肺がん組織ブロックの細胞型組成と細胞状態（Visium HD 25 µm、CU071）

takara_demo · 11,684 bins · 2026-08-13T22:11:04Z

要旨

11,684個の組織ビンを11の細胞集団に分類し、すべての集団に細胞型を付与した。内訳は、扁平上皮4,936ビン（42.2%）、形質細胞1,501ビン（12.8%）、リンパ球混合集団1,362ビン（11.7%）、線維芽細胞／間質1,284ビン（11.0%）、骨髄系／マクロファージ1,233ビン（10.6%）、内皮429ビン（3.7%）、気管支粘膜下腺319ビン（2.7%）、軟骨235ビン（2.0%）、繊毛性気道表面上皮216ビン（1.8%）、形質細胞様樹状細胞117ビン（1.0%）、マスト細胞52ビン（0.4%）であった。本切片では、粘膜下腺と軟骨板を保持した気管支壁が認められ、その周囲を扁平上皮からなる腫瘤と高密度の形質細胞浸潤が取り囲んでいた。

上皮区画（4,936ビン）は10のサブクラスタに分けられた。各サブクラスタを最も高いスコアを示した状態に割り当てると、増殖性2,708ビン（上皮区画の54.9%）、角化型1,340ビン（27.1%）、EMT様814ビン（16.5%）、基底細胞様74ビン（1.5%）となった。このうち、離散した集団として分離したのは基底細胞様のみであり、残る3つの状態は上皮区画内で連続的に変化しており、明確に分離可能な部分集団を形成しなかった。そのため、これらは細胞クラスではなく、プログラムの強度として報告する。なお、EMT様として集計された814ビンは、実際には上皮細胞の状態を反映したものではなく、間質細胞およびリンパ球の混入によるものであった（EMT様に関する節）。

リンパ球区画では、CD8Aが1,362ビン中189ビン（13.9%）で検出された。この189ビンのうち、細胞傷害性パネルに含まれる遺伝子を2つ以上検出したビンは81%、疲弊パネルの遺伝子を2つ以上検出したビンは40%、両方を満たしたビンは38%であった。疲弊を示すビンのほぼすべてが細胞傷害性も示しており、両プログラムは独立した集団として分離しなかった。また、CD8特異的な部分集団も分離されず、この区画を特徴づける主要な軸は、細胞傷害性と疲弊の違いではなく細胞周期であった。

制約

本データはVisium HDによる25 µm解像度の空間トランスクリプトームデータであり、解離した単一細胞データではない。解析単位は25 µm四方の組織ビンであり、1つのビンに複数の細胞が含まれる場合がある。したがって、本報告書に示す百分率はすべて「該当する遺伝子が検出されたビンの割合」を示すものであり、細胞の割合を示すものではない。T細胞を含め、いかなる細胞集団の割合としても解釈できない。この細胞の混在は単なる解析上の制

約ではなく、本報告書における重要な所見であるEMT様シグナルの主要因でもある（EMT様に関する節）。

本データは、1名の患者から得られた1個の組織ブロックに由来する2枚のキャプチャ切片から構成される。いずれも同一ブロック由来であり、生物学的反復は存在しない。したがって、本解析で得られた知見をこの1組織ブロックを超えて一般化することはできない。

本解析では悪性度を評価しておらず、本納品物のみから悪性度を評価することもできない。解析パイプラインにはコピー数推定が含まれておらず、H&E画像も提供されていない（利用可能なのは座標情報のみである）。したがって、以下に示す上皮に関する所見は、すべて本切片内で観察された転写状態を記述したものであり、腫瘍細胞と正常細胞を判定したものではない。依頼書では上皮区画に「悪性」という属性が前提として設定されているが、その属性を本解析結果から検証することはできない。

上皮の4状態の判定は、上皮区画を10のサブクラスタに分割した結果に基づく。ただし、この細分割は1回のクラスタリングによって得られたものであり、検討したいずれの分割数（2〜10）についても、事前に設定した再現性基準を満たさなかった。再現性基準を緩和した解析や再クラスタリングは行っていない。したがって、上皮4状態の判定は再現性が確認されたクラスタ構造に基づくものではなく、各サブクラスタとマーカー発現との対応関係に基づく判定である。

Materials

ヒト肺腫瘍切片を、25 µm 分解能の空間トランスクリプトーム法で測定したデータ。品質管理を通過した測定単位は 11,684、各単位が 18,129 遺伝子のカウントを保持する。

Software

PACKAGE	VERSION	USED FOR
R	4.4.2	解析環境
Seurat	5.4.0	品質管理・正規化・グラフ構築・クラスタリング・UMAP・差次発現
SeuratObject	5.3.0	オブジェクトモデル
sctransform	0.4.3	分散安定化正規化
glmGamPoi	1.18.0	SCTransform のバックエンド
scanpy	1.10.4	発現オブジェクトの Python 側処理
anndata	0.10.9	h5ad の入出力

PACKAGE	VERSION	USED FOR
CellTypist	1.7.1	参照に基づく細胞型分類（Human_Lung_Atlas.pkl）
decoupler	1.8.0	PanglaoDB／Human Protein Atlas／CancerSEA に対する過剰代表解析
scipy	1.13.1	Kruskal-Wallis 検定および順位統計量
pandas	2.2.3	表形式データの処理
numpy	1.26.4	数値計算
matplotlib	3.11.0	作図

Procedure

qc	検出遺伝子数 200 以上、ミトコンドリア転写産物 20% 以下のビンを保持した。ミトコンドリア遺伝子の存在は必須としていない。ダブレット検出（scrublet）は有効とした。11,852 ビンのうち 11,684 ビンを保持した（除外の内訳は最小遺伝子数で 1.13%、ミトコンドリア割合の上限で 0.27%）。
normalisation	SCTransform（glmGamPoi バックエンド）で正規化した。可変遺伝子は 3,000、共変量の回帰除去は行っていない。
reduction	PCA を 50 成分算出し、SCTransform 残差の 30 次元ノイズ低減線形埋め込みをグラフ構築・クラスタリング・埋め込みに用いた。統合およびバッチ補正は適用していない（batch_key = null、harmony オフ）。生物学的検体が 1 つであるためである。
clustering	30 次元上で共有最近傍グラフを構築した（k = 20、SNN 剪定 0.0667、annoy 距離はユークリッド）。クラスタリングは Louvain（アルゴリズム 1）である。納品分割は 11,684 ビンを 12 コミュニティに固定した。調整版ではクラスタ 2 と 11 が 1 クラスタとなり、11 クラスタとなる。100 ビン未満のクラスタは除去せず、規模を注記して報告する。
embedding	同じ 30 次元埋め込み上で UMAP を計算した（n_neighbors = 15、min.dist = 0.30、コサイン距離）。
annotation	クラスタごとのマーカー遺伝子は、SeuratのFindAllMarkersを用いて同定した。検定にはWilcoxon順位和検定を用い、正のマーカーのみを対象とした（min.pct = 0.10、log2 fold-change threshold = 0.25、SCT assay、各クラスタ最大500ビン）。 細胞型の判定には、性質の異なる3系統の情報を用いた。それぞれ異なる要因で誤判定が生じうるため、結果は個別に評価した。 (i) CellTypistによるアノテーション。 CellTypist 1.7.1の学習済みモデルHuman_Lung_Atlas（61クラス）を各ビンに適用し、ビン単位の予測結果を多数決によってクラスタごとのラベルに集約した。この方法は、発現プロファイル全体を既存の参照データで学習された細胞型と照合するものである。そのため、対象となる細胞型が参照に含まれている場合には有効である一方、参照に存在しない細胞型を正しく判定することはできない。実際、本モデルは健常肺アトラスに基づいており、軟骨細胞のクラスを含まない

ため、軟骨細胞に特徴的な発現が明瞭であっても軟骨細胞というラベルを返すことはできない。
(ii) エンリッチメント解析に基づくアノテーション。 PanglaoDBに収録された164細胞型のマーカー集合を用い、2種類の方法で評価した。試料の臓器による事前の絞り込みは行わず、臓器名をラベルに併記することで、候補となる細胞型をあらかじめ除外しないようにした。1つ目はdecoupler 1.8.0による過剰代表解析（min_n = 5）で、各細胞型のマーカーがクラスタ内にどの程度含まれるかを評価した。2つ目は、クラスタ間で標準化したマーカー遺伝子の平均発現量であり、各細胞型のマーカーが他のクラスタと比較してどの程度高く発現しているかを評価した。両者は評価している内容が異なるため、結果が一致しない場合がある。特に、少数の高特異的な遺伝子によって特徴づけられる細胞型では、過剰代表解析では低く評価される一方、標準化発現量では高く評価されることがある。本解析では軟骨クラスタで実際にこの差が認められた。
2系統の結果が一致しない場合には、本組織で直接観察されたマーカー遺伝子の発現を優先して判定した。参照分類器やマーカーデータベースはいずれも外部データから構築された固定の分類体系であり、そこに含まれない細胞型を適切に表現できないためである。
なお、上皮の4状態、および細胞傷害性と疲弊の区別については、対応する肺がん由来の参照データを用いた分類器がないため、参照分類器による判定は行わず、一貫してマーカー遺伝子およびマーカー集合に基づいて評価した。

The settings above are those that produced the delivered result.

