

Supplementary Figure 1: Annotated Sequence of the chicken Z chromosome.

Chart of the Z chromosome sequence showing protein-coding genes, CpG islands, non-coding transcripts, intrachromosomal similarity, G+C content, interspersed repeats, and sequenced clones. All sequence features, BACs, and fosmids are drawn to scale. (a) The positions of all predicted protein-coding genes of the chicken Z chromosome are shown in black. Plus (+) strand above, minus (-) strand below. CpG islands, as predicted by cpghplot, are shown in grey. (b) The positions of non-coding transcripts of the chicken Z chromosome are shown in black. Plus (+) strand above, minus (-) strand below. (c) Intrachromosomal similarity is plotted in windows of 5kb every 1kb, using the percent identity of the best blast hit from a repeat-masked non-overlapping window. All values greater than 90% are shown. (d) G+C content is plotted in sliding windows of 10kb every 5kb. (e) Interspersed repeat content calculated from the RepeatMasker .out file. Repeat content is expressed as percentage of nucleotides in windows of 100kb every 5 kb. DNA transposons are plotted in red, LINE elements are plotted in green, Endogenous Retroviruses (ERV) are plotted in blue. (f) BAC and fosmid clones from the finished sequence of the chicken Z chromosome. 5 physical gaps in the clone map are indicated by gaps in clone coverage. The remaining gaps are spanned by clones (shown in grey) whose sequences were not finished at the time of sequence assembly (June 2009).

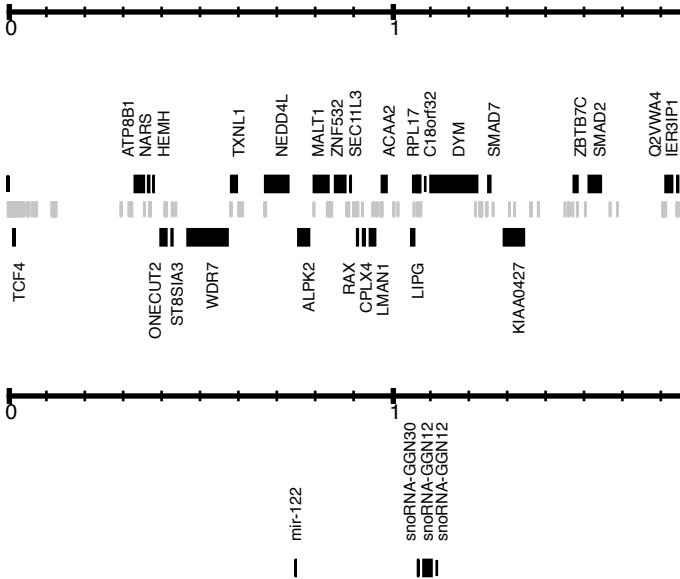
Points of interest along the chicken Z chromosome sequence include:

The Male Hypermethylated Region (MHM), located at 26Mb. This region is especially high in endogenous retroviruses (ERVs) and G+C content. The tandem array of the non-coding MHM transcript is located in this region.

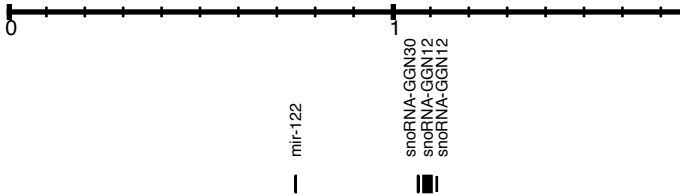
The Z amplicon, located from 73 Mb to the end of the chromosome. This region of high intrachromosomal similarity is low in interspersed repeats and elevated in G+C content. The Z amplicon is interrupted by two single-copy islands at 81 Mb and 83 Mb, which contain protein coding genes with orthologs on human autosomes 5 and 9.

Scale (1Mb)

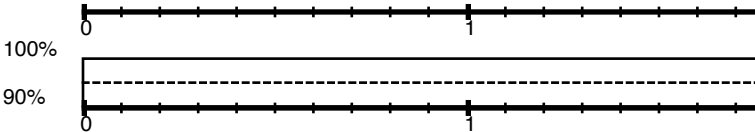
a Coding genes +
CpG islands
Coding genes -



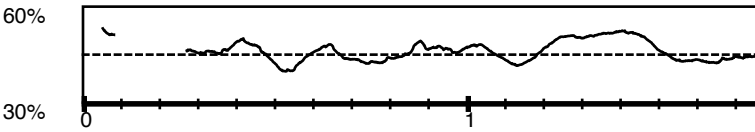
b Non-coding transcripts +
Non-coding transcripts -



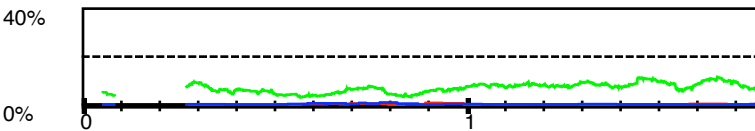
c Intrachromosomal
similarity (% identity)



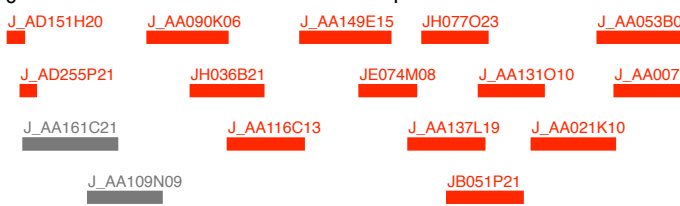
d G+C content



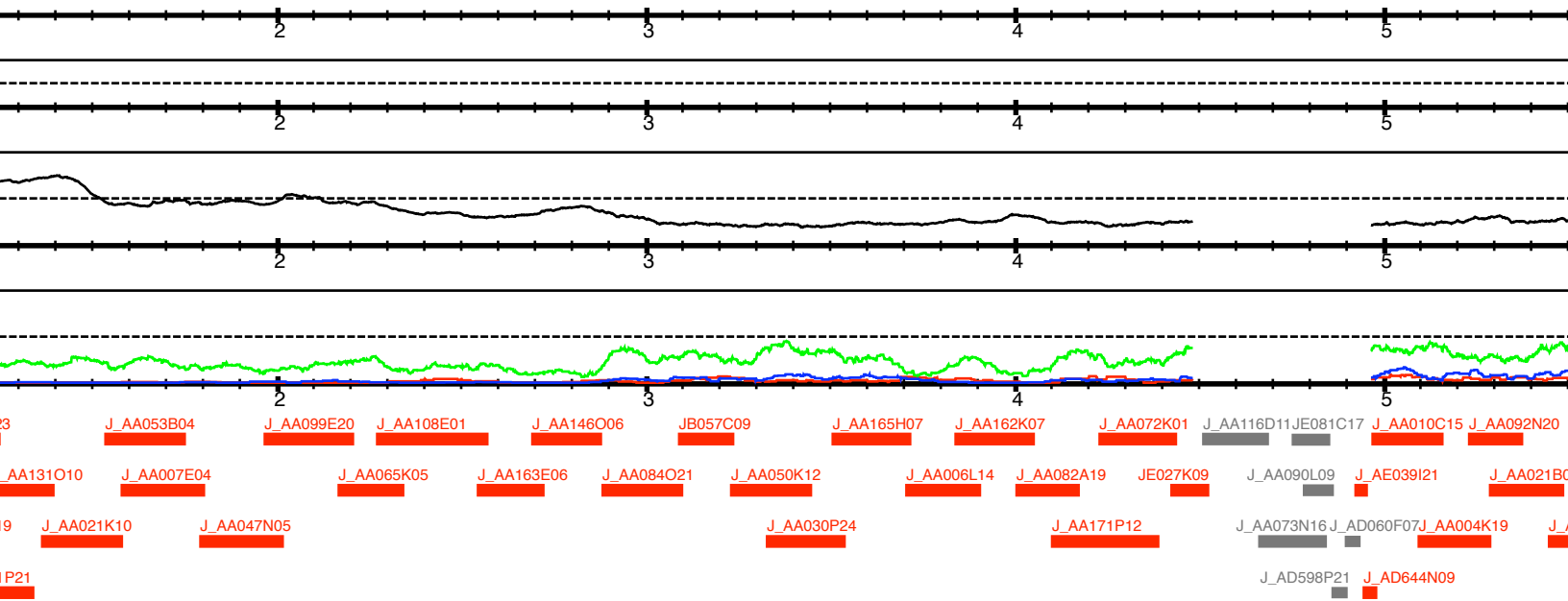
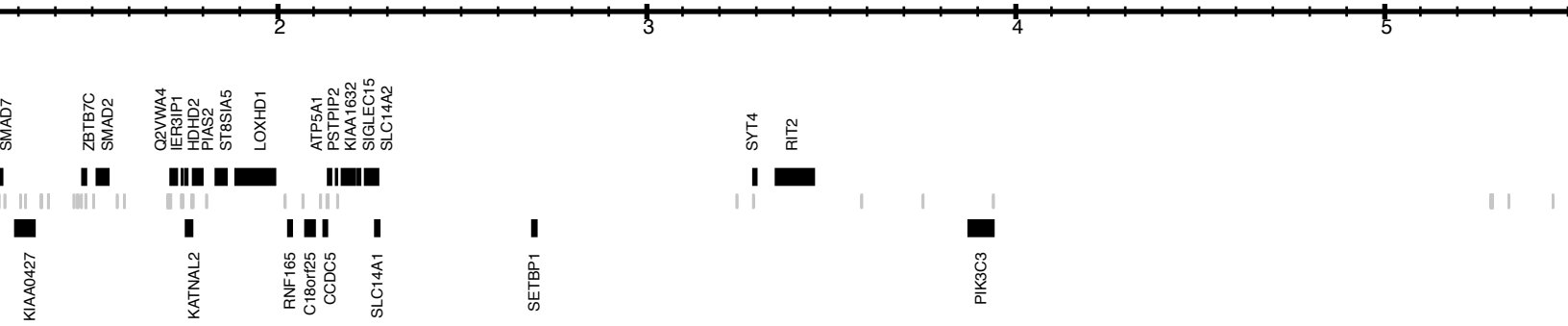
e DNA/LINE/ERV
densities

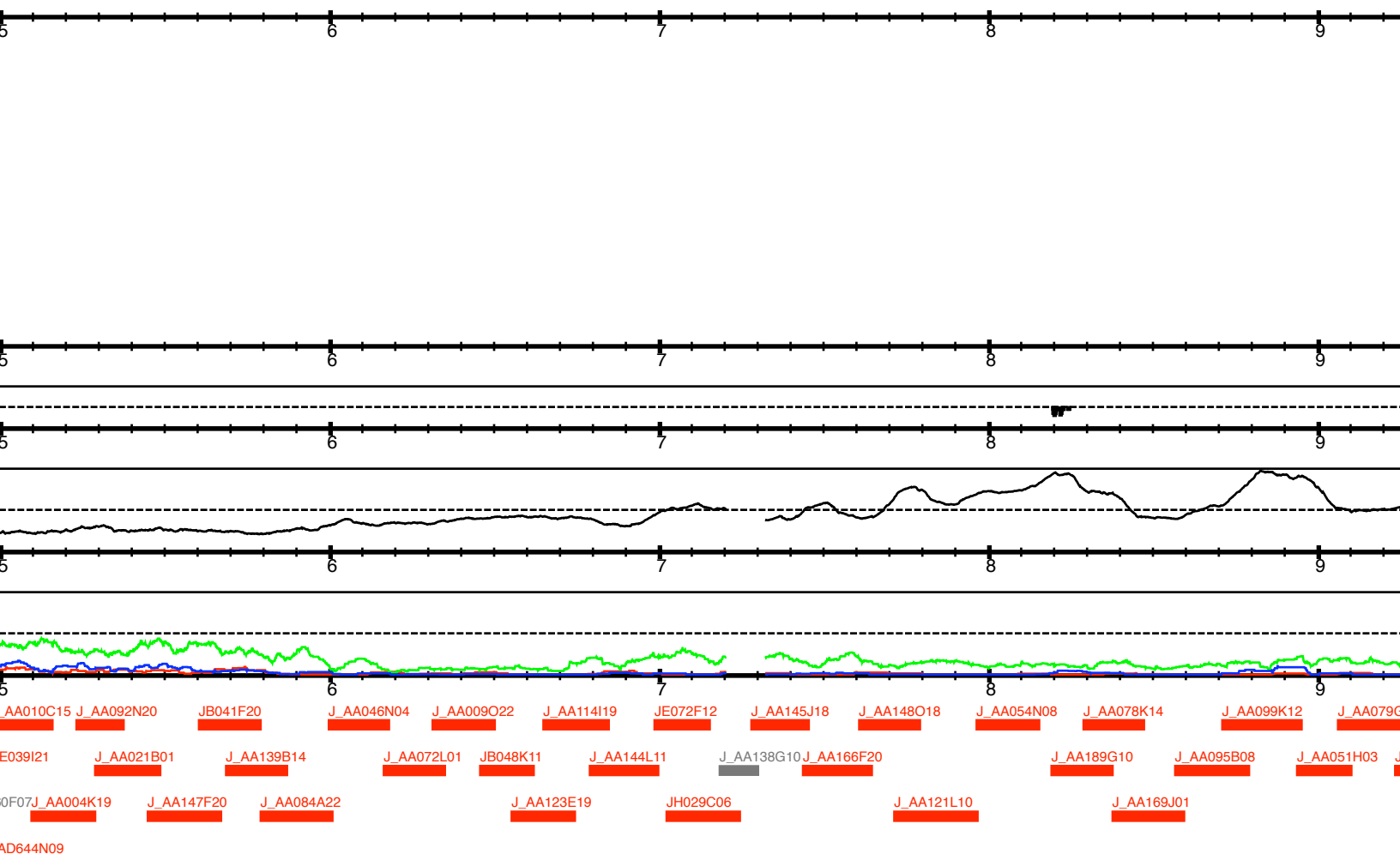
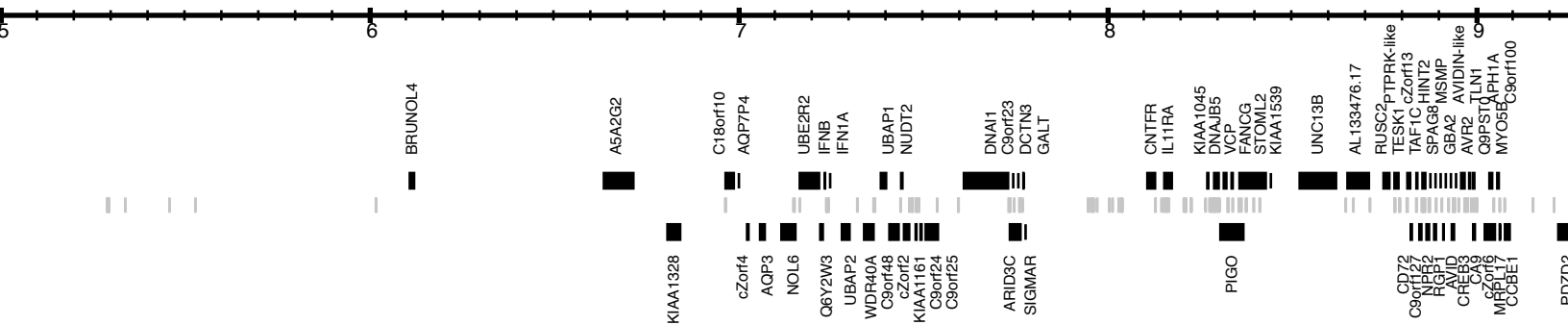


