Supplementary Material

Supplementary Table S1. The detail information of GEO datasets.

GEO	Platform	CD	Control
GSE93624	GPL11154	210	35
GSE57945	GPL11154	218	48
GSE123141	GPL16791	9	9
GSE179285	GPL6480	47	31
GSE126124	GPL6244	37	19
GSE16879	GPL570	73	13
GSE3365	GPL96	59	42
GSE193677	GPL16791	1151	455

Supplementary Table S2. Primer sequence for qPCR.

Gene	Forward Primer	Reverse Primer
FPR1	GCTGGCTACATCGTTCTGGATG	GTTTCATGCGGAAACCAGCCAC
CD86	ACGTATTGGAAGGAGATTACAGCT	TCTGTCAGCGTTACTATCCCGC
IL1β	TGGACCTTCCAGGATGAGGACA	GTTCATCTCGGAGCCTGTAGTG

Supplementary Table S3. The AUC of each model.

Combined model	GSE93624	GSE57945	GSE179285	GSE16879	GSE126124
Lasso+Stepglm[both]	0.965	0.928	0.969	0.943	0.952
SVM	0.836	0.616	0.500	0.528	0.526
glmBoost+SVM	0.824	0.607	0.500	0.528	0.526
Ridge	0.965	0.928	0.972	0.925	0.954
Lasso+SVM	0.824	0.619	0.500	0.528	0.526
glmBoost+Ridge	0.965	0.928	0.972	0.925	0.954
Enet[alpha=0.1]	0.965	0.929	0.970	0.933	0.952
glmBoost+Enet[alpha=0.1]	0.966	0.929	0.970	0.932	0.953
Enet[alpha=0.2]	0.966	0.929	0.970	0.932	0.952
Enet[alpha=0.3]	0.965	0.929	0.970	0.933	0.952
glmBoost+Enet[alpha=0.3]	0.966	0.929	0.970	0.934	0.952
glmBoost+Enet[alpha=0.2]	0.966	0.929	0.970	0.933	0.952
Enet[alpha=0.4]	0.965	0.929	0.970	0.933	0.952
glmBoost+Enet[alpha=0.4]	0.965	0.929	0.970	0.933	0.952
Lasso+glmBoost	0.965	0.927	0.969	0.947	0.952
Enet[alpha=0.5]	0.965	0.929	0.969	0.936	0.953
glmBoost	0.965	0.927	0.969	0.947	0.952
glmBoost+Enet[alpha=0.5]	0.965	0.929	0.970	0.935	0.953
Enet[alpha=0.6]	0.965	0.929	0.970	0.935	0.953
glmBoost+Enet[alpha=0.6]	0.965	0.929	0.969	0.936	0.953
glmBoost+Enet[alpha=0.7]	0.966	0.929	0.969	0.938	0.952
glmBoost+Enet[alpha=0.8]	0.965	0.929	0.969	0.941	0.952
Enet[alpha=0.8]	0.965	0.929	0.969	0.941	0.952
Enet[alpha=0.9]	0.965	0.928	0.969	0.943	0.952
Lasso	0.965	0.927	0.969	0.945	0.950
Enet[alpha=0.7]	0.966	0.929	0.969	0.938	0.952
glmBoost+Enet[alpha=0.9]	0.965	0.928	0.969	0.943	0.952
glmBoost+Lasso	0.965	0.928	0.969	0.945	0.950
Lasso+plsRglm	0.965	0.928	0.973	0.920	0.954
glmBoost+plsRglm	0.965	0.928	0.973	0.920	0.954
glmBoost+Stepglm[forward]	0.965	0.928	0.969	0.943	0.952
Lasso+Stepglm[forward]	0.965	0.928	0.969	0.943	0.952
RF+SVM	0.836	0.616	0.500	0.528	0.526
Stepglm[forward]	0.965	0.928	0.969	0.943	0.952
plsRglm	0.965	0.928	0.973	0.920	0.954
RF+Ridge	0.965	0.928	0.972	0.925	0.954
RF+Enet[alpha=0.1]	0.965	0.929	0.970	0.933	0.952
RF+plsRglm	0.965	0.928	0.973	0.920	0.954
RF+Stepglm[forward]	0.965	0.928	0.969	0.943	0.952
RF+Enet[alpha=0.2]	0.966	0.929	0.970	0.932	0.953
RF+Enet[alpha=0.3]	0.966	0.929	0.970	0.934	0.952

