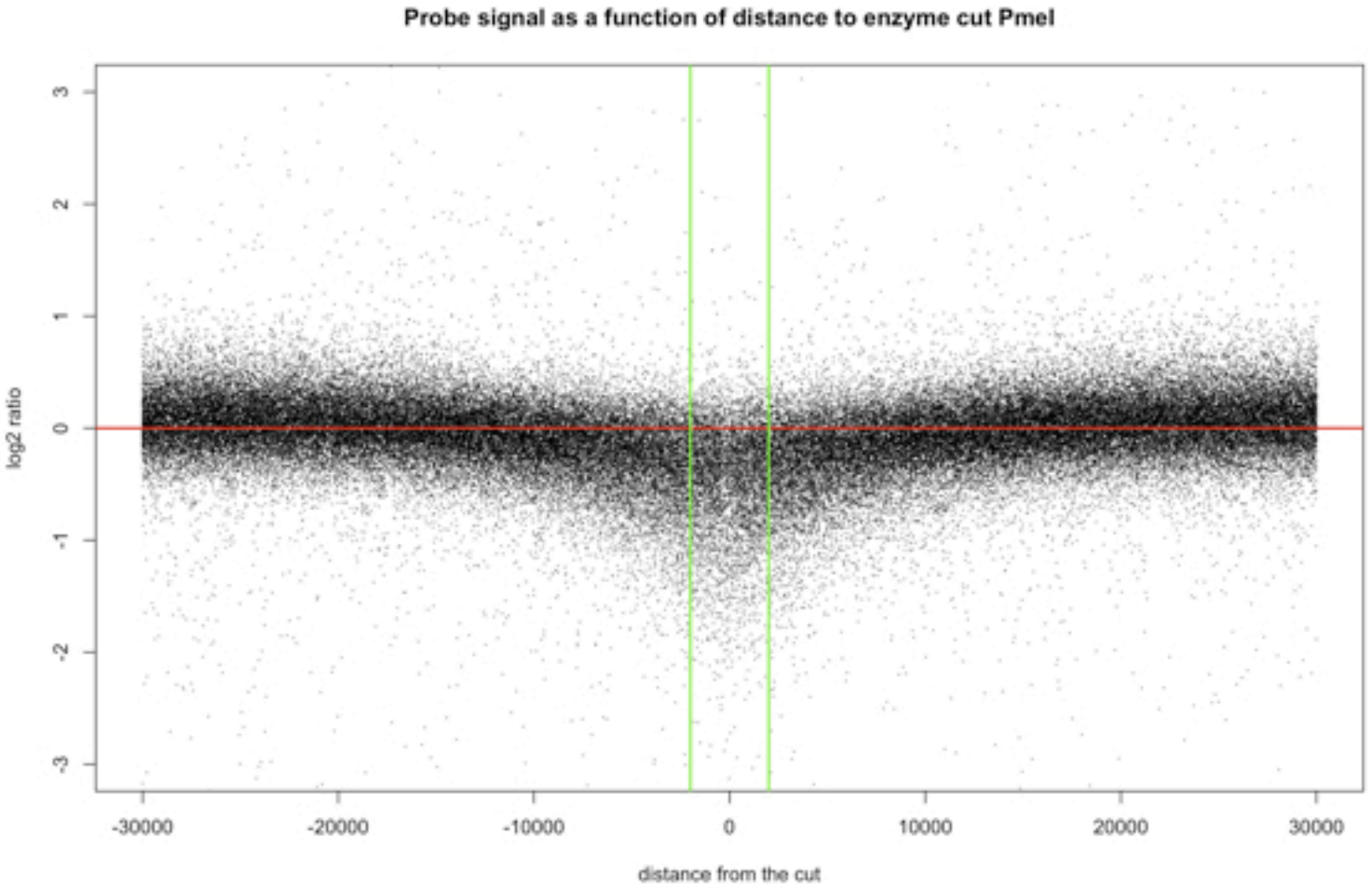
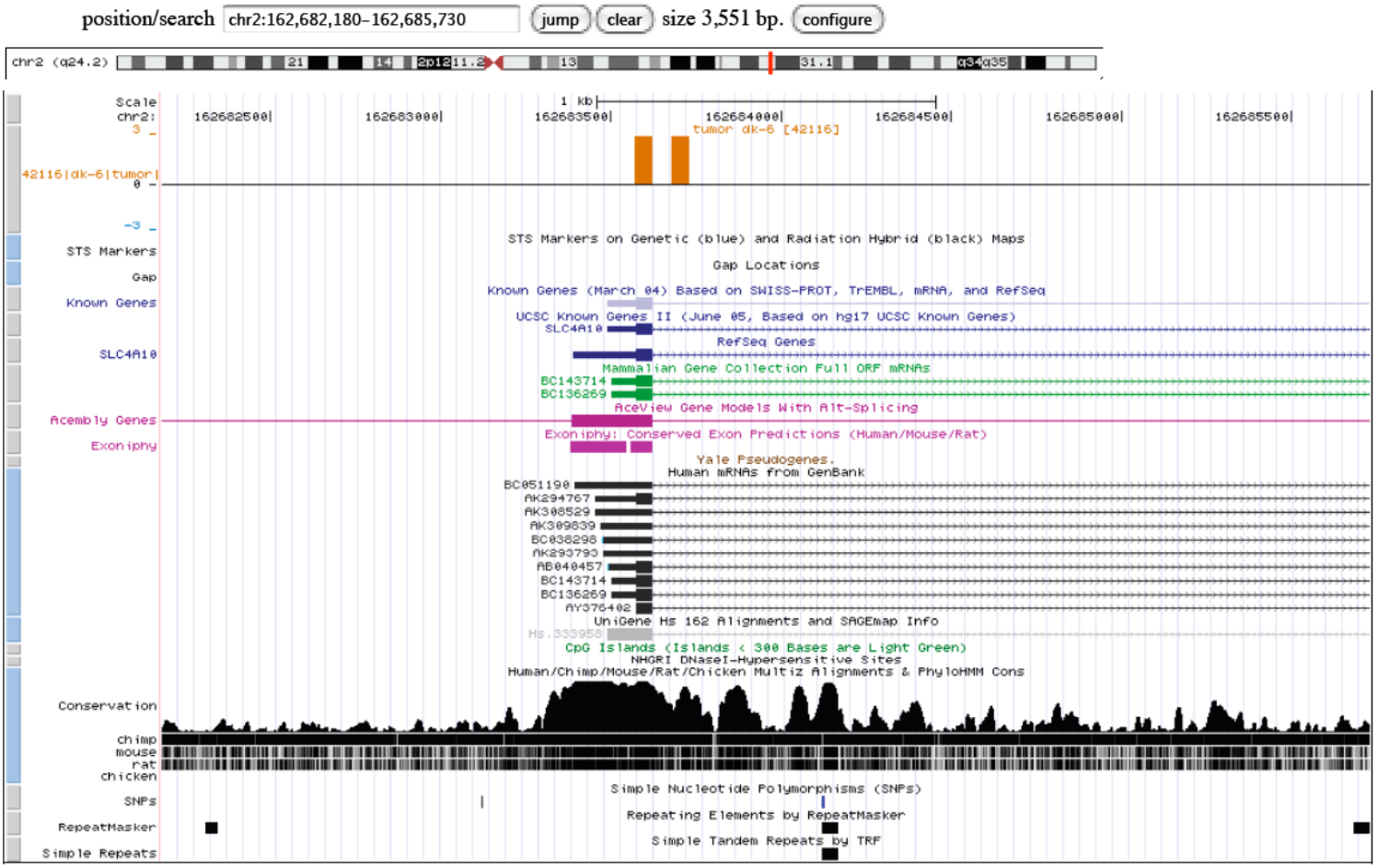

Probe signal as a function of distance to enzyme cut PmeI
log2 ratio
3
2
1
0
-1
-2
-3
-30000
-20000
-10000
0
10000
20000
30000
distance from the cut

Figure S2A.

H. sapiens solute carrier family 4, sodium bicarbonate transporter, member 10 mRNA.



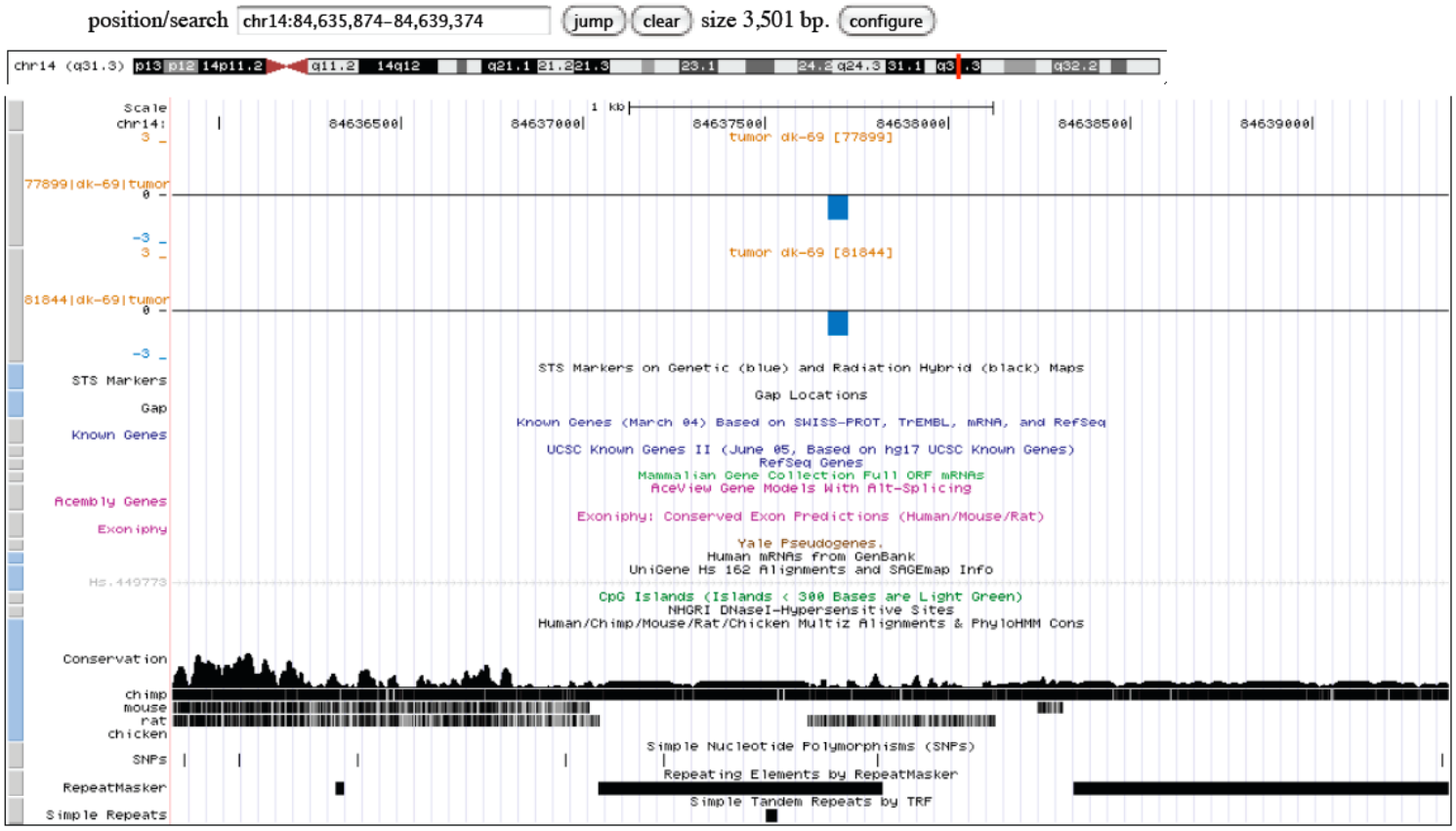

```
[chr2:162682180-162685730] NM_002058, (3[kb] 1c~50.0[b])
A >>> a    NM_022058
INFO  ---------------------------A-------------------------------------------a     INFO
GENE  ---------------------------->>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>     GENE
PROBE ---------------------------o--o-------------------------------------------     PROBE
CPG   ----------------------------*****-------------------------------------      CPG
REP   ------------------------------------------------------------------------S     REP
ACI   -------------------|---------|||-----------------||--|-------------------     ACI
HhaI  ------------------------------------------------|----------------------      HhaI
McrBC -----------------|-||------|||||||||--|------|--||||-||---|-|------------     McrBC1
```

```
BISULFITE SEQUENCING RESULTS FROM 7 CLONES

NM_002058 <18>                C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-
c>t                           T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-
HhaI+AciI                                 |   |   |
McrBC                          |        |   | | | | | | | | | | |
NM_002058-4-D8-M13R [0.72]      -O-O-O-@-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-
NM_002058-2-B8-M13R [0.71]      -O-O-O-O-O-O-O-O-O-O-@-O-O-O-O-O-O-O-O-
NM_002058-3-C8-M13R [0.71]      -O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-@-@-O-
NM_002058-5-E8-M13R [0.71]      -O-O-O-@-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-
NM_002058-6-F8-M13R [0.70]      -O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-O-
NM_002058-1-A8-M13R [0.70]      -@-O-@-@-@-O-O-O-O-O-@-O-O-O-O-O-O-O-
NM_002058-7-G8-M13R [0.66]      -O-O-O-@-O-O-O-O-O-O-O-O-O-O-O-O-O- -
```

Figure S2B.

AluSq element within MLT1C flanks



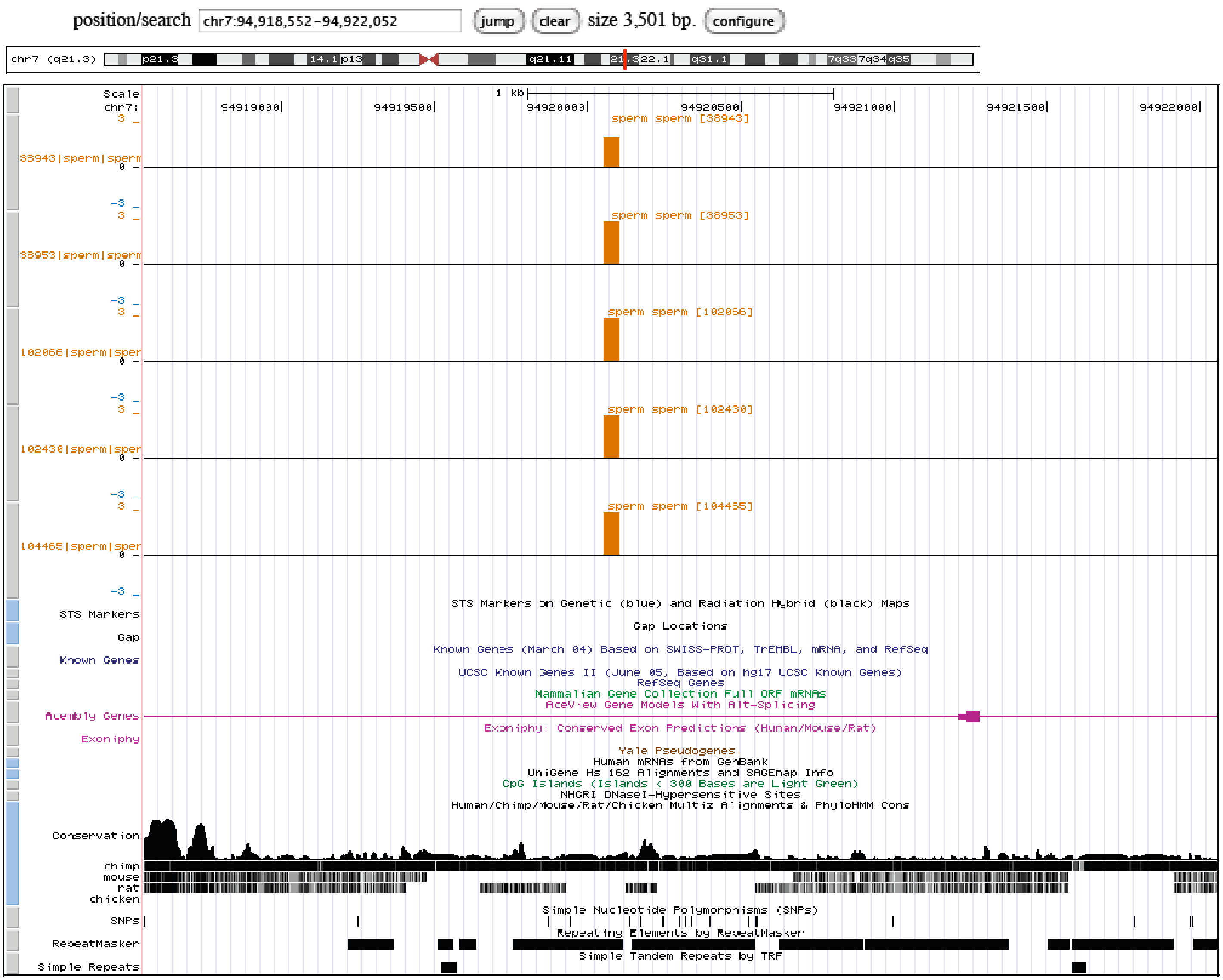

```
[chr14:84635874-84639374] flanks, (3[kb] 1c~50.0[b])
INFO  ------------------------------------------------------------------------      INFO
GENE  ------------------------------------------------------------------------      GENE
PROBE ----------------------------------o-------------------------------------      PROBE
CPG   ----------------------------******-------------------------------------      CPG
REP   ------------------------TTTAAAAAATTTTTT-----------11111111111111111111      REP
ACI   -----------|----|----------|--||-------------------------------------      ACI
McrBC1-----|----|||-----|---||||---|||||--|||-|---------------|-----------||-      McrBC1


BISULFITE SEQUENCING RESULTS FROM 8 CLONES


probe64983 <314>              -C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-
c>t                           -T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-
HhaI+AciI                        |       |           |           |   |
McrBC                             | |   |   | |   | | | | | | | | | | |   |   | | | | |   |
probe64983_H01.R-AS [0.11]    - - - - - - - -O-O-@-@-@-@-O-@-O-@-@-@-@-@-@-@-@-@- -@- - - - - -
probe64983_1A.R-AS [0.11]     - - - - - - - - - -@-@-@-@-@-@-@-@-@-@-@-@-@-@-@-@- -@- - - - - -
probe64983_E01.R-AS [0.11]    - - - - - - - - - -O-@-@-O-@-@-O-O-O-O-@-@-@-O-@-@- -@- - - - - -
probe64983_G01.R-AS [0.11]    - - - - - - - - - - -@-@-@-O-O-@-@-@-@-@-@-@-@-@-@- -@- - - - - -
probe64983_B01.R-AS [0.10]    - - - - - - - - - -@-@-@-@-@-@-@-@-@-@-O-@-@-@-@-@- -@- - - - - -
probe64983_D01.F-AS [0.10]    - - - - - - - - - - -O-@-@-@-@-@-O-@-@-@-@-@-@-@-@-O- - - - - - - -
probe64983_C01.R-AS [0.10]    - - - - - - - - - - - -@-@-O-O-@-@-@-@-@-@-@-@-@-@- -@- - - - - -
probe64983_D01.R-AS [0.09]    - - - - - - - - - - -O-@-@-@-@-@-O-@-@-@-@-@-@-@-@- - - - -@- - - -
```

Figure S2C.

AluY element with L2 and MER5A1 flanking elements




```
[chr7:94918552-94922052] sequences, (3[kb] 1c~50.0[b])
INFO   ----------------------------------------------------------------------  INFO
GENE   ----------------------------------------------------------------------  GENE
PROBE  ------------------------------o---------------------------------------  PROBE
CPG    --------------------------------*********-----------------------------  CPG
REP    -------------===----==---2222222==AAAAAA--22222-TTTTTTTTT---2-AAAAAA-==  REP
ACI    ----------------|-----------------||-||----------------------|-------  ACI
HhaI   -----------|----------------------|-|-|-----------------------------  HhaI
McrBC1 --|--|--|--|----||---|---------||-||||||-|------|-------------|--||---       McrBC1


BISULFITE SEQUENCING RESULTS FROM 6 CLONES

probe302302 <321>    -C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-C-
c>t                  -T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-
HhaI+AciI               |             | |               |   | |       | | |               | |       | |             |
McrBC                 | | |   | | | | | | |   | | |   | |   | | |   | | |   | |   | | |   | | |   | |       | |   |
probe302302_F05.R-AS - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -O- - - - - - - -
probe302302_C05.R-AS - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -@- - - - - - - -
probe302302_G05.F-AS - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -O- - - - - - - -
probe302302_D05.F-SE - - - - - - - - - - - - - - -O-O-O-O-O-O-O-O-O- -O-O-O-O-O-O- - - - - - - - - - - - - - - - -
probe302302_F05.F-SE - - - - - - - - - - - - - - -O-O- - -O-O-O-O-O- -O-O-O- - - - - - - - - - - - - - - - - - - -
probe302302_C05.F-SE - - - - - - - - - - - - - - -O-O-O- -O-O-O-O- - - - - - - - - - - - - - - - - - - - - - - - -
```

Figure S3A.

