

Supplementary Material

Supplementary Figures

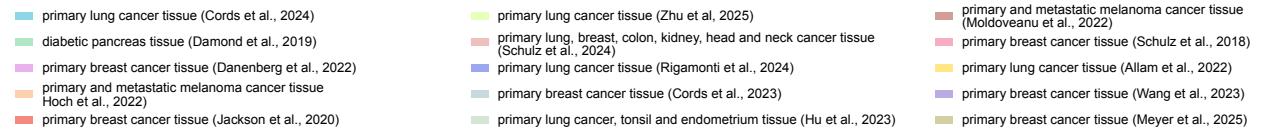


Figure S1: Color legend for the 15 datasets used in this study.

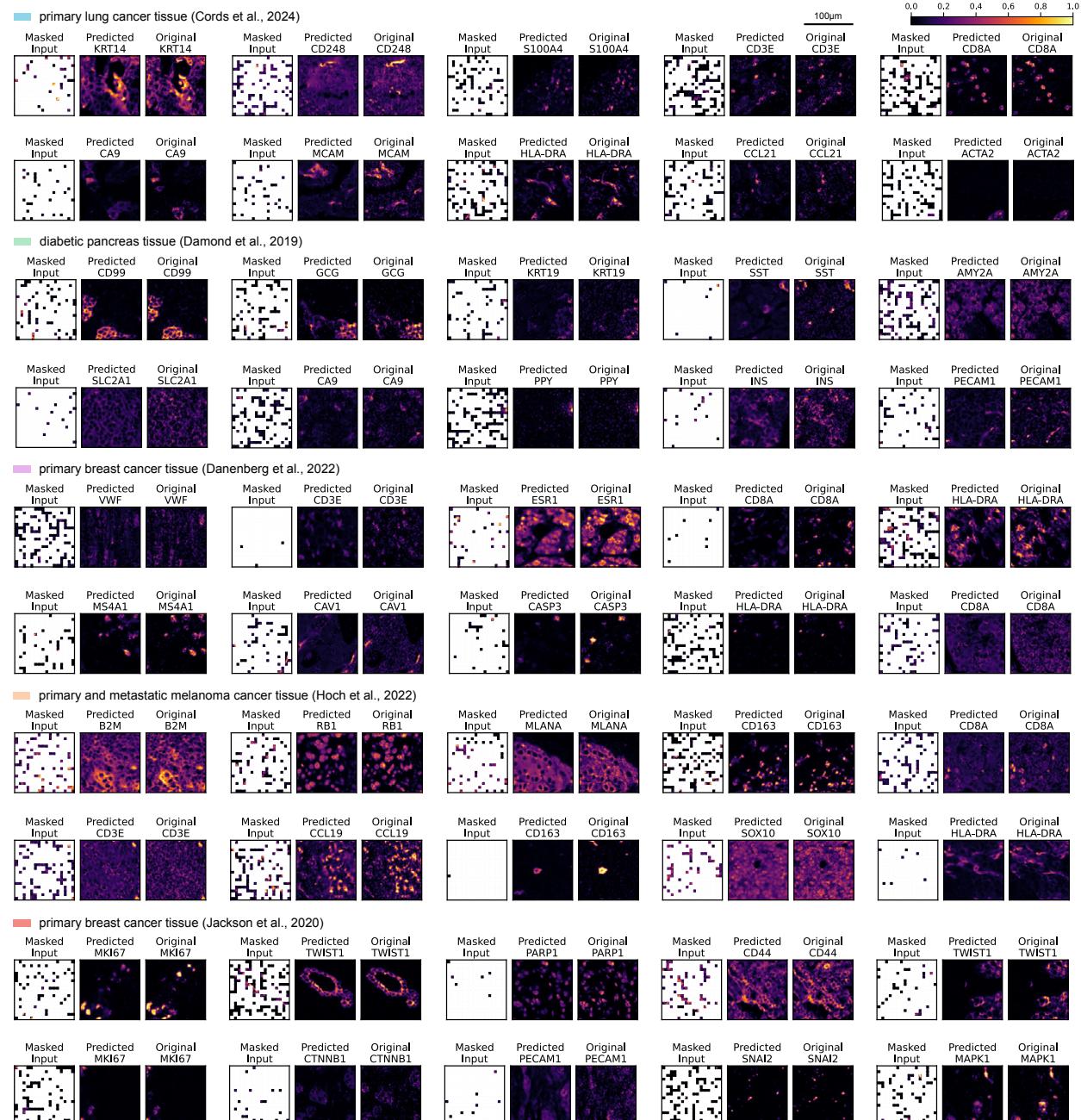


Figure S2: Examples of image reconstruction results using independent masking, with masking ratios between 60% and 100%, for Cords et al.³⁸, Damond et al.⁸⁰, Danenber et al.⁹, Hoch et al.⁴³ and Jackson et al.³⁹.

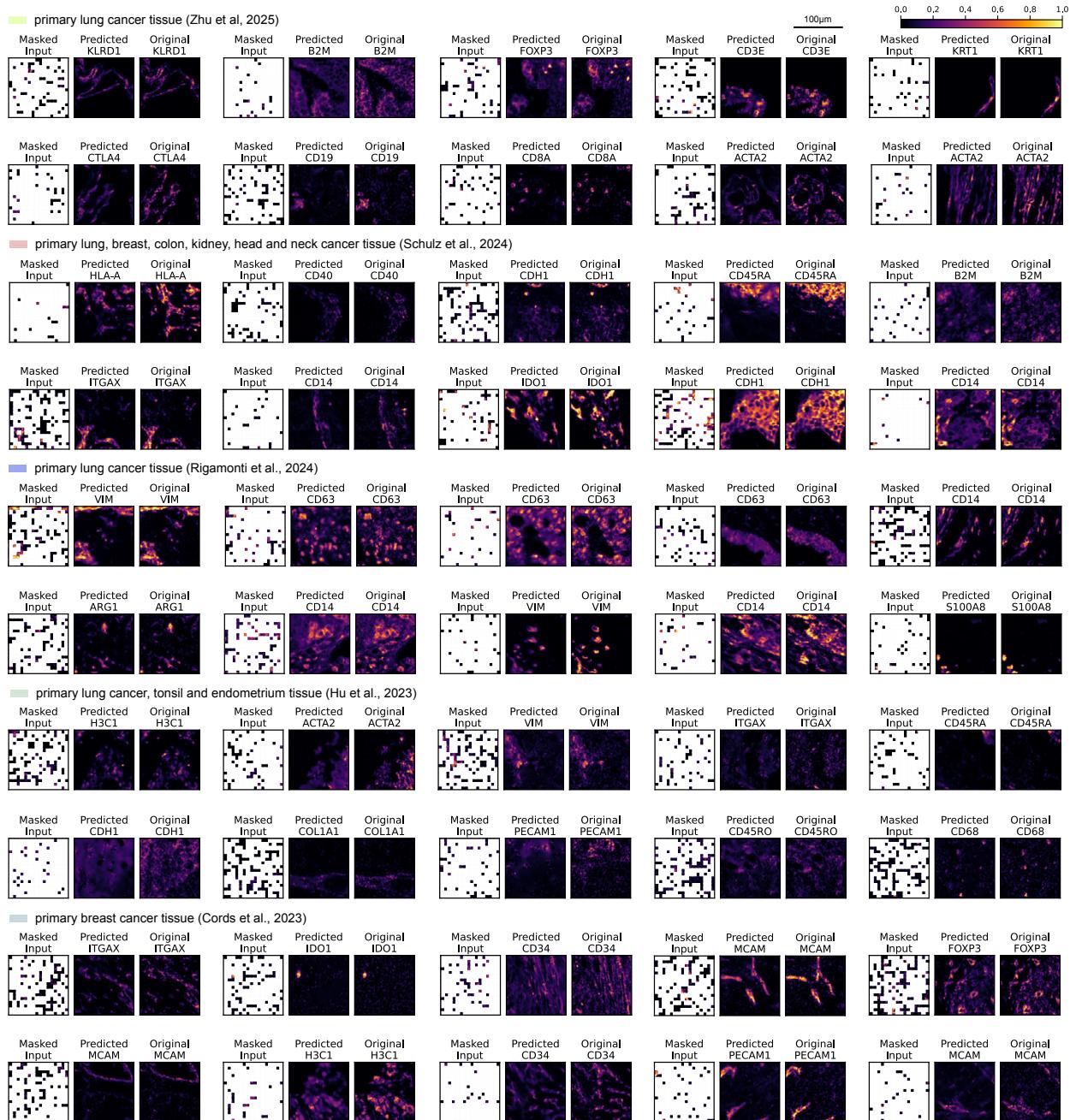


Figure S3: Examples of image reconstruction results using independent masking, with masking ratios between 60% and 100%, for Zhu et al.⁷³, Schulz et al.⁷⁶, Rigamonti et al.¹², Hu et al.⁷⁴ and Cords et al.⁷⁷.

