GMR
The Original

ARTIFICIAL INTELLIGENCE-ASSISTED PROFILING OF CIRCULATING TUMOR CELL GENETIC HETEROGENEITY OF EPITHELIAL-MESENCHYMAL TRANSITION DYNAMICS UNDER THE INFLUENCE OF 5-AZACYTIDINE AND CYCLOPHOSPHAMIDE REGIMENS IN BREAST CANCER PATIENTS

Ajit Kumar Saxena[1], Shalini[2], Prachi and Manoj Kumar[3]

[1]Depaertment of Patho;ogy /Laboratory Medicine Patna Bihar india
[2]Department of Zoology, B R D Ambedekar University, Muzzaffrpur, Bihar India
[3]Department of General Surgery, All India Institute of Medical Sciences, Patna, Bihar India
***Corresponding Author:** Dr. Ajit Kumar Saxena, Email: draksaxena1@rediffmail.com

ABSTRACT
Circulating tumor cells (CTCs) represent a critical non-invasive liquid biopsy target for early breast cancer (BC) diagnosis and prognostic evaluation. Epithelial-mesenchymal transition (EMT) regulators—specifically Sox4, EpCAM, CK19, and Vimentin—govern key processes during malignant transformation, cellular invasion, and distant dissemination. In this investigation, the differential effects of two distinct antineoplastic therapeutics—the DNA methyltransferase inhibitor 5-Azacytidine (5-AZA) and the alkylating agent Cyclophosphamide (CP)—were analyzed regarding the genetic and phenotypic signatures of CTCs retrieved from BC patient cohorts. Employing flow cytometric immunophenotyping with PE-conjugated antibodies targeting CK19 and Vimentin, alongside PCR-based DNA Copy Number Variation (DNACNV) profiling for Sox4, Oct4, Nanog, CK19, and Vimentin, comprehensive datasets were gathered across 149 to 187 clinically confirmed BC patients and matched control groups. An artificial intelligence (AI) framework was deployed to integrate these multipara metric experimental metrics, modeling CTC heterogeneity, pharmacologically induced genetic susceptibility, and age-stratified cellular responses. Experimental results showed significant down-regulation of CK19 ($p < 0.001$) and Vimentin ($p < 0.001$) following CP treatment, reaching maximal cytostatic impact at 72 hours. Conversely, 5-AZA induced Sox4 over-expression ($p < 0.01$), which acted in synergy with Nanog signaling to facilitate metastatic progression. EpCAM maintained high diagnostic sensitivity (98.98%) across both pharmacological protocols. Machine learning models successfully validated these biological trends, stratifying metastatic risk, pinpointing optimal therapeutic windows, and ranking prognostic markers, while highlighting that the youngest age demographic (25–35 years) exhibited the highest CTC frequency (42.86%).

KEYWORDS: Circulating Tumor Cells; Artificial Intelligence; EMT Markers; 5-Azacytidine; Organophosphate; DNA Copy Number Variations; Breast Cancer; Flowcytometery

1. INTRODUCTION

Breast cancer remains a major global public health challenge characterized by extensive molecular and histological heterogeneity. World Health Organization statistics highlight an expanding epidemiological burden, particularly among postmenopausal females. The clinical management of breast cancer is complicated by dynamic alterations in epithelial-mesenchymal transition (EMT) markers expressed on circulating tumor cells (CTCs). These circulating cells function as critical cellular intermediates connecting primary mammary lesions to distant organ sites. CTCs enter peripheral blood during both early and advanced stages of malignant progression, offering an essential substrate for minimally invasive liquid biopsy. The EMT process is driven by key transcription factors such as Sox4, cell adhesion molecules like EpCAM, and intermediate filament proteins including CK19 and Vimentin. Together, these proteins enable epithelial cancer cells to acquire a motile, mesenchymal phenotype that enhances invasive capacity and promotes chemoresistance.

Sox4, encoded on chromosome 6p22.3, is a 47 kDa Sry-related High Mobility Group (HMG) box transcription factor operating upstream of kRAS-mediated oncogenic signaling and TGF-β developmental pathways. The functional interplay between Sox4 and pluripotency factors such as Nanog (and to a lesser degree Oct4) plays a vital role in maintaining cancer stem cell self-renewal and initiating aggressive metastatic cascades. Disruption of this coordinated signaling network—such as through Oct4 dysregulation—can alter metastatic capacity.

