



HAL
open science

Open for connections: HiCAR reveals the interactions of accessible DNA

Benoit Moindrot, Daan Noordermeer

► **To cite this version:**

Benoit Moindrot, Daan Noordermeer. Open for connections: HiCAR reveals the interactions of accessible DNA. *Cell Genomics*, 2022, 2 (4), pp.100121. <10.1016/j.xgen.2022.100121>. <hal-03818673>

HAL Id: hal-03818673

<https://hal.science/hal-03818673v1>

Submitted on 18 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Preview

Open for connections: HiCAR reveals the interactions of accessible DNA

Benoit Moindrot¹ and Daan Noordermeer^{1,*}¹Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France*Correspondence: daan.noordermeer@i2bc.paris-saclay.fr<https://doi.org/10.1016/j.xgen.2022.100121>

Hi-C is a powerful technology for exploring 3D genome organization on a genome-wide scale, yet it can be financially and computationally challenging. In a recent issue of *Molecular Cell*, Wei et al.¹ introduce HiCAR, which simplifies Hi-C by targeting the 3D interactions of accessible regions only.

Hi-C (high-throughput chromosome conformation capture) has rapidly become the gold standard for the exploration of 3D genome organization.² The capacity of this assay to analyze genome-wide 3D organization at high resolution has unearthed different levels of chromosome organization, which notably include Hi-C compartments (compartment A/B) and TADs (topologically associating domains).^{2–4} Although instrumental, high-resolution Hi-C studies require excessively large numbers of sequencing reads (i.e., many billions to generate sub-kb interaction maps), making Hi-C studies expensive and computationally challenging. To circumvent these limitations, diverse derivatives of Hi-C have emerged that focus the analysis on sites with specific characteristics. Examples include pre-selected sites in the genome (capture Hi-C) and sites bound by specific proteins or histone modifications (HiChIP and PLAC-seq). In a recent issue of *Molecular Cell*, Wei et al.¹ expand the Hi-C toolbox by developing the HiCAR assay (Hi-C on accessible regulatory DNA), which explores the 3D interactions of accessible regions in the genome, thereby preferentially targeting regulatory regions.¹

HiCAR uses a sophisticated library preparation protocol that incorporates the Tn5 transposase to target accessible DNA within crosslinked chromatin, where it will insert a complex DNA adaptor with a 5'-TA overhang (Figure 1A, tagmentation in step 1). Next, the material is digested without preference for accessible chromatin by the frequent cutting *CviQI* restriction enzyme. The resulting 5'-TA

overhangs permit proximity ligation between tagged fragments and the *CviQI*-digested sites (Figure 1A, step 2; “open-versus-all” ligation). After removal of crosslinks and DNA repair, short linear DNA fragments are generated using the frequent cutting *NlaIII* enzyme. Because of the tagmentation strategy, only junctions between an accessible region and a captured partner will contain the Illumina adaptor sequences, separated by a recognition site for the infrequent cutting *PmeI* restriction enzyme (Figure 1A, step 3). The short fragments are again ligated (step 4) and digested using *PmeI*, which results in the selective linearization of fragments containing open-versus-all ligation junctions (Figure 1A, step 5). The presence of Illumina adaptors at both extremities next permits library amplification and paired-end sequencing, with R2 reads representing accessible (tagmented) fragments and R1 reads corresponding to their interactions. Using this setup, HiCAR data enabled robust detection of Hi-C compartments, TADs, and chromatin loops with similar sensitivity as Hi-C despite considerably fewer reads.

In recent years, four other Hi-C derivatives have been developed to target 3D interactions of accessible regions in the genome (Figure 1B). These assays can be divided into two categories, based on their method for enrichment of accessible DNA. A first strategy, similar to HiCAR, uses enzymes that specifically fragment accessible chromatin. In the Trac-looping assay, the Tn5 transposase—here assembled as a tetramer—is used to link two accessible regions in the genome,

thereby circumventing the need for proximity ligation (Figure 1B).⁵ Alternatively, the NicE-C assay uses the Nt.CviPII nicking enzyme to fragment accessible chromatin, followed by conventional proximity ligation (Figure 1B).⁶ The second general strategy is based on the preferential enrichment of accessible DNA in two-phase separation systems. By subjecting a proximity-ligated Hi-C library, prior to de-crosslinking, to either a phenol-chloroform DNA extraction (OCEAN-C) or silica-gel adsorption (HiCoP), accessible ligation junctions are enriched (Figure 1B).^{7,8} These four assays strongly favor the capture of interactions between two accessible regions (“open-versus-open”). In contrast, the sophisticated use of dedicated adaptor sequences and the multiple rounds of DNA digestion in the HiCAR assay uniquely allow the exploration of open-versus-all interactions. HiCAR thus provides more comprehensive views of genome organization of accessible regions in the genome.

