The Effect of Nuclear Perturbations on the 3D Organization of the Genome

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Rosela Golloshi

December 2020

DEDICATION

This dissertation is dedicated to my parents Dhionis and Rajmonda Golloshi and my late aunt Mariana Spahija for all their continued support and never-ending love.

ACKNOWLEDGEMENTS

All the hard work I have been able to dedicate through graduate school would not have been possible without the help of so many, both individuals and organizations. I will forever be grateful for the love and support throughout this challenging but fulfilling period in my life.

To my dearest parents Dhionis and Rajmonda Golloshi: Thank you for providing such a wonderful life for me and my siblings. Thank you for instilling in me the importance of hard work and always encouraging me to pursue my dreams. Your dedication and sacrifices that you continue to make for us, have been the biggest driving force in the completion of my PhD. You are my best friends and my biggest role models.

To my advisor and mentor Dr. Rachel Patton McCord: Words cannot express how grateful I am for your guidance and mentorship throughout my tenure. Thank you for creating such a positive environment in the lab and sharing your passion for science. I am grateful for the time you dedicated, that ensured I received the proper training. Thank you for encouraging me to gain exposure and experience by attending workshops, seminars, and meetings (national and international). And, to always seize the opportunity to present our work, whenever it came along. I genuinely appreciate the advice and opportunities you have provided for me, that enabled be to grow as a scientist. Thank you for always being so lenient even when my mistakes came at a great cost to you (I will hopefully pay you back one day). I would also like to thank you for giving me the freedom to explore my

thoughts and ideas in my research. Above all, I am grateful for your presence in my life, your drive and passion for science ground me and make me remember how much I love to do science.

To my other mentor Dr. Keerthi Krishnan: First, thank you for pushing and challenging me into always giving my best and realizing my potential. Thank you for being such a strong and independent role model in my life. I have appreciated the welcoming environment you created as soon as you stepped foot at UTK; thank you for all the volleyball gatherings and letting me destress with Saha. It means the world to me.

To the rest of my committee members, Dr. Mariano Labrador, Dr. Bruce McKee, Dr. Francisco Barrera and Dr. Christopher Baker: Thank you for dedicating time out of your lives to provide constructive feedback on my work. To Dr. Labrador and Dr. McKee, thank you for your support in ensuring a successful future career for me. I appreciate you all.

To the McCord lab members, past and present: Thank you for all the memories you were able to provide and all the help in the lab. I owe a great deal of gratitude to my lab partner Dr. Jacob Tyler Sanders for being present, since the beginning, and have always been a constant support through graduate school struggles. I am also thankful for my other lab partner, Priyojit Das, for his willingness to always help me figure out my computational problems at any hour of the day (literally). Finally, to Dr. Rebeca San Martin, even though

we may not have seen eye to eye some days, I have an immense gratitude for your advice, mentorship, and friendship.

To one my closest friends, Jyotirmoy Mondal and Dr. Gabriel Fuente: I could not begin to imagine graduate school without you. Thank you for all your love and support through every challenge life has thrown at me while in graduate school. Your friendship is so special to me and I will forever cherish it.

To the BCMB staff, Miss Cheryl, LaShel, Chrissy and Lori: I am forever grateful for your constant support and love. Thank you for always making my life easier and being understanding when I forgot to meet administrative deadlines.

To the 2015 BCMB cohort and the rest of my BCMB friends: I am very thankful for all the times we shared and unforgettable memories we made throughout the years.

To my brother and sister, Krisi and Eleni: Thank you for loving me and always being there. Thank you for all the headaches you have provided to distract me from the graduate school stresses. To my sister, I am very grateful for all your love and support and I cannot imagine what I would have done without you.

Finally, to my fiancé, Cashif Pritchard: I am forever indebted to you for all the immense support you have provided for me throughout all these years. Thank you for always

cheering me on and making me believe that I was good enough to do this. Above all, thank you for being here despite all the challenges we have had to endure. We made it!

ABSTRACT

Cells in our body experience constant mechanical forces that influence biological functions such as growth and development. The nucleus has been implicated as a key mechanosensor and can directly influence chromatin organization and epigenetic alterations leading to gene expression changes. However, the mechanism by which such mechanical forces lead to genomic alterations and expression of mechanosensitive genes is not fully understood. The work presented in this dissertation investigates the effect of mechanical and epigenetic perturbations on the 3D genome organization. To investigate this 3D genome folding, we use Chromosome Conformation Capture followed by high throughput sequencing (Hi-C) (Chapter-1) which identifies a hierarchical organization of the genome to ensure proper gene regulation. We then investigate how this non-random 3D genome organization is affected when cancer cells undergo compressive and tensile forces induced by migration through constricted spaces (Chapter-2) and the pre-existing structural patterns that may enable such dramatic nuclear deformations (Chapter-3). Chapter 4 includes a detailed characterization of structural patterns associated with global decompaction of the chromatin, which has been shown to modulate nuclear mechanical properties. Lastly, Chapter-5 characterizes the effect of mechanical stretching of cells on the 3D genome structure. Overall, these findings provide evidence about the role of the 3D genome organization in ability of cells to withstand mechanical perturbations by not only regulating transcriptional networks but also aiding as a protective barrier from such forces that could jeopardize the nuclear integrity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ATTACHMENTS

Phase contrast live cell imaging in a 2D culture dish of a Top-5 cell
(Movie_S1_Top5_2DMigration.avi)

Phase contrast live cell imaging in a 2D culture dish of Bottom-5 cells
(Movie_S2_Bottom5_2DMigration.avi)

3D rendering of Top-5 nucleus using Leica Sp8 software
(Movie_S3_Top5_3D_nuclear_rendering_final.avi)

3D rendering of Bottom-5 nucleus using Leica Sp8 software
(Movie_S4_Bottom5_3D_nuclear_rendering_final.avi)

Live time-lapse imaging of Top-5 cells migrating through 3D collagen matrices
(Movie_S5_Top5_3D_Collagen_Migration.avi)

Live time-lapse imaging of Bottom-5 cells migrating through 3D collagen matrices
(Movie_S6_Bottom5_3D_Collagen_migration.avi)

3D rendering of nucleus of collagen embedded Top-5 cell
(Movie_S7_Top5_3DCollagen_nuclear_rendering_final.avi)

3D rendering of nucleus of collagen embedded Bottom-5 cell
(Movie_S8_Bottom5_3DCollagen_nuclear_rendering_final.avi)

INTRODUCTION

Cells in our body are constantly exposed to mechanical forces such as tensile forces in myocytes (Magid and Law, 1985), compressive forces in chondrocytes (Freeman et al., 1994), shear forces in endothelial cells (Dewey et al., 1981) and traction forces in migrating cells (Lange and Fabry, 2013). However, the process by which individual cells detect the mechanical signals and convert them into biochemical signals and subsequently gene expression changes (a process referred to as mechanotransduction) is not fully understood. The ability of these cells to sense and respond to these forces influences the biological function of the cells such as embryogenesis (Le et al., 2016), development and maintaining tissue homeostasis (Barnes et al., 2017).

Of particular interest to this dissertation work, cancer cells experience many of these forces during tumor growth and invasion. As a tumor grows, cancer cells are exposed to more rigid environments due to deposition of extracellular matrix (ECM) in the surrounding tissue or the increase of cell-cell collision due to rapid proliferation of these cells (Jaalouk and Lammerding, 2009, Paszek and Weaver, 2004, Paszek et al., 2005). Subsequently, when cancer cells disseminate from the primary tumor and invade to surrounding tissue and metastasize, they encounter various forces such as compressive and tensile forces when squeezing through the ECM, shear forces when they travel through blood vessels and forces originating from changes in tissue mechanics when they colonize to secondary sites (Friedl and Wolf, 2010).

It has been shown that these mechanical forces can influence cell morphology and behavior (Reidy and Bowyer, 1977, Tucker and Meats, 1976) from the contribution of

mechanosensitive elements such as ion-channels, cell-cell junctions, adhesion and cytoskeletal components that often activate Rho-family GTPases or the mitogen-activated protein kinase-extracellular signal-regulated kinase (MAPK-ERK), induce nuclear translocation of transcription factors YAP/TAZ and MKL1 and ultimately induce expression of genes known to be involved in mechanoresponse (Sun et al., 2016, Murthy et al., 2017, Panciera et al., 2017).

Aside from the role the cytoplasmic components in mechanosensing, it is now evident that the nucleus can also respond directly to mechanical stress by signal propagation through the cytoskeleton and that this response can cause deformation of the nucleus, chromatin and nuclear bodies and can modulate gene expression (Bissell et al., 1982, Wang et al., 2009, Guilluy et al., 2014, Tajik et al., 2016). Specifically, the compressive forces experienced by cancer cells migrating through constrictive environment jeopardizes the integrity of the nuclear envelope and has been shown to induce rupture of the nucleus. Such nuclear rupture induces loss of DNA leading to genome instability, DNA damage and genomic aberrations (Denais et al., 2016, Irianto et al., 2017a, Raab et al., 2016) all of which could influence progression of cancer.

The role of nuclear lamina and chromatin in maintaining nuclear integrity

The nucleus is the largest and stiffest organelle of the cell and is comprised of the nucleus interior, which houses chromatin, and the surrounding nuclear envelope (Caille et al., 2002, Lammerding, 2011). The nuclear envelope is composed of outer and inner nuclear membrane (ONM and INM) (Dreger et al., 2001) and nuclear pore complexes (NPC) that

regulate the entry of large molecules inside the nucleus (Jovanovic-Talisman and Zilman, 2017).

The nuclear lamina, positioned underneath the INM, is comprised of a filamentous network of proteins including A-type and B-type lamins and their respective binding partners (Ho and Lammerding, 2012). The major A-type lamin isoforms are Lamin A and C encoded by the LMNA gene in mammalian somatic cells. Mutations in the LMNA gene lead to various diseases such as muscular dystrophy and cardiomyopathy which are often referred to as laminopathies such as Hutchinson Gilford Progeria Syndrome (HGPS) (Bonne et al., 1999, Fatkin et al., 1999, Genschel and Schmidt, 2000). Intriguingly, mechanically active tissues such as cardiac and skeletal muscles are the most affected by mutations in the LMNA gene. Consequently, cells that lack Lamin A/C or harbor mutations in this gene exhibit defects in nuclear stability and force transmission from cytoplasm to the nucleus (Lee et al., 2007, Zwerger et al., 2013, Robijns et al., 2016).

Furthermore, when mechanically stimulated, Lamin A/C-deficient/mutated cells do not activate genes important in mechanoresponse (Bertrand et al., 2014). This evidence points to the role of the nucleus and in particular Lamin A/C in mechanotransduction.
In fact, it has been shown that forces from the cytoskeleton are transmitted to the nuclear envelope through the Linker of Nucleoskeleton and Cytoskleton (LINC) complex (Alam et al., 2015) (Figure 1). The LINC complex is comprised of nesprins located in the ONM whose C-terminal KASH (Klarsicht, ANC-1, Syne Homology) domain interacts with SUN

domain proteins located in the INM. SUN proteins have been shown to bind to the nuclear lamina, NPCs and chromatin while nesprins interact with all major cytoskeletal filaments (Crisp et al., 2006, Padmakumar et al., 2005, Ketema et al., 2007, Warren et al., 2005). The interaction between nesprins and SUN proteins is thought to mediate the connection between the cytoplasm and nucleoplasm (Haque et al., 2006). An intact LINC complex is required for the effective force transmission to the nucleus and consequently depletion of nesprin or SUN proteins impairs force transmission and expression of mechanosensitive genes (Lombardi et al., 2011, Banerjee et al., 2014, Tajik et al., 2016).

More recently, besides gene expression changes due to mechanotransduction, chromatin has also been shown to be involved in mechanosensing and modulating nuclear stiffness in response to mechanical stress (Miroshnikova et al., 2017, Stephens et al., 2017). One mechanism by which force transmission into the nucleus can induce gene expression changes is by modulation of chromatin organization. The 2-meter chromatin fiber is non randomly packaged inside the nucleus and influences DNA replication, repair, and gene expression (Marchal et al., 2019, Krijger and de Laat, 2016, Spielmann et al., 2018). Post translational modification of the chromatin fiber by different histone marks informs two major chromatin states: heterochromatin and euchromatin (van Steensel and Belmont, 2017, Strom et al., 2017, Larson et al., 2017). Heterochromatic regions have been shown to be tethered to the nuclear lamina and the genes found within these regions exhibit transcriptionally inactive signatures (van Steensel and Belmont, 2017). Meanwhile, euchromatic regions have been shown to be internalized in the nucleus and genes within

these domains exhibit active transcriptional states (Gaspar-Maia et al., 2011). Changes in chromatin state have been shown to alter the physical properties of the nucleus, where increasing levels of heterochromatin decreases nuclear deformability while increasing levels of euchromatin by decondensing agents increases the deformability of the nucleus (Stephens et al., 2017, Stephens et al., 2018).

Mounting evidence also points to the effect that extracellular physical forces have on nuclear mechanics and chromosome structure. Prolonged force application leads to an increase in heterochromatin and subsequently transcription repression (Le et al., 2016). Furthermore, alteration of cytoskeletal organization by culturing cells in micropatterned substrates alters nuclear shape and chromosome distribution accompanied by gene expression changes (Ramdas and Shivashankar, 2015, Wang et al., 2017). Intriguingly, application of force on the cell surface is sufficient to induce chromatin stretching and expression of transgenes in the stretched region, while LINC complex disruption abolishes this effect (Tajik et al., 2016). In addition, nuclei aspirated into narrow micropipettes exhibit stretching of chromatin domains that is sometimes irreversible (Irianto et al., 2017b). Despite such connections between chromatin structure, nucleus deformation, and nucleus mechanics, it is not fully understood how the 3D organization of the genome is affected by such mechanical perturbations.

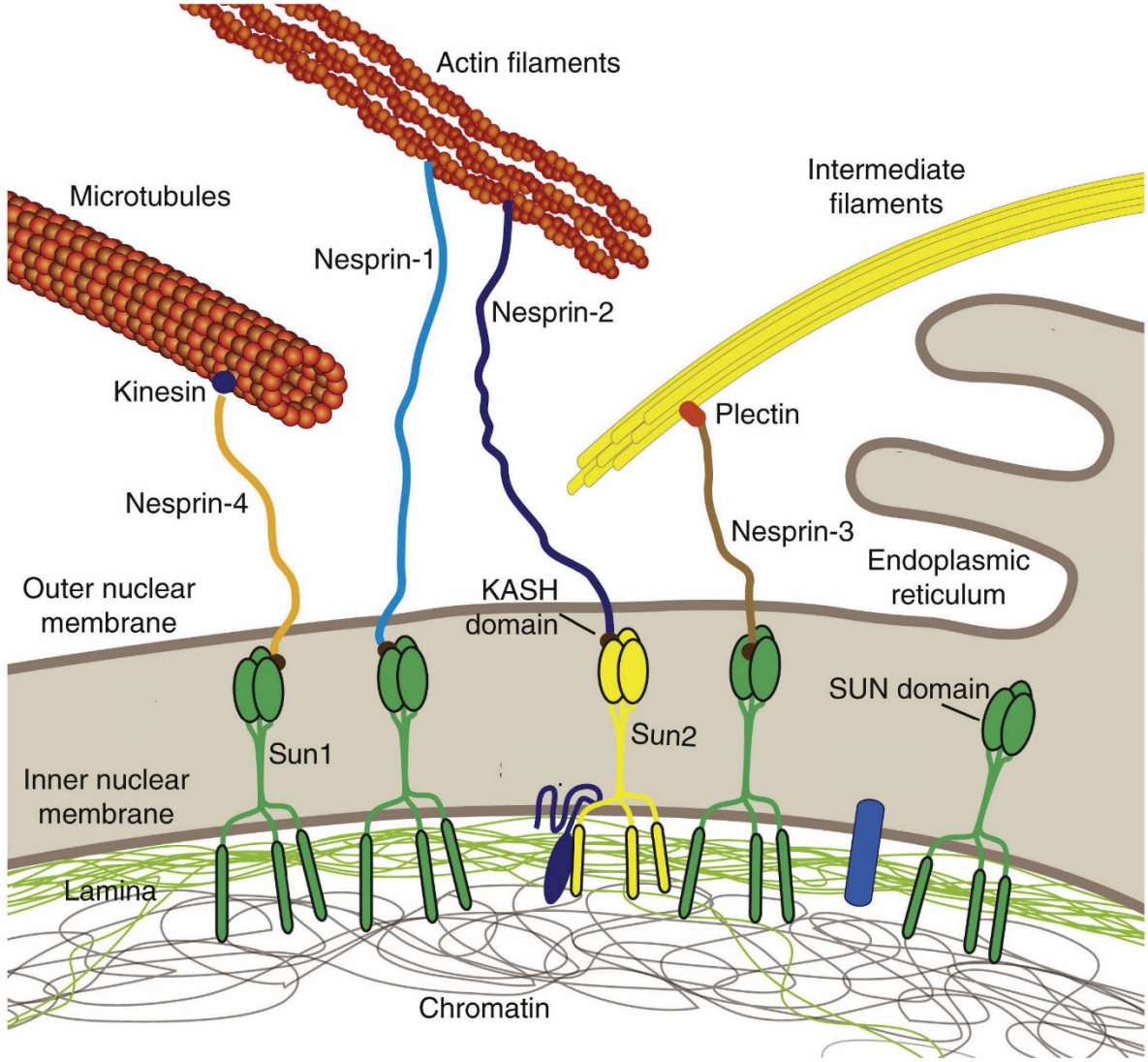


Figure 1. Interactions between major filaments of the cytoskeleton with the nucleus through the LINC complex proteins.

(Reprinted from (Isermann and Lammerding, 2013). Copyright 2013, with permission from Elsevier.)

The 3D genome is non-randomly organized in the nucleus

The eukaryotic DNA measures ~2 meters in length and it must tightly package in a nucleus ranging between 5-10μm in diameter. To achieve this efficient packing, 147 base pairs of DNA wrap around four core histone proteins which include H2A, H2B, H3 and H4 (Kornberg, 1974). Additionally, further compaction of DNA is achieved from the linker histone protein, H1, and its variant H5, which lead to the first layer of higher order chromatin organization (Laybourn and Kadonaga, 1991, Cutter and Hayes, 2017). Protruding N-terminal tails of histones are subject to post translational modifications which are important regulators of chromatin dynamics and gene expression (Campos and Reinberg, 2009). These biological mechanisms that regulate gene expression through structural changes of chromatin are one type of epigenetics. The epigenetic state of the chromatin fiber is highly plastic in part due to the contribution of chromatin modifying enzymes classified into “writers”, “readers” and “erasers” (Biswas and Rao, 2018). Epigenetic writers are made up of a group of enzymes that add modifications to the histone tails or DNA, while readers translate these modifications into a transcriptional output. Finally, erasers (histone demethylases and acetylases) remove the modifications from histones and DNA and revert to the original state of the chromatin. The dynamic exchange between these enzymes influences the accessibility of chromatin region for transcriptional regulation and its higher order folding in 3D.

Recent advances in high resolution microscopy imaging and genomic techniques have characterized in higher resolution the building blocks of this 3D organization of chromatin.

Particularly, development of techniques utilizing chromosome conformation capture (3C) (Dekker et al., 2002) have revealed a non-random organization of the 3D chromosome structure. Integrating chromosome conformation capture with high-throughput sequencing (Hi-C) using crosslinking and proximity ligation, captures the physical interaction frequency between two genomic regions (Golloshi et al., 2018, Lieberman-Aiden et al., 2009). Information generated from Fluorescence in-situ Hybridization (FISH) and Hi-C methods have revealed that chromosomes maintain their spatial territories, which are characterized by their association with components of the nucleus such as nucleoli or nuclear lamina (Cremer and Cremer, 2001) (Figure 2). Association with the nuclear lamina or post-translational histone modifications of the chromatin fiber inform the segregation of the chromatin into active and inactive regions which we refer to as A and B compartments, respectively. This compartmentalization is a hallmark of cell type identity (Schmitt et al., 2016), informs replication timing (Marchal et al., 2019), and changes coordinately with gene expression in cancer and other diseases (Barutcu et al., 2015, McCord et al., 2013). Maintaining proper expression of the genes within each compartment is mediated not only by the histone modification state and associated reader proteins, but also by proteins such as CCCTC-binding factor (CTCF) and cohesin which have been shown to facilitate interactions between promoters and enhancers through loop extrusion. Proper spatial organization of these regulatory regions is accomplished through structures called Topologically Associating Domains (TADs) (Fudenberg et al., 2016, Nora et al., 2012, Nora et al., 2017, Schwarzer et al., 2017). Segregation of TADs

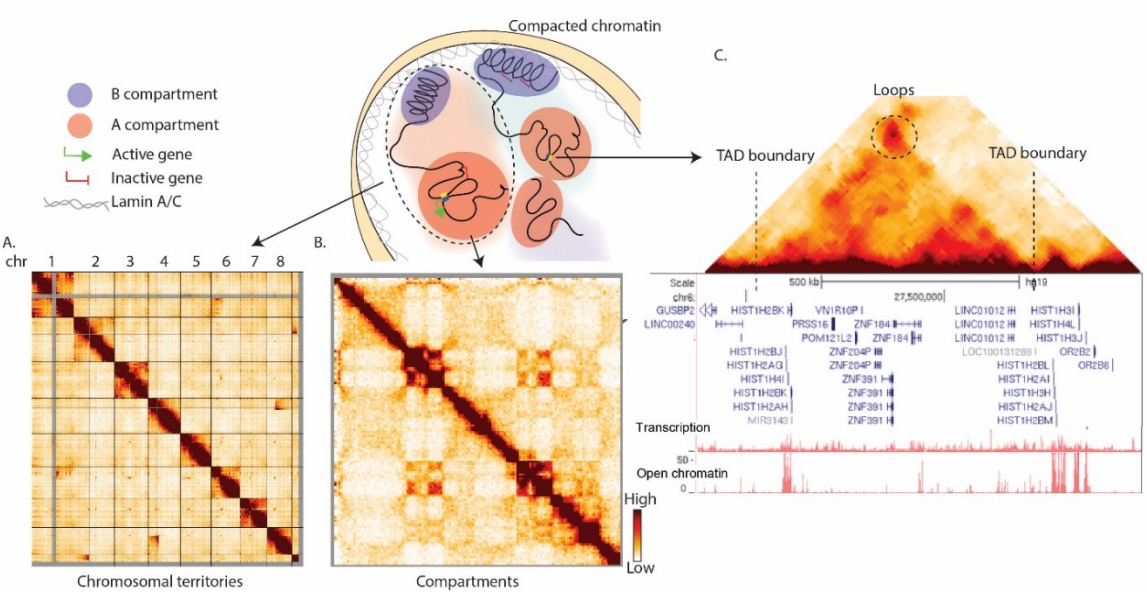


Figure 2. 3D organization of the genome inside the nucleus as revealed by Hi-C.

(A) Genome wide contact frequency maps binned at 2.5 Mb reveal high interaction frequency within each chromosome. (B) Hi-C contact frequency heatmap binned at 250 kb reveals spatial segregation of compartments (plaid pattern). (C) Hi-C contact frequency map representing TADs and regulatory loops which associate with transcriptional regulation of specific genes found within TADs and the compaction state of the chromatin.

is regulated by binding of boundary elements such as CTCF and is critical to maintain appropriate gene expression patterns. Consequently, TAD boundary deletion leads to human limb malformations (Lupianez et al., 2015). The 3D structure is important during cell division and it is subject to constant changes (Golloshi et al., 2017) and therefore maintenance of this organization is important for proper gene regulation and its disruption is shown to be correlated with diseases including cancer progression (Hnisz et al., 2016, Valton and Dekker, 2016).

Disruption of nuclear integrity and 3D genome organization in cancer

To metastasize to secondary sites, cancer cells disseminate from the primary tumor, invade the surrounding tissue, intravasate through blood vessels and colonize to distant organs (Lambert et al., 2017, Gupta and Massague, 2006, Fidler, 2003, Chambers et al., 2002). Migrating and invading through these environments often poses a spatial barrier where cells must undergo dramatic squeezing and deformation to successfully metastasize.

Measurements of tumor environments in vivo have indicated that the pore sizes cancer cells have to pass through range between ~2-3μm in diameter which is representative of only ~ 20% of the nuclear diameter of cancer cells (Harada et al., 2014). While nuclear atypia has been used for decades to predict metastatic grade of cancer (Kadota et al., 2012), the contribution of constricted migration during cancer metastasis to the observed changes in nuclear phenotypes remain elusive.

A major contributing factor during this constricted migration is the stiffness of the nucleus which is partially influenced by Lamin A/C expression as described previously (Harada et al., 2014, Swift et al., 2013). Decreased expression of Lamin A/C has been evident in mouse embryonic fibroblasts (MEF) and is associated with their high migration efficiency through constrictions (Davidson et al., 2014). In contrary, migration efficiency of neutrophils is impaired when Lamin A/C expression is elevated (Rowat et al., 2013). Moreover, aberrant expression of Lamin A/C is also evident in cancer cells. Low levels of Lamin A/C correlate with an increase in invasiveness of some cancer cells (Wazir et al., 2013, Broers et al., 1993). However, colorectal and skin basal cell carcinoma exhibit high expression of Lamin A/C and its downregulation decrease migration speed (Venables et al., 2001). While increase in Lamin A/C expression leads to increase in nucleus stiffness, decrease in its expression threatens the integrity of the nuclear envelope and increases nuclear rupture events (Denais et al., 2016, Irianto et al., 2017a, Raab et al., 2016). In fact, nuclear envelope rupture of cancer cells which is also associated with formation of nuclear blebs have been observed during constricted migration through microfluidic or transwell migration devices. During these rupture events, nuclear proteins mislocalize to the cytoplasm while cytoplasmic components enter the nucleus (Xia et al., 2018). cGas, a cyclic guanosine monophosphate-adenosine monophosphate, is one of the cytoplasmic proteins shown to bind to ruptured sites in the nucleus. Entrance of this protein in the nucleus triggers inflammatory response through the STING pathway as observed by RNA expression assays (Chen et al., 2016).

In addition to Lamin A/C aberrations, chromatin integrity has also been shown to be affected during cancer progression. While cancer has been long thought as a genetic disease arising from gene mutations, emerging evidence suggests that the mutational signatures of cancer are non-randomly distributed and are influenced by chromatin structure, transcriptional and replication timing activity (De and Michor, 2011b, De and Michor, 2011a, Schuster-Bockler and Lehner, 2012, Perera et al., 2016, Pedersen and De, 2013, Barutcu et al., 2015, Taberlay et al., 2016, Zhou et al., 2019).

As previously stated, the 3D organization of the genome and its location with respect to the nuclear lamina, plays an important role in DNA replication and response to DNA damage and repair (Marchal et al., 2019, Krijger and de Laat, 2016, Spielmann et al., 2018). It has been previously shown that translocation frequencies and the downstream effects of these translocations are influenced by the 3D genome structure and the positioning of the chromosomes in the nucleus (Zhang et al., 2012, Hnisz et al., 2016). Furthermore, it has been previously reported that association with the nuclear lamina contributes to variation in mutation rates (Ananda et al., 2011). Interestingly, genomic regions that associate with the nuclear lamina use alternative end-joining mechanisms to repair DNA double-stranded breaks which is prone to more errors compared to homologous-recombination (HR) mediated repair (Lemaitre et al., 2014). As a result, chromosome 18, located more frequently in the nuclear periphery and chromatin regions associated with the lamina, exhibits higher mutation rates than regions of the genome located in the interior of the nucleus across 366 cancer genomes analyzed representative

of 6 different cancer types (Smith et al., 2017). Additionally, TAD boundaries have been reported to be disrupted in cancer, which causes disruption of chromatin loops leading to aberrant expression of genes located within these insulated domains (Weischenfeldt et al., 2017, Kaiser et al., 2016).

Taken together, this evidence suggests a complex role for Lamin A/C and its interactions and the influence of the 3D organization of the genome in the progression of cancer.

Targeting the epigenetic state of cancer as a therapeutic alternative to hinder cancer progression

While the chromatin modifier enzymes are important to maintain regulation of gene expression under normal conditions, misregulation of their expression and activity are involved in many diseases including cancer. Many cancers exhibit aberrant DNA methylation and aberrant expression of histone modifier enzymes such as demethylases (DNMT) and histone deacetylases (HDACs) (Berdasco and Esteller, 2010). In fact, HDACs have been found to be overexpressed in many cancers which could lead to inappropriate expression of oncogenes or tumor suppressors, a signature of cancer. Compared to DNA mutations, these epigenetic alterations are more reversible and dynamic and therefore have become an attractive target for cancer therapy.

Four classes of HDACs have been reported, including Class I (comprised of HDAC 1, 2 ,3 and 8), Class II (comprised of HDAC 4, 5, 6,7, 9 and 10), Class III (comprised of seven mammalian sirtuins SIRT1-7) and Class IV (comprised of HDAC 11) (de Ruijter et al.,

2003). Apart from regulating histone modification, HDACs also regulate the post-translational acetylation status of many non-histone proteins, including transcription factors, chaperones, and signaling molecules, resulting in changes in protein stability, protein-protein interactions, and protein-DNA interactions (Glozak et al., 2005). For example, maintaining the acetylation status of RUNX3, a tumor suppressor and transcription factor, is important for its stability and transcriptional activity. Furthermore, deacetylation of HSP90 by HDAC6 is essential for the stability and function of other proteins such as Bcr-Abl, c-Raf, and AKT (Bali et al., 2005). Various HDAC inhibitors (HDACi) have been developed and described to induce cell cycle arrest, differentiation, and apoptosis in cell lines (Mariadason, 2008, Litvak et al., 1998). Most of these have potent antitumor activities in vivo and in vitro (Marks et al., 2000, Marks et al., 2001). However, it is not clear how inhibition of these chromatin remodelers affects the 3D organization of the genome.

This dissertation aims to address the effect of nuclear perturbations, due to mechanical stressors or epigenetic perturbation, on the 3D organization of the genome. To study how the 3D genome is organized within the nucleus, chromatin conformation capture followed by high-throughput sequencing (Hi-C) will be primarily used throughout this dissertation (Chapter 1). The first project addresses artifacts that can occur during Hi-C library preparation and important precautions to increase the yield and complexity of the Hi-C libraries (Chapter 1). The next project explores the 3D genomic changes that are associated with ability of melanoma cells to withstand mechanical forces induced by

constricted migration (Chapter 2). The following project then investigates the heterogeneity of 3D genome structure within melanoma cells and how it influences ability to undergo constricted migration (Chapter 3). Chapter 4 characterizes 3D genome structural changes that occurs upon treatment with cancer epigenetic therapeutics (Trichostatin A, TSA). The final project (Chapter 5) examines the effect of mechanical stretching in the 3D organization of the genome. Taken together, these projects outline the role of 3D genome organization in response to nucleus perturbations.

References

ALAM, S. G., LOVETT, D., KIM, D. I., ROUX, K. J., DICKINSON, R. B. & LELE, T. P. 2015. The nucleus is an intracellular propagator of tensile forces in NIH 3T3 fibroblasts. *J Cell Sci,* 128**,** 1901-11.

ANANDA, G., CHIAROMONTE, F. & MAKOVA, K. D. 2011. A genome-wide view of mutation rate co-variation using multivariate analyses. *Genome Biol,* 12**,** R27.

BALI, P., PRANPAT, M., BRADNER, J., BALASIS, M., FISKUS, W., GUO, F., ROCHA, K., KUMARASWAMY, S., BOYAPALLE, S., ATADJA, P., SETO, E. & BHALLA, K. 2005. Inhibition of histone deacetylase 6 acetylates and disrupts the chaperone function of heat shock protein 90: a novel basis for antileukemia activity of histone deacetylase inhibitors. *J Biol Chem,* 280**,** 26729-34.

BANERJEE, I., ZHANG, J., MOORE-MORRIS, T., PFEIFFER, E., BUCHHOLZ, K. S., LIU, A., OUYANG, K., STROUD, M. J., GERACE, L., EVANS, S. M., MCCULLOCH, A. & CHEN, J. 2014. Targeted ablation of nesprin 1 and nesprin 2 from murine myocardium results in cardiomyopathy, altered nuclear morphology and inhibition of the biomechanical gene response. *PLoS Genet,* 10**,** e1004114.

BARNES, J. M., PRZYBYLA, L. & WEAVER, V. M. 2017. Tissue mechanics regulate brain development, homeostasis and disease. *J Cell Sci,* 130**,** 71-82.

BARUTCU, A. R., LAJOIE, B. R., MCCORD, R. P., TYE, C. E., HONG, D., MESSIER, T. L., BROWNE, G., VAN WIJNEN, A. J., LIAN, J. B., STEIN, J. L., DEKKER, J., IMBALZANO, A. N. & STEIN, G. S. 2015. Chromatin interaction analysis reveals

changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol,* 16**,** 214.

BERDASCO, M. & ESTELLER, M. 2010. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell,* 19**,** 698-711.

BERTRAND, A. T., ZIAEI, S., EHRET, C., DUCHEMIN, H., MAMCHAOUI, K., BIGOT, A., MAYER, M., QUIJANO-ROY, S., DESGUERRE, I., LAINE, J., BEN YAOU, R., BONNE, G. & COIRAULT, C. 2014. Cellular microenvironments reveal defective mechanosensing responses and elevated YAP signaling in LMNA-mutated muscle precursors. *J Cell Sci,* 127**,** 2873-84.

BISSELL, M. J., HALL, H. G. & PARRY, G. 1982. How does the extracellular matrix direct gene expression? *J Theor Biol,* 99**,** 31-68.

BISWAS, S. & RAO, C. M. 2018. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur J Pharmacol,* 837**,** 8-24.

BONNE, G., DI BARLETTA, M. R., VARNOUS, S., BECANE, H. M., HAMMOUDA, E. H., MERLINI, L., MUNTONI, F., GREENBERG, C. R., GARY, F., URTIZBEREA, J. A., DUBOC, D., FARDEAU, M., TONIOLO, D. & SCHWARTZ, K. 1999. Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat Genet,* 21**,** 285-8.

BROERS, J. L., RAYMOND, Y., ROT, M. K., KUIJPERS, H., WAGENAAR, S. S. & RAMAEKERS, F. C. 1993. Nuclear A-type lamins are differentially expressed in human lung cancer subtypes. *Am J Pathol,* 143**,** 211-20.

CAILLE, N., THOUMINE, O., TARDY, Y. & MEISTER, J. J. 2002. Contribution of the nucleus to the mechanical properties of endothelial cells. *J Biomech,* 35**,** 177-87.

CAMPOS, E. I. & REINBERG, D. 2009. Histones: annotating chromatin. *Annu Rev Genet,* 43**,** 559-99.

CHAMBERS, A. F., GROOM, A. C. & MACDONALD, I. C. 2002. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer,* 2**,** 563-72.

CHEN, Q., SUN, L. & CHEN, Z. J. 2016. Regulation and function of the cGAS-STING pathway of cytosolic DNA sensing. *Nat Immunol,* 17**,** 1142-9.

CREMER, T. & CREMER, C. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet,* 2**,** 292-301.

CRISP, M., LIU, Q., ROUX, K., RATTNER, J. B., SHANAHAN, C., BURKE, B., STAHL, P. D. & HODZIC, D. 2006. Coupling of the nucleus and cytoplasm: role of the LINC complex. *J Cell Biol,* 172**,** 41-53.

CUTTER, A. R. & HAYES, J. J. 2017. Linker histones: novel insights into structure-specific recognition of the nucleosome. *Biochem Cell Biol,* 95**,** 171-178.

DAVIDSON, P. M., DENAIS, C., BAKSHI, M. C. & LAMMERDING, J. 2014. Nuclear deformability constitutes a rate-limiting step during cell migration in 3-D environments. *Cell Mol Bioeng,* 7**,** 293-306.

DE RUIJTER, A. J., VAN GENNIP, A. H., CARON, H. N., KEMP, S. & VAN KUILENBURG, A. B. 2003. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J,* 370**,** 737-49.

DE, S. & MICHOR, F. 2011a. DNA replication timing and long-range DNA interactions predict mutational landscapes of cancer genomes. *Nat Biotechnol,* 29**,** 1103-8.

DE, S. & MICHOR, F. 2011b. DNA secondary structures and epigenetic determinants of cancer genome evolution. *Nat Struct Mol Biol,* 18**,** 950-5.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

DENAIS, C. M., GILBERT, R. M., ISERMANN, P., MCGREGOR, A. L., TE LINDERT, M., WEIGELIN, B., DAVIDSON, P. M., FRIEDL, P., WOLF, K. & LAMMERDING, J. 2016. Nuclear envelope rupture and repair during cancer cell migration. *Science,* 352**,** 353-8.

DEWEY, C. F., JR., BUSSOLARI, S. R., GIMBRONE, M. A., JR. & DAVIES, P. F. 1981. The dynamic response of vascular endothelial cells to fluid shear stress. *J Biomech Eng,* 103**,** 177-85.

DREGER, M., BENGTSSON, L., SCHONEBERG, T., OTTO, H. & HUCHO, F. 2001. Nuclear envelope proteomics: novel integral membrane proteins of the inner nuclear membrane. *Proc Natl Acad Sci U S A,* 98**,** 11943-8.

FATKIN, D., MACRAE, C., SASAKI, T., WOLFF, M. R., PORCU, M., FRENNEAUX, M., ATHERTON, J., VIDAILLET, H. J., JR., SPUDICH, S., DE GIROLAMI, U., SEIDMAN, J. G., SEIDMAN, C., MUNTONI, F., MUEHLE, G., JOHNSON, W. & MCDONOUGH, B. 1999. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med,* 341**,** 1715-24.

FIDLER, I. J. 2003. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer,* 3**,** 453-8.

FREEMAN, P. M., NATARAJAN, R. N., KIMURA, J. H. & ANDRIACCHI, T. P. 1994. Chondrocyte cells respond mechanically to compressive loads. *J Orthop Res,* 12**,** 311-20.

FRIEDL, P. & WOLF, K. 2010. Plasticity of cell migration: a multiscale tuning model. *J Cell Biol,* 188**,** 11-9.

FUDENBERG, G., IMAKAEV, M., LU, C., GOLOBORODKO, A., ABDENNUR, N. & MIRNY, L. A. 2016. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep,* 15**,** 2038-49.

GASPAR-MAIA, A., ALAJEM, A., MESHORER, E. & RAMALHO-SANTOS, M. 2011. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol,* 12**,** 36-47.

GENSCHEL, J. & SCHMIDT, H. H. 2000. Mutations in the LMNA gene encoding lamin A/C. *Hum Mutat,* 16**,** 451-9.

GLOZAK, M. A., SENGUPTA, N., ZHANG, X. & SETO, E. 2005. Acetylation and deacetylation of non-histone proteins. *Gene,* 363**,** 15-23.

GOLLOSHI, R., SANDERS, J. T. & MCCORD, R. P. 2018. Iteratively improving Hi-C experiments one step at a time. *Methods,* 142**,** 47-58.

GUILLUY, C., OSBORNE, L. D., VAN LANDEGHEM, L., SHAREK, L., SUPERFINE, R., GARCIA-MATA, R. & BURRIDGE, K. 2014. Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. *Nat Cell Biol,* 16**,** 376-81.

GUPTA, G. P. & MASSAGUE, J. 2006. Cancer metastasis: building a framework. *Cell,* 127**,** 679-95.

HAQUE, F., LLOYD, D. J., SMALLWOOD, D. T., DENT, C. L., SHANAHAN, C. M., FRY, A. M., TREMBATH, R. C. & SHACKLETON, S. 2006. SUN1 interacts with nuclear lamin A and cytoplasmic nesprins to provide a physical connection between the nuclear lamina and the cytoskeleton. *Mol Cell Biol,* 26**,** 3738-51.

HARADA, T., SWIFT, J., IRIANTO, J., SHIN, J. W., SPINLER, K. R., ATHIRASALA, A., DIEGMILLER, R., DINGAL, P. C., IVANOVSKA, I. L. & DISCHER, D. E. 2014. Nuclear lamin stiffness is a barrier to 3D migration, but softness can limit survival. *J Cell Biol,* 204**,** 669-82.

HNISZ, D., WEINTRAUB, A. S., DAY, D. S., VALTON, A. L., BAK, R. O., LI, C. H., GOLDMANN, J., LAJOIE, B. R., FAN, Z. P., SIGOVA, A. A., REDDY, J., BORGES-RIVERA, D., LEE, T. I., JAENISCH, R., PORTEUS, M. H., DEKKER, J. & YOUNG, R. A. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science,* 351**,** 1454-1458.

HO, C. Y. & LAMMERDING, J. 2012. Lamins at a glance. *J Cell Sci,* 125**,** 2087-93.

IRIANTO, J., XIA, Y., PFEIFER, C. R., ATHIRASALA, A., JI, J., ALVEY, C., TEWARI, M., BENNETT, R. R., HARDING, S. M., LIU, A. J., GREENBERG, R. A. & DISCHER, D. E. 2017a. DNA Damage Follows Repair Factor Depletion and Portends Genome Variation in Cancer Cells after Pore Migration. *Curr Biol,* 27**,** 210-223.

IRIANTO, J., XIA, Y., PFEIFER, C. R., GREENBERG, R. A. & DISCHER, D. E. 2017b. As a Nucleus Enters a Small Pore, Chromatin Stretches and Maintains Integrity, Even with DNA Breaks. *Biophys J,* 112**,** 446-449.

ISERMANN, P. & LAMMERDING, J. 2013. Nuclear mechanics and mechanotransduction in health and disease. Curr Biol, 23, R1113-21.

JAALOUK, D. E. & LAMMERDING, J. 2009. Mechanotransduction gone awry. *Nat Rev Mol Cell Biol,* 10**,** 63-73.

JOVANOVIC-TALISMAN, T. & ZILMAN, A. 2017. Protein Transport by the Nuclear Pore Complex: Simple Biophysics of a Complex Biomachine. *Biophys J,* 113**,** 6-14.

KADOTA, K., SUZUKI, K., COLOVOS, C., SIMA, C. S., RUSCH, V. W., TRAVIS, W. D. & ADUSUMILLI, P. S. 2012. A nuclear grading system is a strong predictor of survival in epitheloid diffuse malignant pleural mesothelioma. *Mod Pathol,* 25**,** 260-71.

KAISER, V. B., TAYLOR, M. S. & SEMPLE, C. A. 2016. Mutational Biases Drive Elevated Rates of Substitution at Regulatory Sites across Cancer Types. *PLoS Genet,* 12**,** e1006207.

KETEMA, M., WILHELMSEN, K., KUIKMAN, I., JANSSEN, H., HODZIC, D. & SONNENBERG, A. 2007. Requirements for the localization of nesprin-3 at the nuclear envelope and its interaction with plectin. *J Cell Sci,* 120**,** 3384-94.

KIRBY, T. J. & LAMMERDING, J. 2018. Emerging views of the nucleus as a cellular mechanosensor. *Nat Cell Biol,* 20**,** 373-381.

KORNBERG, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science,* 184**,** 868-71.

KRIJGER, P. H. & DE LAAT, W. 2016. Regulation of disease-associated gene expression in the 3D genome. *Nat Rev Mol Cell Biol,* 17**,** 771-782.

LAMBERT, A. W., PATTABIRAMAN, D. R. & WEINBERG, R. A. 2017. Emerging Biological Principles of Metastasis. *Cell,* 168**,** 670-691.

LAMMERDING, J. 2011. Mechanics of the nucleus. *Compr Physiol,* 1**,** 783-807.

LANGE, J. R. & FABRY, B. 2013. Cell and tissue mechanics in cell migration. *Exp Cell Res,* 319**,** 2418-23.

LARSON, A. G., ELNATAN, D., KEENEN, M. M., TRNKA, M. J., JOHNSTON, J. B., BURLINGAME, A. L., AGARD, D. A., REDDING, S. & NARLIKAR, G. J. 2017. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature,* 547**,** 236-240.

LAYBOURN, P. J. & KADONAGA, J. T. 1991. Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. *Science,* 254**,** 238-45.

LE, H. Q., GHATAK, S., YEUNG, C. Y., TELLKAMP, F., GUNSCHMANN, C., DIETERICH, C., YEROSLAVIZ, A., HABERMANN, B., POMBO, A., NIESSEN, C. M. & WICKSTROM, S. A. 2016. Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nat Cell Biol,* 18**,** 864-75.

LEE, J. S., HALE, C. M., PANORCHAN, P., KHATAU, S. B., GEORGE, J. P., TSENG, Y., STEWART, C. L., HODZIC, D. & WIRTZ, D. 2007. Nuclear lamin A/C deficiency

induces defects in cell mechanics, polarization, and migration. *Biophys J,* 93**,** 2542-52.

LEMAITRE, C., GRABARZ, A., TSOUROULA, K., ANDRONOV, L., FURST, A., PANKOTAI, T., HEYER, V., ROGIER, M., ATTWOOD, K. M., KESSLER, P., DELLAIRE, G., KLAHOLZ, B., REINA-SAN-MARTIN, B. & SOUTOGLOU, E. 2014. Nuclear position dictates DNA repair pathway choice. *Genes Dev,* 28**,** 2450-63.

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., SANDSTROM, R., BERNSTEIN, B., BENDER, M. A., GROUDINE, M., GNIRKE, A., STAMATOYANNOPOULOS, J., MIRNY, L. A., LANDER, E. S. & DEKKER, J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science,* 326**,** 289-93.

LITVAK, D. A., EVERS, B. M., HWANG, K. O., HELLMICH, M. R., KO, T. C. & TOWNSEND, C. M., JR. 1998. Butyrate-induced differentiation of Caco-2 cells is associated with apoptosis and early induction of p21Waf1/Cip1 and p27Kip1. *Surgery,* 124**,** 161-9; discussion 169-70.

LOMBARDI, M. L., JAALOUK, D. E., SHANAHAN, C. M., BURKE, B., ROUX, K. J. & LAMMERDING, J. 2011. The interaction between nesprins and sun proteins at the nuclear envelope is critical for force transmission between the nucleus and cytoskeleton. *J Biol Chem,* 286**,** 26743-53.

LUPIANEZ, D. G., KRAFT, K., HEINRICH, V., KRAWITZ, P., BRANCATI, F., KLOPOCKI, E., HORN, D., KAYSERILI, H., OPITZ, J. M., LAXOVA, R., SANTOS-SIMARRO,

F., GILBERT-DUSSARDIER, B., WITTLER, L., BORSCHIWER, M., HAAS, S. A., OSTERWALDER, M., FRANKE, M., TIMMERMANN, B., HECHT, J., SPIELMANN, M., VISEL, A. & MUNDLOS, S. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell,* 161**,** 1012-1025.

MAGID, A. & LAW, D. J. 1985. Myofibrils bear most of the resting tension in frog skeletal muscle. *Science,* 230**,** 1280-2.

MARCHAL, C., SIMA, J. & GILBERT, D. M. 2019. Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol*.

MARIADASON, J. M. 2008. HDACs and HDAC inhibitors in colon cancer. *Epigenetics,* 3**,** 28-37.

MARKS, P., RIFKIND, R. A., RICHON, V. M., BRESLOW, R., MILLER, T. & KELLY, W. K. 2001. Histone deacetylases and cancer: causes and therapies. *Nat Rev Cancer,* 1**,** 194-202.

MARKS, P. A., RICHON, V. M. & RIFKIND, R. A. 2000. Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. *J Natl Cancer Inst,* 92**,** 1210-6.

MCCORD, R. P., NAZARIO-TOOLE, A., ZHANG, H., CHINES, P. S., ZHAN, Y., ERDOS, M. R., COLLINS, F. S., DEKKER, J. & CAO, K. 2013. Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res,* 23**,** 260-9.

MIROSHNIKOVA, Y. A., NAVA, M. M. & WICKSTROM, S. A. 2017. Emerging roles of mechanical forces in chromatin regulation. *J Cell Sci,* 130**,** 2243-2250.

MURTHY, S. E., DUBIN, A. E. & PATAPOUTIAN, A. 2017. Piezos thrive under pressure: mechanically activated ion channels in health and disease. *Nat Rev Mol Cell Biol,* 18**,** 771-783.

NORA, E. P., GOLOBORODKO, A., VALTON, A. L., GIBCUS, J. H., UEBERSOHN, A., ABDENNUR, N., DEKKER, J., MIRNY, L. A. & BRUNEAU, B. G. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell,* 169**,** 930-944 e22.

NORA, E. P., LAJOIE, B. R., SCHULZ, E. G., GIORGETTI, L., OKAMOTO, I., SERVANT, N., PIOLOT, T., VAN BERKUM, N. L., MEISIG, J., SEDAT, J., GRIBNAU, J., BARILLOT, E., BLUTHGEN, N., DEKKER, J. & HEARD, E. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature,* 485**,** 381-5.

PADMAKUMAR, V. C., LIBOTTE, T., LU, W., ZAIM, H., ABRAHAM, S., NOEGEL, A. A., GOTZMANN, J., FOISNER, R. & KARAKESISOGLOU, I. 2005. The inner nuclear membrane protein Sun1 mediates the anchorage of Nesprin-2 to the nuclear envelope. *J Cell Sci,* 118**,** 3419-30.

PANCIERA, T., AZZOLIN, L., CORDENONSI, M. & PICCOLO, S. 2017. Mechanobiology of YAP and TAZ in physiology and disease. *Nat Rev Mol Cell Biol,* 18**,** 758-770.

PASZEK, M. J. & WEAVER, V. M. 2004. The tension mounts: mechanics meets morphogenesis and malignancy. *J Mammary Gland Biol Neoplasia,* 9**,** 325-42.

PASZEK, M. J., ZAHIR, N., JOHNSON, K. R., LAKINS, J. N., ROZENBERG, G. I., GEFEN, A., REINHART-KING, C. A., MARGULIES, S. S., DEMBO, M., BOETTIGER, D., HAMMER, D. A. & WEAVER, V. M. 2005. Tensional homeostasis and the malignant phenotype. *Cancer Cell,* 8**,** 241-54.

PEDERSEN, B. S. & DE, S. 2013. Loss of heterozygosity preferentially occurs in early replicating regions in cancer genomes. *Nucleic Acids Res,* 41**,** 7615-24.

PERERA, D., POULOS, R. C., SHAH, A., BECK, D., PIMANDA, J. E. & WONG, J. W. 2016. Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature,* 532**,** 259-63.

RAAB, M., GENTILI, M., DE BELLY, H., THIAM, H. R., VARGAS, P., JIMENEZ, A. J., LAUTENSCHLAEGER, F., VOITURIEZ, R., LENNON-DUMENIL, A. M., MANEL, N. & PIEL, M. 2016. ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. *Science,* 352**,** 359-62.

RAMDAS, N. M. & SHIVASHANKAR, G. V. 2015. Cytoskeletal control of nuclear morphology and chromatin organization. *J Mol Biol,* 427**,** 695-706.

REIDY, M. A. & BOWYER, D. E. 1977. Scanning electron microscopy of arteries. The morphology of aortic endothelium in haemodynamically stressed areas associated with branches. *Atherosclerosis,* 26**,** 181-94.

ROBIJNS, J., MOLENBERGHS, F., SIEPRATH, T., CORNE, T. D., VERSCHUUREN, M. & DE VOS, W. H. 2016. In silico synchronization reveals regulators of nuclear ruptures in lamin A/C deficient model cells. *Sci Rep,* 6**,** 30325.

ROWAT, A. C., JAALOUK, D. E., ZWERGER, M., UNG, W. L., EYDELNANT, I. A., OLINS, D. E., OLINS, A. L., HERRMANN, H., WEITZ, D. A. & LAMMERDING, J. 2013. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. *J Biol Chem,* 288**,** 8610-8.

SCHMITT, A. D., HU, M., JUNG, I., XU, Z., QIU, Y., TAN, C. L., LI, Y., LIN, S., LIN, Y., BARR, C. L. & REN, B. 2016. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep,* 17**,** 2042-2059.

SCHUSTER-BOCKLER, B. & LEHNER, B. 2012. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature,* 488**,** 504-7.

SCHWARZER, W., ABDENNUR, N., GOLOBORODKO, A., PEKOWSKA, A., FUDENBERG, G., LOE-MIE, Y., FONSECA, N. A., HUBER, W., HAERING, C. H., MIRNY, L. & SPITZ, F. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature,* 551**,** 51-56.

SMITH, K. S., LIU, L. L., GANESAN, S., MICHOR, F. & DE, S. 2017. Nuclear topology modulates the mutational landscapes of cancer genomes. *Nat Struct Mol Biol,* 24**,** 1000-1006.

SPIELMANN, M., LUPIANEZ, D. G. & MUNDLOS, S. 2018. Structural variation in the 3D genome. *Nat Rev Genet,* 19**,** 453-467.

STEPHENS, A. D., BANIGAN, E. J., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2017. Chromatin and lamin A determine two different mechanical response regimes of the cell nucleus. *Mol Biol Cell,* 28**,** 1984-1996.

STEPHENS, A. D., LIU, P. Z., BANIGAN, E. J., ALMASSALHA, L. M., BACKMAN, V., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2018. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol Biol Cell,* 29**,** 220-233.

STROM, A. R., EMELYANOV, A. V., MIR, M., FYODOROV, D. V., DARZACQ, X. & KARPEN, G. H. 2017. Phase separation drives heterochromatin domain formation. *Nature,* 547**,** 241-245.

SUN, Z., GUO, S. S. & FASSLER, R. 2016. Integrin-mediated mechanotransduction. *J Cell Biol,* 215**,** 445-456.

SWIFT, J., IVANOVSKA, I. L., BUXBOIM, A., HARADA, T., DINGAL, P. C., PINTER, J., PAJEROWSKI, J. D., SPINLER, K. R., SHIN, J. W., TEWARI, M., REHFELDT, F., SPEICHER, D. W. & DISCHER, D. E. 2013. Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. *Science,* 341**,** 1240104.

TABERLAY, P. C., ACHINGER-KAWECKA, J., LUN, A. T., BUSKE, F. A., SABIR, K., GOULD, C. M., ZOTENKO, E., BERT, S. A., GILES, K. A., BAUER, D. C., SMYTH, G. K., STIRZAKER, C., O'DONOGHUE, S. I. & CLARK, S. J. 2016. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res,* 26**,** 719-31.

TAJIK, A., ZHANG, Y., WEI, F., SUN, J., JIA, Q., ZHOU, W., SINGH, R., KHANNA, N., BELMONT, A. S. & WANG, N. 2016. Transcription upregulation via force-induced direct stretching of chromatin. *Nat Mater,* 15**,** 1287-1296.

TUCKER, J. B. & MEATS, M. 1976. Microtubules and control of insect egg shape. *J Cell Biol,* 71**,** 207-17.

VAN STEENSEL, B. & BELMONT, A. S. 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell,* 169**,** 780-791.

VENABLES, R. S., MCLEAN, S., LUNY, D., MOTELEB, E., MORLEY, S., QUINLAN, R. A., LANE, E. B. & HUTCHISON, C. J. 2001. Expression of individual lamins in basal cell carcinomas of the skin. *Br J Cancer,* 84**,** 512-9.

WANG, N., TYTELL, J. D. & INGBER, D. E. 2009. Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. *Nat Rev Mol Cell Biol,* 10**,** 75-82.

WANG, Y., NAGARAJAN, M., UHLER, C. & SHIVASHANKAR, G. V. 2017. Orientation and repositioning of chromosomes correlate with cell geometry-dependent gene expression. *Mol Biol Cell,* 28**,** 1997-2009.

WARREN, D. T., ZHANG, Q., WEISSBERG, P. L. & SHANAHAN, C. M. 2005. Nesprins: intracellular scaffolds that maintain cell architecture and coordinate cell function? *Expert Rev Mol Med,* 7**,** 1-15.

WAZIR, U., AHMED, M. H., BRIDGER, J. M., HARVEY, A., JIANG, W. G., SHARMA, A. K. & MOKBEL, K. 2013. The clinicopathological significance of lamin A/C, lamin B1 and lamin B receptor mRNA expression in human breast cancer. *Cell Mol Biol Lett,* 18**,** 595-611.

WEISCHENFELDT, J., DUBASH, T., DRAINAS, A. P., MARDIN, B. R., CHEN, Y., STUTZ, A. M., WASZAK, S. M., BOSCO, G., HALVORSEN, A. R., RAEDER, B., EFTHYMIOPOULOS, T., ERKEK, S., SIEGL, C., BRENNER, H., BRUSTUGUN, O. T., DIETER, S. M., NORTHCOTT, P. A., PETERSEN, I., PFISTER, S. M., SCHNEIDER, M., SOLBERG, S. K., THUNISSEN, E., WEICHERT, W., ZICHNER, T., THOMAS, R., PEIFER, M., HELLAND, A., BALL, C. R., JECHLINGER, M., SOTILLO, R., GLIMM, H. & KORBEL, J. O. 2017. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat Genet,* 49**,** 65-74.

XIA, Y., IVANOVSKA, I. L., ZHU, K., SMITH, L., IRIANTO, J., PFEIFER, C. R., ALVEY, C. M., JI, J., LIU, D., CHO, S., BENNETT, R. R., LIU, A. J., GREENBERG, R. A. & DISCHER, D. E. 2018. Nuclear rupture at sites of high curvature compromises retention of DNA repair factors. *J Cell Biol,* 217**,** 3796-3808.

ZHANG, Y., MCCORD, R. P., HO, Y. J., LAJOIE, B. R., HILDEBRAND, D. G., SIMON, A. C., BECKER, M. S., ALT, F. W. & DEKKER, J. 2012. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell,* 148**,** 908-21.

ZHOU, Y., GERRARD, D. L., WANG, J., LI, T., YANG, Y., FRITZ, A. J., RAJENDRAN, M., FU, X., SCHIFF, R., LIN, S., FRIETZE, S. & JIN, V. X. 2019. Temporal dynamic reorganization of 3D chromatin architecture in hormone-induced breast cancer and endocrine resistance. *Nat Commun,* 10**,** 1522.

ZWERGER, M., JAALOUK, D. E., LOMBARDI, M. L., ISERMANN, P., MAUERMANN, M., DIALYNAS, G., HERRMANN, H., WALLRATH, L. L. & LAMMERDING, J. 2013. Myopathic lamin mutations impair nuclear stability in cells and tissue and disrupt nucleo-cytoskeletal coupling. *Hum Mol Genet,* 22**,** 2335-49.

GOLLOSHI, R., SANDERS, J. T. & MCCORD, R. P. 2017. Genome organization during the cell cycle: unity in division. *Wiley Interdiscip Rev Syst Biol Med,* 9.

HNISZ, D., WEINTRAUB, A. S., DAY, D. S., VALTON, A. L., BAK, R. O., LI, C. H., GOLDMANN, J., LAJOIE, B. R., FAN, Z. P., SIGOVA, A. A., REDDY, J., BORGES-RIVERA, D., LEE, T. I., JAENISCH, R., PORTEUS, M. H., DEKKER, J. & YOUNG, R. A. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science,* 351**,** 1454-1458.

MARCHAL, C., SIMA, J. & GILBERT, D. M. 2019. Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol*.

VALTON, A. L. & DEKKER, J. 2016. TAD disruption as oncogenic driver. *Curr Opin Genet Dev,* 36**,** 34-40.

CHAPTER I: ITERATIVELY IMPROVING HI-C EXPERIMENTS ONE STEP AT A TIME

A version of this paper was originally published by Rosela Golloshi, Jacob Sanders and Rachel Patton McCord

Golloshi, R., Sanders, J. T. & McCord, R. P. 2018. Iteratively improving Hi-C experiments one step at a time. Methods, 142, 47-58.

My contributions included: (1) performing experiments, (2) data collection and analysis and (3) writing manuscript and making figures.

Abstract

The 3D organization of eukaryotic chromosomes affects key processes such as gene expression, DNA replication, cell division, and response to DNA damage. The genome-wide chromosome conformation capture (Hi-C) approach can characterize the landscape of 3D genome organization by measuring interaction frequencies between all genomic regions. Hi-C protocol improvements and rapid advances in DNA sequencing power have made Hi-C useful to study diverse biological systems, not only to elucidate the role of 3D genome structure in proper cellular function, but also to characterize genomic rearrangements, assemble new genomes, and consider chromatin interactions as potential biomarkers for diseases. Yet, the Hi-C protocol is still complex and subject to variations at numerous steps that can affect the resulting data. Thus, there is still a need for better understanding and control of factors that contribute to Hi-C experiment success and data quality. Here, we evaluate recently proposed Hi-C protocol modifications as well as often overlooked variables in sample preparation and examine their effects on Hi-C

data quality. We examine artifacts that can occur during Hi-C library preparation, including microhomology-based artificial template copying and chimera formation that can add noise to the downstream data. Exploring the mechanisms underlying Hi-C artifacts pinpoints steps that should be further optimized in the future. To improve the utility of Hi-C in characterizing the 3D genome of specialized populations of cells or small samples of primary tissue, we identify steps prone to DNA loss which should be considered to adapt Hi-C to lower cell numbers.

Introduction

The billions of DNA bases in the eukaryotic genome must be efficiently packed to fit inside the micron-sized nucleus and to perform necessary cellular functions including gene expression as well as DNA replication and repair (Dekker and Mirny, 2016, Dileep et al., 2015, Zhang et al., 2012). Abnormal 3D packaging of chromatin may disrupt cellular homeostasis and lead to diseases (Lupianez et al., 2015, McCord et al., 2013).

Many techniques have been devised to understand how chromatin is packed into the nucleus and how that packaging affects genome function. Light and electron microscopy led to observations of densely packed heterochromatin and open euchromatin as well as the small scale wrapping of DNA around nucleosomes (Olins and Olins, 1974, Heitz, 1928). Fluorescence microscopy approaches such as fluorescence in situ hybridization (FISH) (Shachar et al., 2015) and in vivo techniques such as bacterial operator arrays (Robinett et al., 1996) and CRISPR imaging (Ma et al., 2015) have allowed visualization and tracking of the nuclear arrangements of specific genome loci. In recent years, the

development of the chromosome conformation capture (3C) technique (Dekker et al., 2002), and variations such as 4C (Simonis et al., 2006, Zhao et al., 2006), 5C (Dostie et al., 2006), and Hi-C (Lieberman-Aiden et al., 2009), have accelerated the progress of the genome organization field and sparked the interest of researchers with numerous biological problems and across disciplines from genetics to physics, mathematics, and computer science.

