Supplemental Material

Supplemental data: chimpanzee, baboon, macaque, marmoset and lemur BAC end sequences and quality files and human-chimpanzee, human-baboon, human-marmoset and human-lemur genomic sequence alignments can be found at http://bfgl.anri.barc.usda.gov/Alusite.

Table S1. Quality properties of BAC end sequences

Species	Lemur	Marmoset	Baboon	Macaque	Chimpanzee
Sequences	6,533	5,173	7,303	5,504	5,969
Total Length	3,798,199	3,825,700	3,670,302	2,873,380	2,784,861
Q20 Length	3,226,168	3,455,512	3,082,773	2,403,177	2,291,611
Q25 Length	2,924,161	3,285,240	2,735,360	2,165,983	2,004,188
Q30 Length	2,451,304	3,076,576	2,206,983	1,817,114	1,607,733
Q35 Length	1,302,038	2,867,752	1,132,322	982,631	828,244
Q40 Length	15,205	2,315,339	15,353	10696	10,467
Average Length	581.39	739.55	502.57	522.05	466.55
Average Q20 Length	493.83	667.99	422.12	436.62	383.92
Average Q25 Length	447.60	635.07	374.55	393.53	335.77
Average Q30 Length	375.22	594.74	302.20	330.14	269.35
Average Q35 Length	199.30	554.37	155.05	178.53	138.76
Average Q40 Length	2.33	447.58	2.10	1.94	1.75

Table S2. Summary of 9 newly identified lineage-specific *Alu* subfamilies

Assigned Names	Sequence Count	Figure	Label	Mutations
*Alu*MacYa4	43	Figure S5	BES_NJ_BM1	205 G->A
*Alu*MacYa4	42	Figure 2	12 BES_MS_BM1	205 G->A
*Alu*Ta14	16	Figure S7 A	18 ALN_MS_R1	100 T->A, 212 G->A, 257 C->T, 272 G->C
*Alu*L	55	Figure 4 B	15	
*Alu*L5	57	Figure 4 B	16	70 G->C, 101 A->C, 118 A->T, 222 A->T, 231 T->G
*Alu*L6	96	Figure 4 B	17	101 A->C, 118 A->T, 222 A->T, 231 T->G, 233 A->C, 243 T->G
*Alu*L9	63	Figure 4 B	18	52 A->T, 64 C->T 70 G->C, 101 A->C, 118 A->T, 222 A->T, 231 T->G, 233 A->C, 243 T->G
*Alu*La	76	Figure 4 B	19	
*Alu*La7a	18	Figure 4 B	21	36 A->C, 96-101 AAGAGC->del,
*Alu*La7b	19	Figure 4 B	20	36-37 AC->del, 152 A->ins, 161 C->ins, 162 G->A, 187 G->T, 249 T->C

ALN: genomic sequence alignments; BES: BAC end sequences; NJ: neighboring joining; MS: minimum spanning tree; ins: insertion and del: deletion.

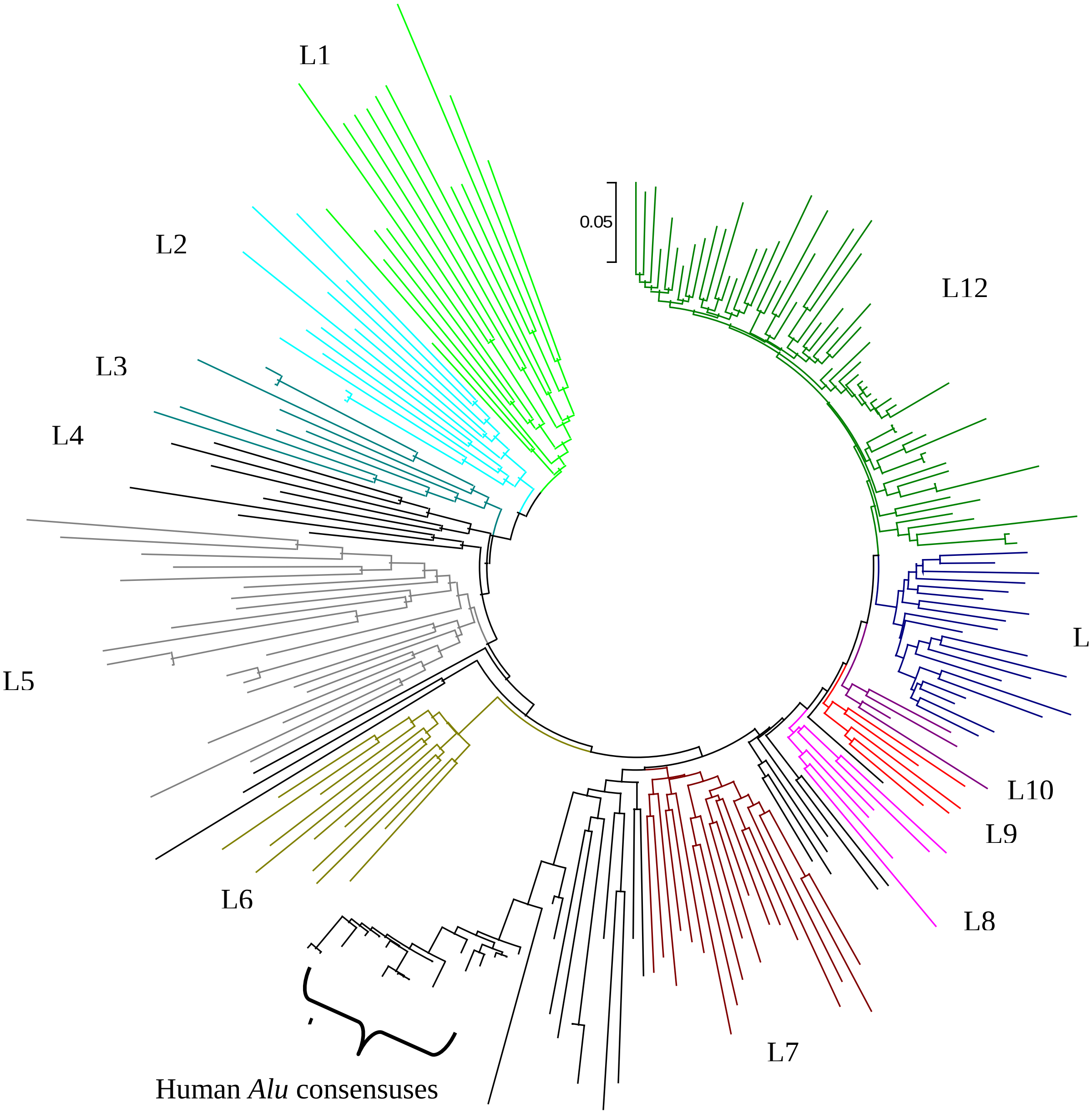


Figure S3. Neighbor-joining tree of 210 *Alu* elements from lemur BAC end sequences.
This tree includes 210 lemur *Alu* elements were extracted from BES. Known human consensus sequences for *Alu* subfamilies were added for comparison. Different colors represent putative *Alu* subfamilies/groups: BES_NJ_L1-12.

