



Fig. S8. Characterization of mast cells

a t-SNE plot showing 1,962 mast cells (MCs) colour-coded by 4 clusters, namely C1_CPA3, C2_CMA1, C3_ICAM1 and C4_AIF1, which were generated using the unaligned clustering method. **b** Plot showing biweight midcorrelation score of CCs from different cancer types. The OvC dataset shows poor correlation with other cancers, indicating CCA was not applicable likely due to too few cells derived from OvC. **c** Fraction of singlet and doublet cells predicted by DoubletFinder for mast cell subclusters. **d** Box plot showing number of detected genes for mast cell subclusters. **e** Heatmap for differential expressed genes of each mast cell subcluster. Historically, MCs were classified based on expression of proteases in secretory granules, including tryptases (*TPSAB1*, *TPSB2*), chymase (*CMA1*) and carboxypeptidase A3 (*CPA3*)⁹⁹. We frequently observed C1_CPA3s and C2_CMA1s, while C3_ICAM1s were rare. Tryptases were highly expressed in C1-C3 clusters, but C1_CPA3s were $CPA3^{high}/CMA1^{low}$, while *vice versa* C2_CMA1s and C3_ICAM1s were $CPA3^{low}/CMA1^{high}$. C3_ICAM1s represented activated MCs that additionally express immune homing factors (*ICAM1*), chemokines and cytokines (*CCL2*, *CCL4*, *CCL4L2* and *IL13*), histamine enzyme (*HDC*) and pro-angiogenic molecules (*VEGFA*). C4_AIF1s had higher predicted doublet rate (**c**), higher detected gene number (**d**), and higher expression of the macrophage markers (**e**), therefore represented macrophage doublets. **f** Fraction of cells for mast cell phenotypes per cancer type. **g** Fraction of cells for mast cell phenotypes from different cancer types (left), and sample origins (right). C3_ICAM1s were underrepresented in malignant tissue ($FDR=3.3 \times 10^{-23}$), suggesting that tumours do not favour mast cell activation. **h** Heatmap showing transcription factor activity (AUC score) calculated by SCENIC for each mast cell phenotype. It revealed that MC activation in C3_ICAM1s is determined by NF- κ B signalling (*NFKB1*, *NFKB2* and *REL*), while MIFT, a mediator of early MC development¹⁰⁰, was active in C1_CPA3s.