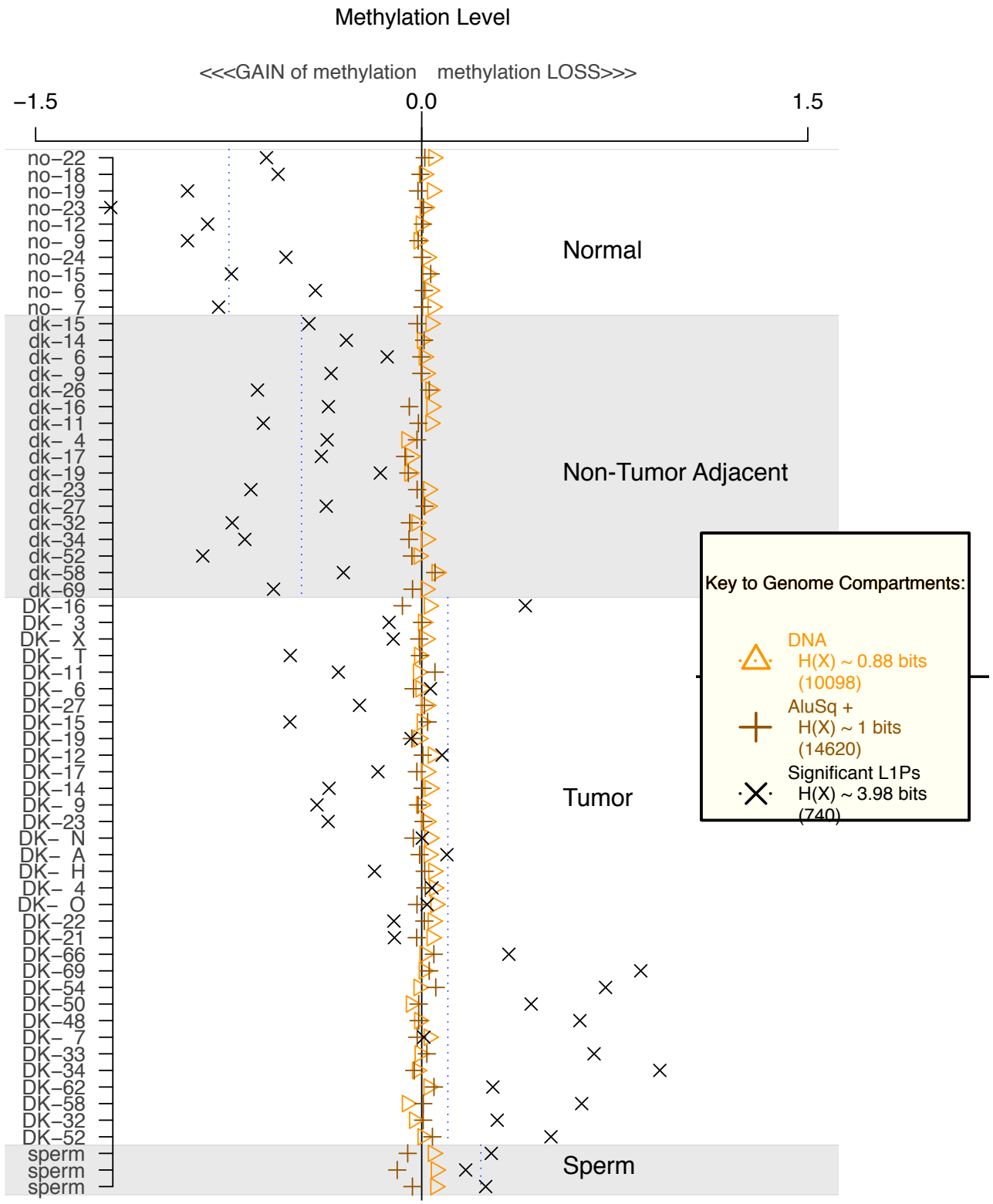

Methylation Level
<<<GAIN of methylation methylation LOSS>>>
−1.5
0.0
1.5
no−22
no−18
no−19
no−23
no−12
no− 9
no−24
no−15
no− 6
no− 7
dk−15
dk−14
dk− 6
dk− 9
dk−26
dk−16
dk−11
dk− 4
dk−17
dk−19
dk−23
dk−27
dk−32
dk−34
dk−52
dk−58
dk−69
DK−16
DK− 3
DK− X
DK− T
DK−11
DK− 6
DK−27
DK−15
DK−19
DK−12
DK−17
DK−14
DK− 9
DK−23
DK− N
DK− A
DK− H
DK− 4
DK− O
DK−22
DK−21
DK−66
DK−69
DK−54
DK−50
DK−48
DK− 7
DK−33
DK−34
DK−62
DK−58
DK−32
DK−52
sperm
sperm
sperm
Normal
Non-Tumor Adjacent
Tumor
Sperm
Key to Genome Compartments:
DNA
H(X) ~ 0.88 bits
(10098)
AluSq +
H(X) ~ 1 bits
(14620)
Significant L1Ps
H(X) ~ 3.98 bits
(740)

Figure S3B.

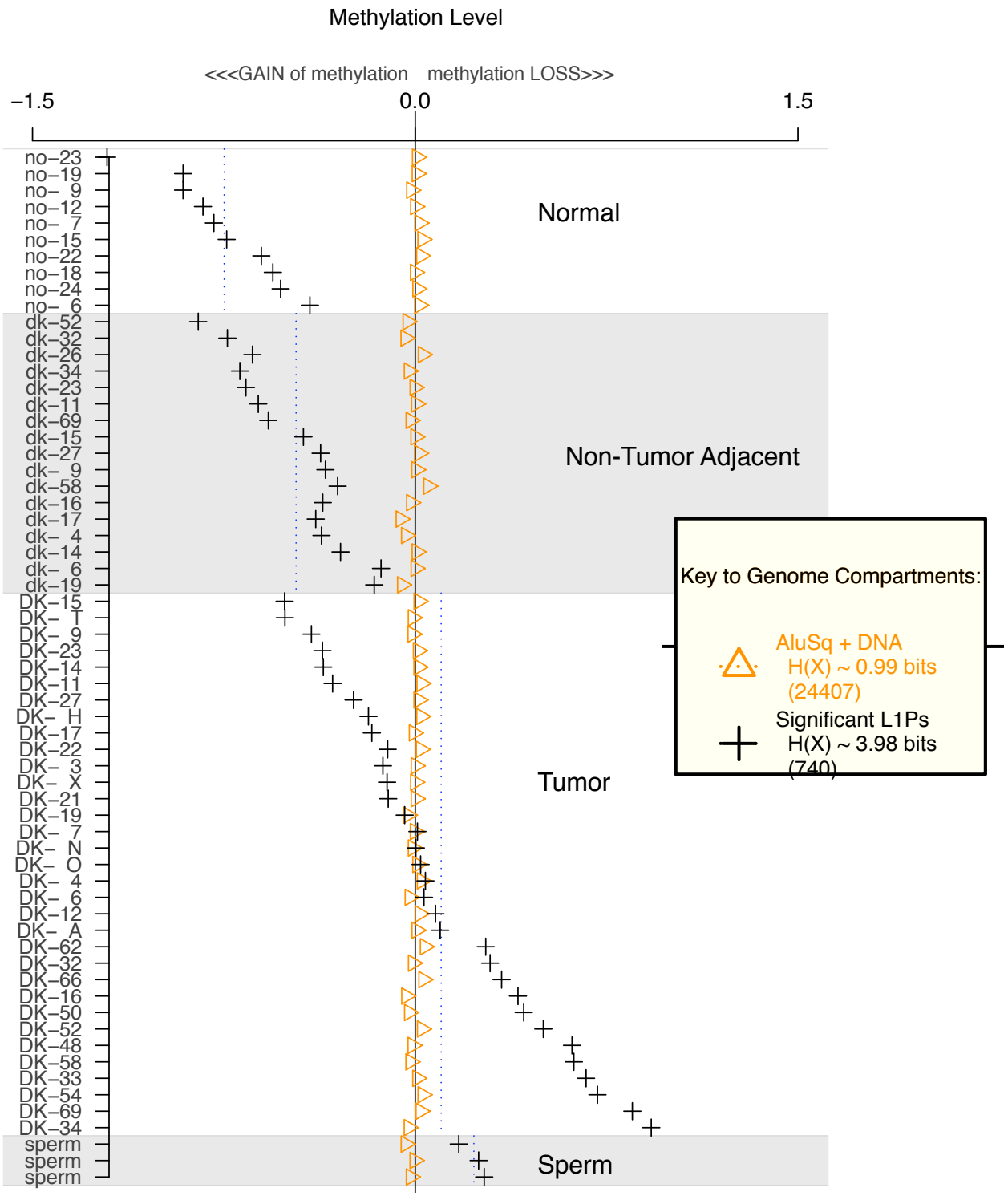

Methylation Level
<<<GAIN of methylation methylation LOSS>>>
−1.5
0.0
1.5
no−23
no−19
no− 9
no−12
no− 7
no−15
no−22
no−18
no−24
no− 6
dk−52
dk−32
dk−26
dk−34
dk−23
dk−11
dk−69
dk−15
dk−27
dk− 9
dk−58
dk−16
dk−17
dk− 4
dk−14
dk− 6
dk−19
DK−15
DK− T
DK− 9
DK−23
DK−14
DK−11
DK−27
DK− H
DK−17
DK−22
DK− 3
DK− X
DK−21
DK−19
DK− 7
DK− N
DK− O
DK− 4
DK− 6
DK−12
DK− A
DK−62
DK−32
DK−66
DK−16
DK−50
DK−52
DK−48
DK−58
DK−33
DK−54
DK−69
DK−34
sperm
sperm
sperm
Normal
Non-Tumor Adjacent
Tumor
Sperm
Key to Genome Compartments:
AluSq + DNA
H(X) ~ 0.99 bits
(24407)
Significant L1Ps
H(X) ~ 3.98 bits
(740)

Figure S4A.

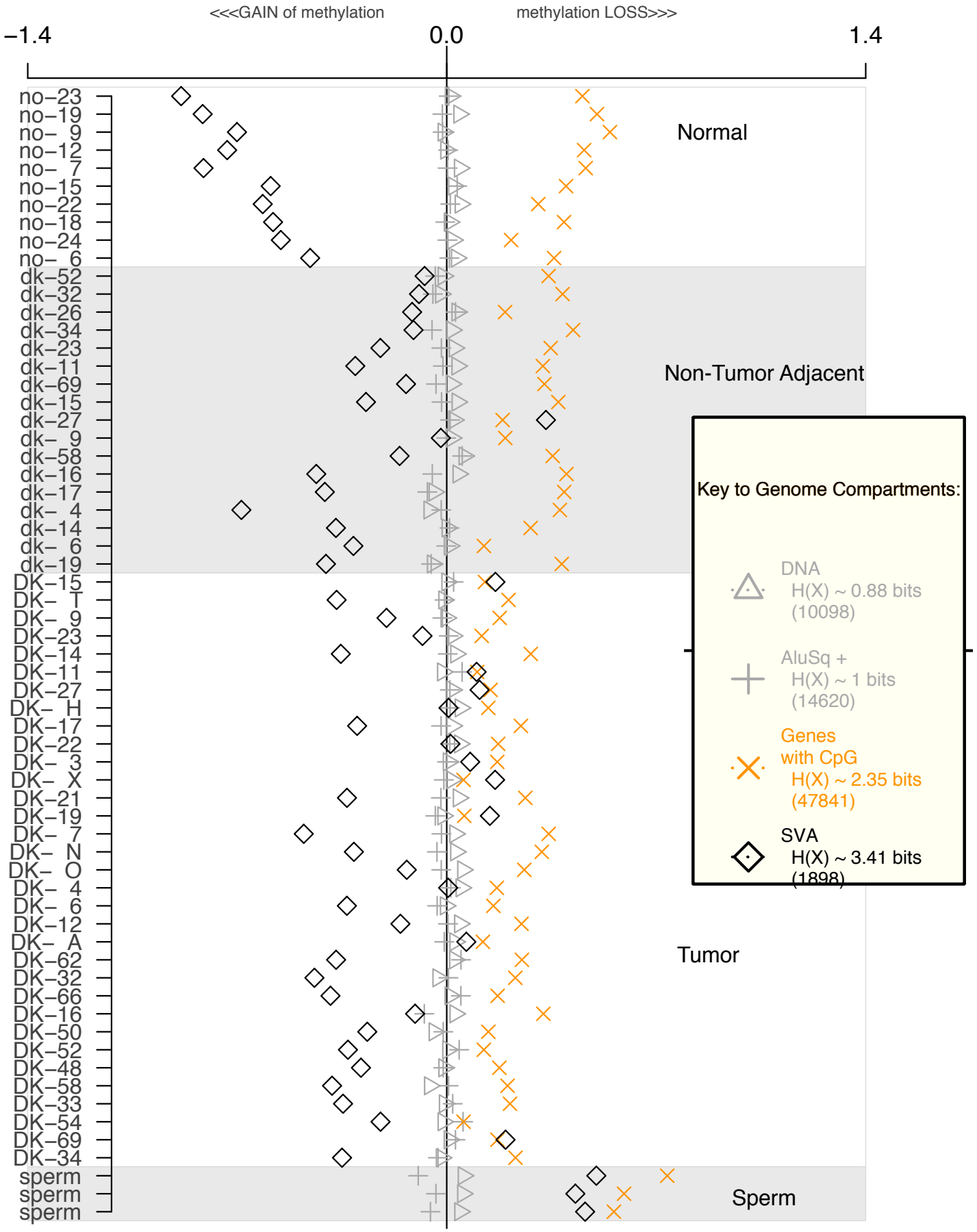

<<<GAIN of methylation
methylation LOSS>>>
−1.4
0.0
1.4
no−23
no−19
no− 9
no−12
no− 7
no−15
no−22
no−18
no−24
no− 6
dk−52
dk−32
dk−26
dk−34
dk−23
dk−11
dk−69
dk−15
dk−27
dk− 9
dk−58
dk−16
dk−17
dk− 4
dk−14
dk− 6
dk−19
DK−15
DK− T
DK− 9
DK−23
DK−14
DK−11
DK−27
DK− H
DK−17
DK−22
DK− 3
DK− X
DK−21
DK−19
DK− 7
DK− N
DK− O
DK− 4
DK− 6
DK−12
DK− A
DK−62
DK−32
DK−66
DK−16
DK−50
DK−52
DK−48
DK−58
DK−33
DK−54
DK−69
DK−34
sperm
sperm
sperm
Normal
Non-Tumor Adjacent
Tumor
Sperm
Key to Genome Compartments:
DNA
H(X) ~ 0.88 bits
(10098)
AluSq +
H(X) ~ 1 bits
(14620)
Genes
with CpG
H(X) ~ 2.35 bits
(47841)
SVA
H(X) ~ 3.41 bits
(1898)

Figure S4B.

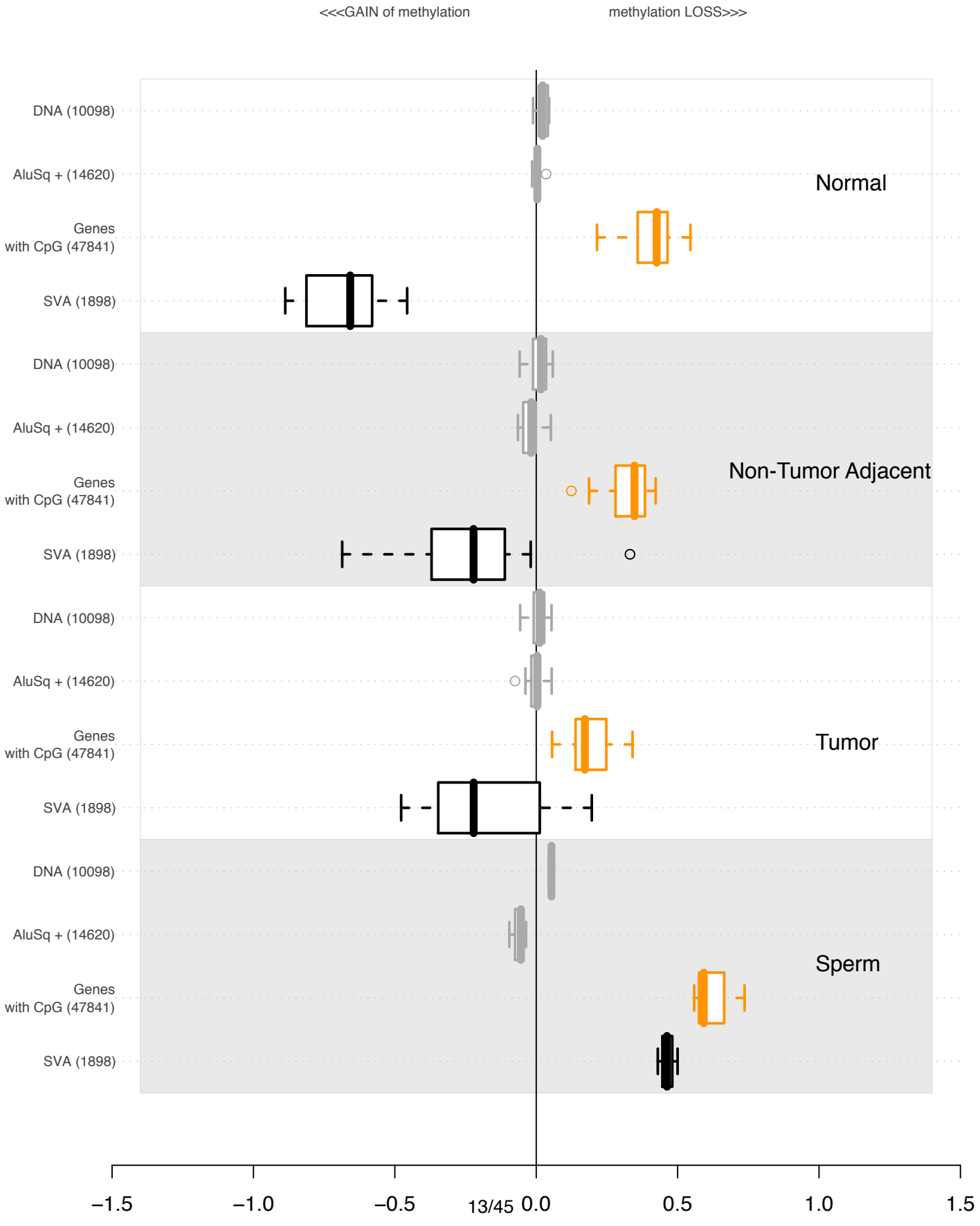

<<<GAIN of methylation
methylation LOSS>>>
DNA (10098)
AluSq + (14620)
Genes with CpG (47841)
SVA (1898)
Normal
Non-Tumor Adjacent
Tumor
Sperm
−1.5
−1.0
−0.5
13/45 0.0
0.5
1.0
1.5

Figure S5.

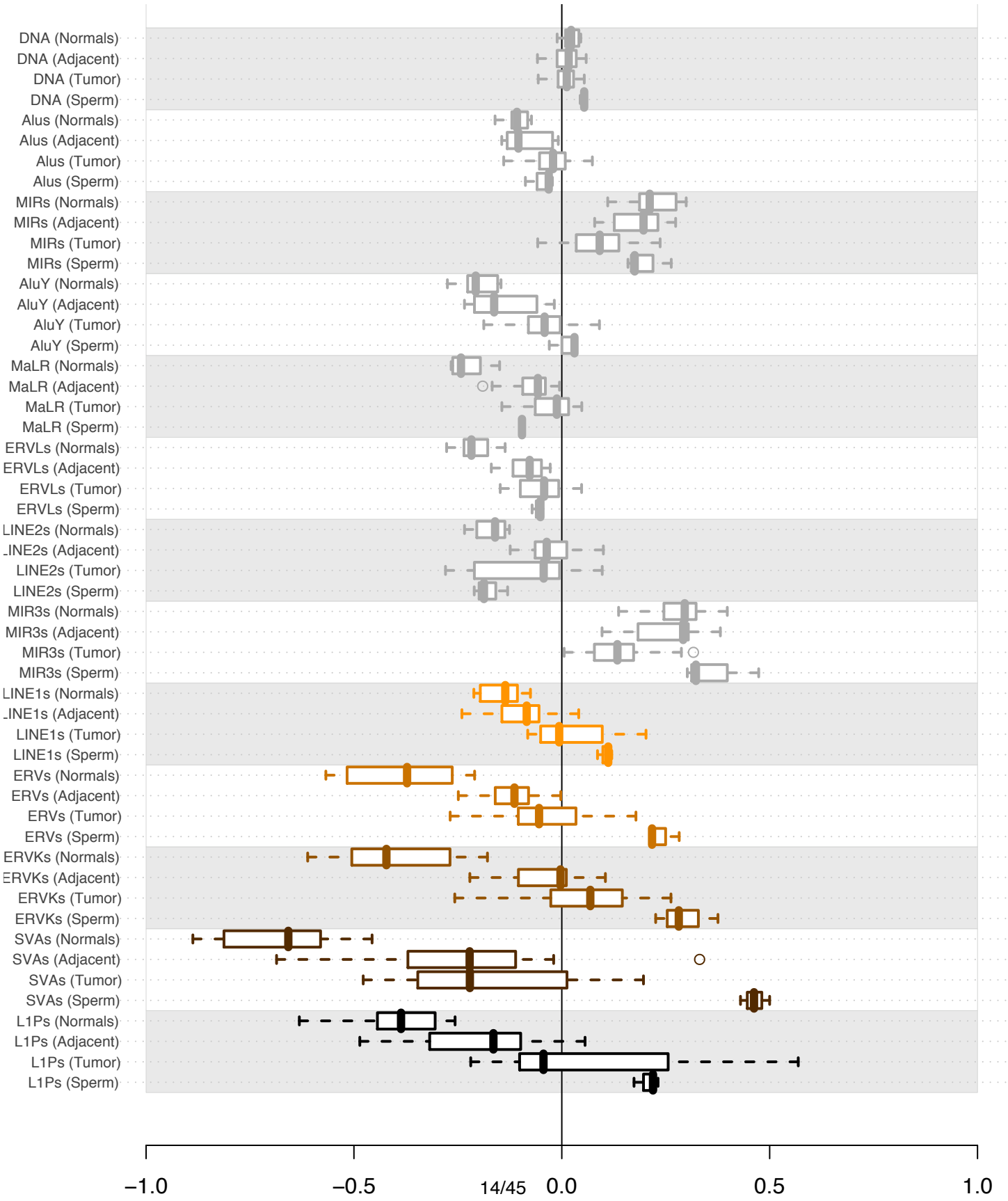

DNA (Normals)
DNA (Adjacent)
DNA (Tumor)
DNA (Sperm)
Alus (Normals)
Alus (Adjacent)
Alus (Tumor)
Alus (Sperm)
MIRs (Normals)
MIRs (Adjacent)
MIRs (Tumor)
MIRs (Sperm)
AluY (Normals)
AluY (Adjacent)
AluY (Tumor)
AluY (Sperm)
MaLR (Normals)
MaLR (Adjacent)
MaLR (Tumor)
MaLR (Sperm)
ERVLs (Normals)
ERVLs (Adjacent)
ERVLs (Tumor)
ERVLs (Sperm)
LINE2s (Normals)
LINE2s (Adjacent)
LINE2s (Tumor)
LINE2s (Sperm)
MIR3s (Normals)
MIR3s (Adjacent)
MIR3s (Tumor)
MIR3s (Sperm)
LINE1s (Normals)
LINE1s (Adjacent)
LINE1s (Tumor)
LINE1s (Sperm)
ERVs (Normals)
ERVs (Adjacent)
ERVs (Tumor)
ERVs (Sperm)
ERVKs (Normals)
ERVKs (Adjacent)
ERVKs (Tumor)
ERVKs (Sperm)
SVAs (Normals)
SVAs (Adjacent)
SVAs (Tumor)
SVAs (Sperm)
L1Ps (Normals)
L1Ps (Adjacent)
L1Ps (Tumor)
L1Ps (Sperm)
−1.0
−0.5
14/45
0.0
0.5
1.0

Figure S6.