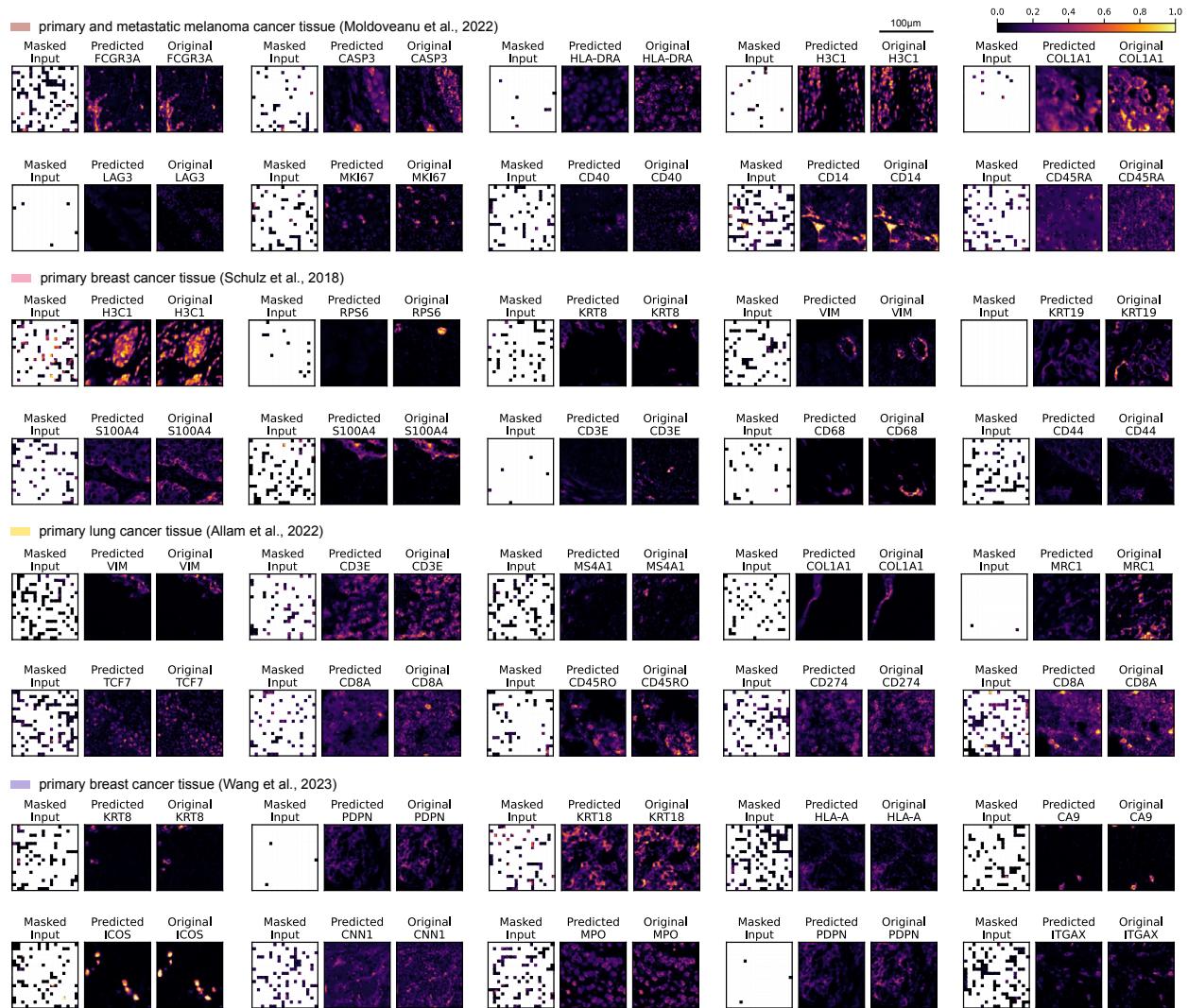


Figure S4: Examples of image reconstruction results using independent masking, with masking ratios between 60% and 100%, for Moldoveanu et al.⁷⁹, Schulz et al.⁷⁸, Allam et al.⁷⁵ and Wang et al.⁴.

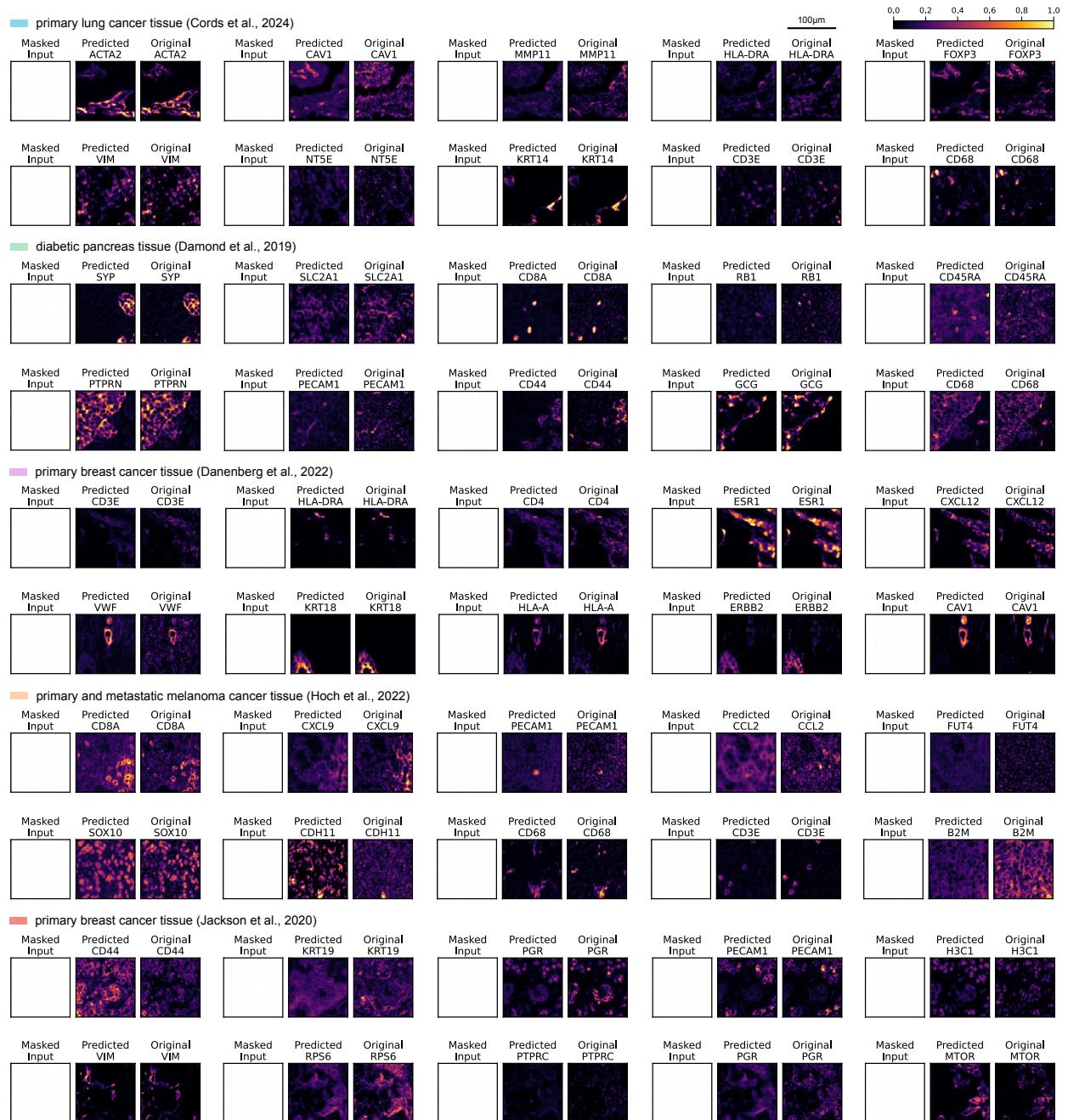


Figure S5: Examples of image reconstruction results after masking an entire marker for Cords et al.³⁸, Diamond et al.⁸⁰, Danenberberg et al.⁹, Hoch et al.⁴³ and Jackson et al.³⁹. Each row depicts different channels of the same tissue niche.

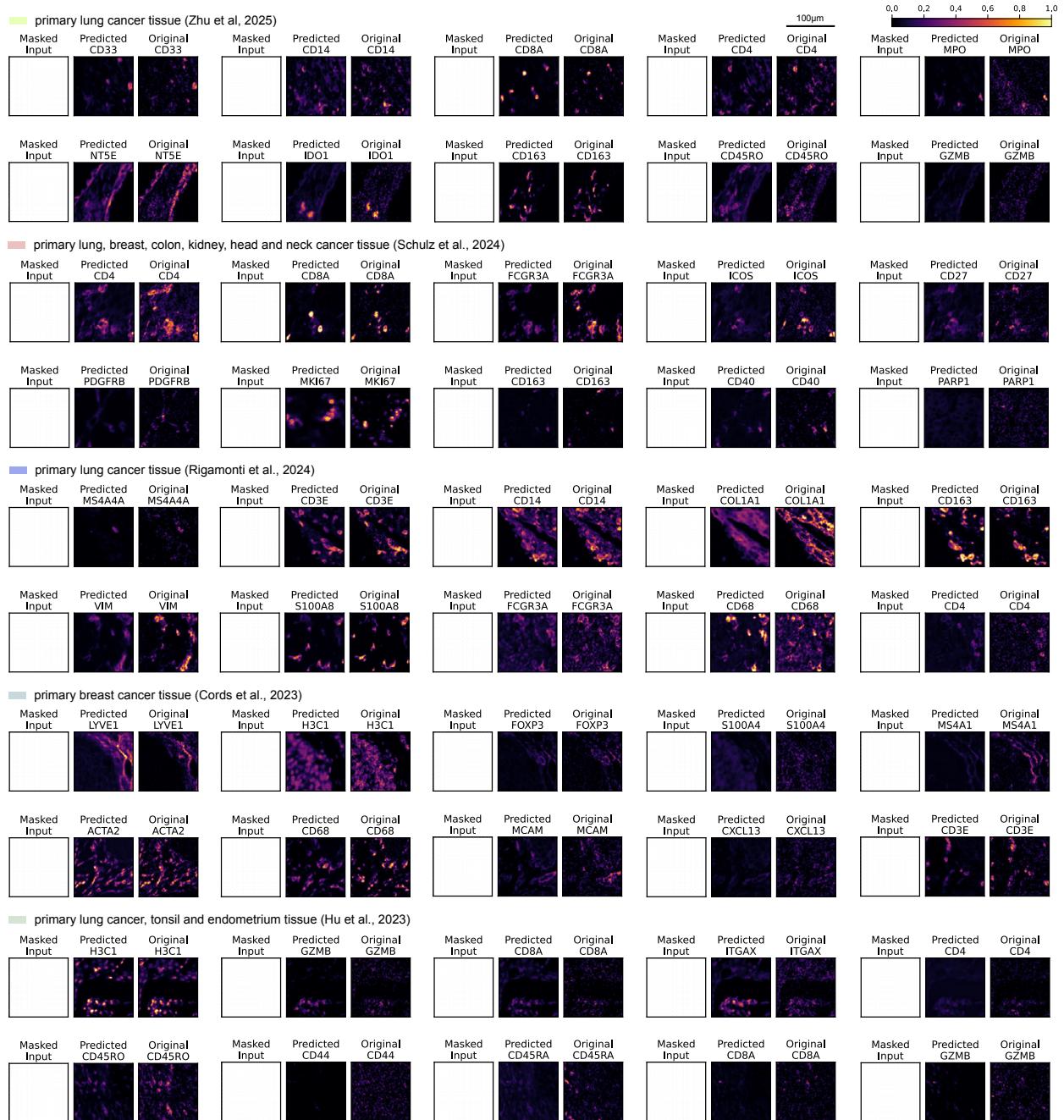


Figure S6: Examples of image reconstruction results after masking an entire marker for Zhu et al.⁷³, Schulz et al.⁷⁶, Rigamonti et al.¹², Hu et al.⁷⁴ and Cords et al.⁷⁷. Each row depicts different channels of the same tissue niche.

