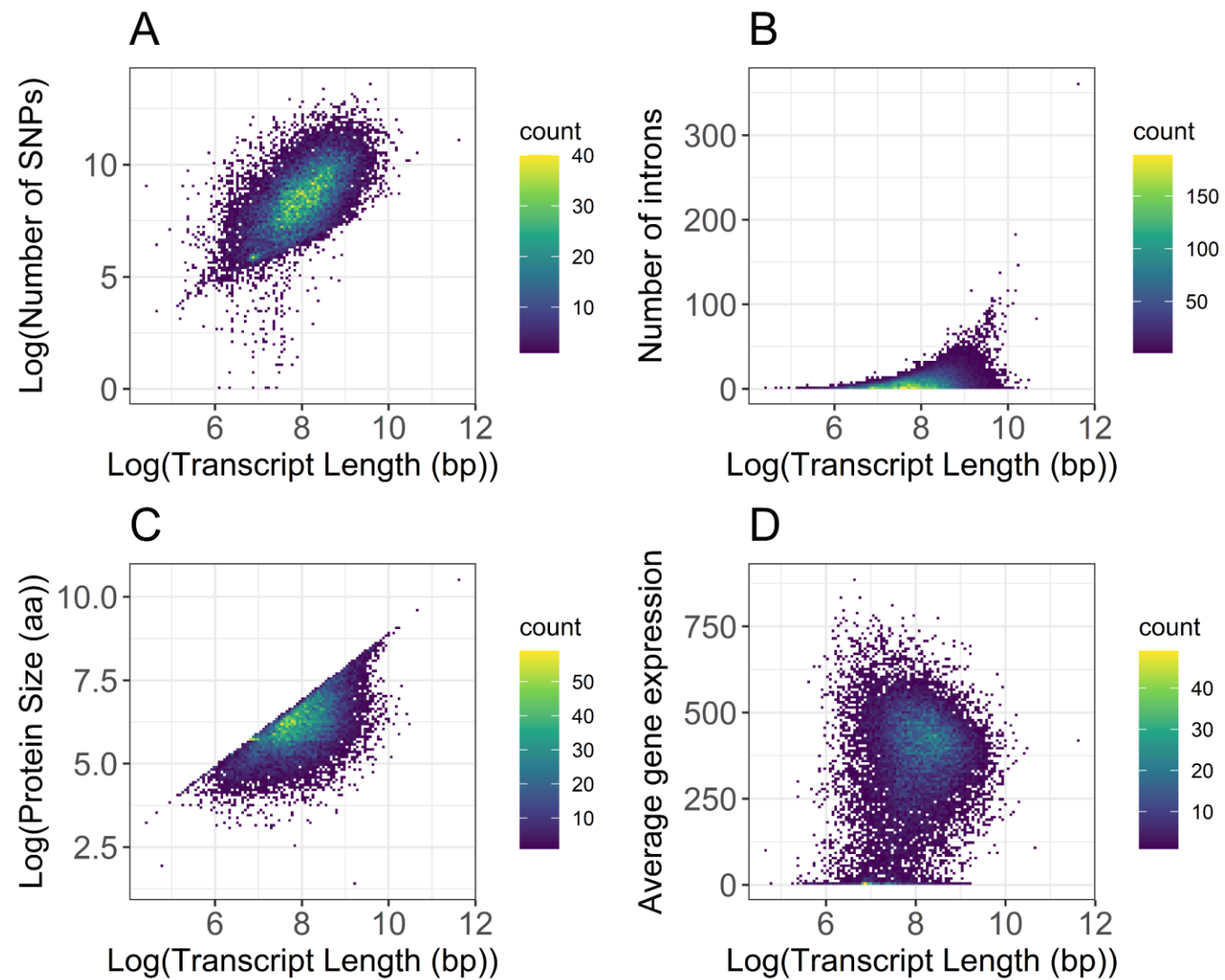


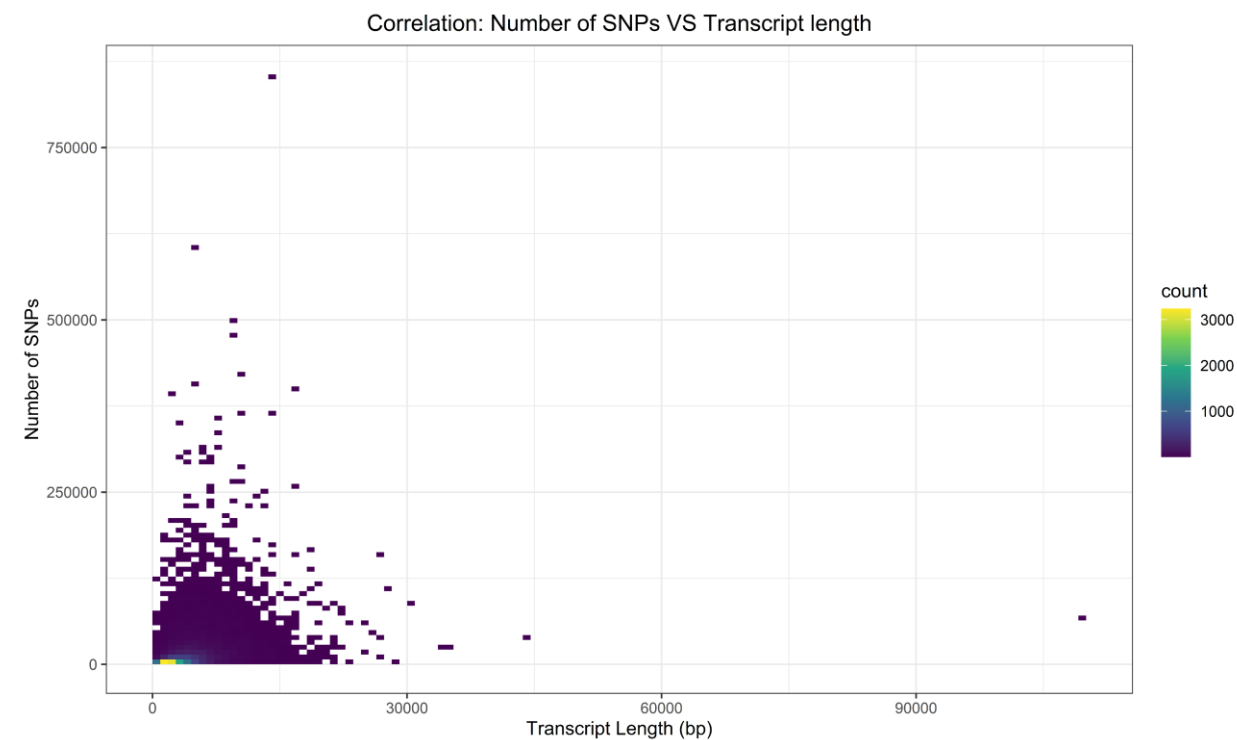
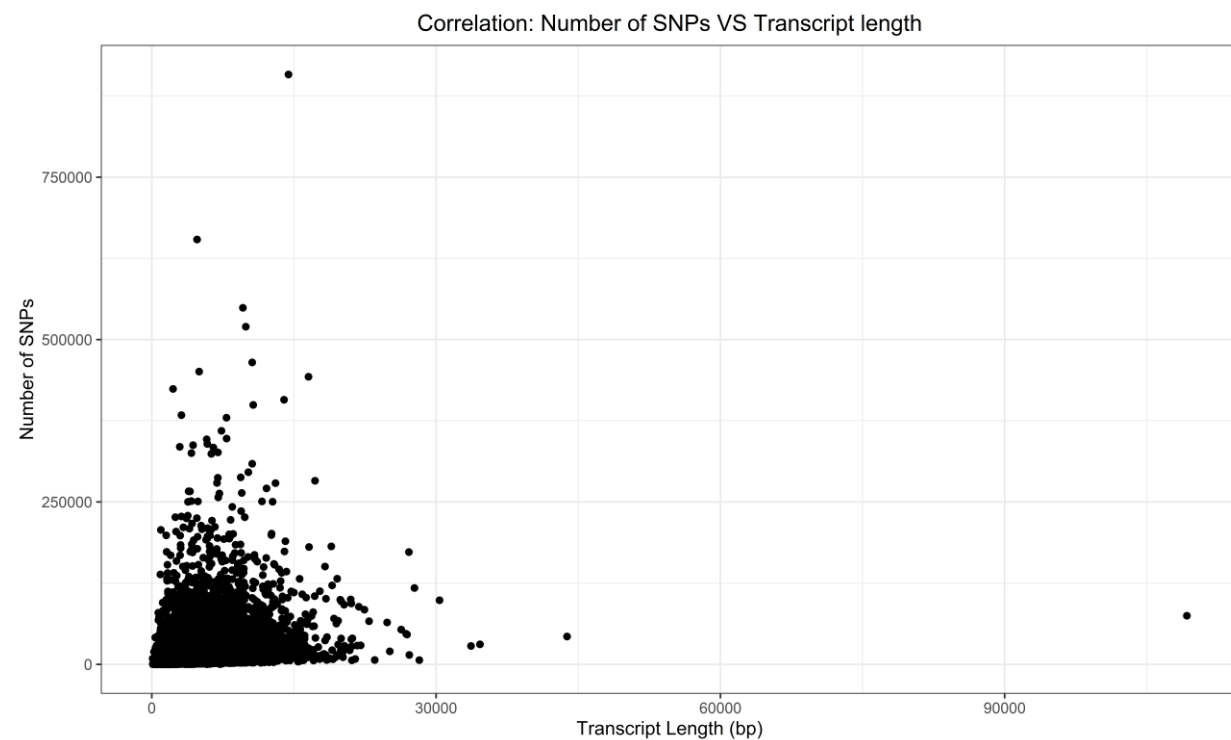
Supplementary Figure 4A.