To understand how antineoplastic therapies impact CTC diversity, two mechanistic classes of chemotherapy were evaluated. Cyclophosphamide is a classical alkylating prodrug that undergoes hepatic transformation via cytochrome P450 enzymes into phosphoramide mustard and acrolein. These reactive metabolites generate interstrand DNA crosslinks and structural chromosomal aberrations, inducing time- and dose-dependent

cytostasis. In contrast, 5-Azacytidine acts as a cytosine nucleoside analogue and DNA methyltransferase inhibitor that reverses promoter hypermethylation, leading to the re-expression of epigenetically silenced EMT drivers and altering marker profiles on CTCs.. This synthesized study integrates empirical data across three patient cohorts to establish a machine learning framework capable of modeling CTC diversity under dual therapeutic exposure. By fusing flow cytometric enumeration, PCR-DNACNV band density metrics, and clinical demographics, the machine learning architecture classifies metastatic risk, identifies optimal therapeutic timeframes, and prioritizes prognostic markers for early breast cancer management.

2. MATERIALS AND METHODS:

2.1 Patient Cohorts and Specimen Collection

Venous blood samples (1.0–3.0 mL) were collected in sterile EDTA tubes from three separate patient cohorts diagnosed with breast cancer at the Department of Surgery and Radiation Oncology, AIIMS Patna, India:

- **Cohort 1:** 149 breast cancer patients and 100 control subjects evaluated for CP-mediated CK19 and Vimentin modifications.
- **Cohort 2:** 187 breast cancer patients and matched controls evaluated for 5-AZA-induced Sox4 and EpCAM dynamics.
- **Cohort 3:** 38 breast cancer cases and 38 healthy controls evaluated for DNACNV profiles of stemness markers (Sox4, Oct4, Nanog).

Ethical approval was granted by the Institute Research Committee and Ethics Committee of AIIMS Patna (IEC/2020/610), and written informed consent was secured from all participants or legal guardians.

2.2 CTCs Isolation and Short-Term Culture

Short-term lymphocyte cultures (n = 150) were established using RPMI-1640 medium supplemented with 10% fetal bovine serum, phytohemagglutinin (PHA), and antibiotics. Cultured cell populations were exposed to CP (1.0 µg/mL) or biologically equivalent doses of 5-AZA for 24, 48, or 72 hours. CTCs were isolated from peripheral blood mononuclear cells (PBMCs) via Ficoll density-gradient centrifugation (400 × g for 30 min at 4°C). Harvested cell pellets were washed twice in phosphate-buffered saline (PBS) and preserved at −20°C. Viability and cellular morphology were confirmed using phase-contrast light microscopy following 1% Giemsa staining prior to immunophenotyping.

2.3 FlowCytometric Immunophenotyping

Flow cytometric profiling was performed on a Beckman Coulter DxFLEX platform equipped with a 488 nm argon laser. Phycoerythrin (PE)-conjugated monoclonal antibodies were utilized, including anti-CK19 (clone A-3, Catalog No. sc-376126, Santa Cruz Biotechnology, USA; 200 µg/mL) and anti-Vimentin (clone VI/RE/1, REF MA1-19656, Thermo Fisher, USA; 100 µg/mL), along with validated anti-Sox4 and anti-EpCAM antibodies. Fixation and permeabilization were carried out using the PerFix-nc kit (Beckman Coulter, France). For each marker assay, a minimum of 50 µL CTC suspension was incubated for 60 minutes, washed, and resuspended in 500 µL PerFix-nc Buffer-3. Gating strategies effectively excluded contaminating lymphocytes and erythrocytes using CD326/SSC and CK19/SSC dot plots, and data were processed using CytExpert software (v2.2).