1. クラスタリング

採用した分割では、11,684ビンを12クラスターに分類した。この段階ではクラスタリングのみであり、アノテーション情報を使用していない。クラスタリング最適化の結果、「基準を満たす分割のうち、クラスタ数が5以上で最も細かい分割」を採用した。

品質評価については、1バッチのデータであったため、バッチ補正は実施しなかった。

Fig. 1 に埋め込み上の分割を、Fig. 2 にクラスタごとの規模を示す（100 ビンの報告下限を線で示した）。細胞アノテーションが不要な場合はこの版を用いる。

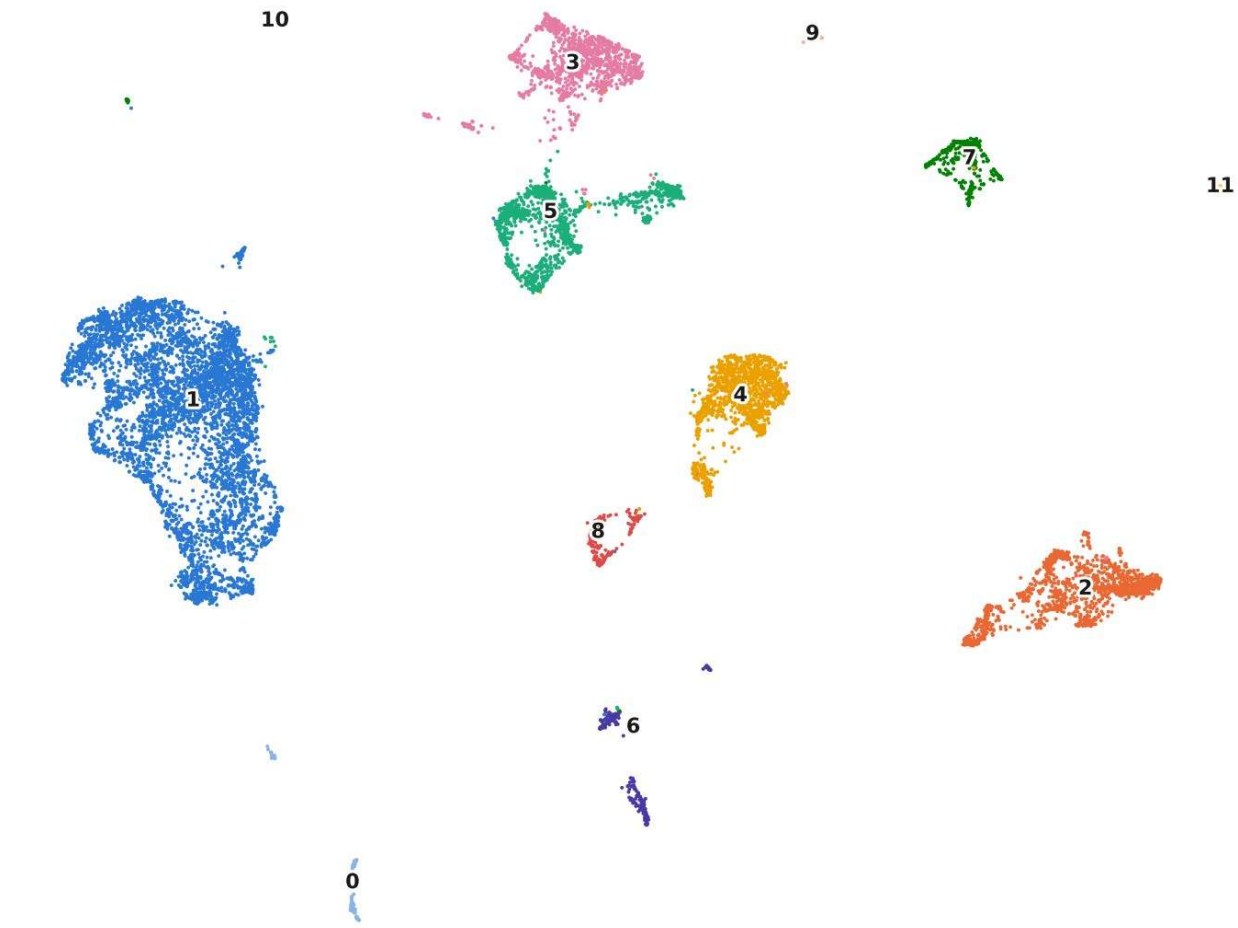


Fig. 1 全捕捉単位の二次元埋め込み。各点が1捕捉単位であり、集団中央の数字はクラスタ番号を示す。埋め込み上の近接は発現の類似性を反映するものであり、組織内の物理的距離ではない。

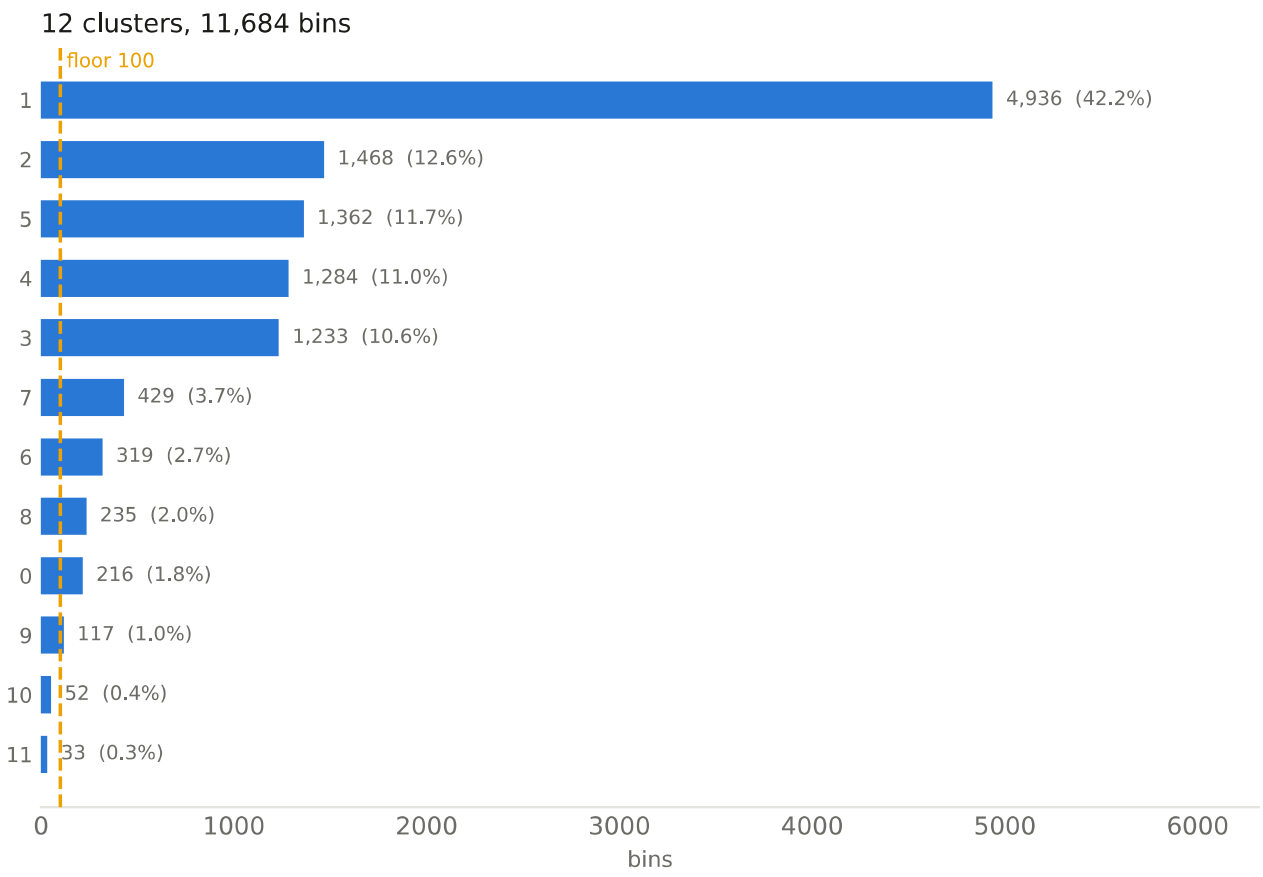


Fig. 2 クラスタごとの単位数（規模順）。破線は、統計的な再分割を論じうる最小クラスタ規模を示す。

品質管理を通過した 11,684 個の組織ビンを 11 クラスタに分割した。最小はマスト細胞の 52 ビンであり、規模図に示した 100 ビンの報告下限を下回るため、以降も規模を併記して報告する。

2. 細胞型アノテーション

クラスタ 2 と 11 は単一の形質細胞集団であり、統合して納品する。統合前の 12 クラスタ分割も v1 および v2 として併せて納品する。 納品分割に細胞型を付したものを Fig. 3 に示す。

推奨する出力結果。第1節ではこの版の結果を示す。全11クラスタで構成され、クラスタ11（33ビン、IgM優位の形質細胞）はクラスタ2（IgG優位の形質細胞）に統合した。両クラスタを直接比較したところ、調整済みp値 < 0.05を満たす遺伝子は5遺伝子のみであった。このうちIGHM、IGHG1、JCHAINは、いずれも免疫グロブリンのアイソタイプの違いを反映するものであり、残る2遺伝子も33ビン中4〜5ビンでしか検出されなかった。このため、両者を独立した集団として扱う根拠は乏しいと判断し、統合した。他のクラスタ境界はすべて維持した。

