反捲積轉錄體分析揭示三陰性乳癌接受新輔助化學免疫治療後與非病理完全緩解相關的腫瘤細胞內在 Luminal-like 特徵

Tumour-Cell-Intrinsic Luminal-like Features Associated with Non-pCR in Triple-Negative Breast Cancer after Neoadjuvant Chemoimmunotherapy

Internship site：Taipei Veterans General Hospital

Authors: **Hsin-Ti Tsai[1], Yi-Chen Yang[1]** · Supervisors: **Tzu-Pin Lu, PhD[1], Chi-Cheng Huang, MD, PhD[2], Ping-Hsing Tsai, PhD[2]**

1 Department of Public Health, National Taiwan University · 2 Comprehensive Breast Health Center, Taipei Veterans General Hospital

Introduction

Triple-negative breast cancer (TNBC) lacks the **oestrogen receptor (ER)**, **progesterone receptor (PR)** and **human epidermal growth factor receptor 2 (HER2)** – that is, **ER−/PR−/HER2−** – so neither endocrine nor anti-HER2 therapy applies. Standard care adds **pembrolizumab** to neoadjuvant chemotherapy (KEYNOTE-522); **pathological complete response (pCR)** is the strongest survival surrogate[1,2], yet a third fail to reach it, with no pre-treatment marker.

Bulk RNA-seq measures the *sum* of all cell types; **BayesPrism**[3] recovers the *per-compartment profile*.

Objective. Is the non-pCR programme intrinsic to the cancer cells, and what is it?

Materials and Methods

1 CLINICAL COHORT

COHORT

n = 72 patients with triple-negative breast cancer

ER− / PR− / HER2− confirmed case by case · Taipei Veterans General Hospital

↓

TREATMENT · NEOADJUVANT

KEYNOTE-522–based regimen

Paclitaxel + carboplatin → doxorubicin + cyclophosphamide — All 72

Pembrolizumab (anti-PD-1 checkpoint inhibitor) — Subset

Pembrolizumab was reimbursed only recently, so exposure is not uniform across the cohort.

↓

ENDPOINT · PATHOLOGICAL RESPONSE

pCR = no residual invasive cancer in breast or nodes

pCR 36 | non-pCR 36

Groups are exactly balanced, so no group-size imbalance enters the comparison.

2 DATA GENERATION

Pre-treatment tumour tissue

Treatment-naive specimen taken before any systemic therapy

↓

RNA extraction · whole-transcriptome sequencing

Bulk RNA-seq, normalised to transcripts per million (TPM)

↓

Expression matrix · 25,333 genes × 72 samples

Each value is one mixed signal – the sum over all cell types present in the tumour

3 BAYESPRISM DECONVOLUTION

The single-cell reference acts as a prior; each gene's bulk reads are assigned to cell types by posterior probability (Gibbs sampling). Adapted from Chu et al., Nat Cancer 2022;3:505–517.

INPUT 1 · PRIOR

Single-cell reference φ

Wu 2021 atlas · TNBC, treatment-naive

15,149 cells · 9 types · S × G

INPUT 2 · OBSERVED

Bulk counts X_n

72 pre-treatment tumours

one mixed signal each · N × G

θ cell-type fractions

how much of each cell type is present

φ

BayesPrism

Z cell-type expression

what each cell type is expressing

θ-weighted sum of Z within each compartment

Cancer epithelial

Basal · LumA/B · Her2 · Cycling

cell-intrinsic programmes

Immune

T · B · myeloid · plasmablast

activation and effector state

CAF / stroma

myCAF · iCAF

fibrosis and ECM remodelling

Three compartment expression matrices (72 samples × genes) – the input to all downstream analyses.

Statistics and validation

- **Deconvolution.** BayesPrism[3] with the Wu 2021 breast atlas[4] as prior (three further atlases as dictionary controls, see footer).
- **Batch.** An **earlier 54** (pCR 32 / 22; **40 on ≥8 cycles of pembrolizumab**) and a **later 18** (pCR 4 / 14, no drug records); **pCR rates differ (59% vs 22%)**. They are merged **only after ComBat batch-effect correction**, and every comparison is **stratified by cohort (van Elteren)**, never a pooled Wilcoxon / Mann–Whitney U.
- **Scoring.** ssGSEA of **279** sets (**229 MSigDB C2:CGP** sets named for breast / ESR1 / ERBB2 + **all 50 Hallmarks**; **by set name only – all tumours are ER−/PR−/HER2−**) in **4 layers** (bulk + cancer / immune / CAF); 165 genes in 17 **hand-curated** programmes scored in the cancer compartment against **20,000** expression-matched nulls.
- **Effect size. Stratified** Cliff's **δ**: **positive = higher in pCR, negative = higher in non-pCR**, 0 = no difference.
- **Validation.** Deconvolved scores are not directly observable, so two routes. **Cross-platform**: against the **same patients'** routine pathology (AR stain, PD-L1 CPS, Ki-67 by correlation; TIL% as a discriminator of response). **External**: rescore them in **GSE25066**[5] (112 TNBC, chemotherapy only) and test direction agreement.

Results

The response signal localises to the cancer compartment

- *Proportions* barely separate responders: **none of 28 cell types survives FDR** (best δ = +0.40); of 6 compartments **only endothelial** does (δ = +0.35, FDR = 0.026, a vascular signal). *Expression* does (δ = **+0.68**, p = 1.3×10^{-5}): of 45 signature × compartment tests, 6 pass FDR < 0.05, **5 in the cancer compartment**.

