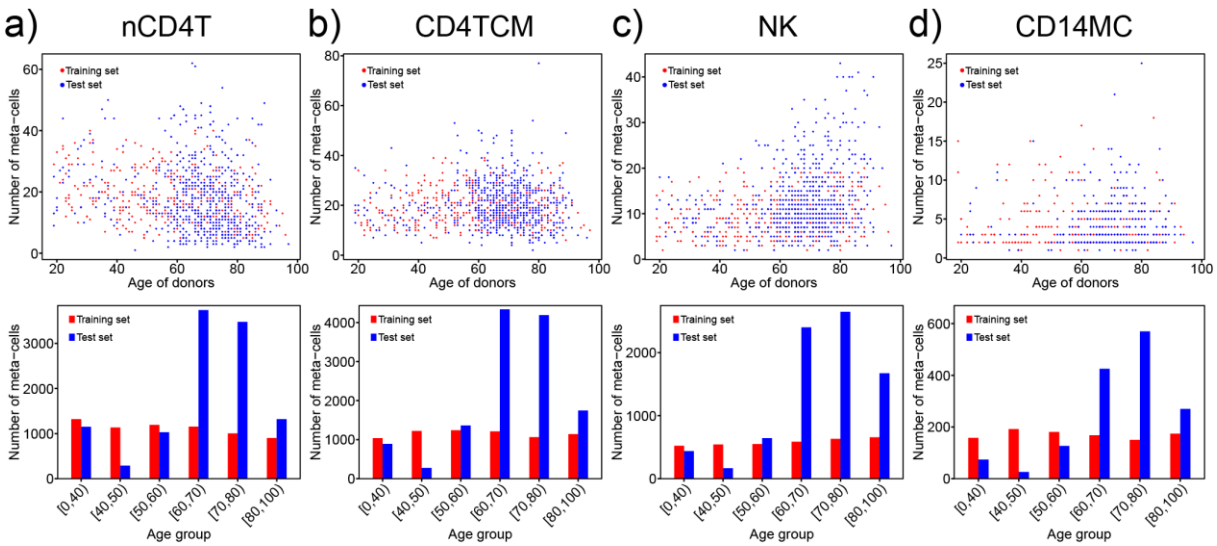
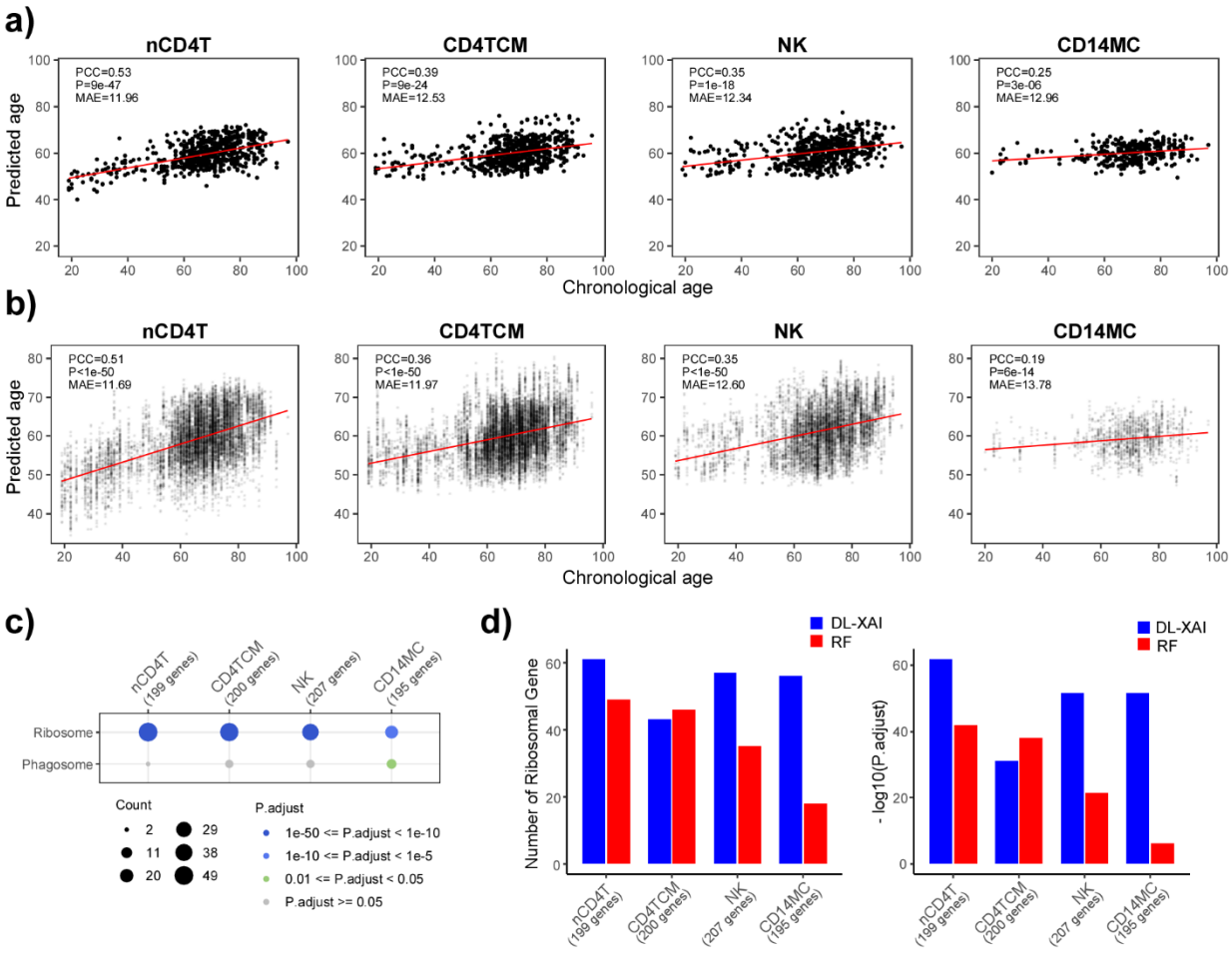
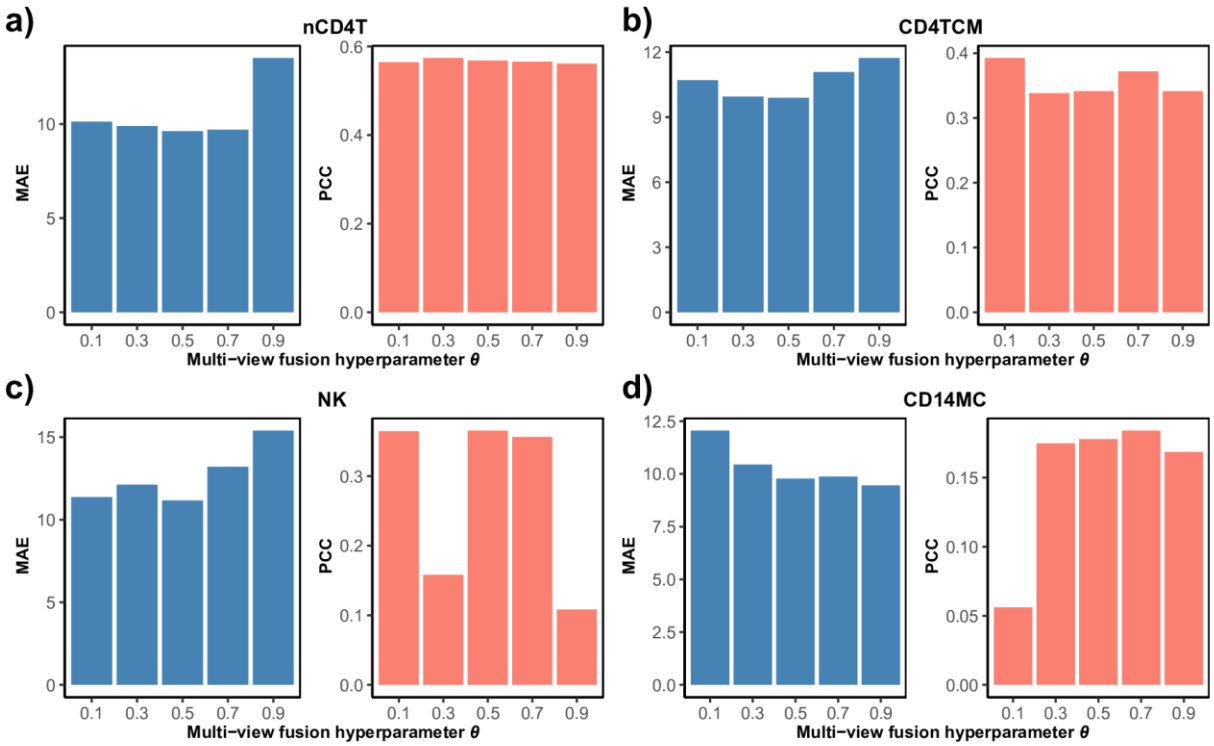
Supplementary Figures