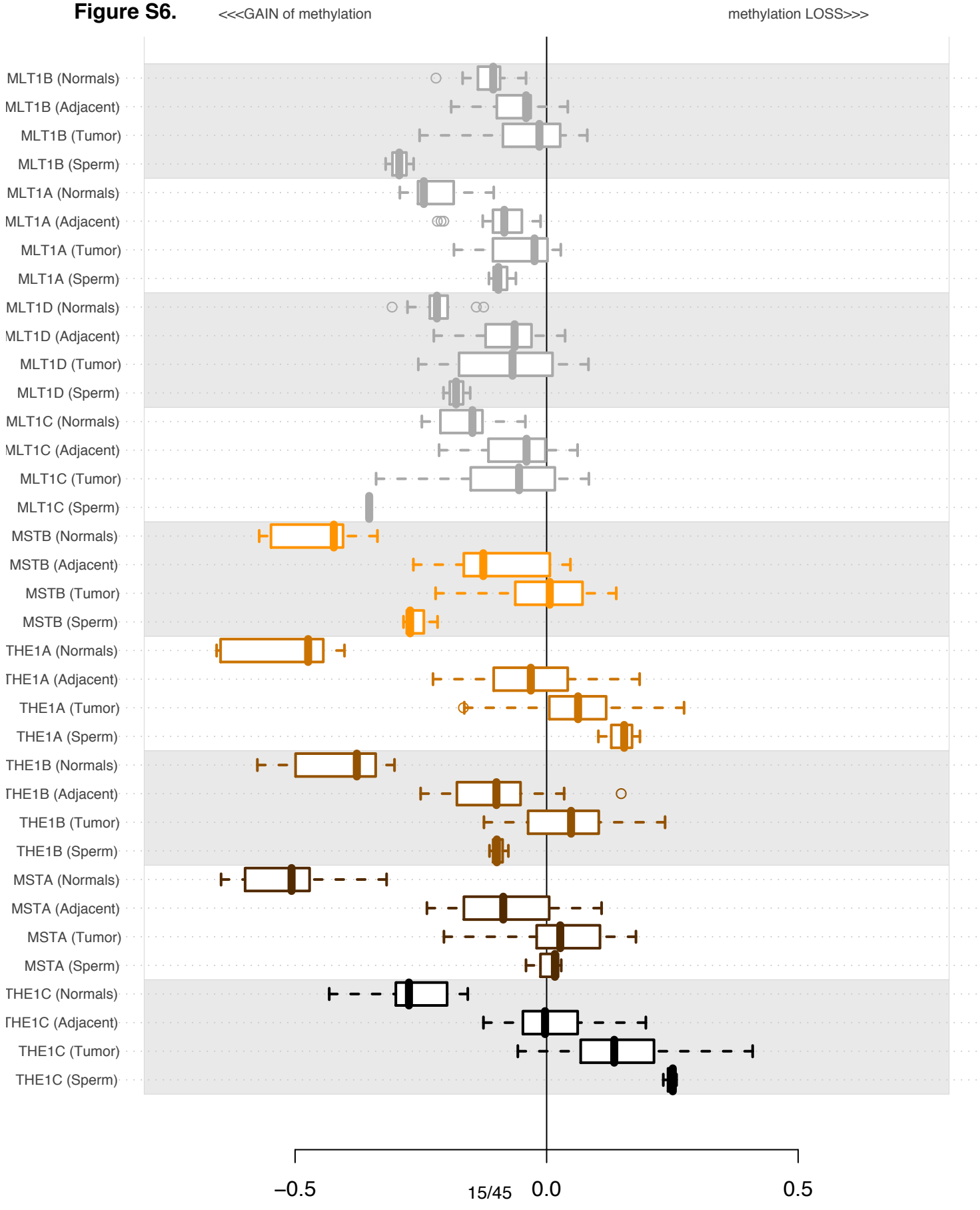

<<<GAIN of methylation
methylation LOSS>>>
MLT1B (Normals)
MLT1B (Adjacent)
MLT1B (Tumor)
MLT1B (Sperm)
MLT1A (Normals)
MLT1A (Adjacent)
MLT1A (Tumor)
MLT1A (Sperm)
MLT1D (Normals)
MLT1D (Adjacent)
MLT1D (Tumor)
MLT1D (Sperm)
MLT1C (Normals)
MLT1C (Adjacent)
MLT1C (Tumor)
MLT1C (Sperm)
MSTB (Normals)
MSTB (Adjacent)
MSTB (Tumor)
MSTB (Sperm)
THE1A (Normals)
THE1A (Adjacent)
THE1A (Tumor)
THE1A (Sperm)
THE1B (Normals)
THE1B (Adjacent)
THE1B (Tumor)
THE1B (Sperm)
MSTA (Normals)
MSTA (Adjacent)
MSTA (Tumor)
MSTA (Sperm)
THE1C (Normals)
THE1C (Adjacent)
THE1C (Tumor)
THE1C (Sperm)
−0.5
15/45
0.0
0.5

Figure S7.

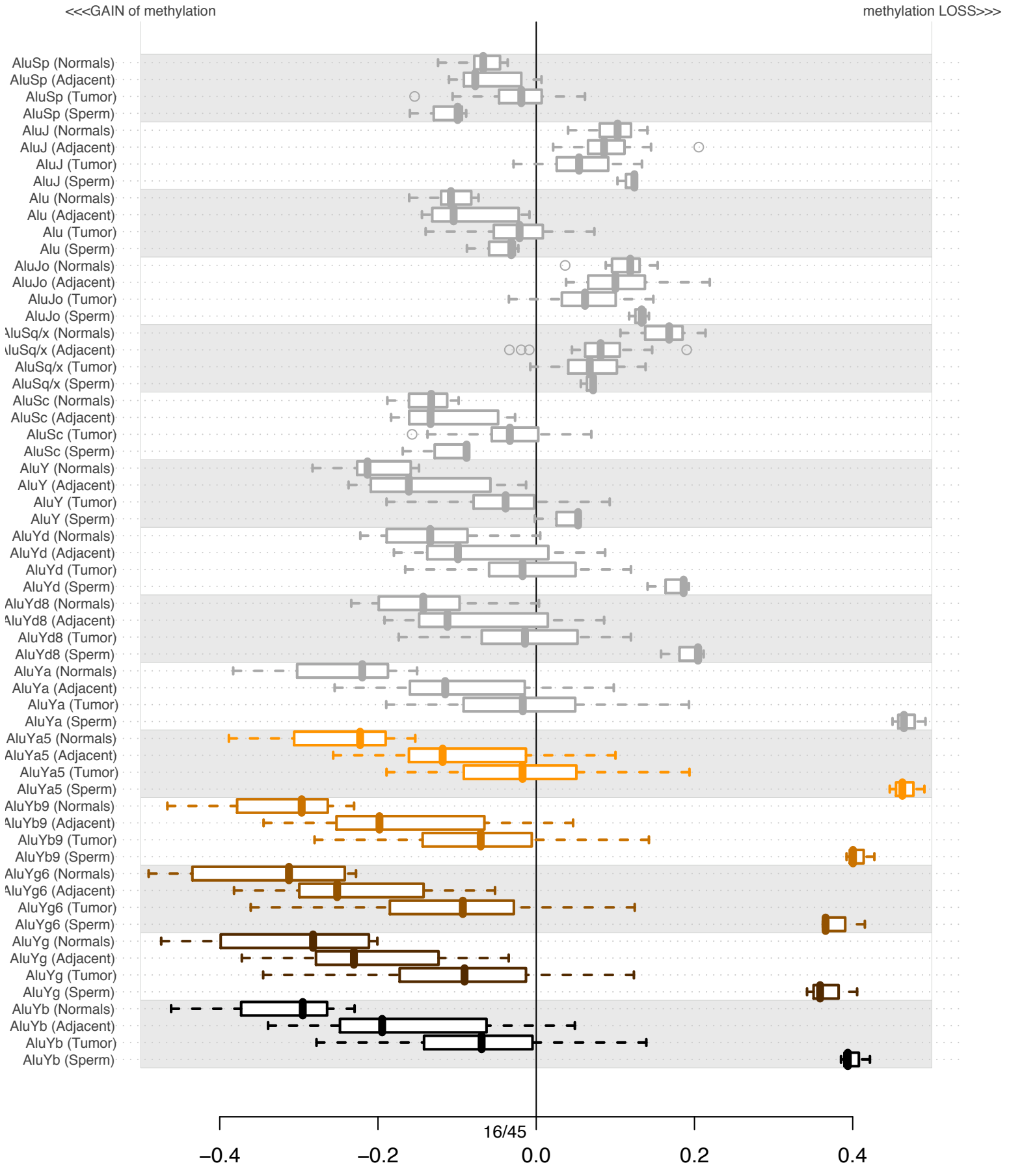

<<<GAIN of methylation
methylation LOSS>>>
AluSp (Normals)
AluSp (Adjacent)
AluSp (Tumor)
AluSp (Sperm)
AluJ (Normals)
AluJ (Adjacent)
AluJ (Tumor)
AluJ (Sperm)
Alu (Normals)
Alu (Adjacent)
Alu (Tumor)
Alu (Sperm)
AluJo (Normals)
AluJo (Adjacent)
AluJo (Tumor)
AluJo (Sperm)
AluSq/x (Normals)
AluSq/x (Adjacent)
AluSq/x (Tumor)
AluSq/x (Sperm)
AluSc (Normals)
AluSc (Adjacent)
AluSc (Tumor)
AluSc (Sperm)
AluY (Normals)
AluY (Adjacent)
AluY (Tumor)
AluY (Sperm)
AluYd (Normals)
AluYd (Adjacent)
AluYd (Tumor)
AluYd (Sperm)
AluYd8 (Normals)
AluYd8 (Adjacent)
AluYd8 (Tumor)
AluYd8 (Sperm)
AluYa (Normals)
AluYa (Adjacent)
AluYa (Tumor)
AluYa (Sperm)
AluYa5 (Normals)
AluYa5 (Adjacent)
AluYa5 (Tumor)
AluYa5 (Sperm)
AluYb9 (Normals)
AluYb9 (Adjacent)
AluYb9 (Tumor)
AluYb9 (Sperm)
AluYg6 (Normals)
AluYg6 (Adjacent)
AluYg6 (Tumor)
AluYg6 (Sperm)
AluYg (Normals)
AluYg (Adjacent)
AluYg (Tumor)
AluYg (Sperm)
AluYb (Normals)
AluYb (Adjacent)
AluYb (Tumor)
AluYb (Sperm)
16/45
−0.4
−0.2
0.0
0.2
0.4

Figure S8.

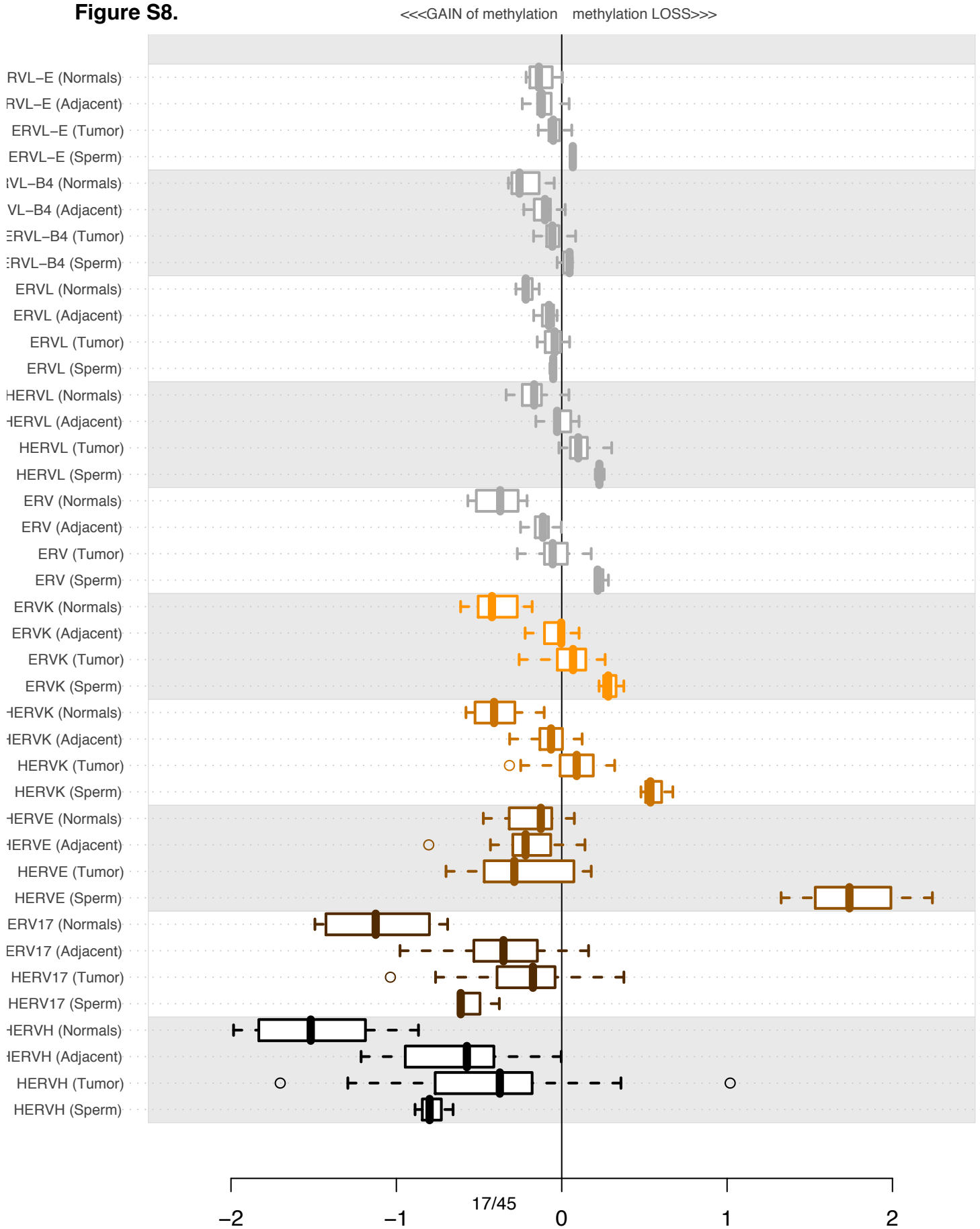

<<<GAIN of methylation methylation LOSS>>>
RVL–E (Normals)
RVL–E (Adjacent)
ERVL–E (Tumor)
ERVL–E (Sperm)
VL–B4 (Normals)
VL–B4 (Adjacent)
RVL–B4 (Tumor)
RVL–B4 (Sperm)
ERVL (Normals)
ERVL (Adjacent)
ERVL (Tumor)
ERVL (Sperm)
HERVL (Normals)
HERVL (Adjacent)
HERVL (Tumor)
HERVL (Sperm)
ERV (Normals)
ERV (Adjacent)
ERV (Tumor)
ERV (Sperm)
ERVK (Normals)
ERVK (Adjacent)
ERVK (Tumor)
ERVK (Sperm)
HERVK (Normals)
HERVK (Adjacent)
HERVK (Tumor)
HERVK (Sperm)
HERVE (Normals)
HERVE (Adjacent)
HERVE (Tumor)
HERVE (Sperm)
ERV17 (Normals)
ERV17 (Adjacent)
HERV17 (Tumor)
HERV17 (Sperm)
HERVH (Normals)
HERVH (Adjacent)
HERVH (Tumor)
HERVH (Sperm)
17/45
−2
−1
0
1
2

Figure S9.

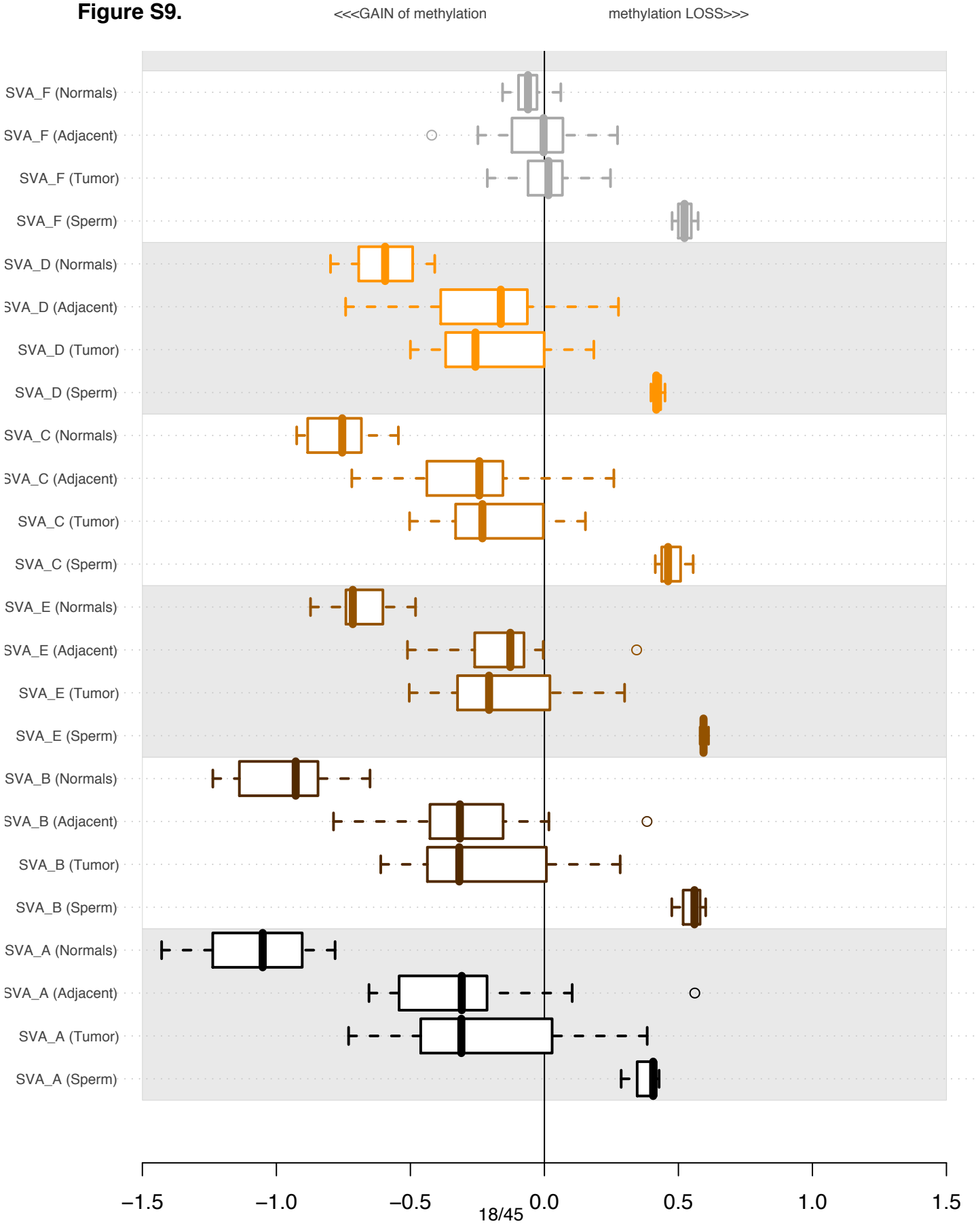

<<<GAIN of methylation
methylation LOSS>>>
SVA_F (Normals)
SVA_F (Adjacent)
SVA_F (Tumor)
SVA_F (Sperm)
SVA_D (Normals)
SVA_D (Adjacent)
SVA_D (Tumor)
SVA_D (Sperm)
SVA_C (Normals)
SVA_C (Adjacent)
SVA_C (Tumor)
SVA_C (Sperm)
SVA_E (Normals)
SVA_E (Adjacent)
SVA_E (Tumor)
SVA_E (Sperm)
SVA_B (Normals)
SVA_B (Adjacent)
SVA_B (Tumor)
SVA_B (Sperm)
SVA_A (Normals)
SVA_A (Adjacent)
SVA_A (Tumor)
SVA_A (Sperm)
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

Figure S10.