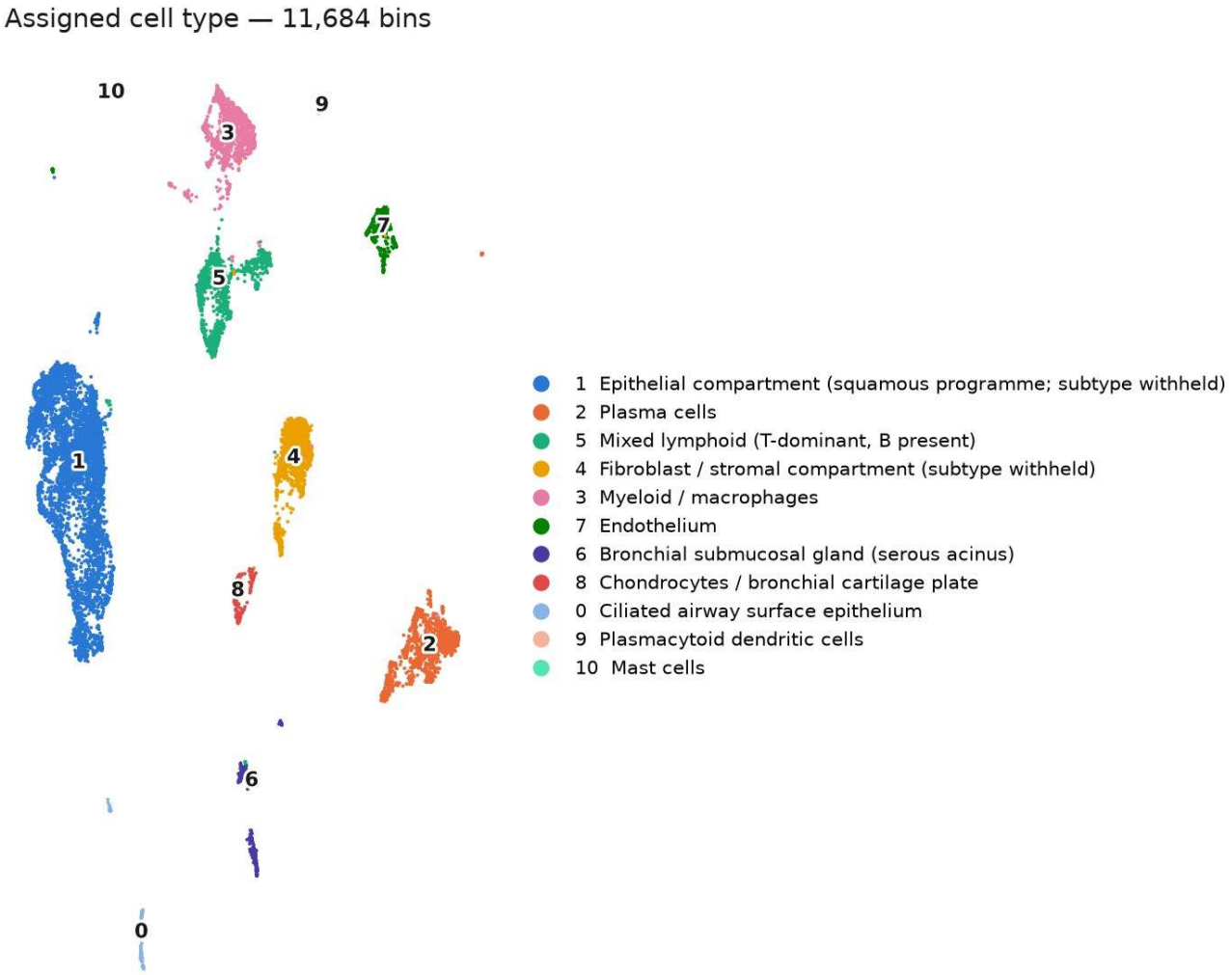


Fig. 3 Fig. 1 と同一座標上に示した納品分割。細胞型により着色し、クラスタ番号を集団中央に併記したため、Fig. 1 と対応させて読むことができる。

3. アノテーションの根拠

Fig. 4 に、事前定義したパネルの全遺伝子を納品した全クラスタに対して示す。以下のクラスタごとの判定は、いずれもこの図と照合して確認できる。

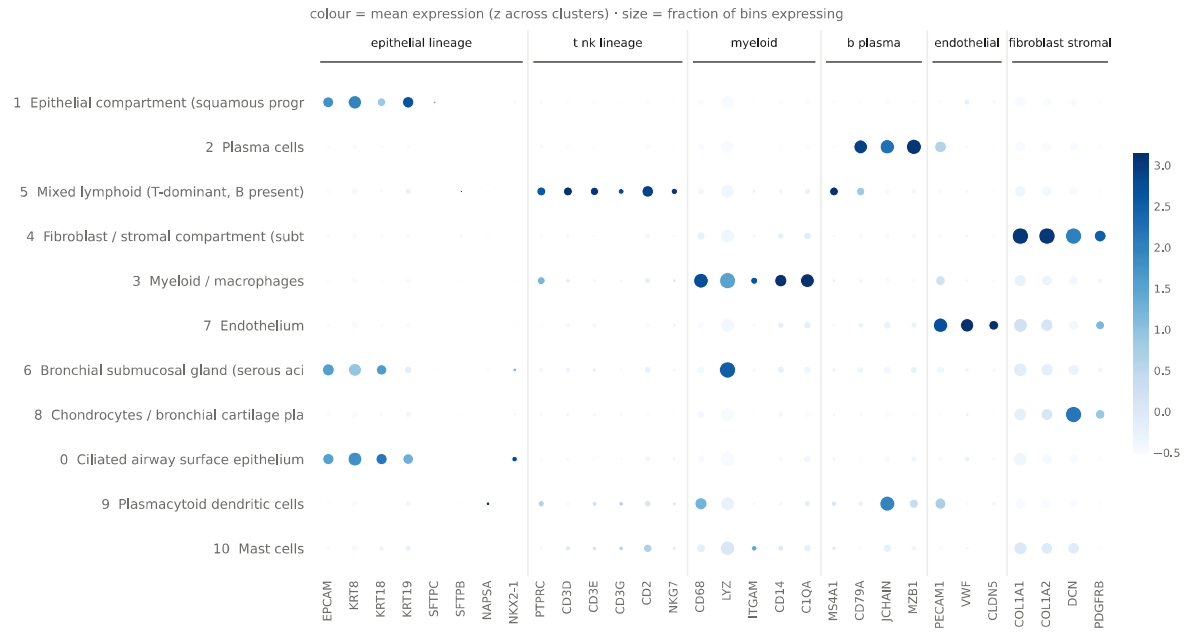


Fig. 4 事前定義した区画パネルの全遺伝子を、納品した全クラスタに対して示す。遺伝子はパネル順を保持し、各パネルを軸上で括弧と名称により明示した。したがって各クラスタの割り当ては、それを支持した遺伝子のみではなくパネル全体に対して検証できる。色はクラスタ間で標準化した平均発現量、点の大きさは当該遺伝子が検出された単位の割合である。単一遺伝子のみの括弧は、そのパネルが特徴空間に存在しない遺伝子を含むことを示す。

▼ Cluster 0: 線毛気道表面上皮

CellTypistはこのクラスタをBasal resting、DecoupleRによるマーカーエンリッチメント解析はLuminal epithelial cellsと判定した。いずれも上皮系という点では一致するものの、サブタイプの判定は本クラスタの遺伝子発現と整合しなかった。そこで、直接観察されたマーカー遺伝子の発現に基づいて命名した。

本クラスタでは、軸糸構造に関連するTPPP3、CAPS、C20orf85、RSPH1、PIFO、TMEM190[1]が特徴的に発現していた。このうち4遺伝子は組織内の他の領域では検出率0%であり、大きな上皮クラスタと比較してlog2 fold changeが8を超えていた。一方、線毛形成のマスター転写因子であるFOXJ1は検出されなかったため、典型的な線毛上皮マーカーが一式そろっているとはせず、上記遺伝子の発現を根拠とした。本クラスタは線毛気道表面上皮と命名した。クラスタ6および8とともに気管支壁を構成する。Fig. 5 に埋め込み上の位置と、本クラスタを最も特徴づける遺伝子を示す。

CELLTYPIST Basal resting · DECOUPLER Luminal epithelial cells

支持されたマーカー遺伝子群: 軸糸／線毛の構造遺伝子（TPPP3, CAPS, C20orf85, RSPH1, PIFO, TMEM190, DNALI1）

TPPP3 ビンの 41.2%（組織の他の部分では 2.0%） · **CAPS** 35.2%（2.6%） · **C20orf85** 17.6%（0.0%） · **RSPH1** 13.0%（0.0%） · **PIFO** 11.6%（0.0%） · **TMEM190** 10.2%（0.0%）

報告文献: [1]

