Supplementary figures

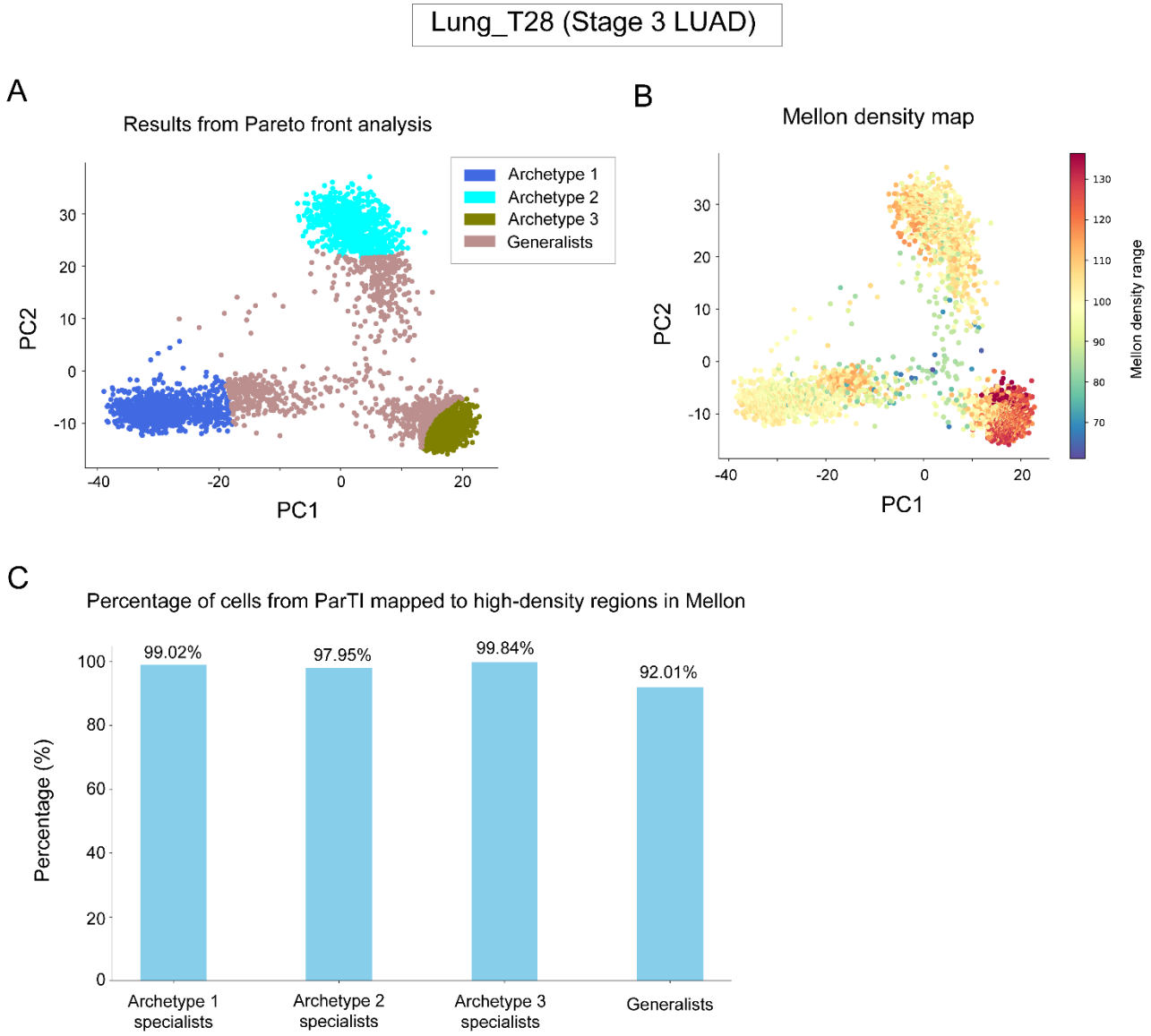

Figure S1 – Overlap between specialists and generalists identified in Pareto front analysis and high-density regions in cell-density based stability analysis. (A) Plot showing classification of cells to specific archetypes and generalists in the Pareto front analysis by ParTI for a stage 3 LUAD sample (Lung_T28). **(B)** Plot showing the density map for the same cells from the same sample obtained using Mellon. **(C)** Bar plot showing the overlap between specialists and generalists obtained using ParTI and the high-density regions – signifying stable cell states - identified using Mellon.

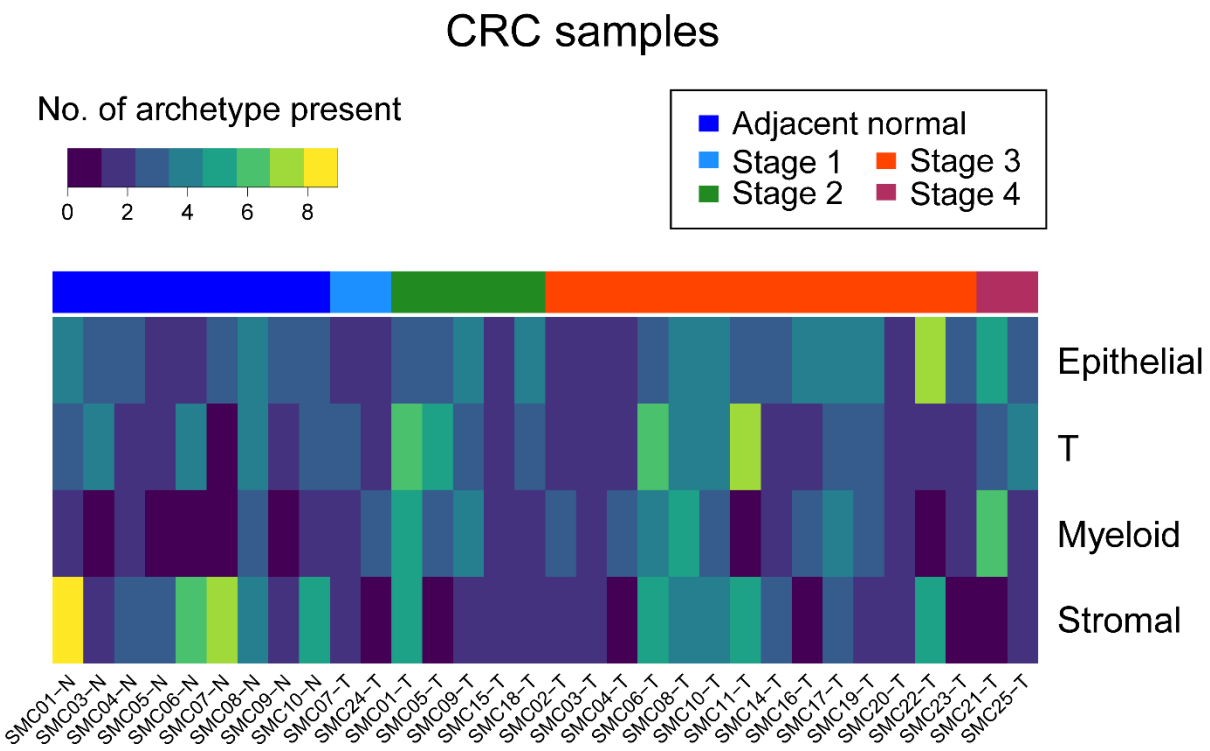

Figure S2 – **Functional diversity in each cell type in individual patient samples in CRC.** Heatmap showing the number of archetypes in which each cell type (Epithelial, T, Myeloid and Stromal) was part of in each patient sample in CRC, thereby, depicting their functional diversity. The colour bar above the heatmap shows stages for cancer samples along with adjacent normal samples.

