

2 **Supporting Information for**

3 **Vision-Based Deep Learning of Biomedical Tabular Data via Self-Supervised Cartographic** 4 **Representation**

5 **Sakib Mostafa, Yuming Jiang, James Zou, Tarik F. Massoud, Ash A. Alizadeh, Maximilian Diehn, Lei Xing, and Md. Tauhidul**
6 **Islam**

7 **Corresponding Authors: Md. Tauhidul Islam and Lei Xing**

8 **E-mail: tauhid@stanford.edu and lei@stanford.edu**

9 **This PDF file includes:**

10 **Figs. S1 to S32**

11 **Tables S1 to S3**

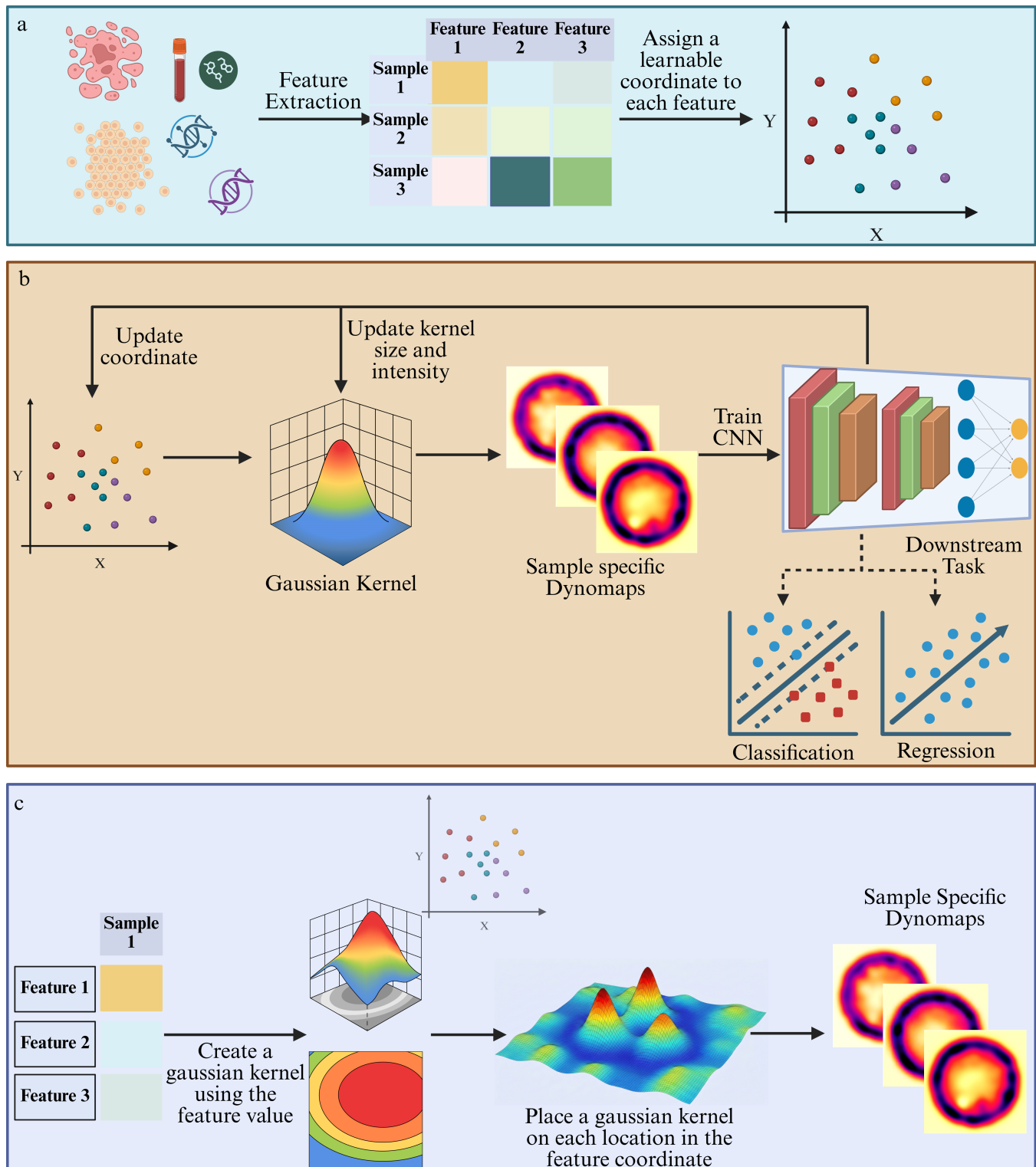


Fig. S1. Detailed illustration of Dynomap construction from learned feature coordinates. **a**, Tabular biomedical measurements are first organized into a feature matrix, where each column corresponds to a biological variable and each row represents a sample. Dynomap assigns every feature a learnable coordinate in a continuous two-dimensional space, establishing a spatial layout that reflects task-dependent relationships among variables. **b**, During model training, both the feature coordinates and rendering parameters are updated jointly with the predictive model. Feature values are projected onto the spatial grid through Gaussian kernels whose location corresponds to the learned feature coordinate and whose intensity reflects the feature magnitude. The resulting spatial maps are used as inputs to a convolutional neural network to perform downstream tasks such as classification or regression. **c**, Detailed rendering process for a single sample. Each feature value generates a Gaussian kernel centered at its learned coordinate. The superposition of all kernels produces a continuous two-dimensional surface, which is discretized into an image representation referred to as a Dynomap. This process preserves feature identity while translating tabular measurements into a spatial representation that can be analyzed using vision-based models.

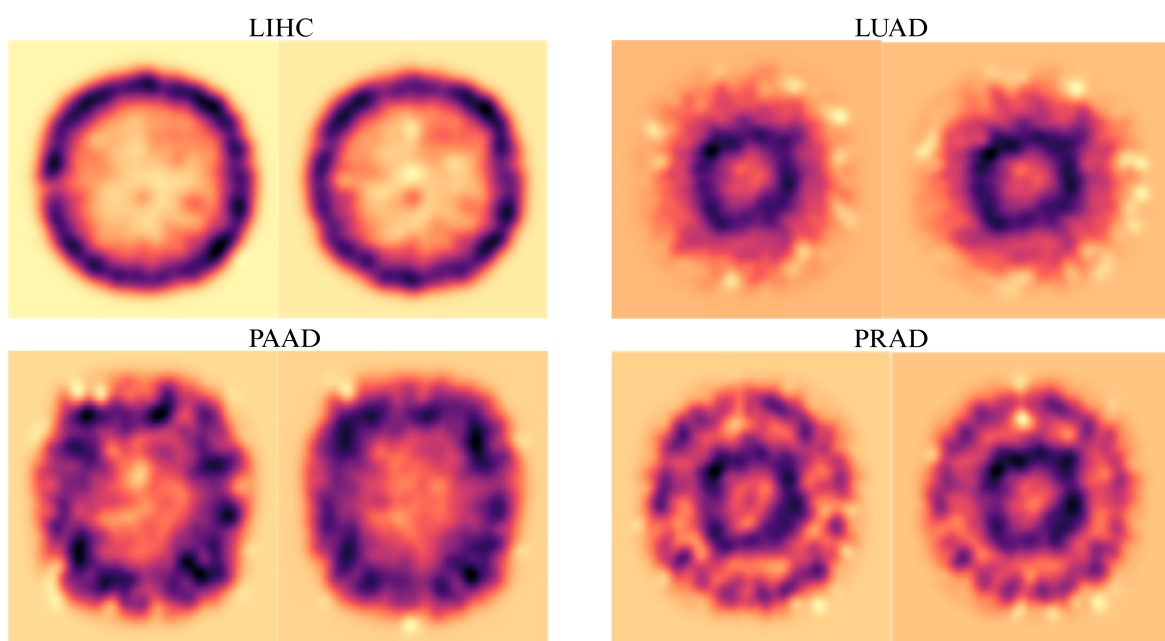


Fig. S2. Representative Dynomap image embeddings across RARE-Seq classes. Dynomap-generated image representations for individual RARE-Seq cfRNA samples, shown as class-wise examples across tasks. Within each class, samples exhibit visually consistent spatial organization, whereas different classes display distinct spatial configurations. These qualitative patterns illustrate that Dynomap maps tabular expression profiles into structured image-like representations without relying on a predefined gene layout. Color intensity denotes the normalized painted signal on the learned spatial canvas, with higher intensity corresponding to stronger localized contribution.

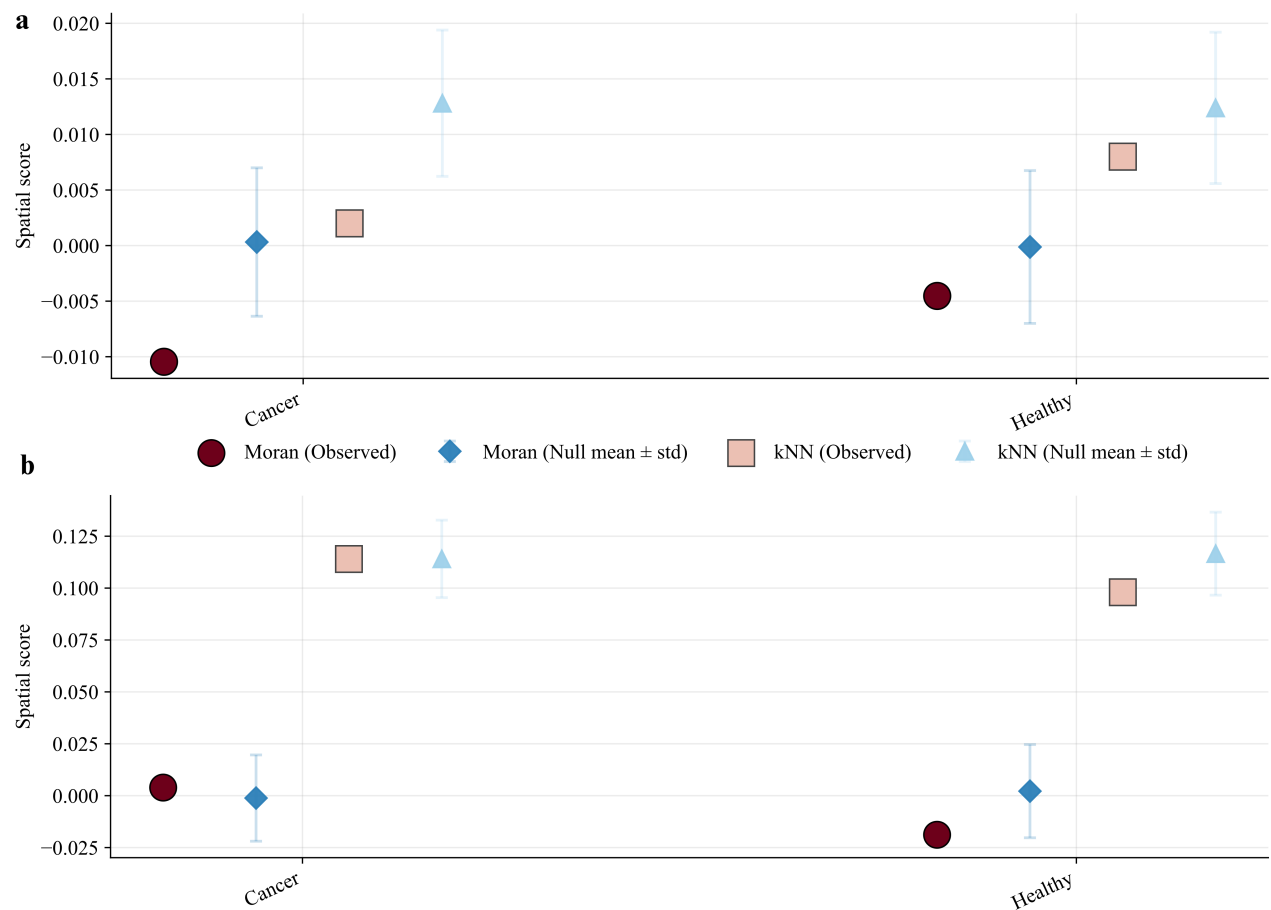


Fig. S3. Spatial organization of class-specific gene attributions in RARE-Seq binary classification. Spatial autocorrelation and neighborhood consistency of Integrated Gradients (IG) attribution maps for healthy versus cancer prediction, quantified using Moran's I and kNN purity. **a**, Observed Moran's I values for cancer- and healthy-associated genes compared with null distributions generated by spatial permutation. Cancer-associated attributions exhibit stronger deviation from null expectations, indicating non-random spatial organization. **b**, kNN-based neighborhood purity for the same attribution maps. Cancer-associated genes show elevated local neighborhood consistency relative to null expectations, whereas healthy-associated attributions display weaker spatial structure. Together, these results demonstrate that Dynomap organizes cancer-relevant genes into spatially coherent regions even in a binary classification setting.

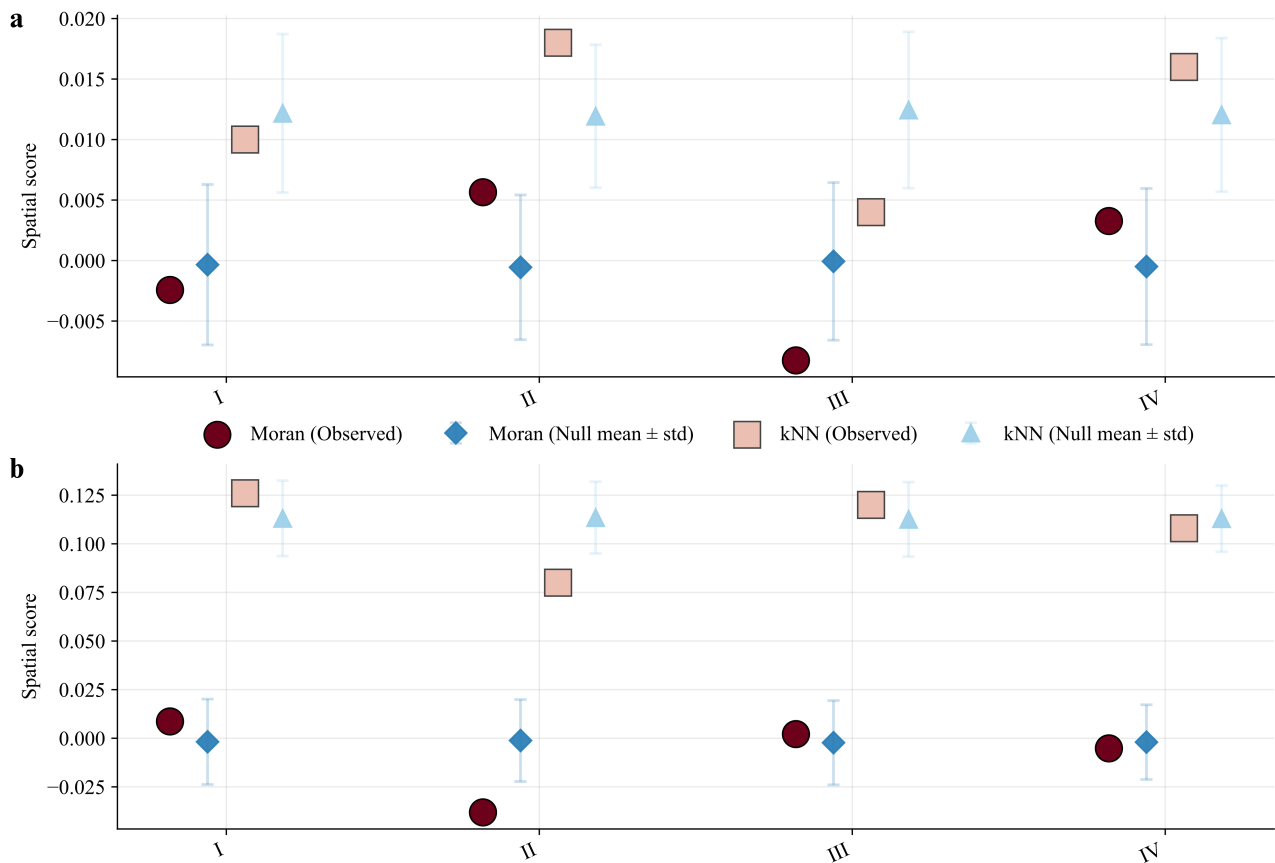


Fig. S4. Stage-specific spatial structure of gene attributions in RareSeq. Spatial statistics of Integrated Gradients (IG) attribution maps for cancer stage prediction (Stages I–IV), assessed using Moran's I and kNN purity. **a**, Observed Moran's I values across stages compared with null distributions, revealing heterogeneous spatial organization of stage-specific gene importance. Stages II and IV show stronger spatial autocorrelation than expected under random layouts. **b**, kNN purity analysis indicates elevated local neighborhood consistency for multiple stages, confirming that genes emphasized for the same stage tend to cluster spatially. These results indicate that Dynomap adapts its spatial organization to disease progression, redistributing predictive gene signal in a stage-dependent manner.