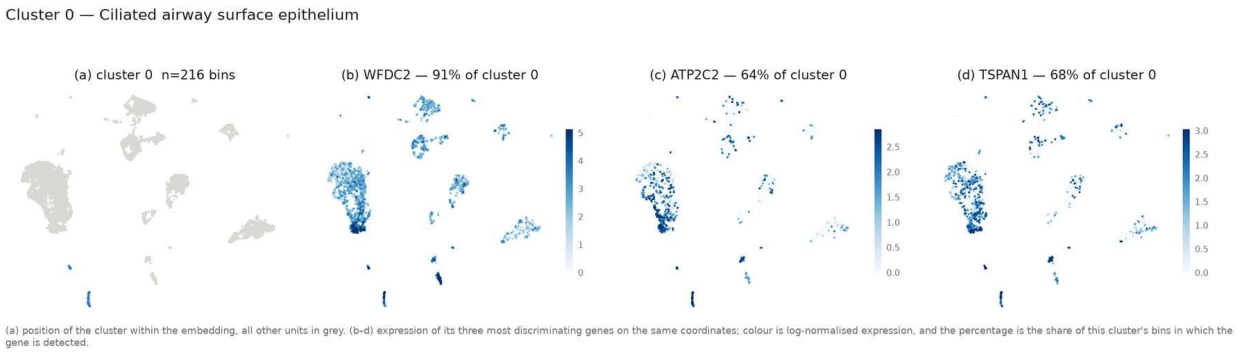


Fig. 5 クラスタ 0（線毛気道表面上皮）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 1: 上皮区画、扁平上皮プログラム（サブタイプ保留）

CellTypistはこのクラスタをBasal resting、DecoupleRによるマーカーエンリッチメント解析はLuminal epithelial cellsと判定した。KRT5、KRT6A、TP63 [2]の発現がその主な根拠であり、肺胞系マーカーはほとんど認められなかった（SFTPCは全クラスタで0.0%、NAPSA 0.3%、NKX2-1 1.1%）。以上から、本クラスタは腺系ではなく扁平上皮と判定した。

一方、サブタイプについては特定しなかった。CellTypistは健常肺アトラスを参照しており、本検体のようながん組織ブロックへの適用には限界がある。また、本クラスタの上位マーカーに対するデータベース上の一致も、角化細胞／皮膚（1.157）、肝芽細胞／肝臓（1.014）など、本組織とは直接関係しない細胞・臓器が上位を占めた。さらに、同一ブロック内に気管支壁が含まれることから、KRT5、KRT6A、TP63などのケラチン関連マーカーのみでは、正常気管支の基底上皮と扁平上皮化生を区別できない。したがって、本クラスタは上皮区画として命名した。Fig. 6 に埋め込み上の位置と、最も特徴づける遺伝子を示す。

CELLTYPIST Basal resting · DECOUPLER Luminal epithelial cells

支持されたマーカー遺伝子群: 基底／扁平上皮のケラチンとその転写因子（KRT5, KRT6A, KRT14, TP63）; 肺胞上皮の遺伝子。検定したうえで不在（SFTPC, NAPSA, NKX2-1）

KRT5 ビンの 69.6% · **TP63** 60.3% · **KRT6A** 78.0% · **SFTPC** 全クラスタで 0.0% · **NAPSA** 0.3% · **NKX2-1** 1.1%

報告文献: [2]

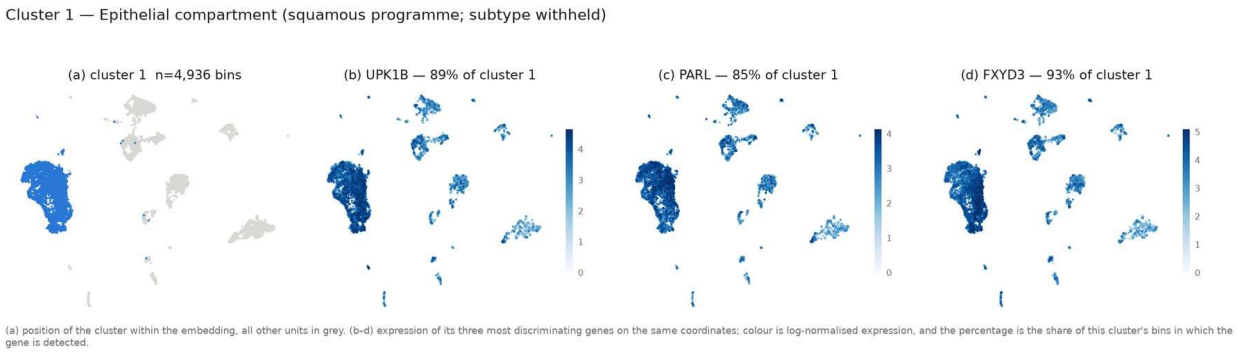


Fig. 6 クラスタ 1（上皮区画、扁平上皮プログラム（サブタイプ保留））の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 2: 形質細胞

CellTypistおよびDecoupleRによるマーカーエンリッチメント解析はいずれも、このクラスタをPlasma cellsと判定した。マーカーデータベースでも、独立に形質細胞が最上位となった。MZB1、JCHAIN、IGHG1、IGHM [3]などの分泌型免疫グロブリン関連遺伝子が特徴的に発現し、B／形質細胞パネルのスコアは1.269、次点は0.000であった。以上から、本クラスタを形質細胞と命名した。

本クラスタには、旧クラスタ11（33ビン）として分離されていたIgM優位の形質細胞も統合した。両クラスタを直接比較して有意差を示した遺伝子は5遺伝子のみであり、そのうち3遺伝子は免疫グロブリンのアイソタイプの違いを反映していた。このため、IgM優位とIgG優位の違いは独立した細胞型の境界とはせず、形質細胞集団内のアイソタイプ組成の違いとして扱った。Fig. 7 に位置と特徴遺伝子を示す。

CELLTYPIST Plasma cells · DECOUPLER Plasma cells

支持されたマーカー遺伝子群: 形質細胞の分泌プログラムと免疫グロブリン遺伝子（MZB1, JCHAIN, IGHG1, IGHM）

IGHM 統合した旧クラスタ 11 のビンの 100%（クラスタ 2 は 19.8%） · **JCHAIN** 100%（71.3%） · **IGHG1** 72.7%（96.7%） · **b_plasma panel** z 1.269、次点との差 1.269

報告文献: [3]

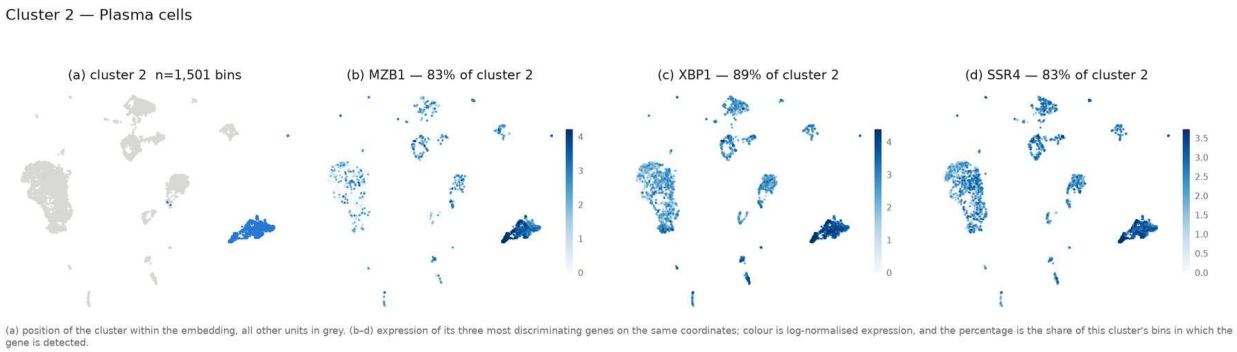


Fig. 7 クラスタ 2（形質細胞）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 3: 骨髄系／マクロファージ

CellTypistはこのクラスタをMonocyte-derived Mph、DecoupleRによるマーカーエンリッチメント解析はMacrophagesと判定し、両者で骨髄系／マクロファージ系列という点で一致した。CD68、LYZ、ITGAM、CD14、C1QA [4]からなる骨髄系マーカーパネルでも明瞭なシグナルが得られ、スコアは2.642、次点は0.093、差は2.549と、本分割の中でも特に大きかった。以上から、本クラスタを骨髄系／マクロファージと命名した。

なお、LYZは全クラスタの62〜97%のビンで検出されており、本組織ではLYZ単独で骨髄系を識別することはできない。今回の判定はLYZ単独ではなく、上記のマーカー遺伝子群全体の発現に基づく。Fig. 8 に位置と特徴遺伝子を示す。

CELLTYPIST Monocyte-derived Mph · DECOUPLER Macrophages

支持されたマーカー遺伝子群: マクロファージ系列と補体の遺伝子（CD68, C1QA, CD14, ITGAM）

myeloid panel z 2.642、次点 0.093、差 2.549 · **n** 1,233 ビン（10.6%） · **LYZ** この組織では情報を持たない。全クラスタでビンの 62〜97% に検出

報告文献: [4]

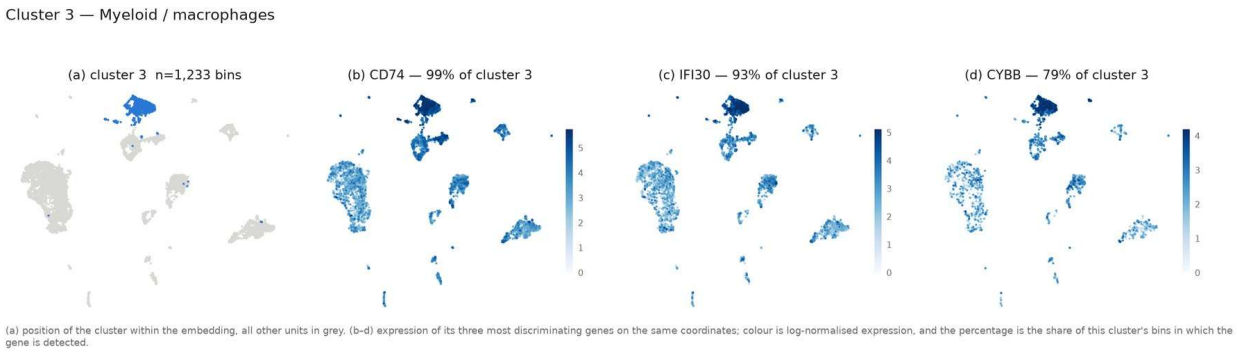


Fig. 8 クラスタ 3（骨髄系／マクロファージ）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 4: 線維芽細胞／間質区画（サブタイプ保留）

