



HAL
open science

deconvPDAC : a quantifier of PDAC tumor cellular heterogeneity

Yasmina Kermezli, Lucie Lamothe, Jerome Cros, Magali Richard, Yuna Blum

► To cite this version:

Yasmina Kermezli, Lucie Lamothe, Jerome Cros, Magali Richard, Yuna Blum. deconvPDAC : a quantifier of PDAC tumor cellular heterogeneity. Journées Ouvertes en Biologie, Informatique et Mathématiques (JOBIM 2024), Jun 2024, Toulouse, France. . hal-04669313

HAL Id: hal-04669313

<https://hal.science/hal-04669313v1>

Submitted on 8 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

deconvPDAC : a quantifier of PDAC tumor cellular heterogeneity

Yasmina Kermezli¹, Lucie Lamothe^{1,+}, Jerome Cros², Magali Richard^{1,*} and Yuna Blum^{3,*}



¹ Laboratory TIMC-IMAG, UMR5525, Univ.Grenoble-Alpes, CNRS,38000, Grenoble,France;

² Beaujon Hospital, Dpt of Pathology, Univ. Paris-INSERM U1149, Paris, France;

³ IGDR, UMR 6290, CNRS, Université de Rennes 1, Rennes, France

*corr: magali.richard@univ-grenoble-alpes.fr & yuna.blum@univ.rennes1.fr; + presenting author



Introduction

- The pancreatic ductal adenocarcinoma (PDAC) **intra-tumor heterogeneity** is a major pathological feature that can confer aggressiveness and chemoresistance
- Accuracy of bulk deconvolution algorithms to infer tumor composition is currently limited.
- In this study, we provide a single-cell quantifier of PDAC tumor cell heterogeneity including:
 - A list of **robust integrative cell-type markers**.
 - Guidelines to **assign single cell identities**.
 - A **benchmark of deconvolution tools** applied to PDAC bulk transcriptome deconvolution tools.

deconvPDAC Pipeline

In this study, we characterized a pool of **796,000 single cell RNA-Seq** from normal and tumoral pancreas from **12 different published studies**.