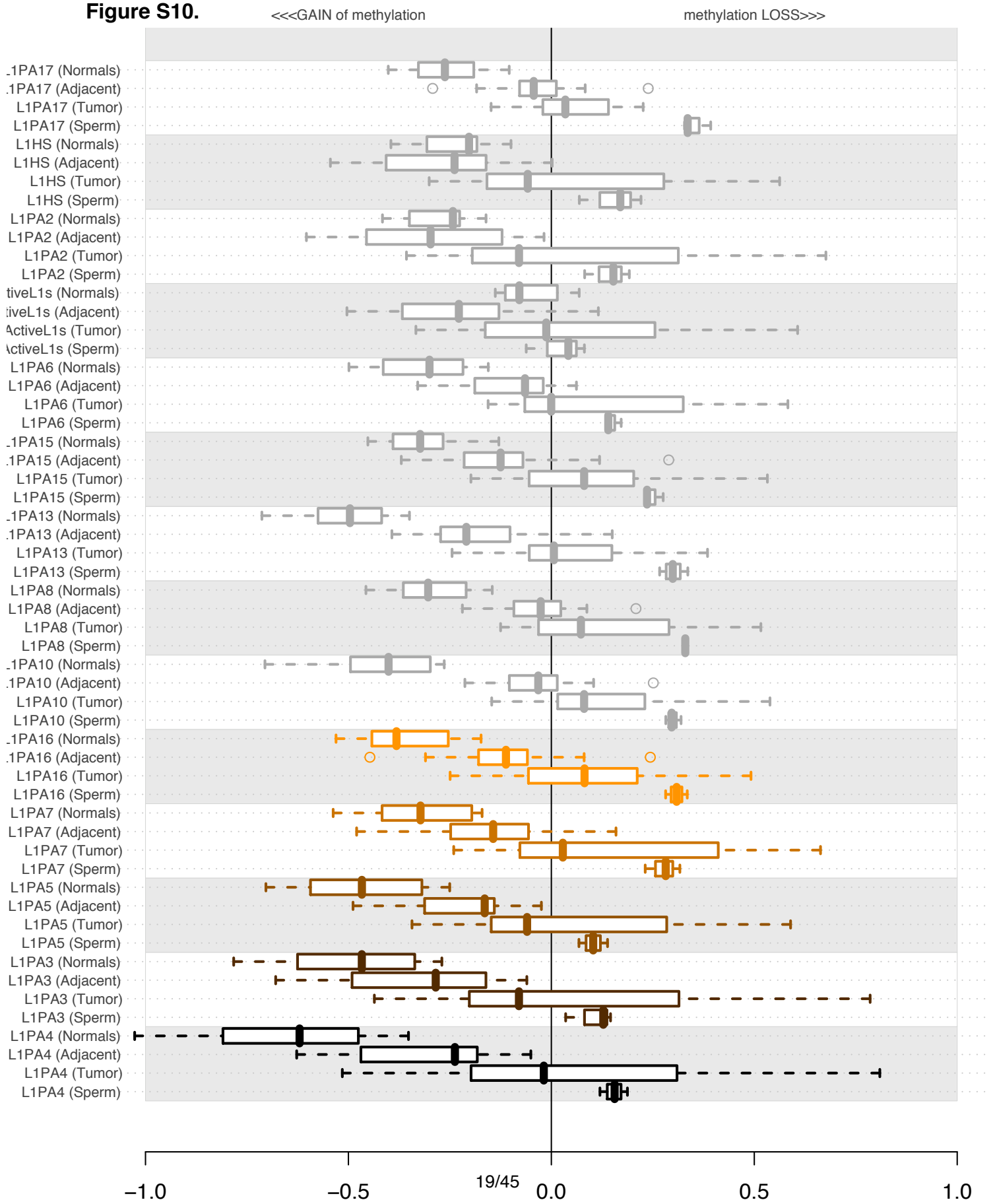
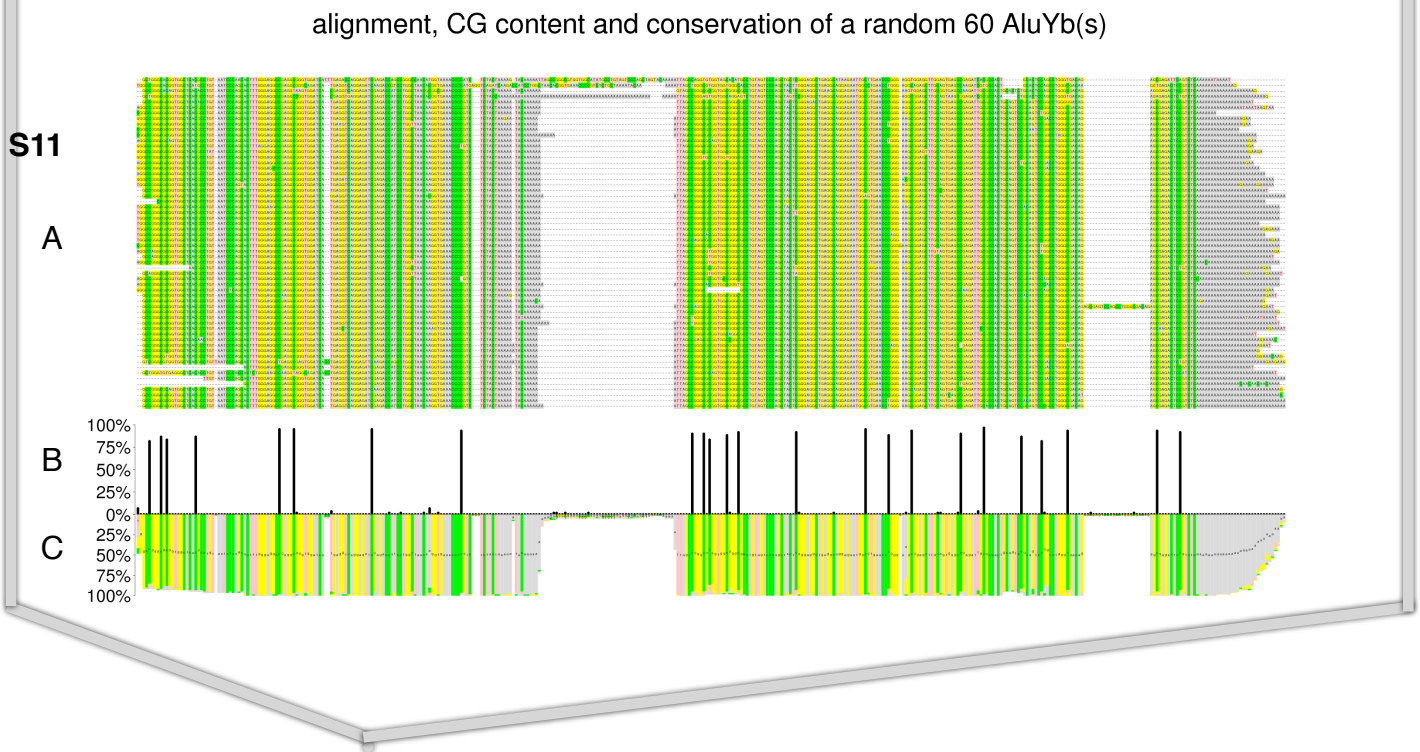

<<<GAIN of methylation
methylation LOSS>>>
L1PA17 (Normals)
L1PA17 (Adjacent)
L1PA17 (Tumor)
L1PA17 (Sperm)
L1HS (Normals)
L1HS (Adjacent)
L1HS (Tumor)
L1HS (Sperm)
L1PA2 (Normals)
L1PA2 (Adjacent)
L1PA2 (Tumor)
L1PA2 (Sperm)
L1PA6 (Normals)
L1PA6 (Adjacent)
L1PA6 (Tumor)
L1PA6 (Sperm)
L1PA15 (Normals)
L1PA15 (Adjacent)
L1PA15 (Tumor)
L1PA15 (Sperm)
L1PA13 (Normals)
L1PA13 (Adjacent)
L1PA13 (Tumor)
L1PA13 (Sperm)
L1PA8 (Normals)
L1PA8 (Adjacent)
L1PA8 (Tumor)
L1PA8 (Sperm)
L1PA10 (Normals)
L1PA10 (Adjacent)
L1PA10 (Tumor)
L1PA10 (Sperm)
L1PA16 (Normals)
L1PA16 (Adjacent)
L1PA16 (Tumor)
L1PA16 (Sperm)
L1PA7 (Normals)
L1PA7 (Adjacent)
L1PA7 (Tumor)
L1PA7 (Sperm)
L1PA5 (Normals)
L1PA5 (Adjacent)
L1PA5 (Tumor)
L1PA5 (Sperm)
L1PA3 (Normals)
L1PA3 (Adjacent)
L1PA3 (Tumor)
L1PA3 (Sperm)
L1PA4 (Normals)
L1PA4 (Adjacent)
L1PA4 (Tumor)
L1PA4 (Sperm)
−1.0
−0.5
19/45
0.0
0.5
1.0

Figure S11


alignment, CG content and conservation of a random 60 AluYb(s)
A
B
C
100%
75%
50%
25%
0%
25%
50%
75%
100%

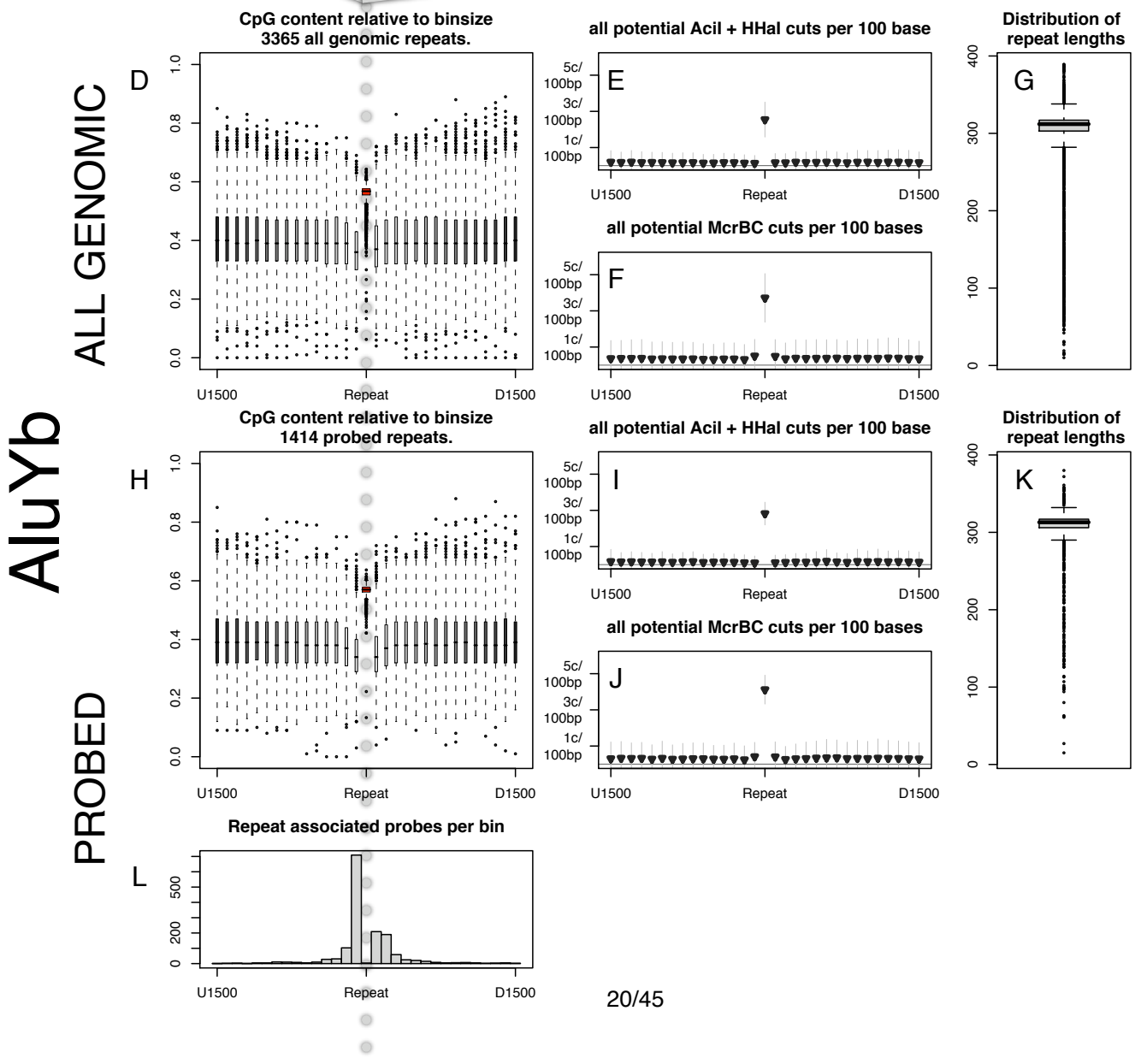

AluYb
ALL GENOMIC
PROBED
D
CpG content relative to binsize
3365 all genomic repeats.
E
all potential AciI + HHaI cuts per 100 base
F
all potential McrBC cuts per 100 bases
G
Distribution of
repeat lengths
H
CpG content relative to binsize
1414 probed repeats.
I
all potential AciI + HHaI cuts per 100 base
J
all potential McrBC cuts per 100 bases
K
L
Repeat associated probes per bin
U1500
Repeat
D1500
5c/
100bp
3c/
100bp
1c/
100bp

Figure S12

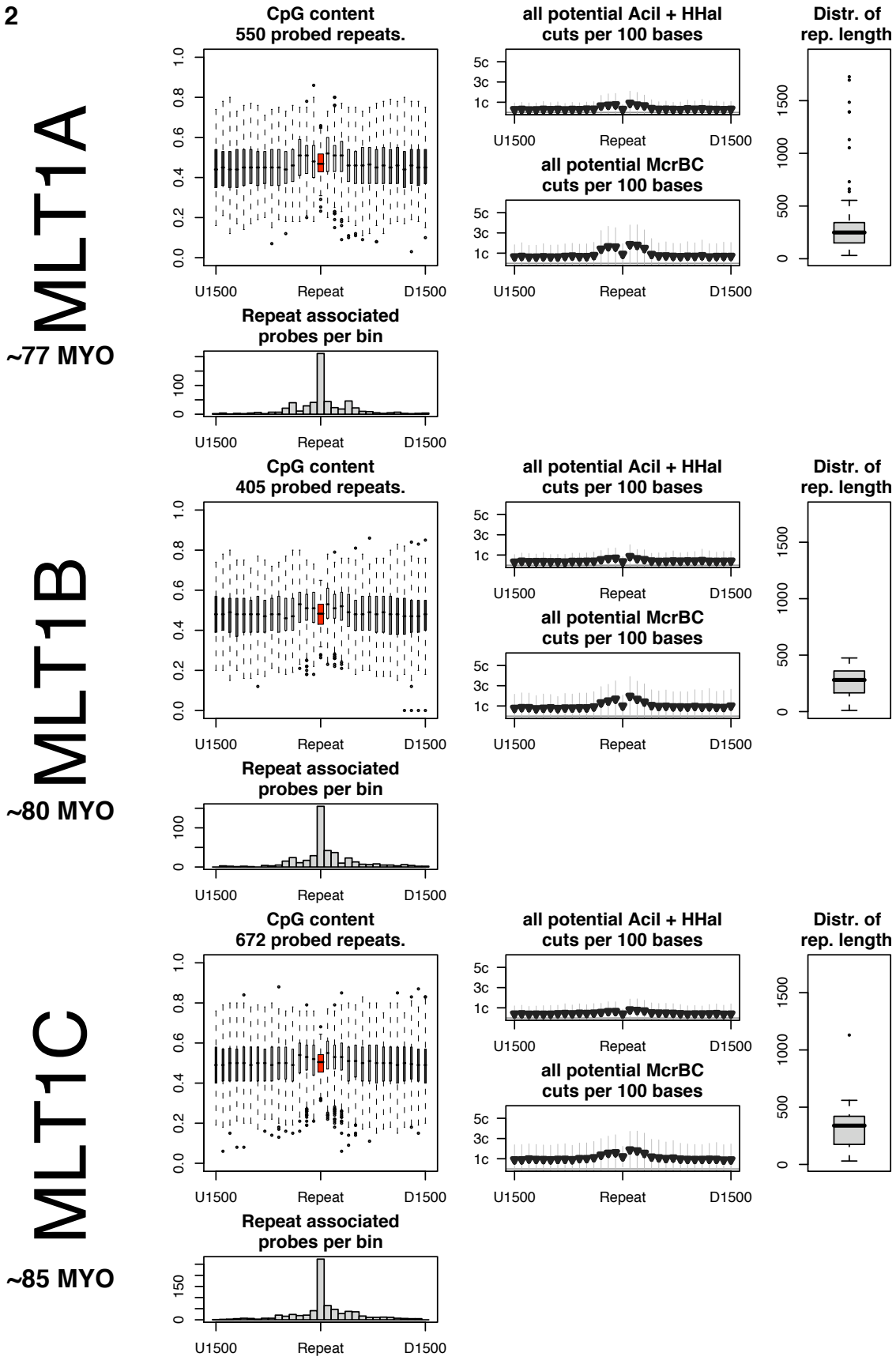

MLT1A
~77 MYO
CpG content
550 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
MLT1B
~80 MYO
CpG content
405 probed repeats.
MLT1C
~85 MYO
CpG content
672 probed repeats.
U1500
Repeat
D1500

Figure S12

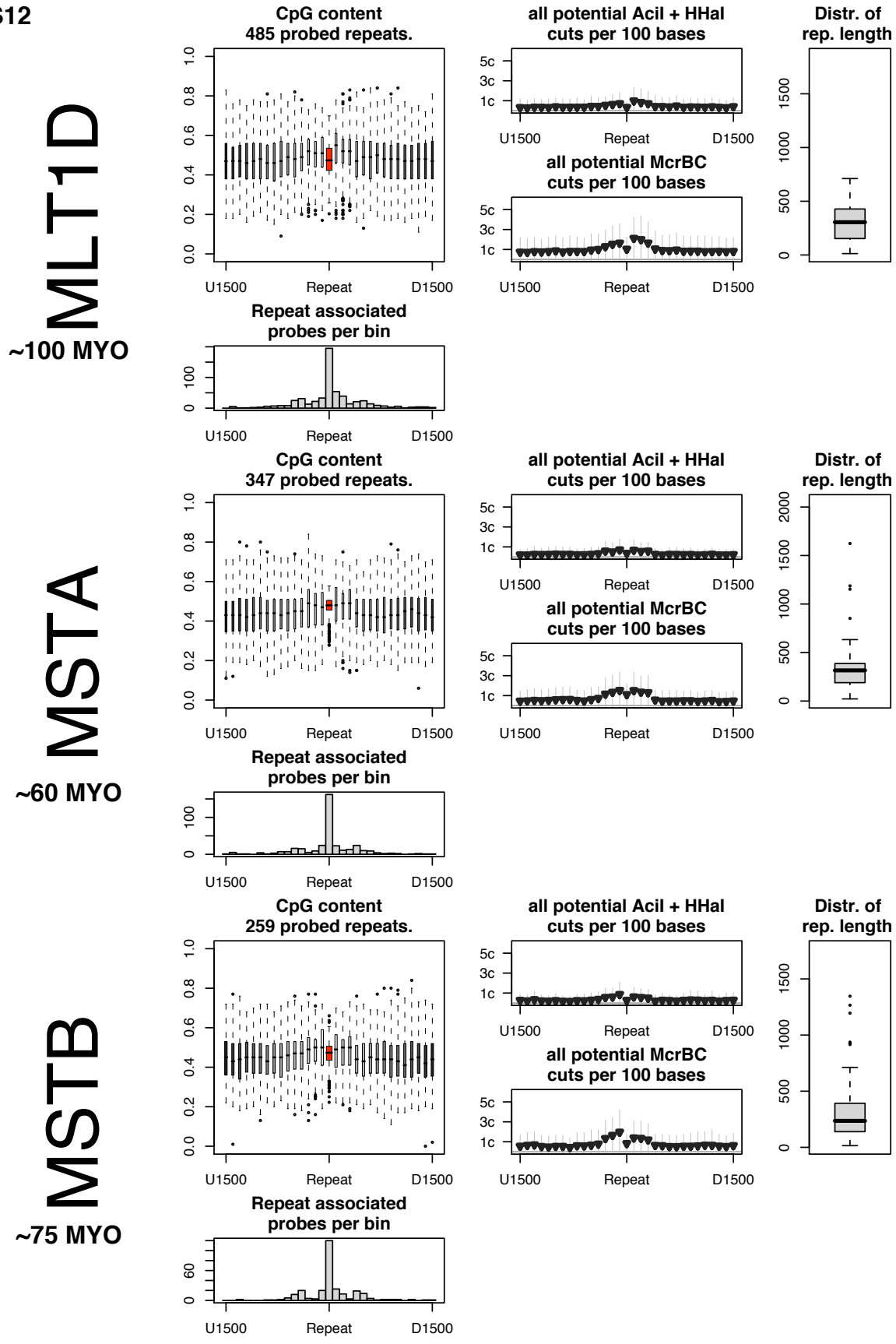

MLT1D
~100 MYO
CpG content
485 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
MSTA
~60 MYO
CpG content
347 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
MSTB
~75 MYO
CpG content
259 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
U1500
Repeat
D1500

Figure S12

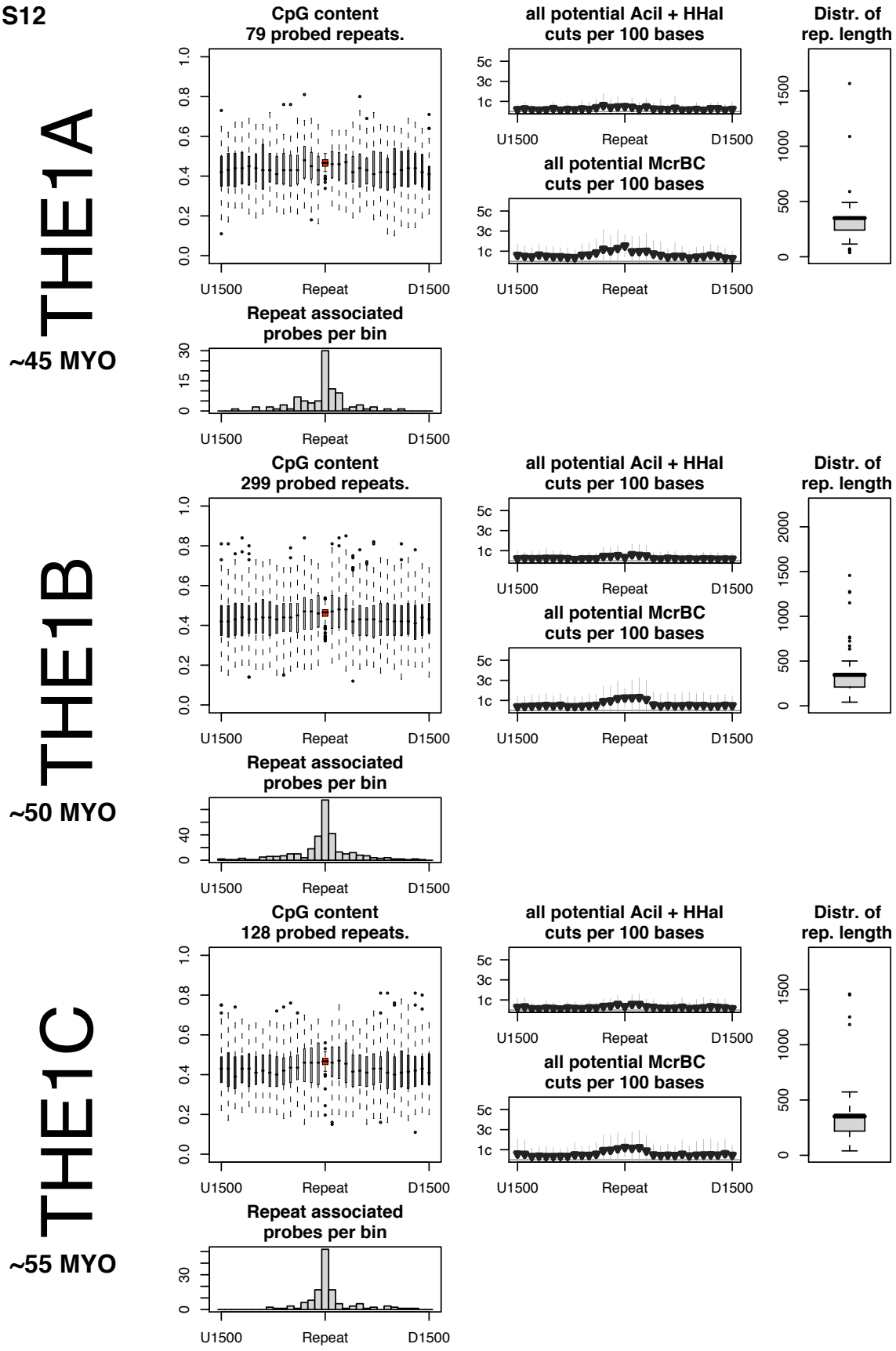

THE1A
~45 MYO
CpG content
79 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
THE1B
~50 MYO
CpG content
299 probed repeats.
THE1C
~55 MYO
CpG content
128 probed repeats.
U1500
Repeat
D1500

