

Molecular Divergence Shapes Lung- and Liver-Tropic Metastasis in HER2-Positive Breast Cancer

Batul Al-zubeidy

University of Southern California

Edgar Gonzalez

University of Southern California

Aaron Baugh

University of Southern California

Marqus Dela Cruz

University of Southern California

Cheol Park

University of Southern California

Matthew Jacobo

University of Southern California

Isaac Chan

The University of Texas Southwestern Medical Center

Michael Press

University of Southern California

Evanthia T. Roussos Torres

`evanthia.roussosstorres@med.usc.edu`

University of Southern California

Article

Keywords:

Posted Date: March 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8991205/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. ERT is a paid consultant for Avix, Inc. and Synaptical Inc. M.F.P has a potential conflict of interest involving affiliations or financial relationships

with Novartis Pharmaceuticals Corp., TORL BioTherapeutics LLC, Jazz Pharmaceuticals, TRIO-US, and Iambic Therapeutics; declares a consultant role for Novartis Pharmaceuticals Corp.; reports ownership interest in TORL BioTherapeutics LLC and reports fee-for-service agreements from Novartis Pharmaceuticals Corp., TORL BioTherapeutics LLC, Jazz Pharmaceuticals, TRIO-US, and Iambic Therapeutics; and supported in part by grants from the Breast Cancer Research Foundation and the Dr. Miriam & Sheldon Adelson Medical Research Foundation and holds the Harold E. Lee Chair in Cancer Research. All other authors declare that no competing interests exist.

Version of Record: A version of this preprint was published at npj Breast Cancer on August 8th, 2026. See the published version at <https://doi.org/10.1038/s41523-026-01016-w>.

Abstract

Metastatic colonization of distant organs remains the primary cause of breast cancer specific mortality, yet the underlying mechanisms governing organ-specific tropism remain poorly understood. To better delineate the underlying mechanisms governing the establishment of successful niches in visceral organs, animal models recapitulating the pathogenesis of human breast cancer are necessary. Through *in vivo* selection, we established a liver-tropic HER2⁺ breast cancer murine model, NT2.5LV, which achieved 90% hepatic penetrance within 10 weeks. Comparative whole-genome DNA and bulk RNA sequencing of NT2.5LV, the lung-tropic counterpart NT2.5LM, and the parental primary breast tumor cell line NT2.5 revealed conserved genomic alterations, including *ErbB2* amplification, and *Cdkn2a/b* deletion despite their distinct organ-tropic phenotypes. Transcriptomic analysis showed that NT2.5LM incorporates DNA damage-immune adaptation with homologous recombination deficiency leading to cGAS-STING pathway activation coupled with immune checkpoint remodeling, and lung-enriched integrin-mediated adhesions. Conversely, NT2.5LV activated proliferative and metabolic programs with MMP-mediated ECM remodeling suitable for the hepatic microenvironment. These organ-specific strategies were validated in two independent clinical cohorts; the MET500 and AUORA datasets. Integrated multi-omic and immunohistochemistry analyses demonstrated that NT2.5LV recapitulates key features of human HER2⁺ breast tumors, including HER2 overexpression. Therapeutic testing showed that NT2.5LM's concurrent *Bcl2* overexpression confers PARP inhibitor resistance despite homologous recombination deficiency, a vulnerability reversed by *Bcl2* and PARP inhibition. The HER2-targeted tyrosine kinase inhibitor neratinib reduced primary and metastatic tumor burden *in vivo*, establishing NT2.5LV as a clinically relevant preclinical model to assess HER2-targeted therapies.

Introduction

Cancer metastasis accounts for more than 90% of cancer-related mortality and remains the major barrier to curative treatment in solid tumors.^{1,2} HER2-positive (HER2⁺) breast cancers comprise approximately 15–20% of cases and exhibit aggressive clinical behavior with distinct organ-specific metastatic tropism, most commonly involving liver (30–40%), lung (20–30%), brain (25–30%), and bone (20–25%).^{3,4} Various epidemiological studies identified liver metastases to be disproportionately enriched with HER2-positive disease, occurring at 3.13-fold higher odds in HR⁺HER2⁺ tumors and 4.75-fold higher odds in HR[−]HER2⁺ tumors compared with other subtypes.^{5,6} While treatment of metastatic breast cancer does not vary significantly across visceral involvement, organ-specific dissemination patterns have important prognostic implications. In patients with stage-IV disease, liver involvement is associated with substantially poorer prognosis and reduced responsiveness to HER2-targeted therapy, with shorter progression-free survival compared to other metastatic sites.³ Although HER2-targeted therapies, such as trastuzumab, pertuzumab, and tyrosine kinase inhibitors (TKIs), revolutionized early-stage disease treatment, metastatic HER2⁺ disease remains incurable. Hence, organ-specific colonization mechanisms and variable therapeutic responses among involved organs highlight the need to define site-specific metastatic programs to enable precision therapies.

A century after the “seed and soil” hypothesis first introduced by Stephen Paget, the principle that successful metastasis depends on compatibility between disseminated cancer cells and the organ microenvironment remains a core principle of cancer biology.^{2,7} Various clinical and molecular studies have demonstrated that organ-specific metastasis involves complex interactions requiring metabolic adaptation, immune evasion, extracellular matrix (ECM) remodeling, premetastatic-niche formation, and expression of organ-tropic gene signatures.^{8–12} However, it remains unclear whether organ-specific colonization reflects distinct genomic alterations, epigenomic modifications or transcriptional adaptations. Further studies are necessary to understand the complex interactions between tumor cells and their tumor microenvironment in the primary and metastatic sites to better design systemic and site-specific therapeutic targets.

However, addressing these important challenges require experimental systems that compare isogenic tumor cell populations with divergent organ tropism.¹³ Although the MMTV-HER2/*neu*N transgenic mouse model develops HER2-positive mammary adenocarcinomas, its translational utility for studying organ-specific metastasis is limited by prolonged latency (> 6 months), variable penetrance, and predominant metastatic lung tropism.^{14,15} These key limitations reflect a broader challenge given that most genetically engineered mouse models (GEMMs) progress to lung-predominant metastasis, failing to recapitulate the clinical diversity of HER2-positive metastases in patients.¹⁶ The NT2.5 breast and the lung-tropic NT2.5LM cell lines, both derived from MMTV-*neu*, provide accelerated growth kinetics and enhanced distant metastases upon orthotopic injection in immunocompetent *neu*-N transgenic FVB/N mice.^{17,18} However, a comparable liver-tropic variant has been lacking, limiting comparison of organ-specific metastatic programs within an isogenic framework. Here, we report the development of NT2.5LV, a liver-tropic HER2-positive derivative that facilitates direct molecular, immune and therapeutic analyses of the hepatic and pulmonary tumor microenvironments (TMEs).

Results

Development of NT2.5LV, a reproducible liver-tropic breast cancer mouse model

To generate a liver-tropic model, we performed iterative *in vivo* selection by orthotopic injection of 2.5×10^5 NT2.5 parental cells into the fourth mammary fat pad (MFP) of 8-week-old syngeneic MMTV-HER2/*neu* virgin female mice (Fig. 1a). We resected the primary tumors at day 32–40 and monitored for metastases. At day 80, the mice developed ascites, jaundice and lethargy, and were subsequently euthanized. Liver metastases were detected in less than 20% of examined mice. We harvested and dissociated the liver lesions into single-cell suspensions, cultured ex-vivo, and re-injected into the 4th MFP (Passage 1). Through 2 additional iterative passages, the penetrance of hepatic metastasis increased from < 20% to 95%. Coupled with liver metastases, most of the mice inoculated with NT2.5LV developed lung metastases by day 55–65. The resulting cell line, designated NT2.5LV, was isolated

following three iterative passages, and exhibited liver and lung metastases across > 10 subsequent passages.

Histopathological examination, reviewed by a board-certified pathologist, confirmed metastatic colonization of the liver and lung. Well-defined nodules were identified in the liver by day 20 and in the lung by day 22. We observed macrometastases by day 41 and 50, respectively (Fig. 1b). Our systematic organ survey (n = 20 mice) confirmed widespread metastasis to liver (95%), lung (90%), spleen (55%), adrenal glands (10%), ovary (5%), and colon (5%), recapitulating the multi-organ dissemination pattern of advanced human breast cancer (Fig. 1c).

To determine the epithelial nature and breast cancer markers of NT2.5LV, we performed immunohistochemical (IHC) profiling of the primary and metastatic tumors. IHC profiling demonstrated HER2 positive staining in primary tumor, liver, and lung metastases (Fig. 1d). NT2.5LV lacked estrogen receptor (ER) but expressed the luminal lineage marker GATA3 and epithelial markers cytokeratin 5/6 (Ck5/6). Ki67 proliferation index varied between 20–60% across tumor regions and metastases.

Correlation of tumor mutational burden with genomic architecture

To investigate the role of genomic alterations in metastatic dissemination, we performed whole-genome sequencing (WGS) of NT2.5 parental, NT2.5LM and NT2.5LV cell lines. We identified somatic single-nucleotide variants (SNVs), dinucleotide polymorphisms (DNPs), and small insertions/deletions (indels) using GATK Mutect2, and annotated genomic alterations using ANNOVAR (Methods). Tumor mutational burden (TMB), calculated as the number of amino acid-altering mutations per megabase (Mb), increased in NT2.5LM (2.730) and NT2.5LV (2.477) relative to NT2.5 parental cell line (1.014; Fig. 2a,

Supplementary Table 1), consistent with increased TMB in clinical metastases (MET500 and MSKTMet).^{20,21} Venn diagram analysis of protein-coding variants (SNVs, DNPs, and indels) revealed substantial overlap between the metastatic derivatives (Fig. 2b). All three cell lines shared 29 conserved mutations, likely representing early genomic alterations in parental NT2.5. The metastatic cell lines, NT2.5LM and NT2.5LV shared an additional 84 mutations not present in NT2.5, while each cell line shared only 1–2 unique mutations with the parental NT2.5 (**Supplementary Table 1**).

Consistent with early clonal divergence, SNV burden was 2-2.5-fold higher in NT2.5LM and NT2.5LV compared with the parental NT2.5 (Fig. 2c). Similarly, exonic indels increased in NT2.5LM and NT2.5LV, suggesting genome-wide instability (Fig. 2d). To determine the clinical relevance of identified mutations, we used OncoKB cancer-specific gene set. It revealed shared *ErbB2* in all cell lines, and metastatic-specific *Ercc5* and *Fat4* (tumor suppressors), with the proportion of OncoKB-annotated mutations increasing from 25% in parental NT2.5 to 75% and 65% in NT2.5LM and NT2.5LV, respectively (**Supplementary Fig. 1**).¹⁹

We performed copy-number variation (CNV) profiling using CNVkit package, which revealed conserved genomic architecture across the parental and metastatic lines (Fig. 2e). The three cell lines exhibited

focal amplification of *ErbB2* on chromosome 11 (3.3–6.7 copies), consistent with the MMTV-HER2/*neu* transgenic origin of this model. Homozygous deletion of the tumor suppressor locus harboring *Cdkn2a* (p16^{INKA4a}) and *Cdkn2b* (p15^{INK4b}) on chromosome 4 was identified. We noted broad deletions on chromosomes 7 and 13. Importantly, no cell line-specific focal amplifications or deletions were noted.

Clonal evolution trajectory reveals shared metastatic trunk with organ-specific branches

To reconstruct the evolutionary trajectory from the parental NT2.5 to NT2.5LM, and NT2.5LV, we used PyClone-VI. The analysis identified 18 distinct clones (**Supplementary Fig. 2a-c**) with Clone 9 representing the ancestral population harboring *ErbB2* overexpression, mutations in *Ksr2* and *Pclo*, and homozygous deletion in *Cdkn2a/b*.

The ancestral clone bifurcated into a parental-specific lineage (Clone 7) and “metastatic trunk” (Clones 0 and 3) shared by NT2.5LM and NT2.5LV. Clones 0 and 3 acquired *Ercc5*, which is involved in nucleotide excision repair, and *Fat4* involved in the Hippo/YAP pathway implicated in multiple cancers.^{20,21} NT2.5LM-specific clones (Clones 6 and 12) expanded to 40–70% and included alterations in *Sesn2* (mTOR pathway regulator), *Atm* (DNA damage response kinase), and *Ptptr* (receptor tyrosine phosphatase). Conversely, NT2.5LV-specific clones (Clones 2 and 10) reached 40–100% prevalence (**Supplementary Fig. 2c**).

We mapped identified genomic alterations to known cancer drivers and therapeutic targets (**Supplementary Fig. 2d**). Clone 9 harbors *ErbB2* overexpression, the target of approved HER2-directed therapies. Clones 0 and 3 acquired missense mutation in *Ercc5*. *ERCC5*-deficient tumors have been shown to have increased sensitivity to PARP inhibitors.²² NT2.5LM-specific clones revealed *Atm* and *Sesn2* mutations. ATM loss creates ATR dependency and thus targeted by ATR inhibitors, while *Sesn2* is targeted by mTOR antagonists.^{23,24} Despite high genomic variations, NT2.5LV-specific clones lacked well characterized druggable tumor drivers beyond the shared mutations.

Organ-tropic metastatic cell lines exhibit divergent transcriptional programs

To identify whether transcriptional programs underlie organ-tropic metastasis, we performed bulk RNA sequencing on NT2.5, NT2.5LM, and NT2.5LV (n = 3, 2, 3 biological replicates, respectively). Principal component analysis revealed unique transcriptional profile for examined cell lines (**Supplementary Fig. 3a**). Differential gene expression analysis using limma-voom package identified approximately 4000 and 2,000 differentially expressed genes (DEGs) in NT2.5LM and NT2.5LV relative to parental NT2.5, respectively (FDR < 0.05 and |log2FC| > 0.5; **Supplementary Fig. 1a-c**, **Supplementary Fig. 3b**, **Supplementary Fig. 4a-c**, and **Supplementary Table 2**). Direct comparisons of NT2.5LM vs. NT2.5LV revealed 2,113 DEGs (**Supplementary Fig. 2a**). Analysis of 40 curated cancer-associated genes including

those involved in cell cycle, HER2/EGFR, and EMT pathways revealed distinct expression pattern (**Supplementary Fig. 3c**). To identify cell line-specific programs, we derived gene signatures as described by Minn et al. (Fig. 3a).²⁵ The NT2.5LV-specific gene signature was enriched for matrix metalloproteinases (MMPs), while NT2.5LM-specific one was dominated by interferon-stimulated genes (*Irf9*, *Isg15*, *Usp18*). Gene enrichment analysis using Hallmark GSEA uncovered divergent transcriptional programs (Fig. 3b). To investigate the molecular basis of site-specific colonization, we compared transcriptional programs between NT2.5LM and NT2.5LV across three axes: ECM engagement, proteolytic remodeling, and metabolic adaptation.

