Impact of Rapamycin on Genome Organization

A Thesis Submitted to the

College of Graduate and Postdoctoral Studies

In Partial Fulfilment of the Requirements

For the Degree of Master of Science

In the Department of Biochemistry

University of Saskatchewan

Saskatoon

By

Joshua A. D. Pickering

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dead of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Head of the Department of Biochemistry

107 Wiggins Rd

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5E5

Canada

Dean

College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan S7N 5C9

Canada

Abstract

Rapamycin, a putative anti-aging drug, has been shown to extend the lifespan of several model organisms. Rapamycin's core effect is to inhibit a major nutrient sensing pathway of the cell, mechanistic Target of Rapamycin Complex 1 (mTORC1). This inhibition was previously thought to induce similar beneficial effects that parallel caloric restriction – the restriction of calories in an organism without causing malnutrition. Previous work has shown that Rapamycin induces significant changes in gene expression of normal human dermal fibroblasts, as well as inducing repositioning of chromosome within the nuclear volume. However, questions remained on how other chromosomes repositioned and if changes were occurring at sub-chromosomal levels. By utilizing fluorescence *in situ* hybridization (FISH) and Chromosome Conformation Capture (3C), we analyzed genome organization under several different growth conditions including Rapamycin treatment. Despite significantly altering the transcript levels of several *CXCL* (chemokine C-X-C motif ligand) genes during treatment with Rapamycin, we observed only minimal movement of chromosome 4 which contains these genes. However, when we analyzed the sub-chromosomal organization of the locus containing these genes, we discovered altered topology in response to both Rapamycin treatment and reversible growth arrest through induction of quiescence. We also observed that interactions occurred between *LIF* (Leukemia inhibitory factor) and two putative enhancers. Furthermore, we investigated a locus containing several *KRTAP* (keratin associated proteins) on chromosome 17 and discovered they also altered local genome topology in response to our treatments.

Acknowledgements

This project would not have been possible without continued support from and discussions with the members of the Eskiw lab, notably Dr. Christopher H. Eskiw, Dr. Zach Belak, Zoe Gillespie, and Kira Wang. I would also like to thank my committee members, Dr. Luo, Dr. Bonham, Dr. Lukong, for the time they invested into advisory meetings as well as all the helpful advice provided. I would also like to thank Dr. Steve Robinson for serving on my defense committee.

Furthermore, I would like to thank the member of the Biochemistry departments for their funding, support, events, and organization that facilitated my time in the program. Additionally, I thank the Food and Biosciences department for welcoming me to their events during the course of my program. Finally, I would like to thank my family for their support during my pursuit of a Master's Degree.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2DD – Primary human foreskin fibroblasts

3C – Chromosome Conformation Capture

3D – Three-dimensional

AMPK – AMP-activated protein kinase

AREG – Amphiregulin

AMP – Adenosine Monophosphate

ATP – Adenosine Tri-phosphate

BAF - BRG1- or HBRM-associated factors

Bp - Base pair

BRG1 – Brahma-related gene 1

BSA – Bovine Serum Albumin

BTM – Basal Transcription Machinery

CASTER – cellular arginine sensor for mTORC1

Cat – Catalogue

CDS – Coding DNA Sequences

CRCs – Chromatin Remodeling Complexes

CR – Caloric Restriction

CRM – Caloric Restriction mimetic

CT – Chromosome Territory

CTCF – CCCTC-binding factor

CTD – Carboxyl Terminal Domain

CXCL – Chemokine C-X-C Motif Ligand

DBD – DNA-Binding Domain

DDR – DNA Damage Response

DE – Digestion Efficiency

DMEM – Dulbecco's Modified Eagle Medium

DMSO – Dimethyl Sulfoxide

DNA – Deoxyribonucleic acid

DSB – double strand break

DTT – Dithiothreitol

EFEMP2 – EGF-containing fibulin-like extracellular matrix protein 2

EGFR - Epidermal growth factor receptor

eILF4E – eukaryotic translation factor 4E

EREG – Epiregulin

Erk5 – Mitogen activated protein kinase 7

EtOH – Ethanol

FBS – Fetal Bovine Serum

FISH – fluorescence in situ hybridization

FKBP12 – FK-binding protein 12

g – Gravity

GAPDH - Glyceraldehyde 3-phosphate dehydrogenase

GATOR – GTP activating proteins toward Rags complex

H – Histone (including; H1, H2A, H2B, H3, H4, H2A.X and H2A.Z)

H33342 – Hoechst 33342

H3K9 – Histone 3, Lysine residue 9

HATs – Histone acetyl transferases

HDACs – Histone deacetylases

HGPS – Hutchinson-Gilford Progeria Syndrome

HindIII – *HindIII* Restriction Enzyme

Hox – Homeobox

HS – Hypersensitivity (In regards to digestion by DNAse)