Figure S13

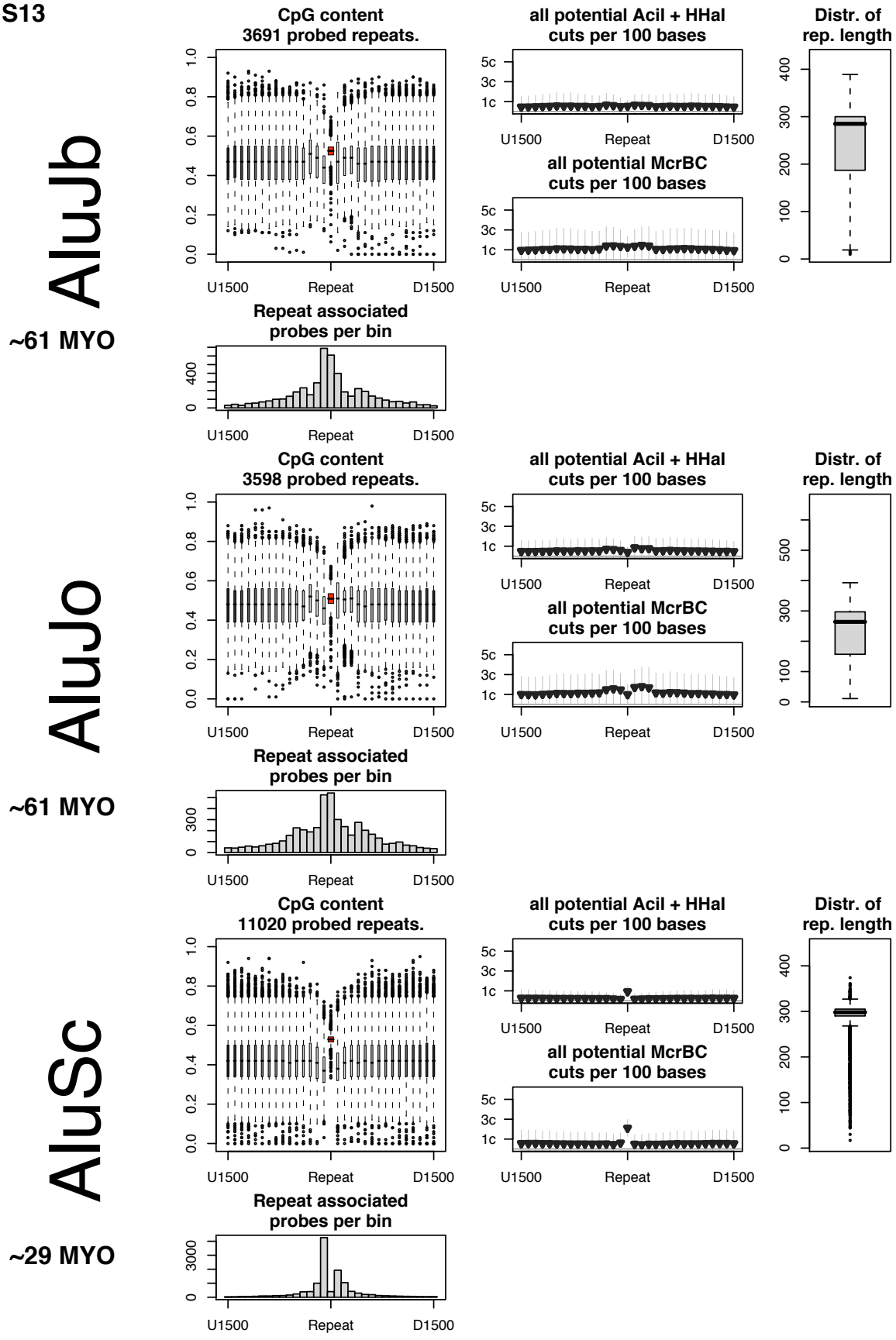

AluJb
~61 MYO
CpG content
3691 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
AluJo
~61 MYO
CpG content
3598 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
AluSc
~29 MYO
CpG content
11020 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
U1500
Repeat
D1500

Figure S13

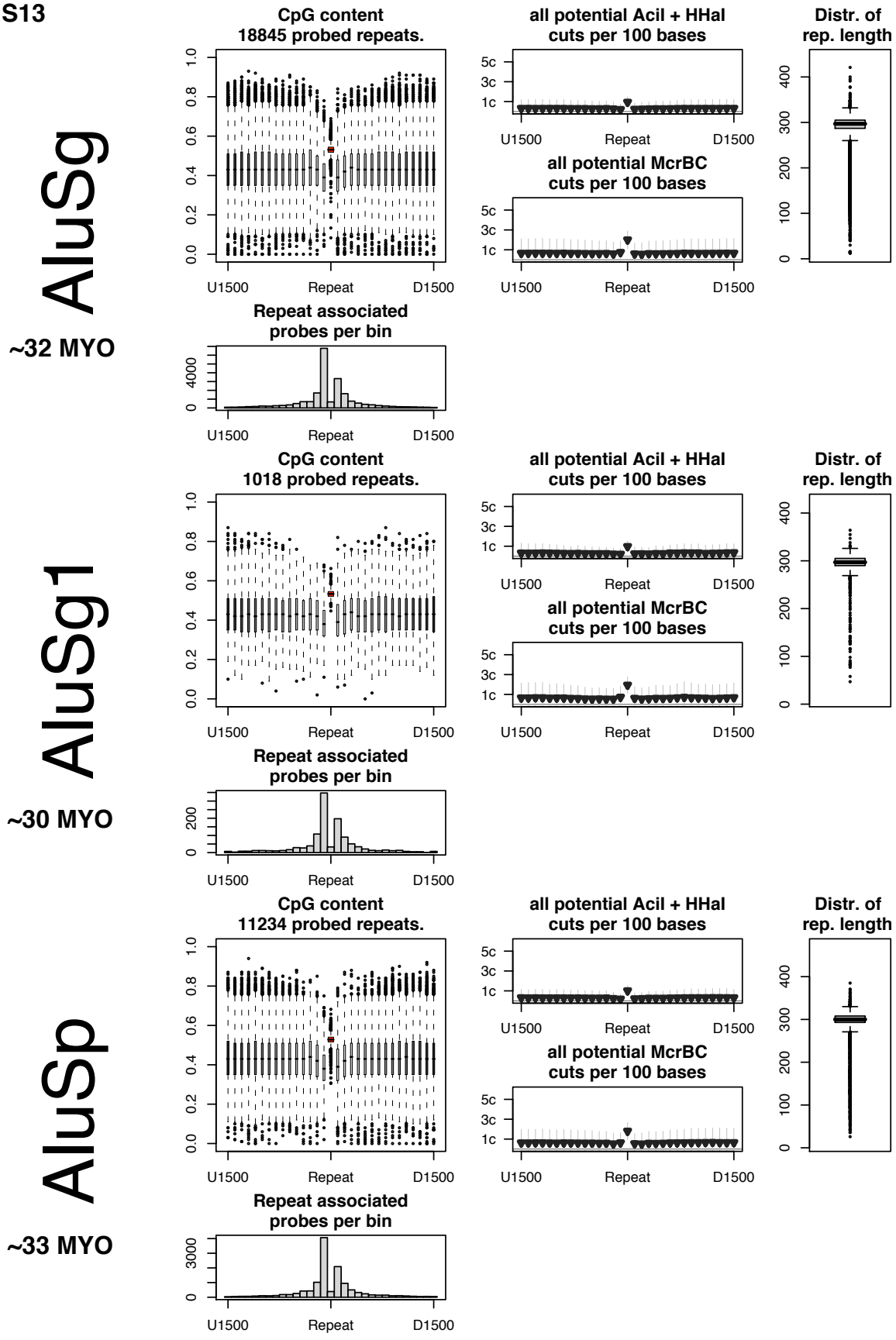

AluSg
~32 MYO
CpG content
18845 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
AluSg1
~30 MYO
CpG content
1018 probed repeats.
AluSp
~33 MYO
CpG content
11234 probed repeats.
U1500
Repeat
D1500

Figure S13

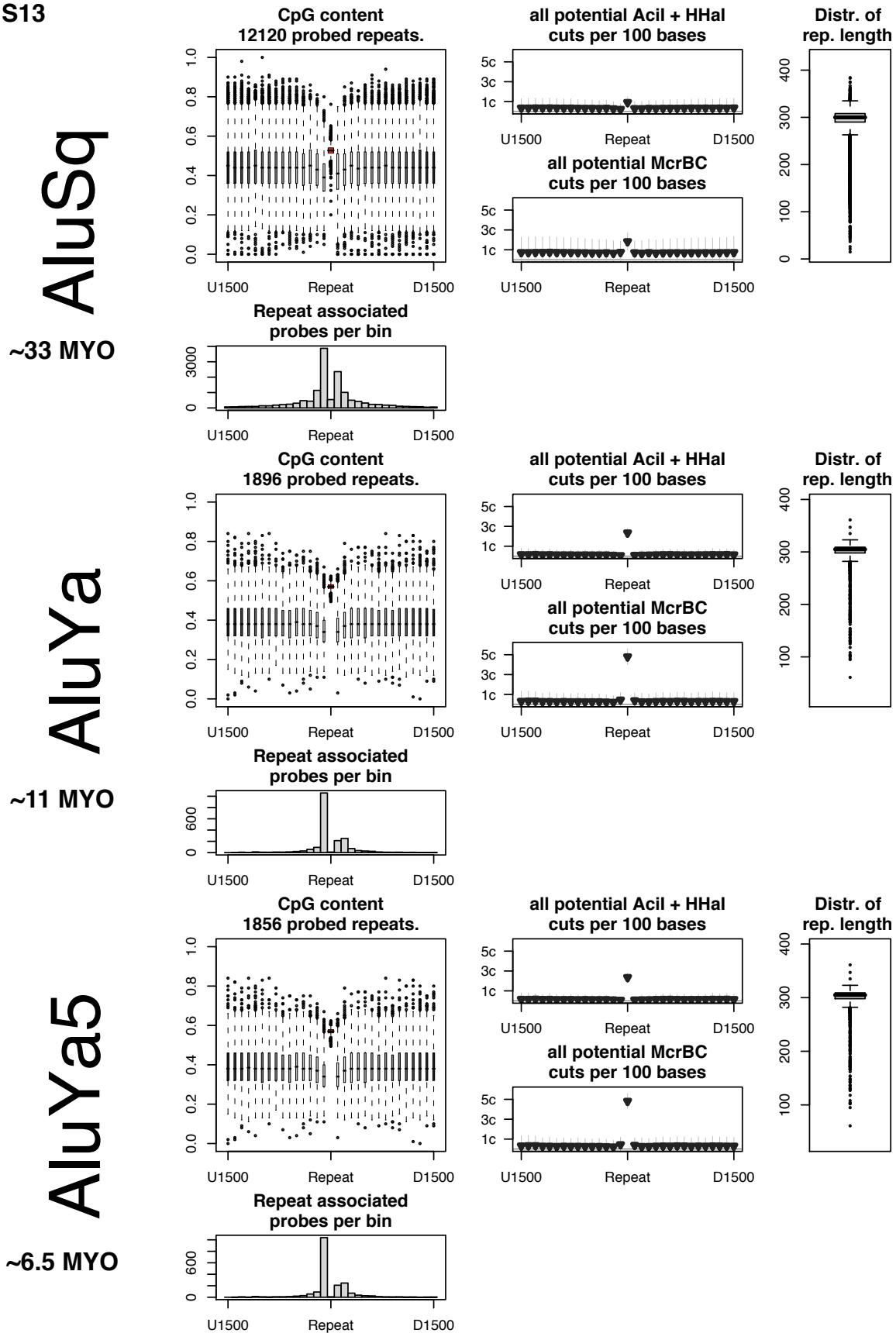
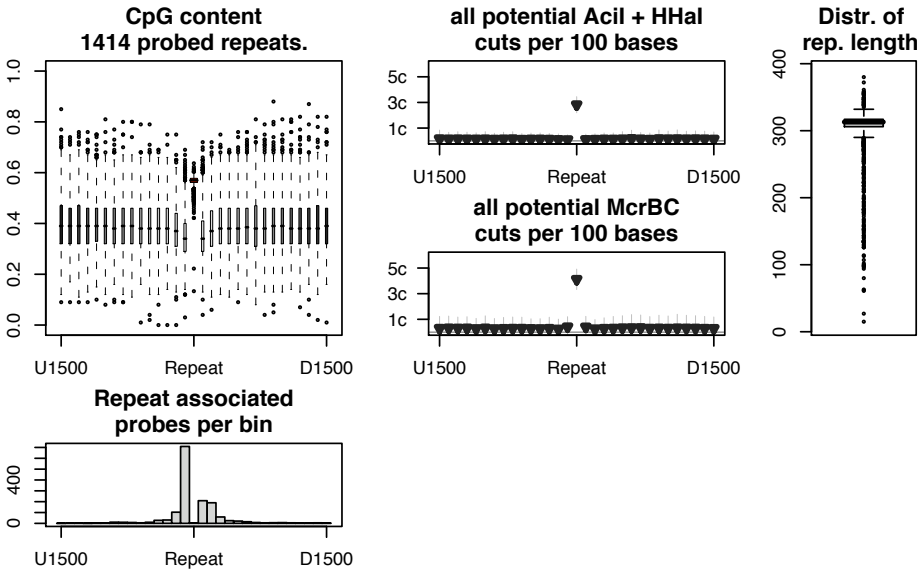

AluSq
~33 MYO
CpG content
12120 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
AluYa
~11 MYO
CpG content
1896 probed repeats.
AluYa5
~6.5 MYO
CpG content
1856 probed repeats.
U1500
Repeat
D1500

Figure S13.

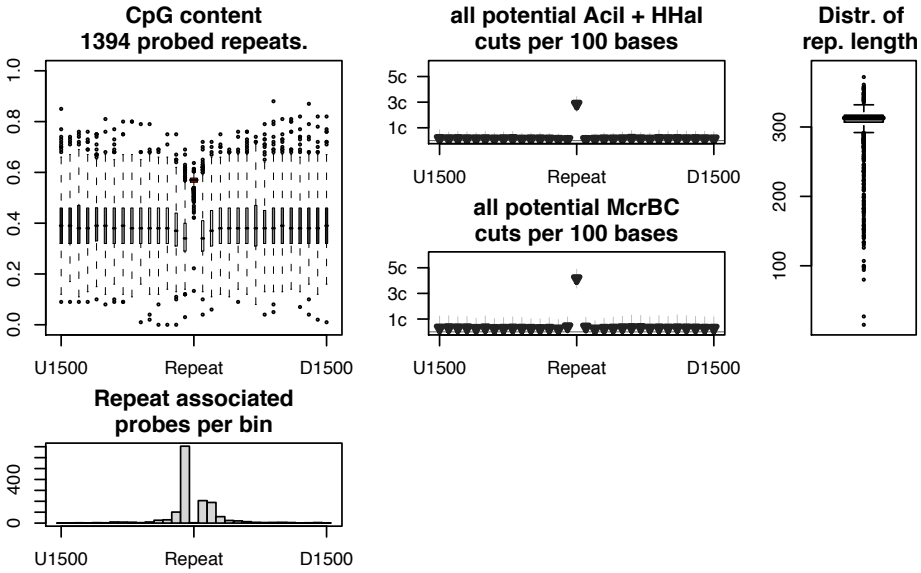
AluYb

~5 MYO



AluYb9

~3 MYO



AluYd

~15 MYO

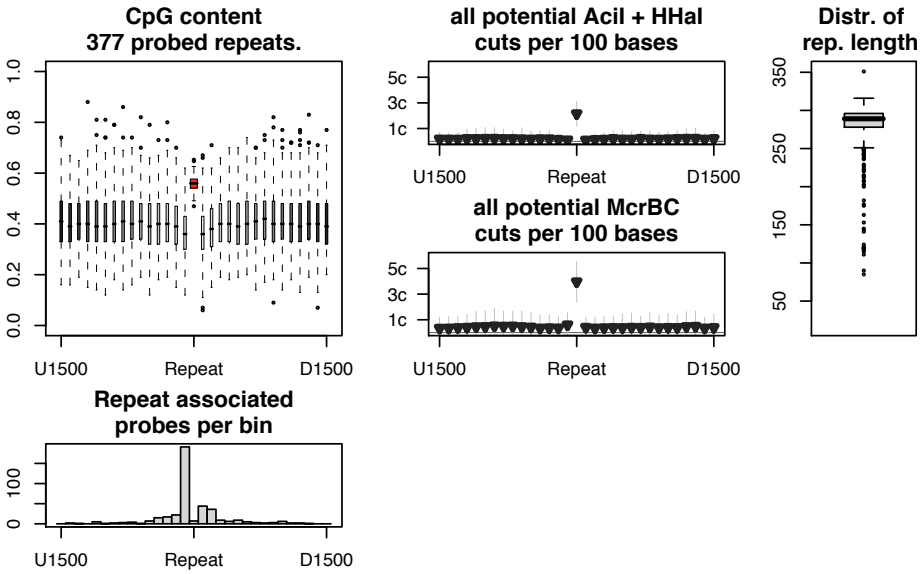


Figure S13

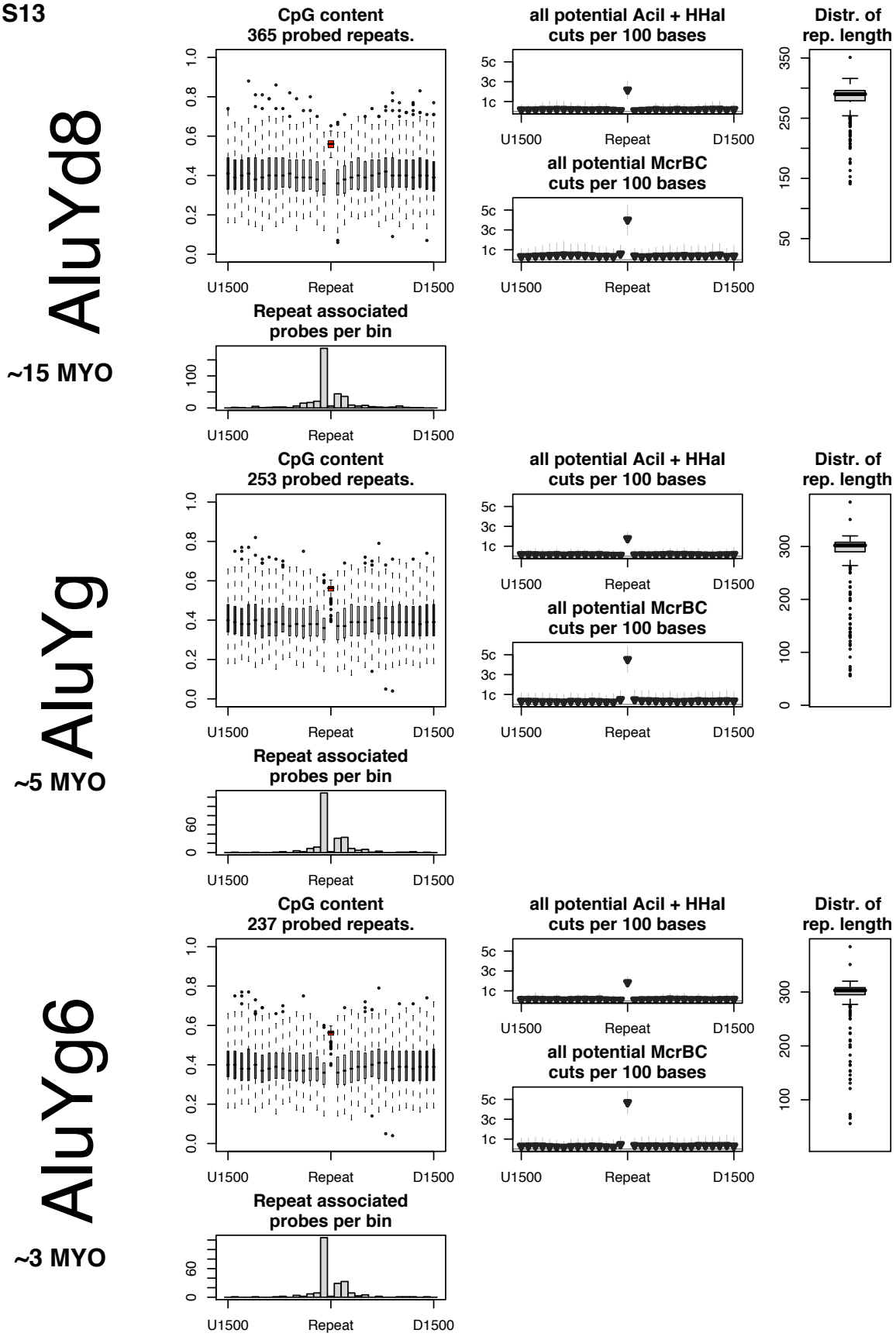

AluYd8
~15 MYO
CpG content
365 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
AluYg
~5 MYO
CpG content
253 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
AluYg6
~3 MYO
CpG content
237 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
U1500
Repeat
D1500

Figure S14

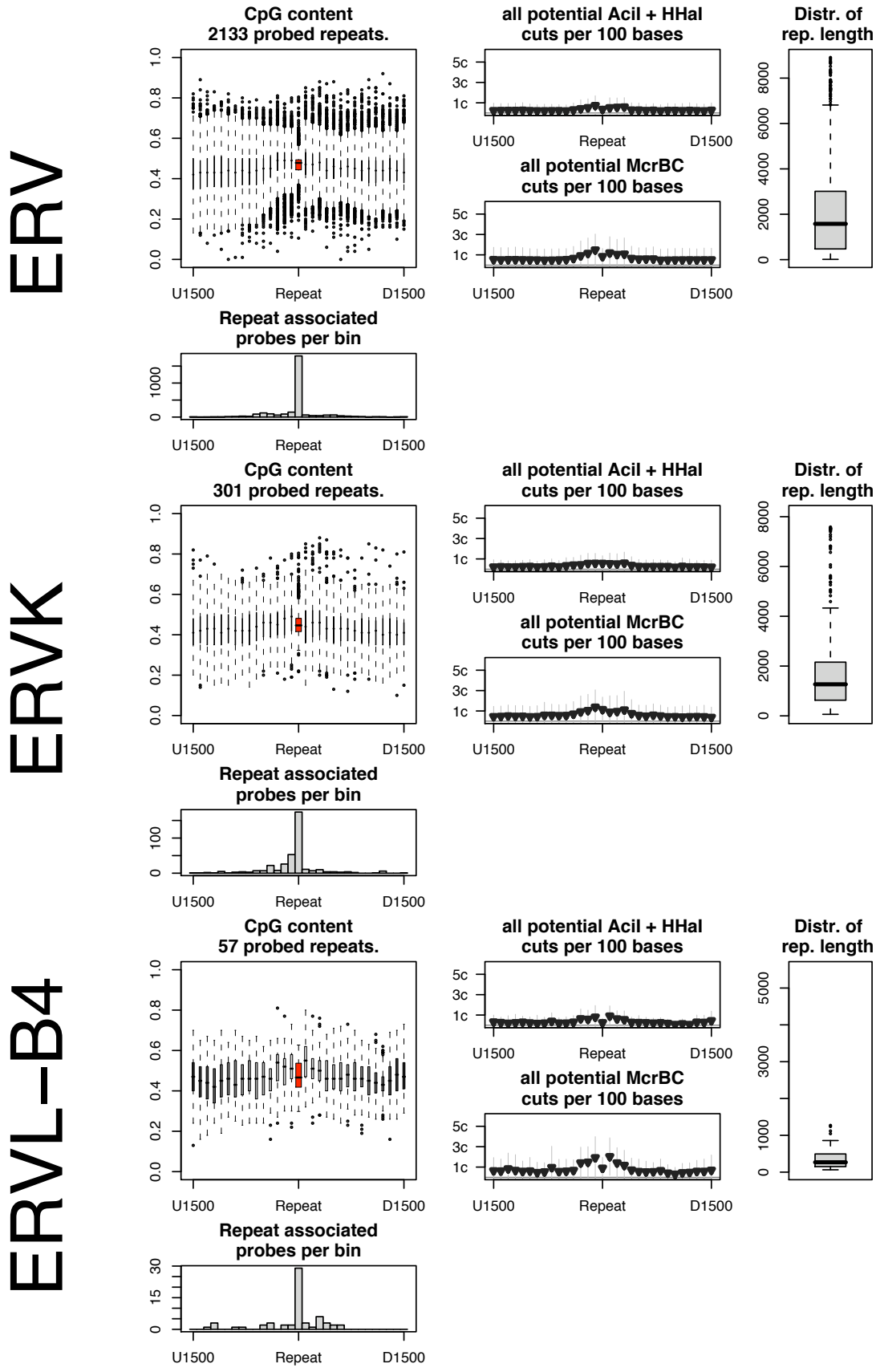

ERV
CpG content
2133 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
ERVK
CpG content
301 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
ERVL–B4
CpG content
57 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
U1500
Repeat
D1500