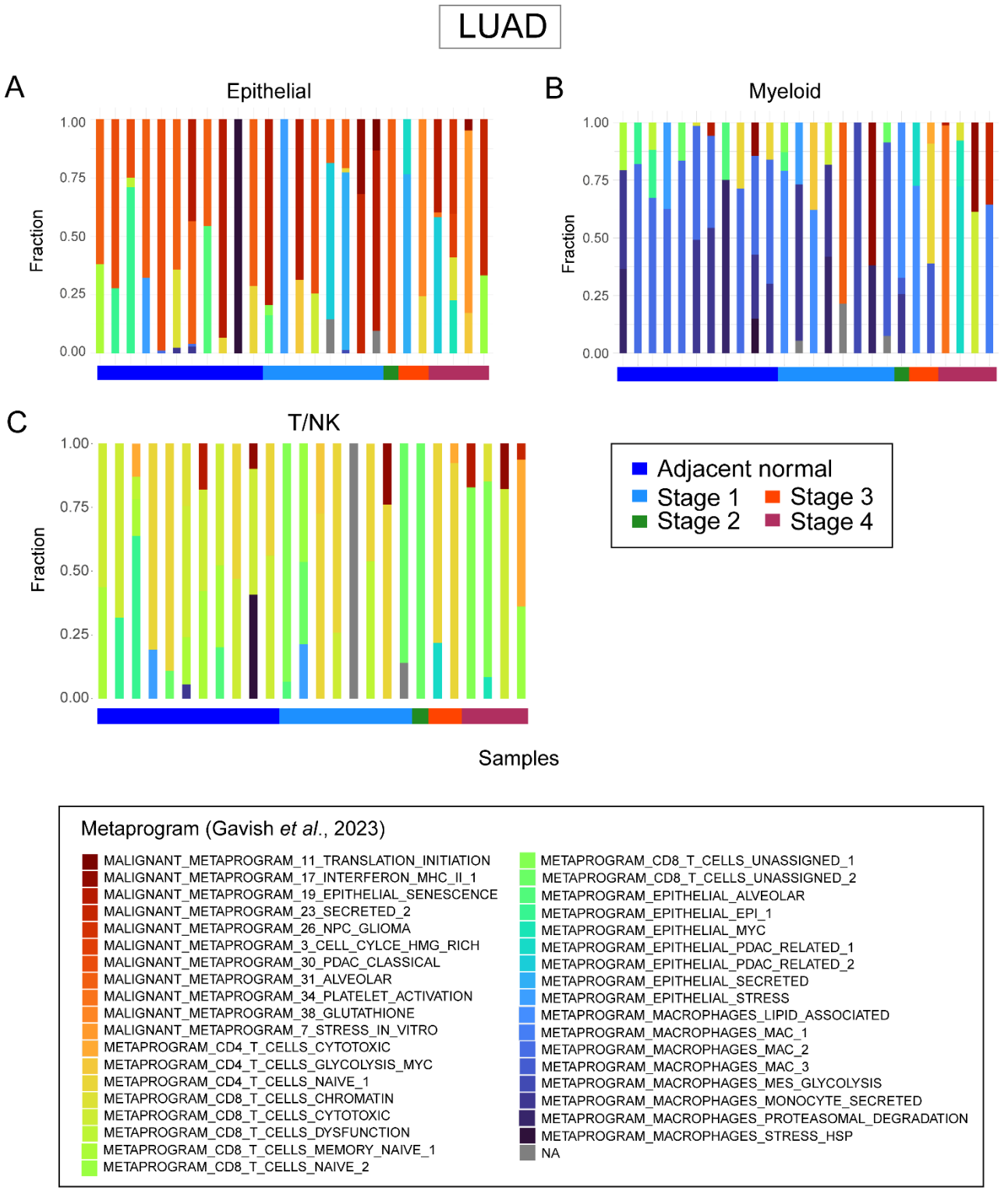

Figure S3 – **Changes in top-ranked functional programs with cancer progression in LUAD.** The stacked bar plots show the top-ranked programs for a cell type in individual samples. As a cell type within a sample could be present in multiple archetypes, there were multiple top-ranked programs seen for a cell type across samples. The stacked bar plots in (A), (B) and (C) show the top-ranked program seen in epithelial, myeloid and T/NK cells in LUAD samples, respectively. The colour bars at below the stacked bar plots show stages for cancer samples along with adjacent normal samples.

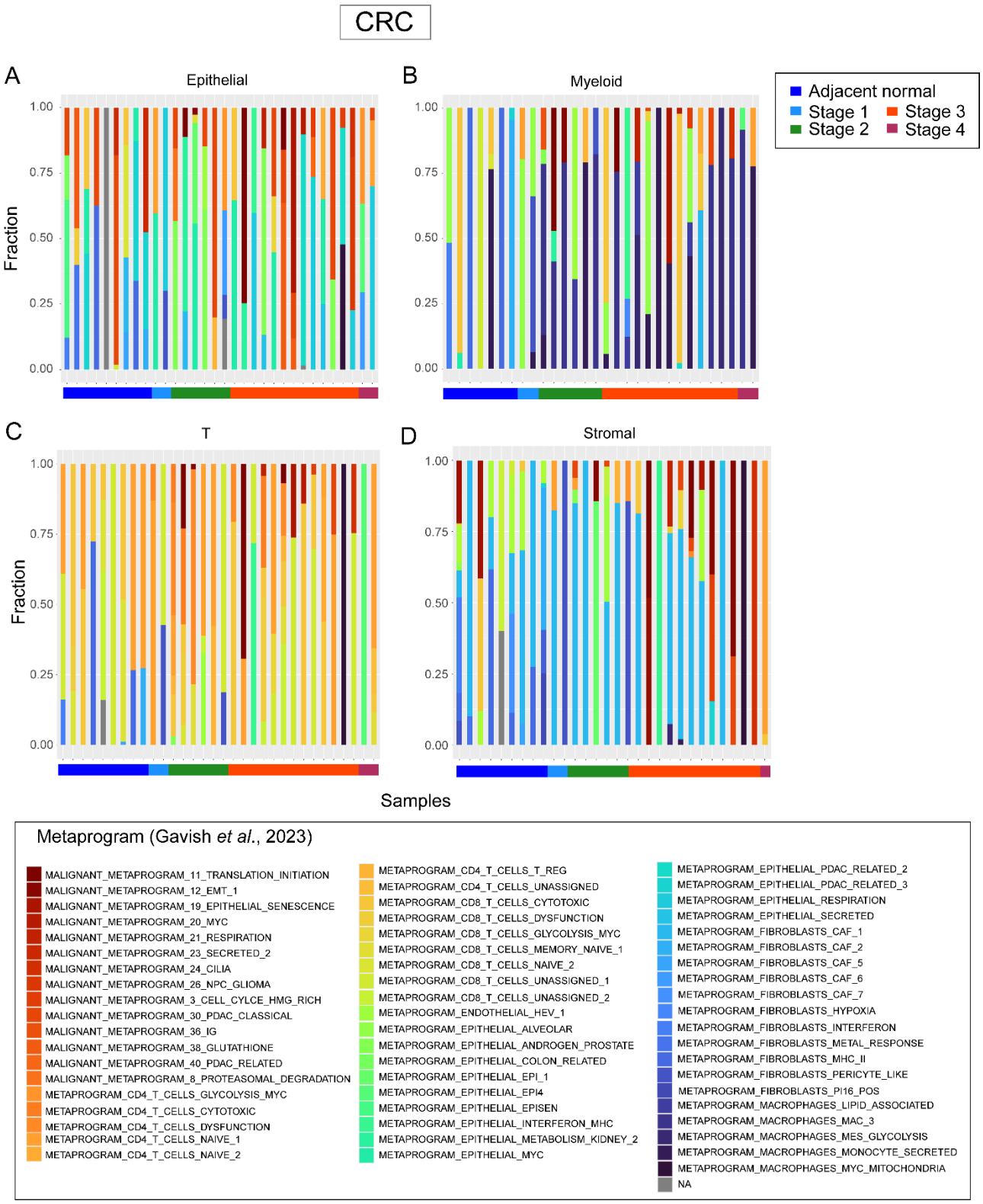

Figure S4 – **Changes in top-ranked functional programs with cancer progression in CRC.** The stacked bar plots show the top-ranked programs for a cell type in individual samples. As a cell type within a sample could be present in multiple archetypes, there were multiple top-ranked programs seen for a cell type across samples. The stacked bar plots in (A), (B), (C) and (D) show the top-ranked program seen in epithelial, myeloid, T and stromal cells in CRC samples, respectively. The colour bars at below the stacked bar plots show stages for cancer samples along with adjacent normal samples.

