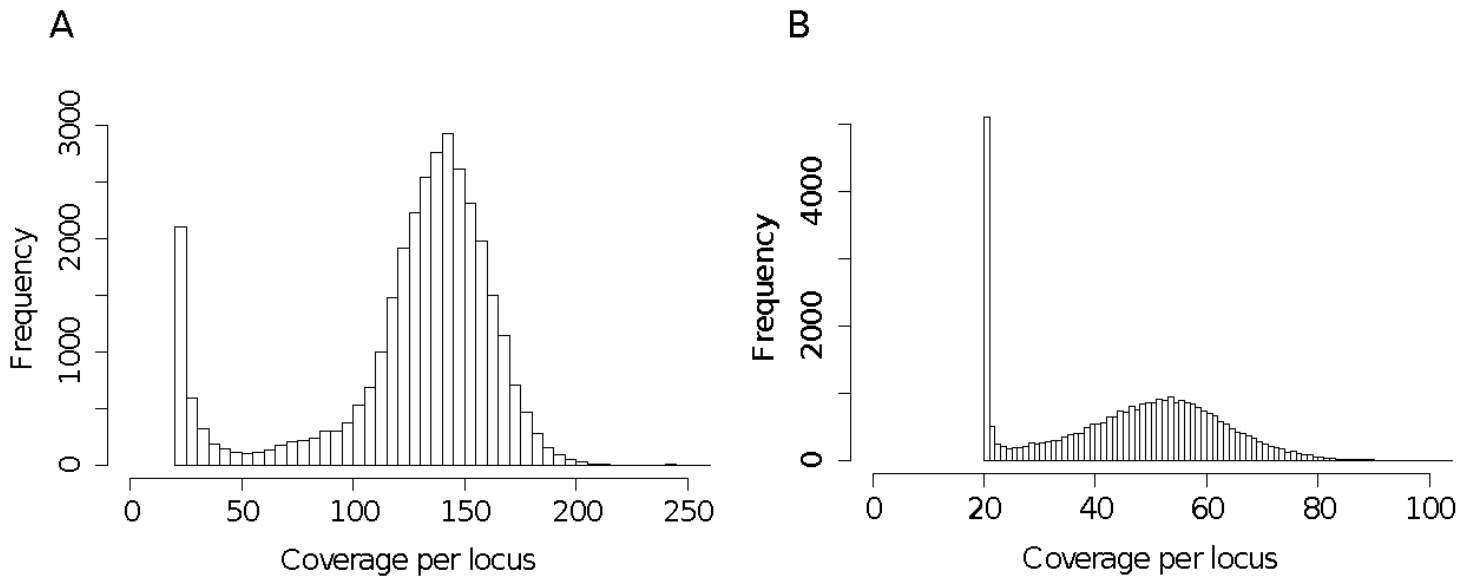**Supplemental Figure 1. Identification of putative introgression candidates based on divergence.** The figure shows the distribution of the proportion of the total external branch length accounted for by two neighboring external branches for each of the 30,907 trees.
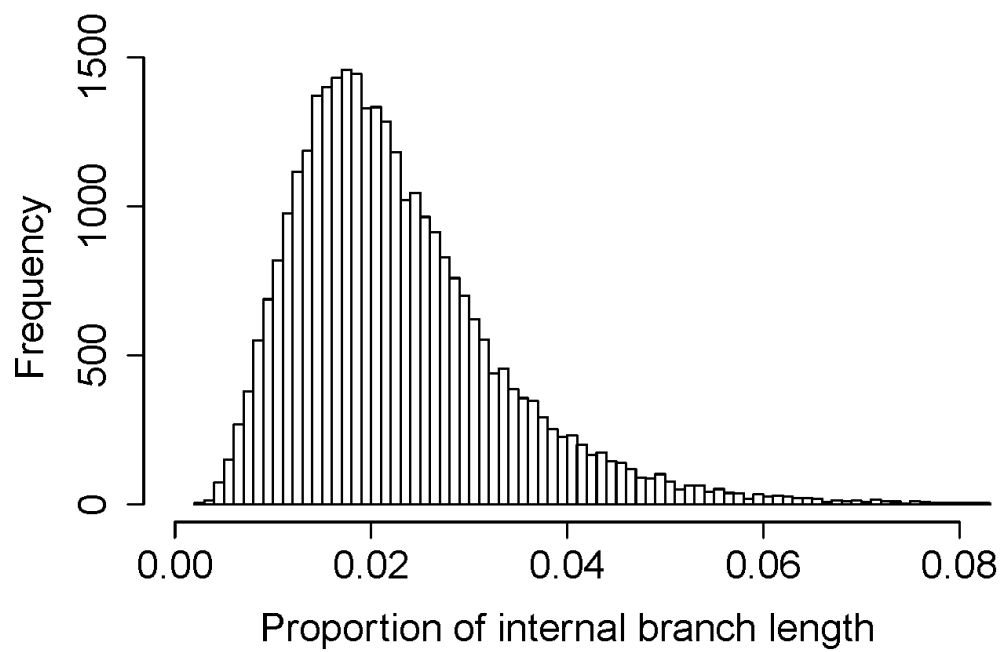
Supplemental Figure 2. Distribution of the sequence coverage per locus. The histograms show the average coverage (x-axis, in reads per base pair) of each locus assembled for (A) *Anopheles farauti* 4 and (B) *Anopheles punctulatus.