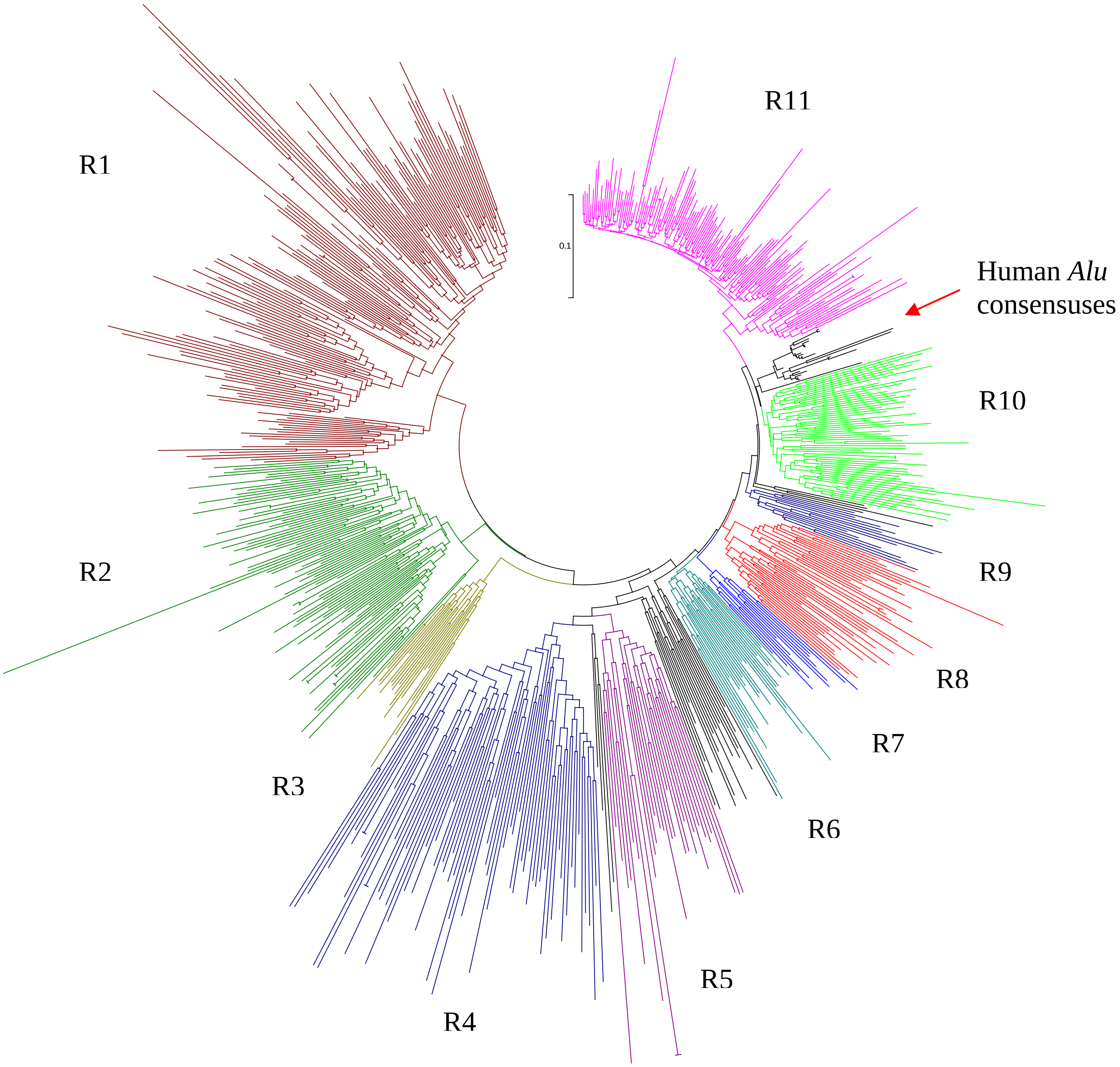


Figure S4. Neighbor-joining tree of 718 *Alu* elements from marmoset BAC end sequences. This tree includes 718 marmoset *Alu* elements were extracted from BES. Known human consensus sequences for *Alu* subfamilies were added to provide an orientation. Different colors represent putative *Alu* subfamilies/groups: BES_NJ_R1-11.

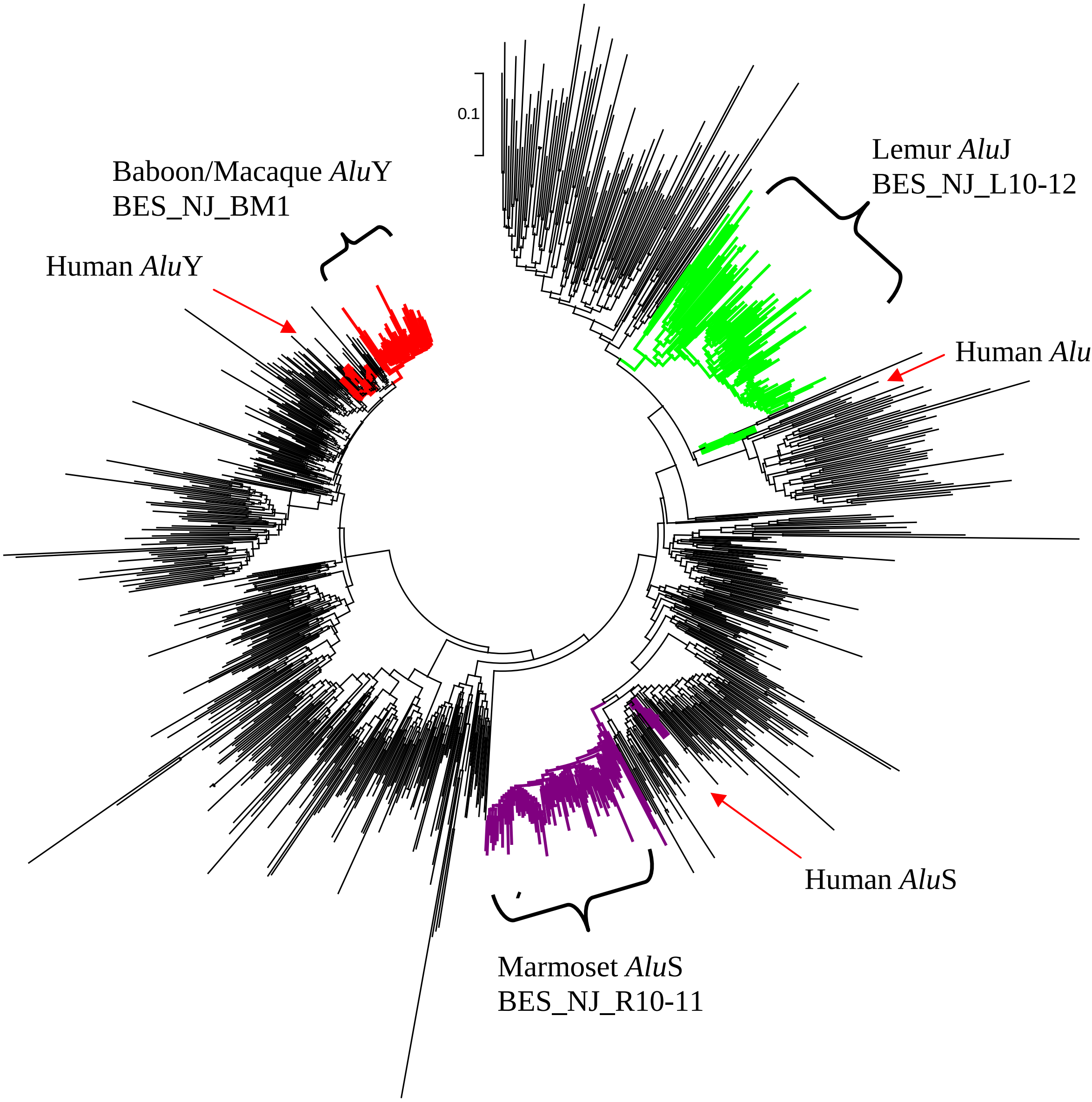


Figure S5. Neighbor-joining tree of selected *Alu* elements from primate BAC end sequences.
This tree includes selected *Alu* repeats from Figure S1 and S2 (BES_NJ_L10-12, BES_NJ_R10-11 and BES_NJ_BM1). Additional selected *Alu* elements were added from primate BES. Totally this tree contains 980 *Alu* elements, including human 150, chimpanzee 150, baboon 289, and macaque 92, marmoset 181 and lemur 118. Known human consensus sequences for *Alu* subfamilies were added to provide an orientation. All lineages with brackets were supported by the bootstrap values >50% with n = 1,000 replicates.