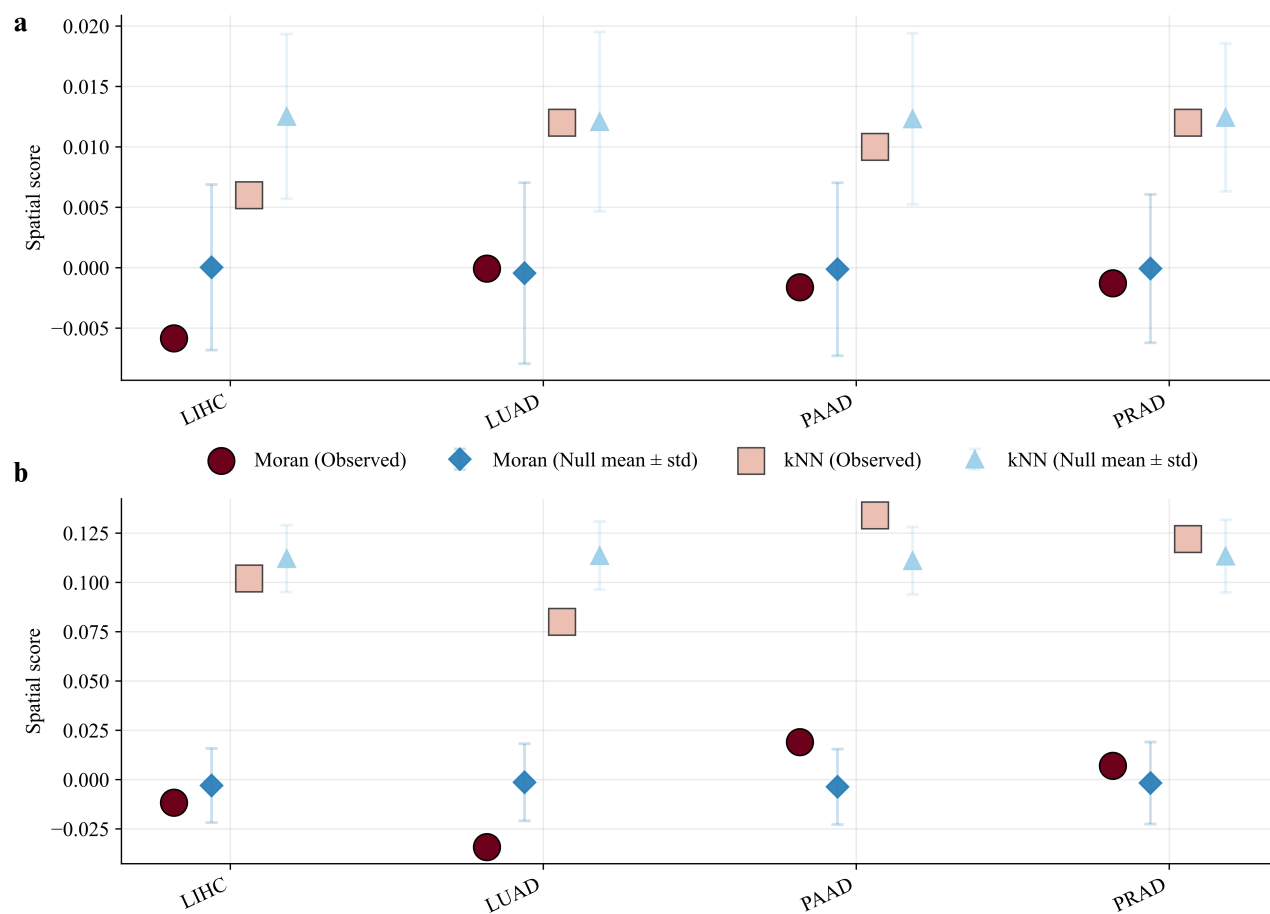


Fig. S5. Subtype-dependent spatial clustering of gene attributions in RareSeq. Quantification of spatial autocorrelation and neighborhood consistency for subtype-specific Integrated Gradients (IG) attribution maps across multiple cancer types. **a**, Moran's I values for liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), pancreatic adenocarcinoma (PAAD), and prostate adenocarcinoma (PRAD), shown alongside null expectations. Subtypes exhibit distinct autocorrelation profiles, reflecting task-specific spatial organization. **b**, kNN purity scores demonstrate consistently elevated neighborhood consistency across subtypes relative to null distributions, indicating robust spatial clustering of subtype-relevant genes. These patterns show that Dynomap induces subtype-specific spatial structure rather than relying on a shared set of globally dominant genes.

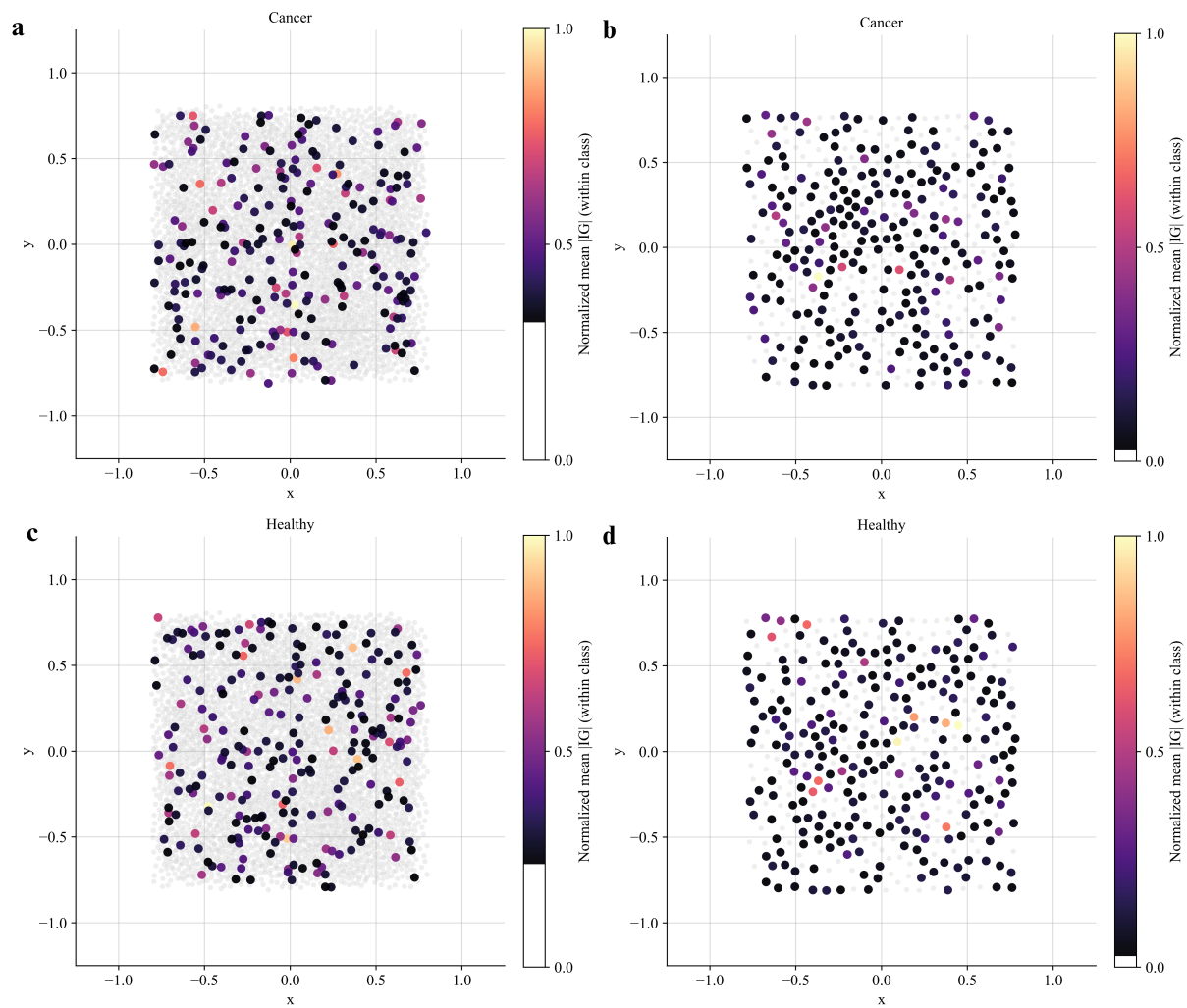


Fig. S6. Class-specific feature attribution projected onto the Dynomap layout for RARE-Seq binary classification. **a,b,** Cancer-class layouts showing gene positions colored by normalized mean Integrated Gradients (IG) values within class. **c,d,** Corresponding healthy-class layouts. Gray points indicate all genes in the layout, while colored points highlight genes with higher class-specific attribution.

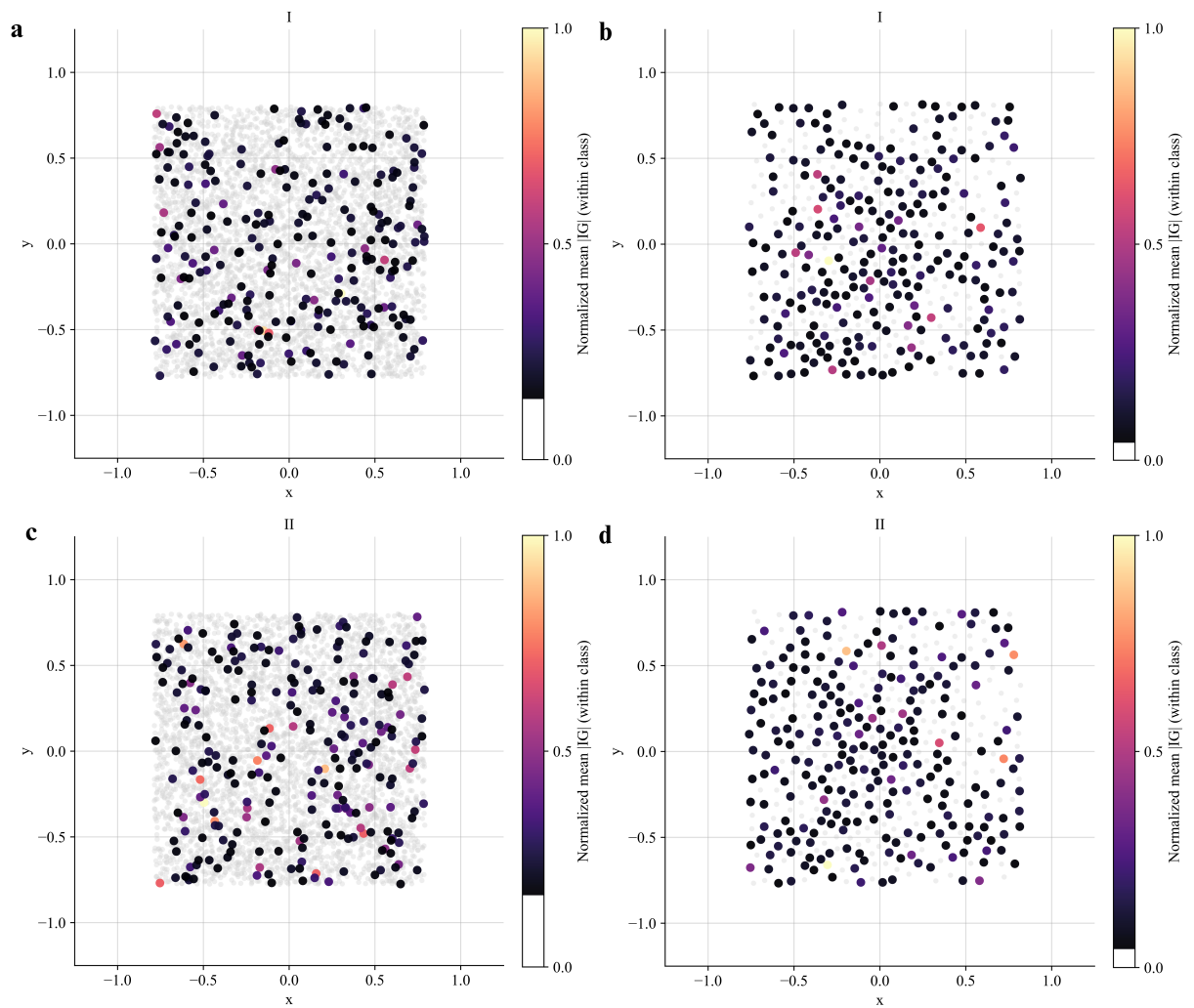


Fig. S7. Stage-specific Dynomap layouts for RARE-Seq stages I and II. **a,b,** Stage I layouts showing IG-weighted gene importance. **c,d,** Stage II layouts. Gene positions are fixed across stages, enabling direct comparison of how discriminative regions shift with disease progression.

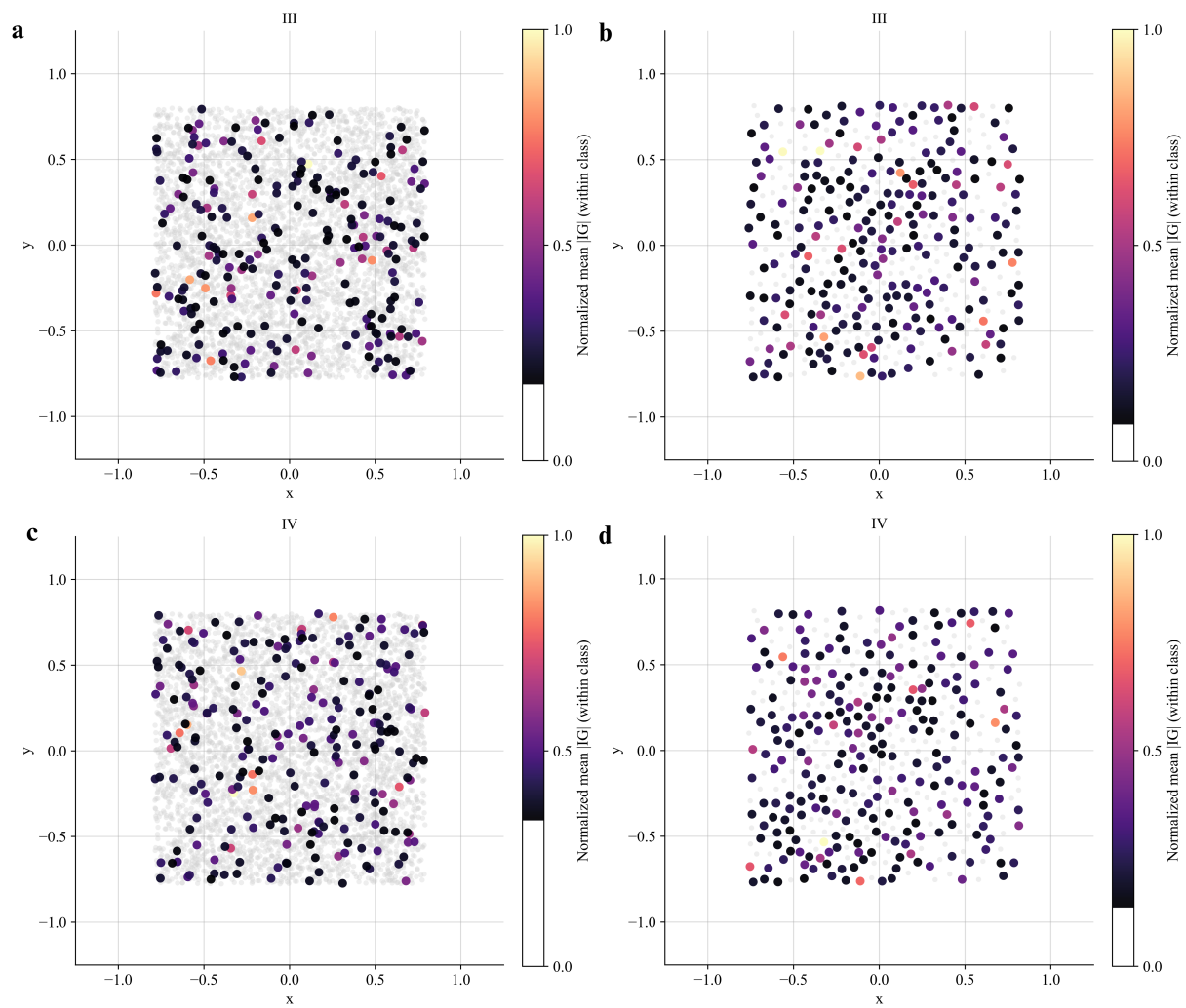


Fig. S8. Stage-specific Dynomap layouts for RARE-Seq stages III and IV. a,b, Stage III layouts colored by normalized IG values. c,d, Stage IV layouts. Later-stage disease is associated with altered spatial emphasis patterns while preserving the underlying gene layout geometry.