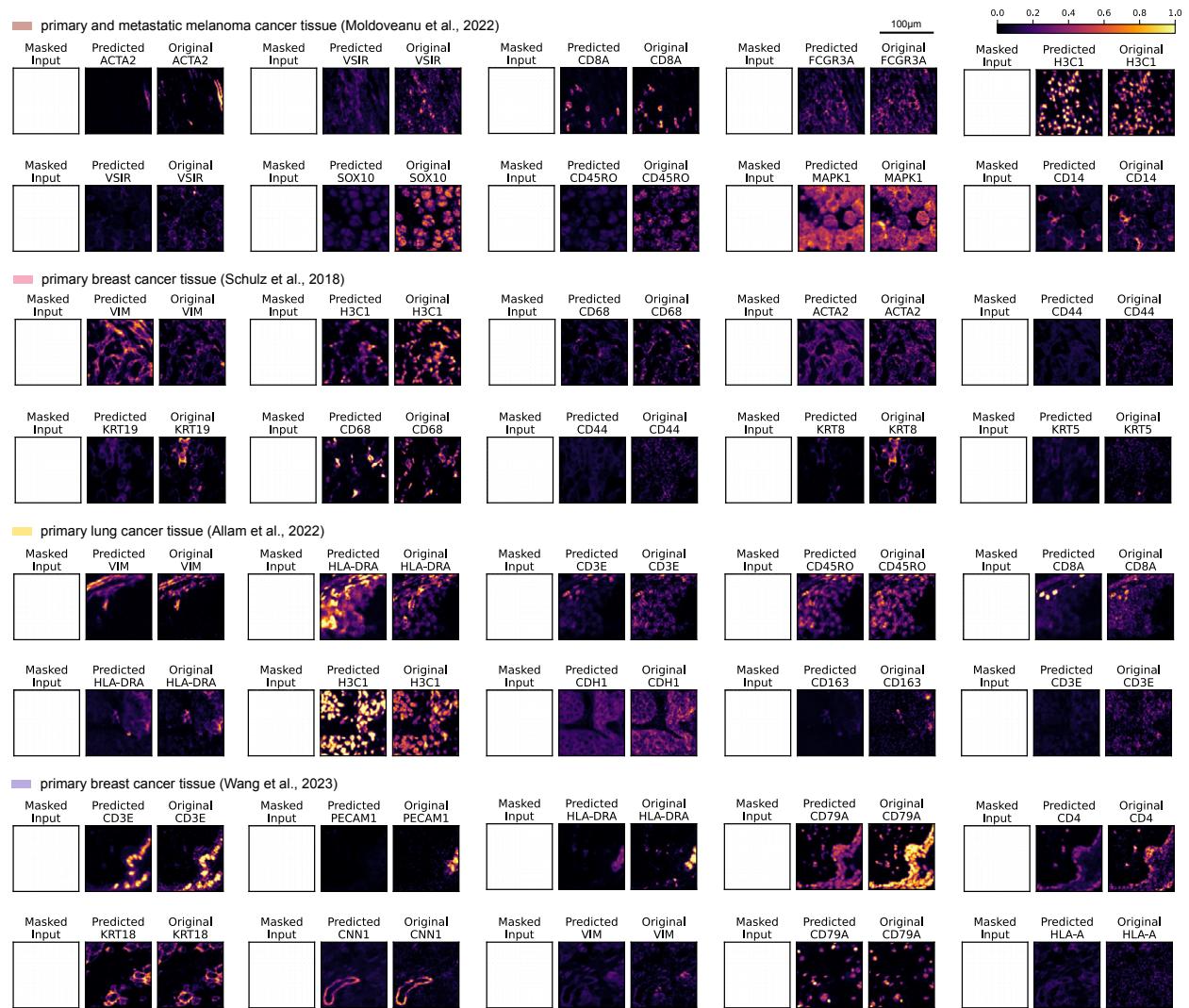


Figure S7: Examples of image reconstruction results after masking an entire marker for Moldoveanu et al.⁷⁹, Schulz et al.⁷⁸, Allam et al.⁷⁵ and Wang et al.⁴. Each row depicts different channels of the same tissue niche.

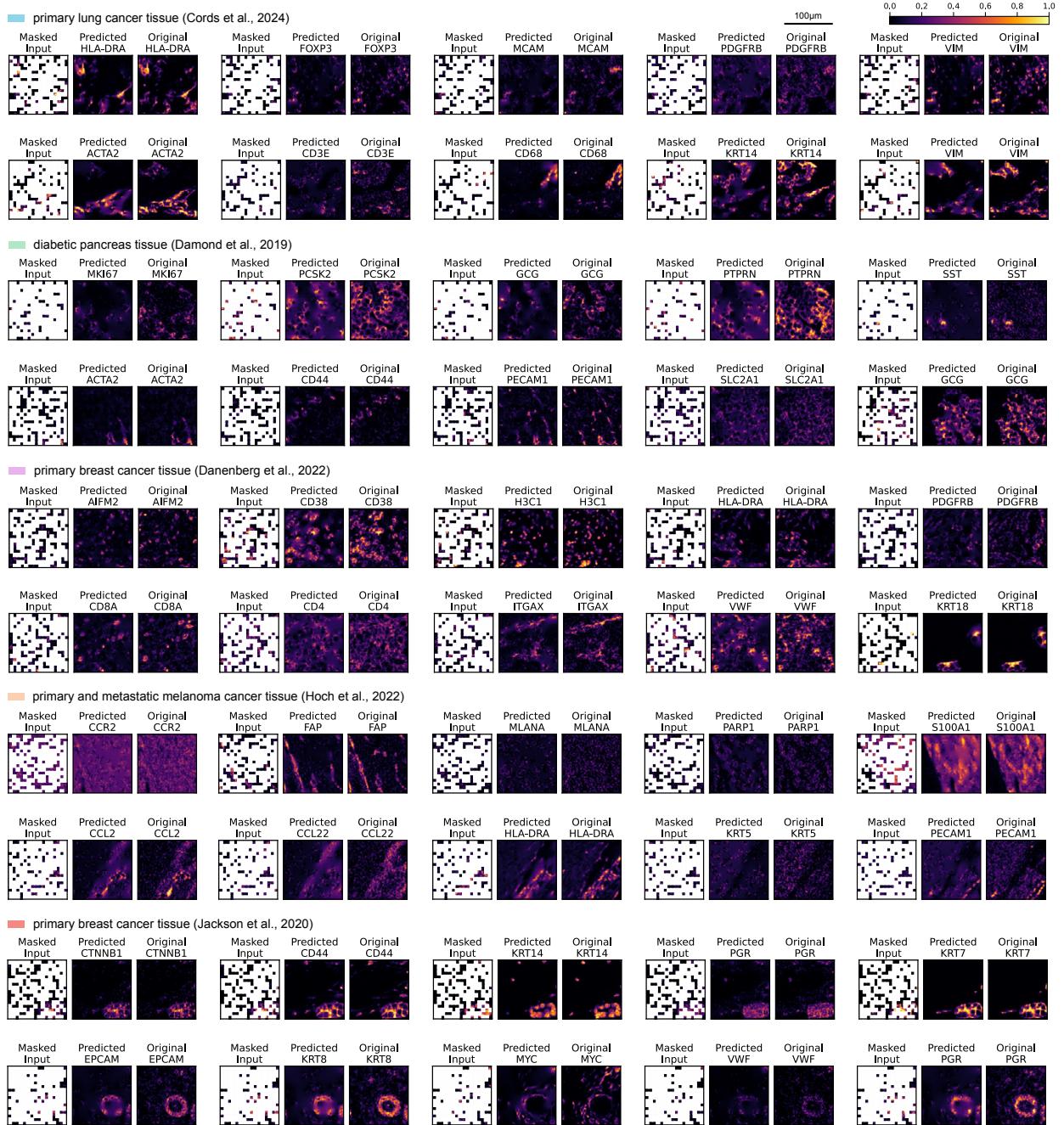


Figure S8: Examples of reconstruction results using niche masking, with masking ratio between 60% and 100%, for Cords et al.³⁸, Damond et al.⁸⁰, Danenberge et al.⁹, Hoch et al.⁴³ and Jackson et al.³⁹. Each row depicts different channels of the same tissue niche.