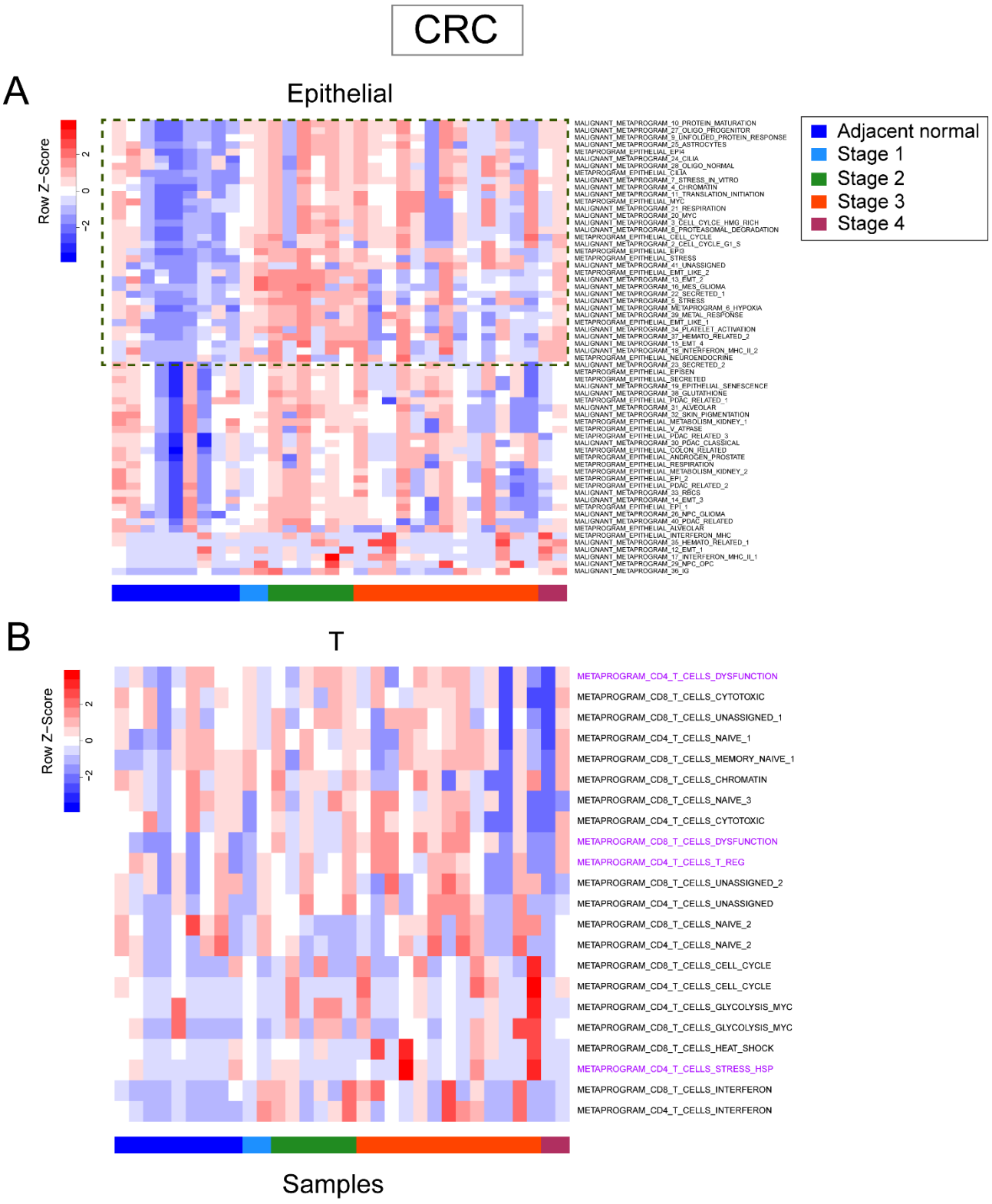

Figure S5 – **Changes in functional utilization in epithelial and T cells in CRC samples with cancer progression. (A)** Heatmap showing changes in utilization of functional metaprograms (as defined by Gavish et al., 2023) in epithelial cells in CRC samples with cancer progression. The metaprograms highlighted in red show some malignant metaprograms whose utilization increased with cancer progression. **(G)** Heatmap showing changes in utilization of functional metaprograms in T/NK cells in LUAD samples with cancer progression. The metaprograms in purple show functions associated with regulatory T cells and T cell exhaustion. The colour bars below the heatmaps show stages for cancer samples along with adjacent normal samples.

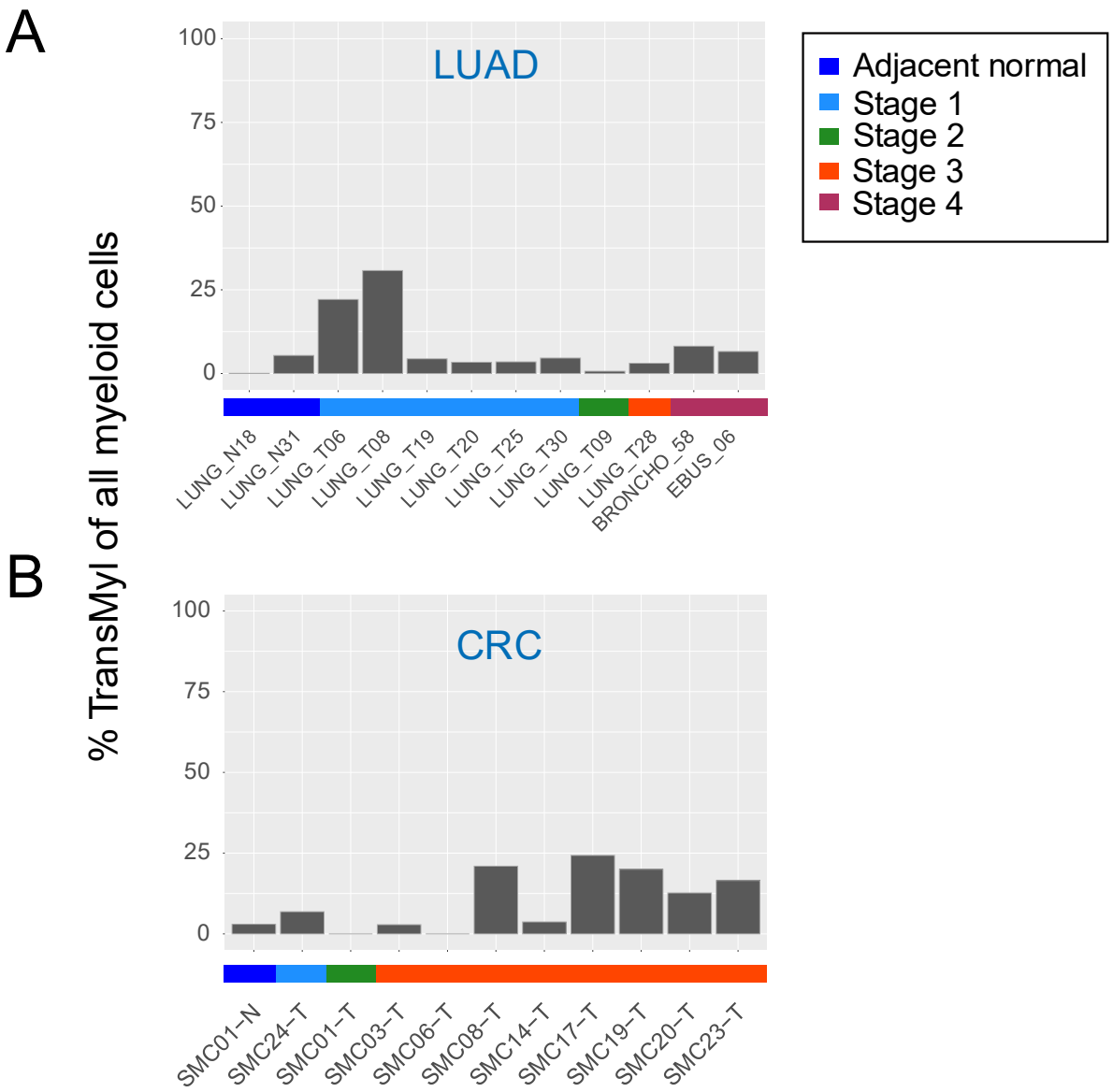

Figure S6 – Percentage of transitory myeloid cells out of total myeloid population in samples that harboured transitory myeloid-T/NK (or T) functional states. (A) Percentage of transitory myeloid cells out of the total myeloid population in individual LUAD samples that harboured transitory myeloid-T/NK functional states. **(B)** Percentage of transitory myeloid cells out of the total myeloid population in individual CRC samples that harboured transitory myeloid-T functional states. The colour bars below the bar plots show stages for cancer samples along with adjacent normal samples.

