Trends in Genetics

CellPress

Spotlight

Uncovering dark matter in cancer by identifying epigenetic drivers

Jun Zhong [1,*] and Laufey T. Amundadottir[1,*]

The complex relationship between chromatin accessibility, transcriptional regulation, and cancer transitions presents a daunting puzzle. Terekhanova *et al.* created a pan-cancer epigenetic and transcriptomic atlas at single-cell resolution, yielding important insights into the underlying chromatin architecture of cancer transitions and novel discoveries with the potential to advance precision medicine.

Generating a more comprehensive landscape of the complex and often cancer type-specific molecular events driving cancer has important implications for precision oncology [1]. While large-scale whole-genome sequencing (WGS) initiatives, such as those performed by The Cancer Genome Atlas (TCGA) [2] and the International Cancer Genome Consortium (ICGC) [3], have uncovered somatic mutations contributing to tumorigenesis, the enigma persists. Although important for personalized treatment in many cases, somatic mutations fall short of providing a comprehensive understanding of cancer transitions. Epigenetic features, notably chromatin remodeling, have a pivotal role in transcriptional regulation [4]. While previous studies have provided an initial characterization of the chromatin regulatory landscape in primary human cancers [5], the advent of single-cell technologies has significantly improved resolution, allowing an in-depth examination of the epigenome and transcriptome [6].

In a recent study, Terekhanova *et al.* [7] constructed a single-cell multi-omic atlas spanning 225 samples across 11 cancer types. Integration of single-nucleus assay for transposase-accessible chromatin with sequencing (snATAC-seq) with single-cell/nucleus RNA sequencing (sc/snRNA-seq) unveiled detailed relationships between chromatin accessibility and gene transcription within individual cells. This comprehensive study provides invaluable insights into epigenetic drivers across various cancers, elucidating transitions from normal adjacent-to-tumor tissues (NAT) to primary tumors, and subsequently to metastatic cancer. Analysis of snATAC-seq data, encompassing over a million nuclei, identified tissue- and cancer cell-specific differentially accessible chromatin regions (DACRs). Enhancer regions, comprising 53% of DACRs, were estimated to have a significant role in gene expression changes, often surpassing the influence of promoters. Most accessible chromatin region (ACR)-to-gene links identified were novel. In particular, 397 linked ACRs (mainly enhancers) gained accessibility in most primary pancreatic ductal adenocarcinoma (PDAC) samples, including for *ASAP2*, a gene that encodes a GTPase-activating protein, the expression of which is associated with an unfavorable prognosis. Notably, ACR-to-gene associations included transcription factor (TF) genes, such as *KLF6* and *PPARG*, the enhancers of which exhibited increased accessibility in PDAC. Another unfavorable prognostic marker for PDAC, *FLNB*, was linked to five enhancer regions, indicating extensive epigenetic regulation. Several identified enhancers were linked to *EN1*, *VIM*, and *VEGFA* in basal breast cancer (BRCA) tumors.

To better understand the underlying transcriptional regulation implicated in cancer development, 'regulons' or networks of regulatory connections between TFs and their target genes were identified [8]. This revealed 258 regulons characterized by consistent gene expression patterns between TFs and their targets, with each regulon encompassing from 20 to 4310 target genes (median: 372). Of note, 46 regulons were cancer cell specific. To further substantiate select findings, a cleavage under targets and release using nuclease (CUT&RUN) assay was conducted in U251 glioblastoma (GBM) cells, showing direct binding of NRF1 at the promoters of target genes.

The study comprised 52 metastatic samples from six tumor types, including nine paired primary–metastatic samples, allowing for the identification of epigenetic alterations associated with metastatic potential. Notably, *ELF3* and *GATA6* emerged as top-tier TFs exhibiting significantly decreased motif accessibility across regulatory elements in PDAC. Validation in a genetically engineered mouse model confirmed a decline in *GATA6* activity in PDAC metastases, underscoring the potential role of GATA6 in the metastatic cascade of PDAC.

To resolve interactions between epigenetic and genetic drivers, a comprehensive examination of somatic mutations and copy-number variations (CNVs) was undertaken in 176 tumor samples, leveraging whole-exome sequencing (WES) data. Specifically, snATAC-seq analysis revealed significantly elevated *TERT* promoter (*TERTp*) accessibility in tumors harboring the hotspot C228T and C250T mutations, and increased *TERT* expression compared with normal-derived cells. Moreover, the expression of 30 oncogenes was intricately tied to enhancer accessibility. The most robust connections were identified for pivotal oncogenes, including *EGFR*, *KRAS*, *ERBB2*, *CTNNB1*, and *MET*.

This work highlights the importance of epigenomic programs with substantial clinical implications. The evaluation of regulon activity, as determined by bulk RNA-seq expression data in patients with GBM and PDAC from TCGA, underscored a noteworthy association between heightened *PITX3* and *KLF6* regulon activity, respectively, and unfavorable progression-free and overall survival outcomes. Furthermore, differentially expressed genes (DEGs) and DACRs were

identified as potential drug targets, including already known examples, such as *ESR1* in BRCA and uterine corpus endometrial carcinoma (UCEC), and *VEGFA* in clear cell renal cell carcinoma (ccRCC) and colorectal cancer (CRC). Additionally, known drug targets were enriched in cancer types wherein these targets are not currently used in clinical practice, including *EGFR* accessibility in ccRCC, *TOP1* expression in UCEC, multiple myeloma (MM) and ccRCC, and *FGFR2* expression in GBM, ccRCC, and basal BRCA. These results suggest important targets that could be used therapeutically in novel cancer types, and warrant further preclinical validation.

