

1 細胞解析による浸潤性小葉がんの腫瘍内不均一性の解明

Evaluation of intratumoral heterogeneity in invasive lobular carcinoma by single-cell RNA-seq

国立がん研究センター研究所 病態情報学ユニット

(現所属先：大阪国際がんセンター研究所 腫瘍増殖制御学部) 中山 淳

要約

乳腺末梢組織から発症する浸潤性小葉がん (Invasive Lobular Carcinoma、以下 ILC) などの特殊型乳がんは、生物学的には乳管がん (Invasive Ductal Carcinoma、以下 IDC) と全く異なる。しかしながら、現在の ILC 治療は IDC に準じており、ILC の生物学的特性や分子メカニズムに基づいた治療法は確立されていない。また、ILC は多発性 (別の小葉内で複数同時に発生する特性) のがんであり、これは同一検体内でそれぞれの小葉内にはがん幹細胞 (Tumor Initiating Cell, Cancer stem cell) が存在する可能性を示唆する。本研究の目的は、組織空間トランスクリプトーム解析を用いて ILC における腫瘍内細胞不均一性を明らかにし、細胞間相互作用の解明や ILC がん細胞の増殖・生存に関わる分子経路を同定することである。国立がん研究センターバイオバンクから得られた ILC 臨床検体に対して、組織空間トランスクリプトーム解析を行い、ILC 組織中のがん幹細胞マーカーの発現多様性を評価した。結果として、ILC 組織は間質細胞マーカーである COL2A1 の発現が高く、間質リッチな腫瘍であることがわかった。4 検体から得られたスポットを統合解析した結果、17 個のクラスターに分類することができ、検体に特異的なクラスターの抽出および検体横断的なクラスターの抽出に成功した。さらにはがん幹細胞マーカーの発現を評価した結果、がん幹細胞マーカー遺伝子である CD44、HMGA1、ALDH1A1 は同一スポットから発現が確認できているものもあるが、全く異なるスポット部位で発現していることも確認された。このことから、ILC 組織中においてもがん幹細胞マーカー遺伝子の発現に多様性があることを明らかにした。

緒言

乳がんは年間9万人以上が罹患し、早期発見によって5年生存率は90%を超えるが、壮年期に限れば女性の死亡原因の第1位である。乳がん全体の約80%を占める乳管に由来する乳管がん (Invasive Ductal Carcinoma、以下 IDC) であるが、約20%は乳腺の末梢組織から発症する浸潤性小葉がん (Invasive Lobular Carcinoma、以下 ILC) などの特殊型乳がんである。約10%が ILC であり、残り10%は粘液がん・アポクリンがん・化生がんなどに分類される。

ILC は比較的高齢者に発生し、画像検査でとらえられにくいので、進行例が多く見られる。また、晩期の再発症例が多いがんである。早期発見が難しく、病理的には E-Cadherin 陰性でがん細胞がばらばらに浸潤するため境界がわかりにくく、多中心発生

をするので、乳房温存手術を避けることが多く、進行すると腹膜、卵巣、消化管等特異な転移様式を取る特徴がある¹。

生物学的には IDC と全く異なるにもかかわらず、現在の ILC 治療は IDC に準じており、ILC の生物学的特性や分子メカニズムに基づいた治療法は確立されていない。浸潤性小葉がんの分子特性の理解とその治療法の確立は重要な課題の1つである。また、ILC は多発性 (別の小葉内で複数同時に発生する特性) のがんであり、これは同一検体内でそれぞれの小葉内にはがん幹細胞 (Tumor Initiating Cell, Cancer stem cell) が存在する可能性を示唆する。

本研究の目的は、組織空間トランスクリプトーム解析を用いて ILC における腫瘍内細胞不均一性を明らかにし、細胞間相互作用の解明や ILC がん細胞の

増殖・生存に関わる分子経路を同定することである。また、ILCに特徴的ながん幹細胞集団とその多様性を明らかにすることで、ILCに最適化した治療の確立へと繋げる。

方法

バイオバンク検体の解析とcDNAライブラリの構築

国立がん研究センターバイオバンクに登録されているILC臨床検体を供与いただき、解析に使用した(国立がん研究センター生命倫理委員会承認済み2019-018)。国立がん研究センターバイオバンクからILC4症例を抽出し、Cryostatを用いて凍結ブロックから組織の切り出しを行った。得られた腫瘍組織切片をVisium Spatial Gene Expression Slide (10x Genomics)の指定領域(6.5mm x 6.5mmのDNAプローブ領域)に配置し、組織透過処理、RNAの抽出、cDNAライブラリの調製を行った。ライブラリ調製は、企業が公開している推奨プロトコルに準じて行った。得られたcDNAライブラリについて、次世代シーケンサーDNBSEQ-400(Azenta, 委託解析)にてシーケンスを行い、リードを獲得した。