HUVEC – Human Umbilical Vein Cells

IFN- γ – Interferon gamma

IL – Interleukin

INO80 – Inositol requiring 80

Inr – initiator element

kB – kilobase

KEGG – Kyoto Encyclopedia of Gene and Genome

KLF – Kruppel-like factor

KRTAP – Keratin associated proteins

LAD – Lamin Associated Domains

LCR – Locus Control Region

LIF – Leukemia inhibitory factor

LMNA/C – Lamin A/C

LMNB1 – Lamin B1

lncRNA – long non-coding RNA

M – Molar

MAD – Matrix Associated Domains

mM – Millimolar

mRNA – Messenger RNA

mTOR – Mechanistic Target of Rapamycin

mTORC – Mechanistic target of Rapamycin complex

NAD – Nucleolar Associated Domains

NaOAC – Sodium Acetate

NB1 hTERT – NB1 human fibroblasts with human Telomerase expression

NDR – nucleosome-depleted regions

NF-kB – Nuclear factor kappa-light-chain-enhancer of activated B cells

NLS – Nuclear localization signal

nM – Nanomolar

NOR – nucleolar organizing regions

OR – Odorant Receptor

PRC2 - Polycomb repressive complex 2

TAD – Topologically associated domain

PTM – Post translational modification

qRT-PCR – quantifiable real time PCR

RAG1/2 – Recombinase Activating Genes 1 and 2

Raptor – Regulator protein associated with mTOR

REDD1 – Regulated in development and DNA damage responses

Ref – Reference #

Rheb – Ras homolog enriched in brain

Rictor – Rapamycin insensitive companion of mTOR

RNA – Ribonucleic acid

RNA Pol II – RNA polymerase II

RNA-Seq - RNA-Sequencing

rRNA – Ribosomal RNA

RSS – Recombinant Signal Sequence

RT – Room Temperature

S6K – Ribosomal protein S6 kinase beta

SAMD4a – Sterile Alpha Motif Domain Containing 4A

SESN2 - Sestrin2

siRNA – small interfering RNA

SSC – Saline Sodium Citrate

STAT5A/B – Signal transducer and activator of transcription

TAD – Topologically associating domains

TBE – Tris Borate EDTA

TBP – TATA binding protein

TF – Transcription factor

TFEB – Transcription factor EB

TFII – General Transcription Factor

T_h2 – T-Helper 2

TNF-AIP2 – Tumor Necrosis Factor Alpha-induced Protein 2

TNF-α - TNF alpha

tRNA – transfer RNA

TSC – tuberous sclerosis protein

TSS – Transcriptional Start Site

ULK1 – Unc-51 like autophagy activating kinase 1

UTR – Un-transcribed regions

UVRAG – UV radiation resistance associated gene

V(D)J – Variable Diversity and Joining

Wnt6 – Wingless-type MMTV integration site family, member

1.0 Review of the Literature

1.1 Introduction

Organization of the genome is fundamental for living organisms. The nucleus is a highly organized environment where chromosomes occupy defined volumes, known as chromosome territories (CT). Additionally, organization occurs below sub-chromosomal levels and play roles in the regulation of nuclear processes such as replication, repair and transcription. This folding of the genome has implications in many cellular processes, such as ageing and differentiation. However, it is unknown how somatic cells respond to external stimuli to refold their genome (or even if this is required) to facilitate the expression or repression of the appropriate genes. The drug Rapamycin disrupts cellular nutrient sensing and has been ascribed as an anti-ageing drug, with treatment in several model organisms resulting in increased lifespan (Bjedov et al., 2010; Ehninger et al., 2014; Harrison et al., 2009). We hypothesize that treating cultured primary dermal fibroblasts with Rapamycin will be informative on nutrient response and the ageing process. This project has several focuses based around genome organization, including: analysis of chromosome territory reorganization by fluorescence *in situ* hybridization (FISH), and analysis of long-range chromatin interactions using Chromosome Conformation Capture (3C). As 3C is a new technique we will be identifying internal controls and optimizing the protocol. The protocol will then be used for the analysis of genomic interactions within a region of chromosome 4 based on previous finding from RNA-Sequencing (RNA-Seq) data.

1.2mTOR and Rapamycin

The mechanistic Target of Rapamycin (mTOR) pathway, named after its inhibitor Rapamycin, functions as a centralized nutrient sensing hub. mTOR is responsible for transducing up-stream nutritional signals to downstream effectors through multiple protein pathways and has a significant role

in many functions. The mTOR pathway guides cell growth/proliferation, the synthesis of proteins, lipids, and nucleotides, as well as being a regulator of autophagy (cell *self-eating*). Activation of the mTOR pathway promotes cell proliferation/growth and inhibits autophagy as reviewed in Saxton and Sabatini, 2017. Conversely, during inhibition of the pathway using Rapamycin, cellular proliferation and growth are limited and inhibition of autophagy by mTOR is reduced.

Rapamycin (**Figure 1-1**) was discovered in 1975 by Sehgal *et al* (1975)*.,* eleven years after a Canadian expedition to collect soil samples from Easter Island*.* This macrocyclic lactone was extracted from the bacterium *Streptomyces hygroscopicus* and was discovered to have antibacterial properties. This antibacterial property was due to the ability of Rapamycin to selectively inhibit the ubiquitous mTOR pathway (Loewith and Hall, 2011). Since this discovery, the drug has been repurposed as an immunosuppressant during kidney transplantations (Abraham and Wiederrecht, 1996), and was suggested as a tool to slow cancer growth – inspiring the creation of Rapamycin paralogs, or 'Rapalogs', better tailored for treatment of cancers (temsirolimus, everolimus) (Graziani, 2009; Li et al., 2014). However, the impact of Rapamycin as well as the Rapalogs has not been investigated at the genomic level, and with this information, new functions and applications of Rapamycin may be discovered.

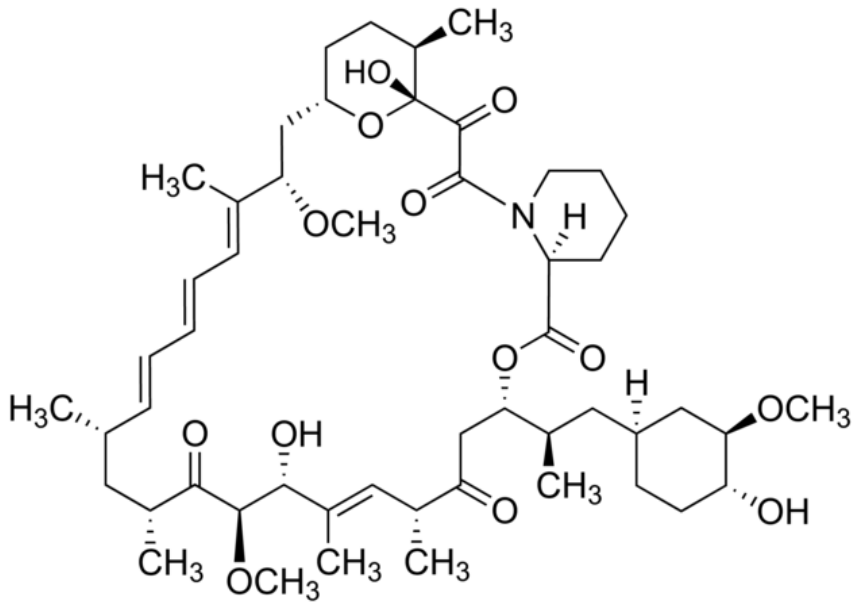


Figure 1-1: Chemical Structure of Rapamycin. Rapamycin is a macrocyclic lactone discovered from a soil sample found on Easter Island (Rapa Nui) in the Pacific Ocean. (Credit: https://www.sigmaaldrich.com)

1.2.1 The mTOR Complexes: TORC1 and TORC2

The mTOR pathway contains two major complexes, known as mTOR Complex 1 (mTORC1) and mTOR Complex 2 (mTORC2). Both complexes are responsible for regulating the health of cells; however, this is accomplished differently in each case. mTORC1 regulates the status of the cell and when active positively regulates translation, nucleotide synthesis, lipid synthesis, glucose metabolism, growth and proliferation of cells while downregulating autophagy (as reviewed in (Saxton and Sabatini, 2017)). mTORC2 acts on downstream effectors to positively regulate ion transport, glucose metabolism, cell migration and cytoskeletal rearrangement, and to negatively regulate autophagy as reviewed in Oh and Jacinto, 2011. mTOR function has a significant impact on the status of the cell; however, how this complex regulation relates to maintaining cellular health and cellular ageing has yet to be established.