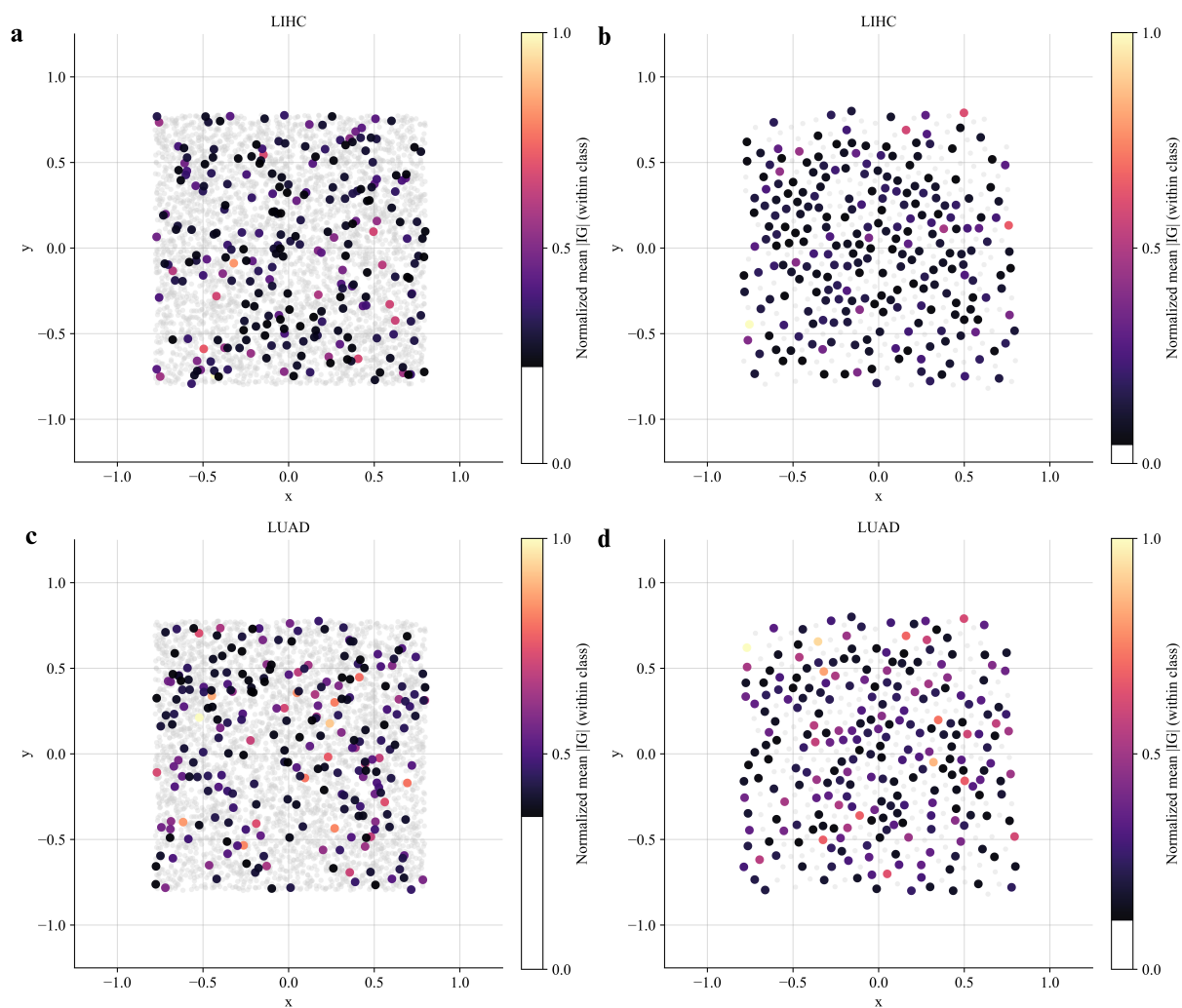


Fig. S9. Subtype-specific Dynomap layouts for RARE-Seq (subset 1). a,b, IG-weighted gene layouts for two molecular subtypes. Colored points represent genes with higher subtype-specific attribution, while gray points denote the full gene set. Distinct spatial emphasis patterns are observed across subtypes within a shared layout.

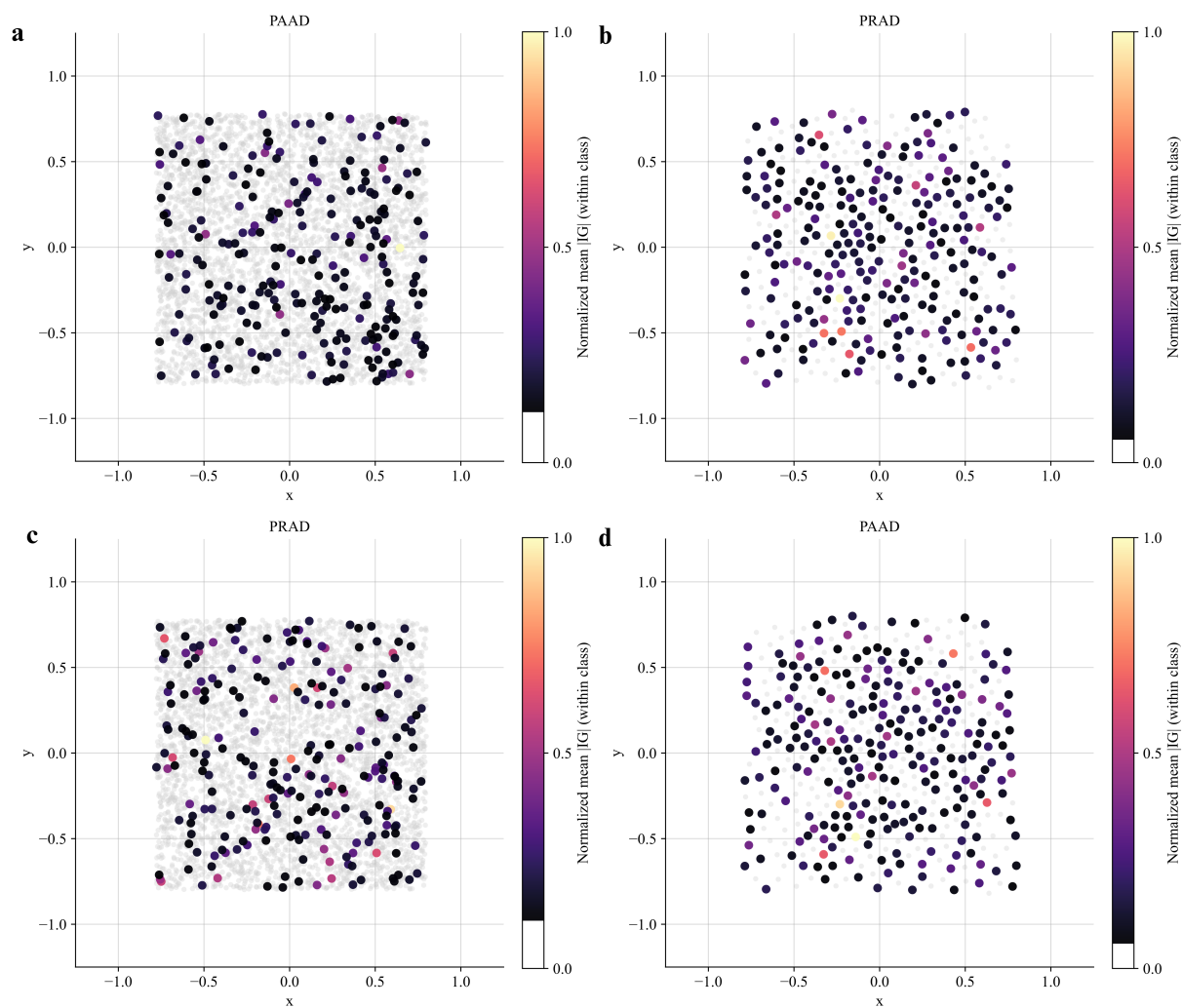


Fig. S10. Subtype-specific Dynomap layouts for RARE-Seq (subset 2). **a,b,** Additional subtype layouts highlighting class-specific attribution patterns. **c,d,** Corresponding complementary subtypes. The reuse of a common spatial scaffold allows subtype distinctions to emerge through localized attribution rather than global rearrangement.

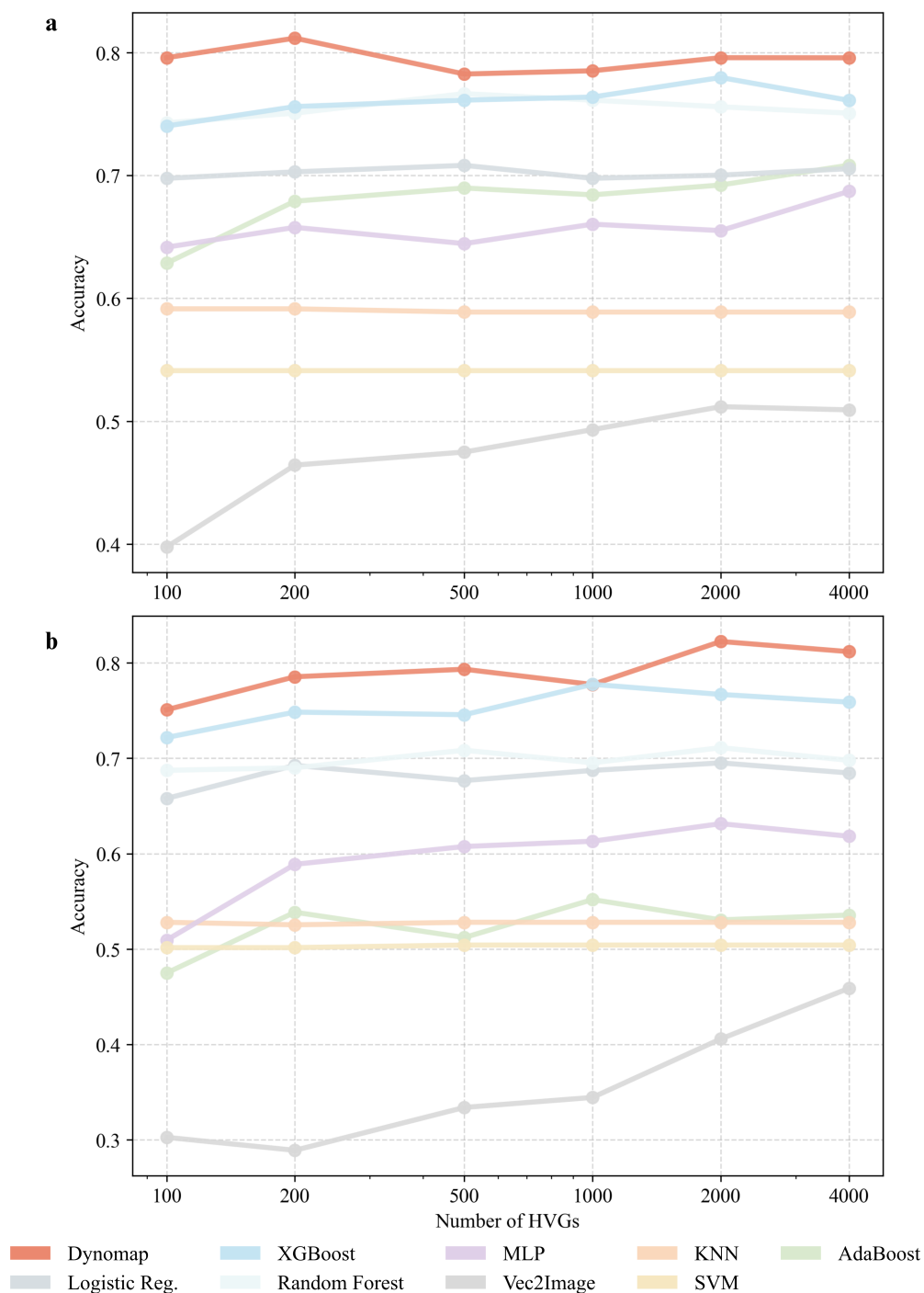


Fig. S11. Effect of highly variable gene (HVG) selection on classification performance in RareSeq. Classification accuracy is shown as a function of the number of HVGs used to construct the input representation. **Top**, Binary cancer versus healthy classification. **Bottom**, Multiclass subtype classification. Dynomap maintains consistently higher accuracy across a wide range of HVG counts compared with baseline methods, demonstrating robustness to feature dimensionality and stability of the learned representation under varying gene selection regimes.

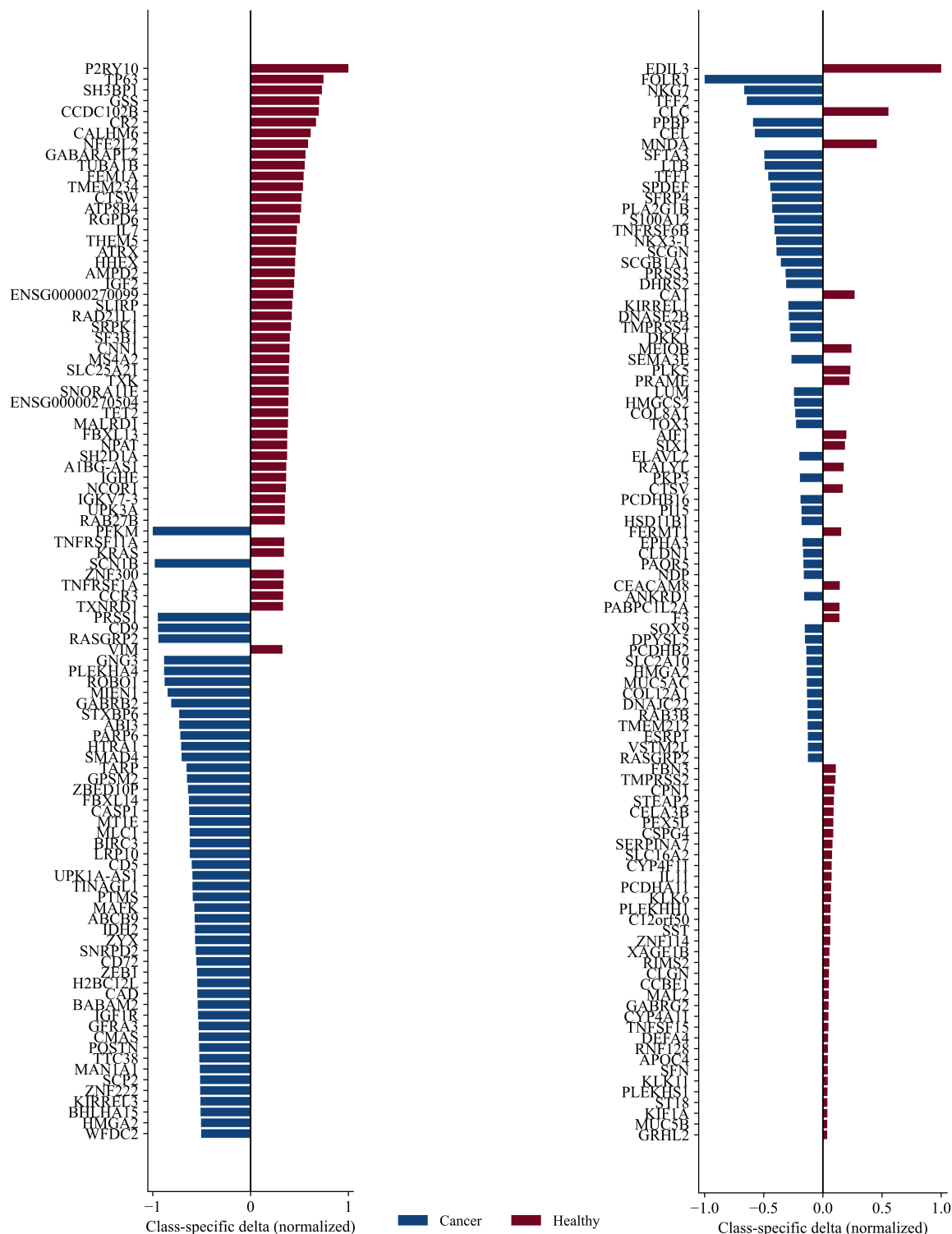


Fig. S12. Differential gene importance for binary healthy vs. cancer detection. Horizontal bars indicate the normalized difference in Integrated Gradients (IG) attribution strength between classes. **a**, Differential attribution using the full gene set shows distinct molecular signatures for healthy (red) and cancer (blue) predictions. **b**, Differential attribution using the 622-gene signature panel reveals a more concentrated set of discriminative features. The shift in specific top-ranked genes across settings demonstrates that Dynomap adapts its feature selection strategy based on the available genomic space.

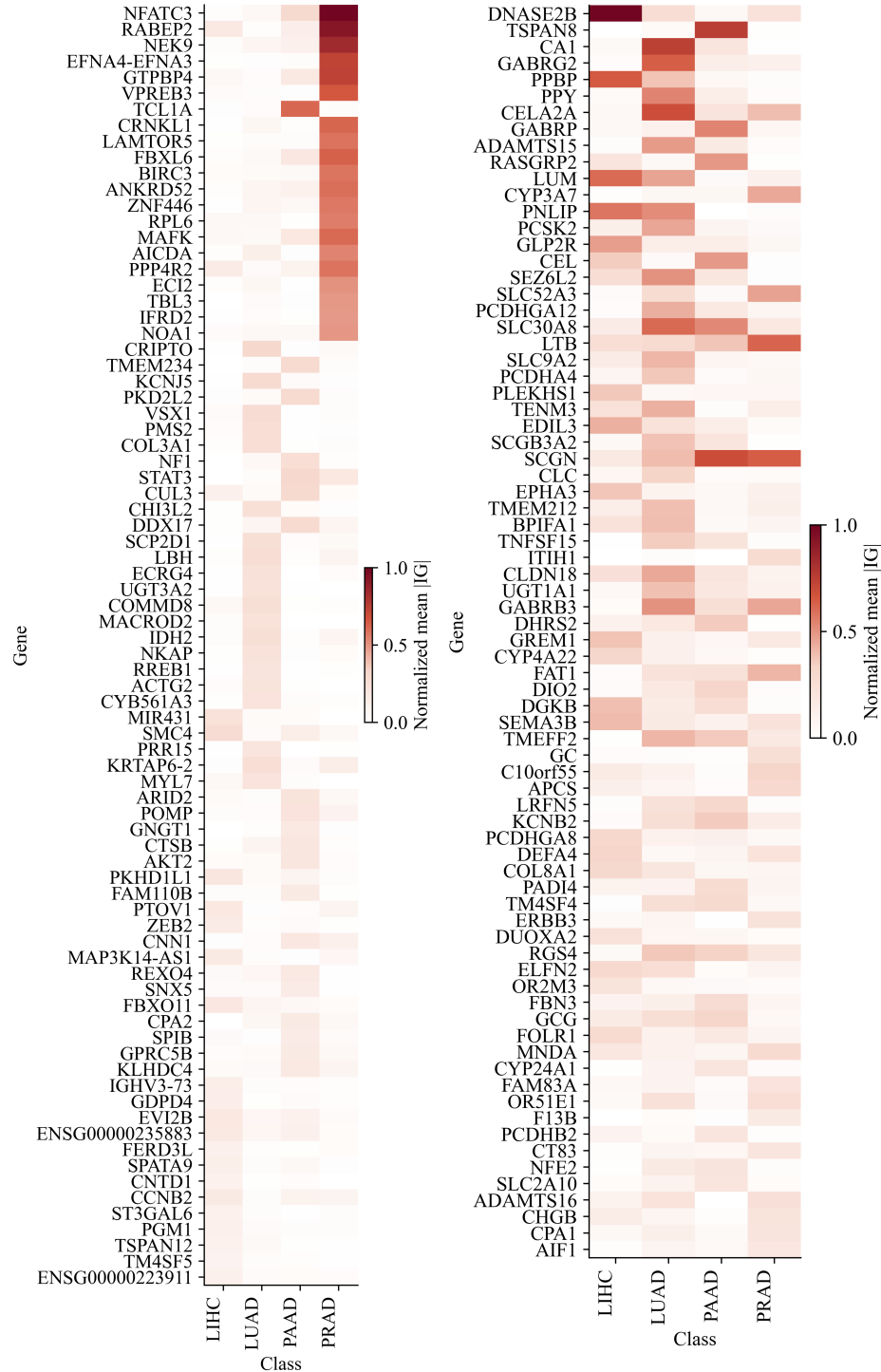


Fig. S13. Subtype-specific gene attribution across cfRNA profiles. Heatmaps of normalized Integrated Gradients (IG) show distinct transcriptional programs utilized for multiclass cancer subtype classification. **a.** Gene utilization for subtype prediction (LIHC, LUAD, PAAD, PRAD) using the full gene set. **b.** Subtype-specific attribution patterns using the 622-gene signature panel. While top predictive genes differ between feature regimes, the model consistently identifies non-overlapping gene sets for each tumor class, supporting the emergence of class-separable spatial representations.