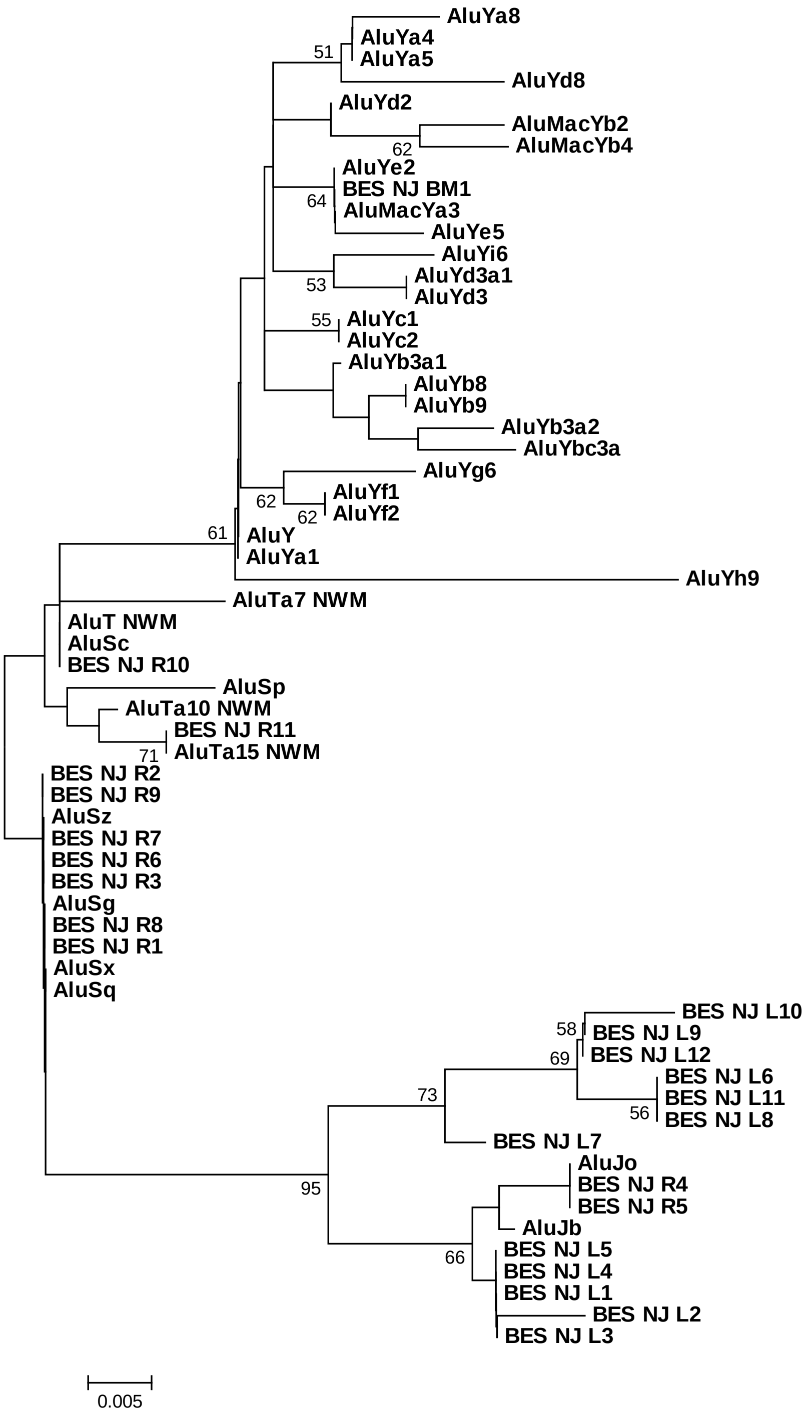


Figure S6. Neighbor-joining tree of *Alu* consensus sequences derived from BAC end sequences. We used the NJ trees (Figure S1 and S2) to group related *Alu* elements. *Alu* consensus sequences were derived from each group, including lemur (BES_NJ_L1-12), marmoset (BES_NJ_R1-11) and baboon/macaque (BES_NJ_BM1). We then constructed a NJ tree using all derived *Alu* consensus sequences with known primate *Alu* consensus sequences. All lineages are labeled when the bootstrap values >50% with n = 1,000 replicates.