1.2.1.1 Up-stream regulation of mTORC1

As a nutrient sensing hub, mTORC1 receives signals from many up-stream inputs (**Figure 1-2**). The signal from many up-stream effectors act on a complex of Tuberous sclerosis proteins 1 and 2 (TSC1/2) (as reviewed in (Huang and Manning, 2008)). When activated, this TSC complex inhibits the activity of Ras homolog enriched in brain (Rheb). This inhibition prevents Rheb from activating mTORC1. This TSC complex integrates positive growth signals (inactivates TSC, activates mTORC1) and negative growth signals (activates TSC, inhibits mTORC1) (Inoki et al., 2006). Upon activation, epidermal growth factor receptor (EGFR) transduces a signal through the MAPK (Mitogen-activated protein kinase) cascade and inhibits TSC and activates mTORC1. AMP-activated protein kinase (AMPK) acts as an energy sensor for mTORC1 activity by detecting the ratio of AMP to ATP and inhibiting mTORC1 activity by activating TSC in the absence of adequate ATP (Gwinn et al., 2008; Inoki et al., 2006). p53 (tumor protein 53) acts up-stream of AMPK and detects DNA damage, indirectly inhibiting mTORC1 activity in the presence of DNA damage (Akeno et al., 2015; Feng et al., 2005; Hasty et al., 2013). Hypoxic conditions are capable of inhibiting mTORC1 activity through regulated in development and DNA damage responses 1 (REDD1/DDIT4), which activates TSC, resulting in inhibition of Rheb and mTORC1 (Brugarolas et al., 2004). Cytoplasmic arginine and leucine regulate mTORC1 activity through the RAGulator complex which in turn is regulated through the CASTOR/GATOR pathway (Chantranupong et al., 2016). CASTOR, cellular arginine sensor for mTORC1, directly senses cytoplasmic arginine levels, resulting in modulation of mTORC1 function. GATOR, GTPase activating proteins toward Rags complex, directly senses cytoplasmic leucine, also modulating mTORC1 function. Increased levels of arginine result in the activation of mTORC1. Leucine is also associated with this pathway and increased leucine levels repress Sestrin2, also resulting in activation of mTORC1 (Wolfson et al., 2016). In summary, mTORC1 senses for adequate energy, amino acids (leucine and arginine) (Jewell et al., 2015; Jewell et al., 2013), growth signals, while also modulating the activity of mTORC1 to accommodate for stressors

such as DNA damage and hypoxic conditions (as reviewed in (Saxton and Sabatini, 2017; Zoncu et al., 2011). mTORC1 is a sensing hub for many signals and is able to monitor the environment around the cell to allow for appropriate proliferation and growth to occur. mTORC1 also has the ability to slow cell proliferation when the situation is unfavorable or if adequate signals are given to change the stance of cells to repair and maintenance from growth/proliferation.

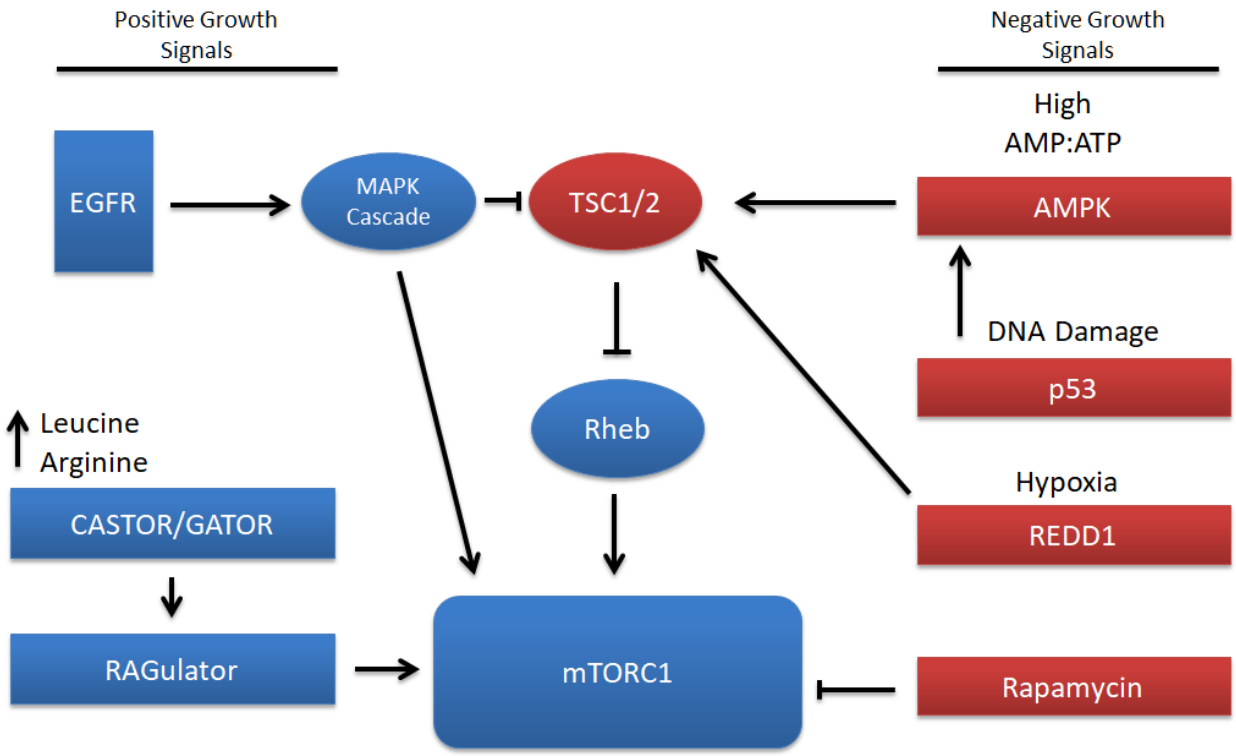


Figure 1-2: Upstream Regulation of mTORC1. A summary of the multiple upstream effectors of mechanistic Target of Rapamycin Complex 1 (mTORC1). Positive Growth signals (activators) of mTORC1 are shown in blue. Negative growth signals (in-activators) of mTORC1 are shown in red. mTORC1 is a central nutrient sensing hub that mediates cellular response according to nutritional availability and environmental conditions.

1.2.1.2 mTORC1 control of translation

mTORC1 is responsible for modulating translation rates through ribosomal protein-S6-kinase beta (S6K). Active mTORC1 phosphorylates S6K which phosphorylates ribosomal subunits to initiate translation (Holz and Blenis, 2005; Magnuson et al., 2012). Active mTORC1 also promotes translation indirectly by activating eukaryotic translation factor 4E (elF4E), which promotes 5' cap-dependent translation by recruiting ribosomes. Both of these mechanisms indicate the vital role active mTORC1 has in promoting translation (Wraught and Gingras, 1999) and provide further evidence that mTORC1 has significant control and involvement in major cellular processes.

1.2.1.3 *mTORC1* control of autophagy

Active mTORC1 suppresses autophagy (defined as cell *self-eating*) (Jung et al., 2010). Autophagy has been shown to be central for protein homeostasis in times of cellular stress. Furthermore, autophagy is involved in the elimination of damaged organelles such as mitochondria (mitophagy) and is linked to health- and life-span extension. When mTORC1 is activated by adequate nutrient presence, it inhibits autophagy by inhibiting several proteins responsible for driving autophagy (ULK1: Unc-51 like autophagy activating kinase 1, TFEB: Transcription Factor EB, UVRAG: UV Radiation Resistance Associated Gene) and Proteasome assembly (Erk5: Mitogen activated protein kinase 7) (Martina et al., 2012; Noda, 2017). While mTORC1 is not activated, decreased phosphorylation of ULK1 leads to increased ULK1 activity resulting in phosphorylation of downstream activators of autophagy. Therefore, mTORC1 serves a central role that balances protein translation (active mTORC1) and protein recycling and organelle degradation (inactive mTORC1) by autophagy.

1.2.2 Rapamycin and its inhibition of mTORC1

Despite the similar composition of the two mTOR complexes, Rapamycin selectively inhibits mTORC1 due to its preferential binding of the FK-binding protein 12 (FKBP12), resulting in partial blocking of the active site of mTOR upon binding the FKBP12-Rapamycin pair to the complex (Dumont, 1995). mTORC2, though similar, is not directly inhibited by this mechanism due to its utilization of Rapamycin insensitive companion of mTOR (Rictor) in its complex rather than regulatory protein associated with mTOR (Raptor). Rictor uniquely occludes the binding site (an FRB homology domain on mTOR) to the FKBP12/Rapamycin pair (Reviewed in Saxton and Sabatini, 2017). However, there is discussion on whether long term treatment of cells or patients would lead to inhibition of mTORC2 as well (Sarbassov et al., 2006). Due to its specificity, Rapamycin is a powerful tool for investigating the effects of mTORC1 on the genome and cellular status.

