

Supplementary Material

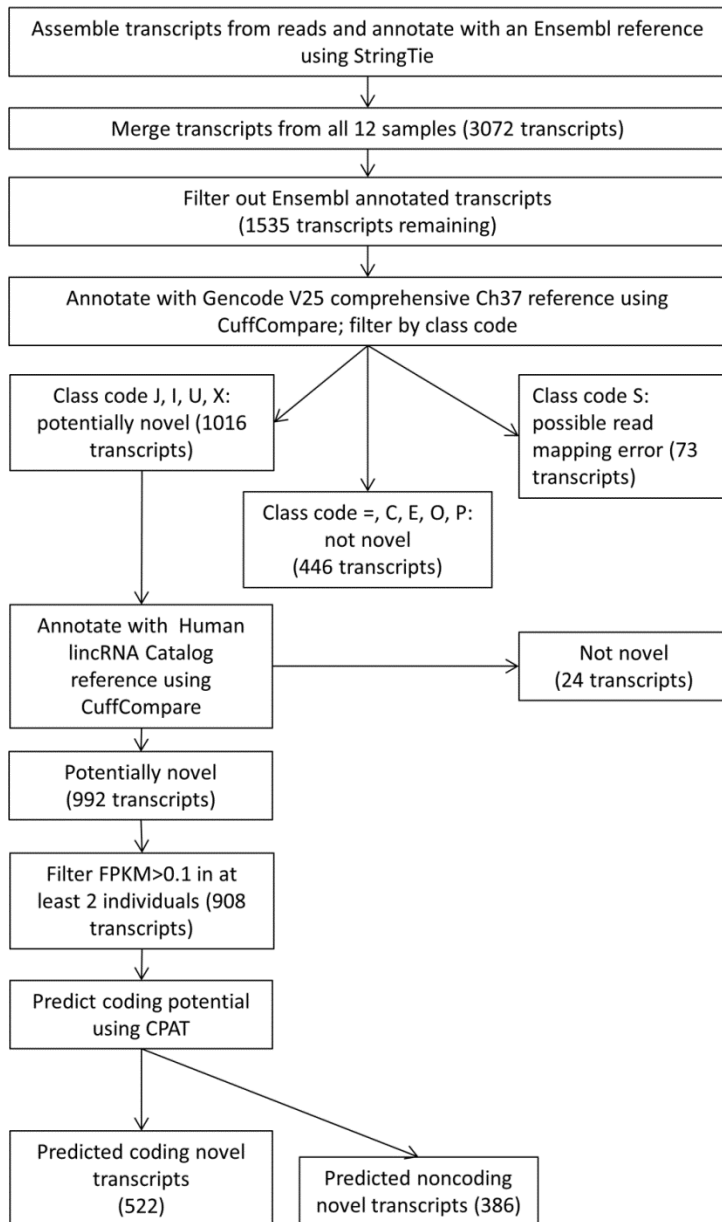


Figure S1: A flow chart depicting the workflow used for identifying novel transcripts from assembled transcripts. In total, 908 novel transcripts were identified. Novel transcripts were not present in the Ensembl, Comprehensive Gencode Release 25, or the Human lincRNA Catalog reference transcriptomes. All novel transcripts were identified in a minimum of two independent samples, and met a relative transcription level threshold (FPKM > 0.1).