組織空間トランスクリプトーム解析

次世代シーケンサーにて取得したfastqファイルに対して、Spaceranger(10x Genomics)解析を行い、スポットごとの発現量を取得した。取得した発現量をSeurat²を用いて解析を行った¹。SCTransform法³を用いて、サンプルごとに正規化を行い、統合解析を行った。解析プログラムはSeuratの標準解析プログラムに準じて行った(https://satijalab.org/seurat/articles/spatial_vignette)。

結果

ILC臨床検体の組織空間トランスクリプトーム統合

表1 ILC臨床検体の臨床情報と分子プロファイル

| Sample | Age | ER ratio | PR ratio | HER2 |
|----------|-----|----------|----------|------|
| Patient1 | 49 | >95% | >95% | 2+ |
| Patient2 | 69 | 90% | 70% | 0 |
| Patient3 | 69 | 0% | 0% | 0 |
| Patient4 | 60 | >90% | 70% | 2+ |

解析

国立がん研究センターバイオバンクから得られた検体の臨床情報と分子プロファイルを表1に示す。Luminalサブタイプ(ER・PR陽性)1検体、LuminalHER2サブタイプ(ER・PR・HER2陽性)2検体、トリプルネガティブサブタイプ(ER・PR・HER2陰性)1検体とサブタイプ横断的に検体を入手することができた(表1)。

これらILC4検体の腫瘍組織切片に対して、空間トランスクリプトーム解析を行い、得られた遺伝子発現情報をRにロードし、Seuratを用いて統合した(図1)。統合した結果、17個のクラスターに分類することができ、検体に特異的なクラスターの抽出および検体横断的なクラスターの抽出に成功した。

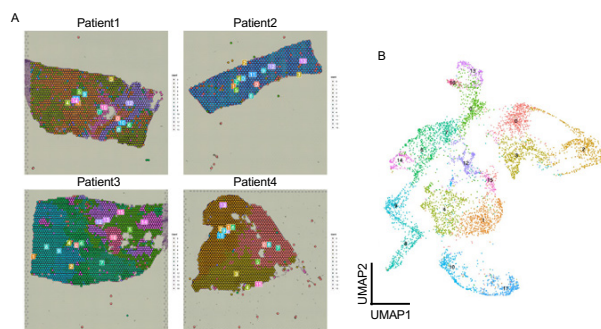


図1. ILC検体の空間トランスクリプトーム解析と統合UMAP

A. 4検体の病理切片像とクラスタリング解析のスポットティング。計17クラスターに分類することができ、検体特異的なクラスターと検体横断的なクラスターを取得することに成功した。
B. 17クラスターのUMAPプロット。

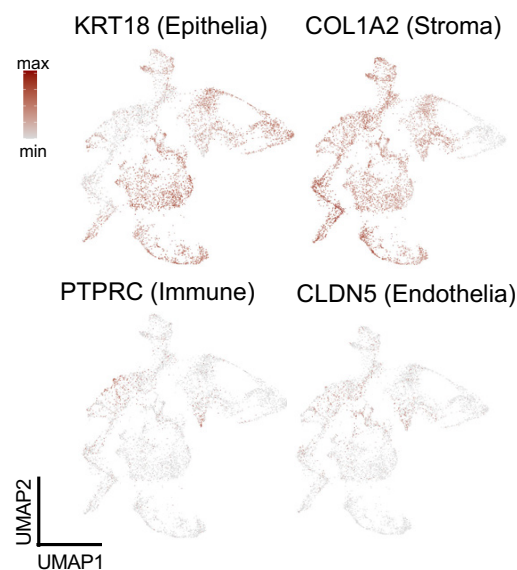


図2. 主要マーカー遺伝子の発現プロファイル

UMAPプロット上で、細胞タイプを分類する腫瘍マーカー遺伝子の発現プロファイルを表示した。上皮細胞マーカー(KRT18)、間質細胞マーカー(COL2A1)、免疫細胞マーカー(PTPRC)、内皮細胞マーカー(CLDN5)の発現をgray-redスケールで表示した。