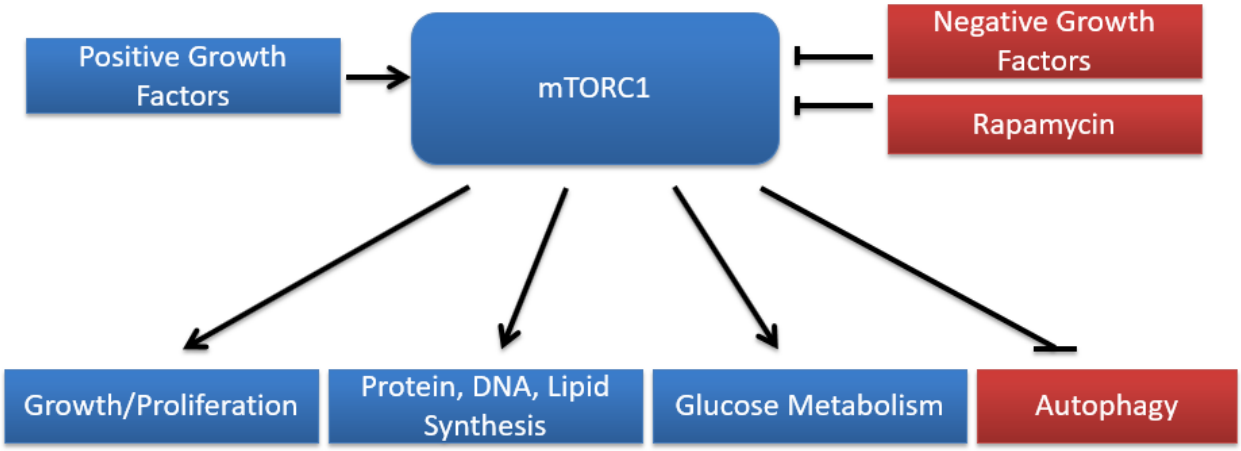


Figure 1-3: Rapamycin inactivates growth and proliferation of cells and promotes autophagy. mTORC1 serves as a nutrient sensing hub responsible for driving growth and proliferation when adequate nutrition is available. By treating with Rapamycin, cells bypass nutrient sensing and switch from promoting proliferation and growth (blue) to promoting maintenance and repair (red).

1.2.3 Rapamycin impacts gene expression patterns of normal dermal fibroblasts

Earlier work from the Eskiw Lab (Gillespie et al., 2015) characterized genome wide gene expression changes in normal primary fibroblast cells following treatment with 500 nM Rapamycin. Following RNA-sequencing, it was revealed that treatment with Rapamycin induced the expression of many genes regulated by the Signal Transducer and Activator of Transcription 5A/B (STAT5A/B) regulation, including several cytokine genes as well as the *Leukemia Inhibitory Factor* (*LIF*). Using Kyoto Encyclopedia of Gene and Genome (KEGG) terms, a link to mTOR was discovered through the significantly enriched PI3K-AKT signaling terms. Furthermore, rather than driving a specific cytokine pathway, treatment with Rapamycin provoked a wide array of enrichment across cytokines. How Rapamycin functions, and the rationale for its putative anti-aging properties, may lay in some unexamined connection between these up-regulated genes.

1.2.4 Rapamycin parallels Caloric Restriction

Caloric Restriction (CR) is defined as a reduction of calorie intake without causing malnutrition. CR treatment been shown to ameliorate ageing phenotypes in several model organisms (Colman *et al*., 2009; Sohal and Forster., 2014; Gilmore and Redman., 2018; Arlia-Ciommo *et al*., 2018; Fontana and Partridge, 2015); however, caloric restriction diets are difficult to follow. As a result, there is interest in discovering drugs that mimic the effects of CR without the need for strict dietary plans; called CR mimetics (CRM).. Several CRM drug candidates have been identified, including Rapamycin (Kitada and Koya, 2013). Rapamycin is thought to parallel CR by downregulation of the mTORC1 pathway. CR lowers mTORC1 pathway activity by reducing the amount of cellular energy (AMP:ATP ratio) sensed upstream of mTORC1 (Blagosklonny., 2010). Functionally, Rapamycin treatment uncouples molecular nutrient/energy sensing which induces a perceived state of lower cellular energy and nutrient

availability, regardless of actual nutrient and energy presence. Rapamycin, and the class of drugs with similar structure to rapamycin (called rapalogues) would therefore be a candidates to reduce the ageing phenotype without requiring a restriction of caloric intake.

1.3 Genome Organization

The human genome consists of 3.6 billion nucleotide base pairs (bp) encoded on 23 pairs of chromosomes: 22 autosomal pairs and a pair of sex chromosomes. These chromosomes range in size from 250 million to 50 million bp. This linear genetic sequence is formed from 4 distinct nucleotides; Thymine (T), Adenine (A), Cytosine (C), and Guanidine (G) and forms a duplex with a complementary strand through hydrogen-binding. Hydrogen-bonding occurs between pairs of these nucleotides, containing one purine (A or G) and one pyrimidine (T or C) per pair (AT or GC), and dictates the formation of the canonical deoxyribonucleic acid (DNA) double-helix (Watson and Crick, 1953). This double-helix has a diameter of 20 Angstrom (Å) or 2 nanometers (nm), rotates roughly every 10 residues, and contains two types of grooves; the major groove and the minor groove. The unique characteristics and three-dimensional structure of DNA allows a variety of functions to be possible, including annealing and denaturation of complementary strands as required, protein binding at major grooves of DNA, and sequence repair based on the complementary nature of the genetic code (as reviewed in (Travers and Muskhelishvili, 2015).

DNA is the 'blueprint' for life and its sequences code for genes and regulatory elements. By combining these elements, genes can be regulated and produced in response to stimuli, or during development. To date, many factors that affect disease state and phenotypic difference have been attributed to portions of the genome previously thought to have no function, positioned outside of the coding region of genes, and were previously named 'junk DNA'. A transition has been made from interpreting non-coding sequences as 'junk' to new and complicated roles in genomic function (as

reviewed in (Palazzo, 2014)). When accounting for ~2% of genetic code being protein coding genes, it can be estimated that ~98% of the genome is non-coding material – 3.53 billion of 3.6 billion base pairs.