(STEP 1) Identification of 15 PDAC cell types across several sc-datasets

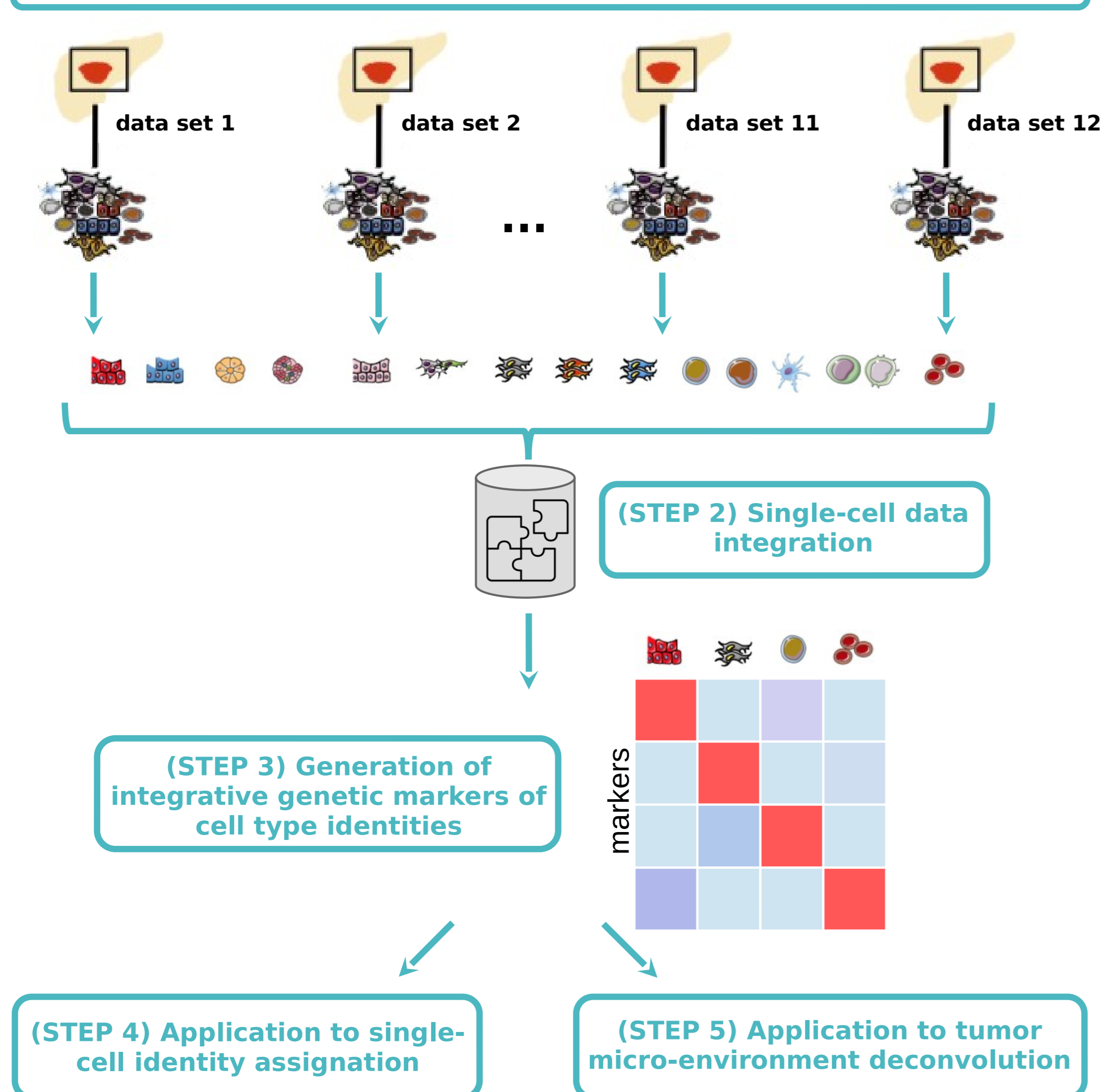


Figure 1: Pipeline of the project. Steps 1 to 3 lead to identification of robust integrative markers for PDAC cell types. Step 4 and 5 are potential applications of this new marker set.

2024 update : the project is currently being remade using extended datasets. Following figures were realized in 2021 using only 4 studies

Step 1 & 2 : Cell identification & integration

Single cell type assignment based on a dedicated **literature gene markers curation** and a subsequent analysis of a recent PDAC scRNA-seq data-sets. Identification of **15 independent cell types**.