As high throughput sequencing becomes increasingly powerful and inexpensive, many researchers have adopted Hi-C as a method to capture a snapshot of the 3D nuclear organization genome-wide. The ability to represent 3D genome organization in terms of matrices of contact frequencies allows the characterization of structures across a variety of length scales, from the positioning of whole chromosome territories to the folding of small-scale enhancer-promoter loops (Zhang et al., 2012, Gibcus and Dekker, 2013, Rao et al., 2014). Even when more focused resolution is desired, many techniques begin with an approach similar to Hi-C and then isolate a subpopulation of interactions of interest. This is the case with approaches such as Capture-C (Hughes et al., 2014, Jager et al., 2015), Hi-ChIP (Mumbach et al., 2016), and T2C (Kolovos et al., 2014). Further, it has recently become clear that Hi-C is a very useful tool not only for identifying 3D genome folding patterns, but also for de novo whole genome sequence assembly (Mascher et al., 2017, Kaplan and Dekker, 2013, Burton et al., 2013) and translocation detection (Harewood et al., 2017). As basic researchers from many fields, commercial ventures (such as Phase Genomics and Dovetail Genomics), and clinical researchers adopt Hi-C

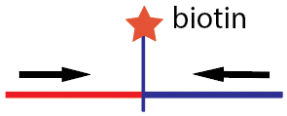
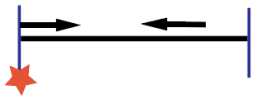
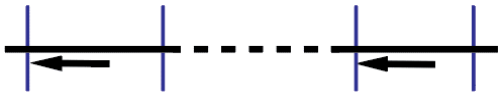
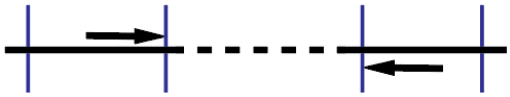
based techniques, it is more important than ever to optimize the methodological details of the Hi-C protocol and understand how they affect the resulting data and interpretation. Several excellent discussions of Hi-C theory and optimizations of Hi-C protocols have been published in recent years (Rao et al., 2014, Belaghzal et al., 2017, Lajoie et al., 2015, Schmitt et al., 2016), and numerous other insights about key steps of the Hi-C protocol have been discovered while researchers optimize this technique in various systems. Yet, Hi-C is still a technique that often requires troubleshooting and can sometimes unexpectedly result in low quality datasets even when the best current protocols are followed. Rather than a comprehensive guide or update to every step of the Hi-C approach, here, we will describe our observations about certain key variables in the Hi-C protocol and the resulting data. We will discuss the effect of certain protocol modifications on Hi-C data quality and utility, propose explanations for previously unreported artifacts that can arise in Hi-C libraries, describe considerations for applying Hi-C to smaller populations of cells, and discuss key areas of uncertainty that will be ripe candidates for future careful optimization. To increase the clarity of our references to sequenced Hi-C interaction molecule types, we have included a terminology definition table (Table 1).

Results

Early steps in the Hi-C protocol with downstream implications

Initial sample preparation: an overlooked variable

Table 1. Definitions of Hi-C molecule types.

	Hi-C Term	Depiction	Definition
Interaction types for all reads	Valid Pair		Reads from different restriction fragments and facing toward restriction site
	Single Side Reads		Only one side of the interacting pair maps to the genome
	Unmapped Reads		Neither side of the interacting pair maps to the genome
	Dangling End		Reads facing inward in the same fragment
	Self Circle		Reads facing outward in the same fragment
Classification of Valid Pairs	Bottom Strand		Reads both on the (-) strand from non-adjacent fragments
	Top Strand		Reads both on the (+) strand from non-adjacent fragments
	Inward		Reads on opposite strands from non-adjacent fragments facing inward
	Outward		Reads on opposite strands from non-adjacent fragments facing outward
	Cis		Interactions within the same chromosome
	Trans		Interactions between different chromosomes

The key initial steps of Hi-C involve the covalent crosslinking of chromatin regions interacting in 3D space, followed by restriction enzyme digestion, marking digested ends with biotin, and proximity ligation of fragments from interacting regions. As we will discuss below, digestion, biotin fill-in, and ligation conditions certainly can have a large impact on Hi-C library quality, but another potential source of variation in Hi-C is the handling of the starting material before restriction enzyme digestion takes place. For cells in culture, variables such as exactly how the formaldehyde is added, how adherent cells are removed from culture dish, and how nuclei are isolated and permeabilized can have downstream implications for Hi-C library quality (Figure 3A). Even if optimized kits such as the Arima Hi-C approach (https://arimagenomics.com/products/) are employed, these early steps still need to be evaluated for new cell types and experimental conditions.

Higher reproducibility and quality of Hi-C datasets occurs when digestion, biotinylation, and ligation steps occur within a permeabilized nucleus that is not lysed into the solution. This type of approach has been variably called the “in situ” (Rao et al., 2014) or “in nucleus” (Nagano et al., 2015) Hi-C protocol. Despite minor variations in exact protocols, the basic idea remains the same: Avoiding high concentrations of SDS, nuclei lysis, and dilution to large volumes during digestion and ligation increases Hi-C reproducibility and decreases the capture of random background interactions (Rao et al., 2014, Nagano et al., 2015). The generally accepted explanation for this improvement has been that intact nuclei constrain the movement and random collisions of crosslinked complexes, but other factors could be at work as well. Other work has suggested that high concentrations of SDS and subsequent Triton sequestration of SDS are more likely to lead to aggregates

A) Hi-C experiments for cells in culture begin with decisions about initial steps such as formaldehyde crosslinking, cell harvest, and cell lysis and permeabilization. B) Different approaches to adding formaldehyde may affect the state of the nucleus. In each case, the final concentration of formaldehyde is 1%. (First Panel Top) MDA-MB-231 cells directly under the site of 37% formaldehyde addition show a notable shape change. (First Panel Bottom) Cell morphology changes are less apparent when formaldehyde is pre-diluted to 1% in 1xHBSS. (Second Panel Top) GM12878 nuclei immobilized on a poly-D-lysine coverslip show a dramatic shape change when directly under the site of 37% formaldehyde addition. (Second panel middle) Nuclear shape change is still observed when 1% formaldehyde is pre-diluted in water (Second panel bottom) Nuclear structure is preserved when formaldehyde is pre-diluted in 0.25x PBS. C) The appearance of DNA varies for HepG2 cells harvested by either scraping after crosslinking or enzymatic removal followed by crosslinking. In each case, crosslinks were reversed in harvested cells for 1 hour at 65^{o}C with proteinase K and then DNA was visualized on a 0.8% agarose gel. “Scraped” DNA presented as sticky threads and partly floated out of the well before entering the gel. D) B16-F1 mouse melanoma cells retain extensive cytoplasmic debris in the same nucleus purification protocol that effectively purifies GM12878 nuclei. E) GM12878 nuclei can lyse when 1x PBS buffer is replaced with water, an example to show that unusual buffer conditions could contribute to higher or lower rates of intact nuclei entering the downstream Hi-C protocol. F) Hi-C protocols performed simultaneously on two different cell types can result in differing proportions of informative interaction pairs (Valid Pairs). While unligated HindIII fragments (HindIII dangling ends) are consistent between the cell types, one experiment suffered from more biotin incorporation at nicks resulting in non-canonical dangling ends. These non-canonical ends are not are recalcitrant to additional biotin removal from DNA ends with the exonuclease activity of T4 DNA polymerase.

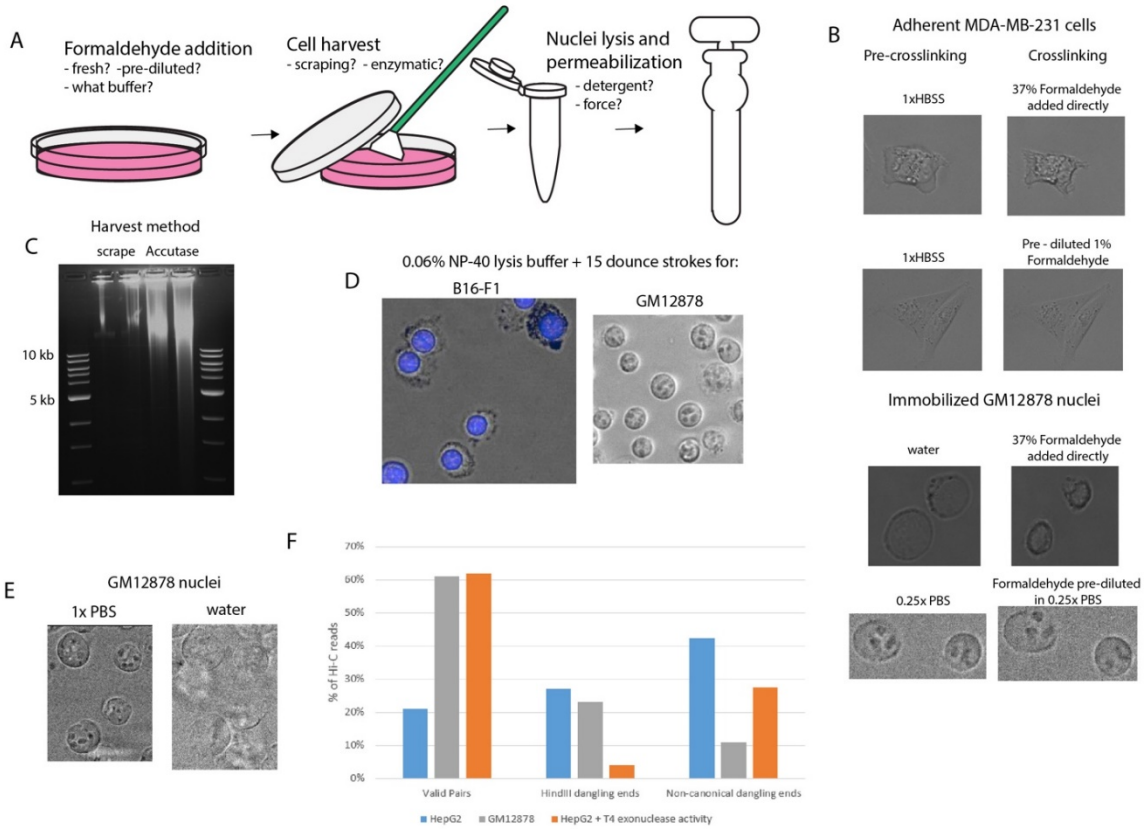


Figure 3. Variable factors in early steps of the Hi-C protocol.

of material that reduce digestion, fill-in, and ligation efficiency (Gondor et al., 2008). The effect of nuclei lysis on the quality of Hi-C libraries supports the idea that the initial state of the biological material going into a Hi-C experiment can affect the likelihood of generating a high-quality Hi-C library. Here, we will discuss aspects of the initial Hi-C sample preparation that may affect whether nuclei stay intact, whether random chromatin damage is generated, and whether chromatin is efficiently digested and recovered from nuclei. Unfortunately, even when a theoretically consistent protocol is used, there can be variation in what actually happens to a given cell type during the initial steps of Hi-C. Such variations may lead to downstream effects such as a low ratio of cis to trans interactions (which may reflect high background noise) and a high percentage of dangling ends (sequenced fragments that do not contain chimeric interactions; see Table 1), reducing the fraction of valid interaction pairs among sequenced reads. In general, for the areas described below, we advocate a thorough documentation of the methods used and their effects on the appearance of nuclei and chromatin. By better linking such metadata to downstream Hi-C library results, the field can gain a better understanding of how to standardize or optimize these steps of the protocol.

Formaldehyde Crosslinking

The effects of variations in formaldehyde concentration and crosslinking time on Hi-C results have been considered previously (Naumova et al., 2013, Belmont, 2014). It has also been noted that the presence of serum in the cell culture media can inhibit effective crosslinking (Belton et al., 2012). Other work has suggested that the preparation of fresh formaldehyde immediately before each experiment, to prevent uncontrolled formaldehyde

polymerization, is helpful to increase reproducibility (Gondor et al., 2008). In addition to these previous observations, we have noticed that changes in cell morphology can occur depending on how formaldehyde is added (Figure 3B). When concentrated formaldehyde (37%) is added directly to adherent cells, an instantaneous and irreversible shrinkage of the cells nearest to the site of formaldehyde addition is observed (Figure 3B). This effect of adding concentrated formaldehyde directly into the sample is even more noticeable for isolated nuclei immobilized on a poly-D-lysine coated dish (Figure 3B). (Note that isolated nuclei are often used for Hi-C experiments on cells with thick cell walls or from complex tissues (Dekker et al., 2002, Dong et al., 2017)). Evident, though less dramatic, alterations in nucleus appearance were also observed when formaldehyde pre-diluted to 1% in water was added to nuclei already incubated in water. This result is a cautionary note for experiments in which cells need to be in an unusual buffer related to the treatment of interest. The best preservation of pre-existing nuclear appearance was achieved by replacing cell culture media with a pre-mixed 1% solution of formaldehyde in either serum-free cell culture media, or other buffered solutions (PBS, HBSS, HEPES). There is no clear evidence at this time about how changes in nuclear appearance relate to the chromosomal interactions detected by Hi-C, but changes in appearance after localized addition of highly concentrated formaldehyde suggest that certain cells are experiencing a much higher effective concentration of formaldehyde than others. This would likely lead to heterogeneity of crosslinked complex size across the cell population and to some cells having crosslinked networks impenetrable to digestion enzymes (Hoffman et al., 2015).

To avoid this, we recommend that formaldehyde be pre-diluted in an appropriate buffer before addition to the cells destined for Hi-C.

Cell Harvest

The harvest of crosslinked cells is a common step to many genomic experiments (including ChIP-Seq and others), but there are still considerations in this step that are often unexplored. We have observed that different cell types vary greatly in how easily they can be recovered by manual scraping from culture dishes after crosslinking. Differentiated myotubes can require trypsinization during scraping after crosslinking to aid release from the dish (McCord et al., 2011) while cells like HepG2 detach very easily (unpublished observations).

Consequently, different cellular preparations may yield variable results: aggregates of cells that are later more difficult to permeabilize, physically damaged cells, or loss of cellular material. One option for increasing the reproducibility and recovery of material is to gently enzymatically detach cells from the plate using, for example, Accutase, before crosslinking and then to perform crosslinking steps in a single cell suspension. Our results indicate that this approach can lead to better estimation of cell numbers going into the experiment and fewer cellular aggregates that may be hard to permeabilize. Visualizing DNA recovered from cells that were harvested with Accutase before crosslinking was easier than cells crosslinked and scraped, but this in part reflects the fact that DNA from these cells was more fragmented, while DNA from scraped cells was so intact that it became sticky and floated out of the gel (Figure 3C). Also, since cellular attachments to

different substrates can affect nuclear structure (Makhija et al., 2016), cells may not retain their native 3D genome conformation after being detached. So, crosslinking adherent cells in their native state is often preferable, but we recommend evaluating crosslinked plates under the microscope during scraping to document the amount of scraping needed to retrieve the cells and the appearance of the cells after harvest.

Nuclei Isolation / Permeabilization

Many Hi-C protocols recommend that nuclei be released by dounce homogenization before restriction digestion begins. But, like cell scraping, the results of this technique vary with douncing speed and cell type. Indeed, we have observed that different cell types can look quite different after the same nuclei isolation steps. While GM12878 nuclei are cleanly isolated after incubation in a detergent-containing buffer and 15 strokes of the dounce homogenizer, mouse melanoma cells still retain a large amount of debris around their nuclei after the same treatment (Figure 3D). If similar variation in cell lysis across cell types occurs during early Hi-C steps, this could result in nuclei that are too inaccessible or excessively ruptured going into the later steps of the Hi-C protocol. Relatedly, changes in buffer conditions can cause dramatic effects on isolated nuclei, again, depending on cell type. We have observed that upon addition of 1 drop of pure water to isolated nuclei from GM12878 cells, the nuclei can rupture (Figure 3E), while a HeLa epithelial-type cancer cell nucleus is minimally affected by the same treatment. While it would be unusual for a researcher to place cells or nuclei into pure water, other buffer components more likely to be overlooked (like the addition of EDTA (Takata et al., 2013)) could cause a similar effect. In practice, resulting variations are often noted in the

visible properties of Hi-C nuclei before the digestion step: some nuclei settle quickly or are more clumpy, granular, or sticky. But such differences are rarely reported or considered. So that the field can develop a more complete understanding of what upstream steps lead to these differences, and how they affect the quality of the resulting Hi-C library, we recommend documenting microscopic images of cells and nuclei to link their appearance before and after crosslinking, and before digestion to their downstream Hi-C success.

DNA damage affects Hi-C library quality

It is possible that variations in the early physical harvesting steps described above may impact the initial integrity of chromatin. In both long read Sanger sequencing of isolated Hi-C library molecules and high throughput sequencing results, we have observed that different replicates and cell types can have varying proportions of breaks in the DNA at sites that do not correspond to restriction enzyme sequences (discussed in detail as “random breaks” in Imakaev et al. (Imakaev et al., 2012)). While the portion of these random breaks that end up forming ligated Hi-C interaction products likely result from off target restriction enzyme activity, many of these non-canonical breaks correspond to unligated but biotinylated molecules that end up being uninformative reads (“dangling ends”) at the end of the Hi-C experiment, reducing the number of interaction pairs recovered. In particular, these non-canonical dangling ends also seem to be recalcitrant to attempts to remove biotin at fragment ends (Figure 3F) and may therefore arise from biotin incorporation at nicks in the middle of fragments. These may arise from factors in the experiment other than off-target restriction enzyme activity. For example, if the nuclei

rupture during cell harvesting or nuclei isolation, this may expose the chromatin to DNA damaging agents or physically damage the DNA. Notably, the example in Figure 3F of extensive random breaks in the final Hi-C library was the end product of the “Accutase” sample shown in Figure 3C, in which the DNA appeared to be a bit more fragmented after initial harvesting. However, a conventional 0.8% gel is not the best approach for visualizing such large DNA fragments. In the future, pulsed-field gel electrophoresis (PFGE) (Bryant, 2012) could greatly aid the investigation of what initial conditions lead to the recovery of the most intact DNA.

Future optimization of cell preparation needed

Poorly controlled variations in the steps described above can have a negative impact on the downstream quality of the Hi-C library. Insufficient cell recovery during cell harvest steps can lead to low amounts of DNA and low library complexity. Incomplete nucleus isolation or permeabilization may affect the efficiency of downstream steps such as digestion, biotin fill-in, and ligation, in which enzymes must access the chromatin. As previously described, DNA recovery and digestion efficiency can be checked on a gel before subsequent steps, though this check would likely be more efficient using PFGE (Belaghzal et al., 2017, Belton et al., 2012). With data from a variety of Hi-C experiments, we have observed that when samples from otherwise similar experiments have varying degrees of DNA digestion, those with less well-digested DNA have a smaller proportion of valid pair Hi-C interaction reads, higher fractions of dangling ends, and higher levels of background ligation noise at the end of the experiment (Appendix A).

All the variables discussed above apply in particular to Hi-C being performed on cells in culture. Controlling the consistent initial preparation of crosslinked material before digestion is even more challenging when whole primary tissues are the source of biological material. Numerous upstream approaches to prepare diverse biological materials have been developed (Mascher et al., 2017, Schwarzer et al., 2017, Gabdank et al., 2016, Crane et al., 2015), but many of these require similar downstream steps discussed above. To increase the success of the complex preparations of diverse biological materials for future Hi-C experiments, future work should explore protocols that reduce variability in ambiguous steps such as scraping and douncing. There is also a great need for more informative quality control tests of initial Hi-C steps that will allow researchers to assay whether a given preparation of material is suitable (good chromatin accessibility and material recovery with low DNA damage and nuclei lysis) for the expensive downstream steps of digestion, biotin fill-in, and ligation, not to mention library preparation and sequencing.

Currently, Hi-C data gathered by the 4D Nucleome Consortium (Dekker et al., 2017) is archived along with a variety of metadata about the experiment (https://data.4dnucleome.org/). We encourage researchers to provide such details about both intermediate phases and final library statistics about Hi-C experiments, including cell appearance at various stages described above as well as DNA integrity and digestion efficiency information. To correlate these early checkpoints with final Hi-C results, we encourage all researchers to report statistics informative about the quality of libraries

including the percentage of dangling ends, the cis/trans interaction ratio, the percent of reads that are PCR duplicates, and the percentage of reads that represent mid-range distance (25 kb – 10 Mb) interactions (poor quality libraries are often enriched for both extremely proximal and extremely distant interactions).

Utility of Hi-C 2.0 Biotin Fill-in conditions for HindIII libraries

Recently, Belaghzal et al. have developed an optimized in nucleus Hi-C protocol called Hi-C 2.0 which uses the DpnII enzyme and is designed for single aliquots of 5 million cells (Belaghzal et al., 2017). Some changes in the protocol from previously published approaches (Rao et al., 2014, Belton et al., 2012) might seem to be minor or just specific to the use of biotin-dATP and DpnII. Here, we have compared an in-nucleus version of the Belton et al. Hi-C protocol (referred to as Hi-C 1.0) (Belton et al., 2012) with the modifications of the Hi-C 2.0 method (Figure 4). We find that the technical variations in Hi-C 2.0 notably improve HindIII Hi-C library quality in a direct comparison using aliquots of 5 million mouse melanoma cells (B16-F1) identically prepared in initial steps (see Appendix B for protocols used). In particular, the modified Hi-C 2.0 protocol involves performing biotin-dCTP fill-in at 23oC for 4 hours instead of 37oC for 2 hours. When applied to a HindIII Hi-C library preparation, these modifications resulted in a nearly 10% increase in valid interaction pairs due to a decrease in unligated dangling ends (Figure 4A). Notably, the decrease in dangling end molecules corresponded to a reduction of the bias toward inward facing reads among valid pairs. This indicates that a fraction of the excess inward facing interaction pairs are actually undigested dangling ends, which are then also reduced when dangling ends are removed.

A) Hi-C library statistics for aliquots of 5 million B16-F1 mouse melanoma cells prepared with: Hi-C 1.0 with HindIII (top), Hi-C 2.0 with HindIII (middle), and Hi-C 2.0 with DpnII (bottom). Left: proportions of reads that are valid interaction pairs as compared to unmapped or single fragment reads. Middle: proportions of mapping orientations of valid interaction pairs. Hi-C 2.0 gives less inward bias, indicating a reduction of biases arising from undigested fragments. Right: Fraction of intrachromosomal (cis) and interchromosomal (trans) interaction pairs. B) Concordance between interaction maps resulting from different Hi-C protocol approaches as calculated by GenomeDISCO. C) Similarity of 40 kb binned and iteratively corrected (Imakaev et al., 2012) heatmaps from the 3 parallel experiments for a 7.5 Mb region of chr19. D) Insulation score profiles (sliding window size of 500 kb, heatmap resolution of 40 kb) for a section of chromosome 19 for each Hi-C protocol. Overall TAD boundaries (major dips in the insulation score) are conserved in position and strength between the Hi-C protocols at this resolution.

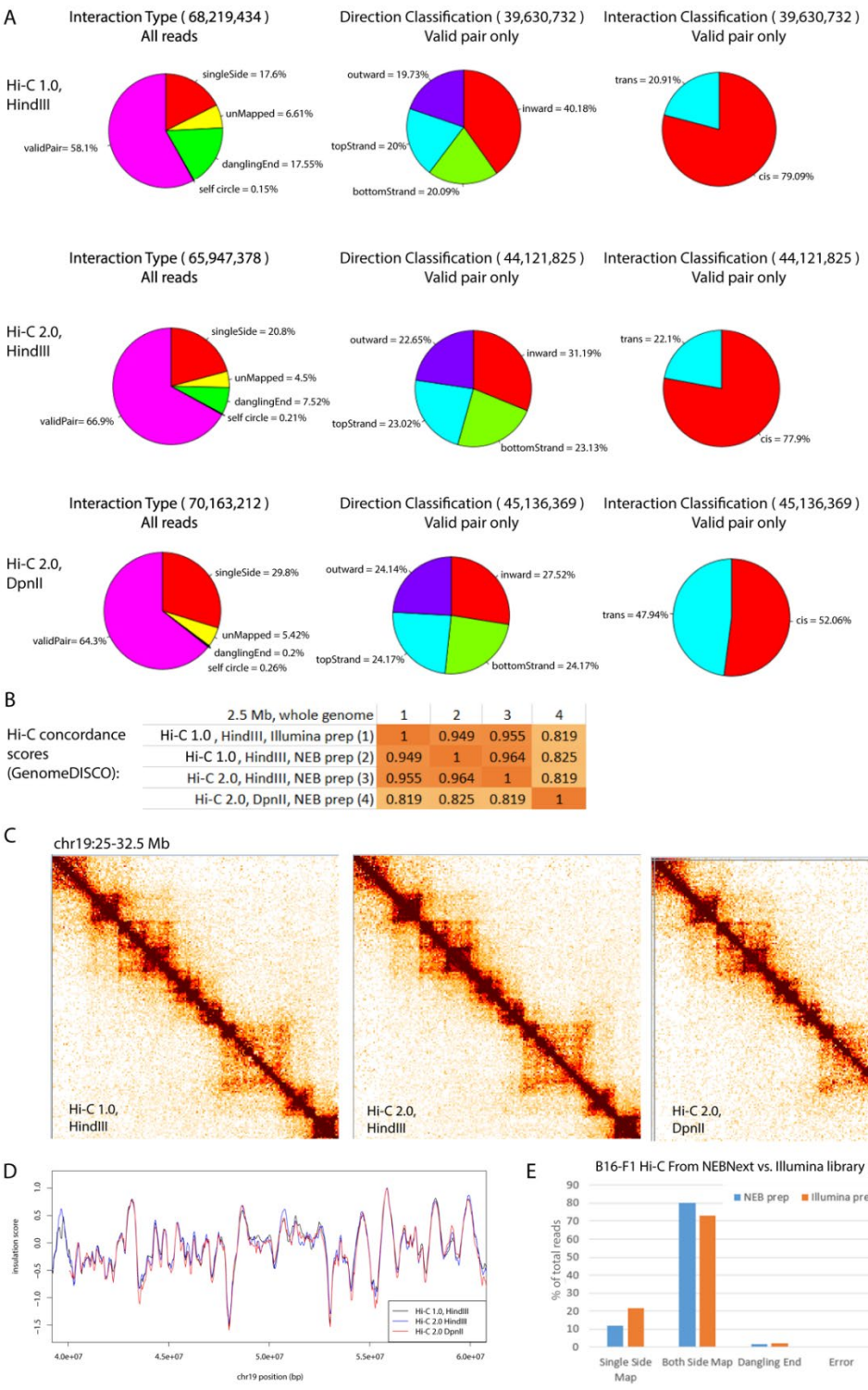


Figure 4. Hi-C 2.0 (Belaghzal et al., 2017) protocol changes benefit HindIII Hi-C libraries as well as DpnII libraries.

Dangling ends are reduced even further when DpnII with biotin-dATP is used instead of HindIII in the Hi-C 2.0 protocol (Figure 4A). However, an increase in interchromosomal ("% trans") interactions is observed, which could indicate higher background noise in this condition. This increase in the fraction of interchromosomal interactions when the cells were derived from the same initial preparation steps shows that nuclear lysis (which would be the same in each of these aliquots) is not the only factor that can contribute to higher levels of potential background ligation. Overall, all of these approaches lead to reproducible final Hi-C libraries according to concordance scores calculated by GenomeDISCO (Ursu et al., 2018) and visual comparisons of the resulting heatmaps at 40 kb resolution (Figure 4B and C). TAD boundary locations and strengths at 40 kb resolution are also overall similar according to an insulation score profile (Figure 4D) (Crane et al., 2015). Thus, the major gain in these particular optimizations will lie primarily in the fraction of useful reads gained from a certain number of cells, allowing the generation of high resolution heatmaps from smaller numbers of total sequenced reads.

Approaches and Artifacts during Hi-C library preparation for sequencing

After the "3C-like" phase of the Hi-C experiment is complete (digestion, biotin fill-in, and ligation), the Hi-C DNA follows some steps that are common to nearly all high throughput sequencing experiments but with variations and factors unique to Hi-C that require special consideration. Even with a high-quality library at the beginning of this stage, variations in library preparation can matter, and problems with low quality libraries become exacerbated during preparations for sequencing.

Reproducible NEBNext library preparation of captured interactions

Like some previously published protocols (Rao et al., 2014), we find it useful to perform streptavidin bead capture of biotinylated interactions immediately after size selection and before any end repair. We have adapted the NEBNext® Ultra™ II DNA Library Prep protocol for use on beads, resulting in lower reaction volumes, fewer steps, and reduced time than some previous library preparation approaches (Belton et al., 2012) (Appendix B). In our hands, this protocol is easier and more reproducible specifically in cases of indexed adaptor ligation than adapting the Illumina protocol to Hi-C, possibly due to the divergent design of NEB adaptors. We find that substantial dilution of adaptors (at least 1:15) is necessary to avoid excess adaptors in the final Hi-C library. In the end, the final Hi-C library statistics are very similar when the same ligated Hi-C material is used for an Illumina library preparation with streptavidin bead capture immediately before adaptor ligation as compared to an NEBNext library preparation (Figure 4E). NEBNext library preparation in fact resulted in a slightly higher percentage of sequencing reads for which both paired reads align to the genome.

Disadvantages of over-sonication to short DNA fragment length

Given that ligated interacting DNA molecules are often multiple kilobases in length, fragmentation of DNA to smaller sizes for sequencing on the Illumina platform is essential. However, some evidence suggests that it is better to leave some larger DNA fragments rather than trying to achieve average sizes of 200 bp, as has sometimes been recommended (Belaghzal et al., 2017, Belton et al., 2012). Smaller fragments reduce the overall mappability of the Hi-C library after sequencing, in part because shorter

sequences are less likely to be uniquely mappable, particularly in repetitive regions (Figure 5A). For short fragments, it is more likely that one of the two interacting fragments will be too short to be aligned uniquely to the genome before a ligation junction is encountered (Figure 5A). This phenomenon results in a higher proportion of “single side mapped” reads. Also, evidence further discussed below suggests that shorter molecules are more likely to contain end-repair related artifacts.

Dangers of chimera formation and end repair-based artifacts

In all high throughput genome-wide sequencing approaches, the complexity of DNA sequences in the same tube raises the possibility that small regions of homology between molecules could hybridize and form chimeras during PCR amplification. However, in techniques other than Hi-C, chimeric reads can easily be excluded due to their impossible genomic positions (paired reads from different chromosomes, reads from the same DNA strand, etc.). In Hi-C, by contrast, informative paired sequence reads should come from all orientations and locations in the genome, making these possible artifacts harder to detect or exclude. We have found evidence of one particular type of artifact that we propose occurs due to complementary base pairing during end repair. In the paired end sequencing results of some Hi-C experiments, we found an unusually low rate of mapping for Read 2 (81% mappability for Read 1 vs. 45% for Read 2). Strangely, we discovered 10-15 bases of the 5’ end of the library molecule had been copied artificially onto the 3’ end, resulting in 10-20 identical bases at the beginning of paired end Read 1 and Read 2. Thus, Read 2 would not map unless the first bases were excluded (Figure 5B and C).

A) Fragment size distributions of the final Hi-C library (ligated DNA insert with ligated adaptors) from an Agilent Bioanalyzer trace. The shorter fragments with a narrower distribution have a higher percentage of single side mapped reads. This is likely attributable to the higher chance of encountering a ligation junction in the first few base pairs of the shorter fragment (blue line, bottom panel), resulting in one unmappable interaction partner. B) Sample reads demonstrating the artifact in which a piece of the 5' end of a Hi-C interaction molecule is artificially attached to the 3' end. As a result, paired end Read 1 and 2 are identical for the first 10-15 bp and Read 2 is unmappable without 5' trimming. C) A Hi-C experiment on GM12878 nuclei showed an extremely high amount of copying from the 5' to 3' end of interaction library molecules, resulting in a large % of the total read pairs with 10-20 bases identical between Read 1 and Read 2. This corresponds to only 45% alignment rate of Read 2. This effect is observed to a lesser, but notable extent in other experimental data, such as published HFF1 metaphase Hi-C (Naumova et al., 2013). Other published datasets, such as pro B cell Hi-C data (Zhang et al., 2012) show only a peak at 13 bp, which corresponds to the amount of identity between adaptor dimers. A Hi-C library prepared by tagmentation (Gabdank et al., 2016) (SRR3105477) shows none of this copying artifact. D) Most library molecules that experience this copying artifact have 4-6 bp of complementary sequence between the two ends of the molecule. The histogram shown represents the length of identical sequence between Read 1 and Read 2 that maps on both sides and is immediately adjacent to unmappable copied sequence on Read 2 (blue box from B). E) Reads containing these artifacts (blue) tend to arise from shorter DNA molecules than non-artifact reads (red). F) Model of explanation of how short complementary regions ("microhomologies", blue boxes) could anneal and then elongate during end repair, copying the 5' end of the top strand and appending it to the 3' end of the top strand. During sequencing, this chimeric 3' end is sequenced as Read 2 as shown. G) In a Hi-C library with large numbers of noisy interchromosomal read pairs ("High % trans"), a high percentage of these reads contained no Hi-C junction sequence, suggesting that they may arise from something other than ordinary Hi-C ligation. When the number of interchromosomal pairs is low, on the other hand, most of these "trans" reads contain a canonical Hi-C junction sequence.

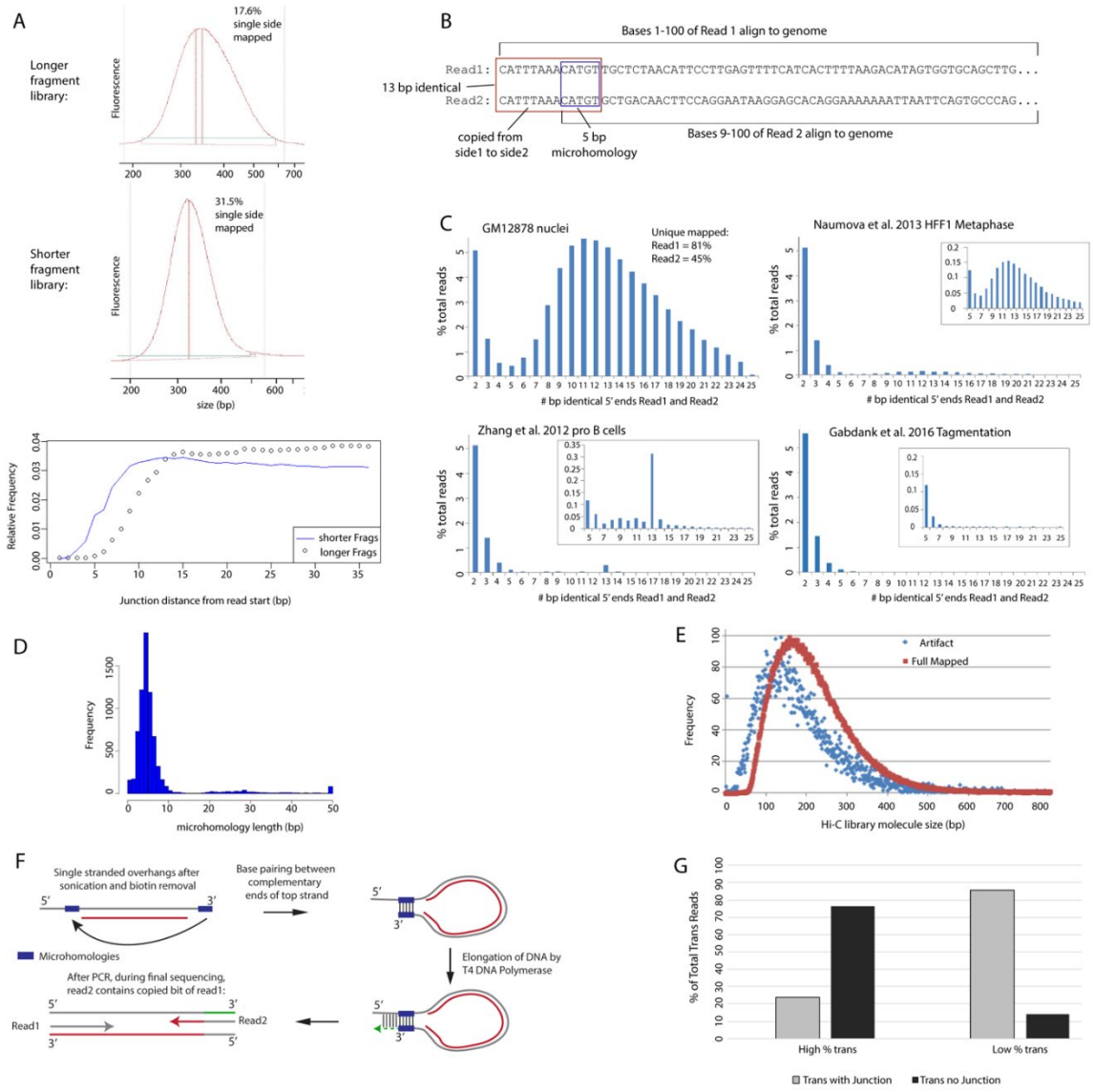


Figure 5. Artifacts arising from Hi-C sequencing library preparation.

Further investigation revealed that molecules with this artifact had 4-6 bp of microhomology within the mappable portions of Read 1 and Read 2 (Figure 5A and D). This indicates that the original ligated Hi-C fragments contained a complementary sequence between the 5' and 3' ends that existed before the copying artifact occurred (Appendix C). This artifact is most common in libraries that have undergone extensive biotin removal steps and sonication to small fragment sizes (Figure 5E), conditions that would tend to expose single stranded overhangs. We propose the following explanation for this phenomenon, based on the evidence we have collected: during the room temperature end repair step, short complementary single stranded 5' and 3' ends of the Hi-C molecule could anneal to each other, and then the T4 DNA polymerase would be able to elongate the DNA from this "priming", copying the 5' sequence (Read 1 when sequenced) onto the 3' end of the molecule (which becomes Read 2) (Figure 5F).

In libraries in which this phenomenon occurs, a percentage of Read 2 sequences will not align to the genome, but instead contain identical sequences at the beginning of Read 1 and Read 2 that span a smooth distribution of lengths between 7-25 base pairs. In some libraries, identical sequences between Read 1 and 2 of exactly 13 bp are observed (Figure 5C and Appendix C), but this can be explained by the sequencing of adaptor sequence at both ends of the molecule, which can occur in adaptor dimers or if more than one adaptor is ligated to the same library molecule end. In cases not explained by adaptor dimers, no enriched sequence motifs are noted in the copied or complementary DNA pieces, but these sequences are AT-rich (Appendix C), which would correspond to

greater fragility (M. Manghi, 2016) and likelihood of ssDNA ends. Stable base pairing resulting from a few complementary bases may sound unbelievable, but, interestingly, a very similar mechanism is known to happen in vivo during certain types of DNA repair (Ottaviani et al., 2014). If there are 3' single stranded overhangs on damaged DNA ends, microhomologies of 1-4 bp can anneal and undergo fill-in synthesis by DNA polymerase (Lieber, 2010, Crespan et al., 2012). Microhomology annealing can be a small bias in non-homologous end joining repair or microhomology mediated repair pathways. This artifact is minor in most Hi-C libraries but suggests that care must be taken not to over-sonicate the DNA or to overdo the biotin removal recessing of DNA ends.

Future library preparation protocols could also consider variations to the end repair steps to avoid the possibility of low temperature priming and extension from small identical regions. Notably, we have not observed this artifact in Hi-C libraries prepared with tagmentation for adaptor incorporation (Gabdank et al., 2016) (Figure 5C) or in 5C libraries, in which no random fragmentation is performed and adaptors are added by PCR (Smith et al., 2016) (Appendix C). On the other hand, consistent with the proposal that this artifact occurs during the end repair step common to many techniques other than Hi-C, we have observed the very same end copying in sequencing data from a ChIP-Seq library (Teif et al., 2017) (Appendix C).

Fortunately, the direct copying of sequence from Read 1 onto Read 2 is easily detectable, and the resulting sequence either does not map or is discarded as two reads within the

same fragment. But, chimeras that are more difficult to detect could form in a similar way either during the end repair or final PCR steps. As previously documented (Belton et al., 2012), concatenations of Hi-C library molecules cause an upward shift in overall Hi-C library molecular weight after too many rounds of PCR amplification. But, such chimeras may form before this shift is visibly detectable. Indeed, in libraries with more noisy random interchromosomal ("trans") ligations, a higher proportion of these interchromosomal interaction reads lack the canonical Hi-C junction sequence (Figure 5G). Such random interaction noise without true ligation junctions could arise from PCR chimeras formed in an attempt to amplify the few real interactions that exist in a poor library. This observation again indicates that increases in random background signal may not always stem from actual random ligation in solution due to nuclei lysis but also can be caused by other factors.

Considerations in sequencing and mapping

How much can be gained by deeper sequencing: Issues of library complexity

The potential number of pairwise genomic interactions in a mammalian genome is on the order of 1011 – 1012, so large numbers (hundreds of millions to a billion) of paired sequence reads are often needed to thoroughly sample this interaction space at high resolution. However, deeper sequencing will not lead to better Hi-C maps for all samples. The input number of cells and the efficiency of digestion and ligation limit the number of possible captured interactions. If digestion with a 6-cutter enzyme and interaction ligation were perfectly efficient, 5 million cells could indeed lead to a library of 1012 captured interactions. However, inefficiencies and losses occur at many steps, so that the

interaction complexity in practice is often much more limited. Thus, while a highly complex library can provide better interaction maps with additional sequencing, the interaction map resulting from a low complexity library will barely change even when the number of reads is increased to 1 billion (Figure 6A). A high proportion (more than a few percent) of PCR duplicate molecules in a Hi-C library can also signal a lack of complexity. But, the increased prevalence of optical duplicates on patterned HiSeq 3000 and 4000 flowcells (Wingett, 2017) adds uncertainty to this potential measure of complexity.

How finely can the interactions be binned?

Structural information derived from read location within restriction fragments

Interactions detected by sequencing Hi-C products are often mapped to the midpoint of each represented restriction enzyme, based on the idea that the crosslinked interaction could have occurred anywhere across the given fragment. We have found, however, that sometimes one end of the fragment is favored for an interaction over the other end, and that this bias can contain biologically significant information about the position of the specific interaction. For example, for interactions of between long fragments that are within 40 kb of each other along the linear genome, more ligations are observed between the fragment ends that are closer together (Figure 6B). This shows that the natural decay of interaction frequency along the chromatin polymer is revealed in a restriction fragment end preference for sequence reads. We therefore argue that it is valid to bin sequence reads to constant bin sizes, taking into account exactly where each read lands rather than only which restriction fragment it belongs to. Employing this approach, however,

A) 10 kb binned and corrected heatmaps in a 4 Mb region of chr22 with varying input read numbers for poor complexity and good complexity replicates of a GM12878 Hi-C experiment. The poor complexity Hi-C library (left) gives no additional interaction resolution or detail even when input reads are increased to 1 billion (upper triangle) vs. 335 million (lower triangle). The high complexity library (right) on the other hand gives much more detailed information at 863 million reads (upper triangle), while only 27 million reads (lower triangle) are needed to give the amount of detail captured by 355 million reads of the poor complexity library. B) Three scenarios of proximal interactions are considered for their effect on interacting strand/side bias: “Frag1 long”: With increasing upstream restriction fragment length, positive strand interactions (“top”, green line) are increasingly favored compared to negative strand interactions (“bottom”, purple line). “Frag2 long”: with increasing downstream fragment length, the opposite pattern holds. “Either/both long”: if either fragment can be equally long, no strand bias is seen, only a preference for inward facing reads (closest on the linear genome) vs. outward facing reads (farthest apart on the linear genome). C) Distributions of HindIII (left) and DpnII (right) restriction fragment lengths in the mouse genome (mm9).

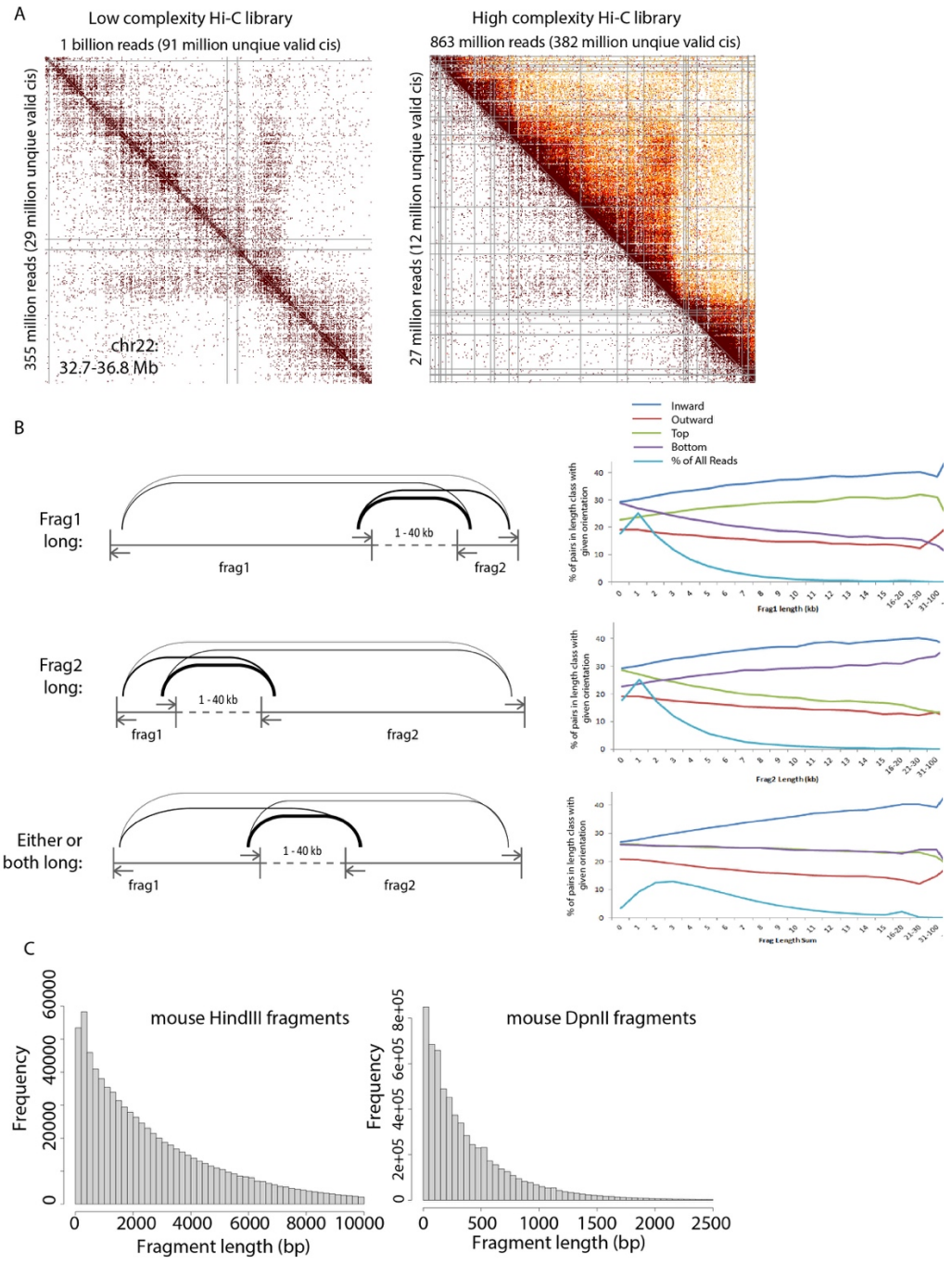


Figure 6. Considerations in Hi-C data alignment and binning.

absolutely requires that resulting interaction maps be carefully corrected for artifacts by a fragment-position naïve method such as iterative correction (Imakaev et al., 2012).

Comparing restriction fragment size distributions to a chosen bin size

When choosing a high-resolution bin size for Hi-C data generated with a given restriction enzyme, it is also important to consider the restriction fragment size distribution to limit the number of bins containing no fragment ends. Publications often quote the average or median restriction fragment size, but we note that this can be a misleading value. HindIII and DpnII fragments follow a non-normal distribution (Figure 6C) in the human genome, so the average or median size does not imply that most fragments are close to this size. In fact, while the average HindIII fragment size in the mouse genome is 3.2 kb, nearly 30% of fragments are less than 1 kb in size, so there are numerous interactions that can be detected at a higher resolution than the average fragment size. Conversely, the median DpnII fragment size in the mouse genome is 260 bp, but nearly 30% of DpnII fragments are larger than 500 bp, so using a bin size of 200-300 bp to match the median fragment size will result in numerous bins with no information.

Considerations for applying Hi-C to lower input cell numbers

Interpreting Hi-C data always requires the consideration that the data represents the average structure across a population of cells (Lajoie et al., 2015). Single cell Hi-C approaches have made it possible to capture interactions that occur in individual cells (Nagano et al., 2017), but these datasets are inherently limited in their information, since a maximum of 4 interactions (2 ends of each fragment on 2 alleles) can be captured from

each genomic location in any given cell. Therefore, there is a need for continued development of protocols effective for small populations of cells, particularly as Hi-C is applied to more primary tissues or to capture genome structure during rare cellular events. We have shown earlier that results between replicate batches of 5 million cells are highly reproducible (Figure 4), but many applications will require lower cell numbers than this. We have noted several experimental considerations that will be important as Hi-C is adapted to lower numbers of cells.

Steps prone to substantial DNA loss

With small numbers of cells, recovery of all possible DNA is important at every step. We have found unexpected losses of DNA at certain steps of the Hi-C protocol, which would be unacceptable if working with low cell numbers.

Ethanol precipitation

Ethanol precipitation performed in large volumes as recommended by many Hi-C protocols can be inefficient. Indeed, we find that substantial amounts of Hi-C DNA may still be in the ethanol supernatant after the precipitation of ligated Hi-C DNA molecules. In two independent experiments, we have performed another round of DNA precipitation on the ethanol supernatant from the first DNA purification attempt and have recovered microgram quantities of additional DNA (10-40% of the original amount recovered) (Figure 7A). This indicates that in the first round of ethanol precipitation, substantial

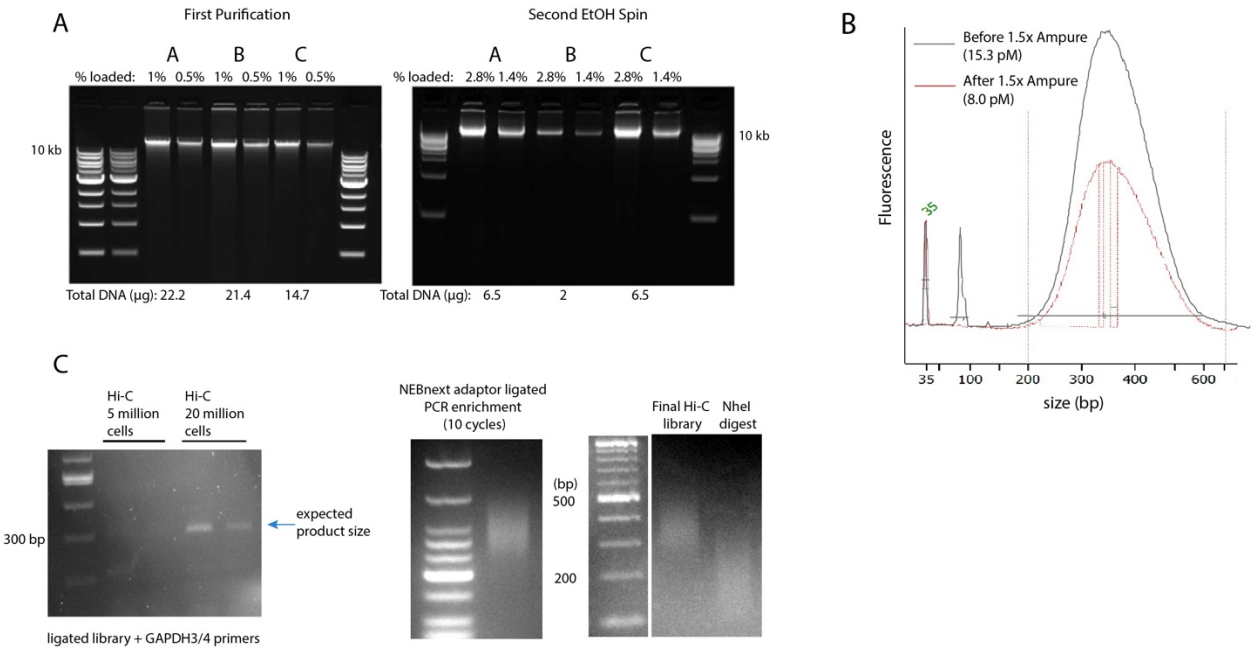


Figure 7. Steps prone to substantial DNA loss.

A) After reversal of crosslinks, three different GM12878 Hi-C libraries (A, B, and C) were purified and ethanol precipitated as recommended (Belton et al., 2012), yielding 14-22 micrograms of DNA. The ethanol supernatant removed from the DNA pellet was then re-centrifuged at the same speed as in the first purification, and another 2 – 6 micrograms (10-40% more) of DNA was recovered. B) Bioanalyzer traces before (black) and after (red) 1.5x Ampure XP purification are shown, scaled according to their 35 bp standard peak. Size-selection of Hi-C library DNA with 1.5x Ampure XP beads successfully removes the ~80 bp excess adaptor peak (black), but reduces the overall concentration of DNA by nearly half. C) For the HindIII Hi-C 2.0 experiment on 5 million B16-F1 cells, no PCR product was observed for a neighboring interaction primer pair in the GAPDH gene region (primers described in (Belton et al., 2012)) (left). However, the final Hi-C library could be amplified with 10 cycles of PCR (middle), and digested well with NheI, as expected for a high quality Hi-C library (right). This final library produced good Hi-C interaction data (see Figure 6) despite the failed quality control neighboring interaction PCR. Note: some lanes in gels have been excised for clarity, but all lanes shown in contiguous images were run on the same gel.

quantities of DNA were lost. Indeed, as 4C experiments are adapted to lower cell numbers, smaller volume DNA precipitations and minimizing tube transfers have been recommended (Davies et al., 2016).

Reversing crosslinks

Many Hi-C protocols recommend reversing formaldehyde crosslinks at 65 degrees C overnight. We have observed that this lengthy high temperature step can result in DNA degradation, presumably from trace amounts of DNase enzymes. Extreme caution must be taken to avoid any DNase contamination before this step and minimizing the length of this incubation may help reduce the likelihood of DNA degradation.

AMPure purification

AMPure XP beads are a powerful and convenient way to purify a given size range of sequencing libraries and to remove adaptor and primer dimers. However, this step can also result in substantial DNA loss in the desired size range as well. We have observed a 50% loss of Hi-C library material after AMPure purification was performed to remove excess adaptor dimers (Figure 7B). Troubleshooting guidelines provided with AMPure beads should be considered to avoid loss of material. Further, since any purification step will involve some DNA loss, protocols should be optimized to minimize the number of purification steps required. For example, increasing the dilution of adaptors in the adaptor ligation step can minimize the need for purification to remove adaptor dimers.

Difficulties with assessing quality for Hi-C with low cell numbers

Previously published Hi-C protocols often recommend checking the quality of a ligated Hi-C library after by performing a targeted PCR on one or a few neighboring interaction products (Belaghzal et al., 2017, Belton et al., 2012). However, we find that this control is less likely to be a reliable indicator of Hi-C library success when smaller input cell numbers are used. Though interaction patterns are reproducible as cell numbers are decreased from 20 million to 5 million (Figure 4B), the likelihood of detecting with PCR any given single interaction out of a small sample of the Hi-C library decreases with lower cell numbers. We have observed cases in which the quality control PCR of one interaction amplified no detectable product, but the final Hi-C library and sequenced results were high quality (Figure 7C).

Indeed, even with larger numbers of cells, a PCR digest control can be unreliable. In a recent publication of Hi-C in barley, for example, two variations of Hi-C were used. Effective NheI digestion of the single interaction PCR product band suggested that one protocol worked much better than the other, but NheI digestion of the complex final library reveals that both experiments are similarly high quality (Mascher et al., 2017). In general, different quality control metrics may be needed to evaluate Hi-C experiments with low cell numbers. Experiments adapting 4C to low cell numbers have found, for example, that short range interaction profiles from low cell numbers can be reproducible even while weak long-range interactions are harder to detect reproducibly (Davies et al., 2016).

Conclusions

The complex nature of Hi-C experiments means that there are many steps, from cell harvesting to sequence mapping, which can affect the quality of the downstream genome interaction maps. Protocol improvements are continuing to increase the power, information, and resolution that can be gleaned from Hi-C experiments. Adjustments to the protocol such as those suggested in Hi-C 2.0, and ideas suggested here, such as avoiding over-sonication, can notably increase the recovery of informative interaction pairs. Artifacts and noise in the Hi-C data can arise from variables in the early cell preparation steps or as a result of the complex mixtures subjected to end repair and PCR steps. Some of these sources of artifacts occasionally lead to Hi-C experiments with lower numbers of valid interacting pairs or higher background noise. Fortunately, as we have shown here, many features of Hi-C maps are reproducible despite protocol variations. As more scientists apply Hi-C to study 3D genome structures in diverse systems, careful attention to and documentation of the details of successful and problematic Hi-C experiments will lead to continued insights into how to optimize this important approach.

Acknowledgements

We thank Yang Xu for his help in calculating Hi-C dataset reproducibility. We thank Job Dekker, Johan Gibcus, Jon Belton, Houda Belaghzal, Natalia Naumova, Noam Kaplan, and Bryan Lajoie for helpful discussions about concepts discussed here. Comparative Hi-C protocol experiments were in part supported by a Ralph E. Powe Junior Faculty Enhancement Award from Oak Ridge Associated Universities.

References

BELAGHZAL, H., DEKKER, J. & GIBCUS, J. H. 2017. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods,* 123**,** 56-65.

BELMONT, A. S. 2014. Large-scale chromatin organization: the good, the surprising, and the still perplexing. *Curr Opin Cell Biol,* 26**,** 69-78.

BELTON, J. M., MCCORD, R. P., GIBCUS, J. H., NAUMOVA, N., ZHAN, Y. & DEKKER, J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods,* 58**,** 268-76.

BRYANT, H. E. 2012. DNA double-strand break damage and repair assessed by pulsed-field gel electrophoresis. *Methods Mol Biol,* 920**,** 315-21.

BURTON, J. N., ADEY, A., PATWARDHAN, R. P., QIU, R., KITZMAN, J. O. & SHENDURE, J. 2013. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol,* 31**,** 1119-25.

CRANE, E., BIAN, Q., MCCORD, R. P., LAJOIE, B. R., WHEELER, B. S., RALSTON, E. J., UZAWA, S., DEKKER, J. & MEYER, B. J. 2015. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature,* 523**,** 240-4.

CRESPAN, E., CZABANY, T., MAGA, G. & HUBSCHER, U. 2012. Microhomology-mediated DNA strand annealing and elongation by human DNA polymerases lambda and beta on normal and repetitive DNA sequences. *Nucleic Acids Res,* 40**,** 5577-90.

DAVIES, J. O., TELENIUS, J. M., MCGOWAN, S. J., ROBERTS, N. A., TAYLOR, S., HIGGS, D. R. & HUGHES, J. R. 2016. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat Methods,* 13**,** 74-80.

DEKKER, J., BELMONT, A. S., GUTTMAN, M., LESHYK, V. O., LIS, J. T., LOMVARDAS, S., MIRNY, L. A., O'SHEA, C. C., PARK, P. J., REN, B., POLITZ, J. C. R., SHENDURE, J., ZHONG, S. & NETWORK, D. N. 2017. The 4D nucleome project. *Nature,* 549**,** 219-226.

DEKKER, J. & MIRNY, L. 2016. The 3D Genome as Moderator of Chromosomal Communication. *Cell,* 164**,** 1110-1121.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

DILEEP, V., AY, F., SIMA, J., VERA, D. L., NOBLE, W. S. & GILBERT, D. M. 2015. Topologically associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication-timing program. *Genome Res,* 25**,** 1104-13.

DONG, P., TU, X., CHU, P. Y., LU, P., ZHU, N., GRIERSON, D., DU, B., LI, P. & ZHONG, S. 2017. 3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments. *Mol Plant,* 10**,** 1497-1509.

DOSTIE, J., RICHMOND, T. A., ARNAOUT, R. A., SELZER, R. R., LEE, W. L., HONAN, T. A., RUBIO, E. D., KRUMM, A., LAMB, J., NUSBAUM, C., GREEN, R. D. & DEKKER, J. 2006. Chromosome Conformation Capture Carbon Copy (5C): a

massively parallel solution for mapping interactions between genomic elements. *Genome Res,* 16**,** 1299-309.

GABDANK, I., RAMAKRISHNAN, S., VILLENEUVE, A. M. & FIRE, A. Z. 2016. A streamlined tethered chromosome conformation capture protocol. *BMC Genomics,* 17**,** 274.

GIBCUS, J. H. & DEKKER, J. 2013. The hierarchy of the 3D genome. *Mol Cell,* 49**,** 773-82.

GONDOR, A., ROUGIER, C. & OHLSSON, R. 2008. High-resolution circular chromosome conformation capture assay. *Nat Protoc,* 3**,** 303-13.

HAREWOOD, L., KISHORE, K., ELDRIDGE, M. D., WINGETT, S., PEARSON, D., SCHOENFELDER, S., COLLINS, V. P. & FRASER, P. 2017. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. *Genome Biol,* 18**,** 125.

HEITZ, E. 1928. Das Heterochromatin der Moose. *Jahrb Wiss Botanik* 69**,** 762–818.

HOFFMAN, E. A., FREY, B. L., SMITH, L. M. & AUBLE, D. T. 2015. Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J Biol Chem,* 290**,** 26404-11.

HUGHES, J. R., ROBERTS, N., MCGOWAN, S., HAY, D., GIANNOULATOU, E., LYNCH, M., DE GOBBI, M., TAYLOR, S., GIBBONS, R. & HIGGS, D. R. 2014. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet,* 46**,** 205-12.

IMAKAEV, M., FUDENBERG, G., MCCORD, R. P., NAUMOVA, N., GOLOBORODKO, A., LAJOIE, B. R., DEKKER, J. & MIRNY, L. A. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods,* 9**,** 999-1003.

JAGER, R., MIGLIORINI, G., HENRION, M., KANDASWAMY, R., SPEEDY, H. E., HEINDL, A., WHIFFIN, N., CARNICER, M. J., BROOME, L., DRYDEN, N., NAGANO, T., SCHOENFELDER, S., ENGE, M., YUAN, Y., TAIPALE, J., FRASER, P., FLETCHER, O. & HOULSTON, R. S. 2015. Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat Commun,* 6**,** 6178.

KAPLAN, N. & DEKKER, J. 2013. High-throughput genome scaffolding from in vivo DNA interaction frequency. *Nat Biotechnol,* 31**,** 1143-7.

KOLOVOS, P., VAN DE WERKEN, H. J., KEPPER, N., ZUIN, J., BROUWER, R. W., KOCKX, C. E., WENDT, K. S., VAN, I. W. F., GROSVELD, F. & KNOCH, T. A. 2014. Targeted Chromatin Capture (T2C): a novel high resolution high throughput method to detect genomic interactions and regulatory elements. *Epigenetics Chromatin,* 7**,** 10.

LAJOIE, B. R., DEKKER, J. & KAPLAN, N. 2015. The Hitchhiker's guide to Hi-C analysis: practical guidelines. *Methods,* 72**,** 65-75.

LIEBER, M. R. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem,* 79**,** 181-211.

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., SANDSTROM, R., BERNSTEIN, B., BENDER, M. A., GROUDINE, M., GNIRKE,

A., STAMATOYANNOPOULOS, J., MIRNY, L. A., LANDER, E. S. & DEKKER, J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science,* 326**,** 289-93.

LUPIANEZ, D. G., KRAFT, K., HEINRICH, V., KRAWITZ, P., BRANCATI, F., KLOPOCKI, E., HORN, D., KAYSERILI, H., OPITZ, J. M., LAXOVA, R., SANTOS-SIMARRO, F., GILBERT-DUSSARDIER, B., WITTLER, L., BORSCHIWER, M., HAAS, S. A., OSTERWALDER, M., FRANKE, M., TIMMERMANN, B., HECHT, J., SPIELMANN, M., VISEL, A. & MUNDLOS, S. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell,* 161**,** 1012-1025.