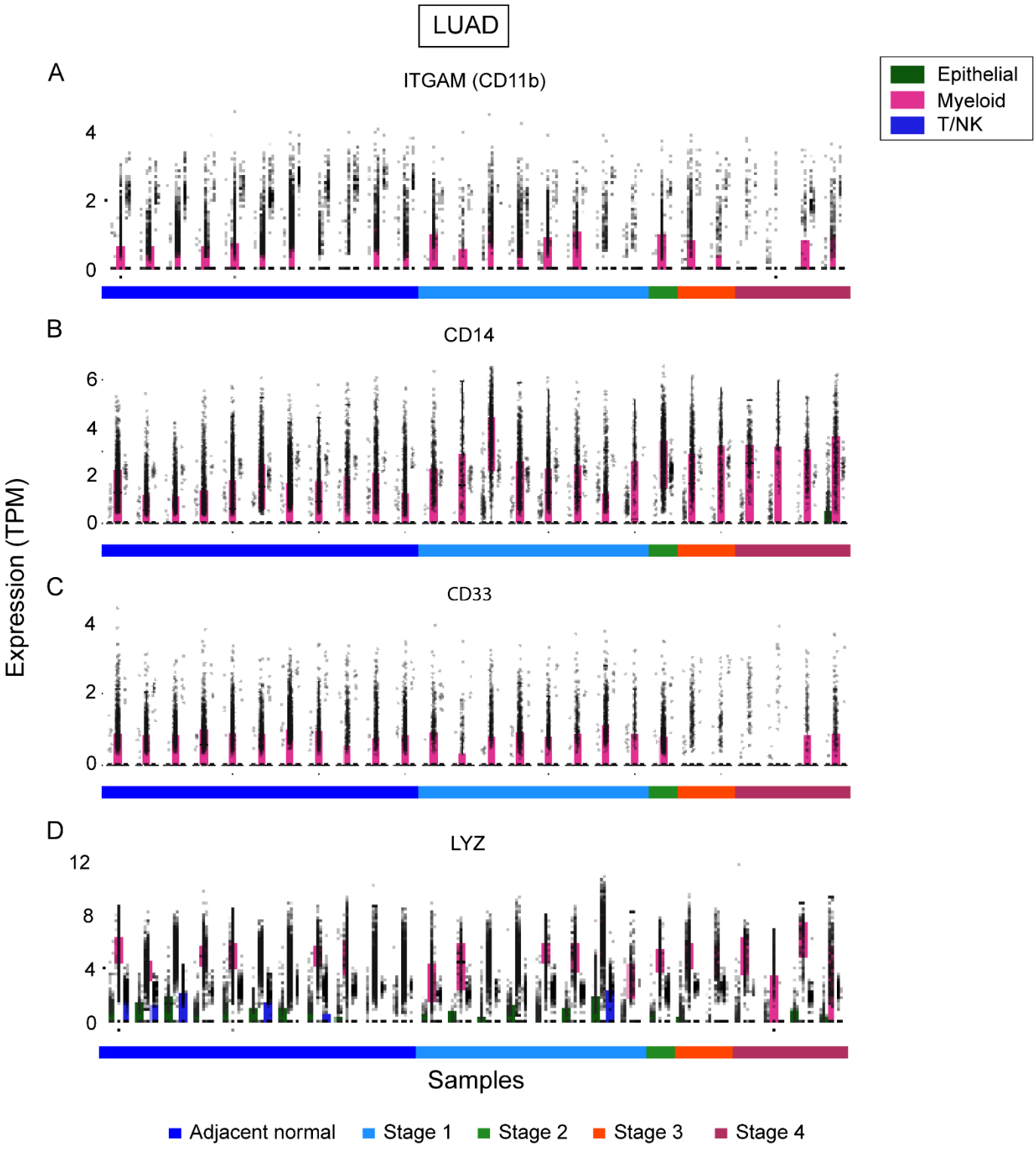

Figure S7 – Expression of ITGAM (CD11b), CD14, CD33 and LYZ genes in adjacent normal and tumor samples in LUAD. Bar plots showing expression of ITGAM (A), CD14 (B), CD33 (C) and LYZ (D) genes in epithelial, myeloid and T/NK cells in individual samples in LUAD. The colour bars below the bar plots show stages for cancer samples along with adjacent normal samples.

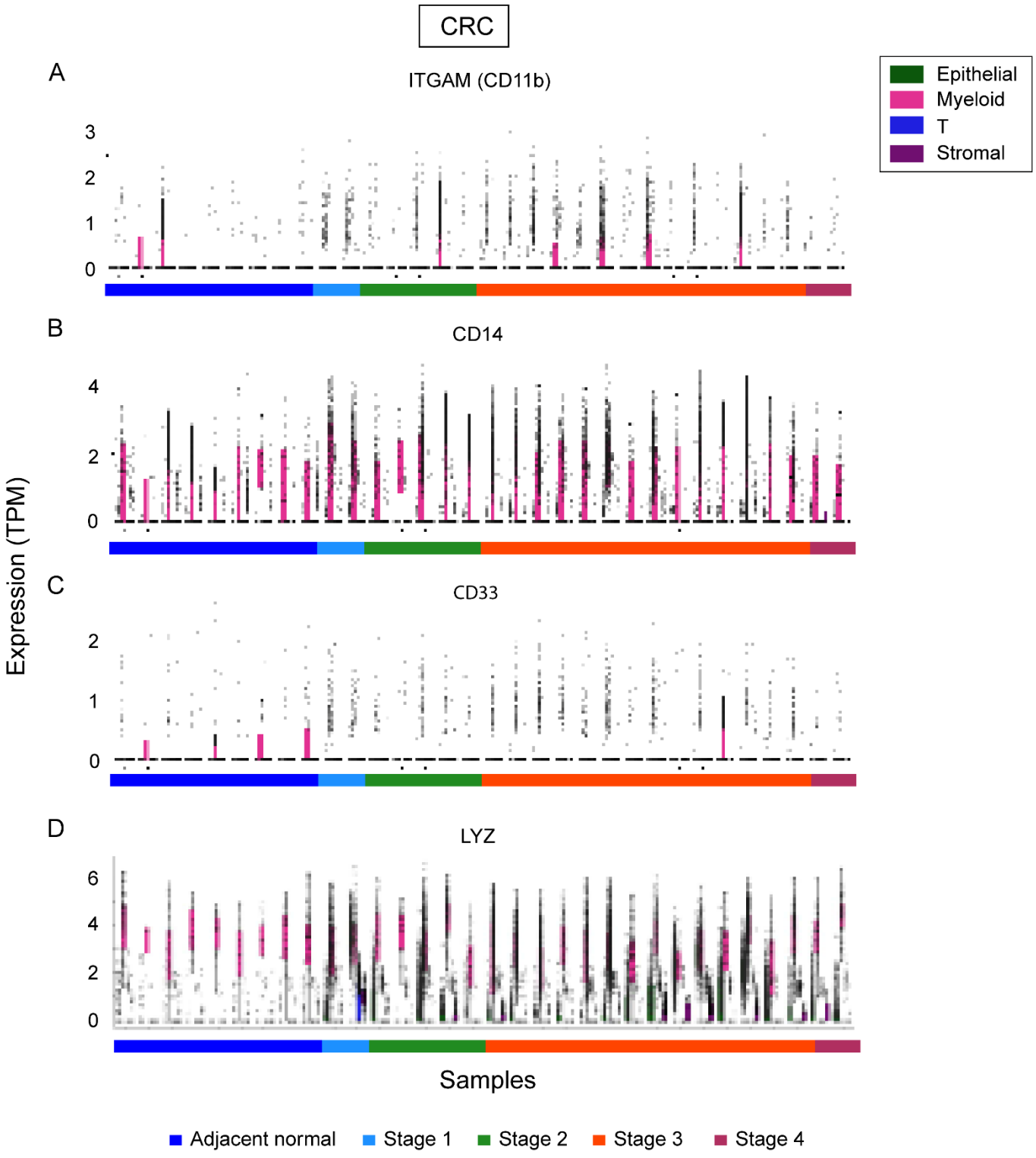

Figure S8 – **Expression of ITGAM (CD11b), CD14, CD33 and LYZ genes in adjacent normal and tumor samples in CRC.** Bar plots showing expression of ITGAM (A), CD14 (B), CD33 (C) and LYZ (D) genes in epithelial, myeloid, T and stromal cells in individual samples in CRC. The colour bars below the bar plots show stages for cancer samples along with adjacent normal samples.