* Note that any base pairs with < 20X are not shown.
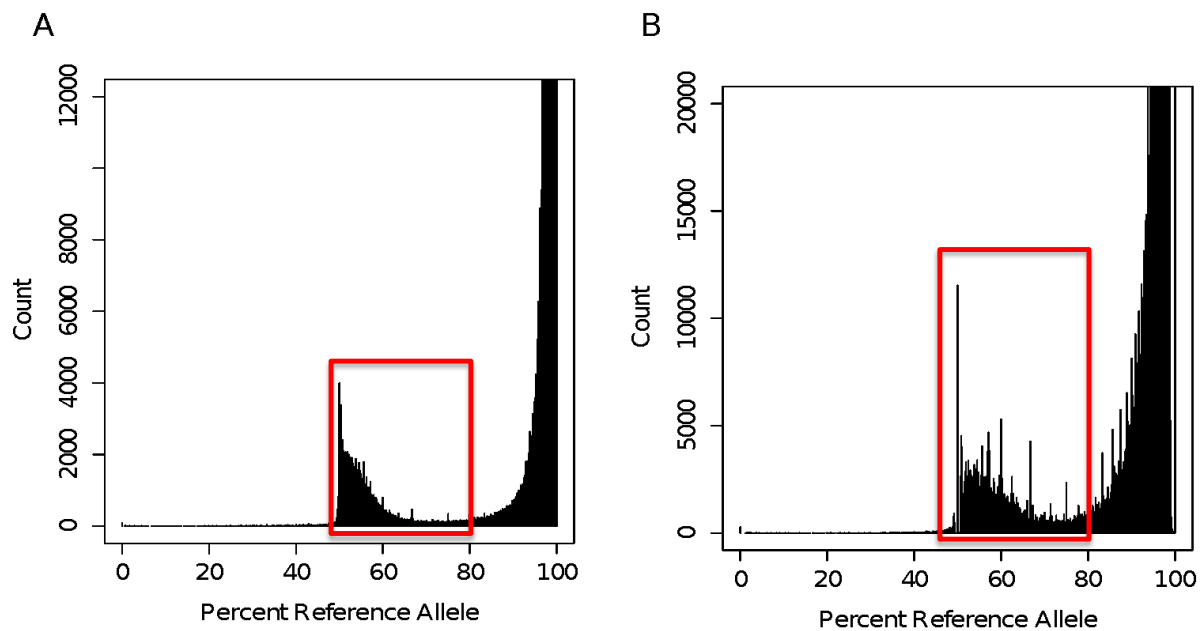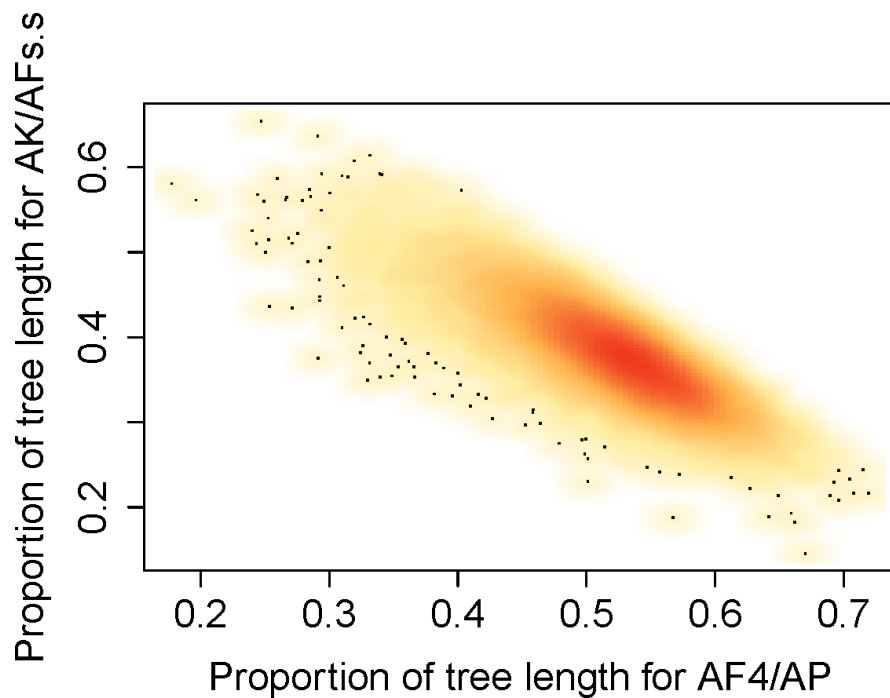
Supplemental Figure 3. Distribution of the reference allele frequencies. The histograms show the proportion of reads carrying the reference allele (x-axis, in percent) at each nucleotide position of the (A) *Anopheles farauti* 4 and (B) *Anopheles punctulatus* assemblies. The red boxes indicate sites that were identified as heterozygous in each species.