Correlation between the Number of SNPs per gene and Gene Length and 2D density plot. Gene Length the Number of SNPs was obtained using the biomaart.



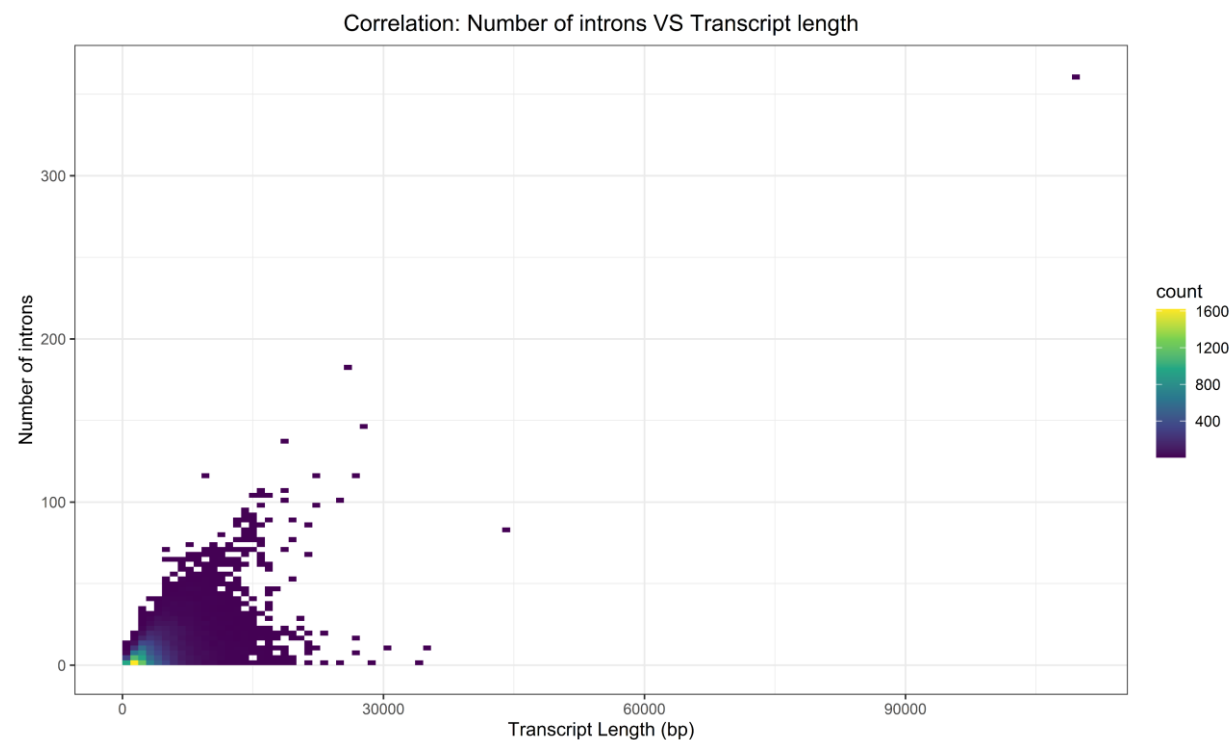
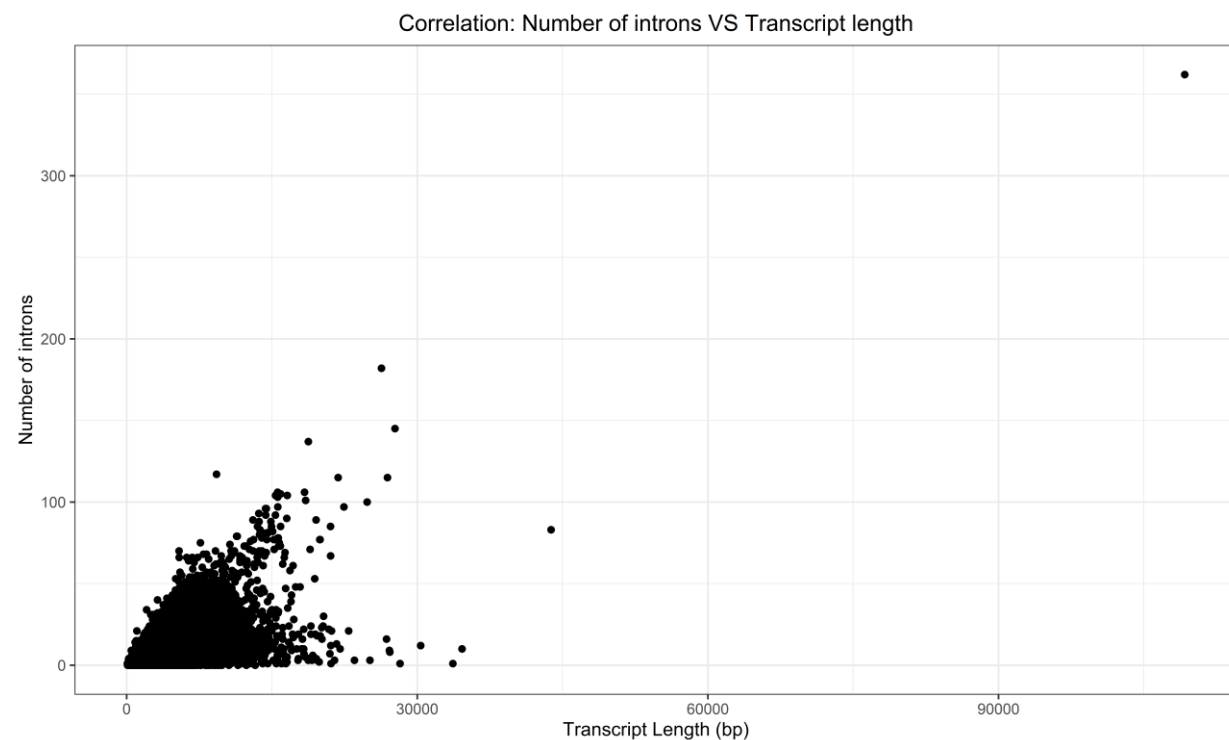
Supplementary Figure 4B.

2D density plot of the correlation analysis between Transcript Length (bp) and several other gene characteristics present in Fig 4. A: Correlation with the log transformed number of SNPs; B: Correlation with the log transformed number of introns; C: Correlation with the log transformed protein size; D: Correlation with the log transformed average gene expression.