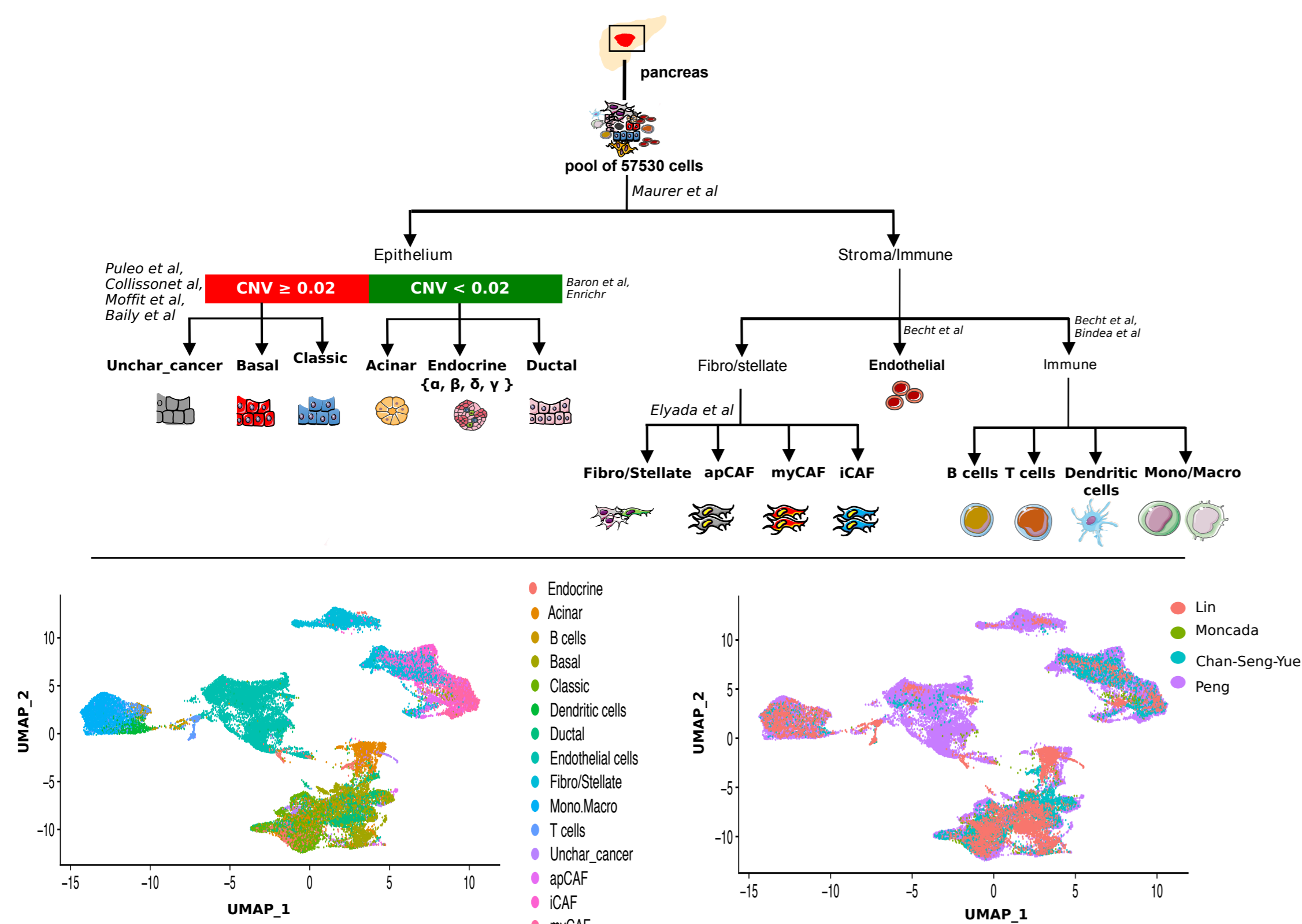


Figure 2: Cell-types identification and integration. Top panel: decision tree for cell type assignment. Bottom panel: Clustering of single cells data after integration of all studies together. Left panel: UMAP of all identified cell-types, Right panel: UMAP of cells colored by their study of origin.

Step 3: Integrative cell-type specific markers

Extraction of **robust integrative gene markers**.