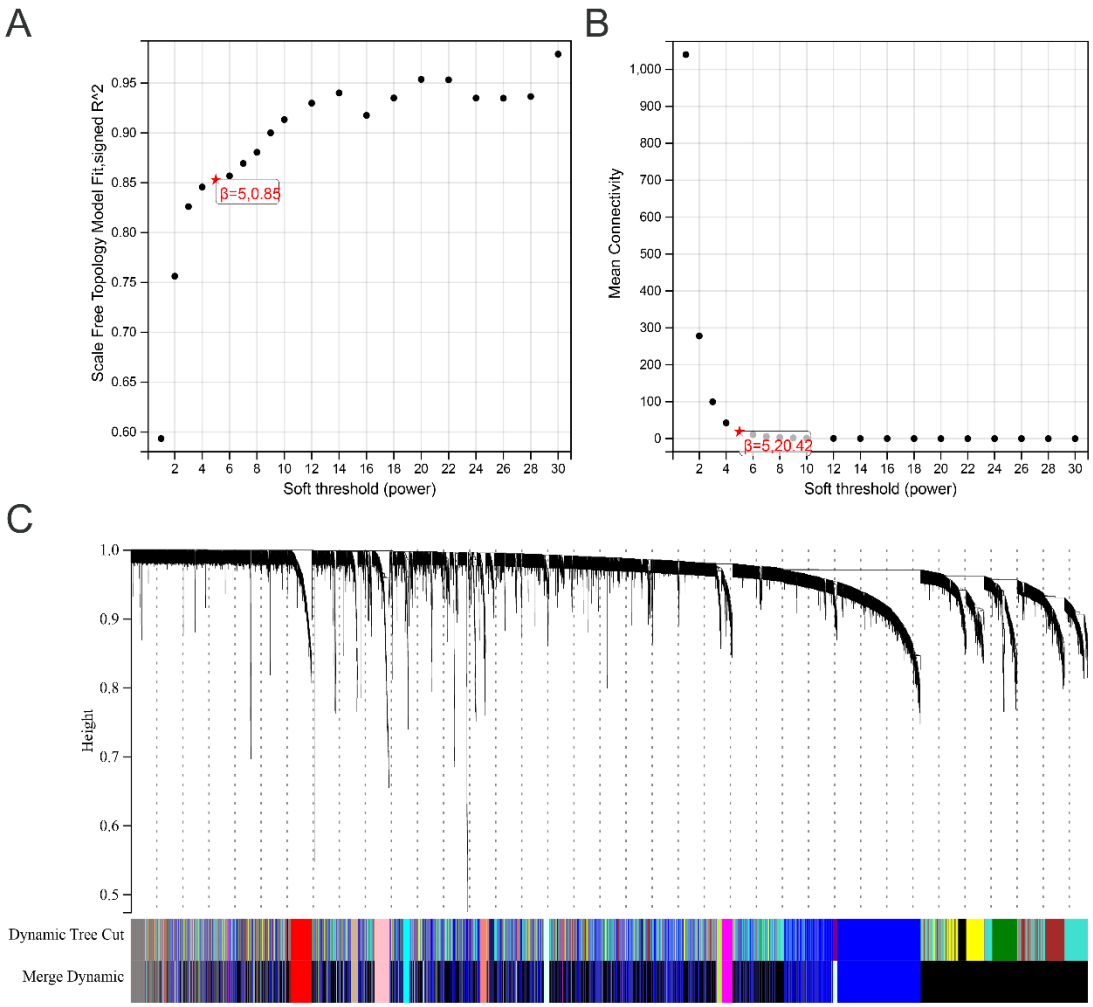
RF+Enet[alpha=0.6]	0.966	0.929	0.970	0.934	0.952
RF+Lasso	0.965	0.927	0.969	0.946	0.952
RF+Enet[alpha=0.7]	0.965	0.929	0.969	0.939	0.952
RF+Enet[alpha=0.5]	0.965	0.929	0.970	0.935	0.953
RF+glmBoost	0.965	0.927	0.969	0.947	0.952
RF+Enet[alpha=0.9]	0.965	0.928	0.969	0.943	0.952
RF+Enet[alpha=0.4]	0.966	0.929	0.970	0.934	0.952
RF+Enet[alpha=0.8]	0.965	0.929	0.969	0.941	0.952
RF+Stepglm[both]	0.965	0.928	0.969	0.943	0.952
RF+Stepglm[backward]	0.965	0.928	0.969	0.943	0.952
Stepglm[both]+Ridge	0.965	0.928	0.972	0.925	0.954
Stepglm[backward]+Ridge	0.965	0.928	0.972	0.925	0.954
Stepglm[both]+plsRglm	0.965	0.928	0.973	0.920	0.954
Stepglm[backward]+plsRglm	0.965	0.928	0.973	0.920	0.954
Stepglm[both]+Enet[alpha=0.9]	0.965	0.928	0.969	0.943	0.952
Stepglm[backward]+Enet[alpha=0.9]	0.965	0.928	0.969	0.943	0.952
Stepglm[both]+Enet[alpha=0.1]	0.966	0.929	0.970	0.932	0.953
Stepglm[backward]+Enet[alpha=0.1]	0.966	0.929	0.970	0.932	0.953
Stepglm[both]+Enet[alpha=0.8]	0.965	0.929	0.969	0.941	0.952
Stepglm[backward]+Enet[alpha=0.8]	0.965	0.928	0.969	0.942	0.952
Stepglm[both]+Enet[alpha=0.2]	0.966	0.929	0.970	0.932	0.953
Stepglm[backward]+Enet[alpha=0.2]	0.966	0.929	0.970	0.932	0.953
Stepglm[both]+Lasso	0.965	0.927	0.969	0.945	0.950
Stepglm[backward]+Lasso	0.965	0.928	0.969	0.945	0.950
Stepglm[both]+Enet[alpha=0.6]	0.965	0.929	0.969	0.936	0.953
Stepglm[backward]+Enet[alpha=0.6]	0.965	0.929	0.969	0.936	0.953
glmBoost+GBM	0.989	0.931	0.965	0.933	0.932
Stepglm[both]+Enet[alpha=0.7]	0.966	0.929	0.969	0.938	0.952
Stepglm[backward]+Enet[alpha=0.7]	0.965	0.929	0.969	0.939	0.952
Lasso+Stepglm[backward]	0.965	0.928	0.969	0.943	0.952
Stepglm[both]	0.965	0.928	0.969	0.943	0.952
Stepglm[backward]	0.965	0.928	0.969	0.943	0.952
glmBoost+Stepglm[both]	0.965	0.928	0.969	0.943	0.952
glmBoost+Stepglm[backward]	0.965	0.928	0.969	0.943	0.952
Stepglm[both]+Enet[alpha=0.4]	0.965	0.929	0.970	0.933	0.952
Stepglm[backward]+Enet[alpha=0.4]	0.965	0.929	0.970	0.935	0.953
Stepglm[both]+Enet[alpha=0.3]	0.966	0.929	0.970	0.932	0.953
Stepglm[backward]+Enet[alpha=0.3]	0.965	0.929	0.970	0.933	0.952
Stepglm[both]+glmBoost	0.965	0.927	0.969	0.947	0.952
Stepglm[backward]+glmBoost	0.965	0.927	0.969	0.947	0.952
Stepglm[both]+Enet[alpha=0.5]	0.966	0.929	0.970	0.935	0.952
Stepglm[backward]+Enet[alpha=0.5]	0.966	0.929	0.969	0.938	0.952
glmBoost+RF	0.993	0.895	0.927	0.947	0.936

