

Supplemental Figures

Supplemental Figure 1: Using log likelihood difference to select optimal number of chromatin states. (A) The corresponding log likelihoods for HMMs fit on the observed data (red) and the average log-likelihood of the models fit on the permuted data (blue) as a function of the total number of chromatin states. (B) The difference between the two curves in (A). The final model includes 15 states (indicated by the red vertical line), which is where the log likelihoods and their difference begin to taper.

Supplemental Figure 2: Proportion of the truncated genome assigned to each combinatorial histone modification pattern out of a possible 32. The red horizontal line represents 0.95. Note that 95% of the genome is covered by 16 patterns and 94% is covered by 15 patterns.

Supplemental Figure 3: Comparison of the performance for different model configurations. Error bars indicate the standard error of the estimated log-likelihood.

Supplemental Figure 4: The proportion of different bin-level states in the truncated genome. Colors of each state correspond to domain membership as in Figure 1.

Supplemental Figure 5: The proportion of different bin-level states in the intergenic regions. Colors of each state correspond to domain membership as in Figure 1.

Supplemental Figure 6: The proportion of different bin-level states in lamina-associated domains. Colors of each state correspond to domain membership as in Figure 1.

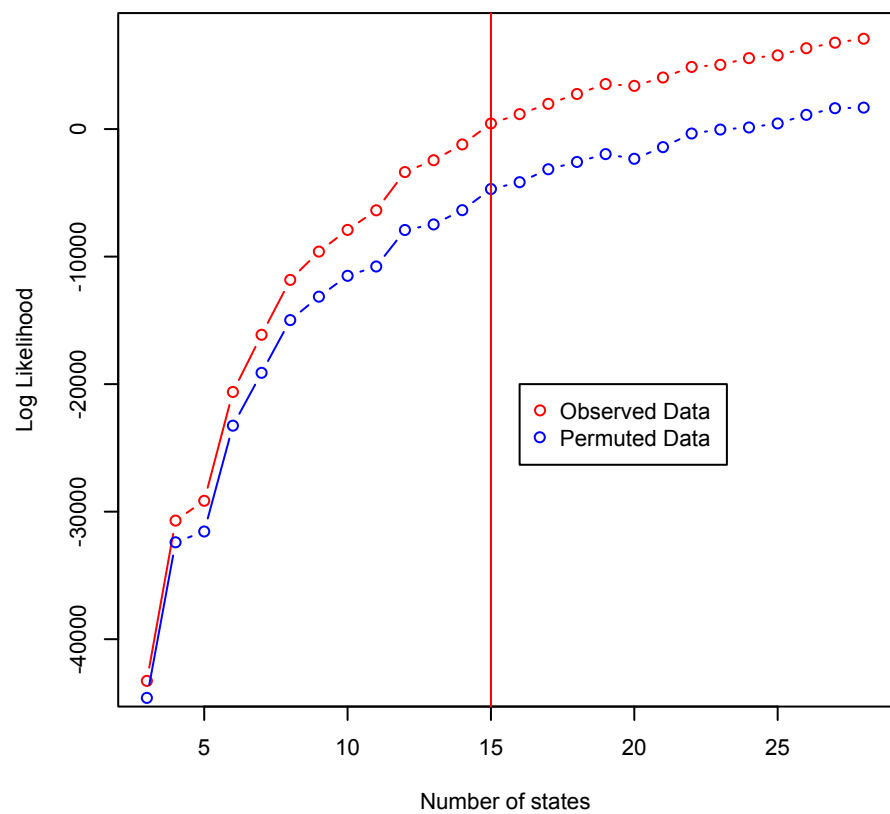
Supplemental Figure 7: The average RNAseq counts for each bin-level state in the intergenic regions. Vertical lines represent the mean value for each domain-type (red – active; blue non-active; yellow – null). The bottom track is the average over the whole intergenic regions. Colors of each state correspond to domain membership as in Figure 1.

Supplemental Figure 8: The average RNAseq counts corresponding to lincRNAs for each bin-level state in the intergenic regions. Vertical lines represent the mean value for each domain-type. Vertical lines represent the mean value for each domain-type (red – active; blue non-active; yellow – null). The bottom track is the average over the whole intergenic regions.