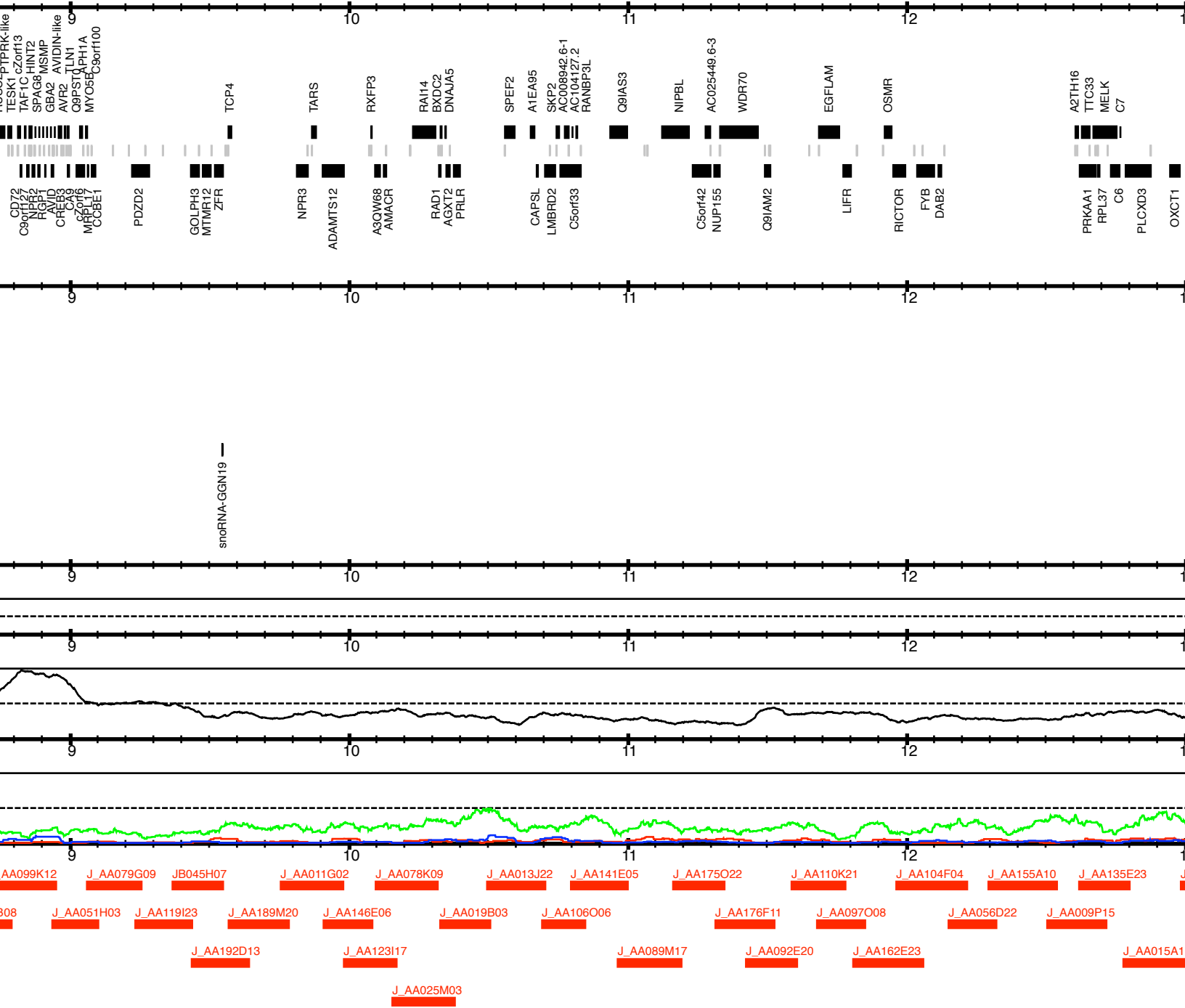
f Sequenced
BACs and Fosmids

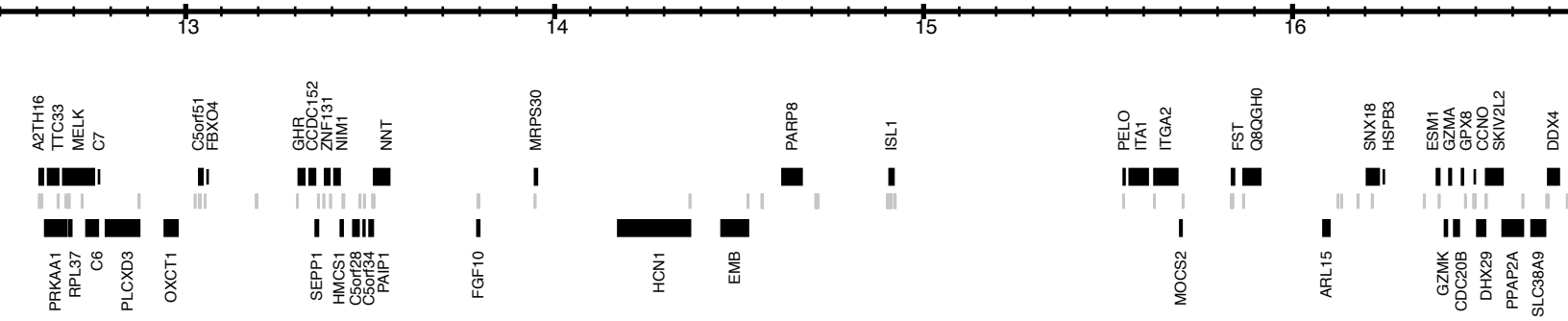


Supplementary Figure 1: Annotated Sequence of the Chicken Z Chromosome

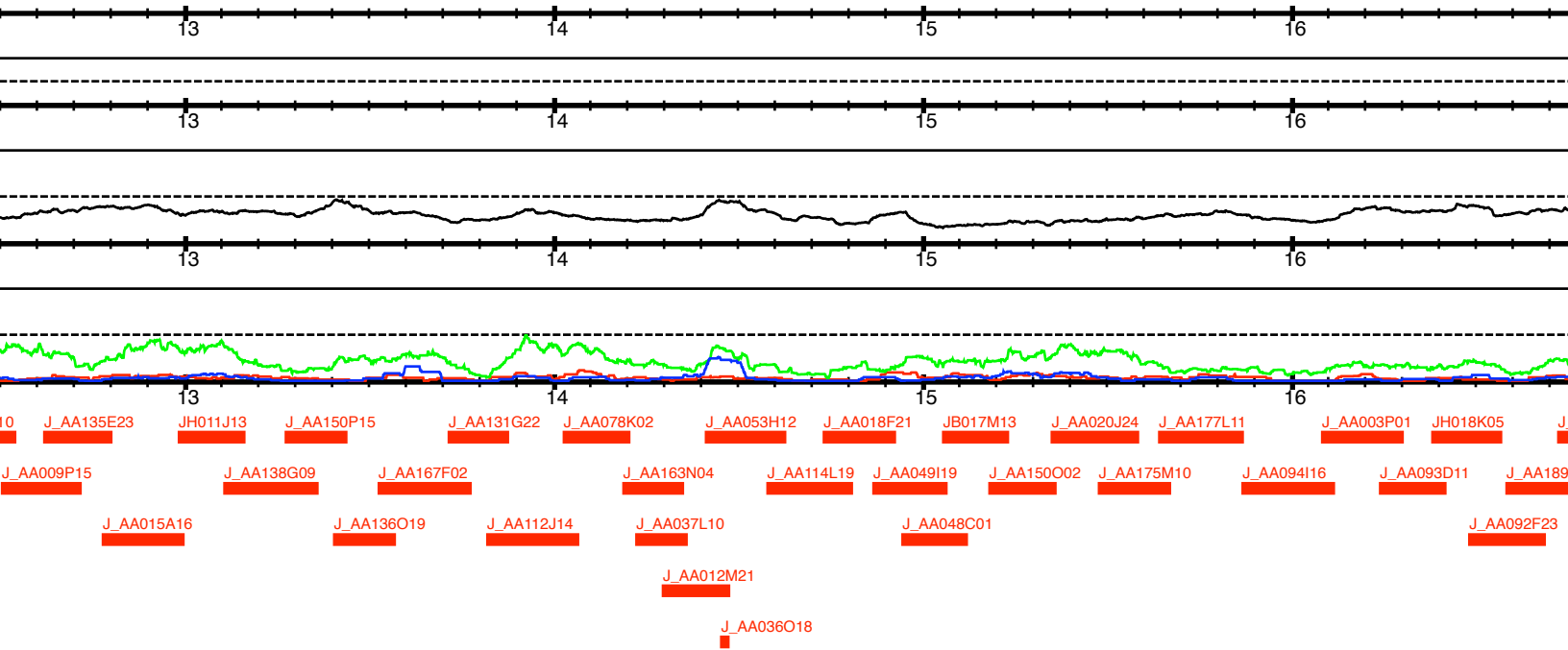


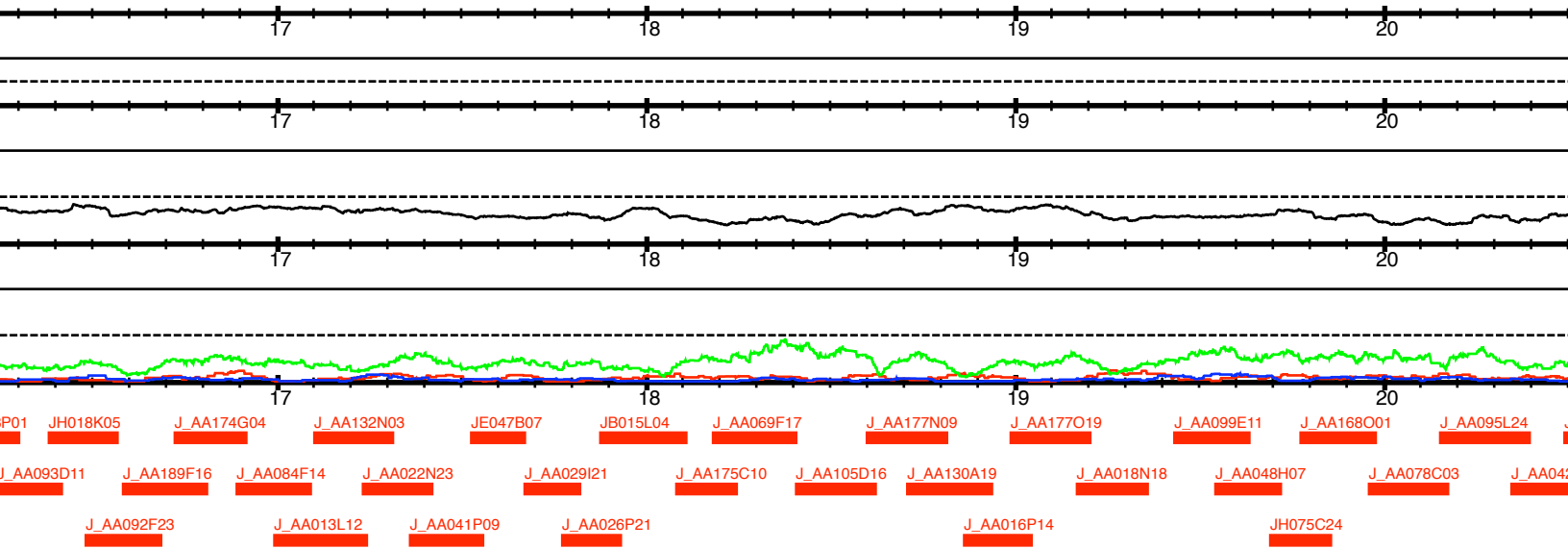
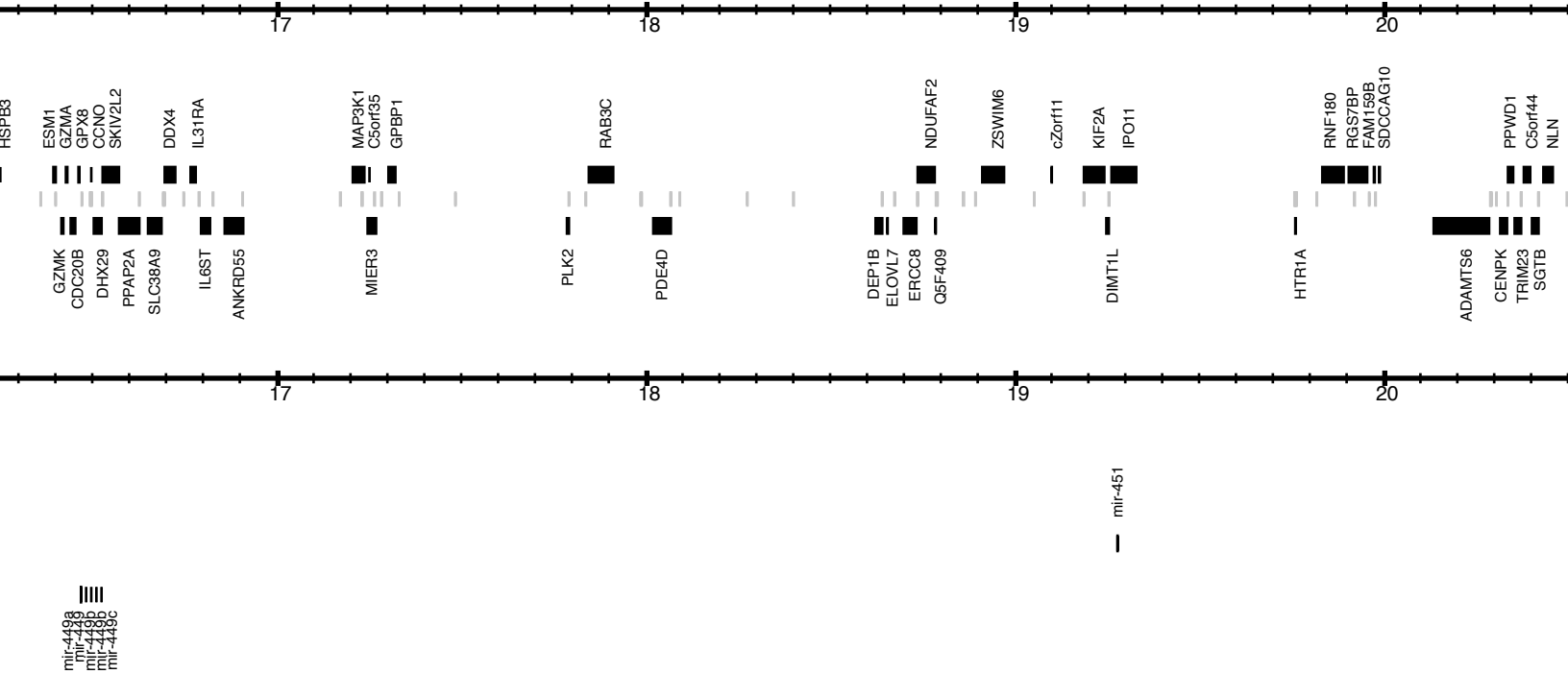