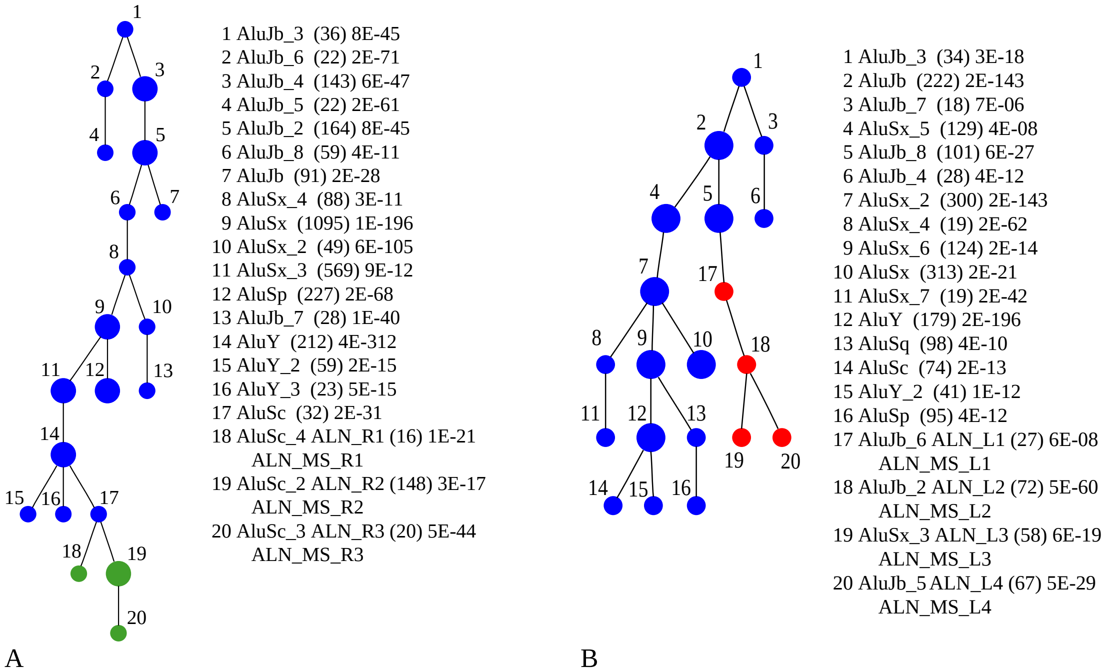
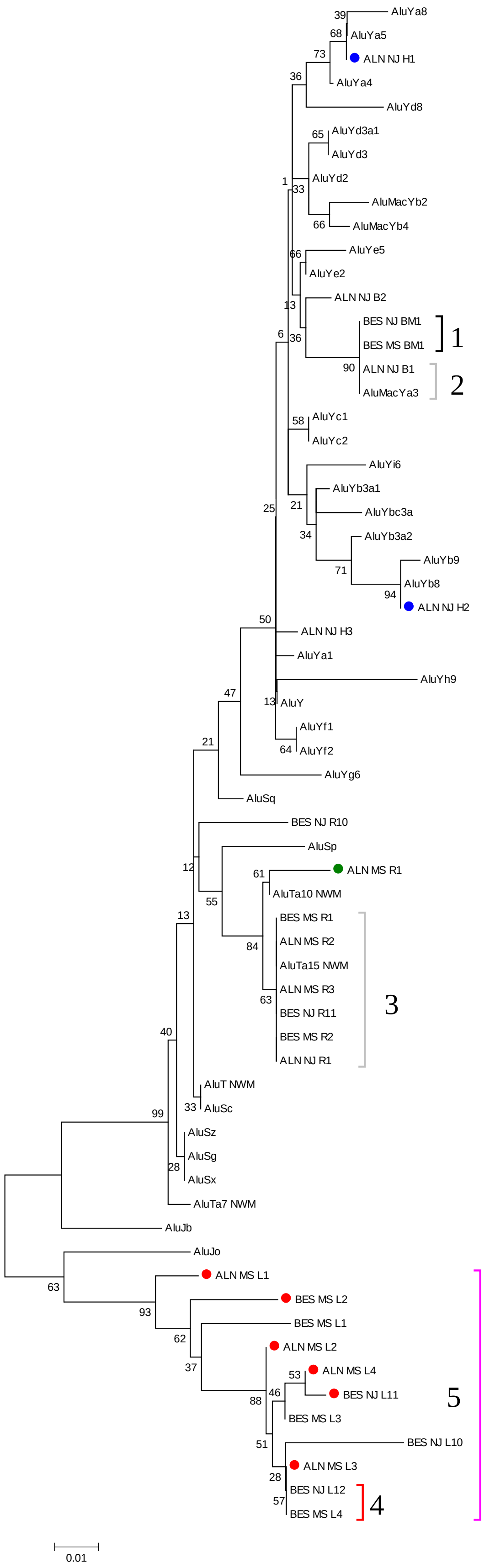
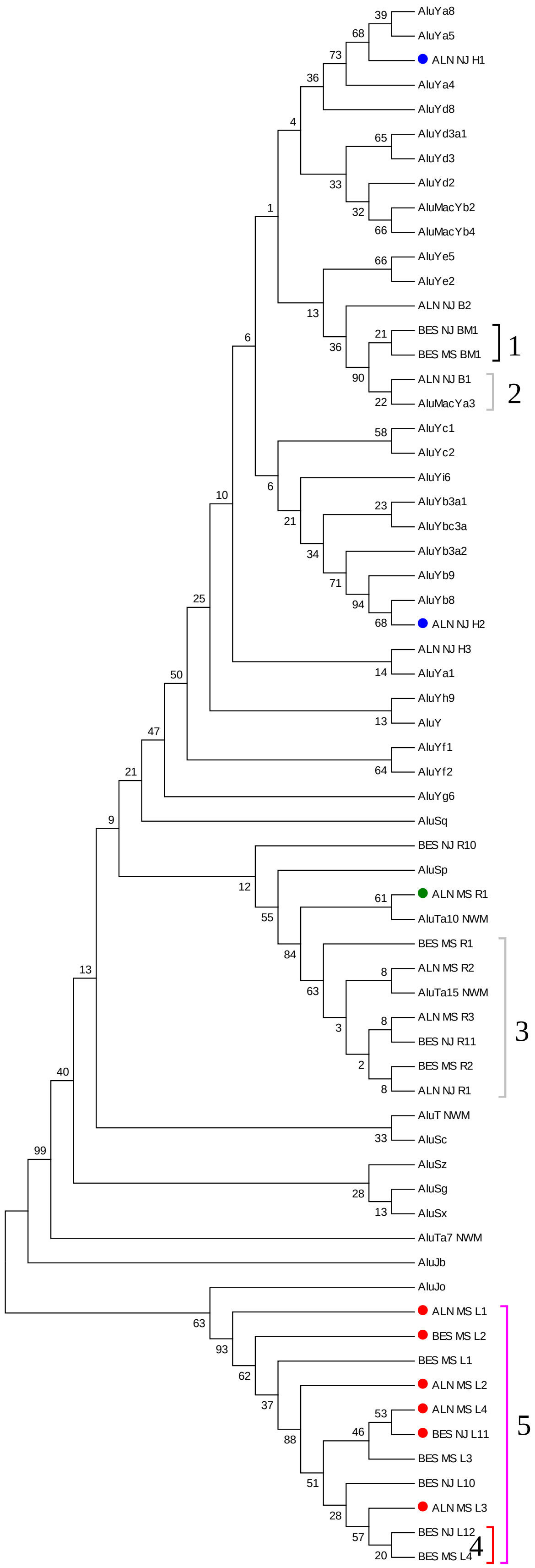


Figure S7. Minimum spanning trees of the *Alu* subfamilies identified from genomic sequence alignments. *Alu* subfamilies belong to or shared with human are labeled blue. A. Human-marmoset comparison: 3 marmoset specific subfamilies (Nodes 18-20, ALN_MS_R1-3) are labeled green and B. Human-lemur comparison: 4 lemur specific subfamilies (Nodes 17-20, ALN_MS_L1-4) are labeled red.



39
AluYa8
68
AluYa5
73
● ALN NJ H1
36
AluYa4
AluYd8
65
AluYd3a1
1
AluYd3
33
AluYd2
AluMacYb2
66
AluMacYb4
66
AluYe5
AluYe2
13
ALN NJ B2
6
36
BES NJ BM1
BES MS BM1
1
90
ALN NJ B1
AluMacYa3
2
58
AluYc1
AluYc2
AluYi6
AluYb3a1
25
21
AluYbc3a
34
AluYb3a2
AluYb9
71
AluYb8
94
● ALN NJ H2
50
ALN NJ H3
AluYa1
AluYh9
13
AluY
AluYf1
64
AluYf2
47
AluYg6
21
AluSq
BES NJ R10
AluSp
12
61
● ALN MS R1
55
AluTa10 NWM
BES MS R1
13
84
ALN MS R2
AluTa15 NWM
ALN MS R3
63
BES NJ R11
3
BES MS R2
40
ALN NJ R1
AluT NWM
99
33
AluSc
AluSz
AluSg
28
AluSx
AluTa7 NWM
AluJb
AluJo
● ALN MS L1
63
● BES MS L2
93
BES MS L1
62
● ALN MS L2
37
53
● ALN MS L4
46
● BES NJ L11
5
88
BES MS L3
51
BES NJ L10
● ALN MS L3
28
BES NJ L12
57
BES MS L4
4
0.01



39
AluYa8
68
AluYa5
73
● ALN NJ H1
36
AluYa4
4
AluYd8
65
AluYd3a1
33
AluYd3
1
AluYd2
32
AluMacYb2
66
AluMacYb4
66
AluYe5
AluYe2
13
ALN NJ B2
6
21
BES NJ BM1
36
BES MS BM1
1
90
ALN NJ B1
22
AluMacYa3
2
58
AluYc1
AluYc2
10
6
AluYi6
21
23
AluYb3a1
AluYbc3a
34
AluYb3a2
25
71
AluYb9
94
AluYb8
68
● ALN NJ H2
50
ALN NJ H3
14
AluYa1
47
AluYh9
13
AluY
21
AluYf1
64
AluYf2
AluYg6
9
AluSq
BES NJ R10
12
AluSp
61
● ALN MS R1
55
AluTa10 NWM
84
BES MS R1
8
ALN MS R2
63
AluTa15 NWM
13
3
8
ALN MS R3
BES NJ R11
3
2
BES MS R2
8
ALN NJ R1
40
AluT NWM
33
AluSc
99
AluSz
28
AluSg
13
AluSx
AluTa7 NWM
AluJb
AluJo
63
● ALN MS L1
93
● BES MS L2
62
BES MS L1
37
● ALN MS L2
53
● ALN MS L4
88
46
● BES NJ L11
5
BES MS L3
51
BES NJ L10
28
● ALN MS L3
57
BES NJ L12
20
BES MS L4
4

Figure S8. Neighbor-joining trees of primate lineage-specific *Alu* elements.
On the left is the NJ tree in the branch style. It is the same as Figure 4 A. On the right is a topology version of this tree. All branches are labeled with dots and brackets when the bootstrap values > 50% with n=1,000 replicates: human: blue; baboon/macaque: black; marmoset: green and lemur: red and pink. The gray brackets (2 and 3) and two blue dots (ALN_NJ_H1, ALN_NJ_H2) indicate our results confirm earlier published consensus sequences. All other brackets and dots are putative new *Alu* subfamilies we identified.