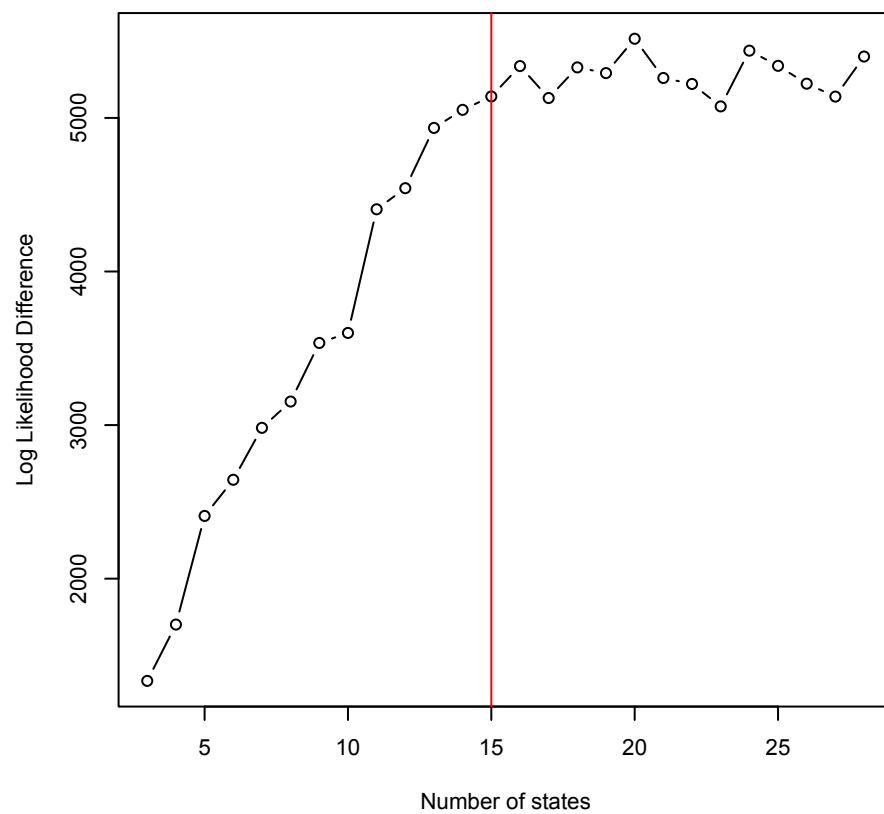
Supplemental Figure 9: The enrichment of active domains for different subtypes of pseudogenes. "IC C" – "IG_C_pseudogene"; "IG J" – "IG_J_pseudogene"; "IG V" – "IG_V_pseudogene"; "TR V" – "TR_V_pseudogene"; "DEL" – "deleted"; "PRO" – "processed_pseudogene"; "PS" – "pseudogene"; "TR P" – "transcribed_processed_pseudogene"; "TR UN" – "transcribed_unprocessed_pseudogene"; "U" – "unitary_pseudogene"; "UN" – "unprocessed_pseudogene"