RF	0.993	0.894	0.926	0.946	0.935
Lasso+GBM	0.985	0.931	0.965	0.937	0.932
RF+GBM	0.988	0.931	0.964	0.937	0.930
GBM	0.984	0.932	0.967	0.942	0.930
Stepglm[both]+SVM	0.824	0.619	0.500	0.528	0.526
Stepglm[backward]+SVM	0.836	0.616	0.500	0.528	0.526
Lasso+RF	0.993	0.895	0.926	0.947	0.936
Stepglm[both]+GBM	0.987	0.932	0.963	0.937	0.927
Stepglm[backward]+GBM	0.986	0.931	0.966	0.943	0.930
Stepglm[both]+RF	0.993	0.895	0.927	0.947	0.936
LDA	0.965	0.927	0.968	0.946	0.949
glmBoost+LDA	0.965	0.927	0.968	0.946	0.949
RF+LDA	0.965	0.927	0.968	0.946	0.949
Stepglm[both]+LDA	0.965	0.927	0.968	0.946	0.949
Stepglm[backward]+LDA	0.965	0.927	0.968	0.946	0.949
Lasso+LDA	0.965	0.927	0.968	0.946	0.949
Stepglm[backward]+RF	0.993	0.895	0.925	0.946	0.936
XGBoost	0.991	0.917	0.933	0.963	0.935
Lasso+XGBoost	0.988	0.914	0.934	0.954	0.940
glmBoost+XGBoost	0.978	0.891	0.907	0.954	0.911
RF+XGBoost	0.993	0.916	0.933	0.943	0.930
Stepglm[both]+XGBoost	0.990	0.915	0.933	0.963	0.935
Stepglm[backward]+XGBoost	0.990	0.915	0.933	0.963	0.935
NaiveBayes	0.967	0.925	0.972	0.926	0.954
Lasso+NaiveBayes	0.967	0.925	0.972	0.926	0.954
glmBoost+NaiveBayes	0.967	0.925	0.972	0.926	0.954
RF+NaiveBayes	0.967	0.925	0.972	0.926	0.954
Stepglm[both]+NaiveBayes	0.967	0.925	0.972	0.926	0.954
Stepglm[backward]+NaiveBayes	0.967	0.925	0.972	0.926	0.954