In the cancer compartment, non-pCR runs a luminal-like metabolic programme

Of the 77-gene resistance panel scored in the cancer compartment, only **LAR / androgen** and **lipogenesis** pass programme-level permutation (FDR < 0.02): all 8 LAR and all 7 lipogenesis genes point to non-pCR, **AR** strongest (δ = **−0.39**, p = 0.0029), while every other programme – efflux pumps, stemness, EMT, DNA repair, hypoxia – falls inside the null (Figure 1). Only 6 genes reach p < 0.05 against 3.9 expected, none surviving correction.

Widening to **17 programmes** adds **cholesterol synthesis** (z = +3.9) and **OXPHOS** (+3.1) on the non-pCR side; the cancer-compartment **IFN-γ response / antigen presentation** (−8.0) and **basal / squamous** (−3.4) mark the pCR end.

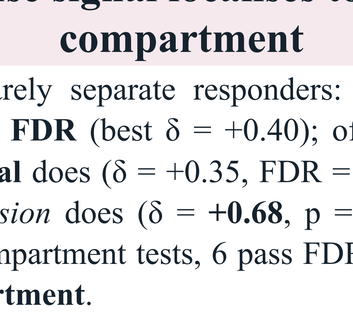

Figure 1 Seventy-seven resistance / stemness genes in the cancer compartment

n = 72. **Solid dark points** = programme at permutation FDR < 0.02.

Results (continued)

The same axis dominates an unbiased scan of 279 gene sets

Scored blind, **21 of 279** sets reach FDR < 0.25 and **12 belong to the ER/luminal axis, all pointing the same way**. van 't Veer ESR1_DN is strongest (bulk δ = **+0.53**, cancer compartment **+0.63**); Smid BASAL_DN likewise points to non-pCR.

Cross-platform: same-patient pathology

AR immunohistochemistry correlates with bulk AR mRNA (ρ = **+0.60**, p = 0.0006, n = 29) and with the cancer-compartment luminal (+0.48) and AR (+0.53) programmes, inversely with basal (−0.39): a routine stain measures the axis.

The immune side mirrors it: **stromal TIL%** (immune *quantity*) does not discriminate (δ = +0.04, p = 0.80), whereas immune-compartment cytolytic activity (immune *quality*) adds on top of it (likelihood-ratio p = **0.0013**); the reverse test is null (p = 0.95).

External replication: direction holds, effect size does not

In GSE25066 (112 TNBC, chemotherapy alone) direction agrees for **20 of 28** signatures (binomial p = 0.036), though no set reaches FDR < 0.25. The **AR programme is the informative failure** – internally δ = −0.26, externally −0.01. This may reflect the different treatment settings, but with pembrolizumab exposure not uniform here, **immunotherapy-specificity cannot be established**.

Discussion and Conclusion

- **An ER-negative cancer running an ER-driven programme.** All 72 tumours are ER-negative on IHC, yet the programme separating non-pCR is the ER/luminal one, running *inside the cancer cells* – the transcriptional **LAR subtype** of Lehmann *et al.*[6], recovered without assigning subtypes.
- **Two ends of one axis.** The pCR end is basal/squamous plus the **cancer-compartment IFN-γ-response / antigen-presentation programme** – the cancer cell's response to immune signalling, not IFN-γ of cancer-cell origin. Here pCR is only weakly associated with the **amount** of immune infiltrate, more clearly with the **cancer-cell-intrinsic programme**. The AR stain being routine, luminal/AR-high, basal-low TNBC is testable.
- **Limitations.** n = 72, single institution, discovery scale; no gene set reaches FDR < 0.05 and no gene survives correction; pembrolizumab exposure is not uniform.

Conclusion. Failure to achieve pCR is marked by a **tumour-cell-intrinsic luminal-like programme** – ER/luminal target genes, androgen-receptor signalling, lipid and cholesterol metabolism – mirrored by loss of the basal and cancer-compartment IFN-γ-response / antigen-presentation programmes: present **before treatment**, measurable with a routine AR stain.

References

1. Schmid P, *et al.* Event-free survival with pembrolizumab in early triple-negative breast cancer. *N Engl J Med* 2022;386:556–67.
2. Schmid P, *et al.* Pembrolizumab for early triple-negative breast cancer. *N Engl J Med* 2020;382:810–21.
3. Chu T, Wang Z, Pe'er D, Danko CG. Cell type and gene expression deconvolution with BayesPrism. *Nat Cancer* 2022;3:505–17.
4. Wu SZ, *et al.* A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet* 2021;53:1334–47.
5. Hatzis C, *et al.* A genomic predictor of response and survival following taxane–anthracycline chemotherapy. *JAMA* 2011;305:1873–81.
6. Lehmann BD, *et al.* Identification of human triple-negative breast cancer subtypes. *J Clin Invest* 2011;121:2750–67.

Deconvolution: BayesPrism (Chu *et al. Nat Cancer* 2022) · **Reference dictionaries:** Wu *et al. Nat Genet* 2021 (primary); Bassez 2021, Pal 2021, Qian 2020 (controls); CAF dropped from the dictionary (C+I+CAF → C+I): δ +0.61→+0.48, cancer fraction +26.7 points · **Regimen:** KEYNOTE-522 · **Gene sets:** MSigDB v2024.1.Hs (≥10 genes/set).

Protocol 11333 · Taipei Veterans General Hospital n = 72 · 2026