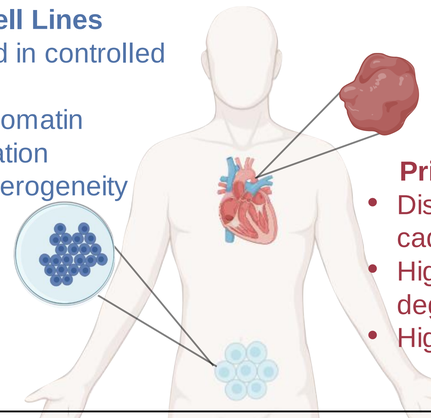
Dual Cross-Linking Improves Ability to Map Chromatin Loops

Shaiza Pasha[1,3], Ragini Mahajan[2,3], Muhammad Shamim[2,3], Erez Lieberman Aiden[1,2,3], Greg Morrison[1,2]
[1]University of Houston, [2]Rice University, [3]Baylor College of Medicine

Introduction

- **Hi-C** is a method that maps the 3D architecture and folding pattern of the genome, generating annotations of features, like DNA **loops** and domains, that bring distal elements in proximity and **influence their regulation**.
- Cross-linking samples is the first step in Hi-C, which is akin to "shooting glue through the nucleus". The kinds and amounts of cross-linkers used can affect the quality of contacts observed in the Hi-C data.
- While Hi-C has been generally robust, the quality of the data can be dependent on the inherent state of the input sample. In this study, we compare **how cross-linking variations affect Hi-C data generated from standard cell-lines versus novel primary tissues, and find if dual cross-linking can improve ability to map loops.**

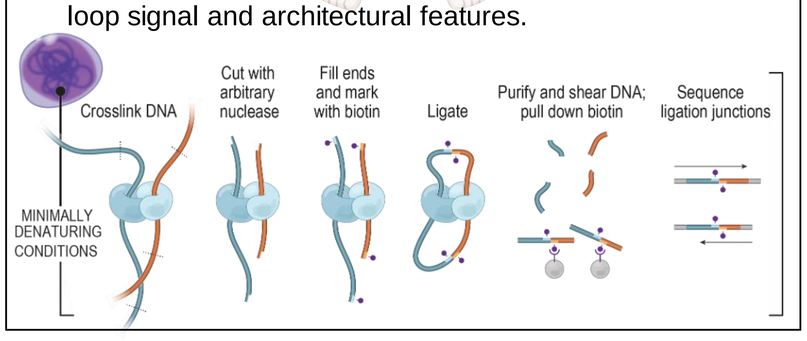
Cell Lines
- Cultured in controlled settings
- Low chromatin degradation
- Low heterogeneity



Primary Tissue
- Dissected from cadavers/patients
- Higher chromatin degradation
- High heterogeneity

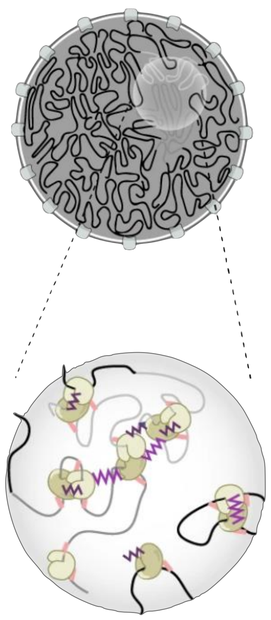
Methods

- We test cross-linking with **1% Fa** for 10 minutes (control) or **1% Fa + 3mM DSG** for 30 minutes on GM12878 cultured cells and human aorta tissue.
- We then process samples using **Intact Hi-C**, generating contact maps and quality statistics using the Juicer 2 pipeline. We then compare results between GM12878 and aorta tissue to see differences in quality statistics, loop signal and architectural features.

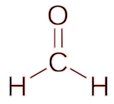


Results

Cross-Linking Chemistry

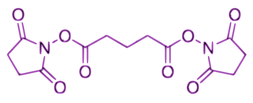


Formaldehyde (Fa)