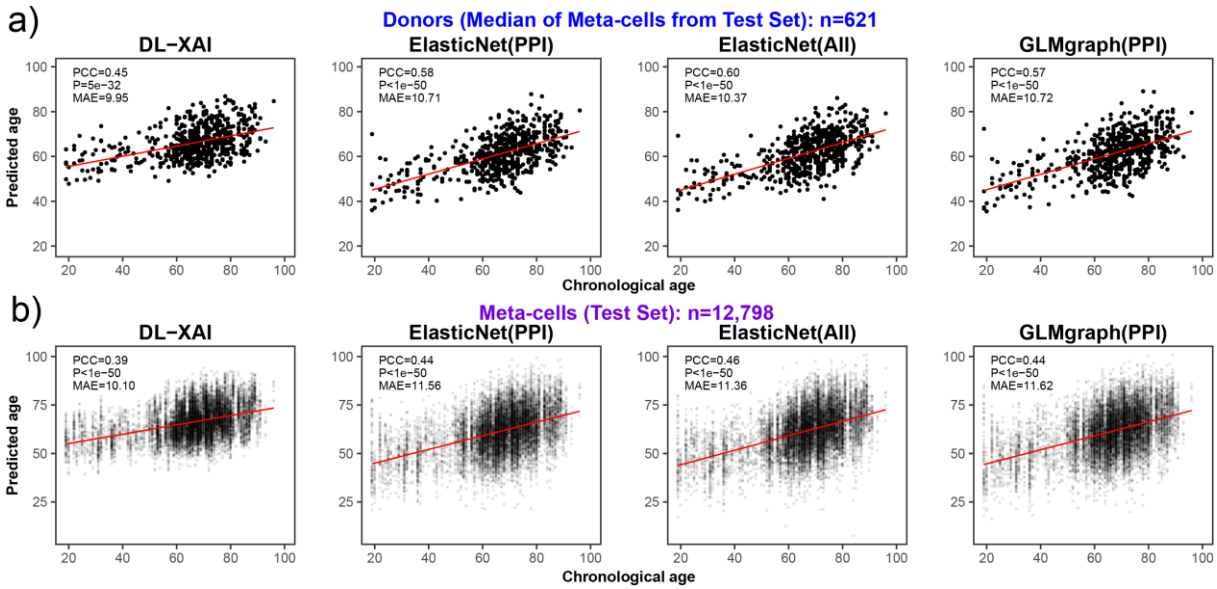
SI fig.S1: Distribution of meta-cell numbers per age-group and training/test set partition. a) Scatterplot of the number of naïve CD4+ T meta-cells (nCD4T) (y-axis) against donor-age (x-axis), with meta-cells used in training colored in red, and those for testing colored in blue, and barplot below displays the number of naïve CD4+ T meta-cells per age-group for training and test sets. **b)** As in a) but for CD4+ central memory T (CD4TCM) meta-cells. **c)** As a) but for NK meta cells. **d)** As a) but for CD14+ monocyte (CD14-MC) meta-cells.