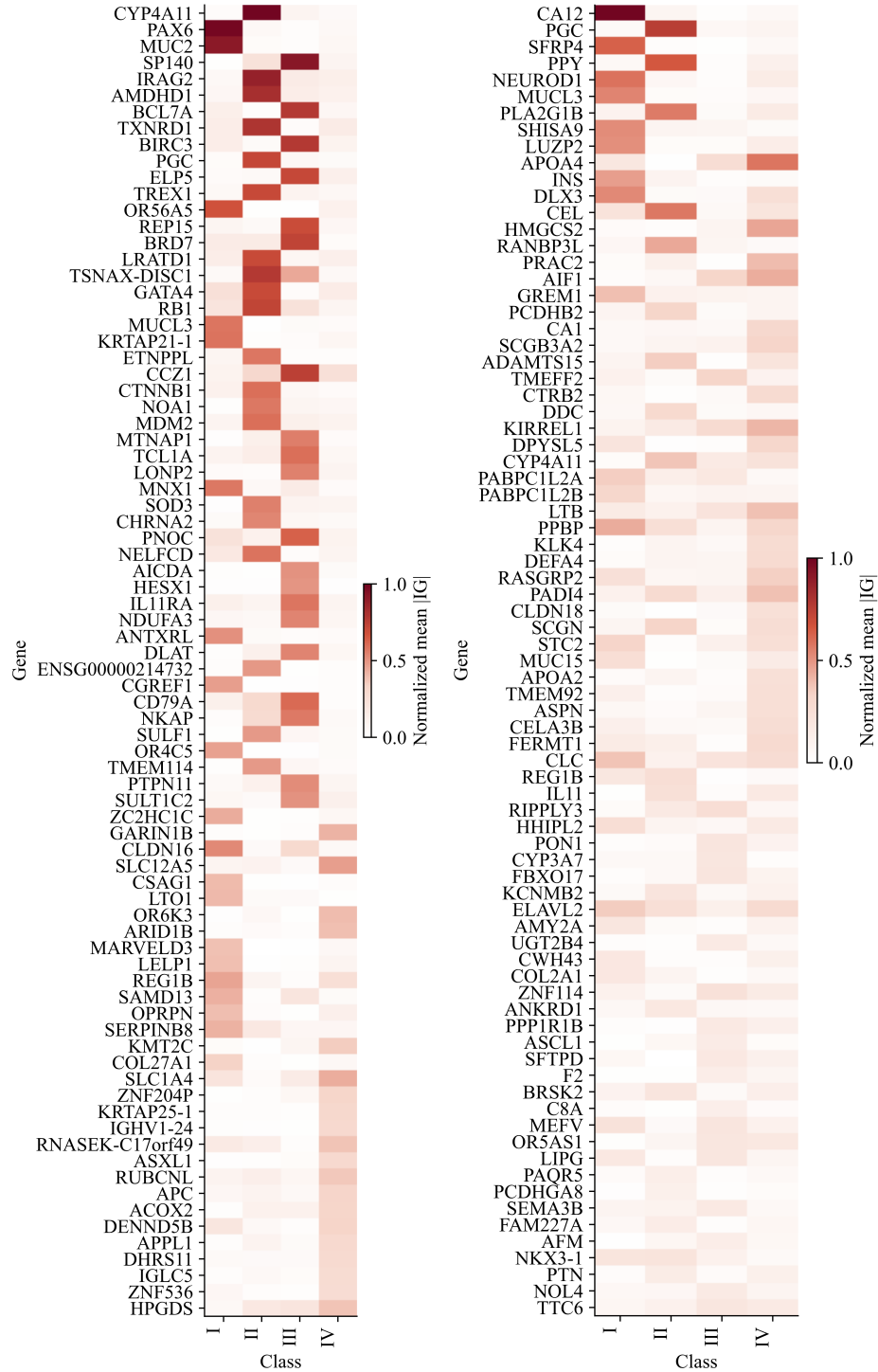


Fig. S14. Gene-level attribution for liquid biopsy cancer stage prediction. Task-adaptive feature importance profiles across clinical stages (I, II, III, and IV). **a.** Normalized mean Integrated Gradients (|IG|) for top predictive genes using the full gene set. **b.** Stage-specific attribution using the 622-gene signature panel. The model identifies a more compact set of stage-informative genes, reflecting a redistribution of predictive signal when the feature space is restricted. The distinct patterns across columns confirm that Dynomap captures subtle signals necessary for fine-grained clinical stratification.

Table S1. Tabula Muris SHAP gene list (part 1 of 3; first 400 genes).

Fcer1g	Ubb	Cd79a	Cd79b	Cd3g	Bst2	Il1b	Vpreb3
Cfh	Cst3	Fxyd5	Tmsb10	Laptn5	Lsp1	Gnas	Gng11
S100a4	Alox5ap	Ly6c1	Plvap	Ctsl	Irf8	H2_K1	Tpm2
Klf2	Lyz2	Sepp1	Ccl3	Vim	Socs3	Lyz1	Igf2bp2
Krt14	Ramp2	Ckb	Oaz1	Gm1821	Marcks1	Arpc2	Arpc3
Hbb_b2	Car2	Scgb1a1	Prss34	Ccl9	Defb1	Tyrbp	Gzma
AW112010	Nkg7	Ifitm1	Tg	H2_D1	Emp3	Hmgb2	Lgals1
Dek	Ly6c2	Coro1a	Slc12a1	Fbln1	Ccl6	Ctsd	Hmgn2
S100a8	Krt19	Shisa5	Rps15a_ps6	Clic1	Rps20	B2m	Cmtm7
Igj	Crip1	Igf2bp7	Ifitm3	Fabp4	Sparcl1	Selk	Cd81
Jund	Capzb	Krt7	Serp1	Cd52	A130040M12Rik	Ctsb	Sln
Myl4	Nppa	Ifitm6	Retnlg	S100a9	x1100001G20Rik	Laptn4a	Rhob
Tmsb4x	Cxcl15	Sftpb	Sftpc	Ly6a	Atf1b1	Ly6e	Ppap2b
Tuba1a	Clec3b	Sod3	Pcolce2	Tnfaip6	Dntt	Ptprcap	Gcc10
Sh3bgrl3	Sh3bgrl3	Slc34a1	Erff1	Mt1	Btg1	Hspa1a	Ccnl1
Meg3	Vcam1	Id3	Ndufa2	x1110003E01Rik	Upk1a	Upk1b	Ifitm2
Cfl1	Rpl37	Gadd45b	Zfp36	Jun	Rps6	Postn	Rpl13
Rps16	Dpt	Ppp1r15a	Rpl7	Srsf5	Npc2	Cryab	Gatm
Pabpc1	Rps26	Atp5k	Cyc1	Tceb2	Sec61g	x2410015M20Rik	Actr3
Ubl5	Mdh2	Vdac1	Atox1	Ap1p2	Cox6c	Lamp1	Cdc42
x1190002H23Rik	Cks2	Ccl7	Cyr61	Aqp3	Perp	Cd9	Atp5d
Tbca	Mif	Pim3	Eif6	Trim29	Jup	Krt6a	Vsig8
Rab25	Fabp5	Ly6d	Rtn4	Nfkbia	D14Ert449e	Rpl9	Rpl27a
Rps10	Eif4a1	Lgals3	x2010001M09Rik	Chchd10	Ndufc2	Psap	Ier2
Sepw1	Psm8	Rhoa	Gabarap	Arpc5	Tmed10	Mal	Tpt1
Lum	Csrp1	H3f3b	Cd24a	S100a11	Cotl1	Taldo1	Ndufa7
Actb	Rpl37a	Rplp2	Serinc3	Tmem176a	Gdi2	H2_Aa	Rps15a_ps4
Btg2	Ccl17	Tmem176b	Atf4	Fosb	Gadd45g	Rnase4	Cbr2
Ptpn18	Rps2	Rps11	Rps17	Atp5b	Rpl8	Rplp1	Txn1
Rpl18a	Rps8	Sat1	Ndufb10	Chchd2	Ccdc72	Cox8a	Eif4g2
Atp5a1	Btf3	Ndufv3	Myl12a	Rnasek	Foxq1	Ivl	S100a6
Wfdc2	Rplp0	Pdlim1	Rpl19	Sdc1	Igf2bp3	Gltscr2	Krt5
Hspa8	Rpl15	S100a10	Mrfap1	Pkm2	Psm3	Sfn	Csnk1a1
Rbm3	Rps5	Txndc17	Cldn5	H2_Ab1	Cd74	H2_Eb1	Lmna
Cd164	Krt8	H3f3a	Romo1	Hmgn1	Rps28	H2afv	Mdh1
Eif5a	Rpl23	Rps24	Aldh2	Rpl22l1	Tomm6	Fos	Gpx1
Arpc1b	Shfm1	Klf9	Cox7a2l	Rpl27	Rpl18	x1500012F01Rik	Eef1a1
Atf3	Cct4	Adamts1	Pa2g4	Atp6v0e	Timp3	Rpl32	Rpl36al
Rps15	Rpl41	Rps27	Hnrnpk	Ran	Pcbp1	Usmg5	Cycs
Cnbp	Eif2s2	Map1lc3a	Myl6	Cyba	Sec11c	Hint1	Birc5
Anp32e	Cenpa	Eif1	Zfp706	Atp1a1	Bag1	Rpl5	Rpl35
Ifi27l2a	Atp6v0b	Slc38a2	Ier5	Pnrc1	Tpm4	Vtn	Apoa1
Akr1c6	Gnmt	Srp9	Nfe2l2	x2310003F16Rik	Gsr	Hp	Cyp2f2
S100a1	x1810037l17Rik	Rac2	Brp44	Nr4a1	Egr1	Arf1	Ndufb4
Hist1h2bc	Hras1	Sepx1	Apoe	lfrd1	Prdx5	Umod	Ctla
Hnrnpf	Rpl4	Rps27a	Use1	Tmed2	Bsg	Tomm7	Tpm3
Uqcrc1	Ctla2a	Gmfg	Gas6	Rpl35a	Pcbp2	Sgk1	Rps9
Psca	x2700060E02Rik	Rps14	Gnb2l1	Rpl29	Eef2	Lsm4	Dusp1
Junb	Nme2	Rpl11	C3	Igf2bp4	Tmem123	Bgn	Wif1
Myl9	Elane	Ms4a3	Anxa1	Glrx	H2afx	Mki67	Pfn1

Table S2. Tabula Muris SHAP gene list (part 2 of 3; next 400 genes).

Park7	Rpl17	Sub1	Ube2i	Erp29	Mcpt8	Icam1	Tsc22d1
Atp5e	Fxyd3	Rps18	Ndufb9	x3110003A17Rik	Dmkn	Uqcrfs1	Rps19
Rps25	Ltb	Hnrnpab	Plac9	Dnaja1	Itm2b	Id1	Ier3
Rarres2	Csrp2	Mmp3	Calm1	Gnai2	Pfdn5	Sumo2	Ctnnb1
Atp5j	Cox4i1	Rpsa	Slc25a4	Cct7	Ddx5	Atp5h	Rpl23a
Snrpd2	Tagln	Chmp2a	Ndufa4	Ccl5	Nbl1	Rps21	Uqcrb
Rpl10	Arpc1a	Rpl10a	Akr1a1	Etfb	Pomp	Srgn	Ubc
Ftl1	Rpl28	Rps27l	Stmn1	Tubb5	H2afy	Hist1h2ae	Ssr4
Ctsq	Top2a	Aes	Rpl21	Rpl3	Reep5	Cct5	Tmbim6
Rps4x	Ywhaq	Atp6v1g1	Arl6ip1	Aqp1	Ldhd	Cd36	Cxcl12
Cd14	C1qa	C1qb	Ndufa8	Edf1	Hprt	Fkbp1a	Top1
Mfap5	Mmp2	Ccl11	Pi16	Col6a1	Ugdh	Dcn	Serpinf1
Mt1	Myoc	Smoc2	Pcolce	Cebpd	Serping1	Cxcl1	Fgl2
Zfp36l1	Rps29	Tmem134	D8ErtD738e	Slc25a5	Smc4	x2410006H16Rik	Chmp4b
Hk2	Ahnak	Sdc4	Atp5o	Dynlrb1	Dad1	Aldoa	Map1lc3b
Hdgf	Uqcrh	Cox6b1	Nme1	Rpl7a	Rps12	Rps7	Nhp2l1
Polr1d	Cdkn1a	Ywhaz	Rps3	Maff	Ndufb8	Klf6	Mgst3
Rps23	Gnb2	Hspe1	Rpl12	Pgk1	Crlf1	Hist1h1c	Klf5
Phb2	Arf5	Avp1	Pebp1	Psmb5	Crip2	Ptma	Prdx1
Psmb6	Snrpb	Rpl6	Uba52	Cox6a1	Nap1l1	Hnrnpa2b1	Ndufa12
Rps3a	Zfand5	Eef1b2	Eif3h	x1110008F13Rik	Vamp8	Xist	Rpl31
H2afj	Tst	Dstn	Dbi	Slc25a3	Pycard	Rnh1	Pir
Ppp1r14b	Serbp1	x2610528A11Rik	Rpl14	Arf6	Dnajb1	Cstb	Emp2
Rpl22	Fosl1	Pkp1	Ovol1	Hbegf	Nhp2	Ndufa4l2	Capg
Calml3	Them5	Capns2	Gpx2	S100a14	Ptgr1	Lypd3	Krt16
x2310033E01Rik	Slpi	Spink5	Actg1	Eef1d	Nedd8	Ndufa3	Plac8
Fau	Ndufb11	Sqstm1	Acta2	Atp5l	Pdia3	Spint2	Gsto1
Ost4	Rbx1	Ppp1ca	Rps13	Polr2l	Actn4	Mgp	Swi5
Hsp90aa1	Snrpg	Limd2	Pkp3	Srsf3	Slc3a2	Arpp19	Tuba1b
Hsp90b1	Sec61b	Glxr3	Rac1	Ccnd3	Ltf	Ngp	Ap3s1
Tkt	Gpi1	Chi3l3	Camp	Fcnb	Atpif1	Minos1	Ndufc1
Prdx2	Gm11428	Hmox1	Atp5f1	Ccl4	Napsa	Ctss	x1500032L24Rik
Alas2	Pglyrp1	Eif3i	Nop10	Clic4	Lgals7	Metap2	Ppia
Eif3m	Tubb4b	Atp5g3	Gja1	Ezr	Ldha	Hspd1	Anxa8
Gm94	Ube2d3	Uqcrq	Thy1	Tm4sf1	Glul	F3	Krt18
Snrpe	Anp32b	H2afz	Cltb	Gnb1	Fth1	Lpl	S100a16
Cyb5	Odc1	Tacstd2	Psma7	Rpl36a	Timp2	Mrpl52	Serf2
Srsf2	Higd1a	Ndufb5	Tuba4a	Myeov2	Cd8a	Phlda1	Cd8b1
Rpl31_ps12	Skp1a	Tmem59	Blvrb	Uqcr10	Psma2	Rps15a	Hspa5
Pcna	Oat	Ncl	Ybx1	Gpx3	x2200002D01Rik	Gsta4	Hes1
Sparc	Cpe	Fmod	Eif3k	Chad	Prg4	Srsf7	Pdcd4
Serpine2	Gcat	Ppp1r2	Rpl39	Cox5a	Tpi1	Rab10	Chit1
Ndufs6	Uqcr11	Ndufa6	Ucp2	Ndufa13	Pgam1	Id2	Atp5g2
Rcan1	Arhgdib	Hnrnpu	x2810417H13Rik	Krtcap2	Hist1h1e	Rrm2	Manf
Timm13	Tspo	Naca	Fam162a	Sfr1	Nrp1	Acsm2	Apob
Hpx	Rabac1	Cdo1	Atp6v1f	Acaa1b	Fxyd2	Akr1c21	Slc27a2
Hpd	Apoh	Rgn	Serpina1b	Fabp1	Mup2	Ndrgr1	Upk3a
Gclc	Col3a1	Dynl1	Hmgb1	Gng5	Ralbp1	Ucma	Col11a1
Comp	Pdzk1ip1	Ndufb7	Calm4	Anxa5	Apoc4	Gm13889	Krt15
Eif5	Glycam1	Ube2c	Psma3	Rpl24	Rpl38	Eef1g	Gapdh

Table S3. Tabula Muris SHAP gene list (part 3 of 3; remaining 288 genes).