CellTypistはこのクラスタをPeribronchial fibroblasts、DecoupleRによるマーカーエンリッチメント解析はFibroblastsと判定し、いずれも線維芽細胞系列を支持した。COL1A1、COL1A2、DCN、PDGFRB [4]からなる線維芽細胞／間質マーカーパネルでも明瞭なシグナルが得られ、スコアは2.658、次点は−0.193、差は2.851であった。以上から、本クラスタを線維芽細胞／間質と命名した。

一方、サブタイプについてはクラスタ1と同様の理由から特定しなかった。実際、本クラスタの上位マーカーをデータベースと照合すると、最上位は膵星細胞（1.962）であり、組織特異的なサブタイプ名をそのまま採用する根拠には乏しかった。

また、本クラスタではEMT関連転写因子が適切に検出されていた。このことは、クラスタ1でEMT状態を支持するシグナルが乏しいという結果が、単なる検出不足による陰性所見ではないことを示している。すなわち、EMT関連シグナルは本データ上で検出可能であり、その主な由来は上皮区画ではなく間質区画であると解釈できる。Fig. 9 に位置と特徴遺伝子を示す。

CELLTYPIST Peribronchial fibroblasts · DECOUPLER Fibroblasts

支持されたマーカー遺伝子群: 線維性コラーゲンと間質の中核遺伝子（COL1A1, COL1A2, DCN, PDGFRB）

fibroblast_stromal panel z 2.658、次点 -0.193、差 2.851 · **n** 1,284 ビン（11.0%） · **COL1A1** 単独では情報を持たない。全クラスタでビンの 36〜99% に検出

報告文献: [4]

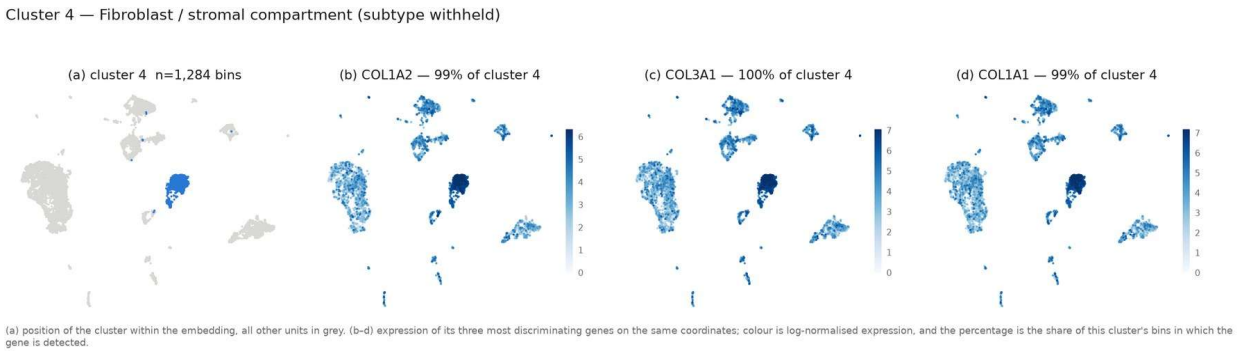


Fig. 9 クラスタ 4（線維芽細胞／間質区画（サブタイプ保留））の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 5: 混合リンパ系集団（T 優位、B も存在）

CellTypistはこのクラスタをCD4 T cells、DecoupleRによるマーカーエンリッチメント解析はT cellsと判定した。CD3D、CD3E、CD3G、CD2 [5]からなるT系列マーカーパネルはいずれも他のクラスタより高く、パネルスコアは2.928、次点は0.654であった。

一方、このクラスタではB細胞マーカーであるMS4A1（26.7%）とCD79A（25.5%）も、T細胞マーカーであるCD3D（27.5%）と同程度の割合で検出された。25 µm四方のビンにはT細胞とB細胞が同時に含まれうるため、T細胞のみからなる集団とはみなさず、本クラスタをリンパ球混合集団と命名した。したがって、後述するCD8組成の割合についても、分母はT細胞ではなく、このリンパ球混合集団の全ビンである。Fig. 10 に位置と特徴遺伝子を示す。

CELLTYPIST CD4 T cells · DECOUPLER T cells

支持されたマーカー遺伝子群: T 細胞受容体複合体と汎 T 細胞遺伝子（CD3D, CD3E, CD3G, CD2, TRAC）; 同一ビンで共検出される B 系列遺伝子（MS4A1, CD79A）

CD3D ビンの 27.5% · **CD3E** 24.4% · **CD3G** 10.7% · **CD2** 48.2% · **PTPRC** 28.6% · **MS4A1** 26.7% · **CD79A** 25.5% · **CD8A** 13.9%。12 クラスタ中の最高値（他クラスタは 1.3〜4.3%）

報告文献: [5]

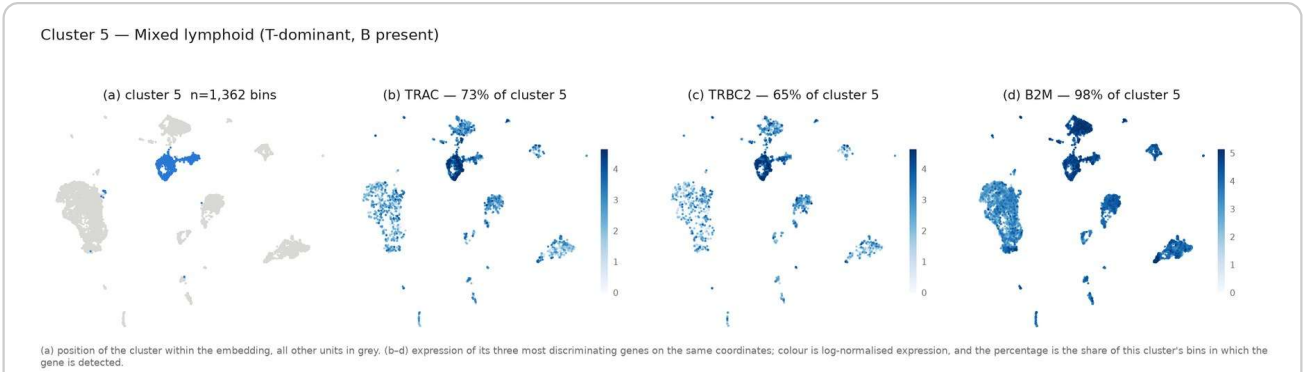


Fig. 10 クラスタ 5（混合リンパ系集団（T 優位、B も存在））の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 6: 気管支粘膜下腺、漿液腺房

CellTypistはこのクラスタをSMG serous (bronchial)、DecoupleRによるマーカーエンリッチメント解析はLuminal epithelial cellsと判定した。

本クラスタでは、腺分泌に関連するAZGP1、PIGR、ZG16B、LTF、BPIFB1、MUC5B [6]が特徴的に発現しており、6遺伝子すべてで他のクラスタを大きく上回った。これらの発現に基づき、本クラスタを気管支粘膜下腺の漿液腺房と命名した。クラスタ0および8とともに気管支壁を構成する。Fig. 11 に位置と特徴遺伝子を示す。

CELLTYPIST SMG serous (bronchial) · DECOUPLER Luminal epithelial cells
支持されたマーカー遺伝子群: 気道粘膜下腺の分泌産物（AZGP1, PIGR, ZG16B, LTF, BPIFB1, MUC5B）

AZGP1 ビンの 90.9%（他クラスタの最大は 31.1%） · **PIGR** 93.1%（17.4%） · **ZG16B** 61.4%（5.8%） · **LTF** 65.8%（23.8%） · **BPIFB1** 89.7%（40.4%） · **MUC5B** 82.1%（33.9%）

報告文献: [6]

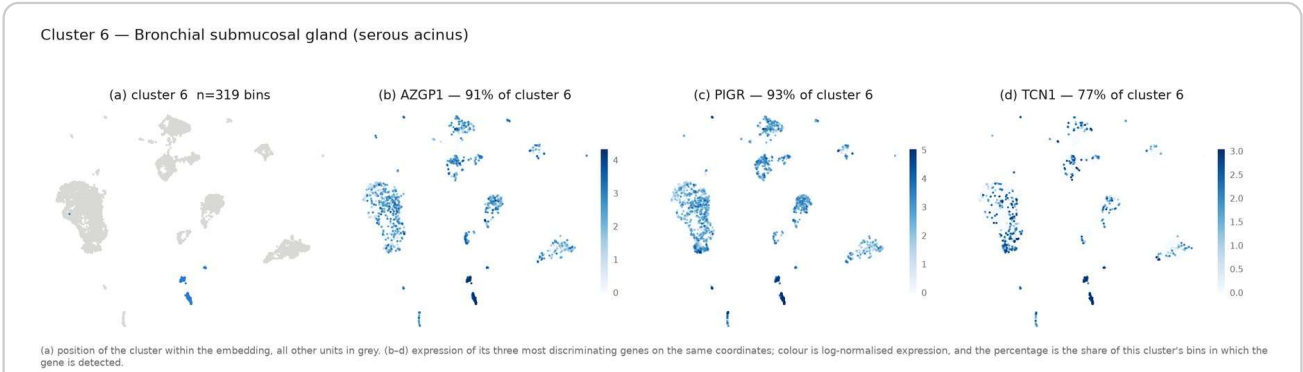


Fig. 11 クラスタ 6（気管支粘膜下腺、漿液腺房）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 7: 内皮