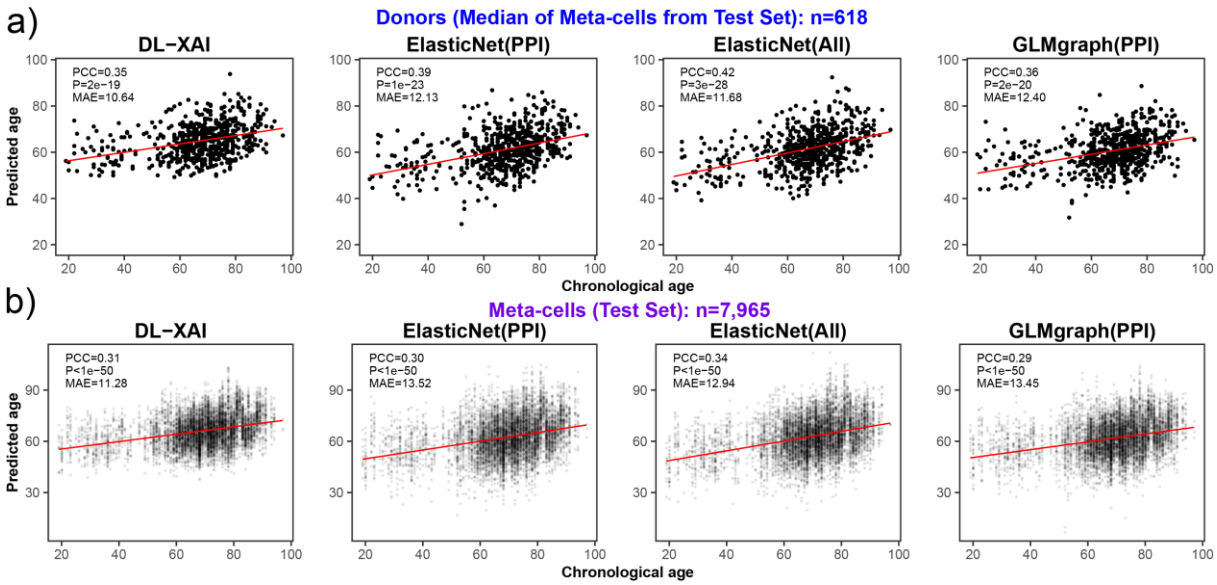
SI fig.S2: Random Forest model predictions and enrichment analysis. a) Scatterplot of donors' predicted age (y-axis) against donors' chronological age (x-axis) for 4 different immune cell-types. The predicted age of a given donor (taking the median prediction over all meta-cells of a given donor) was obtained with a Random Forest model. PCC represents Pearson correlation coefficient, MAE represents mean absolute error, P-value was obtained from Pearson correlation analysis. **b)** Scatterplot of predicted age of meta-cells (y-axis) against meta cells' chronological age (x-axis) for the 4 immune cell-types. The predicted age was obtained with a Random Forest model. PCC represents Pearson correlation coefficient, MAE represents mean absolute error, P-value was obtained from Pearson correlation analysis. **c)** Enrichment analysis result for top genes selected by Random Forest model in each of the 4 immune cell-types, highlighting the top 2 enriched terms. **d)** Barplot side-by-side comparison of the enriched ribosomal gene count (y-axis, left panel) and -$\log_{10}$[adjusted P-value] (y-axis, right panel) for the DL-XAI and Random Forest (RF) models in each of the 4 immune cell-types. P value was obtained from KEGG enrichment analysis and has been adjusted by Benjamini-Hochberg (BH) procedure.



