Oncology 2025
The Cancer Control Challenge:
Better Understanding It to Better Face It

TRANSCRIPTIONAL SIGNATURES OF FIELD CANCERIZATION IN GASTRIC CANCER

Jéssica Manoelli Costa da Silva[1]; Williams Barra[4,5]; Ronald Silva Mourão[1]; Daniel Avelar Costa[3]; Diego Pereira[1]; Valeria Cristiane Santos da Silva[2]; Rubem Silva[2]; Samia Demachki[4]; Fabiano Cordeiro Moreira[5]; Paulo Pimentel de Assumpção[5].

[1] Programa de Pós Graduação em Genética e Biologia Molecular, Universidade Federal do Pará, Pará, Brazil.
[2] Programa de Pós Graduação em Oncologia e Ciências Médicas, Universidade Federal do Pará, Pará, Brazil.
[3] Programa de Pós Graduação em Biotecnologia, Universidade Federal do Pará, Pará, Brazil.
[4] Complexo Hospitalar da UFPA - Hospital Universitário João de Barros Barreto, Pará, Brazil.
[5] Núcleo de Pesquisa em Oncologia, Universidade Federal do Pará, Pará, Brazil.

Introduction: Gastric cancer (GC) remains one of the most prevalent and lethal malignancies worldwide, particularly in regions like Pará, Brazil. While histopathological classifications, such as those by Lauren and the WHO, provide valuable clinical guidance, they do not fully capture the molecular complexity of GC. An often-overlooked aspect is the use of peritumoral tissue as a reference for normality, despite mounting evidence that these areas may already harbor early molecular changes. The field cancerization theory posits that histologically normal tissues adjacent to tumors can exhibit early transcriptional alterations, playing a pivotal role in early carcinogenesis. **Objectives:** To identify transcriptional alterations shared between tumor and peritumoral tissues but absent in healthy gastric mucosa, characterize their biological relevance, and assess their prognostic value in gastric cancer. **Methods:** RNA-seq data from 72 paired tumor and peritumoral samples (CAAE:47580121.9.0000.5634) and 68 normal gastric tissues (PRJNA1054173) were analyzed. A deep autoencoder with dense ReLU-activated layers was trained over 800 epochs. SHAP (SHapley Additive exPlanations) was applied to identify genes with the highest relevance to each tissue type. Coexpression patterns were assessed in tumor and peritumoral contexts. Functional network analysis was performed using the STRING database. Prognostic value was evaluated using two approaches: (i) Kaplan–Meier analysis from public GC cohorts (KM Plotter), and (ii) internal survival analysis of 46 patients from the Pará cohort using the survminer R package. **Results:** SHAP-based selection yielded 10,609 relevant genes for tumor and 9,483 for peritumoral tissues, with 8,871 shared genes. Among these, 138 genes showed stable expression between tumor and peritumoral tissues (Wilcoxon $p > 0.9$), but were differentially expressed compared to normal samples ($|\text{log2FC}| > 1$; adj. $p < 0.05$), suggesting early molecular establishment in the transformation field. From these, three COSMIC-listed cancer driver genes—*FAT4, MDM4*, and *NDRG1*—were highlighted. *FAT4* (LFC = 2.00), linked to Hippo signaling, was overexpressed in tumors; *MDM4* (LFC = 1.71), a p53 inhibitor, may promote genomic instability; and *NDRG1* (LFC = −2.00), a known EMT suppressor, was repressed. STRING network analysis revealed mitochondrial and metabolic enrichment among

the 138 genes, with the three drivers displaying low connectivity, supporting their role as phenotypic transition regulators. Coexpression analysis showed a negative correlation between *FAT4* and *NDRG1* in tumors (r = −0.36), absent in peritumoral tissue (r = +0.07), and a positive correlation between *FAT4* and *MDM4* only in peritumoral samples (r = +0.35). Public survival data indicated that high expression of *FAT4* and *MDM4* was significantly associated with poor prognosis (HR = 1.43; p < 0.001), whereas *NDRG1* was not. Internally, in the study cohort (n = 46), high expression of *MDM4* was also associated with worse overall survival (p = 0.013), while *FAT4* and *NDRG1* showed no significant association. **Conclusion:** These results reinforce that peritumoral tissue represents a molecularly active transitional field in gastric carcinogenesis. *FAT4* and *MDM4* emerge as potential markers of early transformation and poor prognosis.

Keywords: Field Cancerization; Gastric Adenocarcinoma; Deep learning.