2.4 PCR-Based DNA Copy Number Variation (DNACNV) Profiling

Genomic DNA extracted from PBMCs was quantified using spectrophotometry. Polymerase chain reaction (PCR) mixtures (25 µL total volume) contained 5× Green GoTaq reaction buffer, 10 mM dNTP mix, 10 pmol of gene-specific primers, and 0.2 µL GoTaq DNA Polymerase (5 U/µL; Promega). Primer sequences were confirmed via NCBI-BLAST searches. Amplification comprised 35 cycles with initial denaturation at 95°C and marker-specific annealing temperatures. Amplicons were separated on 1.5% agarose gels, stained with ethidium bromide, and documented using a GelDoc gel documentation system (Bio-Rad, USA). Image Lab 5.1 software was used to determine Densitometric band intensities as indirect metrics for relative gene copy number variations.

Target primer sequences and expected amplicon lengths were as follows:

- **CK19 (573 bp):** Forward 5'-ATTCCGCTCCGGGCACCGATCT-3', Reverse 5'-CGCTGATCAGCGCCTGGATATGCG-3'.
- **Vimentin (323 bp primary; 225 bp isoform):** Forward 5'-GACAATGCGTCTCTGGCACGTCTT-3', Reverse 5'-TCCTCCGCCTCCTGCAGGTTCTT-3'.
- **Sox4 (57.2°C annealing):** Forward 5'-GGTCTCTAGTTCTTGCACGCTC-3', Reverse 5'-CGGAATCGGCACTAAGGAG-3'.
- **Oct4 (60°C annealing):** Forward 5'-GACCATCTGCCGCTTTGAG-3', Reverse 5'-CCCCCTGCCCCCATTCCTA-3'.
- **Nanog (56°C annealing):** Forward 5'-CTGTGATTTGTGGGCCTGAA-3', Reverse 5'-TGTTTGCCTTTGGGACTGGT-3'.

2.5 Statistical Analysis

Statistical calculations were performed using SPSS version 27.0. Two-tailed Chi-square tests evaluated significant differences between CTC cohorts and healthy controls at a 95% confidence interval. Odds ratios (OR), standard deviations (SD), and standard errors (SE) were derived for each marker group. Significance levels were established at $p < 0.05$ (*) and $p < 0.001$ ().

2.6 Machine Learning Predictive Framework Design

An integrated machine learning workflow was established to analyze the compiled dataset using the following structure:

- **Input Layer:** Quantitative parameters incorporated flow cytometric CTC percentages for CK19, Vimentin, Sox4, and EpCAM across drug regimens and timeframes; DNACNV gel band optical densities; categorical age brackets (25–35, 35–45, 45–55, 55–65, 65–75 years); therapeutic agent (CP vs 5-AZA) and exposure duration (24, 48, 72 h); and diagnostic status (case vs control).
- **Model Architecture:** Machine learning models, including Random Forest (RF), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) networks, were implemented to predict metastatic risk scores, therapeutic sensitivity indices, EMT staging (epithelial, hybrid, or mesenchymal), and ideal drug treatment windows. SHAP (SHapley Additive exPlanations) algorithms provided feature importance ranking, and 5-fold cross-validation prevented model overfitting.
- **Validation:** Predictions generated by the machine learning framework were cross-referenced against empirical Chi-square outputs to verify model accuracy.

3. RESULTS

3.1 Flow Cytometric Enumeration of CTCs Under CP and 5-AZA Treatment

Quantification of CK19-positive CTCs following Cyclophosphamide exposure revealed a time-dependent, continuous decrease compared to control levels (3243.66 ± 863.84; 17.51%). CK19 CTC counts dropped to 2753.50 at 24 h ($p = 0.034$), 2720.10 at 48 h ($p = 0.039$), and reached a minimum of 1945.50 at 72 h ($p = 0.0001$), demonstrating potent cytostatic action.For Vimentin-expressing CTCs (control baseline: 9528 ± 4002.40; 12.19%), significant reductions were noted at 24 h (4565.50; $p = 0.0001$) and 72 h (2997.50; $p = 0.0001$). However, at the 48-hour timepoint, a non-significant decrease was observed (1055.99; $p = 0.2338$), suggesting a transient phase of cellular resistance or cell cycle phase-dependent evasion.

In contrast, 5-Azacytidine exposure resulted in significant Sox4 over-expression (OR = 18.0; $p < 0.018$), with up-regulated Sox4 DNACNV levels in 75% of breast cancer cases. EpCAM exhibited consistent diagnostic sensitivity reaching 98.98% post-treatment. Nanog demonstrated significant upregulation ($p < 0.001$) in cooperation with Sox4, whereas Oct4 exhibited no statistically significant variation relative to control groups. The flow cytometric data was summarizes in table-1 CTCs parameters and DNACNVs trends under both drug exposures..