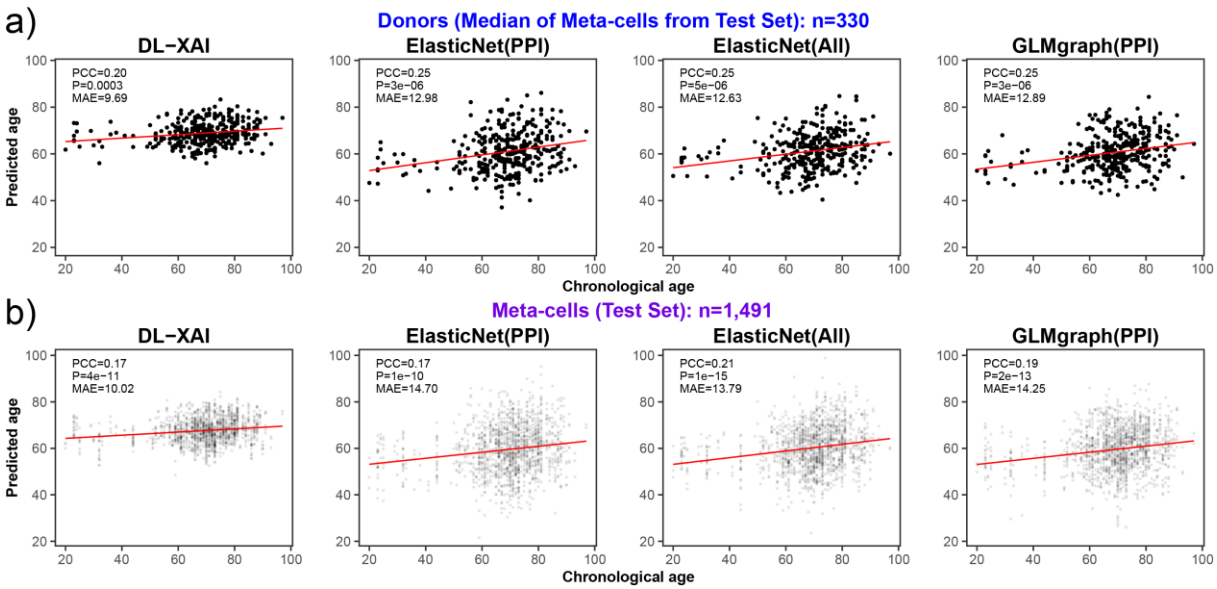
SI fig.S3: Effect on model performance of different multi-view fusion hyperparameter values in different immune cell types. a) Analysis of the impact of hyperparameter tuning for various multi-view fusion strategies on model performance in naïve CD4+ T cell scRNA-seq dataset. PCC represents Pearson correlation coefficient, MAE represents mean absolute error. **b)** As a) but for the central memory CD4+ T-cells. **c)** As a) but for natural killer (NK) cells. **d)** As a) but for CD14+ monocytes.



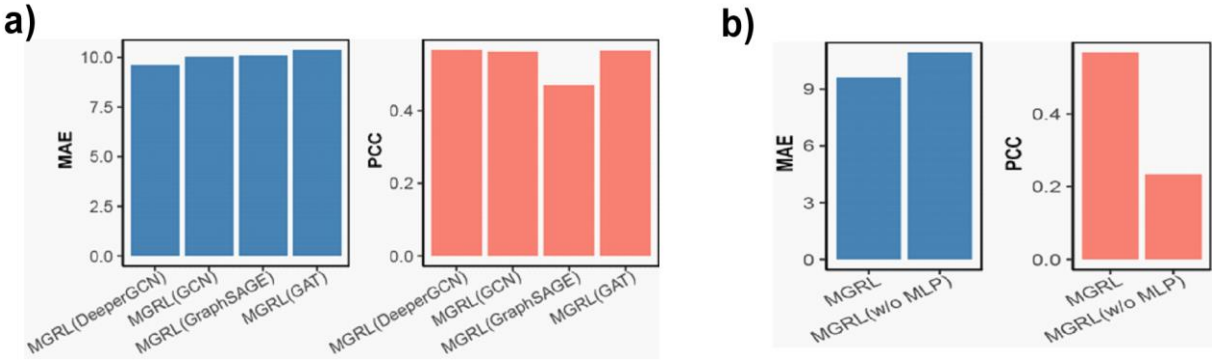
SI fig.S4: Predicted age versus chronological age for CD4+ central memory T (CD4TCM) meta-cells. a) Scatterplot of donors' predicted median age over CD4TCM meta-cells (y-axis) against donors' chronological age (x-axis). The predicted age of a given donor (taking the median prediction over all meta-cells of a given donor) obtained by DL-XAI clock, the elastic net models with PPI restricted and all genes, and GLMgraph model, respectively. **b)** Scatterplot of predicted age of CD4TCM meta cells (y-axis) against chronological age (x-axis), the predicted age obtained by DL-XAI clock, the elastic net models with PPI restricted and all genes and GLMgraph model, respectively. PCC represents Pearson correlation coefficient, MAE represents mean absolute error. P-value was obtained from Pearson correlation analysis.



