Phylogenetic and Transcriptomic Analyses Reveal Specific Adaptations for Progression of Breast Cancer

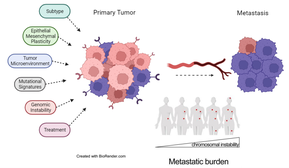
Mariano Russo[1], Praveen-Kumar Raj-Kumar[1,2], Jianfang Liu[1], Anupama Praveen-Kumar[1], Brenda Deyarmin[1], Albert J. Kovatich[2,3], Aik Choon Tan[4], Hai Hu[1,2], Craig D. Shriver[2,5], Xiaoying Lin[1,2]

[1]Chan Soon-Shiong Institute of Molecular Medicine at Windber, Windber, PA; [2]Murtha Cancer Center Research Program, Department of Surgery, Uniformed Services University of the Health Sciences, Bethesda, MD
[3]The Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, USA; [4]Huntsman Cancer Institute, University of Utah Health Sciences Center, Salt Lake City, Utah, USA; [5]Department of Surgery, Walter Reed National Military Medical Center, Bethesda, MD

Authors' contact:
Mariano Russo: m.russo@wriwindber.org; Xiaoying Lin: x.lin@wriwindber.org

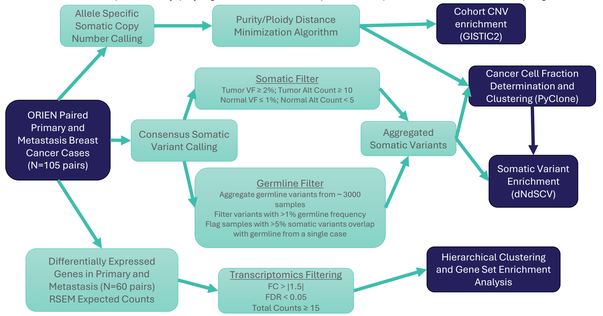
Introduction

- Metastatic progression is thought to result from evolutionary selection of subclones within the primary tumor. Previous studies have demonstrated an overall pattern of increased genomic alterations concurrent with increased clonality in metastasis compared to primary tumors, and determined the potential for distal metastasis relates to capacity of subclonal expansions in primary tumors[1–4].
- Ongoing mutational processes during tumor evolution may favor specific subclones for metastatic seeding. Alternatively, high metastatic clonality could also result from evolutionary constraints due to anticancer therapy.
- Using the Oncology Research Information Exchange Network (ORIEN) datasets, we aim to characterize the genomic progression from primary to metastatic tumors and identify seeding clones[5]. The genomic composition of seeding clones across tumors may give us unique insight into metastatic progression.
- Our overall objective is to map the composition of primary tumor progression into metastasis as it relates to driving factors and its reflection in gene expression (illustrated below).

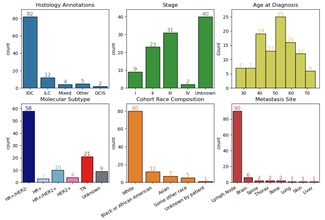


Methods and Cohort Overview

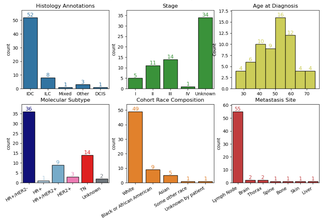
Paired primary and metastasis tumors from 105 individuals in the ORIEN cohort were analyzed with WES and RNASeq to identify phylogenetic and transcriptomic adaptations that drive tumor progression.