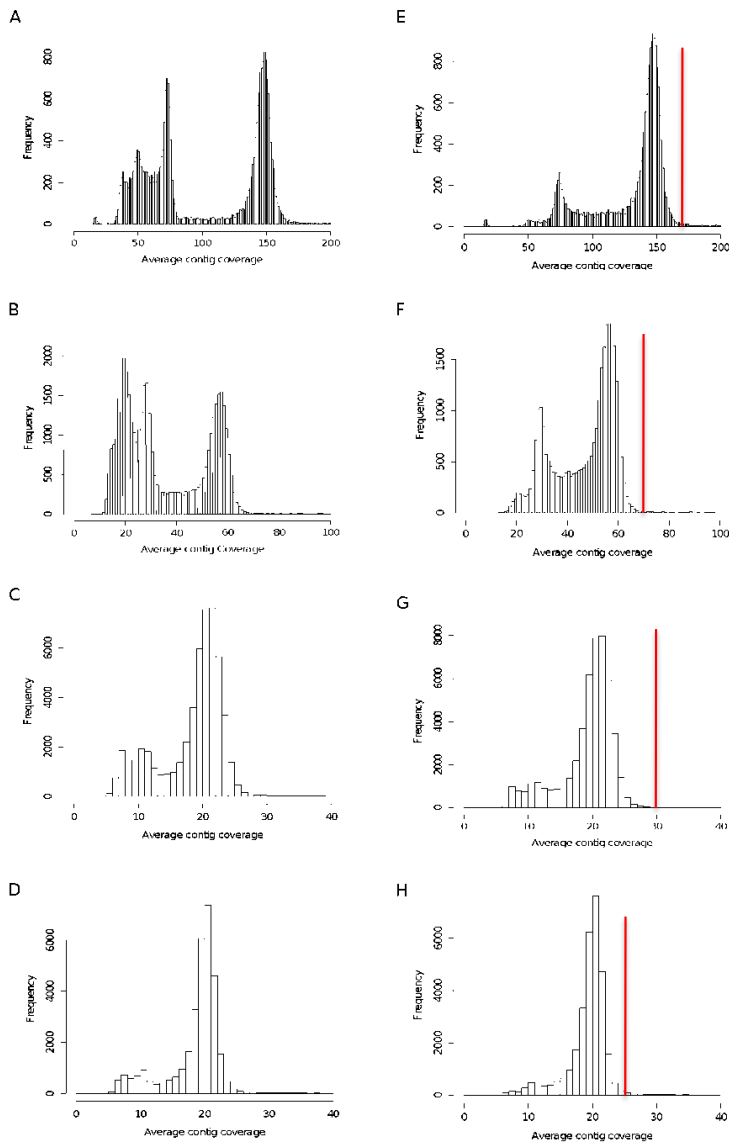
Supplemental Figure 4. Distribution of the sequence coverage per contig for the initial assemblies (A-D) and after removing redundant contigs (E-H). The histograms display the data for *Anopheles farauti* 4 (A and E), *Anopheles punctulatus* (B and F), *Anopheles koliensis* (C and G) and *Anopheles farauti* s.s. (D and H). The red line represents the cut off used to filter out putative paralogs from the assembly. See legend of Supplemental Figure 2 for details.