Table 1. Comparative flow cytometric CTCs Frequencies and DNACNV profiles of EMT markers under CP and 5-AZA exposure in breast cancer patients.

EMT Marker	Drug / Condition	Mean Frequency / Expression	SD	p-value	95% Confidence Interval / Notes
CK19	Control Baseline	3243.66 (17.51%)	863.84	—	—
CK19	5-AZA (Early)	Up-regulated	Variable	$p < 0.05$	Re-expressed post-demethylation
CK19	CP (24 h)	2753.50 (16.58%)	551.5	0.034*	942.17 – 38.16
CK19	CP (48 h)	2720.10 (15.93%)	719.0	0.039*	1019.35 – 27.98

CK19	CP (72 h)	1945.50 (20.54%)	21.5	0.0001**	1679.26 – 917.06
Vimentin	Control Baseline	9528 (12.19%)	4002.40	—	—
Vimentin	5-AZA (Early)	Significant Reduction	Variable	$p < 0.001$	Modulated expression
Vimentin	CP (24 h)	4565.50 (14.51%)	1885.5	0.0001**	6913.76 – 301.1
Vimentin	CP (48 h)	1055.99 (10.38%)	682.1	0.2338 (NS)	719.64 – 2861.64
Vimentin	CP (72 h)	2997.50 (15.00%)	1073.5	0.0001**	8358.08 – 4702.91
Sox4	Control Baseline	~1618.9	216.0	—	—
Sox4	5-AZA Exposure	Up-regulated	368.6	$p < 0.05$	OR = 18.0 (511–0.53 CI)
EpCAM	Control Baseline	Baseline level	—	—	—
EpCAM	5-AZA Exposure	98.98% Sensitivity	—	$p < 0.001$	Stable diagnostic marker

***p < 0.05; p < 0.001; NS = Not Significant.**

3.2 PCR-DNACNV Profiling of Stemness and EMT Markers

Densitometric assessment of agarose gel electrophoresis bands provided relative copy number variations across targeted loci. Sox4 exhibited a non-linear expression trajectory—showing elevated band intensity in early disease stages (cases 3–5) followed by attenuated intensity in advanced stages (cases 6–8)—reflecting a multi-step metastatic pathway. Nanog demonstrated concurrent upregulation ($p = 0.001$), reinforcing its cooperativity with Sox4. Oct4 displayed variable expression without reaching statistical significance ($p = 0.88$), indicating an independent regulatory mechanism.

CK19 (573 bp) displayed continuous, time-dependent reduction in band intensity upon CP exposure ($p = 0.0001$, OR = 188). Evaluation of Vimentin revealed a primary band at 323 bp alongside a secondary truncated isoform band at 225 bp. The emergence of this 225 bp isoform correlated with altered chemo-sensitivity, suggesting that isoform switching may influence drug binding in specific CTCs subpopulations. Table 2 showing details the DNACNVs characteristics across all evaluated gene targets.

Table-2. PCR-based DNACNV band characteristics and expression profiles for EMT and pluripotency markers in breast cancer patients.

Gene Target	Amplicon Size	Observed Expression Pattern	Odds Ratio (OR)	Statistical Significance	Biological Significance
Sox4	~350 bp	Elevated in 75% of BC cases	18.0	p = 0.018*	Initial surge driving metastatic onset
Oct4	~300 bp	Elevated in 75% of BC cases	2.25	p = 0.88 (NS)	Uncoupled from Sox4-Nanog axis
Nanog	~250 bp	Elevated in 50% of BC cases	0.75	p = 0.001**	Synergizes with Sox4 to maintain stemness
CK19	573 bp	Sequential decrease under CP	188.0	p = 0.0001**	High tissue-specific susceptibility
Vimentin	323 bp	Reduced post-CP exposure	—	p = 0.05*	Major mesenchymal filament
Vimentin Isoform	225 bp	Increased drug-binding subpopulation	—	p < 0.05	Truncated variant altering therapy sensitivity

*p < 0.05; p < 0.001; NS = Not Significant.