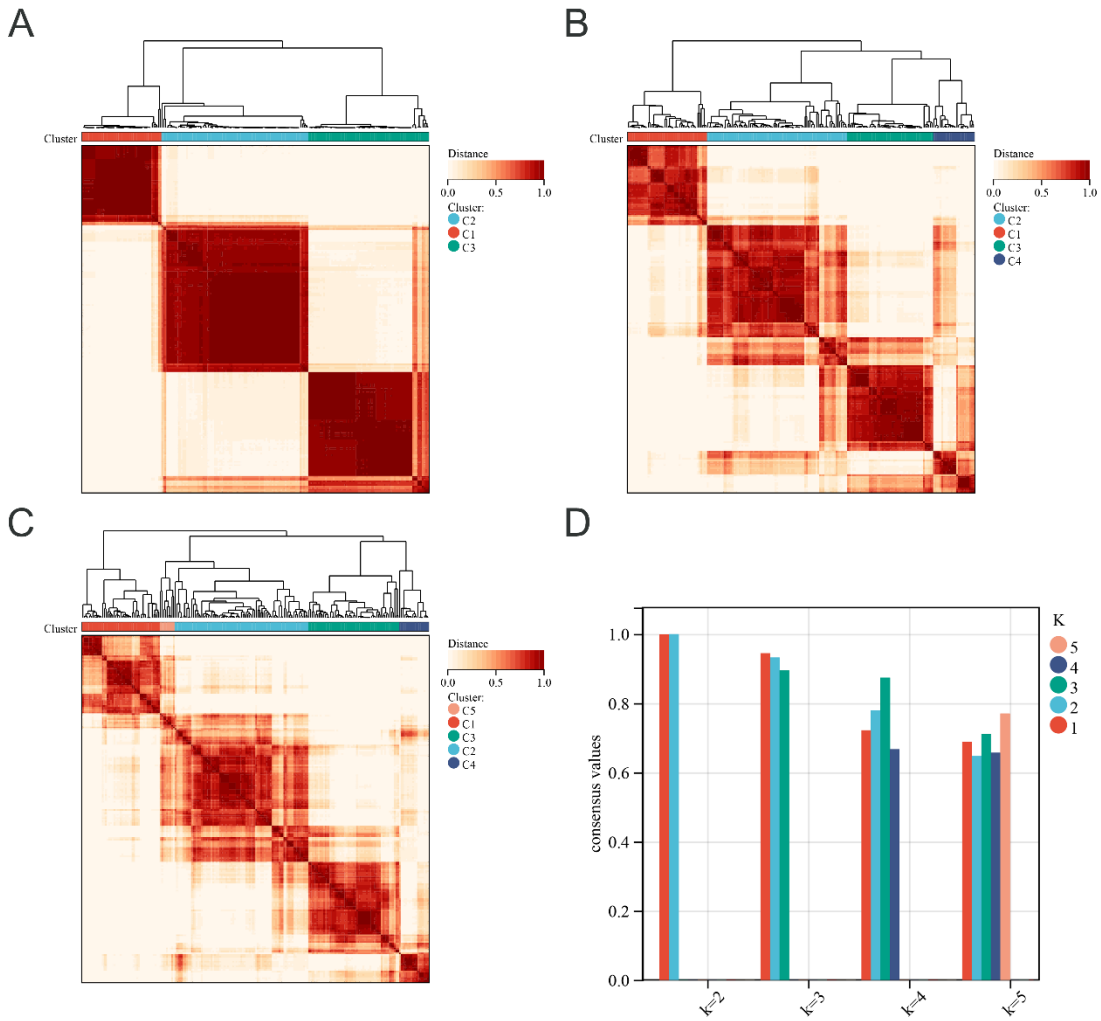
Supplementary Table S4. The published signatures and biomarkers from CD studies