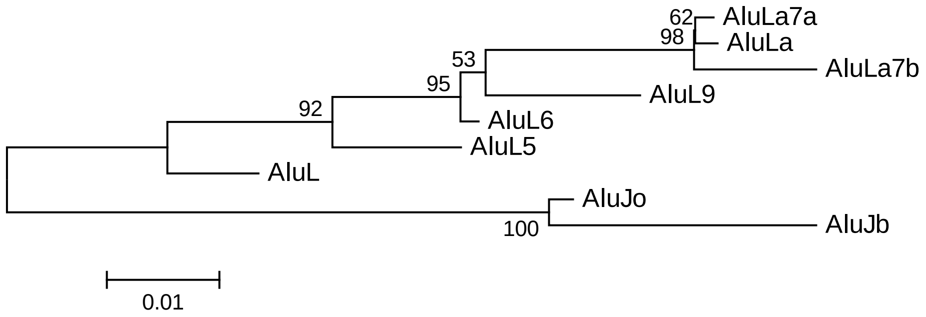


Figure S9. Neighbor-joining trees of 7 lemur lineage-specific *Alu* elements.
Human *Alu*Jo and *Alu*Jb consensus sequences were added to provide an orientation. All branches are labeled when the bootstrap values > 50% with n=1,000 replicates. Note: The topology of the phylogenetic tree lacks bootstrap support due to the small number of changes between examined consensus sequences.

Figure S10. Aligned consensus sequences of putative lemur *Alu* subfamilies.

```
                     10        20        30        40        50        60
             ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo        GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGAGGA
AluJb        ............................................................
ALN_MS_L1    ........A......................T......C.........A...........
ALN_MS_L2    ...............................T......C.....................
ALN_MS_L3    ...............................T......C.................T...
ALN_MS_L4    ...............................T......C.....................
BES_NJ_L10   ....T..T.T..........A..........T......C..............A......
BES_NJ_L11   ....R....Y.....................T......C.................T...
BES_NJ_L12   ...............................T......C.................T...
BES_MS_L1    ....................A..........T......C........T............
BES_MS_L2    ........A......................T......C.....................
BES_MS_L3    ...............................T......C.....................
BES_MS_L4    ...............................T......C.................T...

                     70        80        90       100       110       120
             ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo        TTGCTTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTACA
AluJb        .CA.....................................G.T..A..............
ALN_MS_L1    .C.......GT.....................A....G.--.T........A........
ALN_MS_L2    .C.......GT.....................A....G.--..................T
ALN_MS_L3    .C...C...GT.....................A....G.--..................T
ALN_MS_L4    .C.T......T.....................A....G.--..................T
BES_NJ_L10   .C.......GT...K.....R...........A....G.--........T.R.......T
BES_NJ_L11   .C.T......T.....................A....G.--..................T
BES_NJ_L12   .C...C...GT.....................A....G.--..................T
BES_MS_L1    .C.......GT.....................A....G.--..........A.......T
BES_MS_L2    .C........T.........A...........A....G.--..................T
BES_MS_L3    .C........T.....................A....G.--..................T
BES_MS_L4    .C...C...GT.....................A....G.--..................T

                    130       140       150       160       170       180
             ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo        AAAAATACAAA--------------------------------------AATTAGCCGGG
AluJb        ...........--------------------------------------...........
ALN_MS_L1    .......G...--------------------------------------...........
ALN_MS_L2    .......G...--------------------------------------...........
ALN_MS_L3    .......G...GAAATTATCTGGCCAACTAAAAATATATATAC-AAAAA...........
ALN_MS_L4    .......G...--------------------------------------...........
BES_NJ_L10   .......G...GAAATTATCTGGACAGCTAAAATTATATATATAGAAAA.......Y...
BES_NJ_L11   .......G...GAAATTAGCTGGACAACTAAAAHTATATATA--GAAAA...........
BES_NJ_L12   .......G...GAAATTATCTGGCCAACTAAAA-TATATATA--GAAAA...........
BES_MS_L1    .......G...--------------------------------------...........
BES_MS_L2    .......G...--------------------------------------........A..
BES_MS_L3    .......G...--------------------------------------...........
BES_MS_L4    .......G...GAAATTATCTGGCCAACTAAAA-TATATATA--GAAAA...........
```

```
                  190       200       210       220       230       240
           ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo      CGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTG
AluJb      ............................................................
ALN_MS_L1  .A..........................................................
ALN_MS_L2  .A.........AT...............................................
ALN_MS_L3  .A.........AT...................................T..........A
ALN_MS_L4  .A.........AT...............................................
BES_NJ_L10 .A.........AT...................T...............A.....Y....R
BES_NJ_L11 .A.........AT...................................A.....T.....
BES_NJ_L12 .A.........AT...................................T..........A
BES_MS_L1  .A......T..AT...............................................
BES_MS_L2  T..........AT..........................................A....
BES_MS_L3  .A.........AT...............................................
BES_MS_L4  .A.........AT...................................T..........A

                  250       260       270       280       290       300
           ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo      AGCCCAGGAGTTCGAGGCTGCAGTGAGCTATGATCGCGCCACTGCACT----CCAGCCTG
AluJb      .....G....G.................CG..................----........
ALN_MS_L1  ............T....T................GAT...........CT----....C.
ALN_MS_L2  ............T....T...T........G.C.GA......G.....CT----....C.
ALN_MS_L3  ............T....T...T........G.C.GA......G.....CAN-N-NAG.CC
ALN_MS_L4  ............T....T...T........G.C.GA......G.....CT----....C.
BES_NJ_L10 ............T....T...T........G.C.GA...Y..G.....CT----....A.
BES_NJ_L11 ............T....T...T........G.C.GA......G.....CT----....Y.
BES_NJ_L12 ............T....T...T........G.C.GA......G.....CACT.T....C.
BES_MS_L1  ............T....T................GA............CT----....A.
BES_MS_L2  ............T....T................GAT...........CT----....CA
BES_MS_L3  ............T....T...T........G.C.GA......G.....CT----......
BES_MS_L4  ............T....T...T........G.C.GA......G.....CA---A....C.

                  310       320       330
           ....|....|....|....|....|....|
AluJo      GGC-GACAGAGCGAGACCCTGTCTCAAAAA
AluJb      ...-..........................
ALN_MS_L1  ...-A.......A....T............
ALN_MS_L2  ...-A......T.....T............
ALN_MS_L3  ..GCA...A..T.....T............
ALN_MS_L4  ...-A......T.....T............
BES_NJ_L10 ...-A......Y.....T............
BES_NJ_L11 ...-A......T.....T............
BES_NJ_L12 ...-A...A..T.....T............
BES_MS_L1  ...-A.......A....T............
BES_MS_L2  ...-A............T............
BES_MS_L3  ...-A......T.....T............
BES_MS_L4  ...-A...A..T.....T............
```

Figure S11. Aligned consensus sequences of lemur *Alu*L subfamilies.

```
                   10        20        30        40        50        60
          ....|....|....|....|....|....|....|....|....|....|....|....|
ALUL      GGCCGGGCACGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGA
ALUL5     ........G...................................................
ALUL6     ........G...................................................
ALUL9     ........G...............................................T...

                   70        80        90       100       110       120
          ....|....|....|....|....|....|....|....|....|....|....|....|
ALUL      TCGCTTGAGGTCAGGAGTTCGAGACCAGCCTGAGCAAGAGAGAGACCCCGTCTCTACAAA
ALUL5     .........C..............................C................T..
ALUL6     ........................................C................T..
ALUL9     ...T.....C..............................C................T..

                  130       140       150       160       170       180
          ....|....|....|....|....|....|....|....|....|....|....|....|
ALUL      AAATAGAAAAATTAGCCGGGCGTGGTGGCGCACGCCTGTAGTCCCAGCTACTCGGGAGGC
ALUL5     .................A...A......T...T...........................
ALUL6     .....................A..........T...........................
ALUL9     .....................A.......A..T...........................

                  190       200       210       220       230       240
          ....|....|....|....|....|....|....|....|....|....|....|....|
ALUL      TGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCAGTGAGCTATGATGATGCC
ALUL5     .........................................T........G.........
ALUL6     .........................................T........G.C...C...
ALUL9     ..............T..........................T........G.C...C...

                  250       260       270       280
          ....|....|....|....|....|....|....|....|....|
ALUL      ACTGCACTCTAGCCCGGGCAACAGAGTGAGACTCTGTCTCAAAAA
ALUL5     ..........................C..................
ALUL6     ..G..........................................
ALUL9     ..G..........................................
```