This study also provides a compelling approach poised for continued exploration. In tandem with this unique resource, future studies with greater statistical power will be important to broaden our knowledge of dynamic epigenetic signals in cancer transitions. Moreover, augmenting such a data set with additional omics information, such as chromatin interactions (e.g., H3K27ac HiChIP), could provide direct evidence linking enhancers to their target genes. Additionally, incorporating WGS data offers an avenue to explore the interaction between epigenetic and noncoding somatic drivers. Finally, comprehensive functional validation will be imperative to affirm the clinical relevance of novel findings.

This impressive study helps us envision a future in which cancer management will be advanced by rigorous multi-omic analyses of tumor enhancer, promoter, and regulon activity, DNA methylation, gene expression, and somatic mutations, in conjunction with assessment of rare and common inherited variants. This could facilitate more precise tumor classification and improve personalized treatment recommendations through innovative approaches, such as liquid biopsies (Figure 1). A recent proof-of-concept study in advanced cancers demonstrated the potential of plasma epigenomic profiling to unlock clinically actionable insights, currently accessible only through direct tissue sampling [9]. This new large-scale single-cell multi-omic atlas and rapidly evolving technologies (e.g., liquid biopsies and artificial intelligence [10]) hold the promise of transformative breakthroughs that could significantly improve cancer management, ushering us toward a future where cancer is not only fully understood, but also conquered.

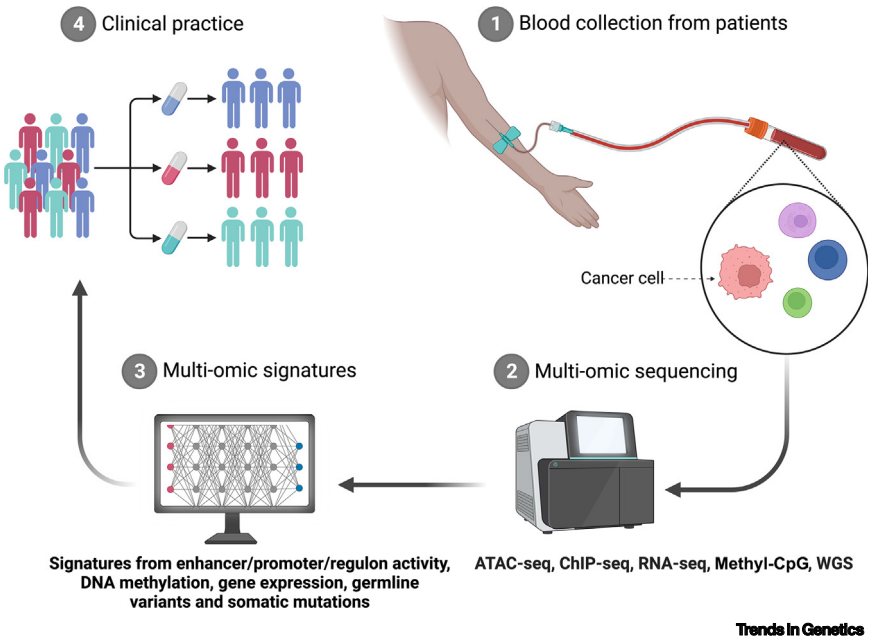

Figure 1. A possible future approach to personalized cancer management with rigorous multi-omic analyses. (1) Peripheral blood drawn from patients to isolate normal and circulating tumor DNA. (2) Multi-omics sequencing. (3) Detection of multi-omic signatures, such as enhancer, promoter, and regulon activity, DNA methylation, gene expression, as well as germline variants and somatic mutations. (4) Providing personalized medical solutions for patients based on these signatures. Figure created with BioRender (BioRender.com). Abbreviations: ATAC-seq, assay for transposase-accessible chromatin with sequencing; RNA-seq, RNA sequencing; WGS, whole-genome sequencing.

Acknowledgments

The authors are supported by the Intramural Research Program (IRP) of the Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute (NCI), and US National Institutes of Health (NIH). The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Declaration of interests

The authors declare no competing interests.

[1]Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

*Correspondence:
jun.zhong@nih.gov (J. Zhong) and
amundadottirl@mail.nih.gov (L.T. Amundadottir).

https://doi.org/10.1016/j.tig.2023.12.005

Published by Elsevier Ltd.

References

1. Li, Y. *et al.* (2023) Pan-cancer proteogenomics connects oncogenic drivers to functional states. *Cell* 186, 3921–3944
2. Hutter, C. and Zenklusen, J.C. (2018) The Cancer Genome Atlas: creating lasting value beyond its data. *Cell* 173, 283–285
3. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium (2020) Pan-cancer analysis of whole genomes. *Nature* 578, 82–93
4. de Miguel, F.J. *et al.* (2023) Mammalian SWI/SNF chromatin remodeling complexes promote tyrosine kinase inhibitor resistance in EGFR-mutant lung cancer. *Cancer Cell* 41, 1516–1534
5. Corces, M.R. *et al.* (2018) The chromatin accessibility landscape of primary human cancers. *Science* 362, eaav1898
6. Zhang, K. *et al.* (2021) A single-cell atlas of chromatin accessibility in the human genome. *Cell* 184, 5985–6001
7. Terekhanova, N.V. *et al.* (2023) Epigenetic regulation during cancer transitions across 11 tumour types. *Nature* 623, 432–441
8. Aibar, S. *et al.* (2017) SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* 14, 1083–1086
9. Baca, S.C. *et al.* (2023) Liquid biopsy epigenomic profiling for cancer subtyping. *Nat. Med.* 29, 2737–2741
10. Zhong, J. *et al.* (2023) Artificial intelligence and improved early detection for pancreatic cancer. *Innovation* 4, 100457