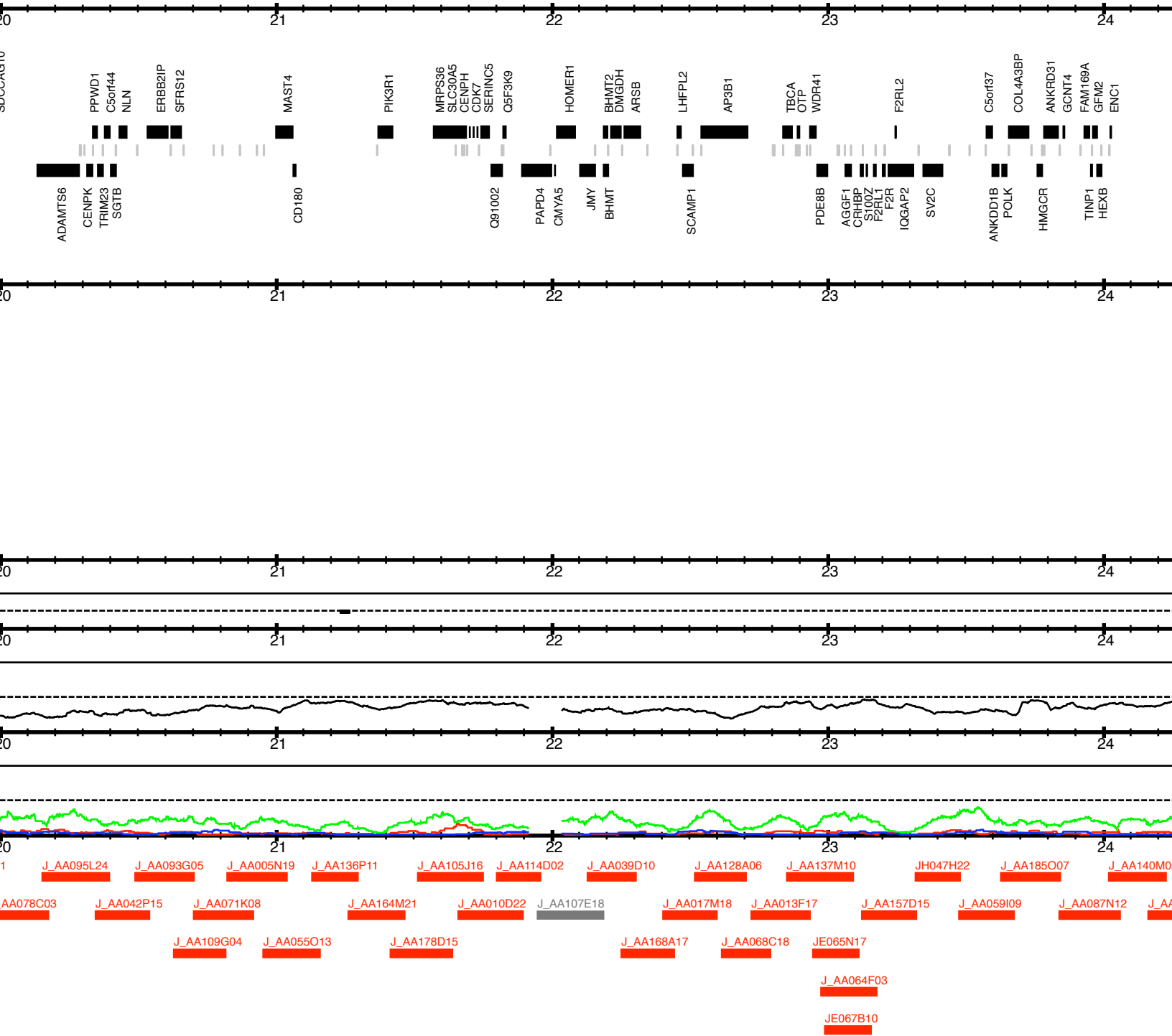


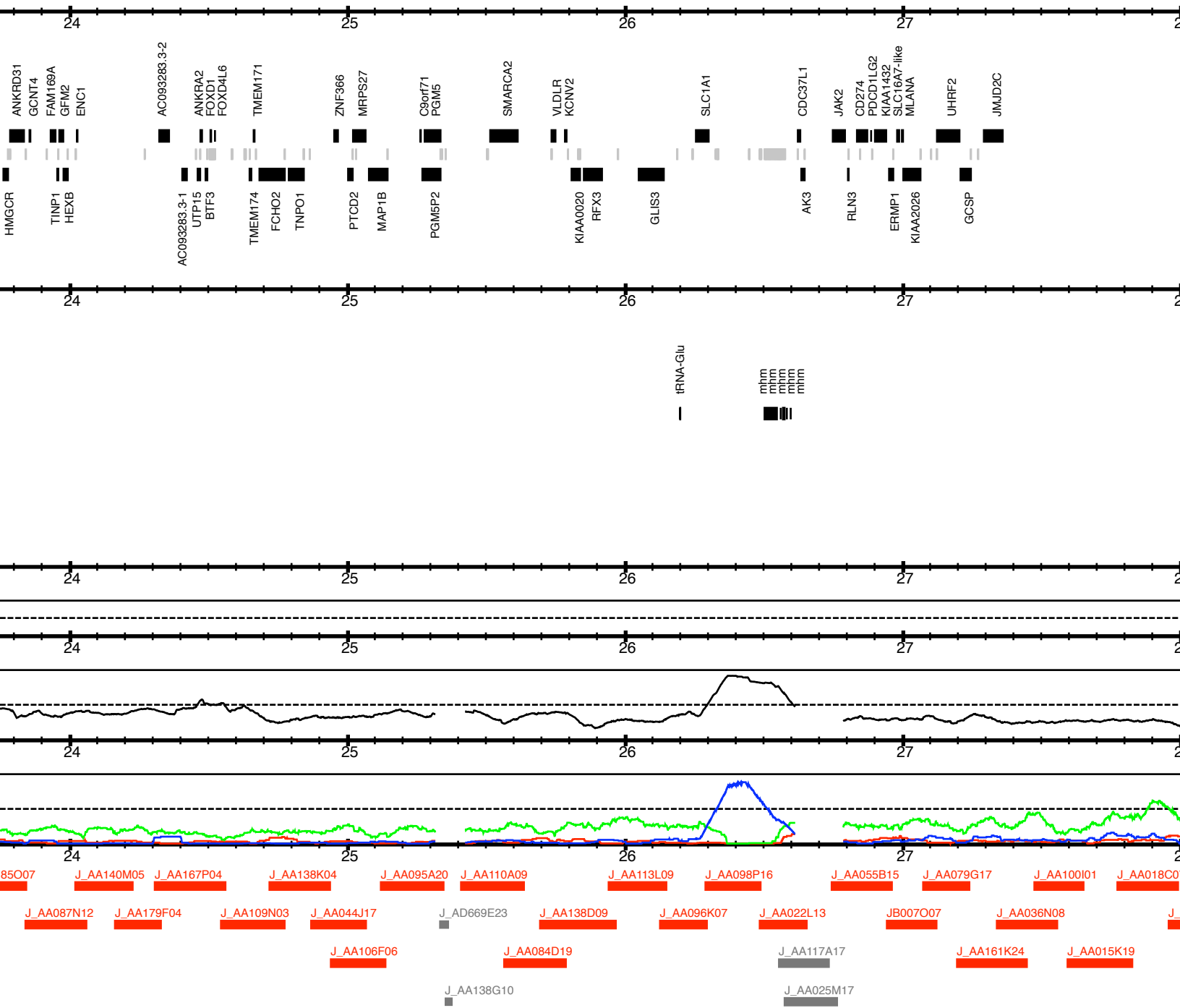


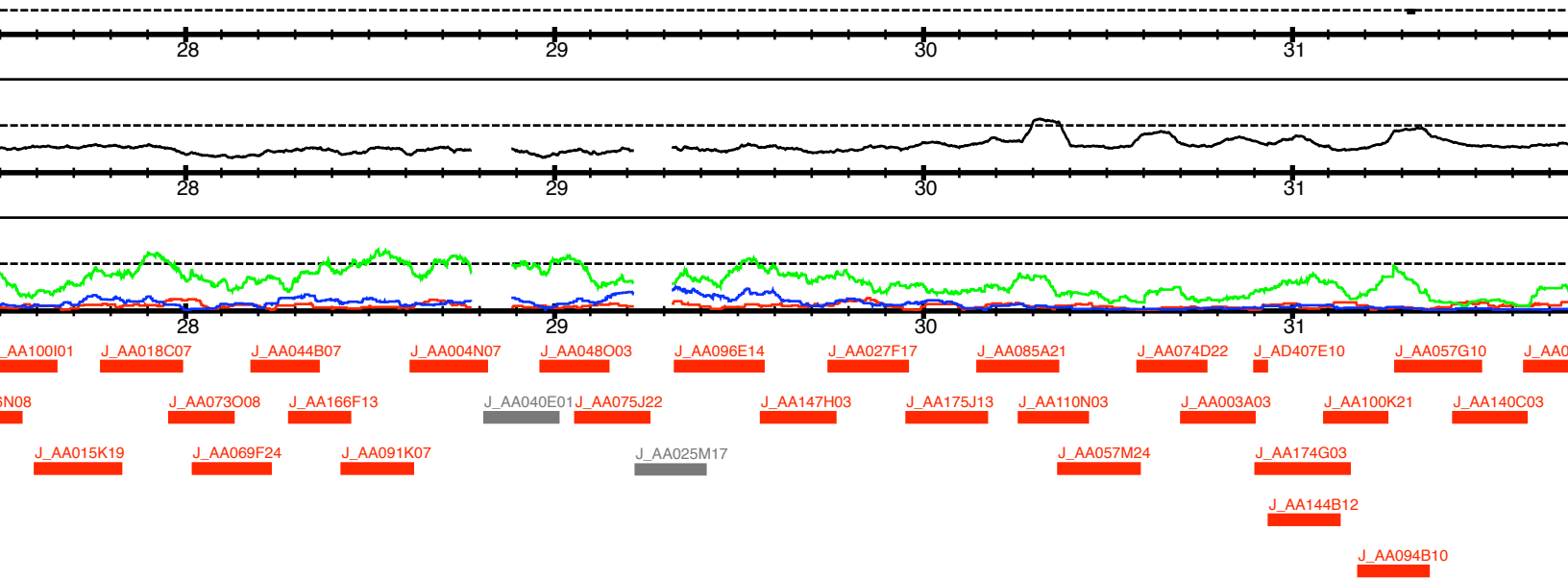
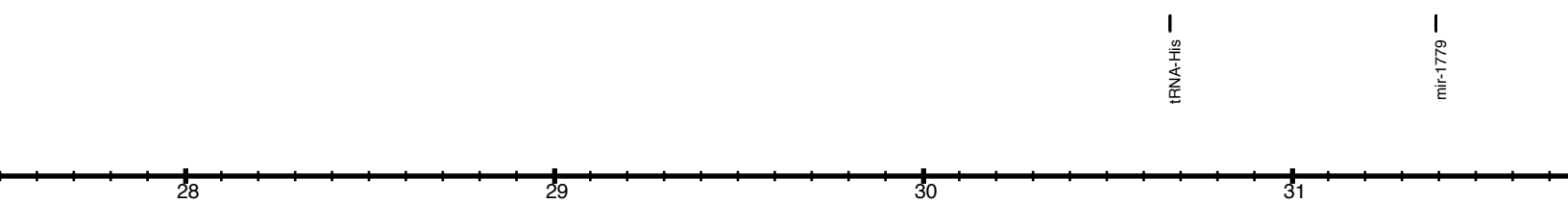
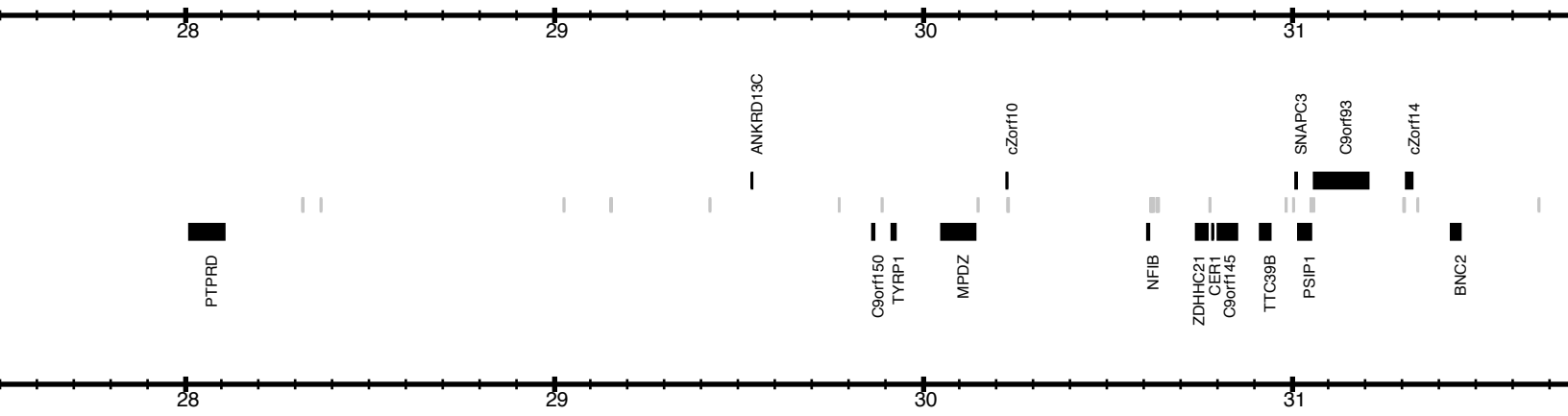
|||||
miR-4498
miR-4499
miR-4496
miR-4495
miR-4496

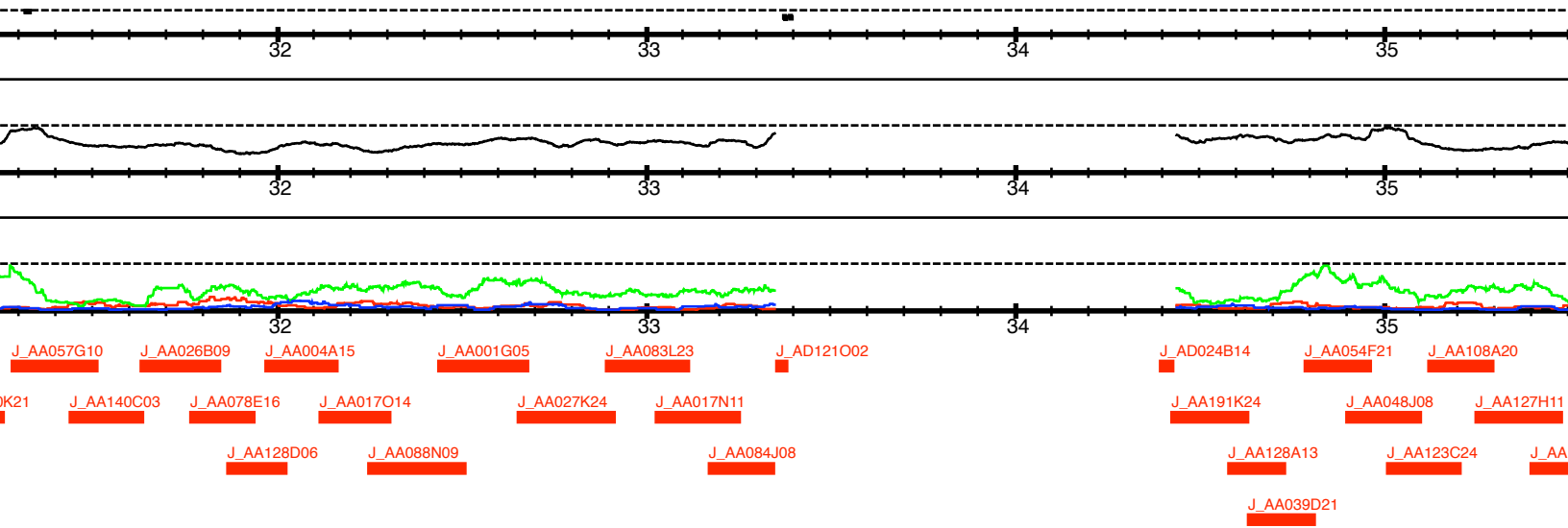
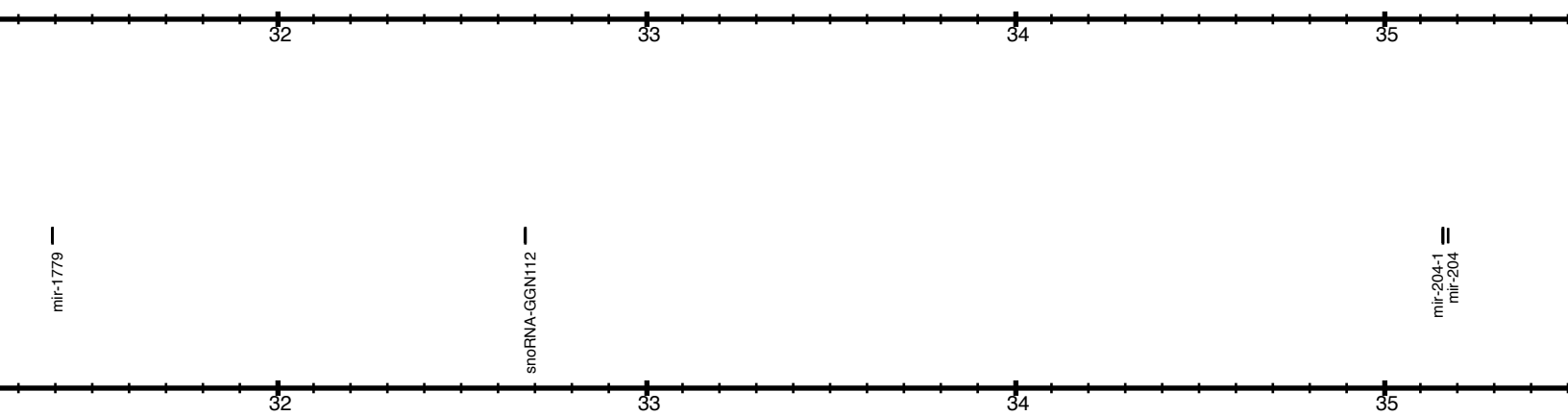
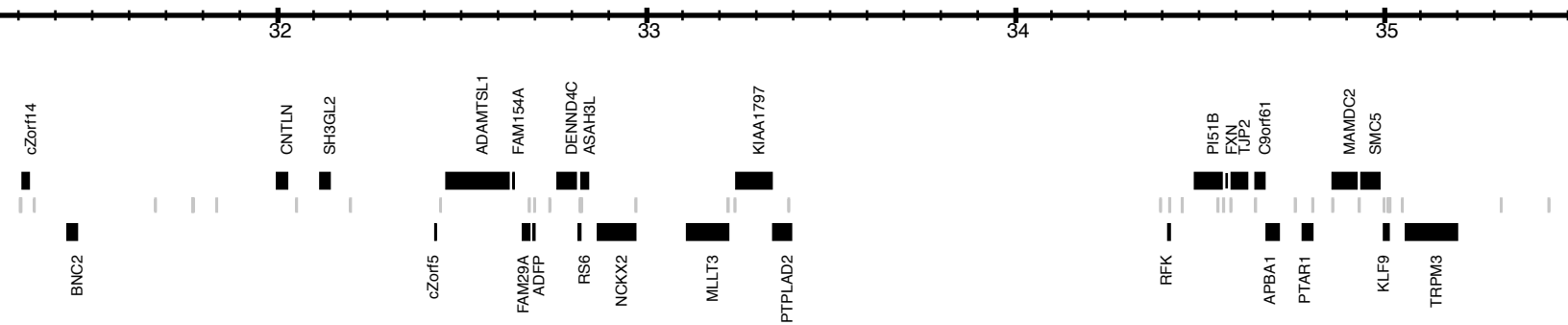


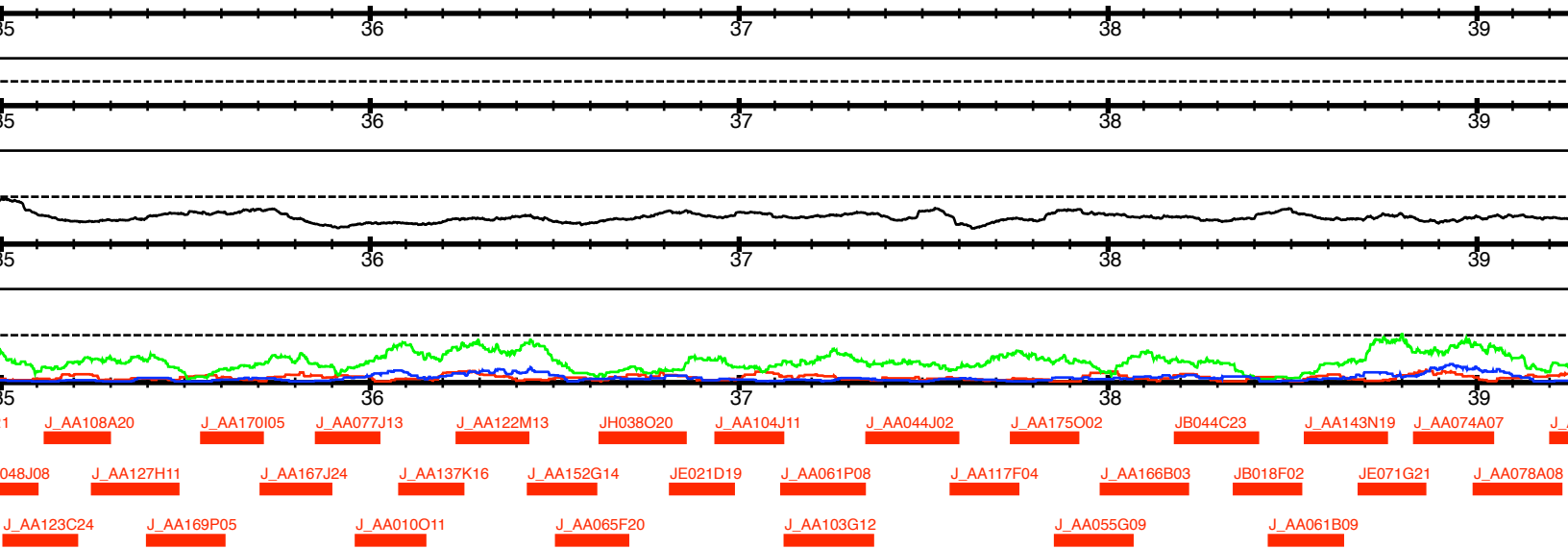
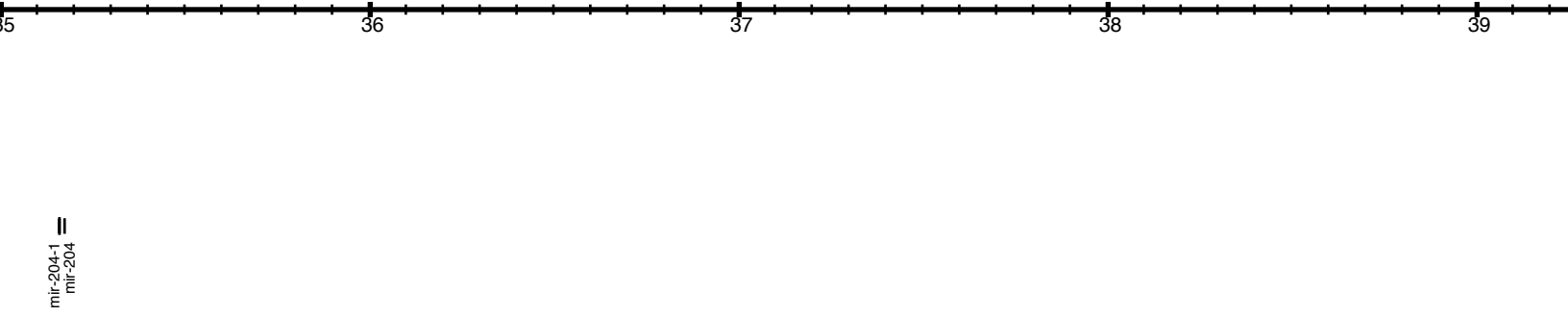
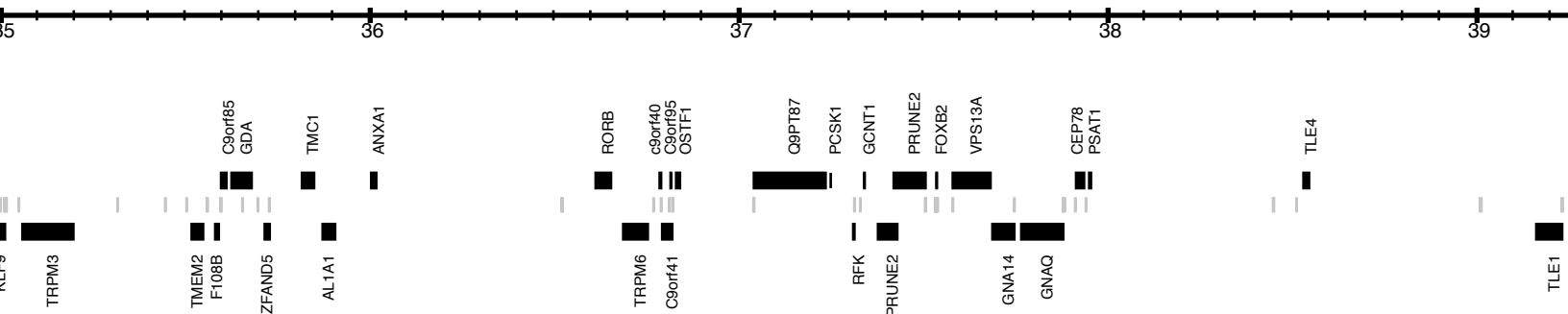