ECM engagement and remodeling

NT2.5LV exhibited integrin reprogramming consistent with active engagement of the collagen- and fibronectin-rich liver ECM. It upregulated the laminin receptor *Itga7* (1.8 log₂FC) and the ECM-binding integrins *Itgb3* (1.1 log₂FC) relative to NT2.5 (**Supplementary Fig. 5a; Supplementary Table 5**). Moreover, NT2.5LV increased expression of *Mmp3* (~ 20-fold), *Mmp9* (~ 5.5-fold) and *Mmp24* (Fig. 3c, **Supplementary Table 3**). These data suggest that liver-tropic cells actively involved in ECM remodeling and stabilizing metastatic niches. In contrast, NT2.5LM upregulated *Itga8* (3.1 log₂FC), *Itgb8* (4.5 log₂FC), and the immune-associated *Itgb2* (2.5 log₂FC), while downregulating the collagen-binding *Itga1* (-7.3 log₂FC) and *Itga2* (-2.6 log₂FC). In contrast to NT2.5LV's proteolytic program, NT2.5LM exhibited features of epithelial barrier integration, with higher expression of tight junction proteins (*Cldn1*, *Cldn4*) and vascular adhesion molecules (*Ceacam1*, *Mcam*; **Supplementary Table 3**) consistent with a cohesive growth pattern.

Consistent with ECM-remodeling, NT2.5LV exhibited significantly higher invasive capacity than NT2.5LM in basement membrane Transwell assay (33.4 ± 1.9% vs. 24.3 ± 2.7%, P < 0.048; Fig. 3d) representing 37% increase in invasive capacity. MDA-MB-231 demonstrated robust invasion and served as a positive control given its highly aggressive triple-negative phenotype. This functional result confirms that NT2.5LV's proteolytic transcriptional program translates into enhanced basement membrane invasion.

Metabolic reprogramming

NT2.5LV exhibited glycolytic shift with marked upregulation of 11 of 19 glycolysis genes, including *Ldha* (log₂FC = + 2.08), *Pgk1* (log₂FC = + 2.01), and *Pkm* (log₂FC = + 1.8). This was accompanied by increased TCA-cycle gene expression, such as *Idh1* with 0.96 log₂FC, and active fatty acid oxidation program, highlighted by a ~ 6.2-fold increase in *Cpt1b*, the rate-limiting enzyme for mitochondrial fatty acid import. In contrast, NT2.5LM upregulated one-carbon metabolism (*Mthfd2*, *Phgdh*, and *Psat1*), which supply serine and glycine for de novo purine and pyrimidine biosynthesis.²⁶ Furthermore, it upregulated *Gpt* expression, which is involved in glutamine metabolism. This in turn may enhance the anaplerotic generation of α-ketoglutarate and TCA-cycle intermediates (**Supplementary Table 4**).²⁷

Clinical validation

To determine whether cell line-specific transcriptional signatures recapitulate human disease, we analyzed two independent clinical cohorts. In the MET500 cohort (n = 159 breast cancer metastases), liver metastases (n = 64) versus lymph node metastases (n = 36) showed significant upregulation of *LDHA* (2.08-fold, $p = 9.9 \times 10^{-5}$), *MMP7* (10.4-fold, $p = 2 \times 10^{-5}$) and *MMP10* (4.1-fold, $p = 0.001$) (Fig. 3f and **Supplementary Fig. 5b-d**).²⁸ The MET500 cohort had only 3 breast-to-lung metastases, which limited our comparison. In patient-matched primary tumor-metastasis pairs from the AURORA cohort (GSE209998), liver metastases showed significant upregulation of hallmark glycolysis genes (paired t-test, $P = 0.019$) and fatty acid metabolism ($P = 0.007$; Fig. 3g).²⁹ In contrast, lung metastases showed no significant change ($P = 0.261$ and $P = 0.0676$, respectively).

Metastatic derivatives show selective immune checkpoint remodeling

To determine whether organ-tropic metastatic cell lines exhibit immune checkpoint remodeling, we profiled the expression of cGAS-STING genes and T-cell checkpoint ligands across NT2.5-panel. Both NT2.5LM and NT2.5LV upregulated the cytosolic DNA sensor (cGAS) and the downstream signaling effectors (*Irf7* and *Mx1*) relative to parental NT2.5, consistent with constitutive cytosolic DNA sensing (**Supplementary Fig. 6a**). However, NT2.5LM showed significantly higher induction of interferon-stimulated effectors (*Oasl1*, *Oasl2*, *Isg15*, *Oas1a*, *Ifit1*, and *Ifit3*) than NT2.5LV. In contrast, while NT2.5LV showed similar cGAS upregulation as NT2.5LM, suggesting that both lines sense cytosolic DNA, NT2.5LV exhibited a more attenuated interferon signaling cascade.

This transcriptional divergence extended to immune checkpoint ligand expression, as shown in **Supplementary Fig. 6b**. NT2.5LM upregulated PD-L1 (Cd274; logCPM = 2.0) relative to both NT2.5 and NT2.5LV (logCPM = 0.5). Given that PD-L1 is a direct transcriptional target of interferon-JAK-STAT1 signaling, the observed upregulation is concordant with NT2.5LM's upregulated ISG molecular program.³⁰ Similarly, NT2.5LM upregulated *Ceacam1*, which functions as both a cell-cell adhesion molecule and T-cell inhibitory ligand.³¹ Together, these findings suggest that NT2.5LM tumor cells couple cGAS-STING-induced interferon overexpression with specific gain of inhibitory ligands, mainly PD-L1 and *Ceacam1*.

Metastatic breast cancer undergoes progressive mesenchymal-to-epithelial transition

To assess the epithelial-mesenchymal phenotypes of the three cell lines, we analyzed the RNA expression of 315 EMT gene signatures reported by Tan et al.³² Of these, we identified 291 genes (133 epithelial and 158 mesenchymal) in our dataset. NT2.5LM showed significant upregulation of epithelial markers (78/133; Fig. 4a-c, **Supplementary Table 5**). Both NT2.5LM and NT2.5LV downregulated the

EMT transcription master regulator, *Zeb1* (27-fold lower in NT2.5LM and 2.6-fold decrease in NT2.5LV). NT2.5LV upregulated certain mesenchymal markers (*Samsn1*, and *Srgn*) and epithelial markers (*Nqo1*).

To investigate this epithelial shift in clinical samples, we analyzed the TCGA primary breast tumors (n = 1,097), MET500, and AURORA metastatic tumors. Breast cancer metastases (MET500) to the lymph nodes and liver exhibited higher epithelial score compared to unmatched TCGA primary breast tumors (Fig. 4d). To determine whether this trend is breast cancer-specific or a generalized metastatic principle, we analyzed pan cancer liver and lymph node metastases. Pan-cancer analysis revealed that liver metastases (n = 243) showed higher epithelial scores compared to lymph node (n = 199) across cancer types (median = 0.347 vs 0.281; Fig. 4e). EMT-signature scores based on Hallmark Gene set in patient-matched primary tumors and metastases revealed that liver metastases had significantly lower EMT signature relative to primary breast tumors (P = 0.005; Fig. 4f).

Extended PARP inhibition reveals BCL2-mediated resistance in HRR-deficient breast cancer cell lines

To uncover the broad transcriptional phenotypes of each tumor cell line, we analyzed various pathway gene-sets by calculating the mean log2FC across all genes for each pathway (**Supplementary Fig. 10a**). It revealed that NT2.5LM downregulated its homologous recombination repair (HRR) genes, whereas NT2.5LV did not (Fig. 6a). Similar trends were seen in DNA replication and mismatch repair pathways (**Supplementary Fig. 10a**). Conversely, NT2.5LM upregulated cGAS-STING pathway genes, while NT2.5LV upregulates MYC gene sets consistent with its proliferative phenotype. Given the downregulation of cell cycle and DNA repair pathways in NT2.5LM, we examined the HRR-pathway genes essential for double-stranded DNA break repair. NT2.5LM downregulated 16 (76%) genes relative to NT2.5 (mean log2FC of -0.93, Fig. 6a, **Supplementary Table 4**).

Homologous recombination deficiency (HRD) generates genomic instability that leads to micronuclei formation and cytosolic DNA accumulation, which activates the cGAS-STING innate immune sensing pathway.^{33,34} Both NT2.5LM and NT2.5LV upregulate cGAS expression (log₂FC = + 1.69 and + 0.28, respectively; **Supplementary Table 4**). While NT2.5LM increased the expression of *Mx1*, *Isg15*, and *Ifi1* genes relative to NT2.5, NT2.5LV decreased its expression of *MX1* and *Isg15*.

Despite NT2.5LM's HRR-deficient transcriptional profile, extended pathway analysis revealed a concurrent and paradoxical shift in apoptosis-related gene expression. The anti-apoptotic *Bcl2* was significantly upregulated relative to NT2.5 (+ 2.3 logCPM; Fig. 6b), associated with overexpression of pro-apoptotic BH3-only proteins including *Hrk* (+ 7.2logCPM), *Bik* (+ 3.2), *Bmf* (+ 3.2), and downregulation of extrinsic pathway mediators *TNF* (-4.2) and *Bok* (-1.3; **Supplementary Fig. 10a**) among others.

To uncover transcriptional vulnerabilities, we performed in vitro dose response assays using olaparib (6-day) and cisplatin (72-hour) using six breast cancer cell lines. The human *BRCA1*-mutant line, MDA-MB-436 demonstrated the highest olaparib sensitivity (IC₅₀ = 0.75 μM) confirming vulnerability in HRR-deficient cells (Fig. 6c, **Table 1**). However, NT2.5LM showed significant resistance (IC₅₀ = 49 μM) despite

harboring HRR deficiency. Notably, cisplatin induced universal dose-dependent killing regardless of HRR status ($IC_{50} = 1-2 \mu M$; **Supplementary Fig. 10c**), while olaparib-cisplatin combination did not enhance efficacy.

Given that Bcl2 overexpression can confer PARP inhibitor resistance by blocking apoptotic execution downstream of DNA damage, we tested whether *Bcl2* inhibition could restore sensitivity to olaparib. ABT-199 monotherapy showed modest efficacy in NT2.5LM ($IC_{50} = 25 \mu M$) but minimal to modest effect in NT2.5LV and NT2.5, respectively (Fig. 6d, **Table 1**). Combination treatment with ABT-199 and olaparib ($25 \mu M$) induced synergistic tumor killing exclusively in NT2.5LM ($CI = 0.8$; Bliss EOB = + 0.090 where $CI < 1$ and $EOB > 0$ depicts synergy) even at low ABT-199 concentrations.

HER2-targeted therapy diminishes primary tumor growth and metastatic potential

NT2.5-based cell lines harbor *ErbB2* overexpression conferring HER2-dependent proliferation and potential sensitivity to HER2-targeted therapies. However, the NT2.5 *ErbB2* transgene encodes rat *ErbB2* under control of the mouse MMTV promoter.¹⁴ Thus we performed protein sequence alignment of the NT2.5LV *ErbB2* (inferred from transcriptome sequence) with the human *ERBB2* (UniProt P04626) to assess conservation of therapeutic target epitopes.³⁵

The extracellular domain (aa 23–652) revealed 74% conservation with variation of critical binding sites of pertuzumab (domain II) and trastuzumab (domain IV; Fig. 5a, **Supplementary Fig. 7**).³⁶ In vitro treatment with trastuzumab ($0.1-100 \mu g/mL$ for 72 hours) failed to reduce NT2.5LV viability, consistent with altered binding-site in the protein domain IV (**Supplementary Fig. 8c**). However, *in vivo* treatment with anti-HER2 monoclonal antibody revealed a modest 54% reduction in tumor volume (410 vs 880 mm^3 ; Fig. 5b). The combination treatment of anti-HER2 antibody and paclitaxel demonstrated significant synergy with 82% reduction in tumor volume. The intracellular tyrosine kinase domain (aa 676–1255) exhibited 92% conservation, including ATP-binding and TKI-binding residue sites. Neratinib, an irreversible pan-HER TKI, binds conserved Cys805 and Val773 sites, while lapatinib, a reversible EGFR/HER2 TKI, binds Leu726, Ala751, and Arg849 residues.^{37–39}

Based on amino acid sequence conservation, we hypothesized that TKIs would retain biological activity against tumor cells. Dose-response assays confirmed potent neratinib sensitivity with IC_{50} of 0.33, 0.22, and $0.16 \mu M$ for NT2.5LV, NT2.5LM and NT2.5, respectively (Fig. 5c; **Supplementary Fig. 8a**). For comparison, the triple-negative 4T1 line exhibited IC_{50} of $15.08 \mu M$, demonstrating > 45-fold selectivity for HER2-amplified NT2.5 lines. 4T1 appeared to retain low-level TKI sensitivity attributable to *ERBB2* copy number gain ($\log_2=0.088$) and *EGFR/ERBB2* expression, rather than complete target absence (**Supplementary Fig. 9a-c**). Lapatinib showed similar trends but with higher IC_{50} values (Fig. 5d; **Supplementary Fig. 8b**). The enhanced efficacy of neratinib relative to lapatinib is likely due to its irreversible inhibition.³⁹

To evaluate Neratinib's efficacy *in vivo*, *NeuN* mice injected with orthotopic 2.5×10^5 NT2.5LV cells in the fourth MFP were randomized to vehicle vs. neratinib (40mg/kg/day) when tumor size reached 6-7mm. Neratinib significantly suppressed tumor growth relative to vehicle (two-way ANOVA, $p < 0.01$; Fig. 5e). The treated tumors weighed 59% less than controls (430 mg vs. 1050 mg; unpaired t-test, $p = 0.048$; Fig. 5f).