CellTypistはこのクラスタをEC venous systemic、DecoupleRによるマーカーエンリッチメント解析はEndothelial cellsと判定し、いずれも内皮系列を支持した。PECAM1、VWF、CLDN5 [4]からなる内皮マーカーパネルでも明瞭なシグナルが得られ、スコアは3.013、次点は0.305、差は2.708と、本分割の中でも特に大きかった。以上から、本クラスタを内皮と命名した。

一方、CellTypistの参照データは健常肺アトラスであるため、分類器が返した静脈系というサブタイプまでは採用しなかった。Fig. 12 に位置と特徴遺伝子を示す。

CELLTYPIST EC venous systemic · DECOUPLER Endothelial cells

支持されたマーカー遺伝子群: 内皮の細胞間結合と Weibel-Palade 体の遺伝子（PECAM1, VWF, CLDN5）

endothelial panel z 3.013、次点 0.305、差 2.708 · **n** 429 ビン（3.7%）

報告文献: [4]

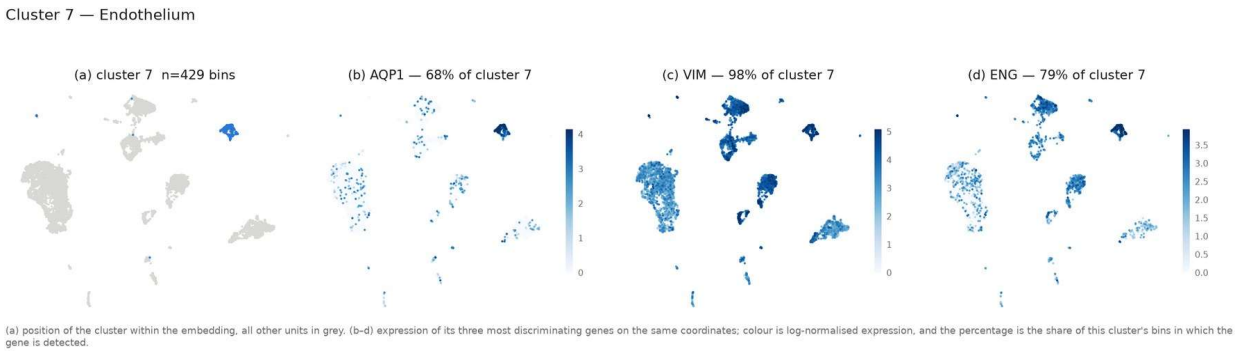


Fig. 12 クラスタ 7（内皮）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 8: 軟骨細胞／気管支軟骨板

CellTypistはこのクラスタをAdventitial fibroblasts、DecoupleRによるマーカーエンリッチメント解析はFibroblastsと判定したが、いずれも本クラスタの実態とは一致しなかった。これは、参照アトラスに軟骨細胞のクラスが含まれておらず、分類器が軟骨細胞という判定を返せないことに起因していると考えられる。

軟骨基質に関連するACAN、COL2A1、COL9A3、HAPLN1、COMP、CHAD、CYTL1、SOX9 [7]はいずれも本クラスタで最も高く発現していた。これらの結果に基づき、本クラスタを軟骨板と命名した。クラスタ0および6とともに、気管支壁を構成する主要な組織成分である。Fig. 13 に位置と特徴遺伝子を示す。

CELLTYPIST Adventitial fibroblasts · DECOUPLER Fibroblasts

支持されたマーカー遺伝子群: 軟骨の細胞外基質とその転写因子（ACAN, COL2A1, COL9A3, HAPLN1, COMP, CHAD, CYTL1, SOX9）

ACAN ビンの 73.2%（他クラスタの最大は 4.2%）・**CHAD** 76.6%（9.7%）・**COL2A1** 49.8%（5.6%）・**COL9A3** 37.0%（1.9%）・**HAPLN1** 45.1%（2.0%）・**COMP** 65.1%（21.5%）・**SOX9** 27.7%（16.6%）

報告文献: [7]

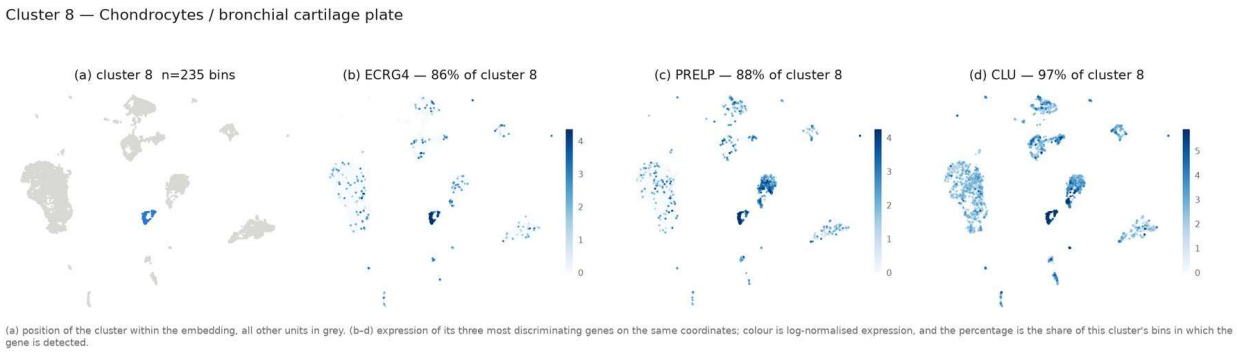


Fig. 13 クラスタ 8（軟骨細胞／気管支軟骨板）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 9: 形質細胞様樹状細胞

CellTypistはこのクラスタをPlasmacytoid DCs、DecoupleRによるマーカーエンリッチメント解析はPlasmacytoid dendritic cellsと判定し、マーカーデータベースでも同様に形質細胞様樹状細胞が最上位となった。

一方、解析計画で定義した区画パネルには樹状細胞のクラスが含まれていない。そのため、パネルスコアは最上位でも0.262、次点が0.203、差が0.059と小さく、この結果は区画パネルに対応するクラスが存在しないことを反映している。そこで、命名には区画パネルではなく、形質細胞様樹状細胞のマーカーであるCLEC4CとLILRA4 [5]の発現を用いた。以上から、本クラスタを形質細胞様樹状細胞と命名した。

なお、細胞傷害性パネルに含まれるGZMBはリンパ系クラスタではなく本クラスタで最も高頻度に検出され、88.9%のビンで認められた。Fig. 14 に位置と特徴遺伝子を示す。

CELLTYPIST Plasmacytoid DCs・DECOUPLER Plasmacytoid dendritic cells

支持されたマーカー遺伝子群: 形質細胞様樹状細胞の受容体（CLEC4C, LILRA4）

n 117 ビン（1.0%）・**compartment panel** 最上位 0.262、次点 0.203、差 0.059。解析計画のパネルに pDC クラスが存在しない・**GZMB** ビンの 88.9%。データセット全体での最高値

報告文献: [5]

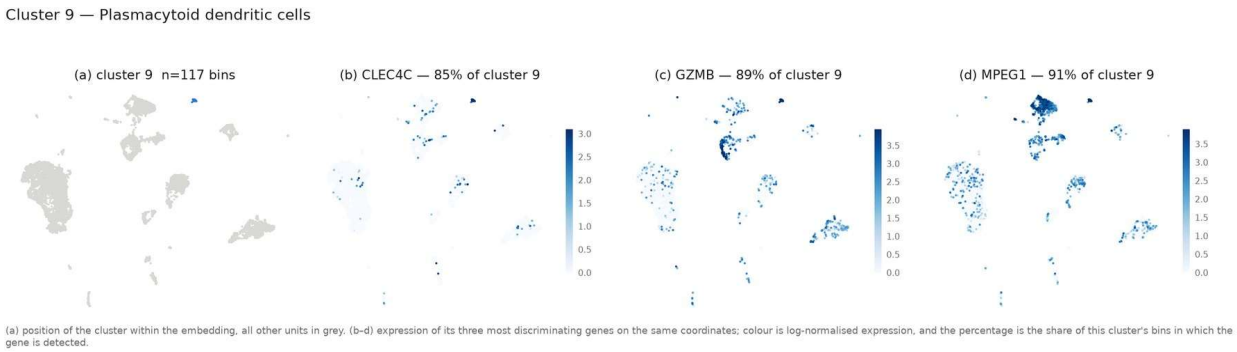


Fig. 14 クラスタ 9（形質細胞様樹状細胞）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

▼ Cluster 10: マスト細胞

CellTypistはこのクラスタをMast cells、DecoupleRによるマーカーエンリッチメント解析はFibroblastsと判定し、両者の結果は一致しなかった。このように2つの解析で異なる判定が得られるた。