Figure S12. Aligned consensus sequences of lemur *Alu*La subfamilies.

```
                     10        20        30        40        50        60
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluLa       GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGA
AluLa7a     ...................................C........................
AluLa7b     ...................................--.......................

                     70        80        90       100       110       120
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluLa       TCGCTCGAGGTCAGGAGTTCGAGACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAA
AluLa7a     ...................................------...................
AluLa7b     ........................................T...................

                    130       140       150       160       170       180
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluLa       AAATAGAAAGAAATTATCTGGCCAACTAAAA-TATATATA-GAAAAAATTAGCCGGGCAT
AluLa7a     ...............................-........-...................
AluLa7b     ...............................A........CA..................

                    190       200       210       220       230       240
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluLa       GGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGC
AluLa7a     ............................................................
AluLa7b     ......T.....................................................

                    250       260       270       280       290       300
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluLa       CCAGGAGTTTGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTCAAA--GCCCGGGCA
AluLa7a     ................................................NNN.........
AluLa7b     ........C........................T..............NNN.........

                    310       320
            ....|....|....|....|....|
AluLa       ACAAAGCGAGACTCTGTCTCAAAAA
AluLa7a     ......T..................
AluLa7b     ......T..................
```

Figure S13. Aligned consensus sequences of putative marmoset *Alu* subfamilies.

```
                        10        20        30        40        50        60
                ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo           ........................................................A...
AluJb           ........................................................A...
AluSx           GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGA
AluSp           ............................................................
AluSq           ........................................................T...
AluT_NWM        ............................................................
AluTa7_NWM      ............................................................
AluTa10_NWM     ........................R...............................T...
AluTa15_NWM     ........................................................T...
ALN_NJ_R1       ...................A....................................T...
ALN_MS_R1       ........................A...............................T...
ALN_MS_R2       ...................A....................................T...
ALN_MS_R3       ...................A....................................T...
BES_NJ_R2       ...................M....................................T...
BES_NJ_R1       R.YYR...RYR.....................................R...Y...YA..
BES_MS_R1       ........................................................T...
BES_MS_R2       ...................A....................................T...
AluY            ............................................................

                        70        80        90       100       110       120
                ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo           .TG.T...CC.......................G......A.C..G.............A
AluJb           ....T...CC.......................G.........................A
AluSx           TCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT
AluSp           ............G...................A.........A.................
AluSq           ............................................................
AluT_NWM        ....--.......A...A.........T................................
AluTa7_NWM      ....--.......A...A.........T................................
AluTa10_NWM     ....--.......A...A.........T.....T..........................
AluTa15_NWM     ....--.......A...A.........T.....T..........................
ALN_NJ_R1       ....--.......A...A.........T.....T..........................
ALN_MS_R1       ....--.......A...A.........T.....T.....A....................
ALN_MS_R2       ....--.......A...A.........T.....T..........................
ALN_MS_R3       ....--.......AC..A.........T.....T..........................
BES_NJ_R2       ....--.......A...A.........T.....T..........................
BES_NJ_R1       ...T--.......R...A.........T................................
BES_MS_R1       ....--.......A...A.........T.....T..........................
BES_MS_R2       ....--.......AC..A.........T.....T..........................
AluY            ....--...........A.........T......T....C....................

                       130       140       150       160       170       180
                ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo           ...........--...............................G...............
AluJb           ...........--...............................G...............
AluSx           AAAAATACAAA--AATTAGCCGGGCGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGGG
AluSp           ...........--......................AT.......................
AluSq           ...........--.....................G.........................
AluT_NWM        ...........--.......T.......................G...............
AluTa7_NWM      ...........AA.......T.......................G...............
AluTa10_NWM     ...........--.......T....A..................G...............
AluTa15_NWM     ...........-A.......T....A..........T....................A..
ALN_NJ_R1       ...........-A.......T....A..........T....................A..
ALN_MS_R1       ...........--.......T....A......T...........G...............
ALN_MS_R2       ...........-A.......T....A..........T....................A..
ALN_MS_R3       ...........-A.......T....A..........T....................A..
BES_NJ_R2       .....H.....-A.......T....A..........T.......R............A..
BES_NJ_R1       ...........-A.......T....A.......A..Y.......G............R..
BES_MS_R1       ...........-A.......T....A..........T....................A..
BES_MS_R2       ...........--.......T....A..........T....................A..
AluY            ...........-A.....................G.........G...............
```

```
                        190       200       210       220       230       240
                ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo           ...............G.........G...A....TTC....C..........TAT.....
AluJb           ...............G.........G.........TC....C............T.....
AluSx           AGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCGAGATCG
AluSp           .............................................G..............
AluSq           ............................................................
AluT_NWM        ..................T.........................................
AluTa7_NWM      ..................T.....................................G...
AluTa10_NWM     ..................T..........A...............G..............
AluTa15_NWM     ..................T..C.......A...............G..............
ALN_NJ_R1       ..................T..C.......A...............G..............
ALN_MS_R1       ..................T..........A.A.............G..............
ALN_MS_R2       ..................T..C.......A...............G..............
ALN_MS_R3       ..................T..C.......A...............G..............
BES_NJ_R2       ..................T..C.......A...............G..............
BES_NJ_R1       ...................R.........A......R.......................
BES_MS_R1       ..................T..C.......A...............G..............
BES_MS_R2       ..................T..C.......A...............G..............
AluY            ..................G..G..................C...................

                        250       260       270       280       290
                ....|....|....|....|....|....|....|....|....|....|....|.
AluJo           .......................------....-.........C.T..........
AluJb           .......................------....-.........C.T..........
AluSx           CGCCACTGCACTCCAGCCTGGGC------GACA-GAGCGAGACTCCGTCTCAAAAA
AluSp           .....T.................------A...A......A...............
AluSq           .......................------A...A......A...............
AluT_NWM        .....................CG-------...-......................
AluTa7_NWM      .....................CGCCTGGT...G-...T.......T..........
AluTa10_NWM     .....T................T---AACA.G.-.--...A...............
AluTa15_NWM     .....T................T---AACA.G.-.--...A...............
ALN_NJ_R1       .....T................T---AACA.G.-.--...A...............
ALN_MS_R1       T....T..........T.....T---AACA...-.--...A...............
ALN_MS_R2       .....T................T---AACA.G.-.--...A...............
ALN_MS_R3       .....T................T---AACANNN-......A...............
BES_NJ_R2       Y....T................T---AACA.G.-.--...A...............
BES_NJ_R1       .....................SY---RACA-G.-.--.A.......R.........
BES_MS_R1       .....T................T---AACA.G.-.--...A...............
BES_MS_R2       .....T................T---AACA.G.-.--...A...............
AluY            .......................------....-......................
```