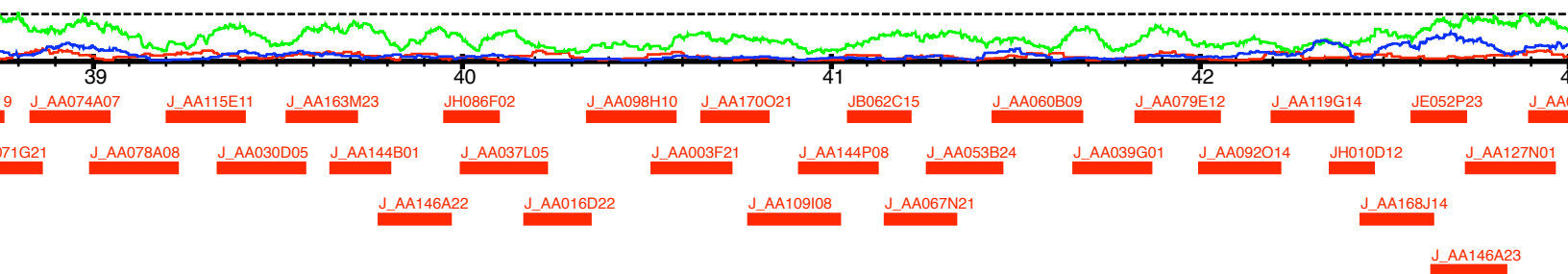
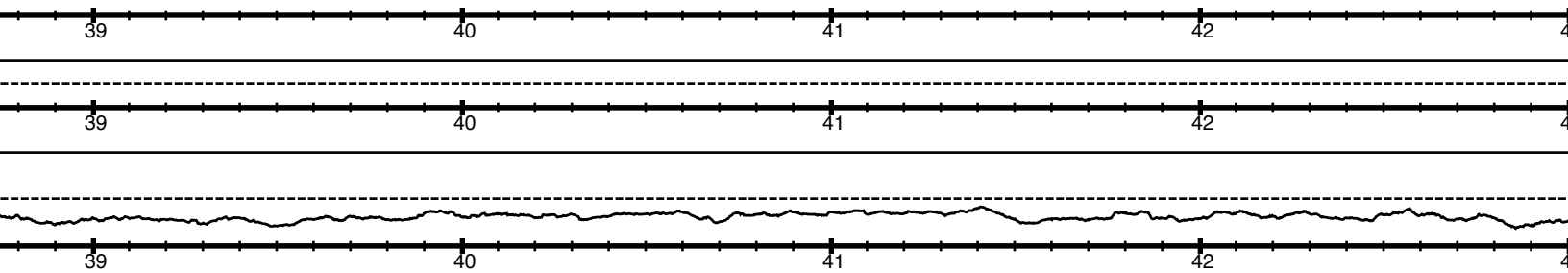
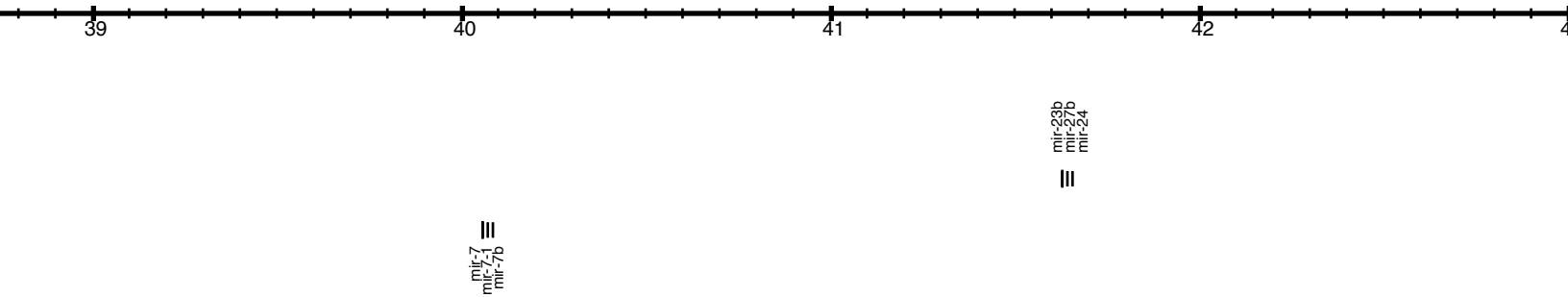
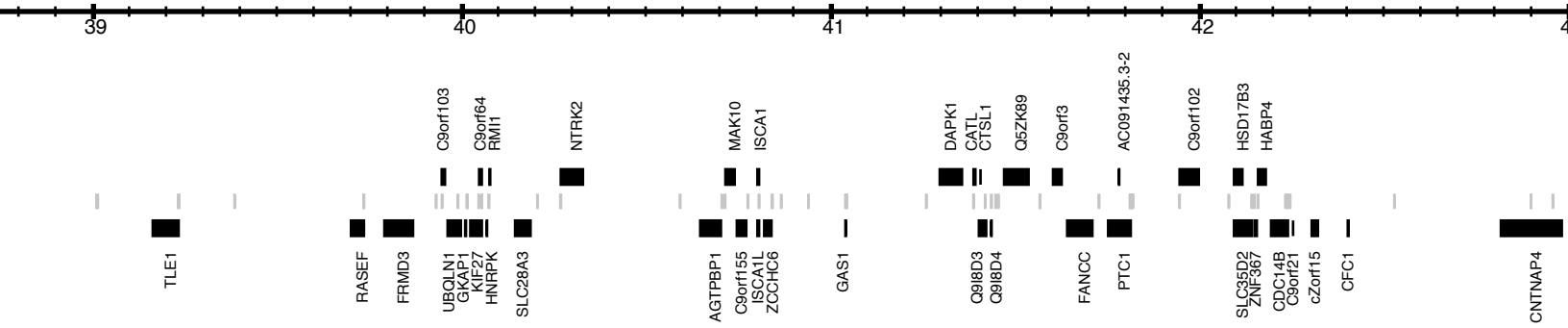


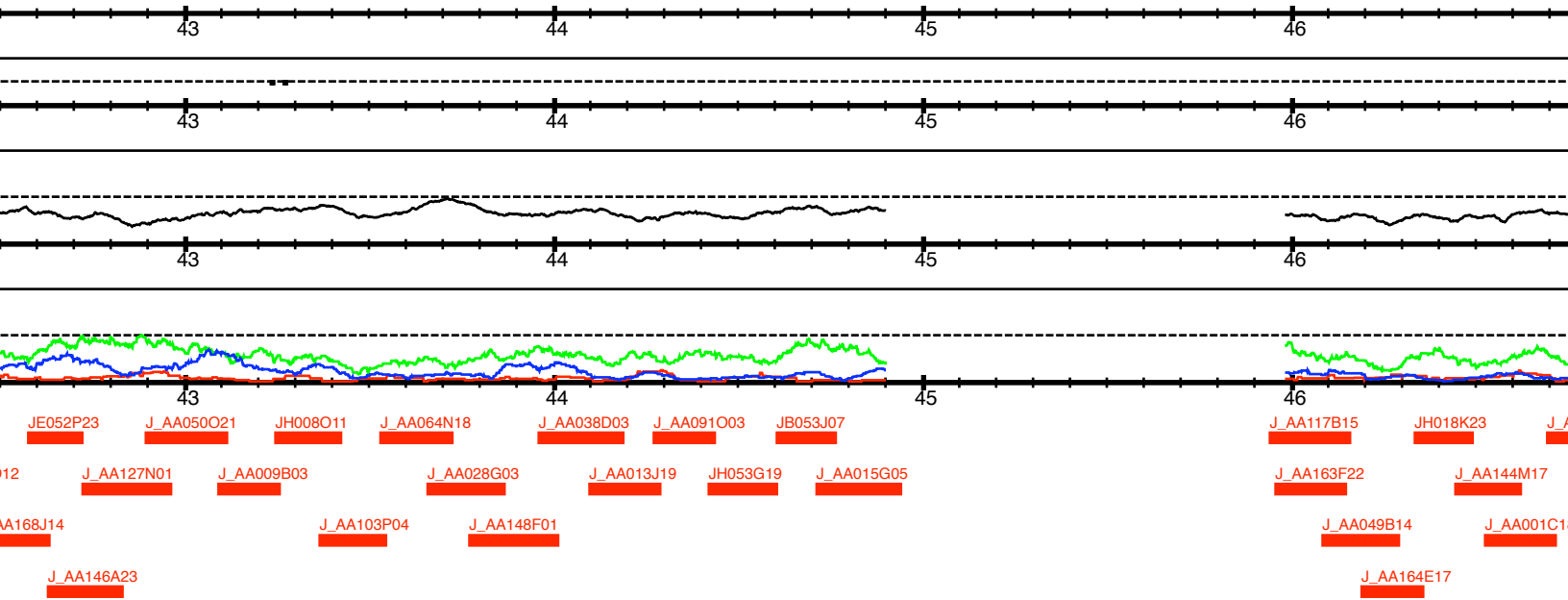
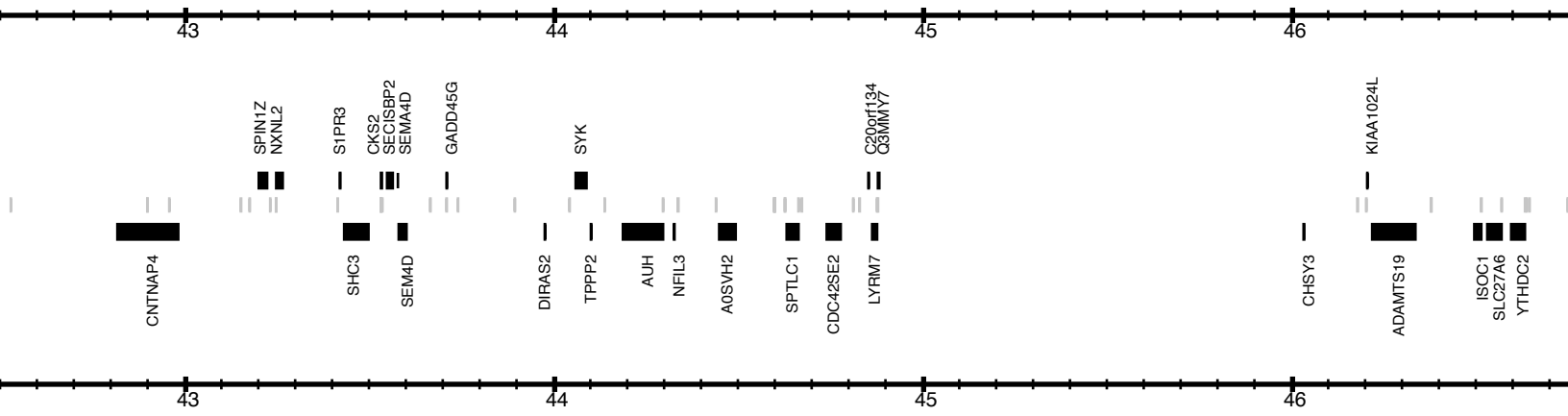


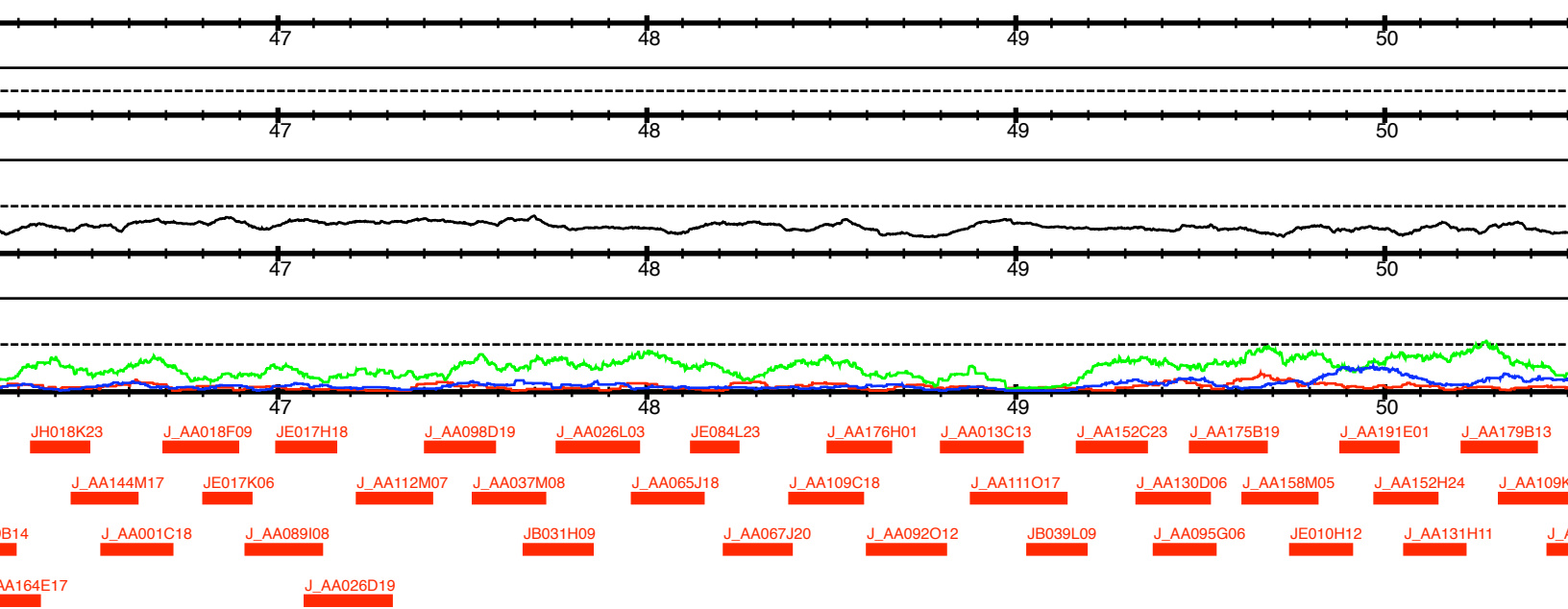
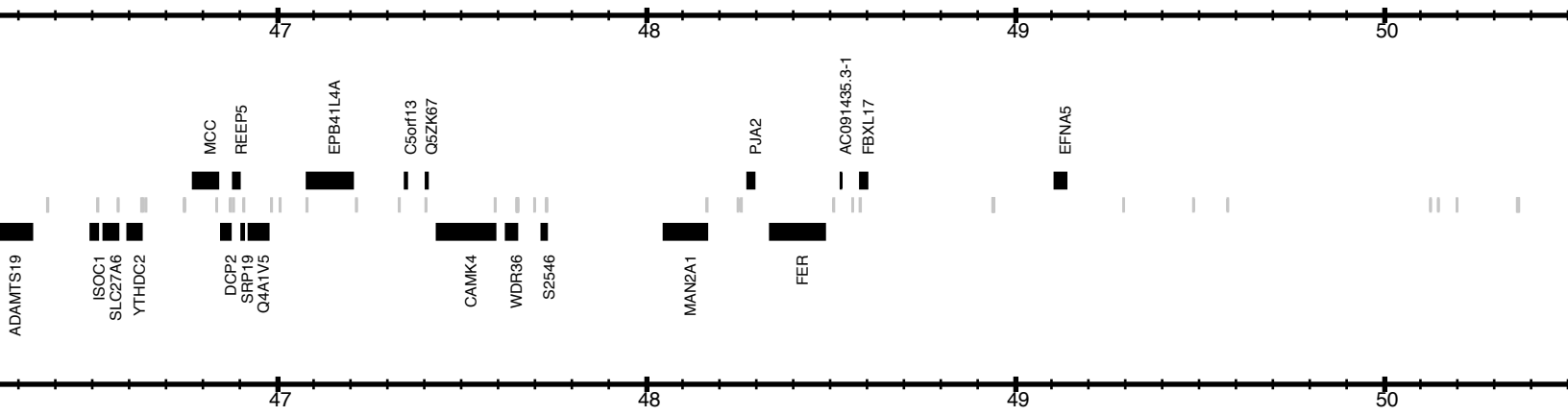


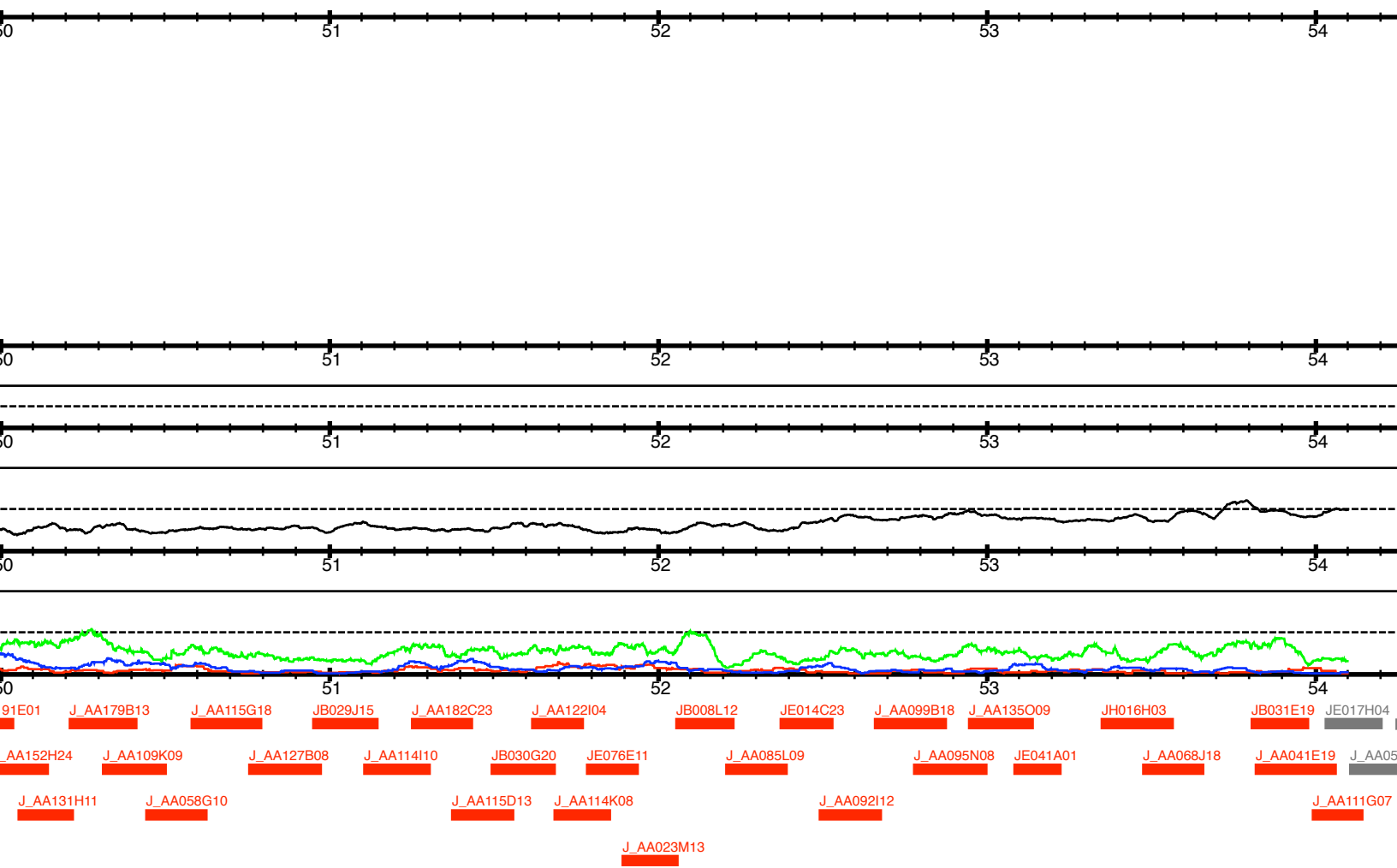
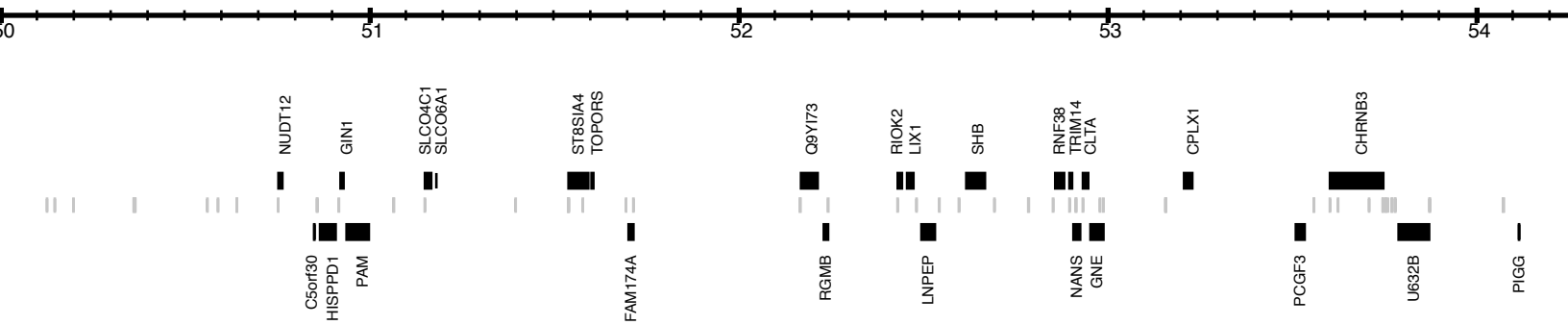