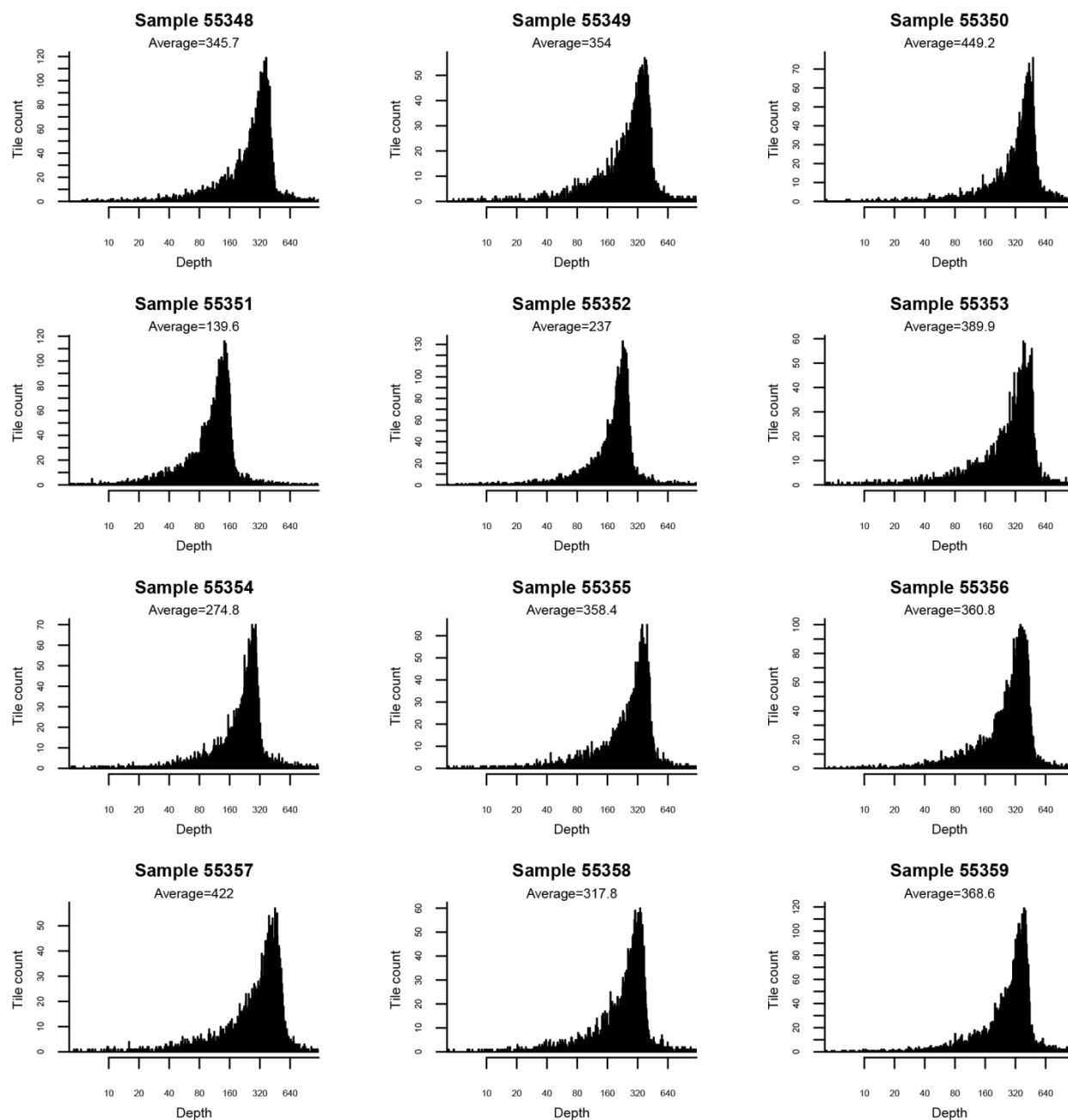


Figure S2: Distribution, mean, and median DNA read depth in the MHC in each sample. Read depth is shown per tiled region sequenced, with the number of tiles presented on the y axis