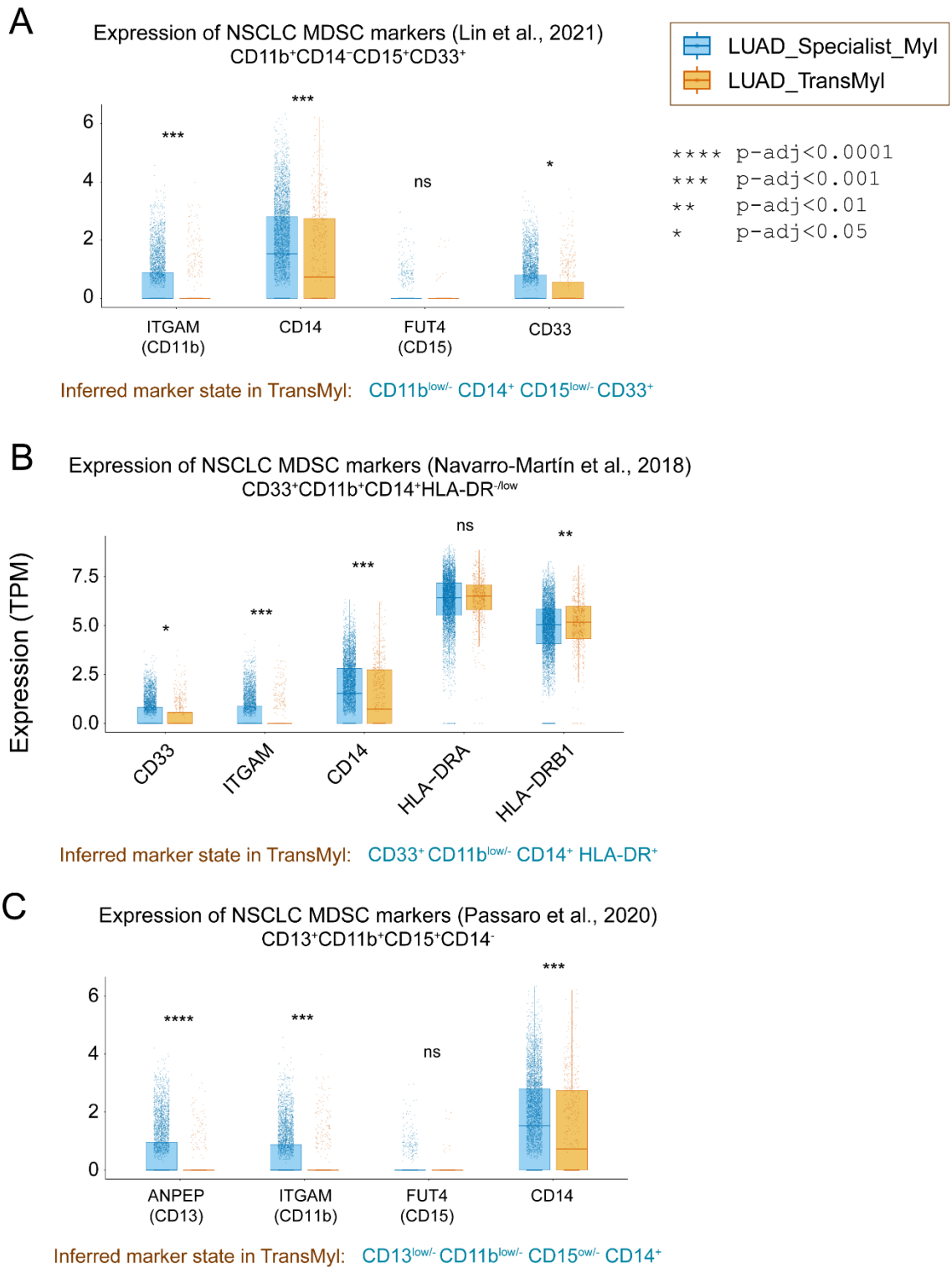

Figure S9 – Test for expression of NSCLC-specific MDSC markers in transitory myeloid cells in LUAD samples. Boxplots show expression of genes associated with NSCLC-specific MDSC markers described in Lin et al. (2021) (A), Navarro-Martin et al. (2018) (B) and Passaro et al. (2020) (C) in transitory and specialist myeloid cells in LUAD samples.

