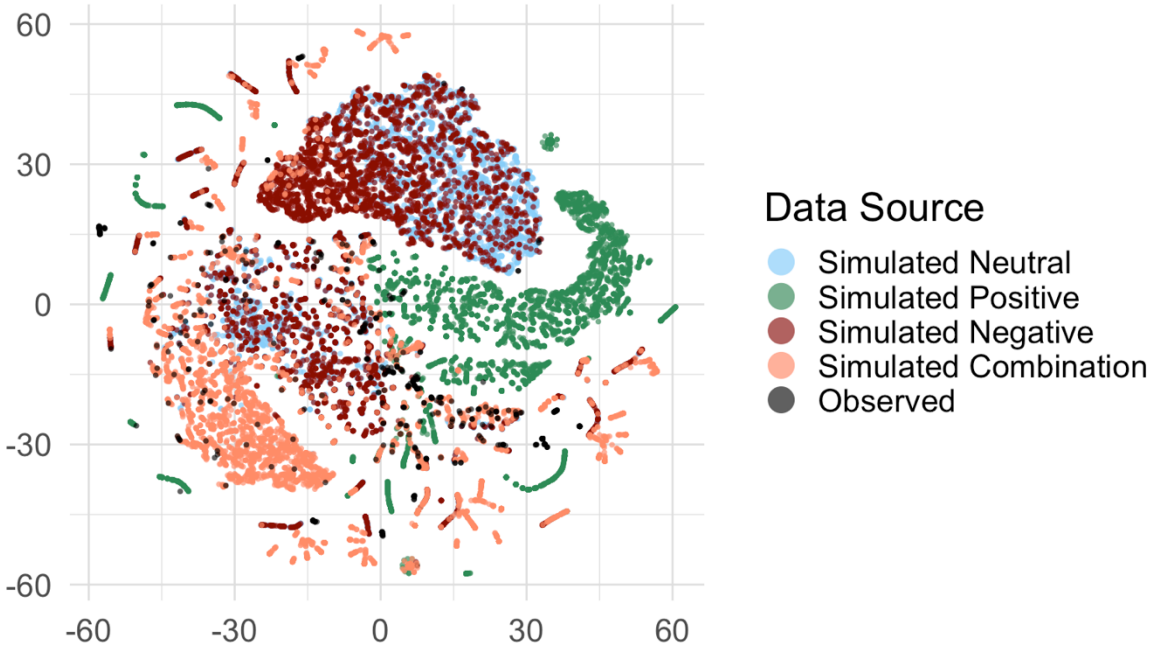
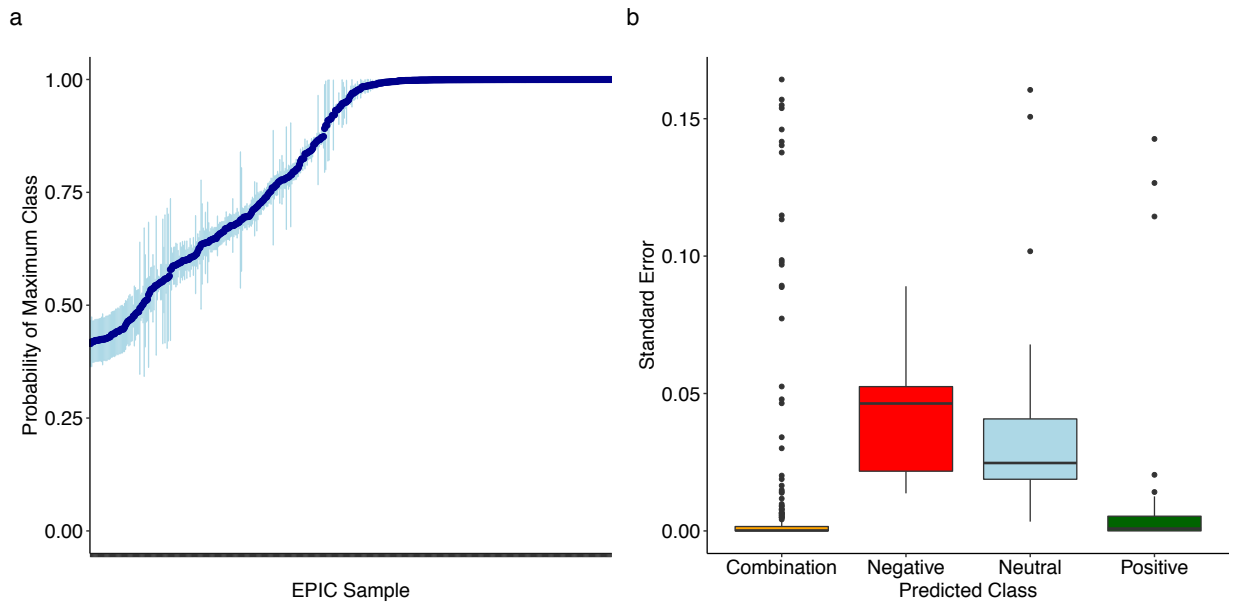
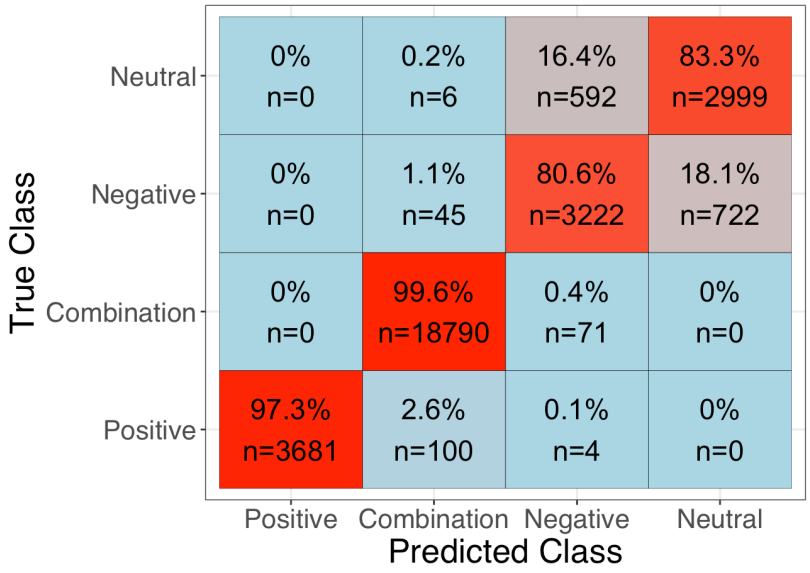
Supplementary Information



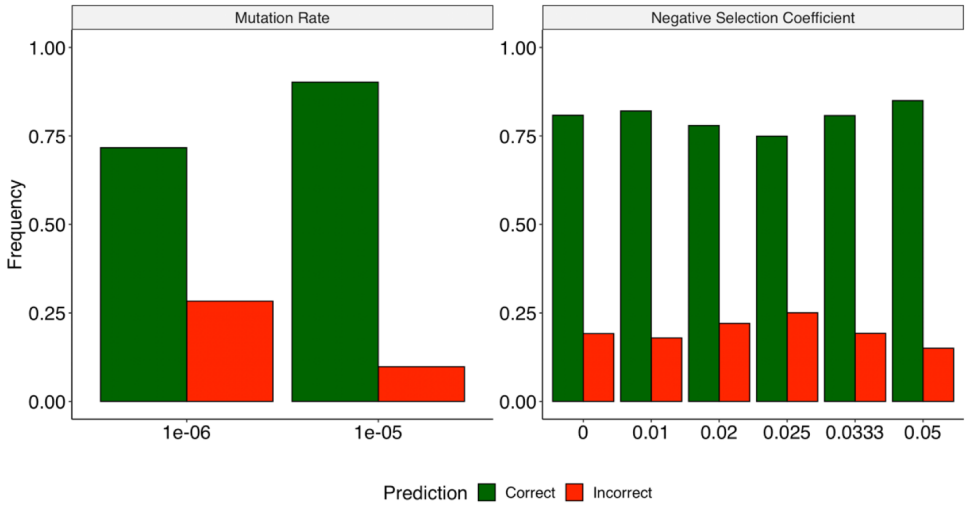
Supplementary Figure 1. Distribution of summary statistics from simulated data and observed data. t-Distributed Stochastic Neighbor Embedding ((t-SNE) mapping the distribution of 16 summary statistics for simulated and observed blood populations. Points were sampled from four different evolutionary classes (total sample size: n = 20,000) and are shown in colour (positive = green, negative = red, combination = orange, neutral = blue). Colour intensity was reduced to 70% to increase visibility. Observed summary statistics from the EPIC cohort are shown in black). Source data are provided as a Source Data file.



