Integrating Environmental Risk Factors and mRNA Expression Profiles as Prognostic Biomarkers in Triple-Negative Breast Cancer

Amjila Bam[1], Yawen Hu[2], Xiao-Cheng Wu[1], Nubaira Rizvi[1], Luis Del Valle[2], Fokhrul Hossain[2], Jovanny Zabaleta[2], Augusto Ochoa[2], Lucio Miele[2], Edward Trapido[1], Qingzhao Yu[1]

LSU Health NEW ORLEANS School of Public Health

publichealth.lsuhsc.edu

Background

- TNBC is an aggressive breast cancer subtype that comprises approximately 10-15% of all breast cancer cases and is associated with poor prognosis and substantial heterogeneity in clinical outcomes.
- While DNA-level alterations define tumor potential, mRNA expression reflects the tumor's active biological state and is highly responsive to cellular signaling and external stressors.
- Growing evidence links environmental exposures, including air pollution, water contamination, transportation-related emissions, and social vulnerability, to TNBC risk, delayed diagnosis, and aggressive disease features, particularly in structurally disadvantaged communities.
- However, most existing studies evaluate exposures in isolation and do not account for cumulative environmental burden.
- To address this gap, we integrated mRNA expression profiles from 253 TNBC tumors with multi-exposure Environmental Justice Index (EJI) metrics.
- Using a mutual information (MI) based pathway framework, we identified environmentally sensitive transcriptomic programs associated with TNBC stage and tumor grade, revealing proliferative, inflammatory, metabolic, and non-canonical hormonal signaling pathways that may mediate the biological impact of environmental injustice in TNBC.

Methodology and Data Analysis

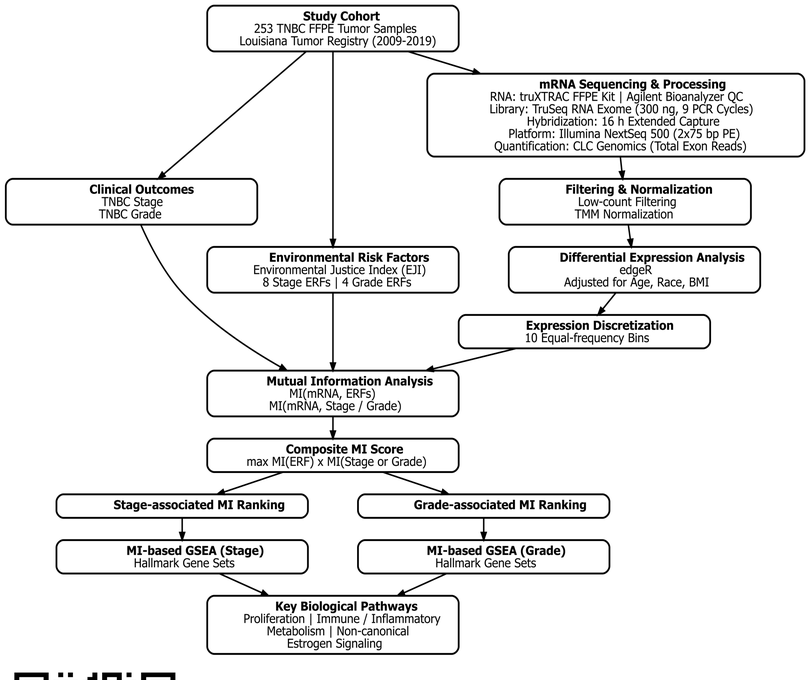

Figure 1. Flowchart Illustrating the MI Based Workflow for Integrating Environmental Risk Factors and TNBC Transcriptomic Profiles