Supplemental Figure 1

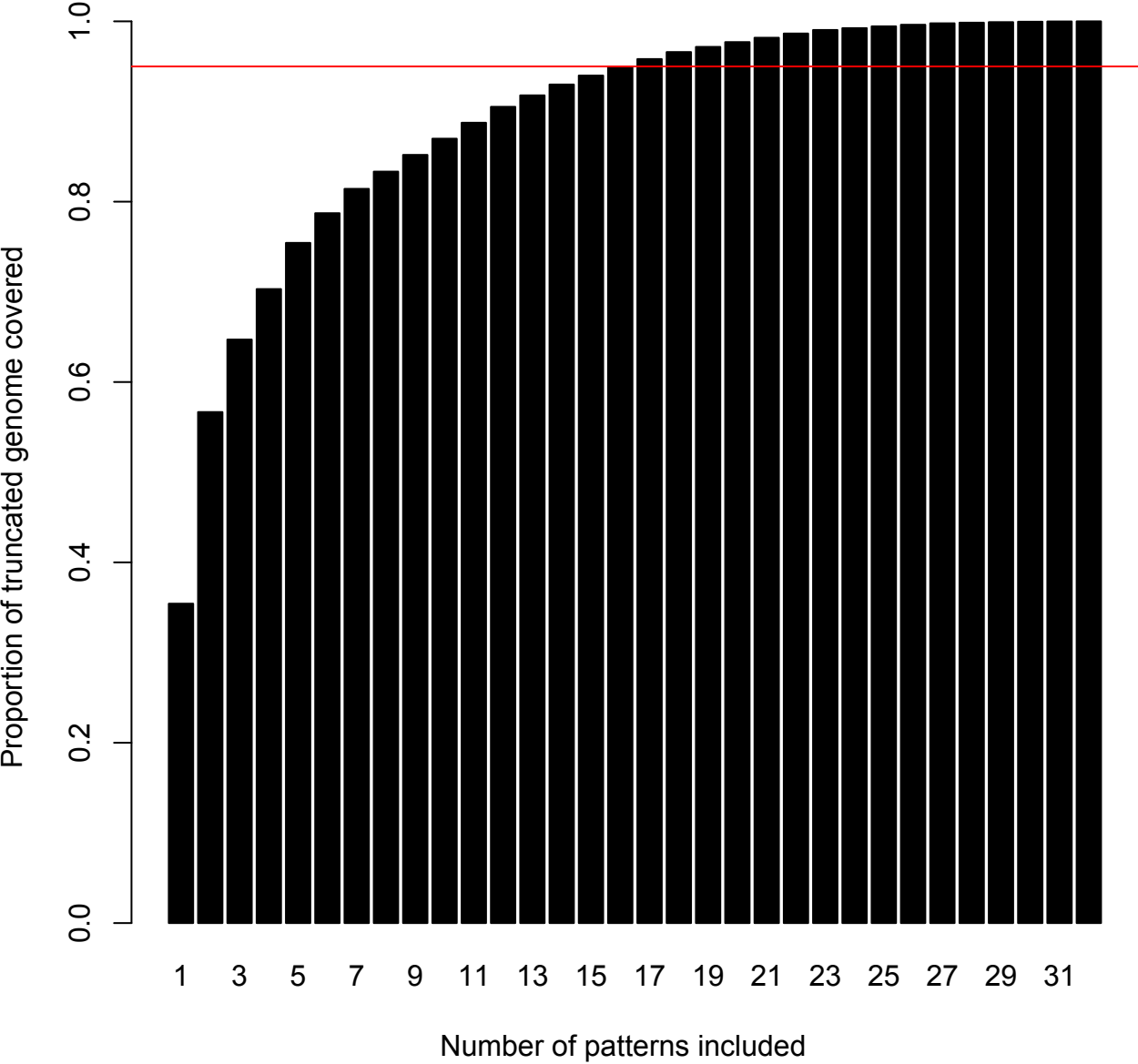