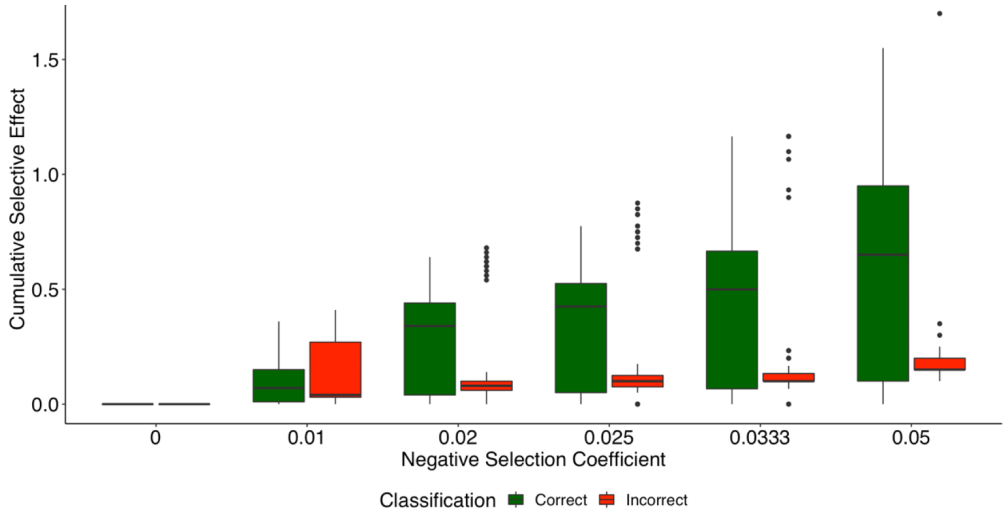
Supplementary Figure 2. Uncertainty associated with class predictions in EPIC cohort. A benefit of our ensemble-based approach is that, for each blood cell population, each DNN emits a softmax probability distribution across the four overarching evolutionary classes. In a conventional classification task, the class with the highest probability will be selected as the best fit. However, as we are employing an ensemble-based approach, we obtain a distribution of predictions for each population so as to measure the uncertainty associated with each prediction. To obtain the best fit evolutionary class for each individual, we calculated the mean and standard error for each softmax probability across the four evolutionary classes and accepted the class with the maximum softmax probability as the class of best fit. a) Here, we show the mean (dark blue) and standard error (light blue) softmax probability for all EPIC participants. We find that for approximately half of the participants (n= 215 (51.5%)) we obtain average probability distributions of over 99%, with a maximum standard error of 0.008, indicating that we can predict overarching evolutionary classes with high certainty. b) All participants predicted with a high degree of certainty are classified as evolving under either beneficial or mixture models of evolution. Participants classified as evolving under negative or neutral classes of evolution typically exhibit increased levels of predictive uncertainty. Each boxplot illustrates the distribution of standard error associated with each prediction across evolutionary classes. The midline represents the medians, the upper and lower bounds the interquartile ranges, and the whiskers extend to 1.5 times the interquartile range. Boxplots are coloured according to evolutionary class (positive(n=28): green, negative(n=63): red, combination(n=228): orange, neutral(n=158): blue). Source data are provided as a Source Data file.