Acp5	Rap1b	Hspb8	Col2a1	Il6	Tomm20	Dsp	Klf4
Malat1	Epsti1	Ccl2	Has1	Neat1	Prelid1	Calr	Hspb1
P4hb	Ppib	Fdps	Mup3	Cxcl14	Cxcl2	x2010107E04Rik	Hrsp12
Miox	Cat	Col8a1	Cxcl10	Mgst1	C1qc	Rpl13a	Mpo
Lcn2	Prtm3	Eif3e	Gas5	Impdh2	Npm1	Tagln2	Cox7b
Hsp90ab1	Tcp1	Procr	Psmb1	Trf	Emp1	x1500015O10Rik	Hist1h2ao
Myl7	Tpm1	Rab11a	Akr1b8	Rpl36	Scgb3a2	Sbpl	Scgb3a1
Tff2	Ndufa1	Serpinb5	Eif3f	Prdx6	Areg	Cldn4	Hnrnpa3
Eln	Tnfrsf12a	Tmem27	Acy3	Aldob	Cyp4b1	Guca2b	Ranbp1
Anxa2	Klk1	Tppp3	Psmb2	Calm2	Ttc36	Tcn2	Tecr
Hilpda	Gstm1	Ttr	Cyp2e1	Scd1	Esd	Sep15	Egf
Wfdc15b	Sprr1a	x1110032A04Rik	Bpifa1	Mia1	Sftpa1	Fgg	Gltp
Aldh3a1	Dsg1a	Timp1	Ccl19	Fgb	Sult2b1	Cd63	Homer2
Rbp4	Fmo5	Sbsn	Snprf	Cox5b	Scp2	Atp5j2	Cox7c
Cma1	Retnla	Ctnnbip1	Erh	Serpinh1	Fn1	Ywhae	Set
Apoc3	Fga	Cyp3a11	Mup20	Spink3	Mrpl33	Apoa2	Alb
Serpina1d	Hn1	Thbs1	Fam25c	Gjb2	Barx2	Krt24	Gem
Ube2s	Spp1	Gc	Krtdap	Car3	Gsta3	Ahsg	Krt10
Serpina1c	Serpina3k	Apoc1	Kdm6b	Dut	Gpx4	Tubb2a	Csf3
Epcam	Expi	Clu	S100g	Tmem14c	Atp5c1	Col1a2	Fgfbp1
Sprr1b	Cnfn	Selenbp1	Dapl1	Cst6	Krt13	Krt4	Ccl20
Krt6b	x2310002L13Rik	Ly6g6c	Krt23	Tgm3	Atp5g1	Col1a1	Nupr1
Lmo1	Tesc	Inmt	Psapl1	Efh2d2	Apod	Calb1	Col10a1
Krt17	Chi3l1	Serpinb2	Mfge8	Calca	Reg3g	Selp	Krt35
Igfbp5	Serpinb3a	Elovl4	x2310002J15Rik	Cldn7	Ctsk	Prr9	Msmg
x2310079G19Rik	Lce3c	Serpina1a	Rn45s	Serpinb12	Defb4	Rbp2	Klk14
Ptx3	Adh7	Gstp1	Sox9	Csn3	Lrrc15	Kap	x8430408G22Rik
Atp2a2	Myh6	Ttn	Ada	Lor	Krt76	Slurp1	Crcr1
Darc	Krt36	Acan	Cytl1	Basp1	Krt84	Tchh	Chi3l4
Lipf	x3110079O15Rik	G0s2	Bpifb1	Slc14a2	Krt85	Sprr2i	Otor
Dcpp2	Krt33b	Dcpp3	Krtap3_3	Mfap4	x1600029D21Rik	Car1	Aqp2
Lyg1	Serpina1e	Krtap13_1	Serpinb3c	Pf4	Pbbp	Sprr3	Ppa1
Krt34	Krt81	Bhmt	Krt31	Dcpp1	Krtap3_1	Krt42	x5430421N21Rik
Krtap13	Krtap9_3	Psors1c2	Krt33a	Krt86	Krtap3_2	x2310034C09Rik	x2310061N02Rik
Cyp2c69	Hamp	Klk10	Ptms	Cox7a2	Mt4	Pam	Pgls
Aldh1a1	Capns1	Sod1	Cebpb	Fstl1	Gsn	Ifi205	Igfbp6

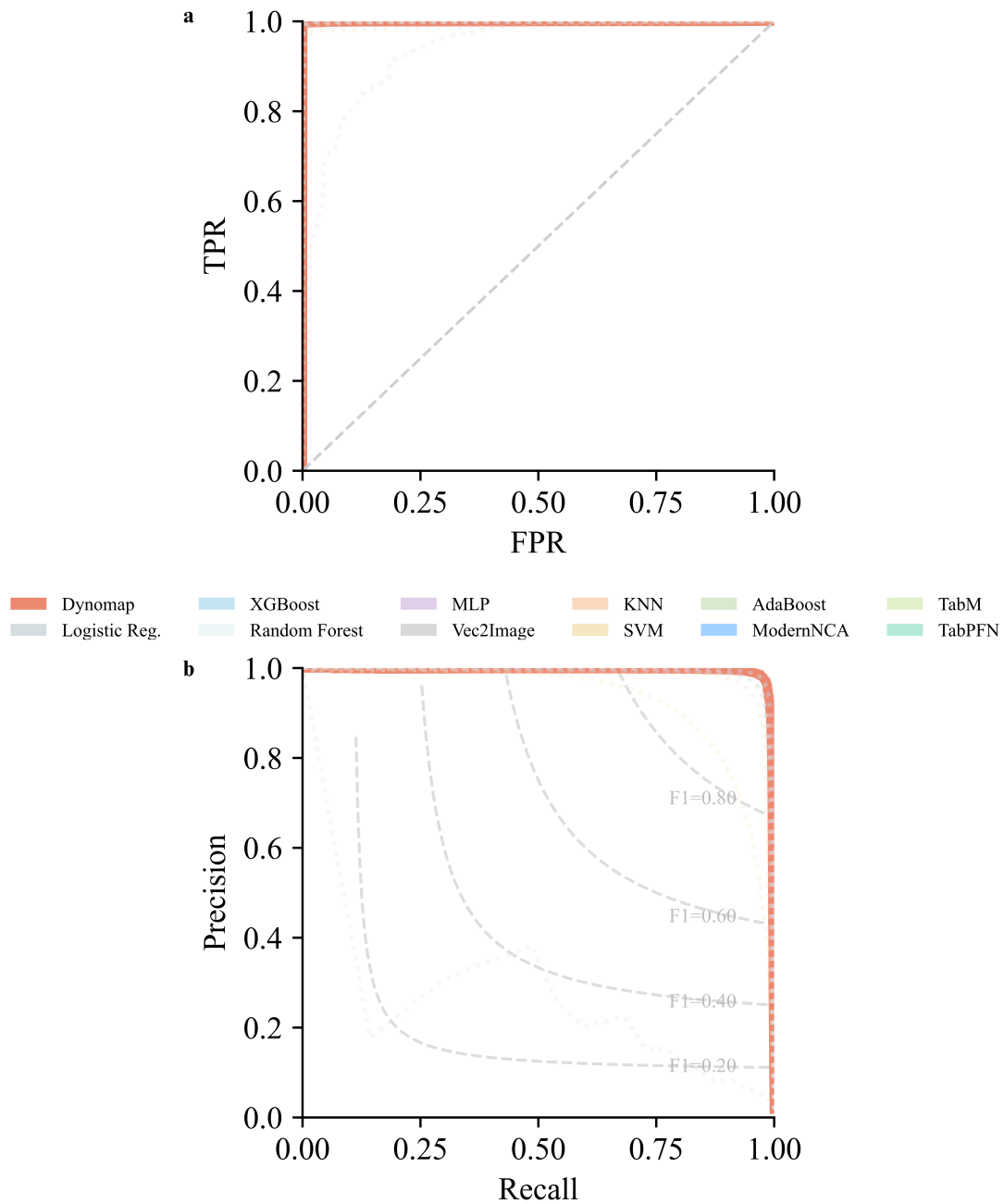


Fig. S15. Precision-Recall and ROC analysis for Tabula Muris cell type classification. **a**, Receiver Operating Characteristic (ROC) curve showing the True Positive Rate (TPR) against the False Positive Rate (FPR). Dynomap (solid orange line) achieves near-perfect area under the curve (AUC), significantly outperforming baseline methods. **b**, Precision-Recall curves across various cell types. Dynomap maintains high precision across the entire recall range, demonstrating robust classification even for rare or heterogeneous cell populations.

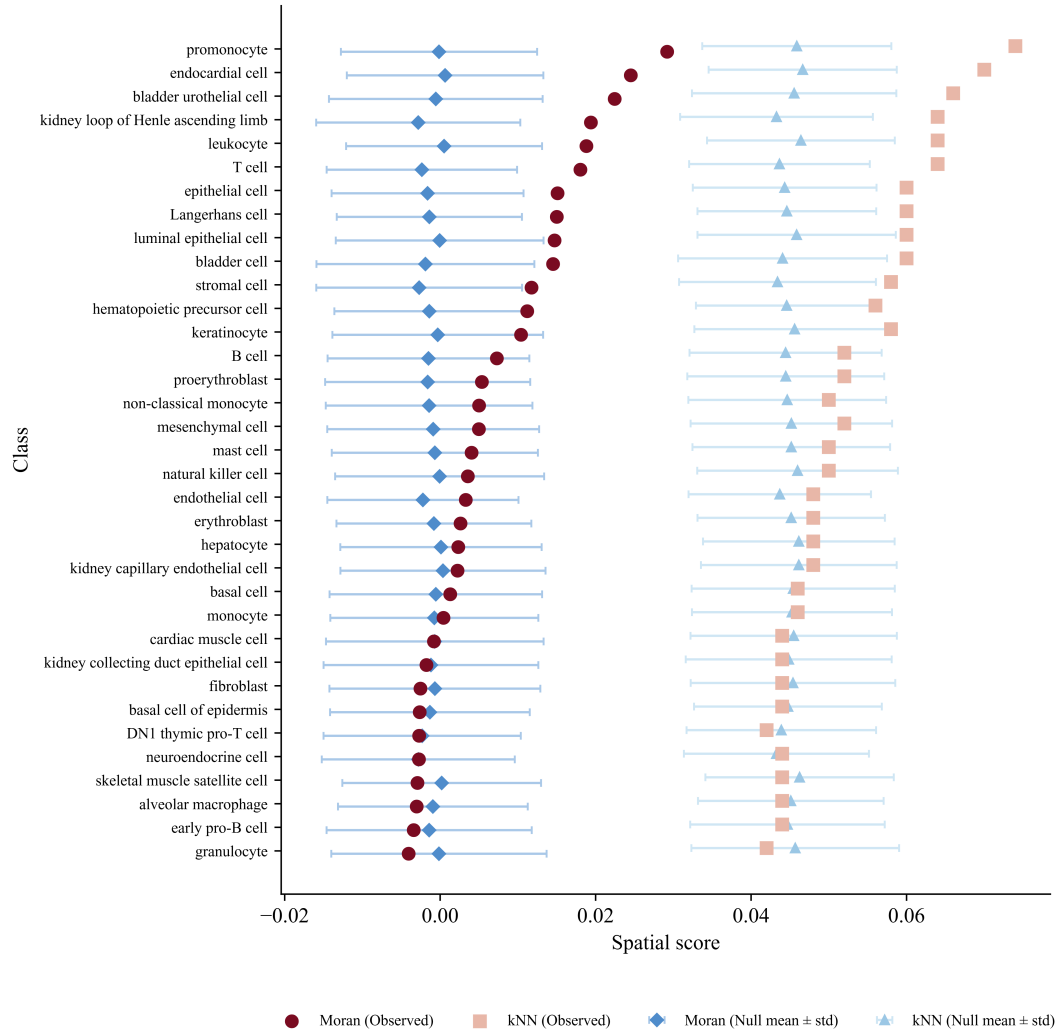


Fig. S16. Quantification of spatial organization across cell lineages in the Tabula Muris atlas. Observed Moran's I (spatial autocorrelation, dark red circles) and kNN purity (neighborhood consistency, light orange squares) are compared against null distributions (blue diamonds and triangles with error bars). For nearly all cell types, including immune and stromal populations, the observed scores significantly exceed null expectations. This confirms that cell-type-specific transcriptional programs are organized into coherent spatial neighborhoods within the learned Dynamap layout.

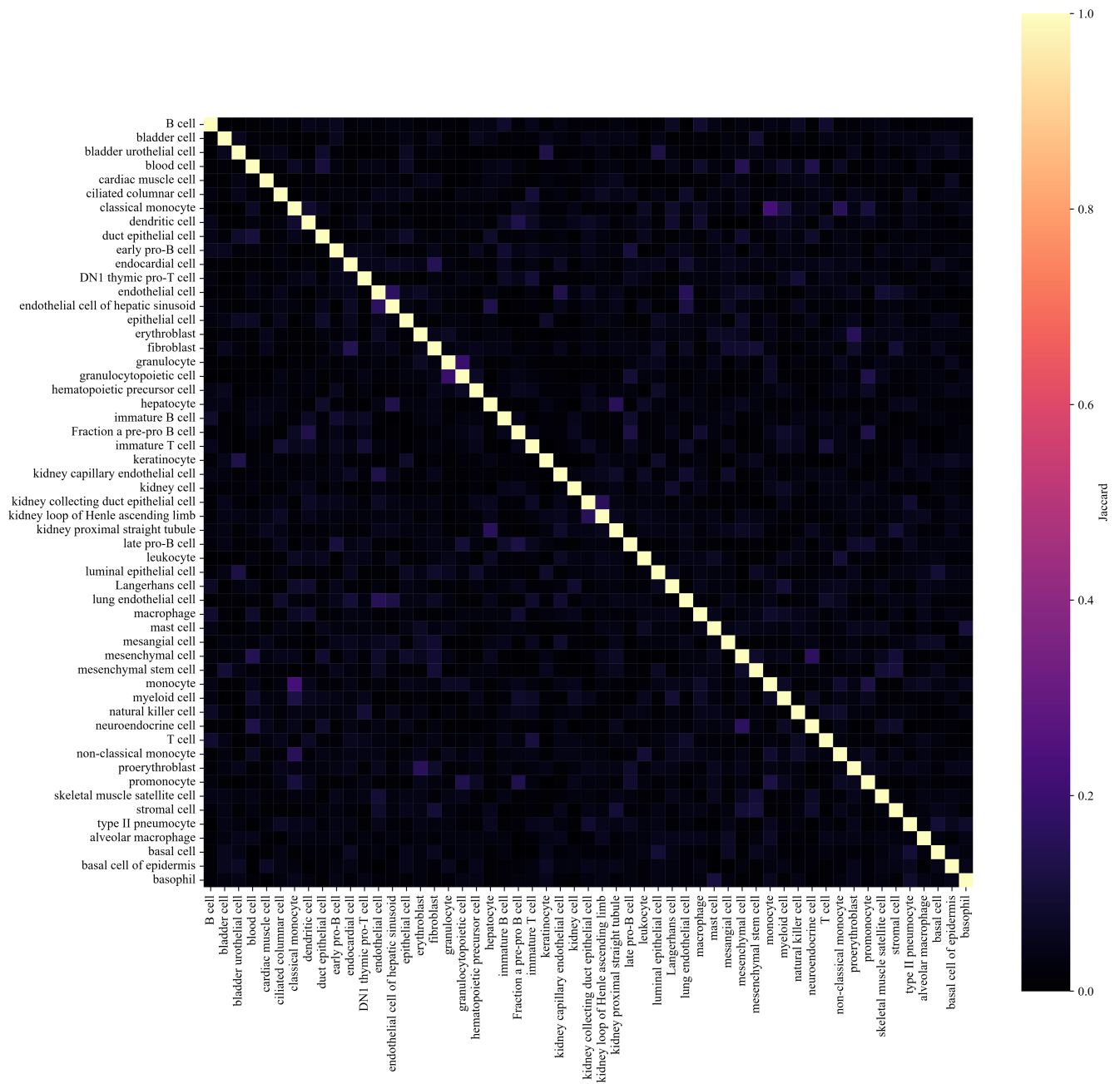


Fig. S17. Jaccard similarity matrix of learned spatial neighborhoods for cell types. Heatmap values represent the overlap between the top predictive feature sets of different cell types. The strong diagonal indicates highly stable and distinct spatial neighborhoods for individual classes. Sparse off-diagonal signals reveal biologically expected similarities between related cell lineages, such as different subtypes of immune or endothelial cells.