Figure S14

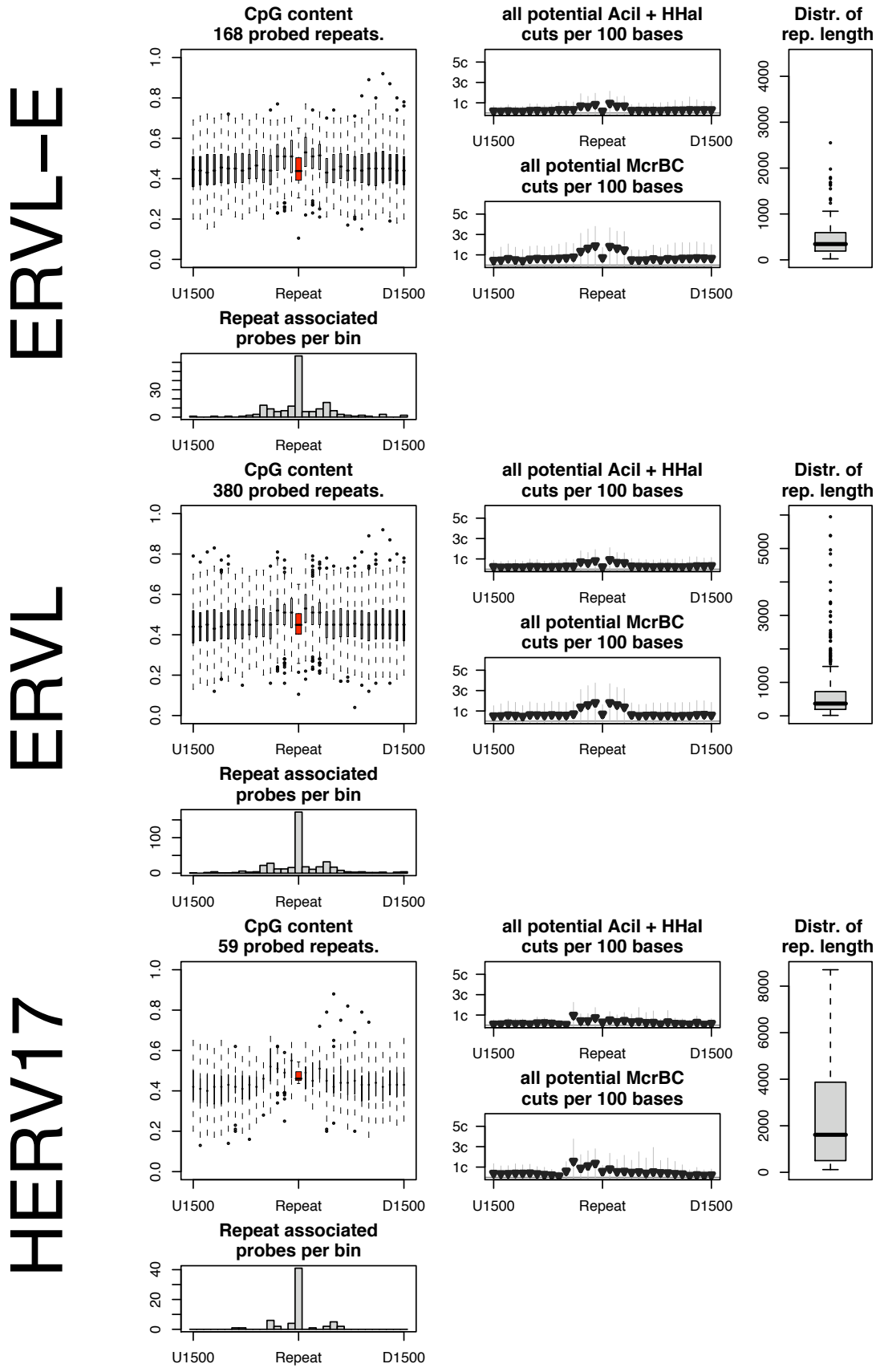

ERVL–E
ERVL
HERV17
CpG content
168 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
CpG content
380 probed repeats.
CpG content
59 probed repeats.
U1500
Repeat
D1500

Figure S14

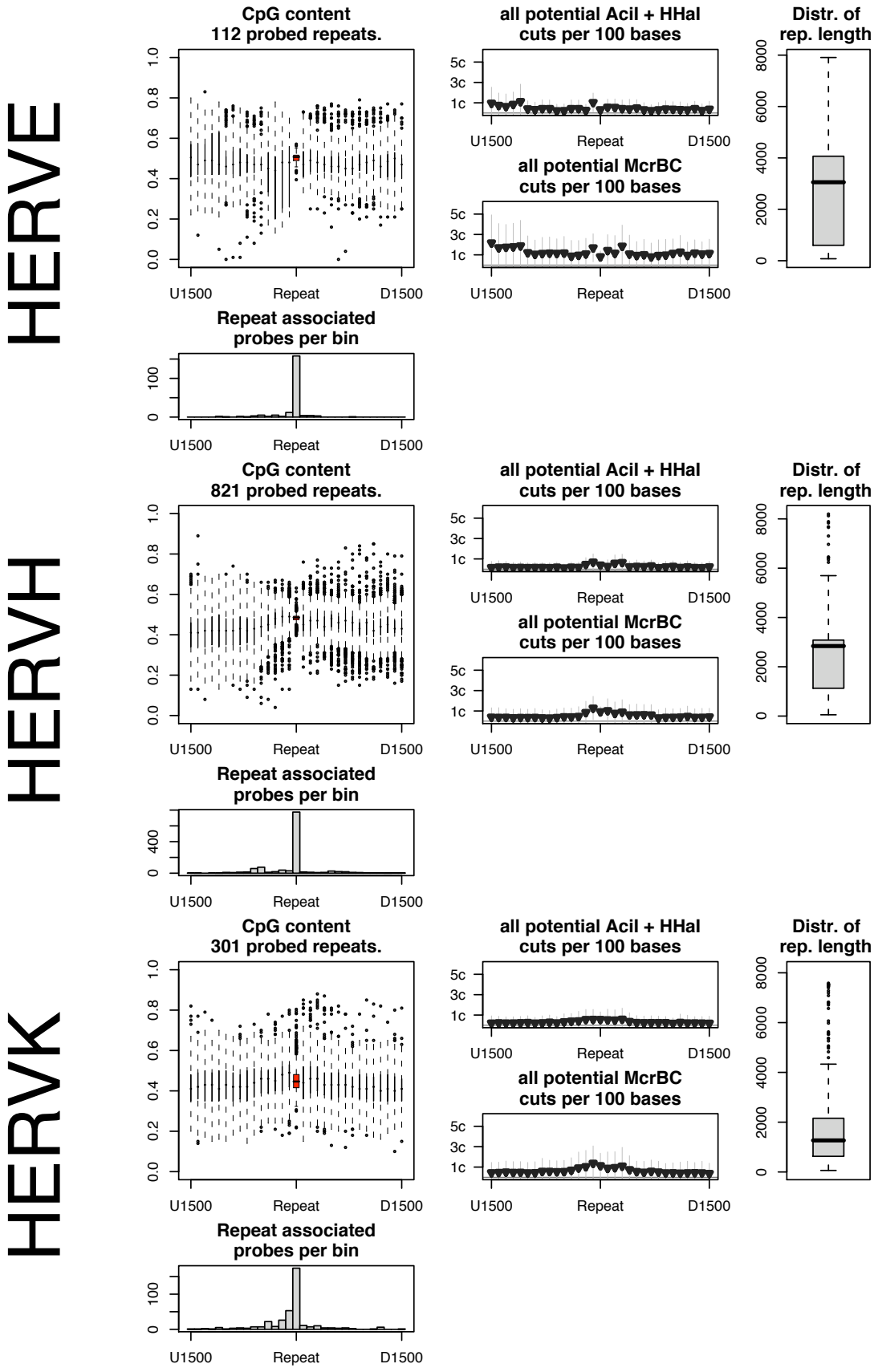

HERVE
HERVH
HERVK
CpG content
112 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
CpG content
821 probed repeats.
CpG content
301 probed repeats.
U1500
Repeat
D1500

Figure S14

HERVL

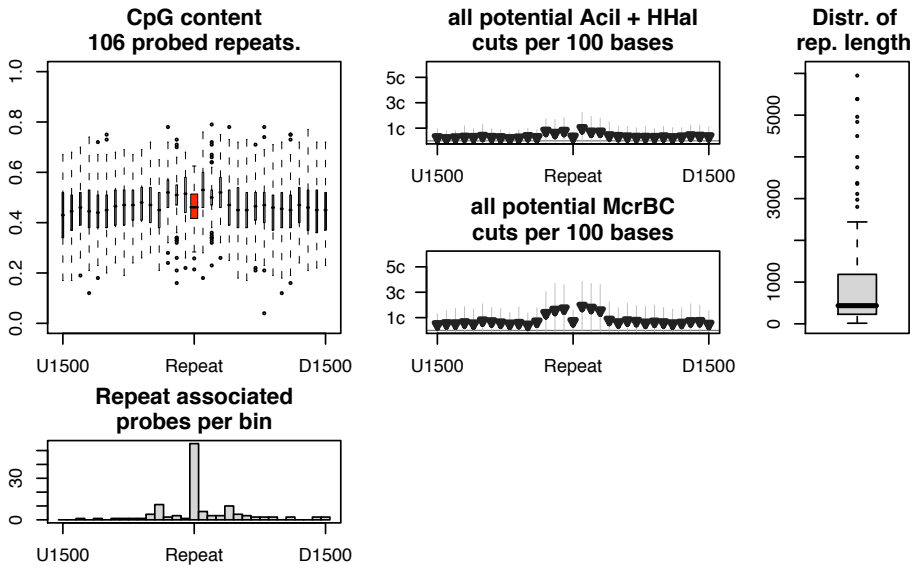


Figure S15

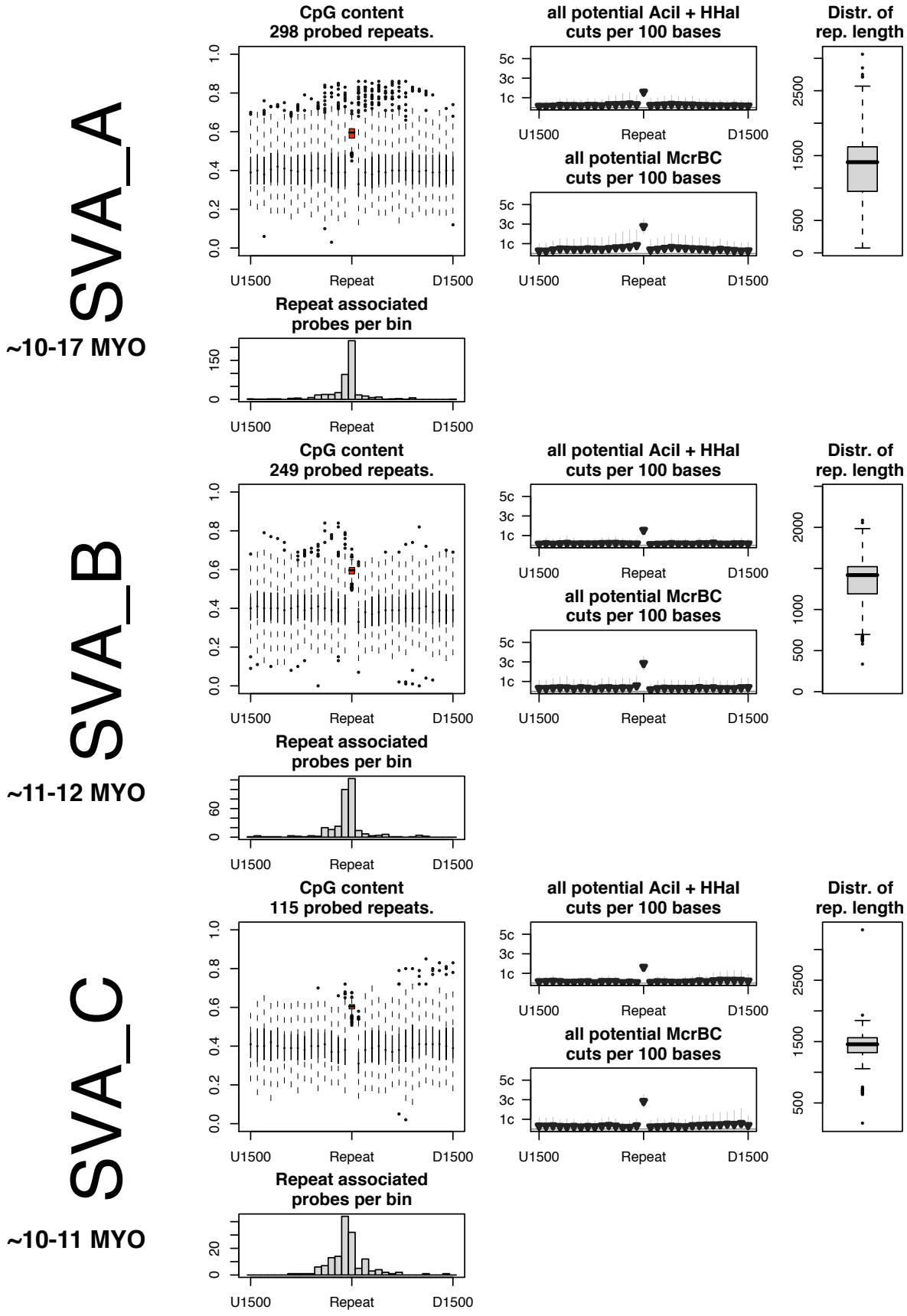

SVA_A
~10-17 MYO
CpG content
298 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
SVA_B
~11-12 MYO
CpG content
249 probed repeats.
SVA_C
~10-11 MYO
CpG content
115 probed repeats.
U1500
Repeat
D1500

Figure S15

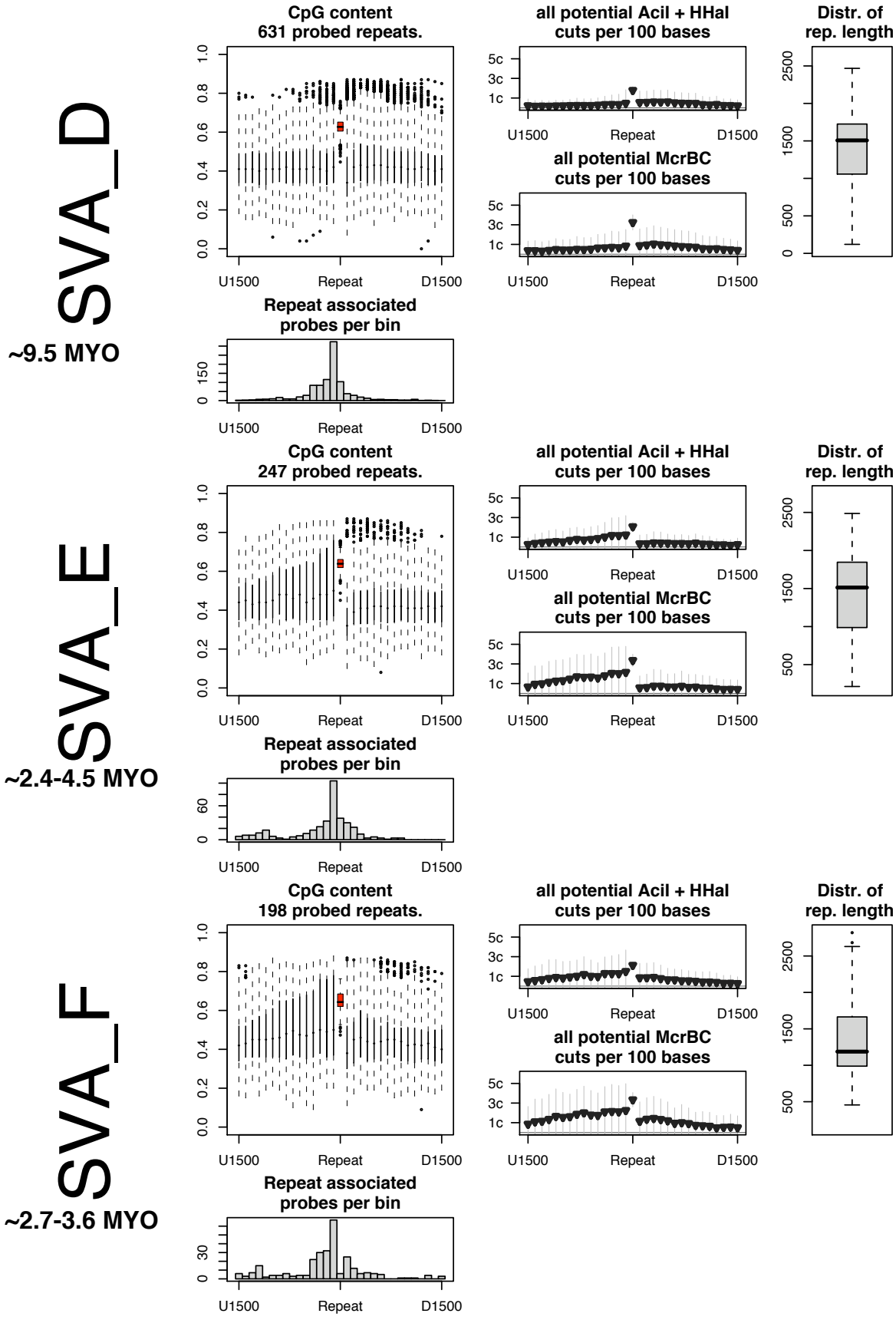

SVA_D
~9.5 MYO
CpG content
631 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
SVA_E
~2.4-4.5 MYO
CpG content
247 probed repeats.
SVA_F
~2.7-3.6 MYO
CpG content
198 probed repeats.
U1500
Repeat
D1500

Figure S16

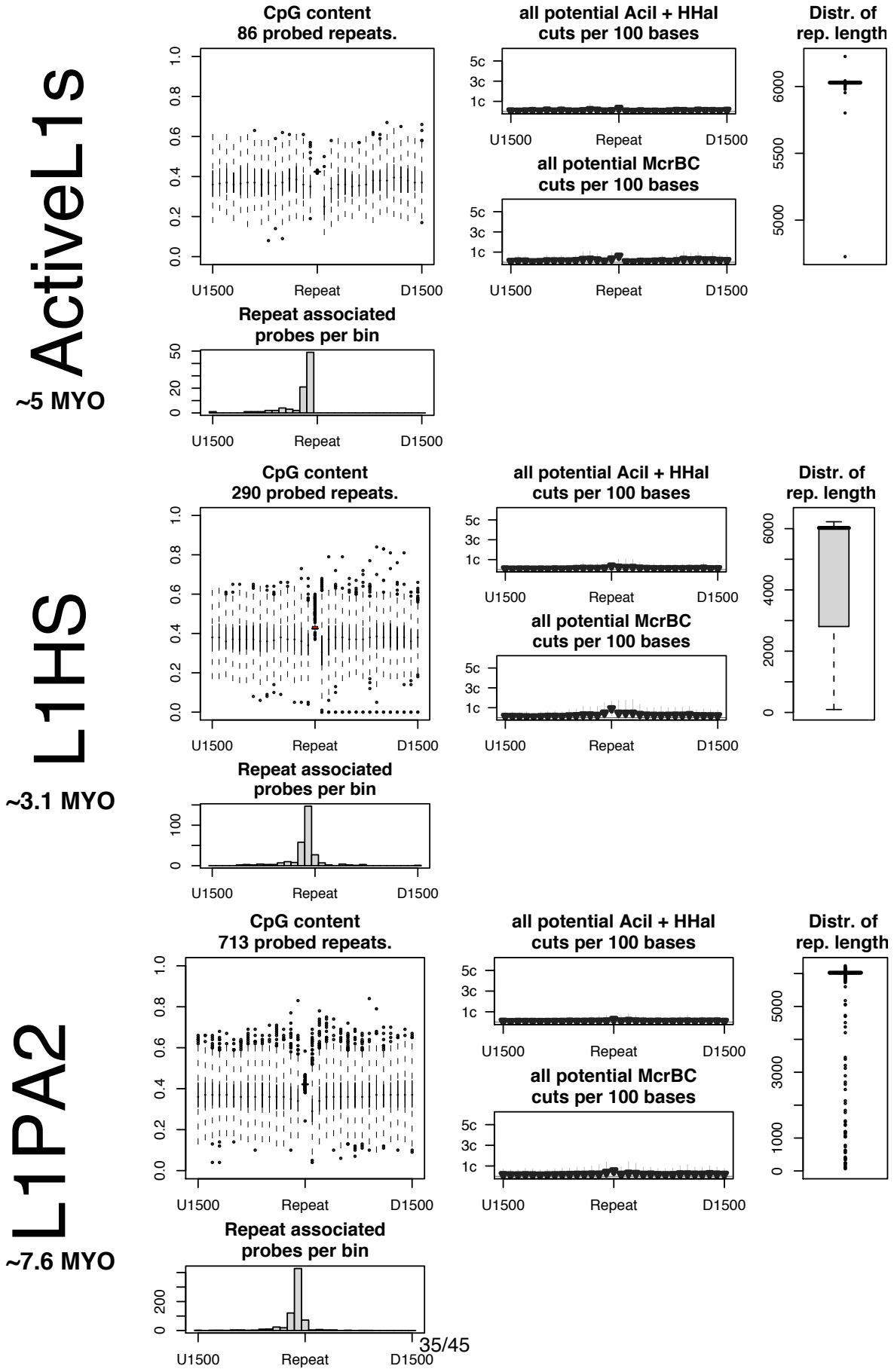

ActiveL1s
~5 MYO
CpG content
86 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
L1HS
~3.1 MYO
CpG content
290 probed repeats.
L1PA2
~7.6 MYO
CpG content
713 probed repeats.
U1500
Repeat
D1500