Using HiCAR in human cells, Wei et al. find that up to 75% of tagged (R2) reads overlap known accessible sites (as previously detected by ATAC-seq).¹ Around 60% of these reads overlap sites with regulatory characteristics (promoters and enhancers) or binding sites of the CTCF protein (a protein associated with TAD structuration³). In contrast, the R1 reads are essentially not enriched for overlap with ATAC-seq peaks, thus confirming the capacity of HiCAR to detect interactions anywhere in the genome.

By focusing on HiCAR interactions that overlap with regulatory regions, the



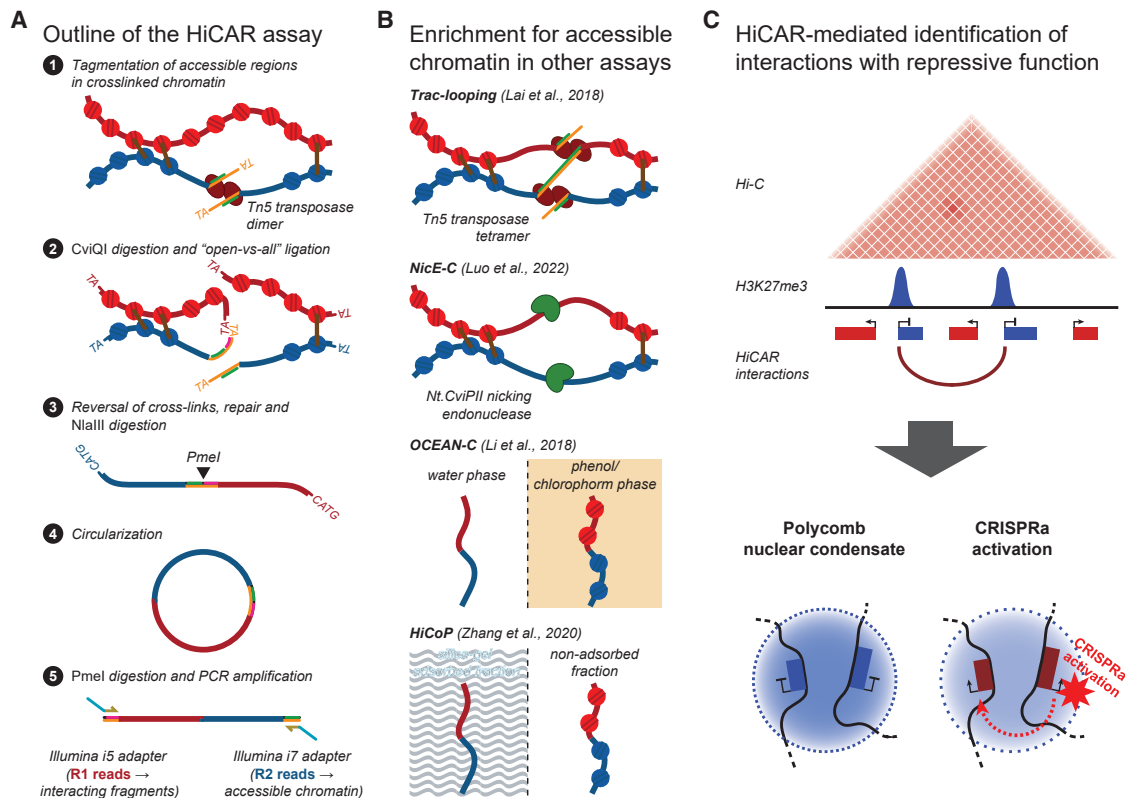


Figure 1. The open-versus-all HiCAR method and its identification of repressive promoter-promoter contacts

(A) Outline of the experimental steps in the HiCAR assay to target interactions between accessible DNA and other sites in the genome. HiCAR is based on a multi-step protocol that permits the detection of contacts between accessible regions, which are specifically enriched by Tn5 transposition, and any other site in the genome (*open-versus-all*). The engineered transposed linker contains the Illumina adaptor sequences (orange and pink) separated by a PmeI restriction site. In the HiCAR sequencing output, tagmented fragments will be in the R2 reads (blue) and their interacting fragments in the R1 reads (red).

(B) Principle of enrichment for accessible chromatin in other *open-versus-open* assays. Enrichment is either achieved by using DNA fragmenting enzymes that favor accessible chromatin (Trac-looping and NicE-C) or by using preferential enrichment of accessible DNA in a two-phase separation system (OCEAN-C and HiCoP).