M. MANGHI, N. D. 2016. Physics of base-pairing dynamics in DNA. *Physics Reports,* 631**,** 1-41.

MA, H., NASERI, A., REYES-GUTIERREZ, P., WOLFE, S. A., ZHANG, S. & PEDERSON, T. 2015. Multicolor CRISPR labeling of chromosomal loci in human cells. *Proc Natl Acad Sci U S A,* 112**,** 3002-7.

MAKHIJA, E., JOKHUN, D. S. & SHIVASHANKAR, G. V. 2016. Nuclear deformability and telomere dynamics are regulated by cell geometric constraints. *Proc Natl Acad Sci U S A,* 113**,** E32-40.

MASCHER, M., GUNDLACH, H., HIMMELBACH, A., BEIER, S., TWARDZIOK, S. O., WICKER, T., RADCHUK, V., DOCKTER, C., HEDLEY, P. E., RUSSELL, J., BAYER, M., RAMSAY, L., LIU, H., HABERER, G., ZHANG, X. Q., ZHANG, Q., BARRERO, R. A., LI, L., TAUDIEN, S., GROTH, M., FELDER, M., HASTIE, A.,

SIMKOVA, H., STANKOVA, H., VRANA, J., CHAN, S., MUNOZ-AMATRIAIN, M., OUNIT, R., WANAMAKER, S., BOLSER, D., COLMSEE, C., SCHMUTZER, T., ALIYEVA-SCHNORR, L., GRASSO, S., TANSKANEN, J., CHAILYAN, A., SAMPATH, D., HEAVENS, D., CLISSOLD, L., CAO, S., CHAPMAN, B., DAI, F., HAN, Y., LI, H., LI, X., LIN, C., MCCOOKE, J. K., TAN, C., WANG, P., WANG, S., YIN, S., ZHOU, G., POLAND, J. A., BELLGARD, M. I., BORISJUK, L., HOUBEN, A., DOLEZEL, J., AYLING, S., LONARDI, S., KERSEY, P., LANGRIDGE, P., MUEHLBAUER, G. J., CLARK, M. D., CACCAMO, M., SCHULMAN, A. H., MAYER, K. F. X., PLATZER, M., CLOSE, T. J., SCHOLZ, U., HANSSON, M., ZHANG, G., BRAUMANN, I., SPANNAGL, M., LI, C., WAUGH, R. & STEIN, N. 2017. A chromosome conformation capture ordered sequence of the barley genome. *Nature,* 544**,** 427-433.

MCCORD, R. P., NAZARIO-TOOLE, A., ZHANG, H., CHINES, P. S., ZHAN, Y., ERDOS, M. R., COLLINS, F. S., DEKKER, J. & CAO, K. 2013. Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res,* 23**,** 260-9.

MCCORD, R. P., ZHOU, V. W., YUH, T. & BULYK, M. L. 2011. Distant cis-regulatory elements in human skeletal muscle differentiation. *Genomics,* 98**,** 401-11.

MUMBACH, M. R., RUBIN, A. J., FLYNN, R. A., DAI, C., KHAVARI, P. A., GREENLEAF, W. J. & CHANG, H. Y. 2016. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods,* 13**,** 919-922.

NAGANO, T., VARNAI, C., SCHOENFELDER, S., JAVIERRE, B. M., WINGETT, S. W. & FRASER, P. 2015. Comparison of Hi-C results using in-solution versus in-nucleus ligation. *Genome Biol,* 16**,** 175.

NAGANO, T., WINGETT, S. W. & FRASER, P. 2017. Capturing Three-Dimensional Genome Organization in Individual Cells by Single-Cell Hi-C. *Methods Mol Biol,* 1654**,** 79-97.

NAUMOVA, N., IMAKAEV, M., FUDENBERG, G., ZHAN, Y., LAJOIE, B. R., MIRNY, L. A. & DEKKER, J. 2013. Organization of the mitotic chromosome. *Science,* 342**,** 948-53.

OLINS, A. L. & OLINS, D. E. 1974. Spheroid chromatin units (v bodies). *Science,* 183**,** 330-2.

OTTAVIANI, D., LECAIN, M. & SHEER, D. 2014. The role of microhomology in genomic structural variation. *Trends Genet,* 30**,** 85-94.

RAO, S. S., HUNTLEY, M. H., DURAND, N. C., STAMENOVA, E. K., BOCHKOV, I. D., ROBINSON, J. T., SANBORN, A. L., MACHOL, I., OMER, A. D., LANDER, E. S. & AIDEN, E. L. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell,* 159**,** 1665-80.

ROBINETT, C. C., STRAIGHT, A., LI, G., WILLHELM, C., SUDLOW, G., MURRAY, A. & BELMONT, A. S. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol,* 135**,** 1685-700.

SCHMITT, A. D., HU, M. & REN, B. 2016. Genome-wide mapping and analysis of chromosome architecture. *Nat Rev Mol Cell Biol,* 17**,** 743-755.

SCHWARZER, W., ABDENNUR, N., GOLOBORODKO, A., PEKOWSKA, A., FUDENBERG, G., LOE-MIE, Y., FONSECA, N. A., HUBER, W., HAERING, C. H., MIRNY, L. & SPITZ, F. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature,* 551**,** 51-56.

SHACHAR, S., VOSS, T. C., PEGORARO, G., SCIASCIA, N. & MISTELI, T. 2015. Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping. *Cell,* 162**,** 911-23.

SIMONIS, M., KLOUS, P., SPLINTER, E., MOSHKIN, Y., WILLEMSEN, R., DE WIT, E., VAN STEENSEL, B. & DE LAAT, W. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet,* 38**,** 1348-54.

SMITH, E. M., LAJOIE, B. R., JAIN, G. & DEKKER, J. 2016. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *Am J Hum Genet,* 98**,** 185-201.

TAKATA, H., HANAFUSA, T., MORI, T., SHIMURA, M., IIDA, Y., ISHIKAWA, K., YOSHIKAWA, K., YOSHIKAWA, Y. & MAESHIMA, K. 2013. Chromatin compaction protects genomic DNA from radiation damage. *PLoS One,* 8**,** e75622.

TEIF, V. B., MALLM, J. P., SHARMA, T., MARK WELCH, D. B., RIPPE, K., EILS, R., LANGOWSKI, J., OLINS, A. L. & OLINS, D. E. 2017. Nucleosome repositioning during differentiation of a human myeloid leukemia cell line. *Nucleus,* 8**,** 188-204.

URSU, O., BOLEY, N., TARANOVA, M., WANG, Y. X. R., YARDIMCI, G. G., NOBLE, W. S. & KUNDAJE, A. 2018. GenomeDISCO: A concordance score for chromosome conformation capture experiments using random walks on contact map graphs. *bioRxiv*, 181842.

WINGETT, S. 2017. *Illumina Patterned Flow Cells Generate Duplicated Sequences* [Online]. Available: https://sequencing.qcfail.com/articles/illumina-patterned-flow-cells-generate-duplicated-sequences/ [Accessed].

ZHANG, Y., MCCORD, R. P., HO, Y. J., LAJOIE, B. R., HILDEBRAND, D. G., SIMON, A. C., BECKER, M. S., ALT, F. W. & DEKKER, J. 2012. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell,* 148**,** 908-21.

ZHAO, Z., TAVOOSIDANA, G., SJOLINDER, M., GONDOR, A., MARIANO, P., WANG, S., KANDURI, C., LEZCANO, M., SANDHU, K. S., SINGH, U., PANT, V., TIWARI, V., KURUKUTI, S. & OHLSSON, R. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet,* 38**,** 1341-7.

CHAPTER II: CONSTRICTED MIGRATION IS ASSOCIATED WITH STABLE 3D GENOME STRUCTURE DIFFERENCES IN MELANOMA CELLS

A version of this chapter was originally published by Rosela Golloshi, Trevor F. Freeman, Priyojit Das, Thomas Isaac Raines, Rebeca San Martin, Christopher Playter, Delaney Thurston and Rachel Patton McCord.

Golloshi, R., Freeman, T. F., Das, P., Raines, T. I., Martin, R. S., Playter, C., Thurston, D. & McCord, R. P. 2020. Constricted migration is associated with stable 3D genome structure differences in melanoma cells. bioRxiv, 856583.

My contributions to this work include: (1) conceived the project, (2) designed the experiments, (3) wrote the paper, (4) performed genomic, cell culture, and imaging experiments and (5) bioinformatics data analysis.

Abstract

To spread from a localized tumor, metastatic cancer cells must squeeze through constrictions that cause major nuclear deformations. Since chromosome structure affects nucleus stiffness, gene regulation and DNA repair, here we investigate how confined migration affects or is affected by 3D genome structure. Using melanoma (A375) cells, we identify phenotypic differences in cells that have undergone multiple rounds of constricted migration. These cells display a stably higher migration efficiency, elongated morphology, and differences in the distribution of Lamin A/C and heterochromatin. Using Hi-C, we observe differences in chromosome spatial compartmentalization specific to cells that have passed through constrictions and related alterations in expression of genes associated with migration and metastasis. These sequentially constricted cells also

show more nuclear deformations and altered behavior in a 3D collagen matrix. Our observations reveal a relationship between chromosome structure changes, metastatic gene signatures, and the altered nuclear appearance of aggressive melanoma.

Introduction

Despite significant improvements in the diagnosis and treatment of cancer, most patients with advanced metastatic disease face a terminal illness incurable by current therapeutic methods. The dissemination of cancer cells from the primary tumor and the formation of new tumor colonies at distant sites involves a complex multi-step invasion-metastasis cascade (Lambert et al., 2017, Gupta and Massague, 2006, Chambers et al., 2002, Fidler, 2003). To metastasize, cancer cells must squeeze through constrictions of the extracellular matrix or endothelial lining that are much smaller than their nucleus, causing major nuclear deformations. While the cell membrane and cytoplasm are quite elastic, the ability of the cell's nucleus to withstand large deformations and squeeze through these small spaces is limited by its size and stiffness, posing challenges to confined migration (Davidson et al., 2014, Fu et al., 2012, Wolf et al., 2013, Friedl et al., 2011). Nuclear stiffness and deformability depend on two major components: lamin proteins (Lamin A/C and Lamin B) and the chromatin (Stephens et al., 2017, Harada et al., 2014, Davidson et al., 2014).

Lamin A/C expression and its stoichiometric relation to Lamin B contributes to nuclear mechanical properties (Harada et al., 2014, Swift et al., 2013). Low expression levels of Lamin A/C lead to decreased nuclear stiffness and increased deformability (Lammerding

et al., 2006). Likely resulting from increased nuclear flexibility, decreased expression of Lamin A/C has been shown to promote the ability of mouse embryonic fibroblasts (MEFs) to undergo constricted migration (Davidson et al., 2014). Conversely, neutrophils show decreased constricted migration ability when expression of Lamin A/C is increased (Davidson et al., 2014, Rowat et al., 2013). Given this relation between lamin content and confined migration, it is not surprising that low levels of Lamin A/C are correlated with the increased aggressiveness of some cancers hose localized to more accessible euchromatin tend to be active or poised for quick transcriptional activation (Gaspar-Maia et al., 2011, van Steensel and Belmont, 2017). Chromatin state has also been shown to influence nuclear physical properties: increasing heterochromatin increases nuclear stiffness while increasing euchromatin decreases stiffness and increases nuclear deformability (Stephens et al., 2017, Stephens et al., 2018). Such global changes in chromatin state can also influence confined migration. For example, increasing euchromatin or decreasing heterochromatin inhibits mouse melanoma cell migration (Gerlitz and Bustin, 2010, Maizels et al., 2017, Segal et al., 2018, Panagiotakopoulou et al., 2016, Krause et al., 2019). Mounting evidence also points to the effect that extracellular physical forces have on nuclear mechanics and chromosome structure. Nuclei aspirated into narrow micropipettes exhibit stretching of chromatin domains that is sometimes irreversible (Irianto et al., 2017b). External cellular forces can also affect gene expression through direct chromatin deformations (Tajik et al., 2016). Despite such connections between chromatin structure, nucleus deformation, and nucleus mechanics,

it is not fully understood whether the 3D organization of the genome affects or is affected by confined migration.

Ever-improving microscopic techniques and the development of chromosome conformation capture approaches have revolutionized our understanding of 3D genome architecture in the past decade (Dekker et al., 2002, Bickmore and van Steensel, 2013, Rowley and Corces, 2018, Pombo and Dillon, 2015, Abbas et al., 2019). Chromosomes are folded at different length scales into loops, topologically associating domains (TADs), active and inactive (A/B) compartments, and chromosomal territories. How these layers of structure are affected by physical nucleus shape changes such as the kind experienced during constricted migration remains unknown. Previous work has shown that 3D chromosome organization changes correlate with the progression of cancer (Taberlay et al., 2016, Barutcu et al., 2015, Zhou et al., 2019) and that both the formation and downstream effect of chromosomal translocations in cancer progression is influenced by 3D genome structure (Zhang et al., 2012, Hnisz et al., 2016). In addition, some changes have been noted in the 3D genome structures of neutrophils that have undergone constricted migration (Jacobson et al., 2018). Taken together, previous evidence leads us to hypothesize that chromosome structure may influence a cancer cell's ability to undergo confined migration, that changes in chromosome structure could be caused by nuclear deformation, and that such chromosome structure changes could influence gene regulation and cancer cell phenotype. In this study, we seek to unite these previous ideas

and investigate whether certain 3D genome structures of cancer cells facilitate or are changed by constricted migration.

Here, we use invasive human melanoma cells (A375) to investigate the properties of the nucleus and 3D chromosome structure that accompany confined migration. We find that A375 cells that have migrated numerous times through tight constrictions show increased migration efficiency and are phenotypically distinct from cells that do not pass through constrictions. These repeatedly constricted cells exhibit specific and stable differences in 3D genome organization, particularly at the level of A/B compartmentalization, lamin localization, and gene expression patterns. Our results suggest that 3D genome architecture correlates with the ability of cancer cells to undergo constricted migration. Our experimental data and modeling suggest that 3D genome and nucleus differences could arise from both initial heterogeneity in the melanoma cell population and changes induced by constricted migration itself.

Results

Melanoma cells exhibit an increased migration efficiency and cell morphology differences after multiple rounds of constricted migration

To investigate the relationship between 3D genome organization and constricted migration, we performed sequential rounds of migration of A375 cells through 5μm transwell filters (Figure 8A, see Materials and Methods for a detailed experimental design). The diameter of A375 nuclei ranges from 8-16μm (Figure 9A), so passing

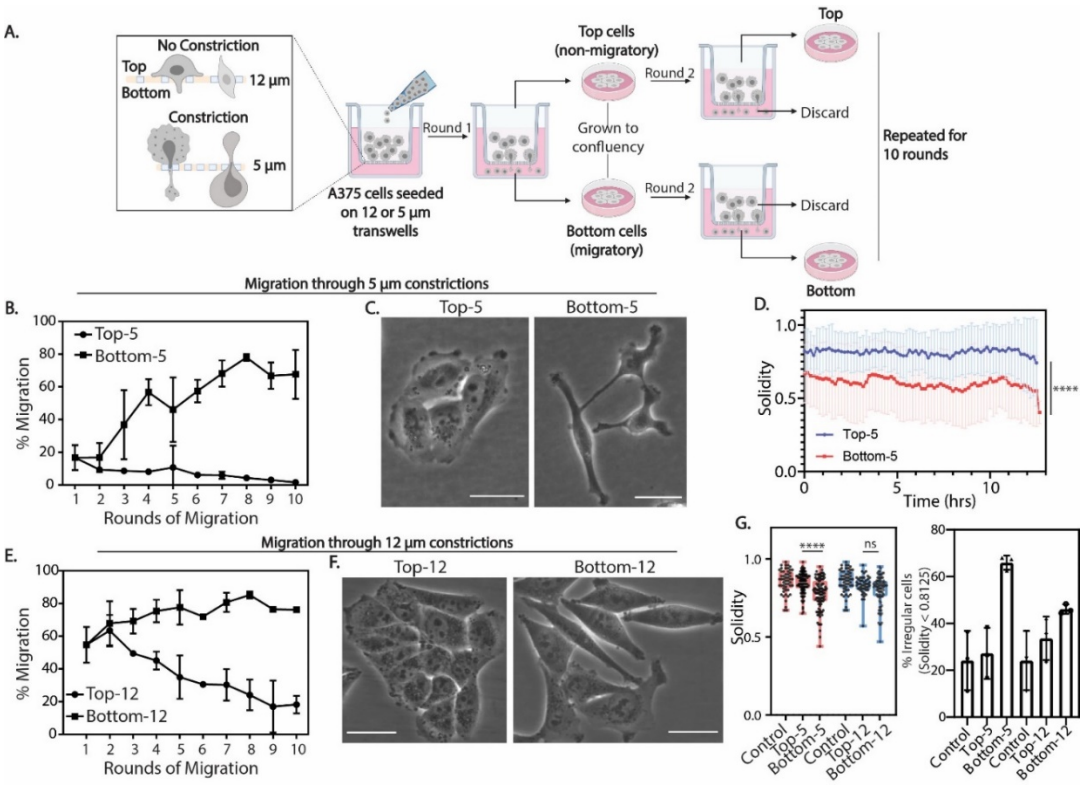


Figure 8. **A375 cells exhibit morphological differences and increase in migration efficiency after sequential rounds of constricted migration.**

(A) Layout of sequential transwell migration experiment with A375 cells. (B) Migration efficiency (% of cells that migrated through filter pores in each round) of A375 cells during 5µm constricted migration. Error bars = mean ± SD for n = 3 biological replicates. (C) Phase contrast imaging of A375 cells that have not (Top-5) or have (Bottom-5) undergone 5µm constricted migration. (D) Solidity measurements of A375 cells undergoing live-cell migration for 13 hours on a 2D surface (1 indicates ideally round cell shapes and lower values indicate more irregularities and protrusions). Error bars = mean ± SD for n = 76 Top-5 cells and n = 77 Bottom-5 cells ($p < 0.0001$, two-tailed t-test). (E) Migration efficiency of A375 cells undergoing 12µm transwell migration. (F) Phase contrast imaging of A375 cells that have not (Top-12) or have (Bottom-12) undergone 12µm transwell migration. (G) (Left panel) Quantification of solidity measurements between A375 cells that have undergone migration through 5µm and 12µm pores. Box and whiskers plot showing all data points in Control (n = 52 cells), Top-5 (n = 99 cells), Bottom-5 (n = 98 cells), Top-12 (n = 43 cells) and Bottom-12 (n = 82 cells) (**** $p < 0.0001$, ns $p = 0.0726$, two-tailed t-test). (Right panel) % of A375 cells displaying irregular morphology (Solidity < 0.8125, first quartile of Control). Scale bars (white) indicate a length of 25µm.

(A) Minor axis (minimum diameter) measurements of A375 cell nuclei based on confocal z-stacks. N=85. (B) Migration efficiency of each biological replicate for 10 rounds of 5μm transwell migration in Top (left) and Bottom (right) cells. (C) Phase contrast images of A375 cells exposed to 5μm transwell migration for 10 and 20 rounds. (D) Aspect ratio (left graph) and solidity (right graph) measurements of cells in A375 subpopulations that have undergone 10 rounds (Top-5 and Bottom-5) and 20 rounds (Top20-5 and Bottom20-5) of migration through 5μm constrictions (**** $p < 0.0001$; Control (n = 52), Top-5 (n = 99), Top20-5 (n = 93), Bottom-5 (n = 98), Bottom20-5 (n = 48); two – tailed t-test). (E) Migration efficiency of A375 cells after 5 passages of continuous culture (left graph) and after one cycle of freeze - thawing (right graph) in two replicates. (F) Confocal images of Phalloidin (green) and DAPI (blue) immunostained cells that have undergone constricted migration. Graphs represent area of the cell with phalloidin staining passing a threshold (*** $p = 0.0003$, **** $p < 0.0001$; two-tailed t-test) and fraction of cells that display phalloidin staining area < 15 μm2. (G) Migration efficiency of each replicate for 10 rounds of 12-μm transwell migration. (H) Phase contrast images of A375 cells exposed to 12μm transwell migration. Scale bars (white) indicate a length of 50μm.

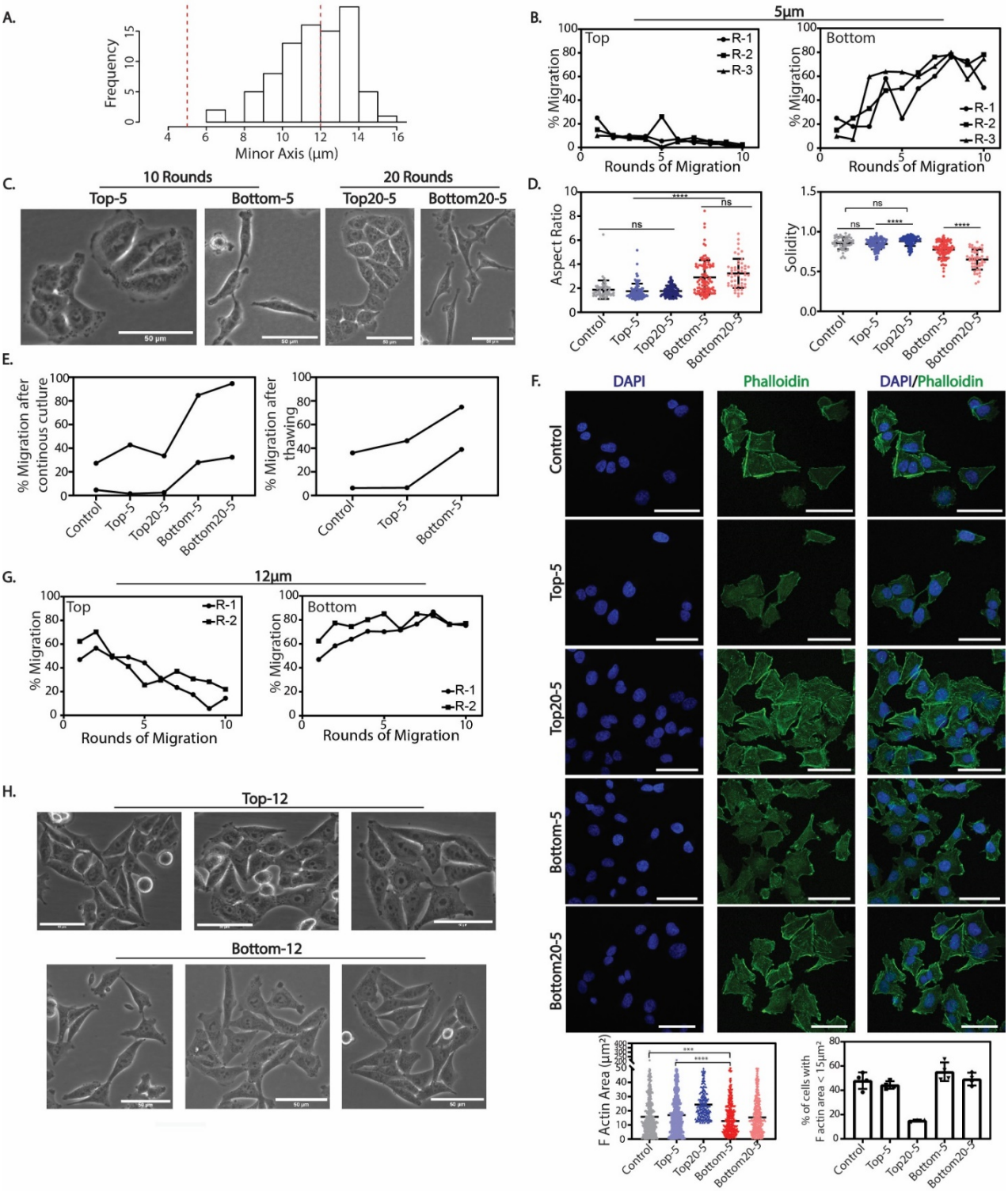


Figure 9. A375 cells exhibit persistent increase in migration efficiency and morphological changes after 5μm constricted migration.

through these 5μm pores requires nuclear constriction and leads to nuclear deformations. Interestingly, cells that successfully undergo constricted migration through 5μm pores in each round migrate more efficiently in subsequent rounds of the same Transwell assay. The population of cells that have passed through the 5μm pores for 10 consecutive rounds (Bottom-5) migrate at high efficiency (~70%) as compared to the initial population (20%) (Figures 8B and 9B). In contrast, the population of cells that failed to undergo constricted migration through 5μm pores in the first round, shows a progressive decrease in migratory efficiency, approaching 0% migration after 10 consecutive rounds (Top-5) (Figures 8B and 9B).

These differences in migration efficiency correspond to other notable differences in cell phenotype. Phase contrast imaging of the A375 subpopulations shows that Bottom-5 cells exhibit a decrease in cell-cell adhesions and an elongated cell body (suggesting a more mesenchymal phenotype) when compared to Top-5 cells (Figures 8C, 9C and 9D). We extended the sequential transwell assay up to 20 rounds and observe similar elongated cell bodies in A375 cells that have undergone 20 rounds of 5μm transwell migration (Bottom20-5) (Figure 9C and 9D). These morphological characteristics and migration efficiency differences are maintained even after continuous culture (5 passages) or multiple freeze-thaw cycles, indicating stable phenotypes among Top-5 and Bottom-5 cells (Figure 9E). Live cell imaging of Top-5 and Bottom-5 cells migrating on a 2D surface shows that Bottom-5 cells maintain their elongated shapes and display lobopodia/lamellipodia-like migration, while Top-5 cells display amoeboid-like migration

over a 13-hr migration period (Figure 8D and Movie S1 and S2). We then investigated whether loss of cell-cell adhesion in Bottom-5 cells corresponded to differences in focal adhesions by phalloidin stain (Figure 9F). Bottom-5 cells exhibit a decrease in the area covered by bright F-actin stain compared to Control and Top cells (Top-5 and Top20-5) (Figure 9F, Focal adhesion area quantification). This suggests a decrease in focal adhesion size, which has been previously reported to correlate with increase in migration speed and cell motility (Kim and Wirtz, 2013).

These results indicate that the morphological and/or genetic differences associated with passing through constrictions could promote migration in consecutive rounds. These phenotypes are consistent with previous studies showing that the degree of invasiveness is often correlated with an increase of cell shape irregularity (Lyons et al., 2016, Machesky, 2008).

It has been previously reported that A375 cells are heterogeneous in their metastatic ability (Kozlowski et al., 1984), so it is possible that the differences we observe are associated only with ability to migrate and are not specific to passing through a constriction. We repeated the sequential migration experiment using transwell filters with 12μm pores (Figures 8E and 9G). Migration through 12μm pores does not pose a major spatial barrier for the nucleus since the minimum diameter of the majority of A375 cells is smaller than the 12μm diameter of the pore (Figure 9A). Accordingly, A375 cells migrate

at a much higher rate (~ 55-70%) in the first round of the 12μm transwell assay (Figures 8E and 9G).

Cells that were successful at migrating through 12μm pores for 10 rounds (Bottom-12) displayed only a minimal increase in migratory efficiency as compared to the more dramatic increase shown by Bottom-5 cells after each consecutive round (Figures 8E and 9G). In contrast, cells that failed to migrate through the 12μm pores (Top-12), displayed a significant decrease in migration efficiency after 10 sequential rounds of migration (Figures 8E and 9G).

Phase contrast imaging revealed that Top-12 cells display high cell-cell adhesion and more epithelial like phenotypes (similar to Top-5) when compared to Bottom-12 cells, which show a decrease in cell-cell adhesion and more elongated phenotypes (Figure 8F and Figure 9H). However, when quantifying solidity or aspect ratio (Figure 8G) for all A375 sub-populations, it becomes clear that Bottom-5 cells exhibit more dramatic elongated phenotypes (~65% of cells with solidity < 0.8125) compared to Bottom-12 cells (~ 45% of cells with solidity < 0.8125) which show no significant difference compared to the Control, Top-5 and Top-12 cells (Figure 8G). When challenged to migrate through 5-μm pores, Bottom-12 cells showed a migration efficiency similar to the control population (~ 3-20 %). These unconstricted migration results indicate that there is heterogeneity in the initial A375 population, with some cells unable to migrate even through large pores (Top-12). But, simply being able to migrate is not enough to guarantee efficient migration through

constricted spaces, indicating that additional features are associated with constricted migration. Given the large degree of nuclear squeezing necessary to undergo constricted migration, we next turned to investigating the organization of the nucleus and chromosome structure of these different subgroups of A375 cells.

Irregular distribution of Lamin A/C and peripheral localization of heterochromatin correlate with an increased constricted migration proficiency

Changes in nuclear morphology are consistently used as a diagnostic tool in cancer progression (Kadota et al., 2012) and previous studies report nuclear envelope rupture and blebbing as a result of constricted migration (Denais et al., 2016, Raab et al., 2016) (Irianto et al., 2017a, Davidson and Lammerding, 2014). Thus, we next investigated whether nuclear morphology changes accompany the phenotype differences we observed among cell populations after sequential constricted migration. We used immunofluorescence imaging to observe nucleus morphology, heterochromatin distribution (H3K9me3), and Lamin A/C levels and distribution 18 hours after sequential constricted migration. Analyzing maximum projected images, we find that Bottom-5 nuclei more often exhibit elongated morphologies (~ 40% of population) than Control (25%), Top-5 (20%) and Bottom-12 nuclei (~ 15%) (Figure 10A). When statistically comparing these cell populations, only Bottom-12 has a significantly different mean aspect ratio compared to Bottom-5. This observation indicates the inherent variability in nuclear morphology that already exists in the original population and that variation is still seen in each subpopulation after migration. The differences after constricted migration lie in the extreme tail of the distribution rather than being observed in the average.

(A) Aspect ratio quantification of A375 maximum projected nuclei exposed to transwell migration (*** p = 0.0002, **** p < 0.0001, two-tailed t-test; Control = 59, Top-5 = 63, Bottom-12 = 80, Bottom-5 = 64, Bottom20-5 = 52 nuclei). (B) Nuclear shape measurements of 3D reconstructed A375 nuclei. The number of nuclei quantified remains the same as in A. (C) Lamin A/C mean intensity in A375 cells 18 hrs after migrating through 5μm transwell chambers. Number of nuclei quantified remains the same as in A. (D) Nuclear shape measurements of 3D reconstructed nuclei at 3 and 5hrs after migrating through 5μm (Bottom-5; n = 23 for 3hrs and n = 26 for 5hrs) and 12μm (Bottom-12; n = 19 for 3hrs and n = 22 for 5hrs) transwell pores. ** p = 0.0043; two-tailed t-test. (E) Top Panel: DAPI (blue) and Lamin A/C (green) immunostained nuclei of A375 cells that have migrated through 5 and 12μm transwell pores for 3 and 5 hrs. Bottom Panel: Cumulative distribution plots of Lamin A/C coefficient of variance in maximum projected nuclei. Number of nuclei remains the same as in D. (F) Lamin A/C mean intensity of A375 cells that have undergone migration through 5 and 12μm transwell chambers for 3 and 5 hours. Number of nuclei quantified is the same as D. ** p = 0.0058; two-tailed t-test. (G-J) Western blot analysis of A375-Control, Top-5 and Bottom-5 cells probing for Lamin A/C, Lamin B1, phospho-Lamin A/C (Ser22) and H3K9me3. Quantification shown above each blot. No significant differences found in pairwise comparisons by two-tailed t-test. All scale bars (white) in this figure = 5μm.

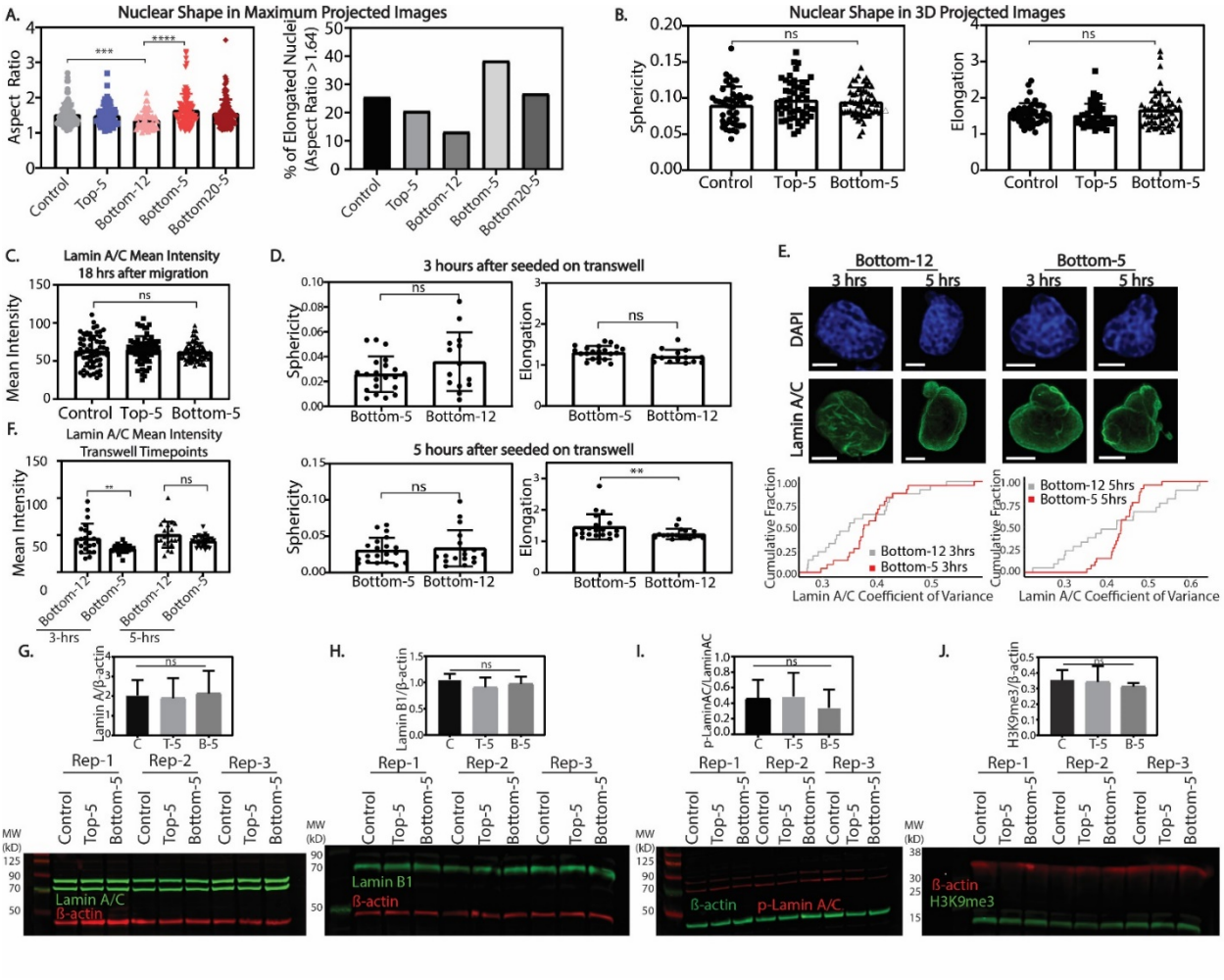


Figure 10. Changes in nuclear morphology after constricted migration are not dependent on Lamin A/C or H3K9me3 protein levels.

We also investigated nuclear morphology features of 3D projections (sphericity and elongation). These also show no significant differences in mean, but, again, there are a larger number of Bottom-5 cells that are outliers from the central tendency with more elongated nuclei (Figure 10B). These results are in agreement with previous reports showing that the morphological changes with constriction are reversible (Irianto et al., 2017a), but indicate that outliers with elongated nuclei are more likely to be observed among the cell population that has gone through constricted migration.

Given their previously reported roles in mechanosensing and modulating nucleus stiffness, we next investigated whether differences in Lamin A/C and constitutive heterochromatin (H3K9me3) levels or distribution exist among A375 subpopulations with different migratory ability (Figure 11A). High resolution confocal microscopy of immunostained nuclei showed no difference in the overall level of Lamin A/C in the cell subpopulations that had undergone constricted migration vs. those that had not (Figure 10C). The lack of difference in Lamin A/C protein levels was confirmed by western blot (Figure 10G). These unchanged Lamin A/C levels echo recent observations reported in mechanically stretched cells (Nava et al., 2020) and are in contrast to other reports that show an association between lower levels of Lamin A/C and increased constricted migration ability (Harada et al., 2014). We find that, like total Lamin A/C levels, Lamin B1 protein levels and H3K9me3 levels do not change across the subpopulations (Figure 10H and J). Lamin A/C can be regulated not only in abundance, but also by phosphorylation,

(A) A375 nuclei stained with DAPI (blue), Lamin A/C (green) and H3K9me3 (red) in the indicated migration conditions. Images shown represent a maximum projection of z-stacks. (B) Snapshots of 3D reconstructed z-stacks of A375 nuclei stained with Lamin A/C. For full 3D visualization see Movie S3 and S4. White arrows indicate areas devoid of Lamin A/C. (C) Fraction of A375 cells displaying irregular patterns of Lamin A/C, defined by visual inspection, in Control (n = 59), Top-5 (n = 63), Bottom-12 (n = 80) and Bottom-5 (n = 64) cells. Two replicates shown in the figure; lines connect A375 conditions within the same replicate. (D) Cumulative distribution plots of Lamin A/C coefficient of variance in all indicated A375 subpopulations. Kolmogorov-Smirnov test p-values indicated in each plot. (E) Central orthogonal slices of Lamin A/C stained nuclei in Top-5 and Bottom-5 cells. (F) A375 nuclei stained with DAPI (blue) and H3K9me3 (red) for indicated A375 cell subpopulations. Images represent maximum projection of z-stacks. (G) Radial intensity distribution of H3K9me3 signal in the maximum projected nuclei of A375 subpopulations. Top graph: H3K9me3 radial intensity distribution in Control (n = 59), Top-5 (n = 63) and Bottom-5 (n = 64) nuclei (** $p = 0.0069$, two-tailed t-test). Bottom graph: H3K9me3 radial intensity distribution in Control (n = 59), Bottom-12 (n = 80) and Bottom-5 (n = 64) nuclei (**** $p < 0.0001$, ** $p = 0.0069$; two-tailed t-test). Error bars = mean ± SD. (H) Intensity line scans across major and minor axes of Lamin A/C and H3K9me3 central slices in A375 Control (n = 59), Top-5 (n = 63) and Bottom-5 (n = 64) nuclei. Error bars = mean ± SD. All scale bars (white) in this figure = 5μm.

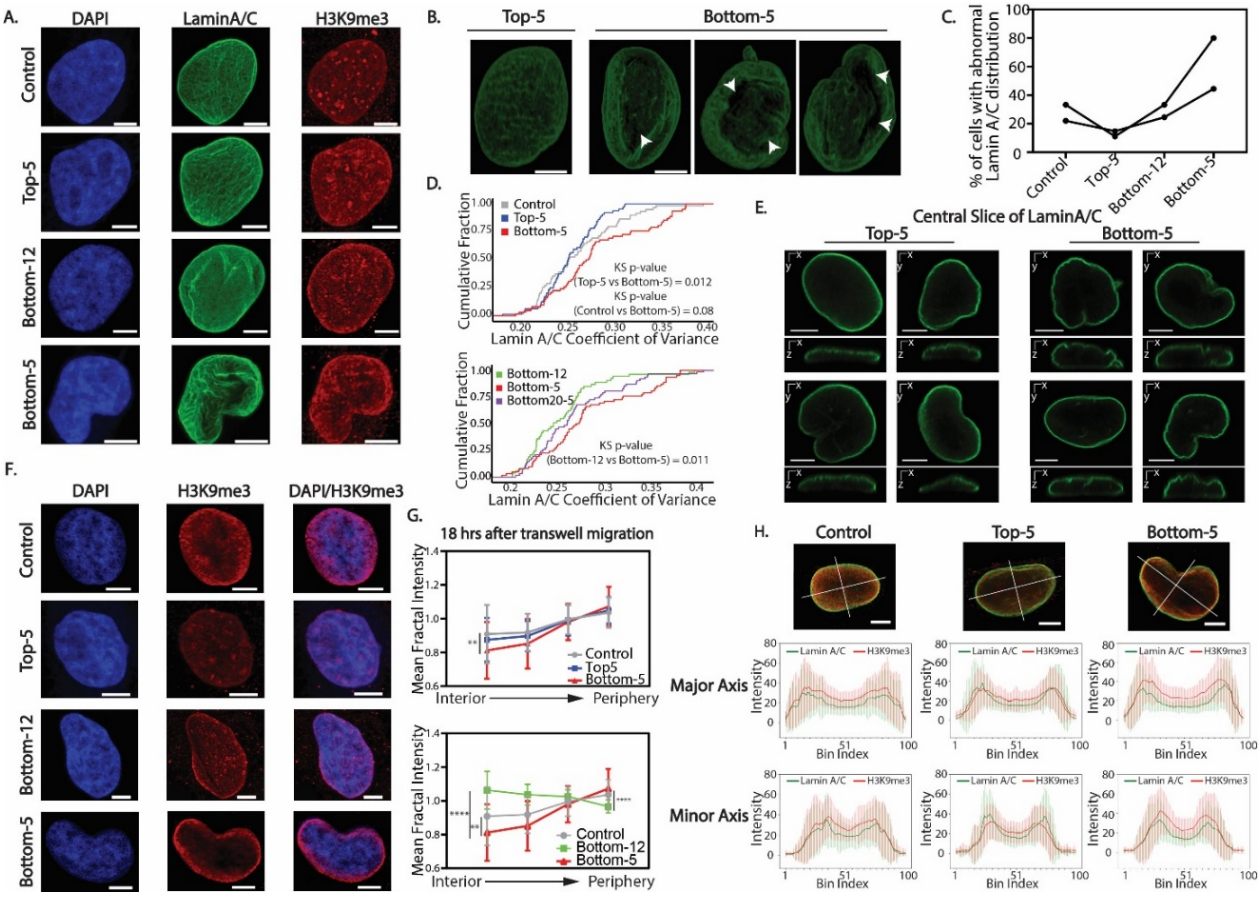


Figure 11. Irregular distribution of Lamin A/C and heterochromatin upon constricted migration.

which can affect its solubility and distribution (Torvaldson et al., 2015). But we also see no difference in levels of Lamin A/C phosphorylated at Ser22 (Figure 10L).

Though total levels of Lamin A/C were unchanged, we noted that the distribution of Lamin A/C appeared to be different after constricted migration. A375 Control, Top-5 and Bottom-12 nuclei showed a uniform distribution of Lamin A/C throughout the nuclear envelope (Figure 11A, middle panel). In contrast, Bottom-5 nuclei display an altered distribution of Lamin A/C with certain regions of the nuclear envelope devoid of it (Figures 2A and 2B, indicated by white arrows, Movie S3 and S4). By visual inspection, we found that ~ 50% of Bottom-5 nuclei show an abnormal distribution of Lamin A/C when compared to ~ 25 % of Bottom-12 nuclei (Figure 11C) suggesting that differences in Lamin A/C distribution are specific to constricted migration.

To quantify this difference more systematically, we compared the coefficient of variance of the Lamin A/C signal across each nucleus between the A375 subpopulations. We find that Bottom-5 nuclei exhibit higher Lamin A/C coefficient of variance when compared to Control, Top-5 and Bottom-12 cells (Figure 11D). This variation in Lamin A/C pixel intensity could arise due to an increase in nuclear envelope invaginations and wrinkling in Bottom-5 nuclei as observed from central orthogonal slices of Lamin A/C stained nuclei (Figure 11E).

Like Lamin A/C, we find that the distribution of heterochromatin is altered in A375 cells that have undergone constricted migration. Heterochromatic foci are randomly distributed throughout the nucleus in Control, Top-5, and Bottom-12 cells (Figure 11A) with some foci localized to the nuclear periphery. In contrast, heterochromatin foci in Bottom-5 nuclei are more dispersed and the majority of H3K9me3 signal is localized to the nuclear periphery (Figures 11A and 11F). Indeed, quantifying the radial distribution of H3K9me3 signal in the nucleus reveals that Bottom-5 nuclei display a lower H3K9me3 intensity in the interior of the nucleus when compared to Control and Top-5 nuclei (Figure 11G, top graph). These differences are very evident when compared to Bottom-12 cells which display a random distribution of H3K9me3 foci across the nucleus (Figure 11G, bottom graph). Alterations in H3K9me3 and Lamin A/C distributions are also evident in line scans across central slices of the nucleus (Figure 11H). Bottom-5 cells exhibit a significant increase in heterochromatin localization to the nuclear periphery when compared to Control and Top-5 cells along both the major and minor axes. Interestingly, this analysis also reveals a lower intensity of Lamin A/C at the nuclear periphery of Control and Bottom-5 cells as compared to Top-5 cells, specifically across the minor (narrower) axis of the nucleus.

All effects on nuclear morphology and Lamin A/C distribution discussed thus far are observed persisting 18 hours after the final round of constricted migration. To determine whether transient effects of constricted migration on nuclear morphology and Lamin A/C distribution were different or more dramatic, we performed transwell migration of Bottom-

12 and Bottom-5 cells for 3 and 5 hours after seeding on the transwell chambers (pore sizes 12μm and 5μm respectively). After 3 hours of migration, there is no difference in the average nucleus sphericity and elongation between constricted (5μm) and unconstricted (12μm) migrating cells (Figure 10D, top panel). In contrast, at 5 hours, the nuclei of Bottom-5 cells are on average more elongated compared to Bottom-12 nuclei (Figure 10D, bottom panel). The Lamin A/C distribution of Bottom-5 cells is more variable than Bottom-12 at both 3 and 5 hours (Figure 10E). While we see no difference in overall Lamin A/C levels among the subpopulations 18 hours after migration, there could be a minor transient difference in Lamin A/C- we observe a lower Lamin A/C intensity in Bottom-5 cells that have undergone migration for 3 hours as compared to Bottom-12 cells, but this difference is no longer present at 5 hours into the transwell migration experiment (Figure 10F). Overall, our microscopic analyses imply that, in A375 cells, constricted migration is associated with alterations in the distribution of Lamin A/C and heterochromatin that are visible at even 18 hours after migration, but is not dependent on changes in total Lamin A/C or heterochromatin levels.

Given the variability in Lamin A/C expression across different cell lines and cancer types, we repeated these sequential migration and imaging experiments in MDA-MB-231 cells (Figure 12). These cells exhibit a higher initial ability to migrate through 5μm constrictions when compared to A375 cells. MDA-MB-231 cells that migrate through 5μm pores for 10 rounds (Bottom-5) show an increase in migration efficiency, like A375 cells, though the change is not as dramatic (Figure 12A). When challenged to migrate through large 12

(A) Transwell migration efficiency of MDA-MB-231 cells through 5 and 12μm pore sizes for 10 transwell migration rounds. Error bars = mean ± SD of 5μm (3 biological replicates) and 12μm (2 biological replicates) transwell assays. (B) Phase contrast images of MDA-MB-231 cell subpopulations. Scale bars = 50μm. (C) Left: Aspect ratio of MDA-MB-231 cells for Control (n = 142), Top-5 (n = 393) and Bottom-5 (n = 621) (**** $p < 0.0001$; two-tailed t-test). Right: Fraction of MDA-MB-231 cell subpopulations with an aspect ratio > than 3rd quartile of Control. (D) Maximum projected images of MDA-MB-231 nuclei immunostained with DAPI (blue), Lamin A/C (green) and H3K9me3 (red). Scale bars = 5μm. (E) Left graph: Aspect ratio of MDA-MB-231 nuclei in Control (n = 49), Top-12 (n = 47), Top-5 (n = 49), Bottom-12 (n = 46) and Bottom-5 (n = 47) (** $p = 0.0038$; two-tailed t-test). Right graph: Fraction of MDA-MB-231 nuclei that have an aspect ratio > 3rd quartile of control. (F) Cumulative distribution plot of Lamin A/C coefficient of variance in Control, Top-5 and Bottom-5 MDA-MB-231 cells. Number of nuclei quantified remains the same as in E. KS-test shows no significant difference between conditions. (G) Fraction of MDA-MB-231 nuclei displaying irregularities in Lamin A/C distribution by visual inspection. (H) Line scan intensity for H3K9me3 (red) and Lamin A/C (green) across major and minor axis of central slice images of MDA-MB-231 cells. Number of nuclei quantified remains the same as in E. Error bars = mean ± SD. (I-J) Western blot analysis and quantification of MDA-MB-231 Control, Top-5 and Bottom-5 cells probing for Lamin A/C and H3K9me3. Error bars = mean ± SD. No significant differences found in pairwise comparisons by two-tailed t-test

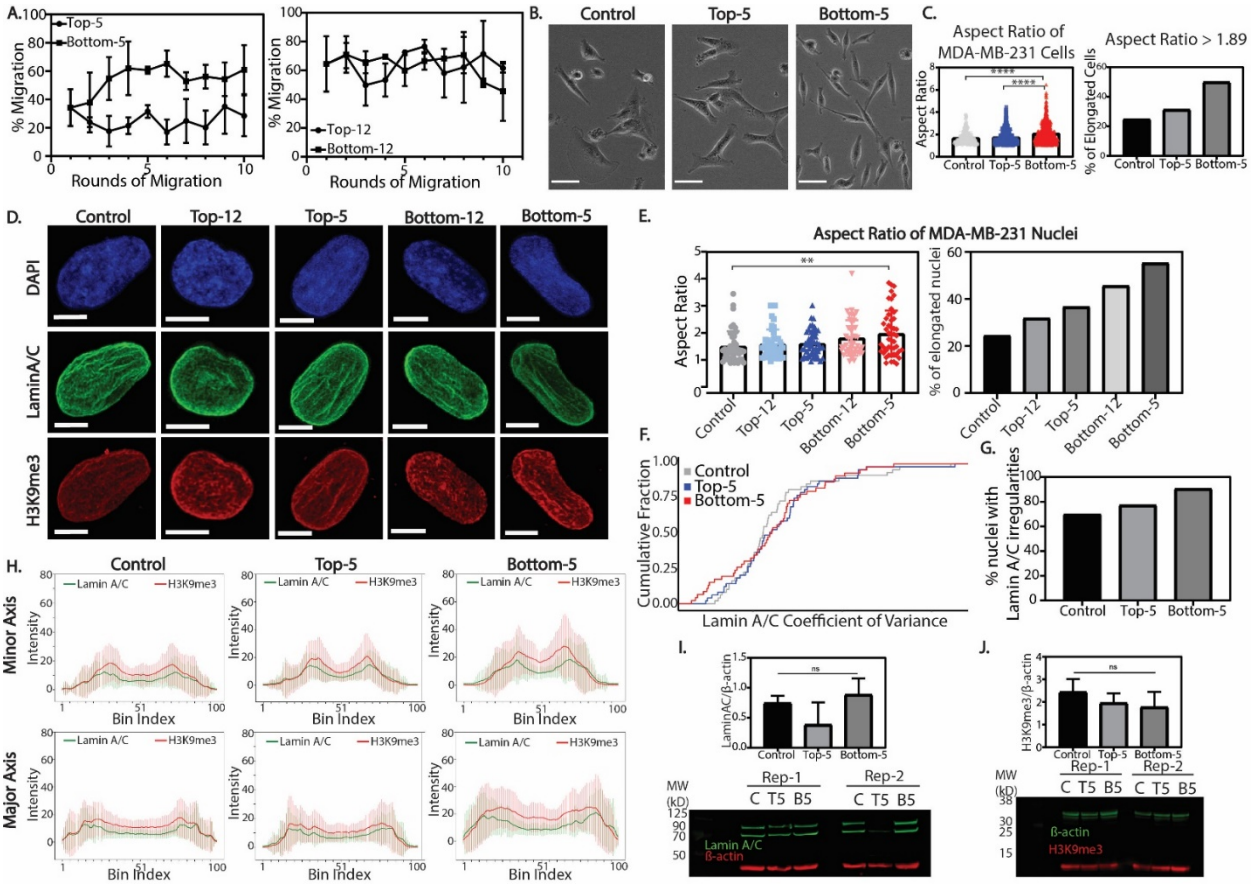


Figure 12. Sequential rounds of constricted migration result in two distinct populations of MDA-MB-231 cells.

micron pores, MDA-MB-231 Top-12 and Bottom-12 cells show no significant separation in their migratory efficiency, in contrast to A375 cells, which had a subpopulation that could not migrate through even large pores. This indicates that there is less initial heterogeneity among the MDA-MB-231 cells in their ability to migrate. Indeed, whether or not they have undergone constricted migration, MDA-MB-231 cells exhibit more mesenchymal phenotypes with elongated cell bodies and low cell-cell adhesions (Figure 12B). However, like A375 cells, the MDA-MB-231 cells that undergo constricted migration for 10 rounds (Bottom-5) and increase their migratory efficiency also show more cell elongation (Figure 12C).

We then investigated whether Lamin A/C and heterochromatin organization in the nucleus were affected by constricted migration in MDA-MB-231 cells (Figure 12D). The nuclei of MDA-MB-231 cells are generally elongated compared to A375 cells. However, about 60% of the Bottom-5 MDA-MB-231 nuclei are as elongated as the top 25% of Control nuclei (Figure 12E).

As in A375 cells, the increased migratory efficiency of MDA-MB-231 cells is not related to a change in overall Lamin A/C or H3K9me3 levels (Figure 12I and 12J). Differences in Lamin A/C distribution in MDA-MB-231 cells are more subtle than in A375 cells: there are no significant differences among Control, Top-5 and Bottom-5 cells in Lamin A/C intensity variance (Figure 12F), but a somewhat higher percentage (90%) of Bottom-5 cells display visibly wrinkled nuclei compared to Control and Top-5 (75-80%). Our results support

previous observations in which elongated cell and nuclear shape correlated with higher migration propensity in MDA-MB-231 cells (Baskaran et al., 2020). Like A375 cells, MDA-MB-231 Bottom-5 cells have a higher H3K9me3 intensity at the nuclear periphery across the major and minor axis when compared to Control and Top-5 cells (Figure 12H).

In both these cancer cell types, constricted sequential migration can result in a more highly migratory subpopulation that exhibits some nucleus morphology changes and differs in the distribution, but not the levels of Lamin A/C or H3K9me3. Given the role of lamins and heterochromatin in organizing the 3D genome structure, we next asked whether differences are found in the 3D organization of chromosomes in sequentially constricted cells.

Genomic regions change their compartmentalization after rounds of constricted migration

To measure changes in chromosomal contacts in A375 cells, we performed genome-wide chromosome conformation capture (Hi-C) on Control, Top-5, Bottom-5, Top-12 and Bottom-12 cells (see Materials and Methods and Table S1 for processing and statistics). Strikingly, using 250kb binned matrices, we observe regions of the genome that have a different pattern of interactions in Bottom-5 cells compared to all other conditions (Figure 13A, example indicated by black arrow). Since these changes primarily affected the plaid pattern that indicates spatial compartmentalization, we next performed Principal Component Analysis (PCA) to classify regions according to their A/B compartment identity. We identified regions of the genome that have a consistent compartment identity among all unconstricted conditions (Control, Top-12, Top-5, Bottom-12) but have a

(A) 250Kb binned Hi-C interaction heatmaps of chr1 (chr1:181,785,969 - 235,044,680) across all A375 subpopulations. Arrow indicates visible differences in heatmaps of Bottom-5 cells when compared to the rest. (B) PC1 track of compartment identity in the same region of chr1 across all A375 subpopulations. Compartment tracks are binned at 250Kb. Boxed region highlights one of several compartment switches in this region: the same switch highlighted by arrows in (A). (C) Correlation between A375 Control PC1 with A375 cells that did not undergo transwell migration (top three panels) and A375 cells that did undergo transwell migration (bottom three panels). Percentage of bins that met the criteria for “compartment switch” from B to A (PC1 < -0.01 to > 0.01) or A to B (PC1 < -0.01 to > 0.01) indicated on each panel. (D) Distribution of PC1 values across all A375 subpopulations for the regions that switched their compartment identity from B to A compartment (top panel) and A to B compartment (bottom panel) in Bottom-5 cells. (E) Clustering of all A375 subtypes by their compartment PC1. (F) Enrichment of different LAD types in regions that switch their compartments from B to A and A to B in Bottom-5 (compartment comparison: Bottom-5 vs Top-5) (grey) and Bottom-12 (compartment comparison: Bottom-12 vs Top-12) (black) cells.

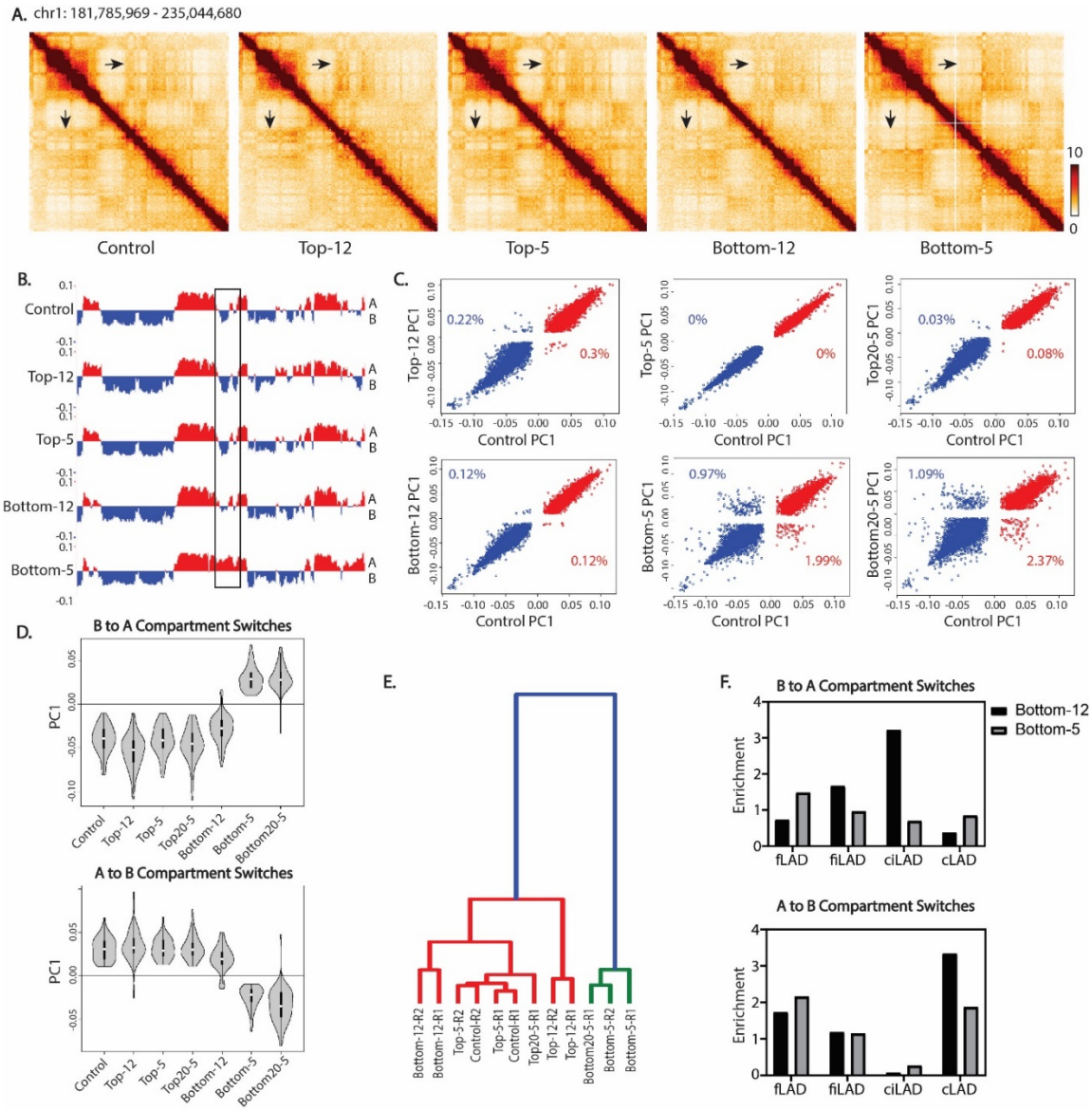


Figure 13. Compartment identity switches observed after constricted migration.

different compartment identity (A to B or B to A) in Bottom-5 cells (Figure 13B, highlighted region, for example). We determined that 1% of 250 kb regions across the genome switched from the B (typically more heterochromatic) to A (typically more euchromatic) compartment and about 2% of the 250Kb bins switched from A to B compartment in Bottom-5 cells compared to Control (Figure 13C). We observed a comparatively very small number of compartment identity switches in all the other conditions (Figure 13C) implying that 3D genome structure differences could be specific to cells that have undergone constricted migration. We also extended migration through 5μm pores up to 20 rounds (Bottom20-5) and observe similar compartment patterns and compartment identity switches, reinforcing the idea that these specific rearrangements of the genome are related to constricted migration (Figure 13C and 14).

To investigate whether any other cell subpopulations showed shifts in compartmentalization at regions that switched compartments in Bottom-5, we investigated the distribution of PC1 values of these regions in all the A375 conditions (Figure 13D). Regions of the genome that are in the B compartment in Control cells but in the A compartment in Bottom-5 cells show a negative distribution of PC1 values in non-migrating cells (Top-12 and Top-5), consistent with a B compartment identity, similar to the Control. As expected, the PC1 distribution has a positive mean in cells that have undergone constricted migration (Bottom-5 and Bottom20-5) (Figure 13D, top panel). However, while Bottom-12 cells exhibit a negative PC1 mean (reflecting a B compartment

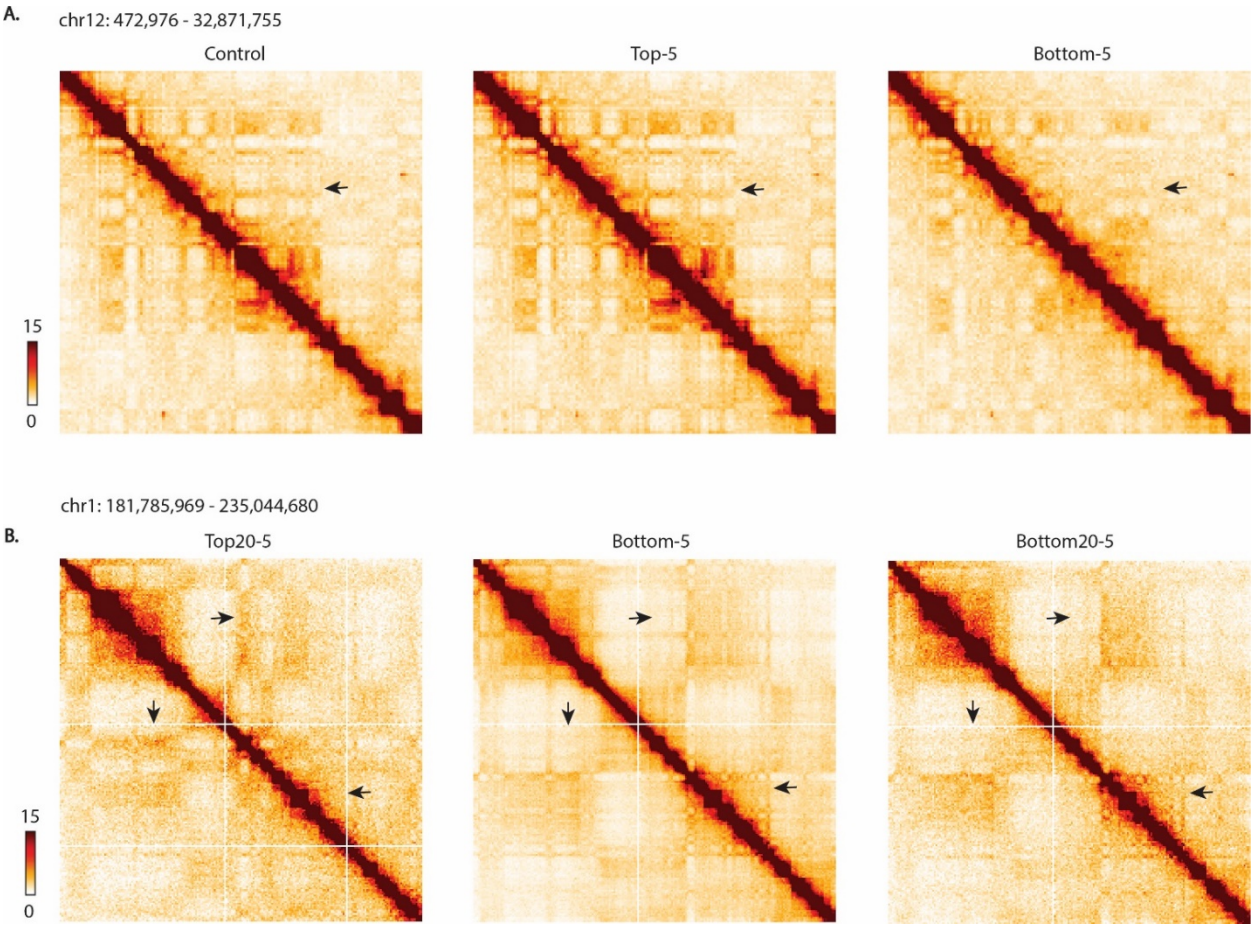


Figure 14. Compartment identity switches are also observed after 20 rounds of 5-μm constricted migration.