At necropsy, we performed systematic organ survey, which revealed decreased visceral metastases in neratinib-treated mice. Vehicle-treated mice developed macroscopic liver metastases in 3/5 (60%), lung metastases in 3/5 (60%), iliac lymph node metastases in 4/5 (80%), and splenic metastases in 1/5 (20%; Fig. 5g). In contrast, neratinib-treated mice developed iliac lymph node metastases in 2/5 (40%) without any liver, lung, or splenic metastases (0/5). Histological analysis of iliac lymph nodes from vehicle- and Neratinib-treated mice confirmed metastatic adenocarcinoma cells infiltrating and compressing the cortical lymphoid follicles (Fig. 5h).

Discussion

Studies presented here provide evidence that site-specific colonization of our murine cell lines NT2.5LV and NT2.5LM are driven primarily by transcriptional programs rather than genomic alterations outside key cancer-specific drivers. Despite sharing conserved *ErbB2* overexpression, *Cdkn2a/b* homozygous deletion, and similar TMB, NT2.5LM and NT2.5LV express divergent transcriptional programs tailored to organ-specific colonization. NT2.5LM adopts a DNA damage-immune signaling program characterized by HRR suppression, cGAS-STING activation, and high cell-cell adhesion, whereas NT2.5LV employs a metabolic, proliferative program featuring Warburg-like glycolysis, fatty acid oxidation, MMP-mediated ECM remodeling, and rapid proliferation.

The NT2.5-model system offers unique benefits for preclinical metastasis research. Unlike xenograft model systems, this model is characterized by: (1) immunocompetent host (FVB/N *NeuN* mice), which preserves the tumor-immune-stromal interactions necessary to evaluate immune evasion and immunotherapies. (2) High organ-tropic fidelity with 80-90% site-specific colonization within 8-10 weeks, enabling mechanistic and therapeutic studies. (3) Isogenic model permitting investigations of organ-specific adaptations without the confounding heterogeneity of patient-derived models. (4) Recapitulates human HER2⁺ breast cancer features HER2 overexpression (15-20% of cases), *CDKN2A/B* deletion (16% of metastases), elevated TMB, and multi-organ dissemination.²⁸ Integrated genomic and transcriptomic characterization of NT2.5LV and NT2.5LM revealed that genomic evolution and transcriptional plasticity shape organotropism. NT2.5LV's colonization of multiple organs mirrors the complexity of advanced breast cancer. Collectively, the model's preserved immune context, predictable organ tropism, and multi-site colonization enable comprehensive investigation of primary and metastatic TMEs.

Organ-specific metastasis depends not only on environmental factors but also on intrinsic cancer-cell traits that enable dissemination and colonization.⁴⁰ Consistent with MET500 analyses showing that metastases exhibit higher TMB without organ-specific mutations, our WGS demonstrated conserved

genomic architecture in NT2.5LM and NT2.5LV, with transcriptional, rather than genomic, divergence underlying organ tropism.⁴¹ Furthermore, single-cell RNA (scRNA) sequencing revealed that primary breast tumors harbor extensive transcriptional heterogeneity with diverse subclones suggesting evolutionary branching prior to metastasis.⁴² The NT2.5 models extend these clinical observations, with NT2.5LV exhibiting the proliferative-metabolic rewiring and NT2.5LM the inflammatory-immune engagement.²⁸ These transcriptional signatures were found to be weakly associated with biopsy site suggesting that metastatic tropism emerges from transcriptional adaptation to organ-specific selective pressures rather than predetermined genomic programs.

This transcriptional branching may be explained by several non-mutually exclusive mechanisms. First, epigenetic reprogramming via DNA methylation and chromatin remodeling occurs in response to organ-specific microenvironmental cues.⁴³ Disseminated cancer cells encounter organ-specific metabolite gradients, oxygen tensions, and pH fluctuations in different organs. Such stimuli are likely to reprogram the cell transcriptional state through metabolite-sensing transcription factors such as PPAR α (lipid sensing), HIF-1 α (hypoxia sensing), and LXR/RXR (cholesterol sensing).⁴⁴ These factors are likely to drive NT2.5LV's hypermetabolic phenotype and NT2.5LM's cGAS-Sting/IFN-associated immune signature. Second, post-transcriptional regulation may be mediated by RNA-binding proteins, protein-protein interactions and noncoding RNAs. Third, selection of pre-existing NT2.5 subclones may enrich tumor cells with transcriptional programs compatible with liver- or lung-microenvironment.

Distinguishing whether adaptation mechanisms or clonal selection drive organ-specific metastasis requires scRNA sequencing of pre-injection NT2.5 cells with lineage tracing. Our WGS and bulk RNA sequencing analysis suggests both mechanisms contribute to site-specific colonization, as clonal analysis revealed shared "metastatic trunk" mutations with organ-specific transcriptional branches.

NT2.5LM's downregulation of HRR-related genes (*Brca1*, *Palb2*, and *Rad51*) raises a fundamental question: how do genomically unstable cells survive and colonize distant sites? Our findings suggest that NT2.5LM transforms this liability into a strategic advantage by activating the cGAS-STING pathway, which canonically is associated with immune-mediated cell death.^{33,34} Impaired DNA repair generates cytosolic DNA fragments and micronuclei, sensed by cGAS, triggering STING-dependent type I interferon signaling normally reserved for pathogen detection.⁴⁵ However, NT2.5LM not only tolerates this immune activation but appears to paradoxically depend on it, co-opting the resulting interferon response (*Isg15*, *Mx1*, *Ifit1*) to facilitate rather than suppress lung colonization.

Our data revealed that both lung- and liver-tropic derivatives upregulated cGAS expression, suggesting constitutive cytosolic DNA sensing in both cell lines. However, the downstream interferon response diverged between NT2.5LM amplified its ISG profile and NT2.5LV attenuated it. This dissociation between DNA sensing (cGAS-mediated) and interferon amplification suggests that cGAS activation alone is insufficient. NT2.5LM's successful immune co-option requires this sustained downstream signaling maintained by Bcl2-mediated survival program, which tolerates its constitutive genomic stress.

In contrast, NT2.5LV lacks this anti-apoptotic rewiring and suppresses the interferon signaling program to evade immune detection, relying on its hypermetabolic and proliferative strategies for colonization.

The cGAS-STING activation likely promotes lung colonization through multiple mechanisms, which will be examined in future studies. First, NT2.5LM employs immune evasion mechanisms via immune checkpoint (PD-L1 and *Ceacam1*) upregulation and chemokine secretion (Cxcl10, Ccl5). This selective strategy of inhibitory ligands coupled with amplified ISGs, suggests that NT2.5LM may exploit the cGAS-STING system to establish immune tolerance rather than induce immune-mediated cell death. Second, disseminated tumor cells activate survival signaling through *NF-κB* and *STAT1*, which could confer resistance to apoptosis.⁴⁶ Third, tumor cells may remodel the TME via inflammasome-mediated *IL-1β* and *IL-18* secretion.⁴⁷ NT2.5LM's unique upregulation of canonical interferon-stimulated genes (*Cxcl10*, *Isg15*, *Mx1*), coupled with NT2.5LV's downregulation of such programs, indicates that cGAS-STING activation is NT2.5LM-specific adaptation rather than a universal metastatic phenotype. This strategy likely succeeds in the lung given that it is an immune-enriched environment with alveolar macrophages, dendritic cells, and tissue-resident memory T cells.⁴⁸ NT2.5LM's interferon signature also likely pre-adapts tumor cells to this hostile environment by upregulating immune checkpoints prior to arrival. Thus, if this system is required for NT2.5LM to establish immune-tolerant niches in the lung, then cGAS-STING inhibition during dissemination may hinder lung colonization.

In addition to immune co-option, NT2.5LM enhances cell-cell adhesion through claudins (*Cldn1*, *Cldn4*), *Vcam1*, and *Ceacam1*, consistent with a cohesive growth pattern. Its integrin profile reveals lung-specific adaptations with upregulation of *Itga8* (normally expressed in lung tissue), *Itgb8* (*TGF-β* activation), and *Itgb2* (immune-cell interactions) while downregulating collagen-binding integrins (*Itga1*, *Itga2*). This shift from ECM engagement to claudin- and integrin-based cell-cell adhesion may protect metastatic niches from anoikis and insulate them from immune infiltration.⁴⁹

NT2.5LV exhibits a hypermetabolic, invasive, and proliferative signature adapted for hepatic microenvironment. Hepatocytes maintain high glycolytic flux, secrete glucose and store abundant triglyceride reserves.⁵⁰ They also form fenestrated sinusoids that allow direct contact with circulating cells, partially explaining the widespread hepatic metastases across cancer types.⁵¹ NT2.5LV adapts to this unique microenvironment through coordinated upregulation of glycolysis genes (11 of 19) and lipid metabolic enzymes (9 of 21) relative to both NT2.5 and NT2.5LM. The upregulation of *Cpt1b*, the rate-limiting enzyme for mitochondrial fatty acid import, allows NT2.5LV to exploit hepatic lipid stores for energy production via fatty acid oxidation, while sustained *Ldha*, *Pgk1*, and *Pkm* expression suggests aerobic glycolysis consistent with a Warburg-like phenotype rather than a response to glucose limitation.⁵² Notably, oxidative phosphorylation appeared comparable between NT2.5LM and NT2.5LV, indicating that organ-specific substrate utilization may be dependent on available nutrients rather than mitochondrial capacity. This metabolic profile was validated in the AURORA cohort, where patient-matched liver metastases rather than lung-metastases showed significant upregulation of both glycolysis and fatty acid metabolism. Consistent with our findings, pan-cancer analysis showed that

liver metastases from lung, esophageal and adrenocortical carcinomas upregulate adipogenesis and fatty acid metabolism pathways.⁴¹ Such metabolic reprogramming across various cancer types shows that targeting *Cpt1* (etomoxir, perhexiline) or glycolysis enzymes (2-deoxyglucose) may provide broad therapeutic benefit for liver metastases. In contrast, NT2.5LM upregulated one-carbon metabolism genes, which may be suitable for well-oxygenated pulmonary environment.

In addition to metabolic reprogramming, NT2.5LV demonstrates extensive ECM remodeling through MMP upregulation. NT2.5LV increased *Mmp3* expression by 20-fold relative to NT2.5LM, and 5.5-fold for *Mmp9*, making them among the most differentially expressed genes in the dataset. Functionally, *Mmp3* cleaves collagen III, IV, V, and IX, as well as laminin, fibronectin, and proteoglycans, thereby facilitating cellular transmigration across the space of Disse.⁵³ Moreover, *Mmp9* degrades denatured collagen and activates latent *TGF-β*, reshaping the ECM to favor of invasive growth.^{54,55} This proteolytic phenotype translated into functional invasive capacity, with NT2.5LV showing 37% higher basement invasion than NT2.5LM in Transwell assay. This MMP-driven hepatic colonization program is conserved in clinical breast-to-liver metastases as demonstrated by the MET500 cohort, where liver metastases showed significant upregulation of *MMP7* and *MMP10* relative to LN metastases. Furthermore, NT2.5LV uniquely upregulates *Itga7*, a laminin-binding integrin that facilitates basement membrane adhesion and is associated with hepatocellular carcinoma progression.⁵⁶ This dual strategy of MMP-mediated degradation coupled with selective integrin-mediated adhesion enables NT2.5LV to navigate the unique hepatic sinusoids, where matrix degradation alone would be insufficient without basement membrane attachment.

While EMT has long been viewed as essential for tumor cell dissemination, our data suggest mesenchymal-to-epithelial (MET) reversion may be necessary for successful colonization. Both NT2.5LM and NT2.5LV exhibited epithelial enrichment (78 and 71 of 133 epithelial genes, respectively) with ZEB1 suppression (27-fold reduction in NT2.5LM and 2.6-fold reduction in NT2.5LV relative to parental NT2.5). These findings align with emerging research revealing that EMT facilitates tumor cell dissemination and chemoresistance but epithelial reversion is necessary for distant-organ proliferation.^{57–59} Genetic deletion of EMT transcription factors (*Zeb1*, *Snail*, *Twist*) in pancreatic cancer mouse models paradoxically increased hepatic metastases through stabilization of epithelial programs.⁶⁰ Similarly, epithelial (CD49f⁺/EpCAM⁺/E-cadherin^{high}) populations from patient-derived xenografts of triple negative breast cancer harbor superior tumor-initiating capacity compared with mesenchymal-rich cells.⁵⁸ Extending these findings to human breast cancers, our breast and pan-cancer analyses of MET500 cohort revealed progressive epithelial gene enrichment from TCGA primary breast tumors through lymph node to liver metastases, with liver lesions showing the highest epithelial scores. This transcriptional pattern indicates that EMT is a dynamic process reflecting cancer cell plasticity to shift between different states depending on selective environmental pressures regardless of primary tumor site. Hence, inhibiting epithelial colonization machinery by targeting claudin-based tight junctions, E-cadherin adhesion complexes, and organ-specific integrins, may more effectively diminish metastatic outgrowth than disrupting EMT-associated dissemination.

In contrast, transcriptional biomarkers failed to predict therapeutic response in some contexts. Despite HRR deficiency in NT2.5LM (76% downregulation of pathway genes), PARP inhibitors showed limited efficacy ($IC_{50} = 49 \mu M$), well above maximum plasma concentration.⁶¹ This finding aligns with clinical observations demonstrating that genomic alterations do not consistently predict durable response to PARP inhibition in the absence of *BRCA1* or *BRCA2* mutations. Multiple resistance mechanisms have been described, including compensatory PARP1 overexpression, 53BP1/shieldin loss restoring HR proficiency, and replication fork protection via FANCD2/RADX.^{62–65}

Our data uncover a distinct resistance mechanism involving anti-apoptotic rewiring through *Bcl2* overexpression given its constitutive genomic stress. NT2.5LM demonstrated the highest TMB (2.733) alongside HRD, consistent with ongoing genomic damage. This instability engages the cGAS-STING cytosolic DNA sensing pathway. However, rather than inducing cell death, NT2.5LM simultaneously upregulates the anti-apoptotic gene, *Bcl2*, and pro-apoptotic BH3-only proteins (*Hrk*, *Bik*, *Bmf*, *Puma*, and *Noxa*); consistent with BH3 priming.⁶⁶ We propose that *Bcl2* overexpression could enable NT2.5LM to survive downstream effects from constitutive cGAS-STING activation and accumulated pro-apoptosis signaling, thus uncoupling genomic instability from cell death. Future studies will evaluate this mechanism further.