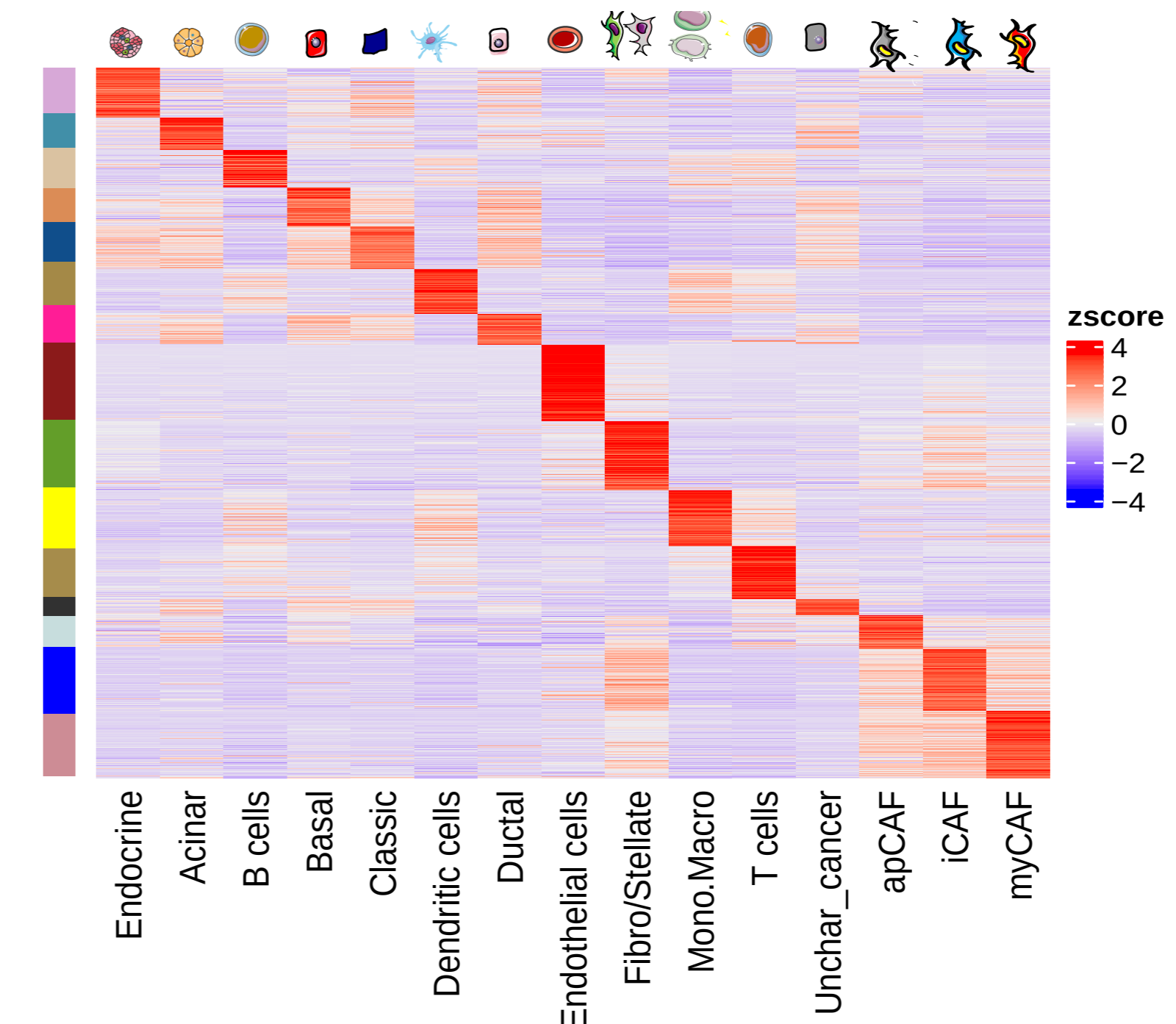


Figure 3: Robust integrative cell-type markers. Heatmap of the newly generated cell-types markers across averaged genetic profiles of integrated cell types. Markers were identified by applying the SCRAN workflow (Lun et al, 2016) after bootstrap sampling.