本クラスタでは、マスト細胞の顆粒プロテアーゼおよび受容体であるCPA3、KIT、MS4A2 [8] が特徴的に発現し、いずれも次点のクラスタを1桁以上上回った。これらの発現に基づき、本クラスタをマスト細胞と命名した。なお、代表的なマスト細胞トリプターゼであるTPSAB1およびTPSB2は、今回測定された特徴量空間に含まれていなかったため、陰性ではなく未測定として扱う。

本クラスタは52ビンで構成され、100ビンとして設定した報告下限を下回る。Fig. 15 に位置と特徴遺伝子を示す。

CELLTYPIST Mast cells · DECOUPLER Fibroblasts

支持されたマーカー遺伝子群: マスト細胞の顆粒プロテアーゼと受容体（CPA3, KIT, MS4A2。TPSAB1／TPSB2 は未測定）

CPA3 ビンの 98.1%（他クラスタの最大は 4.6%） · **KIT** 84.6%（6.9%） · **MS4A2** 80.8%（1.8%） · **n** 52 ビン（0.45%）。報告下限を下回る

報告文献: [8]

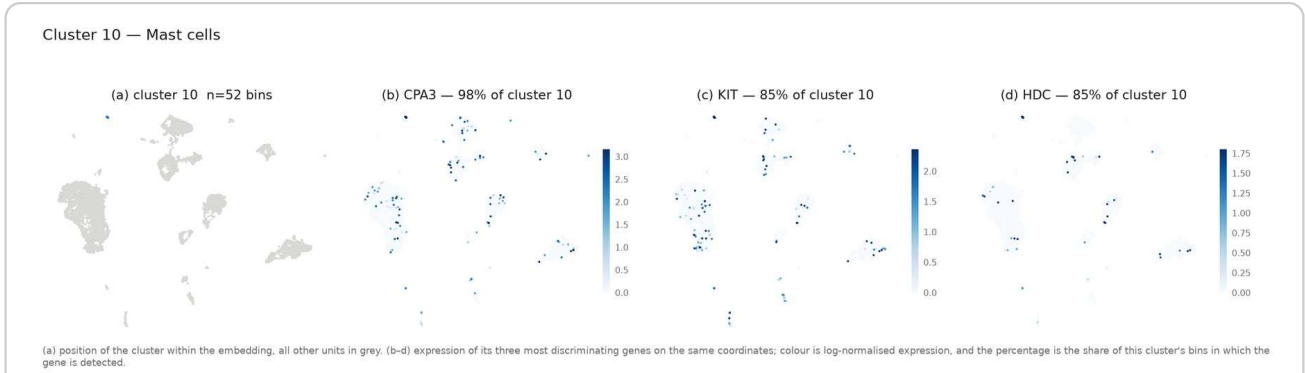


Fig. 15 クラスタ 10（マスト細胞）の発現所見。(a) 埋め込み上での当該クラスタの位置（他の単位は灰色）。(b–d) 判別力上位3遺伝子の同一座標上での発現。各見出しの百分率は、当該クラスタの単位のうちその遺伝子が検出された割合を示す。

上記の割り当てを Fig. 16 に総括する。各クラスタの特徴遺伝子を同一軸上に配置しており、組織全体を通じて各割り当ての特異性を比較できる。

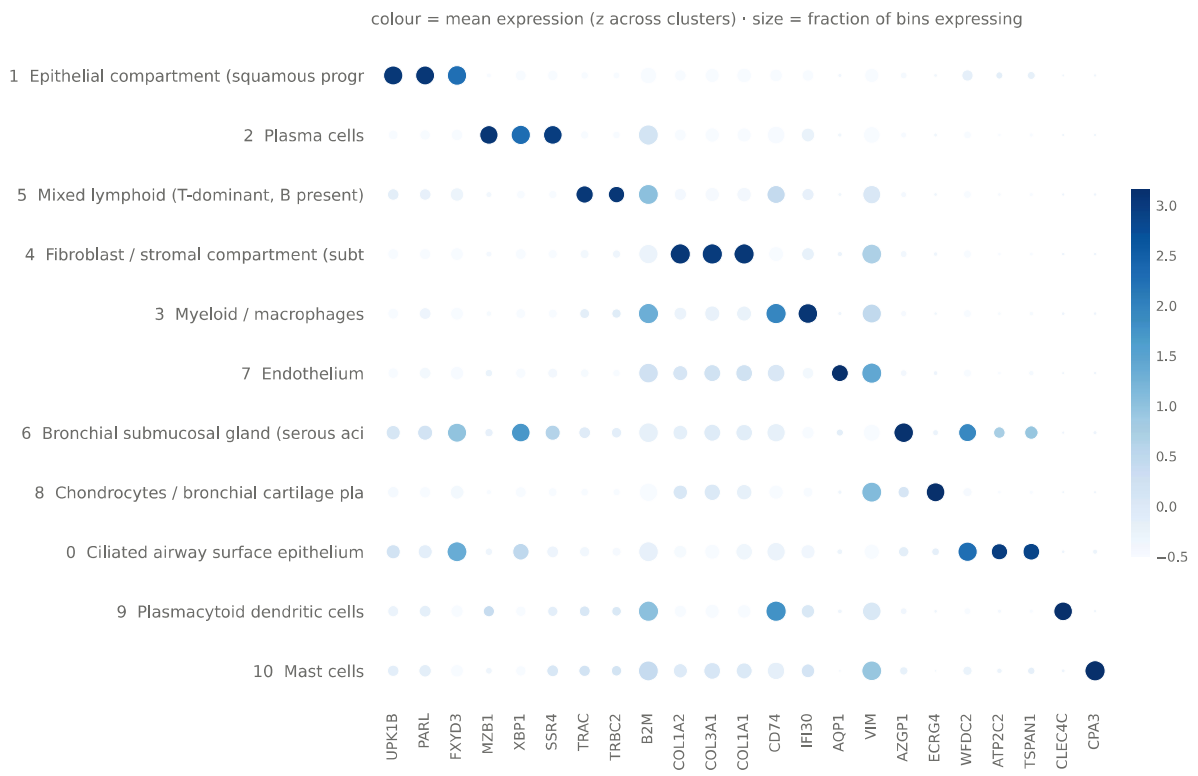


Fig. 16 上記の所見の総括。行は各クラスタ、列はそれを特徴づける遺伝子である。色はクラスタ間で標準化した平均発現量であり、全体で高発現の遺伝子が色調を占有しない。点の大きさは当該遺伝子が検出された単位の割合を表す。

4. 結果

▼ Q1 — 上皮区画は 4 状態のうち 3 つを保持し、離散した集団として分離するのは 1 つのみ

依頼書では、上皮区画について、①基底細胞様・増殖性・角化型・EMT様の4状態が存在するか、②それぞれがどの程度を占めるか、③各状態を特徴づける遺伝子は何か、の3点が問われている。そこで、上皮区画（4,936ビン、組織全体の42.2%）のみを抽出して再クラスタリングし、得られた10のサブクラスタを、事前に定めた4つのマーカーパネルで採点した。結果をFig. 17に示す。各サブクラスタに最も高いスコアを示した状態、その構成比、および同一座標上における各パネルの発現を示している。

構成比。 4状態のうち、明確なビン数を示せるのは基底細胞様のみである。基底細胞様は74ビンで、上皮区画の1.5%を占め、独立した集団として分離した。一方、増殖性・角化型・EMT様は明確な集団として分離しなかったため、それぞれに属するビン数を直接定義することはできない。

この3状態について示せるのは、10のサブクラスタを最も高いスコアを示したパネルに割り当てた場合の分布である。その結果、増殖性2,708ビン（54.9%）、角化型1,340ビン（27.1%）、EMT様814ビン（16.5%）となった。ただし、これは各サブクラスタを便宜的に最上位のパネルへ割り当てた結果であり、その状態にある細胞数を示すものではない。どのパネルにも明確に合致しないサブクラスタであっても、最上位のパネルには必ず割り当てられる。EMT様に分類された814ビンはその典型例であり、後述するように、実際には上皮のEMT状態を反映したものではなかった。

離散した集団か、連続的な勾配か。 4状態のうち、独立した集団として分離したのは基底細胞様のみであった。基底細胞様は小規模な単一サブクラスタとして他の上皮区画から分離し、事前に設定した検定基準もすべて満たした。一方、増殖性・角化型・EMT様のシグナルは検出できるものの、明確な集団境界は形成しなかった。Fig. 17の(c)〜(e)に示すように、特定のサブクラスタだけで強く発現するのではなく、隣接する複数のサブクラスタにまたがって連続的に強度が変化している。したがって、これら3状態は独立した細胞集団ではなく、上皮区画内に連続的に分布するプログラムとして扱う。この違いは解析上も重要であり、離散した集団であれば集団として抽出できる一方、連続的な状態はスコアの強度として評価する必要がある。

特徴遺伝子。 各状態について上皮区画の残りとの発現差解析を行い、事前に定義したマーカーパネルに含まれ、かつ検定基準を満たした遺伝子は以下のとおりであった。

- 基底細胞様 — KRT14、KRT5、NGFR、ITGB4、ITGA6（パネル8遺伝子中5遺伝子）
- 増殖性 — PCNA、BIRC5、STMN1、CENPF、UBE2C、TOP2A、CCNB1（8遺伝子中7遺伝子）
- 角化型 — CSTA、KRT6A、KRT13、SPRR1B、SPRR3、IVL、KRT16 (9遺伝子中7遺伝子。SPRR1Aは本特徴量空間では未測定）
- EMT 様 — FN1、MMP2、VIM（10遺伝子中3遺伝子）