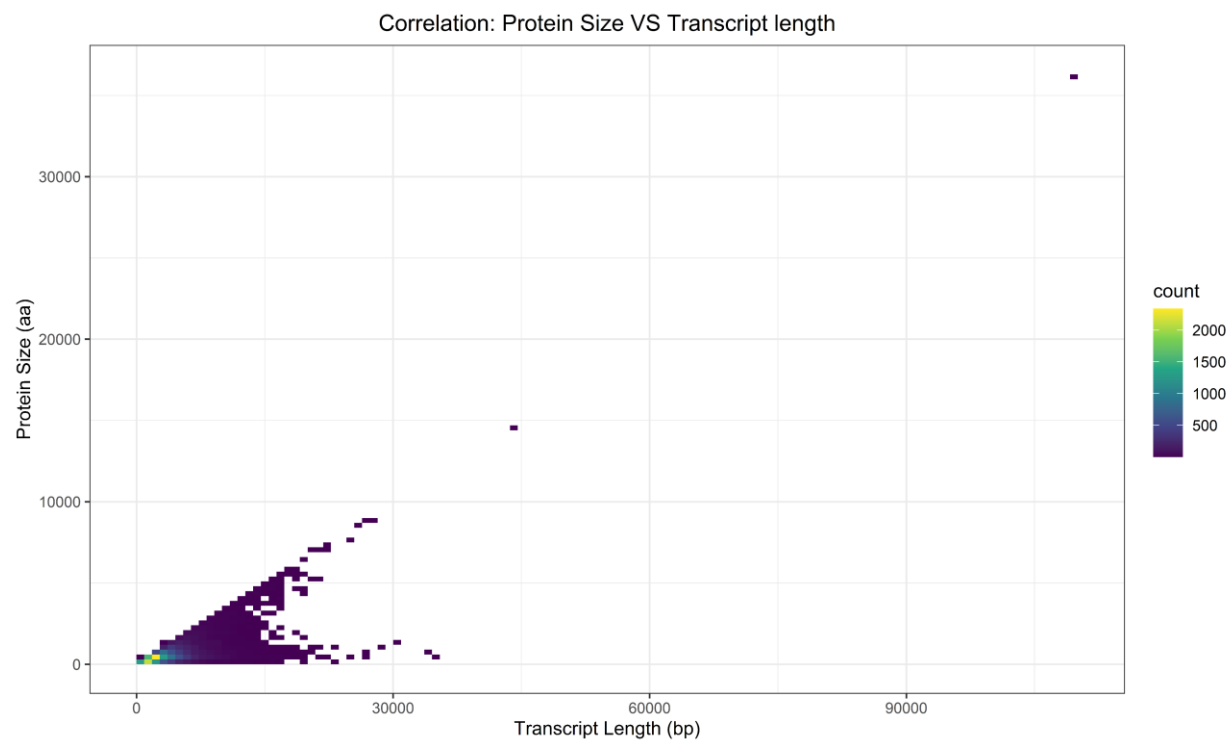
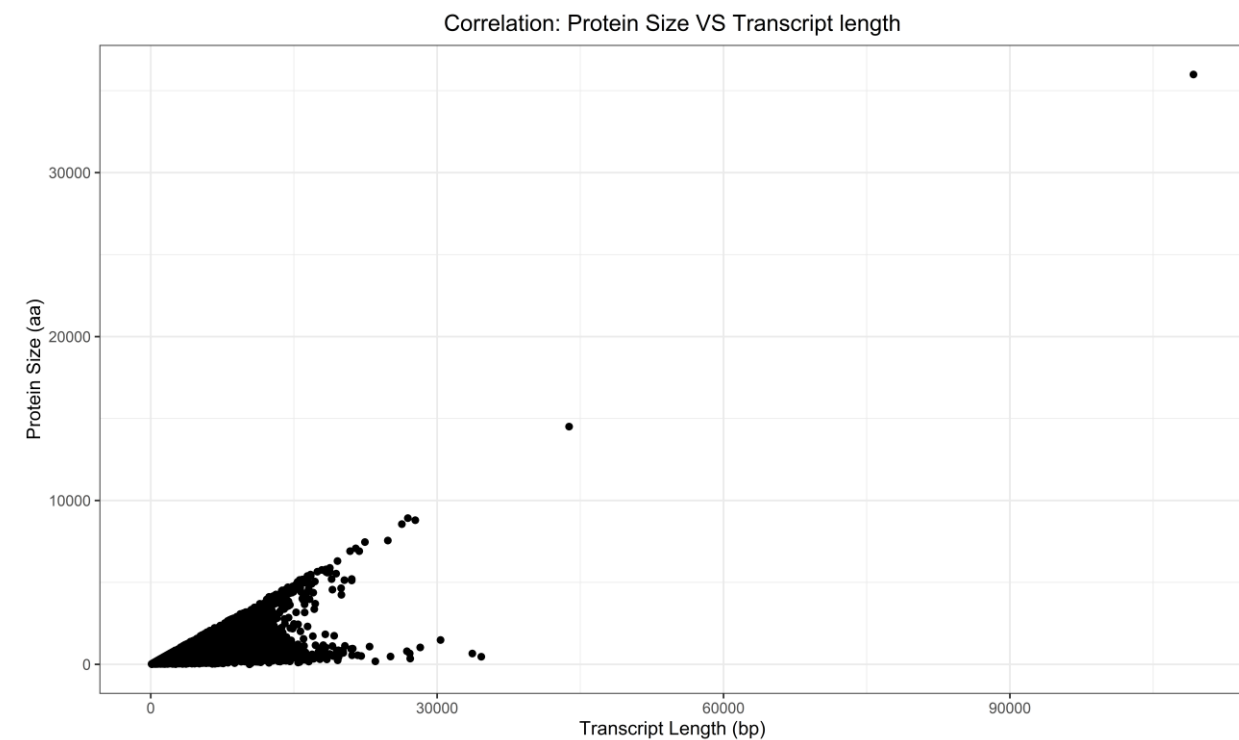
Supplementary Figure 4C.

Correlation between the Number of SNPs and Transcript Length and 2D density plot. Number of SNPs and Transcript Length were obtained using biomaart.



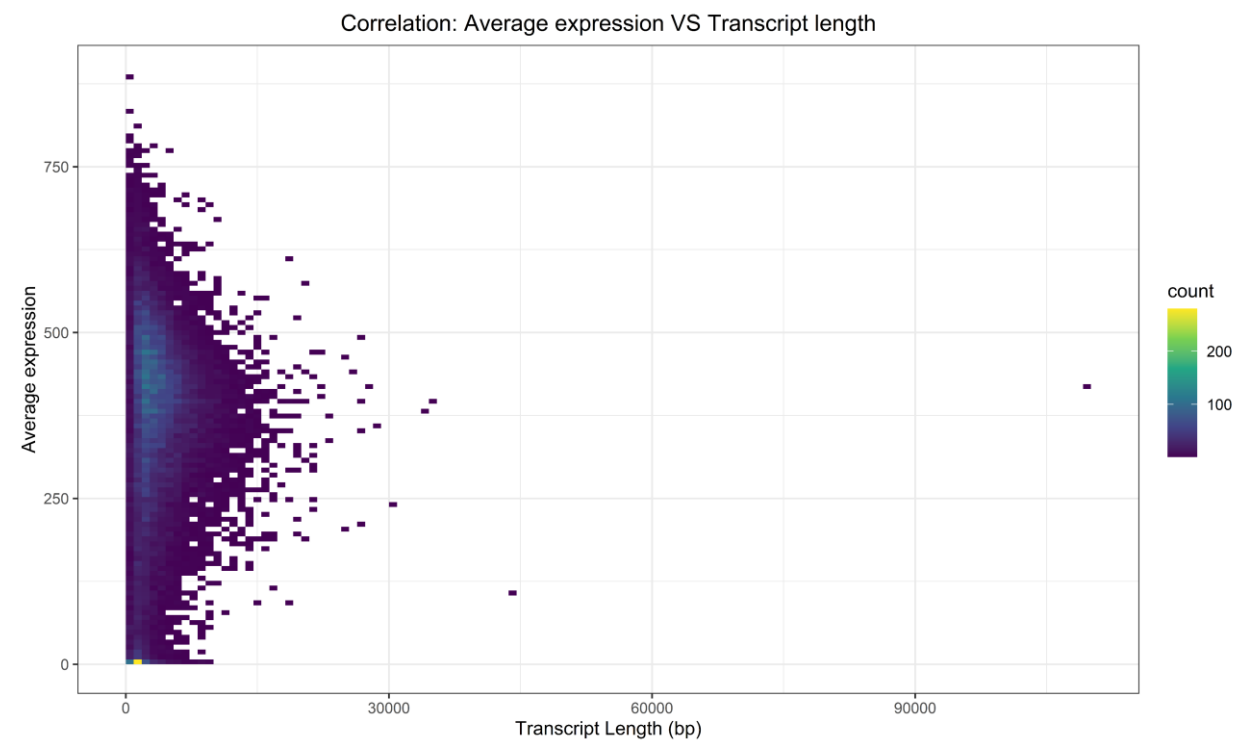
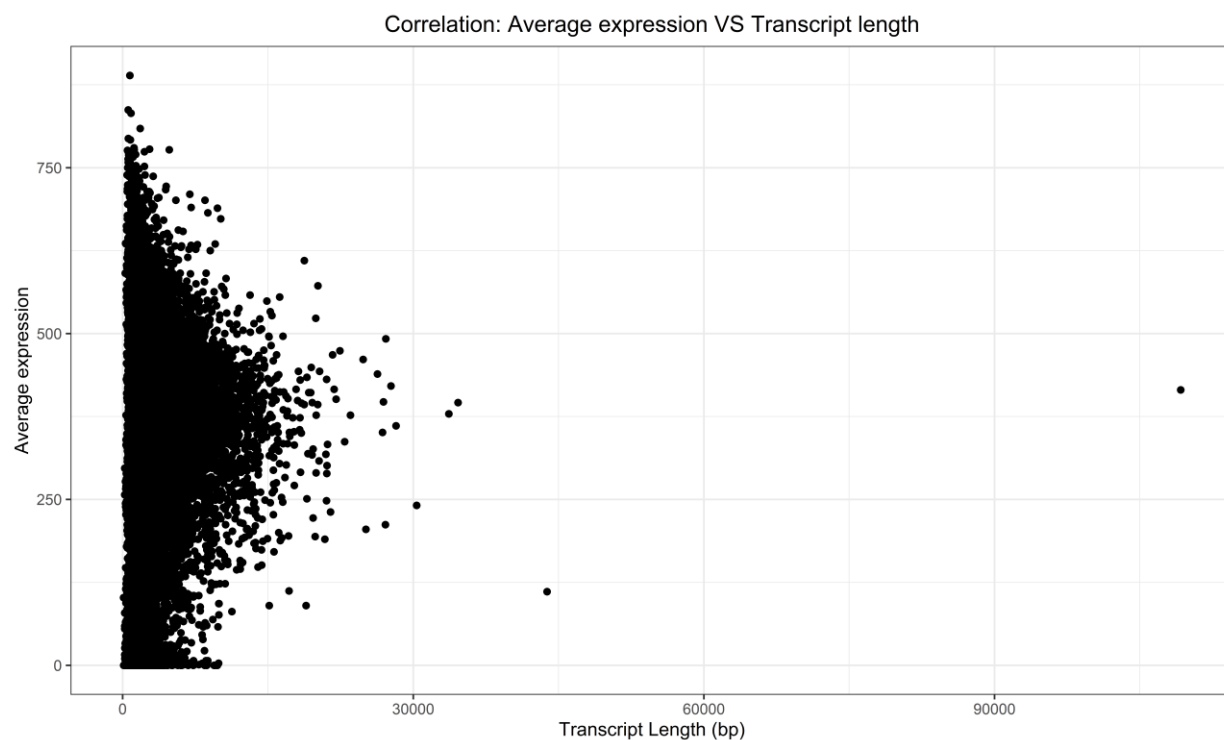
Supplementary Figure 4D.