Figure S16

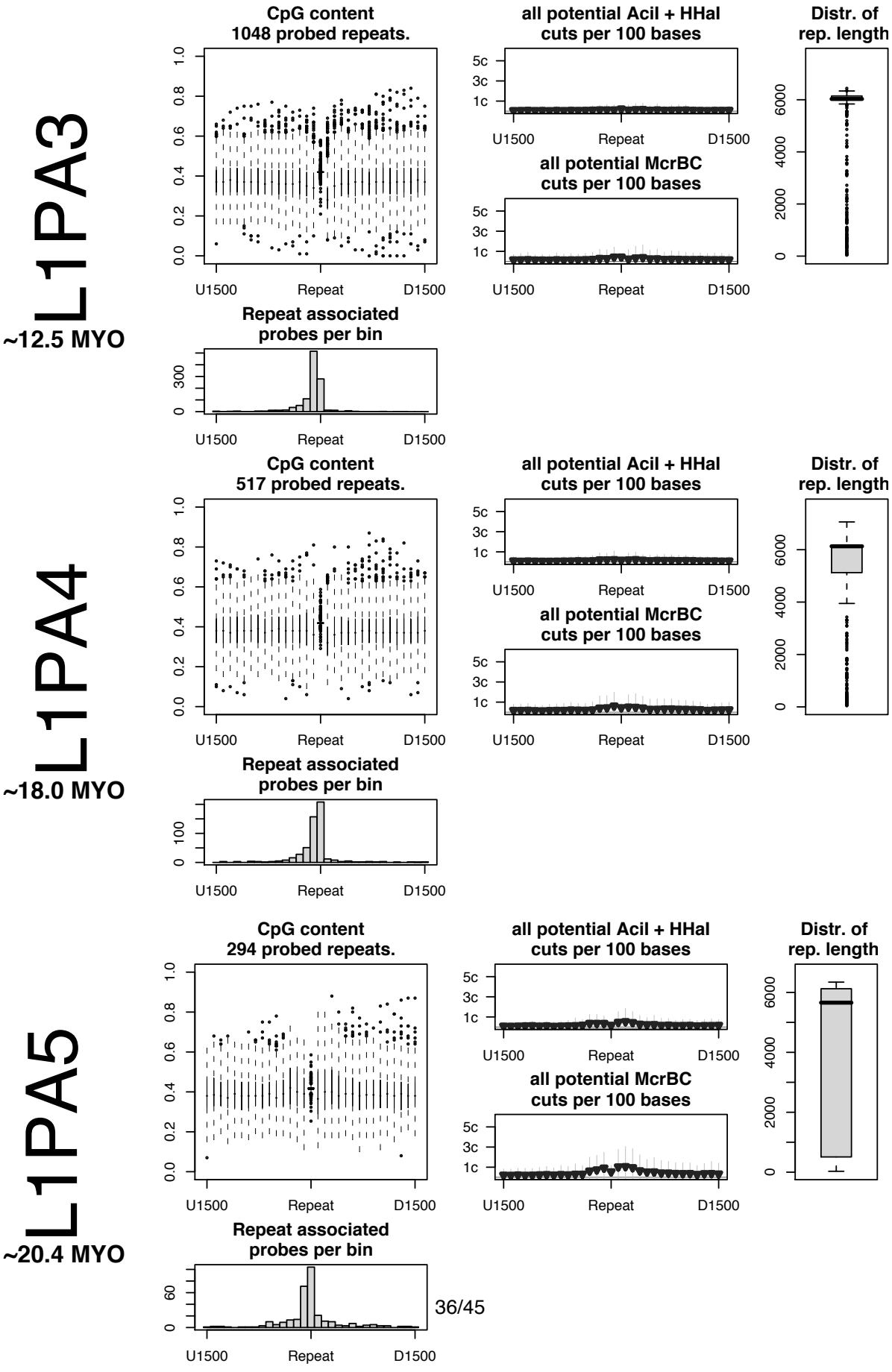

L1PA3
~12.5 MYO
CpG content
1048 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
L1PA4
~18.0 MYO
CpG content
517 probed repeats.
L1PA5
~20.4 MYO
CpG content
294 probed repeats.
U1500
Repeat
D1500
36/45

Figure S16

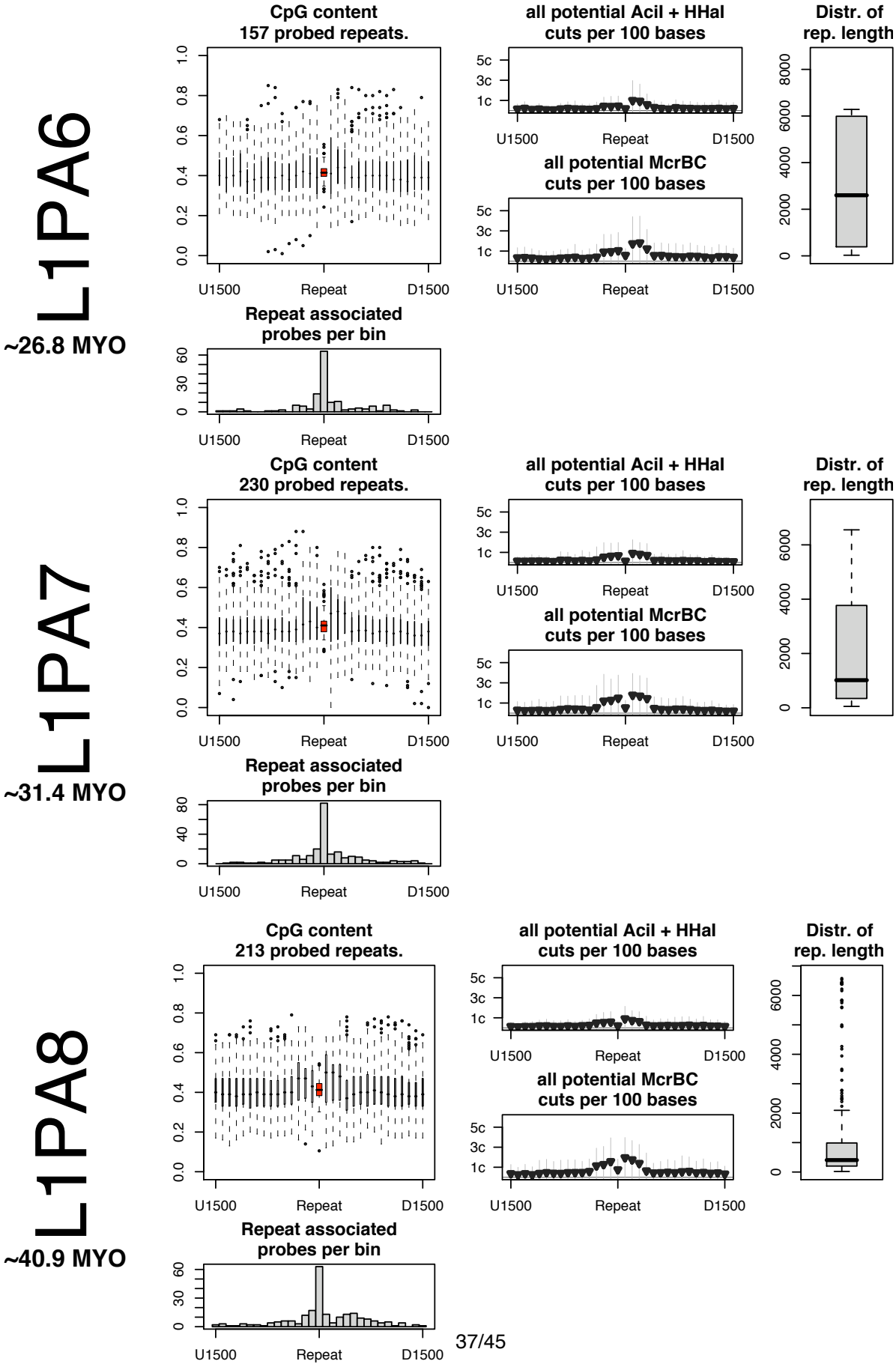
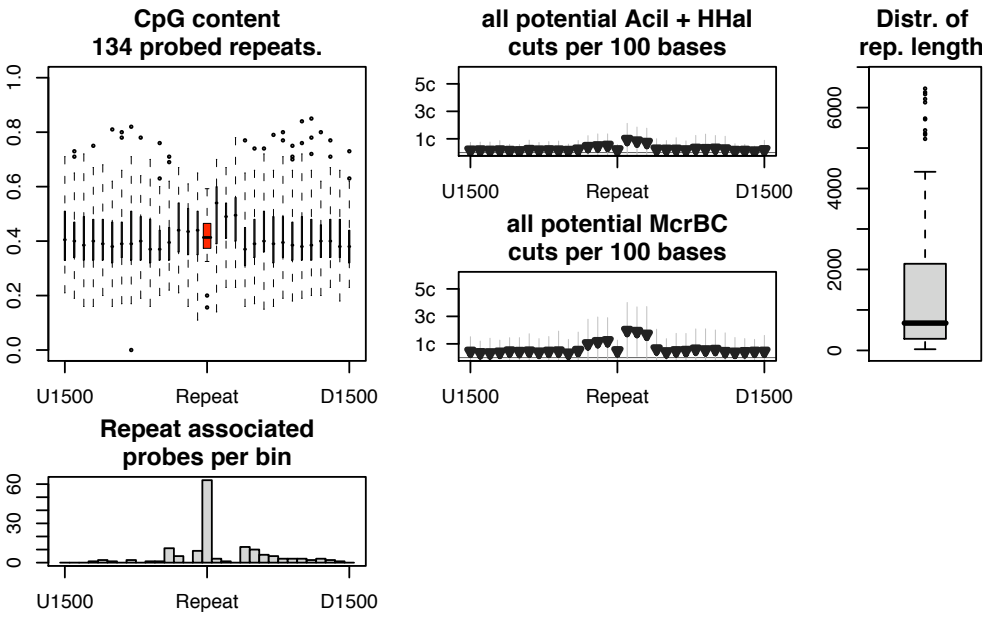

L1PA6
~26.8 MYO
CpG content
157 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
L1PA7
~31.4 MYO
CpG content
230 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
L1PA8
~40.9 MYO
CpG content
213 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
U1500
Repeat
D1500

Figure S16

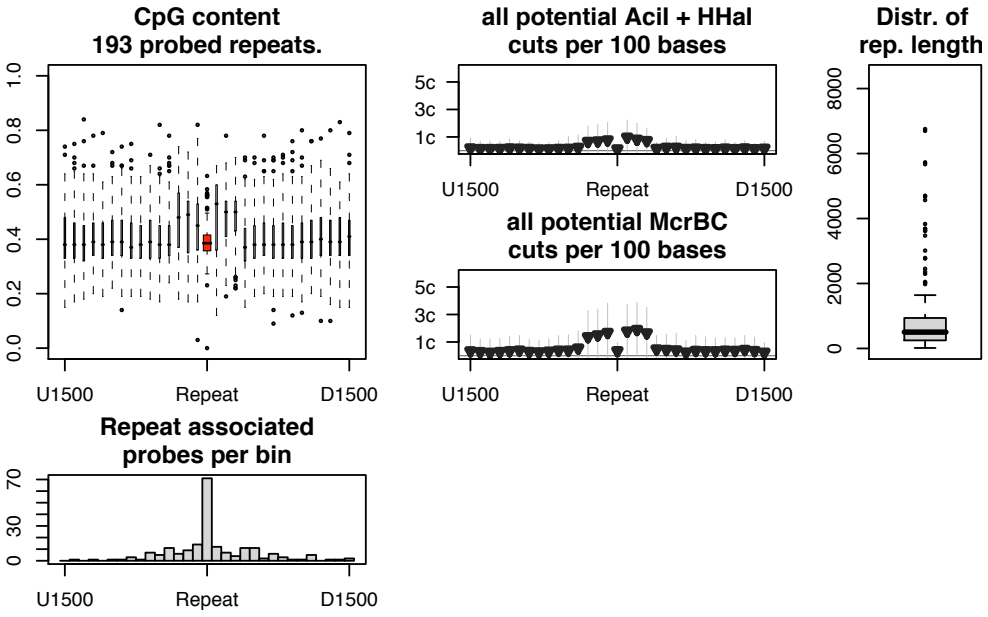
L1PA10

~46.4 MYO



L1PA13

~60 MYO



L1PA15

~70.5 MYO

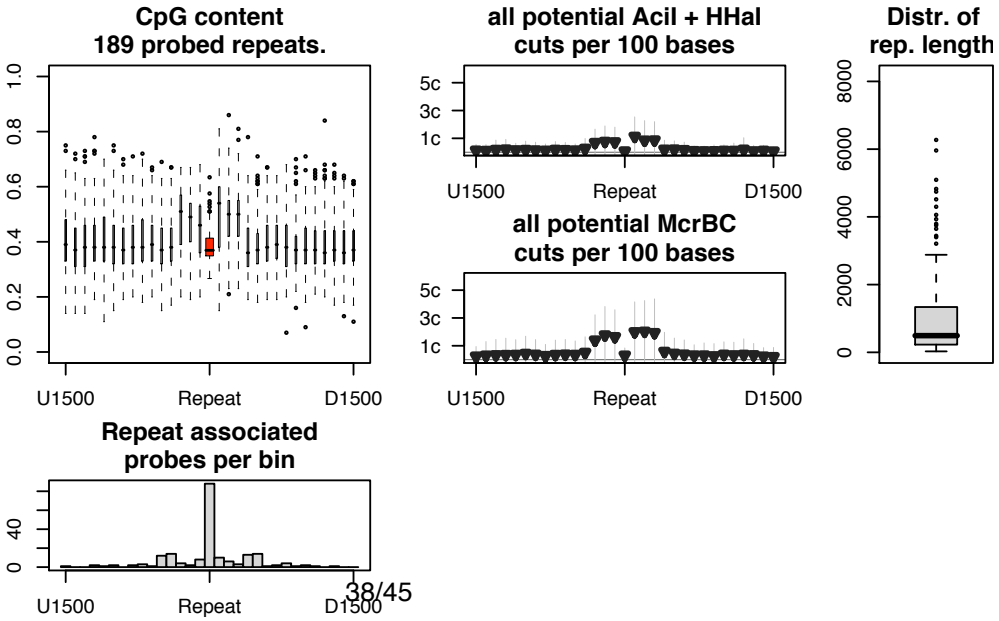


Figure S16

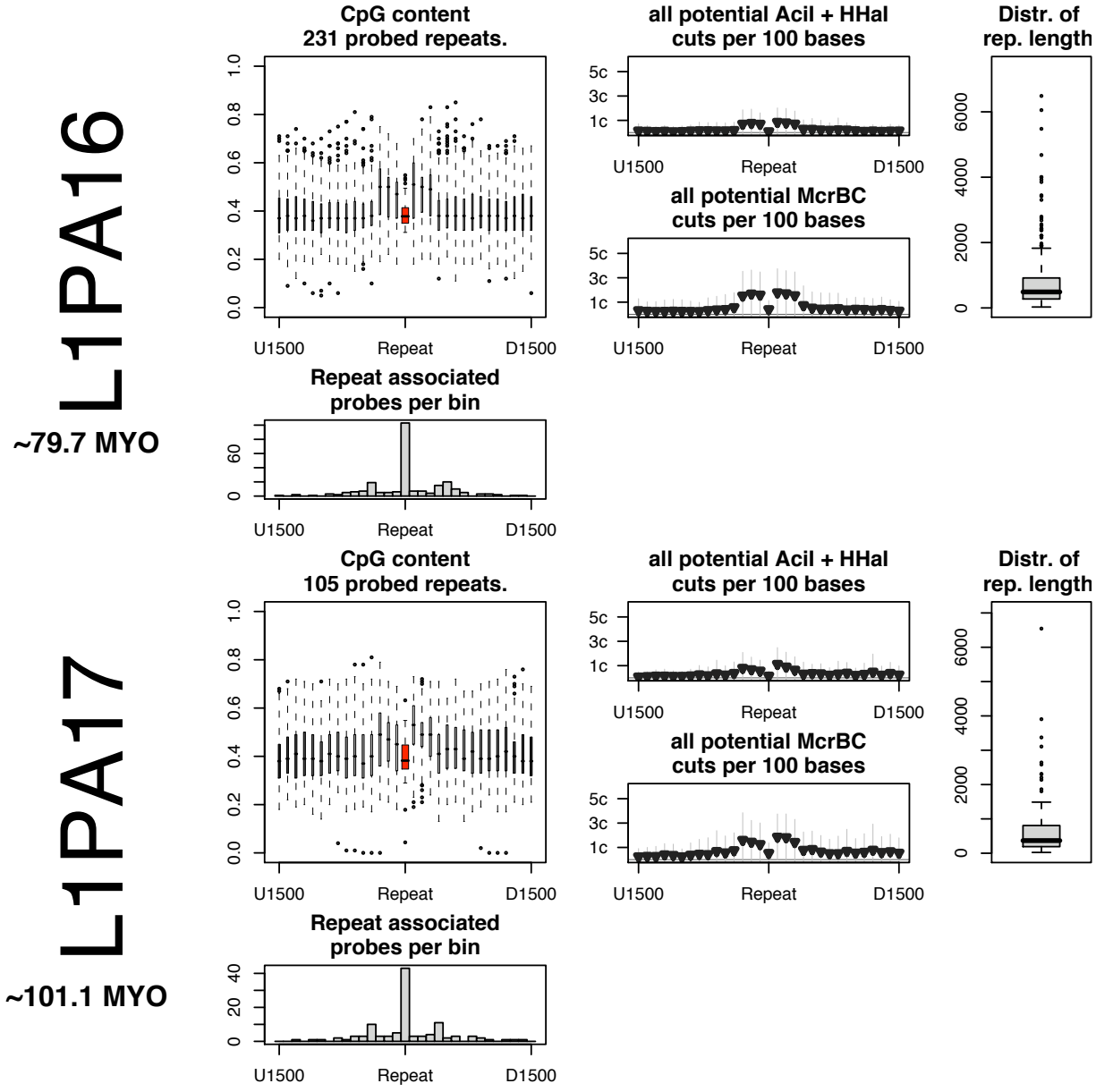

L1PA16
~79.7 MYO
CpG content
231 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
L1PA17
~101.1 MYO
CpG content
105 probed repeats.
all potential AciI + HHaI
cuts per 100 bases
all potential McrBC
cuts per 100 bases
Distr. of
rep. length
Repeat associated
probes per bin
U1500
Repeat
D1500

Figure S17

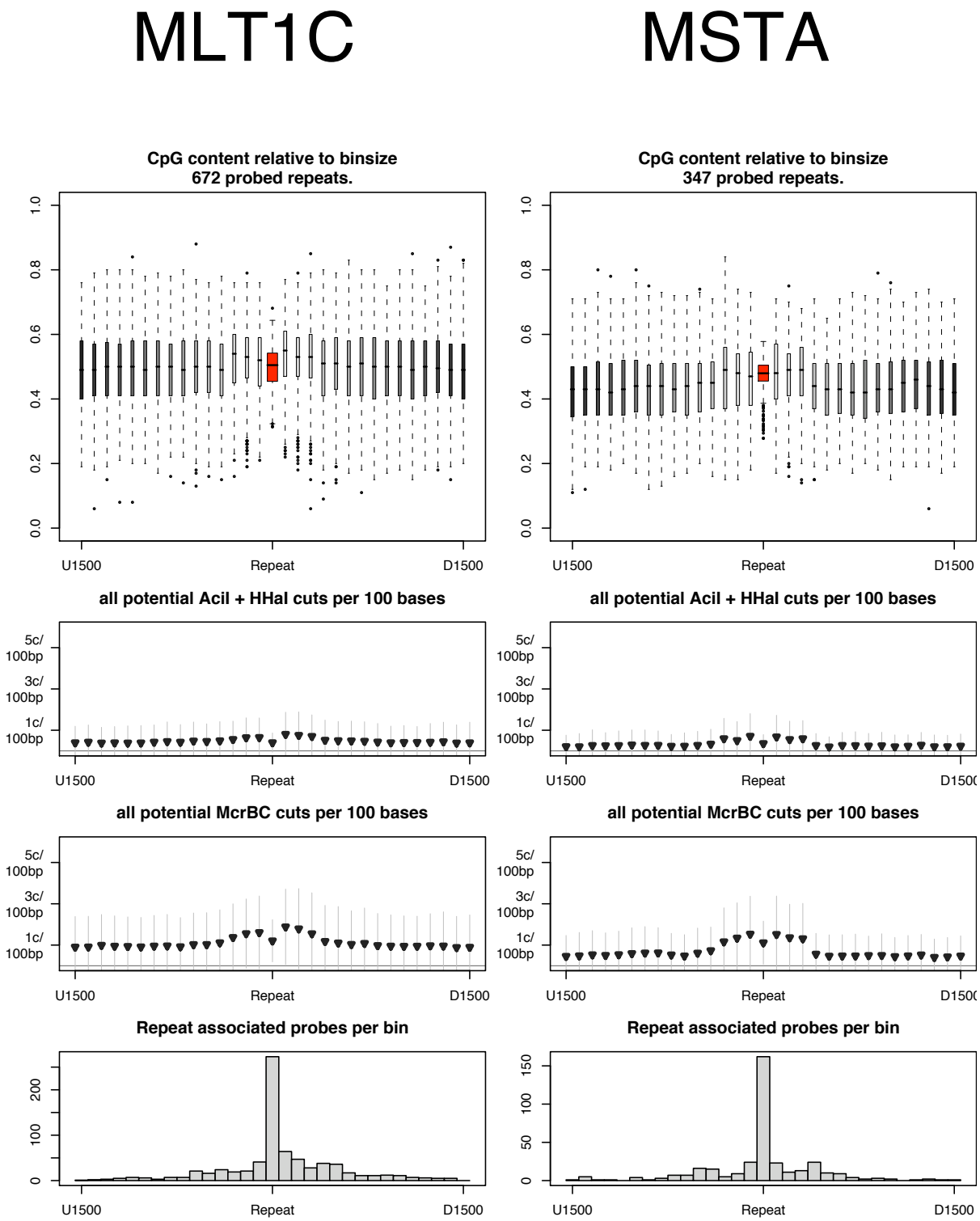

MLT1C
MSTA
CpG content relative to binsize
672 probed repeats.
CpG content relative to binsize
347 probed repeats.
0.0
0.2
0.4
0.6
0.8
1.0
U1500
Repeat
D1500
all potential AciI + HHaI cuts per 100 bases
5c/ 100bp
3c/ 100bp
1c/ 100bp
all potential McrBC cuts per 100 bases
Repeat associated probes per bin
0
100
200
50
150