(A)



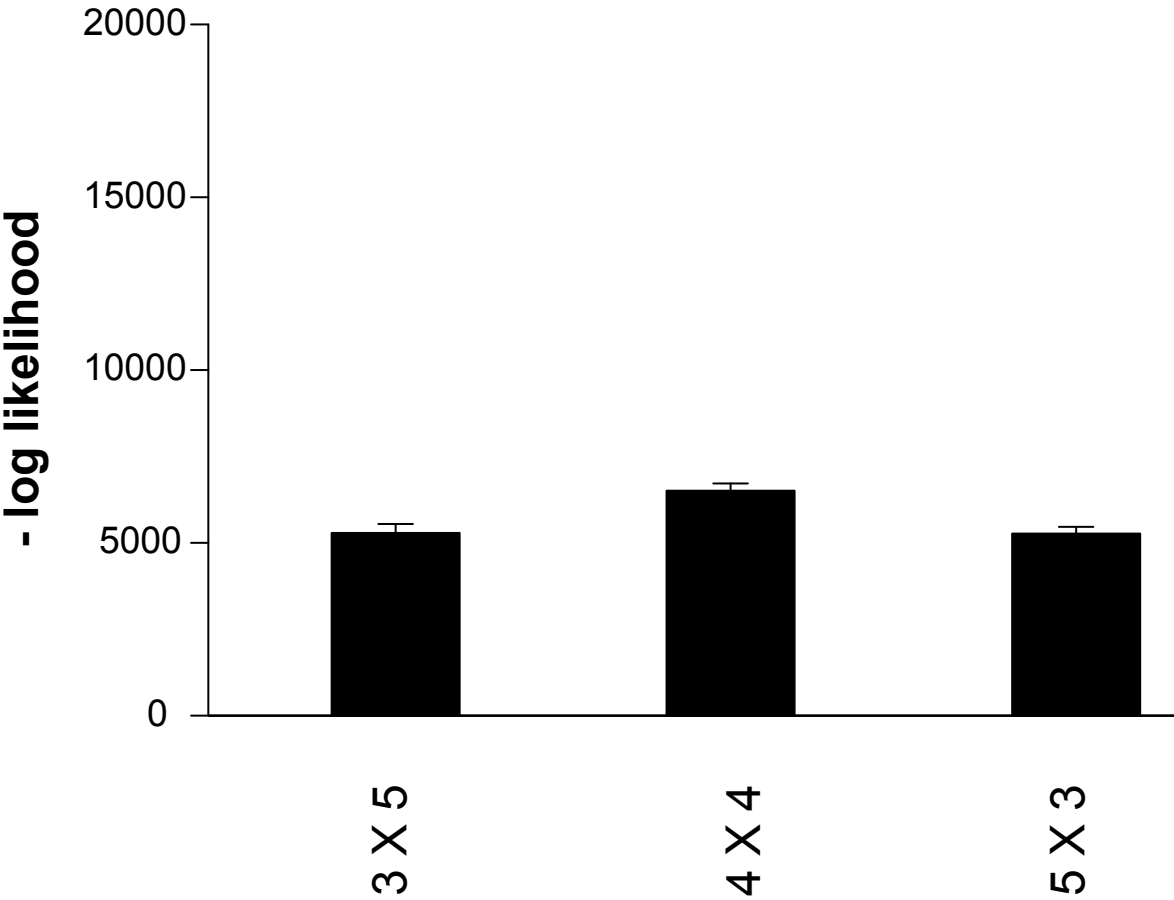
(B)