Supplemental Figure 5. Distribution of the internal branch lengths. The figure shows the distribution of the proportion of the internal branch length for each of the 30,907 trees.

Supplemental Figure 6. Distribution of the external branch lengths. The density plot shows the proportion of each tree total length accounted for by the length of the external branches separating AK and AFs.s. (y-axis, in percent) and by the length of the external branches separating AF4 and AP.
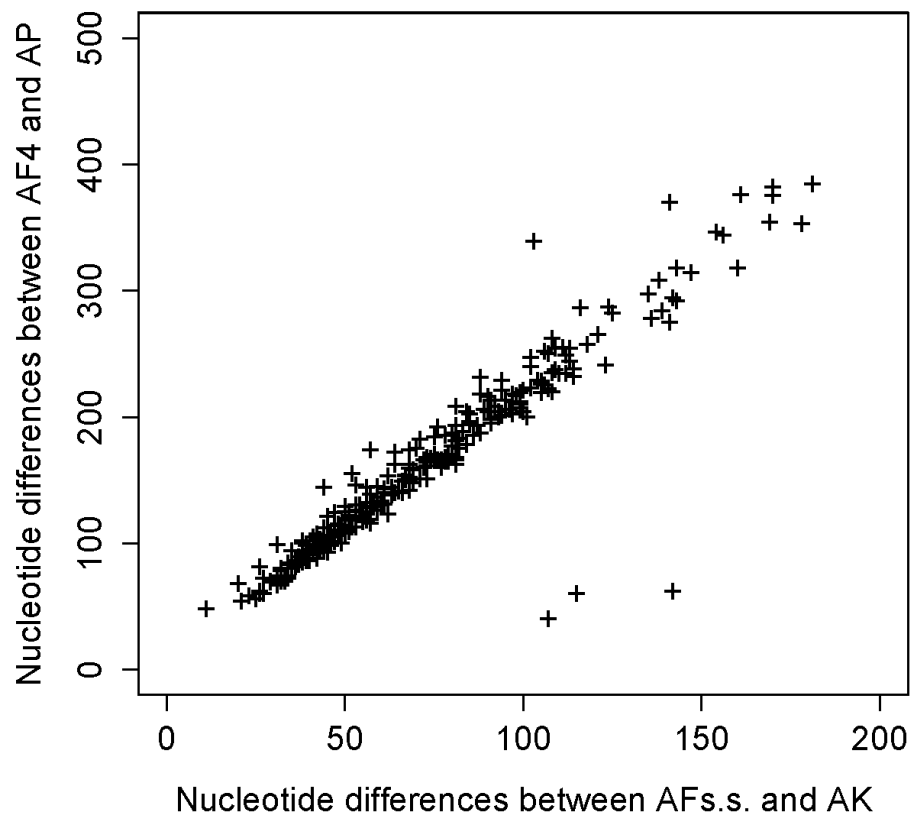
Supplemental Figure 7. Distribution of nucleotide differences between species. The plot shows the number of nucleotide differences of each introgression candidate locus between *Anopheles farauti* 4 and *Anopheles punctulatus* (y-axis) and *Anopheles farauti s.s.* and *Anopheles koliensis* (x-axis). Note that the three outliers in the bottom right corner are driven by one disproportionally long branch length in either AFs.s. or AK (not by two short branches) and are not indicative of introgression (but may indicate positive selection on one of the four lineages).