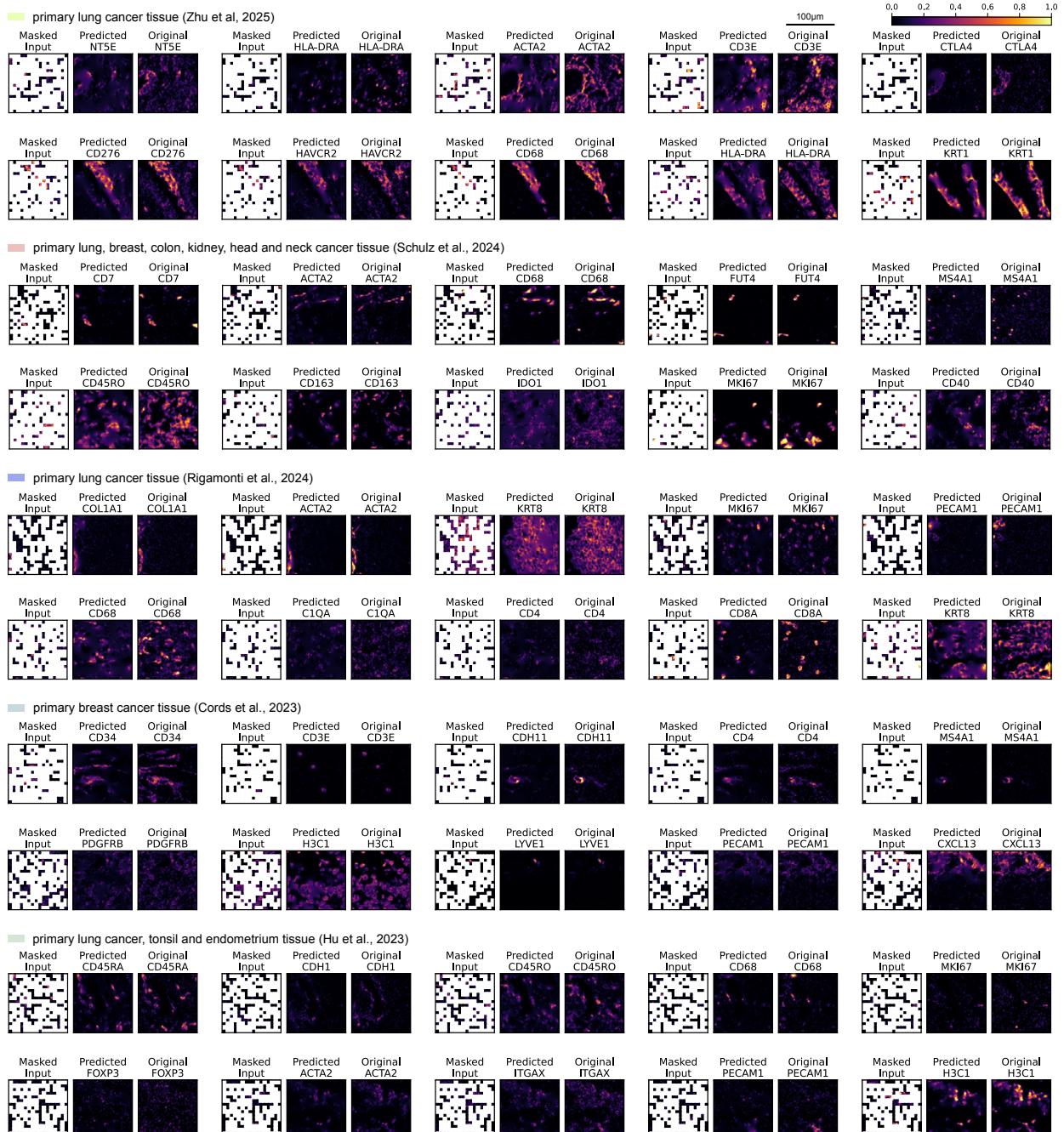


Figure S9: Examples of reconstruction results using niche masking, with masking ratio between 60% and 100%, for Zhu et al.⁷³, Schulz et al.⁷⁶, Rigamonti et al.¹², Hu et al.⁷⁴ and Cords et al.⁷⁷. Each row depicts different channels of the same tissue niche.

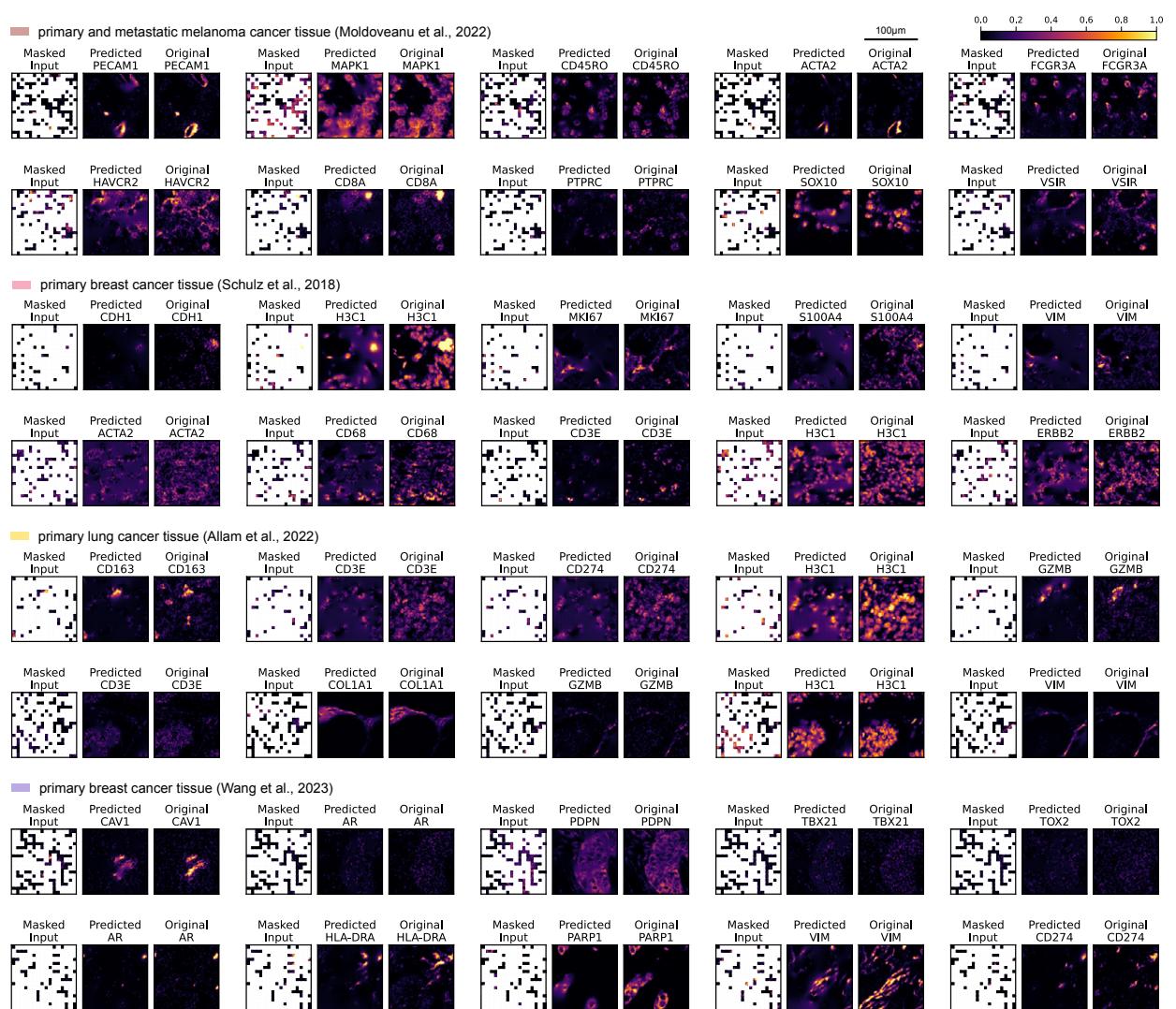


Figure S10: Examples of reconstruction results using niche masking, with masking ratio between 60% and 100%, for Moldoveanu et al.⁷⁹, Schulz et al.⁷⁸, Allam et al.⁷⁵ and Wang et al.⁴. Each row depicts different channels of the same tissue niche.

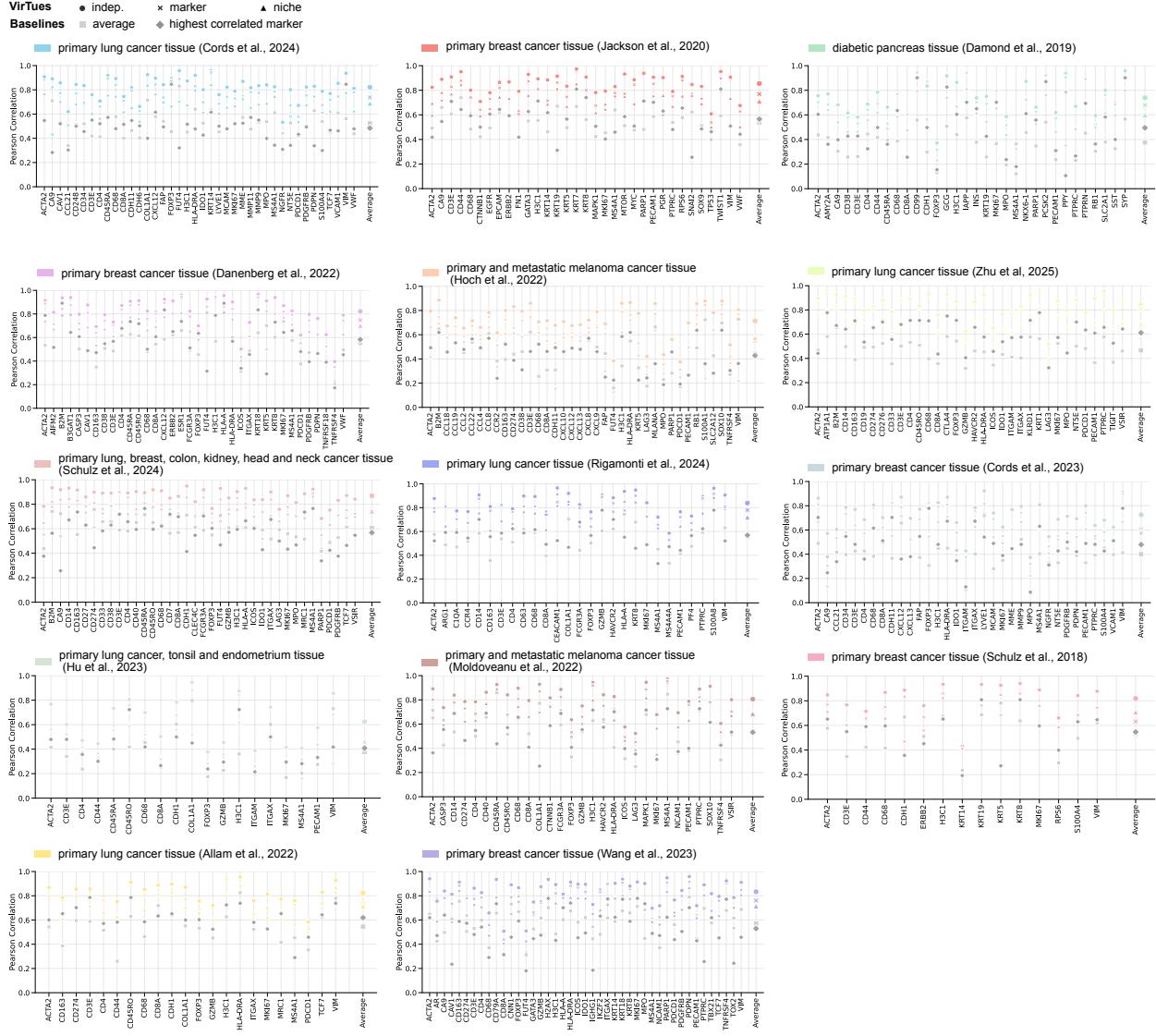


Figure S11: Overview of marker-wise reconstruction performance, measured by Pearson correlation, for independent (circles), marker (crosses), and niche (triangles) masking schemes. For comparison, light gray squares represent the correlation obtained by predicting the mean channel intensity of visible pixels for all masked pixels under independent masking. Dark gray diamonds indicate the correlation of each marker with the highest correlated other marker.