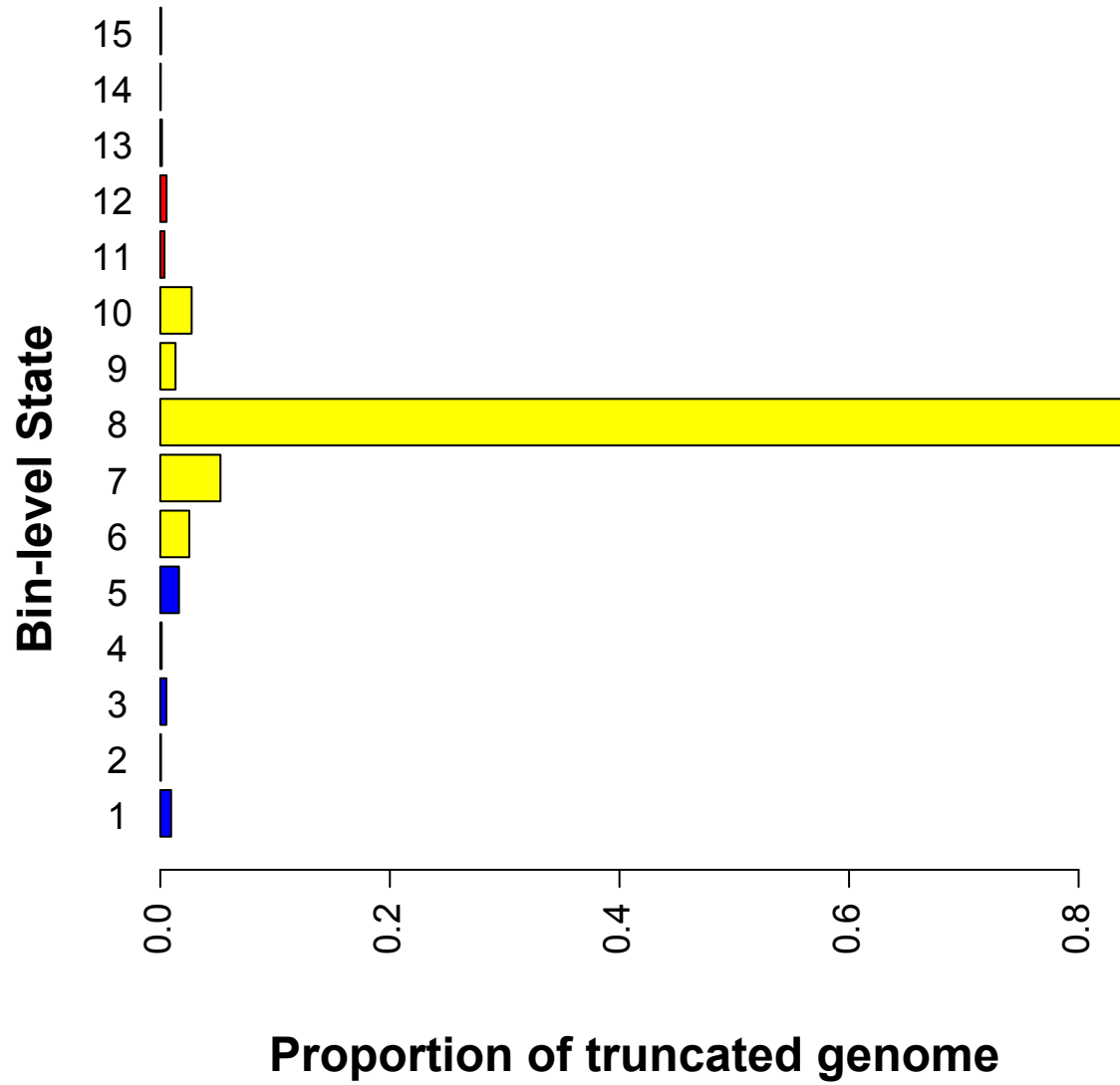
Supplemental Figure 2



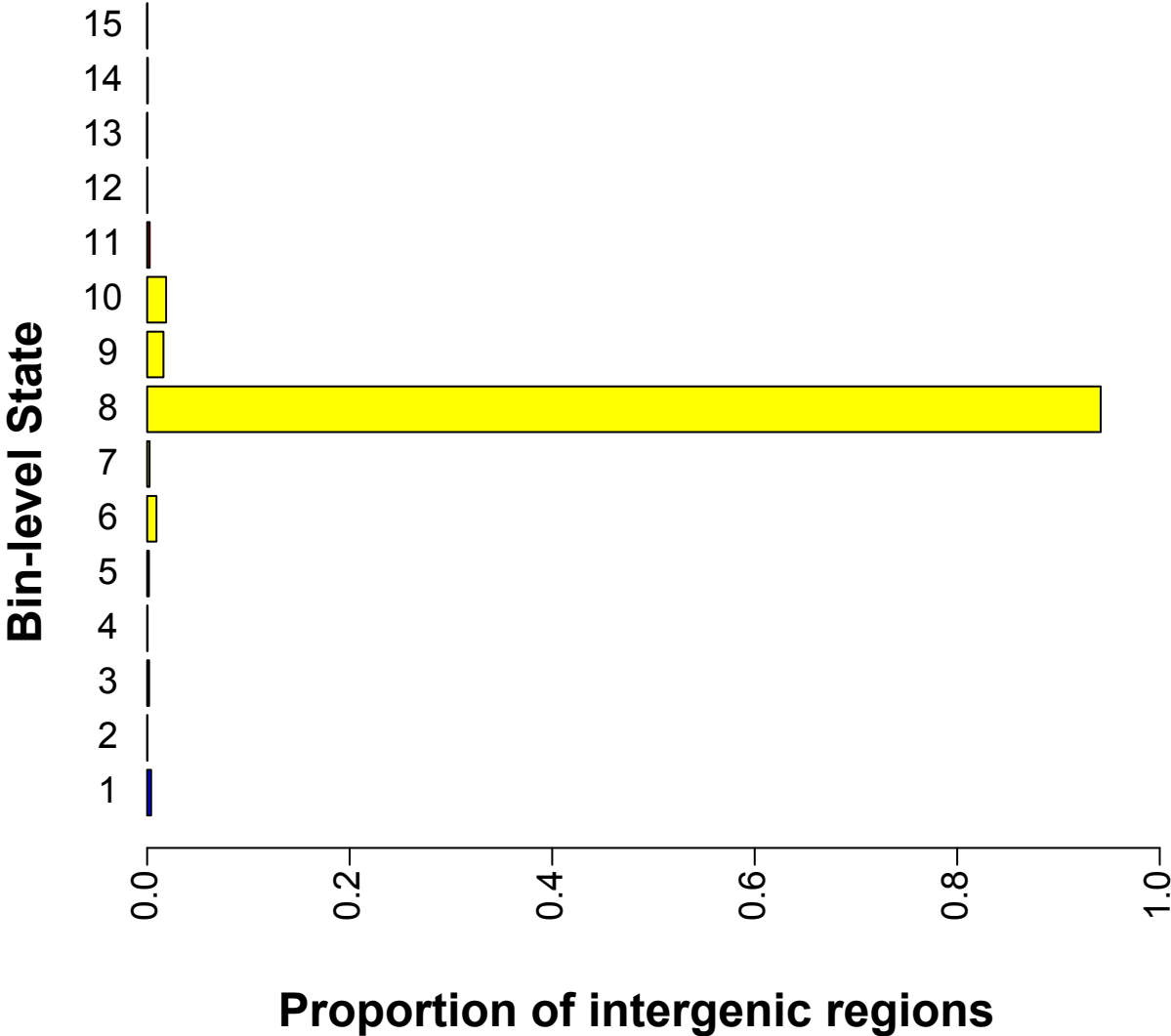