Ensembl	Model	PMID	Author	Gene name
ENSG00000167984	Wu-FI	36685554	Wu	NLRC3
ENSG00000154451	Wu-FI	36685554	Wu	GBP5
ENSG00000005100	Wu-FI	36685554	Wu	DHX33
ENSG00000140368	Wu-FI	36685554	Wu	PSTPIP1
ENSG00000109320	Wu-FI	36685554	Wu	NFKB1
ENSG00000136810	Wu-FI	36685554	Wu	TXN
ENSG00000103313	Wu-FI	36685554	Wu	MEFV
ENSG00000077150	Wu-FI	36685554	Wu	NFKB2
ENSG00000174130	Wu-FI	36685554	Wu	TLR6
ENSG00000137752	Wu-FI	36685554	Wu	CASP1
ENSG00000110218	Wu-FI	36685554	Wu	PANX1
ENSG00000173039	Wu-FI	36685554	Wu	RELA
ENSG00000096384	Wu-FI	36685554	Wu	HSP90AB1
ENSG00000105483	Wu-FI	36685554	Wu	CARD8
ENSG00000232810	Dai-IJMS	37047025	Dai	TNF
ENSG00000023445	Dai-IJMS	37047025	Dai	BIRC3
ENSG00000135046	Dai-IJMS	37047025	Dai	ANXA1
ENSG00000050730	Dai-IJMS	37047025	Dai	TNIP3
ENSG00000134285	Dai-IJMS	37047025	Dai	FKBP11
ENSG00000085265	Chen-JTM	36932401	Chen	FCN1
ENSG00000011422	Zhang-CMMM	36643579	Zhang	PLAUR
ENSG00000122884	Ye-FIG	37428395	Ye	P4HA1
ENSG00000140650	Ye-FIG	37428395	Ye	PMM2
ENSG00000169245	Huang-FI	37622114	Huang	CXCL10
ENSG00000108821	Huang-FI	37622114	Huang	COL1A1