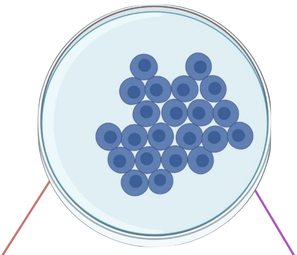
DNA-protein interactions

DSG

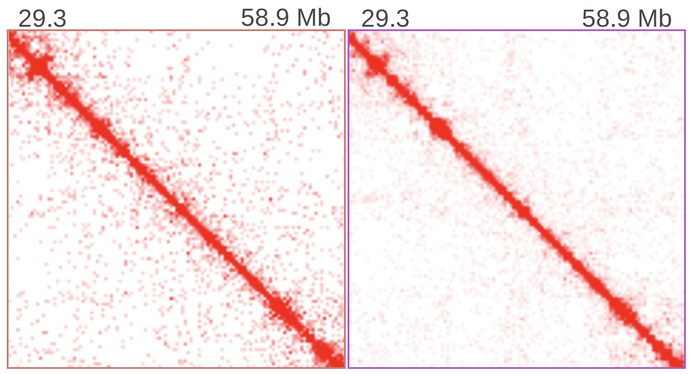


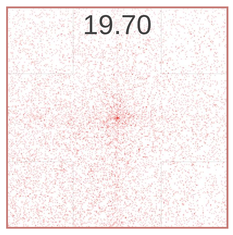
protein-protein interactions

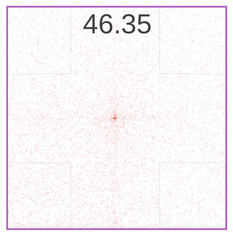
GM12878 Cell-Line



Fa | Fa + DSG



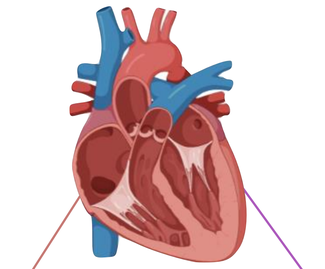




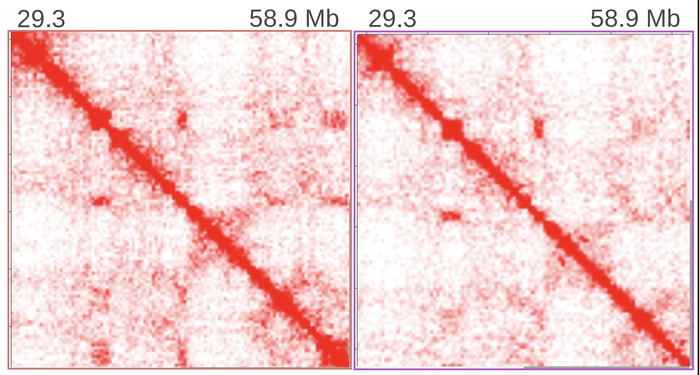
Sequenced Reads	5M
Inter-chromosomal	10.99%
Long Range	23.54%
APA	19.7
DNA Conc. (ng/uL)	14.7

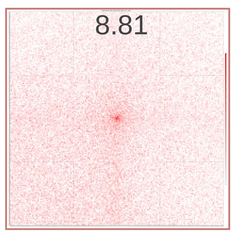
Sequenced Reads	5M
Inter-chromosomal	4.51%
Long Range	17.08%
APA	46.35
DNA Conc. (ng/uL)	73.8

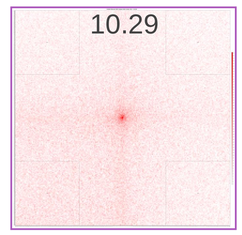
Primary Aorta Tissue



Fa | Fa + DSG







Sequenced Reads	13M
Inter-chromosomal	20.25%
Long Range	29.65%
APA	8.809
DNA Conc. (ng/uL)	10.8

Sequenced Reads	13M
Inter-chromosomal	12.65%
Long Range	27.48%
APA	10.295
DNA Conc. (ng/uL)	24.6

Discussion

- Dual cross-linking on GM12878 cells improves noise statistics, showing a lower % of inter-chromosomal contacts than its Fa-only counterpart. This trend is also observed in the aorta tissue. Along with a lower inter-chromosomal %, we observe higher DNA yield in the Fa+DSG samples, both in GM12878 and aorta tissue. We suspect that this could be due to greater chromatin stability as a result of the double cross-linking.
- Aggregate Peak Analysis (APA) of GM12878 vs Aorta tissue shows an increase in APA score in Fa+DSG samples, signaling improvement of looping signal in both, GM12878 cells and aorta tissue.
- Our findings suggest that dual cross-linking **improves overall quality** of both tissues and cell-lines comparably in terms of noise levels and DNA yields, while also **improving loop signal** recognition**.**
- For future work, we will test this method on a broader panel of cell lines and human tissues to gauge robustness for new standards.

Acknowledgements and References

Acknowledgments: Thank you to everyone at The Aiden Lab in The Department of Human and Molecular Genetics at Baylor College of Medicine, The Encode Project at Stanford University, The Center for Theoretical and Biological Physics at Rice University, and The College of Natural Sciences & Mathematics at The University of Houston.
References: Oksuz, A.B. et al., *Nat. Methods*, **18** 1046–1055 (2021), Rao, S.S.P. et al., *Nature Cell Biology*, **159** 1665-1680 (2014), Images made by BioRender & Keynote.