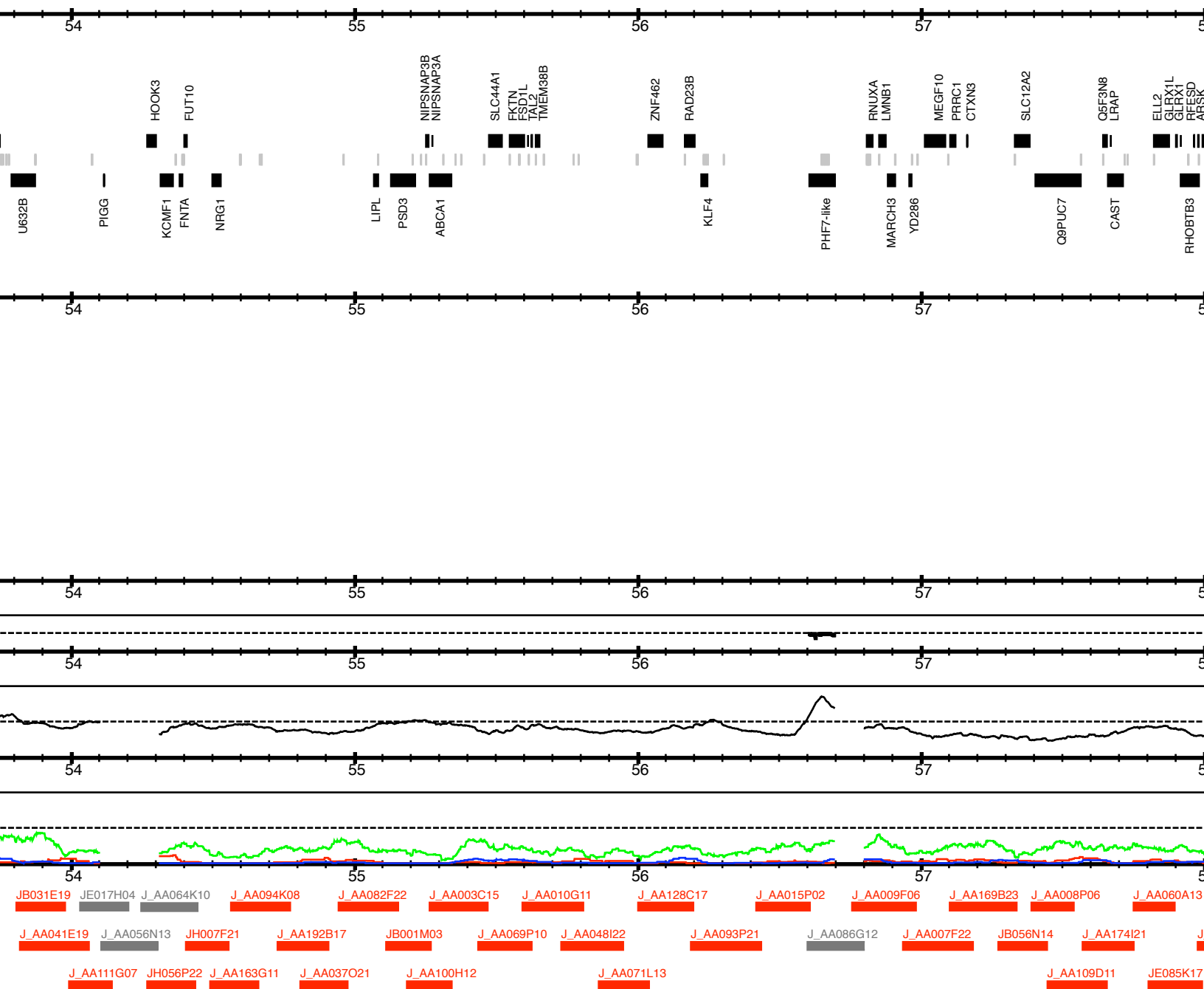


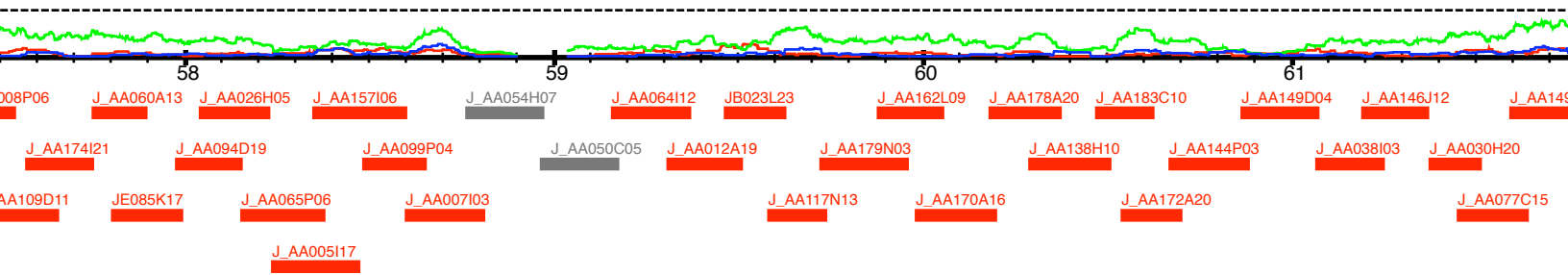
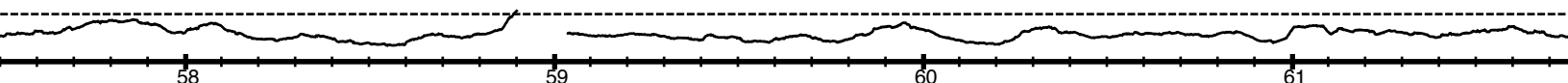
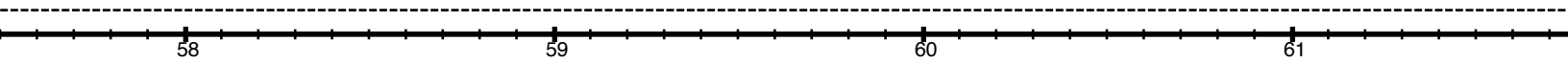
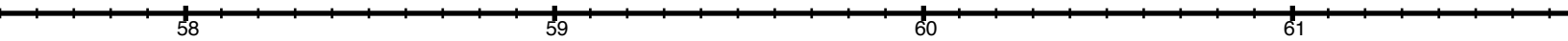
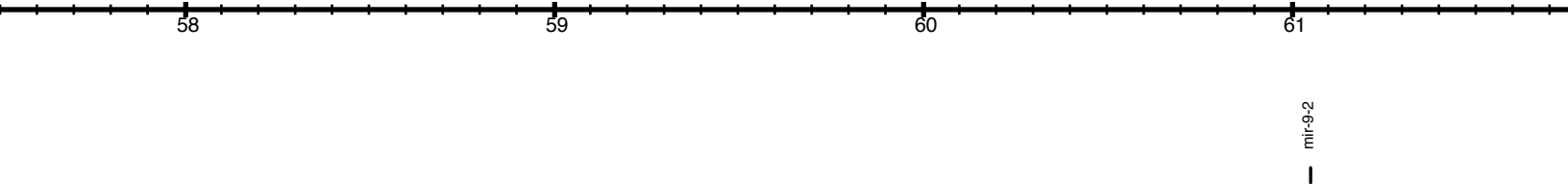
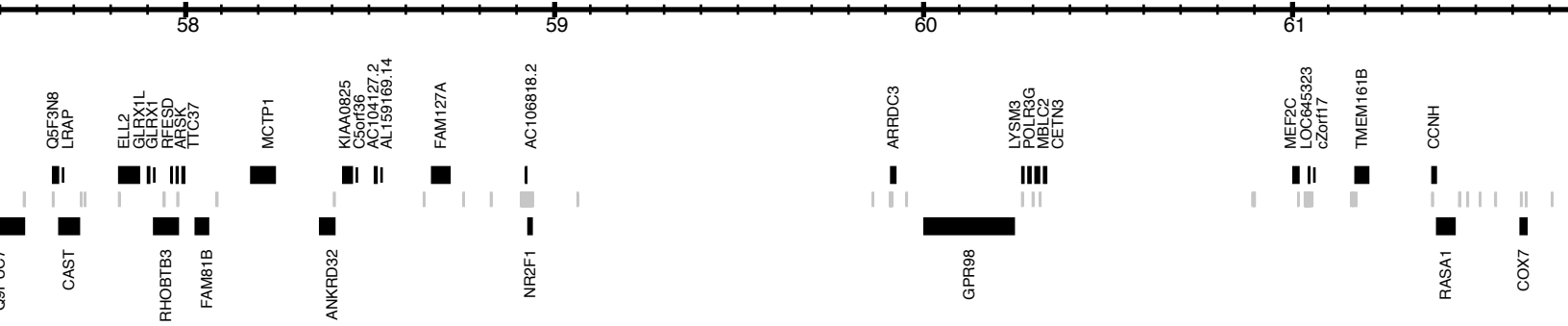