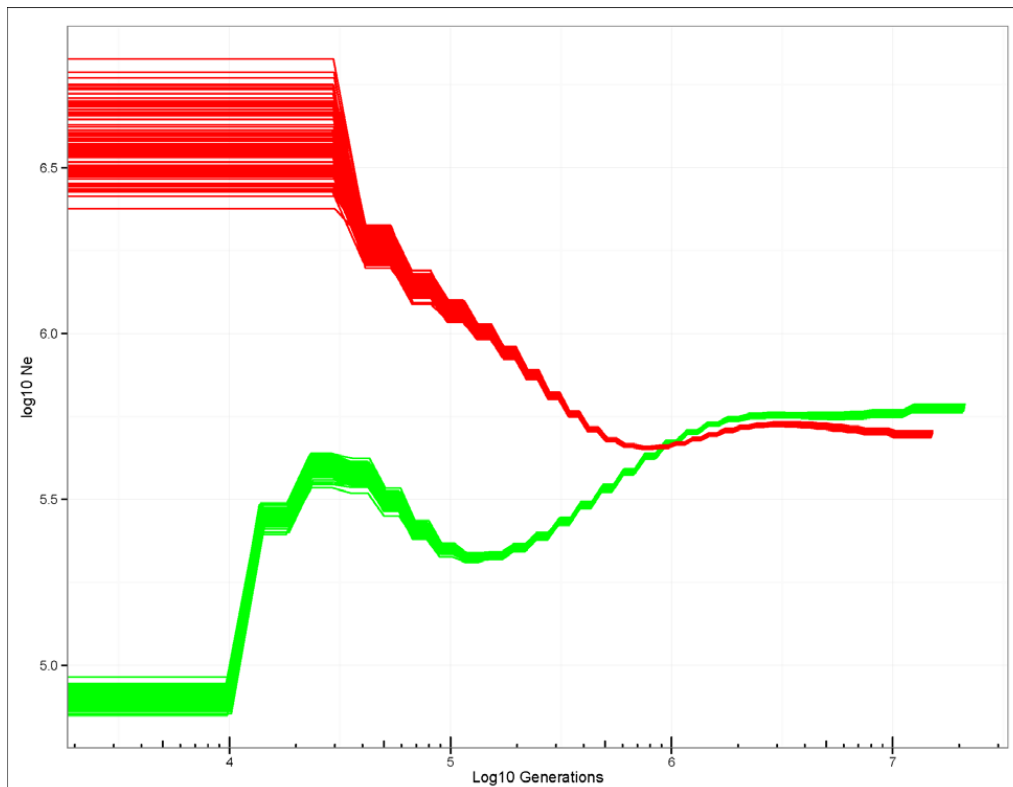
Supplemental Figure 8. Demographic history of AF4 and AP. The figure shows the estimated historical variations in effective population size for AF4 (green) and AP (red) based on PSMC analyses. The composite green (AF4) and red (AP) lines represent 100 bootstrap replicates. The y-axis represents the effective population size (in log10 of Ne) and the x-axis represents the pairwise sequence divergence (in log10 generations), assuming a mutation rate of 2.8×10^{-9} per base pair per generation (Keightley *et al.* 2014). Note that this mutation rate is from *Drosophila* and may not accurately reflect the life history of *Anopheles* mosquitoes.
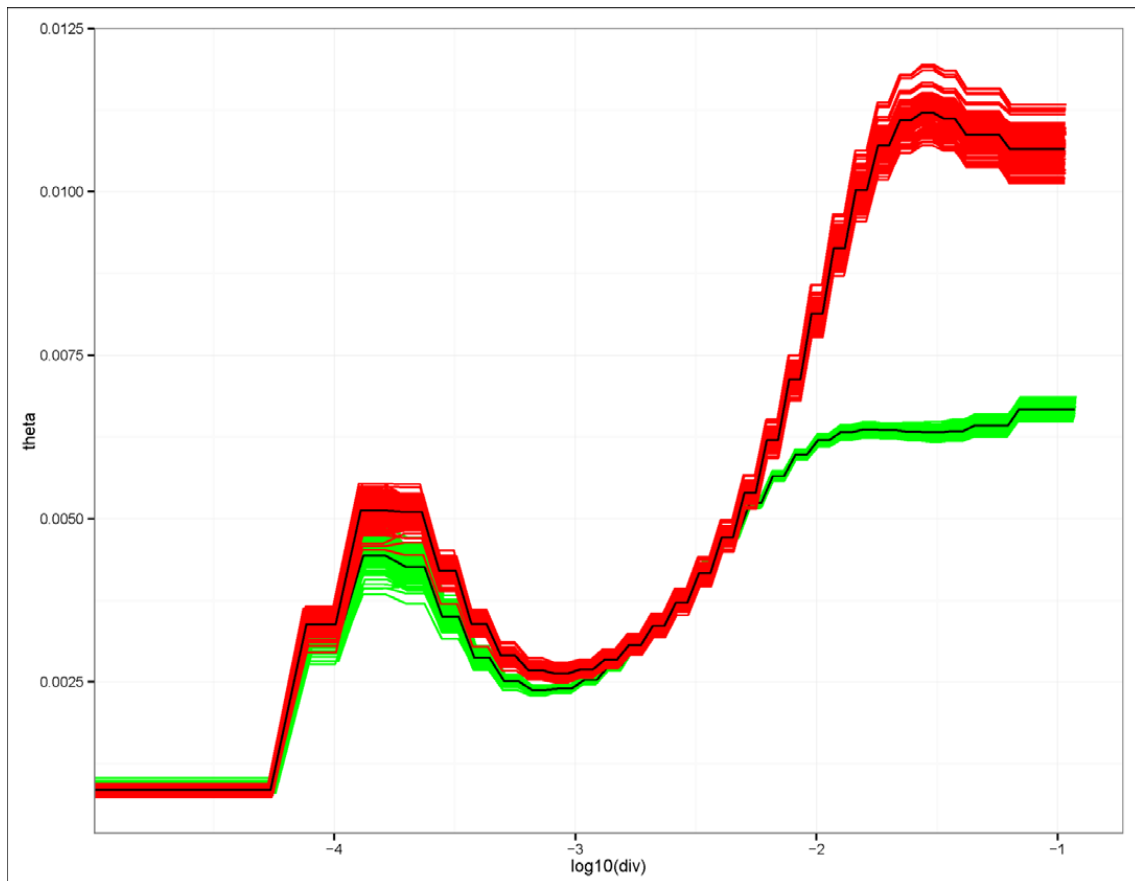
Supplemental Figure 9. Influence of sequence coverage on the estimated demographic history of AF4. The figure shows the estimated historical variations in effective population size for AF4 based on the entire dataset (131X, in green) or the down-sampled dataset (51X, in red). The composite green and red lines represent 100 bootstrap replicates.