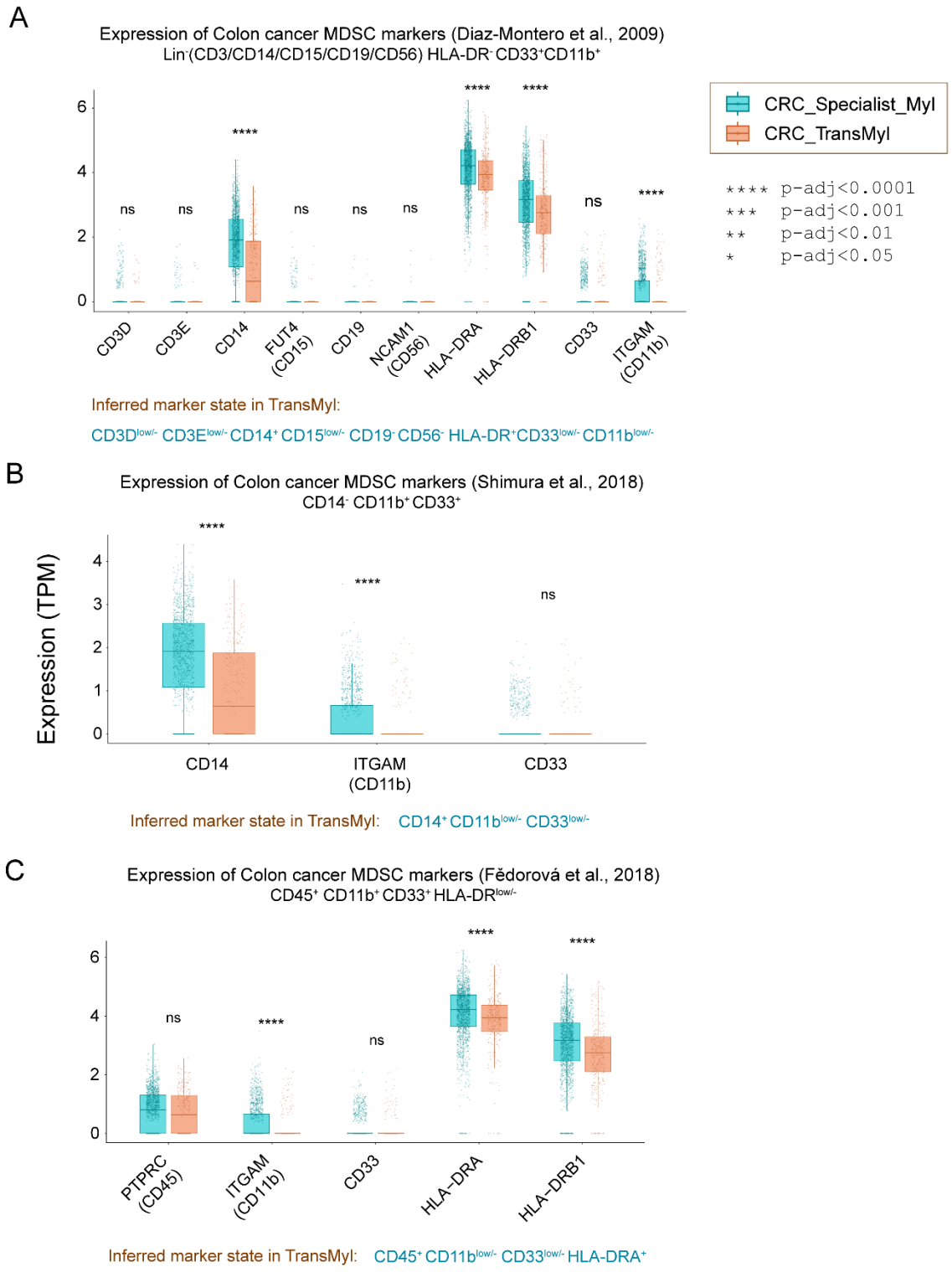

Figure S10 – Test for expression of colon cancer-specific MDSC markers in transitory and specialist myeloid cells in CRC samples. Boxplots show expression of genes associated with NSCLC-specific MDSC markers described in Diaz-Montero et al. (2009) (A), Shimura et al. (2018) (B) and Fědorová et al. (2018) (C) in transitory and specialist myeloid cells in CRC samples.

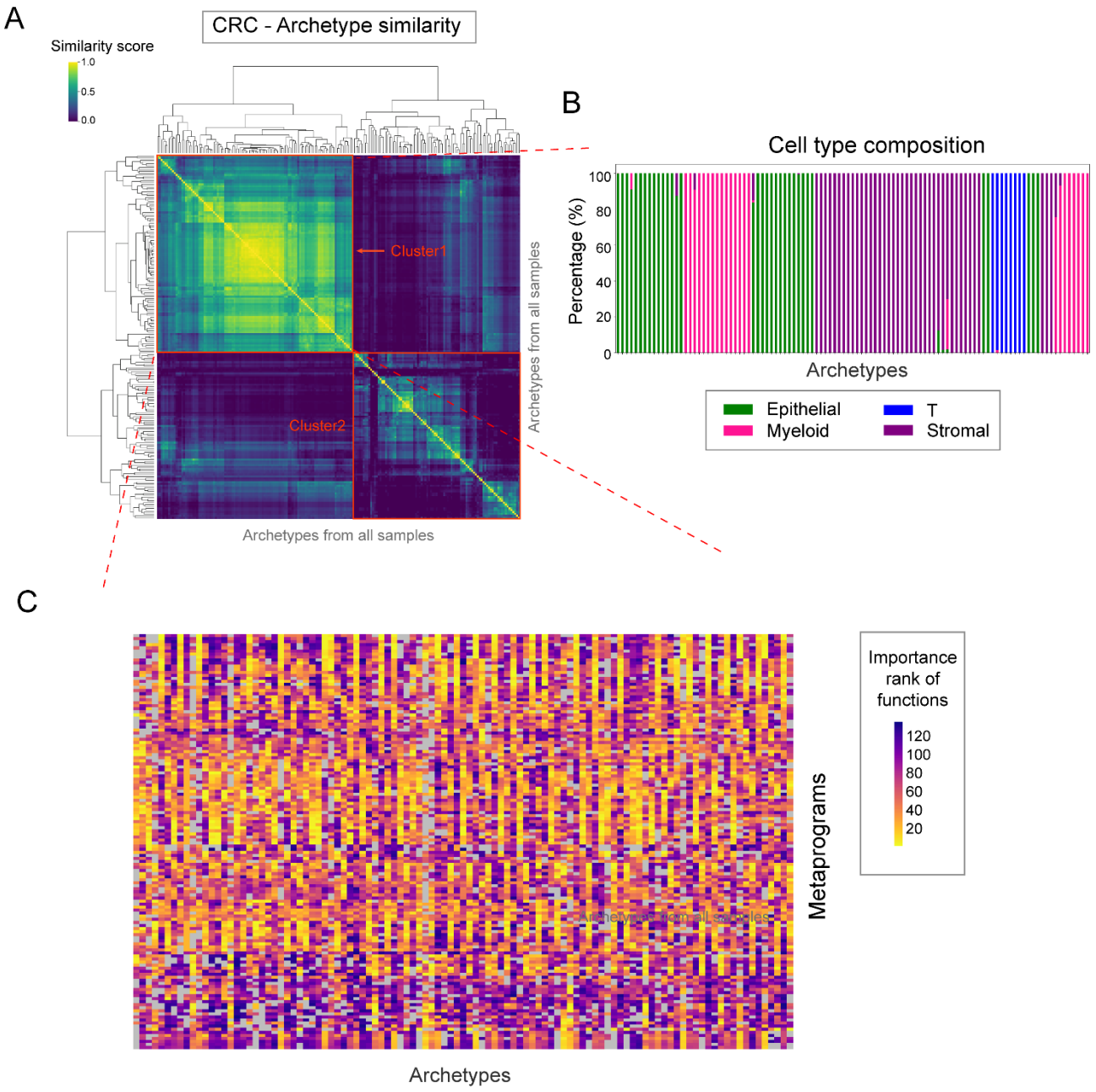

Figure S11 – **Different cell types converge to similar functional profiles in CRC samples. (A)** Heatmap depicting pairwise similarity scores between archetypes of all samples obtained after alignment.