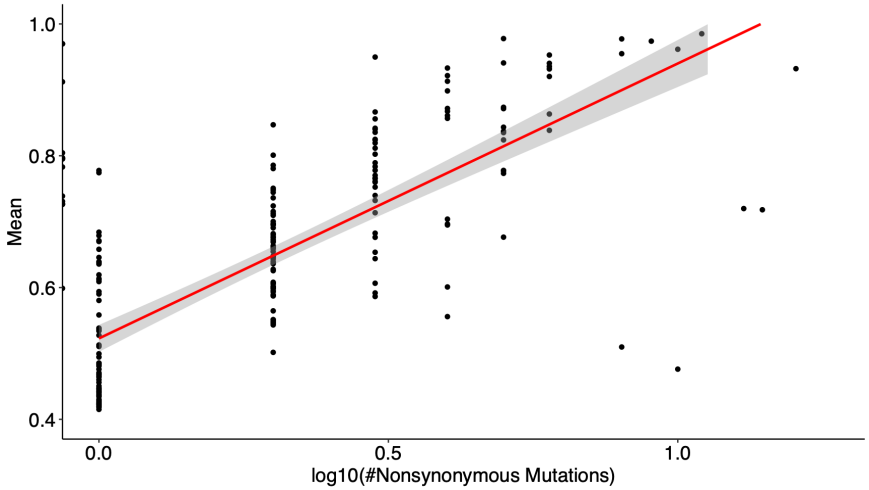
Supplementary Figure 3. Classification performance on novel parameter combinations. The y-axis represents the true evolutionary class, and the x-axis represents the predicted evolutionary class. Classification accuracy ranges from blue (low accuracy) to red (high accuracy). We tested our model's performance on a novel set of simulated data generated from novel parameter combinations. We find that we are able to achieve similar degrees of accuracy in evolutionary class predictions with positive and combination classes are predicted with 97.3% and 99.6%, respectively. Similarly, we observe a similar minor reduction in accuracy in discriminating between neutral and negative evolutionary classes with our classifier achieving accuracies of (80.6%) and (83.3%), respectively. Source data are provided as a Source Data file.



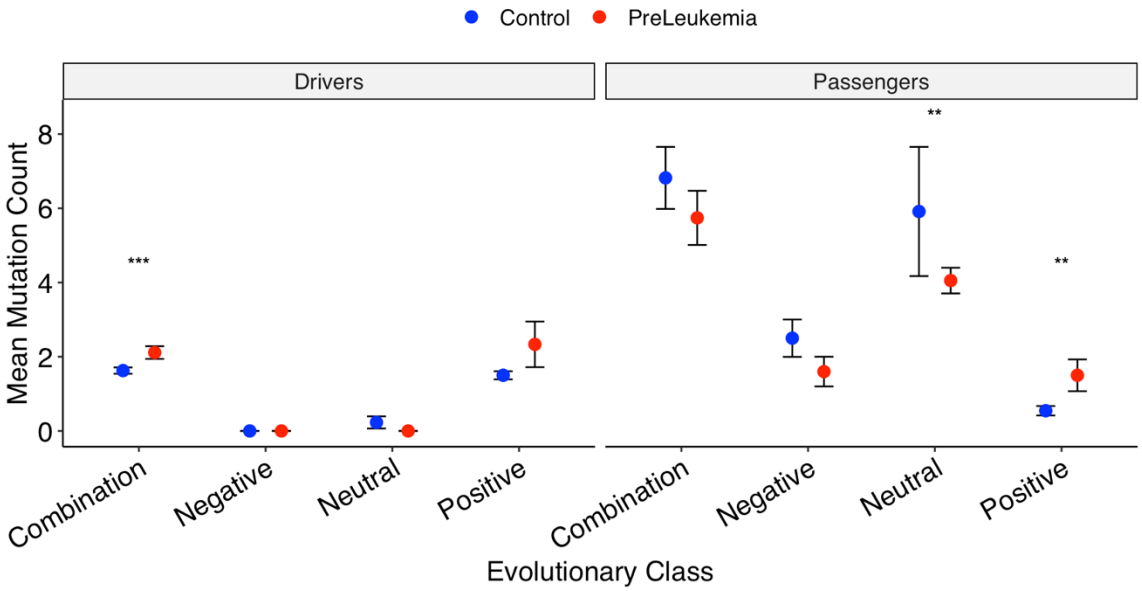
Supplementary Figure 4. Impact of parameters on predictive accuracy**.** We evaluated the distribution of parameters under which simulations were performed for neutral and negative classes of evolution. We show the proportion of the parameters where the evolutionary classes were correctly classified (green) compared to the proportion of parameters where incorrectly classified (red). Note, we only considered mutation rate and the coefficient of negative selection as the other two parameters are not present in the neutral and negative classes (probability of a mutation being beneficial, and coefficient of positive selection). We observe an increase in the proportion of correctly classified populations in simulations performed with a higher mutation rate (left panel). Similarly, simulations performed with a lower mutation rate are more commonly misclassified. This is in keeping with our expectation that populations simulated with a lower mutation rate will have fewer mutations which corresponds with a decrease in information to inform the DNN ensemble. We do not observe an obvious trend in the proportion of simulations misclassified across coefficients of negative selection. Source data are provided as a Source Data file.