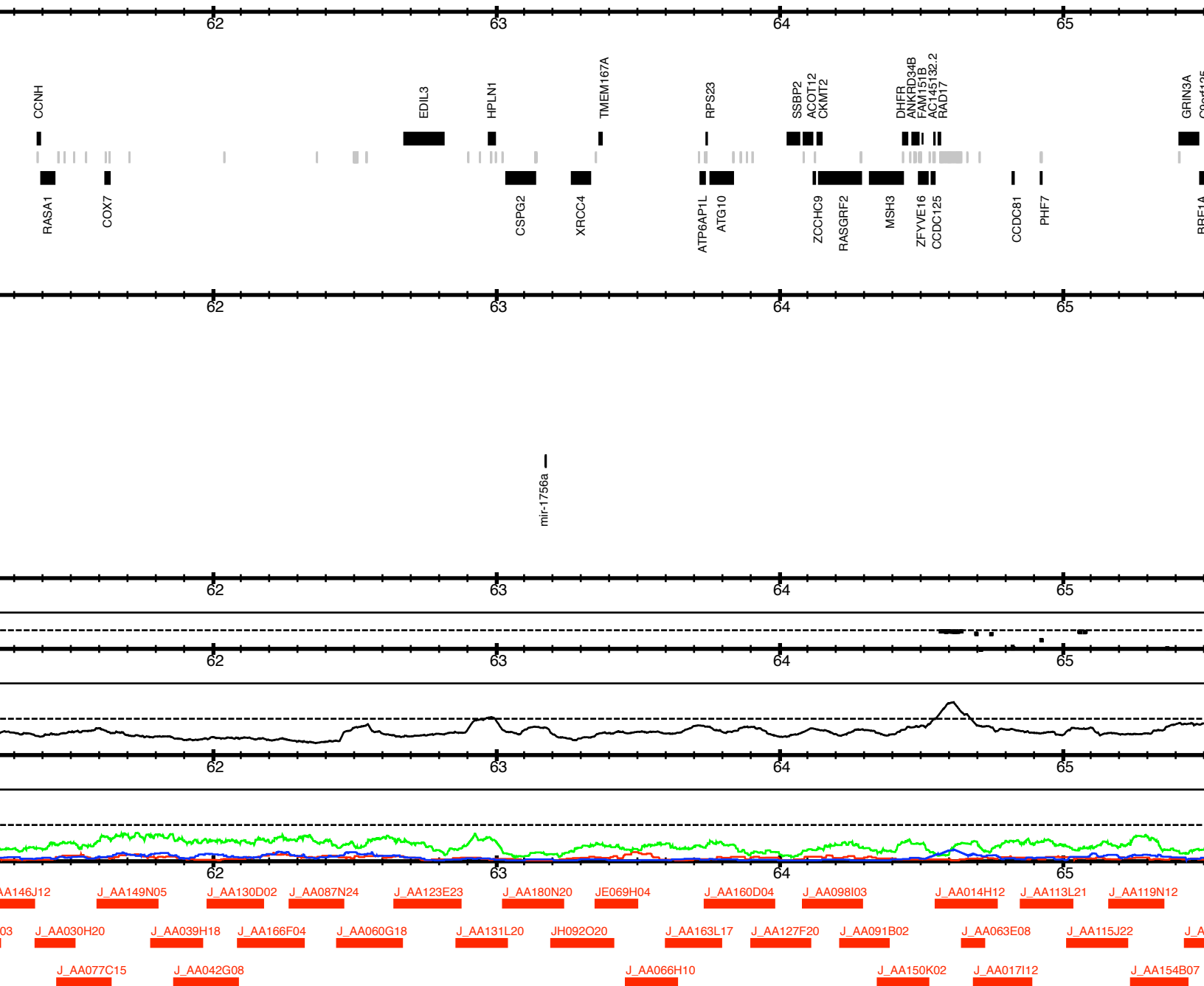


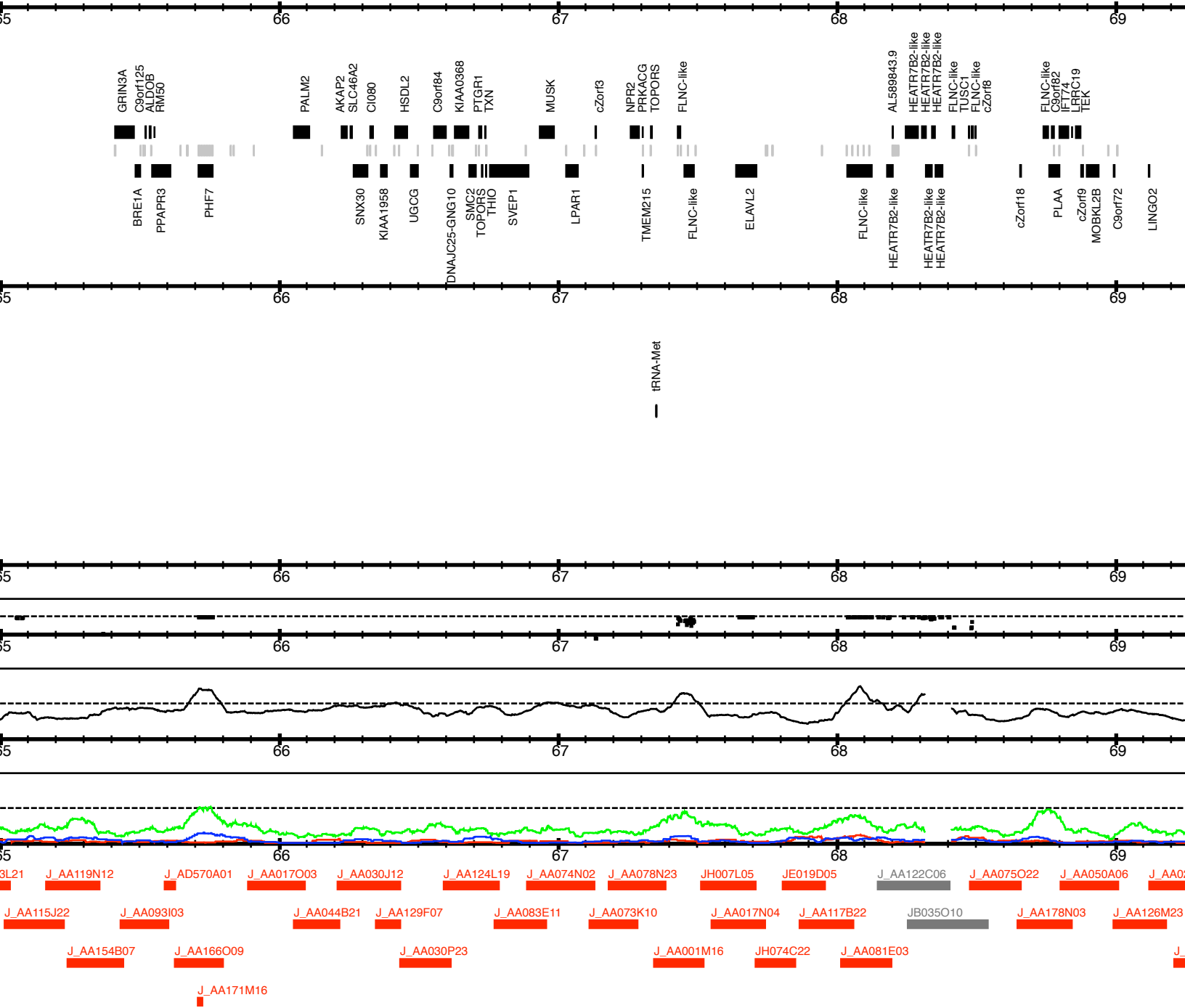


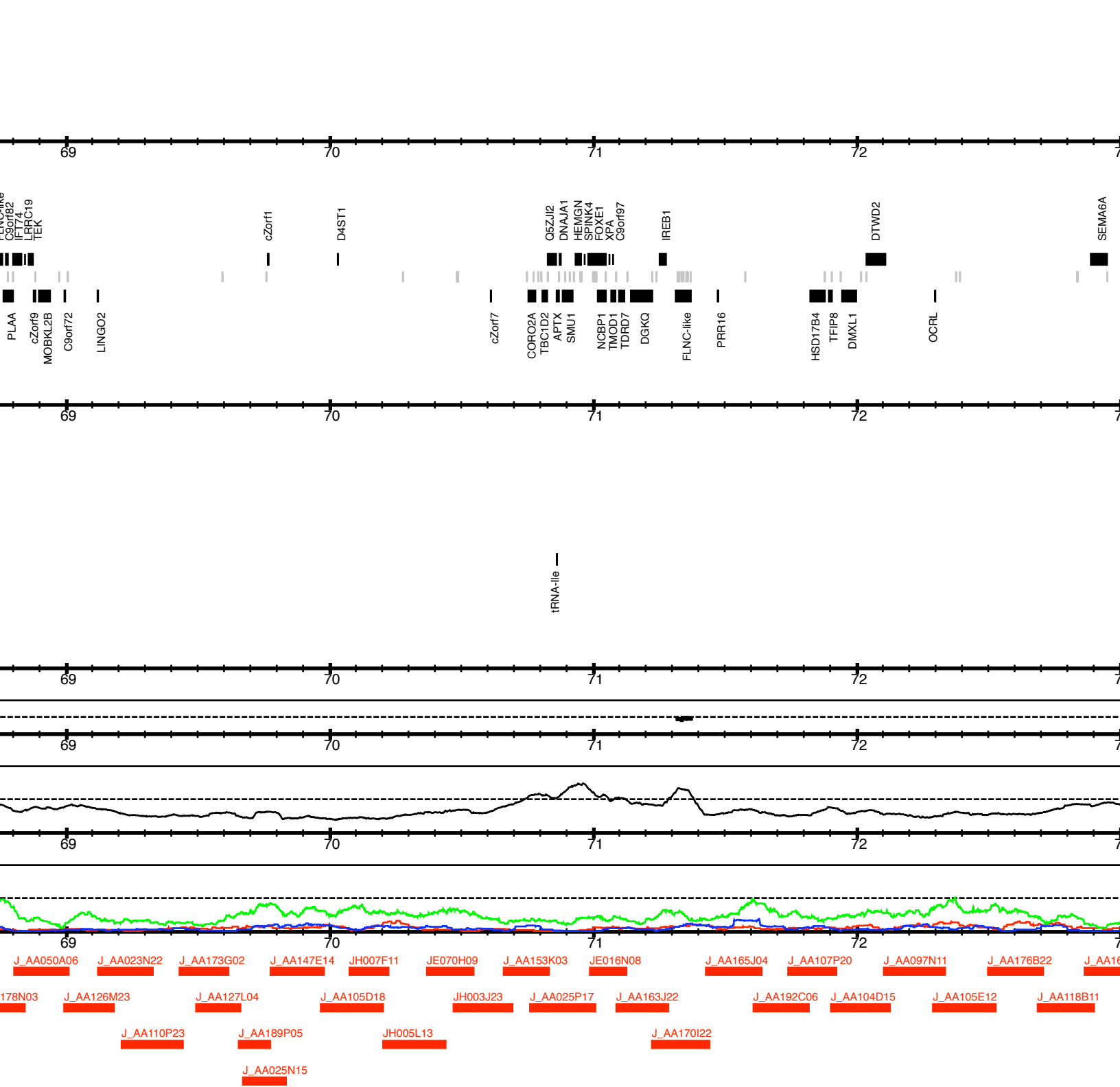


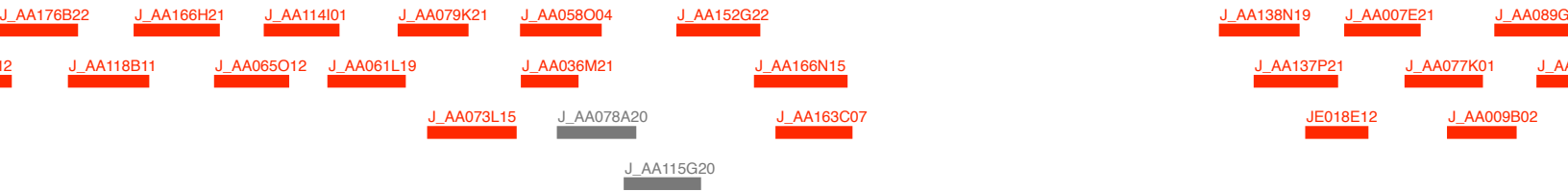
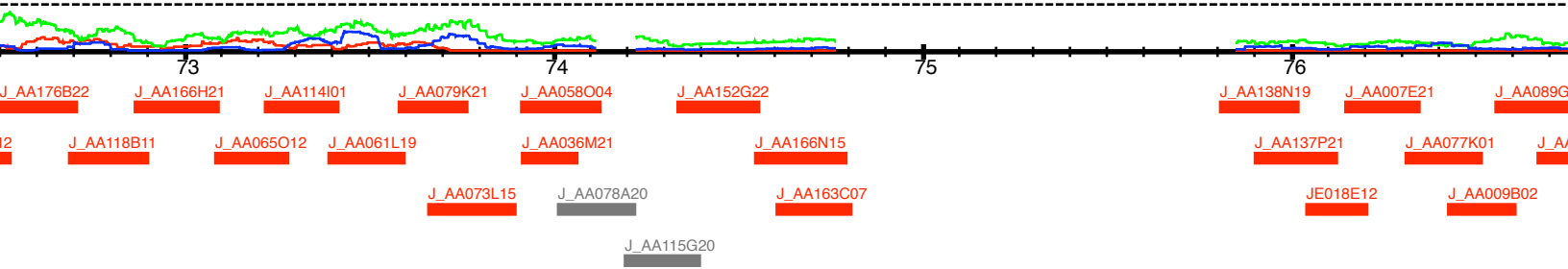
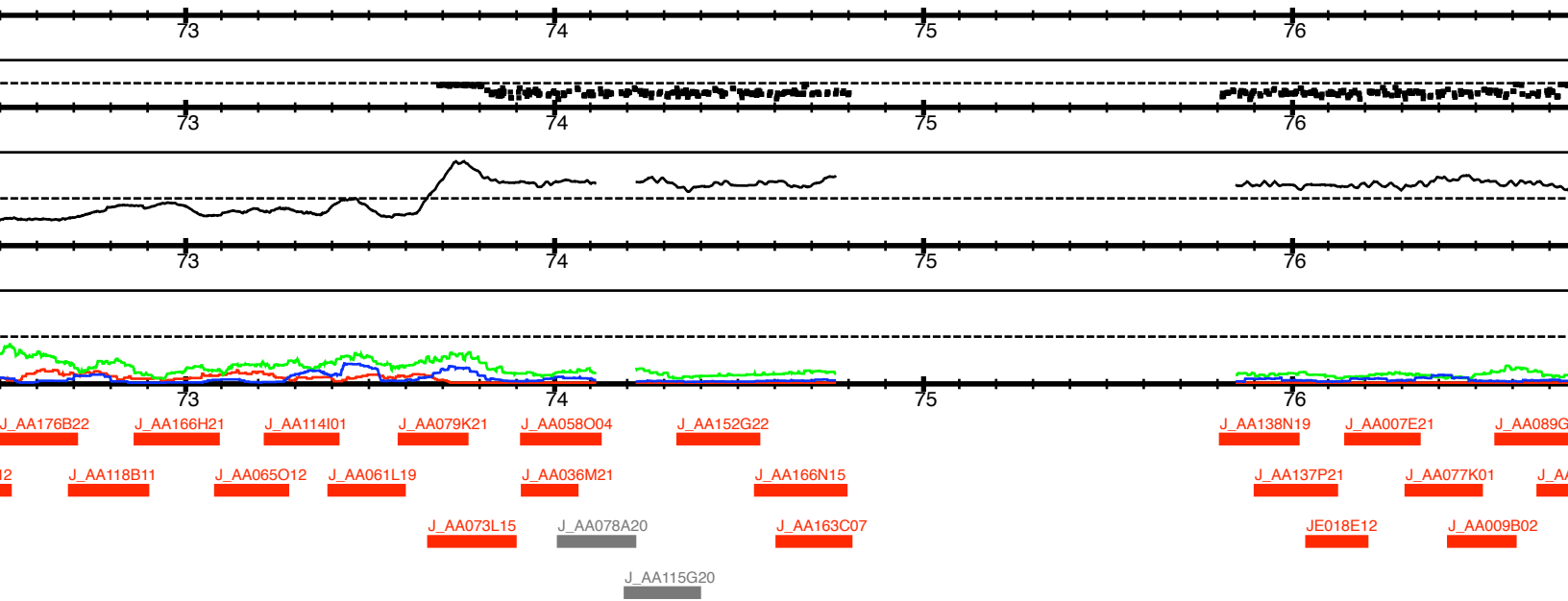
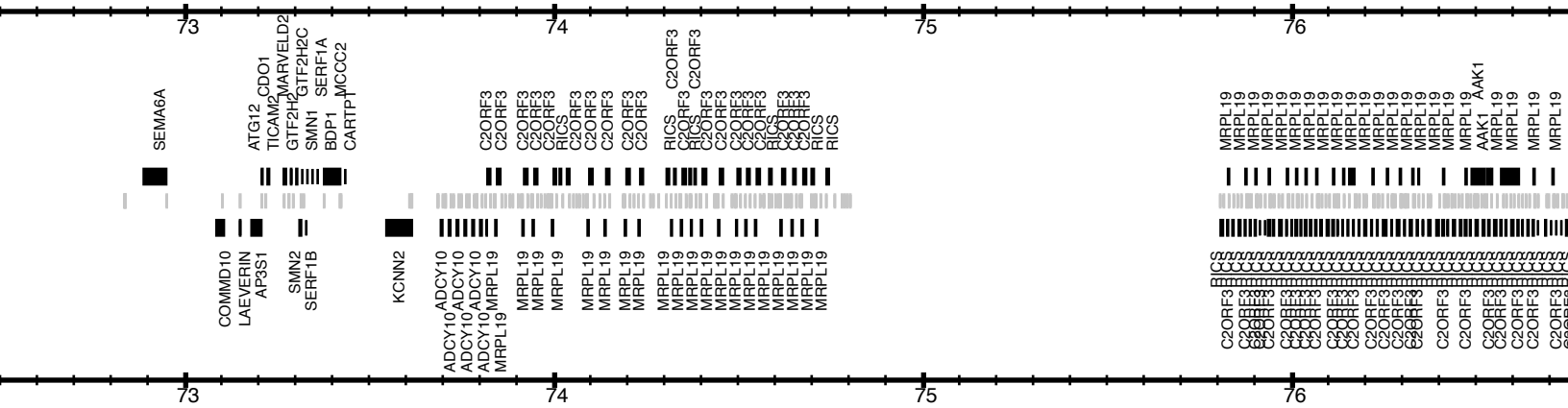