Figure S14. **Aligned consensus sequences of putative baboon/macaque *Alu* subfamilies.**

```
                    10        20        30        40        50        60
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo       ........................................................A...
AluSx       ............................................................
AluY        GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGA
AluYa1      ............................................................
ALN_NJ_H3   ........................................................T...
AluYa5      ............................................................
AluYa8      ............................................................
ALN_NJ_H1   ............................................................
AluYb9      ........................................................T...
AluYb8      ........................................................T...
ALN_NJ_H2   ........................................................T...
AluMacYb2   ............................................................
AluMacYa3   ...................A...............................A........
AluMacYb4   ...................A........................................
ALN_NJ_B1   ...................A...............................A........
ALN_NJ_B2   ...................A....................................T...
BES_NJ_BM1  ...................A...............................A........
BES_MS_BM1  ...................A...............................A........

                    70        80        90       100       110       120
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo       .TG.TT...CC......T.........G.....GC....TA.C..G.-............
AluSx       ....CT...........T.........G......C....T.......-............
AluY        TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAA-CCCCGTCTCTAC
AluYa1      ....--.........................................-............
ALN_NJ_H3   ....--.........................................A....A.......
AluYa5      ....--........................C.....A.........-............
AluYa8      ....--........................C.....A.........-............
ALN_NJ_H1   ....--........................C.....A.........-............
AluYb9      ...T--.................................A.......-............
AluYb8      ...T--.................................A.......-............
ALN_NJ_H2   ...T--.................................A.......-............
AluMacYb2   ....--A....................------------........-............
AluMacYa3   ....--.........................................-............
AluMacYb4   ....--A....................------------.A......-............
ALN_NJ_B1   ....--.........................................-............
ALN_NJ_B2   ....--.................................T.......-............
BES_NJ_BM1  ....--.........................................-............
BES_MS_BM1  ....--.........................................-............

                   130       140       150       160       170       180
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo       A.....-......-.....................C........................
AluSx       ......-......-.....................C.........A..............
AluY        TAAAAA-TACAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG
AluYa1      ......-.....................................................
ALN_NJ_H3   ......A.....................................................
AluYa5      ......-.....................A............................T..
AluYa8      .....C-..........-..........A....................T.......T..
ALN_NJ_H1   ......-.....................A............................T..
AluYb9      ......-....................C.............................G..
AluYb8      ......-....................C................................
ALN_NJ_H2   ......-....................C................................
AluMacYb2   ......-....................C...T.T..........................
AluMacYa3   ......A.........C..........A................................
AluMacYb4   ......-....................C..............................A.
ALN_NJ_B1   ......A.........C..........A................................
ALN_NJ_B2   ......-.........C...........................................
BES_NJ_BM1  ......A.........C..........A................................
BES_MS_BM1  ......A.........C..........A................................
```

```
                    190       200       210       220       230       240
            ....|....|....|....|....|....|....|....|....|....|....|....|
AluJo       ................G..C..T...G...A....TTC...GC..........TAT....
AluSx       ...................C..T..................G..................
AluY        GAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATC
AluYa1      ............................................................
ALN_NJ_H3   ..............................A......A......................
AluYa5      ............................................................
AluYa8      ............................................................
ALN_NJ_H1   ............................................................
AluYb9      ..................................A........................T
AluYb8      ..................................A........................T
ALN_NJ_H2   ..................................A........................T
AluMacYb2   ............................................................
AluMacYa3   .....................................................T......
AluMacYb4   .......................G....................................
ALN_NJ_B1   .....................................................B......
ALN_NJ_B2   -..........G................................................
BES_NJ_BM1  ........................A............................T......
BES_MS_BM1  ........................A............................T......

                    250       260       270       280       290
            ....|....|....|....|....|....|....|....|....|....|....|..
AluJo       .............-------........................C.T..........
AluSx       .............-------.....................................
AluY        GCGCCACTGCACT-------CCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
AluYa1      C............-------.....................................
ALN_NJ_H3   A............-------.....................................
AluYa5      C............-------.....................................
AluYa8      C............-------.....................................
ALN_NJ_H1   C............-------.....................................
AluYb9      ...........G.CCGCAGT..G..................................
AluYb8      ...........G.CCGCAGT..G..................................
ALN_NJ_H2   ...........G.CCGCAGT..G..................................
AluMacYb2   .............-------.....................................
AluMacYa3   CG...........-------.....................................
AluMacYb4   .............-------...........A....C.T..................
ALN_NJ_B1   CG...........-------..........-.G........................
ALN_NJ_B2   .............-------..........H....C.....................
BES_NJ_BM1  CG...........-------.....................................
BES_MS_BM1  CG...........-------.....................................
```