Matched Primary and Metastasis Tumors Cohort (N=105)



Differential Expression Analysis Cohort (N=60)



Results I. Phylogenetic and Genomics Analysis

Phylogenetic Characterization and Cluster Classification of Primary and Metastatic Tumors

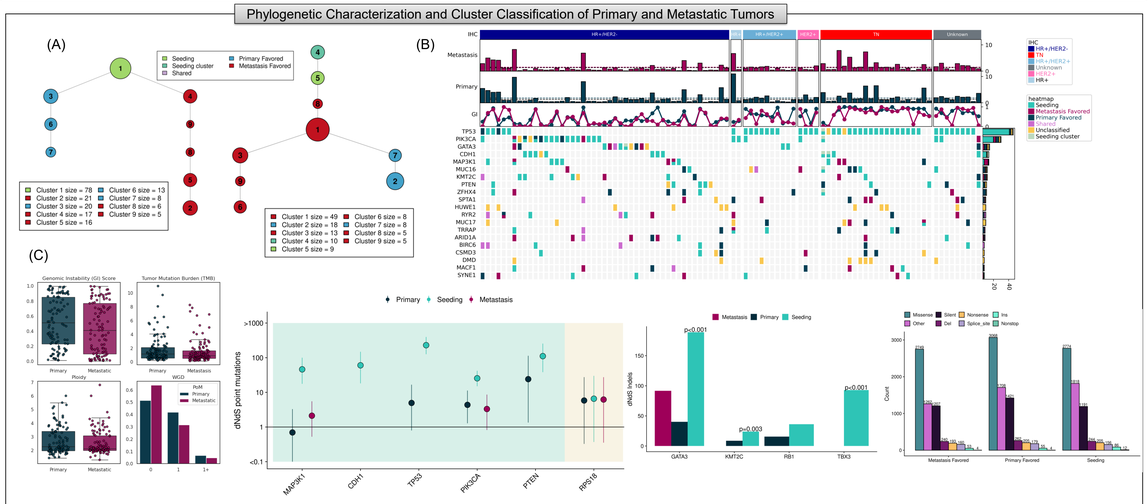


Figure 1. Phylogenetic characterization of primary breast cancer and metastasis. (A) Examples of phylogenetic trees for two different breast cancer patients. Cluster size refers to the number of variants categorized within each cluster. Seeding clones correspond to the highest cancer cell fraction metastasis clone also present in the primary. (B) Somatic mutations in primary breast cancer vs metastasis. The top 20 most frequently mutated driver genes are displayed. The top panels show tumor mutation burden (TMB) or Genomic Instability (GI) per patient sample. The right panel demonstrates the number of mutations identified for each gene and their cluster classification, N = 105. (C) Comparison of primary, regional metastasis (lymph nodes), or distant metastasis for GI, TMB, ploidy, and whole genome duplications (WGD). (D) Positive selection of small variants in driver genes associated with different stages of metastatic progression.

Copy Number Alterations in Primary and Metastasis Tumors

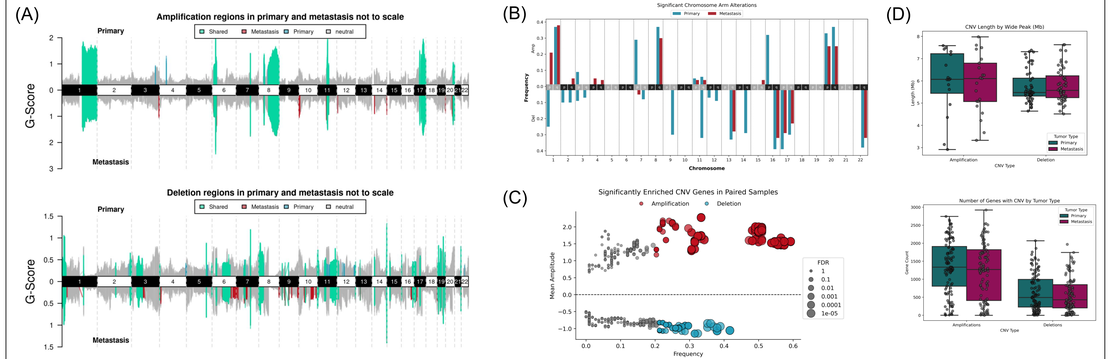


Figure 2. Conserved copy number alterations in primary and metastasis tumors. (A) Genomic representation of significantly altered copy number regions from gistic analysis. Shared regions indicate statistical enrichment in primary and metastasis tumors. (B) Frequency of broad copy number alterations (arm-level) in primary and metastasis. (C) Enrichment of amplified or deleted genes in significantly altered regions based on Poisson means test with FDR adjustment (<0.05). 387 amplified genes and 116 deleted genes were significantly enriched for CNVs in the cohort. (D) Comparison of CNV region length or overall number of genes identified in primary and metastasis tumors.