Step 4: Single cell identity assignment

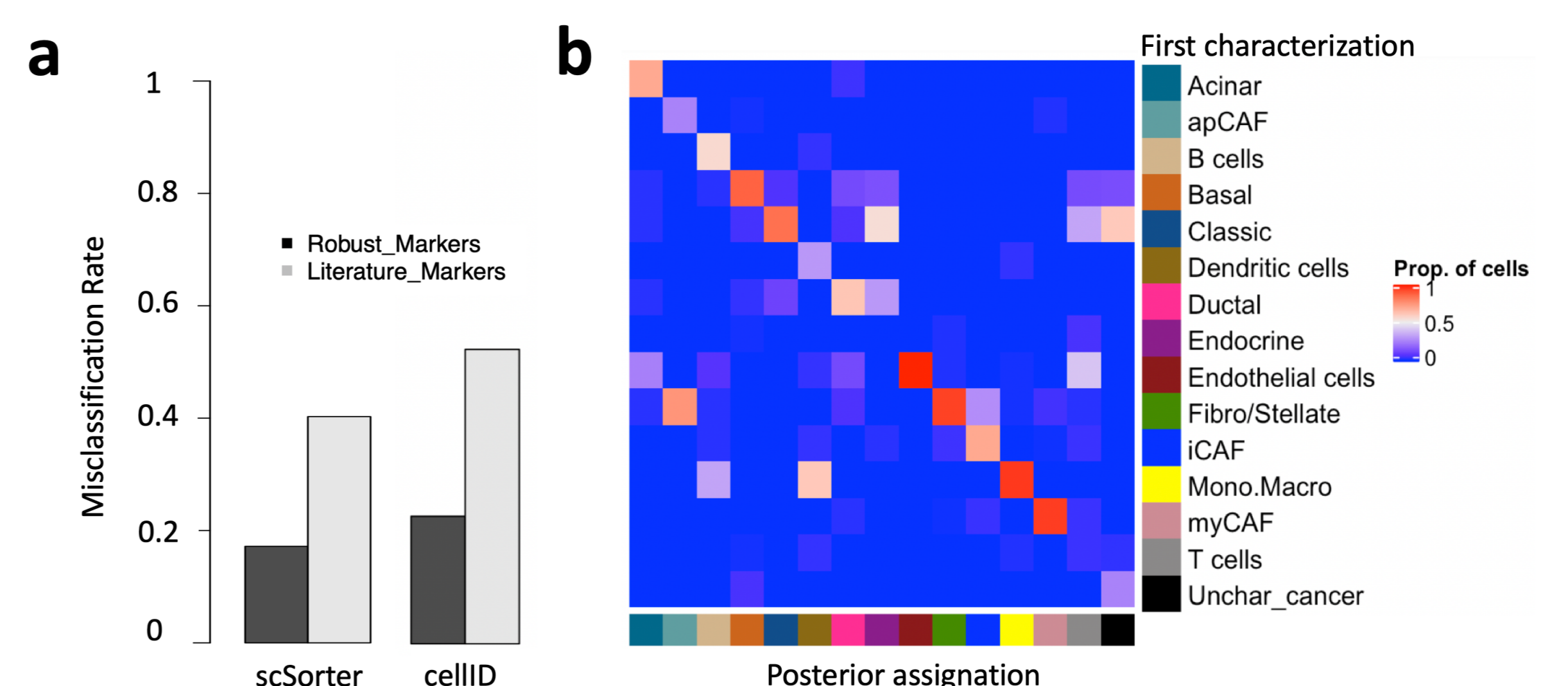


Figure 4: Posterior cell assignment on Peng's dataset. a) Misclassification rate using our robust list of markers or the markers from the literature and the tools scSorter (Guo et al, 2021) or cellID (Cortal et al 2021). b) Representation of the contingency table showing the proportion of true and false discoveries for each cell population using cellID and our robust markers.