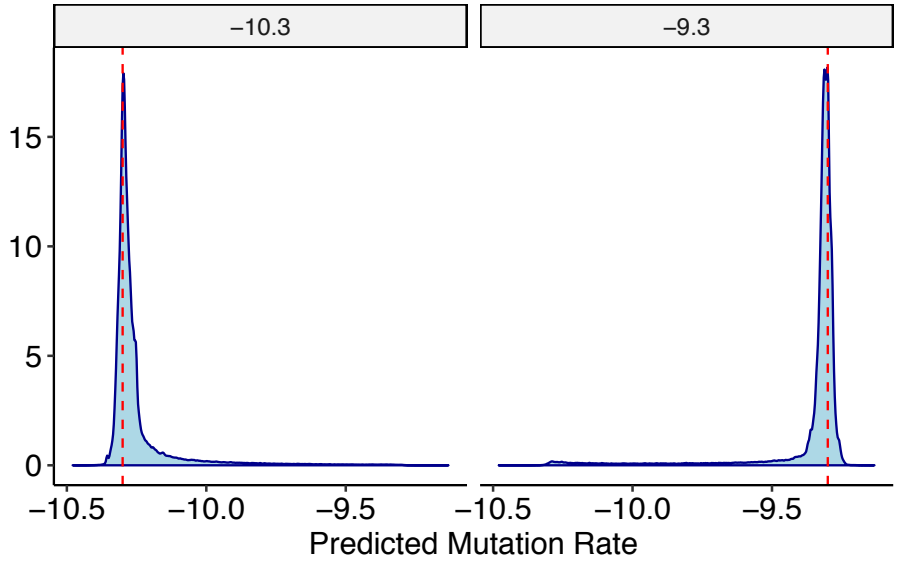
Supplementary Figure 5. Impact of cumulative selective effect on accuracy. We investigated the effects of the coefficient of negative selection on our ability to accurately classify neutral and negative classes of evolution. Specifically, we explored the cumulative selective effect, the number of mutations subject to a given selection coefficient, in our simulated populations. Cumulative selective effect was calculated by taking the product of the number of nonsynonymous mutations in a population and the coefficient of negative selection for each simulated population. Boxplots are coloured according to classification accuracy (correctly classified: green, incorrectly classified: red). Each boxplot illustrates the distribution of the cumulative selective effect for each selection coefficient, the midline represents the medians, the upper and lower bounds the interquartile ranges, and the whiskers extend to 1.5 times the interquartile range. We find that in the instances where we can correctly classify our simulated populations (n = 14,550, green), we observe a higher cumulative selective effect compared to the instances where we are not able to correctly classify our simulated populations (n = 3,426, red). This is in keeping with the expectation that the number of mutations segregating in a population is critical to informing evolutionary class predictions. Source data are provided as a Source Data file.