Correlation between the number of introns and Transcript Length and 2D density plot. Number of introns and Transcript Length were obtained using biomaart.



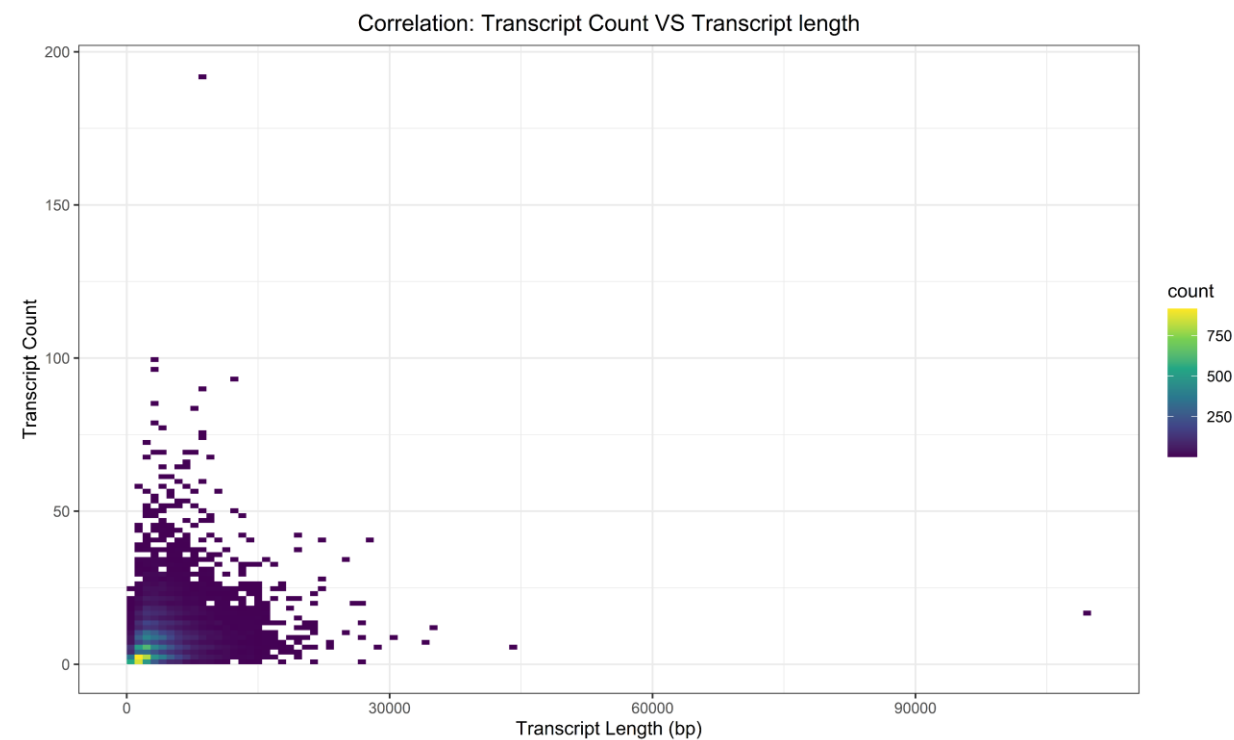
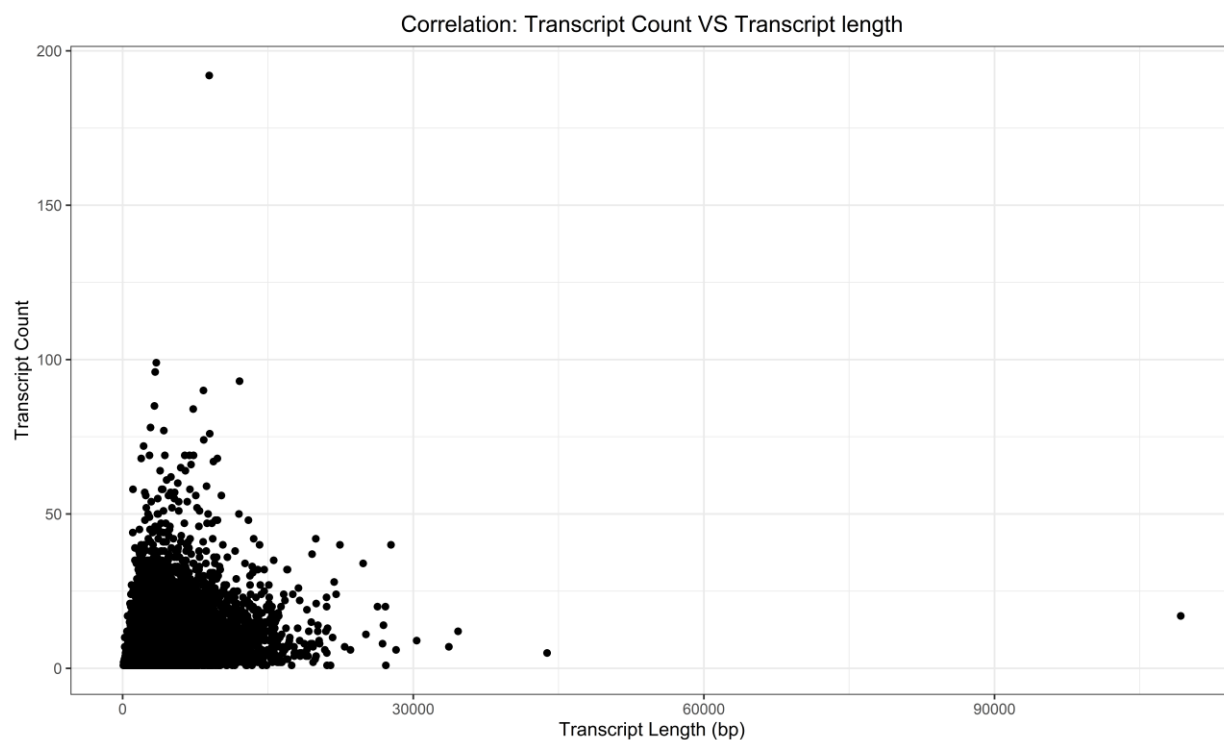
Supplementary Figure 4E.