(A) Hi-C interaction heatmap of a chr12 region binned at 250Kb in A375-Control, Top-5 and Bottom-5. (B) Hi-C interaction heatmaps of chr1 (same chromosomal location as Figure 13A) in A375 cells that have undergone 20 rounds of 5-μm transwell migration. Black arrows indicate regions that exhibit changes in compartment profiles upon migration through 5-μm constrictions.

identity similar to Control), it is less negative when compared to Control and Top cells (Figure 13D, top panel). A similar pattern is observed in regions of the genome that switch compartment identity from A to B type. While the mean of PC1 distribution is positive for Control and Top cells and negative for Bottom-5 and Bottom20-5 cells, Bottom-12 cells exhibit a less positive mean of PC1 distribution (Figure 13D, bottom panel). This observation that unconstricted migrating cells have a slight tendency toward the structure differences seen in constricted migration could mean that these changes are somewhat induced by migration in general even with minor nucleus deformations, and/or that a sub-population of cells within the unconstricted migrating set would be able to undergo constrictions, and this subset with a "Bottom-5-like" structure influences the average interaction signal.

Not only are compartment changes a visually apparent difference in constricted migrating cells, but also genome-wide hierarchical clustering analysis of the compartment eigenvectors segregates conditions according to whether they have undergone constricted migration or not, further linking changes in chromatin compartmentalization to migratory ability (Figure 13E). Since compartmentalization relates to the spatial segregation of heterochromatin and euchromatin, these changes may relate to our microscopic observations of increased H3K9me3 peripheral localization after confined migration (Figure 11). To investigate this link further, we quantified whether genomic regions that change compartment identity in Bottom cells (when compared to their respective Top: Bottom-5 vs Top-5 and Bottom-12 vs Top-12) correspond to certain lamin

associated domain (LAD) types. We used a previously established LAD atlas across a cohort of nine different cell types that classifies genomic regions as: cLADs (constitutive LADs; regions associated with the nuclear lamina in all cell types), ciLADs (constitutive inter LADs; regions located in the interior of the nucleus across all cell lines), fLADs and fiLADs (facultative LADs and facultative inter LADs; regions that are sometimes associated with the nuclear lamina across the cell types) (Kind et al., 2015, Carolyn de Graaf, 2019).We observed both B to A and A to B switches in constricted migration (Bottom-5) were enriched in fLADs (Figure 13F). The high incidence of compartment changes in fLAD regions might relate to their ability of these fLADs, representing their ability to transition between lamina-association and dissociation. Additionally, Bottom-12 and Bottom-5 presented with an enrichment of cLADs in the regions that switch from A to B compartment, suggesting a possible further spatial consolidation of heterochromatin regions corresponding to the observed coalescence of H3K9me3 at the nuclear periphery. Interestingly, we observed an over-representation of ciLADs in regions switching their identity from B to A compartment between Top-12 and Bottom-12 (Figure 13F, top panel). Given that ciLADs are regions that do not associate with the nuclear lamina in all the examined cell types, it was unusual to find these regions in the B-compartment in the non-migratory cells.

We have identified 96 bins that are classified as ciLADs residing in a B compartment in Top-12 cells but are found in the A compartment in all other conditions. Gene ontology enrichment analysis identified functions correlating to cell motility and the regulation of

actin cytoskeleton for the genes found in these regions. Additionally, we observed an over-representation of cLADs in regions that switched their compartment identity from A in the Top-12 to B in Bottom-12, encompassing 253 bins found in the A compartment in Top-12 and in the B compartment in all the other conditions. The genes in these regions were associated with melanogenesis and neuronal pathways. Taken together, these observations suggest that Top-12 cells (the least migratory in the constricted migration spectrum) harbor a 3D genome structure that separates them from the highly migratory cells (Bottom-5 and Bottom20-5) and likely reflects maintenance of melanocyte identity.

Gene expression changes after constricted migration reflect metastatic potential

To investigate potential differences in gene expression in cells that have undergone constricted migration, we performed RNA-Seq on all A375 subpopulations. We identified 977 and 1473 genes that were differentially expressed in Bottom-5 and Bottom20-5 relative to Control, respectively (Table 3). Hierarchical clustering of all conditions, starting with the subset of genes differentially expressed in Bottom-5 or Bottom20-5, revealed that cells that undergo constricted migration cluster separately from all other conditions (Figure 15A and Figure 16A). Overall, we observe three major gene clusters: cluster 1 included 482 genes that are upregulated in Bottom-5 and Bottom20-5 and downregulated in other subsets. Pathway enrichment analysis identified upregulated pathways related to metastasis, such as TGF-beta, TNF-alpha, EGFR and Integrin β-4 signaling pathways (Figure 15B). On the other hand, cluster 2 included 368 genes that were downregulated in Bottom-5 and Bottom20-5 but upregulated in all other cells. Interestingly, enriched

(A) Hierarchical clustering of RNA-Seq datasets based on differentially expressed genes of Bottom-20 cells relative to Control (Log2FC > 1). (B) Pathway enrichment analysis of identified gene clusters from (A) using BioPlanet 2019 package from Enrichr. (C) Log2FC of expression in the indicated condition vs. Control for genes located in regions that switch compartments from A to B (top panel) and B to A (bottom panel) in Bottom-5 cells. (D) Enrichment of regions that switched from A to B compartment based on the presence or absence of genes within the bin. (E) 20Kb binned Hi-C matrices of upregulated (FBXL7) and downregulated (CCDC149) genes in Bottom-5 cells. Tracks over the matrices represent ATAC-Seq (black) and RNA-Seq (red). The tracks have been smoothed with 40Kb bins. Black arrows indicate known CTCF motif directionality sites.

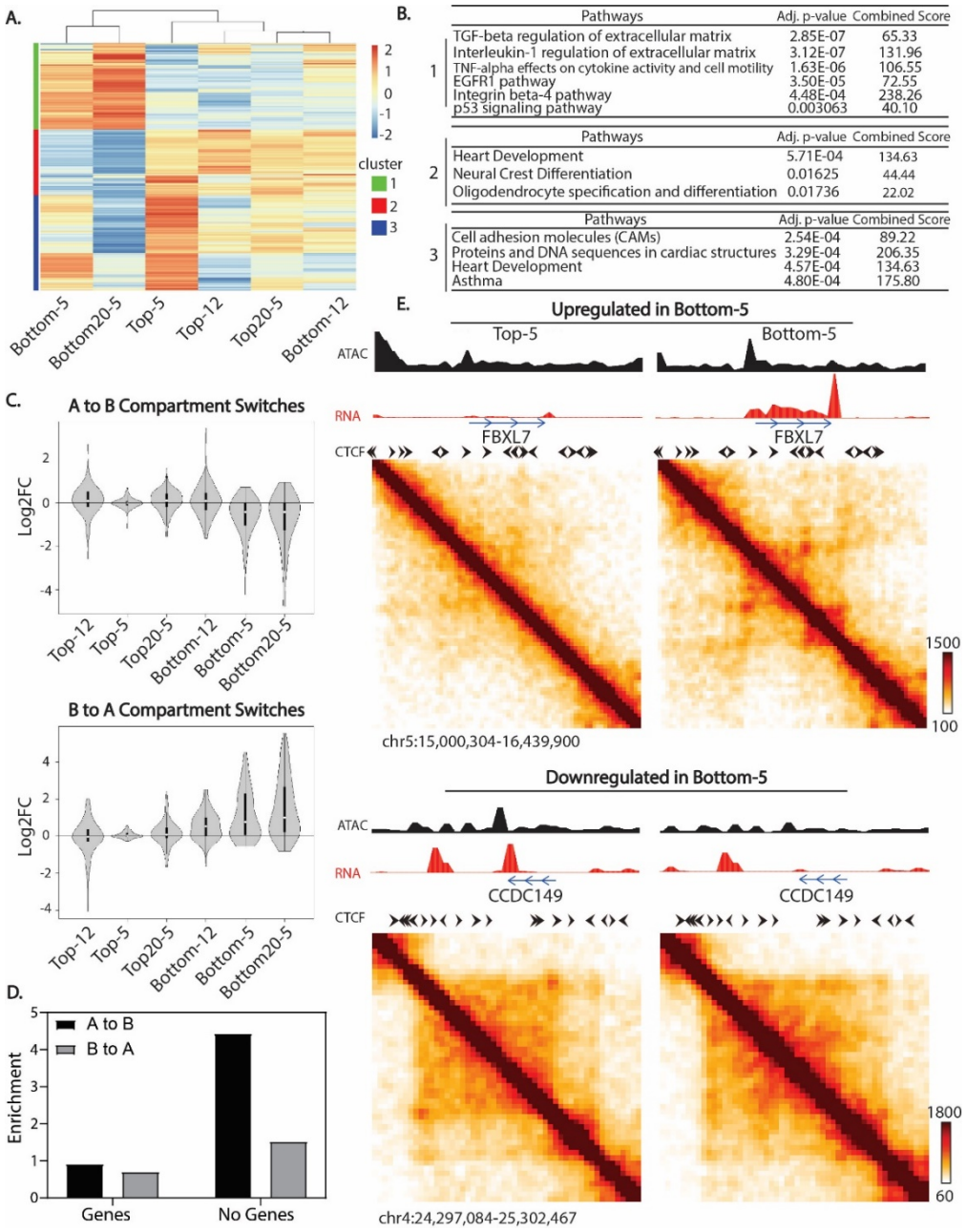


Figure 15. Changes in gene expression correlate with metastatic ability and chromosome structure changes.

downregulated pathways are involved in developmental pathways such as neural development. Since melanoma cells originate from neural crest, these observations suggest that this cluster relates to downregulation of melanocyte specific gene expression in constricted migrating cells (Figure 15B). Finally, cluster 3 included 533 genes that were downregulated in Bottom20-5 and upregulated in Top-5 (Figure 15A). Some of these genes are involved in cell adhesion, which correlates with the differences we observe in cell-cell contacts between these subsets (Figure 8C).

We next searched for other potential links between observed phenotypic differences (Figure 8 and Figure 16B) and changes in gene expression. We first compared cells which have migrated through non-constricted spaces to cells that do not migrate at all (Bottom-12 vs Top-12, Figure 16B). The majority of genes that were upregulated in Bottom-12 cells were related to TGF-beta, EGFR, integrin and p-53 signaling, all pathways related to cancer cell migration and metastasis. To assess whether we observe constriction-specific changes in gene expression, we next compared cells which have migrated through 5μm constrictions to cells which migrated without constriction (Bottom-5 vs Bottom-12, Figure 16B). Interestingly, genes that were more highly expressed in constricted vs. unconstricted migration are related to focal adhesion pathways and regulation of actin cytoskeleton. These changes correlate with the increase in protrusions and changes in cell shape observed for Bottom-5 cells in Figure 8 and 9 (Figure 16B). Similarly, when we compared cells that underwent 5μm constricted migration (Bottom-5) with the ones that were not able to migrate, even through large pores (Top-12), we found

(A) Hierarchical clustering of all A375 subset of cells based on genes that are differentially expressed in Bottom-5 cells relative to Control (Log2FC >1). (B) Gene ontology analysis of differentially expressed genes comparing indicated sub-populations of A375 cells. Cell illustration indicates morphology differences we observe within each sub-population of A375 cells. (C) Hi-C contact matrices binned at 20Kb and smoothed at 40Kb bin size of upregulated genes (PLEKHA6, left two panels) and downregulated genes (PKDCC, right two panels) in Bottom-5 cells. Black track represents ATAC-Seq signal while the red track represents RNA-Seq signal. Black arrows indicate known CTCF motif directionality. (D) Correlation plots of differentially expressed genes between mechanically stretched cells (Nava et al., 2020) and A375 cells that have undergone constricted migration (10 rounds of 5 micron constriction, left; 20 rounds of 5 micron constriction, right).

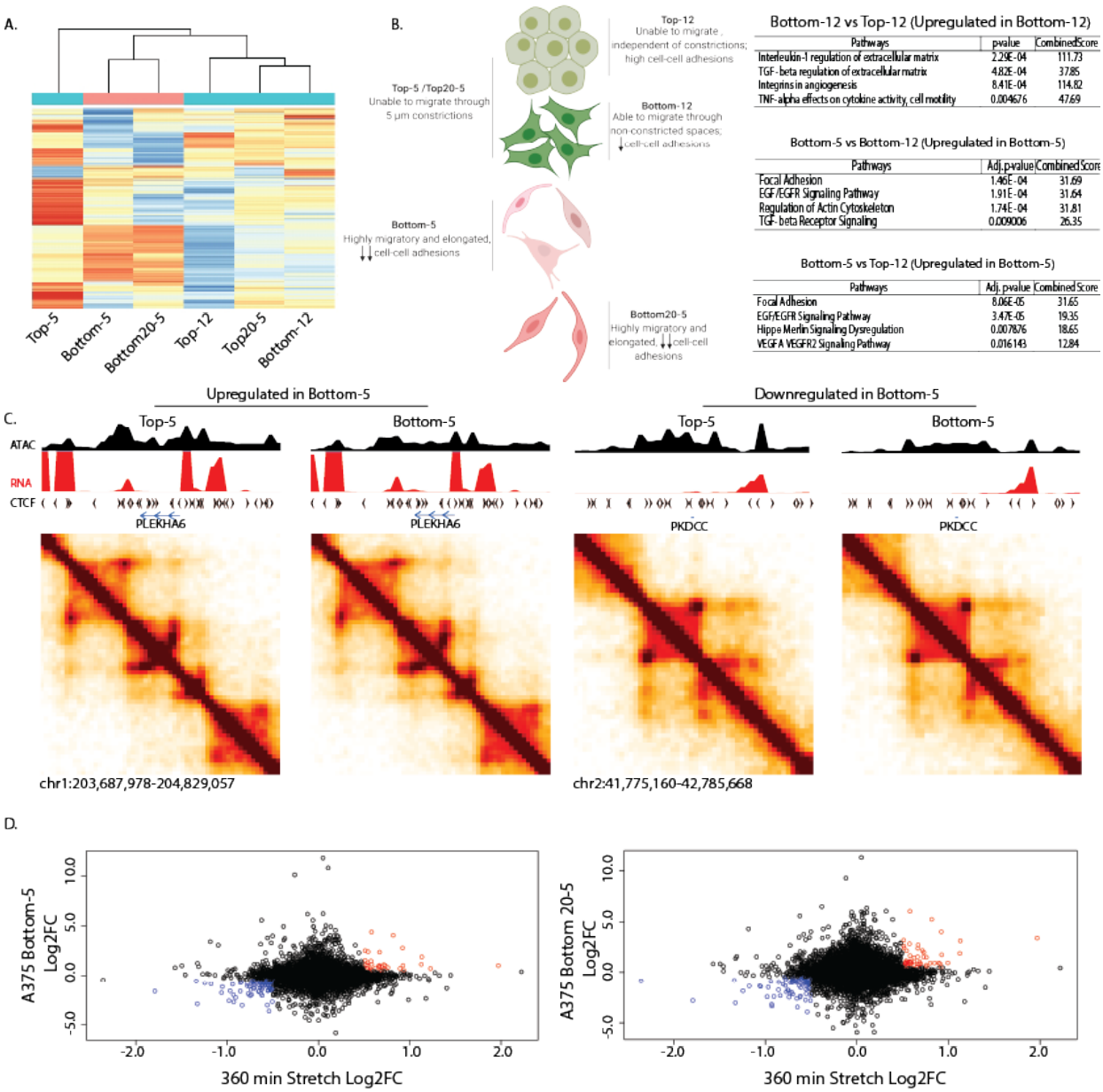


Figure 16. Gene expression profile separately clusters cells that have undergone constricted migration (Bottom-5 and Bottom-20) from all other A375 subpopulations.

upregulated genes that are related to cancer invasiveness pathways such as EGF and VEGF. Additionally, many genes upregulated in Bottom-5 vs.Top-12 were involved in focal adhesion pathways which correlates with the more punctate, focal, staining of F-actin by phalloidin in Figure S1F (Figure 16B). These gene expression changes also correlate with A/B compartment changes observed upon constricted migration (Fig. 4C). Genes found in regions that switch their compartment identity from A to B in Bottom-5 or Bottom20-5 cells display an overall decrease in expression when compared to all other conditions (Figure 15C, top panel). Similarly, genes found in regions that switch their identity from B to A compartment in Bottom-5 and Bottom20-5 cells display an overall increase in expression (Figure 15C, bottom panel). These observations suggest that some of the compartment changes after confined migration directly relate to changes in gene regulation. We cannot definitively say whether certain key migration genes were, for example, already active and in the A compartment in a subset of A375 control cells, and these were selected by sequential migration, or whether some such genes became active after a compartment switch was induced by confined migration. However, we can conclude that at least part of the underlying biological function of the compartment switching we observe in confined migrating cells involves changes in gene expression.

Gene expression changes do not account for all compartment changes, however. Notably, we found that among the genomic bins that switch from A to B and B to A compartment, a larger than expected fraction contain no genes (4-fold and 2-fold enrichment over random expectation, respectively) (Figure 15D). This observation

indicates that a substantial portion of structural changes after migration must be explained by something other than gene regulation changes. We hypothesize that these regions might change as a result of or to accommodate the physical strain on the chromatin fiber during constricted migration.

To investigate whether the observed gene expression changes relate to local genome structure changes, we zoomed in to regions of the genome encompassing some of the highest upregulated or downregulated genes after constricted migration (Figure 15E). We initially explored genes that were upregulated (log2FC > 1) in cells that had undergone constricted migration (Bottom-5 and Bottom20-5). When zooming in to the FBXL7 gene (~ 8-fold upregulated in Bottom-5 and Bottom20-5), we observed the emergence of a domain-like structure in Bottom-5 cells and an increase in chromatin accessibility (Figure 15E, top panel). Similarly, when exploring genes that were downregulated after constricted migration (Bottom-5 and Bottom20-5), we found genes whose local 3D structure was altered in Bottom-5 or Bottom20-5 cells with the disappearance of the loop present in Top-5 and a decrease in chromatin accessibility (SOD3 and CCDC149 3-15 fold downregulated; Figure 15E, bottom panel). However, we also encountered cases in which the genes were differentially expressed such as PLEKHA6 (log2FC > 2 in Bottom-5 and Bottom20-5) and PKDCC (log2FC = -2.5 in Bottom-5 and Bottom-20) but their local genomic structures remained similar with modest changes in chromatin accessibility (Figure 16C). We observe that the loss or formation of loops or domains associated with changes in transcription occurs in regions that had few or no other strong architectural

loops, such as the kind mediated by CTCF. Genes that change expression without notable 3D looping changes, in contrast, were located in regions with many other loops and TAD boundaries. These observations support the idea that transcription driven domains can arise separately from CTCF-based domains (Rowley et al., 2017, Zhang et al., 2020). Further, the transcription-associated CCDC149 loop occurs between an existing TAD boundary and a promoter, and could be explained by a collision between opposing direction loop extrusion and RNA polymerase activity, as suggested in other systems (Brandao et al., 2019).

Previous work has observed gene expression changes induced by mechanical stress, such as stretching (Nava et al., 2020, Le et al., 2016). We observe little correlation between differentially expressed genes in cells that have undergone constricted migration and cells that have been mechanically stretched (Figure 16D). This points to differences between a cell being passively deformed and one that is actively coordinating its own migration through a tight constriction.

Inter-compartment interactions increase after constricted migration

While the switching of compartment identity after confined migration is a striking phenomenon, and clearly apparent by visual inspection of the contact maps, quantitatively, only ~ 3% of the genomic bins (250Kb bins) were affected by a complete compartment switch according to our thresholds. To evaluate how constricted migration may impact the rest of the chromosome contacts, we evaluated contacts not involved in

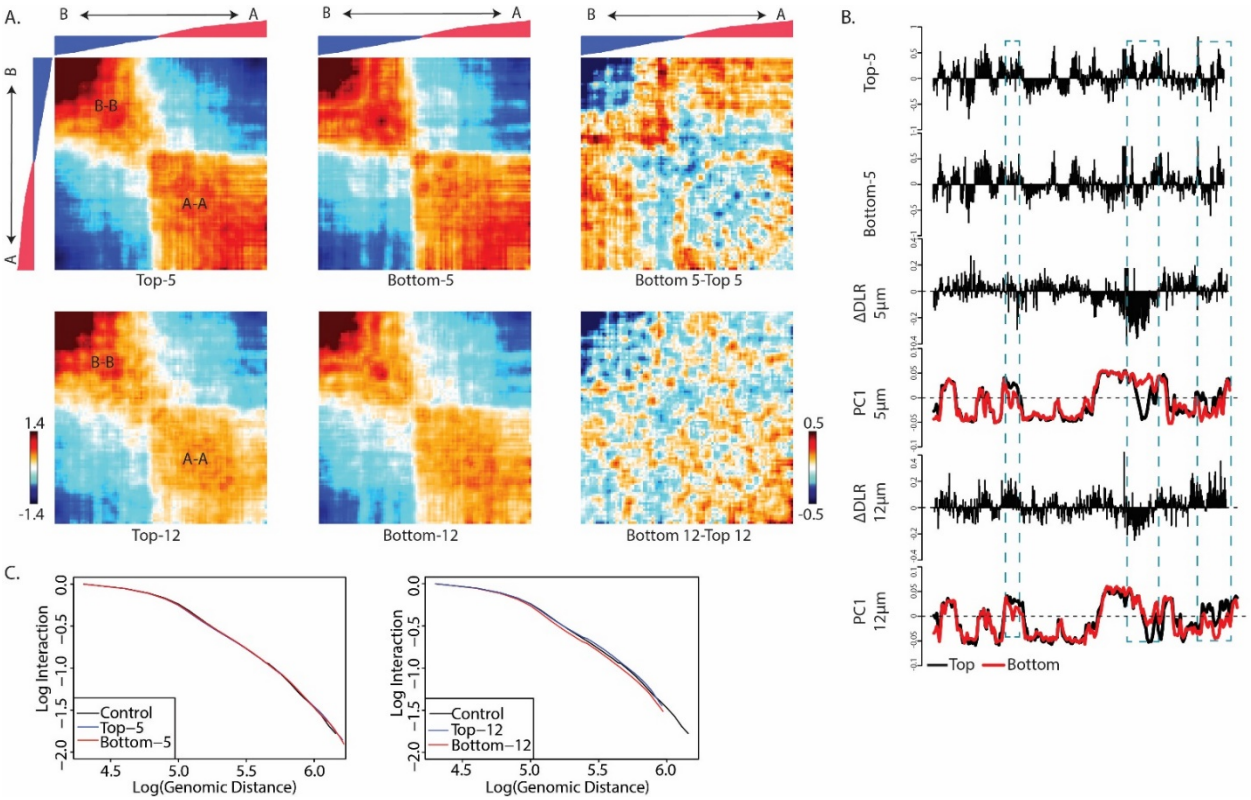


Figure 17. Changes in compartment strength associated with constricted migration.

(A) Compartmentalization saddle plots of chr10 display interaction Z-score (interactions normalized by generic distance decay) or Z score difference (right column) between all bin pairs, reordered by PC1 values derived from A375 Control HiC data. Interactions between the strongest B regions are in the upper left corner while A-A interactions are in the lower right corner. The reordered Control PC1 values are displayed over the saddle plots. (B) Along chromosome 1 (chr1:170,000,001–225,000,000), from top to bottom: Distal to local (DLR; 100 kb bins) ratio for Top-5 cells; DLR of Bottom-5 cells; Delta DLR (Bottom-5 – Top5). Positive DLR values indicate a gain of distal interactions (>3Mb) while negative values indicate a loss of distal and gain of local interactions (<3Mb); PC1 values (250 kb bins) of Top-5 (black) and Bottom-5 (red) samples; Delta DLR (Bottom12-Top12); PC1 values of Top-12 (black) and Bottom-12 (red). Positive PC1 values represent A-compartment identity and negative PC1 values B-compartment identity. (C) Whole genome scaling plots at 20Kb bin size comparing A375 cells that have been exposed to 5-μm transwell migration (top plot) and 12-μm transwell migration (bottom).

compartment identity switches by reordering the 250Kb binned matrices from the strongest B to strongest A bins based on PC1 values (Figure 17A) (Imakaev et al., 2012). We observe an evident separation between the B and A compartment in all subpopulations of A375 arising from 5μm (Figure 17A top panel and Figure 18A) and 12μm (Figure 17A bottom panel and Figure 18A) transwell migration experiments. However, the relative strength of inter and intra-compartment contacts changes between Top and Bottom cells (Figure 17A, middle panels). Bottom-5 cells exhibit a loss in the strongest B-B interactions and an overall increase in interactions between the strongest B and strongest A compartments. This phenomenon has been previously reported in neutrophils after passage through small pores, suggesting it may be a general feature of confined migration genomic structural changes (Jacobson et al., 2018). This change in interaction frequency correlates with changes in gene expression. Regions that exhibit a strong B-B interaction loss display an overall increase in expression of genes found in these regions. Additionally, the expression of genes found in regions that display a gain in B-B interactions show an overall decrease in expression correlating with strengthening of the B-B interactions (Figure 18B). Interestingly, though there is increased mixing of the strongest A and B compartments after confined migration, there is actually increased segregation of regions that are weakly compartmentalized (eigenvector values near 0) in cells passed through constrictions. It appears that the B compartment bins that were previously ambiguous or near compartment boundaries in Top cells are more strongly associated with the rest of the B compartment after confined migration. We also observe

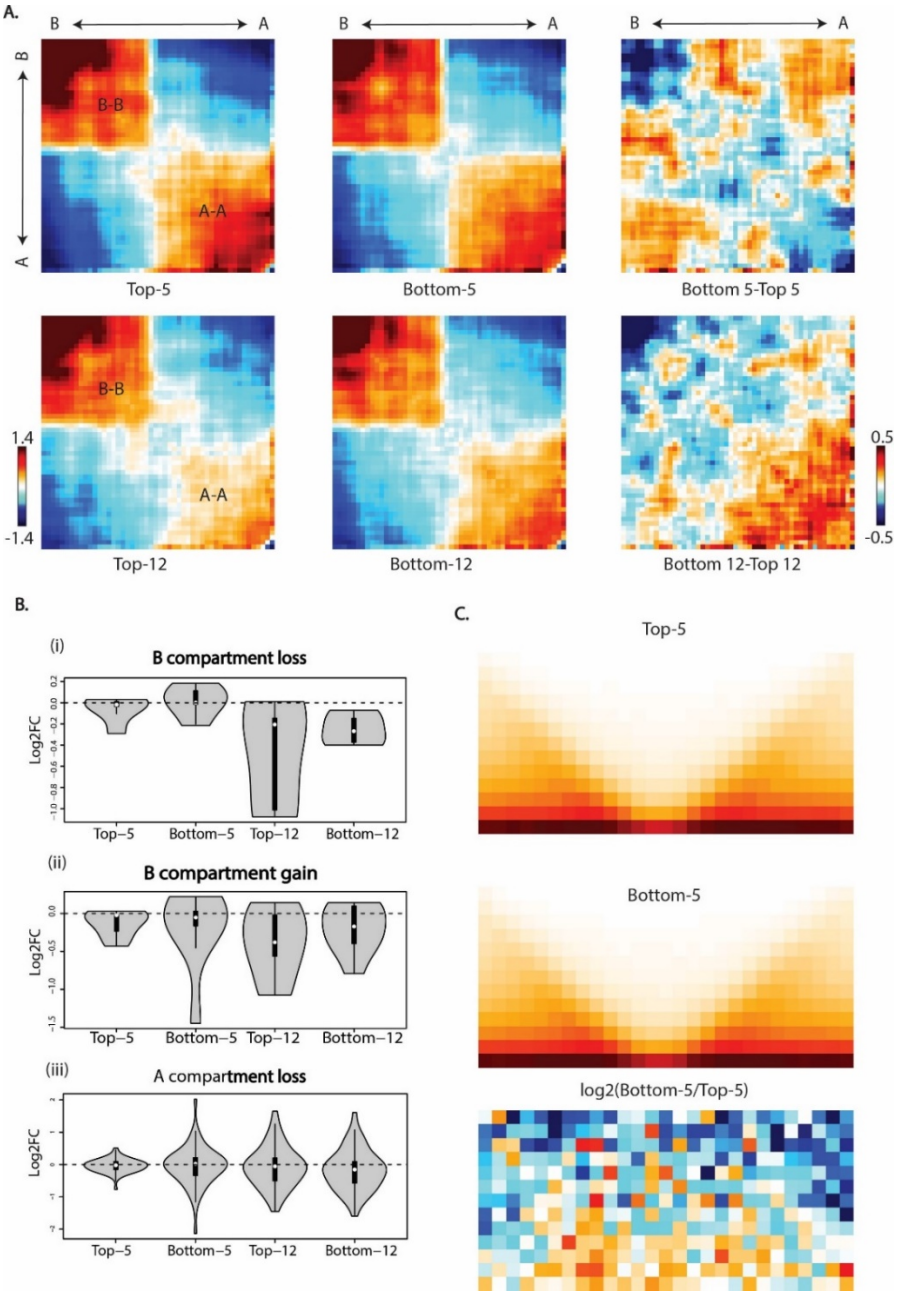


Figure 18. Changes in compartment interaction strength correlates with overall gene expression pattern.

(A) Saddle plots binned and smoothed using 500Kb bin size for chr18 reordered from strongest B to strongest A compartment. (B) Gene expression changes corresponding to strong B-B interaction loss (i), moderate B-B interaction gain (ii) and A-A interaction loss (iii). (C) Aggregate contact maps (40 kb bins) at TAD boundaries called by the Insulation Score method of cworld-dekker with strength greater than 1 of Top-5 (Top panel) and Bottom-5 (middle panel).

a loss of the strongest B-B contacts in Bottom-12 as compared to Top-12, though this effect is not as strong as in the constricted migration case. In contrast to Bottom-5 cells, Bottom-12 cells exhibit an increase of interactions between A-A contacts relative to Top-12. To understand how these differences in compartment interaction frequencies relate to interactions with linearly local or distal chromosome regions, we calculated the distal-to-local ratio (DLR) of Hi-C interaction frequencies (Figure 17B). Here, distal is defined as interactions with regions more than 3 Mb away along the chromosome. This measure has previously been associated with changes in chromosome compaction (Heinz et al., 2018) and is related to a measure used to quantify loss of local chromosome structure with chromosome fiber digestion (Belaghzal et al., 2019). We find that in general, where regions along the chromosome become stronger A compartments, distal interactions decrease relative to local interactions (negative DLR value). Where regions increase their B compartmentalization or lose A identity, distal interactions increase (positive DLR). This could suggest that increased B compartmentalization often occurs when a region is brought closer to other heterochromatin regions far away on the chromosome, perhaps by co-localization at the nuclear lamina, while loss of B/increased A compartmentalization involves breaking distal contacts and increasing local contacts. (highlighted region, Figure 17B).

Although the DLR shows that individual regions do have different local and distal interaction patterns, the average decay of interactions with distance does not change among any of the A375 subpopulations (Figure 17C). This local interaction scaling reflects

local chromatin compaction and average local loop size and density, and thus we show that constricted migration does not overall change these local chromosome features. Moreover, we investigated whether constricted migration is associated with changes in TAD boundaries, and we observe no clear differences between any of the conditions (Figure 18C). These observations demonstrate that the major differences with constricted migration occur in larger scale chromosome structures and are in line with previous reports that TAD-scale interactions are independent from compartment-scale interactions such that one can change without dramatically affecting the other (Nuebler et al., 2018).

Global genomic rearrangements after constricted migration

Constricted migration has previously been shown to result in DNA damage and genome instability (Irianto et al., 2017a). Since Hi-C can detect structural aberrations such as translocations, we examined interchromosomal interactions in 2.5 Mb binned matrices to search for potential genomic aberrations after confined migration (Figure 19).

At first glance, it is evident that translocations inherent to A375 are present in all sub-populations of cells (Figure 19A and 20). However, a new translocation between chr5 and chr13 (t(5;13)) is evident only in cells that had undergone constricted migration (Bottom-5 and Bottom20-5) (Figure 19B and 20). To narrow the specific breakpoint region of this translocation, we focused on Hi-C contact frequency between chr5 and chr13 (Figure 20). While in Top cells (Top-12, Top-5 and Top20-5) contact frequencies between chr5 and 13 remain uniform across the whole chromosome, in Bottom-5 cells we observe a high

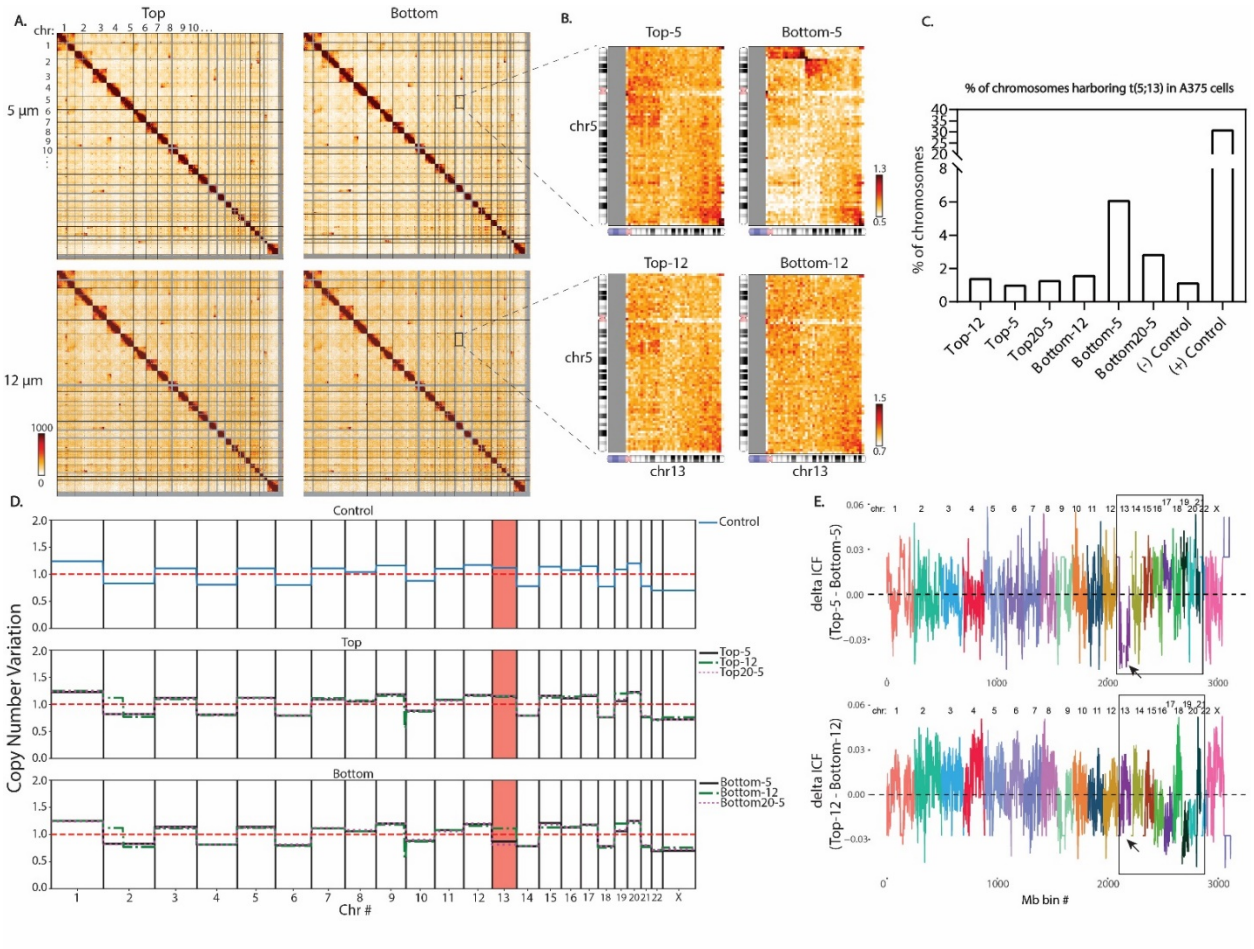


Figure 19. Global genomic rearrangements after constricted migration.

(A) Whole genome 2.5Mb Hi-C interaction frequency maps for cells that have undergone 5 and 12-µm transwell migration. (B) Zoom-in of inter-chromosomal interactions between chr5 and chr13 for cells that have undergone 5 and 12-µm transwell migration. (C). Estimation of the % of chromosomes bearing the t(5;13) based on the frequency of the proposed translocation interaction divided by the average neighboring bin cis interactions. (D) Copy number variation among all A375 subtypes inferred from total raw Hi-C counts. Red dashed line represents the mean copy number level while the other lines represent the copy number of A375 cell subpopulations. The region highlighted in red shows that chr13 exhibits copy number loss after 10 rounds of constricted migration. (E) Plots represent the difference in inter-chromosomal fraction (delta ICF) in cells that have undergone migration through 5-µm constrictions (top plot) and cells that have undergone migration through 12-µm constrictions (bottom plot) using 1Mb binned Hi-C contacts. The boxed region highlights regions of difference between constricted and unconstricted migration.

frequency of interaction between chromosome 5 and chromosome 13 followed by a sudden drop of interactions, indicating the translocation breakpoint (Figure 20). Interestingly, the translocation breakpoint falls in a gene poor region of both chromosomes.

We next used Hi-C contacts to estimate what fraction of the chromosomes in the population contain this translocation (Figure 19C, See Materials and Methods for detailed analysis approach). We found that this translocation was only present in 6.13% of chromosomes in Bottom-5 and 2.89% in Bottom-20. Since A375 cells have an average of 3 copies of each chromosome per cell, this can be extrapolated to this translocation occurring on one chromosome copy in ~18% of the Bottom-5 cells and 9% of the Bottom20-5 cells (Figure 19C). Interestingly, this translocation was not detected in the cells that did not undergo constricted migration (Top-12, Top-5, Top20-5 and Bottom-12) as the fraction of reads coming from the translocation site matched the random background level of interchromosomal interactions in the Hi-C contact map (Figure 19C, negative control). The absence of this translocation in non-migratory cells (Top-12, Top-5 and Top20-5) and in Bottom-12 cells suggests that the translocation event may have arisen due to the physical stress on the nucleus during constricted migration. However, it is also possible that this translocation was present at undetectable levels in the heterogeneous control A375 population, and we are selecting the cells with this translocation during sequential migration.

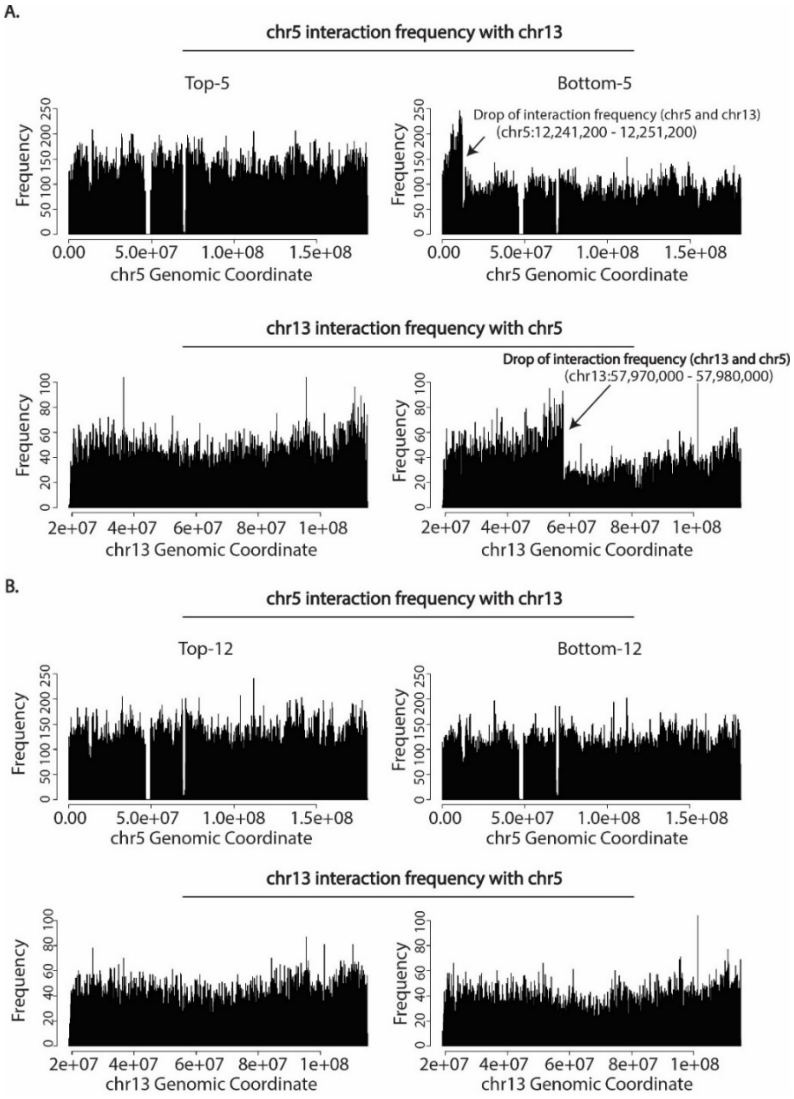


Figure 20. Translocation between chr5 and chr13 is specific to migration through 5µm constrictions.

(A) Top panel: Contact frequency of chr5 genomic coordinates with chr13 in Top-5 cells (left) and Bottom-5 cells (right). Bottom panel: Contact frequency of chr13 genomic coordinates with chr5 in Top-5 cells (left) and Bottom-5 cells (right). Black arrows indicate a decrease in contact frequency between chr5 and chr13 (Top panel) and chr13 and chr5 (Bottom panel) indicating the end of the translocated region. (B) Top panel: Contact frequency of chr5 genomic coordinates with chr13 in Top-12 cells (left) and Bottom-12 cells (right). Bottom panel: Contact frequency of chr13 genomic coordinates with chr5 in Top-12 cells (left) and Bottom-12 cells (right). A drop in contact interaction frequency is not observed due to undetectable t(5,13).

In addition to translocations, changes in chromosome copy number have been previously reported after constricted migration (Irianto et al., 2017a). We therefore followed a published approach (Servant et al., 2018) to use raw read counts from our Hi-C data to investigate possible copy number variations (CNV) due to constricted migration in A375 subpopulation of cells (Figure 19D). The control population of A375 cells already exhibits non-uniform copy number across chromosomes and chromosome regions (Figure 19D, top panel). When comparing the subpopulations of A375, most of the chromosomes maintained their copy numbers independent of constricted migration. However, cells that have undergone constricted migration (Bottom-5 and Bottom20-5) exhibit notable copy number loss in chr13 (Figure 19D, bottom panel).

Beyond revealing translocations, interchromosomal contact data can reveal changes in patterns of whole chromosome territory interactions between conditions. So, we next sought to explore how the interchromosomal fraction (ICF) of interactions (Heinz et al., 2018) across the genome varies among the cells that have undergone constricted migration (Figure 19E). Interestingly, we observe a decrease in delta ICF of chr13 (Figure 19E, top panel, indicated by black box) in Bottom-5 cells. This is not likely to be explained purely by copy number loss at chr13, as this would decrease both inter and intra-chromosomal contacts rather than changing the ratio between the two. Additionally, smaller chromosomes (chr14-chr22) display an overall increase in delta ICF in Bottom-5 cells (Figure 19E, top panel, indicated by black box) but a decrease of ICF in Bottom-12 cells (Figure 19E, bottom panel, indicated by black box). This measurement could indicate

that smaller chromosomes differ in their relative compactness and intermixing depending on constricted migration.

Cells that have undergone constricted migration exhibit altered nuclear morphology in 3D collagen matrices

To investigate whether the cells that had undergone multiple rounds of constricted migration show metastatic behavior in more biologically-relevant environments, we embedded A375 Control, Top-5 and Bottom-5 cells expressing Dendra2-H4-GFP (histone-4) in 3D collagen gel matrices as described previously (Denais et al., 2016). Bottom-5 cells embedded in collagen with a density of 2.42 mg/mL, displayed long cytoplasmic protrusions and nuclear deformations compared to Control and Top-5 which exhibited rounder cell and nuclear morphologies (Figure 21A and 22A). Bottom-5 cells showed protrusions after only 5 hours after embedding in collagen and these protrusions increased after 24 and 48 hours of incubation in the 3D collagen matrix (Figure 21B). In contrast, only a small fraction of cells that did not undergo constricted migration in the transwell (Top-5 and Control) exhibited cell protrusions even after 48 h incubation in 3D collagen matrices (Figure 21B). With increasing incubation time in the 3D collagen matrices, the nuclei of the Bottom-5 cells that are extending protrusions also appear to be more elongated than Control and Top-5 cells, suggesting that Bottom-5 cells are migrating through the collagen fibers and deforming their nuclei (Figure 21C). Indeed, Bottom-5 cells display highly motile behavior in 3D collagen matrices compared to Top-5 cells (Figure 21D, Movies S5 and S6). When migrating through collagen, Bottom-5 cells travel longer distances at a higher speed (Figure 21E, Movie S6). The minimal nucleus

(A) Phase contrast images of A375 cells 48 hours after embedded in 3D collagen matrices show differences in their cell morphology. Dendra2-H4-GFP labels the nucleus (green). (B) Fraction of Control, Top-5 and Bottom-5 cells displaying projections at indicated timepoints after embedding in 3D collagen matrices. Each biological replicate is connected by a line. (C) Aspect ratio quantification of A375 nuclei at timepoints after collagen embedding for Control, Top-5 and Bottom-5 cells. ($N > 20$ for each condition; * $p = 0.0168$, ** $p = 0.0016$, **** $p < 0.0001$; two tailed t-test). (D) Live cell image tracking of Control, Top-5 and Bottom-5 cells migrating in 3D collagen matrices. (E) Quantification of accumulated distance and velocity of Control (n=17), Top-5 (n=17) and Bottom-5 (n=24) migrating in 3D collagen matrices. (F) Nuclei of Top-5 (left) and Bottom-5 cells (2 different cells, right) stained with DAPI (blue) and Lamin A/C (green). Scale bars (white) = 5μm. For 3D reconstruction visualization, see Movie S5 and S6. (G) Left panel: Measurements of nucleus elongation by aspect ratio in Control (n = 54), Top-5 (n = 52) and Bottom-5 (n = 53). Top-5 vs Control ($p = 0.8232$); Bottom-5 vs Top-5 (**$p = 0.0054$); Bottom-5 vs Control (**$p = 0.0031$) as calculated by two-tailed t-test. Right panel: Percent of cells with elongated nuclei in A375 Control, Top-5 and Bottom-5 in 3D collagen (Aspect Ratio > 2.5). (H) Percent of nuclei that display irregularities in Lamin A/C distribution in Control, Top-5 and Bottom-5 cells embedded in 3D collagen by visual inspection. Each biological replicate connected by a line. (I) Cumulative distribution plot of Lamin A/C coefficient of variance for Control, Top-5 and Bottom-5 cells embedded in 3D collagen matrices. Number of nuclei quantified remains the same as in G.

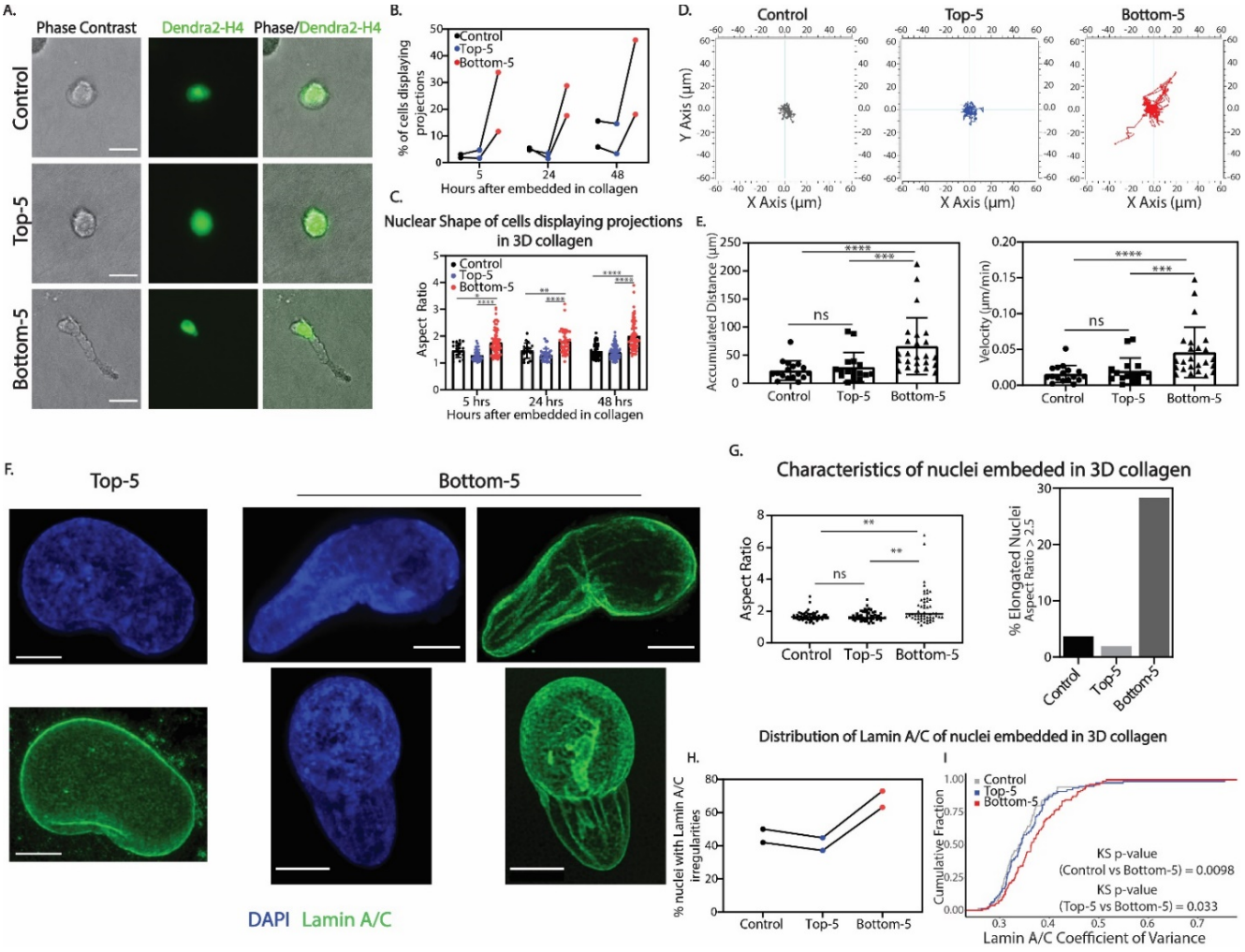


Figure 21. Sequentially constricted cells show elongated nuclei in 3D collagen matrix model.

diameter of Bottom-5 cells migrating through the collagen fiber meshwork ranges from 3-11μm (Figure 22B). These nucleus constriction sizes suggest that the collagen meshwork produces constrictions comparable in size with transwell pore sizes through which cells must squeeze their nuclei.

Analyzing the nuclear morphology differences at higher resolution, we observe that Bottom-5 cells display larger deformations of the nucleus in collagen than Control and Top-5 cells (Figures 21F, 22C, Movie S7 and S8). About 30% of Bottom-5 cells exhibit elongated nuclear morphology (defined by aspect ratio > 2.5) in collagen while only 1-3% of Control and Top-5 cells exhibited elongated nuclei (Figure 21G).

Changes in nuclear morphology also correlated with changes in distribution of Lamin A/C in Bottom-5 nuclei embedded in 3D collagen. While nuclear wrinkling and therefore abnormal Lamin A/C patterns were evident in all A375 subpopulations, a higher fraction of Bottom-5 nuclei exhibit these irregularities (Figure 21H). Bottom-5 cells exhibit a higher Lamin A/C coefficient of variance in the nucleus when compared to Control and Top-5 cells (Figure 21I). Overall, quantification of nuclear shape differences in collagen distinguished cells that had undergone constricted migration in transwell chambers from cells that did not, indicating that the ability to deform their nucleus and migrate in a 3D environment is a key distinguishing factor among the A375 cell subpopulations.

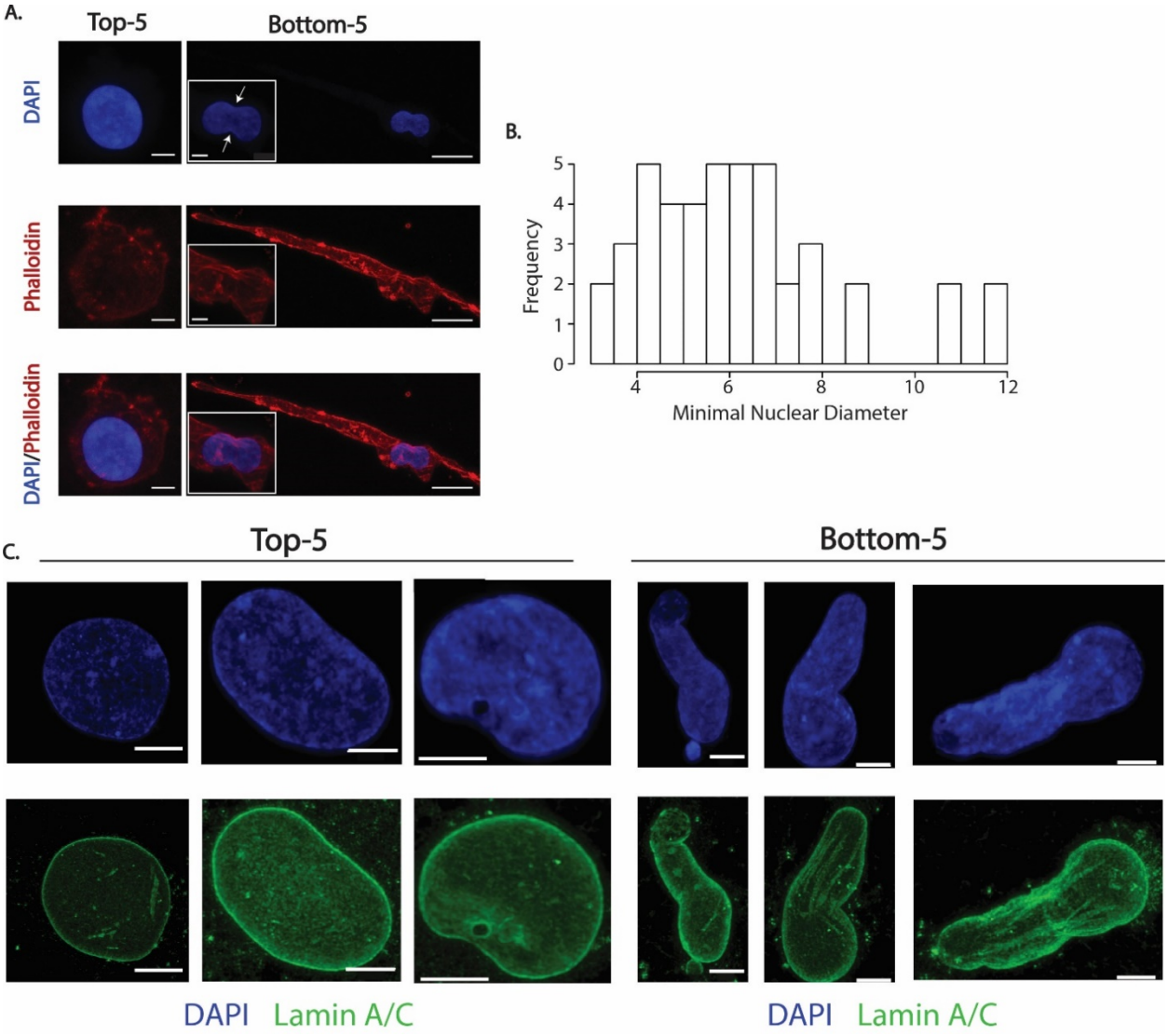


Figure 22. A375 cells that have undergone sequential transwell migration deform their nuclei during migration through 3D collagen matrices

(A) Maximum projection confocal images of Top-5 and Bottom-5 cells stained with DAPI (blue) and Phalloidin (red). (Insets) Zoom in of nucleus of Bottom-5 cell. Scale bars (white) indicate 5μm for Top-5 and Bottom-5 insets; 20μm for Bottom-5 zoomed out images. White arrows in insets indicate the minimal nuclear diameter. (B) Distribution of minimal nuclear diameter of all elongated Bottom-5 nuclei embedded in 3D collagen. (C) Maximum projection confocal images of Top-5 (left panel) and Bottom-5 (right panel) nuclei visualizing DAPI (blue) and Lamin A/C (green). Scale bars (white) indicate a length of 5 μm.

Agent based modeling reveals that heterogeneity and constricted migration induced changes could both contribute to sequential migration effects

As described above, the presence of the subpopulation of A375 cells that cannot migrate at all (Top-12) demonstrates an inherent heterogeneity in migratory ability within the population, as previously reported in this cell line (Kozlowski et al., 1984). However, it is not clear whether the increase in migratory efficiency and associated 3D genome and nucleus morphology alterations in Bottom-5 cells are induced by repeated constricted migration or represent a pre-existing population selected by the constriction (Figure 23A). To evaluate whether either or both of these options could explain our observed migration results, we used agent-based modeling (ABM) to simulate a variety of initial population compositions and changes with migration (Figure 23B; see Materials and Methods for more detail). This model reports what fraction of cells would go through a transwell filter at each round of migration based on probabilistic choices of migration and division (Figure 23C). If the initial population contains a mixture of moderate and super-migrators, the migration efficiency of Top cells does not decline over the rounds of migration, which does not replicate the results of A375 cells, but does in some ways mimic the behavior of MDA-MB-231 cells (Figures 23B and 23D, first panel). The A375 experimental data can be recapitulated if the initial starting population is composed of three heterogeneous subpopulations (non-migrators, moderate migrators and super migrators) with highly migratory cells representing 0.5% of the population (Figure 23D, second panel). If this scenario is true, then our Bottom-5 specific 3D genome and nucleus structures were pre-

(A) A375 cell subpopulations observed after sequential constricted migration. (B) For comparison with model results, results of sequential transwell migration through 5μm constrictions are recapitulated from Figure 1 for A375 (left) and from Figure S3 for MDA-MB-231 (right) cells. (C) Choices a cell can make at each timestep during each simulated migration round. (D) Parameters used in each agent-based model (left columns) and results of each model (rightmost column).

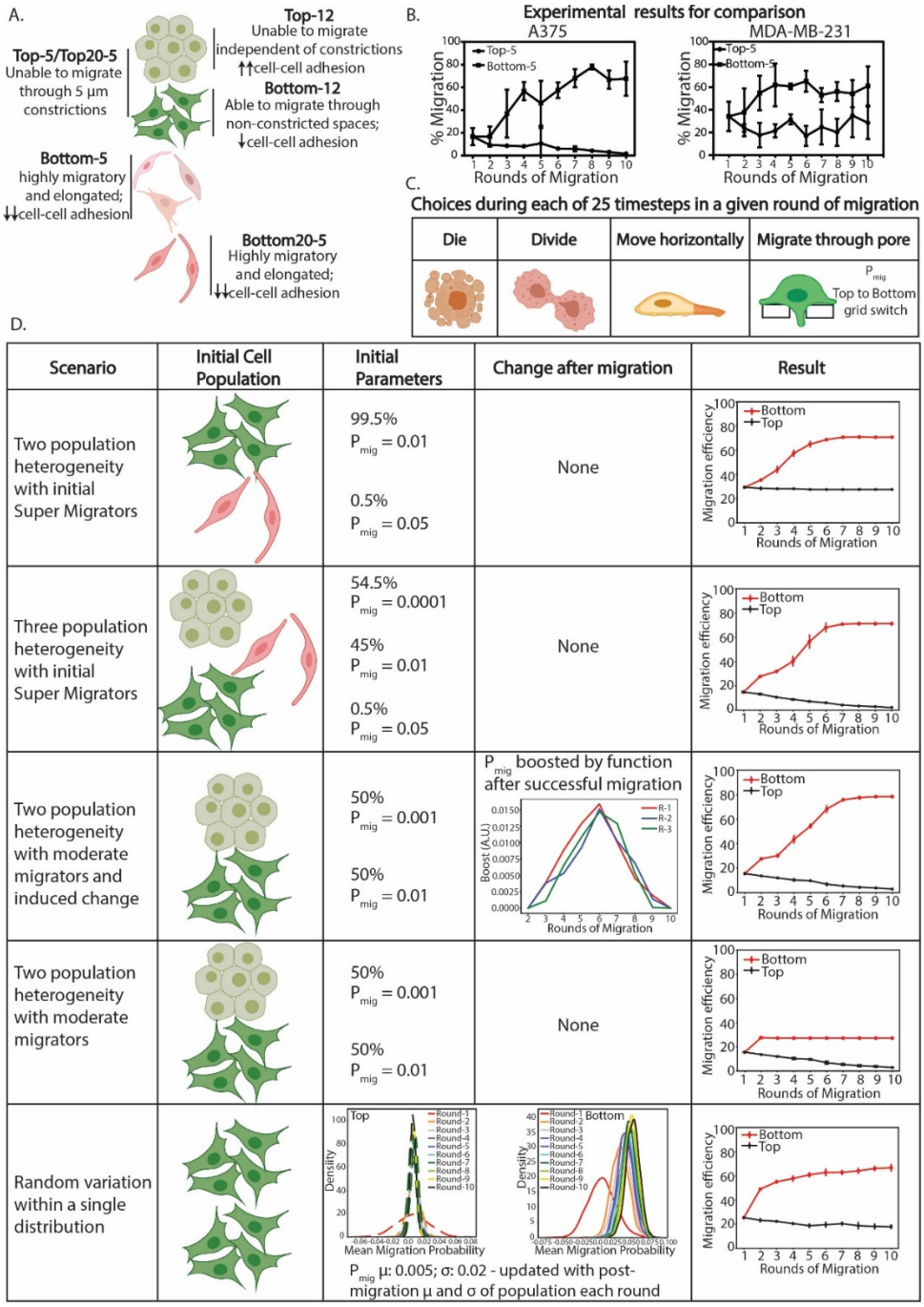


Figure 23. Agent based modeling of cell populations undergoing constricted migration.

existing in a small fraction of the initial population and primed these cells for confined migration. However, we are also able to recapitulate our experimental data with a scenario in which there are only non-migrators and moderate migrators in the initial population but then cells that squeeze through constrictions experience an induced change that alters their phenotype (boosting function shown in Figure 23D, third panel). In contrast, this same mixed population of non-migrators and moderate migrators without induced changes cannot recapitulate the increase in migration efficiency we observe (Figures 23D, fourth panel). Overall, our modeling results show that A375 cells must have an initial non-migratory subpopulation, but that the increase in migratory efficiency with constriction could either be a selection or induction process or a combination of both.

