Harnessing Multi-Omic Technologies to Reveal Novel Non-Coding Regulatory Elements in Cancer Genomes

Coupling CRISPR, single-cell RNA sequencing, and Hi-C technologies revealed novel non-coding regulatory elements that contribute to oncogenic transcriptional programming and cell proliferation of ER+ breast cancer cells.

Breast cancer is the most common cancer among women, with more than 2 million new cases reported in 2020, according to the World Health Organization[1]. In about 70% of breast cancer cases, the transcription factor Estrogen Receptor Alpha (*ERS1*) is a prime oncological driver contributing to tumor growth and disease progression[2]. Estrogen Receptor (ER) can bind directly to DNA or alter transcriptome programming by engaging other transcription factors (TF), producing genes such as *CCND1*, which regulates CDK kinases and can cause alterations to the cell cycle.

Through RNA analysis, ER was found to regulate approximately 3,000 genes; however, RNA analysis alone cannot elucidate which of these genes or regulatory elements contribute to cell proliferation or other functions necessary for an oncogenic phenotype.

Challenge: Mapping the Cistrome of Cancer Cells

To overcome the limitations of RNA analysis, scientists have developed an approach that couples CRISPR-Cas9, scRNA-seq, and Hi-C technologies to help screen genes and transcriptional regulatory elements (TREs) and reveal their mechanisms as well as contributions to cancer progression.

Researchers from Novartis Institute for Biomedical Research in Basel, Switzerland, harnessed the power of Arima Hi-C and CROP-seq to perform a comprehensive survey, screening TREs associated with ER+ breast cancer. By combining CRISPR screens with epigenetic profiling, the team was able to map the cistrome of cancer cells.

Multi-Omic Approach to Understanding Estrogen Receptor

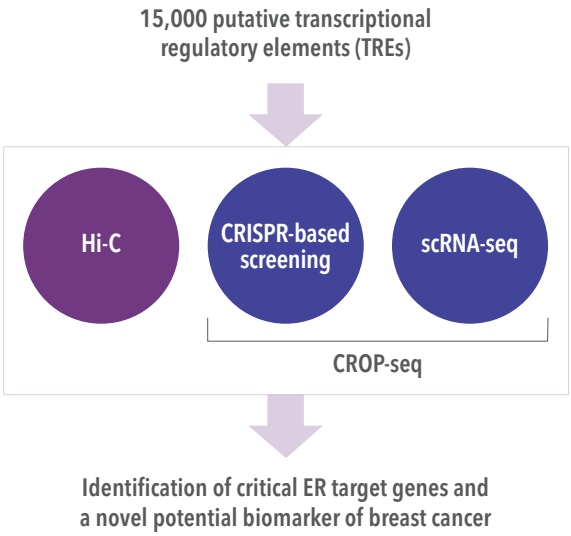
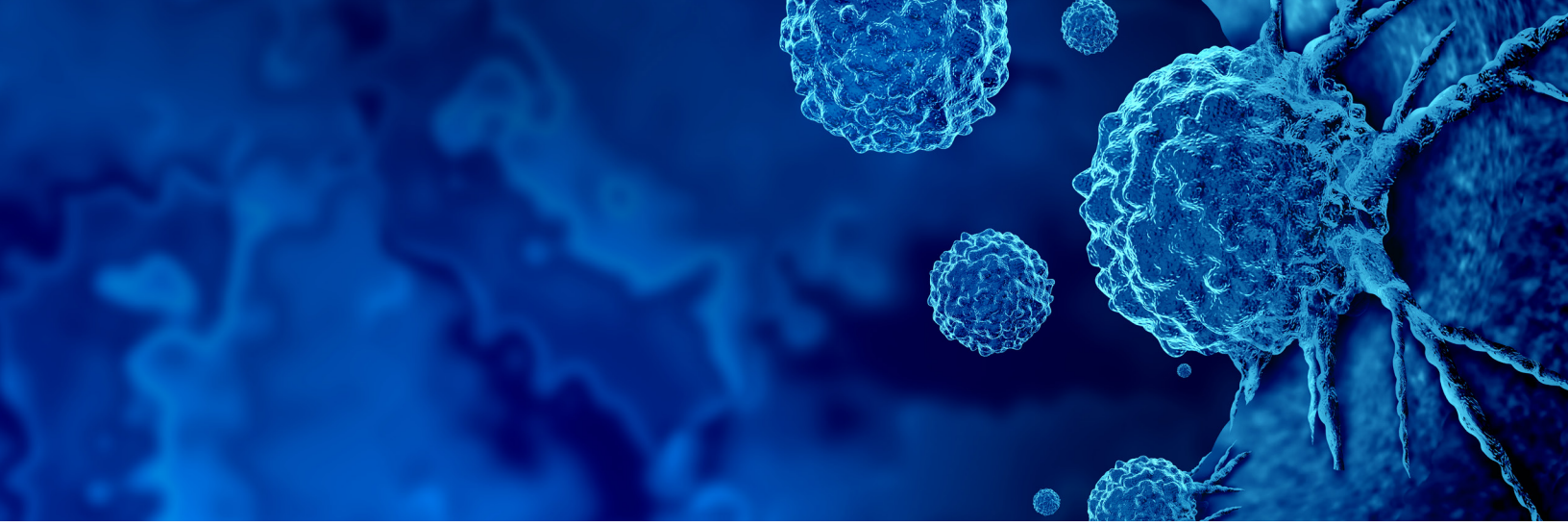