File S15. Fasta files of 9 lineage-specific *Alu* consensus sequences and 26 putative *Alu* subfamilies.

```
#9 newly identified Alu consensus sequences
#1 baboon/macaque Alu consensus sequence
>AluMacYa4
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACGGGCGGATCACGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAACTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCC
CAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTAAACCCGGGAGGCGGAGCTTGCAGTGAGCTGAGATCCGGCCACTGCAC
TCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
#1 marmoset Alu consensus sequence
>AluTa14
GGCCGGGCGCGGTGGCTCACGCCTATAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAAGAGATCGAGAC
CATCCTGGTCAACAAGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGTGCGCGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCAGAAGGCGGAGGTTGCGGTGAGCCGAGATCGTGCCATTGCACTC
CAGTCTGGGTAACAACAGCGAAACTCCGTCTCAAAAA
#7 new lemur Alu consensus sequences
>AluL
GGCCGGGCACGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGCTTGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGAGAGACCCCGTCTCTACAAAAAATAGAAAAATTAGCCGGGCGTGGTGGCGCACGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCAGTGAGCTATGATGATGCCACTGCACTC
TAGCCCGGGCAACAGAGTGAGACTCTGTCTCAAAAA
>AluL6
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGCTTGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTC
TAGCCCGGGCAACAGAGTGAGACTCTGTCTCAAAAA
>AluL5
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGCTTGAGCTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCAGGCATGGTGGTGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCTGTGAGCTAGGATGATGCCACTGCACTC
TAGCCCGGGCAACAGAGCGAGACTCTGTCTCAAAAA
>AluL9
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGATCGTTTGAGCTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCGGGCATGGTGGCACATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATTGCTTGAGCCCAGGAGTTTGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTC
TAGCCCGGGCAACAGAGTGAGACTCTGTCTCAAAAA
>AluLa
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGATCGCTCGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTATCTGGCCAACTAAAATATATATAGAAAAAA
TTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGCCCAGGAGTTTG
AGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTCAAAGCCCGGGCAACAAAGCGAGACTCTGTCTCAAAAA
>AluLa8
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCCCTCTGGGAGGCCGAGGCGGGTGGATCGCTCGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTATCTGGCCAACTAAAATATATATAGAAAAAATTAGCC
GGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGCCCAGGAGTTTGAGGTTG
CTGTGAGCTAGGCTGACGCCACGGCACTCANNNAGCCCGGGCAACAAAGTGAGACTCTGTCTCAAAAA
>AluLa7b
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCTCTGGGAGGCCGAGGCGGGTGGATCGCTCGAGGTCAGGAGTTCGAGAC
CAGCCTGAGCAAGAGTGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTATCTGGCCAACTAAAAATATATATACAAAAAAA
TTAGCCGGGCATGGTGGCTCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGCCCAGGAGTCTG
AGGTTGCTGTGAGCTAGGCTGATGCCACGGCACTCANNNAGCCCGGGCAACAAAGTGAGACTCTGTCTCAAAAA

#26 putative Alu subfamilies
>ALN_NJ_H1
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGTCAGGAGATCGAGAC
CATCCCGGCTAAAACGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGTAGTGGCGGGCGCCTGTAGTCCC
AGCTACTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCCCGCCACTGCACT
CCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
>ALN_NJ_H2
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCATGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACAAGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCC
AGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAAGCGGAGCTTGCAGTGAGCCGAGATTGCGCCACTGCAGT
CCGCAGTCCGGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
>ALN_NJ_H3
```

```
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACACGGTGAAAACCCCATCTCTACTAAAAAATACAAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGTAGT
CCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCAGGAGGCAGAGCTTGCAGTGAGCCGAGATCACGCCACTGC
ACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
>ALN_NJ_B1
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACGGGCGGATCACGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAACTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCC
CAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCBGAGATCCGGCCACTGCAC
TCCAGCCTGGGGGCAGAGCGAGACTCCGTCTCAAAAA
>ALN_NJ_B2
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAACTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCC
AGCTACTCGGAGGCTGAGGCGGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACTGCACTC
CAGCCTGGGHGACACAGCGAGACTCCGTCTCAAAAA
>BES_NJ_BM1
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACGGGCGGATCACGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAACTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCC
CAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTAAACCCGGGAGGCGGAGCTTGCAGTGAGCTGAGATCCGGCCACTGCAC
TCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
>BES_MS_BM1
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACGGGCGGATCACGAGGTCAGGAGATCGAGAC
CATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAACTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCC
CAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTAAACCCGGGAGGCGGAGCTTGCAGTGAGCTGAGATCCGGCCACTGCAC
TCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA
>ALN_NJ_R1
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAAGAGATCGAGAC
CATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCTGGGCATGGTGGCGCGTGCCTGTAATCCC
AGCTACTCAGGAGGCTGAGGCAGGAGAATTGCCTGAACCCAGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGCCATTGCACT
CCAGCCTGGGTAACAAGAGCGAAACTCCGTCTCAAAAA
>ALN_MS_R1
GGCCGGGCGCGGTGGCTCACGCCTATAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAAGAGATCGAGAC
CATCCTGGTCAACAAGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGTGCGCGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCAGAAGGCGGAGGTTGCGGTGAGCCGAGATCGTGCCATTGCACTC
CAGTCTGGGTAACAACAGCGAAACTCCGTCTCAAAAA
>ALN_MS_R2
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAAGAGATCGAGAC
CATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCTGGGCATGGTGGCGCGTGCCTGTAATCCC
AGCTACTCAGGAGGCTGAGGCAGGAGAATTGCCTGAACCCAGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGCCATTGCACT
CCAGCCTGGGTAACAAGAGCGAAACTCCGTCTCAAAAA
>ALN_MS_R3
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAACAGATCGAGAC
CATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCTGGGCATGGTGGCGCGTGCCTGTAATCCC
AGCTACTCAGGAGGCTGAGGCAGGAGAATTGCCTGAACCCAGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGCCATTGCACT
CCAGCCTGGGTAACANNNGAGCGAAACTCCGTCTCAAAAA
>BES_NJ_R10
RGYYRGGCRYRGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCRAGGYGGGYAGATCATGAGGTCARGAGATCGAGAC
CATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCTGGGCATGGTGGCACGYGCCTGTAGTCCC
AGCTACTCRGGAGGCTGAGGCAGGAGAATCRCTTGAACCCAGGAGGCRGAGGTTGCAGTGAGCCGAGATCGCGCCACTGCACT
CCAGCCTGGSYRACAGAGCAAGACTCCRTCTCAAAAA
>BES_NJ_R11
GGCCGGGCGCGGTGGCTCAMGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAAGAGATCGAGAC
CATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAAAAAHACAAAAAATTAGCTGGGCATGGTGGCGCGTGCCTGTARTCCC
AGCTACTCAGGAGGCTGAGGCAGGAGAATTGCCTGAACCCAGGAGGCGGAGGTTGCGGTGAGCCGAGATCGYGCCATTGCACT
CCAGCCTGGGTAACAAGAGCGAAACTCCGTCTCAAAAA
>BES_MS_R1
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAAGAGATCGAGAC
CATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCTGGGCATGGTGGCGCGTGCCTGTAATCCC
AGCTACTCAGGAGGCTGAGGCAGGAGAATTGCCTGAACCCAGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGCCATTGCACT
CCAGCCTGGGTAACAAGAGCGAAACTCCGTCTCAAAAA
>BES_MS_R2
GGCCGGGCGCGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAACAGATCGAGAC
CATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCGCGTGCCTGTAATCCCA
```

```
GCTACTCAGGAGGCTGAGGCAGGAGAATTGCCTGAACCCAGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGCCATTGCACTC
CAGCCTGGGTAACAAGAGCGAAACTCCGTCTCAAAAA
>ALN_MS_L1
GGCCGGGCACGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCAAGGCGGGAGGATCGCTTGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGTGAGACCCCATCTCTACAAAAAATAGAAAAATTAGCCGGGCATGGTGGCGCGCGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCAGTGAGCTATGATGATGCCACTGCACTC
TAGCCCGGGCAACAGAGCAAGACTCTGTCTCAAAAA
>ALN_MS_L2
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGCTTGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTC
TAGCCCGGGCAACAGAGTGAGACTCTGTCTCAAAAA
>ALN_MS_L3
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGATCGCTCGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTATCTGGCCAACTAAAAATATATATACAAAAA
AATTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGCCCAGGAGTT
TGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTCANNNAGCCCGGGCAACAAAGTGAGACTCTGTCTCAAAAA
>ALN_MS_L4
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGTTTGAGCTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTC
TAGCCCGGGCAACAGAGTGAGACTCTGTCTCAAAAA
>BES_NJ_L10
GGCCTGGTGTGGTGGCTCACACCTGTAATCCTAGCACTCTGGGAGGCCGAGGCAGGAGGATCGCTTGAGGTCAGKAGTTCRAG
ACCAGCCTGAGCAAGAGCGAGACCTCRTCTCTACTAAAAATAGAAAGAAATTATCTGGACAGCTAAAATTATATATATAGAAA
AAATTAGCYGGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGAAGGATYGCTTRAGCCCAGGAGT
TTGAGGTTGCTGTGAGCTAGGCTGACGCYACGGCACTCTAGCCAGGGCAACAGAGYGAGACTCTGTCTCAAAAA
>BES_NJ_L11
GGCCRGGCGYGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGATCGTTTGAGCTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTAGCTGGACAACTAAAAHTATATATAGAAAAA
ATTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGAAGGATTGCTTGAGCCCAGGAGTTT
GAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTCTAGCCYGGGCAACAGAGTGAGACTCTGTCTCAAAAA
>BES_NJ_L12
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGATCGCTCGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTATCTGGCCAACTAAAATATATATAGAAAAAA
TTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGCCCAGGAGTTTG
AGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTCACTCTAGCCCGGGCAACAAAGTGAGACTCTGTCTCAAAAA
>BES_MS_L1
GGCCGGGCGCGGTGGCTCACACCTGTAATCCTAGCACTCTGGGAGGCTGAGGCGGGAGGATCGCTTGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCATCTCTACTAAAAATAGAAAAATTAGCCGGGCATGGTGGTGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCAGTGAGCTATGATGACGCCACTGCACTC
TAGCCAGGGCAACAGAGCAAGACTCTGTCTCAAAAA
>BES_MS_L2
GGCCGGGCACGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGCTTGAGCTCAGGAGTTCAAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCAGGTGTGGTGGCGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCACTTGAGCCCAGGAGTTTGAGGTTGCAGTGAGCTATGATGATGCCACTGCACTC
TAGCCCAGGCAACAGAGCGAGACTCTGTCTCAAAA
>BES_MS_L4
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGTGGATCGCTCGAGGTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAGAAATTATCTGGCCAACTAAAATATATATAGAAAAAA
TTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGTAGGATCGCTTAAGCCCAGGAGTTTG
AGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTCAAAGCCCGGGCAACAAAGTGAGACTCTGTCTCAAAAA
>BES_MS_L3
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTCTGGGAGGCCGAGGCGGGAGGATCGCTTGAGCTCAGGAGTTCGAG
ACCAGCCTGAGCAAGAGCGAGACCCCGTCTCTACTAAAAATAGAAAAATTAGCCGGGCATGGTGGCGCATGCCTGTAGTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTGAGGTTGCTGTGAGCTAGGCTGACGCCACGGCACTC
TAGCCTGGGCAACAGAGTGAGACTCTGTCTCAAAAA
```