Correlation between Protein size and Transcript Length and 2D density plot. Protein size and Transcript Length were obtained using biomart.



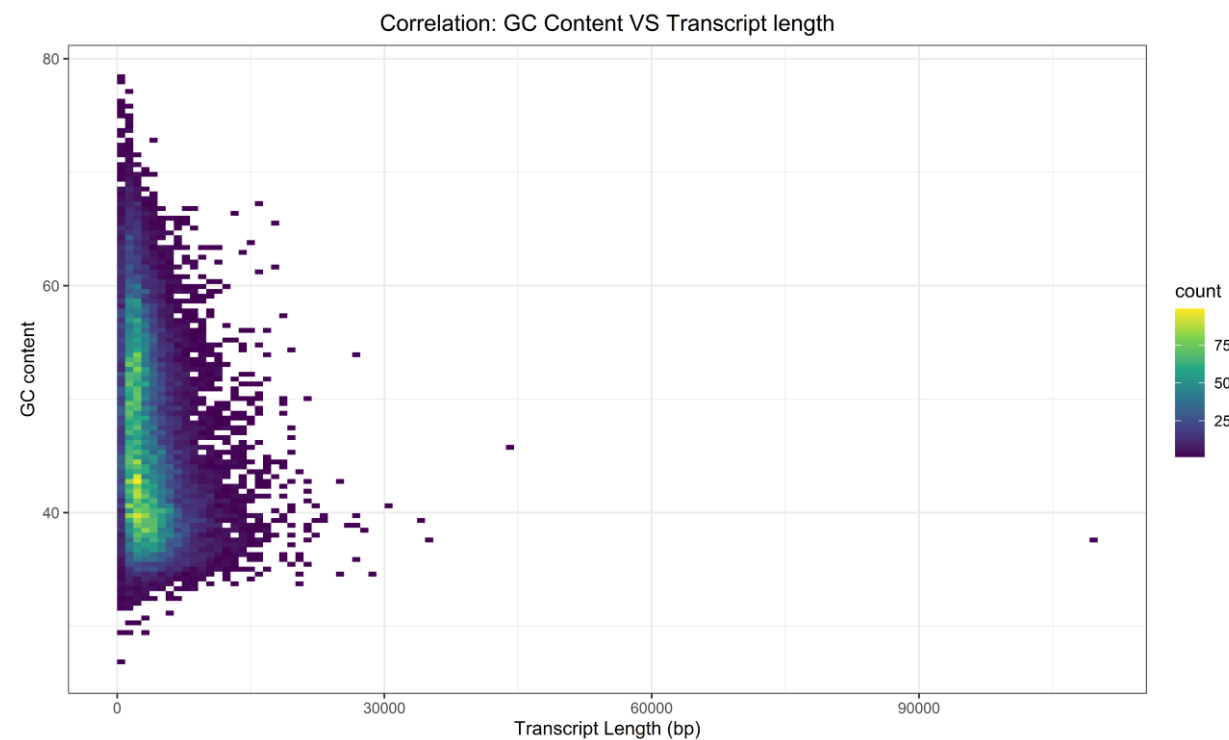
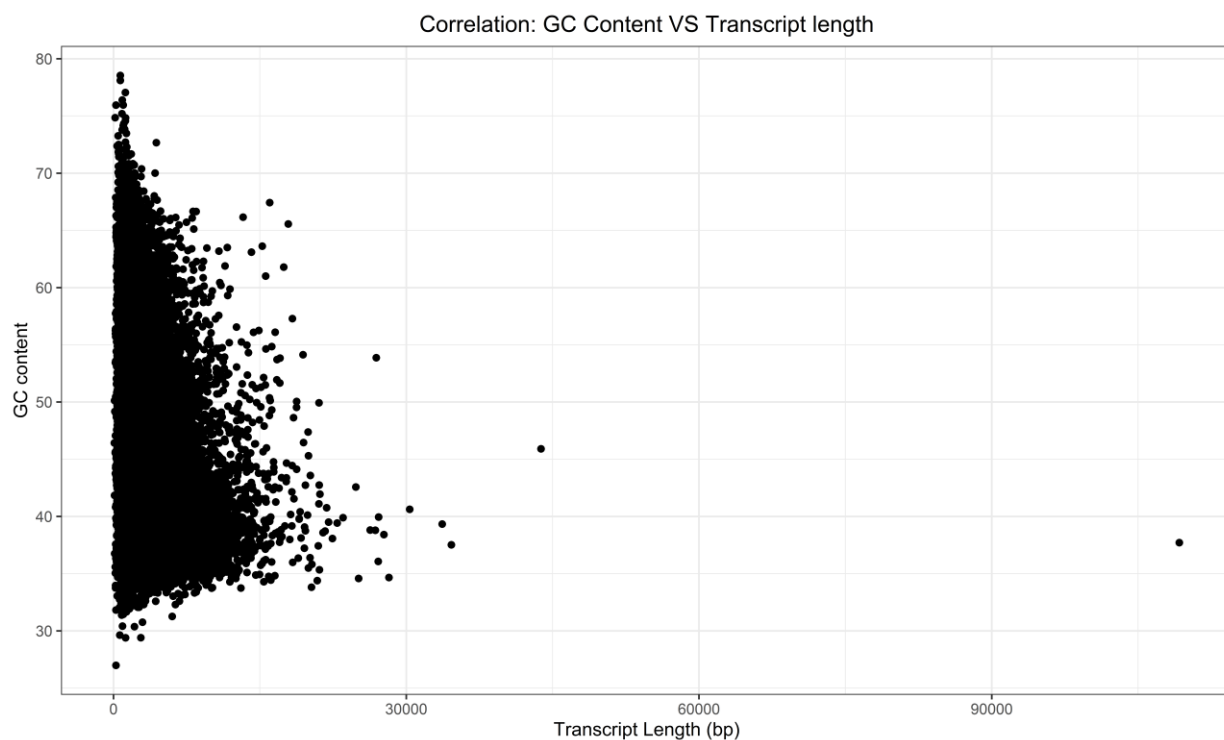
Supplementary Figure 4F.

Correlation between the Average Gene Expression and Transcript Length and 2D density plot. Average Gene Expression was obtained from the UCSC Genome browser and Transcript Length was obtained using biomart.