Figure 1. Lopes et al. employed a novel framework including Arima Hi-C technology for characterizing how oncogenic transcription factors can engage specific transcriptional regulatory elements to impose their pathogenic program[3].

"Hi-C is a powerful tool for conducting genome-wide analyses. We used Hi-C to narrow down the window of the scRNA-seq data and were able to successfully dissect a transcriptional regulatory network and identify the function of non-coding elements of cancer cells."

– Giorgio G. Galli, PhD, Group Leader, Oncology Drug Discovery Biology, Novartis Institutes for Biomedical Research



Approach: Dissecting the Transcriptional Regulatory Networks of ER+ Breast Cancer Cells

The Drug Discovery Team at Novartis performed a cell proliferation–based CRISPRi screening to identify critical regulatory elements contributing to *ESR1* and *CCND1* expression. In addition, Hi-C experiments were used to confirm DNA-DNA interactions between genetic elements and *ESR1* or *CCND1*. Based on this, researchers identified several critical transcriptional regulatory elements which contribute to *ESR1* and *CCND1* expression in ER+ breast cancer cells.

Previous ChIP-seq analysis revealed 15,000 ER binding sites, the majority of which have never been assessed for their functional contributions to oncogenic programming and cell proliferation. To that end, researchers designed a comprehensive CRISPR library and screened the ER binding sites for functional significance. Notably, GATA3 was enriched at enhancers, while H3K27ac was only enriched at promoters. However, the CRISPR screen revealed that only a couple of elements –1.65% of the tested regions – were required for ER-driven oncogenic programming.

To further dissect the transcriptional regulatory networks of ER+ breast cancer cells, researchers created 3D genome maps using the Arima-HiC kit coupled with CROP-seq and functional genomics analysis. CROP-seq (CRISPR droplet sequencing) combines the power of CRISPR technologies with the high-resolution functional analysis of scRNA-seq. Arima-HiC kits complement these methods by capturing the 3D genome folding of cells. Together, these methods reveal how genes interact with each other and how misfolding and misregulation can lead to cancer cell growth. This approach of combining CROP-seq and Hi-C data is a powerful high throughput method to obtain the functional contributions of TREs on oncogenic programming.

Through this multi-omic technique, the Novartis team was able to unveil the network of critical transcription factors of ER. The results demonstrate that perturbing GATA3_+1.1Mb down-regulates GATA3 expression, a known cofactor of ER, and thus contributes to ER+ breast cancer proliferation. Hi-C analysis also discovered that *TFAP2C* is linked to a region ~30kb upstream of its transcription start site – this region is denoted as a transcriptional enhancer of *TFAP2C,* or TET. Lower levels of *TFAP2C* are associated with a better patient prognosis and have been found to mainly regulate metabolic genes and a small subset of critical ER targets.