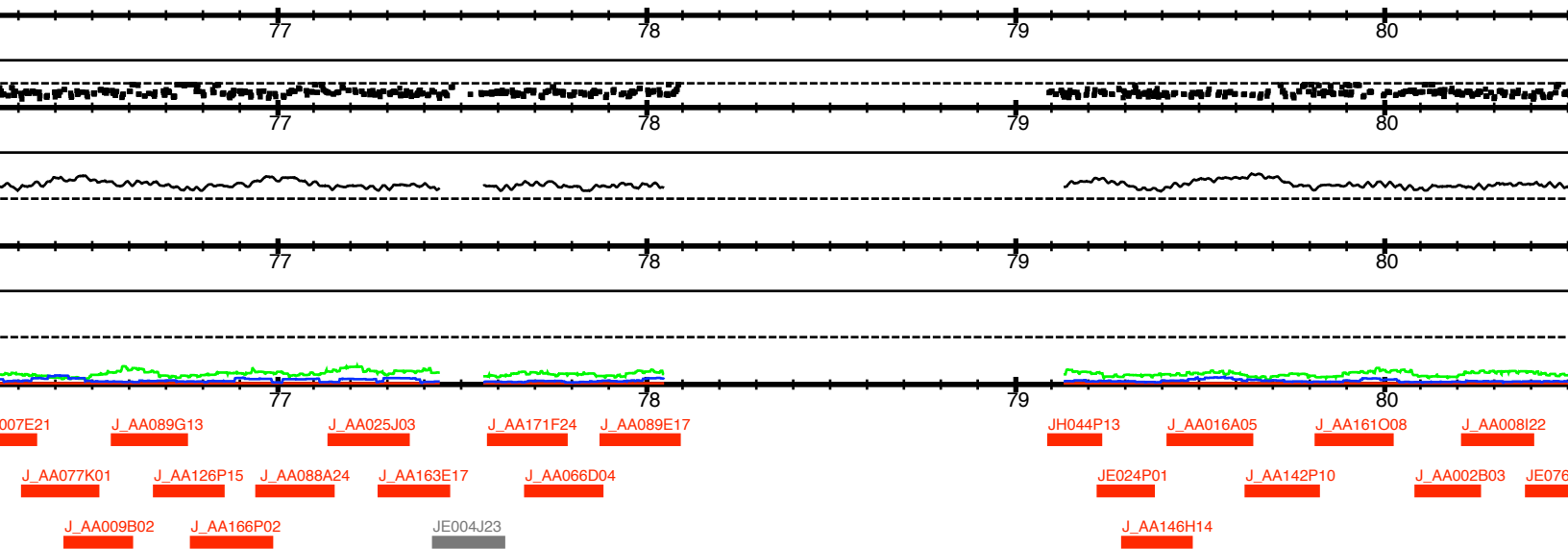


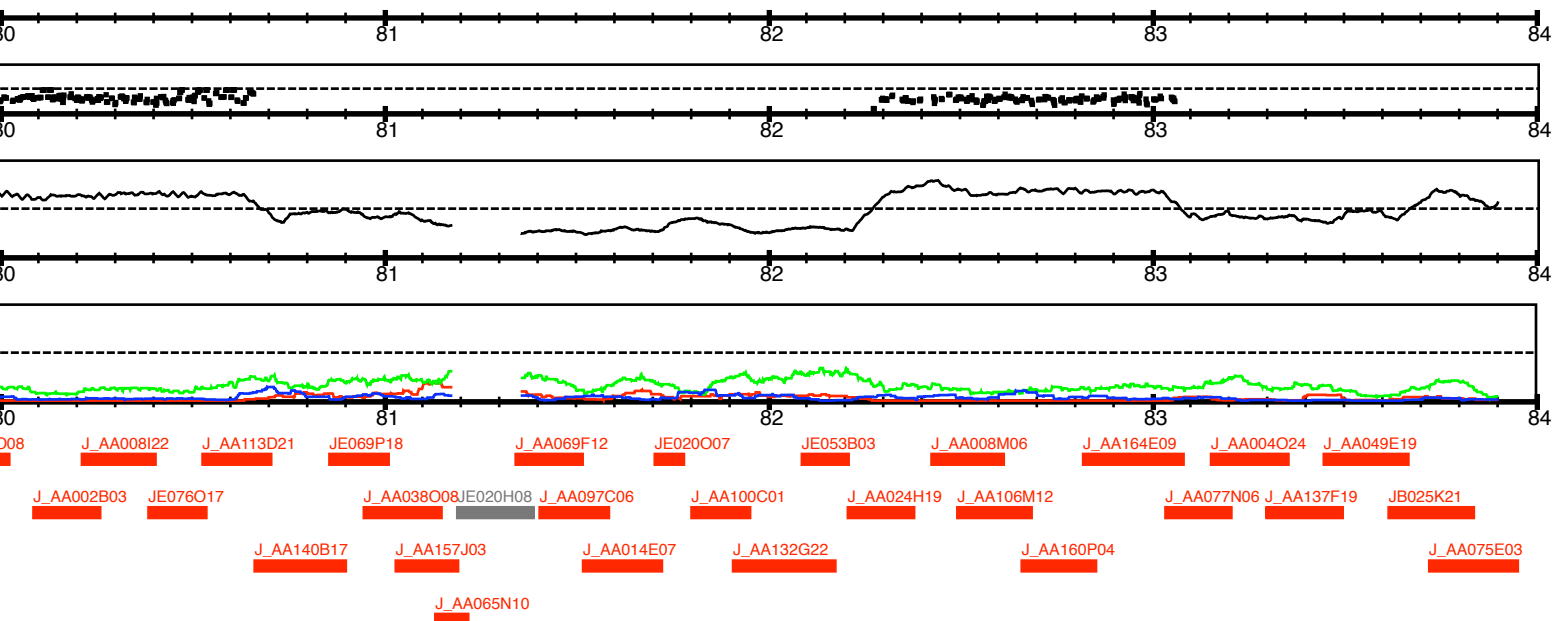


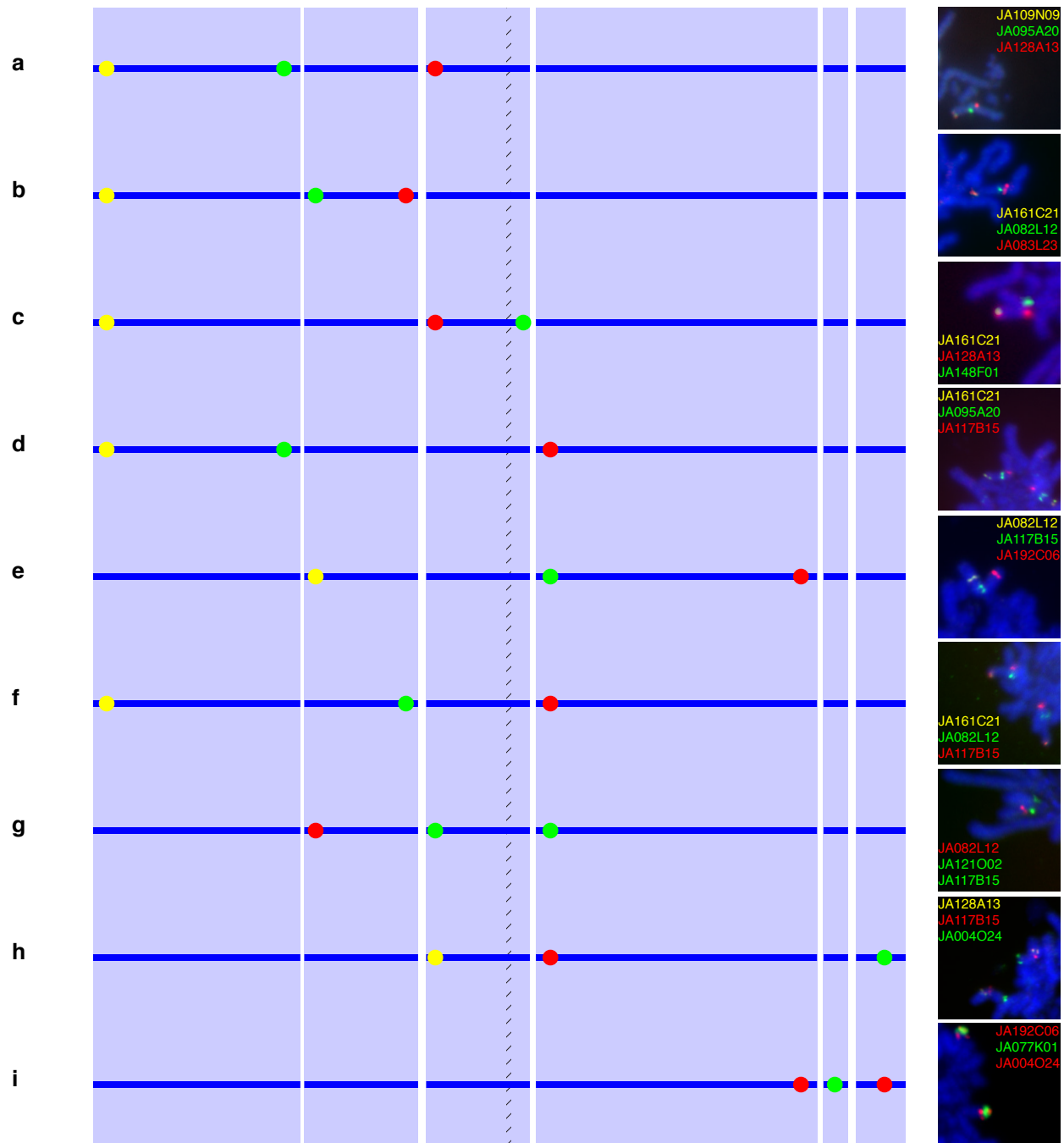






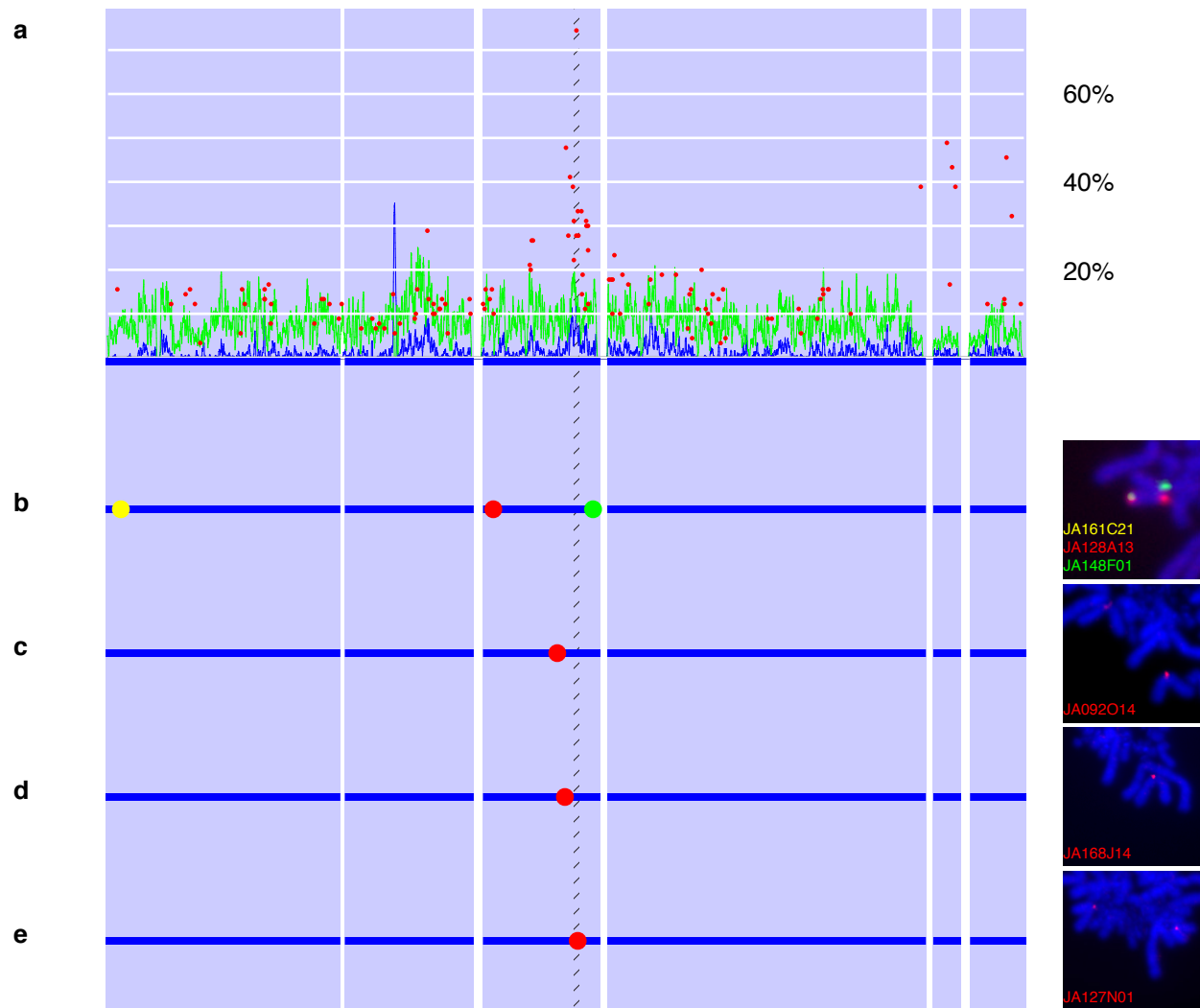






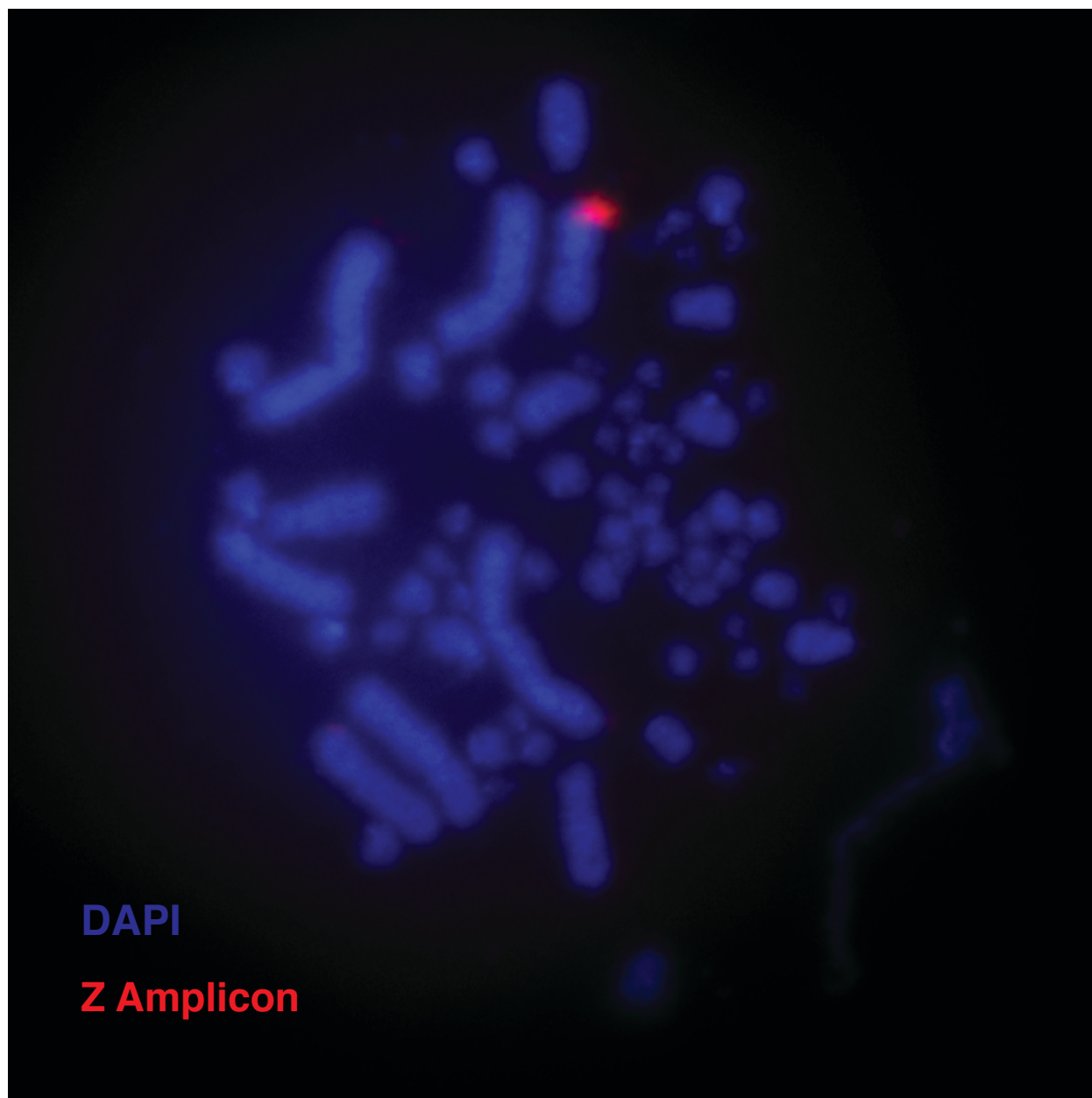
Supplementary Figure 2: Order and orientation of Z chromosome contigs by DNA-FISH.

Schematic diagrams describing the locations of BAC probes (colored circles) within Z chromosome sequence contigs (blue lines). Each schematic is accompanied by a representative FISH image. Centromere position is shown by hashed lines. (a-e) Establish the orientation of the first four contigs and their position relative to the centromere. The first two contigs are on the short arm, the third contains the centromere, and the fourth contig is on the long arm. (f-i) Establish the order of contigs. BAC probe JA077K01 illuminates the Z amplicon, which is flanked by single-copy sequences represented by JA192C06 and JA004O24.



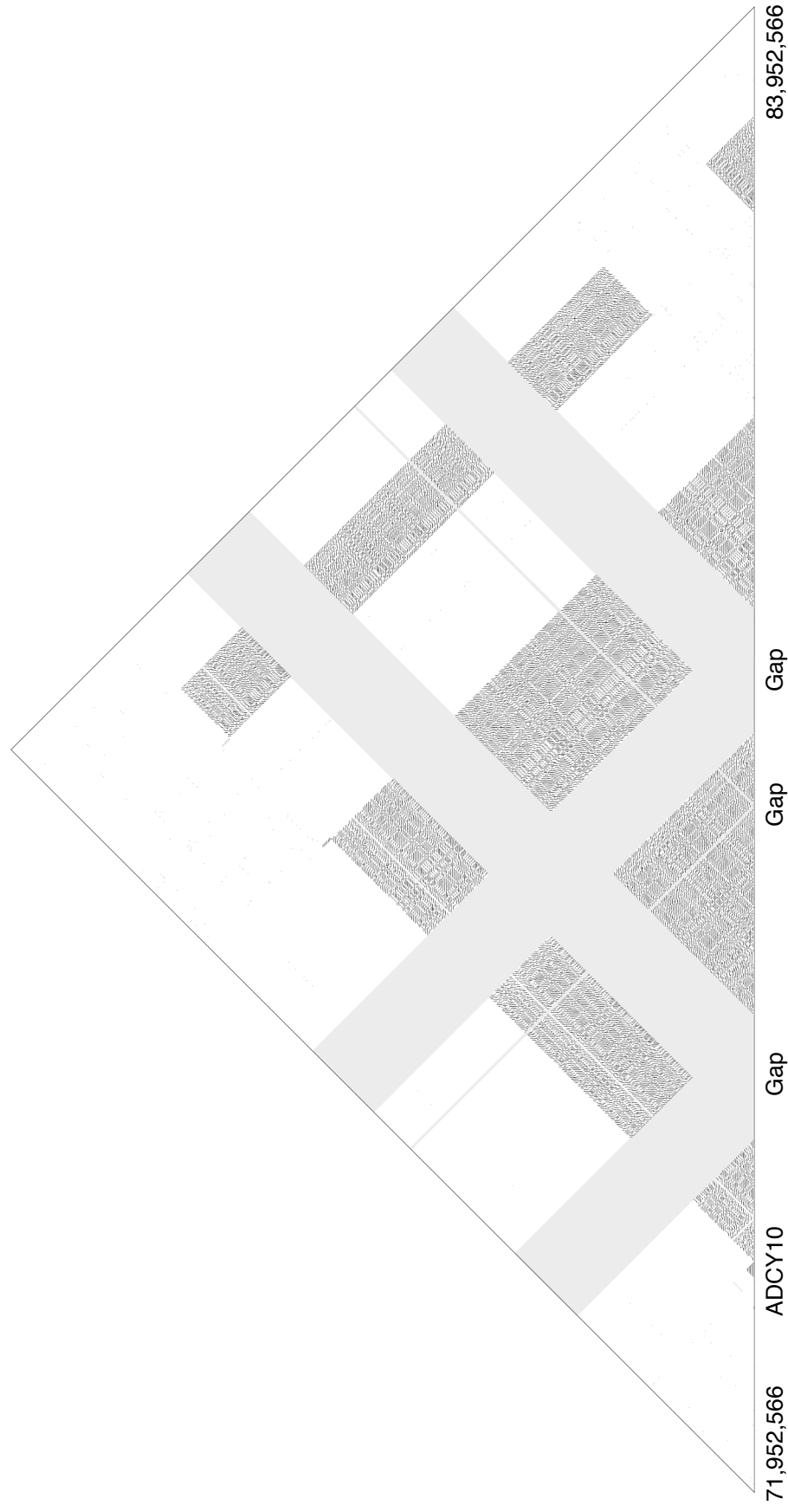
Supplementary Figure 3: Location of the Z chromosome centromere.

(a) Graph showing the retention frequency of Z-linked markers on ChickRH6 (red dots), LINE density (green line) and endogenous retrovirus density (blue line). Retention frequency peaks in the third contig on the Z chromosome. This peak coincides with a local peak in LINE and ERV density. (b-e) Schematic diagrams describing the locations of BAC probes (colored circles) within Z chromosome sequence contigs (blue lines). Each schematic is accompanied by a representative FISH image. Centromere position is shown by hashed lines. (b) Establishes the position of the centromere within the third contig. (c-d) Place BACs JA092O14 and JA168J14 on the short arm adjacent to the centromere. (e) Shows JA127N01 at a position indistinguishable from the centromere. A marker (WIBR04610) designed from a BAC end sequence of JA127N01 (GenBank: CC247852.1) is the marker with the highest retention frequency on the RH panel.



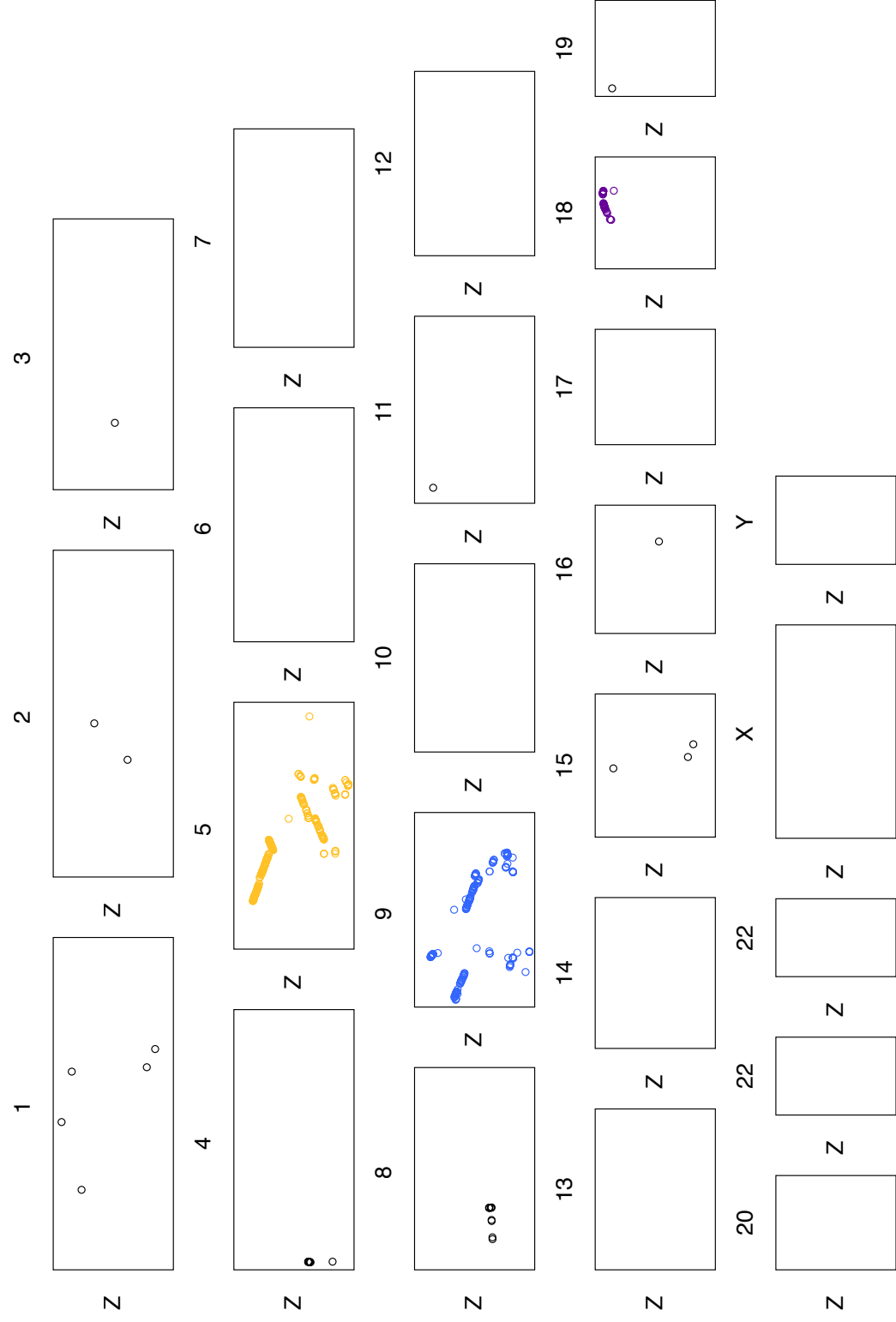
Supplementary Figure 4: Z amplicon FISH

Fluorescence in situ hybridization of Z amplicon containing BAC CH261-77N6 (red) to chicken chromosomes (blue). CH261-77N6 hybridizes specifically to distal Z long arm.



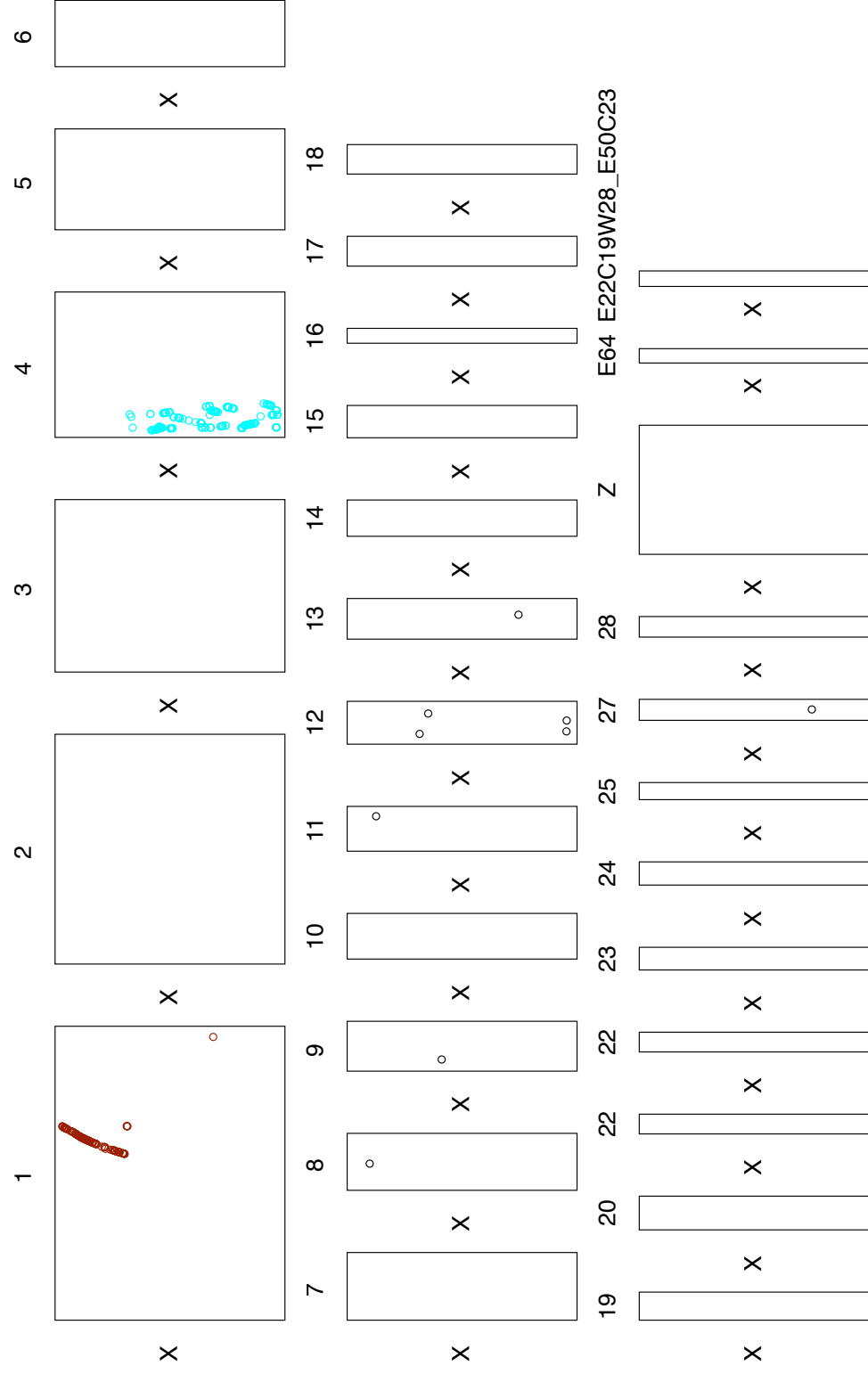
Supplementary Figure 5: Triangular dot plot of the Z amplicon region

This dot plot spans the entire Z amplicon region. Each dot represents a perfect match of 100 base pairs. Tandem repeats appear as horizontal lines, inverted repeats as vertical lines. Gaps in the sequence are shaded gray; the sequence contains two physical gaps where further chromosome walking was not possible, represented as 1Mb blocks, and a single gap in clone assembly estimated at 35kb. The orientation of sequences between the two physical gaps has not been determined. The Z amplicon array is interrupted by 2 large single-copy islands containing genes with orthologs on human autosomes 5 and 9. The ADCY10 array is visible as a small tandem array before the main body of the Z amplicon, with copies present at the beginning and end of the first single-copy island.