Figure S18

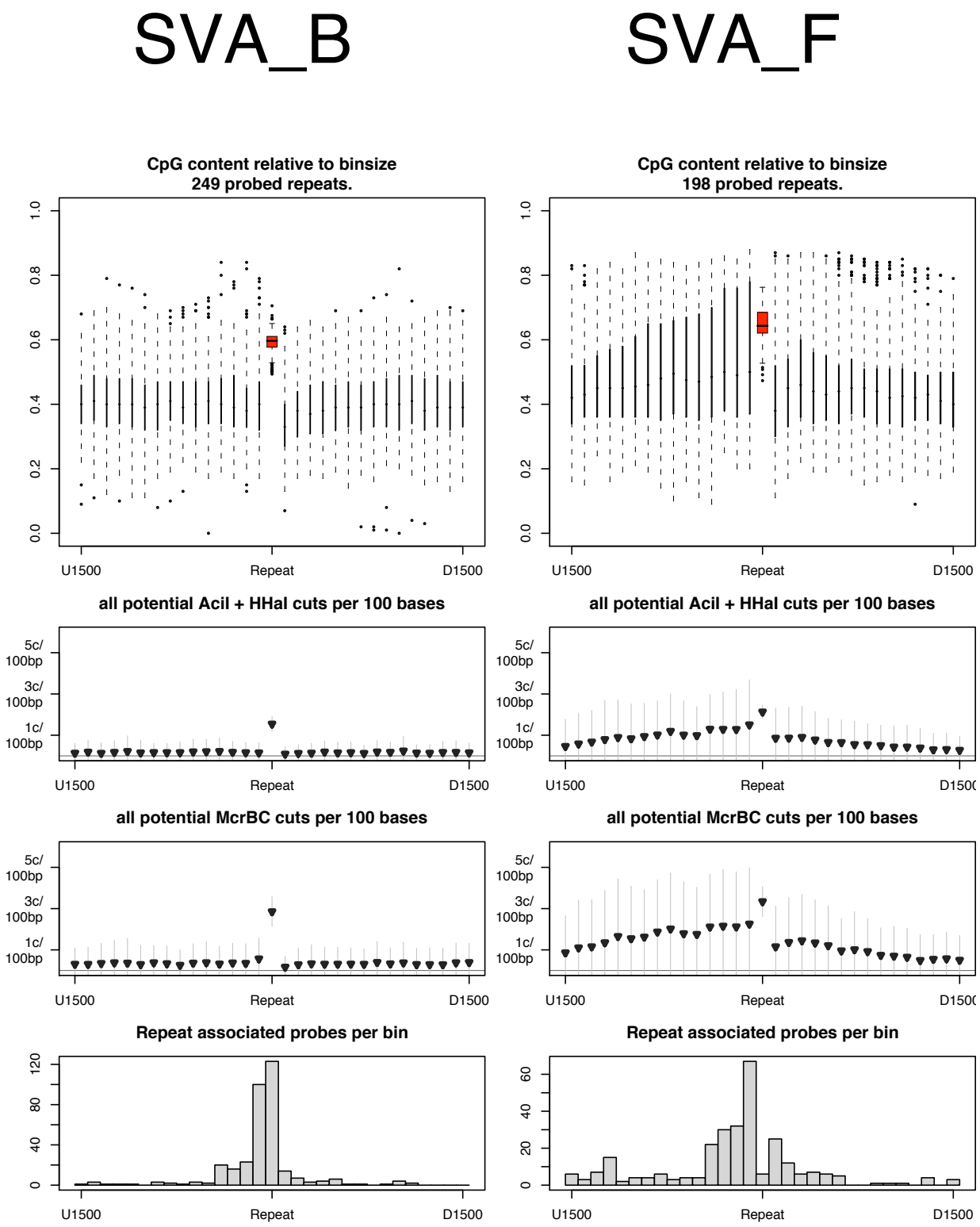

SVA_B
SVA_F
CpG content relative to binsize
249 probed repeats.
CpG content relative to binsize
198 probed repeats.
0.0
0.2
0.4
0.6
0.8
1.0
U1500
Repeat
D1500
all potential AciI + HHaI cuts per 100 bases
5c/
100bp
3c/
100bp
1c/
100bp
all potential McrBC cuts per 100 bases
Repeat associated probes per bin
0
40
80
120
20
60

Figure S19

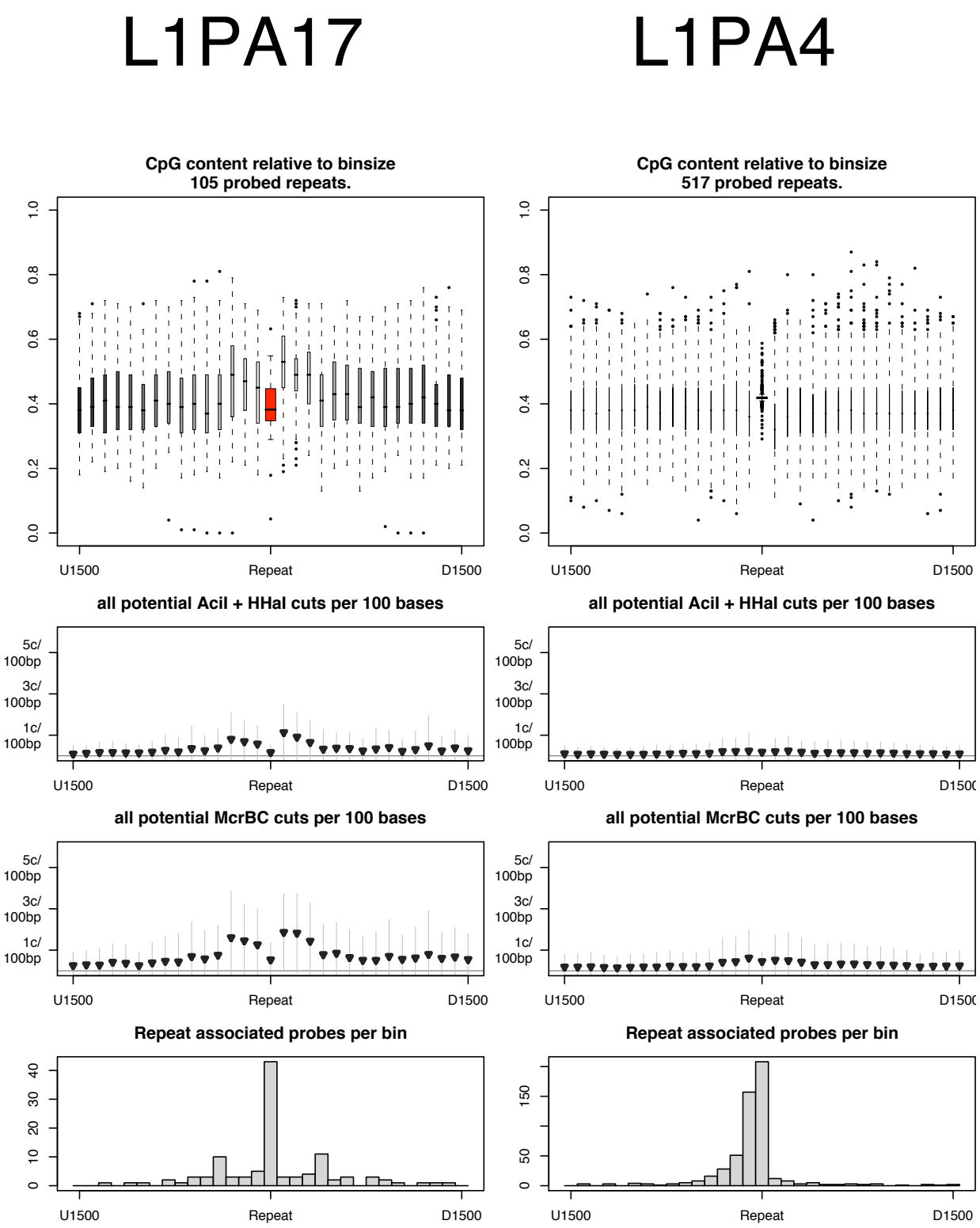

L1PA17
L1PA4
CpG content relative to binsize
105 probed repeats.
CpG content relative to binsize
517 probed repeats.
0.0
0.2
0.4
0.6
0.8
1.0
U1500
Repeat
D1500
all potential AciI + HHaI cuts per 100 bases
5c/
100bp
3c/
100bp
1c/
100bp
all potential McrBC cuts per 100 bases
Repeat associated probes per bin
0
10
20
30
40
50
150

Figure S20

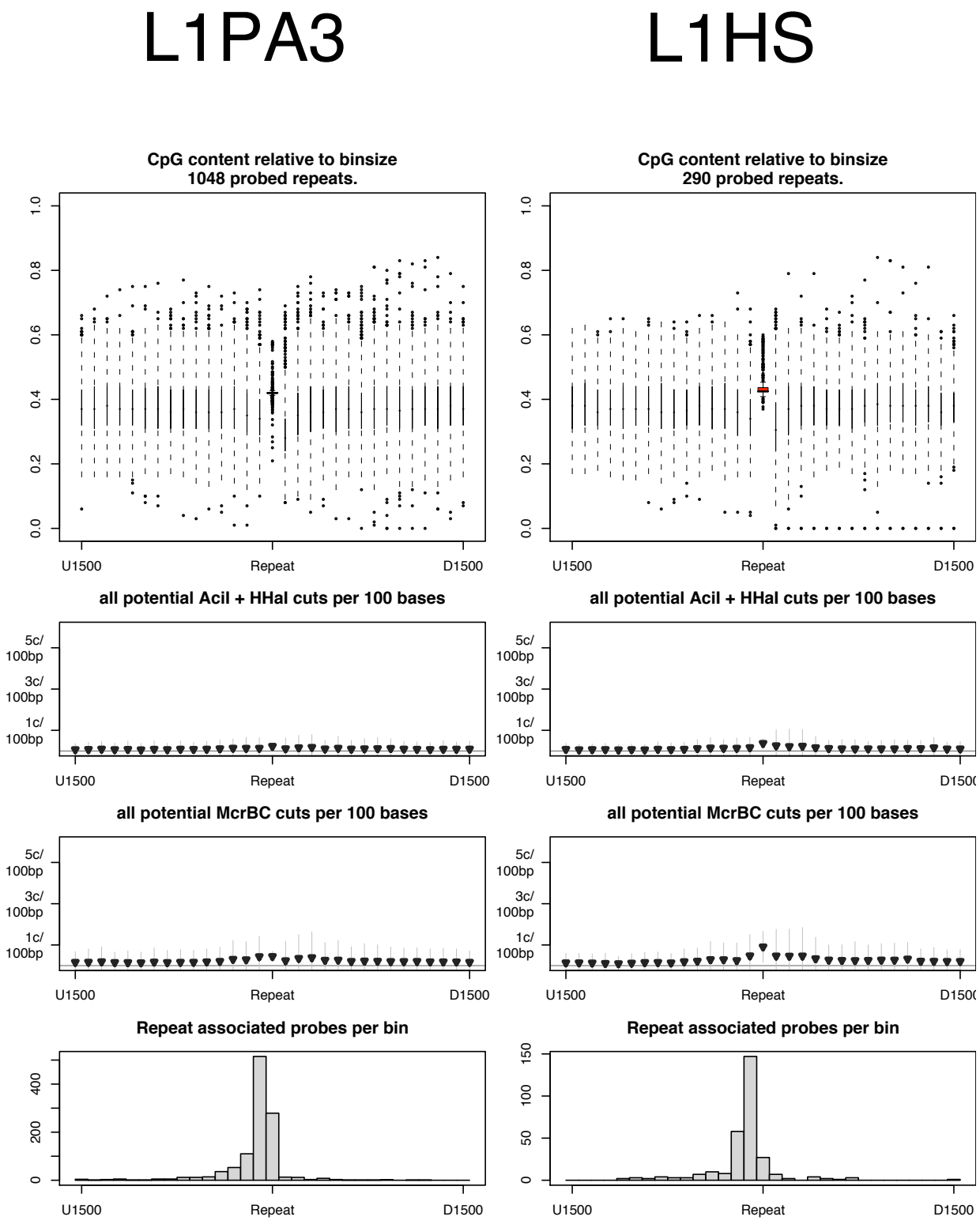

L1PA3
L1HS
CpG content relative to binsize
1048 probed repeats.
CpG content relative to binsize
290 probed repeats.
0.0
0.2
0.4
0.6
0.8
1.0
U1500
Repeat
D1500
all potential AciI + HHaI cuts per 100 bases
5c/ 100bp
3c/ 100bp
1c/ 100bp
all potential McrBC cuts per 100 bases
Repeat associated probes per bin
0
200
400
50
100
150

Figure S21

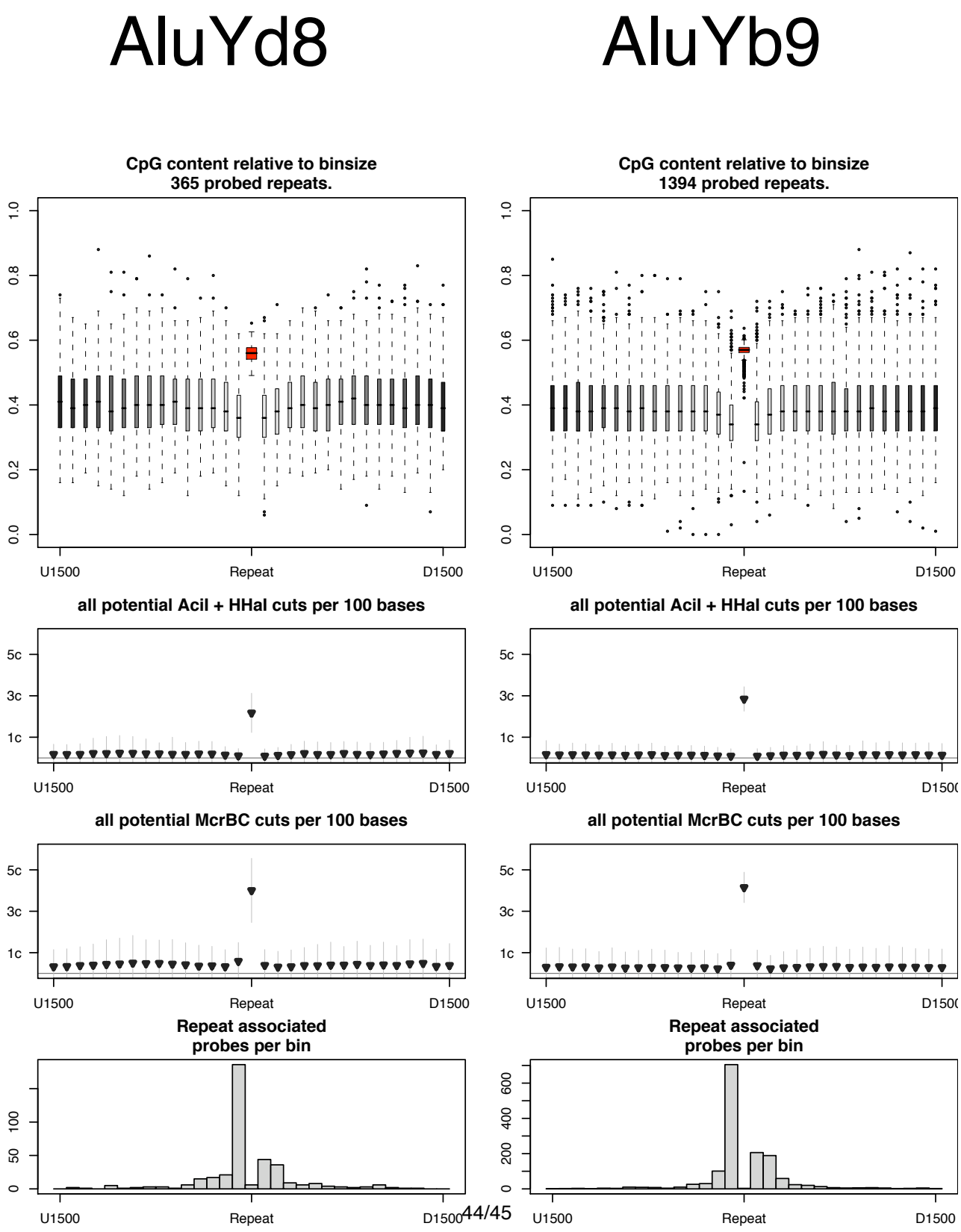
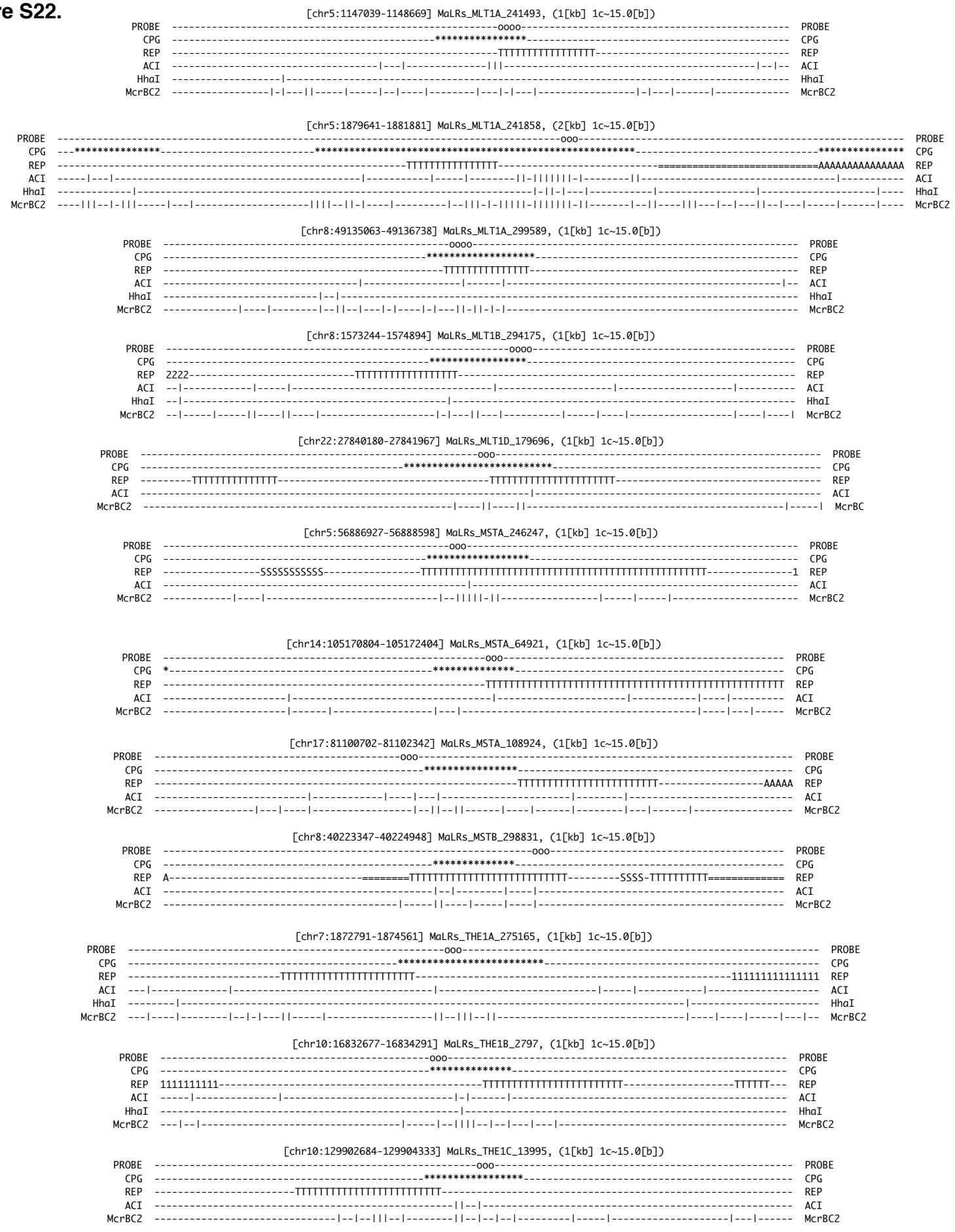

AluYd8
AluYb9
CpG content relative to binsize
365 probed repeats.
CpG content relative to binsize
1394 probed repeats.
0.0
0.2
0.4
0.6
0.8
1.0
U1500
Repeat
D1500
all potential AciI + HHaI cuts per 100 bases
1c
3c
5c
all potential McrBC cuts per 100 bases
Repeat associated
probes per bin
0
50
100
200
400
600

Figure S22.


[chr5:1147039-1148669] MaLRs_MLT1A_241493, (1[kb] 1c~15.0[b])
[chr5:1879641-1881881] MaLRs_MLT1A_241858, (2[kb] 1c~15.0[b])
[chr8:49135063-49136738] MaLRs_MLT1A_299589, (1[kb] 1c~15.0[b])
[chr8:1573244-1574894] MaLRs_MLT1B_294175, (1[kb] 1c~15.0[b])
[chr22:27840180-27841967] MaLRs_MLT1D_179696, (1[kb] 1c~15.0[b])
[chr5:56886927-56888598] MaLRs_MSTA_246247, (1[kb] 1c~15.0[b])
[chr14:105170804-105172404] MaLRs_MSTA_64921, (1[kb] 1c~15.0[b])
[chr17:81100702-81102342] MaLRs_MSTA_108924, (1[kb] 1c~15.0[b])
[chr8:40223347-40224948] MaLRs_MSTB_298831, (1[kb] 1c~15.0[b])
[chr7:1872791-1874561] MaLRs_THE1A_275165, (1[kb] 1c~15.0[b])
[chr10:16832677-16834291] MaLRs_THE1B_2797, (1[kb] 1c~15.0[b])
[chr10:129902684-129904333] MaLRs_THE1C_13995, (1[kb] 1c~15.0[b])
PROBE
CPG
REP
ACI
HhaI
McrBC2