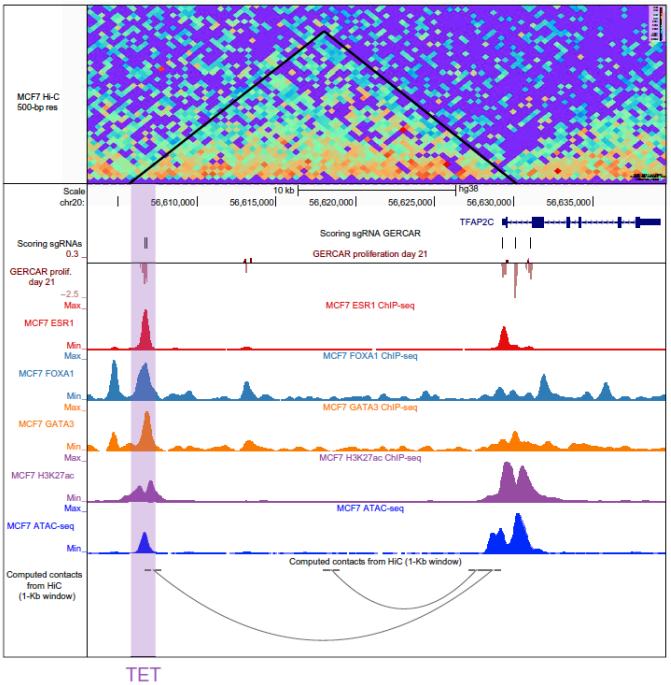


Figure 2. Hi-C 3D genome maps help reveal *TFAP2C* as a critical target gene of ER through interactions between *TFAP2C* and the Transcriptional Enhancer of *TFAP2C* (TET) [3].

To further uncover the critical role of *TFAP2C* in ER+ breast cancer, a knockdown was created using shRNAs targeting *TFAP2C* and *ESR1*, which strongly impacted cell growth. RNA-seq analysis showed 258 genes commonly regulated by *TFAP2C* and *ESR1*. The results demonstrate that ER triggers the expression of *TFAP2C* and other TFs that play a critical role in the oncogenic programming of ER+ breast cancer cells.

Impact: Improved Understanding of Transcriptional Regulation in Cancer

In this study, Novartis scientists investigated the transcriptional regulatory networks of the ER cistrome at both the genome-scale and single-cell level using a multi-omic approach, harnessing the power of CRISPR screens, scRNA-seq, and Hi-C technologies. Importantly, this study observed that only a small subset of ER binding sites are implicated in cell fitness. Systematic analyses of regulatory networks revealed the interaction between ER and downstream TFs previously unknown, such as *TFAP2C*, and the contribution of TREs enriched in GATA3 and H3K27Ac signals.

Combining Hi-C, CRISPR, and RNA-seq is a powerful approach to characterizing oncogenic transcriptional regulatory elements and revealing unknown factors within transcriptional networks. Furthermore, this approach can be applied in other models to advance understanding of transcriptional regulation in cancer and influence more robust diagnosis techniques, future drug targets, and novel treatment methods.

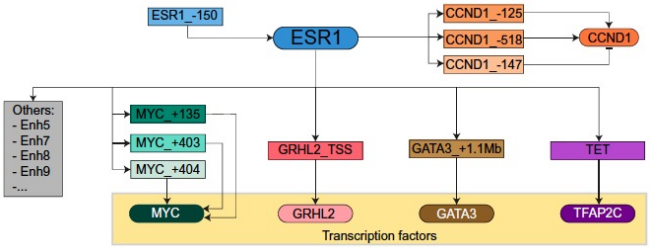
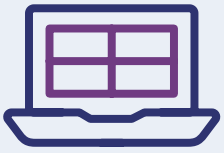


Figure 4. The newly expanded understanding of the estrogen receptor regulatory network in breast cancer cells[3].



Learn More

Webinar: Next Generation Epigenomics for Oncology: Dissecting Noncoding Regulatory Elements in Cancer Genomes

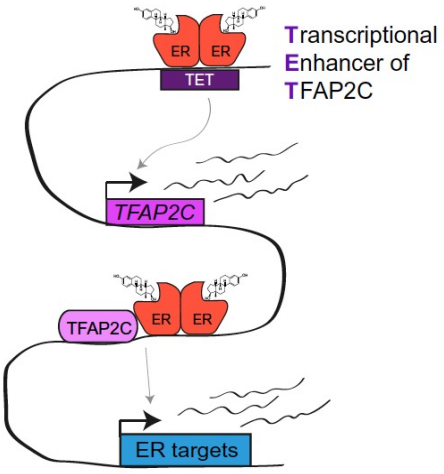
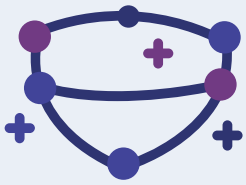


Figure 3. Model depicting the cross-talk between Estrogen Receptor, TET, and *TFAP2C* in activating downstream target genes[3].



Technology

Arima genome-wide HiC+ enables 3D genome mapping to identify how spatial relationships in DNA structure can impact gene regulation and disease processes.

References

1. Breast Cancer. (2021). WHO Fact Sheet.
2. Lei, J. T., et al. (2019). ESR1 alterations and metastasis in estrogen receptor positive breast cancer. Journal of Cancer Metastasis and Treatment. 5,38.
3. Lopes, R., et al., (2021). Systematic dissection of transcriptional regulatory networks by genome-scale and single-cell CRISPR screens. *Science advances*, 7(27), eabf5733.

arimagenomics.com | info@arimagenomics.com

CS-002-022322