In humans, DNA forms a complex with histone (H) proteins to form chromatin. Chromatin is a complex of DNA and octamers of histone proteins containing 2 copies of H2A, H2B, H3, and H4 wrapped by 146 base pairs of DNA (as reviewed in (Kouzarides, 2007)). This association results in the 'beads on a string' 11 nm fiber. Chromatin is classified into either euchromatin or heterochromatin depending on the level of DNA condensation and presence of additional factors. Euchromatin is less compacted and more 'open' and contains expressed genes, whereas heterochromatin is highly compacted and resultantly 'closed' and contains genes that are inactive (as reviewed in (Margueron and Reinberg, 2010)). The 'closed' state of heterochromatin denies RNA polymerase II and other factors access to the contained genes, subsequently silencing them. Heterochromatin is further divided into either constitutive or facultative heterochromatin. Constitutive heterochromatin occurs in gene poor regions such as the centromeres or telomeres on chromosomes and is generally conserved between cell types. Whereas facultative heterochromatin varies between cell types and its formation plays a role in silencing gene expression and establishing cell type (Dixon et al., 2015; Kramer, 2016; Shchuka et al., 2015). For example, one of the two X chromosome in female somatic cells undergoes silencing by heterochromatin formation (Cheng and Disteche, 2004). Both facultative and constitutive heterochromatin are enriched with the histone linker (H1), which forms a chromatosome when combined with the nucleosome, and potentially further condenses the 11 nm fiber into the more highly compact 30 nm solenoid or zig-zag structure.

Chromatin structure is highly dynamic and is regulated at several levels, such as through the exchange of histones and the introduction of functionally unique histone variants, and by chromatin remodeling complexes (CRC). The exchange of histones, also called 'nucleosome dynamics', facilitates the maintenance of chromatin and histone exchange/turnover within transcriptionally active regions.

This also allows the introduction of histone variants that differ by only a few amino acids into the nucleosome. Examples include the chromatin remodeler recruiting H2A.Z, which is integrated into nucleosomes near nucleosome-depleted regions (NDRs) and facilitates the activation of genes by antagonizing DNA methylation (Venkatesh and Workman, 2015; Zilberman et al., 2008). H3.3 which is associated with active genes, and macroH2A which is involved in the inactivation of the X-chromosome (as reviewed in (Venkatesh and Workman, 2015)). Another actor regulating chromatin structure lays in the chromatin remodeling complexes, such as Inositol requiring 80 (INO80) or Brahma-related gene 1 (BRG1- or hBRM)-associated factors (BAF) (Kapoor et al., 2013). These complexes utilize ATP, as the association between the positively charged histones and negatively charged DNA is energetically favourable and resistant to alteration, to slide and/or evict nucleosomes from a target region through several different mechanisms. Other factors such as TATA binding protein (TBP) act in tandem with remodeling complexes to facilitate nucleosome sliding (as reviewed in (Juven-Gershon and Kadonaga, 2010)).

Above the linear sequence and chromatin organization of the genome is the epigenetic code. The epigenetic code is a series of post-translational modifications (PTMs) applied to N-terminal peptide tails of the histones (as reviewed in (Bannister and Kouzarides, 2011)). These modifications may occur on several different residues such as lysine (K) and arginine (R) and include acetylation, methylation, phosphorylation, and ubiquitination (as reviewed in (Kouzarides, 2007)). The result of these modifications is dependent on the specific residue modified as well as the type of modification and the local genomic elements. For example, acetylation of lysine residues regulated by histone acetyl transferases (HATs) and histone deacetylases (HDACs) interrupts the positive charge of the lysine residues and weakens the interaction with the negatively charged DNA. Methylation states can occur as mono-, di-, or tri-methylated at lysine or arginine residues and rather than altering the charge status of the residue, the methylation facilitates the binding of proteins. Depending on the specific lysines or

arginines being modified, methylation states may activate or repress transcription. For example, H3K9 acetylation is associated with actively transcribed regions and tri-methylated H3K9 is associated with a region of repressed expression and heterochromatin. Furthermore, sections of DNA known as CpG islands (a repeated CG motif) can undergo methylation to silence transcription of nearby genes and block the binding of proteins (Saxonov et al., 2006). CpG islands are usually positioned near the transcription start site of genes and are capable of undergoing methylation at C residues to inhibit transcription.

The human genome is compartmentalized into the nucleus; a double lipid bilayered organelle that provides a highly controlled environment for the genome. The nucleus is known to be highly organized and the work of many labs has opened this field into many directions, from genome topology in the nucleus to characterizing accumulation of factors at precise nuclear locations (Albiez et al., 2006; Bolzer et al., 2005; Cremer and Cremer, 2010; Cremer et al., 2018; Dundr, 2012; Guo et al., 2011; Osborne et al., 2004; Solovei et al., 2009; Spann et al., 2002; Ulianov et al., 2016; Zorn et al., 1976). While the systems of nuclear organization are still being investigated, several aspects have emerged on how the organization of such a fundamental organelle is orchestrated. One such example is the nucleolus. The nucleolus serves as a prototypical example of nuclear organization. The acrocentromeric chromosomes (13, 14, 15, 21 and 22) containing the rRNA genes organize to a distinct nuclear location. Nucleolar organization can coalesce into one of several nucleoli depending on cell type and disease state. Upon coalescence of the nucleolar organizing regions (NORs) a functional nucleolus is formed which facilitates ribosome biogenesis (Floutsakou et al., 2013). This provides an example of nuclear organization and also an example of a conserved biological process that utilizes organization of chromatin for a functional purpose.

Nuclear lamina represent another component of nuclear organization (as reviewed in (Dechat et al., 2009)). Lamins are proteins that form a network at the inner face of the nuclear membrane,

providing a scaffolding site that serves as an anchorage point for heterochromatin domains. The significance of this scaffolding is seen in the pre-mature ageing disease Hutchinson-Gilford Progeria Syndrome (HGPS), caused by a point mutation in the *Lamin A/C* gene, leading to the expression of a truncated form of the Lamin A/C protein (Liu et al., 2005). The aberrant accumulation of this protein in the lamina results in dys-regulation of several essential nuclear functions including, transcriptional regulation and DNA repair (Buchwalter and Hetzer., 2017; Weaton *et al*., 2017).

The nucleus is a crowded organelle requiring spatiotemporal organization of many factors (Hancock, 2014; Spitz and Furlong, 2012). It potentially contains between several hundred to thousands of varieties transcription factors, most of which are currently uncharacterized, that must be able to locate and bind the correct genomic loci (Lambert et al., 2018; Vaquerizas et al., 2009). It also must orchestrate mRNAs to be properly localized to and processed by splicing machinery prior to export from the nucleus (as reviewed in (Katahira, 2015)). The nuclear environment facilitates the compaction and regulation of DNA by the formation euchromatin and heterochromatin; a dynamic and highly tailored process. To this end the nuclear environment must also facilitate the integration and renewal of histones and histones variants with functional consequences, such as histone tagging of double strand breaks (DSB) which are then repositioned to distinct nuclear sites for repair (Marnef and Legube, 2017; Redwood et al., 2011). Historically it has been suggested in textbooks that the genome is condensed to allow it to fit inside of the nuclear volume, additionally, it could be said that the compaction of the genome also allows room for other factors to be present (Dehghani et al., 2005).