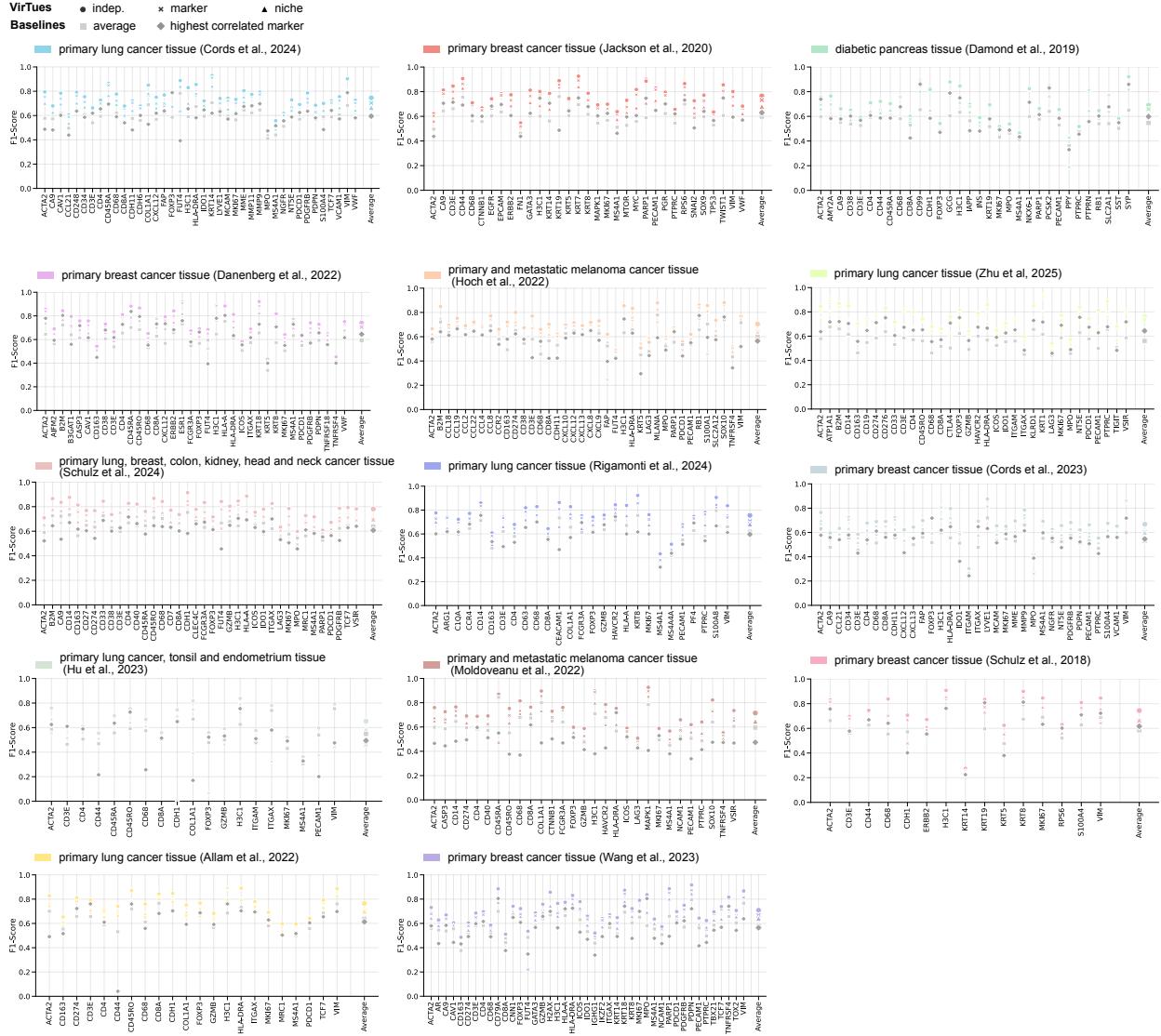


Figure S12: Overview of marker-wise reconstruction performance, measured by F1-scores of pixel-wise marker positivity, for independent (circles), marker (crosses), and niche (triangles) masking schemes. Positivity is determined by thresholding at the mean intensity of each marker. For comparison, light gray squares represent the performance obtained by predicting the mean channel intensity of visible pixels for all masked pixels under independent masking. Dark gray diamonds indicate the performance when predicting for each marker the highest correlated other marker.

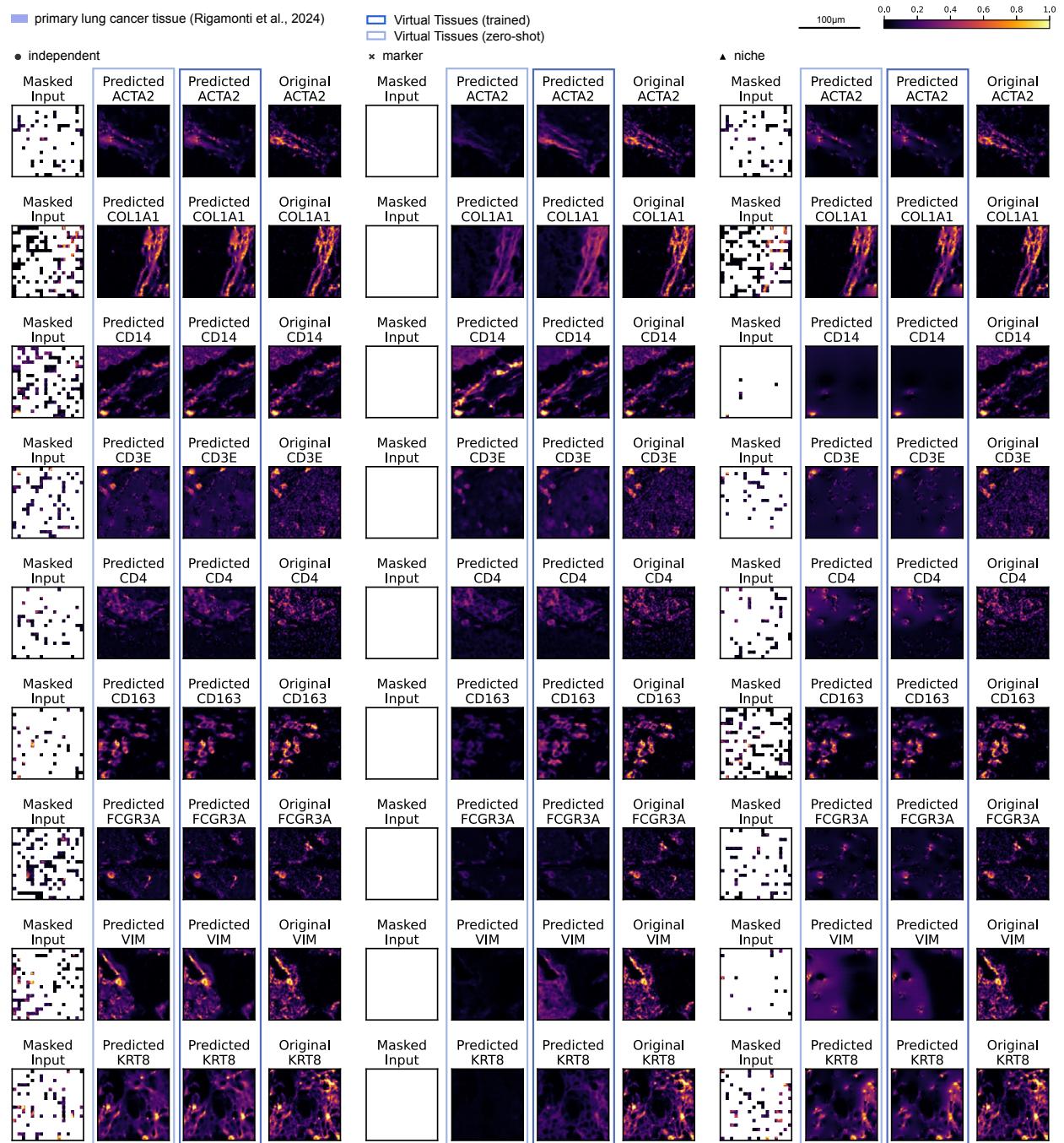


Figure S13: Examples of zero-shot reconstruction results using independent (left), marker (middle) and niche (right) masking for markers of Rigamonti et al.¹² observed during pretraining. For comparison, the same samples are also reconstructed using a model instance pretrained on Rigamonti et al.¹².