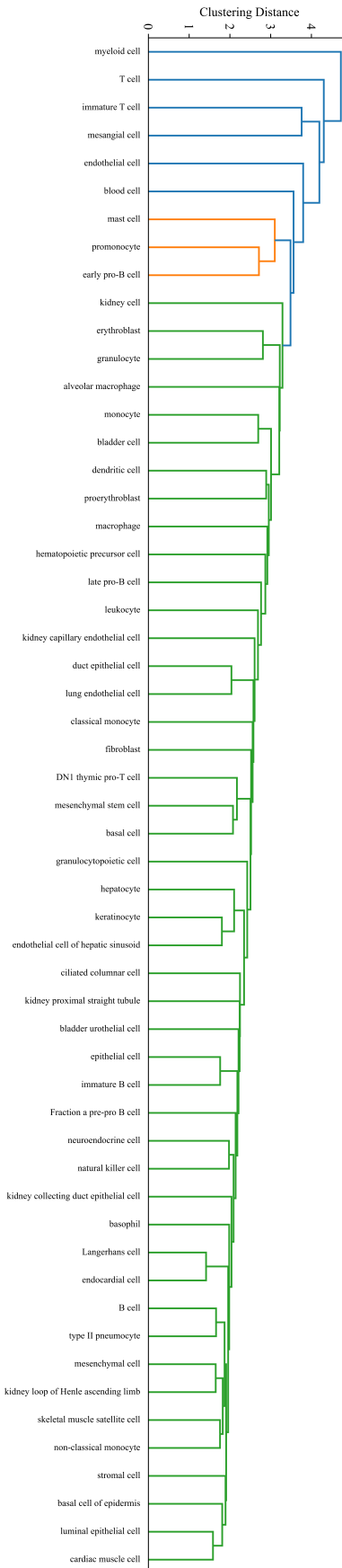


Fig. S18. Hierarchical clustering of cell type attribution profiles. This dendrogram illustrates the relationships between cell types based on their Integrated Gradients (IG) feature importance. Closely related cell types, such as various B cell stages or lung-specific populations, cluster together. This demonstrates that the Dynomap framework captures a hierarchical representation of biological identity directly from single-cell transcriptomic data.

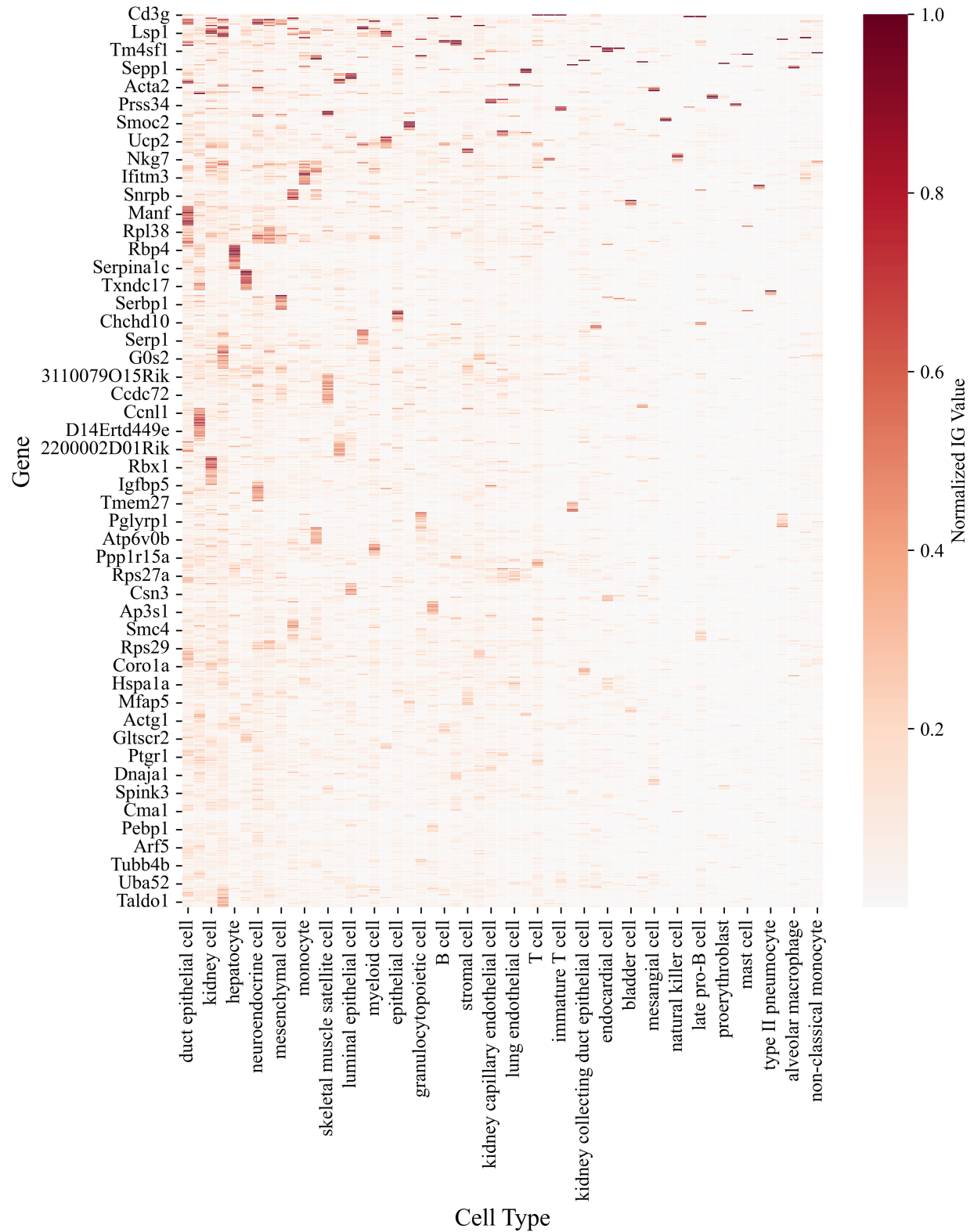


Fig. S20. Comprehensive feature attribution heatmap for single-cell lineages. The heatmap displays normalized Integrated Gradients (IG) values for top predictive genes (rows) across dozens of cell types (columns). Distinct vertical bands highlight marker genes uniquely utilized by the model for specific lineages, such as *Cd3g* for T cells and *Hepatocyte*-specific markers. The sparse structure confirms that the model relies on specialized transcriptional signatures to distinguish between complex cellular identities.

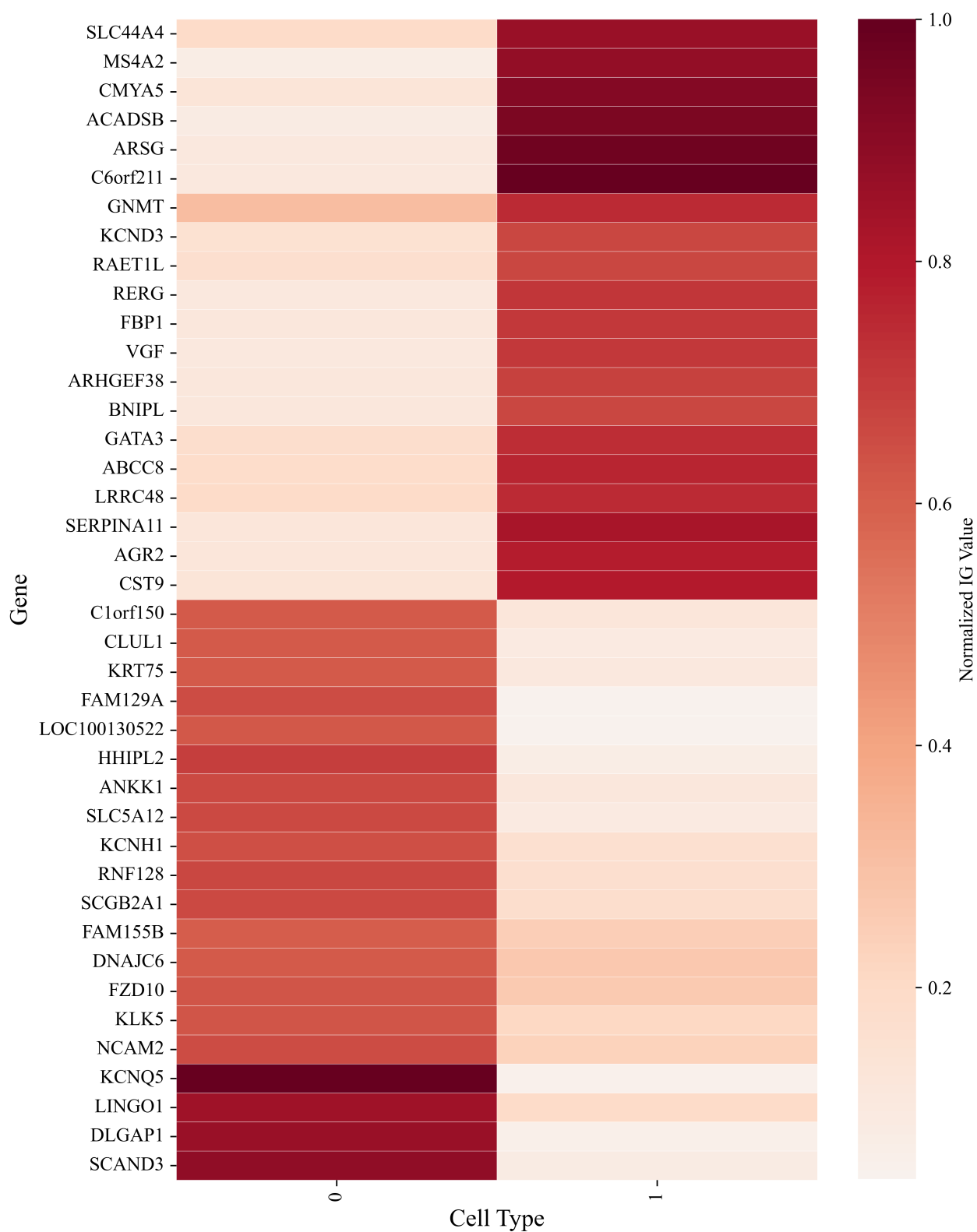


Fig. S21. Class-specific gene attribution for breast cancer risk stratification. Heatmap of normalized Integrated Gradients (IG) for the top 50 predictive genes in the binary task (Aggressive vs. Non-aggressive). The model utilizes distinct sets of genes to identify aggressive phenotypes, including *C6orf211* and *ARSG*, while non-aggressive tumors are characterized by a different molecular signature featuring *KCNQ5* and *DLGAP1*. The clear vertical separation of high-attribution scores confirms that Dynomap learns non-overlapping transcriptional programs for clinical risk assessment.

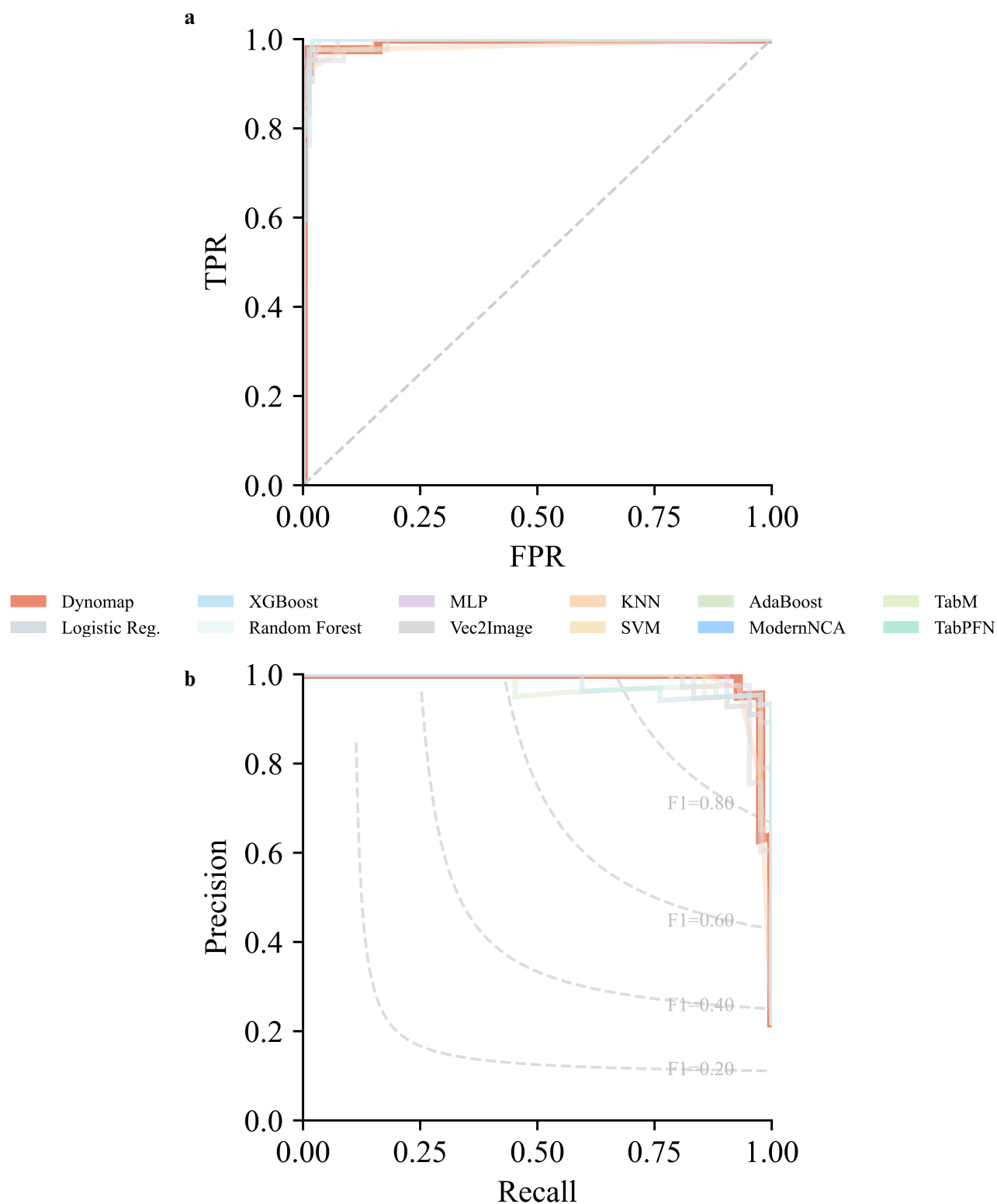


Fig. S22. Predictive performance for binary risk stratification in TCGA-BRCA. **a**, Receiver Operating Characteristic (ROC) curve and **b**, Precision-Recall (PR) curve for the Aggressive vs. Non-aggressive classification task. Dynomap (solid line) achieves high area under the curve (AUC) values, demonstrating robust discrimination even in the presence of tumor heterogeneity. The high precision maintained across recall levels indicates that the model's risk assessments are highly reliable for aggressive cases.

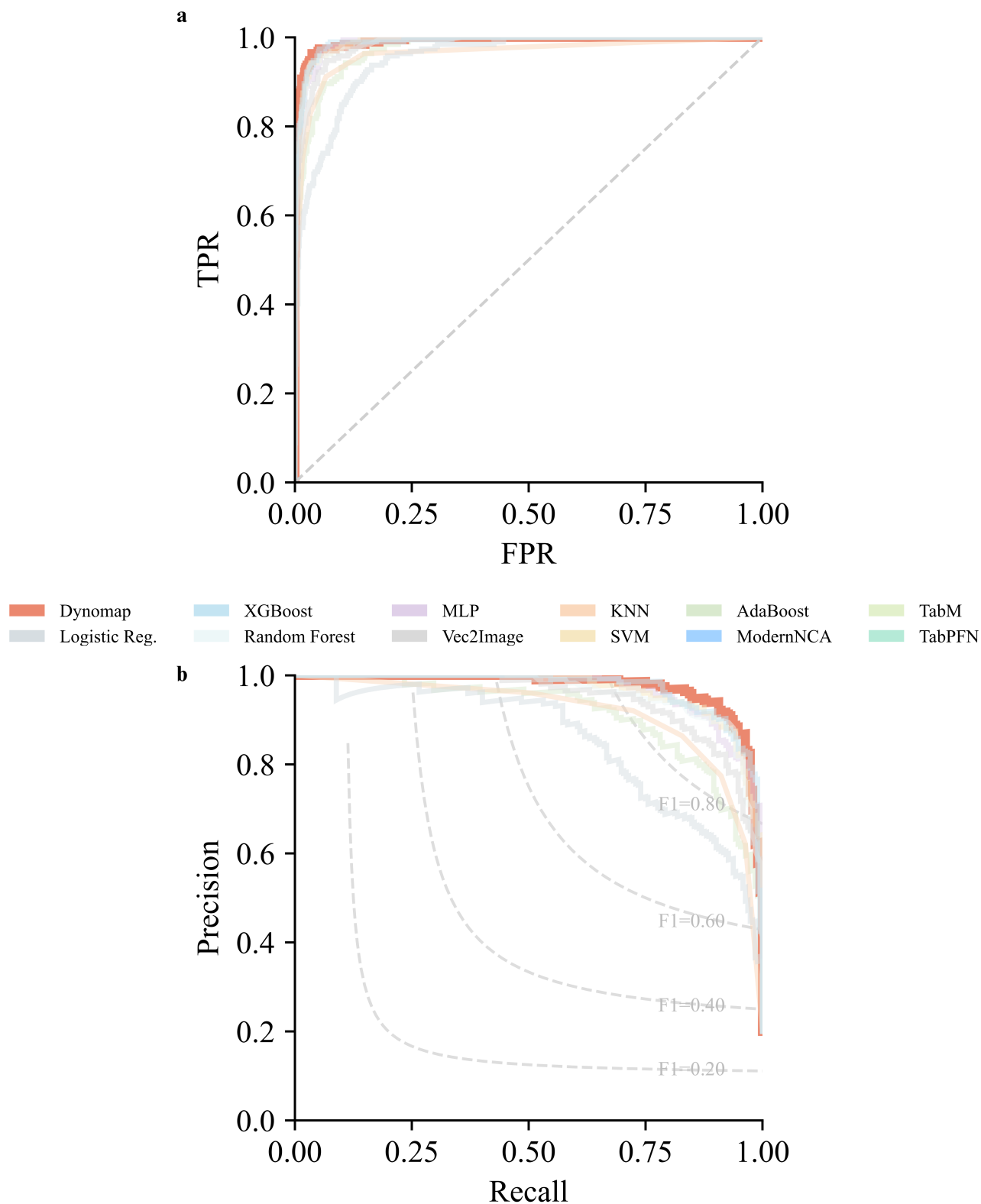


Fig. S23. Multiclass performance for breast cancer molecular subtype prediction. **a**, ROC and **b**, Precision-Recall curves for five intrinsic subtypes (LumA, LumB, Basal, HER2, and Normal-like). Dynomap maintains superior precision-recall trade-offs across all subtypes compared to baseline averages. Performance is particularly strong for the Basal and HER2-enriched subtypes, which exhibit the most distinct molecular signatures in the learned spatial representation.