Discussion

Constricted migration is an important phenomenon in cancer metastasis and presents a major challenge to nuclear mechanics. Major avenues of previous research have separately investigated the role of the nucleus in confined migration, the impact of migration on linear DNA integrity, and the role of 3D genome structure in cancer. Our present results bridge these fields, revealing differences in spatial chromosome compartmentalization specific to cells that have passed through numerous constrictions smaller than their nucleus. These 3D genome structures are associated with notable differences in cell and nucleus phenotypes both when the cells are migrating on a 2D surface and when they are squeezing through a 3D collagen matrix. Cells proficient at confined migration more often deform their nuclei, elongate their cell body, and have fewer cell-cell attachments.

As described in our agent based modeling results, a question raised by our results is whether the differences in 3D genome structure we see in cells that have undergone confined migration (Bottom-5 and Bottom20-5) result from a selection process on an initially heterogeneous population or a “training process” where confined migration itself causes chromosome structure changes. Our ABM results indicate that either selection, training, or a combination could explain the differences observed in sequentially constricted cells. It is biologically plausible that cells which are initially primed for confined migration in a heterogeneous tumor begin to squeeze through tight junctions, and that then their chromosome structure is further rearranged as a result of the nucleus deformations. Within melanoma lesions, a select group of cells can gain the ability to metastasize, and that further diversification and phenotype change can occur as a result of metastasis (Damsky et al., 2010). Overall, our data suggest that both selection and training could be taking place.

As discussed, our isolation of A375 cells that cannot migrate at all, even given 10 chances to go through a wide pore (Top-12 cells), shows that the initial A375 population is heterogeneous, as expected from previous reports (Kozlowski et al., 1984). The genome structures and gene expression patterns specific to this subset informs us about the properties of cells that cannot migrate. On the other hand, the fact that Bottom-12 cells, which have passed through a large pore, present with genome structure differences compared to Bottom-5 cells shows that we are not simply distinguishing between cells with a pre-existing ability to migrate. There are cellular, nuclear, and genome

organization properties specific to confined migration. Further, some structures and phenotypes of Bottom-5 cells, such as the chromosome 5-13 translocation and extremely elongated cell shapes, are not detectable in the original control population and could be a result of induced phenotype changes or random events that are further reinforced by a selection process. As passage through a constriction physically moves the chromosomal regions leading to the possibility of physical mixing of domains in the nucleus. This provides loci previously more distant with an opportunity to associate. Recent work has demonstrated that attractions between heterochromatic regions are an essential part of the phase separation that organizes chromatin in the nucleus (Falk et al., 2019). During constriction, these domains could cause new associations to occur. This could be an explanation for the strengthening of interactions we observe between previously weak B compartments and stronger B compartment regions. Meanwhile, certain interactions may be pulled apart as suggested by chromatin stretching previously described (Irianto et al., 2017b) and loss of B-B interaction strength with confined migration observed in neutrophils (Jacobson et al., 2018). We note while genome-wide changes like compartment weakening might result directly from nucleus deformation and physical movement of chromatin domains, highly specific local compartment switches would be unlikely to have been pushed or pulled in the same way in every individual cell. Such specific changes could arise either by selection of pre-existing differences in the population, selection of a beneficial change that is induced by deformation in some cells during migration, or induction of a programmed cell response to constricted migration. It is known that mechanical signaling can induce cellular reprogramming and that cellular

reprogramming, such as an epithelial to mesenchymal cell transition (Shivashankar, 2019), can involve genome structure changes.

Our results suggest that the confined migration-specific 3D genome structure relates to several biological processes. First, we observe an overall correlation between spatial compartmentalization changes and gene expression alterations, and the altered genes relate to migratory ability. For example, in Bottom-5 cells, we observe downregulation of a group of genes important for cell-cell/matrix adhesion (CADM3, CADM1, ITGA9). Why would 3D genome changes be needed to accompany gene expression changes? Altered local gene regulation alone can be sufficient to explain temporary changes in cellular behaviors in response to a stress or stimulus, and these changes do not require 3D genome rearrangement (Jin et al., 2013). But our data suggests that migration behaviors of the different A375 sub-populations are stably divergent. From our observations, we propose that the stability of this phenotype is encoded in the 3D genome structure change we measure, since rearrangement of whole genomic to different spatial and chromatin environments is less reversible than transiently regulating individual genes. Such stable differences in 3D genome structure after confined migration could underlie the distinct phenotype and increased in vitro invasiveness of cancer cells derived from metastatic sites compared to those from primary tumors or normal tissue (Oppenheimer, 2006).

Many genomic regions with altered compartmentalization in sequentially constricted cells do not show a connection to gene regulation. In fact, we observe a striking enrichment

of regions with no genes at all among the set of regions that switch from the A to the B compartment in confined migration. This suggests that the genome structure differences we observe could have a role beyond gene regulation and could play a role in modulating the physical properties of the cancer cell nucleus. Indeed, we observe that the nuclei of Bottom-5 cells deform more than Top-5 cells when embedded in a collagen matrix or even when spontaneously migrating on a 2D surface. Previous research has shown that nucleus stiffness can be altered by increasing or decreasing the overall expression of Lamin A/C or by changing overall levels of heterochromatin (Lammerding et al., 2006, Davidson et al., 2014, Stephens et al., 2017, Stephens et al., 2018). But, the apparent change in nucleus pliability observed in the current study cannot be explained by changes in overall Lamin A/C content or heterochromatin levels: Lamin A/C and H3K9me3 levels were unchanged as measured by both immunostaining and western blot. Instead, we observed a redistribution of Lamin A/C and H3K9me3 localization in the nucleus, with individual heterochromatin foci originally in the interior of the nucleus merging at the periphery and Lamin A/C becoming less evenly distributed. These observations in A375 cells, also recapitulated in MDA-MB-231 cells, suggest that changing levels of Lamin A/C and heterochromatin are not always required to increase constricted migration potential and that re-localization of these proteins and chromosome regions could also be important to the ability of the nucleus to undergo physical deformations. The visibly altered heterochromatin localization in sequentially constricted cells may relate to differences in chromosome compartmentalization we observe with Hi-C. We observe that regions with stronger B compartment identity gain more distal interactions, consistent with

the observation that disparate heterochromatin foci come together at the nuclear periphery in sequentially constricted cells. Future work will be required to determine whether the 3D genome changes we observe translate into quantitative differences in nucleus physical properties.

Our data also provide insight about basic 3D genome organization principles. As described above, our data allow conceptual links between changes in contacts observed in Hi-C data and the appearance and localization of chromatin types observed by imaging. Further, our observation that compartments, compartment strength, and whole chromosome interactions change while TADs are unaltered reinforces the idea that these are separate and largely independent layers of genome organization (Mirny et al., 2019, Schwarzer et al., 2017). These changes also suggest that the compartment and whole chromosome level of genome structure may be more important to nuclear mechanics than the TAD and loop level of organization. Meanwhile, our observation that local punctate interactions change around sites where gene expression changes echo previous evidence that collisions between loop extrusion and transcription can influence local contact patterns (Brandao et al., 2019).

Are the 3D genome changes we observe in our constricted migration cell population diagnostic of metastatic potential? Previous reports have sometimes detected gene expression signatures in primary tumors that may be predictive of metastasis (Ramaswamy et al., 2003). Nuclear morphology abnormalities (“nuclear atypia”) are also

already used clinically as a marker of cancer aggressiveness (Kadota et al., 2012). Recent reports have taken advantage of microfluidic engineering to predict metastatic potential of breast cancer cells based on their morphological features (Yankaskas et al., 2019). But, beyond being a marker, the biological significance of such abnormalities and what chromosome structure changes accompany them have not been defined. Here, we provide evidence that chromosome structure changes could link metastatic gene expression signatures with the abnormal nuclear appearance of aggressive cancer. Our results reveal chromosome spatial compartmentalization differences in genomic regions related to metastatic potential in a highly invasive subset of A375 cells. Future work will be needed to investigate whether such changes are also observed in patient metastatic melanoma samples. Our data provides a foundation for future investigation of whether similar structural changes accompany constricted migration in other cancer types and therefore constitute 3D genome structure signatures of metastatic cancer potential.

Acknowledgements

We thank Mariano Labrador, Jacob Sanders, and Yang Xu for insightful discussion. We thank Jan Lammerding for advice about the project. We thank Bas van Steensel for providing the LAD type annotations. BJ1-hTERT cells were provided by Adayabalam Balajee at Oak Ridge Institute of Science and Education. Confocal images were obtained at the University of Tennessee Advanced Microscopy Imaging Core. This research was supported in part by a Ralph E. Powe Junior Faculty Enhancement Award from Oak Ridge Associated Universities to R.P.M. and by NIH NIGMS grant R35GM133557 to R.P.M.

R.G. is supported in part by a Yates Dissertation Fellowship from UTK. D.T. was supported in part by a UTK Summer Undergraduate Research Internship award.

Materials and Methods

Cell Lines and Cell Culture

A375 and MDA-MB-231 cells were obtained from ATCC (CRL-1619 and HTB-26, respectively). Cells were grown using complete DMEM medium (Corning – 10-013-CV; 10% FBS, 1% Pen-Strep, 1% L-Glutamine) at 37^{o}C supplied with 5% CO_2.

Sequential Transwell Migration

For the sequential transwell migration, we used transwells with 12 μm (VWR-10769-224) and 5 μm pore sizes (VWR-10769-236). Briefly, the bottom of the transwells were coated with 40 μL of 10 μg/mL fibronectin for ~ 45 minutes. 24 well plates were prepared for transwell migration assay adding 500 μL of 1x DMEM (Corning) with full supplements per well. A375 cells were detached from culture dishes at 80-90% confluency and aliquoted to 100,000 cells per 100 μL of 1xDMEM. Each transwell was placed into its corresponding well of the 24 well plate and 100 μL of the cell suspension was added to the top of each filter. Cells were incubated at 37^{o}C, 5% CO_2 and allowed to migrate for 24 hours. After the 24-hour incubation, migration efficiency was quantified as follows. First, freely floating cells were removed from the top of the filter (unmigrated; “Top” cells) and from the well beneath the filter (migrated “Bottom” cells) and placed in two separate tubes. Then, 400 μL of trypsin was added into the bottom chamber of the 24 well plate and 200 μL of trypsin was added into the top chamber to detach any

remaining attached cells. Recovered cells after trypsinization were added to the unmigrated or migrated tubes, accordingly. Cells were spun down (1000 rpm, 5min) and counted (using trypan blue) to calculate % migration such as: #bottom/(#top+#bottom). Additionally, a small aliquot of cells was saved for immunofluorescence (IF). The rest of the cells were seeded into wells of a 24-well plate to expand. When cells reached 80-90% confluency, another transwell migration was performed (R2). A375 Top and Bottom cells were detached and counted. Two transwell filters were prepared as previously described (one for Top and one for Bottom). Then 100,000 cells suspended in 100 µL of 1xDMEM without supplements were seeded into the transwell filters. After 24-hour incubation at 37^{o}C and 5% CO_2, cells were trypsinized as previously described to quantify migration efficiency. Only the cells that were always on top (Top of Top) and the cells that were always on the bottom (Bottom of Bottom) were saved and grown for further sequential rounds of migration. This process was repeated for 10 and 20 rounds of migration which lead to the generation of A375-Top5, Bottom-5, Top20-5, Bottom20-5, Top-12 and Bottom-12.

2D Single Cell migration

Live cell imaging

To track single cell movement in 2D, A375-Control, Top-5 and Bottom-5 cells were seeded on wells of a 6-well plate at a 30,000 cells/well density. After cells were attached, live cell imaging was performed using the EVOS FL Auto microscope at 40X magnification. Images were acquired every 10 minutes for a 24-hour time period.

Cell morphology analysis

Parameters that describe cell morphology such as solidity and aspect ratio were quantified using the Shape Descriptors plugin in ImageJ.

Immunocytochemistry

IF staining

Approximately 50,000 – 100,000 cells for each sub-population of A375 were seeded into either poly-D-lysine treated coverslips or 35-mm coverslip bottom dishes. Cells were allowed to attach overnight and then crosslinked with 4% formaldehyde for 10 min followed by three, 5-minute washes with PBS. After washing, cells were permeabilized with permeabilization buffer (10% goat serum, 0.5% Triton in PBS) for 1 hour at room temperature. After the incubation, primary antibodies (Lamin A – sc-376248, Lamin B1 – ab133741, H3K9me3 – ab176916), diluted in antibody dilution buffer (5% Goat serum, 0.25% Triton in PBS) were added and incubated overnight at 4^{o}C. After primary antibody incubation, cells were washed 3 times with PBS for 5 minutes for each wash. Cells were then incubated with secondary antibodies (Alexa Fluor 594 and Alexa Fluor 488) per manufacturer's directions (2 drops of secondary antibody/1mL of PBS – R37117 and R37120) for 30 minutes at room temperature. After secondary antibody incubation, cells were washed 3 times with PBS and sealed using mounting media with DAPI. Slides were treated in the mounting media for 24 hours before imaging. Cells were imaged using a Leica Sp8 Confocal microscope was equipped with a 63x oil immersion objective.

Nuclear morphology analysis

To analyze the diameter (minor axis), aspect ratio and standard deviation of maximum projected nuclei, Image J Shape Descriptor plugin was used. For analysis of 3D projected nuclei, Image J Nucleus J plugin was used with default parameters. Lamin A/C and H3K9me3 radial distribution was quantified using Measure Object Intensity Distribution module in Cell Profiler. The nuclei were separated into four equidistant bins. Mean fractional intensities for each bin (4 total bins; starting from innermost (bin1) to outermost (bin4) radial position) was plotted.

Western Blot

A375 cells were trypsinized as previously described, and approximately 5 million cells were collected by centrifugation (1000 rcf, 5 min, room temperature). The cell pellets were resuspended in 300 ml of RIPA buffer (Thermo Scientific, I89900), supplemented with protease inhibitors/EDTA (GenDepot 50-101-5485), and phosphatase inhibitors (GenDepot 50-101-5488). Cells were then incubated on ice for 10 minutes. DNA was degraded by adding 6µl of micrococcal nuclease (0.5 U/ml stock, Thermo Scientific, FEREN0181) and 12µl of calcium chloride (100 mM stock) to each lysate, followed by incubation at 37°C for 15 min. Nuclease activity was inhibited by subsequent incubations at 68 °C for 15 min, and ice for 10 min. Protein samples were stored at −80 °C until use. Protein concentration was measured using the BCA Protein Assay Kit. Thermo (234225), according to the manufacturer's instructions.

Denatured protein samples were resolved using a mini gel tank (Invitrogen, A25977) for 35 min at 200V, using 4-12% Bis-Tris Plus gels (Invitrogen, NW04120BOX) for all targets.

Protein was transferred to low fluorescence PVDF membranes using the mini Bolt Blotting System (Invitrogen, B1000), along with system-specific reagents. Blotting was performed using the Odyssey TBS blocking system (Licor, 927-400000), according to the manufacturer's protocol. Briefly, membranes were activated with methanol after transfer (1 min, RT), washed twice with milliQ water (5 min), twice with TBS (2 min), and blocked with Odyssey TBS blocking solution for 1 hour at RT. Primary antibodies were diluted in blocking buffer (containing 0.2% Tween), and incubated over night at 4°C as follows: H3K9me3 (1:1000, ab176916), LaminA/C (1:1000, sc-376248) and LaminB1 (1:1000, ab133741). Beta actin was used as a loading control, using the appropriate antibody (1:10,000; rabbit polyclonal PA1-16889, Thermo Fisher or 1:10,000; mouse monoclonal MA1-140, Thermo Fisher) alongside with the targets of interest. Detection was carried out using the following secondary antibodies: goat anti-rabbit (1:10,000; IRDye 680RD [92568071]; Licor) and goat anti-mouse (1:10,000, IRDye 800CW [95-32210]; Licor). Secondary antibodies were diluted in Odyssey TBS blocking buffer containing 0.2% Tween and 0.01% SDS. Membranes were incubated in secondary antibody for 1 hour at RT, followed by three washes (TBST, 5 min each). Images of near infrared fluorescent signal were acquired using an Odyssey scanner (Licor) in both the red and green channels. Signal was quantified using the Image Studio software (Licor).

Hi-C

Library Preparation

Hi-C libraries were prepared for each cell sub-population (A375-Control, Top-5, Bottom-5, Top20-5, Bottom20-5, Top-12 and Bottom-12) as previously described (Golloshi et al.,

2018) including two biological duplicates for each condition. Briefly, 10 million cells were grown in T-75 cell culture flasks to 80-90% confluency and crosslinked with 1% formaldehyde for 10 min. Crosslinked cells were then suspended in lysis buffer to permeabilize the cell membrane and were dounce homogenized. Chromatin was then digested in-nucleus overnight using DpnII restriction enzyme. Digested ends were filled in with biotin-dATP, and the blunt ends of interacting fragments were ligated together. DNA was then purified by phenol-chloroform extraction. For library preparation, the NEBNext Ultra II DNA Library prep kit (NEB) was used for libraries with size ranges from 200-400bp. End Prep, Adaptor Ligation, and PCR amplification reactions were carried out on bead bound DNA libraries.

Sequencing was performed at the Oklahoma Medical Research Foundation Clinical Genomics facility using the Illumina HiSeq 3000 platform with 75 bp paired end reads or a NovaSeq 6000 with 50 bp paired end reads. Sequencing reads were mapped to the human genome (hg19), filtered, and iteratively corrected as previously described (Imakaev et al., 2012) (https://github.com/dekkerlab/cMapping). All Hi-C contact matrices were scaled to a sum of 1 million interactions to allow comparisons between conditions in downstream analysis. For library quality and mapping statistics see Table 2.

Compartment Analysis

Compartment analysis was performed by running principal component analysis using matrix2compartment.pl script in the cworld-dekker pipeline available on GitHub

Table 2. Summary statistics for Hi-C sequencing experiments (Chapter 2).

Sample	Total Raw Reads	Both Sides Mapped	Valid Pairs	Unique Valid Pairs	% Cis Interactions	% Dangling Ends
Control-R1	201,352,157	124,670,760	123,653,686	107,358,325	61.2	0.35
Control-R2	200,881,087	102,710,892	101,397,252	80,394,938	47.7	0.31
Top5-R1	320,786,521	163,004,931	161,788,811	132,650,834	55.6	0.23
Bottom5-R1	386,733,576	193,207,211	191,853,175	152,440,505	60.3	0.21
Top5-R2	164,616,353	84,226,628	83,087,615	68,000,595	48.7	0.28
Bottom5-R2	107,320,517	54,380,776	53,662,580	43,911,557	49.2	0.45
Top20-5-R1	142,531,635	66,090,879	62,567,577	49,317,732	47.7	0.60
Bottom20-5-R1	159,577,275	73,992,021	70,695,806	57,329,882	49.7	0.44
Top12-R1	117,695,787	90,932,905	90,442,670	81,815,020	43.0	0.39
Bottom12R1	110,637,334	85,346,366	84,975,979	76,728,534	37.0	0.30
Top12-R2	145,063,083	110,772,496	110,380,318	98,396,558	40.4	0.26
Bottom12R2	126,055,079	96,662,104	96,161,324	86,277,860	43.5	0.34

(https://github.com/dekkerlab/cworld-dekker). The PC1 value was then used to determine compartment identity for 100 and 250Kb binned matrices. We considered PC1 values greater than 0.01 to be A compartment and less than -0.01 to be B compartment for this analysis and determination of compartment switches.

Saddle Plots

Saddle plots were constructed to investigate changes in interaction frequency between or within compartments in highly migratory and non-migratory melanoma cells. Briefly, PCA analysis was performed on 100Kb binned matrices to assign compartment identity for each bin. For the analysis, interaction zScores, normalizing for interaction decay for genomic distance, were calculated according to the matrix2compartment.pl script in cworld-dekker. Zscore matrices were reordered based on compartment strength (from strongest B to strongest A using eigenvector values). Finally, matrices were smoothed at 500Kb bin size.

Interchromosomal Fraction (ICF)

Ratio of interchromosomal interactions relative to the total number of interactions for a given chromosome as previously described (Heinz et al., 2018).

$$ICF = \frac{Interchromosomal\ Interactions}{Inter + Intra\ chromosomal\ Interactions}$$

$$\text{deltaICF} = ICF_{Bottom} - ICF_{Top}$$

Distal to Local Ratio (DLR)

Log2 ratio of HiC interactions with distance greater than 3Mb relative to local interactions less than 3 Mb away. To find the change in DLR after constricted migration, delta DLR was calculated as $DLR_{Bottom} - DLR_{Top}$ as previously described (Heinz et al., 2018).

$$DLR = log2\frac{Distal\ Interactions\ (> 3Mb)}{Local\ Interactions\ (< 3Mb)}$$

$$deltaDLR = DLRBottom - DLRTop$$

Copy Number Variation Estimation

CNV estimation was performed using the approach available from https://github.com/nservant/cancer-hic-norm (Servant et al., 2018)

Translocation Detection and Frequency Calculation

For the novel translocation visibly evident in 2.5 Mb binned contact maps for Bottom-5 constricted migrated cells, we further pinpointed the specific break site on each chromosome (chr5 and chr13) using raw valid pair counts within the region. First, we filtered all valid pair interactions between chr5 and chr13 to include only regions of chr13 between 38.3 and 57.5 Mb (a range of coordinates upstream of the translocation site, which all interact more than expected with chr5). Then, we plotted a histogram of interaction frequencies between each locus along chr5 and this set of chr13 coordinates. We noted a sharp interaction drop within the region chr5:12.24-12.25 Mb, which represents the translocation breakpoint. We then reversed this process to determine the breakpoint along chr13.

To calculate the percent of chromosomes containing this chr5-13 translocation in each subpopulation, we first found the 2.5 Mb bin with the maximum raw Hi-C counts near the translocation site in Bottom-5 cells. We then calculated an average value for raw Hi-C counts between neighboring bins (one bin off diagonal) in cis on chr7. We then calculated the percent of chromosomes bearing the translocation in any given cell type as the ratio of raw interaction counts at the translocation bin over average interaction counts between neighbors in cis. Given the median triploid karyotype of A375 cells, we then multiplied this % of chromosomes by 3 to estimate a % of cells carrying the translocation.

Distance Decay Scaling Plots

Using 40 and 20 Kb binned iteratively corrected and scaled contact matrices, we extracted contact frequencies between bins at each genomic distance, excluding the diagonal bin (zero distance). A loess fit was then used to find a smooth curve describing interaction decay vs. distance (using the matrix2loess.pl script in cworld-dekker). The interaction frequencies were then normalized to set the maximum value (loess fit interaction value for the minimum distance) for each dataset to 1 and then plotted on a log scale vs. log genomic distance.

RNA-Seq

Library Preparation

For Control, Top-5 and Bottom-5 A375 cells, RNA was extracted, and the library was prepared in our laboratory. Briefly, RNA was extracted using Qiagen RNAeasy plus mini kit (Cat. No. 74134). Cells were lysed and homogenized, spun down in gDNA eliminator

columns to remove any genomic DNA. All samples were washed with ethanol and the total RNA was eluted. The quality and quantity of the RNA were quantified using Agilent Bioanalyzer Nano RNA kit. All the libraries used were characterized by a RIN value between 9 and 10. rRNA was depleted using NebNext rRNA Depletion Kit (Cat. No. E6310S). Total RNA was hybridized to the probe followed by RNaseH and DNase I digestion. RNA was then purified using AmpureXP beads from Beckman Coulter (Ref. A63881) at a 2.2x concentration of beads. RNA libraries were prepared using NebNext Ultra II RNA library prep Kit (Cat. No. E7770G). After rRNA depletion, RNA was fragmented (~200 nt) and primed following first and second strand cDNA synthesis. The double stranded cDNA was then purified using AmpureXP beads at a 1.8X concentration. After purification, cDNA libraries were end prepped, ligated adapter and PCR amplified for 7 cycles to enrichment for adaptor ligated DNA using NebNext Multiplex Oligos for illumine (Cat. No. E7335G). cDNA librariers were then purified using AmpureXP beads at a 0.9X concentration and their quality was checked using Agilent Bioanalyzer high sensitivity DNA kit. cDNA samples were then sent to Oklahoma Medical Research Facility (OMRF) for high throughput sequencing.

For the rest of A375 libraries, Top20-5, Bottom20-5, Top-12 and Bottom-12, RNA was extracted using the Qiagen RNAeasy plus mini kit as described above. Total RNA was then sent for library preparation and sequencing at GeneWiz. For quality of RNA-Seq libraries see Table 3.

Table 3. Summary statistics for RNA sequencing experiments (Chapter 2).

Sample	Total Reads after rRNA removal	Mapped Reads	Percent Mapped Reads
Control-R1	25,802,932	23,454,526	90.9
Control-R2	44,936,964	39,168,478	87.16
Top5-R1	57,420,692	49,189,316	85.66
Top5-R2	40,240,820	34,943,927	86.84
Bottom5-R1	44,784,816	37,982,597	84.81
Bottom5-R2	42,149,066	36,445,763	86.47
Top20-5-R1	39,235,465	36,102,375	92.01
Top20-5-R2	97,958,087	89,841,371	91.71
Bottom20-5-R1	107,282,198	98,413,826	91.73
Bottom20-5-R2	54,332,623	49,381,035	90.89
Top12-R1	47,231,386	42,228,487	89.41
Top12-R2	44,863,196	41,601,169	92.73
Bottom12-R1	43,203,103	38,744,667	89.68
Bottom12-R2	8,436,059	7,655,956	90.75

RNA-Seq Data Analysis

Quality of reads was checked using Fastqc and adapter sequences were trimmed using BBTools (https://github.com/kbaseapps/BBTools). Additionally, quality trimming of the reads was performed and any reads with a lower than 28 quality score were discarded. Reads were then aligned using STAR alignment (https://github.com/alexdobin/STAR) with an average alignment rate of 89%. The aligned reads were then sorted by genomic position and feature counts was performed using htseq 0.11.1(https://github.com/simon-anders/htseq). Differential expression of genes was determined by using DESeq through Galaxy.

ATAC-Seq

ATAC-Seq libraries for Top-5 and Bottom-5 cells was prepared as previously described (Buenrostro et al., 2015). Libraries were generated using Ad1_noMX and Ad2.1-2.3 barcoded primers from (Buenrostro et al., 2015) and were amplified for a total of 12-13 cycles. The quality and quantity of the libraries were measured using Agilent Bioanalyzer High Sensitivity DNA kit. Sequencing was performed at Oklahoma Medical Research Facility using HiSeq3000 with 75 bp paired end reads. For quality of ATAC-Seq libraries see Table 3.

ATAC-Seq Data Analysis

Adapters and reads with a quality score less than 20 were trimmed from the final libraries using Skewer/0.2.2 (https://github.com/relipmoc/skewer). Trimmed reads were aligned to

Table 4. Summary statistics for ATAC sequencing experiments (Chapter 2).

Sample	Total Reads	Reads after trimming	Mapped Reads	% Mapped Reads
Control-R1	169,579,026	169,348,735	137,428,769	81.15
Top5-R1	169,763,417	113,906,483	94,322,265	82.81
Bottom5-R1	213,830,754	115,332,896	108,423,723	94.01

the hg19 genome using bowtie2 with default parameters (https://github.com/BenLangmead/bowtie2). Aligned reads were sorted, indexed and ChrM reads were removed using samtools. Duplicated reads were removed using Picard (https://github.com/broadinstitute/picard/blob/master/src/main/java/picard/sam/markduplicates/MarkDuplicates.java). Finally, peaks were called using MACS2 (https://github.com/taoliu/MACS) with the following parameters: --nomodel --extsize 300 -g hs --keep-dup all.

3D Collagen Matrix

Experiments investigating the morphology and migration of A375 subpopulation of cells in 3D environment were performed using single cell suspension in 3D collagen matrices (A375-Control, A375-Top5 A375-Bottom10 generated from passage of A375 through 5 and 12 μm sized pore transwells) as previously described (Denais et al., 2016). Briefly, 75,000 cells/condition were added into collagen matrices (Rat Tail, Corning 354236) with a density of 2.32 mg/ml. The pH of the collagen gels was normalized to a neutral pH and gels were incubated at 37^0C for 30 min to gel. After the incubation, media with full supplements was supplied to the collagen gels. Cell were allowed to migrate through the collagen for 5 days before the cells were stained and imaged.

Immunofluorescence for cells embedded in 3D collagen matrices

Cells embedded on collagen were stained as previously described (Denais et al., 2016). Briefly, collagen matrices were fixed in 4% formaldehyde for 30 min at 37C. Collagen matrices were then washed with 1xPBS for 10 min for three times. Each collagen matrix was then incubated with blocking buffer (0.5% Triton, 10% Goat Serum in 1xPBS) overnight at 4C. After blocking, samples were then incubated with primary antibody (LaminA – sc376248) at a 1:500 dilution in 1xBlocking Buffer:1xPBS at 4C overnight. After incubation, aspirate primary antibody solution and wash three times with 1xPBS for 10 min each. Samples were then incubated in secondary antibody (2 drops/mL of PBS) for 1hr and room temperature. After secondary antibody incubation, samples were then washed with 1xPBS three times for 10 min each. Samples were incubated with DAPI (1µg/mL) for 30 min at RT. Collagen matrices were then washed three times with 1xPBS for 10 min each. Z-stacks of collagen matrices were taken with Leica SP8 confocal microscope at 40x water objective.

Image analysis for cells embedded in 3D collagen matrices

Aspect ratio and Lamin A/C standard deviation of maximum projected nuclei was quantified using ImageJ plugin Shape Descriptors. Tracking of cells migrating in 3D collagen matrices was performed using ImageJ Manual Tracking plugin. The resulting data was then used in Chemotaxis and Migration tool (Ibidi) to generate accumulated distance and velocity measurements.

Agent Based Modeling of Constricted Migration

We model constricted cell migration using an on-lattice two-dimensional (2D) agent-based model (ABM) system. In this model, cells are represented by agents. The basic building block of this agent-based system is obtained from the collective cell migration ABM (collectiveABM), available on GitHub (https://github.com/zcolburn/cell-sheet-abm) and necessary changes are made to model cell transmigration. Most of the parameter values and the model working procedures of collectiveABM are kept untouched unless otherwise mentioned. Since the collectiveABM models a wound healing assay, we exclude the wound scratch section of the code from our model. We also add two squared 2D grid spaces of size 20x20 to the agent-based system in place of a single 2D grid space, as we are using transwell migration assay in this work for the constricted migration. One of those two grid spaces represent the top compartment, where the cells are seeded at the beginning on each round of sequential migration and another grid space is used as the bottom compartment, where the cells transmigrate from the top compartment. The beginning of each round of sequential migration model starts with 2000 cells seeded on the top compartment. The population of the cells is either considered to be homogenous or heterogeneous depending on the transmigration probability (ptmig). A total of 25 timepoint iterations are implemented to mimic the ~24 hours duration of experimental rounds. After each round of sequential migrations, the proportion of different cells in the bottom or top compartment depending on the sequential transwell migration experiment type is measured and the same proportion is maintained during the cell seeding in the next round. The directionality of the transmigration is only

from top to bottom compartment and once the cells reach the bottom compartment, they stay there. The constricted migration of a cell is modeled by comparing its assigned ptmig with a uniform random number drawn from the range [0,1). If the random number is less than ptmig, then that cell will transmigrate to the bottom compartment from the top. In addition to having models with fixed ptmig values, we also designed agent-based model systems having variable ptmig : 1) the ptmig value of bottom or bottom cells are boosted by random values after each round and 2) each cell has different ptmig, drawn from a normal distribution and after each round, the mean and standard deviation of top or bottom compartment cells depending on the sequential transwell migration experiment type will be used in next round's normal distribution. The efficiency of migration is calculated at the end of each round by dividing the number of cells in the bottom compartment by total cells in the top and bottom compartments combined. For all the different test cases, three independent runs are performed by using different random number seeds.

References

ABBAS, A., HE, X., NIU, J., ZHOU, B., ZHU, G., MA, T., SONG, J., GAO, J., ZHANG, M. Q. & ZENG, J. 2019. Integrating Hi-C and FISH data for modeling of the 3D organization of chromosomes. *Nat Commun,* 10**,** 2049.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-9.

BARUTCU, A. R., LAJOIE, B. R., MCCORD, R. P., TYE, C. E., HONG, D., MESSIER, T. L., BROWNE, G., VAN WIJNEN, A. J., LIAN, J. B., STEIN, J. L., DEKKER, J., IMBALZANO, A. N. & STEIN, G. S. 2015. Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol,* 16**,** 214.

BASKARAN, J. P., WELDY, A., GUARIN, J., MUNOZ, G., SHPILKER, P. H., KOTLIK, M., SUBBIAH, N., WISHART, A., PENG, Y., MILLER, M. A., COWEN, L. & OUDIN, M. J. 2020. Cell shape, and not 2D migration, predicts extracellular matrix-driven 3D cell invasion in breast cancer. *APL Bioeng,* 4**,** 026105.

BELAGHZAL, H., BORRMAN, T., STEPHENS, A. D., LAFONTAINE, D. L., VENEV, S. V., WENG, Z., MARKO, J. F. & DEKKER, J. 2019. Compartment-dependent chromatin interaction dynamics revealed by liquid chromatin Hi-C. *bioRxiv***,** 704957.

BICKMORE, W. A. & VAN STEENSEL, B. 2013. Genome architecture: domain organization of interphase chromosomes. *Cell,* 152**,** 1270-84.

BRANDAO, H. B., PAUL, P., VAN DEN BERG, A. A., RUDNER, D. Z., WANG, X. & MIRNY, L. A. 2019. RNA polymerases as moving barriers to condensin loop extrusion. *Proc Natl Acad Sci U S A,* 116**,** 20489-20499.

BUENROSTRO, J. D., WU, B., CHANG, H. Y. & GREENLEAF, W. J. 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol,* 109**,** 21 29 1-9.

CAROLYN DE GRAAF, J. K., LUDO PAGIE, DAAN PERIC HUPKES, BAS VAN STEENSEL. 2019. *LAD Atlas* [Online]. Available: https://osf.io/dk8pm/ [Accessed].

CHAMBERS, A. F., GROOM, A. C. & MACDONALD, I. C. 2002. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer,* 2**,** 563-72.

DAMSKY, W. E., ROSENBAUM, L. E. & BOSENBERG, M. 2010. Decoding melanoma metastasis. *Cancers,* 3**,** 126-163.

DAVIDSON, P. M., DENAIS, C., BAKSHI, M. C. & LAMMERDING, J. 2014. Nuclear deformability constitutes a rate-limiting step during cell migration in 3-D environments. *Cell Mol Bioeng,* 7**,** 293-306.

DAVIDSON, P. M. & LAMMERDING, J. 2014. Broken nuclei--lamins, nuclear mechanics, and disease. *Trends Cell Biol,* 24**,** 247-56.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

DENAIS, C. M., GILBERT, R. M., ISERMANN, P., MCGREGOR, A. L., TE LINDERT, M., WEIGELIN, B., DAVIDSON, P. M., FRIEDL, P., WOLF, K. & LAMMERDING, J.

2016. Nuclear envelope rupture and repair during cancer cell migration. *Science,* 352**,** 353-8.

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

FALK, M., FEODOROVA, Y., NAUMOVA, N., IMAKAEV, M., LAJOIE, B. R., LEONHARDT, H., JOFFE, B., DEKKER, J., FUDENBERG, G., SOLOVEI, I. & MIRNY, L. A. 2019. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature,* 570**,** 395-399.

FIDLER, I. J. 2003. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer,* 3**,** 453-8.

FRIEDL, P., WOLF, K. & LAMMERDING, J. 2011. Nuclear mechanics during cell migration. *Curr Opin Cell Biol,* 23**,** 55-64.

FU, Y., CHIN, L. K., BOUROUINA, T., LIU, A. Q. & VANDONGEN, A. M. 2012. Nuclear deformation during breast cancer cell transmigration. *Lab Chip,* 12**,** 3774-8.

GASPAR-MAIA, A., ALAJEM, A., MESHORER, E. & RAMALHO-SANTOS, M. 2011. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol,* 12**,** 36-47.

GERLITZ, G. & BUSTIN, M. 2010. Efficient cell migration requires global chromatin condensation. *J Cell Sci,* 123**,** 2207-17.

GOLLOSHI, R., SANDERS, J. T. & MCCORD, R. P. 2018. Iteratively improving Hi-C experiments one step at a time. *Methods,* 142**,** 47-58.

GUPTA, G. P. & MASSAGUE, J. 2006. Cancer metastasis: building a framework. *Cell,* 127**,** 679-95.

HARADA, T., SWIFT, J., IRIANTO, J., SHIN, J. W., SPINLER, K. R., ATHIRASALA, A., DIEGMILLER, R., DINGAL, P. C., IVANOVSKA, I. L. & DISCHER, D. E. 2014. Nuclear lamin stiffness is a barrier to 3D migration, but softness can limit survival. *J Cell Biol,* 204**,** 669-82.

HEINZ, S., TEXARI, L., HAYES, M. G. B., URBANOWSKI, M., CHANG, M. W., GIVARKES, N., RIALDI, A., WHITE, K. M., ALBRECHT, R. A., PACHE, L., MARAZZI, I., GARCIA-SASTRE, A., SHAW, M. L. & BENNER, C. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell,* 174**,** 1522-1536 e22.

HNISZ, D., WEINTRAUB, A. S., DAY, D. S., VALTON, A. L., BAK, R. O., LI, C. H., GOLDMANN, J., LAJOIE, B. R., FAN, Z. P., SIGOVA, A. A., REDDY, J., BORGES-RIVERA, D., LEE, T. I., JAENISCH, R., PORTEUS, M. H., DEKKER, J. & YOUNG, R. A. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science,* 351**,** 1454-8.

IMAKAEV, M., FUDENBERG, G., MCCORD, R. P., NAUMOVA, N., GOLOBORODKO, A., LAJOIE, B. R., DEKKER, J. & MIRNY, L. A. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods,* 9**,** 999-1003.

IRIANTO, J., XIA, Y., PFEIFER, C. R., ATHIRASALA, A., JI, J., ALVEY, C., TEWARI, M., BENNETT, R. R., HARDING, S. M., LIU, A. J., GREENBERG, R. A. & DISCHER,

D. E. 2017a. DNA Damage Follows Repair Factor Depletion and Portends Genome Variation in Cancer Cells after Pore Migration. *Curr Biol,* 27**,** 210-223.

IRIANTO, J., XIA, Y., PFEIFER, C. R., GREENBERG, R. A. & DISCHER, D. E. 2017b. As a Nucleus Enters a Small Pore, Chromatin Stretches and Maintains Integrity, Even with DNA Breaks. *Biophys J,* 112**,** 446-449.

JACOBSON, E. C., PERRY, J. K., LONG, D. S., OLINS, A. L., OLINS, D. E., WRIGHT, B. E., VICKERS, M. H. & O'SULLIVAN, J. M. 2018. Migration through a small pore disrupts inactive chromatin organization in neutrophil-like cells. *BMC Biol,* 16**,** 142.

JIN, F., LI, Y., DIXON, J. R., SELVARAJ, S., YE, Z., LEE, A. Y., YEN, C. A., SCHMITT, A. D., ESPINOZA, C. A. & REN, B. 2013. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature,* 503**,** 290-4.

KADOTA, K., SUZUKI, K., COLOVOS, C., SIMA, C. S., RUSCH, V. W., TRAVIS, W. D. & ADUSUMILLI, P. S. 2012. A nuclear grading system is a strong predictor of survival in epitheloid diffuse malignant pleural mesothelioma. *Mod Pathol,* 25**,** 260-71.

KIM, D. H. & WIRTZ, D. 2013. Focal adhesion size uniquely predicts cell migration. *FASEB J,* 27**,** 1351-61.

KIND, J., PAGIE, L., DE VRIES, S. S., NAHIDIAZAR, L., DEY, S. S., BIENKO, M., ZHAN, Y., LAJOIE, B., DE GRAAF, C. A., AMENDOLA, M., FUDENBERG, G., IMAKAEV, M., MIRNY, L. A., JALINK, K., DEKKER, J., VAN OUDENAARDEN, A. & VAN STEENSEL, B. 2015. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell,* 163**,** 134-47.

KOZLOWSKI, J. M., HART, I. R., FIDLER, I. J. & HANNA, N. 1984. A human melanoma line heterogeneous with respect to metastatic capacity in athymic nude mice. *J Natl Cancer Inst,* 72**,** 913-7.

KRAUSE, M., YANG, F. W., TE LINDERT, M., ISERMANN, P., SCHEPENS, J., MAAS, R. J. A., VENKATARAMAN, C., LAMMERDING, J., MADZVAMUSE, A., HENDRIKS, W., TE RIET, J. & WOLF, K. 2019. Cell migration through three-dimensional confining pores: speed accelerations by deformation and recoil of the nucleus. *Philos Trans R Soc Lond B Biol Sci,* 374**,** 20180225.

LAMBERT, A. W., PATTABIRAMAN, D. R. & WEINBERG, R. A. 2017. Emerging Biological Principles of Metastasis. *Cell,* 168**,** 670-691.

LAMMERDING, J., FONG, L. G., JI, J. Y., REUE, K., STEWART, C. L., YOUNG, S. G. & LEE, R. T. 2006. Lamins A and C but not lamin B1 regulate nuclear mechanics. *J Biol Chem,* 281**,** 25768-80.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-9.

LE, H. Q., GHATAK, S., YEUNG, C. Y., TELLKAMP, F., GUNSCHMANN, C., DIETERICH, C., YEROSLAVIZ, A., HABERMANN, B., POMBO, A., NIESSEN, C. M. & WICKSTROM, S. A. 2016. Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nat Cell Biol,* 18**,** 864-75.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G., DURBIN, R. & GENOME PROJECT DATA PROCESSING, S.

2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-9.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

LYONS, S. M., ALIZADEH, E., MANNHEIMER, J., SCHUAMBERG, K., CASTLE, J., SCHRODER, B., TURK, P., THAMM, D. & PRASAD, A. 2016. Changes in cell shape are correlated with metastatic potential in murine and human osteosarcomas. *Biol Open,* 5**,** 289-99.

MACHESKY, L. M. 2008. Lamellipodia and filopodia in metastasis and invasion. *FEBS Lett,* 582**,** 2102-11.

MAIZELS, Y., ELBAZ, A., HERNANDEZ-VICENS, R., SANDRUSY, O., ROSENBERG, A. & GERLITZ, G. 2017. Increased chromatin plasticity supports enhanced metastatic potential of mouse melanoma cells. *Exp Cell Res,* 357**,** 282-290.

MIRNY, L. A., IMAKAEV, M. & ABDENNUR, N. 2019. Two major mechanisms of chromosome organization. *Curr Opin Cell Biol,* 58**,** 142-152.

NAVA, M. M., MIROSHNIKOVA, Y. A., BIGGS, L. C., WHITEFIELD, D. B., METGE, F., BOUCAS, J., VIHINEN, H., JOKITALO, E., LI, X., GARCIA ARCOS, J. M., HOFFMANN, B., MERKEL, R., NIESSEN, C. M., DAHL, K. N. & WICKSTROM, S. A. 2020. Heterochromatin-Driven Nuclear Softening Protects the Genome against Mechanical Stress-Induced Damage. *Cell,* 181**,** 800-817 e22.

NUEBLER, J., FUDENBERG, G., IMAKAEV, M., ABDENNUR, N. & MIRNY, L. A. 2018. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A,* 115**,** E6697-E6706.

OPPENHEIMER, S. B. 2006. Cellular basis of cancer metastasis: A review of fundamentals and new advances. *Acta Histochem,* 108**,** 327-34.

PANAGIOTAKOPOULOU, M., BERGERT, M., TAUBENBERGER, A., GUCK, J., POULIKAKOS, D. & FERRARI, A. 2016. A Nanoprinted Model of Interstitial Cancer Migration Reveals a Link between Cell Deformability and Proliferation. *ACS Nano,* 10**,** 6437-48.

POMBO, A. & DILLON, N. 2015. Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol,* 16**,** 245-57.

QUINLAN, A. R. & HALL, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26**,** 841-2.

RAAB, M., GENTILI, M., DE BELLY, H., THIAM, H. R., VARGAS, P., JIMENEZ, A. J., LAUTENSCHLAEGER, F., VOITURIEZ, R., LENNON-DUMENIL, A. M., MANEL, N. & PIEL, M. 2016. ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. *Science,* 352**,** 359-62.

RAMASWAMY, S., ROSS, K. N., LANDER, E. S. & GOLUB, T. R. 2003. A molecular signature of metastasis in primary solid tumors. *Nat Genet,* 33**,** 49-54.

ROWAT, A. C., JAALOUK, D. E., ZWERGER, M., UNG, W. L., EYDELNANT, I. A., OLINS, D. E., OLINS, A. L., HERRMANN, H., WEITZ, D. A. & LAMMERDING, J.

2013. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. *J Biol Chem,* 288**,** 8610-8.

ROWLEY, M. J. & CORCES, V. G. 2018. Organizational principles of 3D genome architecture. *Nat Rev Genet,* 19**,** 789-800.

ROWLEY, M. J., NICHOLS, M. H., LYU, X., ANDO-KURI, M., RIVERA, I. S. M., HERMETZ, K., WANG, P., RUAN, Y. & CORCES, V. G. 2017. Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol Cell,* 67**,** 837-852 e7.

SCHWARZER, W., ABDENNUR, N., GOLOBORODKO, A., PEKOWSKA, A., FUDENBERG, G., LOE-MIE, Y., FONSECA, N. A., HUBER, W., HAERING, C. H., MIRNY, L. & SPITZ, F. 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature,* 551**,** 51-56.

SEGAL, T., SALMON-DIVON, M. & GERLITZ, G. 2018. The Heterochromatin Landscape in Migrating Cells and the Importance of H3K27me3 for Associated Transcriptome Alterations. *Cells,* 7.

SERVANT, N., VAROQUAUX, N., HEARD, E., BARILLOT, E. & VERT, J. P. 2018. Effective normalization for copy number variation in Hi-C data. *BMC Bioinformatics,* 19**,** 313.

SHIVASHANKAR, G. V. 2019. Mechanical regulation of genome architecture and cell-fate decisions. *Curr Opin Cell Biol,* 56**,** 115-121.

STEPHENS, A. D., BANIGAN, E. J., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2017. Chromatin and lamin A determine two different mechanical response regimes of the cell nucleus. *Mol Biol Cell,* 28**,** 1984-1996.

STEPHENS, A. D., LIU, P. Z., BANIGAN, E. J., ALMASSALHA, L. M., BACKMAN, V., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2018. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol Biol Cell,* 29**,** 220-233.

SWIFT, J., IVANOVSKA, I. L., BUXBOIM, A., HARADA, T., DINGAL, P. C., PINTER, J., PAJEROWSKI, J. D., SPINLER, K. R., SHIN, J. W., TEWARI, M., REHFELDT, F., SPEICHER, D. W. & DISCHER, D. E. 2013. Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. *Science,* 341**,** 1240104.

TABERLAY, P. C., ACHINGER-KAWECKA, J., LUN, A. T., BUSKE, F. A., SABIR, K., GOULD, C. M., ZOTENKO, E., BERT, S. A., GILES, K. A., BAUER, D. C., SMYTH, G. K., STIRZAKER, C., O'DONOGHUE, S. I. & CLARK, S. J. 2016. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res,* 26**,** 719-31.

TAJIK, A., ZHANG, Y., WEI, F., SUN, J., JIA, Q., ZHOU, W., SINGH, R., KHANNA, N., BELMONT, A. S. & WANG, N. 2016. Transcription upregulation via force-induced direct stretching of chromatin. *Nat Mater,* 15**,** 1287-1296.

TORVALDSON, E., KOCHIN, V. & ERIKSSON, J. E. 2015. Phosphorylation of lamins determine their structural properties and signaling functions. *Nucleus,* 6**,** 166-71.

VAN STEENSEL, B. & BELMONT, A. S. 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell,* 169**,** 780-791.

WOLF, K., TE LINDERT, M., KRAUSE, M., ALEXANDER, S., TE RIET, J., WILLIS, A. L., HOFFMAN, R. M., FIGDOR, C. G., WEISS, S. J. & FRIEDL, P. 2013. Physical limits of cell migration: control by ECM space and nuclear deformation and tuning by proteolysis and traction force. *J Cell Biol,* 201**,** 1069-84.

YANKASKAS, C. L., THOMPSON, K. N., PAUL, C. D., VITOLO, M. I., MISTRIOTIS, P., MAHENDRA, A., BAJPAI, V. K., SHEA, D. J., MANTO, K. M., CHAI, A. C., VARADARAJAN, N., KONTROGIANNI-KONSTANTOPOULOS, A., MARTIN, S. S. & KONSTANTOPOULOS, K. 2019. A microfluidic assay for the quantification of the metastatic propensity of breast cancer specimens. *Nat Biomed Eng,* 3**,** 452-465.

ZHANG, D., HUANG, P., SHARMA, M., KELLER, C. A., GIARDINE, B., ZHANG, H., GILGENAST, T. G., PHILLIPS-CREMINS, J. E., HARDISON, R. C. & BLOBEL, G. A. 2020. Alteration of genome folding via contact domain boundary insertion. *Nat Genet,* 52**,** 1076-1087.

ZHANG, Y., LIU, T., MEYER, C. A., EECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C., MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol,* 9**,** R137.

ZHANG, Y., MCCORD, R. P., HO, Y. J., LAJOIE, B. R., HILDEBRAND, D. G., SIMON, A. C., BECKER, M. S., ALT, F. W. & DEKKER, J. 2012. Spatial organization of the

mouse genome and its role in recurrent chromosomal translocations. *Cell,* 148**,** 908-21.

ZHOU, Y., GERRARD, D. L., WANG, J., LI, T., YANG, Y., FRITZ, A. J., RAJENDRAN, M., FU, X., SCHIFF, R., LIN, S., FRIETZE, S. & JIN, V. X. 2019. Temporal dynamic reorganization of 3D chromatin architecture in hormone-induced breast cancer and endocrine resistance. *Nat Commun,* 10**,** 1522.

CHAPTER III: INVESTIGATING THE ROLE OF INTRINSIC HETEROGENEITY IN DRIVING MELANOMA CELL MIGRATION

The work on this chapter is presented as a follow-up study to Chapter 2. This data is preliminary and not yet submitted for publication.

My contributions include: (1) developing the project, (2) performing experiments with the help of Christopher Playter, (3) performing sequencing experiments and (4) data analysis.

Abstract

During metastasis, cancer cells have to invade through small constrictions of the extracellular matrix (ECM) or blood vessel lining. While the cell membrane is highly elastic and deformable, the nucleus poses a barrier to this constricted migration due to its stiffness influenced by nuclear lamina and chromatin. The chromatin is non-randomly organized in the nucleus to maintain proper gene regulation, DNA replication and repair. This organization of chromatin is altered when neutrophils undergo constricted migration. More recently, it has been reported that specific 3D structural changes have been associated with an invasive phenotype in melanoma cells. However, it is not yet understood whether specific structural features of the genome that promote constricted migration pre-exist in the initial cancer cell population. Here, we identify ITGB4 as a marker for highly migratory melanoma cells. By using Hi-C to analyze the 3D structure of these cells, we identify similarities in compartment identities between ITGB4(+) cells and previously published Bottom-5 cells that have undergone sequential rounds of constricted migration. However, differences within the compartment signatures between Bottom-5 and ITGB4(+) cells point to additional changes that may happen due to physical constriction of the nucleus.

Introduction

Cells in the body experience different types of forces such as compresive and tensile forces when cancer cells have to undergo constricted migration during metastasis (Magid and Law, 1985, Lange and Fabry, 2013). The ability of cells to respond to such forces influences their subsequent fate and is often modulated and influenced by the nucleus mechanics (Barnes et al., 2017, Bissell et al., 1982, Guilluy et al., 2014). The nucleus is the stiffest organelle and its translocation through these constricted spaces is often thought to be the rate limiting step of migration (Davidson et al., 2014). The stiffness of the nucleus is modulated by two major components: the nuclear lamina and chromatin (Harada et al., 2014, Davidson et al., 2014, Stephens et al., 2017). Interrogation of these components such as by modulating the levels of Lamin A/C or changing the state of chromatin influences the stiffness of the nucleus (Lammerding et al., 2006, Davidson et al., 2014, Rowat et al., 2013, Stephens et al., 2017, Stephens et al., 2018).

It has been previously shown that during migration through constricting environments, the nuclear envelope can rupture allowing translocation of cytoplasmic proteins into the nucleus leading to DNA damage (Denais et al., 2016, Irianto et al., 2017a, Raab et al., 2016). Additionally, multiple rounds of constricted migration has been shown to induce chromosomal instability and influence copy number variation (Irianto et al., 2017a). However, how these dramatic events influence the organization of the DNA inside the nucleus has been poorly understood. Chromosome conformation capture techniques, such as Hi-C, have revealed a non-random organization of the chromatin in the nucleus

(Dekker et al., 2002, Bickmore and van Steensel, 2013, Rowley and Corces, 2018). Chromosomes have been shown to maintain their territorries and are spatially segregated into heterochromatin and euchromatin. This segregation is important for cell identity maintenance but also regulation of gene expression. Genes have been shown to be regulated within regulatory neighborhoods called Topologically Associating Domains (TADs) through loop extrusion that promotes interactions between enhancers and promoters or other regulatory factors. Therefore, understanding how this complex organization influences or is influenced by cells undergoing large nuclear deformations during constricted migration will provide further information on how gene expression and how the physical properties of chromosomes might contribute to a metastatic phenotype of cancer cells.

Previous work in our lab has concluded that melanoma cancer cells that migrate through sequential rounds of squeezing exhibit changes in cell and nuclear morphology (Golloshi et al., 2019). These changes are also correlated with changes in the 3D structre of the genome. Melanoma cells that have undergone ten rounds of constricted migration through 5μm transwells (Bottom-5) exhibit changes in compartment identity that are not evident in cells that do not undergo constricted migration or cells that undergo migration without constrictions (Bottom-12). Genes found within these regions that switch compartments represent pathways that have been related to constricted migration and cancer metastasis. Moreover, we observe large genomic rearrangements and the appearance of a new translocation between chr5 and chr13. However, it is not evident

whether the structural and genomic rearrangements we observe are selected from a heterogeneous initial Control population or are caused by constricted migration. Agent based modeling (ABM) of all these melanoma sybtypes we have generated suggested that both factors (heterogeneity and changes induced by constricted migration) could reproduce our data.

This melanoma cell line (A375) has been previously shown to be highly heterogeneous in relation to metastatic behavior and it is therefore important to decipher the structural properties that pre-exist in the initial population that might favor constricted migration.

Here, we identify β4-integrin (ITGB4) as a marker for cells that can undergo constricted migration and provide initial analysis of 3D structural profiles associated with ITGB4 expression and its association with Bottom-5 cells.

Results

ITGβ4 expression is correlated with ability to undergo constricted migration

To investigate whether we could identify a marker that would separate highly migratory from non-migratory A375 cells, we turned to RNA-Seq expression data. We observed an overexpression of ITGB4 in cells that have undergone constricted migration for 10 and 20 rounds as compared to Control cells (Bottom-5 – log2FC ~ 4 and Bottom20-5 – log2FC ~ 2.8) (Figure 24A). Cells that underwent migration without constriction (Bottom-12), exhibited a lower expression of ITGB4 (Figure 24A) indicating that the expression of this

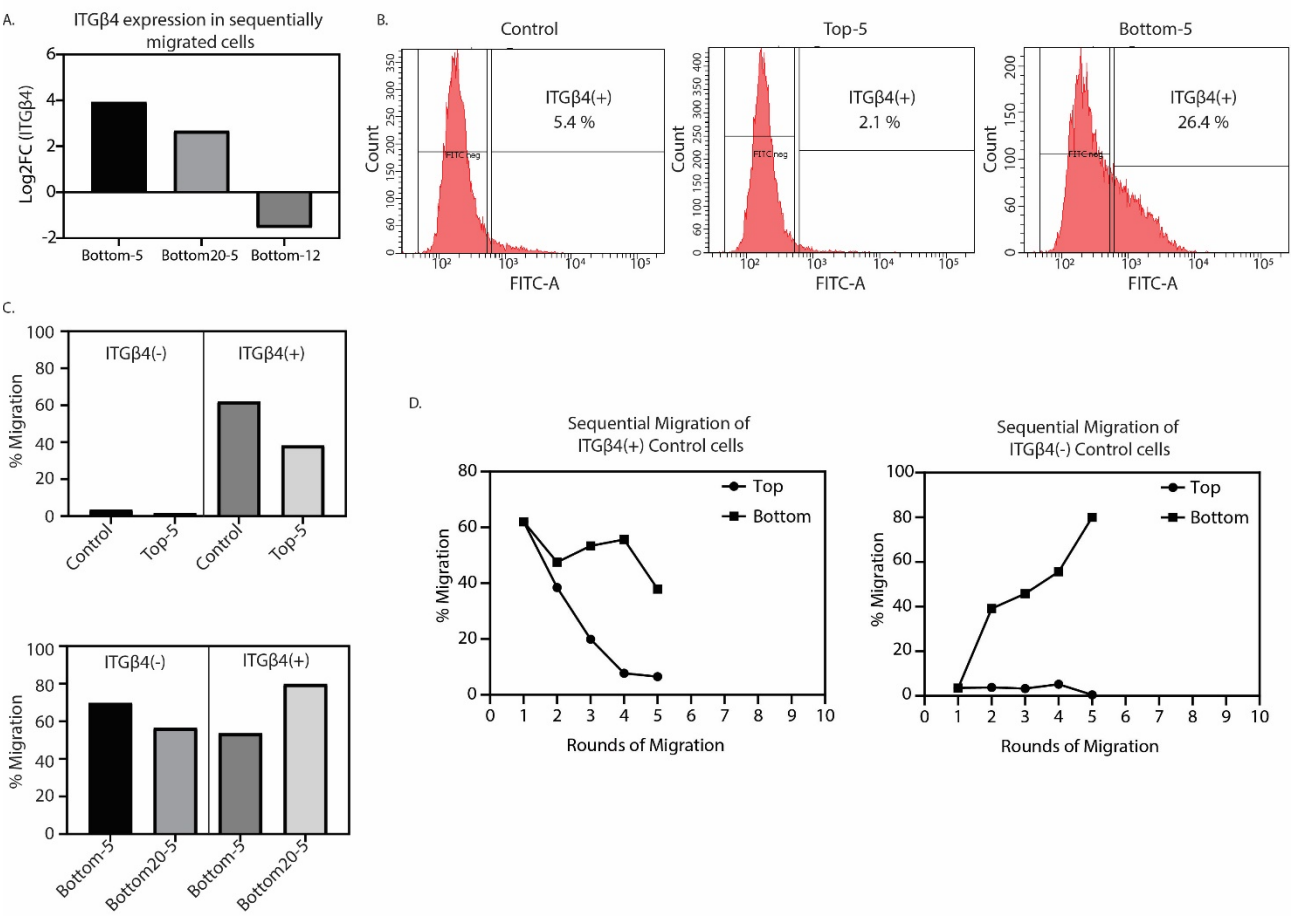


Figure 24. Expression of ITGB4 influences constricted migration but does not completely explain all phenotypes observed.

(A) Log2FC expression of ITGB4 in Bottom-5, Bottom20-5 and Bottom-12 compared to Control cells. (B) FACS sorting analysis based on ITGB4 staining on Control, Top-5 and Bottom-5 cells. (C) Transwell migration through 5μm pores in Control, Top-5, Bottom-5 and Bottom20-5 cells with respect to their ITGB4 expression. (D) Five rounds of sequential transwell migration through 5μm pores of Control ITGB4(+) (left panel) and ITGB4(-) (right panel).

gene could be specific to constricted migration. It has been previously shown that ITGB4 can promote invasion of cancer cells (Li et al., 2017). We then wondered whether expression of this gene is associated with ability of melanoma cells to undergo constricted migration. FACS sorting analysis revealed a small % of cells with higher ITGB4 protein expression among the control population (~ 5.4% ITGB4(+)) and Top-5 (~ 2.1% ITGB4(+)) samples (Figure 24B). In agreement with RNA-Seq data, Bottom-5 cells exhibited a larger ITGB4(+) population (~ 26.4%) (Figure 24B). We then sorted all A375 subpopulation of cells based on their ITGB4 expression status and investigated their migration efficiency through transwells with 5μm pores (Figure 24C). Interestingly, Control and Top-5 cells negative for ITGB4 displayed very low migration efficiency compared to Control and Top-5 cells that stained positive for ITGB4 (Figure 24C, top panel). Cells that had undergone 5μm constricted migration for 10 and 20 rounds (Bottom-5 and Bottom20-5, respectively) displayed a higher migration efficiency through the 5μm pores, independent of ITGB4 expression (Figure 24C, bottom panel).

We then focused on the Control cells to investigate whether inherent differences exist among ITGB4 (+) and ITGB4(-) cells in their behavior when challenged to undergo the sequential transwell migration. As expected, ITGB4(+) Control cells displayed a high migration frequency in the first round (Figure 24D, left panel). These cells displayed variable migration efficiencies after each consecutive round but overall maintained their high migration efficiency (Bottom). Even in this ITGB4(+) Control population, we were able to filter out some cells that were not able to migrate through the constrictions and

whose migration efficiency did not increase with each consecutive round (Top). Furthermore, ITGB4(-) Control cells started out with a low migration efficiency as previously observed (Figure 24D, right panel). To our surprise, the ITGB4(-) cells that made it through the first round of constricted migration increased their migration efficiency after each consecutive round as we had previously observed with the control population (Golloshi et al., 2019).

ITGB4 expression does not fully explain compartment identity switches observed upon sequential migration.