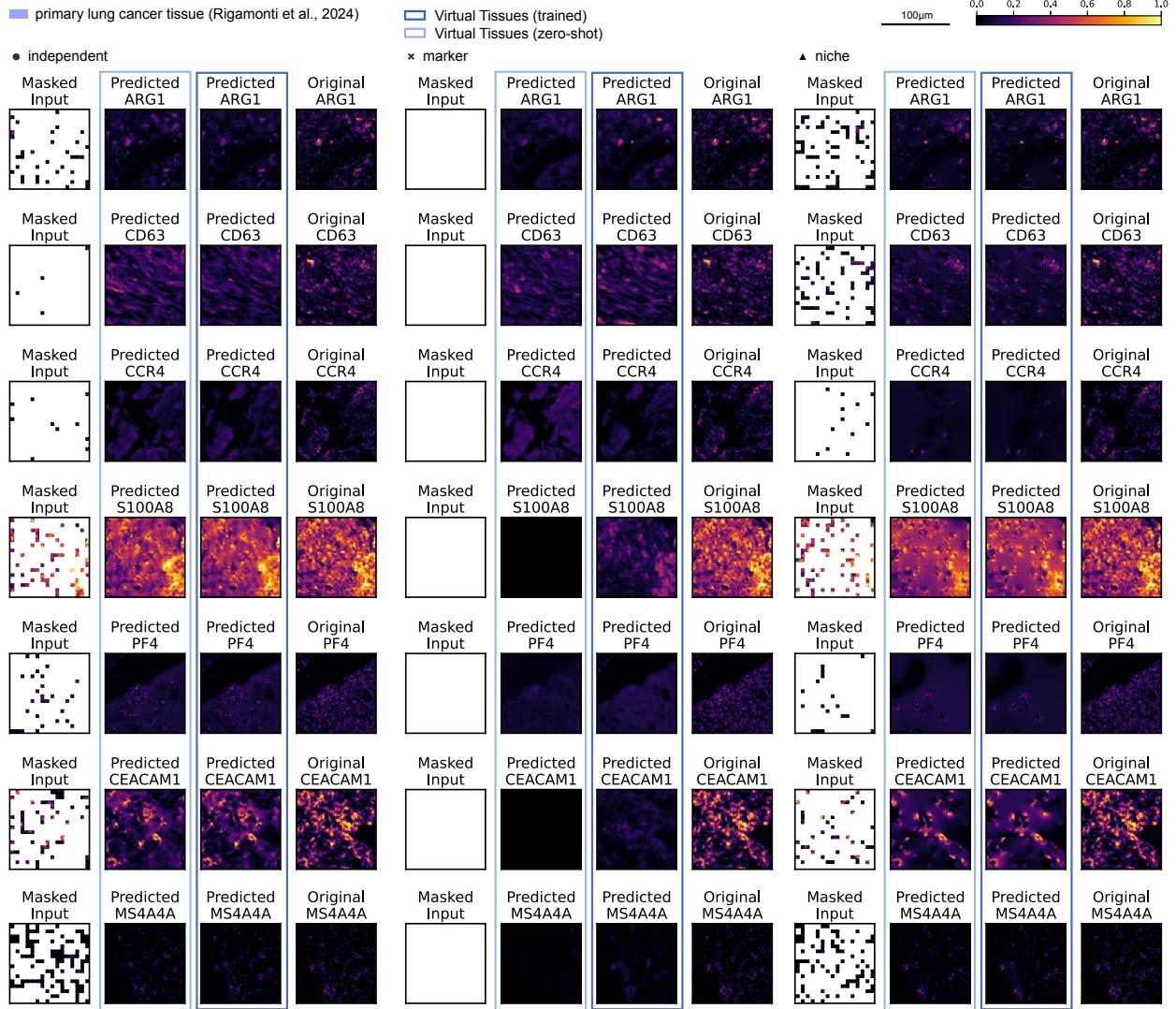


Figure S14: Examples of zero-shot reconstruction results using independent (left), marker (middle) and niche (right) masking for markers of Rigamonti et al. ¹² not observed during pretraining. For comparison, the same samples are also reconstructed using a model instance pretrained on Rigamonti et al. ¹².

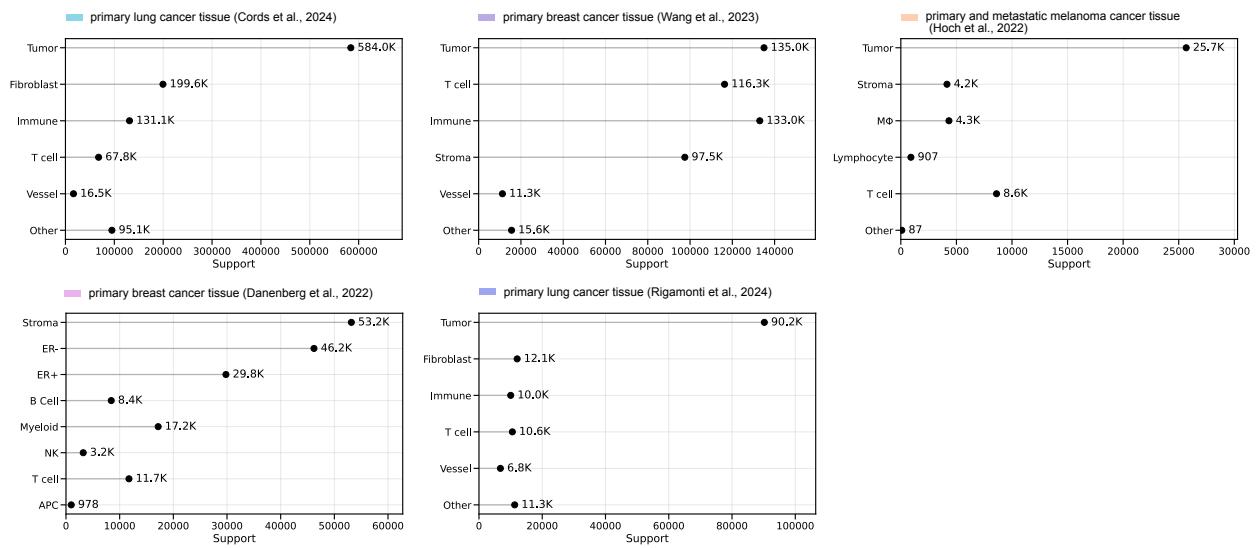


Figure S15: Support distribution for coarse cell types across the test sets of Cords et al. ³⁸, Wang et al. ⁴, Hoch et al. ⁴³, Danenbergs et al. ⁹, and Rigamonti et al. ¹², illustrating the imbalance in cell type representation in cell classification tasks.

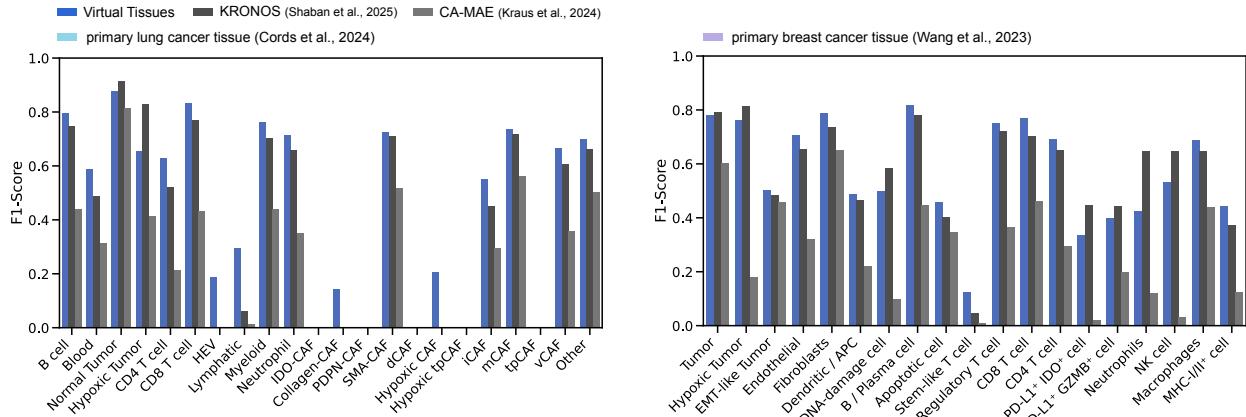


Figure S16: F1-scores for fine-grained cell-type classification on Cords et al. ³⁸ and Wang et al. ⁴.

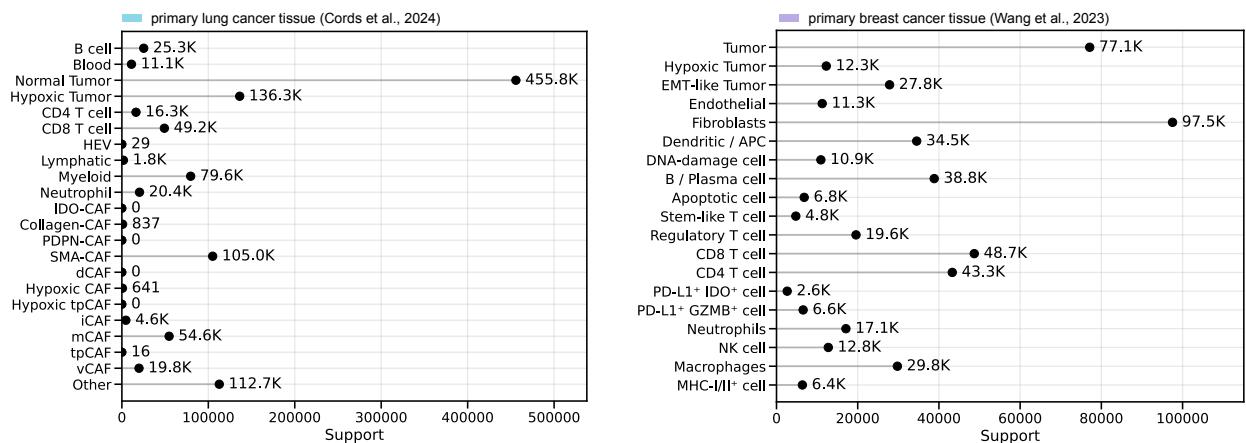


Figure S17: Support distribution for fine-grained cell types across the test sets of Cords et al. ³⁸ and Wang et al. ⁴, highlighting the imbalance in cell type representation in cell classification tasks.