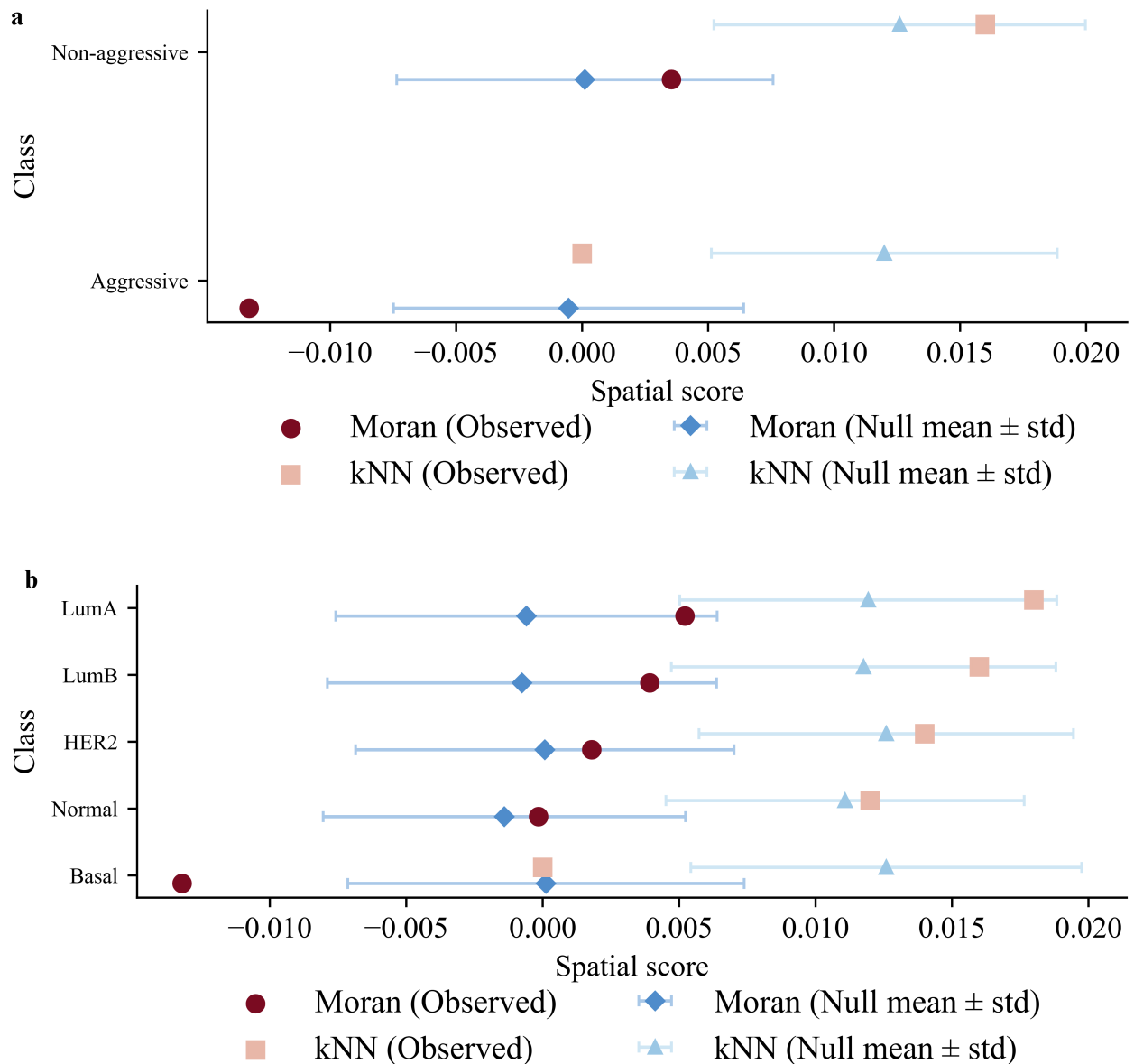


Fig. S25. Spatial quantification of learned representations in TCGA-BRCA. Observed Moran's I (spatial autocorrelation) and kNN purity (neighborhood consistency) for the binary (red) and multiclass (blue) tasks are compared against null distributions. In both experimental settings, the observed metrics significantly exceed random expectations. This demonstrates that Dynomap consistently organizes subtype-specific and risk-associated genes into coherent spatial neighborhoods regardless of the classification objective.

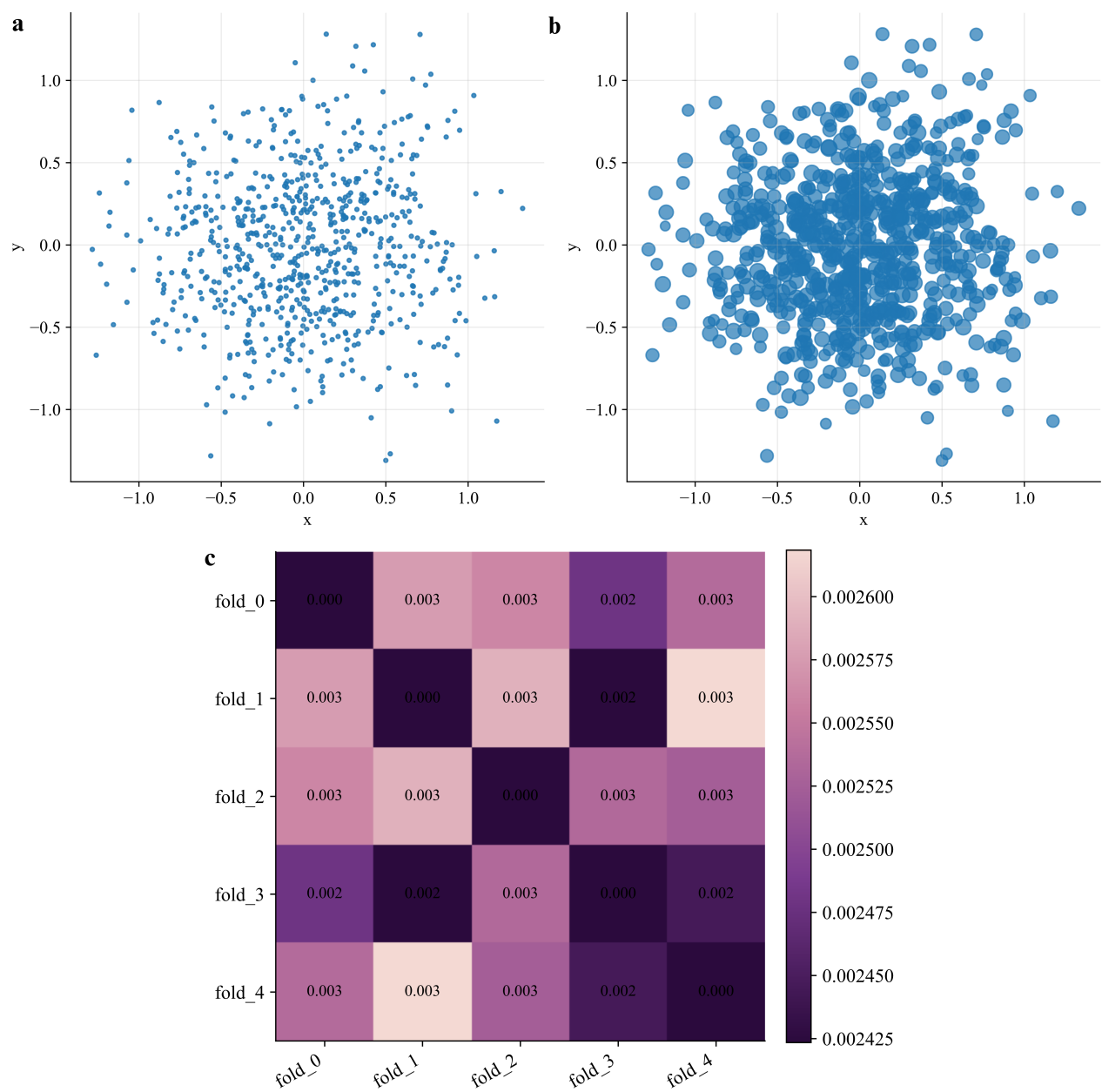


Fig. S26. Reproducibility and stability of learned spatial layouts in Parkinson's disease classification. a, Mean consensus layout across repeated training runs. Feature coordinates are averaged following Procrustes alignment to account for rotational and translational invariance. The resulting configuration demonstrates stable spatial organization of phenomic features across independent model initializations. b, Layout consensus uncertainty. The spatial dispersion of each feature across repeated runs is quantified as coordinate variance after alignment. Most features exhibit low positional variance, indicating convergence toward a consistent spatial configuration. c, Pairwise Procrustes distance matrix between independently trained layouts. Distances are computed after optimal alignment of feature coordinates. Low pairwise distances confirm high structural similarity between learned layouts across runs, demonstrating that the spatial representation is reproducible rather than driven by random initialization.

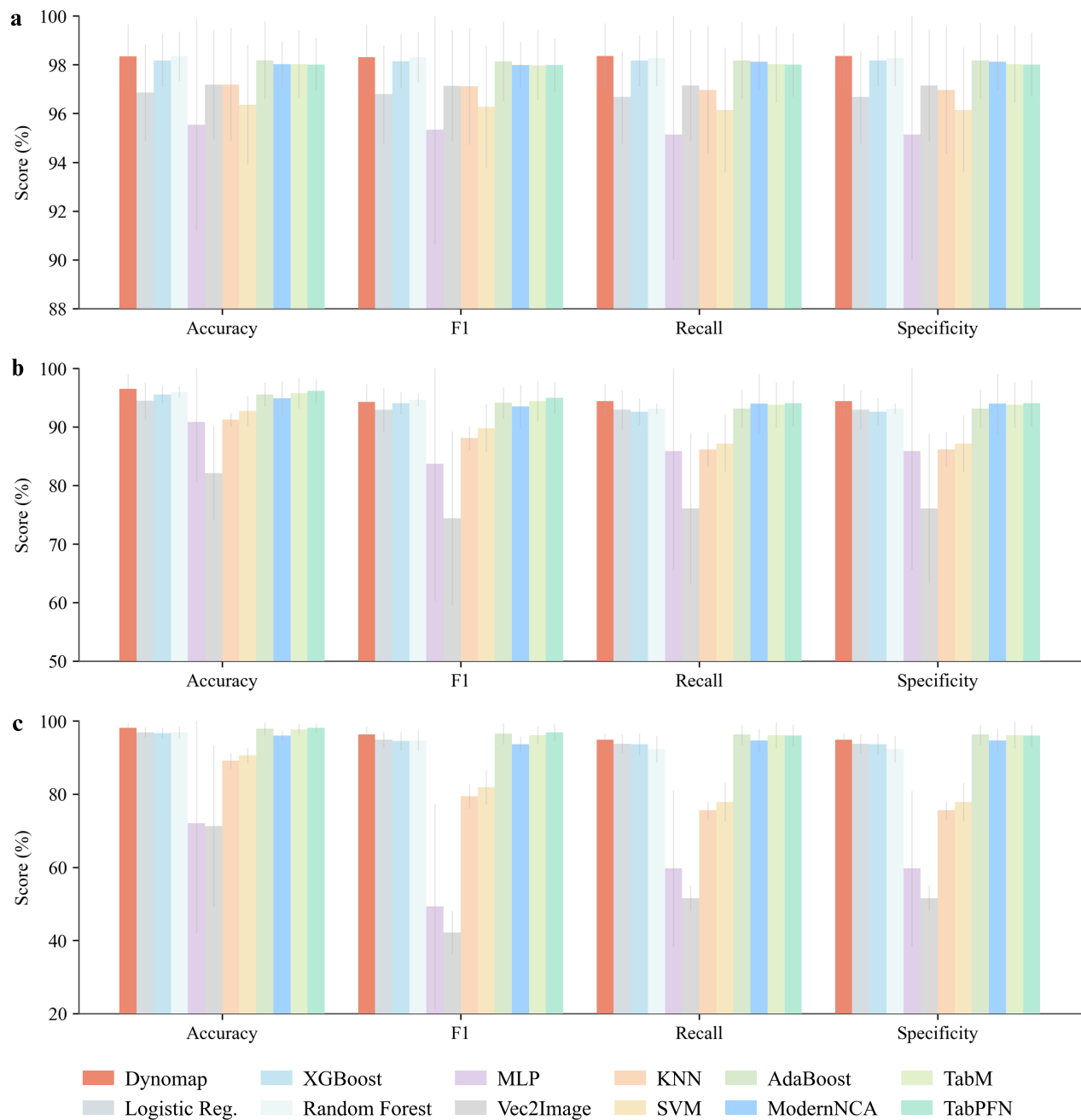


Fig. S27. Benchmark performance across three microarray pairwise cancer classification datasets. **a**, AP Breast vs Kidney (OpenML 1158; 604 samples, 10,935 original features, 4,000 features used). **b**, AP Breast vs Lung (OpenML 1150; 470 samples, 4,000 features used). **c**, AP Breast vs Omentum (OpenML 1127; 421 samples, 4,000 features used). Accuracy, macro-F1, macro-recall, and macro-specificity are reported for Dynomap and eleven baseline methods spanning classical and deep tabular models. Dynomap attains the highest accuracy across all three datasets and the highest macro-F1 on the AP Breast vs Kidney and AP Breast vs Lung tasks. Error bars indicate standard deviation across five cross-validation folds. For binary classification under macro averaging, recall and specificity are mathematically equivalent, and the two panels are therefore identical by construction.

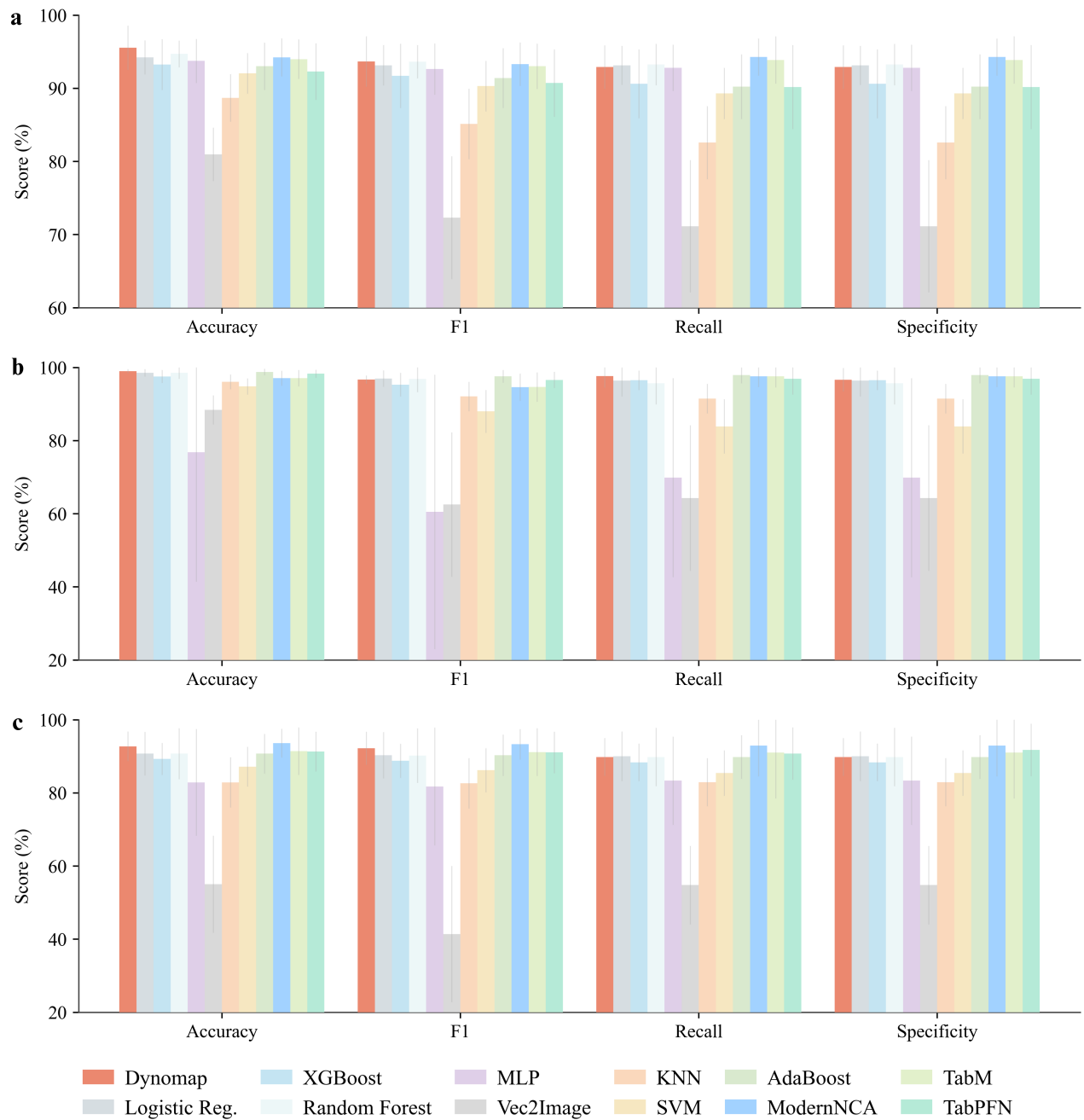


Fig. S28. Benchmark performance across three additional microarray pairwise cancer classification datasets. **a**, AP Colon vs Lung (OpenML 1126; 412 samples, 4,000 features used). **b**, AP Endometrium vs Breast (OpenML 1123; 405 samples, 4,000 features used). **c**, AP Endometrium vs Omentum (OpenML 1151; 138 samples, 4,000 features used). Accuracy, macro-F1, macro-recall, and macro-specificity are reported for Dynomap and eleven baseline methods. Dynomap achieves the highest accuracy and macro-F1 on the AP Colon vs Lung and AP Endometrium vs Breast tasks, and remains competitive with the strongest deep tabular baselines on AP Endometrium vs Omentum despite a small cohort of 138 samples. Error bars indicate standard deviation across five cross-validation folds.