1.3.1 Protein Coding genes

It has been predicted that the human genome contains anywhere between 19,000 to 25,000 protein coding genes that code the human 'proteome' – the sum of all proteins expressed by a cell (Clamp et al., 2007; Kim et al., 2014). From these 19,000 to 25,000 possible genes, up to 300,000 peptide and proteins may be produced. These protein genes include sequences that code for mRNA

consisting of exons which contain codons that encode amino acid sequence, introns which are contained between exons and excised during splicing events, and poly-A tails that regulate mRNA half-life. Intron and exon splicing during pre-mRNA processing allows for protein variants through alternative splicing that selects for different combinations of exons to form the final coding DNA sequence (CDS) (as reviewed by (Bentley, 2014; Shi, 2017)). mRNA additionally contains two un-transcribed regions (UTR), known as the 5' and 3' UTR, which flank the protein coding sequence. The 5' UTR contains the initiation sequence ('Start codon') AUG which codes for the amino acid methionine. The function of the 3' UTR has been implicated in mRNA stability and translation efficiency (Matoulkova et al., 2012). The output from this process is an amino-acid chain or peptide capable of folding, with or without chaperones, into a functional enzyme capable of catalyzing chemical reactions. The amino acids contained in this peptide sequence may include factors such as a nuclear localization signal (NLS) or glycosylation sites that encode where enzymes should localize and operate (as reviewed in (Bauer et al., 2015)). The linear sequence of the genome provides the potential for protein splicing variants (different combinations of exons) to be produced from a single protein coding gene. This splicing is a highly organized, co-transcriptional, three-dimensional event in the nucleus, requiring the co-localization of many factors and resulting in several hundred-thousand proteins being produced from a much lower number of protein coding genes (as reviewed by (Spector and Lamond, 2011)).

1.3.2 Promoters

To regulate the expression of mRNA-producing genes the genome contains promoters and enhancers. Promoters occupy regions up-stream of genes and facilitate the binding of protein factors as well as RNA Polymerases (such as RNA Polymerase II). For protein-coding genes, these regions occur within short distance of the gene and contain the transcriptional start site (TSS), the location where the polymerization of RNA begins (as reviewed in (Blackwood and Kadonaga, 1998; Kugel, 2017)). Many

(~46%) human genes contain a 17 bp sequence known as an initiator element (*Inr*) that provides a binding site for factors that summarily recruit the basal transcription machinery (BTM) (Kugel, 2017; Sikorski and Buratowski, 2009; Smale and Baltimore, 1989). Many human promoters contain a TATA-box or a TATA-like-box – a sequence in the promoter region that contains a conserved TATAA sequence within 50 base pairs up-stream of the TSS (Yanga et al., 2007). The TATA-box and TATA-box-like sequences has been estimated occurring in ~24% of genes; however, this number varies (11.6% - 76%) depending on the method and parameters of estimation (Yanga et al., 2007). The TATA box provides a docking site for the TATA binding protein (TBP). This binding introduces a bend in the DNA and facilitates transcriptionally-permissive unwinding of the DNA strand. For a promoter to be properly utilized, it requires the recruitment of relevant factors at specific sequences. These factors must somehow navigate the chaos of the nucleus to find the relevant DNA-binding domain or mediator protein to facilitate transcriptional activation by the promoter.

1.3.3 Enhancers

Enhancers are regions of nucleotide sequences distal to a gene with motifs that are recognized by DNA-binding proteins, such as transcription factors (Blackwood and Kadonaga, 1998; Pennacchio et al., 2013). Upon introduction of DNA-bending proteins and the binding of transcription factors at the enhancer, enhancers are held in close physical proximity to their relevant promoter by looping of the DNA strand. This looping is maintained by transcription factors, as well as other proteins such as the Mediator complex (as reviewed in (Allen and Taatjes, 2015)). The formation of this three-dimensional folding has dramatic effects on gene expression and increases the likelihood of gene transcription when formed. Enhancers have been attributed roles in differentiation and development, wherein differential enhancer utilization may guide tissue differentiation or developmental variants of genes. Enhancers have also served a role in evolution through the formation of a novel regulation for a gene. Examples of

these types of enhancer utilization can be found in T-helper cell differentiation (Spilianakis and Flavell, 2004; Spilianakis et al., 2005), and the timed expression of the globin genes during development (Carter et al., 2002; Tolhuis et al., 2002). It has been observed that some genes have multiple enhancers available which may provide a role in developmental or in alternate usage when one enhancer is unavailable. These are referred to as 'Shadow Enhancers' and are currently under investigation to discover if their existence is a redundancy or a functional element (Barolo, 2012; Cannavo et al., 2016). Additionally, silencer elements have also been identified and are capable of inhibiting transcription by blocking the binding of RNA Pol II to the promoters of genes. Enhancers, mediated by transcription factors, provide powerful elements of gene regulation. Together they are capable of being tissue specific and able to enhancer the transcription of multiple difference genes. Importantly, these regulatory mechanisms require the three-dimensional organization (looping) of the genome.

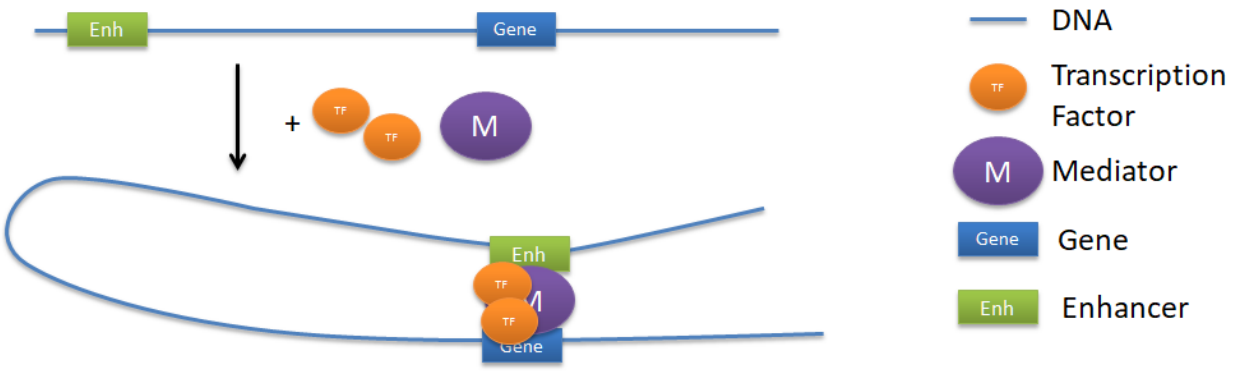


Figure 1-4: Enhancers act on gene promoters. Genomes are capable of 'looping out' sequences of DNA to facilitate contact between different genetic elements, such as enhancers and promoters, to increase transcription. This contact is facilitated by transcription factors that bind to specific DNA motifs and other proteins that mediate or bridge contact between the transcription factors or genetic elements.

1.3.4 Insulators/CTCF Sites

Insulator elements, sometimes referred to as boundary elements, are regions of the genome that 'block' the erroneous interaction between an enhancer and promoter (Barkess and West, 2012). These regions contain sequences recognized by CCCTC-binding factor (CTCF) as well as other factors (Dunn and Davie, 2003; Herold et al., 2012; Kim et al., 2007). CTCF, along with the Cohesion complex, bind these insulator elements and through CTCF dimerization localize them together to form a loop (Nativio et al., 2009; Rubio et al., 2008). The formation of this loop compartmentalizes a section of the genome and resultantly affects gene regulatory mechanisms by promoting internal enhancer-promoter interactions while discouraging external interactions (as reviewed by (Ong and Corces, 2014; Phillips and Corces, 2009)). This topological regulation has been attributed significant roles in complex biological processes such as V(D)J (Variable, diversity, and joining) recombination (discussed later) (Guo et al., 2011). Loss of insulator elements has also been discovered as the cause of certain disease states, where an enhancer acts inappropriately over a deleted boundary and on a nearby gene such as seen in adult onset leukodystrophy (Giorgio et al., 2015). These examples show the importance of insulator elements in mediating genome organization and function. Insulator elements may also provide a role in tissue differentiation and development by regulating enhancer looping that will be required at a later point (Doerks et al., 2002; Vogelmann et al., 2011). Binding of the CTCF protein to the CTCF element is disrupted by CpG methylation on the element and is a known component of tissue differentiation commitment (Phillips-Cremins et al., 2013).