結果より、上皮細胞のマーカー遺伝子であるKRT18が多くのスポットより観察された。これらのスポットには一定の割合で上皮細胞が含有されているものと考えられる。また、ILC腫瘍組織は間質細胞の浸潤がよく見られることで知られているが、間質細胞マーカーであるCOL2A1の発現も多くのスポットで確認された。一方で、免疫細胞マーカーであるPTPRCや内皮細胞マーカーであるCLDN5の発現は、COL2A1ほど強く発現が確認されなかった。すなわち、空間トランスクリプトーム解析結果からも、ILC間質リッチな腫瘍組織であることがわかった。

研究代表者である中山は、これまでに乳がん同一腫瘍組織内に複数種類のがん幹細胞が存在する可能性を提示してきた。具体的にCD44陽性乳がん幹細胞集団とHMGA1陽性乳がん幹細胞集団がそれぞれ同一腫瘍内に存在していることを示した⁴。ILCの多発性性質から、それぞれの小葉内にがん幹細胞 (Tumor Initiating Cell, Cancer stem cell) が存在し、空間的な多様性がある可能性を考え、乳がん幹細胞マーカー遺伝子であるCD44、HMGA1、ALDH1A1の発現多様性を示した【図3】。

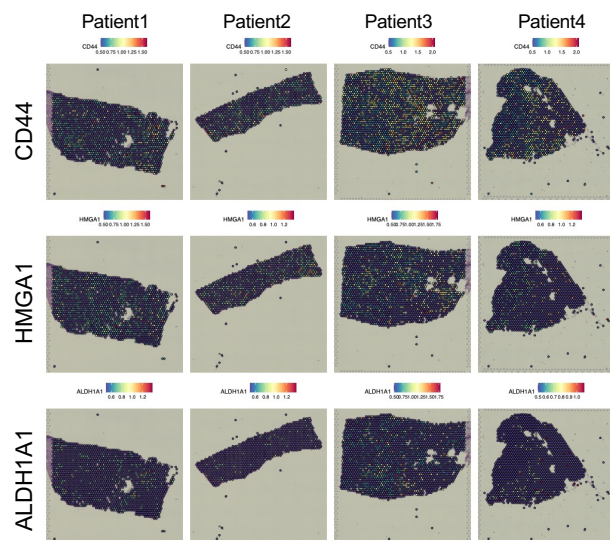


図3. 乳がん幹細胞マーカー遺伝子の空間的発現プロフィール
乳がん幹細胞マーカー遺伝子CD44、HMGA1、ALDH1A1空間的発現プロフィール。

それぞれのマーカー遺伝子は同一スポットから発現が確認できているものもあるが、全く異なるスポット部位で発現していることも確認された。同様にこれらマーカー遺伝子の発現を、UMAPプロットおよびクラスターごとのバイオリンプロットにて確認し

た【図4】。CD44およびHMGA1は多くのクラスターから検出することができた。一方でALDH1A1は限られたクラスターから検出された。

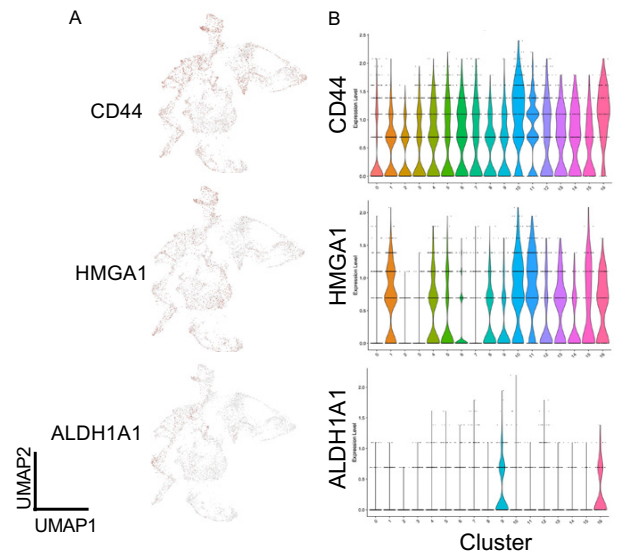


図4. 乳がん幹細胞マーカー遺伝子の発現プロフィール
A. UMAPプロット上で、乳がん幹細胞マーカー遺伝子CD44、HMGA1、ALDH1A1の発現を表示した。
B. クラスターごとの乳がん幹細胞マーカー遺伝子CD44、HMGA1、ALDH1A1発現バイオリンプロット。