We have previously demonstrated that cells that undergo constricted migration exhibit differences in their 3D structure of the genome (Golloshi et al., 2019). However, we have not been able to identify whether the differences we observed are inherently present in the original A375 population or if they are induced by contricted migration. To investigate the 3D genome structural organization of Control ITGB4(+) and ITGB4(-) cells, we performed Hi-C and compared the results with A375 cells that have undergone constricted migration (Bottom-5 cells) and A375 Control cells.

From our previous study, the major structural rearrangements we observed in cells that had undergone 5μm transwell migration for 10 rounds (Bottom-5) were at compartment level. Therefore, we binned the Control ITGB4(+) Hi-C contact frequencies using 250Kb bin sizes to investigate their compartment status in relation to Bottom-5 cells and the original A375 Control population (Figure 25). We aimed to identify whether the

compartment identity switches we had previously observed in Bottom-5 cells were present in the ITGB4(+) subpopulation of cells.

We therefore performed Principal Component Analysis (PCA) to determine compartment identity using 250Kb binned matrices (Principal Component 1 (PC1) is highly correlated with compartment identity). When assessing interaction contact maps, we were able to observe structural differences specific to Bottom-5 cells independent of ITGB4 expression (Figure 25A). Part of the highlighted region in chr1 belongs to a B compartment in Control and ITGB4(+) cells but switches into an A compartment in Bottom-5 cells. To investigate differences in compartment identity across the whole genome, we plotted correlation plots of PC1 between ITGB4(+) cells and Control, Bottom-5 and Bottom-12. Interestingly, ITGB4(+) cells exhibited a higher difference in compartment identity when compared to Control rather than Bottom-5 and Bottom-12 (Figure 25B). Numerous compartment identitiy switches observed in Bottom-5 vs Control are also observed in ITGB4 (+) population, suggesting that some selection is taking place during constricted migration. In fact, when taking a stringent cutoff to consider only the strongest changes in compartment identity, we find that 120 bins from the whole genome switch their compartment identity from active (A) to inactive (B) when ITGB4(+) compartment identity was compared with Control cells. Additionally, 204 genomic bins were observed to change their compartment identity from inactive (B) to active (A). Interestingly, 91 genomic bins were observed to switch their compartment identity from A to B when when

comparing ITGB4(+) and Bottom-5. Moreover, 87 genomic bins were observed to switch their compartment identity from B to A.

To identify what compartment switches were specific to constricted migration and ITGB4(+) population of cells, we compared compartment identity in ITGB4(+) and Bottom-5 in relation to the Control population. When compared to Control PC1 values, ITGB4(+) and Bottom-5 displayed 79 shared bins that changed compartment identity from A to B. Additionally, 138 bins were observed to change their compartment identity from B to A in ITGB4(+) and Bottom-5 cells when compared to Control. Genes observed within these regions include CSMD3, PCDH15, HYDIN, ANK3, CNTNAP2, DDX60, GPC5, CDH6, AXL, MSR1, NETO1 and PTPRO, which have been reported to be asociated with cancer progression (Tan et al., 2015, Fu et al., 2016, Gugnoni et al., 2017, Hsing et al., 2007, Xu et al., 2020, Motiwala et al., 2004).

Given that the association of the chromatin with the nuclear lamina is one factor that drives compartmentalization, we therefore decided to investigate how the lamina associated domains (LADs) were represented in regions that switched compartments. We used a previously established LAD atlas across a cohort of nine different cell types that classifies genomic regions as: cLADs (constitutive LADs; regions associated with the nuclear lamina in all cell types), ciLADs (constitutive inter LADs; regions located in the interior of the nucleus across all cell lines), fLADs and fiLADs (facultative LADs and

(A) Hi-C contact frequency matrices binned at 250Kb bin sizes zoomed to a region of chr1 in Control, ITGB4(+) and Bottom-5 cells. Highlighted region indicates compartment identity switch specific to Bottom-5. Overlaid red and blue tracks indicate PC1 values along that region. (B) Whole genome corelation plots of PC1 values (red - A compartment, blue - B compartment) comparing Control, Bottom-5 and Bottom-12 A375 subpopulations with ITGB4(+) cells. (C) Enrichment of LAD types in regions that switch compartments from B to A (top panel) and A to B (bottom panel) in ITGB4(+) and Bottom-5 cells. Compartment switches were determined by comparing each subset with A375 Control population. (D) Hierarchical clustering of all A375 subtypes by their PC1 compartment identity.

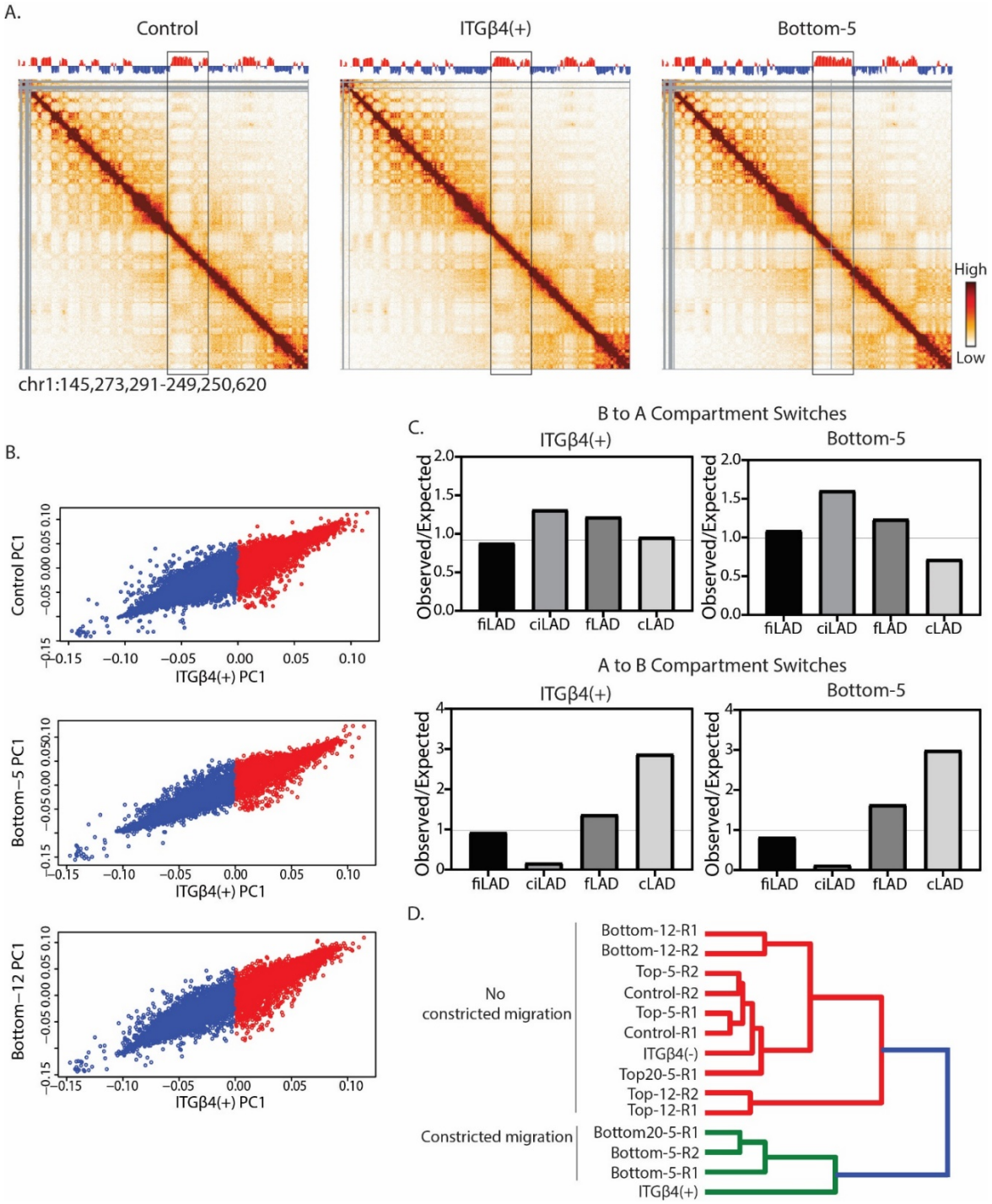


Figure 25. ITGb4 expression does not fully explain compartment identity switches observed upon sequential migration.

facultative inter LADs; regions that are sometimes associated with the nuclear lamina across the cell types) (Carolyn de Graaf, 2019, Kind et al., 2015). In both B to A and A to B compartment identity switches, we observed an overrepresentation of fiLADs which could relate with their dynamic location between the nuclear lamina and the nucleus interior (Figure 25C). Interstingly, we observed an overepresentation of cLADs in regions that switched from A to B compartment identity in both ITGB4(+) and Bottom-5 cells. However, Bottom-5 cells exhibited a higher ratio of cLADs in regions that changed which could correlate with localization of heterochromatin to the nuclear periphery in cells that have undergone constricted migration, as previously observed.

Finally, we performed hierarchical clustering of compartment identity across all A375 subtypes to identify whether expression of ITGB4 relates to the ability of cells to undergo constricted migration (Figure 25D). We observe that most cells that have not undergone constricted migration, including ITGB4(-) cells, cluster separately from sequentially migrated cells. Interestingly, ITGB4(+) cells which have also not undergone constricted migration, cluster together with sequentially migrated cellsbut further separated from them indicating that while there are similarities in compartment identities among these cells, differences due to the constricted migration may exist.

Global genomic rearrangements observed specific to ITGB4 expression

To investigate structural abberrations genome wide in ITGB4(+) cells and Bottom-5 cells, we used 2.5Mb binned Hi-C contact frequency maps that provide a whole genome view of the Hi-C contacts (Figure 26).

By visual inspection of these Hi-C heatmaps of Control, ITGB4(+) and Bottom-5 cells, the strong interaction frequency within each chromosome (diagonal pattern) is evident and in this specific cell line, translocations between different chromosomes (Figure 26A). We then compared all the contact frequencies among the A375 subpopulations by calculating log2(ITGB4(+)/Control) (Figure 26B) and log2(ITGB4(+)/Bottom-5) (Figure 26C). Interestingly, ITGB4(+) cells harbor new translocations when compared to Control (Figure 26B) and Bottom-5 (Figure 26C) cells (translocations indicated by black arrows). A new translocation appears in ITGB4(+) cells between chr4 and chr6 (t(4;6)) when compared to both Control and ITGB4(+) cells. The long arm of chr6 has been previously shown to be prone to translocations or other genetic variation in other cancer types including melanoma (Morrison et al., 2014, Guan et al., 1998). Additionally, another translocation is observed in ITGB4(+) cells between chr 6 and chr 11 (t(6;11)) which is not present in Bottom-5 cells. Such major rearrangements of chromosme 6 can modulate expression of genes that are associated with progression of cancer. Interestingly, we observed major compartment idenetiy switches in chr6 long arm (Figure 26D, top panel) and chr6 short arm (Figure 26D, bottom panel). These observations are important because this particular chromosome contains cancer related genes and further study of how these

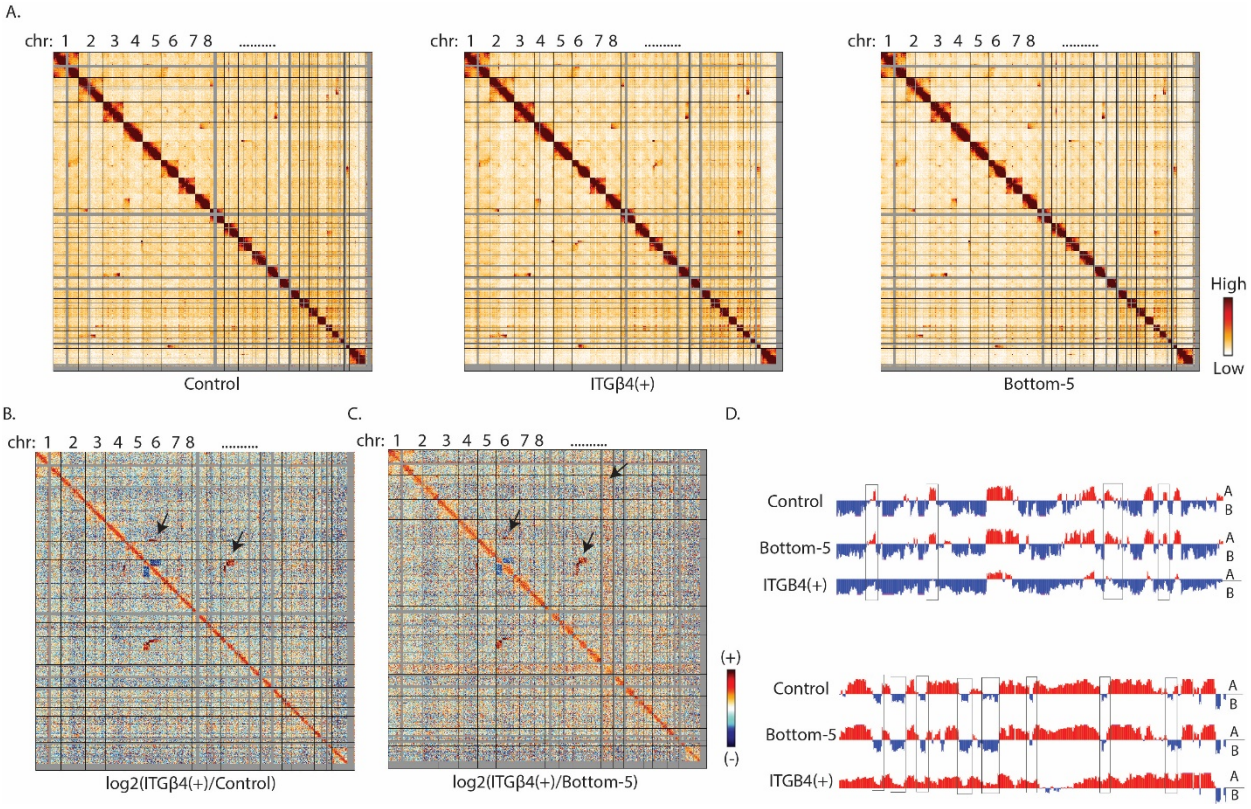


Figure 26. Global genomic rearrangements observed specific to ITGB4 expression.

(A) Whole genome Hi-C contact frequency heatmaps binned at 2.5Mb bin sizes. (B) Heatmaps represent the log2(ITGB4(+)/Control) and log2(ITGB4(+)/Bottom-5) cells (C). Black arrows indicate the new emerged translocations in ITGB4(+) cells. (D) PC1 tracks for long arm of chr6 (top graph) and short arm of chr6 (bottom graph) which undergo major structural rearrangements. Highlighted regions represent compartment identity switches in ITGB4(+) cells.

structural rearrangements modulate gene expression and influence disease progression is necessary.

Intriguingly, in our previous publication we observed a translocation between chr5 and chr13 (t(5;13)) in cells that had undergone constricted migration which is not evident in ITGB4(+) cells. This observation does not rule out the possibility of selection of cells that harbor t(5;13) but also suggests that the emergence of that translocation could be induced by structural aberrations due to constricted migration.

Discussion

Here we have provided initial preliminary data in investigating the heterogeneity within the 3D genome structure of melanoma cells. Based on our gene expression data presented in Chapter 2, we have identified β4-integrin (ITGB4) as an upregulated gene in cells that have undergone constricted migration (Bottom-5 and Bottom20-5) as compared to Control and cells that have undergone migration without constrictions (Bottom-12) (Figure 24A). FACS analysis of Control, Top-5 and Bottom-5 cells revealed some inherent heterogeneity in ITGB4 expression (Figure 24B). Consistent with the idea that there is a small subpopulation of the Control cells that already is primed for constricted migration, we find that ITGB4 positive cells in the Control and Top samples migrate much more through constrictions than ITGB4 negative cells (Figure 24C). However, cells that have passed through 5μm constrictions for 10 (Bottom-5) and 20 rounds (Bottom20-5) exhibit a similar migration efficiency independent of ITGB4

expression (Figure 24C, bottom graph). This suggests that while initial selection from the heterogeneous control population can separate cells already good at migrating, further changes likely happen to those cells that allow them to maintain constricted migratory ability whether or not they retain expression of those initial markers.

Furthermore, the Control ITGB4(-) cells still increase their migration efficiency after subsequent rounds of constricted migration. This suggests that even when we remove important initial heterogeneity, changes possibly due to constriction can still take place.

In terms of 3D genome structure, ITGB4(+) cells exhibit similar organization of the global structure with the exception of two new translocations (t(4;6;11)) emerging that is not evident in Control or Bottom-5 cells. If ITGB4 expression was essential to migration potential and only the cells that express this marker can undergo constricted migration, these translocations should have been evident in Bottom-5 cells. Since these translocations are not evident in Bottom-5 cells, this suggests that other changes could arise due to continuous clonal evolution of these cells. This idea is evident when ITGB4(+) cells are challenged to migrate through 5μm and despite their ITGB4 expression, a subset of these cells are not able to migrate through the 5μm constrictions.

ITGB4(+) cells display similarities in their compartment identity with Bottom-5 cells. However, compartment identity switches exists specific to the Bottom-5 cells that suggest they could have arisen from squeezing of the nucleus. Hierarchical clustering of whole

genome compartment identities cluster cells that have undergone constricted migration separate from cells that have not undergone migration through constrictions. Interestingly, ITGB4(+) cells cluster more closely with Bottom-5 and Bottom20-5 cells further confirming their 3D genome structure similarities.

Taken together, this initial evidence suggests that there exists an inherent heterogeneity within melanoma cells that separate highly migratory cells from non-migratory cells. However, Hi-C data also reveals that ITGB4 expression does not fully explain the variability in 3D genome structure of these migrating cells. There exist structural features that are specific to Bottom-5 cells which may indicate constriction induced changes to the 3D genome structure. Further analysis of gene expression data are necessary to correlate the structural changes we observe with transcriptional regulation of the constricted migration process. Moreover, we observe an increase of migration efficiency in ITGB4(-) cells with each subsequent round of constricted migration. This indicates that besides ITGB4 expression, other changes are taking place that are altering the phenotypic and genotypic properties of these cells to enable their migration through constrictions (Figure 24D). Detailed analysis of 3D genome structure and gene expression changes after each subsequent round of migration would be necessary to tease apart the mechanism by which large nuclear deformations due to constricted migration influence their ability to undergo subsequent steps of this constricted migration.

Materials and Methods

FACS sorting based on ITGB4 expression

To prepare A375 cells for FACS sorting, we used FITC-conjugated anti-integrin beta 4 antibody (ab22486) and followed manufacturers instruction. Briefly, cells were harvested by trypsinization and washed with ice-cold PBS containing 10% Fetal Bovine Serum (FBS). After the wash, 0.5µg/ml of conjugated ITGB4 antibody was added and cells were incubated in the dark for 30min. After incubation, cells were washed three times with 500µL of ice-cold PBS by centrifugation at 400xg for 5min. Expression of ITGB4 was then analyzed and cells with positive and negative ITGB4 expression were sorted using FACS.

Sequential Transwell Migration

For the sequential transwell migration of ITGB4(+) and ITGB4(-) A375 subpopulations, we used transwells with 5µm pore sizes as previously described (VWR-10769-236) (Golloshi et al., 2019). Briefly, the bottom of the transwells were coated with 40 µL of 10 µg/mL fibronectin for ~ 45 minutes. 24 well plates were prepared for transwell migration assay adding 500 µL of 1x DMEM (Corning) with full supplements per well. A375 cells were detached from culture dishes at 80-90% confluency and aliquoted to 100,000 cells per 100µL of 1xDMEM. Each transwell was placed into its corresponding well of the 24 well plate and 100µL of the cell suspension was added to the top of each filter. Cells were incubated at 37^{o}C, 5% CO_2 and allowed to migrate for 24 hours. After the 24-hour incubation, migration efficiency was quantified as follows. First, freely floating cells were removed from the top of the filter (unmigrated; “Top” cells) and from the well beneath the filter (migrated “Bottom” cells) and placed in two separate tubes. Then, 400 µL of trypsin

was added into the bottom chamber of the 24 well plate and 200 µL of trypsin was added into the top chamber to detach any remaining attached cells. Recovered cells after trypsinization were added to the unmigrated or migrated tubes, accordingly. Cells were spun down (1000 rpm, 5min) and counted (using trypan blue) to calculate % migration such as: #bottom/(#top+#bottom). Additionally, a small aliquot of cells was saved for immunofluorescence (IF). The rest of the cells were seeded into wells of a 24-well plate to expand. When cells reached 80-90% confluency, another transwell migration was performed (R2). A375 Top and Bottom cells were detached and counted. Two transwell filters were prepared as previously described (one for Top and one for Bottom). Then 100,000 cells suspended in 100 µL of 1xDMEM without supplements were seeded into the transwell filters. After 24-hour incubation at 37^{o}C and 5% CO_2, cells were trypsinized as previously described to quantify migration efficiency. Only the cells that were always on top (Top of Top) and the cells that were always on the bottom (Bottom of Bottom) were saved and grown for 10 sequential rounds of migration.

Hi-C

Library Preparation

Hi-C libraries were prepared for each cell sub-population (A375-Control, Top-5, Bottom-5, Top20-5, Bottom20-5, Top-12 and Bottom-12) as previously described (Golloshi et al., 2018) including two biological duplicates for each condition. Briefly, 10 million cells were grown in T-75 cell culture flasks to 80-90% confluency and crosslinked with 1% formaldehyde for 10 min. Crosslinked cells were then suspended in lysis buffer to permeabilize the cell membrane and were dounce homogenized. Chromatin was then

digested in-nucleus overnight using DpnII restriction enzyme. Digested ends were filled in with biotin-dATP, and the blunt ends of interacting fragments were ligated together. DNA was then purified by phenol-chloroform extraction. For library preparation, the NEBNext Ultra II DNA Library prep kit (NEB) was used for libraries with size ranges from 200-400bp. End Prep, Adaptor Ligation, and PCR amplification reactions were carried out on bead bound DNA libraries.

Sequencing was performed at the Oklahoma Medical Research Foundation Clinical Genomics facility using the Illumina HiSeq 3000 platform with 75 bp paired end reads or a NovaSeq 6000 with 50 bp paired end reads. Sequencing reads were mapped to the human genome (hg19), filtered, and iteratively corrected as previously described (Imakaev et al., 2012) (https://github.com/dekkerlab/cMapping). All Hi-C contact matrices were scaled to a sum of 1 million interactions to allow comparisons between conditions in downstream analysis.

Compartment Analysis

Compartment analysis was performed by running principal component analysis using matrix2compartment.pl script in the cworld-dekker pipeline available on GitHub (https://github.com/dekkerlab/cworld-dekker). The PC1 value was then used to determine compartment identity for 100 and 250Kb binned matrices. We considered PC1 values greater than 0.01 to be A compartment and less than -0.01 to be B compartment for this analysis and determination of compartment switches.

References

BARNES, J. M., PRZYBYLA, L. & WEAVER, V. M. 2017. Tissue mechanics regulate brain development, homeostasis and disease. J Cell Sci, 130, 71-82.

BICKMORE, W. A. & VAN STEENSEL, B. 2013. Genome architecture: domain organization of interphase chromosomes. Cell, 152, 1270-84.

BISSELL, M. J., HALL, H. G. & PARRY, G. 1982. How does the extracellular matrix direct gene expression? J Theor Biol, 99, 31-68.

CAROLYN DE GRAAF, J. K., LUDO PAGIE, DAAN PERIC HUPKES, BAS VAN STEENSEL. 2019. LAD Atlas [Online]. Available: https://osf.io/dk8pm/ [Accessed].

DAVIDSON, P. M., DENAIS, C., BAKSHI, M. C. & LAMMERDING, J. 2014. Nuclear deformability constitutes a rate-limiting step during cell migration in 3-D environments. Cell Mol Bioeng, 7, 293-306.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. Science, 295, 1306-11.

DENAIS, C. M., GILBERT, R. M., ISERMANN, P., MCGREGOR, A. L., TE LINDERT, M., WEIGELIN, B., DAVIDSON, P. M., FRIEDL, P., WOLF, K. & LAMMERDING, J. 2016. Nuclear envelope rupture and repair during cancer cell migration. Science, 352, 353-8.

FU, T. Y., WU, C. N., SIE, H. C., CHENG, J. T., LIN, Y. S., LIOU, H. H., TSENG, Y. K., SHU, C. W., TSAI, K. W., YEN, L. M., TSENG, H. W., TSENG, C. J., GER, L. P. &

LIU, P. F. 2016. Subsite-specific association of DEAD box RNA helicase DDX60 with the development and prognosis of oral squamous cell carcinoma. Oncotarget, 7, 85097-85108.

GOLLOSHI, R., MARTIN, R. S., DAS, P., RAINES, T. I., THURSTON, D., FREEMAN, T. & MCCORD, R. P. 2019. Constricted migration contributes to persistent 3D genome structure changes associated with an invasive phenotype in melanoma cells. bioRxiv, 856583.

GOLLOSHI, R., SANDERS, J. T. & MCCORD, R. P. 2018. Iteratively improving Hi-C experiments one step at a time. Methods, 142, 47-58.

GUAN, X. Y., ZHANG, H., YANG, J. M., WANG, J., TAETLE, R., MELTZER, P. S. & TRENT, J. M. 1998. Detection of chromosome 6 abnormalities in melanoma cell lines by chromosome arm painting probes. Cancer Genet Cytogenet, 107, 89-92.

GUGNONI, M., SANCISI, V., GANDOLFI, G., MANZOTTI, G., RAGAZZI, M., GIORDANO, D., TAMAGNINI, I., TIGANO, M., FRASOLDATI, A., PIANA, S. & CIARROCCHI, A. 2017. Cadherin-6 promotes EMT and cancer metastasis by restraining autophagy. Oncogene, 36, 667-677.

GUILLUY, C., OSBORNE, L. D., VAN LANDEGHEM, L., SHAREK, L., SUPERFINE, R., GARCIA-MATA, R. & BURRIDGE, K. 2014. Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. Nat Cell Biol, 16, 376-81.

HARADA, T., SWIFT, J., IRIANTO, J., SHIN, J. W., SPINLER, K. R., ATHIRASALA, A., DIEGMILLER, R., DINGAL, P. C., IVANOVSKA, I. L. & DISCHER, D. E. 2014.

Nuclear lamin stiffness is a barrier to 3D migration, but softness can limit survival. J Cell Biol, 204, 669-82.

HSING, A. W., SAKODA, L. C., CHEN, J., CHOKKALINGAM, A. P., SESTERHENN, I., GAO, Y. T., XU, J. & ZHENG, S. L. 2007. MSR1 variants and the risks of prostate cancer and benign prostatic hyperplasia: a population-based study in China. Carcinogenesis, 28, 2530-6.

IMAKAEV, M., FUDENBERG, G., MCCORD, R. P., NAUMOVA, N., GOLOBORODKO, A., LAJOIE, B. R., DEKKER, J. & MIRNY, L. A. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods, 9, 999-1003.

IRIANTO, J., XIA, Y., PFEIFER, C. R., ATHIRASALA, A., JI, J., ALVEY, C., TEWARI, M., BENNETT, R. R., HARDING, S. M., LIU, A. J., GREENBERG, R. A. & DISCHER, D. E. 2017. DNA Damage Follows Repair Factor Depletion and Portends Genome Variation in Cancer Cells after Pore Migration. Curr Biol, 27, 210-223.

KIND, J., PAGIE, L., DE VRIES, S. S., NAHIDIAZAR, L., DEY, S. S., BIENKO, M., ZHAN, Y., LAJOIE, B., DE GRAAF, C. A., AMENDOLA, M., FUDENBERG, G., IMAKAEV, M., MIRNY, L. A., JALINK, K., DEKKER, J., VAN OUDENAARDEN, A. & VAN STEENSEL, B. 2015. Genome-wide maps of nuclear lamina interactions in single human cells. Cell, 163, 134-47.

LAMMERDING, J., FONG, L. G., JI, J. Y., REUE, K., STEWART, C. L., YOUNG, S. G. & LEE, R. T. 2006. Lamins A and C but not lamin B1 regulate nuclear mechanics. J Biol Chem, 281, 25768-80.

LANGE, J. R. & FABRY, B. 2013. Cell and tissue mechanics in cell migration. Exp Cell Res, 319, 2418-23.

LI, X. L., LIU, L., LI, D. D., HE, Y. P., GUO, L. H., SUN, L. P., LIU, L. N., XU, H. X. & ZHANG, X. P. 2017. Integrin beta4 promotes cell invasion and epithelial-mesenchymal transition through the modulation of Slug expression in hepatocellular carcinoma. Sci Rep, 7, 40464.

MAGID, A. & LAW, D. J. 1985. Myofibrils bear most of the resting tension in frog skeletal muscle. Science, 230, 1280-2.

MORRISON, C. D., LIU, P., WOLOSZYNSKA-READ, A., ZHANG, J., LUO, W., QIN, M., BSHARA, W., CONROY, J. M., SABATINI, L., VEDELL, P., XIONG, D., LIU, S., WANG, J., SHEN, H., LI, Y., OMILIAN, A. R., HILL, A., HEAD, K., GURU, K., KUNNEV, D., LEACH, R., ENG, K. H., DARLAK, C., HOEFLICH, C., VEERANKI, S., GLENN, S., YOU, M., PRUITT, S. C., JOHNSON, C. S. & TRUMP, D. L. 2014. Whole-genome sequencing identifies genomic heterogeneity at a nucleotide and chromosomal level in bladder cancer. Proc Natl Acad Sci U S A, 111, E672-81.

MOTIWALA, T., KUTAY, H., GHOSHAL, K., BAI, S., SEIMIYA, H., TSURUO, T., SUSTER, S., MORRISON, C. & JACOB, S. T. 2004. Protein tyrosine phosphatase receptor-type O (PTPRO) exhibits characteristics of a candidate tumor suppressor in human lung cancer. Proc Natl Acad Sci U S A, 101, 13844-9.

RAAB, M., GENTILI, M., DE BELLY, H., THIAM, H. R., VARGAS, P., JIMENEZ, A. J., LAUTENSCHLAEGER, F., VOITURIEZ, R., LENNON-DUMENIL, A. M., MANEL,

N. & PIEL, M. 2016. ESCRT III repairs nuclear envelope ruptures during cell migration to limit DNA damage and cell death. Science, 352, 359-62.

ROWAT, A. C., JAALOUK, D. E., ZWERGER, M., UNG, W. L., EYDELNANT, I. A., OLINS, D. E., OLINS, A. L., HERRMANN, H., WEITZ, D. A. & LAMMERDING, J. 2013. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. J Biol Chem, 288, 8610-8.

ROWLEY, M. J. & CORCES, V. G. 2018. Organizational principles of 3D genome architecture. Nat Rev Genet, 19, 789-800.

STEPHENS, A. D., BANIGAN, E. J., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2017. Chromatin and lamin A determine two different mechanical response regimes of the cell nucleus. Mol Biol Cell, 28, 1984-1996.

STEPHENS, A. D., LIU, P. Z., BANIGAN, E. J., ALMASSALHA, L. M., BACKMAN, V., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2018. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. Mol Biol Cell, 29, 220-233.

TAN, H., BAO, J. & ZHOU, X. 2015. Genome-wide mutational spectra analysis reveals significant cancer-specific heterogeneity. Sci Rep, 5, 12566.

XU, Y., WANG, W., CHEN, J., MAO, H., LIU, Y., GU, S., LIU, Q., XI, Q. & SHI, W. 2020. High neuropilin and tolloid-like 1 expression associated with metastasis and poor survival in epithelial ovarian cancer via regulation of actin cytoskeleton. J Cell Mol Med.

CHAPTER IV: DECONDENSING THE CHROMATIN AND ITS EFFECT ON 3D GENOME ORGANIZATION

The work on this chapter is under preparation for submission by Rosela Golloshi, Priyojit Das and Dr. Rachel Patton McCord.

My contributions to this work included: (1) conceived the project, (2) designed the experiments, (3) performed genomic, cell culture, and imaging experiments, (4) performed bioinformatics and computational data analysis with the help of Priyojit Das and (5) wrote the paper with input from all authors.

Abstract

With the growing understanding of epigenetics and its influence in gene expression and 3D organization of the genome, cancer therapeutics that affect the chromatin state are becoming popular. It has been previously shown that treatments with these agents influences the mechanical properties of the nucleus, but little is known about how they affect the 3D organization of the genome and influence gene expression. Here, we treat melanoma cells with Trichostatin A (TSA), a global decondensing agent. Treatment with this drug has been shown to soften the nuclei and mediate gene expression. We use Chromosome Conformation Capture with High Throughput Sequencing (Hi-C), to investigate the effect of this global decondensing agent in the 3D organization of the genome.

Introduction

Cancer has been traditionally thought to arise from genetic alterations such as amplifications, rearrangements, mutations of the DNA which lead to the aberrant

expression of tumor suppressor and oncogenes. However, it has been established that besides changes to the DNA sequence, epigenetic alterations contribute to initiation and progression of cancer (Baylin and Jones, 2011). Many cancers exhibit aberrant DNA methylation and aberrant expression of histone modifier enzymes such as demethylases (DNMT) and histone deacetylases (HDACs) (Berdasco and Esteller, 2010). In fact, HDACs have been found to be overexpressed in many cancers which could lead to inappropriate expression of oncogenes or tumor suppressors, a signature of cancer. In contrast to mutation in the DNA, these epigenetic alterations are thought to be reversible and therefore have become an attractive target for cancer therapy. Four classes of HDACs have been reported including Class I (comprised of HDAC 1, 2 ,3 and 8), Class II (comprised of HDAC 4, 5, 6,7, 9 and 10), Class III (comprised of seven mammalian sirtuins SIRT1-7) and Class IV (comprised of HDAC 11)(de Ruijter et al., 2003). Apart from regulating histone modification, HDACs also regulate the post-translational acetylation status of many non-histone proteins, including transcription factors, chaperones, and signaling molecules, resulting in changes in protein stability, protein-protein interactions, and protein-DNA interactions (Glozak et al., 2005). For example, maintaining the acetylation status of RUNX3, a tumor suppressor and transcription factor, is important for its stability and transcriptional activity. Furthermore, deacetylation of HSP90 by HDAC6 is essential for the stability and function of other proteins such as Bcr-Abl, c-Raf, and AKT(Bali et al., 2005).Various HDAC inhibitors (HDACi) have been developed and described to induce cell cycle arrest, differentiation, and apoptosis in cell lines (Mariadason et al., 2001, Litvak et al., 1998). Most of these have potent antitumor

activities in vivo and in vitro (Marks et al., 2001, Marks et al., 2000). One of the most effective and best studied HDAC inhibitor is trichostatin A (TSA) (Yoshida et al., 1990, Inoue et al., 2006).

Trichostatin A (TSA), a class I and II HDAC inhibitor, exerts anti-tumor effects in colon carcinoma, breast adenocarcinoma, esophageal squamous cells, prostate and mouse melanoma cells by differential expression of genes related to cell proliferation, cell cycle inhibition, apoptotic pathways and inhibition of cell migration (Sun et al., 2014, Watson et al., 2010, Ma et al., 2015, Gerlitz and Bustin, 2010). It has also been revealed by high resolution microscopy image analysis that treatment with TSA causes a global increase in acetylation levels and decondensation of the chromatin fiber (Toth et al., 2004, Watson et al., 2010). Additionally, nuclei of cells treated with TSA decrease their stiffness in line with the idea that the chromatin fiber is more relaxed upon decondensation (Biggs et al., 2019). It has been previously postulated that the increase in chromatin decondensation could increase the accessibility of regulatory elements to bind to specific genes and therefore alter their expression. However, it is not clear how these changes in chromatin compaction, gene expression and physical properties of the nucleus affect the higher order organization of the chromatin upon treatment with TSA.

The two-meter chromatin fiber is packaged into the nucleus in a complex 3D conformation that is essential for DNA replication, repair, and gene regulation (Marchal et al., 2019, Krijger and de Laat, 2016, Spielmann et al., 2018). Regions of the genome marked with

different histone modifications (such as methylation or acetylation) become spatially compartmentalized into heterochromatin and euchromatin through different mechanisms, such as tethering to nuclear structures like the lamina, or due to phase separation (van Steensel and Belmont, 2017, Strom et al., 2017, Larson et al., 2017). Genes within less-accessible heterochromatin tend to be inactive while those localized to more accessible euchromatin tend to be active or poised for quick transcriptional activation (Gaspar-Maia et al., 2011, van Steensel and Belmont, 2017).

Ever-improving microscopic techniques and the development of chromosome conformation capture approaches have revolutionized our understanding of 3D genome architecture in the past decade (Dekker et al., 2002, Rowley and Corces, 2018, Abbas et al., 2019). Chromosomes are folded at different length scales into loops, topologically associating domains (TADs), active and inactive (A/B) compartments, and chromosomal territories.

Here we investigate the effect of TSA treatment on 3D organization of invasive melanoma cancer cells (A375). In agreement with previous reports, we find that treatment with TSA increases global acetylation levels of A375 cells. Interestingly, this global decompaction of the DNA does not cause any apparent changes in the global 3D organization of the genome. The biggest differences are observed at local scale where loops are disrupted upon TSA treatment correlating with changes in gene expression and local changes in chromatin accessibility.

Results

A375 cells exhibit subtle changes in global 3D genome architecture upon treatment with TSA

To investigate the effect of TSA treatment on global acetylation levels of A375 cells, we treated cells with 0.5 μM TSA for 2 hours and performed immunofluorescence imaging targeting H3K9ac. Treatment with TSA increased global H3K9ac levels on A375 cells (Figure 27A). In addition, we observed an increase on H3K9ac around the edge of the nucleus, which could indicate decondensation of heterochromatin known to be mostly tethered to the nuclear periphery (Figure 27A and B). While treatment with TSA has been shown to inhibit migration of mouse melanoma cells on a 2D surface, we investigated whether this apparent disruption in acetylation levels had an effect on migration of A375 cells through 5μm constrictions (Figure 27C). We observed a decrease in migration efficiency of these cells with increasing concentrations of TSA in agreement with previous reports in other cell lines. However, it is not yet clear how this imbalance of H3K9ac levels caused by TSA affects the higher order structure of DNA.

To investigate the effect of TSA on 3D genome structure, we performed Chromosome Conformation Capture followed by high throughput sequencing (Hi-C). Strikingly, despite the global increase in acetylation levels, treatment of A375 cells with TSA does not cause major rearrangements of the whole genome map (Figure 27D). When we quantify log2 ratio of contacts observed in TSA compared to the control condition, no apparent changes

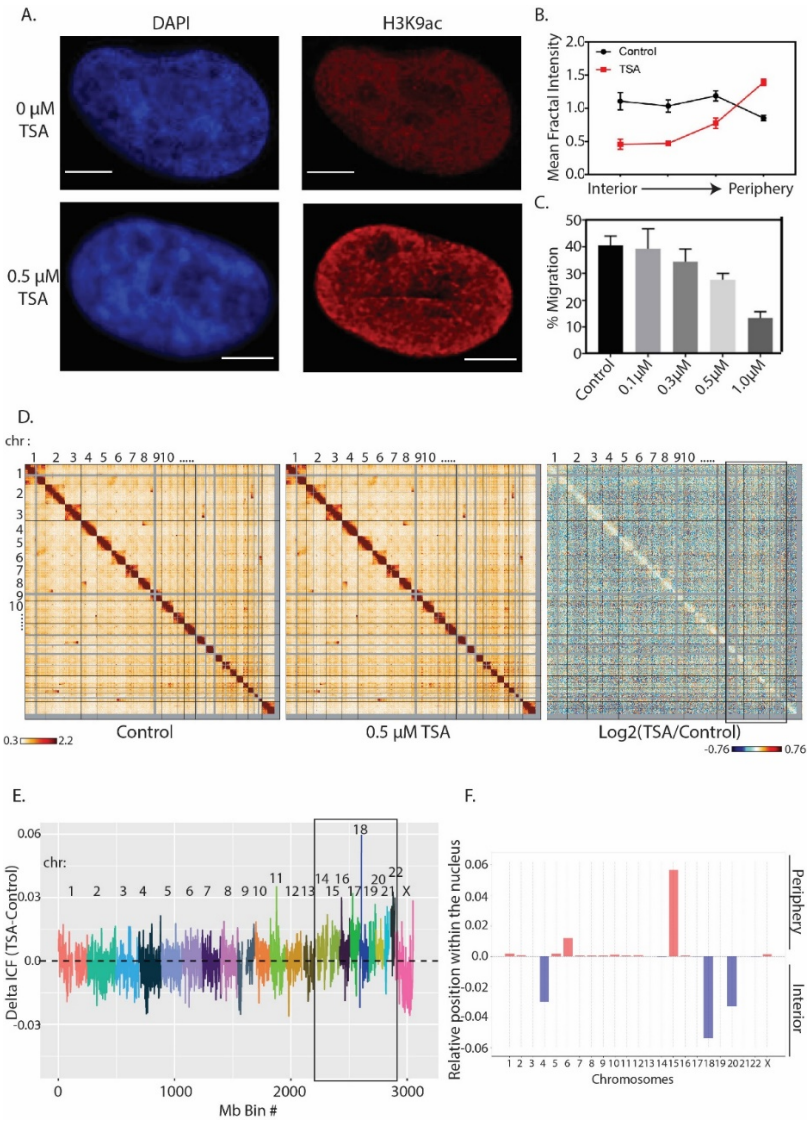


Figure 27. Increase in global acetylation causes only subtle changes in 3D interactions of the whole genome.

(A) Immunofluorescence imaging of A375 cells stained for DAPI (blue) and H3K9ac (Red) after treatment with 0.5 μM TSA for 2 hours. Scale bar indicates 5μm. (B) Radial distribution of H3K9ac signal in the nucleus. (C) Constricted migration efficiency of A375 cells treated with varying concentrations of TSA. (D) Whole genome 2.5Mb binned Hi-C contact interaction frequency maps for Control (left panel), TSA treated (middle) and log2(TSA/Control) (right panel). (E) Difference in interchromosomal fraction (delta ICF) upon treatment with TSA using 1Mb binned HiC contact matrices. (F) Model of relative radial positioning of the chromosomes upon TSA treatment using ellipsoid parameters.

are observed. However, there appears to be slight increase in contacts (increase in orange color when compared to overall blue across the heatmap) between chromosomes 15-22 (Figure 27D, left panel – highlighted region). We can use this whole chromosome data to infer changes in interaction patterns in chromosome territories. So, we calculated interchromosomal fraction (ICF) of interactions across the whole genome upon TSA treatment (Heinz et al., 2018). Interestingly, congruent with observations from log2(TSA/Control) heatmap, smaller chromosomes display an increase in ICF upon treatment with TSA (Figure 27E, highlighted region).

This could indicate that upon treatment with TSA, the smaller chromosomes could be changing their compactness and intermixing with other chromosomes. Additionally, we have used the whole genome Hi-C contact maps to model radial positioning of the chromosomes based on known radial positioning of chromosomes on an ellipsoid nucleus (Figure 27F). We observe that chromosome 6 and chr15 have shifted towards the periphery of the nucleus while chr4, 18 and 20 have shifted more toward the interior of the nucleus. Overall these observations suggest that while TSA treatment does not induce any large scale rearrangements in the contact interaction of A375 cells, changes in specific chromosomes indicate altered compactness and intermixing while changing position in the nucleus possibly due to the global decompaction of the chromatin fiber.

Subtle changes in compartment identity observed after treatment with TSA

Since the changes observed in whole genome contact maps were subtle, we next investigated higher resolution contact maps of individual chromosomes binned at 250Kb

bins (Figure 28A). We then compute principal component analysis (PCA) to determine compartment identity (A compartment – active, euchromatin and B compartment – inactive, heterochromatin) across each bin throughout the genome (Figure 1A, red - A compartment and blue – B compartment tracks) and investigate if there are any compartment identity switches upon treatment with TSA. Interestingly, the majority of the regions maintain their compartment identity upon TSA treatment and 133 bins change their compartment identity from A to B and 411 bins change their compartment identity from B to A (Figure 28B). However, most of the regions that are switching their compartment identity belong to regions classified as weak compartments (small eigen1 values in each direction) instead of strong compartment switches (Figure 28C). These regions are often correlated with regions of the chromatin that are poised or able to easily switch back and forth between different compartment identity. In fact, when we take a threshold based on compartment strength (PC1 < (-0.01) and PC1 > 0.01), only 4 bins fall under the A to B and 27 bins fall under the B to A compartment identity switch.

It has been previously shown that compartment identity of chromosomal regions is often correlated with their association with the lamina referred to as Lamina Associated Domains (LADs). Regions of the genome that associate with the nuclear lamina strongly correlate with inactive compartment while regions associated with the interior of the nucleus correlative with active compartments. Therefore, we sought to investigate the relationship between the regions that switched compartment identity and the lamin

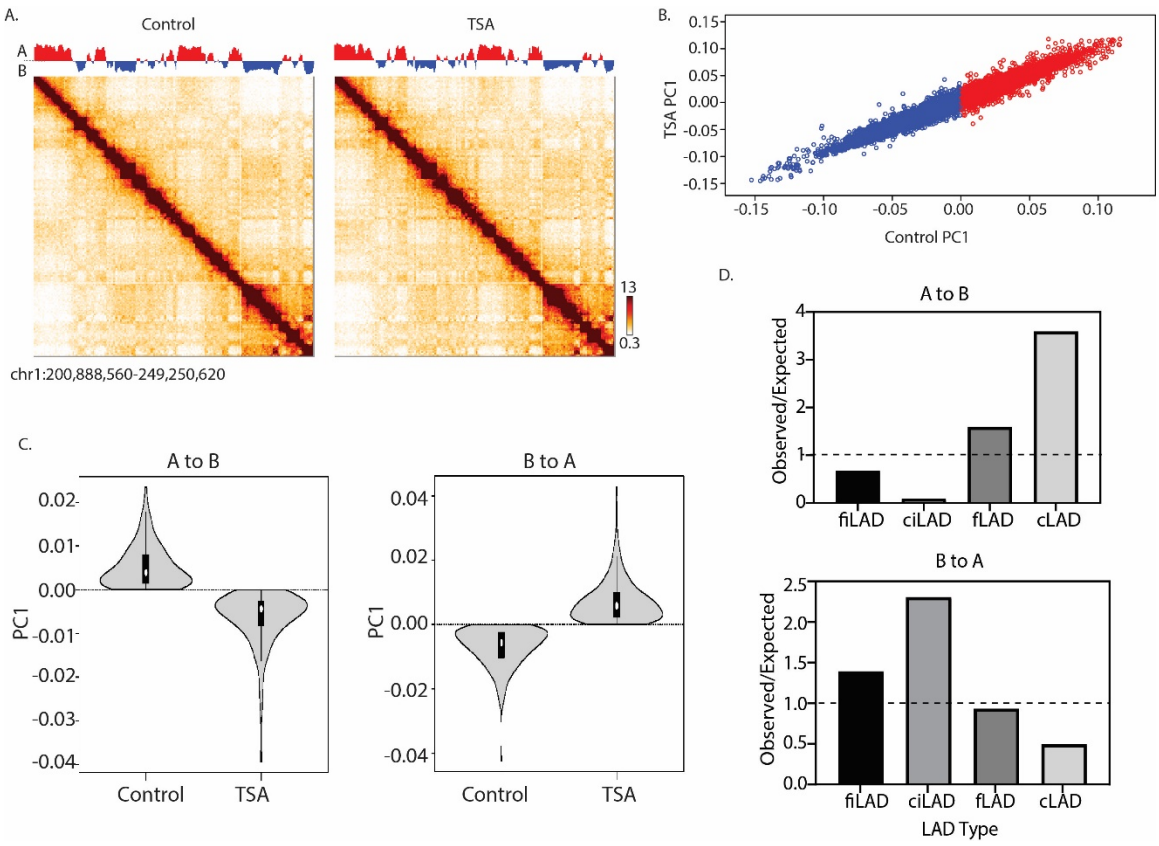


Figure 28. Subtle compartment identity switches observed after treatment with TSA.

(A) 250Kb binned Hi-C interaction heatmaps of chr1 (chr1:200,888,560-249,250,620) in Control and TSA treated cells. Red and blue bar graph over the heatmaps represents compartment identity as calculated by PCA (red – A compartment, blue – B compartment). (B) Correlation between A375 Control PC1 and TSA treated PC1. (C) Distribution of PC1 values between Control and TSA in regions that switched their compartment identity from A to B (left panel) and B to A (right panel). (D) Enrichment of different LAD types in regions that switched compartment identity from A to B (top panel) and B to A (bottom panel).

association types. We used a previously established LAD atlas across a cohort of nine different cell types that classify LADs as: cLADs (constitutive LADs; regions associated with the nuclear lamina in all cell types), ciLADs (constitutive inter LADs; regions located in the interior of the nucleus across all cell lines), fLADs and fiLADs (facultative LADs and facultative inter LADs; regions that are sometimes associated with the nuclear lamina across the cell types) (Kind et al., 2015, Carolyn de Graaf, 2019).

Interestingly, regions of the genome that switch their compartment identity from A to B are over-represented in fLAD (~ 2-fold more than expected) and cLAD (~ 4-fold more than expected) (Figure 28D, top panel). These LAD types represent regions that are often associated with the nuclear lamina and their over-representation could be due to the physical decompaction of the DNA. Moreover, regions of the genome that switched their compartment identify from B to A are overrepresented in fiLADs (~ 1.5-fold more than expected) and ciLADs (~ 2.5-fold more than expected) (Figure 28D, bottom panel). This observation suggests that TSA treatment has a higher effect on regions of the genome that are “poised” and therefore it does not cause major compartment identity shifts.

Increase in long range interactions within and between A compartments observed after TSA treatment

Since most of the bins maintained their compartment identity, we then wondered whether the interaction frequency was affected upon TSA treatment, given its decondensing properties. We used the 250Kb binned matrices and re-ordered the interactions from strongest B to strongest A compartments based on PC1 values (Figure 29A).

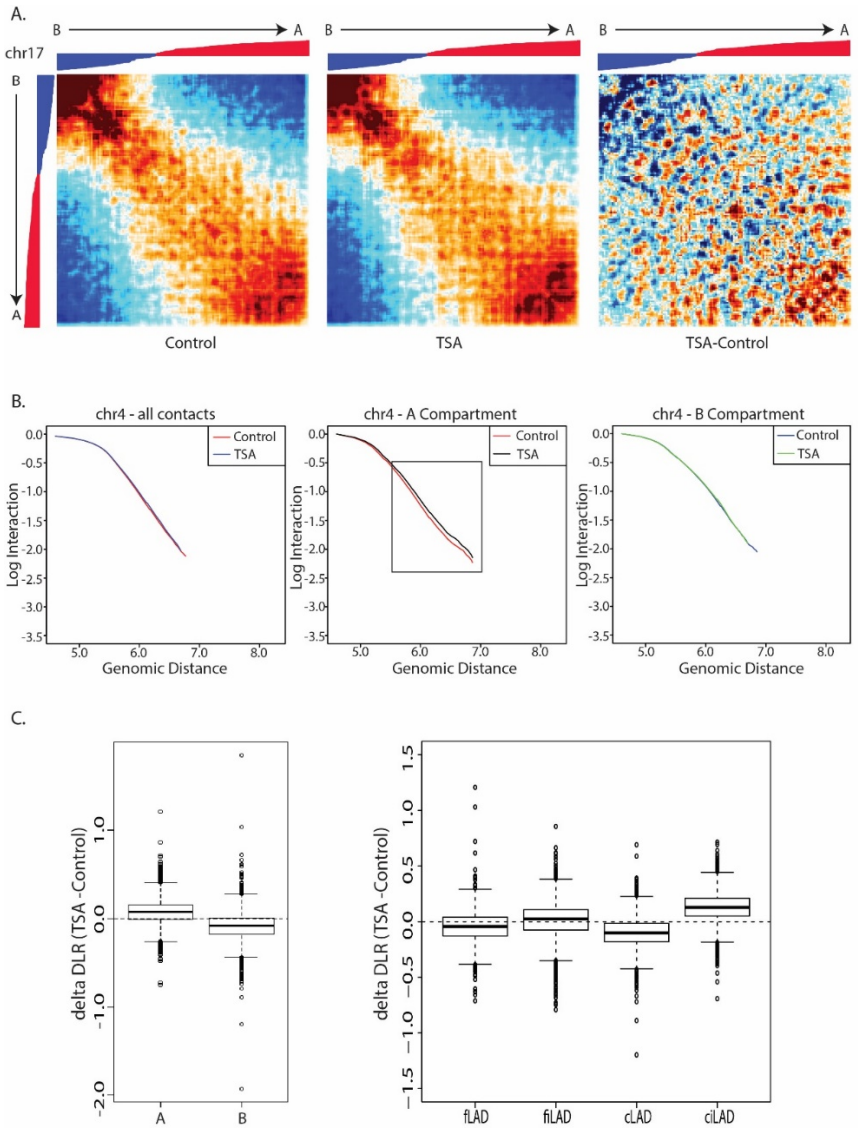


Figure 29. Interaction Hi-C contact frequency increases within A compartments and is correlated to decompaction of the chromatin fiber.

(A) Saddle plots from 250Kb binned Hi-C interaction Z-score matrix (interactions normalized by generic distance decay), reordered by PC1 values derived from A375 Control Hi-C data. Plots are representative of chr17. (B) Scaling plots of chr4 Hi-C interaction contacts using 40Kb binned matrices in A375 Control and TSA treated cells. (C) Distal-to-Local ratio of whole genome contacts within A and B compartment (left panel) and between all types of LADs. Positive deltaDLR (TSA-Control) values indicate a gain in distal interactions (> 3Mb) while negative deltaDLR values indicate a loss in distal and gain in local interactions (< 3Mb) upon treatment with TSA.

The most striking feature of these heatmaps is the distinct separation in contacts between A and B compartment and an overall increase in contact frequency within A and within B compartments, separately (Figure 29A, left and middle panel). This pattern of contact frequencies correlates with the notion that compartments of the same identity interact most frequently within themselves then between other compartment types.

Upon TSA treatment, we observe an increase in interaction frequency between strongest A compartments, correlating with the decompaction of the chromatin fiber. In contrary, a decrease in strongest B-B interactions is observed. We then investigated how interaction contact frequency was affected at higher resolution by using matrices binned at 40Kb (Figure 29B). When investigated all contacts within each chromosome, no apparent change in contact frequency with increasing distance was observed (Figure 29B, left panel). However, when the contacts were separated based on compartment identity (A and B), we observe an increase in A contact frequency at longer genomic distances upon TSA treatment (Figure 29B, middle panel). In contrary, B compartment interaction frequency with respect to genomic distance was not affected, again supporting the evidence that TSA mostly affects active or “poised” regions of the genome.

To investigate whether the differences we observe in interaction frequency between different compartments relate to the linear compaction of the DNA after TSA treatment, we measured distal-to-local ratio (DLR) of Hi-C interaction frequencies. Distal interactions have been defined as all interaction frequencies greater than 3Mb away along the

chromosome and local interactions include interactions that take place within 3Mb away (Heinz et al., 2018). This approach infers information about changes in chromosome compaction and loss of local chromatin structure (Heinz et al., 2018, Belaghzal et al., 2019). In agreement with previous results, we observe an increase in distal interactions in A compartment and a decrease in distal interactions in B compartment (Figure 29C, left panel). This could suggest that the chromatin fiber in genomic regions that fall under A compartment is further decompacted and able to make more interactions with distal sites along the same chromosome. On the other hand, B compartment is known for its signature long range interactions due to heterochromatin formation bringing together distal regions in close proximity by contacting the nuclear lamina or phase separation. Therefore, a decrease in distal interactions in B compartment could suggest slight decompaction of these regions.

We further investigated how the compactness of chromosomal regions varied in relation to their association with the nuclear lamina (Figure 29C, right panel) upon TSA treatment. The biggest change we observe in the decrease of distal interactions within cLAD regions and an increase of distal interactions within ciLAD regions. Since ciLADs involve regions that are most often located in the interior of the nucleus and mostly active, its increase in distal interactions suggest a decompaction of the chromatin fiber in these regions, as expected. On the other hands, cLADs are usually regions that contact the nuclear lamina and exhibit condensed chromatin and already distal interactions and therefore the loss of

distal interaction could indicate that there is a decompaction that is happening in cLADs upon TSA treatment.

Gene expression changes after TSA treatment reflect TSA gene signature in other cell lines

To investigate whether changes in gene expression resulted after exposure to TSA, we performed RNA-Seq analysis on A375 Control and TSA treated cells. We identified 160 genes that were overexpressed (Log2FC >1, adj p-value < 0.05) and 31 genes were downregulated (Log2FC< -1, adj p-value < 0.05) (Figure 30A). Pathway enrichment analysis revealed enrichment of genes associated with blood vessel endothelial cell proliferation and migration related to VEGF pathways which have been previously shown as targets of TSA treatment (Figure 30B, top panel). Of the downregulated genes, pathway enrichment analysis revealed an overrepresentation of pathways related to methyltransferase activity and chromatin silencing complexes which seems to be in line with the chromatin decondensing function of TSA (Figure 30B, middle panel). Additionally, this gene signature we observed has been previously reported on other cell types treated with TSA such as prostate cancer (PC3), breast cancer (MCF7) and leukemia (HL60) cell lines (Figure 30B, bottom panel).

We then sought to investigate whether genes found within the regions that switch compartment identity, exhibited any changes to their expression values. Interestingly, there was no significant difference in mean expression values of the genes found in regions that switch compartment identity. This observation suggests that rather than

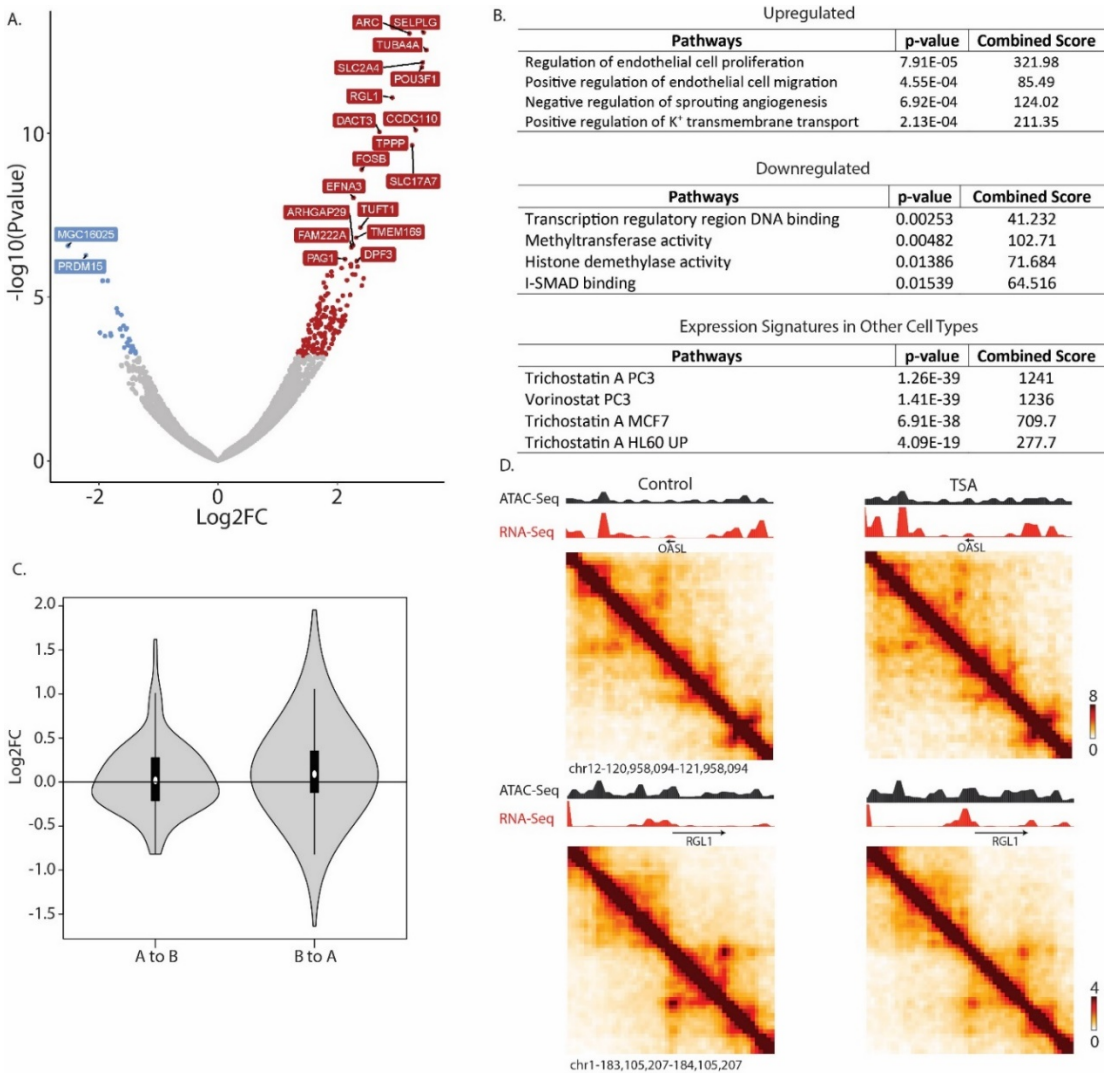


Figure 30. Changes in gene expression relate to TSA signature genes and display alterations of the local structure.

(A) Volcano plot representing RNA-Seq genes upon treatment with TSA. Genes colored in red are upregulated (Log2FC >1, adj. p-value < 0.05) while genes colored in blue are downregulated (Log2FC < -1, adj. p-value < 0.05). (B) Gene ontology analysis of upregulated genes (top panel), downregulated genes (middle panel) and expression signature upon TSA treatment (bottom panel). (C) Violin plot representing expression values of genes that were found in regions of the genome that switched compartment identity. (D) 20Kb binned HiC interaction maps smoothed at 40Kb of upregulated genes that fall under TSA signature in other cell types. Tracks over the matrices represent ATAC-Seq (black) and RNA-Seq (red) signal binned at 40Kb.

larger scale genomic changes, more local scale changes could correlate with changes in gene expression. Therefore, we decided to zoom into individual genes that are part of TSA expression signature in other cell types to investigate whether local structure is affected (Figure 30D). Zooming into *OASL* gene (Log2FC=2.06) revealed subtle differences in the local structure of the chromatin around the gene. The pre-existing strong interactions that resemble loops, seem to get stronger (Figure 30D, top panels). Additionally, an overall increase in chromatin accessibility is observed (ATAC-Seq track, black). In addition, we zoomed in into *RGL1* gene (Log2FC ~ 3.0) which also exhibited difference in the local structure with an overall increase of chromatin accessibility.