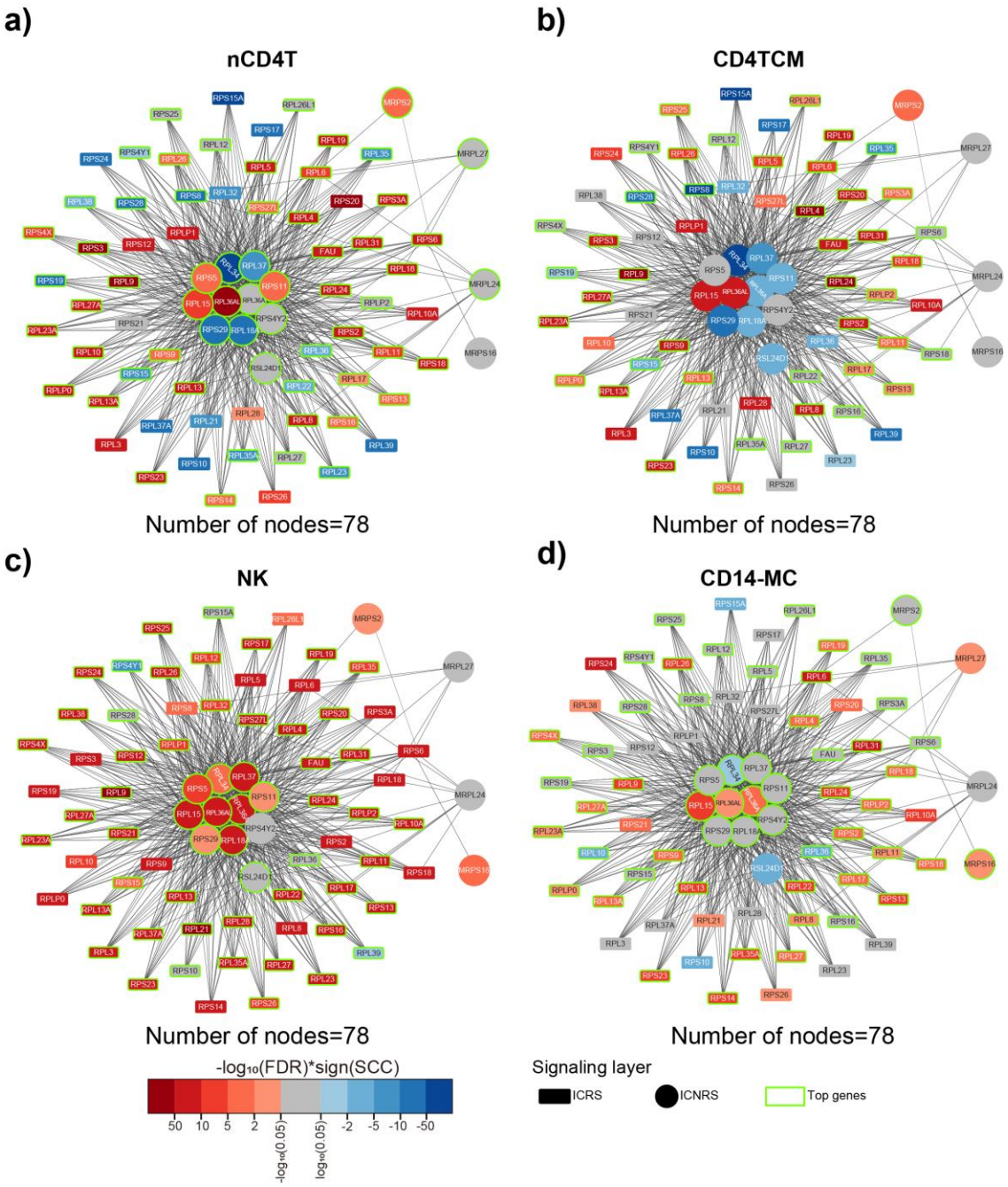
SI fig.S5: Predicted age as a function of the chronological age for natural killer (NK) meta-cells. a) Scatterplot of donors' predicted median age over NK meta-cells (y-axis) against donors' chronological age (x-axis). The predicted age of a given donor (taking the median prediction over all meta-cells of a given donor) obtained by DL-XAI clock, the elastic net models run with PPI-restricted and all genes, and GLMgraph model, respectively. **b)** Scatterplot of predicted age of NK meta-cells (y-axis) against chronological age (x-axis); predicted age obtained by DL-XAI clock, the elastic net models with PPI restricted and all genes, and GLMgraph model, respectively. PCC represents Pearson correlation coefficient, MAE represents mean absolute error. P-value was obtained from Pearson correlation analysis.