This anti-apoptotic dependency creates a pharmacologically targetable vulnerability. ABT-199 and olaparib combination induced synergistic cell death in NT2.5LM (CI = 0.80). NT2.5LM's specificity to this combination supports a mechanistic model in which PARP inhibition generates lethal DNA damage in HRR-deficient cells (BRCA-like), yet *Bcl2* overexpression prevents apoptosis. However, *Bcl2* inhibition can release this block and restore drug lethality. These findings are consistent with emerging evidence that BH3 mimetics can sensitize tumors to DNA-damage agents across multiple cancer types.⁶⁷

These data highlight direct translational implications. Venetoclax (*Bcl2*-antagonist) is FDA-approved for hematological malignancies and is under clinical investigations in solid tumors. Our data show that *Bcl2* overexpression and BH3 priming status may serve as functional biomarkers to identify breast cancer patients who may benefit from combined PARP and *Bcl2* inhibition; a therapeutic vulnerability that HRR status alone may not predict. Moreover, the concurrent cGAS-STING activation in NT2.5LM enhances the possibility that overcoming *Bcl2*-mediated apoptotic blockade could enhance STING-driven anti-tumor immunity, providing rationale for triple combination strategies incorporating immune checkpoint blockade. For NT2.5LV, with modest *Bcl2* overexpression but proliferative MYC-driven phenotype, CPT1 inhibition to block fatty acid oxidation, or glycolysis blockade to disrupt its Warburg-like metabolism may hold greater promise. Hence, these findings demonstrate that organ-specific metastatic programs exhibit distinct therapeutic vulnerabilities and demand precise and strategic targeting modalities.

Our study has several limitations that warrant discussion. First, our findings warrant validation in other HER2⁺ mouse breast tumor models (such as 4T1-HER2, MMTV-PyMT, MMTV-NeuT). Second, our analysis did not address key mechanistic gaps and thus future investigations are necessary. Finally, the rat *ErbB2* sequence may limit direct clinical translation, though conservation of TKI-binding site supports

relevance of the NT2.5 model system for kinase inhibitor and murine monoclonal antibody studies. In summary, our findings suggest that effective treatment of metastatic breast cancer should employ site-specific therapeutic strategies which could be informed by organ-signature transcriptional profiles.

Methods

Cell lines and culture conditions

NT2.5 parental cells were derived from MMTV-*Neu* transgenic mice (FVB/N-Tg(MMTV*neu*)202Mul/J; Jackson Laboratory JAX stock #002376). The NT2.5LM cells were generated through iterative *in vivo* selection for lung tropism, as previously described.¹⁸ NT2.5LV cells were generated in this study through serial passage of liver metastases as described in Results. MDA-MB-231, MDA-MB-436, and MCF7 cell lines were purchased from the American Type Culture Collection (ATCC, Rockville, MD). All NT2.5-based cell lines were cultured in RPMI supplemented with 20% fetal bovine serum (FBS; Gibco), 100 U/mL penicillin, 100 U/mL streptomycin, 0.1% nonessential amino acids, 0.1% glutamine and 10mM N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid (HEPES; STEMCELL Technologies, Vancouver, Canada) at 37°C in a humidified 5% CO₂ incubator. All human cell lines were cultured in DMEM supplemented with 10% fetal bovine serum (FBS; Gibco), 100 U/mL penicillin, and 100 µg/mL streptomycin at 37°C in a humidified 5% CO₂ incubator. All the cell lines used in *in vitro* and *in vivo* were tested negative for mycoplasma contamination (Charles River, Wilmington, MA).

Animal studies

All animal procedures were approved by the University of Southern California Institutional Animal Care and Use Committee and performed in accordance with NIH guidelines. All female FVB/N-Tg(MMTV*neu*) mice (*Neu*N; age 8-12 weeks) were purchased from Jackson Laboratory or bred in-house at USC animal facility. We injected 2.5×10^5 cells in NT2.5LV single-cells in 100 µL 1x phosphate buffered saline (PBS; Corning, Corning, NY) into the fourth mammary fat pad of 7–8-week-old virgin female mice. Primary tumors were resected when reached 15mm in diameter under isoflurane anesthesia with sustained-release buprenorphine (1 mg/kg SC). The tumor volumes were measured by caliper and calculated using the formula: volume = (length × width²)/2. Mice were euthanized when tumors reached end points. At necropsy, organs were harvested, fixed in 10% neutral buffered formalin for 48 hours, and maintained in 70% ethanol.

To evaluate Neratinib efficacy *in vivo*, mice bearing NT2.5LV tumors (6–7 mm diameter; Day 37 post-injection) were randomized to vehicle (0.5% methylcellulose; n = 5) or neratinib (40 mg/kg/day; n = 5; MedChemExpress) administered by oral gavage. We continued the treatment until experimental endpoint (Day 57). Tumor volumes and body weights were recorded every 3–4 days. At necropsy, primary tumors were dissected and weighed, and organs and apparent lymphadenopathy were examined for gross metastases.

To evaluate the efficacy of anti-HER2 monoclonal antibody *in vivo*, mice orthotopically injected with 2.5×10^5 NT2.5LV cells in 100 μ L PBS were randomized into the following groups: a) 100 μ L IgG2a isotype (1mg/mL; BioXCell) weekly intraperitoneal injection (vehicle), b) 100 μ L anti-Her2 monoclonal antibody (1mg/mL; BioXCell) weekly intraperitoneal injection, or c) 300 μ L paclitaxel (1mg/mL; Selleck Chem) twice per week intraperitoneal injection (30mg/kg/mouse/week), or d) combination of paclitaxel and anti-HER2 antibody. Control group included 3 mice, while treatment groups included 5. Treatment began on day 7 post injection. At endpoint, the mice were euthanized at day 36 post injection.

Histology and immunohistochemistry (IHC)

To evaluate the epithelial and tumor markers in NT2.5LV, tumors were formalin-fixed for 24-48 hours, and stored in 70% alcohol. The samples were embedded in paraffin (FFPE), and 3-5 μ m slides were prepared by the USC Clinical Immunohistochemistry and Research Core using standard protocols. All subsequent hematoxylin and eosin (H&E) and immunohistochemistry staining (IHC) were performed at the same core. Briefly, tissue sections were deparaffinized, rehydrated, and subjected to heat-induced antigen retrieval as per antibody protocols.

Whole-genome sequencing and analysis

Genomic DNA was extracted from cell lines using the DNeasy Blood & Tissue Kit (Qiagen). PCR-free whole-genome sequencing libraries were prepared with the TruSeq DNA PCR-Free Library Prep Kit (Illumina) and sequenced by Novogene on a NovaSeq 6000 platform (2×150 bp paired-end) to a target depth of $\sim 30\times$. Raw reads were assessed with FastQC and summarized using MultiQC, and were aligned to the mouse reference genome (GRCm38/mm10) using BWA-MEM (v0.7.17).^{68–70} Somatic single-nucleotide variants (SNVs) and indels were called with GATK Mutect2 (v4.2.0) in tumor–normal mode, using FVB/N genomic DNA as the matched normal reference. Variants were filtered with FilterMutectCalls, normalized with *bcftools norm*, and functionally annotated with ANNOVAR using the mm10 RefGene database.^{71–74} Tumor mutational burden (TMB) was calculated as the number of exonic protein-altering mutations (nonsynonymous SNVs, stopgain/stoploss, frameshift and nonframeshift indels) per megabase of callable exonic sequence (~ 71 Mb).

Copy-number variation (CNV) analysis was performed using CNVkit (v0.9.11).^{75–77} Bin-level $\log_2$ copy ratios were calculated relative to pooled normal references, segmented using circular binary segmentation (CBS), and converted to absolute copy numbers ($CN = 2 \times 2^{\log_2 \text{ratio}}$). Gene-level copy numbers were extracted and plotted in Python 3.12 using matplotlib and pandas. Clonal evolution analysis was performed using PyClone-VI (v0.1.3).⁷⁸ High-confidence somatic SNVs with $\geq 10\times$ coverage were used as the input along with variant allele frequencies (VAFs). PyClone inferred cancer cell fractions (CCF) and clustered mutations into clones. Phylogenetic relationships were reconstructed based on clone prevalence across samples.

RNA sequencing and analysis

We used RNeasy Mini Kit (Qiagen) to extract and isolate total RNA from NT2.5, NT2.5LM, and NT2.5LV cell lines. Biological replicates included three samples each for NT2.5 and NT2.5LV, and two samples for NT2.5LM, harvested at 70-80% confluence. We used Agilent 2100 Bioanalyzer to determine the RNA integrity (RIN > 8.0). Library preparation and sequencing were performed by Novogene using the TruSeq Stranded mRNA Library Prep Kit (Illumina), followed by paired-end sequencing on the NovaSeq 6000 platform (2 × 100 bp; ~30 million reads per sample). The quality of the raw data was assessed using FastQC and MultiQC. The reads were trimmed for adapters. The data was aligned to the mouse reference genome (mm10) with STAR (v2.7.9a), and used RSEM (v1.3.3) for gene level quantification.^{79,80}

We normalized read counts using edgeR (v3.38.0) with trimmed mean of M-values (TMM) normalization with log₂-counts-per-million (logCPM) transformation.⁸¹ Differential gene expression analysis used the limma-voom R package with empirical Bayes moderation of gene-wise variances (shown in **Supplementary Fig. 4**).^{82,83} To account for the small size number, particularly NT2.5LM with n=2, we used linear models to fit each gene, and moderated t-statistics were calculated using the empirical Bayes shrinkage. P-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR). Genes were considered significant if the |log₂FC| > 0.5 and FDR < 0.05.

To assess ERBB family and kinase target gene expression in 4T1 tumor cell line, we utilized publicly available bulk RNA-seq data (SRA accession: SRR17179922). Reads were aligned to mm10 using STAR, and gene expression was quantified as read counts.

Gene set enrichment

Gene set enrichment analysis was performed with fgsea (v1.20.0) using MSigDB Hallmark, KEGG, Reactome, and Gene Ontology Biological Process gene sets.^{84–88} Pathway-level scores were computed as the mean log₂ fold change of all genes within each pathway. Data visualization, including expression heatmaps, bar plots, and enrichment summaries, was performed in Python (NumPy, Pandas, Matplotlib, Seaborn). To define transcriptional programs associated with lung and liver-metastases, normalized gene expression values and differential expression statistics were obtained for the NT2.5LM vs. NT2.5, NT2.5LV vs. NT2.5, and NT2.5LM vs. NT2.5LV. We first applied an expression filter requiring a maximum normalized expression ≥ 5 in at least one of the three lines to exclude low-abundance transcripts. Genes comprising the NT2.5LM-up signature (LM_UP) were defined as those significantly upregulated in NT2.5LM compared with both NT2.5 and NT2.5LV ((FDR) < 0.05 and abs. log₂ FC ≥ 1.585, corresponding to ≥3-FC, for both NT2.5LM vs. NT2.5 and NT2.5LM vs. NT2.5LV). On the other hand, NT2.5LM downregulated genes (LM_DOWN) in NT2.5LM relative to both NT2.5 and NT2.5LV using the same FDR and FC thresholds. The differentially expressed genes in NT2.5LV (LV_UP and LV_DOWN) signatures were defined similar to NT2.5LM.

TCGA primary breast cancer tumor analysis

We used the UCSC Xena data portal to access the gene expression for invasive breast cancer (BRCA) 1,097 primary tumors. The Fragments Per Kilobase Million (FPKM) values were $\log_2$ -transformed ($\log_2(\text{FPKM}+1)$) for subsequent analysis.

MET500 metastatic samples

The gene expression values of metastases from breast-to-liver, breast-to-lymph nodes, pan-cancer liver, and pan-cancer lymph nodes downloaded as $\log_2$ (FPKM+1) with the clinical metadata.

AURORA breast cancer cohort

Gene expression data from patient-matched primary breast tumors and metastases were obtained from the AURORA study dataset (GSE209998). The dataset included 129 samples from 53 breast cancer patients with RNA-sequencing data including 18 matched primary-liver metastasis pairs and 8 matched primary-lung metastasis pairs. Upper quartile normalized (UQN) expression values were used for pathway analysis.

EMT gene expression analysis

We assessed EMT gene expression across the three cell lines using the curated 315 tumor-specific EMT genes reported by Tan et al (2014). We identified 291 genes (of the 315) including 133 epithelial and 158 mesenchymal in our bulk RNA-seq dataset. The gene expression values (TMM-normalized logCPM) were compared across the parental NT2.5 and its metastatic derivatives.

EMT score calculation

We then calculated the EMT scores using the two-sample Kolmogorov-Smirnov (2KS) test as described by Tan et al (2014). This involved calculating the empirical cumulative distribution functions (ECDFs) for epithelial and mesenchymal gene expression values. The EMT score represented the maximum difference between the two ECDFs, and the sign indicates the direction of the EMT balance whereby a positive sign indicates epithelial character while a negative sign indicates a mesenchymal one. Cell lines were classified as epithelial if $\text{ECDF}_{\text{ep}} > \text{ECDF}_{\text{mes}}$, or mesenchymal ($\text{ECDF}_{\text{ep}} < \text{ECDF}_{\text{mes}}$). This method was also applied to the TCGA primary breast tumors, and the MET500 metastatic tumors. The associated p-value from the KS test indicates the statistical significance. To compare the EMT scores between the cell lines, we used two-tailed Student's t-tests given the small sample sizes (NT2.5 and NT2.5LV with n=3 biological replicates, and NT2.5LM with n=2 replicates).

For patient-matched samples in AURORA cohort, EMT signature scores were calculated using HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION gene set (MSigDB2025.1,200 genes) as the mean expression across available pathway genes. Paired two-tailed t-test was used to assess changes in EMT scores across matched primary tumors and metastases.