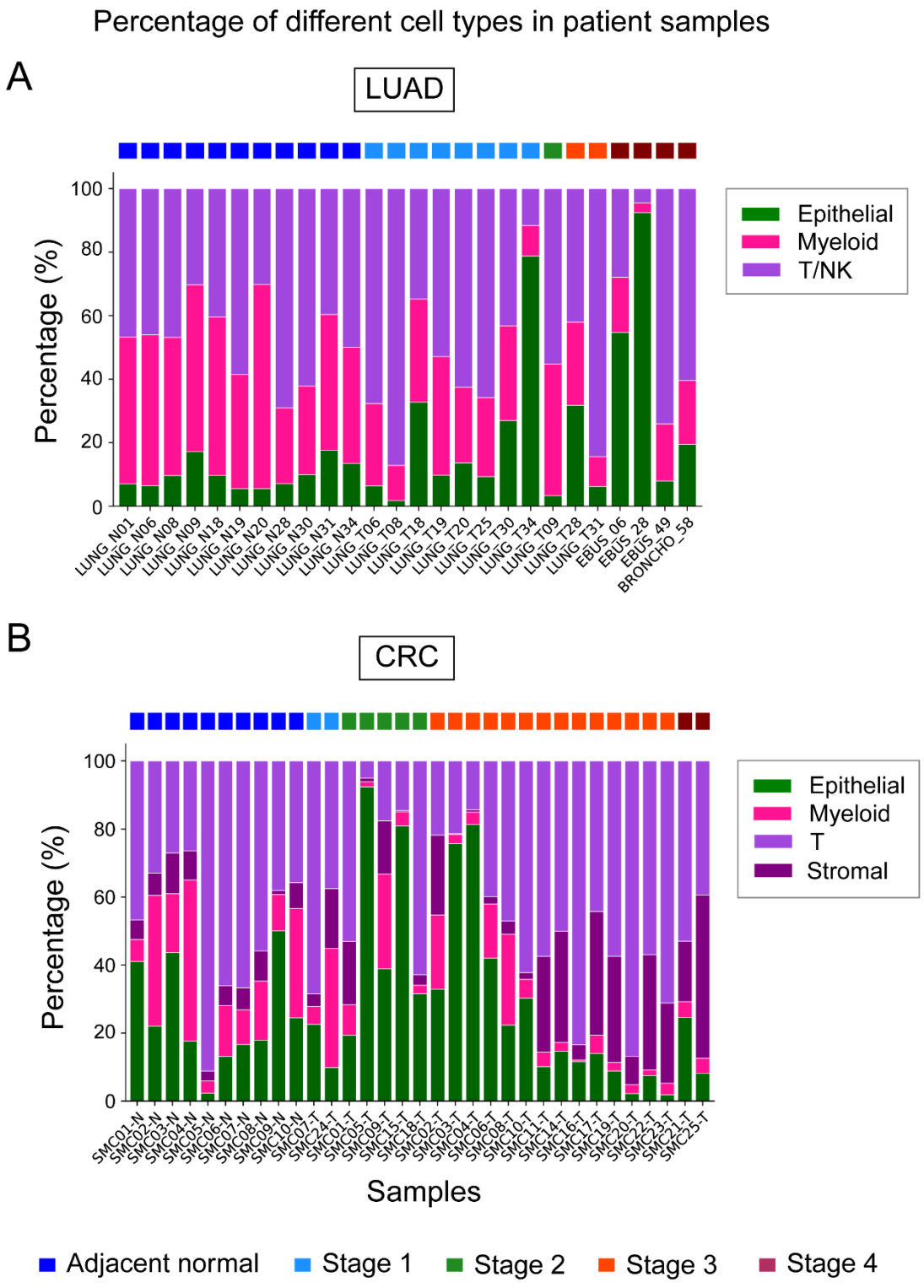

Figure S12 – **Percentage of cell types for patient samples in LUAD and CRC.** **(A)** Stacked bar plot showing percentages of epithelial, myeloid and T/NK cells present in the LUAD samples. **(B)** Stacked bar plot showing percentages of epithelial, myeloid, T and stromal cells present in the CRC samples. The colour bars on top of the bar plots show stages for cancer samples along with adjacent normal samples.

Sample-wise percentage of functional specialist, generalist and transitory cells

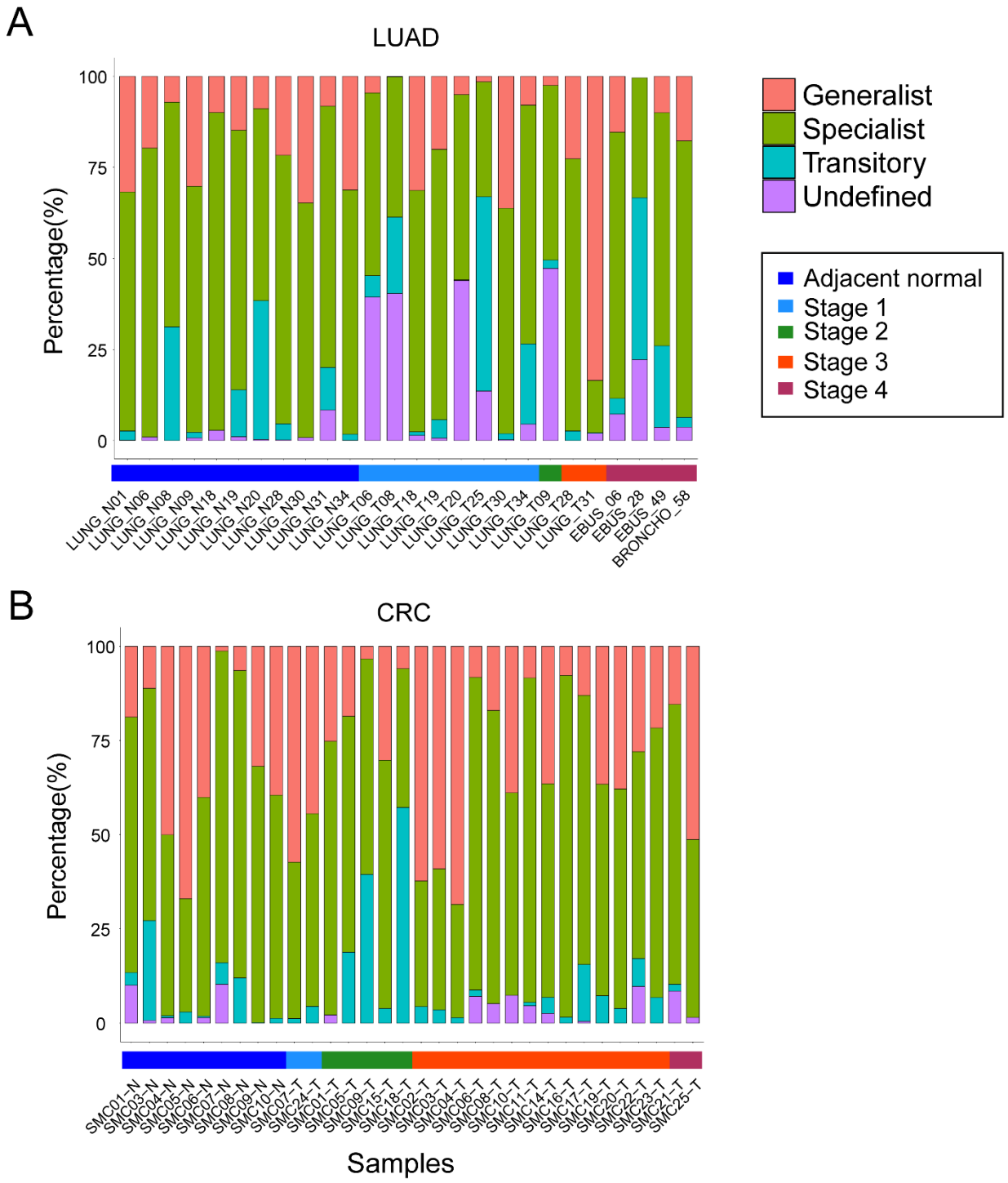

Figure S13 – **Sample-wise percentage of functional specialist, generalist and transitory cells in LUAD and CRC samples. (A)** Percentage of specialist, generalist and transitory cells identified in LUAD samples after Pareto front and cell-density based stability analysis. **(B)** Percentage of specialist, generalist and transitory cells identified in CRC samples after Pareto front and cell-density based stability analysis. The colour bars below the stacked bar plots show stages for cancer samples along with adjacent normal samples.