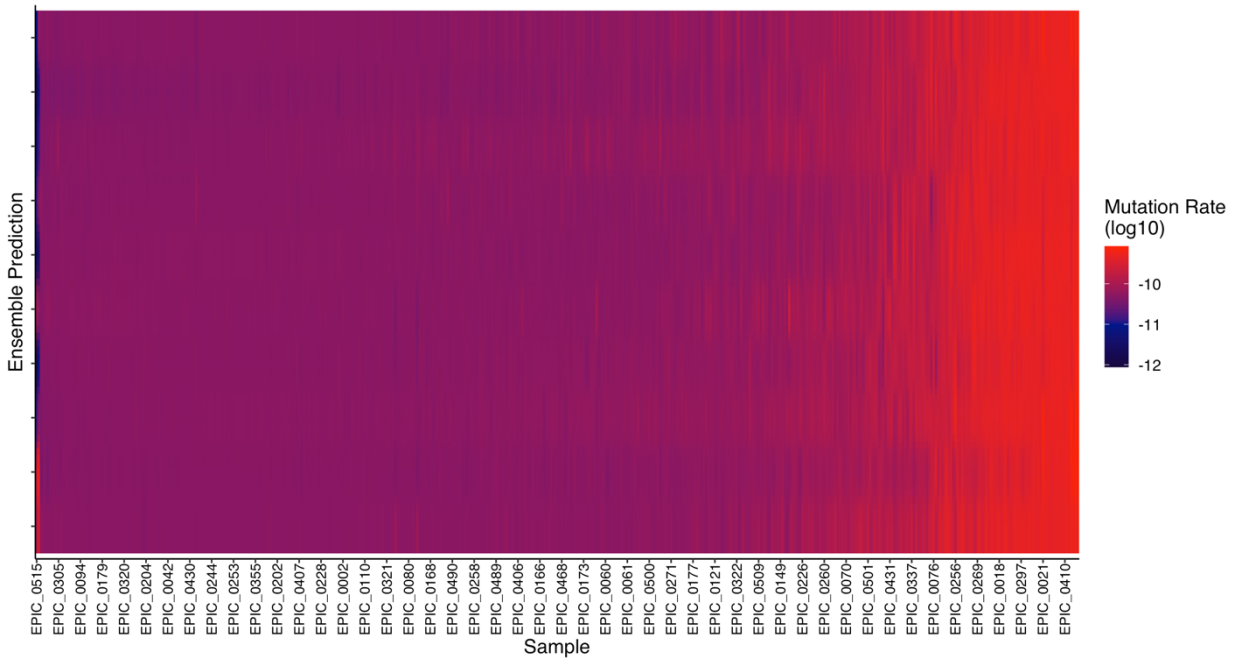
Supplementary Figure 6. Impact of nonsynonymous mutation count on prediction uncertainty. To investigate if higher levels of predictive uncertainty correspond with a reduction in mutations subject to selection, as observed in the simulated populations, we performed a linear regression to evaluate the relationship between the mean softmax probability for the predicted class and the number of nonsynonymous mutations segregating in the mature blood cell pool. The 95% confidence level interval for predictions from the linear model is indicated in grey. In doing so, we observe a linear association between the number of nonsynonymous mutations and our softmax probabilities indicating that we are less certain about our evolutionary class predictions in contexts where we have limited information to inform our predictions. Source data are provided as a Source Data file.



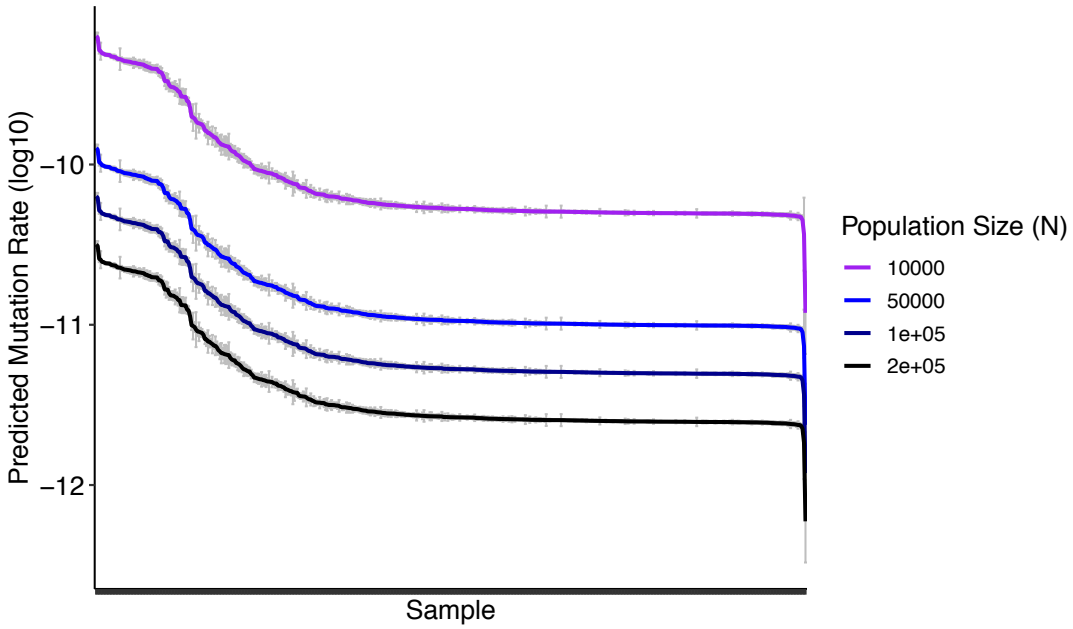
Supplementary Figure 7. Mutational burden across evolutionary classes. We investigated if there are higher numbers of passenger mutations in healthy individuals (blue) compared to cases (red). Genes were annotated as driver or non-driver genes based on if they were listed as one of the 32 driver genes in The Cancer Genome Atlas Acute Myeloid Leukemia project (*DNMT3A, NPM1, FLT3, TET2, RUNX1, IDH2, TP53, IDH1, CEBPA, NRAS, WT1, KIT, PTPN11, KRAS, U2AF1, STAG2, PHF6, ASXL1, RAD21, EZH2, KDM6A, DIS3, SUZ12, CUL1, BCOR, NF1, THRAP3, CHD4, PRPF8, EGFR, MED12, CBFB*). The mean driver and non-driver mutation count was calculated across healthy (n = 385) and preleukemic individuals (n = 92) fitting each evolutionary class. Data are presented as mean values +/- SEM. The level of significance is indicated as follows: ns: p > 0.0, *p-value <= 0.05, **: p-value <= 0.01, ***: p-value <= 0.001, ****: p-value <= 0.0001. We find that preleukemic cases fitting the combination class of evolution have significantly more driver mutations (Two-sided Wilcoxon rank sum test, $W = 3828$, p-value = 8.76e-04, n = 228), and significantly more passenger mutations in the positive class (Two-sided Wilcoxon rank sum test, $W = 29.5$, p-value = 0.03, n = 28). Healthy controls have a significantly higher number of passenger mutations in the neutral class (Two-sided Wilcoxon rank sum test, $W = 890.5$, p-value = 0.02, n = 158). Further, we observe no driver mutations in individuals fitting negative models, preleukemic cases fitting neutral evolution, and only a slight increase in the average number of driver mutations in neutral controls. In almost all classes of evolution, with the exception of positive, we observe a higher number of mutations in non-driver genes in controls compared to preleukemic individuals. Source data are provided as a Source Data file.



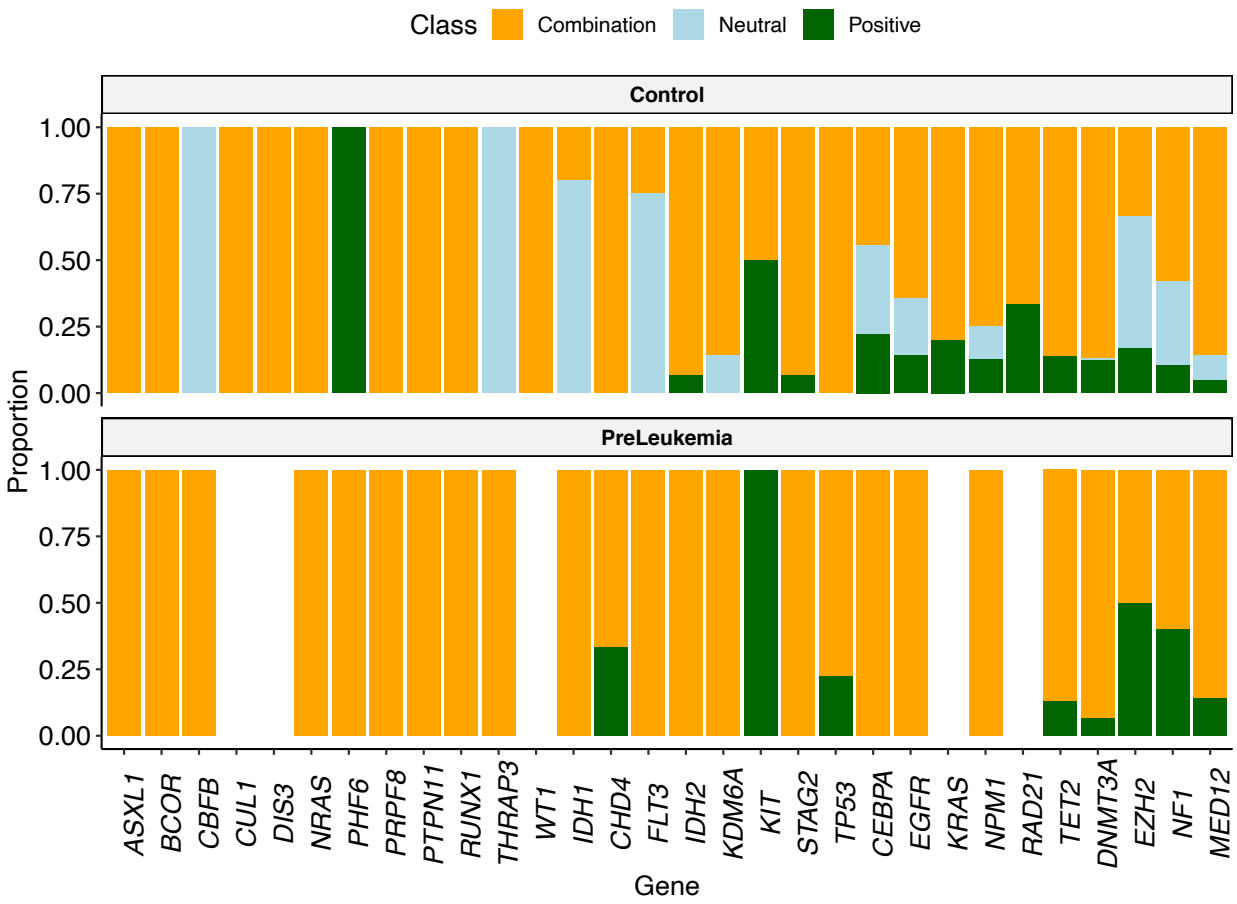
Supplementary Figure 8. Distribution of mutation rate predictions for simulated data. Density distribution of the mutation rate predictions for each true simulated mutation rates (dashed red line). Mutation rates are scaled to a population size (*N*) of 10,000. Mutation rate is plotted on a log10 scale and scaled to the per generation mutation rate. The means of the distribution of estimates for each mutation rate were relatively close to the true simulated mutation rates with the true mutation rate falling within one standard deviation of the mean estimate in both cases (true mutation rate = 4.98e-11 per bp per division, mean estimated mutation rate = 6.06e-11 ± 3.86e-11; true mutation rate = 5.01e-10 per bp per division, mean estimated mutation rate = 4.59e-10 ± 1.00e-10). Source data are provided as a Source Data file.