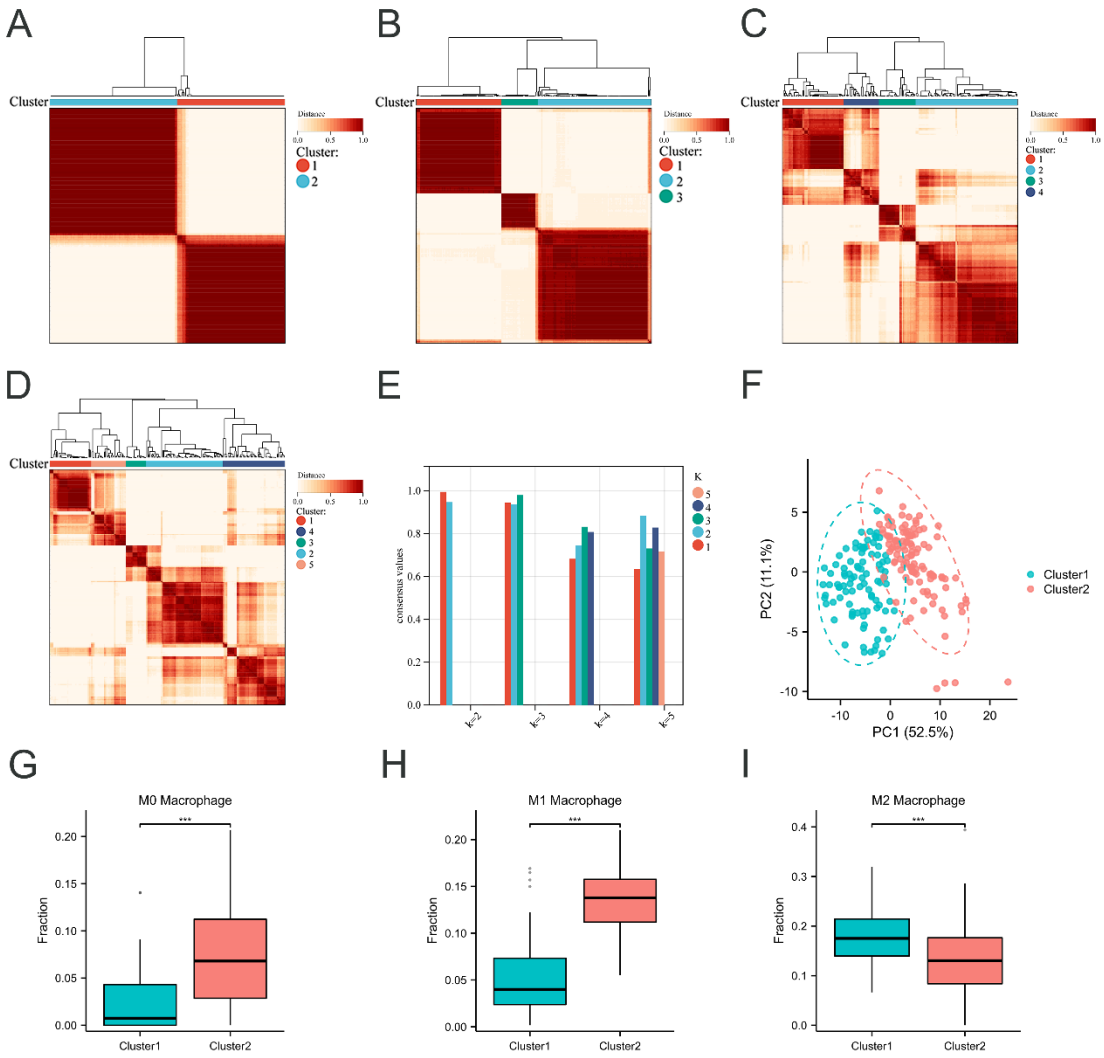
Note: Data from these studies.[30-35]



Supplementary Figure S1. WGCNA in GSE93624. (A-B) WGCNA soft threshold β setting. (C) Resulting gene dendrograms.



Supplementary Figure S2. Clustering for MRGs in GSE93624. (A-C) Consensus clustering matrices in CD merged datasets (k=3-5). (D) The bar plots represent the consensus scores for subtypes with k=2-5.



Supplementary Figure S3. Clustering for MRGs in GSE57945. (A-D) Consensus clustering matrices in CD merged datasets (k=2-5). (E) The bar plots represent the consensus scores for subtypes with k=2-5. (F) PCA for the expression of MRGs to distinguish two clusters. (G-I) The infiltration levels of M0, M1, and M2 macrophages in each cluster. (***p < 0.001).