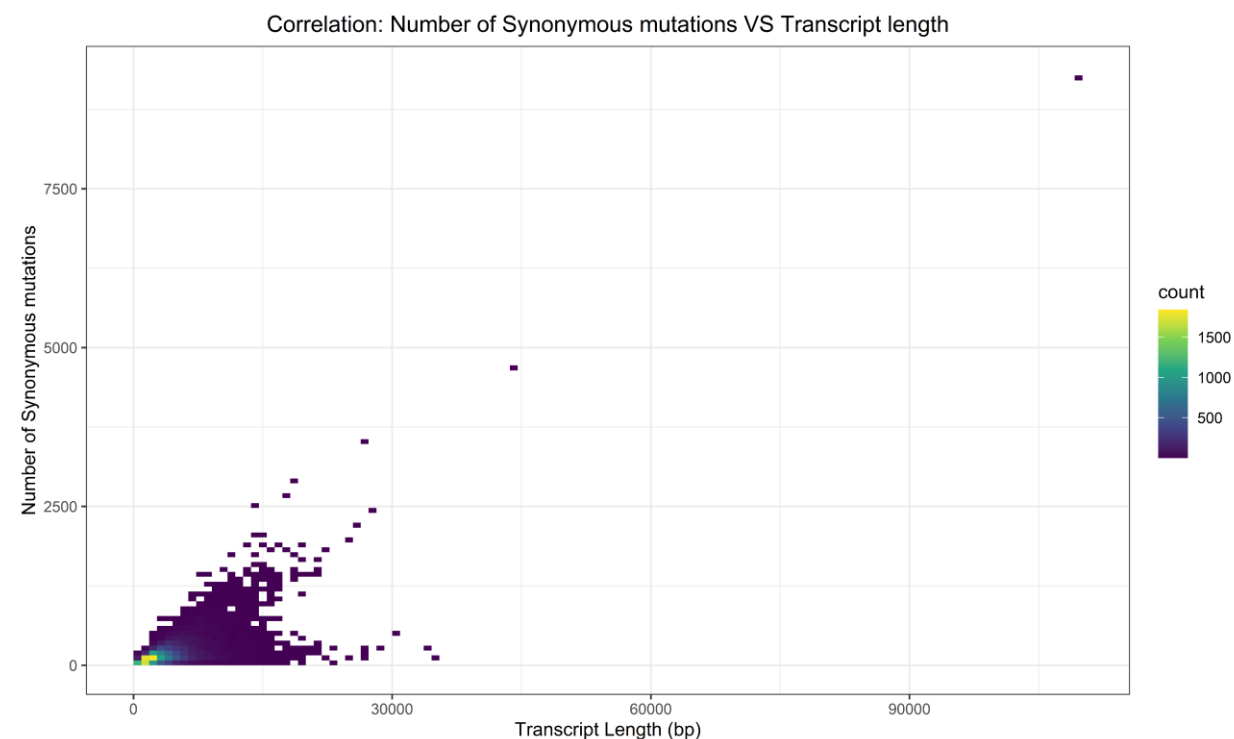
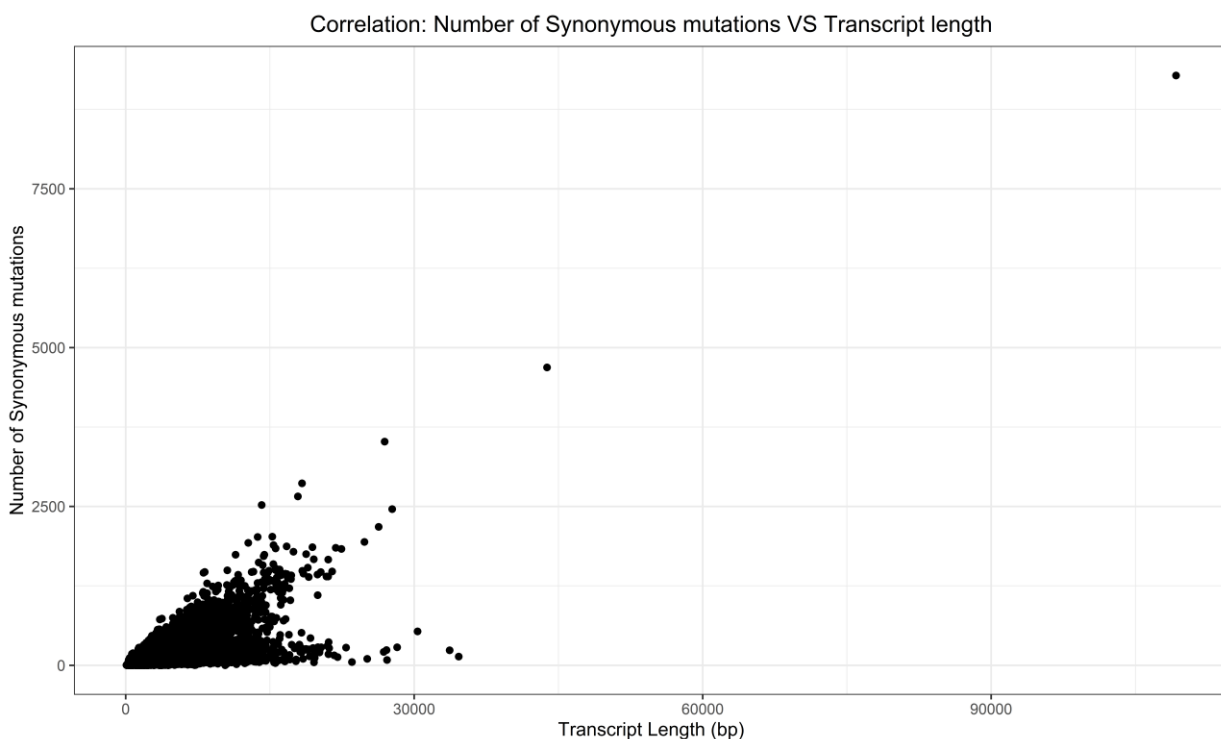
Supplementary Figure 4G.

Correlation between Transcript count and Transcript Length (bp) (Kendall test, $\tau = 0.22$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. Transcript count and Transcript Length for each transcript were obtained using biomart.



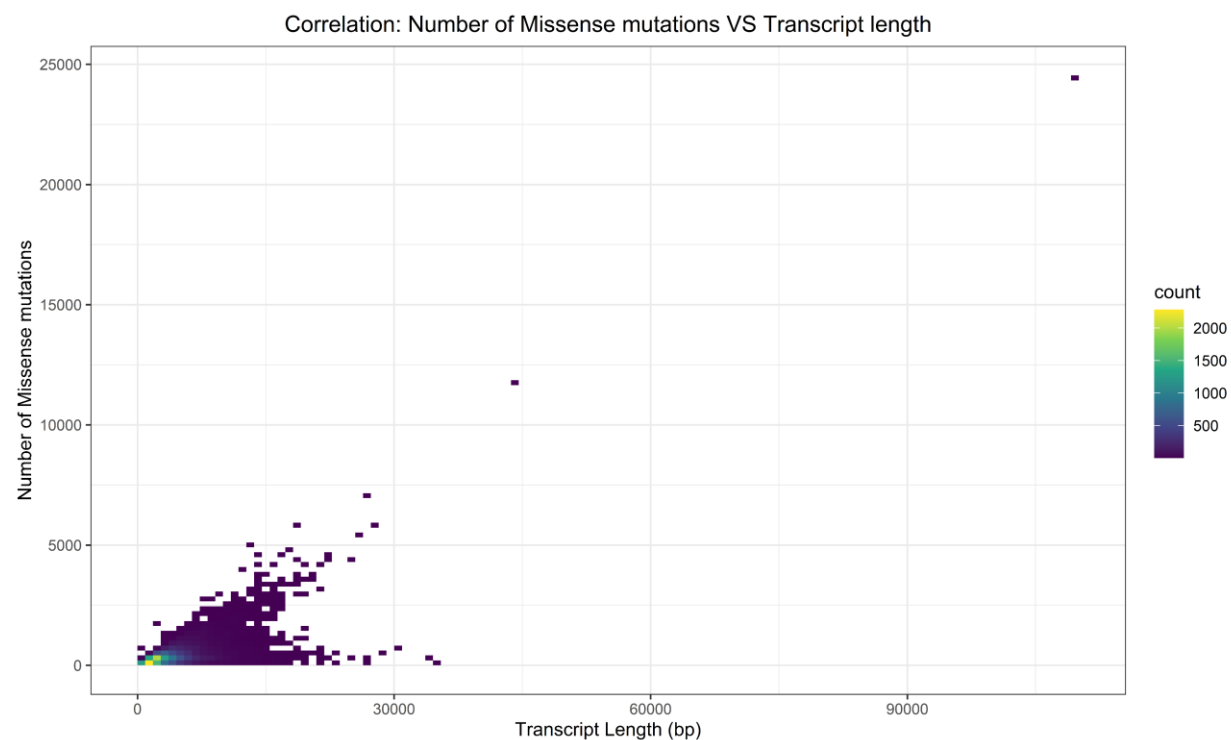
Supplementary Figure 4H.

Correlation between the GC Content and Transcript Length (bp) (Kendall test, $\tau = -0.19$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. GC Content and Transcript Length were obtained using biomart.