(C) Top: HiCAR identifies hundreds of contacts between inactive promoters that are bound by repressive Polycomb group proteins (and the associated H3K27me3 histone modification). Bottom: Targeting of CRISPRa to a repressed gene also activates its interacting partners, possibly due to their co-occupancy of a shared Polycomb nuclear condensate.

authors make a number of interesting observations. In the human GM12878 B-lymphoblastoid cell line, most significant interactions involve active regulatory regions (H3K27ac-marked), although 2% of interactions incorporate poised or bivalent regions (H3K27me3-marked). In human H1 embryonic stem cells, these interactions are markedly more diverse, with both large numbers of active (28%) and poised or bivalent (19%) regions captured. Interactions involving genes are strongly enriched for eQTLs (expression quantitative trait loci), indicating that HiCAR is capable of identifying promoter-enhancer loops. Interestingly, in both cell types, HiCAR identifies several hundreds of interactions between pairs of inactive gene promoters as well. These

inactive promoters are enriched for the binding of repressive Polycomb group proteins, extending previous observations in mouse cells that facultatively repressed genes frequently interact.⁹ Using *in vivo* CRISPRa experiments, the authors confirm that Polycomb-repressed promoters can silence their interacting promoters over distance. The application of HiCAR has therefore revealed a new level of transcriptional regulation mediated by the network of long-range Polycomb interactions, potentially through the creation of nuclear condensates dedicated to repression (Figure 1C).¹⁰

The technical advantage of HiCAR is further supported by its application on 30,000 primary human muscle stem cells, which allowed the detection of muscle-

specific 3D structures at two key regulators of myogenesis. This demonstrates that HiCAR is compatible with clinical biopsies or other material for which cell numbers are limiting. Moreover, the HiCAR protocol allows the extraction of cytoplasmic and nuclear RNA, thus allowing RNA-seq from the same, potentially precious, cellular samples.

Altogether, HiCAR makes for an interesting new addition to the toolbox of Hi-C derivatives, allowing a unique and efficient exploration of *open-versus-all* interactions. Its improved capacity to comprehensively target accessible DNA, as compared to other assays that require either the pre-selection of loci (Capture Hi-C) or interacting proteins (HiChIP and PLAC-seq), or that are limited to

open-versus-open contacts (Figure 1B), allows for more inclusive studies of 3D genome organization and its connections to gene regulation, as confirmed by the identification of a new layer of repressive promoter-promoter interactions.

ACKNOWLEDGMENTS

Work in the Noordermeer lab is supported by funds from the Agence Nationale de la Recherche (grants ANR-18-CE12-0022-02 and ANR-21-CE12-0034-01) and PlanCancer (grant 19CS145-00).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Wei, X., Xiang, Y., Peters, D.T., Marius, C., Sun, T., Shan, R., Ou, J., Lin, X., Yue, F., Li, W., et al. (2022). HiCAR is a robust and sensitive method to analyze open-chromatin-associated genome organization. *Mol. Cell* **82**, 1225–1238.e6. <https://doi.org/10.1016/j.molcel.2022.01.023>.
2. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293. <https://doi.org/10.1126/science.1181369>.
3. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380. <https://doi.org/10.1038/nature11082>.
4. Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* **19**, 789–800. <https://doi.org/10.1038/s41576-018-0060-8>.
5. Lai, B., Tang, Q., Jin, W., Hu, G., Wangsa, D., Cui, K., Stanton, B.Z., Ren, G., Ding, Y., Zhao, M., et al. (2018). Trac-looping measures genome structure and chromatin accessibility. *Nat. Methods* **15**, 741–747. <https://doi.org/10.1038/s41592-018-0107-y>.
6. Luo, Z., Zhang, R., Hu, T., Zhu, Y., Wu, Y., Li, W., Zhang, Z., Yao, X., Liang, H., and Song, X. (2022). NicE-C efficiently reveals open chromatin-associated chromosome interactions at high resolution. *Genome Res.* **32**, 534–544. <https://doi.org/10.1101/gr.275986.121>.
7. Li, T., Jia, L., Cao, Y., Chen, Q., and Li, C. (2018). OCEAN-C: mapping hubs of open chromatin interactions across the genome reveals gene regulatory networks. *Genome Biol.* **19**, 54. <https://doi.org/10.1186/s13059-018-1430-4>.
8. Zhang, Y., Li, Z., Bian, S., Zhao, H., Feng, D., Chen, Y., Hou, Y., Liu, Q., and Hao, B. (2020). HiCoP, a simple and robust method for detecting interactions of regulatory regions. *Epigenetics Chromatin* **13**, 27. <https://doi.org/10.1186/s13072-020-00348-6>.
9. Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., et al. (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* **47**, 1179–1186. <https://doi.org/10.1038/ng.3393>.
10. Tatomosian, R., Kent, S., Brown, K., Yao, T., Duc, H.N., Huynh, T.N., Zhen, C.Y., Ma, B., Wang, H., and Ren, X. (2019). Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J. Biol. Chem.* **294**, 1451–1463. <https://doi.org/10.1074/jbc.RA118.006620>.