EMT gene-level analysis

To obtain the comprehensive EMT gene level analysis across the cell lines (Supplementary Figure 6 and Supplementary Table 5), gene expression levels were first averaged across the biological replicates of each cell line. The Log2FC was calculated by subtracting the NT2.5 expression from the NT2.5LM or NT2.5LV. The differences in expression between the cell lines were assessed using paired Wilcoxon signed-rank tests separately. Additionally, we used the Mann-Whitney U test to compare gene expression distributions of epithelial vs mesenchymal within each cell line. To perform statistical analysis between independent clinical datasets with large sample size (TCGA and MET500) and multiple tissue types, non-parametric Mann-Whitney U tests (two-sided) was used with statistical significance defined as $p < 0.05$.

Drug sensitivity assays

Tumor cells (5,000 per well) were incubated in 96-well plate for 24 hours in 5% CO₂ and 37°C incubator. They were subsequently treated with cisplatin (0.01-50 µM; Sigma -Aldrich, St. Louis, MO), olaparib (0.01-100 µM; Selleck Chemicals, Houston, TX), cisplatin + olaparib combination (0.01-100 µM), neratinib (0.001-50 µM; Selleck Chemicals), lapatinib (0.001-100 µM; Selleck Chemicals), trastuzumab (0.1-200 µM; Selleck Chemicals), ABT-199 (0.1-50 µM; Selleck Chemicals), or vehicle (DMSO for all except cisplatin, or PBS for cisplatin at 0.1% final concentration) for 72 hours at 37°C in a humidified 5% CO₂ incubator. Cells treated with single-agent olaparib were incubated for 6 days with new media supplied with the drug replaced after 96 hours. Cell viability was measured using CellTiter-Glo (Promega, Madison, WI) according to manufacturer's protocol, and read on a ClarioStar plate reader. Experimental conditions were performed with two to four biological replicates. To determine the dose-response curves for each drug, cell viability at each tested concentration was normalized to vehicle control (set at 100% viability). Dose-response curves and IC₅₀ measurements were computed by fitting a four-parametric logistic (4PL) model using Python (SciPy curve_fit), with replicate means plotted ± SEM. The statistical significance between the cell lines at each given concentration of a drug was determined using one-way ANOVA test. For pairwise comparisons, two-sample t-tests with Bonferroni correction for multiple testing were performed. For trastuzumab single-cell-line in vitro experiment, one-sample t-test was used to compare each concentration to control (set at 100% viability), with Bonferroni correction used to compare the multiple concentrations. Significance was defined as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All other values were determined to be not significant (ns). Statistical analyses were performed using Python (SciPy stats module).

Drug combination analysis

Drug combination synergism (ABT-199 and olaparib) was assessed using the Chou-Talalay method and Bliss Excess over Bliss (EOB).⁸⁹⁻⁹¹ The combination index (CI) was calculated using the Chou-Talalay method and results interpreted as followed: CI<1, synergy; CI>1, antagonism). EOB was calculated as the difference between observed and expected effects (expected = $F_{a1} + F_{a2} - (F_{a1} \times F_{a2})$, where $F_a = 1 - \text{Survival Fraction}$). EOB > 0 identifies synergy; EOB<0 indicates antagonism.

Invasion assay

Cell invasion capacity was tested using a 96-well basement membrane invasion assay (Abcam ab235887, Cambridge, MA). Briefly, 8- μ m-pore cell migration inserts were coated with 100 μ L collagen I and incubated overnight. Cells were serum-starved for 18-24 hours prior to assay. 50,000 cells per well in serum-free medium were seeded into upper chambers, while lower chambers contained medium with 10% FBS as chemoattractant. Following 48-hour incubation at 37°C with 5% CO₂. Invaded cells in the lower chamber were dissociated with Cell Dissociation Solution containing fluorescent Cell Dye. Fluorescence was assessed at Ex/Em 530/590 nm using CLARIOstar plate reader. Cell numbers were calculated from cell line-specific standard curves. Invasion index was calculated as (invaded cells/total seeded cells) * 100%.

Statistical analysis

Data are presented as mean $\pm$ SEM unless otherwise stated. Comparisons between two groups were performed using unpaired two-tailed Student's t-tests. Tumor growth curves were analyzed by two-way ANOVA with Šídák post-hoc tests. Fisher's exact test was used for metastasis frequency comparisons. P-values < 0.05 were considered statistically significant. Statistical analyses were performed in GraphPad Prism 9 and Python (NumPy, SciPy, and Matplotlib).

Declarations

Data Availability

WGS and bulk RNA-sequencing from NT2.5, NT2.5LM, and NT2.5LV cell lines will be deposited in the Gene Expression Omnibus (GEO). The WGS of 4T1 will also be deposited in GEO.

The MET500 dataset used for validation is publicly available through UCSC Xena Browser ([https://xenabrowser.net/datapages/?cohort=MET500%20\(expression%20centric\)](https://xenabrowser.net/datapages/?cohort=MET500%20(expression%20centric))) with processed expression data and metadata. The MET500 expression matrix (MET500_log2FPK.txt.gz), metadata (M.meta.plus.txt), and gene mapping files (gencodev233.annotation.gene.probemmap) used in this study were obtained from this repository (Robinson et al., 2017). The TCGA BRCA data are publicly available through the UCSC Xena browser (<https://xenabrowser.net>). The AURORA breast cancer cohort was accessed under accession GSE209998 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE209998>). Public RNA-seq data for 4T1 cells were obtained from GEO under accession GSE190533 (sample GSM5725352, SRA run SRR17179922) and used for characterization of ERBB family gene expression.

All custom code used for data analysis is available from the corresponding author upon reasonable request.

Ethics

For all mouse studies, the University of Southern California adheres to the strictest of guidelines as outlined by the Office of Laboratory and Animal Welfare (OLAW) and maintains a Health Science

Campus-specific Institutional Animal Care and Use Committee (IACUC), which reviews all protocols and rigorously evaluates their adherence to federal guidelines, OLAW/PHS animal health assurance number #21170.

Author Contributions

Conceptualization: B.A., E.R.T.

Methodology: B.A., E.G., A.B., M.D., M.J., C.P., and I.C.

Data Curation: B.A., E.G., A.B., M.D., M.J., and C.P.

Formal analysis and validation: B.A., M.D., A.B., M.P.

Supervision: M.P., E.R.T.

Writing (original draft): B.A., M.D.

Writing review and editing: B.A., A.B., C.P., M.D., E.G., M.J., I.C., M.P., E.R.T.

Funding acquisition: E.R.T., M.P.

All authors read and approved the final manuscript: B.A., E.G., A.B., M.D., M.J., C.P., I.C., M.P., E.R.T.

Competing Interests

ERT is a paid consultant for Avix, Inc. and Synaptical Inc.

M.F.P has a potential conflict of interest involving affiliations or financial relationships with Novartis Pharmaceuticals Corp., TORL BioTherapeutics LLC, Jazz Pharmaceuticals, TRIO-US, and Iambic Therapeutics; declares a consultant role for Novartis Pharmaceuticals Corp.; reports ownership interest in TORL BioTherapeutics LLC and reports fee-for-service agreements from Novartis Pharmaceuticals Corp., TORL BioTherapeutics LLC, Jazz Pharmaceuticals, TRIO-US, and Iambic Therapeutics; and supported in part by grants from the Breast Cancer Research Foundation and the Dr. Miriam & Sheldon Adelson Medical Research Foundation and holds the Harold E. Lee Chair in Cancer Research.

All other authors declare that no competing interests exist.

Acknowledgements

We would like to acknowledge the core facilities at USC Norris Comprehensive Cancer Center, including the Molecular Genomics Core, Translational Pathology Core, Flow Cytometry and Immune Monitoring Core, and Stem Cell Flow Cytometry Core, which were supported by award number P30CA014089 from the National Cancer Institute. The content is solely the responsibility of the authors and does not

necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

This study was also supported by the NCI R01CA283169, 1P20CA290458 (ERT), Tower Cancer Research Foundation (ERT), American Association for Cancer Research (AACR) and the Breast Cancer Research Foundation (BCRF) Grant (ERT), METAVivor Foundation (ERT), Wright Foundation (ERT), Concern Foundation (ERT).

References

1. Gupta, G. P. & Massagué, J. Cancer Metastasis: Building a Framework. *Cell* **127**, 679–695 (2006).
2. Valastyan, S. & Weinberg, R. A. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* **147**, 275–292 (2011).
3. Wang, R. et al. The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer. *BMC Cancer* **19**, 1091 (2019).
4. Metastatic Behavior of Breast Cancer Subtypes | Journal of Clinical Oncology.
https://ascopubs.org/doi/10.1200/JCO.2009.25.9820?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%20pubmed.
5. Ji, L. et al. Risk and prognostic factors of breast cancer with liver metastases. *BMC Cancer* **21**, 238 (2021).
6. Gong, Y., Liu, Y.-R., Ji, P., Hu, X. & Shao, Z.-M. Impact of molecular subtypes on metastatic breast cancer patients: a SEER population-based study. *Sci. Rep.* **7**, 45411 (2017).
7. Paget, S. THE DISTRIBUTION OF SECONDARY GROWTHS IN CANCER OF THE BREAST. *The Lancet* **133**, 571–573 (1889).
8. Brasó-Maristany, F. et al. Gene expression profiles of breast cancer metastasis according to organ site. *Mol. Oncol.* **16**, 69–87 (2022).
9. Elia, I. et al. Breast cancer cells rely on environmental pyruvate to shape the metastatic niche. *Nature* **568**, 117–121 (2019).
10. Ebright, R. Y. et al. Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. *Science* **367**, 1468–1473 (2020).
11. Oskarsson, T. et al. Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat. Med.* **17**, 867–874 (2011).
12. Peinado, H. et al. Pre-metastatic niches: organ-specific homes for metastases. *Nat. Rev. Cancer* **17**, 302–317 (2017).
13. Montero-Calle, A. et al. Metabolic Reprogramming Helps to Define Different Metastatic Tropisms in Colorectal Cancer. *Front. Oncol.* **12**, 903033 (2022).
14. Guy, C. T. et al. Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc. Natl. Acad. Sci.* **89**, 10578–10582 (1992).

15. Muller, W. J., Sinn, E., Pattengale, P. K., Wallace, R. & Leder, P. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. *Cell* **54**, 105–115 (1988).
16. Taneja, P. *et al.* MMTV mouse models and the diagnostic values of MMTV-like sequences in human breast cancer. *Expert Rev. Mol. Diagn.* **9**, 423–440 (2009).
17. Song, H. *et al.* An Immunotolerant HER-2/ *neu* Transgenic Mouse Model of Metastatic Breast Cancer. *Clin. Cancer Res.* **14**, 6116–6124 (2008).
18. Baugh, A. G. *et al.* A new Neu—a syngeneic model of spontaneously metastatic HER2-positive breast cancer. *Clin. Exp. Metastasis* **41**, 733–746 (2024).
19. Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis. Oncol.* **2017**, PO.17.00011 (2017).
20. KATO, M. Function and cancer genomics of FAT family genes. *Int. J. Oncol.* **41**, 1913–1918 (2012).
21. Muniesa-Vargas, A., Theil, A. F., Ribeiro-Silva, C., Vermeulen, W. & Lans, H. XPG: a multitasking genome caretaker. *Cell. Mol. Life Sci. CMLS* **79**, 166 (2022).
22. Mesquita, K. A. *et al.* ERCC1-XPF deficiency is a predictor of olaparib induced synthetic lethality and platinum sensitivity in epithelial ovarian cancers. *Gynecol. Oncol.* **153**, 416–424 (2019).
23. Gong, L., Wang, Z., Wang, Z. & Zhang, Z. Sestrin2 as a Potential Target for Regulating Metabolic-Related Diseases. *Front. Endocrinol.* **12**, 751020 (2021).
24. Leibrandt, R. C., Tu, M.-J., Yu, A.-M., Lara, P. N. & Parikh, M. ATR Inhibition in Advanced Urothelial Carcinoma. *Clin. Genitourin. Cancer* **21**, 203–207 (2023).
25. Minn, A. J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
26. Locasale, J. W. Serine, glycine and the one-carbon cycle: cancer metabolism in full circle. *Nat. Rev. Cancer* **13**, 572–583 (2013).
27. Yoo, H. C., Yu, Y. C., Sung, Y. & Han, J. M. Glutamine reliance in cell metabolism. *Exp. Mol. Med.* **52**, 1496–1516 (2020).
28. Robinson, D. R. *et al.* Integrative Clinical Genomics of Metastatic Cancer. *Nature* **548**, 297–303 (2017).
29. Garcia-Recio, S. *et al.* Multiomics in primary and metastatic breast tumors from the AURORA US network finds microenvironment and epigenetic drivers of metastasis. *Nat. Cancer* **4**, 128–147 (2023).
30. Moon, J. W. *et al.* IFN γ induces PD-L1 overexpression by JAK2/STAT1/IRF-1 signaling in EBV-positive gastric carcinoma. *Sci. Rep.* **7**, 17810 (2017).
31. Chen, Z., Chen, L., Qiao, S.-W., Nagaishi, T. & Blumberg, R. S. Carcinoembryonic Antigen-Related Cell Adhesion Molecule 1 Inhibits Proximal TCR Signaling by Targeting ZAP-70. *J. Immunol. Baltim. Md 1950* **180**, 6085–6093 (2008).
32. Tan, T. Z. *et al.* Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol. Med.* **6**, 1279–1293 (2014).