1.4 Gene Regulation and Transcription

1.4.1 Transcription

Transcription is the process by which a targeted region of DNA is transcribed into RNA, prior to up-stream processes such as protein translation. To accomplish transcription, the human genome encodes at least three functional RNA polymerases with distinct biological roles, as well as various transcription factors. Two of the RNA Polymerases present, RNA Polymerase I and III, are responsible for transcribing rRNA used in ribosome biosynthesis and tRNA which carries specific amino acids prior to their utilization in the growing peptide of protein translation (Cramer et al., 2008; Melnik et al., 2011). The remaining polymerase, RNA Polymerase II is responsible for transcribing protein coding genes into mRNA.

The process of mRNA transcription is divided into three phases; initiation, elongation, and termination. Prior to initiation several components assemble at a site up-stream of a gene known as the core promoter. These factors include general transcription factors (TFIIA, -B, -D, -E, -F, -H) and RNA polymerase II, that together form the basal transcription machinery (BTM) (Goodrich and Tjian, 1994; Shandilya and Roberts, 2012; Villicaña et al., 2014). Following this assembly RNA Polymerase II will initiate and repeatedly produce short abortive transcripts until a long enough transcript is produced and RNA Pol II escapes the promoter to drive the process forward into the elongation phase (as reviewed by (Brookes and Pombo, 2009)). Transcription is terminated upon reading of a polyadenylation signal, subsequent cleavage of the mRNA transcript, and polyadenylation of the mRNA (as reviewed by (Richard and Manley, 2009)). Transcription is only possible when all the necessary factors gather. Considering the number of factors involved it would seem natural and efficient that factors are either collected at a tailored nuclear location, or mechanisms exist to rapidly drive factors to needed locations. Both of these necessitate organization of the nucleus.

1.4.2 Transcription Factors

Transcription factors are proteins containing DNA-binding domains (DBD) as well as other domains that act to promote and/or regulate transcriptional activity, such as adaptor sequences that facilitate binding with other proteins (as reviewed in (Lambert et al., 2018). The DBD provide specificity for transcription factors to bind specific sequence motifs up-stream of the transcriptional start site (TSS) of a gene or at distal enhancers. These sequences are referred to as response elements or more descriptively as transcription factor binding sites (Inukai et al., 2017). Upon transcription factor binding, other factors (such as the mediator complex) are recruited to the gene to prepare for transcription. Transcription factors react to cellular signaling pathways and locate to the nucleoplasm to exert their effect on the expression of relevant genes. These factors may act as either activators or repressors, or pleiotropically, wherein one factor can up-regulate a certain subset of genes while simultaneously down-regulating another subset (Lambert et al., 2018). Additionally, transcription factors may require co-activators or co-repressors to achieve a function on gene regulation. For example, STAT5A transduces up-stream signals from a Janus Kinase (JAK), in the form of a phosphorylation on STAT5A, causing two STAT5A proteins to dimerize and enter the nucleus where it binds a STAT5A response element (Schindler et al., 2007). Transcription factors provide a powerful tool for gene regulation as bind to specific sequences of DNA, activate their effect in response to specific stimuli, and are capable of combining with other factors to further tailor their response on the transcription of a gene. To have any effect on their target of interest, transcription factors must be localized to genes and other relevant genetic elements.

1.5 Genome Topology

Genome topology is a relatively new field. Its investigations are centered how the genome is organizes inside the nuclear volume and how this organization relates to cellular function and responses to stimuli. One of its core techniques, DNA Fluorescence *in situ* hybridization (DNA-FISH), has allowed characterization of chromosome positioning and space occupancy in the nucleus (Bolzer et al., 2005; Cremer and Cremer, 2001; Cremer et al., 2006; Solovei et al., 2009). Investigations into genome topology have further illuminated the nucleus as a highly organized compartment capable of responding to stimuli in a repeatable manner.

1.5.1 Chromosomes and Chromosome Territories

Contrary to the model commonly portrayed in textbooks, featuring unwound chromosomes expanding entropically to fill the nucleus, chromosomes preferentially occupy local regions known as chromosome territories (CTs) (Bolzer et al., 2005; Cremer and Cremer, 2010; Cremer et al., 2006; Zorn et al., 1976). These territories are distinct; however, mingling of chromosome strands may occur inside the territory or at territory boundaries. The position of these territories is non-random and have been observed to change position following treatment (Gillespie et al., 2015; Mehta et al., 2010). Correlating nuclear positioning of chromosome territories to gene expression and function is not clear, but it has been suggested that localization to the periphery of the nucleus correlates to gene-poor chromosomes (low gene density) or decreased gene expression and silencing, while localization to the interior is correlated with gene-rich chromosomes (higher gene density) and increased gene expression (Cremer et al., 2003; Croft et al., 1999). It has been observed that the peripheral localization of genes with low expression levels is consistent with the increased quantities of heterochromatin found at the nuclear periphery (as reviewed in (Bridger and Bickmore, 1998)). However, notions of global radial positioning

correlating to gene expression are challenged by the expression of genes located within peripheral regions, and by complications introduced by gene positioning correlating to tissue differentiation irrelevant of expression (as reviewed in (Mahy et al., 2002)). The investigations made into genome organization and chromosome territory position have highlighted the dynamic ability of the genome to reorganize in response to stimuli.

Figure 1-5: Chromosome territories. Illustration of imagined chromosome territory positioning in the nucleus. Chromosomes occupy distinct territories in the nucleus. Each colored section represents a chromosome and how it may occupy a territory in the nucleus. Chromosome territories can alter position in response to stimuli.

1.5.2 Chromosome Territory Repositioning

Chromosomes occupy specific territories in the nucleus, these are known as chromosome territories (CT). During mitosis the organization of these chromosome territories is lost until re-established early in the G1 phase of the cell cycle. The exact mechanism of chromosome repositioning is unknown, as is the putative nuclear framework to guide and organize the repositioning. However, it has been shown by Mehta and colleagues (Mehta et al., 2010) that disruption of myosin with siRNA (small interfering RNA) interference resulted in no repositioning of chromosomes in response to cellular

stimuli. Under normal conditions, chromosomes are capable of re-positioning in response to stimuli (such as decreased serum levels) within 15 minutes and this process does not require mitosis for re-organization. This work established that chromosomes can move rapidly in response to stimuli and that a process utilizing myosin is involved.

Chromosome territory positioning during interphase is dynamic and repeatable. It has been shown previously in our lab and others that upon induction of quiescence chromosomes 18 and 10 reposition predictably, with chromosome 18 assuming a more internal position, and chromosome 10 a more peripheral position. A similar repositioning of chromosomes has been observed in our lab following treatment with Rapamycin (Gillespie et al., 2015; Mehta et al., 2010). This informs that the movement of a chromosome has a functional purpose and is non-random.