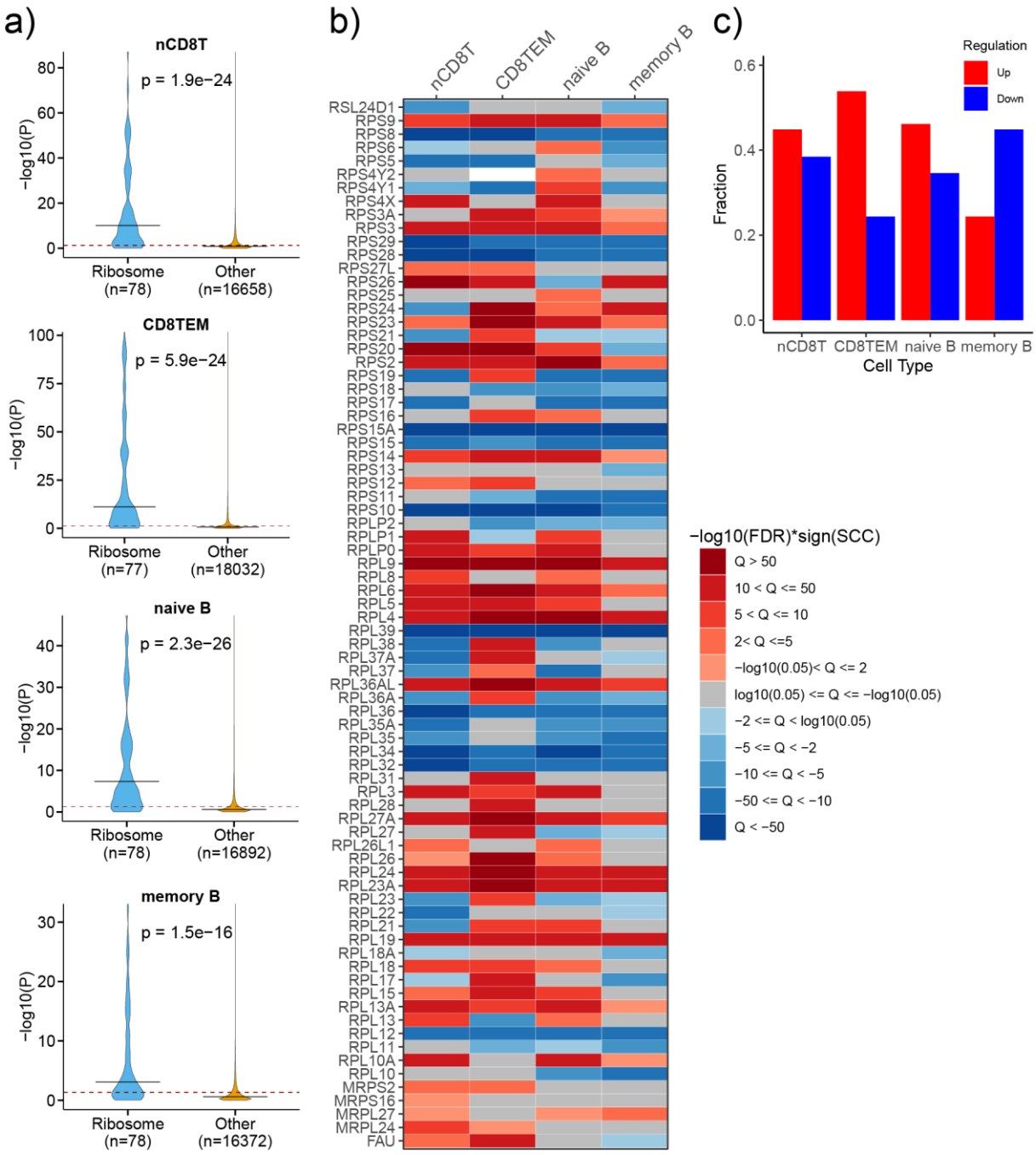
SI fig.S6: Predicted age as a function of the chronological age for CD14+ monocyte (CD14-MC) meta-cells. a) Scatterplot of donors' predicted median age over CD14-MC meta-cells (y-axis) against donors' chronological age (x-axis). The predicted age of a given donor (taking the median prediction over all meta-cells of a given donor) obtained by DL-XAI clock, the elastic net models with PPI-restricted and all genes, and GLMgraph model, respectively. **b)** Scatterplot of predicted age of CD14-MC meta cells (y-axis) against chronological age (x-axis); predicted age obtained by DL-XAI clock, the elastic net models with PPI restricted and all genes, and GLMgraph model respectively. PCC represents Pearson correlation coefficient, MAE represents mean absolute error. P-value was obtained from Pearson correlation analysis.