33. Harding, S. M. *et al.* Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature* **548**, 466–470 (2017).
34. Samson, N. & Ablasser, A. The cGAS–STING pathway and cancer. *Nat. Cancer* **3**, 1452–1463 (2022).
35. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2022).
36. Cho, H.-S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* **421**, 756–760 (2003).
37. Sohrab, S. S. & Kamal, M. A. Screening, Docking, and Molecular Dynamics Study of Natural Compounds as an Anti-HER2 for the Management of Breast Cancer. *Life* **12**, 1729 (2022).
38. The Development of HKI-272 and Related Compounds for the Treatment of Cancer - Wissner - 2008 - Archiv der Pharmazie - Wiley Online Library.
<https://onlinelibrary.wiley.com/doi/10.1002/ardp.200800009>.
39. Rabindran, S. K. *et al.* Antitumor Activity of HKI-272, an Orally Active, Irreversible Inhibitor of the HER-2 Tyrosine Kinase. *Cancer Res.* **64**, 3958–3965 (2004).
40. Obenauf, A. C. & Massagué, J. Surviving at a distance: organ specific metastasis. *Trends Cancer* **1**, 76–91 (2015).
41. Robinson, D. R. *et al.* Integrative Clinical Genomics of Metastatic Cancer. *Nature* **548**, 297–303 (2017).
42. Yates, L. R. *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).
43. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science* **357**, eaal2380 (2017).
44. Gentric, G. *et al.* PML-Regulated Mitochondrial Metabolism Enhances Chemosensitivity in Human Ovarian Cancers. *Cell Metab.* **29**, 156-173.e10 (2019).
45. Chen, Q., Sun, L. & Chen, Z. J. Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nat. Immunol.* **17**, 1142–1149 (2016).
46. Woo, S.-R. *et al.* STING-Dependent Cytosolic DNA Sensing Mediates Innate Immune Recognition of Immunogenic Tumors. *Immunity* **41**, 830–842 (2014).
47. Bakhoum, S. F. *et al.* Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature* **553**, 467–472 (2018).
48. Braciale, T. J., Sun, J. & Kim, T. S. Regulating the adaptive immune response to respiratory virus infection. *Nat. Rev. Immunol.* **12**, 295–305 (2012).
49. Aceto, N. *et al.* Circulating Tumor Cell Clusters are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell* **158**, 1110–1122 (2014).
50. Rui, L. Energy Metabolism in the Liver. *Compr. Physiol.* **4**, 177–197 (2014).
51. Structure and Function of Sinusoidal Lining Cells in the Liver.
<https://journals.sagepub.com/doi/epdf/10.1177/019262339602400114>
[doi:10.1177/019262339602400114](https://doi.org/10.1177/019262339602400114).

52. Pascual, G. *et al.* Targeting metastasis-initiating cells through the fatty acid receptor CD36. *Nature* **541**, 41–45 (2017).
53. Cui, N., Hu, M. & Khalil, R. A. Biochemical and Biological Attributes of Matrix Metalloproteinases. *Prog. Mol. Biol. Transl. Sci.* **147**, 1–73 (2017).
54. Vandooren, J., Van den Steen, P. E. & Opdenakker, G. Biochemistry and molecular biology of gelatinase B or matrix metalloproteinase-9 (MMP-9): The next decade. *Crit. Rev. Biochem. Mol. Biol.* **48**, 222–272 (2013).
55. Radisky, E. S. & Radisky, D. C. Matrix Metalloproteinase-Induced Epithelial-Mesenchymal Transition in Breast Cancer. *J. Mammary Gland Biol. Neoplasia* **15**, 201–212 (2010).
56. Ge, J.-C., Wang, Y.-X., Chen, Z.-B. & Chen, D.-F. Integrin alpha 7 correlates with poor clinical outcomes, and it regulates cell proliferation, apoptosis and stemness via PTK2-PI3K-Akt signaling pathway in hepatocellular carcinoma. *Cell. Signal.* **66**, 109465 (2020).
57. Beerling, E. *et al.* Plasticity between Epithelial and Mesenchymal States Unlinks EMT from Metastasis-Enhancing Stem Cell Capacity. *Cell Rep.* **14**, 2281–2288 (2016).
58. Sikandar, S. S. *et al.* Role of epithelial to mesenchymal transition associated genes in mammary gland regeneration and breast tumorigenesis. *Nat. Commun.* **8**, 1669 (2017).
59. Fischer, K. R. *et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472–476 (2015).
60. Carstens, J. L. *et al.* Stabilized epithelial phenotype of cancer cells in primary tumors leads to increased colonization of liver metastasis in pancreatic cancer. *Cell Rep.* **35**, (2021).
61. Pilla Reddy, V., Bui, K., Scarfe, G., Zhou, D. & Learoyd, M. Physiologically Based Pharmacokinetic Modeling for Olaparib Dosing Recommendations: Bridging Formulations, Drug Interactions, and Patient Populations. *Clin. Pharmacol. Ther.* **105**, 229–241 (2019).
62. Pettitt, S. J. *et al.* Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nat. Commun.* **9**, 1849 (2018).
63. Rondinelli, B. *et al.* EZH2 promotes degradation of stalled replication forks by recruiting MUS81 through histone H3 trimethylation. *Nat. Cell Biol.* **19**, 1371–1378 (2017).
64. Qiu, Z., Oleinick, N. L. & Zhang, J. ATR/CHK1 inhibitors and cancer therapy. *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **126**, 450–464 (2018).
65. Dev, H. *et al.* Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nat. Cell Biol.* **20**, 954–965 (2018).
66. Potter, D. S. *et al.* Dynamic BH3 profiling identifies pro-apoptotic drug combinations for the treatment of malignant pleural mesothelioma. *Nat. Commun.* **14**, 2897 (2023).
67. Merino, D. *et al.* BH3-Mimetic Drugs: Blazing the Trail for New Cancer Medicines. *Cancer Cell* **34**, 879–891 (2018).
68. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

69. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <https://doi.org/10.48550/arXiv.1303.3997> (2013).
70. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
71. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma. Ed. Board Andreas Baxeavanis Al* **11**, 11.10.1–11.10.33 (2013).
72. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
73. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
74. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
75. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
76. Venkatraman, E. S. & Olshen, A. B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657–663 (2007).
77. Olshen, A. B. et al. Parent-specific copy number in paired tumor–normal studies using circular binary segmentation. *Bioinformatics* **27**, 2038–2046 (2011).
78. Gillis, S. & Roth, A. PyClone-VI: scalable inference of clonal population structures using whole genome data. *BMC Bioinformatics* **21**, 571 (2020).
79. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
80. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
81. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
82. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
83. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
84. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
85. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).
86. Jassal, B. et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* **48**, D498–D503 (2020).
87. The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res.* **49**, D325–D334 (2020).
88. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

89. Chou, T.-C. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol. Rev.* **58**, 621–681 (2006).
90. Chou, T.-C. Drug combination studies and their synergy quantification using the Chou-Talalay method. *Cancer Res.* **70**, 440–446 (2010).
91. Bliss, C. I. The Toxicity of Poisons Applied Jointly. *Ann. Appl. Biol.* **26**, 585–615 (1939).

Table

Table 1 is available in the Supplementary Files section.

Figures

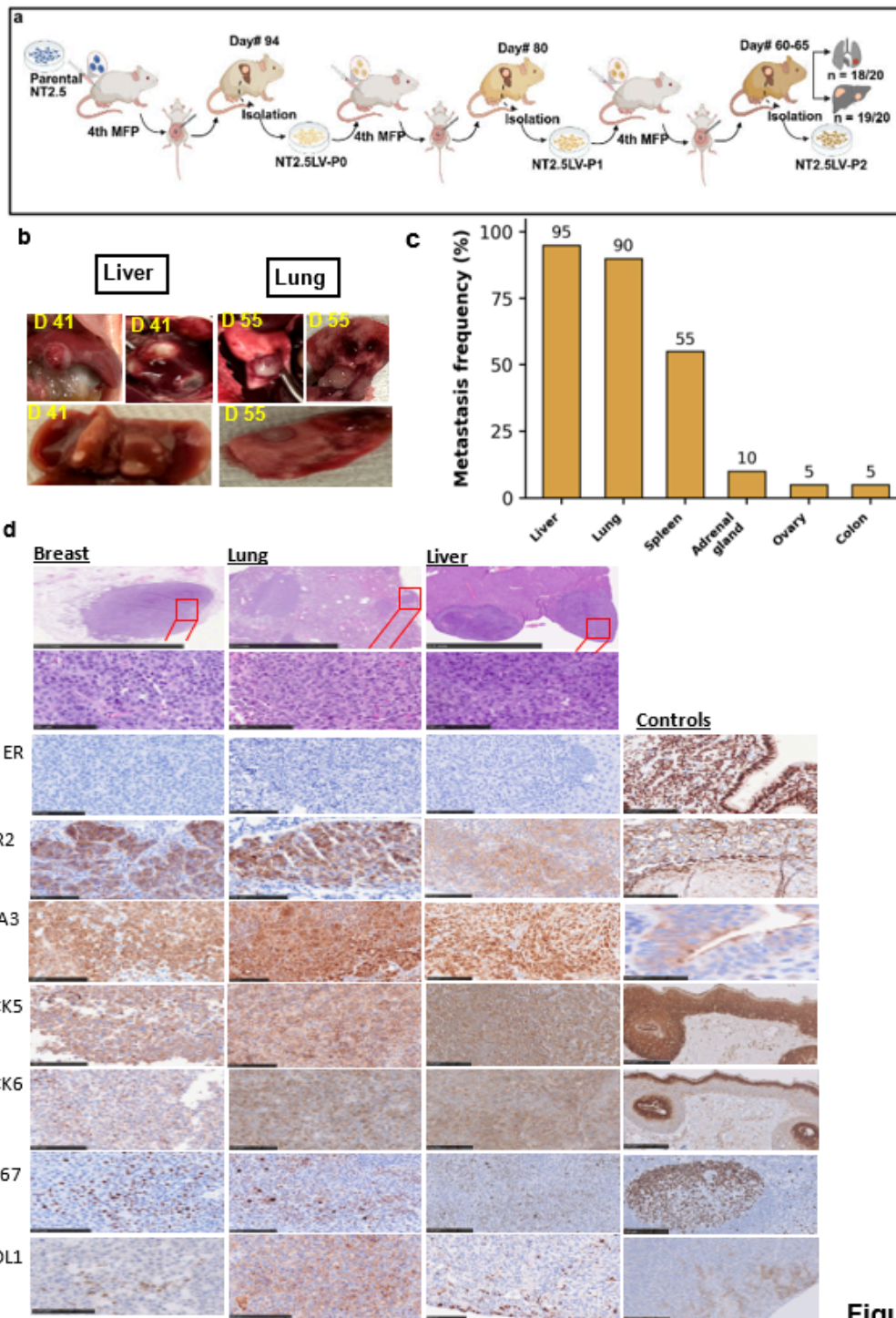


Figure 1

Figure 1

Development and characterization of high-penetrance liver-tropic NT2.5LV.

a Schematic diagram of *in vivo* iterative selection. NT2.5 parental tumor cell line (2.5×10^5) was injected into the fourth mammary fat pad (MFP) of NeuN mice. Primary tumors were resected (Day 44), and liver metastases were harvested on Day 94 (Passage 0, <20% penetrance). After ex vivo culture, cells were re-

injected into the fourth MFP (Passage 1), yielding 45% liver and 75% lung metastases by Day 80. A third round (Passage 2) achieved 95% liver and 90% lung metastases by Day 65. The NT2.5LV cell line was established from Passage 2 and maintained stable metastatic phenotype in subsequent passages.

b Macrometastases depiction in liver and lung. Left panels show gross metastases in the liver on Day 41. Right panels depict gross metastases in the lung on Day 55 post-tumor injection.

c Liver- and lung-tropism. Bar plots showing frequency of spontaneous metastases in NeuN female mice with orthotopic NT2.5LV injection (n=20) at time of euthanasia.

d Immunohistochemical characterization of primary breast tumors, liver and lung metastases. Each panel shows primary breast tumor, lung, liver metastases, and control, respectively. Stained slides show that NT2.5LV lacks estrogen receptor (ER) but exhibited positive staining for ERBB2, CK5, CK6, GATA3, Ki67, and PD-L1 in primary and visceral metastases. Scale bars for H&E in the top row are 2.5mm (1.25x), while in second row 100 μ m (20x). The scale bars for the IHC markers are 100 μ m at resolution of 20x.

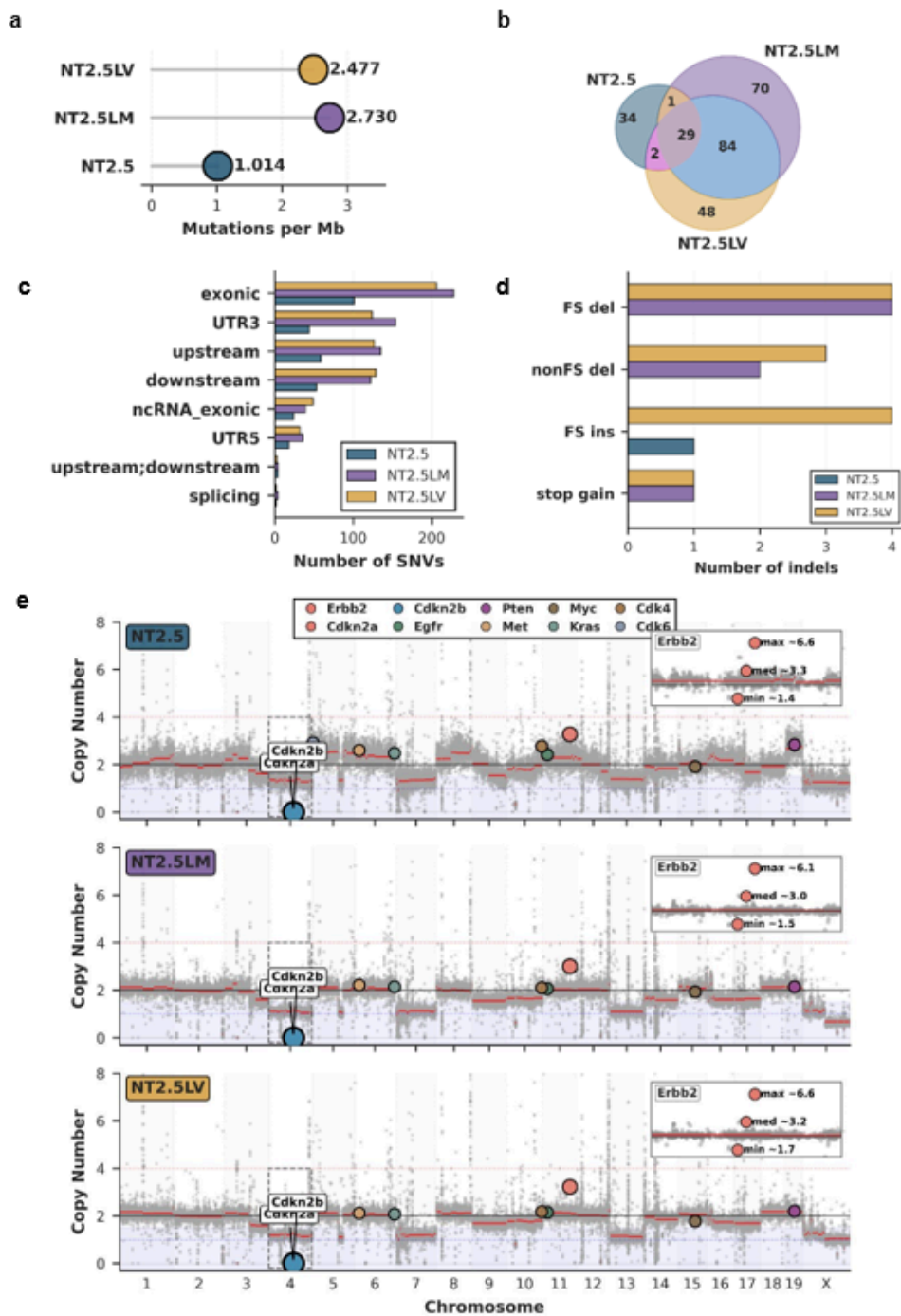


Figure 2

Figure 2

Genomic characterization of NT2.5 parental and metastatic derivatives.