考察

本研究の結果から、ILC臨床検体の空間遺伝子発現プロフィールを明らかにした。また乳がん幹細胞マーカー遺伝子であるCD44、HMGA1、ALDH1A1の発現情報と空間情報を明らかにした。CD44は免疫細胞も多く発現しているため、免疫細胞由来のCD44である可能性が否定できない。しかしながら本解析結果から、ILC組織中の免疫細胞の比率は非常に少なく、上皮細胞および間質細胞が多く含まれるスポットからCD44が検出されており、免疫細胞由来である可能性は低いと考えられる。このことから、ILC組織中においてもがん幹細胞マーカー遺伝子の発現に多様性があることを明らかにした。

しかしながら、技術的制限として、Visiumのスポットは1スポットあたり100 μ mの大きさであり、数10細胞が混じった混成スポットであることが挙げられる。すなわち、厳密に1細胞レベルでがん幹細胞の遺伝子発現を評価できていないと考えられる。この問題を解決するためにはシングルセルレベルでの解析が必要であるが、single-cell RNA-sequencing (scRNA-seq) 解析を行うためにはバイオバンク検体ではなく術後間もないFreshな検体から細胞を採取

することが望ましいため、未だに解析に至っていない。今後はシングルセル解析と統合して本空間トランスクリプトーム解析結果を評価していく必要がある。今後は、空間トランスクリプトーム解析の検体数を増やし、機械学習を用いたより詳細なスポット分類解析を行うことで、統計的議論をよりRobustなものとすることを目標として進める。またILC検体シングルセル解析との統合解析も並行して進めていく予定である。

謝 辞

本研究を遂行するにあたり、支援いただいた神澤医学研究振興財団に感謝申し上げます。また共同研究者である国立がん研究センター研究所 病態情報

学ユニットの山本雄介先生、都倉桃子先生に感謝申し上げます。

参考文献

1. McCart Reed AE, Kalinowski L, Simpson PT, et al., Invasive lobular carcinoma of the breast: the increasing importance of this special subtype. *Breast Cancer Res.* 2021, 23(1):6.
2. Stuart T, Butler A, Hoffman P, et al.. Comprehensive Integration of Single-Cell Data. *Cell.* 2019, 177(7):1888-1902.e21.
3. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 2019, 20(1):296.
4. Nakayama J, Matsunaga H, Arikawa K, et al., Identification of two cancer stem cell-like populations in triple-negative breast cancer xenografts. *Dis Model Mech.* 2022, 15(6):dmm049538.

Abstract

Invasive lobular carcinoma (ILC) is the second common invasive breast cancer, accounting for approximately 10~15% of all invasive breast cancer patients. Treatment of ILC includes surgery, radiation therapy, hormone therapy, and chemotherapy. ILC cells infiltrate into intra-mammary lobules and spread beyond its original lobule as a diffuse phenotype. Most of ILC cells lack E-cadherin activity by mutation or suppression of gene expression, causing epithelial-to-mesenchymal transition (EMT) for invasion in the lobules. In this study, to reveal the spatial significance of ILC tissue, we subjected Visium (10x Genomics) analysis into ILC specimens in the National Cancer Center biobank. Spatial transcriptome analysis identified the expressional differences and cellular diversity in the ILC specimens. Spatial transcriptome analysis showed 17 clusters in integrated ILC specimens. We focused on 3 breast cancer stem cell markers, CD44, HMGA1, and ALDH1A1 in the spatial transcriptome. As a result, the spatial heterogeneity of gene expressional pattern in cancer stem cell markers suggested the existence of multiple cancer stem cell-like populations in ILC specimens. In future, we plan to increase the number of samples and perform the analysis with artificial intelligence.