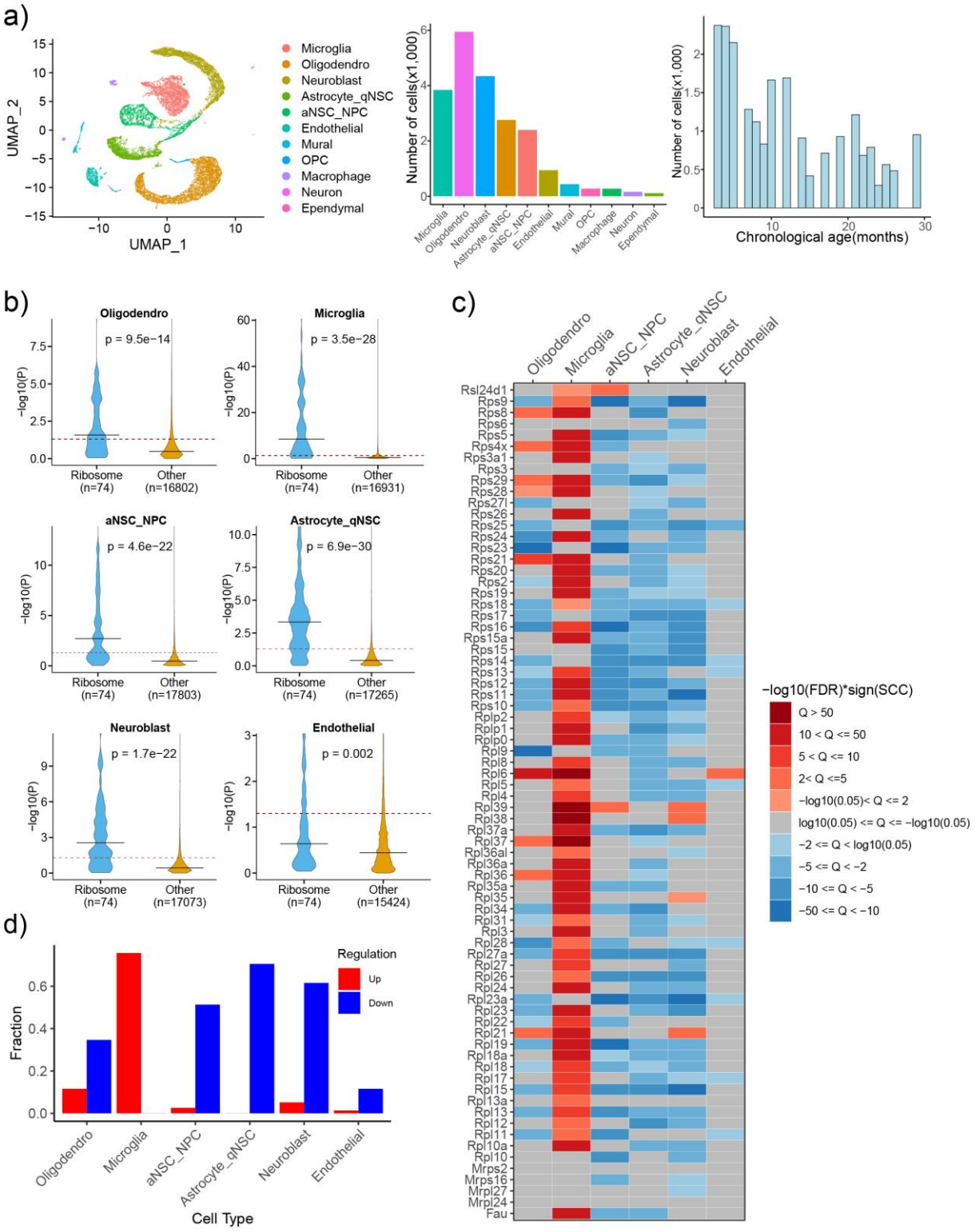
SI fig.S7: Benchmarking of MGRL and DeeperGCN. a) Performance of different graph convolutions in MGRL on the naïve CD4+ T cell scRNA-seq dataset. PCC represents Pearson correlation coefficient, MAE represents mean absolute error. **b)** Performance comparison of MGRL with and without (w/o) MLP on the naïve CD4+ T cell scRNA-seq dataset. PCC represents Pearson correlation coefficient, MAE represents mean absolute error.



SI fig.S8: Ribosome module that changes with age in different cell types. a) Ribosome module that changes with age in naïve CD4+ T (nCD4T) meta-cells. The rectangular node represents that the gene is located in the ICRS signaling layer, the round node represents that the gene is located in the ICNRS signaling layer, and the nodes with fluorescent green borders indicate the top genes in each cell type. The Spearman's rank correlation coefficient (SCC) between gene expression and age was obtained from training set. FDR was obtained by correcting the P value by the Benjamini-Hochberg (BH) method, the P value was obtained from the Spearman correlation analysis. **b)** As in a) but for CD4+ central memory T (CD4TCM) meta-cells. **c)** As a) but for NK meta cells. **d)** As a) but for CD14+ monocyte (CD14-MC) meta-cells.