Local chromatin structure changes and disruption of chromatin loops upon treatment with TSA

Given our observations about disruption of local chromatin structure around the upregulated genes, we then investigated in more detail the effect of TSA treatment in 3D local structure (Figure 31). We investigated the effect of TSA on chromatin accessibility by using Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq). Cells treated with TSA exhibited an increase in local chromatin accessibility (Figure 31A). This observation supports previous microscopy data that TSA overall decondenses the chromatin. We then sought to investigate how this change in accessibility upon TSA treatment affected local 3D structure by investigating TADs. As previously mentioned, TADs have been shown to be important in regulating gene expression since enhancers and promoters have a high probability of interaction when they are found within the same TAD (Lupianez et al., 2015). To visualize the alterations in contacts around TAD

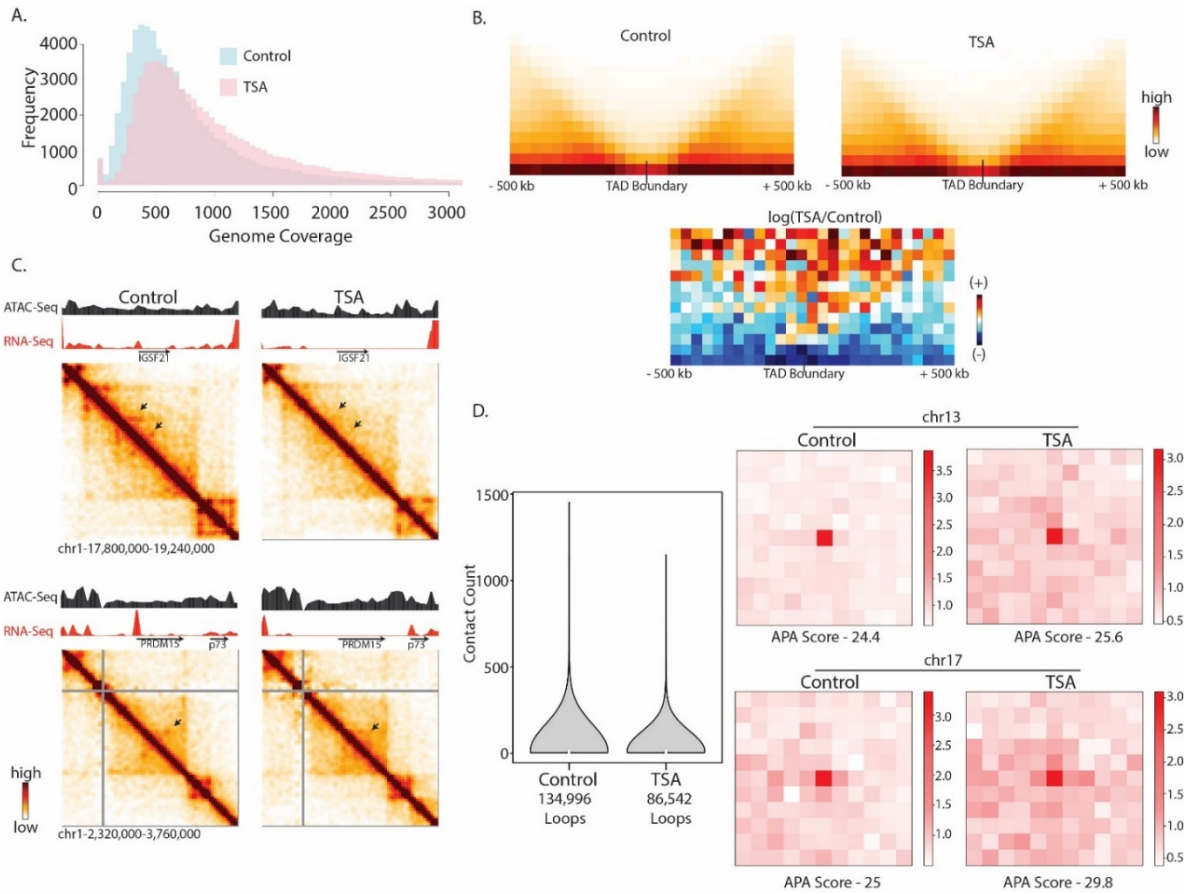


Figure 31. Treatment with TSA increases chromatin accessibility and affects local structure of the genome.

(A) Genome coverage plots of ATAC-Seq signals binned at 40Kb bins. (B) Aggregate contact maps at TAD boundaries binned at 40Kb bins. (C) 20Kb binned HiC contact matrices around the regions that exhibited differential TAD activity. Tracks over the heatmap indicate ATAC-Seq (black) and RNA-Seq (red) signal. Matrices and tracks have been smoothed at 40Kb. Pointing arrows indicate loping interactions in Control which are lost upon TSA treatment. (D) Left panel: Violin plot displaying contact counts on looping regions across Control and TSA samples. Right panels: Aggregate heatmaps of loops called in chr13 and chr17.

boundaries, we plotted the average interaction profile around TAD boundaries (Figure 31B). When we compute the log2ratio (TSA/Control) we observe an increase across TAD boundaries upon TSA treatment which indicates disruption of the initial structure and an increase in interactions between TADs.

To investigate further the changes in TADs, we quantified TAD activity which represents the mean number of interactions within common TADs between Control and TSA treatment. We were able to identify regions of the genome that showed a disruption in interactions within TADs. We zoomed into two different regions of chr1 that exhibited a decrease in TAD activity (Figure 31C). Upon treatment with TSA, we indeed observe a disruption of the general structure around and within the TADs which also affects the expression of genes located in those regions (Figure 31C, RNASeq track). Despite the overall increase in ATAC-Seq signal and therefore increase in chromatin accessibility, the disruption of the structure around the IGSF21 gene correlates with its downregulation (Figure 31C, top panel). Additionally, disruption of the local structure around PRDM15 gene (which is also one of the few downregulated genes upon TSA treatment) correlated with in decrease in expression (Figure 31C, bottom panel). Interestingly, an increase in overall interactions within the tumor suppressor p73 genomic location, is correlated with its increase in expression and could therefore support previous evidence about the antitumor activity of this drug. The decrease in expression of the genes is correlated with loss of looping interaction upon TSA treatment (Figure 31C, pointed arrows).

To investigate how looping interactions were affect across the whole genome, we used a loop calling algorithm that detects loops of 100kb in size (Figure 31D). Surprisingly, we observed a decrease in looping interactions upon TSA treatment (86,542 loops for TSA and 134,996 for Control) (Figure 31D, violin plots). We then piled all the looping interactions for each specific chromosome in Control and TSA treated cells (Figure 31D). Treatment with TSA caused an overall increase in background interaction around the loop anchor as observed in chr13 and chr17 (Figure 31D, right panels) which could be due to the decondensation of the chromatin fiber.

Discussion

Increasing evidence on disruption of the epigenetic landscape and its influence in misregulation of cancer related genes, has attracted the development of epigenetic active drugs that are thought to reverse the chromatin state, such as HDAC inhibitors (HDACi). Extensive research of the affect HDACis exhibit on different types of cancers has revealed its chromatin condensation potency and altering of gene expression. However, it has not been clear how the 3D genome organization is affected upon the decondensation of the chromatin by HDACis.

Here, we show that the overall global structure of the genome is maintained, with each chromosome maintaining their own territories and no evidence of any rearrangements. Additionally, we show that compartment identity is maintained with only subtle compartment identity switches of the regions that exhibit weak compartmentalization.

The major effects of TSA mostly occur at local scale where upregulation of genes that are known to be affected upon TSA treatment in other cell types, also show alterations of their local 3D genome structure. The overall decondensation of the chromatin fiber correlated with an increase in accessibility and disruption of chromatin loops. Disruption of loops could be in part due to the pushing forces exerted on the chromatin upon decondensation. However, the importance of this global decondensation and changing of the physical chromatin structure in the efficacy of the drug is not clear. These drugs exhibit low efficacy and increase in non-specific effects which could be in part due to their global effect on the chromatin fiber. It would therefore be important to consider the changes in 3D genome structure during cancer progression in order to develop therapeutics with high efficacy targets. We and others have shown that during cancer progression, large scale genomic rearrangements take place and would therefore be worth to consider how these regions can be targeted in order to deliver better cancer therapeutics.

Materials and Methods

Cell lines and cell culture

A375 human melanoma cells were obtained from ATCC (CRL-1619). Cells were grown using complete DMEM medium (Corning – 10-013-CV; 10% FBS, 1% Pen-Strep, 1% L-Glutamine) at 37oC supplied with 5% CO2.

Transwell Migration Assay

For the transwell migration assay, we used transwells with 5 µm pore sizes (VWR-10769-236). Briefly, the bottom of the transwells were coated with 40 µL of 10 µg/mL fibronectin

for ~ 45 minutes. 24 well plates were prepared for transwell migration assay adding 500 µL of 1x DMEM (Corning) with full supplements per well. A375 cells were detached from culture dishes at 80-90% confluency and aliquoted to 100,000 cells per 100 µL of 1xDMEM. Each transwell was placed into its corresponding well of the 24 well plate and 100 µL of the cell suspension was added to the top of each filter. Cells were incubated at 37^{o}C, 5% CO2 and allowed to migrate for 24 hours. After the 24-hour incubation, migration efficiency was quantified as follows. First, freely floating cells were removed from the top of the filter (unmigrated; “Top” cells) and from the well beneath the filter (migrated “Bottom” cells) and placed in two separate tubes. Then, 400 µL of trypsin was added into the bottom chamber of the 24 well plate and 200 µL of trypsin was added into the top chamber to detach any remaining attached cells. Recovered cells after trypsinization were added to the unmigrated or migrated tubes, accordingly. Cells were spun down (1000 rpm, 5min) and counted (using trypan blue) to calculate % migration such as: #bottom/(#top+#bottom).

Immunocytochemistry

IF staining

Approximately 50,000 – 100,000 cells for A375 Control and TSA treated were seeded into 35-mm coverslip bottom dishes (MatTek). Cells were allowed to attach overnight and then crosslinked with 4% formaldehyde for 10 min followed by three, 5-minute washes with PBS. After washing, cells were permeabilized with permeabilization buffer (10% goat serum, 0.5% Triton in PBS) for 1 hour at room temperature. After the incubation, primary antibody (H3K9ac), diluted in antibody dilution buffer (5% Goat serum, 0.25% Triton in

PBS) was added and incubated overnight at 4^{o}C. After primary antibody incubation, cells were washed 3 times with PBS for 5 minutes each wash. Cells were then incubated with secondary antibodies (either Alexa Fluor 594 or Alexa Fluor 488) per manufacturer's directions (2 drops of secondary antibody/1mL of PBS – R37117 and R37120) for 30 minutes at room temperature. After secondary antibody incubation, cells were washed 3 times with PBS and sealed using mounting media with DAPI. Slides were treated in the mounting media for 24 hours before imaging.

Cells were imaged using a Leica Sp8 Confocal microscope equipped with a 63x oil immersion objective.

HiC

Library Preparation

Hi-C libraries were prepared for each cell sub-population (A375-Control, Top-5, Bottom-5, Top20-5, Bottom20-5, Top-12 and Bottom-12) as previously described (Golloshi et al., 2018) including two biological duplicates for each condition. Briefly, 10 million cells were grown in T-75 cell culture flasks to 80-90% confluency and crosslinked with 1% formaldehyde for 10 min. Crosslinked cells were then suspended in lysis buffer to permeabilize the cell membrane and were dounce homogenized. Chromatin was then digested in-nucleus overnight using DpnII restriction enzyme. Digested ends were filled in with biotin-dATP, and the blunt ends of interacting fragments were ligated together. DNA was then purified by phenol-chloroform extraction. For library preparation, the NEBNext Ultra II DNA Library prep kit (NEB) was used for libraries with size ranges from

200-400bp. End Prep, Adaptor Ligation, and PCR amplification reactions were carried out on bead bound DNA libraries.

Sequencing was performed at the Oklahoma Medical Research Foundation Clinical Genomics facility using the Illumina HiSeq 3000 platform with 75 bp paired end reads or a NovaSeq 6000 with 50 bp paired end reads. Sequencing reads were mapped to the human genome (hg19), filtered, and iteratively corrected as previously described (Imakaev et al., 2012) (https://github.com/dekkerlab/cMapping). All Hi-C contact matrices were scaled to a sum of 1 million interactions to allow comparisons between conditions in downstream analysis. For Hi-C library quality see Table 5.

Compartment Analysis

Compartment analysis was performed by running principal component analysis using matrix2compartment.pl script in the cworld-dekker pipeline available on GitHub (https://github.com/dekkerlab/cworld-dekker). The PC1 value was then used to determine compartment identity for 100 and 250Kb binned matrices. We considered PC1 values greater than 0.01 to be A compartment and less than -0.01 to be B compartment for this analysis and determination of compartment switches.

Table 5. Summary statistics for Hi-C sequencing experiments (Chapter 4)

Sample	Total Number of Raw Reads	Number of Both Sides Mapped	Valid Pairs	Unique Valid Pairs	% Cis Interactions	% Dangling Ends
Control-R1	201,352,157	124,670,760	123,653,686	107,358,325	61.2	0.35
Control-R2	200,881,087	102,710,892	101,397,252	80,394,938	47.7	0.31
TSA-R1	174,806,387	108,208,928	107,103,301	95,258,083	60.23	0.38
TSA-R2	225,484,425	117,460,789	116,030,441	94,180,623	35.11	0.42

Saddle Plots

Saddle plots were constructed to investigate changes in interaction frequency between or within compartments in highly migratory and non-migratory melanoma cells. Briefly, PCA analysis was performed on 100Kb binned matrices to assign compartment identity for each bin. For the analysis, interaction zScores, normalizing for interaction decay for genomic distance, were calculated according to the matrix2compartment.pl script in cworld-dekker. Zscore matrices were reordered based on compartment strength (from strongest B to strongest A using eigenvector values). Finally, matrices were smoothed at 500Kb bin size.

Interchromosomal Fraction (ICF)

Ratio of interchromosomal interactions relative to the total number of interactions for a given chromosome as previously described (Heinz et al., 2018).

$$ICF = \frac{Interchromosomal\ Interactions}{Inter + Intra\ chromosomal\ Interactions}$$

$$\text{deltaICF} = ICF_{TSA} - ICF_{Control}$$

Distal to Local Ratio (DLR)

Log2 ratio of HiC interactions with distance greater than 3Mb relative to local interactions less than 3 Mb away. To find the change in DLR after constricted migration, delta DLR was calculated as $DLR_{TSA} - DLR_{Control}$ as previously described (Heinz et al., 2018).

$$DLR = log2\frac{Distal\ Interactions\ (> 3Mb)}{Local\ Interactions\ (< 3Mb)}$$

$$deltaDLR = DLRTSA - DLRControl$$

Distance Decay Scaling Plots

Using 40 and 20 Kb binned iteratively corrected and scaled contact matrices, we extracted contact frequencies between bins at each genomic distance, excluding the diagonal bin (zero distance). A loess fit was then used to find a smooth curve describing interaction decay vs. distance (using the matrix2loess.pl script in cworld-dekker). The interaction frequencies were then normalized to set the maximum value (loess fit interaction value for the minimum distance) for each dataset to 1 and then plotted on a log scale vs. log genomic distance.

RNA-Seq

Library Preparation

For Control, Top-5 and Bottom-5 A375 cells, RNA was extracted, and the library was prepared in our laboratory. Briefly, RNA was extracted using Qiagen RNAeasy plus mini kit (Cat. No. 74134). Cells were lysed and homogenized, spun down in gDNA eliminator columns to remove any genomic DNA. All samples were washed with ethanol and the total RNA was eluted. The quality and quantity of the RNA were quantified using Agilent Bioanalyzer Nano RNA kit. All the libraries used were characterized by a RIN value between 9 and 10. rRNA was depleted using NebNext rRNA Depletion Kit (Cat. No. E6310S). Total RNA was hybridized to the probe followed by RNaseH and DNase I digestion. RNA was then purified using AmpureXP beads from Beckman Coulter (Ref. A63881) at a 2.2x concentration of beads. RNA libraries were prepared using NebNext

Ultra II RNA library prep Kit (Cat. No. E7770G). After rRNA depletion, RNA was fragmented (~200 nt) and primed following first and second strand cDNA synthesis. The double stranded cDNA was then purified using AmpureXP beads at a 1.8X concentration. After purification, cDNA libraries were end prepped, ligated adapter and PCR amplified for 7 cycles to enrichment for adaptor ligated DNA using NebNext Multiplex Oligos for illumine (Cat. No. E7335G). cDNA librariers were then purified using AmpureXP beads at a 0.9X concentration and their quality was checked using Agilent Bioanalyzer high sensitivity DNA kit. cDNA samples were then sent to Oklahoma Medical Research Facility (OMRF) for high throughput sequencing.

For the rest of A375 libraries, Top20-5, Bottom20-5, Top-12 and Bottom-12, RNA was extracted using the Qiagen RNAeasy plus mini kit as described above. Total RNA was then sent for library preparation and sequencing at GeneWiz. For RNA-Seq library quality see Table 6.

RNA-Seq Data Analysis

Quality of reads was checked using Fastqc and adapter sequences were trimmed using BBTools (https://github.com/kbaseapps/BBTools). Additionally, quality trimming of the reads was performed and any reads with a lower than 28 quality score were discarded. Reads were then aligned using STAR alignment (https://github.com/alexdobin/STAR) with an average alignment rate of 89%. The aligned reads were then sorted by genomic position and feature counts was performed using htseq 0.11.1(https://github.com/simon-

Table 6. Summary statistics for RNA sequencing experiments (Chapter 4).

Sample	Total Number of Reads	Total number of Mapped Reads	% Mapped Reads
Control-R1	34,469,058	32,948,429	95.59
Control-R2	28,119,492	26,660,183	94.81
TSA-R1	37,499,876	36,069,212	96.18
TSA-R2	27,367,500	25,887,820	94.59

anders/htseq). Differential expression of genes was determined by using DESeq through Galaxy.

ATAC-Seq

ATAC-Seq libraries for Top-5 and Bottom-5 cells was prepared as previously described (Buenrostro et al., 2015). Libraries were generated using Ad1_noMX and Ad2.1-2.3 barcoded primers from (Buenrostro et al., 2015) and were amplified for a total of 12-13 cycles. The quality and quantity of the libraries were measured using Agilent Bioanalyzer High Sensitivity DNA kit. Sequencing was performed at Oklahoma Medical Research Facility using HiSeq3000 with 75 bp paired end reads.

ATAC-Seq Data Analysis

Adapters and reads with a quality score less than 20 were trimmed from the final libraries using Skewer/0.2.2 (https://github.com/relipmoc/skewer). Trimmed reads were aligned to the hg19 genome using bowtie2 with default parameters (https://github.com/BenLangmead/bowtie2). Aligned reads were sorted, indexed and ChrM reads were removed using samtools. Duplicated reads were removed using Picard (https://github.com/broadinstitute/picard/blob/master/src/main/java/picard/sam/markduplicates/MarkDuplicates.java). Finally, peaks were called using MACS2 (https://github.com/taoliu/MACS) with the following parameters: --nomodel --extsize 300 -g hs --keep-dup all.

References

ABBAS, A., HE, X., NIU, J., ZHOU, B., ZHU, G., MA, T., SONG, J., GAO, J., ZHANG, M. Q. & ZENG, J. 2019. Integrating Hi-C and FISH data for modeling of the 3D organization of chromosomes. *Nat Commun,* 10**,** 2049.

BALI, P., PRANPAT, M., BRADNER, J., BALASIS, M., FISKUS, W., GUO, F., ROCHA, K., KUMARASWAMY, S., BOYAPALLE, S., ATADJA, P., SETO, E. & BHALLA, K. 2005. Inhibition of histone deacetylase 6 acetylates and disrupts the chaperone function of heat shock protein 90: a novel basis for antileukemia activity of histone deacetylase inhibitors. *J Biol Chem,* 280**,** 26729-34.

BAYLIN, S. B. & JONES, P. A. 2011. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer,* 11**,** 726-34.

BELAGHZAL, H., BORRMAN, T., STEPHENS, A. D., LAFONTAINE, D. L., VENEV, S. V., WENG, Z., MARKO, J. F. & DEKKER, J. 2019. Compartment-dependent chromatin interaction dynamics revealed by liquid chromatin Hi-C. *bioRxiv***,** 704957.

BERDASCO, M. & ESTELLER, M. 2010. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell,* 19**,** 698-711.

BIGGS, R., LIU, P. Z., STEPHENS, A. D. & MARKO, J. F. 2019. Effects of altering histone posttranslational modifications on mitotic chromosome structure and mechanics. *Mol Biol Cell,* 30**,** 820-827.

CAROLYN DE GRAAF, J. K., LUDO PAGIE, DAAN PERIC HUPKES, BAS VAN STEENSEL. 2019. *LAD Atlas* [Online]. Available: https://osf.io/dk8pm/ [Accessed].

DE RUIJTER, A. J., VAN GENNIP, A. H., CARON, H. N., KEMP, S. & VAN KUILENBURG, A. B. 2003. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J,* 370**,** 737-49.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

GASPAR-MAIA, A., ALAJEM, A., MESHORER, E. & RAMALHO-SANTOS, M. 2011. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol,* 12**,** 36-47.

GERLITZ, G. & BUSTIN, M. 2010. Efficient cell migration requires global chromatin condensation. *J Cell Sci,* 123**,** 2207-17.

GLOZAK, M. A., SENGUPTA, N., ZHANG, X. & SETO, E. 2005. Acetylation and deacetylation of non-histone proteins. *Gene,* 363**,** 15-23.

HEINZ, S., TEXARI, L., HAYES, M. G. B., URBANOWSKI, M., CHANG, M. W., GIVARKES, N., RIALDI, A., WHITE, K. M., ALBRECHT, R. A., PACHE, L., MARAZZI, I., GARCIA-SASTRE, A., SHAW, M. L. & BENNER, C. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell,* 174**,** 1522-1536 e22.

INOUE, K., KOBAYASHI, M., YANO, K., MIURA, M., IZUMI, A., MATAKI, C., DOI, T., HAMAKUBO, T., REID, P. C., HUME, D. A., YOSHIDA, M., AIRD, W. C., KODAMA, T. & MINAMI, T. 2006. Histone deacetylase inhibitor reduces monocyte adhesion to endothelium through the suppression of vascular cell adhesion molecule-1 expression. *Arterioscler Thromb Vasc Biol,* 26**,** 2652-9.

KIND, J., PAGIE, L., DE VRIES, S. S., NAHIDIAZAR, L., DEY, S. S., BIENKO, M., ZHAN, Y., LAJOIE, B., DE GRAAF, C. A., AMENDOLA, M., FUDENBERG, G., IMAKAEV, M., MIRNY, L. A., JALINK, K., DEKKER, J., VAN OUDENAARDEN, A. & VAN STEENSEL, B. 2015. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell,* 163**,** 134-47.

KRIJGER, P. H. & DE LAAT, W. 2016. Regulation of disease-associated gene expression in the 3D genome. *Nat Rev Mol Cell Biol,* 17**,** 771-782.

LARSON, A. G., ELNATAN, D., KEENEN, M. M., TRNKA, M. J., JOHNSTON, J. B., BURLINGAME, A. L., AGARD, D. A., REDDING, S. & NARLIKAR, G. J. 2017. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature,* 547**,** 236-240.

LITVAK, D. A., EVERS, B. M., HWANG, K. O., HELLMICH, M. R., KO, T. C. & TOWNSEND, C. M., JR. 1998. Butyrate-induced differentiation of Caco-2 cells is associated with apoptosis and early induction of p21Waf1/Cip1 and p27Kip1. *Surgery,* 124**,** 161-9; discussion 169-70.

LUPIANEZ, D. G., KRAFT, K., HEINRICH, V., KRAWITZ, P., BRANCATI, F., KLOPOCKI, E., HORN, D., KAYSERILI, H., OPITZ, J. M., LAXOVA, R., SANTOS-SIMARRO, F., GILBERT-DUSSARDIER, B., WITTLER, L., BORSCHIWER, M., HAAS, S. A., OSTERWALDER, M., FRANKE, M., TIMMERMANN, B., HECHT, J., SPIELMANN, M., VISEL, A. & MUNDLOS, S. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell,* 161**,** 1012-1025.

MA, J., GUO, X., ZHANG, S., LIU, H., LU, J., DONG, Z., LIU, K. & MING, L. 2015. Trichostatin A, a histone deacetylase inhibitor, suppresses proliferation and promotes apoptosis of esophageal squamous cell lines. *Mol Med Rep,* 11**,** 4525-31.

MARCHAL, C., SIMA, J. & GILBERT, D. M. 2019. Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol*.

MARIADASON, J. M., VELCICH, A., WILSON, A. J., AUGENLICHT, L. H. & GIBSON, P. R. 2001. Resistance to butyrate-induced cell differentiation and apoptosis during spontaneous Caco-2 cell differentiation. *Gastroenterology,* 120**,** 889-99.

MARKS, P., RIFKIND, R. A., RICHON, V. M., BRESLOW, R., MILLER, T. & KELLY, W. K. 2001. Histone deacetylases and cancer: causes and therapies. *Nat Rev Cancer,* 1**,** 194-202.

MARKS, P. A., RICHON, V. M. & RIFKIND, R. A. 2000. Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. *J Natl Cancer Inst,* 92**,** 1210-6.

ROWLEY, M. J. & CORCES, V. G. 2018. Organizational principles of 3D genome architecture. *Nat Rev Genet,* 19**,** 789-800.

SPIELMANN, M., LUPIANEZ, D. G. & MUNDLOS, S. 2018. Structural variation in the 3D genome. *Nat Rev Genet,* 19**,** 453-467.

STROM, A. R., EMELYANOV, A. V., MIR, M., FYODOROV, D. V., DARZACQ, X. & KARPEN, G. H. 2017. Phase separation drives heterochromatin domain formation. *Nature,* 547**,** 241-245.

SUN, S., HAN, Y., LIU, J., FANG, Y., TIAN, Y., ZHOU, J., MA, D. & WU, P. 2014. Trichostatin A targets the mitochondrial respiratory chain, increasing mitochondrial reactive oxygen species production to trigger apoptosis in human breast cancer cells. *PLoS One,* 9**,** e91610.

TOTH, K. F., KNOCH, T. A., WACHSMUTH, M., FRANK-STOHR, M., STOHR, M., BACHER, C. P., MULLER, G. & RIPPE, K. 2004. Trichostatin A-induced histone acetylation causes decondensation of interphase chromatin. *J Cell Sci,* 117**,** 4277-87.

VAN STEENSEL, B. & BELMONT, A. S. 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell,* 169**,** 780-791.

WATSON, J. A., MCKENNA, D. J., MAXWELL, P., DIAMOND, J., ARTHUR, K., MCKELVEY-MARTIN, V. J. & HAMILTON, P. W. 2010. Hyperacetylation in prostate cancer induces cell cycle aberrations, chromatin reorganization and altered gene expression profiles. *J Cell Mol Med,* 14**,** 1668-82.

YOSHIDA, M., KIJIMA, M., AKITA, M. & BEPPU, T. 1990. Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J Biol Chem,* 265**,** 17174-9.

CHAPTER V: THE EFFECT OF MECHANICAL STRAIN ON 3D GENOME ORGANIZATION

The work presented in this chapter is an ongoing collaboration with Dr. Jan Lammerding at Cornell University and has not yet been submitted for publication.

My contribution to this work included: (1) Performing all Hi-C experiments, (2) computational data analysis of the Hi-C experiments and (3) writing the current manuscript and making the figures.

Abstract

The ability of the nucleus to sense and withstand mechanical forces reflects a biological characteristic important for cell growth and development. The nuclear lamina and chromatin contribute to the stiffness of the nucleus which influences its mechanical properties. Mechanical force application on the nucleus has been shown to induce changes in the structure of chromatin and influence expression of mechanosensing genes. However, it is not clear how these changes in gene expression arise. The chromatin inside the nucleus is non-randomly packed for the proper control of DNA replication, repair and gene expression into chromosomal territories, active and inactive compartments whose gene expression pattern is regulated by regulatory loops and Topologically Associating Domains (TADs). Proper maintenance of this organization is important and has been implicated to be altered in many diseases. Here, we investigate the effect of mechanical stretching on the 3D organization of the genome. We find that the local 3D genomic contacts are not affected upon mechanical stretching. We observe changes in weak compartment identify and small differences in compartment interactions. The regions that are most affected are comprised of genes involved in mechanosensing.

Overall, our work provides initial evidence of an altered 3D genome structure upon mechanical stretching which correlates with transcriptional response upon these events.

Introduction

Mechanical forces such as compressive and tensile forces have been observed across multicellular organisms and are important for growth and development (Zwerger et al., 2011). It has been shown these forces affect the size and shape of cells and induce changes in expression of genes that are key regulators of mechanotransduction (Reidy and Bowyer, 1977, Tucker and Meats, 1976, Sun et al., 2016, Murthy et al., 2017). While the cytoskeletal components and their mechanosensing properties have been widely studied, it has been recently shown that the structure and mechanical properties of the nucleus play an important role in mechanotransduction (Guilluy et al., 2014, Tajik et al., 2016, Wang et al., 2009).

The nucleus responds to mechanical stress by signal propagation from the cytoskeleton through the Linker of Nucleoskeleton and Cytoskleton (LINC) complex which enables a continuous force transmission from the microenvironment (Alam et al., 2015). The mechanical properties of the nucleus and ability to withstand force, are determined by the contribution of the nuclear lamina and the nuclear interior to the stiffness of the nucleus. The nuclear lamina is highly deformable and can undergo large amounts of stretch (Dahl et al., 2004) while the chromatin inside the nucleus represents a more viscoelastic environment (Rowat et al., 2013). Mutations in the nuclear envelope proteins have been shown to contribute to nuclear fragility and have had a clinical impact on multiple diseases

(Simon and Wilson, 2011) such as muscular dystrophy, laminopathies (e.g. Hutchinson Gilford Progeria Syndrome, HGPS) and cardiomyopathies.

Chromatin has also been shown to modulate the mechanical properties of the nucleus by altering nuclear stiffness. The 2-meter chromatin fiber is tightly packaged inside the nucleus to regulate DNA replication, repair and faithful expression of specific genes. This regulation is in part controlled by post translational modification of chromatin with different histone marks that inform two major chromatin states: heterochromatin and euchromatin (Gaspar-Maia et al., 2011, van Steensel and Belmont, 2017). Heterochromatic regions are tethered to the nuclear lamina and regulate gene expression signatures in the cell (van Steensel and Belmont, 2017). Changes in chromatin state, such as increasing of heterochromatin or increasing of euchromatin decrease alter the mechanical properties of the nucleus (Stephens et al., 2017, Stephens et al., 2018). Response to mechanical stress can cause deformation of the nucleus, chromatin and nuclear bodies and can modulate gene expression (Bissell et al., 1982, Guilluy et al., 2014, Tajik et al., 2016).

Ever-improving microscopic techniques and the development of chromosome conformation capture approaches have revolutionized our understanding of 3D genome architecture in the past decade (Dekker et al., 2002, Bickmore and van Steensel, 2013, Rowley and Corces, 2018, Pombo and Dillon, 2015, Abbas et al., 2019). Chromosomes are folded at different length scales into loops, topologically associating domains (TADs), active and inactive (A/B) compartments, and chromosomal territories. How these layers

of structure are affected by physical nucleus and chromatin shape changes induced by mechanical stretching is not yet understood.

Therefore, characterizing the response of the 3D genome organization to mechanical strain will provide useful avenues to further investigate the regulation of such responses in diseased cells. Here, we perform Hi-C experiments on human fibroblast cells (Hffc6) upon mechanical strain and find that the 3D organization is overall robust to such strain. Regions of the genome that are weakly compartmentalized are the most effected and genes found within these regions are important players in mechanotransduction.

Results

Hffc6 skin fibroblasts exhibit subtle changes in global 3D genome architecture after biaxial strain

To induce mechanical strain, Hffc6 cells were starved for a period of 3 days and subjected to strain (4% strain at 1Hz) for 1 hour (ST) or unstrained cells (NST) incubated in the same devices used for straining experiments. Application of mechanical strain leads to elongation of the cytoplasm and nucleus (Isermann et al., 2012), but it is not known how such mechanical stress affects the DNA inside the nucleus.

To investigate the effect of straining in the 3D genome organization, we performed Hi-C in unstrained cells (NST) and 60 min strained cells (ST) (Figure 32A). Interestingly, despite the observed elongation of the nucleus upon mechanical strain, no major

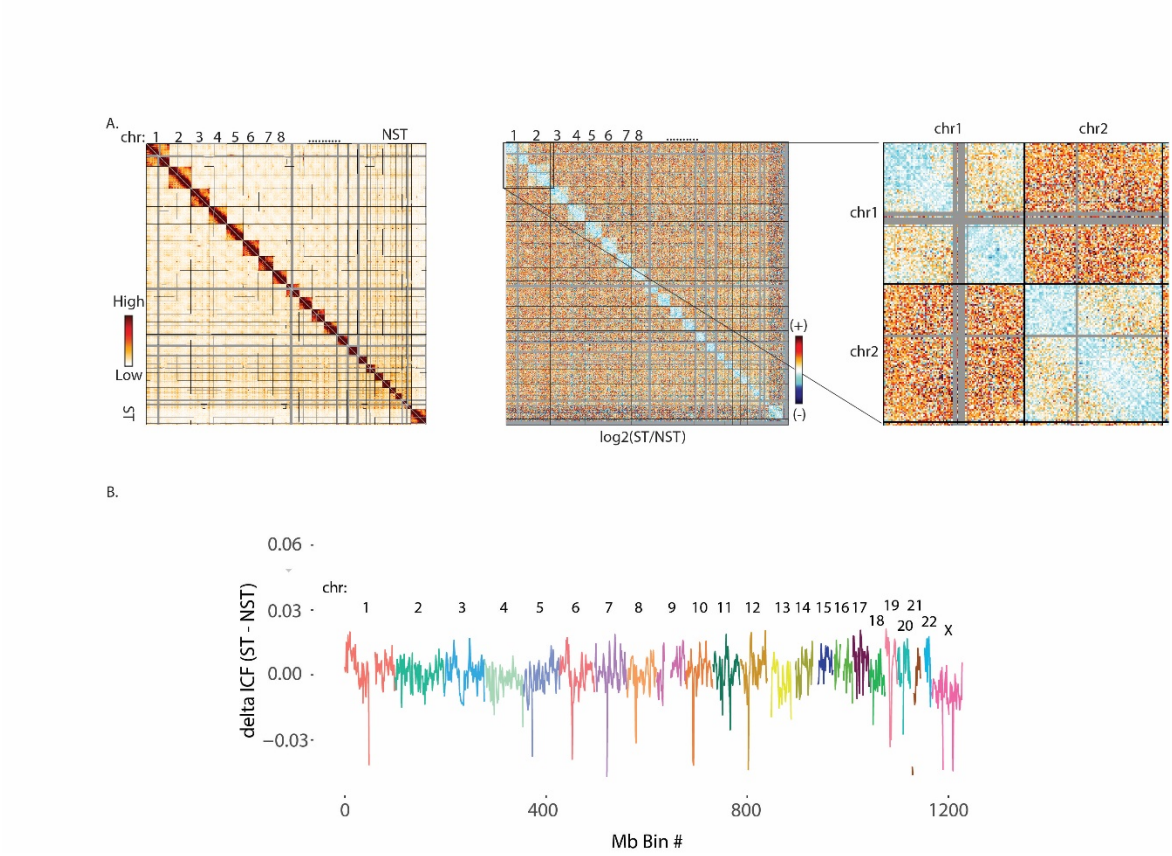


Figure 32. Subtle changes in global 3D genome organization in Hffc6 after 60 min strain.

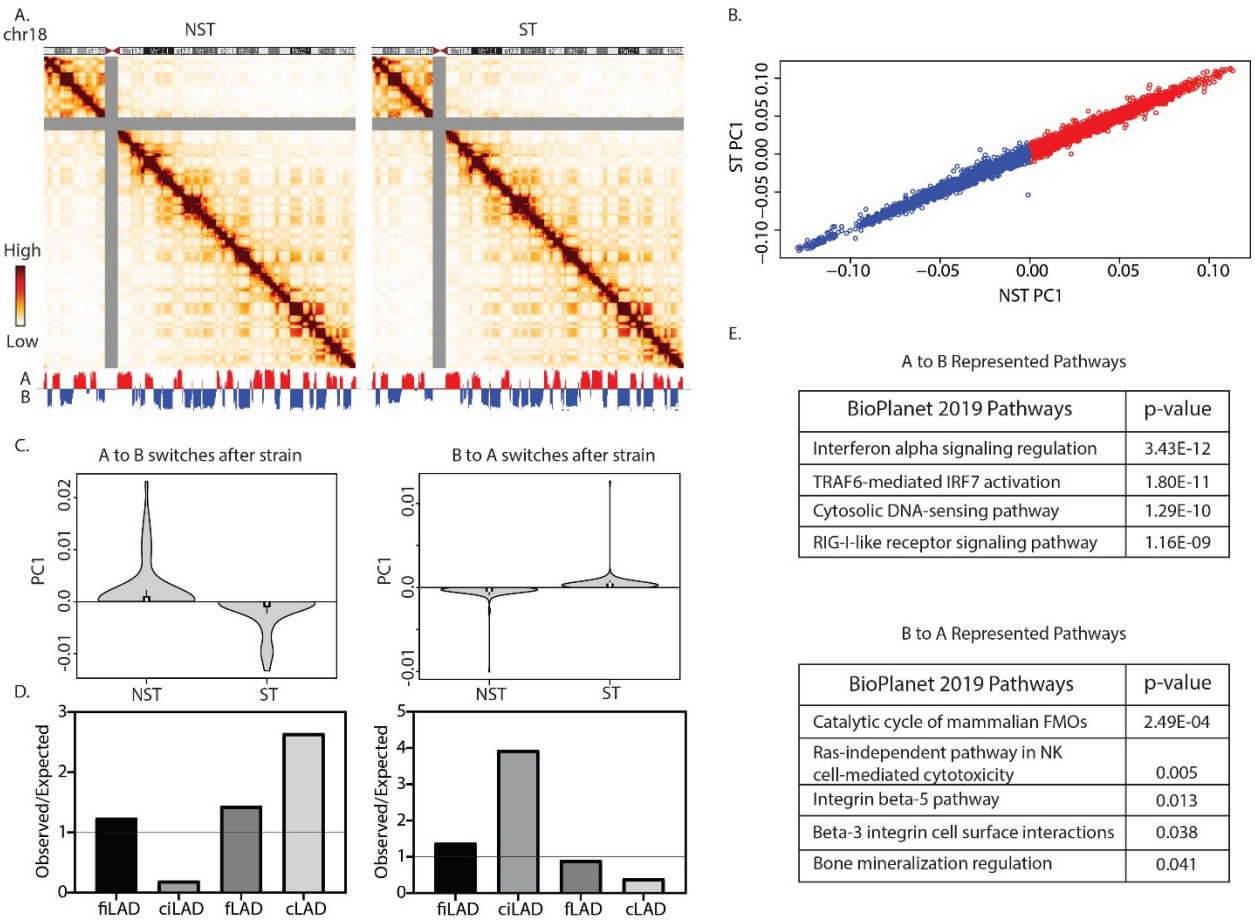
(A) Genome wide Hi-C contact frequency maps binned at 2.5 Mb (left) and log2(ST/NST) contact heatmaps in 2.5 Mb bin interaction matrices. Inset reveals a local decrease of interaction frequencies in chr1 and chr2. (B) The difference in inter-chromosomal fraction of contacts (ICF) in cells that have undergone mechanical strain (ST).

rearrangements of the chromosomes were observed. Investigating Hi-C contact frequency maps binned at 2.5Mb bin sizes revealed that chromosomes maintain their own territories with no major rearrangements in strained cells (lower triangle) when compared to unstrained cells (right triangle) (Figure 32A, left panel). Mathematical comparison of the contact frequencies (log2 ST/NST) suggests an overall increase in intra-chromosomal interactions (Figure 32A, right panel) upon 60 min mechanical strain. This increase in long-range interactions is also evident when zooming in within each chromosome while a decrease in cis interactions is observed across the genome (Figure 32A, inset).

We next sought to explore how the interchromosomal fraction (ICF) of interactions across the genome varies upon 60min mechanical strain (Figure 32B). Interestingly, different ICF patterns are observed among chromosomes. A decrease in ICF was observed in chr4, chr5, chr13, chr18 and chrX. On the other hand, chr14, chr15, chr16, chr17, chr19, chr20 and chr22 exhibit an increase in ICF.

Subtle changes in compartment identity after mechanical strain

To investigate whether changes in compartmentalization are evident in cells that have undergone 60 min mechanical strain, we binned the Hi-C contact frequencies at 250Kb bin sizes which correlates with compartment identity (Figure 33). Overall, no apparent differences were observed in contact frequencies when investigating the heatmaps across all chromosomes (Figure 33A, representative chr18). To correlate the observed heatmap plaid pattern with compartment identity, we performed Principal Component



BioPlanet 2019 Pathways	p-value
Interferon alpha signaling regulation	3.43E-12
TRAF6-mediated IRF7 activation	1.80E-11
Cytosolic DNA-sensing pathway	1.29E-10
RIG-I-like receptor signaling pathway	1.16E-09

BioPlanet 2019 Pathways	p-value
Catalytic cycle of mammalian FMOs	2.49E-04
Ras-independent pathway in NK cell-mediated cytotoxicity	0.005
Integrin beta-5 pathway	0.013
Beta-3 integrin cell surface interactions	0.038
Bone mineralization regulation	0.041

Figure 33. Subtle changes in compartment identity after mechanical strain.

(A) 250Kb binned Hi-C interaction matrices of chr18 in unstrained and strained cells. Line graph represents A (red) and B (blue) compartments for chr18. (B) Correlation plot of PC1 values between unstrained and mechanically strained cells. (C) Violin plots of PC1 values representing regions that switch their compartment identity from A to B (left panel) and B to A (right panel). (D) Enrichment of different LAD types in regions that switch compartment identity from A to B (left) and B to A (right). (E) Enrichment of signature pathways of genes within regions that switch compartment identity using Enrichr.

Analysis (PCA) using the 250Kb binned Hi-C interaction frequencies. Overall, compartment identity remained consistent in Hffc6 cells that underwent mechanical strain when compared to unstrained cells (Figure 33B). However, a small number of the genomic bins exhibited switches in their compartment identity. In fact, 163 genomic bins switched their compartment identity from active (A) to inactive (B). Additionally, 377 genomic bins switched their compartment identity from B to A (Figure 33B). These compartment switches take place in regions that exhibit weak compartment identity (Figure 33C).

To investigate whether regions involved in compartment switching correspond to certain types of Lamin Associated Domains (LADs), we used a previously established LAD atlas across nine different cell types (Figure 33D). Regions of the genome are classified as: cLADs (constitutive LADs; regions associated with the nuclear lamina in all cell types), ciLADs (constitutive inter LADs; regions located in the interior of the nucleus across all cell lines), fLADs and fiLADs (facultative LADs and facultative inter LADs; regions that are sometimes associated with the nuclear lamina across the cell types). Interestingly, regions involved in A to B compartment switches displayed a higher representation of cLADs (Figure 33D, left panel). On the other hand, regions that changed their compartment identity from B to A were over-represented by ciLADs (Figure 33D, right panel).

We then investigated the genes found within these regions that switch compartment identity and the biological pathways they were involved with (Figure 33E). Interestingly, genes found within the regions that switched from A to B compartment identity are associated with interferon signaling pathways (IFNA cluster of genes) which has been shown to be inhibited during mechanical stretching of cardiomyocytes (Rysa et al., 2018) (Figure 33E, top panel). Moreover, when investigating signaling pathways involved with genes that switch their compartment identity from B to A, we find an over-representation of pathways that have been shown to be induced by mechanical stretching of cells (Figure 33E, bottom panel). One of the pathways associated with genes within compartment switches included β5 (VTN, FN1 and MADCAM1) and β3-integrin (VTN, FN1, COL4A5 and COL4A6) pathways which have been previously described for their function in mechanosensing the microenvironment (von Wichert et al., 2003, Kechagia et al., 2019). Finally, bone mineralization pathway (COL4A5 and COL4A6) is represented in the set of genes that switch from B to A compartment and has been previously shown to be induced from mechanical strain applied on osteoblasts (Yu et al., 2016).

Local structure changes are not necessary to induce expression of mechanoresponsive genes

To investigate whether the genes found within the regions that switch compartment identity and important for mechanosensing exhibit differences in their local DNA structure, we zoomed in into these genes using 40Kb binned Hi-C interaction maps (Figure 34). Initially, we zoomed in into the cluster of IFN-α genes found in regions that switched compartment identity from A in unstrained cells to B in mechanically strained cells. We

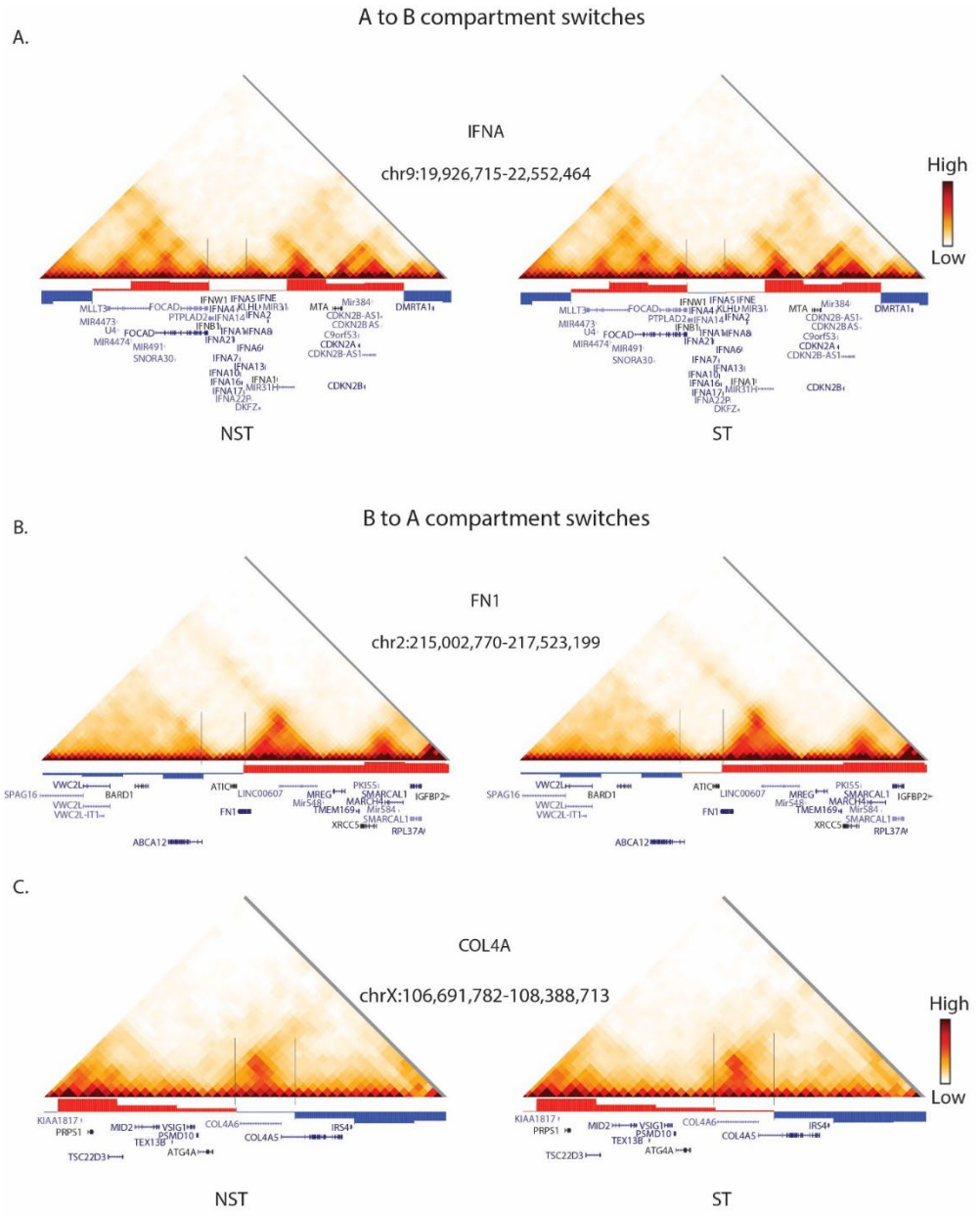


Figure 34. 3D local structure around the genes that switch compartments remains similar.

(A) Hi-C contact frequency heatmaps binned at 40Kb around the IFN-α cluster of genes. (B) Hi-C contact frequency heatmaps binned at 40Kb around the FN1 gene. (C) Hi-C contact frequency heatmaps binned at 40Kb around the COL4A6 and COL4A5 cluster of genes. Red and blue track represent compartment identity at that specific location. Genes listed below the track.

observed a slight difference in interaction frequency by eye by no major rearrangement of the local structure (Figure 34A). We then investigated the genes that were involved in B to A compartment switches. These genes have been shown to be upregulated in cells undergoing mechanical stretching. We first zoomed into the local structure of FN1 gene (chr2:215,002,770-217,523,199) (Figure 34B). Consistent with the previous example, while the compartment identity around this gene changes (red and blue track), there is no major changes within local Hi-C contact frequencies. Additionally, we looked at COL4A6 and COL4A5, located within a B to A compartment switch (Figure 34C). Consistent with the previous examples, while compartment identity switch is observed, no major rearrangements of the local structure is evident.

These observations in Hi-C contacts are not enough to draw conclusions about the transcriptional state of these genes. While changes in Hi-C contacts are not evident, changes in local accessibility of the chromatin fiber could induce transcriptional output and therefore additional experiments that address these scenarios are needed.

Intercompartment interactions are correlated with chromosome positioning within the nucleus

Since the compartment identity switch we observed upon mechanical strain of Hffc6 cells was minimal and most of the regions that switched compartment identity were considered “weak” compartments based on our threshold, we investigated how regions that do not switch compartment identity were affected by mechanical strain. We investigated compartments that were not involved in identity switches by reordering the 250Kb binned

contacts from the strongest B compartments to the strongest A compartment bins based on PC1 values (Figure 35).

At a first glance, reordering by compartment identity strength reveals a clear separation between strongest B from strongest A interacting regions (Figure 35A and 35B) independent of mechanical strain. However, the intra and inter-compartment interaction strength is affected by mechanical strain. We observe an increase in the strongest B-B interactions in chr18 (Figure 35A, third panel) and a decrease in the strongest B-B interactions in chr19 (Figure 35B, third panel). On the other hand, both chr18 and chr19 exhibit an increase in strongest A-A interactions (Figure 35A and 35B). Interestingly, chr18 exhibits a loss of interactions between regions that are weakly compartmentalized or that are found near compartment boundaries.

Given the differences we observe in interaction patterns of inter and intra compartment interactions between chr18 and chr19, we then investigated the average decay of interactions with increasing distance in chr18 and chr19 upon mechanical strain (Figure 35C and 35D, respectively). We investigated the interaction scaling in A compartment and B compartment separately, given their differential response to mechanical force. Surprisingly, the interaction scaling revealed no obvious differences between A or B compartment in chr18 or chr19.

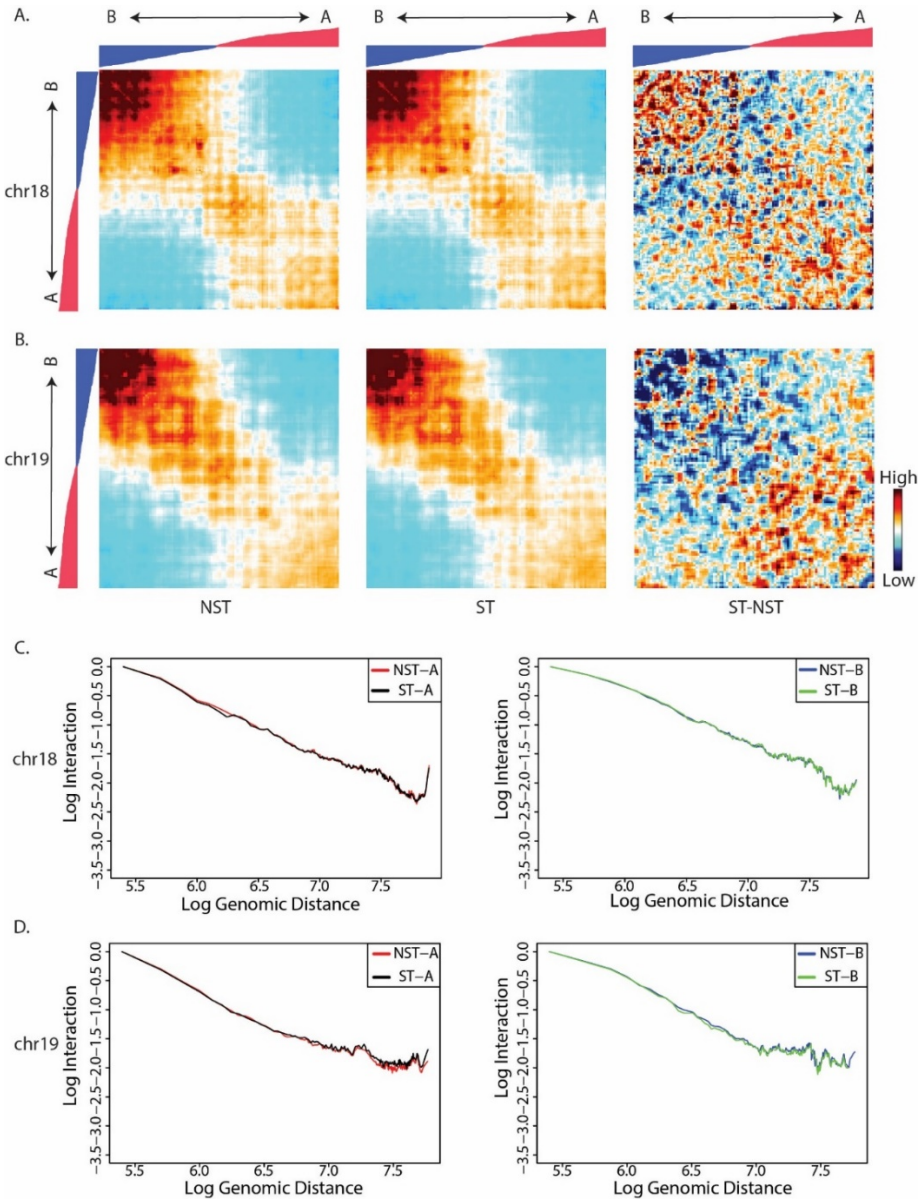


Figure 35. Interaction changes between compartments is dependent on chromosome positioning within the nucleus.

(A) Saddle plots generated from 250Kb binned matrices and smoothed to 500Kb for chr18 reordered from strongest B to strongest A compartment identity. (B) Saddle plots generated from 250Kb binned matrices and smoothed to 500Kb for chr19 reordered from strongest B to strongest A compartment identity. (C) Whole genome scaling plots using Hi-C contact interactions binned at 250Kb bin sizes in chr18 A compartment and B compartment. (D) Whole genome scaling plots using Hi-C contact interactions binned at 250Kb bin sizes in chr19 A compartment and B compartment.

TAD boundary strength increases after mechanical strain

To assess the effect of mechanical strain on the local structure of the chromatin, we investigated Hi-C interactions binned at 40Kb bin sizes (Figure 36). Overall, mechanical straining of the nucleus does not affect the local interactions of the chromatin fiber, independent of gene density or chromosome positioning (Figure 36A).

To assess whether TADs are affected after mechanical stretching, we identified TAD boundaries and calculated the boundary strength (degree of spatial separation between TADs) for each boundary using both the InsulationScore method (Crane et al., 2015). Overall, we observe a strengthening of TAD boundaries upom mechanical stretching (Figure 36B).

To visualize the alterations in contacts around TAD boundaries, we plotted the average interaction profile around TAD boundaries called by the InsulationScore method (Crane et al., 2015) (Figure 36C). We observe an overall decrease in interactions across TAD boundaries that could potentially indicate a protective role of the genome organization in response to mechanical stretching.

Discussion

While mechanical perturbations on the cell and the nucleus in context of the cytoskeleton and nuclear lamina have been widely studied, the effect of such mechanical perturbations

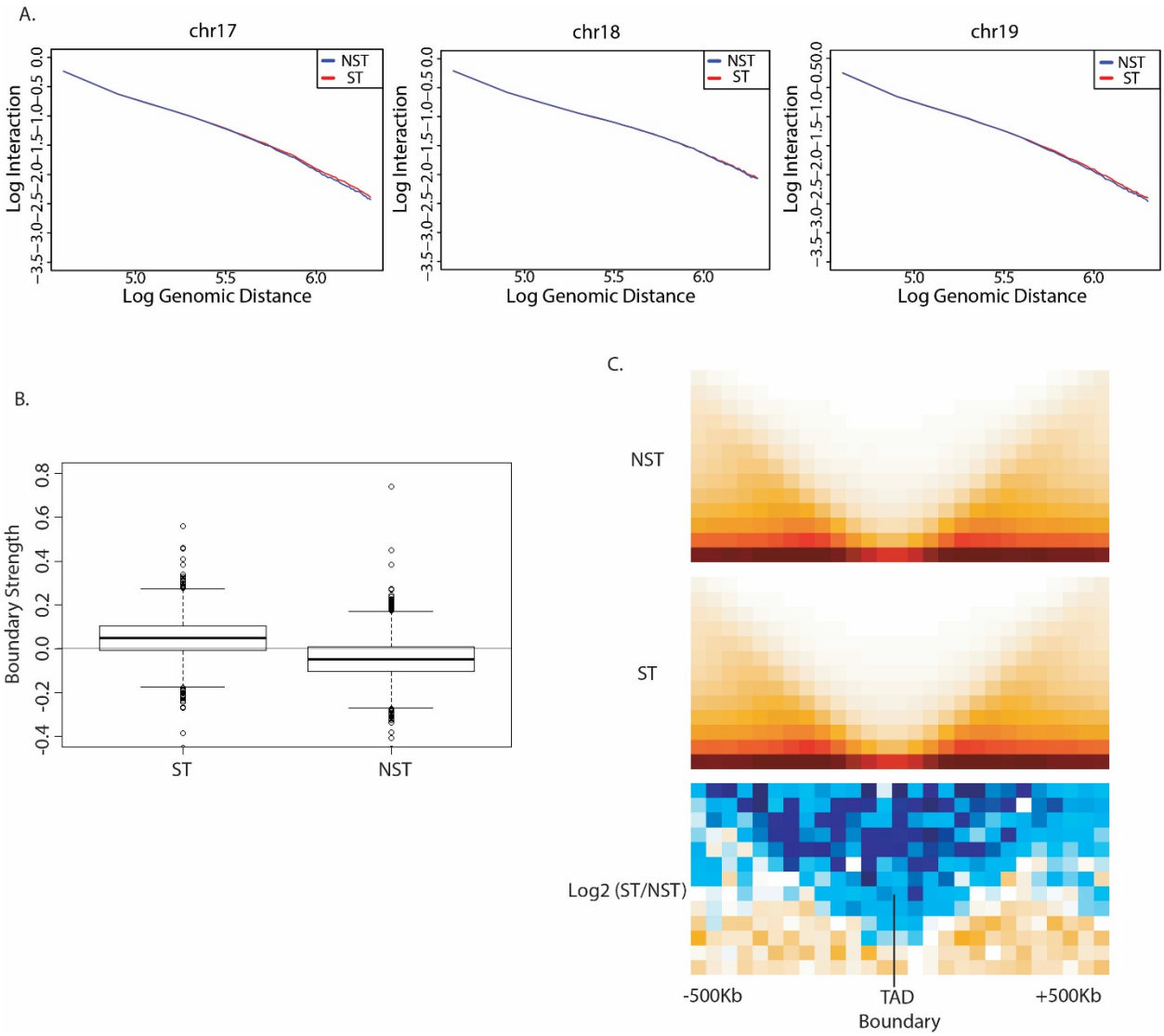


Figure 36. TAD boundaries strengthen after mechanical strain.

(A) Log scaling plots using 40Kb binned matrices for chr17, chr18 and chr19. (B) TAD boundary strength boxplots as calculated by using InsulationScore method. (C) Aggregate contact maps of 40Kb binned matrices at TAD boundaries called by the InsulationScore method with strength greater than 1 in NST (top) and ST (middle) cells. Bottom graph represents log2ratio of ST/NST TAD boundary pileup.

in the 3D organization of the DNA is not well understood. Here, we investigated the response of the 3D genome upon mechanical stretching of fibroblast cells.

We observe compartment identity switches of genomic regions that exhibit weak compartmentalization. Genes found within these genomic regions correspond to genes that are important for mechanotransduction such fibronectin or collagen type IV family of genes. While we do not have direct evidence that suggest switching in compartment identity correlates with modulation of expression of these genes, this observation points to the role of 3D genome structural changes in mechanoresponse. Interestingly, we did not observe any differences in local organization of structures of the genes found within compartment switching. This could imply that larger chromosomal changes such as compartment switches maybe due to liquid-liquid phase separation could potentially influence expression of these genes without modulating local organization of the structure. Another possibility could be that local accessibility of the chromatin fiber is changed that enables binding of the transcriptional machinery to induce expression of mechanosensitive genes.

We also observed that strongest B-B compartment interactions increase in chr18. Given that this chromosome is mainly located within the nuclear periphery, the increase observed could be potentially due to the increase in heterochromatin upon elongated stretching as previously described (Le et al., 2016). Additionally, the loss in the strongest B-B interactions observed in chr19 could be indicative of certain genome regions being

pulled apart as it has been previously observed in neutrophil migration (Irianto et al., 2017, Jacobson et al., 2018).

In addition to these larger genomic changes, we observe some changes at more local scales where the TAD boundaries get stronger upon mechanical stretching indicating a segregation of the genome structure. Stretching of the chromatin fiber could potentially cause this segregation or it could be a protective response to mechanical stretching as it has been previously observed in DNA damage response induced by ionized radiation (Sanders et al., 2019). Such strengthening of TAD boundaries and decreasing contacts between TAD boundaries could be the genome's response to preventing DNA breaks that could potentially lead to deleterious effects.

Future work is needed to clarify the connection between these 3D genomic structural changes with changes in the local accessibility of chromatin states and their role in modulation of mechanosensitive genes.

Materials and Methods

Hi-C

Library Preparation

Hi-C libraries were prepared for each cell sub-population (Hff-NST and HFF-ST) as previously described (Golloshi et al., 2018) including two biological duplicates for each condition. Briefly, 10 million cells were grown in T-75 cell culture flasks to 80-90% confluency and crosslinked with 1% formaldehyde for 10 min. Crosslinked cells were then

suspended in lysis buffer to permeabilize the cell membrane and were dounce homogenized. Chromatin was then digested in-nucleus overnight using DpnII restriction enzyme. Digested ends were filled in with biotin-dATP, and the blunt ends of interacting fragments were ligated together. DNA was then purified by phenol-chloroform extraction. For library preparation, the NEBNext Ultra II DNA Library prep kit (NEB) was used for libraries with size ranges from 200-400bp. End Prep, Adaptor Ligation, and PCR amplification reactions were carried out on bead bound DNA libraries.

Sequencing was performed at the Oklahoma Medical Research Foundation Clinical Genomics facility using the Illumina HiSeq 3000 platform with 75 bp paired end reads or a NovaSeq 6000 with 50 bp paired end reads. Sequencing reads were mapped to the human genome (hg19), filtered, and iteratively corrected as previously described (Imakaev et al., 2012) (https://github.com/dekkerlab/cMapping). All Hi-C contact matrices were scaled to a sum of 1 million interactions to allow comparisons between conditions in downstream analysis. For library quality and mapping statistics see Table 7.