3.3 Age-Stratified CTC Frequency Distribution

Stratification of CTC frequencies across patient age cohorts revealed an inverse relationship with advancing age. Patients in the youngest demographic (25–35 years) demonstrated the highest CTC abundance (42.86%). Subsequent age brackets exhibited lower frequencies: 28.6% in 65–75 years, ~14.3% in 45–55 years, and ~7.0% in 55–65 years.This observed peak in young patients contrasts with traditional paradigms linking CTC burden solely to older age groups. In rural regions such as Bihar, environmental factors—including chronic exposure to organophosphate pesticides and agricultural chemicals—may contribute to early epigenetic dysregulation and somatic mutations, accelerating tumorigenesis in younger women.

3.4 Machine Learning Predictions and Predictive Performance

The integrated AI framework yielded six key predictive insights regarding drug sensitivity, biomarker utility, and patient risk profiles:

● Maximal cytostatic efficacy for CP against CK19-positive CTCs occurs at 72 hours (95% prediction confidence).
● Young patients (25–35 years) display elevated baseline CK19 CTC counts and early drug susceptibility (92% confidence).
● The Vimentin 323/225 bp isoform ratio serves as an effective indicator of drug-antigen interaction dynamics (88% confidence).
● Combined Sox4 and Nanog expression profiles accurately predict metastatic initiation risk (90% confidence).
● Oct4 expression alone is insufficient for predicting pluripotency or metastatic drive (85% confidence).
● Patients aged 65 years and older exhibit reduced overall CTC heterogeneity (89% confidence).Table 3 summarizes the machine learning predictions, associated confidence ratings, and clinical interpretations.

Table-3. Machine learning-predicted outcomes, model confidence scores, and clinical implications derived from integrated CTC datasets.

Feature Input Matrix	**Drug Regimen**	**Predicted Outcome**	**Model Confidence**	**Clinical Implication**
CK19 Frequency + Age (25–35 y)	Cyclophosphamide	Early cytotoxic response	92%	Consider modified dosing schedules in young cohort
Vimentin Isoform Ratio (323/225 bp)	CP + 5-AZA Combination	Enhanced drug-binding affinity	88%	Target specific CTC subsets in combination protocols
Sox4 DNACNV + Nanog Synergy	5-Azacytidine	Elevated metastatic risk	90%	Monitor early EMT activation markers closely
EpCAM Expression + CP (72 h)	Cyclophosphamide	Peak cytostasis at 72 h	95%	Target therapeutic duration between 48 and 72 h
Oct4 Variant Expression	5-Azacytidine	Unstable stemness state	85%	Require joint Sox4-Nanog profiling

Age Demographic ≥ 65 y + CTC Count	CP / 5-AZA	Lower overall CTC diversity	89%	Reduced metastatic spread rate; routine monitoring

Confidence scores reflect cross-validated performance using Random Forest and LSTM models. Feature selection via SHAP confirmed CTC frequency and exposure duration as primary drivers.

Computational Pipeline Workflow

The machine learning workflow follows a four-stage computational process:

1. **Input Ingestion:** Processing flow cytometric CTC counts, drug parameters, age metadata, DNACNV band optical densities, and case/control ratios.
2. **Feature Extraction:** Computing relative marker frequencies, temporal reduction rates, Vimentin isoform ratios, expression indexes, and age-stratified diversity metrics.
3. **Model Training:** Applying Random Forest, Support Vector Machine, and LSTM algorithms paired with SHAP explainability layers and 5-fold cross-validation.
4. **Clinical Inference:** Generating metastatic risk scores, drug sensitivity indices, optimal treatment windows, EMT stage classifications, and biomarker rankings.

4. DISCUSSION

This synthesized investigation combines experimental flow cytometry and PCR-DNACNV data with machine learning algorithms to evaluate CTC dynamics under Cyclophosphamide and 5-Azacytidine treatment. The distinct pharmacological actions of CP (DNA crosslinking) and 5-AZA (epigenetic demethylation) induce divergent marker expression patterns in CTC subpopulations.

The reduction of CK19 and Vimentin following CP administration confirms its cytostatic efficacy, with maximal response observed at 72 hours. The transient non-significant reduction in Vimentin levels at 48 hours points to temporary cellular evasion, potentially mediated by cell-cycle arrest or isoform expression shifts. Identification of the 225 bp Vimentin isoform provides insight into altered chemo-responsiveness, serving as a potential pharmacodynamic marker for therapeutic monitoring.