Results

- Eight community-level environmental risk factors spanning transportation proximity, air quality burden, racial/ethnic minority concentration, and social vulnerability domains were significantly associated with TNBC stage and therefore prioritized for MI based analyses to identify transcriptomic signatures mediating these environmental effects.
- Four community-level environmental risk factors related to impaired water quality and racial/ethnic minority concentration were significantly associated with TNBC grade and prioritized for MI based analyses to identify transcriptomic signatures mediating these environmental effects.
- Across TNBC diagnostic stages, 6,195 mRNAs were differentially expressed, with the greatest transcriptomic impact observed for transportation proximity (681-621 mRNAs), air quality burden (472-444 mRNAs), and social vulnerability domains (435-237 mRNAs), whereas racial/ethnic minority associated factors exhibited a comparatively focused response (64 mRNAs).
- TNBC grade was associated with 4,642 differentially expressed mRNAs, driven primarily by environmental stressors related to impaired water quality (571 mRNAs), whereas racial/ethnic minority associated factors exhibited a comparatively restricted transcriptional impact (13 mRNAs).

MI-GSEA Enriched Hallmark Pathways

Table 1. A comprehensive overview of the biologically relevant and significantly enriched pathways associated with tumor grade and related ERFs selected from the full set of 50 Hallmark pathways.

Hallmark Pathways	NES	Size	p-Adjust (FDR)	Common Gene Count	Common DE Genes
E2F_TARGETS	1.55	200	0	3	PRIM2, LBR, PHF5A
G2M_CHECKPOINT	1.47	198	0	2	PRIM2, LBR
INTERFERON_GAMMA_RESPONSE	1.31	192	0	2	IDO1, GPR18
INTERFERON_ALPHA_RESPONSE	1.46	96	0	1	NUB1
MITOTIC_SPINDLE	1.22	199	0	1	ARHGEF11
MYC_TARGETS_V1	1.24	199	0	1	EPRS1
APICAL_JUNCTION	1.17	198	0.009375	2	ALOX15B, EGFR
MTORC1_SIGNALING	1.15	199	0.019545	2	PSPH, EPRS1
TNFA_SIGNALING_VIA_NFKB	1.14	199	0.02875	2	NFKBIE, CCL20
ESTROGEN_RESPONSE_LATE	1.13	198	0.031923	3	PDZK1, TFF3, TFF1
KRAS_SIGNALING_UP	1.13	200	0.038214	2	CCL20, ATG10
ESTROGEN_RESPONSE_EARLY	1.12	198	0.046875	4	LRIG1, PDZK1, TFF3, TFF1

Table 2. A comprehensive overview of the biologically relevant and significantly enriched pathways associated with tumor stage and related ERFs selected from the full set of 50 Hallmark pathways.