Compartment Analysis

Compartment analysis was performed by running principal component analysis using matrix2compartment.pl script in the cworld-dekker pipeline available on GitHub (https://github.com/dekkerlab/cworld-dekker). The PC1 value was then used to determine compartment identity for 100 and 250Kb binned matrices. We considered PC1 values

Table 7. Summary statistics for Hi-C sequencing experiments (Chapter 5).

Sample	Total Number of Raw Reads	Number of Both Sides Mapped	Valid Pairs	Unique Valid Pairs	% Cis Interaction	% Dangling Ends
NST-R1	446,494,111	232,682,459	228,630,355	172,639,241	81.2	0.58
ST-R2	414,620,689	207,261,409	205,520,811	163,582,474	78.3	0.39

greater than 0.01 to be A compartment and less than -0.01 to be B compartment for this analysis and determination of compartment switches.

Saddle Plots

Saddle plots were constructed to investigate changes in interaction frequency between or within compartments in highly migratory and non-migratory melanoma cells. Briefly, PCA analysis was performed on 100Kb binned matrices to assign compartment identity for each bin. For the analysis, interaction zScores, normalizing for interaction decay for genomic distance, were calculated according to the matrix2compartment.pl script in cworld-dekker. Zscore matrices were reordered based on compartment strength (from strongest B to strongest A using eigenvector values). Finally, matrices were smoothed at 500Kb bin size.

Interchromosomal Fraction (ICF)

Ratio of interchromosomal interactions relative to the total number of interactions for a given chromosome as previously described (Heinz et al., 2018).

$$ICF = \frac{Interchromosomal\ Interactions}{Inter + Intra\ chromosomal\ Interactions}$$

$$\text{deltaICF} = ICF_{ST} - ICF_{NST}$$

Distance Decay Scaling Plots

Using 40 and 20 Kb binned iteratively corrected and scaled contact matrices, we extracted contact frequencies between bins at each genomic distance, excluding the

diagonal bin (zero distance). A loess fit was then used to find a smooth curve describing interaction decay vs. distance (using the matrix2loess.pl script in cworld-dekker). The interaction frequencies were then normalized to set the maximum value (loess fit interaction value for the minimum distance) for each dataset to 1 and then plotted on a log scale vs. log genomic distance.

References

ABBAS, A., HE, X., NIU, J., ZHOU, B., ZHU, G., MA, T., SONG, J., GAO, J., ZHANG, M. Q. & ZENG, J. 2019. Integrating Hi-C and FISH data for modeling of the 3D organization of chromosomes. *Nat Commun,* 10**,** 2049.

ALAM, S. G., LOVETT, D., KIM, D. I., ROUX, K. J., DICKINSON, R. B. & LELE, T. P. 2015. The nucleus is an intracellular propagator of tensile forces in NIH 3T3 fibroblasts. *J Cell Sci,* 128**,** 1901-11.

BICKMORE, W. A. & VAN STEENSEL, B. 2013. Genome architecture: domain organization of interphase chromosomes. *Cell,* 152**,** 1270-84.

BISSELL, M. J., HALL, H. G. & PARRY, G. 1982. How does the extracellular matrix direct gene expression? *J Theor Biol,* 99**,** 31-68.

CRANE, E., BIAN, Q., MCCORD, R. P., LAJOIE, B. R., WHEELER, B. S., RALSTON, E. J., UZAWA, S., DEKKER, J. & MEYER, B. J. 2015. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature,* 523**,** 240-4.

DAHL, K. N., KAHN, S. M., WILSON, K. L. & DISCHER, D. E. 2004. The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. *J Cell Sci,* 117**,** 4779-86.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

GASPAR-MAIA, A., ALAJEM, A., MESHORER, E. & RAMALHO-SANTOS, M. 2011. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol,* 12**,** 36-47.

GOLLOSHI, R., SANDERS, J. T. & MCCORD, R. P. 2018. Iteratively improving Hi-C experiments one step at a time. *Methods,* 142**,** 47-58.

GUILLUY, C., OSBORNE, L. D., VAN LANDEGHEM, L., SHAREK, L., SUPERFINE, R., GARCIA-MATA, R. & BURRIDGE, K. 2014. Isolated nuclei adapt to force and reveal a mechanotransduction pathway in the nucleus. *Nat Cell Biol,* 16**,** 376-81.

HEINZ, S., TEXARI, L., HAYES, M. G. B., URBANOWSKI, M., CHANG, M. W., GIVARKES, N., RIALDI, A., WHITE, K. M., ALBRECHT, R. A., PACHE, L., MARAZZI, I., GARCIA-SASTRE, A., SHAW, M. L. & BENNER, C. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell,* 174**,** 1522-1536 e22.

IMAKAEV, M., FUDENBERG, G., MCCORD, R. P., NAUMOVA, N., GOLOBORODKO, A., LAJOIE, B. R., DEKKER, J. & MIRNY, L. A. 2012. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods,* 9**,** 999-1003.

IRIANTO, J., XIA, Y., PFEIFER, C. R., GREENBERG, R. A. & DISCHER, D. E. 2017. As a Nucleus Enters a Small Pore, Chromatin Stretches and Maintains Integrity, Even with DNA Breaks. *Biophys J,* 112**,** 446-449.

ISERMANN, P., DAVIDSON, P. M., SLIZ, J. D. & LAMMERDING, J. 2012. Assays to measure nuclear mechanics in interphase cells. *Curr Protoc Cell Biol,* Chapter 22**,** Unit22 16.

JACOBSON, E. C., PERRY, J. K., LONG, D. S., OLINS, A. L., OLINS, D. E., WRIGHT, B. E., VICKERS, M. H. & O'SULLIVAN, J. M. 2018. Migration through a small pore disrupts inactive chromatin organisation in neutrophil-like cells. *bioRxiv*.

KECHAGIA, J. Z., IVASKA, J. & ROCA-CUSACHS, P. 2019. Integrins as biomechanical sensors of the microenvironment. *Nat Rev Mol Cell Biol,* 20**,** 457-473.

LE, H. Q., GHATAK, S., YEUNG, C. Y., TELLKAMP, F., GUNSCHMANN, C., DIETERICH, C., YEROSLAVIZ, A., HABERMANN, B., POMBO, A., NIESSEN, C. M. & WICKSTROM, S. A. 2016. Mechanical regulation of transcription controls Polycomb-mediated gene silencing during lineage commitment. *Nat Cell Biol,* 18**,** 864-75.

MURTHY, S. E., DUBIN, A. E. & PATAPOUTIAN, A. 2017. Piezos thrive under pressure: mechanically activated ion channels in health and disease. *Nat Rev Mol Cell Biol,* 18**,** 771-783.

POMBO, A. & DILLON, N. 2015. Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol,* 16**,** 245-57.

REIDY, M. A. & BOWYER, D. E. 1977. Scanning electron microscopy of arteries. The morphology of aortic endothelium in haemodynamically stressed areas associated with branches. *Atherosclerosis,* 26**,** 181-94.

ROWAT, A. C., JAALOUK, D. E., ZWERGER, M., UNG, W. L., EYDELNANT, I. A., OLINS, D. E., OLINS, A. L., HERRMANN, H., WEITZ, D. A. & LAMMERDING, J. 2013. Nuclear envelope composition determines the ability of neutrophil-type cells to passage through micron-scale constrictions. *J Biol Chem,* 288**,** 8610-8.

ROWLEY, M. J. & CORCES, V. G. 2018. Organizational principles of 3D genome architecture. *Nat Rev Genet,* 19**,** 789-800.

RYSA, J., TOKOLA, H. & RUSKOAHO, H. 2018. Mechanical stretch induced transcriptomic profiles in cardiac myocytes. *Sci Rep,* 8**,** 4733.

SANDERS, J. T., FREEMAN, T. F., XU, Y., GOLLOSHI, R., STALLARD, M. A., MARTIN, R. S., BALAJEE, A. S. & MCCORD, R. P. 2019. Radiation-Induced DNA Damage and Repair Effects on 3D Genome Organization. *bioRxiv***,** 740704.

SIMON, D. N. & WILSON, K. L. 2011. The nucleoskeleton as a genome-associated dynamic 'network of networks'. *Nat Rev Mol Cell Biol,* 12**,** 695-708.

STEPHENS, A. D., BANIGAN, E. J., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2017. Chromatin and lamin A determine two different mechanical response regimes of the cell nucleus. *Mol Biol Cell,* 28**,** 1984-1996.

STEPHENS, A. D., LIU, P. Z., BANIGAN, E. J., ALMASSALHA, L. M., BACKMAN, V., ADAM, S. A., GOLDMAN, R. D. & MARKO, J. F. 2018. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol Biol Cell,* 29**,** 220-233.

SUN, Z., GUO, S. S. & FASSLER, R. 2016. Integrin-mediated mechanotransduction. *J Cell Biol,* 215**,** 445-456.

TAJIK, A., ZHANG, Y., WEI, F., SUN, J., JIA, Q., ZHOU, W., SINGH, R., KHANNA, N., BELMONT, A. S. & WANG, N. 2016. Transcription upregulation via force-induced direct stretching of chromatin. *Nat Mater,* 15**,** 1287-1296.

TUCKER, J. B. & MEATS, M. 1976. Microtubules and control of insect egg shape. *J Cell Biol,* 71**,** 207-17.

VAN STEENSEL, B. & BELMONT, A. S. 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell,* 169**,** 780-791.

VON WICHERT, G., JIANG, G., KOSTIC, A., DE VOS, K., SAP, J. & SHEETZ, M. P. 2003. RPTP-alpha acts as a transducer of mechanical force on alphav/beta3-integrin-cytoskeleton linkages. *J Cell Biol,* 161**,** 143-53.

WANG, N., TYTELL, J. D. & INGBER, D. E. 2009. Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. *Nat Rev Mol Cell Biol,* 10**,** 75-82.

YU, H. S., KIM, J. J., KIM, H. W., LEWIS, M. P. & WALL, I. 2016. Impact of mechanical stretch on the cell behaviors of bone and surrounding tissues. *J Tissue Eng,* 7**,** 2041731415618342.

ZWERGER, M., HO, C. Y. & LAMMERDING, J. 2011. Nuclear mechanics in disease. *Annu Rev Biomed Eng,* 13**,** 397-428.

CONCLUSION

The work presented in this dissertation outlines the role of the 3D genome organization in response to mechanical perturbations of the nucleus. While recent studies have reported perturbations of the nuclear shape, chromatin and ultimately changes in expression of mechanosensitive genes, it is not yet clear how the interplay between these components modulate the response to different environments.

Across several of our mechanical perturbation experiments, we observe that the majority of changes affect compartments. We find that when melanoma cells migrate through numerous rounds of constrictions (Chapter 2), they display global changes in their 3D genome structure which is associated with differences in cell and nuclear morphology. Compartment identity switches are the major structural rearrangements that we observe in highly migratory melanoma cells. These compartment switches upon constricted migration are associated with changes in expression of genes involved in the metastatic cascade. Similar, but more subtle compartment identity switches are observed in fibroblast cells that have undergone mechanical stretching. Perhaps, this means that the more subtle changes are the ones that are likely to be induced by the constriction itself while more dramatic differences may be selected, as suggested by ITGB4 expression. Regions of the genome that switch compartment identity after stretching, contain genes that are important in mechanotransduction and therefore could modulate transcriptional signatures. The genomic bins affected by both constricted migration and mechanical stretching are enriched in cLADs, which represent associations with the nuclear lamina. This observation could be related to the redistribution of the heterochromatin to the

nuclear periphery in cells that have undergone constricted migration but could also correlate with previous reports of heterochromatin increase as a protective mechanism to mechanical stressors. Moreover, in both constricted migration and cells that have undergone mechanical stretching, we observe a loss of interactions between the strongest B compartment regions. This could be indicative of stretching of the chromatin fiber and therefore subsequent loss of contacts as the DNA has to fit through tiny constrictions. In addition, many genomic regions with altered compartmentalization in sequentially constricted cells do not show a connection to gene regulation. In fact, we observe a striking enrichment of regions with no genes at all among the set of regions that switch from the A to the B compartment in confined migration. This suggests that the genome structure differences we observe could have a role beyond gene regulation and could play a role in modulating the physical properties of the cancer cell nucleus. Indeed, we observe that the nuclei of Bottom-5 cells deform more than Top-5 cells when embedded in a collagen matrix or even when spontaneously migrating on a 2D surface.

Specific to cells that have passed through constrictions, we observe large genomic rearrangements in the form of a translocation between chr5 and chr13 (t(5;13)) which is not evident in cells that have not undergone constricted migration. Such translocation could be induced by the large nuclear deformation required to undergo constricted migration and the increase in DNA damage. However, the possibility that we are selecting cells that harbor specific phenotypes from the original population is not ruled out.

We were able to identify a subset of cells from the Control population that exhibits high expression of ITGB4 gene. These cells display high migratory phenotypes when challenged to migrate through constricted spaces. We investigated their 3D genome organization and found that ITGB4(+) cells exhibit other translocations that are not present in Control cells. In addition, they also display similarities in compartment identity switches when compared to Bottom-5 cells. However, Bottom-5 cells exhibit specific changes in compartment identity switches with genes involved in calcium ion transport suggestive of the role of mechanical perturbations during constricted migration and modulating the phenotype of these cells.

Interestingly, when we induce global increase in acetylation by treatment with TSA, no global structural rearrangements are observed. Instead, more local changes in the chromatin fiber and especially in A compartment are evident. The overall decondensation of the chromatin fiber correlated with an increase in accessibility and disruption of chromatin loops. Disruption of loops could be in part due to the pushing forces exerted on the chromatin upon decondensation. However, the importance of this global decondensation in the efficacy of the drug is not clear. In addition to the TSA project, we also observe more local 3D structure changes in cells that undergo mechanical stretching. We observe an increase in TAD boundary strength and a decrease in interaction frequency between TADs. In contrast, we observe no changes in TAD boundary strength in constricted migration and a decrease in TAD strength with TSA.

Taken together, our data provide insight about basic 3D genome organization principles and how they are altered as the mechanical properties of the nucleus changes. Our observation that compartments, compartment strength, and whole chromosome interactions change while TADs are less altered reinforces the idea that these are separate and largely independent layers of genome organization. These changes also suggest that the compartment and whole chromosome level of genome structure may be more important to nuclear mechanics than the TAD and loop level of organization.

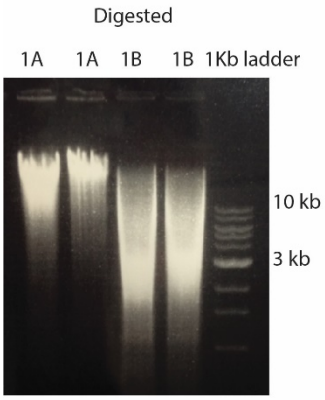
APPENDIX

Appendix A. Digestion Efficiency and Resulting Hi-C Library Statistics

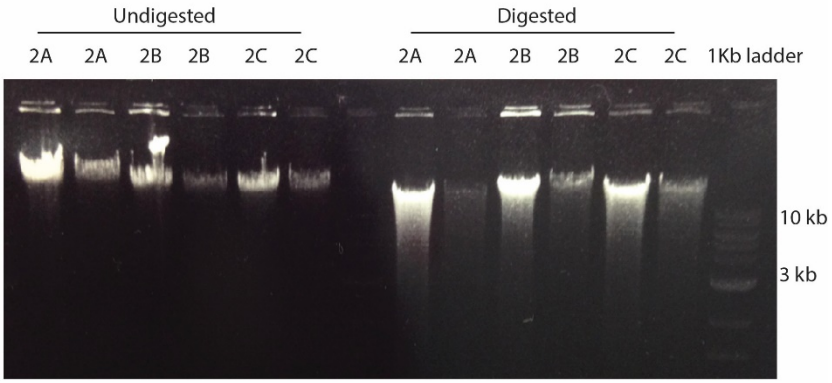
The table below summarizes results from three different problematic Hi-C experiments in which inter-sample variation existed in the extent of HindIII digestion visible on a 0.8% agarose gel (actual gel images shown below the table) immediately after the digestion step of the Hi-C protocol. In each case, the samples within each experiment that had less digestion / more intact DNA resulted in higher percentages of uninformative dangling end molecules in the final Hi-C library. In most cases, more evident digestion was related to a much higher fraction of the molecules ligating to others in cis rather than in trans. The high proportions of trans interactions relate to higher levels of overall background noise. Note that Experiment 3A (starred) corresponds to the low complexity replicate pictured in Figure 6A.

	Digestion gel appearance	Dangling end %	% Cis
Expt 1A	Still quite intact, high MW	69	18
Expt 1B	clearly digested	45	48
Expt 2A	More clearly digested	33	60
Expt 2B	less digested	43	33
Expt 2C	More clearly digested	22	61
Expt 3A*	Still quite intact, high MW	56	64
Expt 3B	More clearly digested	29	64

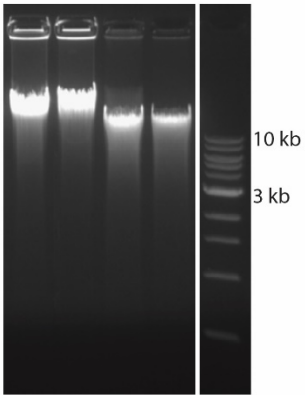
Experiment 1



Experiment 2



Experiment 3



Appendix B. Detailed protocols for Hi-C 1.0 and Hi-C 2.0 variations tested with HindIII digestion.

NEBNext library preparation adaptations detailed (page 10) as well.

This protocol is modeled on the Hi-C 2.0 protocol published by Belaghzal and Gibcus, 2017. The complete Hi-C protocol we used is included, with variations we tested for Hi-C 1.0 and Hi-C 2.0 noted. The 5 million cell aliquot used for the DpnII Hi-C 2.0 was processed according to the protocol below through step 2.3 and then exactly as described in Belaghzal and Gibcus, 2017.

Day 1. CROSSLINKING OF CELLS, CELL LYSIS, AND CHROMATIN DIGESTION

Grow the cells in appropriate culture medium. 5 million cells per Hi-C library are used. The degree of cell confluence, morphology, etc. should be documented carefully before each Hi-C experiment for future reference.

1. **CROSSLINKING ADHERENT CELLS (~90 min):**

Volumes below are for a single T-75 Plate

Prepare 10 mL 1x HBSS + protease inhibitor cocktail per plate, place on ice

1.1) Aspirate the medium, wash with 10 ml 1x HBSS (no serum) per plate.

1.2) Immediately before adding to the plate, mix 275 µl of 37% formaldehyde with 10 mL 1x HBSS to obtain 1% final formaldehyde concentration. Aspirate 1x HBSS wash from cell plate and add this pre-mixed 1% formaldehyde solution.

1.3) Incubate at room temperature (RT) for exactly 10 min on a gently shaking platform.

1.4) To quench the crosslinking reaction, add 554 µL of 2.5 M glycine, mix well

1.5) Incubate for 5 min at RT (rocking platform) and then incubate on ice for at least 15 min to stop crosslinking completely.

1.6) Aspirate liquid off of plates (discard as hazardous waste) and add 10 mL ice cold 1x HBSS + protease inhibitors

1.7) Scrape the cells from the plates with a cell scraper and transfer to a 50 ml tube.

1.8) Centrifuge aliquots of 5 million cells each at 800xg for 10 min, 4oC. Check to make sure all cells are pelleted. If there are visible floating cells, spin again.

1.9) Discard the supernatant by aspiration.

1.10) Cells can be snap-frozen in liquid nitrogen and stored at -80°C for at least 1.5 years or one can continue with cell lysis.

2. CELL LYSIS AND CHROMATIN DIGESTION (WITH HINDIII) (~75 min):

2.1) Resuspend one crosslinked cell aliquot (~5x106cells) in 1 ml of ice-cold lysis buffer (see Belaghzal et al., 2017 or Belton et al., 2012 for buffer formulations) containing 10 µl protease inhibitor cocktail (100x, Thermo)

2.2) Incubate on ice for 15 min.

2.3) Dounce homogenize the cells on ice with pestle A.

2.4) Slowly move pestle up and down 30 times

2.5) Incubate on ice 1 min to let the cells cool down

2.6) Then do 30 more strokes.

2.7) Image lysed cells under the microscope to check for intact nuclei and document cell appearance

2.8) Transfer the lysate to a 1.7ml tube

2.9) Centrifuge for 5 minutes at (Hi-C 1.0: 2,000xg /Hi-C 2.0: 2,500xg) at RT to pellet nuclei.

2.10) Discard the supernatant and then wash the pellet twice by resuspending it in 500 µl of ice cold 1x CutSmart Buffer and then centrifuging the sample for 5 min at (Hi-C 1.0: 2,000xg /Hi-C 2.0: 2,500xg).

2.11) Resuspend the pellet in 1x CutSmart Buffer, so that the total volume of the suspension is 360 uL. (Add 100 uL of buffer first, check total volume, and then add remaining amount necessary to get 360 uL total)

Save 18 µl of lysate as a chromatin integrity control:

Add 50 µl of 1x CutSmart Buffer and 10 µl of Proteinase K (10 mg/ml). Incubate for 30 minutes at 65°C. Purify DNA by single phenol-chloroform extraction without ethanol precipitation. Add 1µl of RNAseA (1 mg/ml) to the aqueous phase and incubate for 15 min at 37°C. Store at -20 oC overnight and run on a 0.75% agarose gel at the same time as the digested DNA sample (see Day 2 below). The sample is good if DNA is either stuck in the well or runs as a single high molecular weight band (>23 kb). After taking 18µl lysate control we have 342 µl remaining total volume.

2.12) Add 38 µl of 1% SDS to the Hi-C tube (380 µl total). Mix carefully by pipetting up and down. Avoid making bubbles (0.1% SDS final).

2.13) Incubate at 65°C for 10 minutes exactly to open chromatin

2.14) Place tubes on ice immediately after

2.15) Add 43 µl of 10% Triton X-100 to the Hi-C-tube (423 µl total) to quench the SDS (1% Triton final). Mix gently by pipetting up and down Avoid making bubbles.

2.16) Add 12 µl of 10x CutSmart Buffer to the Hi-C tube to compensate for added components (435 µl total before restriction enzyme)

2.17) Add 400U (4 µl of 100,000 units/ml) HF HindIII to the Hi-C tube (439 µl total) Mix gently.

2.18) Digest the chromatin overnight at 37°C on a Nutator.

Day 2. BIOTINYLATION, LIGATION, and CROSSLINK REVERSAL

DNA INTEGRITY / DIGESTION CHECK

Incubate tube at 65˚C for 20 mins in order to deactivate the endonuclease enzyme.

After this incubation put the tubes on ice until cooled to room temperature

This inactivation depends on the enzyme used. Please check enzyme datasheet.

Save 10 µl of lysate as a digestion control (*429 uL left; replace with 10 uL 1x CutSmart Buffer*)

Add 50 µl of 1x CutSmart Buffer and 10 µl of Proteinase K (10 mg/ml). Incubate for 30 minutes at 65°C. Purify DNA by single phenol-chloroform extraction without ethanol precipitation. Add 1µl of RNAseA (1 mg/ml) to the aqueous phase and incubate for 15 min at 37°C.

Check undigested and digested aliquots by running on a 0.75% gel, 75-100 V for 1 h.

(*while gel is running, proceed with preparing biotinylation mastermix with everything but expensive Klenow and biotin reagents. If gel results look good, immediately proceed with biotinylation*).

3. BIOTIN FILL-IN

3.1) Prepare a Biotin Fill-in MasterMix as follows:

FILL-IN MIX (ADD IN THIS ORDER)	PER TUBE
milliQ water	2 uL
10x CutSmart Buffer	6 µl
10 mM dATP	1.5 µl
10 mM dGTP	1.5 µl
10 mM dTTP	1.5 µl
0.4 mM biotin-14-dCTP	37.5 µl
5U/ µl Klenow	10.0 µl
TOTAL	60 µL

3.2) Add 60 µl of the Fill-in master mix to the digested Hi-C chromatin *(should be 500 µl total; if some volume decrease occurs during digestion step above, bring total volume up to 500 uL with 1x CutSmart Buffer).*

3.3) Mix gently by pipetting up and down without producing any bubbles

3.4) Incubate the tubes at (Hi-C 1.0: 37°C for 2 hours / Hi-C 2.0: 23°C for 4 hours) in a ThermoMixer (900 RPM mixing; 10 sec every 5 min)

3.5) Prepare ligation mixture (below) during incubation.

3.6) Place on ice immediately after incubation.

4. BLUNT END DNA LIGATION

4.1) Prepare the 665 µl ligation mix during biotinylation:

LIGATION MIX	PER TUBE
Hi-C 1.0: 10x ligation buffer [NEB]	120 µl 240 µl

Hi-C 2.0: 5x ligation buffer [Invitrogen]	
10% Triton X-100	120 µl
10 mg/ml BSA	12 µl
T4 DNA ligase [Invitrogen]	Hi-C 1.0: 10 ul (5 U/uL) Hi-C 2.0: 50 µl (1 U/uL)
Water	To 700 µl
TOTAL	700 µL

Note: we have tried the Hi-C 2.0 protocol with both a total of 50 U and a total of 250 U of ligase and find that additional ligase appears neither to help nor hurt the outcome.

4.2) Add 700 µl of the mix to the Hi-C tube (*1,200 µl total*).

4.3) Incubate all tubes at 16°C for 4 hours in ThermoMixer with interval shaking (900 RPM mixing, 10 sec every 30 min).

5. REVERSE CROSSLINKING

5.1) Add 25 µl of 20 mg/mL proteinase K to the ligated Hi-C tube and incubate at 65°C overnight in the ThermoMixer.

5.2) Add another 25 µl of proteinase K to the Hi-C tube and incubating an additional 2 hours (*1,250 µl total Hi-C).*

Day 3. DNA PURIFICATION

6. DNA PURIFICATION

6.1) Remove tubes from 65 C and allow to cool down.

6.2) Add 2.4 ml saturated phenol:chloroform (1:1, pH=8.0) to 2 x 15 ml tubes (label Hi-C PC1 and Hi-C PC2) for Hi-C

Use fume hood, appropriate gloves, and labcoat when handling phenol:chloroform

6.3) Extract the DNA:

6.4) Transfer Hi-C DNA into Hi-C PC1 tube

6.5) Vortex for 30 seconds to obtain a homogenous, milky solution

6.6) Centrifuge at 1500xg for 5 minutes (room temperature)

6.7) Carefully transfer the aqueous phase to Hi-C PC2. Repeat vortex and spin.

6.8) Transfer the aqueous phase with Hi-C DNA (*~1.2 ml*) to a clean 15 ml tube (adequate for high-speed centrifugation at 16,000 xg)

6.9) Add 1x TLE (pH=8.0) up to 2 ml

6.10) Add 1/10 volume of 3M sodium acetate, pH=5.2 (*200µl*) and vortex briefly.

6.11) Add 2.5 volumes of ice-cold 100% ethanol (*5 ml*) and mix well by inverting the tubes several times.

6.12) Incubate the tubes at -80˚C for 45 min – 1 h

6.13) Centrifuge tubes at 16,000x*g* for 30 min at 4°C.

6.14) Decant the supernatant with extra caution as the pellet detaches from the tube wall quite easily and air dry the DNA pellets very briefly (~1 min).

6.15) The pellet might become invisible while drying out

6.16) Dissolve Hi-C pellet in 450µl of 1x TLE (pH=8.0) and transfer to 2 mL, 30kD Amicon columns.

6.17) Wash the column at least 3 times with the following steps:

6.18) Centrifuge at 14,000xg in tabletop centrifuge for ~5 minutes (room temperature)

6.19) Remove flowthrough (to a temporary save tube). DNA stays in the volume left in the column.

6.20) Add 450 µl TLE

6.21) On the last spin, spin for 10 minutes instead of 5 minutes at 14,000xg

6.22) After the last wash, flip the column onto a new Amicon container tube and spin at 1,000xg for 2 min to recover the DNA from the column

6.23) Add 1 µl of RNAseA (1 mg/ml) to the sample and incubate for 30 mins at 37°C.

DNA samples can now be stored at -20°C.

7. QUALITY CONTROL OF HI-C LIBRARIES

7.1) Make a 1:10 dilution of the Hi-C library and run 2µl and 6 µl on a 0.8% agarose gel for quality control at 100V for 1 hr.

7.2) Quantify amount of DNA by densitometry. Use several dilutions of the NEB 1kb DNA ladder as a standard curve to estimate the DNA concentration more accurately.

7.3) Measure the Biotin Incorporation by a PCR digest:

7.4) Make 1:20 dilution of Hi-C DNA (2.5 uL Hi-C + 47.5 uL H2O)

7.5) Dilute primers to 10 uM in TLE

7.6) Perform a PCR to amplify one neighboring interaction as follows:

INGREDIENT	1 REACTION
Diluted Template DNA	20 uL
MILLIQ	33.9 uL
10X PCR buffer	7.5 µl
25mM dNTP	0.6 µl
50 mM MgSO4	6.0 uL
Primer 1 (10 µM)	3 µl
Primer 2 (10 µM)	3 µl
Taq DNA polymerase	1 µl
Total	75 uL

7.7) Combine 20 uL diluted Hi-C DNA or water (neg ctrl) with 55 uL mastermix in a PCR tube.

7.8) Run the following PCR program:

Step	Temperature	Time
1	95°C	5 minutes
2	95°C	30 seconds
3	65°C	30 seconds
4	72°C	30 seconds
5	Go to 2	34 times
6	95°C	30 seconds
7	65°C	30 seconds
8	72°C	8 minutes

7.9) Run the samples in a 2% gel at 160 V for 45 min.

7.10) Digest the HiC PCR product with HindIII and NheI as follows:

	Control (µl)	HindIII (µl)	NheI (µl)	HindIII/ NheI (µl)
PCR product	16.0	16.0	16.0	16.0
10X Cutsmart buffer	2.0	2.0	2.0	2.0
HindIII	---	1	---	1
NheI	---	---	1	1
MilliQ	2	1	1	---

7.11) Incubate samples at 37°C for at least 30 minutes:

7.12) Combine each with 4 uL 6x dye and load 17 uL on the top row of wells in the 2% gel. Run 1 hr at 175 V.

Successful biotin fill-in of HindIII sites (AAGCTT) followed by blunt-end ligation create sites for the NheI restriction enzyme (GCTAGC).

Successful biotin fill-in of DpnII sites (GATC) followed by blunt-end ligation create sites for ClaI restriction enzyme (ATCGAT).

8. REMOVAL OF BIOTIN FROM UN-LIGATED ENDS

8.1) Quantify DNA from Hi-C ligated library gel (step 7.2). Per 5 ug of DNA recovered, prepare reactions as follows:

INGREDIENT	Per 5 ug
Hi-C DNA sample	**5 µg**
10X NEBuffer 2.1	5 µl
1 mM dATP	1.25 µl
1 mM dGTP	1.25 µl
3,000 U/ml T4 DNA polymerase	5 µl
MILLIQ water	TO 50 µL

8.2) Place each 50 µl reaction in a different well of a PCR strip

8.3) Incubate at 20°C for 4 hours in the PCR machine.

8.4) (Hi-C 1.0: Add 2 µl of 0.5M EDTA to each tube to stop the reaction; Hi-C 2.0: Inactivate the enzyme for 20 mins at 75°C)

8.5) Cool down/ keep at 4°C

8.6) Pool the reactions together and add water up to 500 uL. Perform an Amicon wash:

8.7) Add volume to Amicon column and centrifuge at 14,000xg for ~ 5 min

8.8) Wash twice with 400 µl of milliQ water (spinning 5 min at 14k x g each time)

8.9) Invert the Amicon column onto a fresh tube and centrifuge for 2 minutes at 1,000xg to collect the DNA

Note that Amicon concentration does not remove proteins (they are only inactivated by the heat or EDTA treatment above, and then also by sonication below). Concentration by volume reduction may result in DNA coming out of solution. If this occurs, sonication below will help re-solubilize DNA.

8.10) Add milliQ to a total of 132 µl. Save 2 uL for pre-sonicated gel check and transfer the rest to a Covaris microtube for sonication.

9. DNA SONICATION

9.1) Shear the DNA to a size of 200 – 400 bp using the Covaris M220 sonicator and the following parameters:

20% Duty Factor, 200 cycles/burst, 50 W Peak Incident Power for 110 seconds at 4 degrees C.

(*note that this differs from the Hi-C 2.0 recommended settings for reasons discussed in the manuscript*)

9.2) Check the results of sonication on a 2% gel. Run the gel at 160 V for 45 min.

10. SIZE FRACTIONATION USING AMPURE XP

This step can be omitted for low concentrations of DNA and sonicated DNA can be taken directly to step 11.

The liquid phase of AMpure mixture precipitates DNA onto the AMpure XP magnetic beads. The size of the DNA molecules precipitated depends on the ratio of the volume of AMpure XP liquid (PEG8000) to the volume of the sample when mixed together. Increasing the proportion of AMpure XP liquid decreases the cutoff of the size of the DNA molecules precipitated onto the beads. Here, the 0.7x Ampure step removes fragments above ~400 bp and the 1.2x Ampure captures the remaining fragments between 100-400 bp.

Note that the 0.7x and 1.2x recommendation differs from Hi-C 2.0 in order to capture the larger fragments coming out of sonication conditions above.

10.1) Bring volume of Hi-C sample up to 500ul with 1x TLE

10.2) Allow AMpure XP mixture to come to RT and vortex prior to use. Add 350 µl of AMpure XP mixture to the Hi-C tube. Label the tube as 0.7X *(the ratio of AMpure XP solution volume to Hi-C sample volume is 350/500 = 0.7).*

10.3) Vortex and spin down tubes briefly.

10.4) Incubate tubes for 10 min at RT.

10.5) Place tubes on the Magnetic Particle Separator (MPS) for 5 min at RT.

10.6) During above incubation, prepare a fresh tube with 500 µl AMpure XP mixture; label them as 1.2X.

10.7) Incubate tubes on the MPS for 5 min.

10.8) Remove supernatant

10.9) Resuspend beads in 250 µl AMpure mixture

This step increases the number of beads present in a smaller volume of AMpure XP mixture. This prevents saturation of the beads with DNA, ensuring an efficient precipitation of the DNA.

10.10) Collect the 0.7X supernatants from MPS

10.11) Add the supernatants to the 1.2X tubes.

The ratio of AMpure XP solution volume to the original sample volume is now: (250 + 350)/500 = 1.2. Under these conditions beads bind DNA fragments >100 bp.

10.12) Vortex and spin down tubes briefly.

10.13) Incubate 1.2X tubes for 10 min at RT

10.14) Place 1.2X tubes on the MPS for 5 min at RT

10.15) Discard supernatant from the 1.2X tubes.

10.16) Wash the beads in 0.7X and 1.2X tubes twice with 1 ml fresh 70% ethanol, reclaiming beads against the MPS for 5 min each time.

10.17) Air dry beads on the MPS briefly (too much drying may decrease elution efficiency)

10.18) Resuspend the 0.7X and 1.2X beads in 150 µl of 1x TLE buffer in each tube to elute the DNA

10.19) Incubate 10 min at RT

10.20) Separate AMpure beads from eluate for both 0.7X and 1.2X tubes on the MPS for 5 min

10.21) Keep the eluate

10.22) For the gel, combine:

7 uL 6x dye + 10 uL 1.2x eluate + 10 uL water

2 uL 6x dye + 4.5 uL 0.7x eluate + 4.5 uL water

10.23) Desalt and concentrate the Hi-C sample on Amicon:

10.24) Centrifuge at 14,000xg for ~5 min.

10.25) Add 450 µl of 1x TLE buffer

10.26) Centrifuge again until at 15,000xg for about 5~10min until ~30 µl remains

10.27) Place the Amicon unit upside down onto a new tube provided with the Amicon kit. Collect Hi-C sample by spinning at 1,000xg for 2 min

10.28) Bring volume of the sample to 50 µl with milliQ water

10.29) Check the quality and estimate the quantity of DNA on a 2% agarose gel. Run one lane of the 0.7X sample and a dilution series of the 1.2X sample along with a low molecular weight DNA marker. Quantify the amount of DNA in the 1.2X sample and compare to sonication gel to evaluate degree of DNA loss due to Ampure.

NEBNext Protocol Adapted for Hi-C library Preparation on Beads

At this step, the Hi-C 2.0 protocol recommends end repair and A tailing, then bead pulldown, followed by Illumina adapter ligation. We have adapted the entire NEBNext protocol to work after bead pulldown.

11. BIOTIN PULLDOWN WITH STREPTAVIDIN COATED BEADS

11.1) Vortex the MyOne™ Streptavidin C1 beads and then transfer 20 µl of beads to a 1.7ml LoBind tube.

11.2) Wash the beads with 400 µl of Tween wash buffer (TWB) by pipetting up and down and incubating for 3 min at RT on a rocking platform.

11.3) Reclaim beads against the MPS for 1 min, discard the supernatant.

11.4) Resuspend beads in 400 µl of TWB and transfer to a new LoBind tube. Incubate 3 min again.

11.5) Reclaim beads against the MPS for 1 min, discard the supernatant.

11.6) Resuspend beads in 400 µl of 2X Binding Buffer (BB) and add the DNA (in 400 µl TLE buffer)

11.7) Incubate the sample for 15 min at RT with rotation.

11.8) Reclaim the beads against the MPS for 1 min; discard the supernatant to a “BB sup” save tube.

11.9) Resuspend the beads in 400 µl of 1X BB and transfer them to a new tube

11.10) Reclaim the beads against the MPS for 1 min, discard the supernatant.

11.11) Wash beads with 100 µl of TLE and transfer to a new tube

11.12) Reclaim the beads against the MPS for 1 minute; discard the supernatant

11.13) Finally resuspend the beads in 50 uL TLE buffer.

12. NEBNext End Prep

12.1) Add the following components to a sterile nuclease-free tube:

(green) NEBNext Ultra II End Prep Enzyme Mix	3 µl
(green) NEBNext Ultra II End Prep Reaction Buffer	7 µl
Fragmented DNA	50 µl
Total Volume	60 µl

12.2) Set a 100 µl or 200 µl pipette to 50 µl and then gently pipette the entire volume up and down at least 10 times to mix throughly. Perform a quick spin to collect all liquid from the sides of the tube.

Note: It is important to mix well. The presence of a small amount of bubbles will not interfere with performance.

12.3) Place in a thermocycler, with the heated lid set to ≥ 75°C, and run the following program:

30 minutes @ 20°C

30 minutes @ 65°C

Hold at 4°C

If necessary, samples can be stored at -20°C; however, a slight loss in yield (~20%) may be observed. We recommend continuing with adaptor ligation before stopping.

13. Adaptor Ligation

13.1) Perform a 1:15 or 1:20 dilution (or more if a first attempt gives excess adaptors) of the NEBNext Adaptor for Illumina in 10 mM Tris-HCl, pH 8.0 with 10mM NaCl.

13.2) Add the following components directly to the End Prep Reaction Mixture:

End Prep Reaction Mixture (Step 12.3)	60 µl
(red) NEBNext Ultra II Ligation Master Mix*	30 µl
(red) NEBNext Ligation Enhancer	1 µl
(red) NEBNext Adaptor for Illumina **	2.5 µl
Total volume	93.5 µl

* Mix the Ultra II Ligation Master Mix by pipetting up and down several times prior to adding to the reaction.

Note: The Ligation Master Mix and Ligation Enhancer can be mixed ahead of time and is stable for at least 8 hours @ 4°C. We do not recommend adding adaptor to a premix in the Adaptor Ligation Step.

13.3) Set a 100 µl or 200 µl pipette to 80 µl and then pipette the entire volume up and down at least 10 times to mix thoroughly. Perform a quick spin to collect all liquid from the sides of the tube.

(Caution: The NEBNext Ultra II Ligation Master Mix is very viscous. Care should be taken to ensure adequate mixing of the ligation reaction, as incomplete mixing will result in reduced ligation efficiency. The presence of a small amount of bubbles will not interfere with performance).

13.4) Incubate at 20°C for 15 minutes in a thermocycler with the heated lid off.

13.5) Add 3 µl of (red) USER™ Enzyme to the ligation mixture from Step 13.4.

Note: Steps 13.5 and 13.6 are only required for use with NEBNext Adaptors. USER enzyme can be found in the NEBNext Singleplex (NEB #E7350) or Multiplex (NEB #E7335, #E7500, #E7600, #E7710, #E7730 and #E6609) Oligos for Illumina.

13.6) Mix well and incubate at 37°C for 15 minutes with the heated lid set to ≥ 47°C.

Note: Samples can be stored overnight at -20°C.

13.7) Reclaim the beads against the MPS for 1 min. Discard the supernatant.

13.8) Wash the beads twice with 400 µl of TWB. Add buffer to the beads, mix carefully by pipetting and incubate at RT for 5 min with rotation.

13.9) Reclaim the beads against the MPS for 1 min. Discard the supernatant. (to a save tube)

13.10) Resuspend the beads in 200 µl of 1X BB and transfer to a new tube

13.11) Reclaim the beads against the MPS for 1 minute. Discard the supernatant.

13.12) Wash the beads twice with 200 µl TLE. Resuspend beads then transfer to a new tube

13.13) Reclaim the beads against the MPS for 1 min. Discard the supernatant.

13.14) After the last wash, resuspend the beads in 20 µl of TLE and transfer to a new tube.

Note: Samples can be stored at -20°C.

PCR Enrichment of Adaptor-ligated DNA

Note: Check and verify that the concentration of your oligos is 10 µM.

14. PCR TITRATION and PRODUCTION PCR

14.1) Use the following HiCPCR_6CYC and HiCPCR_3CYC programs for PCRs below:

HiCPCR_6CYC: (~15 min)
• Preheat lid at 100°C
— 98°C for 30 seconds
• 6 cycles of:
— 98°C or 10 seconds
— 65°C for 75 seconds

• 65°C for 5 minutes
**Hold at 4°C

HiCPCR_3CYC: (~10 min)
• Preheat lid at 100°C
— 98°C for 30 seconds
• 3 cycles of:
— 98°C or 10 seconds
— 65°C for 75 second

• 65°C for 5 minutes

14.2) Add the following components to a sterile strip tube:

14.3) Forward and Reverse Primer not already combined

Adaptor Ligated DNA Fragments (Step 13.14)	3 µl
(blue) NEBNext Ultra II Q5 Master Mix	12.5 µl
(blue) Index Primer/i7 Primer*,**	2.5 µl
(blue) Universal PCR Primer/i5 Primer*, ***	2.5 µl
Add water to Total Volume of:	25 µl

* The primers are provided in NEBNext Singleplex (NEB #E7350) or Multiplex (NEB #E7335, #E7500, #E7710, #E7730, #E7600) Oligos for Illumina. For use with Dual Index Primers (NEB #E7600), look at the NEB #E7600 manual for valid barcode combinations and tips for setting up PCR reactions.

** For use with NEBNext Multiplex Oligos (NEB #E7335 or #E7500, #E7710, #E7730) use only one index primer per PCR reaction. For use with Dual Index Primers (NEB #E7600) use only one i7 primer per reaction.

*** For use with Dual Index Primers (NEB #E7600) use only one i5 Primer per reaction.

14.4) Set a 100 µl or 200 µl pipette to 40 µl and then pipette the entire volume up and down at least 10 times to mix thoroughly. Perform a quick spin to collect all liquid from the sides of the tube.

14.5) Run the following PCR programs. (Note that this must be run when you have time to take aliquots every ~10 min for 1 hour)

Set up 5 small tubes for gel loading labeled 6, 9, 12, 15, 18 and combine 2 uL 6x dye + 7 uL water in each.

After each PCR below, take out tube in the last 10 seconds of the 65 C final step, put tube directly on ice, and remove a 3 uL aliquot. Combine each aliquot with dye and water in tubes prepared above.

HiCPCR_6CYC (total 6 cycles)

HiCPCR_3CYC (total 9 cycles)

HiCPCR_3CYC (total 12 cycles)

HiCPCR_3CYC (total 15 cycles)

HiCPCR_3CYC (total 18 cycles)

14.6) Run the titration PCR on a 2% gel. Run the gel at 160 V for 45 min

14.7) Build a calibration curve showing the DNA quantity vs. number of cycles.

Choose an optimal number of cycles and number of PCR reactions for the final amplification of the library for deep sequencing. The cycle number should be chosen so that the PCR amplification is in the linear range and the expected size distribution is preserved (over cycling will shift the size distribution of the library towards higher molecular weight products). The number of PCR reactions should be calculated to produce the amount of DNA desired for sequencing (usually 50-100 ng). It is always better to reduce the number of PCR cycles while increasing the number of PCR reactions. The optimal amount is usually 1-2 cycles below the lowest amount visible on the gel.

Note: One can also use the KAPA Library Quantification kit that uses a qPCR system to accurately estimate the concentration of the library and number of cycles for the final amplification of the library for deep sequencing.

15. PRODUCTION PCR

15.1) Decision from titration: Perform _____ PCR reactions (maximum of 5) with _____ cycles.

15.2) Set up N PCR master mix reactions on ice as follows:

Adaptor Ligated DNA Fragments (Step 3.1.14 or 3.2.11)	3 µl
(blue) NEBNext Ultra II Q5 Master Mix	12.5 µl
(blue) Index Primer/i7 Primer*,**	2.5 µl
(blue) Universal PCR Primer/i5 Primer*, ***	2.5 µl
Add water to Total Volume of:	25 µl

15.3) Run the following PCR program with the selected number of cycles.

• Preheat lid at 100°C
— 98°C for 30 seconds
• N cycles of:

— 98°C or 10 seconds
— 65°C for 75 seconds
• 65°C for 5 minutes

15.4) Pool all PCR reactions together in a 1.7 mL Lo-Bind Tube and incubate on MPS for 2 min.

15.5) Remove supernatant to a new tube (this contains amplified DNA)

15.6) Resuspend beads in N*3 uL TLE and freeze at -20 in case of future need.

15.7) Take a 3 uL “pre-purified” aliquot of supernatant + 2 uL dye + 7 uL H2O for gel check.

15.8) To remove primer dimers, purify the amplified Hi-C library from the supernatant using AMpure XP beads.

15.9) Allow AMpure XP mixture to come to RT and vortex prior the use.

15.10) Add 1.5x volumes (_____ uL) of AMpure XP mixture to the supernatant.

Note: if the Bioanalyzer trace at the end of step 16 indicates high amounts of adapter-dimer or primer-dimer, re-purifying with 1.3x or 1.2x Ampure to remove those products is recommended.

15.11) Vortex and spin down briefly.

15.12) Incubate for 10 min at RT.

15.13) Place on the MPS for 5 min at RT. Discard supernatant to a save tube.

15.14) Wash the beads twice with 1 ml of freshly made 70% ethanol

15.15) Air-dry the beads briefly and then resuspend in 35 µl of TLE buffer

15.16) Incubate the beads for 10 min at RT, tapping the tube every 1-2 min.

15.17) Collect the beads with the MPS for 5 min

15.18) Transfer the supernatant, containing the final Hi-C library, to a new tube.

15.19) Check final HiC library on a 2% agarose gel at 160 V for 45 min.

16. QUALITY CONTROL OF HI-C LIBRARY BY RESTRICTION DIGEST

16.1) Digest a small aliquot of the final Hi-C library with NheI to estimate the portion of molecules with valid biotinylated junctions.

INGREDIENT	1X
10x NEB Cutsmart buffer	1 µl
milliQ water	6.5 uL
NheI-HF	1 µl
Hi-C product	1.5 uL
TOTAL	10 µL

16.2) Allow to digest at 37°C for at least 1 hour

16.3) Combine entire volume of digestion reaction + 2 uL dye and 1.5 uL uncut Hi-C DNA + 8.5 uL H2O + 2 uL 6x dye and run on a 2% agarose gel at 160 V for 45 min

16.4) Quantify approximate % of uncut DNA (unshifted smear) from the gel or a Bioanalyzer analysis. This correlates well to the fraction of non-ligation products (dangling ends) that will be obtained in the final sequencing.

Run 1 uL of final Hi-C library on Agilent DNA Bioanalyzer. Dilute an aliquot of the library to 10 nM in 10 uL with Tris-HCl 10 mM, pH 8.5 + 0.1% Tween 20 for later pooling and sequencing.

Buffer formulations for Hi-C

2.5M Glycine

46.92 g Glycine in 250 mL dH_20

Autoclave

Lysis Buffer

10 mM Tris-HCl pH 8.0

10 mM NaCl

0.2% NP-40 (Igepal CA-630)

In MilliQ H_2O

Filter sterilize

Keep at 4°C

10% SDS

50 g SDS in 500 mL dH_20

Filter sterilize

10% Triton X-100

1 mL 100% TX-100 + 9 mL dH_20

10mg/mL BSA

100 mg BSA

Dissolve in 10 mL of H_2O

Aliquot and store at -20^oC

Phenol Chloroform (1:1)

Add Tris buffer to Phenol, allow to stand overnight at 4 degrees

Add 100 mL Phenol to 100 mL Chloroform + some buffer from top of phenol

Allow to stand for several days at 4 degrees

1x TLE

10 mM Tris

0.1 mM EDTA

Add sterile H_20 up to 100 mL

10X PCR Buffer

600 mM Tris ph 8.9 with H_2SO_4,

Adjust to pH 8.9 with H_2SO_4 (make prior to mixing here)

180 mM $(NH_4)_2SO_4$

Tween Wash Buffer

5 mM Tris-HCl pH8.0

0.5 mM EDTA

1 M NaCl

0.05% Tween

In MilliQ H_2O

Filter Sterilize

2x Binding Buffer

10 mM Tris-HCl pH8.0

1 mM EDTA

2 M NaCl

In sterile H_2O

Filter sterilize

Appendix C. Further Data Describing and Interpreting Sequence Copying Artifact

(Related to Figure 5)

The artifactual copying of a piece of the 5' end of a sequencing library molecule onto the 3' end described in Figure 5 is not only observed after Illumina sequencing, but can be detected when a molecule of a Hi-C library is cloned and sequenced. In the example below, the sequence highlighted in blue aligns to the genome as part of the 5' end fragment but was artificially copied as a reverse complement onto the 3' end fragment. Thus, the 3' end fragment sequence does not fully align to the genome. A 4 base pair microhomology between the two ligated fragments lies just before the non-mapping copied sequence on the 3' end, supporting the idea that this copying occurs due to improper annealing and extension from single stranded microhomologies such as this.

Cloned and sanger sequenced Hi-C library molecule showing copying artifact:

5'-

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGAT

CTCTACCTCTATCATTCACT

AGCTATAAA**(ATCT)**TAGGCAGATTATTTCCTCATTTTTAAAATGTAGGAAATGGATAA

TGTACTTAAGGTTCCCAGTTTCATGCGAGGCACATAGTAAGCACTTAATAGAATATC

ATGGTTCTTTGAATCTTCAAGTTCTATCATACTGACTAAAATACAGAGCATGCTAGC

AAGCTAGCTTTAAATACAACCATGATGACAATCAAGATTGACCCTAGGACCAGCAG

CTGAAAGAAATTAGACTGAACTTTAAAGAAATCAGACCGT**(AGAT)**TTTATAGCTAGT

GAATGATAGAGGTAGAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATC

TCGTATGCCGTCTTCTGCTTG**-3'**

red = Illumina adapters

blue = sequence copied from 5’ to 3’ end

green = HindIII junction

parentheses = 4 bp microhomology between the two fragments

Fragment 5’ of junction

```
CTACCTCTAT CATTCACTAG CTATAAAATC TTAGGCAGAT TATTTCCTCA  50
TTTTTAAAAT GTAGGAAATG GATAATGTAC TTAAGGTTCC CAGTTTCATG  100
CGAGGCACAT AGTAAGCACT TAATAGAATA TCATGGTTCT TTGAATCTTC  150
AAGTTCTATC ATACTGACTA AAATACAGAG CATGCTAGCA AGCT
```

Genomic chr1 (reverse strand):

```
caagtgaaac tgaaccacta aaataccctc tgcccagtgt gtctcatgct  213103540
tatggtgaag tgtctatttc acgataacta aaaggacctg aattcaaata  213103490
CTACCTCTAT CATTCACTAG CTATAAAATC TTAGGCAGAT TATTTCCTCA  213103440
TTTTTAAAAT GTAGGAAATG GATAATGTAC TTAAGGTTCC CAGTTTCATG  213103390
CGAGGCACAT AGTAAGCACT TAATAGAATA TCATGGTTCT TTGAATCTTC  213103340
AAGTTCTATC ATACTGACTA AAATACAGAG CATGCTAGCA AGCTtttatt  213103290
gactgtgtga gtgaataaca aaaactttta aaagccatgt caaaataaat  213103240
aaataaatgt cttacaatga aacattatca cagataatct tttt
```

Fragment 3’ of junction
(note unmapped part is reverse complement
of beginning of 5’ fragment)

```
AGCTTTAAAT ACAACCATGA TGACAATCAA GATTGACCCT AGGACCAGCA  50
GCTGAAAGAA ATTAGACTGA ACTTTAAAGA AATCAGACCG TAGATtttat  100
agctagtgaa tgatagaggt ag
```

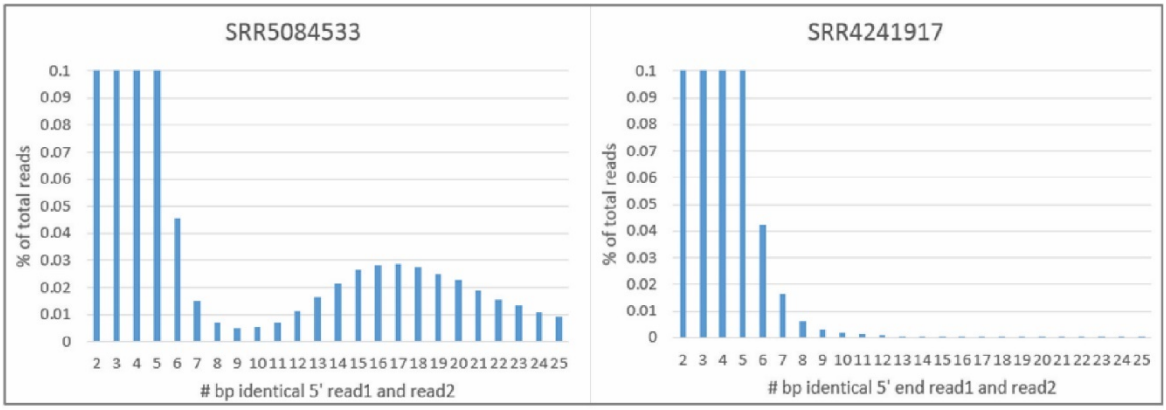
Genomic chr18 :

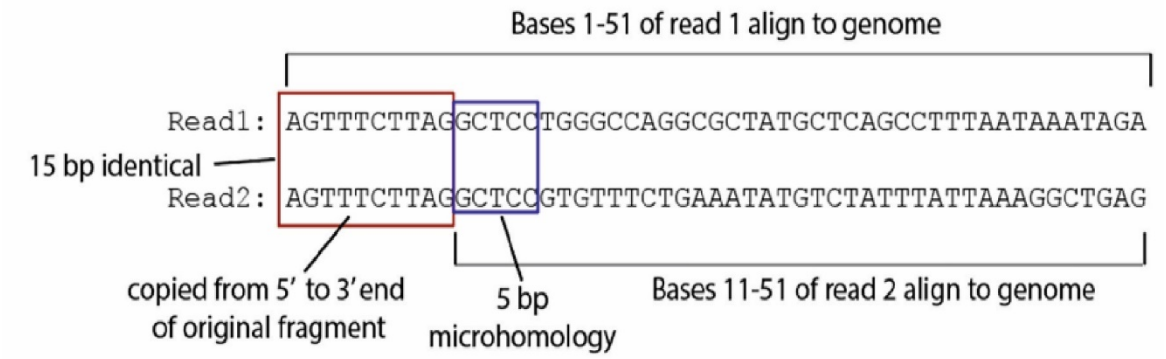
```
agaagcacct tccagaaata catatcagca aaaacaaatc acaaaggaaa  25703321
gacatcagat ttagttccga accacaccca gcacatgaat ggcacccaaa  25703371
AGCTTTAAAT ACAACCATGA TGACAATCAA GATTGACCCT AGGACCAGCA  25703421
GCTGAAAGAA ATTAGACTGA ACTTTAAAGA AATCAGACCG TAGATaaaaa  25703471
agaatttcaa taaaattatt aaacattatc caccaatact gctgtcagcc  25703521
acccatatgt agtccactga gttatctttc agttatttga tttct
```

Some ChIP-Seq Datasets Exhibit Copying Artifact

If the copying artifact indeed occurs during end repair, as proposed in Figure 5, then this artifact should be expected to occur on occasion in other types of sequencing experiments that employ this method of library preparation. Indeed, below, we show an example of a ChIP-Seq library (left; SRR5084533 from GEO Accession: GSE90992 (Teif et al., 2017)) in which a small, but detectable, proportion of reads show copying of the 5’ end fragment (read1) onto the 3’ end (read2) just as we observed in Hi-C. A sample set of paired reads from SRR5084533 shows that these identical bases follow the same pattern as observed in Hi-C libraries.

As with Hi-C, not all libraries prepared with End Repair show this copying artifact (for example, see SRR4241917, right; from GEO accession GSE86903 (von Meyenn et al., 2016)). Both of these two libraries were prepared with sonication followed by end repair, A-tailing, and adapter ligation with NEBNext reagents. We hypothesize that, as with Hi-C libraries, variations in the extent of DNA single stranded overhangs and library size may influence whether this artifact is seen or not.





Lack of Artifact in Libraries that do not include End Repair and Adapter Ligation

As shown in Figure 5C, a Hi-C library prepared by tagmentation does not exhibit the copying artifact that we describe here. Further evidence that sequencing libraries prepared by methods other than end repair and adapter ligation are free from this copying artifact comes from analysis of a 5C (chromosome conformation capture carbon copy) library (from (Smith et al., 2016)). 5C adapters are added by PCR and so there is no sonication step, biotin removal step, or end-repair step. In the case shown below, the high number of 13 bp matched sequences contain the SOLiD P1 adapter sequence, demonstrating that this is an adapter-related artifact, not a copying artifact.

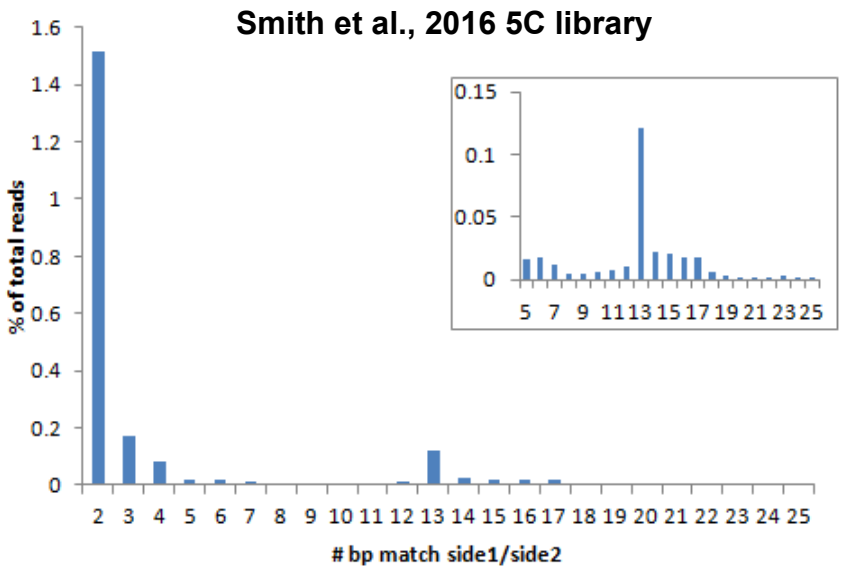
Analysis of Enriched Sequences in Overlap and Microhomologies

When raw FASTQ files are analyzed for sequences that match between read1 and read2 in the paired end set, untrimmed adapter sequence at both ends of the molecule will result in a 12 bp identity, as shown in Figure 5C and expected by the overlap in the TruSeq adapter sequences.

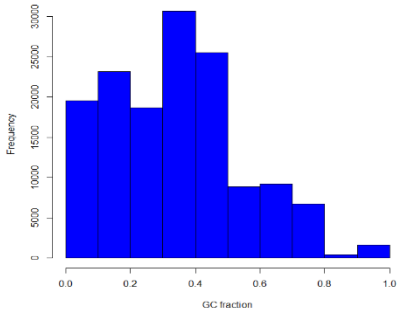
PE Adapter2: GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG

PE Adapter1: GATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT

However, adapter sequence cannot account for the pieces of identical read1/read2 sequences, as described here, that align perfectly on read 1 but not on read2. Thus, we conclude that aberrant adaptor ligation or sequencing does not contribute to the artifact we observe in these cases.



There are no notable enriched sequence motifs among the microhomology seed sequences or copied regions, though we do notice that the short complementary sequences tend to be AT rich, as shown in the histogram at right. We hypothesize that the AT rich DNA is enriched because the A-T base pair is more fragile and likely to be damaged and result in single stranded overhangs after sonication.



Analysis Code Availability:

This copying artifact was analyzed using in house Perl scripts that processed Bowtie2 mapping Samfiles in comparison with original FASTQ files. For others wishing to assess this artifact in their own data, scripts are available upon request.

Appendix B References:

SMITH, E. M., LAJOIE, B. R., JAIN, G. & DEKKER, J. 2016. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. *Am J Hum Genet,* 98**,** 185-201.

TEIF, V. B., MALLM, J. P., SHARMA, T., MARK WELCH, D. B., RIPPE, K., EILS, R., LANGOWSKI, J., OLINS, A. L. & OLINS, D. E. 2017. Nucleosome repositioning during differentiation of a human myeloid leukemia cell line. *Nucleus,* 8**,** 188-204.

VON MEYENN, F., BERRENS, R. V., ANDREWS, S., SANTOS, F., COLLIER, A. J., KRUEGER, F., OSORNO, R., DEAN, W., RUGG-GUNN, P. J. & REIK, W. 2016. Comparative Principles of DNA Methylation Reprogramming during Human and Mouse In Vitro Primordial Germ Cell Specification. *Dev Cell,* 39**,** 104-115.

VITA

Rosela Golloshi was born in Patos, Albania. She attended Turgut Ozal High School in Tirana, Albania and graduated with honors. Upon high school graduation, Rosela decided to move to the United States to continue her college education at University of North Georgia in Dahlonega, Georgia where she received a Bachelor of Science in Chemistry and minored in Biology. At North Georgia, she performed research under the supervision of Dr. Holly Carpenter where she quantitatively compared several extraction methods for an efficient extraction of raspberry ketone from the raspberry fruit. Upon graduation from North Georgia, Rosela joined the team at Aeon Clinical Laboratories where she led and help establish the pharmacogenetics department for pharmacogenetic drug testing. Rosela then chose to attend the University of Tennessee for her doctoral degree under the supervision of Dr. Rachel Patton McCord. In Dr. McCord's laboratory, she studied the role of 3D genome organization in influencing cancer cell constricted migration. She presented her work at multiple conferences including American Association for Cancer Research (AACR), Keystone Symposia for 3D Genome Architecture, American Society for Cell Biology (ASCB) and Women in STEM Symposia at University of Tennessee. At University of Tennessee she received multiple honors including Biochemistry Alumni Graduate Assistant Award, Dr. Cynthia B. Peterson Fellowship and Yates Dissertation Fellowship.