Results II. Transcriptomic Analysis

Hierarchical Clustering of Differentially Expressed Genes

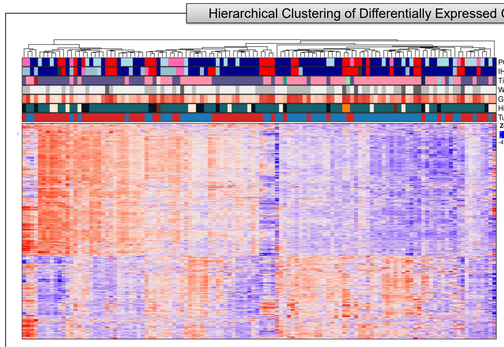


Figure 3. Hierarchical clustering of differentially expressed genes. A total of 1,039 genes (648 upregulated, 391 downregulated) were differentially expressed between primary and metastatic tumors (N = 60). Clustering is a result of mixed contributions from different clinical and molecular factors (not driven by single entity).

Gene Set Enrichment Analysis

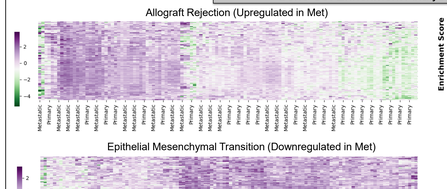


Figure 4. Gene Set Enrichment Analysis (GSEA) in primary vs metastasis tumors. GSEA of normalized gene expression across N=60 primary and metastasis tumors identifies 7 differentially regulated pathways with FDR < 0.2.

Discussion

- We have identified seeding clones in 105 cases of breast cancer from ORIEN cohort with driver mutations which vary across IHC subtypes of breast cancer. dNdS selection modeling identified enrichment of selection for mutations in specific genes nearly exclusively in seeding clusters.
- Copy number analysis demonstrated several regions commonly altered in primary and metastasis suggesting these as fitness contributing early events in tumor progression. CN altered genes consisted mainly of regulatory RNAs and DNA binding genes (such as histones and zinc fingers).
- Differential expression analysis identified 1039 genes altered between primary and metastasis tumors. GSEA characterized functional adaptations in transcriptional programs involving immune signaling and angiogenesis in the progression to metastasis.
- Additional cases and analytical approaches are ongoing to further characterize the contribution of evolutionary selection and gene regulation in the acquisition of metastatic potential for breast cancer.

References

1. Nguyen, B. *et al.* Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. Cell 185, 563-575.e11 (2022).
2. Martínez-Jiménez, F. *et al.* Pan-cancer whole-genome comparison of primary and metastatic solid tumours. Nature 618, 333–341 (2023).
3. Angus, L. *et al.* The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. Nature Genetics 51, 1450–1458 (2019).
4. Biohmer, M., Cheek, D.M., Hung, WT. *et al.* Quantifying cell divisions along evolutionary lineages in cancer. *Nat Genet* **57**, 706–717 (2025).
5. Grigoriadis, K. et al. CONIPHER: a computational framework for scalable phylogenetic reconstruction with error correction. Nature Protocols 19, 159–183 (2024).
6. Dmg, W., Goldberg, D. and Zhou, W. (2023), PyComplexHeatmap: A Python package to visualize multimodal genomics data. iMeta e115.
7. Csardi G, Nepusz T (2006). "The igraph software package for complex network research." *InterJournal*, Complex Systems, 1695.
8. Van Loo, P. *et al.* Allele-specific copy number analysis of tumors. *Proceedings of the National Academy of Sciences* **107**, 16910–16915 (2010).

Disclaimer

The contents of this publication are the sole responsibility of the author(s) and do not necessarily reflect the views, opinions, or policies of Uniformed Services University of the Health Sciences (USUHS), the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., the Department of War (DoW) or the Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. government.

Chan Soon-Shiong Institute of Molecular Medicine at Windber