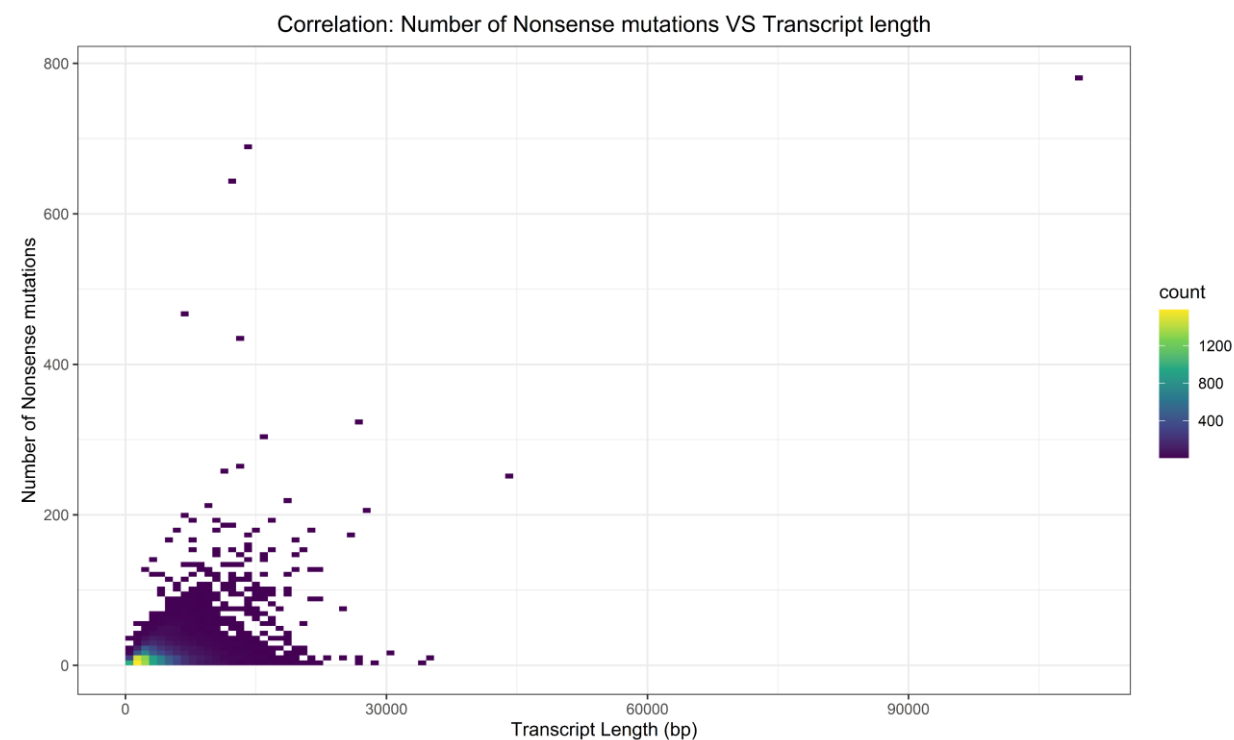
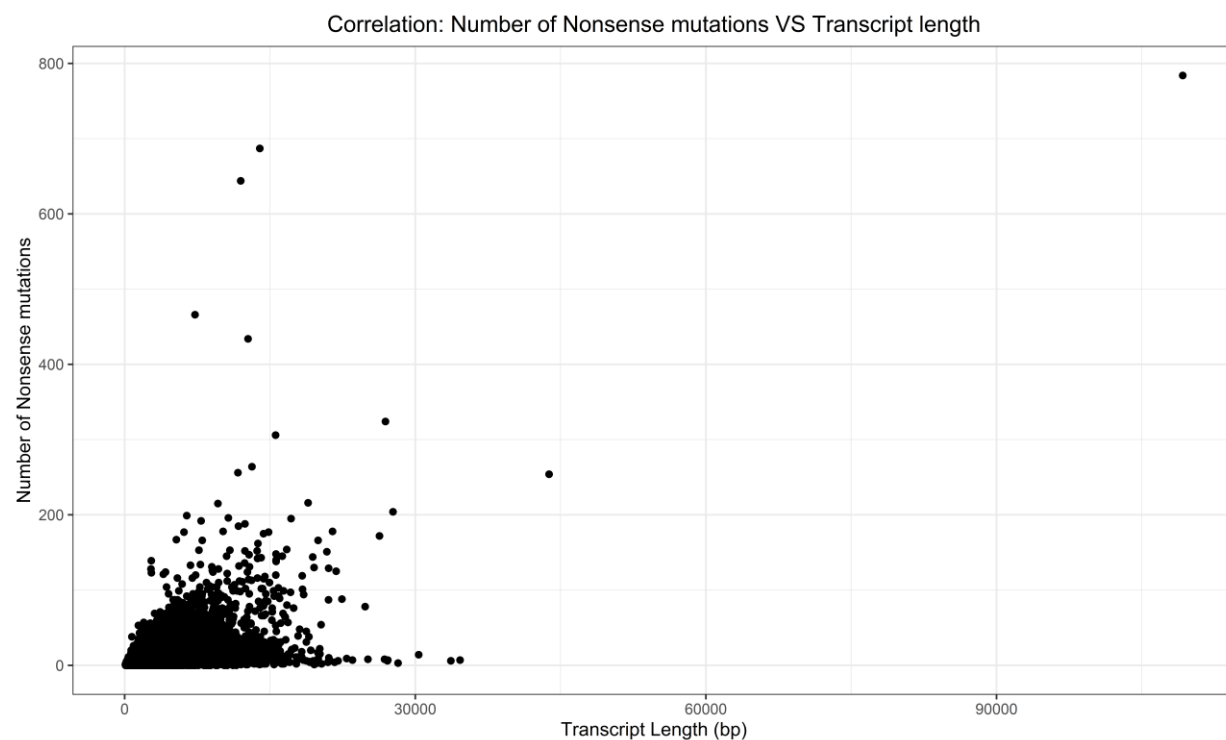
Supplementary Figure 4I.

Correlation between synonymous mutations and Transcript Length (bp) (Kendall test, $\tau = 0.44$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. The number of synonymous mutations and Transcript Length were obtained using biomart.



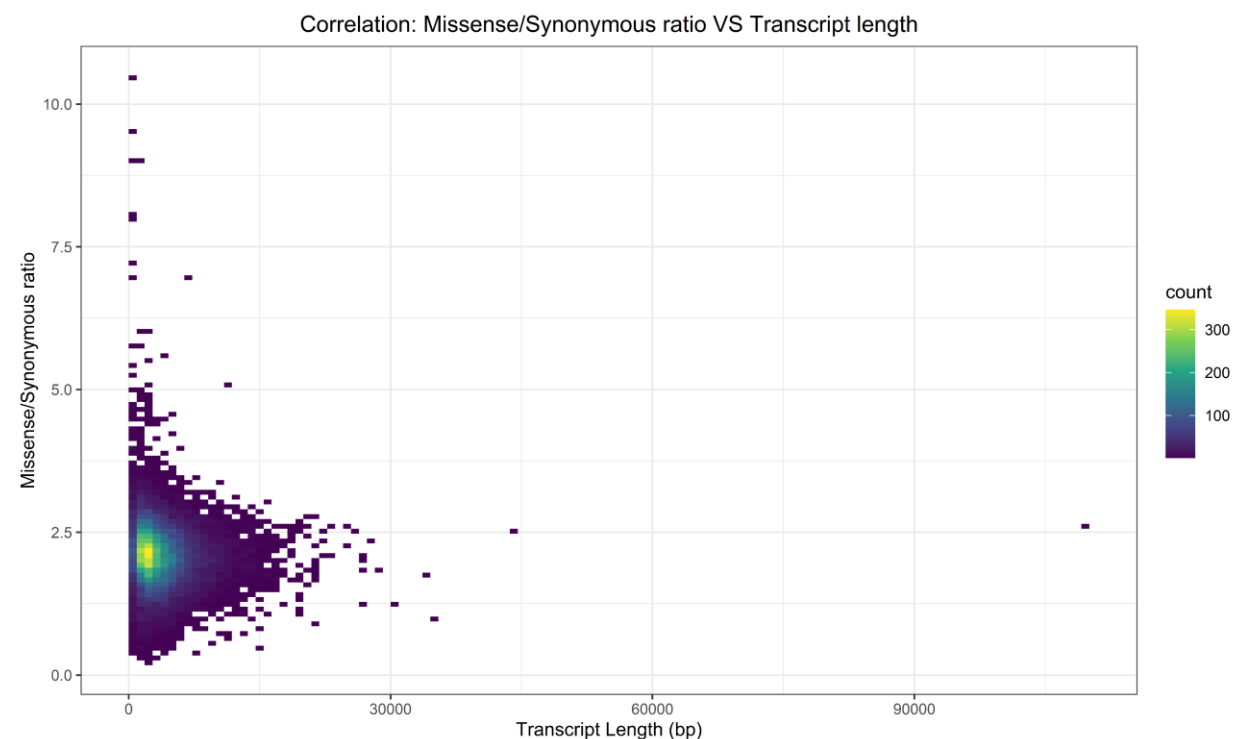
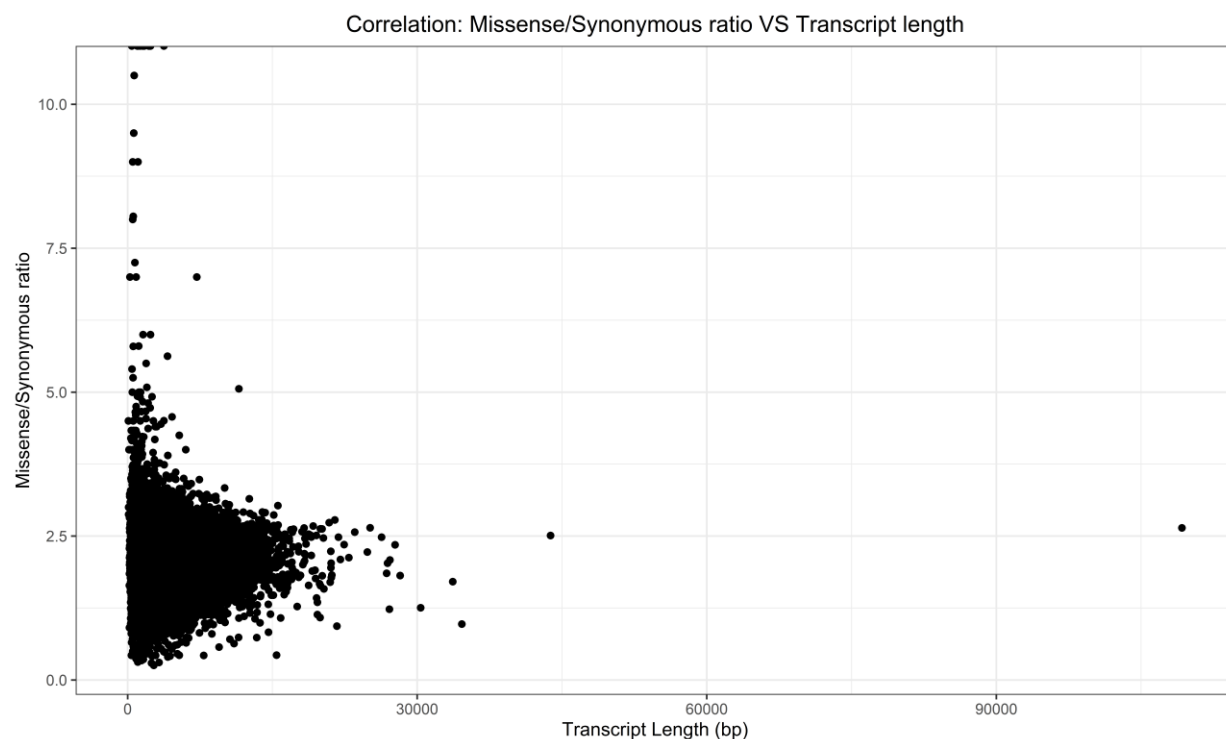
Supplementary Figure 4J.