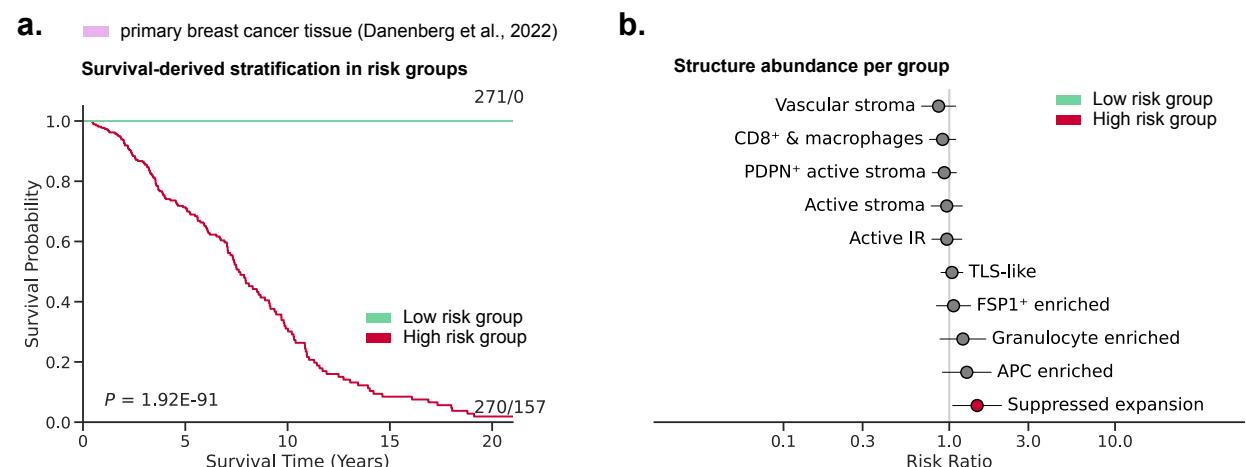


Figure S18: a, Kaplan-Meier survival curves for survival-derived control risk groups on Danenberg et al. ⁹. Risk groups are formed by balanced grouping of (i) survivors and late-censored patients, and (ii) deceased and early-censored patients. b, Risk ratio of each TME structure's occurrence in the high risk control group. Lines indicate the 95% confidence intervals.

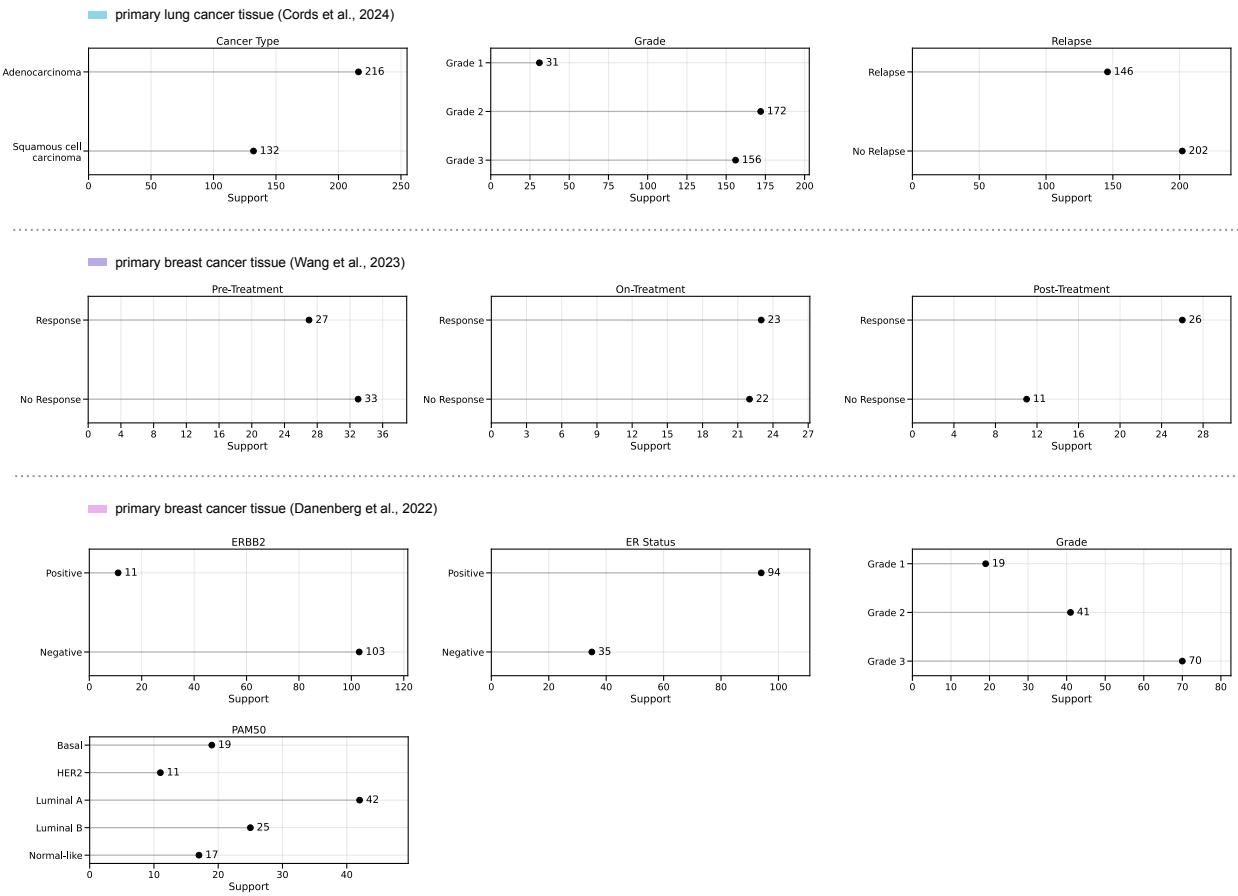


Figure S19: Support distribution of labels for tissue-level classification tasks across the test sets of Cords et al.³⁸, Wang et al.⁴ and Danenberge et al.⁹.

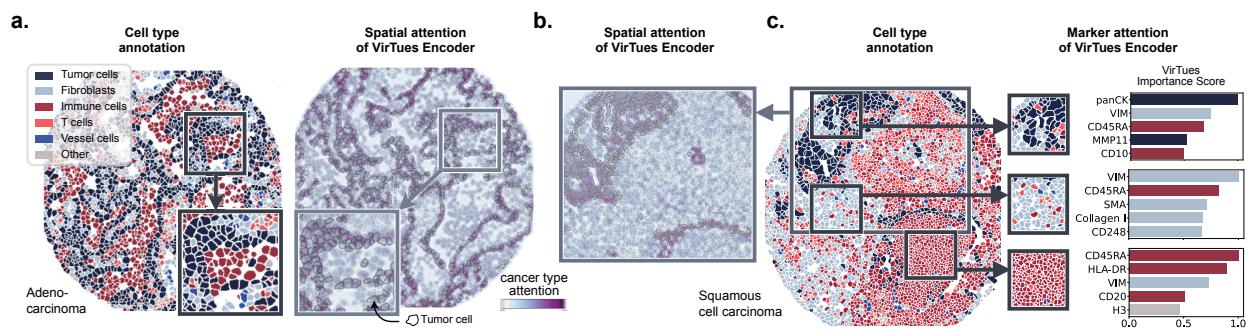


Figure S20: Visualization of spatial and marker attention in Cords et al.³⁸. **a, b,** Spatial attention maps derived from an extended version of VirTues incorporating an additional class token that attends to all patch summary tokens. The model was fine-tuned for cancer subtype prediction and the shown maps represent the class token's attention scores. **c,** Marker attention scores computed across three tissue niches with distinct cell type compositions. The five markers most attended to by other markers (after averaging attention scores across these regions) are shown. Bars are color-coded according to the canonical cell type associated with each marker.

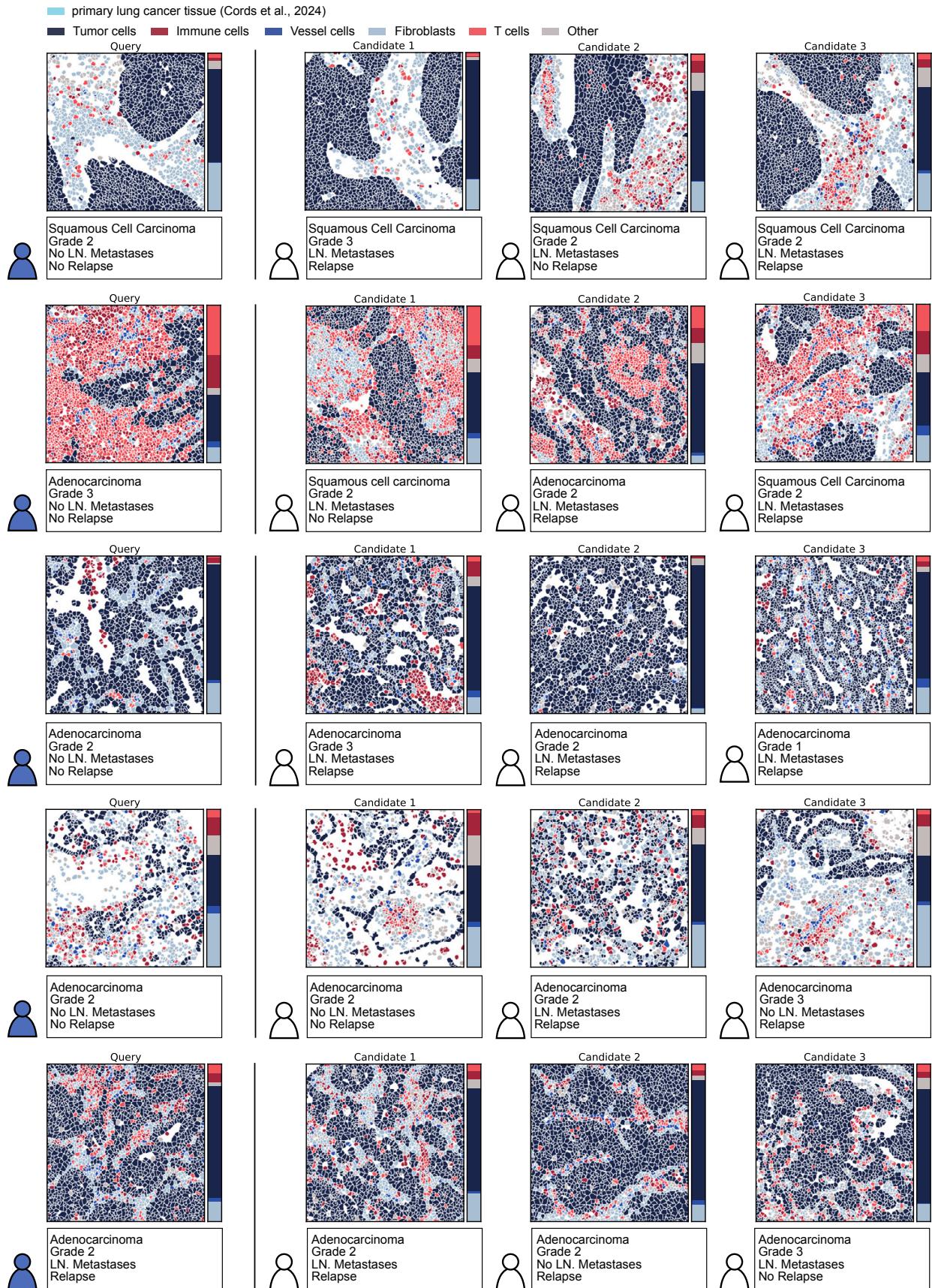


Figure S21: Examples of retrieval results for tissues in Cords et al.³⁸ using the Wasserstein distance of VirTues's niche representation tokens as the similarity measure. Tissues are visualized using their color-coded cell type masks. Colorbars show the proportional cell type compositions.