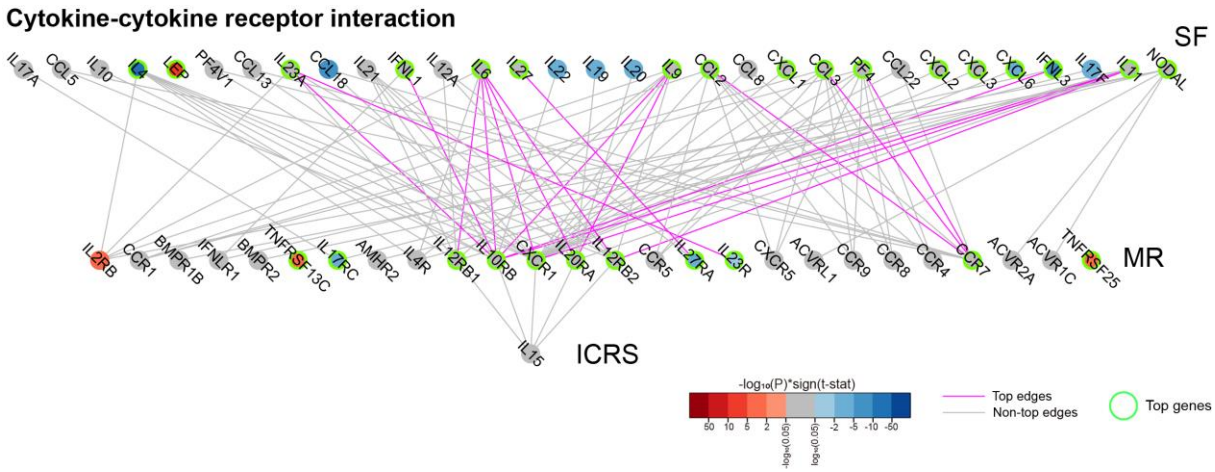


SI fig.S9: Ribosomal module significantly correlates with age in other immune cell types. a) Violin plots display the -log10 transformed P-values with P-values obtained from a Spearman correlation test between gene expression and age, for two groups of genes: Ribosome = the 78 ribosomal genes, Other = all other genes. The P-value in the violin plot indicates whether the difference between the two groups is statistically significant, and is derived from a one-tailed Wilcoxon-rank sum test. **b)** Heatmap of the correlation of 78 ribosomal genes with age in these four cell types. SCC denotes spearman's rank correlation coefficient, P-values have been adjusted using the Benjamini-Hochberg (BH) procedure to yield FDR values. **c)** Bar plot displays the fraction of upregulated ribosomal genes and downregulated ribosomal genes across different cell types.

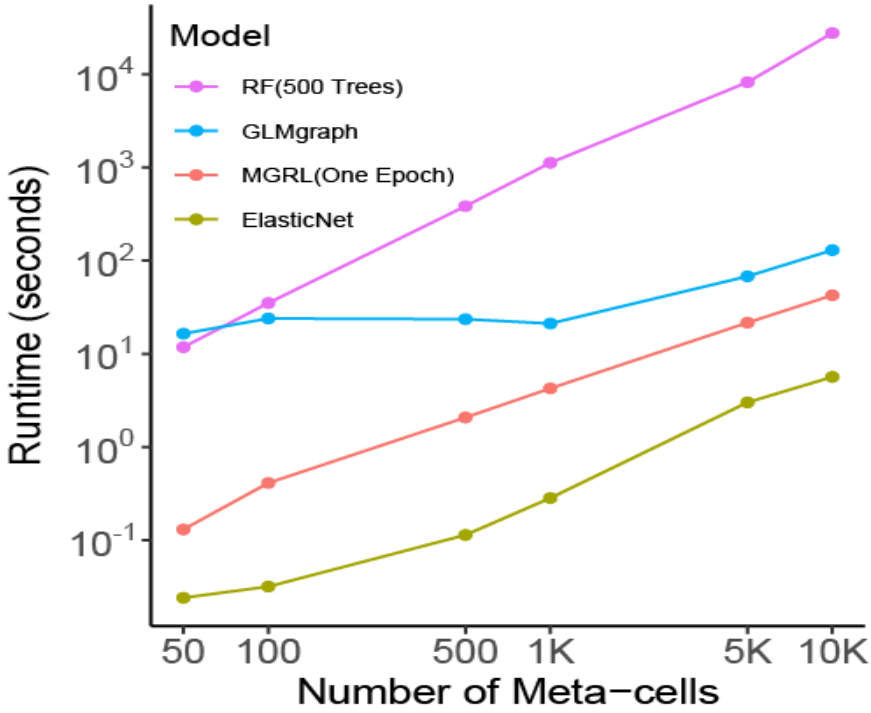


SI fig.S10: Ribosomal module is associated with age in mouse brain. a) The scRNA-seq data of Buckley et al encompassing over 21,458 cells from 28 mice. Left panel: UMAP of the scRNA-Seq data of Buckley et al. Middle panel: Barplot displaying the number of cells per cell-type. Right panel: Barplot displaying the number of cells in each age group. **b)** Violin plots display the -log10 transformed P-values obtained from a Spearman correlation test between gene expression and age, for two groups of genes: Ribosome = the mouse homolog IDs of the 78 ribosomal genes,

Other = all other genes. The P-value indicates whether the difference between the two groups is statistically significant, and is derived from a one-tailed Wilcoxon-rank sum test. **c)** Heatmap of the correlation of ribosomal genes with age in these six cell types. SCC denotes spearman's rank correlation coefficient, P-values have been adjusted using the Benjamini-Hochberg (BH) procedure to yield FDR values. **d)** Bar plot displays the fraction of upregulated and downregulated ribosomal genes across different cell types.



SI fig.S11: Enriched cytokine-cytokine receptor interaction network extracted by DL-XAI. The top genes selected by PGExplainer enriched in the cytokine-cytokine receptor interaction network. The top selected genes are labeled with a bright green border. Magenta line indicates top selected edges, and light gray lines indicate remaining edges in the PPI network. The t-statistic (t-stat) colors the nodes in the network and were obtained from linear regression of promoter DNAm against age in the Reynolds DNAm data set.



SI fig.S12: Comparison of training time required for different models. The plot compares the approximate runtime (y-axis) as a function of the number of meta-cells for four different methods. All methods were run on the same input PPI graph, consisting of 8511 nodes. The time taken to train a single epoch of the MGRL model was measured on a system with two NVIDIA A100 GPUs, while the training times for the other models were measured on a system equipped with an Intel Xeon E5-4660 v4 2.20 GHz, utilizing only a single core. The hyperparameters of these models have been fixed, where the Random Forest was constructed using 500 trees, each with a minimum node size of 5.