Supplemental Figure 3



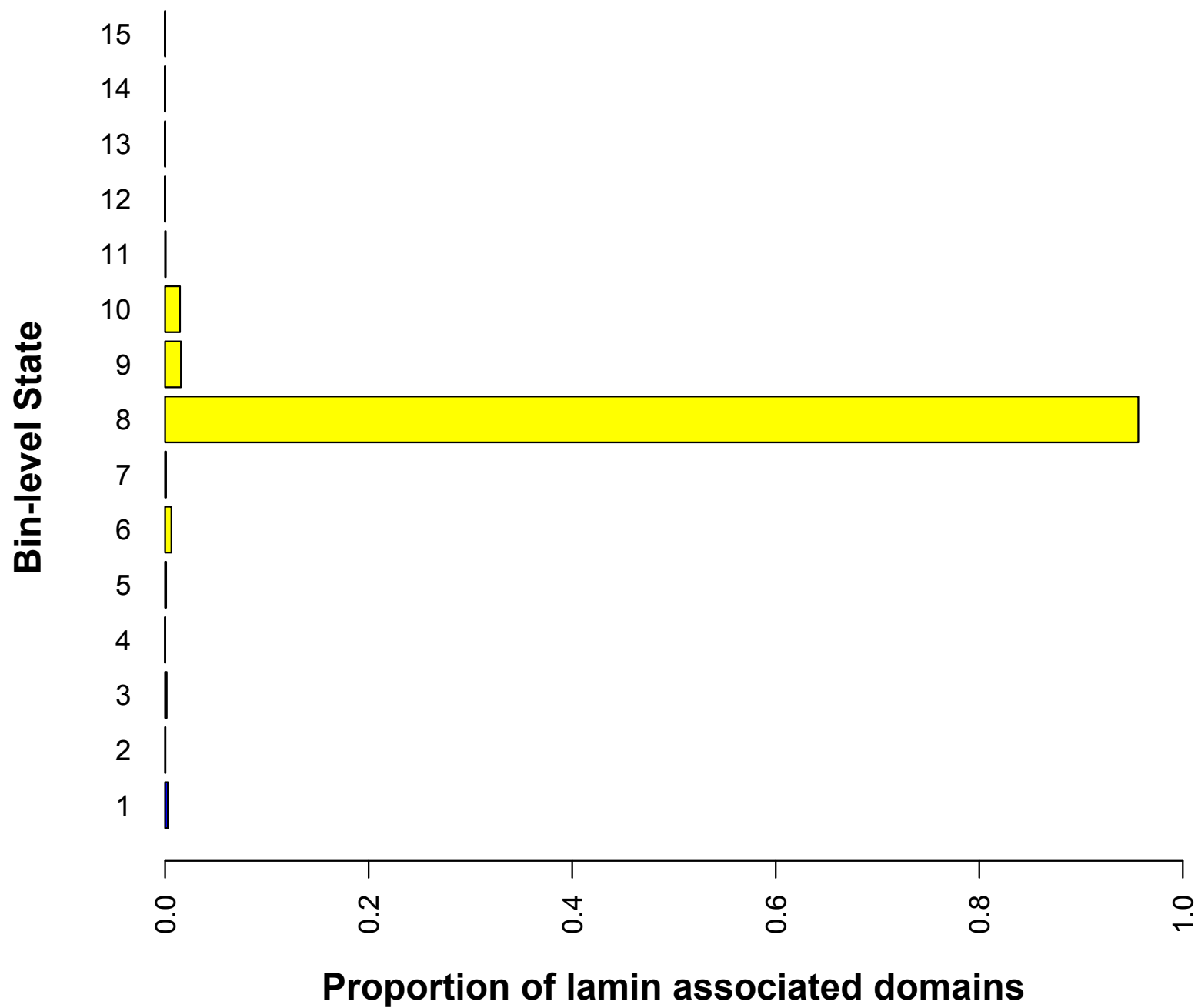
Supplemental Figure 4