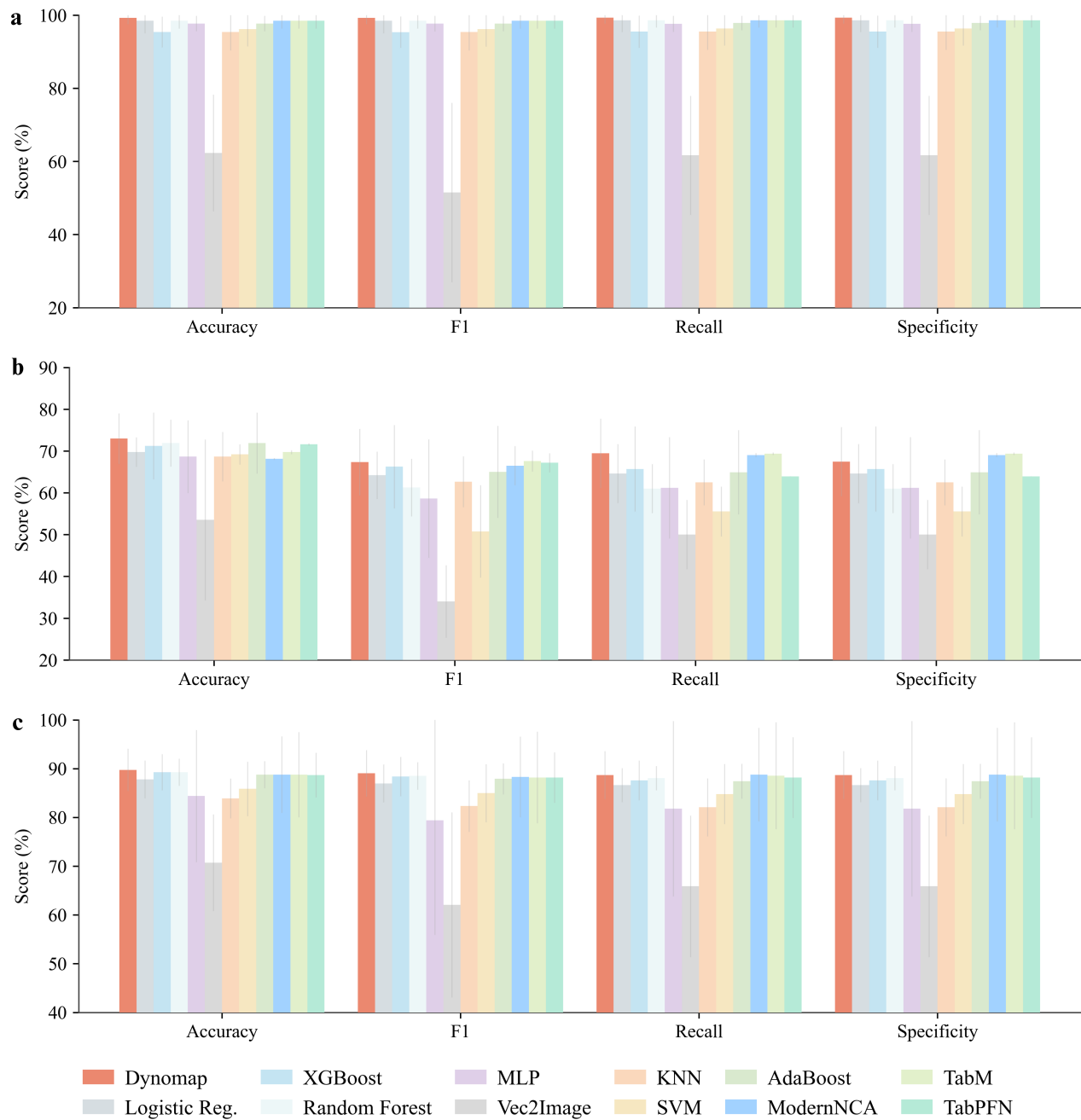


Fig. S29. Benchmark performance on three additional microarray pairwise cancer classification datasets. **a**, AP Endometrium vs Prostate (OpenML 1141; 130 samples, 4,000 features used). **b**, AP Endometrium vs Uterus (OpenML 1164; 185 samples, 4,000 features used). **c**, AP Omentum vs Uterus (OpenML 1124; 201 samples, 4,000 features used). Across all three datasets, Dynomap achieves the highest accuracy. On AP Endometrium vs Uterus, which is the most difficult task in the additional benchmark set, Dynomap reaches 73.0% accuracy compared with 71.9% for the strongest classical baseline and 71.6% for the strongest deep tabular baseline. Error bars indicate standard deviation across five cross-validation folds.

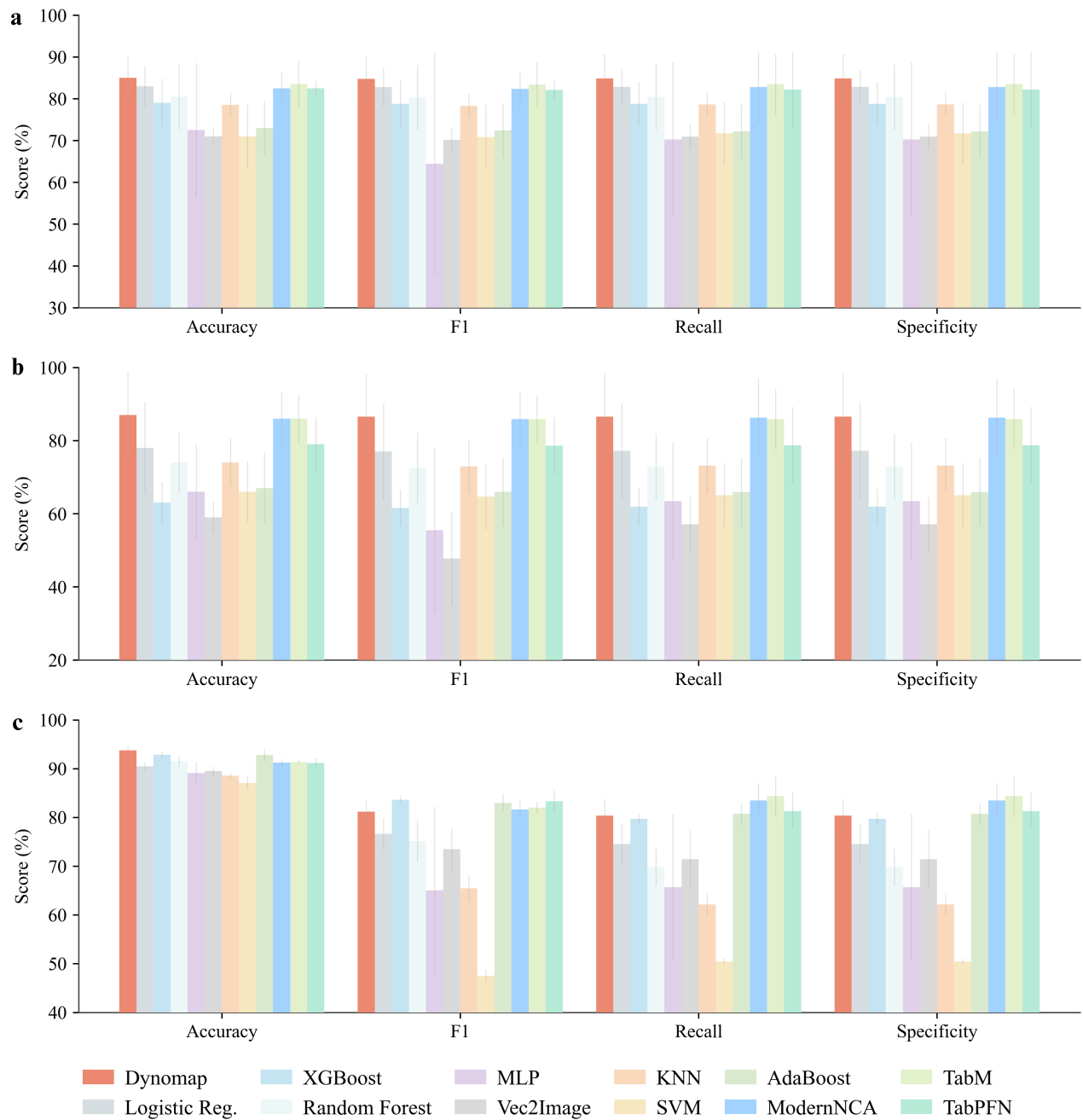


Fig. S30. Benchmark performance on mass spectrometry and class-imbalanced microarray datasets. **a**, Arcene mass spectrometry cancer classification (OpenML 1458; 200 samples, 9,961 original features, 4,000 features used). **b**, Arcene replicate cohort (OpenML 41157; 100 samples, 9,920 original features, 4,000 features used). **c**, OVA Ovary microarray classification (OpenML 1166; 1,545 samples, 10,935 original features, 4,000 features used). The OVA Ovary task is heavily imbalanced toward the negative class, so accuracy and macro-F1 diverge substantially across all models. Dynomap attains the highest accuracy on all three datasets and the highest macro-F1 on both Arcene cohorts. Error bars indicate standard deviation across five cross-validation folds.

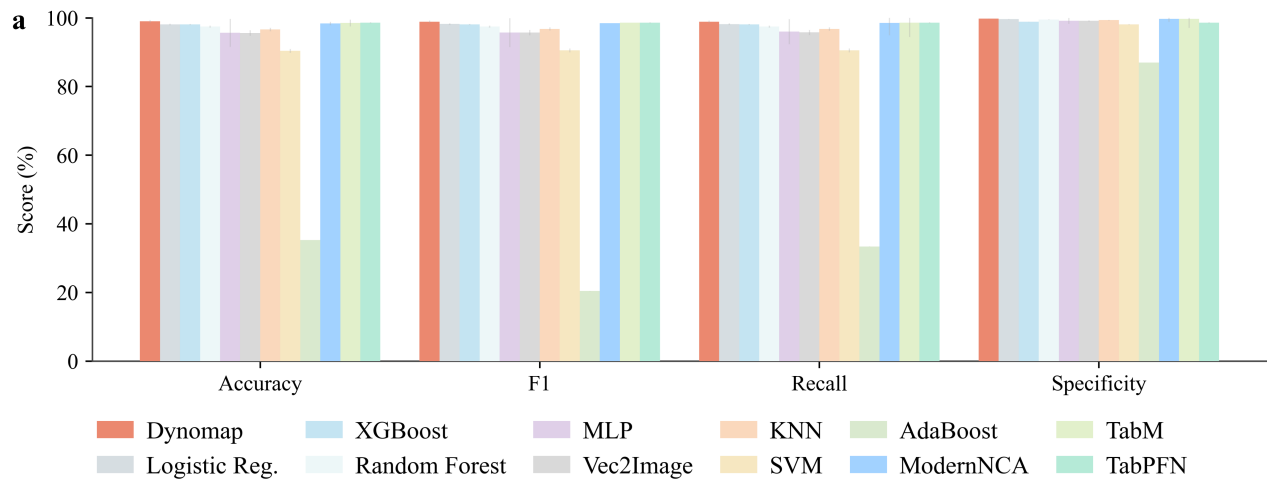


Fig. S31. Benchmark performance on the UCI HAR Smartphone activity recognition dataset. Six-class human activity classification using engineered inertial sensor descriptors from waist-mounted smartphones (10,299 samples, 561 features). Accuracy, macro-F1, macro-recall, and macro-specificity are reported for Dynomap and eleven baseline methods. Dynomap achieves the highest performance across all four metrics, with 98.96% accuracy and 98.82% macro-F1, exceeding the strongest classical baseline (XGBoost, 98.07%) and the strongest deep tabular baseline (TabPFN, 98.54%). The pronounced drop in AdaBoost performance reflects the multiclass setting, in which the base learner fails to recover balanced class margins. Error bars indicate standard deviation across cross-validation folds.

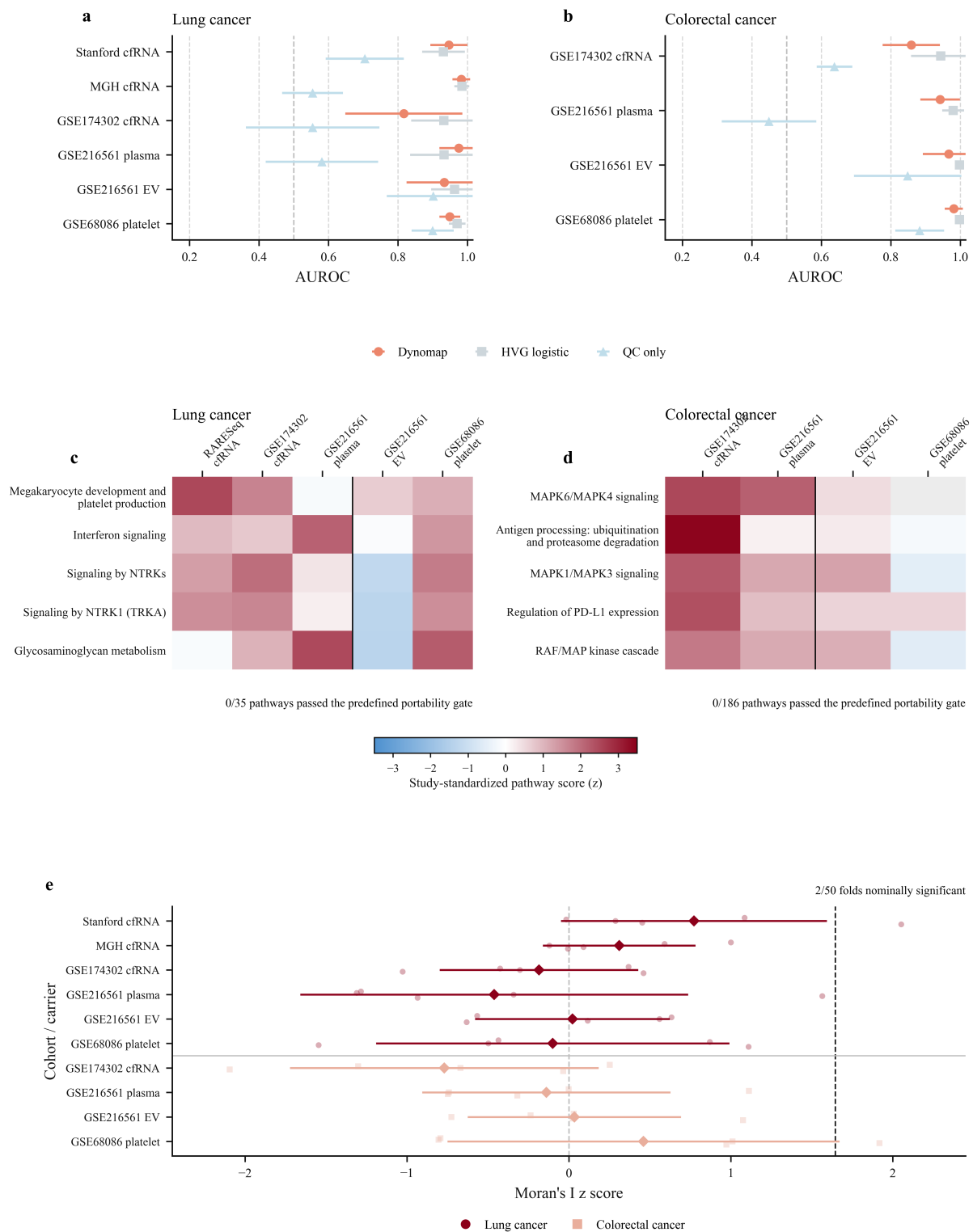


Fig. S32. Application of Dynomap across liquid-biopsy cohorts and RNA carriers. **a,b**, Lung-cancer (**a**) and colorectal-cancer (**b**) versus healthy-control classification across plasma cell-free RNA (cfRNA), extracellular-vesicle (EV) RNA, and tumor-educated platelet RNA datasets. A separate Dynomap model was trained for every cohort, RNA carrier, disease contrast, and cross-validation fold using 1,000 highly variable genes selected from the corresponding training partition. Points show the mean AUROC and error bars show the standard deviation across five donor-disjoint cross-validation folds. The HVG logistic-regression comparator used independently selected training-fold HVGs, and the QC-only logistic model used library size, number of detected genes, and maximum observed expression. Fold standard deviations are descriptive summaries and are not confidence intervals. Repeated libraries from the same RARE-Seq donor were assigned to the same fold. Lung-cancer sample sizes, in display order, were 179 libraries from 165 donors for RARE-Seq Stanford, 117 libraries from 102 donors for RARE-Seq MGH, and 54, 38, 33, and 113 donors for GSE174302 cfRNA, GSE216561 plasma, GSE216561 EV, and GSE68086 platelet RNA, respectively. Colorectal-cancer sample sizes were 73, 42, 33, and 98 donors for GSE174302 cfRNA, GSE216561 plasma, GSE216561 EV, and GSE68086 platelet RNA, respectively. Across all ten cohort- and carrier-specific tasks, independently retrained Dynomap models achieved mean fold AUROCs above 0.80, demonstrating successful application of the framework across heterogeneous liquid-biopsy measurement settings.