Hallmark Pathways	NES	Size	p-Adjust (FDR)	Common Gene Count	Common DE Genes
MYC_TARGETS_V1	1.66	199	0	16	XPO1, HDAC2, MRPL9, PABPC4, CLNS1A, YWHAE, SYNCRIP, HDDC2, CCT7, TCP1, PGK1, HNRNPA1, MRPL23, C1QBP, VBP1, NAP1L1
MTORC1_SIGNALING	1.34	199	0	12	CCT6A, TMEM97, PSPH, HSP90B1, CACYBP, PGK1, PSMD12, NFYC, CFP, NUFIP1, SHMT2, EBP
E2F_TARGETS	1.68	200	0	10	XPO1, SMC3, SYNCRIP, CCNE1, RFC1, SMC4, POP7, NAP1L1, LUC7L3, ZW10
MITOTIC_SPINDLE	1.33	199	0	8	SMC3, PPP4R2, YWHAE, EPB41, UXT, RANBP9, RFC1, SMC4
G2M_CHECKPOINT	1.59	198	0	5	XPO1, SYNCRIP, HSPA8, SMC4, FANCC
OXIDATIVE_PHOSPHORYLATION	1.22	200	0	5	UQCRH, TIMM8B, NDUFS2, MRPS15, COX7A2
DNA_REPAIR	1.32	150	0	3	POLR2J, TAF10, POLR1D
UV_RESPONSE_UP	1.31	156	0	3	MRPL23, CCNE1, PSMC3
PI3K_AKT_MTOR_SIGNALING	1.26	105	0.004167	6	UBE2N, HSP90B1, MAPK8, ARF1, CLTC, EGFR
ESTROGEN_RESPONSE_LATE	1.18	198	0.005385	5	GPER1, SLC16A1, ANXA9, ARL3, XRCC3
PEROXISOME	1.23	104	0.011071	2	GNPAT, CAT
ESTROGEN_RESPONSE_EARLY	1.16	198	0.016667	6	SLC16A1, ANXA9, SH3BP5, ARL3, DEPTOR, RAB17
UNFOLDED_PROTEIN_RESPONSE	1.2	112	0.01875	1	HSP90B1
FATTY_ACID_METABOLISM	1.17	157	0.018824	2	LTC4S, NSDHL
REACTIVE_OXYGEN_SPECIES_PATHWAY	1.28	49	0.020556	3	MPO, NDUFS2, CAT
HEME_METABOLISM	1.15	197	0.023158	13	SNCA, EPB41, SLC22A4, EPB42, CAT, HBB, LMO2, MBOAT2, CA1, SLC4A1, ASNS, GLRX5, CTNS
APOPTOSIS	1.15	160	0.030238	3	EBP, JUN, NEFH
XENOBIOTIC_METABOLISM	1.14	199	0.030682	8	TMEM97, LEAP2, CAT, SHMT2, GNMT, SLC35B1, ACSM1, ARG2
IL2_STAT5_SIGNALING	1.13	198	0.03087	2	CCNE1, ARL4A
P53_PATHWAY	1.14	198	0.032083	5	TGFA, RPS12, HEXIM1, JUN, GLS2

Discussion and Conclusion

- Analysis of 253 TNBC tumors showed that integrating community-level environmental injustice metrics with mRNA sequencing reveals biologically meaningful pathways linked to disease stage and tumor grade.
- TNBC stage was associated with multiple environmental stressors (transportation proximity, air quality, social vulnerability), whereas tumor grade was linked to more specific exposures, particularly impaired water bodies and racial/ethnic minority concentration, suggesting different mechanisms of delay vs. aggressiveness.
- Across stage and grade, cell-cycle and proliferative pathways (E2F, MYC, G2M checkpoint) were consistently enriched, indicating that chronic environmental stress may amplify intrinsic TNBC growth programs.
- High-grade TNBC showed strong enrichment of interferon and TNF-NFκB signaling, consistent with a chronic inflammatory tumor microenvironment potentially driven by sustained environmental exposures.
- Despite triple-negative status, estrogen response pathways (via *GPER1* and related genes) were reproducibly enriched, suggesting ERα-independent, environmentally influenced hormonal signaling in TNBC.
- Stage-associated tumors exhibited enrichment of metabolic and stress-adaptation pathways (PI3K/AKT/mTOR, oxidative phosphorylation, UPR), consistent with cellular adaptation to hostile environmental conditions.
- Findings support a model in which environmental injustice contributes to TNBC disparities through both delayed diagnosis and biologically aggressive tumor programs, highlighting the importance of structural and environmental interventions alongside molecular characterization.

Limitations and Future Work

- Environmental Justice Index (EJI) data were derived from 2022 datasets, whereas TNBC diagnoses occurred between 2009-2019, which may introduce exposure misclassification due to changes in environmental conditions over time.
- The use of the maximum mutual information (MI) value to define the Composite MI Score emphasizes dominant environmental drivers but may underrepresent cumulative or synergistic effects of multiple co-occurring exposures.
- Analyses focused on tumor stage and grade; other clinically relevant endpoints (e.g., survival, treatment response) were not evaluated.

Acknowledgment: This research was funded by National Cancer Institute grant numbers 1R01CA275089, NIMHD 2R15MD012387-02, and NIEHS 5P20CA233374.