a Genomic alteration in coding regions. Tumor mutational burden (TMB) measured as the total genomic alterations in coding regions per megabase (Mb). NT2.5 harbors 1.014 mut/Mb, NT2.5LM exhibits 2.730 mut/Mb, and NT2.5LV harbors 2.477 mut/Mb.

b Shared and unique coding-region mutations in NT2.5-model system. Venn diagram of protein-altering mutations (SNVs, DNPs, and indels). NT2.5, NT2.5LM and NT2.5LV share 29 mutations, while NT2.5LM and NT2.5LV share 84 alterations. NT2.5LM exhibits 70 unique alterations, NT2.5LV harbors 48, and NT2.5 has 34 such changes. NT2.5 and NT2.5LM share only 1 alteration, and NT2.5 and NT2.5LV share 2.

c Elevated single nucleotide variant distribution. Bar plots showing the distribution of single nucleotide variants (SNVs) types among cell lines.

d Exonic indel distribution. Bar plots presenting insertions and deletions (indels) in NT2.5, NT2.5LM and NT2.5LV. Type of shown deletions include frame shift (FS) deletions (del), non-frameshift (nonFS) deletions. Depicted insertions are frameshift (FS) and stop gain terminating transcription.

e Genome-wide copy-number alterations in NT2.5, NT2.5LM and NT2.5LV. Copy number alteration (CNV) plots for NT2.5 (top), NT2.5LM (middle), and NT2.5LV (bottom). Gray points represent gene-level copy numbers. Red lines identify the means of segmented copy-number. Dashed box highlights *Cdkn2a/Cdkn2b* deletion on chromosome 4. Inset in each panel reveal *ErbB2* amplification on chromosome 11, with minimum, median and maximum copy numbers indicated. Colored dots mark key cancer genes, including *ErbB2*, *Cdkn2a/b*, *Pten*, *Myc*, *Met*, *Egfr*, *Cdk4*, *Cdk6*, *Kras*.

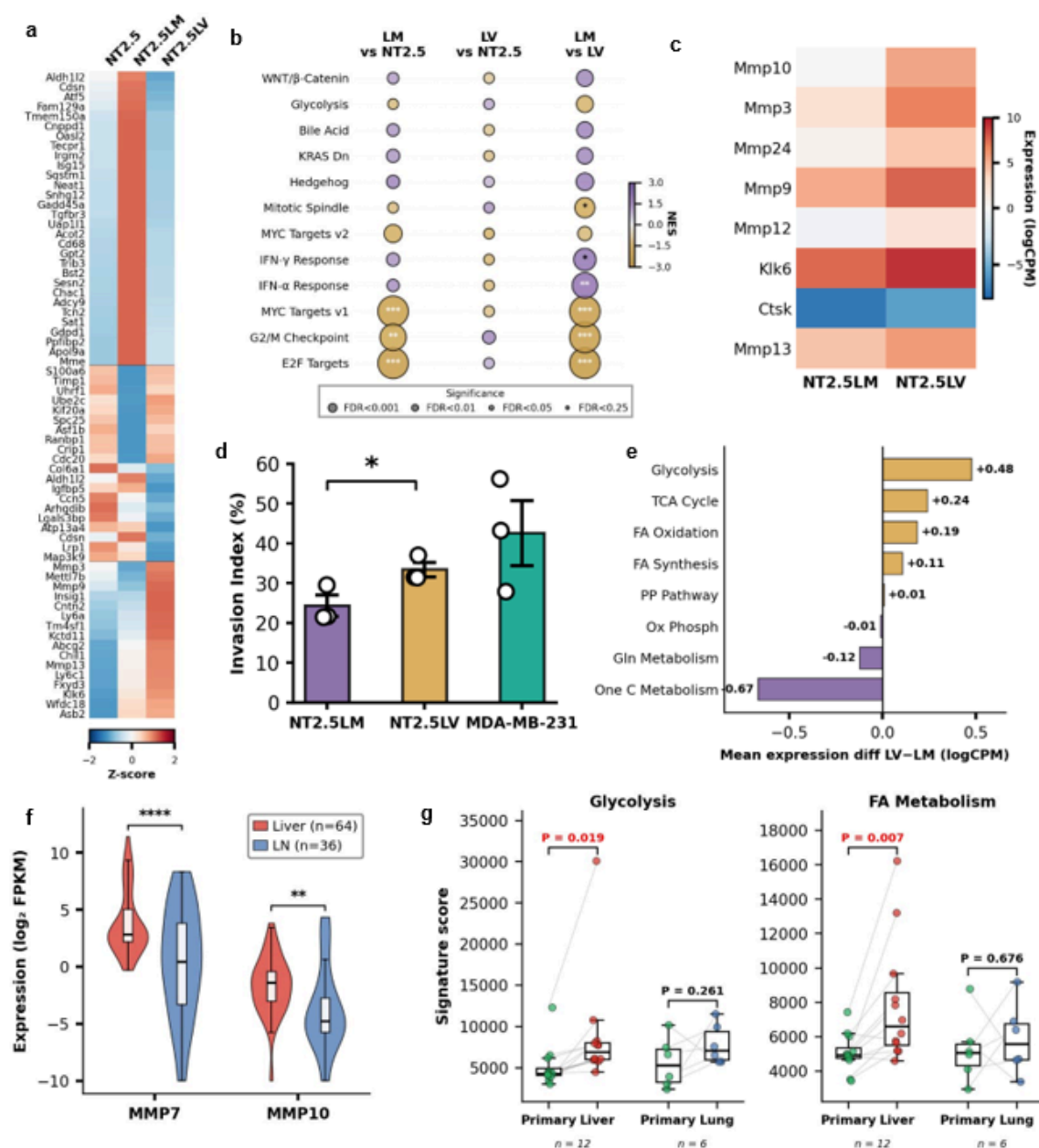


Figure 3

Figure 3

Gene expression distinguishes site-specific metastatic cell lines.

a Cell-line specific gene expression signatures. Heatmap depicting the z-scores normalized gene expression for cell-line specific gene signatures. Genes identified using differential gene expression analysis with significance thresholds: $FDR < 0.05$ and $|\log_2 FC| > 1.585$ identifying 3x FC. Left cluster:

upregulated genes in NT2.5LM (purple) featuring ISG. Right cluster: upregulated genes in NT2.5LV (gold) featuring MMPs and metabolic enzymes. Color scale represents z-score normalized expression (red = high, blue = low). Gene expression data from biological replicates: NT2.5 n = 3, NT2.5LM n = 2, and NT2.5LV n = 3.

b Hallmark pathway enrichment. Bubble plot showing differentially expressed pathways (y-axis) across cell lines (x-axis) by gene set enrichment analysis (GSEA). Bubble size represents enrichment significance ($-\log_{10}\text{FDR}$). Color of each bubble indicates normalized enrichment score (NES denotes color of the cell line whereby gene is upregulated; gold = NT2.5LV, and purple = NT2.5LM). Asterisks (*) denote statistical significance with *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001).

c Gene expression of matrix metalloproteinases (MMPs). Heatmap displaying logCPM expression of MMP genes (y-axis) across cell lines (x-axis). Color scale depicts expression level.

d Baseline invasion capacity of NT2.5 metastatic variants and MDA-MB-231. Invasion indices determined by 96-well basement membrane assay after 48-hour incubation. Invasive capacity of NT2.5LM depicted in purple, NT2.5LV in gold, and MDA-MB-231 in teal were assessed using 8- μm pore Transwell inserts (Abcam). Bars represent mean $\pm$ SEM of 3 independent biological replicates. *p=0.048. Statistical significance was assessed by two-tailed t-test.

e Metabolic pathway expression in NT2.5LV and NT2.5LM. Horizontal bar plots depicting the mean expression difference between NT2.5LV and NT2.5LM (logCPM) of key metabolic pathways (y-axis). Complete list of associated genes found in Supplementary Table 4.

f Validation of MMP upregulation in human breast cancer liver and lymph node (LN) metastases. Violin plots showing *MMP7* and *MMP10* expression ($\log_2\text{FPKM}$) in liver metastases (n = 64, red) and lymph node (LN, n = 36, blue) from the MET500 cohort. Box plots identify the median and quartiles. Two-tailed student's t-test was used to identify statistical significance. Analyzed p-values for *MMP7* = 2.34×10^{-5} , *MMP10* = 0.0432. Asterisks denote statistical significance: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

g Validation of metabolic pathway enrichment in matched primary breast tumor-metastasis pairs from the AURORA breast cancer cohort (GSE209998). Hallmark glycolysis (left) and fatty acid metabolism (right) signature scores were calculated as the mean expression of gene set per sample. Lines connect each patient's primary breast tumor to its matched metastasis. Green, primary tumor; red, liver metastasis; blue, lung metastasis. n, number of matched patient pairs per site. P-values determined by paired two-tailed Student's test.

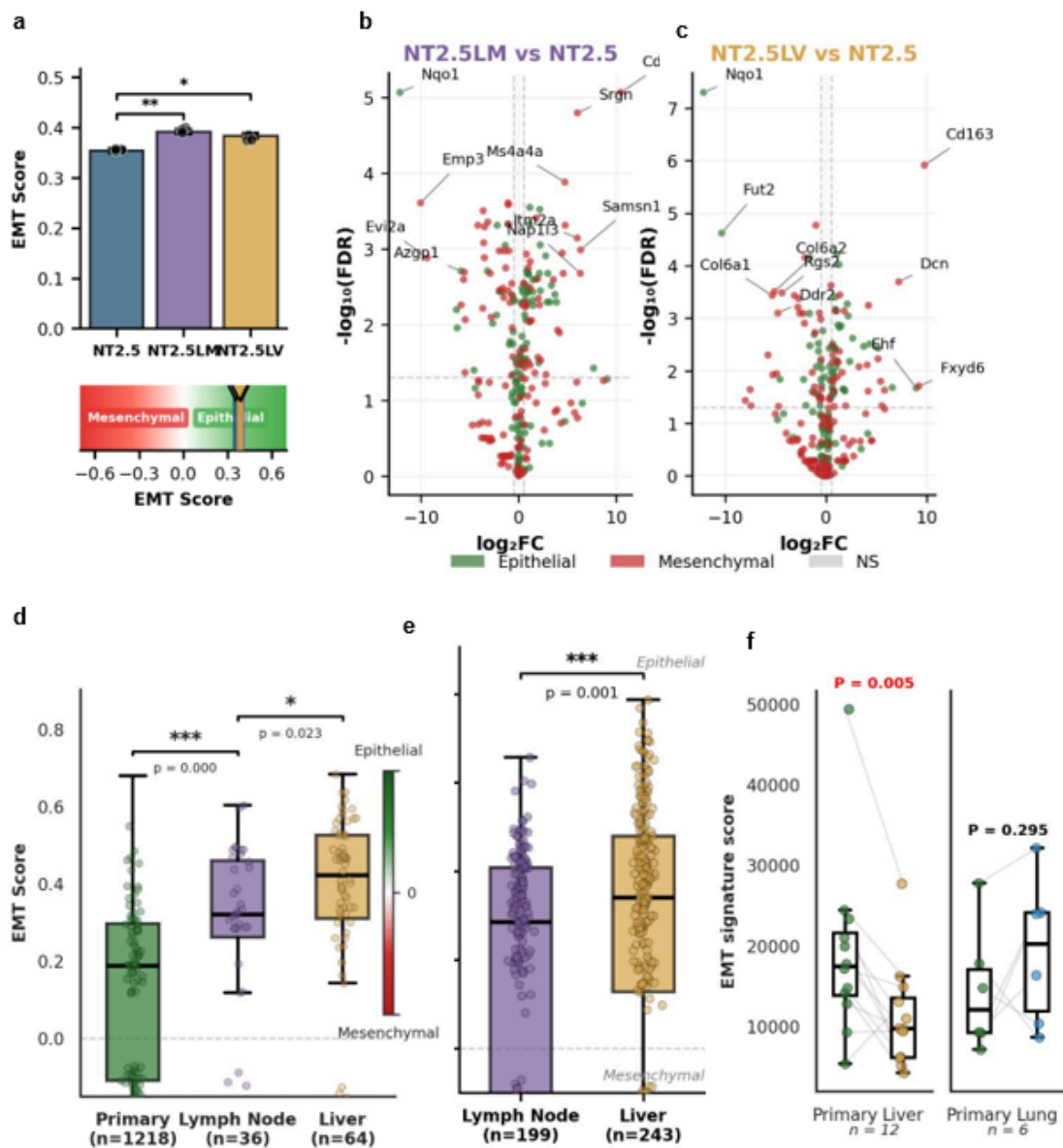


Figure 4

Figure 4

Metastatic breast cancer undergoes mesenchymal-to-epithelial transition

a The EMT scores of NT2.5, NT2.5LM, and NT2.5LV calculated using 2-sample Kolmogorov-Smirnov test with 315 EMT genes. Bar plots represent mean $\pm$ SEM. **p < 0.01, *p < 0.05 (two-tailed t-test). The EMT spectrum gradient show mesenchymal phenotype as red and epithelial phenotype as green.

b and c Differential EMT gene expression. Volcano plots showing differential gene expression of the 315 tumor EMT genes between NT2.5LM vs NT2.5 (b) and NT2.5LV vs NT2.5 (c). Green dots represent epithelial genes, red dots are mesenchymal, and gray dots are not significant. Dashed lines represent $FDR < 0.05$ and $|\log_2 FC| > 0.5$. Top 5 upregulated and downregulated genes are labeled.

d Progressive epithelial selection during the metastatic cascade. TCGA breast primary tumors (green, $n = 1,097$), MET500 breast cancer lymph node metastasis (purple, $n = 36$), and liver metastasis (gold, $n = 64$). $***p = 0.001$, Mann-Whitney U test.

e MET500 pan-cancer validation. Lymph node metastases ($n = 99$) vs liver metastases ($n = 243$) across all cancer types. $***p = 0.001$, Mann-Whitney U test.

f Patient-matched EMT validation (AURORA cohort). EMT signature scores based on Hallmark Gene Set in patient-matched primary breast tumors and metastases. Left panel: liver metastases ($n = 12$ pairs, $**p=0.005$) show decreased EMT score. Right panel: lung metastases ($n=8$ pairs, $p=0.29$) show no significant difference. Box plots show median and quartiles; gray lines connect paired samples. Green dots = primary tumors, gold dots = liver metastases; blue dots = lung metastases. Paired t-test statistical test was performed. TCGA: the Cancer Genome Atlas, MET500: Metastasis 500 consortium, ARURA: Aiming to Understand the Molecular Aberrations in Metastatic Breast Cancer.