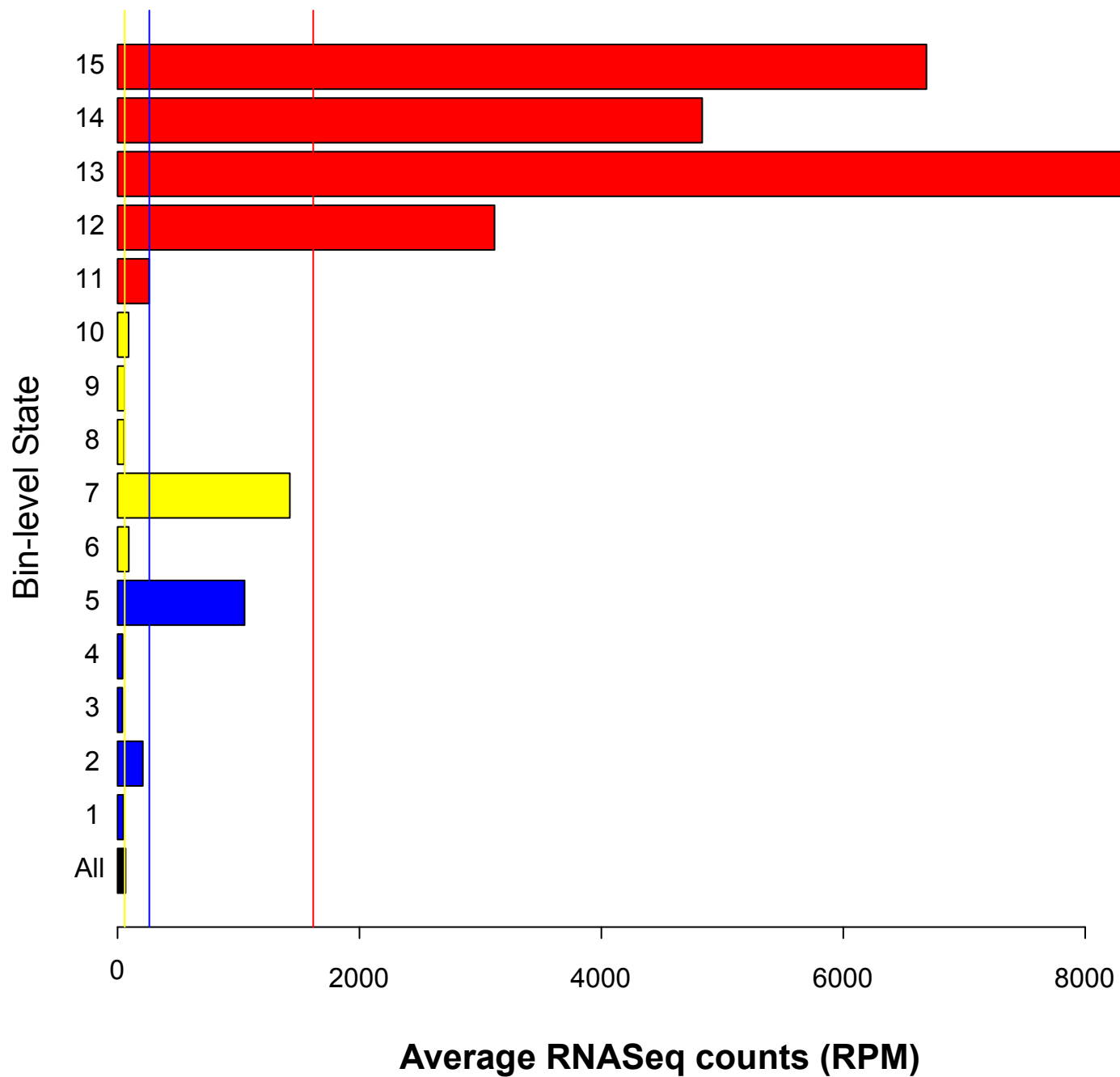
Supplemental Figure 5



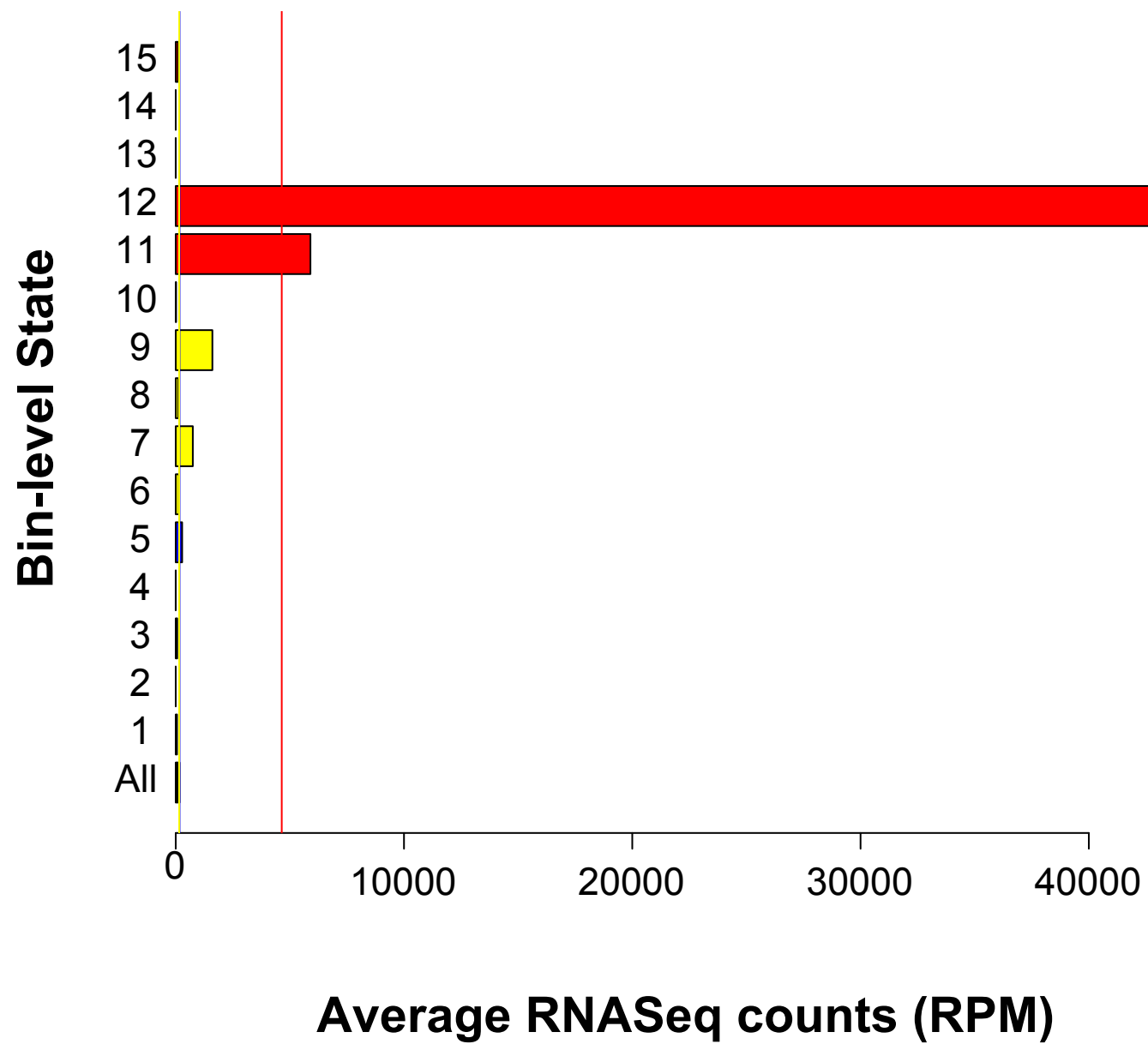
Supplemental Figure 6



Supplemental Figure 7



Supplemental Figure 8



Supplemental Figure 9