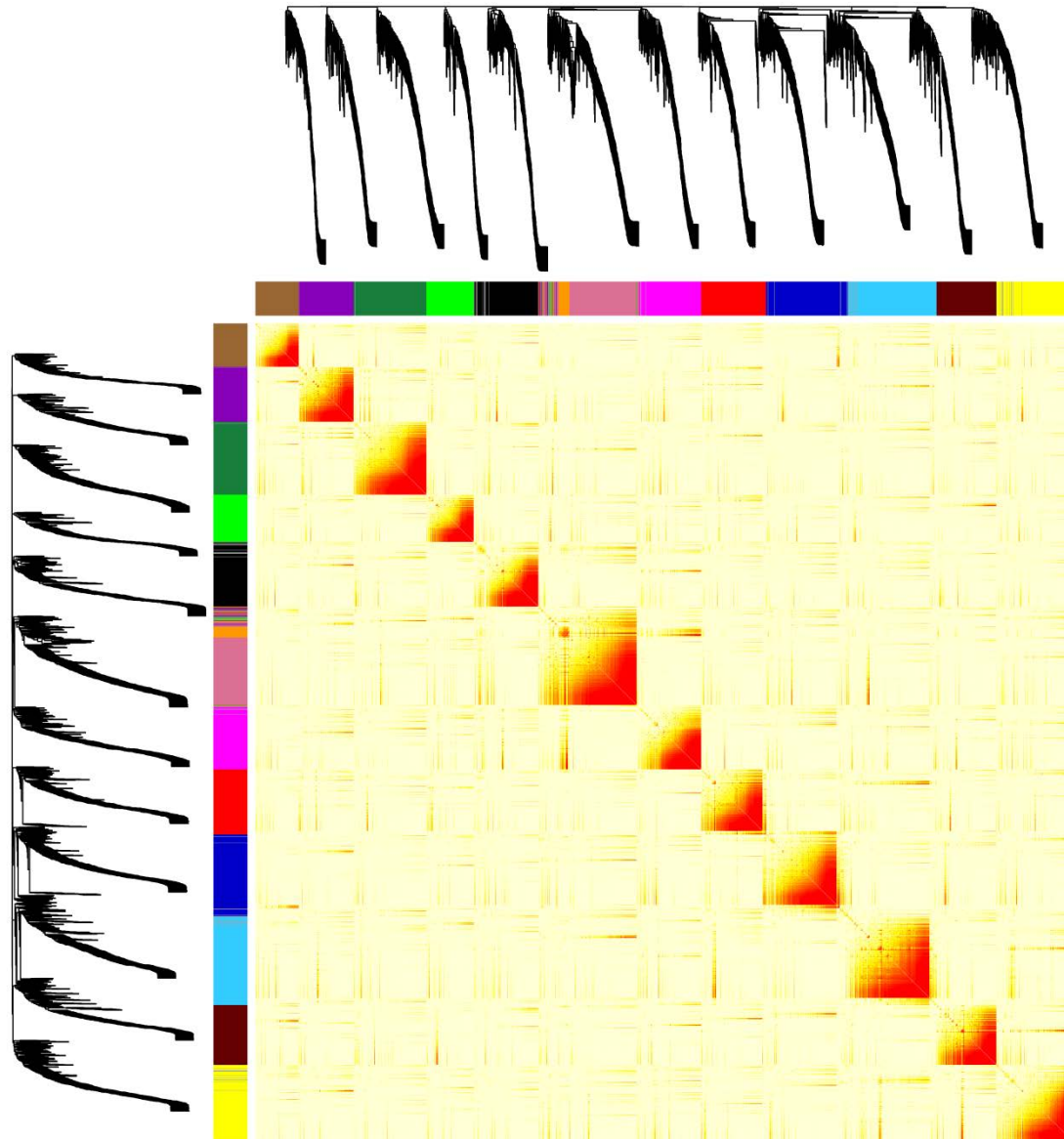


Figure S3: Network heat map depicting expression patterns of all MHC transcripts. Transcripts are aggregated based on co-expression cluster membership, and fourteen discrete colors are used to annotate the clusters on both axes. The precise composition of the co-expression network is seen in the dendrogram before both axes. As demonstrated in this dendrogram, there is a high degree of separation between the clusters in the network. The expression patterns for each transcript are included within the heat map itself (red: high coexpression, yellow: moderate, white: low). In each cluster, transcripts are highly co-expressed, with low or isolated co-expression outside of the assigned cluster.