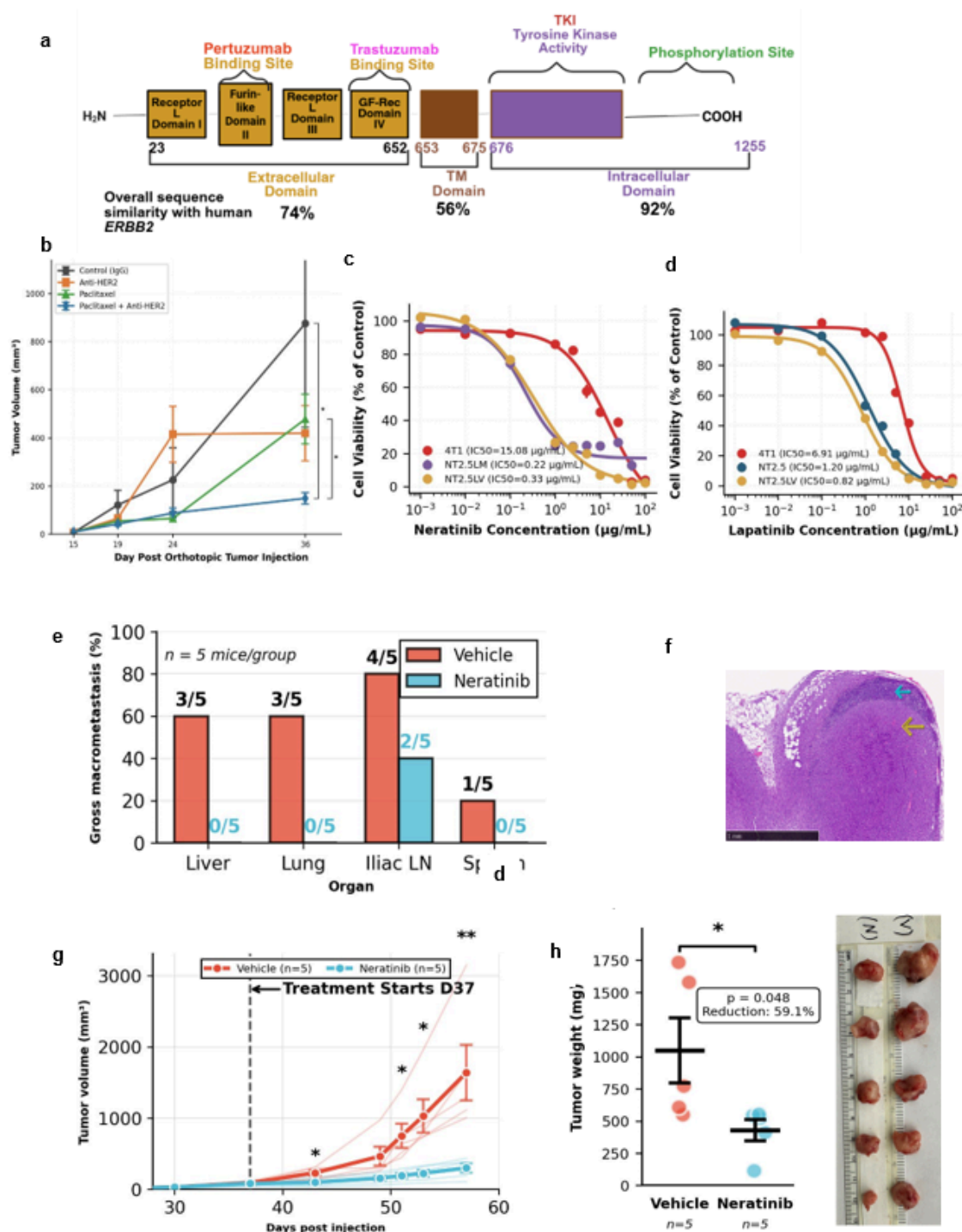


Figure 5

Figure 5

h Iliac lymph node metastasis. Hematoxylin & Eosin image of an iliac lymph node with microscopic metastases isolated from NT2.5LV-tumor-bearing mouse. Tumor cells (yellow arrow) compress the lymphoid cortex (teal arrow) containing smaller lymphocytes).

Conserved HER2 domain supports HER2-monoclonal antibody and TKI-targeted therapies. a

Schematic diagram of *ErbB2* domains. Top layer showing extracellular, transmembrane, and intracellular regions with therapeutic target sites and sequence conservation between human *ERBB2* and NT2.5LV *ErbB2*. **b Efficacy of *in vivo* HER2-targeted therapy with monoclonal anti-HER2 with or without paclitaxel in NT2.5LV.**

Primary breast tumors were measured in the x and y dimensions, and the volume was calculated using $V = 0.5 \times L \times W^2$, where L represents the longest dimension. Data shown represent the mean $\pm$ SEM. Statistical significance was assessed using ANOVA test to compare tumor volume among treatment groups at every time point, while pairwise comparisons with Bonferroni correction was applied for multiple testing (6 time points). Significance brackets indicate comparisons between cell lines: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ns = not significant.

c & d *in vitro* dose-response curves. c shows lapatinib response curve while d shows neratinib response curve. Y-axis depicts cell viability (% of control). X-axis show drug concentration (μ M) in log scale IC_{50} annotated for each cell line.

e *In vivo* neratinib efficacy. Tumor growth curves showing tumor volume (mm^3 , y-axis) measured at various days post tumor injection (x-axis). Dashed line depicts start of treatment (tumor largest diameter 7mm). Asterisks identify statistical significance (* $p < 0.05$, ** $p < 0.01$).

f Tumor weights at experimental endpoint measured in grams. Primary tumor weight assessed in the treatment and control groups. Image to the right shows primary mammary tumors harvested from neratinib- and vehicle-treated groups. Statistical difference identified $p = 0.048$.

g Frequency of distant gross metastases. Bar plots showing the ratio of mice with distant metastases per experimental group.

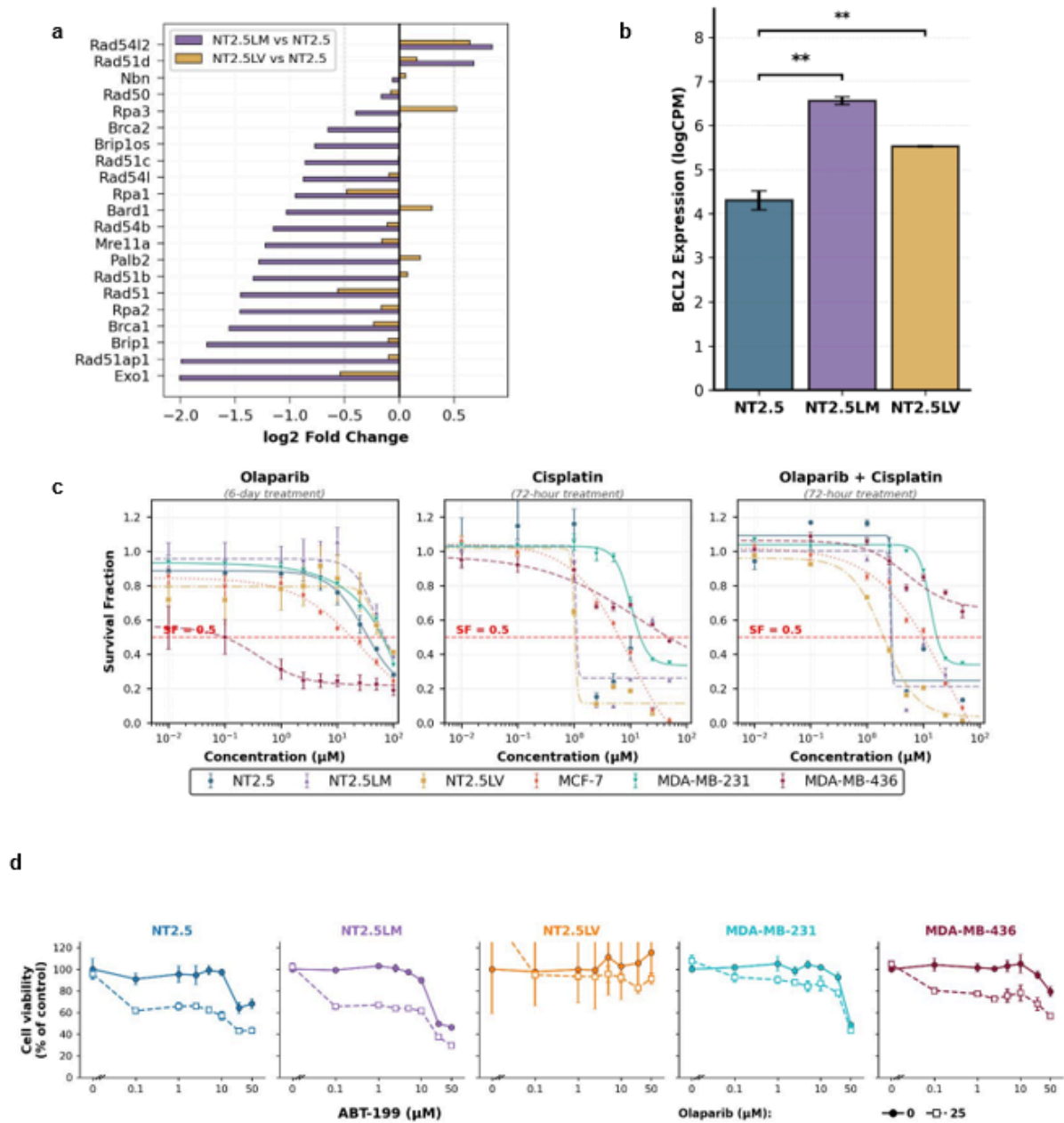


Figure 6

Figure 6

BCL2-Mediated apoptosis evasion drives olaparib resistance in HR-deficient NT2.5LM cells

a Homologous recombination (HR) gene expression. Bar plots show Log₂ fold change (x-axis) in homologous recombination-related genes (y-axis) of metastatic cell lines relative to parental NT2.5. NT2.5LM depicted in purple, and NT2.5LV in gold.

b Dose-response curves for olaparib monotherapy (left; 6-day treatment), cisplatin monotherapy (center; 72-hour), and olaparib and cisplatin (right; 72-hour) across murine (NT2.5, NT2.5LM, NT2.5LV) and human (MCF-7, MDA-MB-231, and MDA-MB-436 harboring BRCA1 mutation) breast cancer cell lines. Cell viability (y-axis) plotted against drug concentration (μM , log scale). Depicted survival curves were fit using 4-parametric logistic curve fitting. Dose concentrations were 0.0-100 μM for olaparib and 0.01-50 μM for cisplatin and combination therapy. Bars colored by cell line; IC_{50} values labeled above bars. Sf=survival fraction. Data represent mean $\pm$ SEM from 2-5 biological replicates.

c BCL2 expression across cell lines measured by RNA-seq (logCPM, mean $\pm$ SEM). NT2.5LM shows approximately 5-fold BCL2 upregulation compared to parental cells (+2.26 logCPM, ***p < 0.001), while NT2.5LV shows 2.3-fold increased (+1.22 logCPM, *p < 0.05). Individual biological replicates shown as overlaid points. Statistical significance assessed by unpaired two-tailed Student's t-test with Bonferroni correction.

d ABT-199 sensitizes NT2.5LM cells to olaparib treatment. Dose-response curves for ABT-199 (0-50 μM) monotherapy (solid circles, solid line) or in combination with 25 μM olaparib (open squares, dashed line) in murine and human breast cancer cell lines. Cells were treated with ABT-199 for 72 hours, with olaparib added 24 hours after ABT-199 treatment. Cell viability was assessed by CellTiter-Glo luminescent assay and normalized to vehicle control. Data are presented as $\pm$ SEM (n=2 biological replicates). NT2.5LM cells, which harbor homologous recombination deficiency yet resistant to olaparib monotherapy, demonstrate marked sensitization to olaparib when combined with Bcl2 inhibition (ABT-199) at concentrations as low as 0.1 μM .

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SynergyanalysisofABT199andolaparibcombinationinvitrotreatmentTable1.docx](#)
- [RevisedSomaticExonicMutationssupptable1.xlsx](#)
- [RNADGELimmaoutputsuptable2.csv](#)
- [MMPintegrinsmetabolicgenessupptable3.xlsx](#)
- [DDRgeneexpressionsupptable4.xlsx](#)
- [EMTGeneExpressionsupptable5.xlsx](#)
- [SupplementaryFigures.docx](#)
- [Supplementarytable1.docx](#)