Correlation between missense mutations and Transcript Length (bp) (Kendall test, $\tau = 0.42$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. The number of missense mutations and Transcript Length were obtained using biomart.



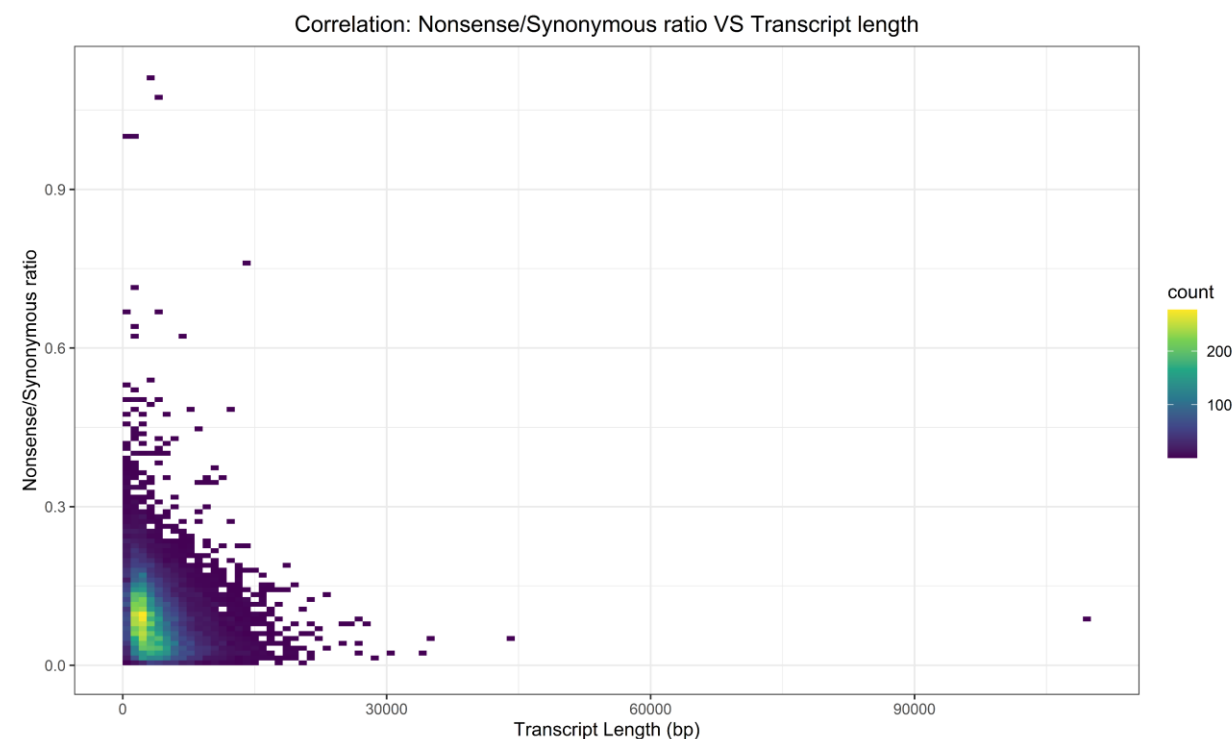
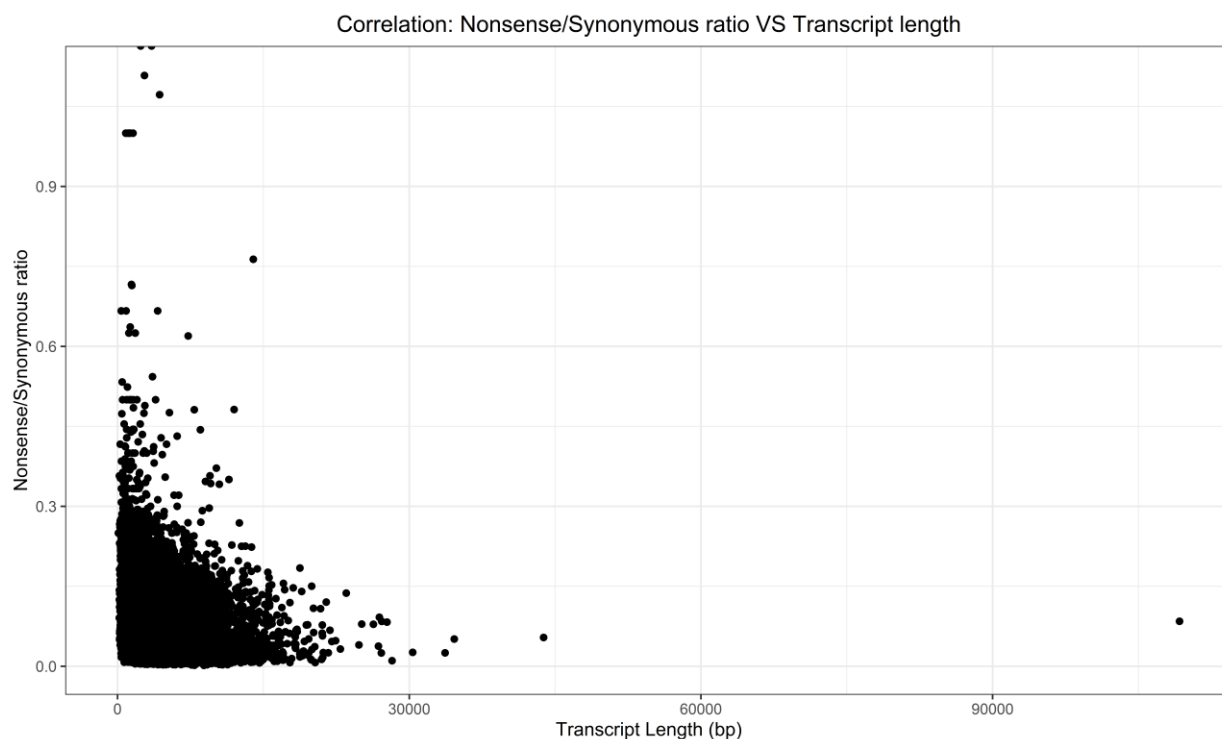
Supplementary Figure 4K.

Correlation between nonsense mutations and Transcript Length (bp) (Kendall test, $\tau = 0.21$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. The number of nonsense mutations and Transcript Length were obtained using biomaRt.



Supplementary Figure 4L.

Correlation between the missense/synonymous ratio of mutations and Transcript Length (bp) (Kendall test, $\tau = -0.07$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. The number of missense and synonymous mutations and Transcript Length were obtained using biomart.



Supplementary Figure 4M.

Correlation between the nonsense/synonymous ratio of mutations and Transcript Length (bp) (Kendall test, $\tau = -0.19$, $p\text{-value} < 2.20\text{E-}16$) and 2D density plot. The number of nonsense and synonymous mutations and Transcript Length were obtained using biomart.