Step 5 : Benchmark of deconvolution using the new single-cell references

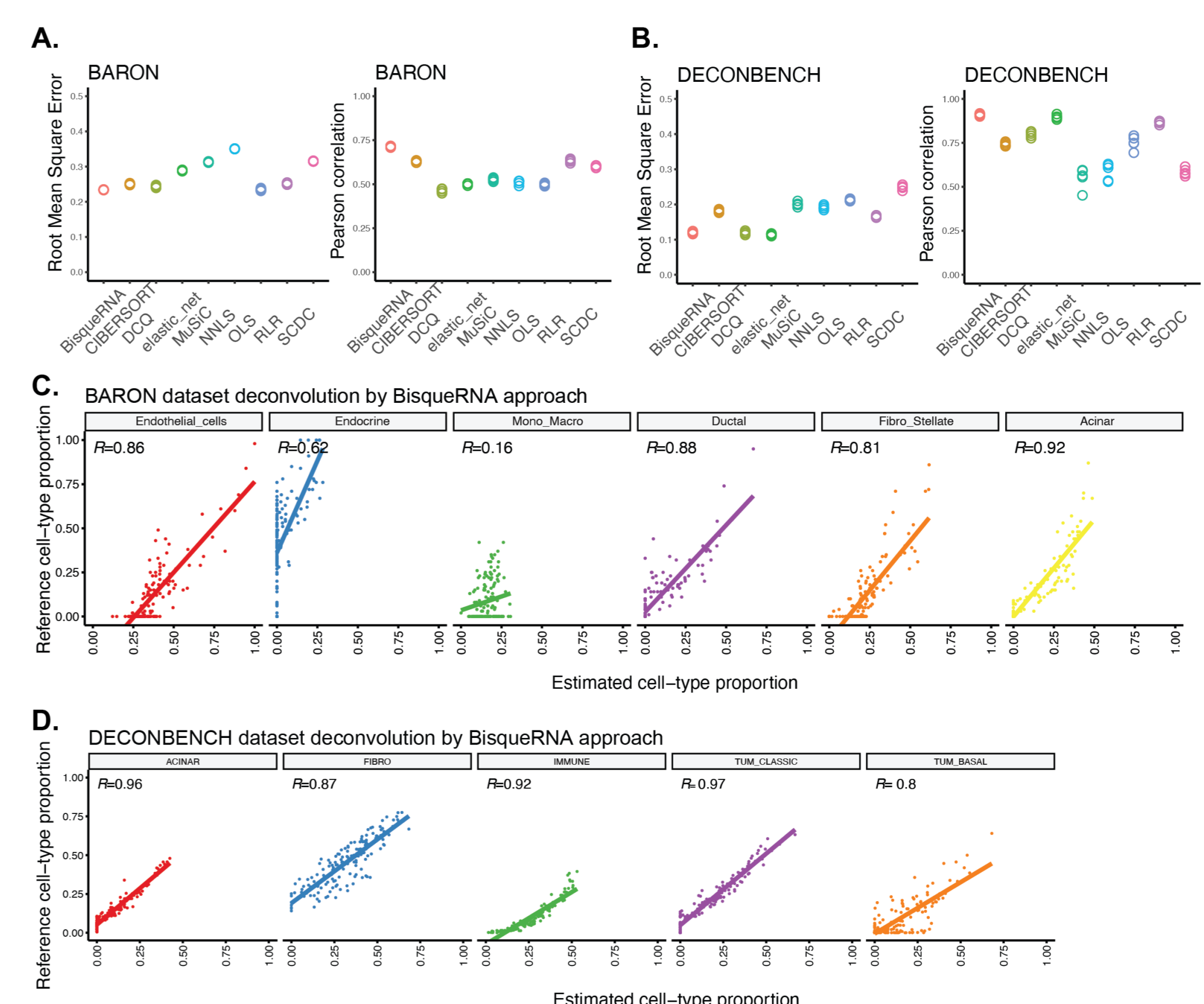


Figure 5: Benchmark of tumor composition estimation. Root mean square error and Pearson correlation of deconvolution methods for 5 independent simulations on (A) pseudo-bulk healthy pancreas (baron dataset) and (B) on in silico simulation of cancer pancreas (deconbench dataset). Correlations between real cell-types proportion and BisqueRNA estimated cell-type proportion for (C) pseudo-bulk healthy pancreas and for (D) in silico simulation of cancer pancreas.

Conclusion

- An **integrative robust set of PDAC cell-type** specific gene markers.
- The use of these signatures can help us understanding and finely quantifying tumor composition and pave the way to more efficient precision medicine.