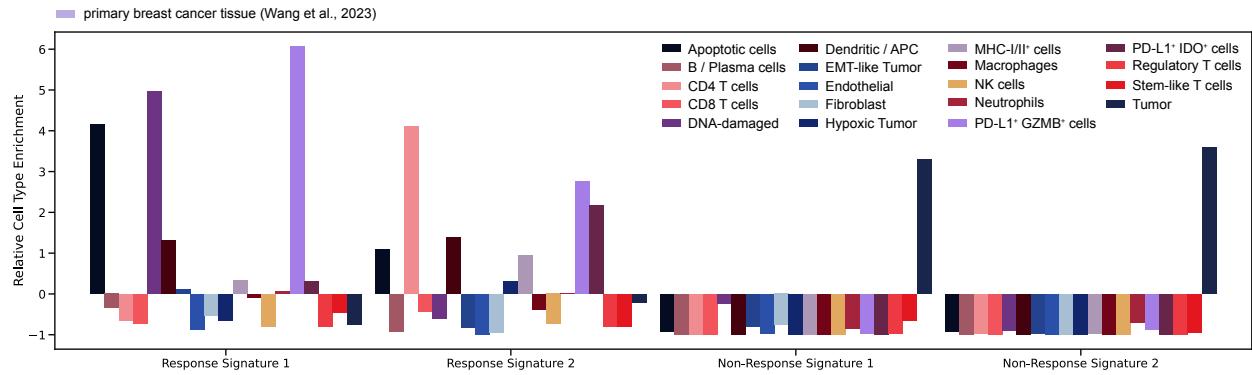


Figure S22: Relative enrichment of cell types in selected response and non-response clusters of Wang et al.⁴ compared to overall proportions.

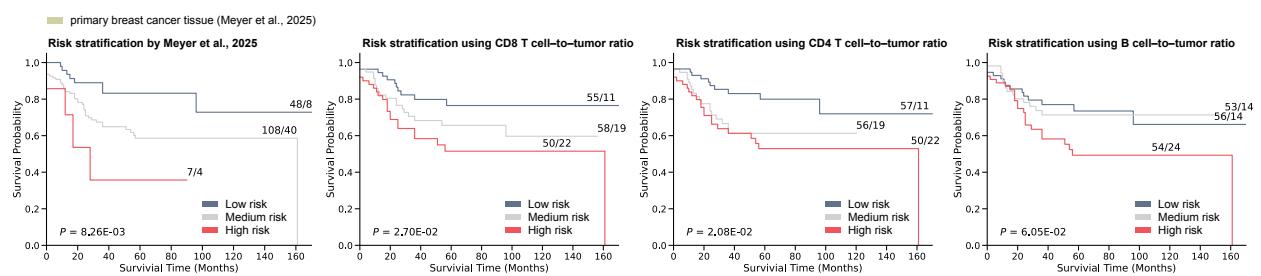


Figure S23: Kaplan-Meier curves of disease-free survival on Meyer et al.³¹ using different baseline risk stratification systems. From left to right: (1) original risk groups from Meyer et al.³¹, (2) risk scores based on tertile-transformed CD8 T cell-to-tumor ratios, (3) CD4 T cell-to-tumor ratios, and (4) B cell-to-tumor ratios. Each plot reports the p-value from a log-rank test comparing the low- and high-risk groups.

Supplementary Tables

Table S1: Overview of IMC datasets. Only images used for the training and evaluation of VirTues are considered. Cell counts are indicated only for datasets with single cell annotations and masks.

Dataset	Tissue Origin	Images	Patients	Crops	Cells
Cords et al. ³⁸	primary lung cancer	1'969	1'048	50'236	5'880'664
Danenberg et al. ⁹	primary breast cancer	688	688	11'837	938'561
Hoch et al. ⁴³	primary and metastatic melanoma cancer	50	35	2'015	325'881
Jackson et al. ³⁹	primary breast cancer	711	284	15'788	1'237'470
Damond et al. ⁸⁰	healthy and diabetic pancreas	843	12	10'013	1'775'386
Zhu et al. ⁷³	primary lung cancer	1'645	62	66'982	-
Schulz et al. ⁷⁶	primary lung, breast, colon, kidney, head and neck cancer	179	179	3'544	-
Rigamonti et al. ¹²	primary lung cancer	158	83	25'586	659'561
Cords et al. ⁷⁷	primary breast cancer	108	12	6'281	-
Hu et al. ⁷⁴	primary lung cancer, tonsil and endometrium tissue	51	12	4'112	-
Moldoveanu et al. ⁷⁹	primary and metastatic melanoma cancer	90	90	4'584	227'592
Schulz et al. ⁷⁸	primary breast cancer	77	77	1'758	-
Allam et al. ⁷⁵	primary lung cancer	26	26	2'073	-
Wang et al. ⁴	primary breast cancer	1'842	279	41'480	2'538'185
Meyer et al. ³¹	primary breast cancer	450	215	13'459	928'993
Total	-	8'887	3'102	259'748	$\geq 14'512'293$

Table S2: Nested subsets of markers from Cords et al.³⁸, as used in the analysis of Fig. 1e, grouped according to their presumed informativeness for general tissue characterization.

Top 10 Markers	Top 11–20 Markers	Top 21–30 Markers	Top 31–40 Markers
FAP	ACTA2	MPO	TCF7
CD68	HLA-DRA	S100A4	MME
CD3E	CA9	H3C1	CD45RA
PDCD1	MS4A1	MCAM	CD248
NT5E	IDO1	CDH11	LYVE1
VIM	MMP9	COL1A1	VWF
FOXP3	PDGFRB	VCAM1	CXCL12
CD8A	CD34	PDPN	CCL21
KRT14	CD4	MMP11	CDH6
MKI67	FUT4	NGFR	CAV1

Table S3: Grouping of the original cell phenotypes available in the metadata of Danenberg et al.⁹ into seven high-level classes used in the cell classification task.

High-Level Class	Original Cell Phenotypes
Stromal Cell	Fibroblasts Myofibroblasts Endothelial Myofibroblasts PDPN ⁺ Fibroblasts FSP1 ⁺
Myeloid	CD15 ⁺ Macrophages Granulocytes MHC ^{high} CD15 ⁺
Natural Killer Cell	MHC I ^{high} CD57 ⁺ CD57 ⁺
ER-Positive Epithelial Cell	CK ⁺ CXCL12 ⁺ ER ^{high} CXCL12 ⁺ HER2 ⁺ Ep Ki67 ⁺ CK8-18 ⁺ ER ^{high} CK ^{low} ER ^{med}
ER-Negative Epithelial Cell	CK ^{med} ER ^{low} CK ^{low} ER ^{low} CK8-18 ^{high} CXCL12 ^{high} Basal cells CK8-18 ^{high} ER ^{low} Ep CD57 ⁺
T Cell	CD8 ⁺ T cell CD4 ⁺ T cell T _{Reg} and T _{Ex} cells
B Cell	B cells CD38 ⁺ lymphocytes
Antigen-Presenting Cell	MHC I ^{high} & II ^{high}

Table S4: Grouping of the original cell phenotypes from the Hoch et al.⁴³ metadata into five high-level classes used in the cell classification task.

High-Level Class	Original Cell Phenotypes
Tumor	Tumor, HLA-DR
T Cell	CD8 ⁺ T cells, CD8 ⁻ T cells
Macrophage	Macrophage
Lymphocyte	CD38, Neutrophil
Stroma	Stroma, Vasculature
Other	Other

Table S5: Grouping of the original cell phenotypes available in the metadata of Wang et al.⁴ into six high-level classes and 19 low-level classes used in the cell classification task.

High-Level Class	Low-Level Class	Original Cell Phenotypes
Immune	Stroma	Fibroblasts PDPN ⁺ Stromal Myofibroblasts
	Dendritic / APC	DCs PD-L1 ⁺ APCs PD-L1 ⁺ IDO ⁺ APCs
	Macrophages	M2 Mac
	B / Plasma cells	CD79a ⁺ Plasma CD20 ⁺ B
	NK cells	CD56 ⁺ NK CD56 ⁺ NE
	Neutrophils	CD15 ⁺ Neutrophils
	Hypoxic Tumor	CA9 ⁺ CA9 ⁺ Hypoxia CK8/18 ^{med} panCK ^{med}
	Tumor cells	CK ^{hi} GATA3 ⁺ CK ^{lo} GATA3 ⁺ Basal AR ⁺ LAR
	DNA-damage cells	pH2AX ⁺ DSB
	EMT-like Tumor	Vimentin ⁺ EMT
T Cell	Apoptotic cells	Apoptosis
	CD4 T cells	CD4 ⁺ TCF1 ⁺ T CD4 ⁺ PD1 ⁺ T
	Regulatory T cells	Treg Helios ⁺
	CD8 T cells	CD8 ⁺ T CD8 ⁺ TCF1 ⁺ T CD8 ⁺ PD1 ⁺ T _{Ex} CD8 ⁺ GZMB ⁺ T
	Stem-like T cells	TCF1 ⁺
Vessel	Endothelial	Endothelial
Other	PD-L1 ⁺ GZMB ⁺ cells	PD-L1 ⁺ GZMB ⁺
	PD-L1 ⁺ IDO ⁺ cells	PD-L1 ⁺ IDO ⁺
	MHC-I&II ⁺ cells	MHC I&II ^{hi}

Table S6: Grouping of the original cell phenotypes available in the metadata of Rigamonti et al.¹² into six high-level classes used in the cell label transfer experiment.

High-Level Class	Original Cell Phenotypes
Tumor	Tumor cells
T cell	CD8+, CD4+, Treg
Immune	B cells, Mf, Ki67+
Fibroblast	aSMA+, Vim
Vessel	Vessels
Other	Other