Demethylation by 5-AZA leads to Sox4 over-expression, reflecting the reactivation of epigenetically silenced transcriptional drivers. The observed bimodal Sox4 expression pattern—initial elevation followed by attenuation—corresponds to sequential stages in the metastatic process. Cooperation between Sox4 and Nanog appears essential for maintaining cancer stem cell characteristics, whereas Oct4 operates independently.EpCAM maintains high diagnostic sensitivity (98.98%) across treatment conditions, supporting its role as a core capture target in CTC detection assays. Meanwhile, dynamic markers like Sox4 and Vimentin offer additional functional insights into metastatic progression and chemo-sensitivity.

The high frequency of CTCs in younger patients (25–35 years) highlights important environmental and epidemiological considerations. In agricultural regions such as Bihar, exposure to environmental toxins and pesticides may accelerate epigenetic dysregulation, driving early-onset breast cancer. Methodological strengths of this study include minimal blood volume requirements (0.5 mL), dual experimental validation (flow cytometry and PCR-DNACNV), multi-cohort integration, and explainable AI modeling using SHAP. Future research should focus on prospective validation across larger geographic cohorts and the integration of next-generation sequencing data.

5. CONCLUSION

- Cyclophosphamide induces time-dependent reductions in CK19-positive and Vimentin-positive CTC populations, achieving maximum cytostatic impact at 72 hours ($p < 0.001$).
- 5-Azacytidine treatment causes Sox4 demethylation and over-expression, acting in synergy with Nanog ($p < 0.01$) to promote metastatic signaling, whereas Oct4 does not independently maintain pluripotency.
- EpCAM exhibits high diagnostic sensitivity (98.98%) across treatment conditions, confirming its utility as an anchor marker for liquid biopsy assays.
- The 225 bp Vimentin isoform correlates with altered drug-binding capacity in CTC subsets, providing a potential biomarker for targeted chemo-regimens.
- Patients in the 25–35 year age demographic display the highest CTC frequency (42.86%), pointing to possible environmental pesticide exposure as an epigenetic accelerant in rural agricultural populations.
- The machine learning framework achieves 85–95% predictive confidence across key clinical metrics, establishing computational modeling as an effective decision-support tool in breast cancer management.

DECLARATIONS.

- **Ethics Approval and Consent to Participate:** Approved by the Institute Ethics Committee of AIIMS Patna (AIIMS/Pat/IRC/2020/610). Written informed consent was secured from all participants.
- **Patient Consent for Publication:** Consent forms were completed by all participating individuals.
- **Competing Interests:** The authors declare no competing financial or non-financial interests.
- **Funding:** Supported by the Ministry of Science and Technology, New Delhi, India (Grant No. DST/TDP/BTDT16/2021) for stem cell marker evaluations. The CP and DNACNV studies received no external funding.
- **Authors' Contributions:** AKS designed the study and Shalini,PhD student analyses the data of the EMT marker and prepared the manuscript. She also performed experimental validation and data analysis. MK clinically diagnosed the breast cancer patients .