EMT様は4状態の中で自身のマーカーパネルとの一致が最も弱く、検定を通過したFN1、MMP2、VIMはいずれも間質細胞で恒常的に発現する遺伝子であった。また、これらのシグナ

ルは上皮ビンに限局していなかった。



Fig. 17 依頼書の 4 状態に対する上皮区画の回答。(a) 上皮区画の 10 サブクラスタを、それぞれ最も高く採点された状態で着色した。(b) 構成比（ビン数と区画に対する百分率）。(c-f) 同一座標上での各事前定義パネルの平均発現。増殖性・角化型・EMT 様は隣接するサブクラスタにまたがって濃淡が変化し、基底細胞様のパネルだけが分離した小集団を標識する。

▼ Q2 — CD8A はリンパ球ビンの 13.9% で検出され、細胞傷害性と疲弊は分離せず同時に検出される

依頼書では、CD8陽性T細胞の構成と、その中にみられる細胞傷害性および疲弊の状態が問われている。そこで、リンパ球区画（1,362ビン、組織全体の11.7%）のみを抽出して再クラスタリングし、事前に定めたマーカーパネルで採点した。Fig. 18 に、リンパ球区画内でのCD8A、T系列パネル、細胞傷害性パネル、疲弊パネルの発現と、CD8A陽性ビンの構成を示す。

構成比。 T系列遺伝子は区画内の87.8%のビンで検出され、この区画がリンパ球を主体とすることを支持した。CD8Aは189ビンで検出され、リンパ球区画全体の13.9%、T系列遺伝子が検出されたビンに限ると15.6%に相当した。この189ビンについて、各パネルで2遺伝子以上の検出を条件とすると、細胞傷害性プログラムを示すビンは81%、疲弊プログラムを示すビンは40%、両方を示すビンは38%、いずれも示さないビンは16%であった。判定条件を1遺伝子以上に緩めると、それぞれ92%、64%、61%となった。

細胞傷害性と疲弊は、この区画を二分する軸ではない。 疲弊プログラムを示すビンは40%であったが、その大部分にあたる38%は細胞傷害性プログラムも同時に示した。細胞傷害性のみを示すビンは43%、疲弊のみを示すビンは2%であった。リンパ球区画を再クラスタリングしても両者は別々の集団として分離せず、主な分割軸は細胞周期であった。この結果には少なくとも2通りの解釈があり、本データのみから両者を区別することはできない。すなわち、同一細胞内で細胞傷害性と疲弊のプログラムが共発現している場合と、25 µmの同一ビン内に細胞傷害性を示す細胞と疲弊した細胞が共存している場合である。1ビンに複数の細胞が含まれうる本データでは、後者も十分に考慮する必要がある。

免疫抑制について何が言えるか。 疲弊プログラムはCD8A陽性ビンの相当数で検出され、依頼書が想定する免疫回避的な表現型とは整合する。一方で、独立した疲弊集団として分離しておらず、細胞傷害性プログラムと排他的でもない。また、Tregマーカーはリンパ球区画の24.3%

のビンで検出され、CTLA4およびTIGITはTregパネルと疲弊パネルの両方に含まれる。したがって、観察された疲弊シグナルをCD8 T細胞のみに帰属させることはできない。

事前に設定した基準を満たしていない点と、その解釈。 本解析では、リンパ球区画を評価する前に、CD8Aが検出されるビンの割合について25%を基準として設定していた。実測値は13.9%であり、この基準を下回った。この基準は、リンパ球区画内でさらにCD8特異的な部分集団を探索するかどうかを判断するためのものであり、上記の構成比そのものを無効にするものではない。構成比はCD8A陽性ビンを直接分母として算出している。したがって、この結果から言えるのはCD8特異的な再分割を行わなかったという点までであり、CD8特異的な部分集団が存在しないことを示すものではない。

細胞傷害性の集計に関する留意点。 GZMBは細胞傷害性パネルに含まれるが、リンパ球区画ではなく形質細胞様樹状細胞クラスタで最も高く発現し、同クラスタの88.9%のビンで検出された。リンパ球区画では、GZMBと細胞傷害性パネルの別の1遺伝子が同一ビンで検出されれば細胞傷害性として計上される。そのため、同一ビン内のT細胞以外の細胞がGZMBの発現に寄与している可能性を考慮すると、81%という値は細胞傷害性を示すCD8 T細胞の割合ではなく、上限寄りの値として解釈する必要がある。

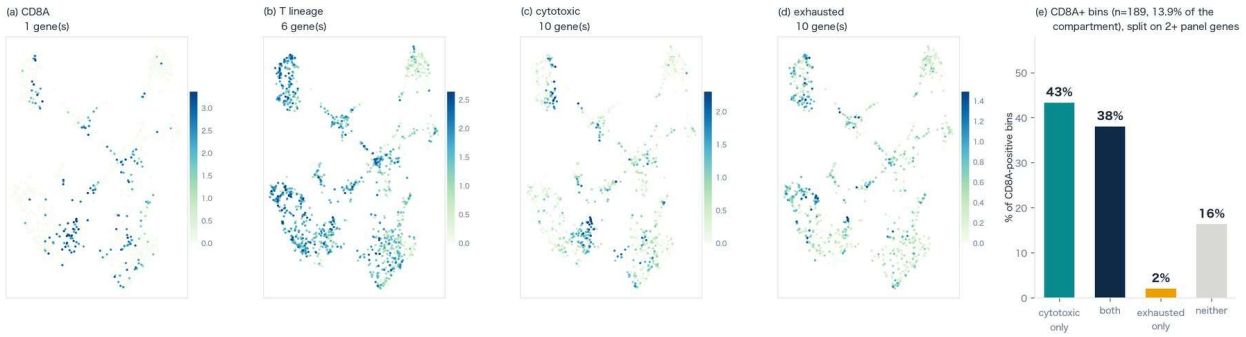


Fig. 18 依頼書の CD8 に関する問いに対するリンパ球区画の回答。(a-d) 区画自身の座標上での CD8A、T 系列パネル、細胞傷害性パネル、疲弊パネル。各パネルに属する遺伝子の平均発現量である。(e) CD8A 陽性ビンの構成。プログラムを有すると判定するには、パネルの 2 遺伝子以上の検出を要求した。

5. 追加のデータにより可能となること

第1は、すでにオブジェクトに含まれている座標情報を用いた空間隣接解析である。未実施の解析の中では最も優先度が高い。「EMT様」と判定されたものの、その後の解析で間質細胞の混入と考えられたビンが、実際に上皮と間質の境界に位置するかを直接検証できるためである。これにより、本解析の最も重要な所見であるEMT様シグナルの由来を空間的に検証できる。

第2は、同一座標に位置合わせしたH&E画像と、コピー数推定である。どの上皮が悪性であるかを評価するには、これらの情報が必要であり、本納品物のみからこの問いに答えることはできない。現在のカウントデータを別の方法で再解析しても、この制約は解消されない。

第3は、患者数を増やすことである。1患者1組織ブロックから得られた本データから示せるのは、この検体内で観察された所見の記述までである。依頼書が含意する患者間や治療群間の比較、あるいは転帰との関連を検証するには、患者識別子によって臨床情報と対応付けられた複数患者からなるコホートが必要である。

引用文献

1. Choksi SP, Lauter G, Swoboda P, Roy S. Switching on cilia: transcriptional networks regulating ciliogenesis. Development 141:1427-1441 (2014). [TPPP3, CAPS, C20orf85, RSPH1, PIFO, TMEM190 as axonemal/ciliary structural genes; FOXJ1 as the master regulator]
2. Rock JR, Onaitis MW, Rawlins EL, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. PNAS 106:12771-12775 (2009). [KRT5, KRT14, TP63, NGFR, ITGA6/ITGB4 as basal identity genes]
3. Shi W, Liao Y, Willis SN, et al. Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. Nature Immunology 16:663-673 (2015). [MZB1, JCHAIN and immunoglobulin heavy chains as the plasma-cell secretory programme]
4. Franzén O, Gan L-M, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database 2019:baz046 (2019); and Uhlén M, et al. Tissue-based map of the human proteome. Science 347:1260419 (2015). [marker databases used for the enrichment calls]
5. Villani A-C, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 356:eaah4573 (2017). [CLEC4C/LILRA4 as plasmacytoid dendritic-cell identity genes; CD3D/E/G, CD2, TRAC as the T-cell receptor complex]
6. Meyer-Hoffert U, Wingertszahn J, Wiedow O. Human leukocyte elastase induces keratinocyte proliferation by epidermal growth factor receptor activation; and Ali M, Lillehoj EP, Park Y, et al. Analysis of the proteome of human airway epithelial secretions. Proteome Science 9:4 (2011). [PIGR, ZG16B, LTF, BPIFB1, MUC5B, AZGP1 as airway submucosal gland secretory products]
7. Lefebvre V, Angelozzi M, Haseeb A. SOX9 in cartilage development and disease. Current Opinion in Cell Biology 61:39-47 (2019). [ACAN, COL2A1, COL9A3, HAPLN1, COMP, SOX9 as the cartilage matrix programme]
8. Dwyer DF, Barrett NA, Austen KF, Immunological Genome Project Consortium. Expression profiling of constitutive mast cells reveals a unique identity within the immune system. Nature Immunology 17:878-887 (2016). [CPA3, KIT, MS4A2 as mast-cell identity genes]