Supplemental Table 1. Sequencing summary for each sample.

Sample	Instrument	Read length	Number of read pairs
An. farauti 4	HiSeq 2000	100	170,493,404
An. punctulatus s.s.	HiSeq 2000	100	64,481,884
An. farauti s.s.	GAIIx	51	44,281,276
An. koliensis	GAIIx	58	37,073,523

Supplemental Table 2. Summary of the initial *de novo* genome assemblies.

Sample	# ctgs	Size (kb)	N50	Median	Max
An. farauti 4	451,824	261,977	13,469	105	331,681
An. punctulatus s.s.	927,145	293,761	8,131	117	97,012
An. farauti s.s.	389,743	193,639	7,880	692	79,463
An. koliensis	484,838	204,191	3,691	311	76,110

Supplemental Table 3. Summary of the *Anopheles. farauti* 4 assembly after sub-sampling sequencing reads to the coverage of *Anopheles punctulatus* s.s.

Sample	Expected Coverage*	Size (kb)	# of Contigs	% Assembled	N50	Max (bp)	Median (bp)
An. farauti 4	131 X	146,386	14,407	63.1	16,229	331,681	6,280
An. farauti 4 sub-sampled	50X	149,618	12,971	66.8	18,553	331,629	7,271
An. punctulatus s.s.	50X	146,190	20,775	62.9	10,258	97,012	5,136

*Based on genome size of *Anopheles gambiae* (260MB)

Supplemental Table 5. Summary of the genetic diversity in *Anopheles farauti* 4 estimated after sub-sampling *Anopheles farauti* 4 reads to the coverage of *Anopheles punctulatus* s.s.

	Length	Heterozygous sites in AF4	Heterozygous sites in AP	Divergence	# of shared polymorphisms
Total	50,812,770	196,731 0.38%	312,996 0.62%	3,969,250 7.81%	492
Per Locus	2,301 [1,000-17,020]	8.91 [0-112]	14.17 [0-138]	179.7 [14-1,543]	0.02 [0-2]
Per Kb	NA	0.009 [0.0-0.1]	0.014 [0.0-0.14]	0.18 [0.01-1.5]	2.20E-05 [0-2.0e-3]

[] represent the minimum and maximum values observed