REFERENCES

1. Cristofanilli M, et al. Circulating Tumor Cells, Disease Progression, and Survival in Metastatic Breast Cancer. *N Engl J Med.* 2004; 351: 781–791.
2. Jansson S, et al. Prognostic Impact of Circulating Tumor Cell Apoptosis and Clusters in Serial Blood Samples from Patients with Metastatic Breast Cancer. *BMC Cancer.* 2016; 16: 433.
3. Cohen SJ, et al. Relationship of Circulating Tumor Cells to Tumor Response, Progression-Free Survival, and Overall Survival in Metastatic Colorectal Cancer. *J Clin Oncol.* 2008; 26: 3213–3221.
4. De Bono JS, et al. Circulating Tumor Cells Predict Survival Benefit from Treatment in Metastatic Castration-Resistant Prostate Cancer. *Clin Cancer Res.* 2008; 14: 6302–6309.
5. Herrmann H, et al. Intermediate Filaments: Primary Determinants of Cell Architecture and Plasticity. *J Clin Invest.* 2009; 119: 1772–1783.
6. Saxena AK, Shalini, Kumar M. Flowcytometric Based Analysis of Sox4 Gene Expression — An Early Transcription Factor Influenced by 5-Azacytidine and Compare with EpCAM in Circulating Tumor Cells Isolated from Breast Cancer Patients. *J Oncology Res Ther.* 2024; 9(03): 1–7.
7. Kumar Saxena A. DNA Copy Number Variation of Sox4 Fail to Maintain Pluripotency Results to Promote Metastasis in Breast Cancer Patients. *J Oncology.* 2024; 4(2): 1164.
8. Saxena AK, Shalini, Manoj K. Impact of Cyclophosphamide to Modulate the Genetic Signature of Epithelial Mesenchymal Transition Markers CK19 & Vimentin in Circulating Tumor Cells Population and DNA Copy Number Variations in Breast Cancer Patients. *Genet Mol Res.* 2026; 25(7s).
9. Anderson JM, et al. Cytokeratin Expression Results in a Drug-Resistant Phenotype to Six Different Chemotherapeutic Agents. *Clin Cancer Res.* 1996; 2: 97–105.
10. Song G, Sun Y, Shen H, Li W. Sox4 Over-expression is a Novel Biomarker of Malignant Status and Poor Prognosis in Breast Cancer Patients. *Tumour Biol.* 2015; 36(6): 4167–4173.
11. Gadisa DA, et al. Toxicity Profile of Doxorubicin-Cyclophosphamide Regimen in Women with Breast Cancer in Ethiopia. *J Oncol Pharm Pract.* 2020; 26(8): 1912–1920.
12. Hung CM, et al. Cyclophosphamide Promotes Breast Cancer Cell Migration through CXCR4 and Matrix Metalloproteinases. *Cell Biol Int.* 2017; 41(3): 345–352.
13. Saxena AK, Singh G. Cyclophosphamide Induced Chromosomal Aberrations and Associated Congenital Malformations in Rat. *In Vitro Cell Dev Biol.* 1998; 34(10): 751–752.
14. Saxena AK, Shalini, Vaishnava V, Singh P, Kumar M. Discordance in the Frequency of Epithelial Mesenchymal Transition Markers Sox4, EpCAM and CK19 Gene Expression in Circulating Tumor Cells in Variety of Cancer Patients. *AS Cancer Biology.* 2024; 8(11): 11–17.
15. Saxena AK, Tiwari M, Singh P, Singh V. Comprehensive Mutational Spectra of Epithelial Mesenchymal Transition Markers — SOX4, EpCAM, CK19. *Int J Cancer Medicine.* 2024; 7(1): 158–182.
16. Davidson S, et al. Cytotoxicity of 5-aza-2'-deoxycytidine in a Mammalian Cell System. *Eur J Cancer.* 1992; 28: 362–368.
17. Kiziltepe T, et al. 5-Azacytidine Induces ATR-Mediated DNA Double-Strand Break Responses, Apoptosis and Synergistic Cytotoxicity. *Mol Cancer Ther.* 2007; 6(6): 1718–1727.
18. Ao Z, et al. Identification of Cancer-Associated Fibroblasts in Circulating Blood from Patients with Metastatic Breast Cancer. *Cancer Res.* 2015; 75: 4681–4687.
19. Saxena AK, Shalini, Singh P, Singh V. Differential Expression of Vimentin Gene Modulated by MTHFR C677T Polymorphism in CTCs Isolated from Breast Cancer Patients. *Genet Mol Res.* 2024; 23(1): 1–10.
20. Zhang J, et al. SOX4 Promotes the Growth and Metastasis of Breast Cancer. *Cancer Cell Int.* 2020; 20(1): 1–11.
21. Konigsberg R, et al. Detection of EpCAM Positive and Negative Circulating Tumor Cells in Metastatic Breast Cancer Patients. *Acta Oncol.* 2010; 20(1): 700–710.
22. Nettersheim D, et al. NANOG Promoter Methylation and Expression Correlation during Normal and Malignant Human Germ Cell Development. *Epigenetics.* 2011; 6(1): 114–122.
23. Prior HM, Walter MA. SOX Genes: Architects of Development. *Mol Medicine.* 1996; 2(4): 405–412.