Supplementary Figure 9. Mutation rate predictions across ensemble of DNNs. Samples are shown along the x-axis with every 10th sample labeled for clarity. Samples are sorted according to mean mutation rate. Mutation rate prediction for each sample is shown for each DNN within the ensemble (y axis). Mutation rates range from black (low mutation rate) to red (high mutation rate). Mutation rates vary by two orders of magnitude across the cohort regardless of health outcome. Source data are provided as a Source Data file.



Supplementary Figure 10. Impact of population size on mutation rate. The expected number of mutations in a population (θ) is a product of the population size N and the per generation mutation rate (μ): $\theta=4N\mu$. Estimates of mutation rate can be extended to account for the range of population size estimates (10,000 – 200,000) which exist for the hematopoietic stem cell population. Here, we show the best fit per generation mutation rate on the (y axis) for each sample (x axis, n = 477). Samples are sorted according to mean mutation rate. Mutation rates have been scaled across varying HSC population size estimates. Source data are provided as a Source Data file.



Supplementary Figure 11. Distribution of mutations in known driver genes across evolutionary classes. Here, we show the frequency at which driver genes harbor mutations across individuals fitting different evolutionary classes. Driver genes are shown along the x axis. Each gene is partitioned according to the frequency at which it is mutated by class (positive = green, combination = orange, neutral = blue). No individuals harbouring mutations in known driver genes were predicted to be under negative selection. Source data are provided as a Source Data file.