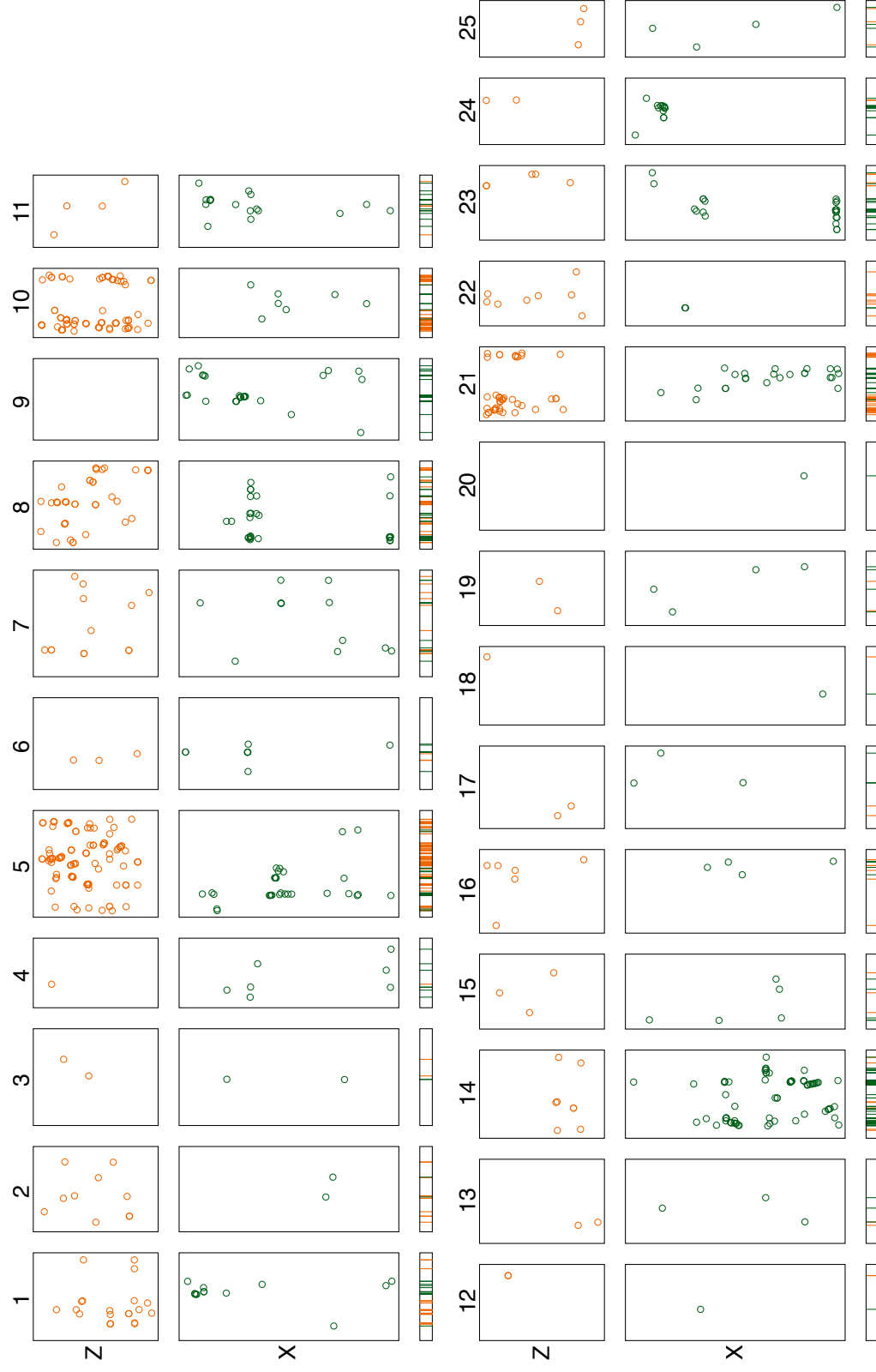
Supplementary Figure 6: Dot plots of chicken Z versus all human chromosomes

Square dot plots show the locations of 1:1 orthologs between two species. The chicken Z is orthologous to portions of human autosomes 5 (yellow), 9 (blue), and 18 (purple). Orthology with other chromosomes is limited to a handful of isolated genes.



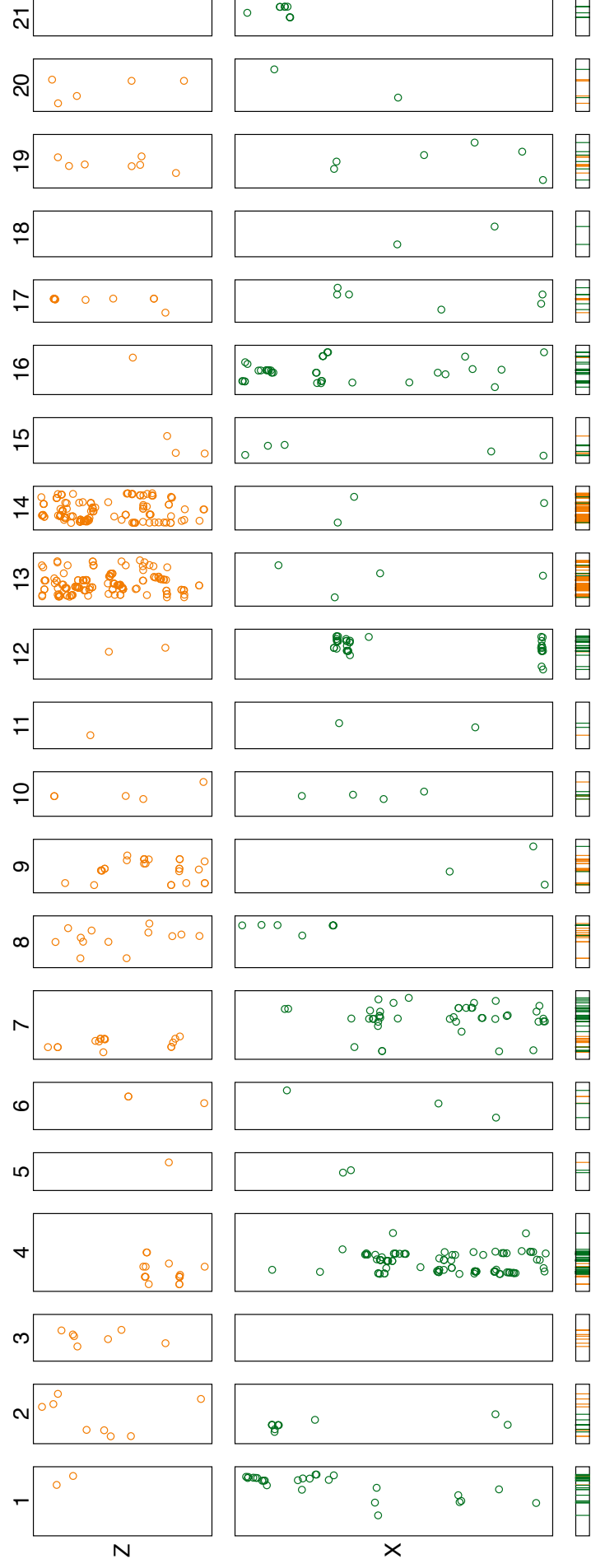
Supplementary Figure 7: Dot plots of human X versus all chicken chromosomes

Square dot plots show the locations of 1:1 orthologs between two species. The human X is orthologous to portions of chicken autosomes 1 (red) and 4 (cyan). Orthology with other chromosomes is limited to a handful of isolated genes.



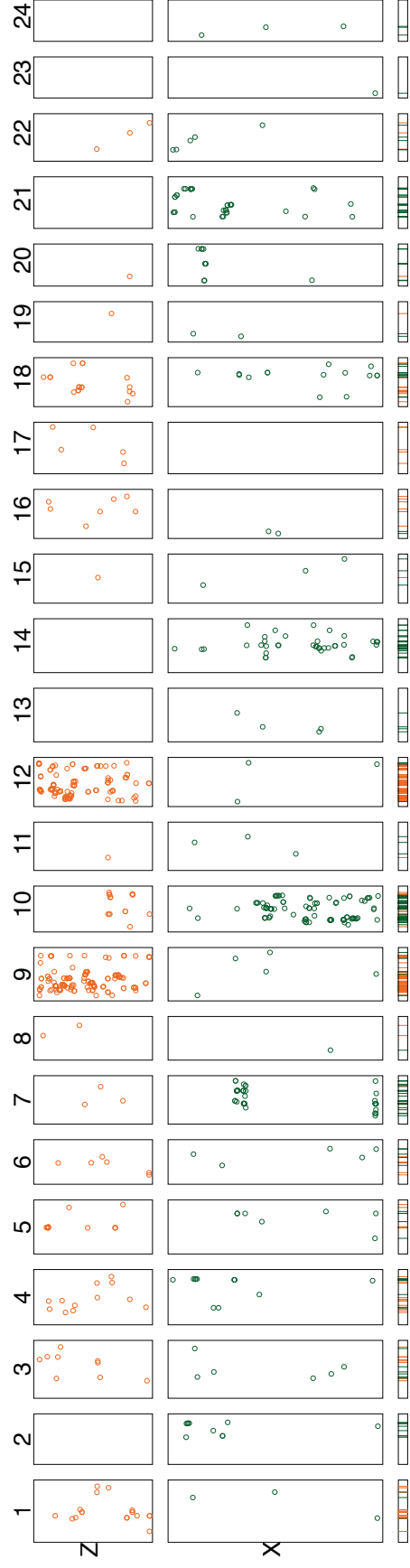
Supplementary Figure 8: Dot plots of chicken Z and human X versus all zebrafish chromosomes

Square dot plots show the locations of 1:1 orthologs between two species. These locations are projected onto bars to show the relative contribution of chicken Z and human X orthologs. Orthologs of genes on the chicken Z (orange) are concentrated on zebrafish chromosomes 5, 10, and 21. Orthologs of genes on the human X (green) are concentrated on zebrafish chromosomes 9 and 14. Other chromosomes contain scattered Z and X orthologs.



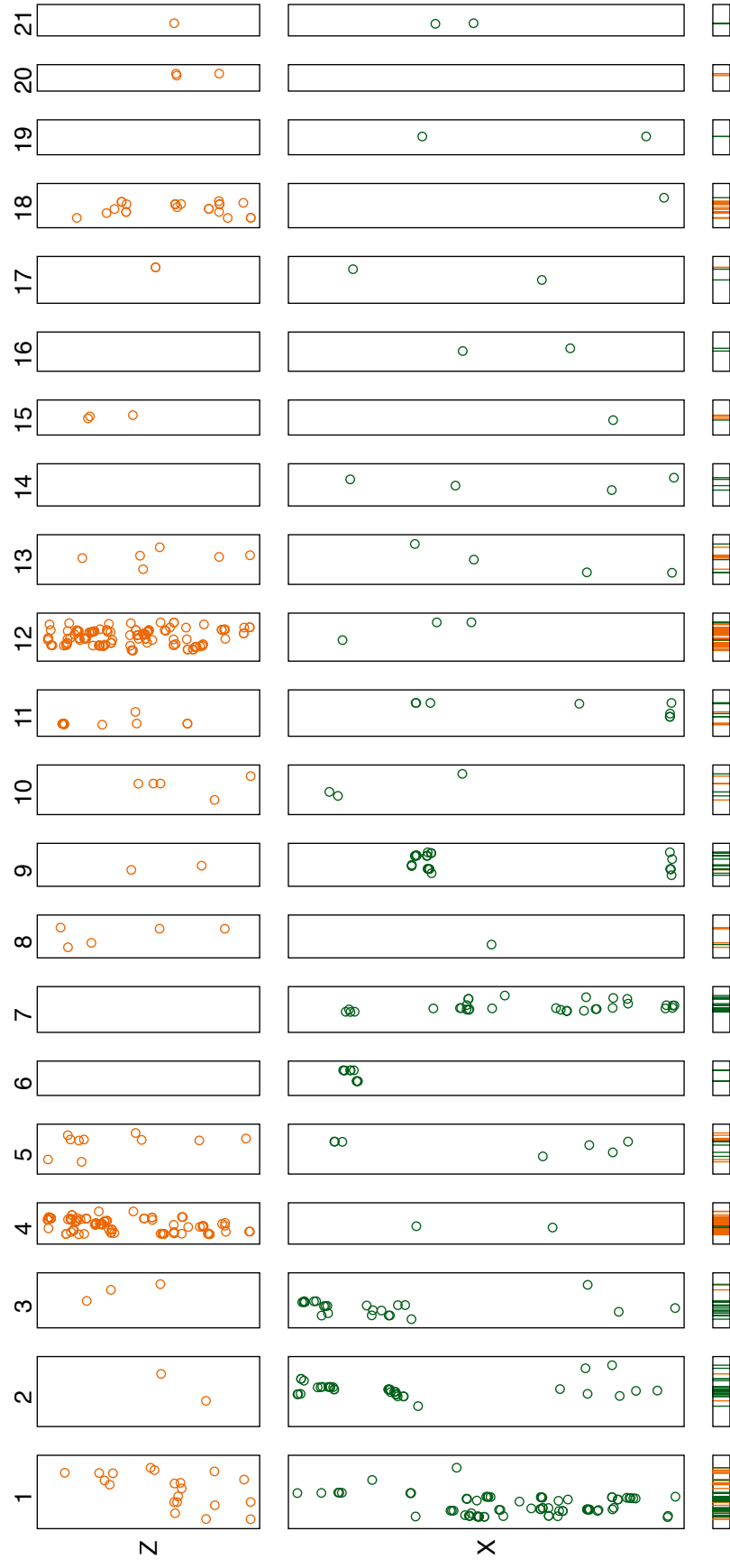
Supplementary Figure 9: Dot plots of chicken Z and human X versus all stickleback chromosomes

Square dot plots show the locations of 1:1 orthologs between two species. These locations are projected onto bars to show the relative contribution of chicken Z and human X orthologs. Orthologs of genes on the chicken Z (orange) are concentrated on stickleback chromosomes 13 and 14. Orthologs of genes on the human X (green) are concentrated on stickleback chromosomes 1, 4, 7, and 16. Other chromosomes contain scattered Z and X orthologs.



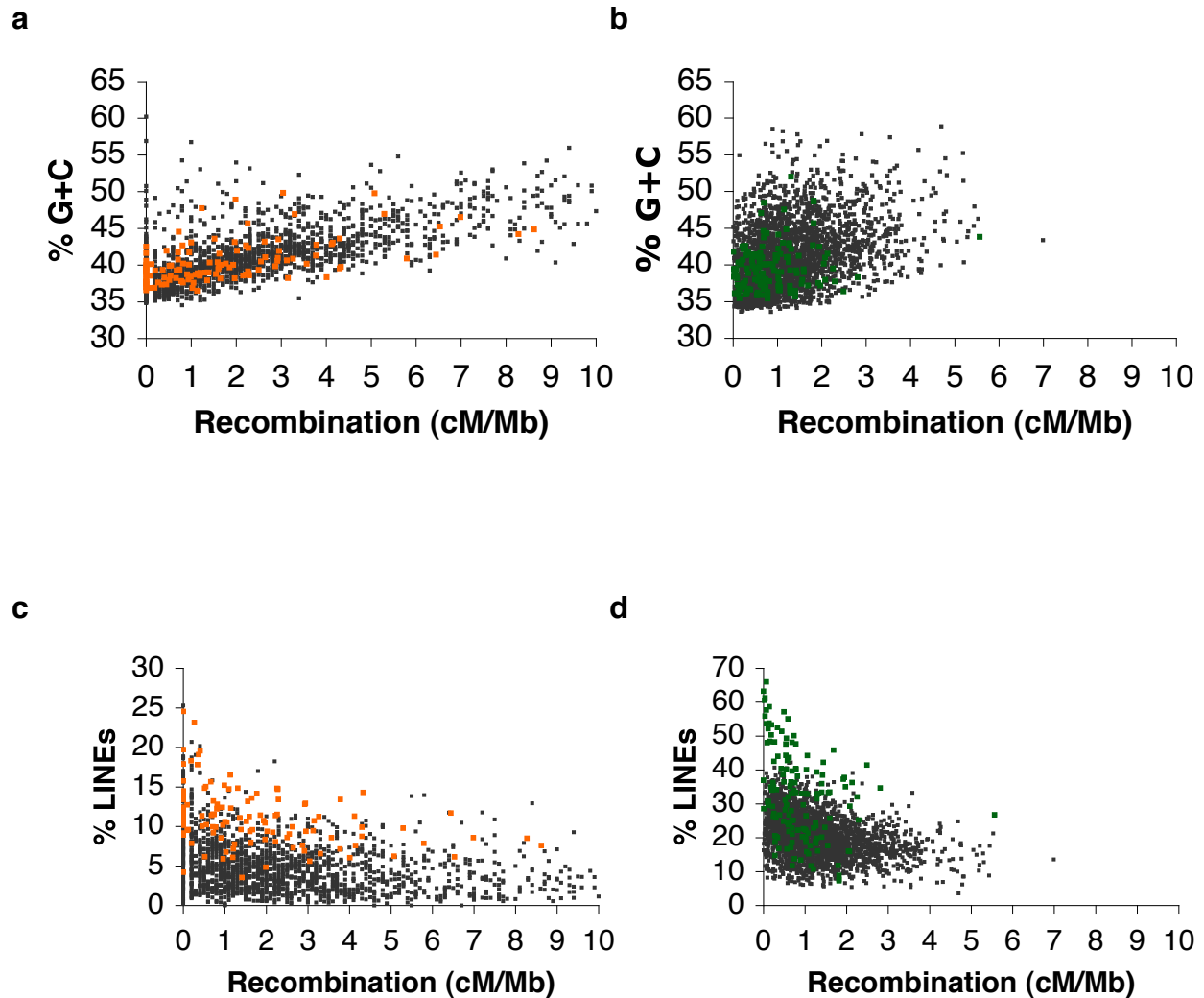
Supplementary Figure 10: Dot plots of chicken Z and human X versus all medaka chromosomes

Square dot plots show the locations of 1:1 orthologs between two species. These locations are projected onto bars to show the relative contribution of chicken Z and human X orthologs. Orthologs of genes on the chicken Z (orange) are concentrated on medaka chromosomes 9 and 12. Orthologs of genes on the human X (green) are concentrated on medaka chromosomes 10, 14, and 21. Other chromosomes contain scattered Z and X orthologs.



Supplementary Figure 11: Dot plots of chicken Z and human X versus all pufferfish chromosomes

Square dot plots show the locations of 1:1 orthologs between two species. These locations are projected onto bars to show the relative contribution of chicken Z and human X orthologs. Orthologs of genes on the chicken Z (orange) are concentrated on pufferfish chromosomes 4 and 12. Orthologs of genes on the human X (green) are concentrated on pufferfish chromosomes 1, 2, 3, and 7. Other chromosomes contain scattered Z and X orthologs.



Supplementary Figure 12: Recombination, G+C content, and LINE density

Each point represents an interval of sequence bounded by adjacent markers on the chicken or human genetic map. G+C content is positively correlated with recombination rate in both the chicken (a) and human (b) genomes. The G+C content of intervals on the Z chromosome (orange) and X chromosome (green) are similar to the G+C content of autosomal intervals (grey) with the same recombination rate.

LINE density is inversely correlated with recombination rate in both chicken (c) and human (d) genomes. Intervals on the Z chromosome (orange) and X chromosome (green) are enriched in LINES compared to autosomes (grey). For additional information, see Supplementary Note 1.