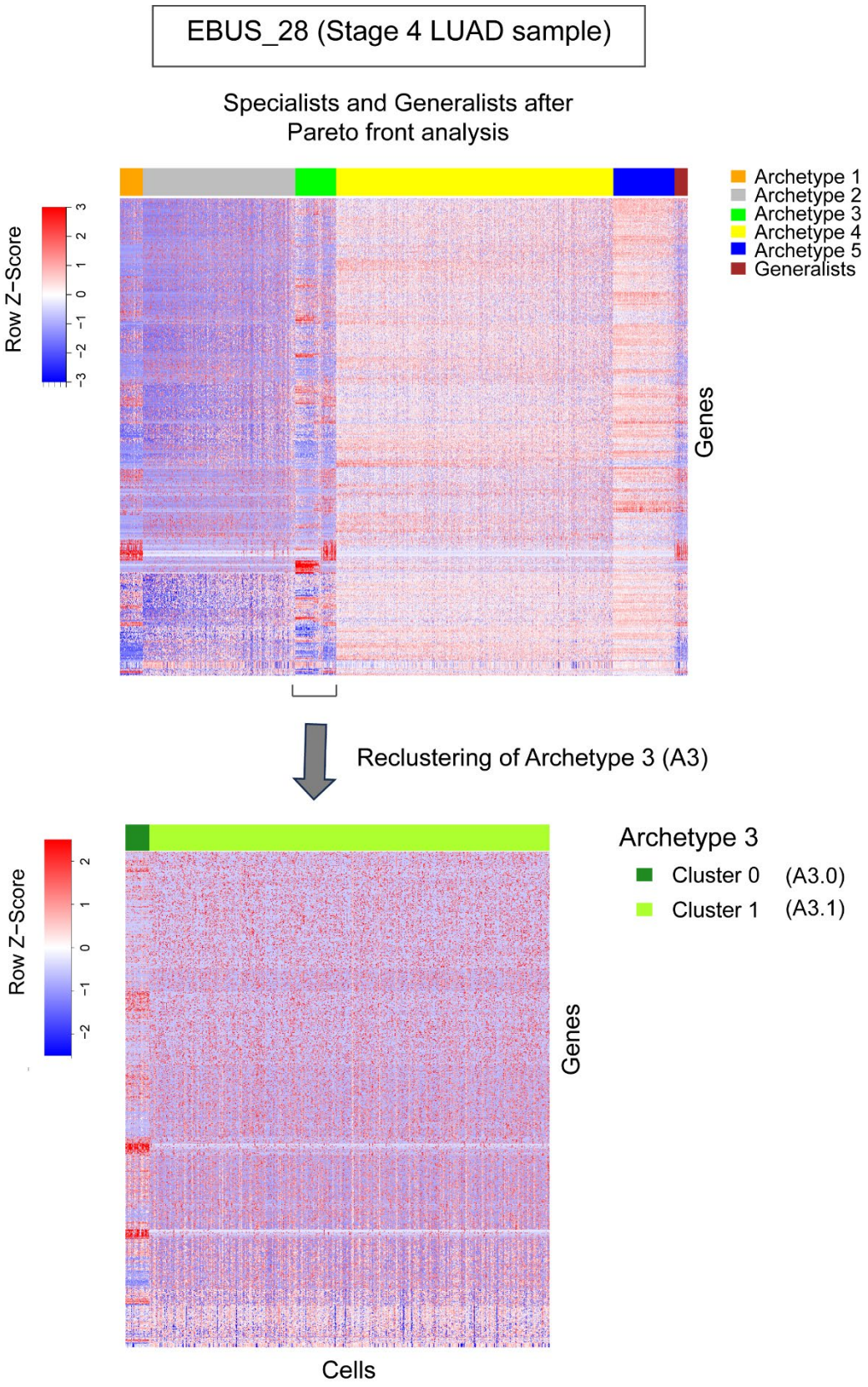

Figure S14 – **An example of refinement of functional specialists through re-clustering.** The archetype 3 of a stage 4 LUAD sample (EBUS_28) contained two sets of cells with distinct expression signatures. There were separated into two functional specialists after re-clustering.

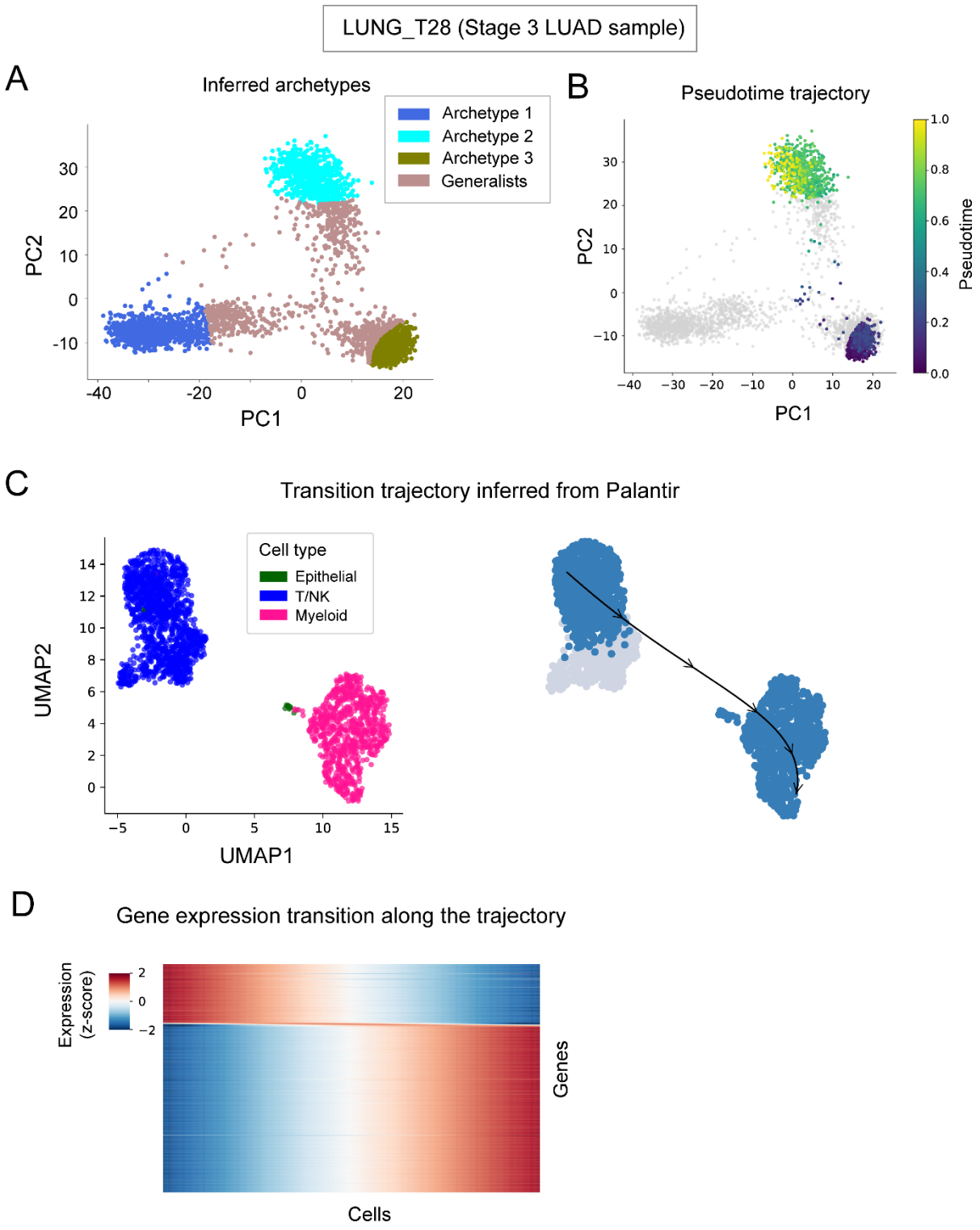

Figure S15 – **Cell state transition trajectory analysis and gene expression change along the trajectory. (A)** The inferred archetypes after Pareto front analysis in a stage 3 LUAD sample (LUNG_T28). **(B)** Pseudo time inferred for the cells in the same sample using *Palantir*. **(C)** A cell state transition trajectory inferred from *Palantir* connecting T/NK and myeloid cell states. **(D)** Gene expression changes along the trajectory as inferred from *Palantir.*

Supplementary references

Diaz-Montero CM, Salem ML, Nishimura MI et al. (2009). Increased circulating myeloid-derived suppressor cells correlate with clinical cancer stage, metastatic tumor burden, and doxorubicin-cyclophosphamide chemotherapy. Cancer Immunol Immunother. 2009 Jan;58(1):49-59. doi: 10.1007/s00262-008-0523-4.

Fědorová L, Pilátová K, Selingerová I et al. (2018). Circulating Myeloid-Derived Suppressor Cell Subsets in Patients with Colorectal Cancer - Exploratory Analysis of Their Biomarker Potential. Klin Onkol. 2018 Winter;31(Suppl 2):88-92.

Lin S, Zhang X, Huang G et al. (2021) Myeloid-derived suppressor cells promote lung cancer metastasis by CCL11 to activate ERK and AKT signaling and induce epithelial-mesenchymal transition in tumor cells. Oncogene. 2021 Feb;40(8):1476-1489. doi: 10.1038/s41388-020-01605-4.

Navarro-Martín A, Galiana IL, Berenguer Frances MA et al. (2018). Preliminary Study of the Effect of Stereotactic Body Radiotherapy (SBRT) on the Immune System in Lung Cancer Patients Unfit for Surgery: Immunophenotyping Analysis. Int J Mol Sci. 2018 Dec 9;19(12):3963. doi: 10.3390/ijms19123963.

Passaro et al. (2020). Gr-MDSC-linked asset as a potential immune biomarker in pretreated NSCLC receiving nivolumab as second-line therapy. Clin Transl Oncol. 2020 Apr;22(4):603-611. doi: 10.1007/s12094-019-02166-z.

Shimura T, Shibata M, Gonda K et al. (2018). Prognostic impact of preoperative lymphocyte-to-monocyte ratio in patients with colorectal cancer with special reference to myeloid-derived suppressor cells. Fukushima J Med Sci. 2018 Aug 29;64(2):64-72. doi: 10.5387/fms.2018-10.