Chromosome territories have the capability to interact with their neighbors. This is known as chromosome territory intermingling and involves the looping out of chromosomal segments from the core of the chromosome territory to an intermediate position between two neighboring chromosomes (as reviewed by (Belyaeva et al., 2017). Boundaries of chromosome territories allow intermingling with strands from other chromosomes to occur on the surface. These strands can extend long distances across the nuclear volume to invade a chromosome territory. Each of these interactions may represent specific functions related to gene expression (discussed in section 1.7). One such function would be the interaction between a promoter and an enhancer which requires two chromosomes to be in close physical proximity in the nuclear volume, as well as precise organization of the genome. An example of this is found in T-helper cells, where regulatory elements on chromosome 11 localize to the promoter of IFN-γ chromosome 10 (Spilianakis et al., 2005). Another function is found in the co-localization of like-regulated genes from different chromosomes. This was observed at transcription factories (discussed in section 1.7), where like-regulated genes and the necessary transcriptional machinery was concentrated. Changes to chromosome territory positioning may facilitate new or different intermingling between

chromosomes, such as bringing an enhancer from one chromosome into proximity of a promoter on a different chromosome. Additionally, it may bring like-regulated genes together in three-dimensional space, such as a shared site of transcription.

1.5.3 Higher Order Genome Topology

With the advent of Chromosome Conformation Capture (3C) (Dekker et al., 2002), it became possible to investigate topology at the sub-chromosomal level. Previously, investigations relied on DNA-FISH that fundamentally lacked the optical resolution to characterize co-localization of elements of interest, such as a promoter and an enhancer (**Figure 1-6**). Since its inception, 3C has spawned a series of derivative techniques, notably Hi-C (Lieberman-Aiden., 2009), that allows topological characterization of entire genomes. These investigations have revealed evolutionarily conserved domains of the genome that encompass up to several genes and other genetic elements (Dixon et al., 2012) and illuminated processes controlling development and cellular differentiation. Furthermore, dysregulation of genome topology has also been tied to several disease states, such as the fusing of the index finger and thumb which was recapitulated in mice (as reviewed in (Lupianez et al., 2016)). Investigation of sub-chromosomal topology paves the way for understanding complicated and currently under-characterized mechanisms of gene regulation.

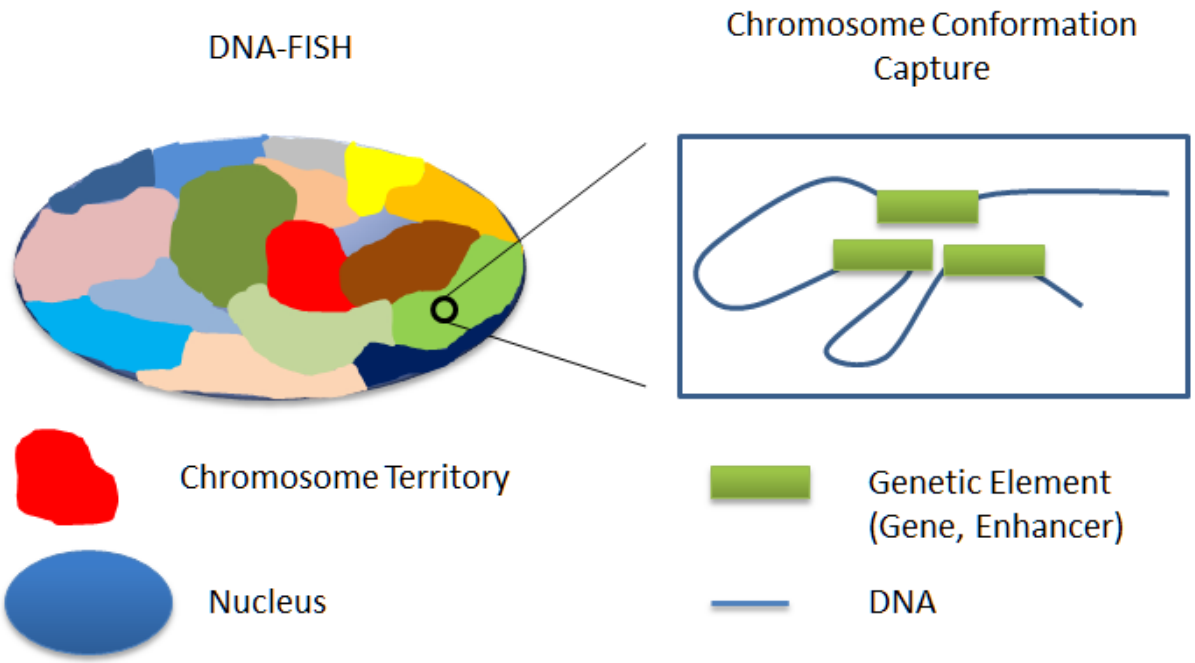


Figure 1-6: From chromosome territory to sub-chromosomal topology. Chromosomes occupy distinct territories in the nucleus. The positioning of these chromosomes has been related to cellular health, status, and transcriptional profile. Two techniques, DNA-FISH and Chromosome Conformation Capture (3C) can be used together to bridge large-scale organizational changes of whole chromosomes to alterations in the topology of genetic elements, such an enhancers and promoters of genes.

1.6 Long Range Interactions

Long range chromatin interactions are interactions between two sections of chromatin that are brought together in three-dimensional space, often 'looping out' sequences contained between the interacting sites (Carter et al., 2002; Spilianakis and Flavell, 2004; Tolhuis et al., 2002). Interactions can be defined as *cis*- (occurring between two points on the same chromosome) and *trans*- (occurring between two points, each on different chromosomes). These interactions are mediated and maintained by proteins capable of targeting distinct sequences of chromatin such as TFs and CTCF. These interactions can be divided into several categories. Firstly, there are interactions that occur between

promoters and up to several elements (enhancers and silencers) that regulate gene expression and result in genome folding (Blackwood and Kadonaga, 1998; Raab and Kamakaka, 2010). Rather than directly affecting gene expression, insulators flank and compartmentalize local regions to form a loop. These 'insulated' loops often contain promoters and enhancers. This insulation effect is seen in the action of the CTCF protein, which binds chromatin at distinct CTCF-sites (a 14 bp sequence). These sites are then brought together in three-dimensional space to form a loop through the dimerization of two opposite-facing CTCFs occupying distal sites of the genome (Holwerda and de Laat, 2013; Ong and Corces, 2014; Phillips and Corces, 2009). The formation of this compartmentalized loop promotes the interaction between internal elements, such as promoters and enhancers, while blocking external enhancer elements from interacting with the region (as reviewed by (Dunn and Davie, 2003; Herold et al., 2012)). Long-range interactions instruct upon proper gene regulation and the loss of insulator elements has been tied to disease states caused by the formation of inappropriate long-range interactions. Long-range interactions promote the efficiency of transcription (as reviewed in (Pennacchio et al., 2013), and in one experimental case the addition of an enhancer element resulted in ~13 fold increase in the expression of the collagen II gene in rats (Horton et al., 1987). These observations connecting long-range interactions and transcription validate the importance of 3D organization and chromatin interactions in regards to facilitating gene expression. Furthermore, cellular health and disease may be interpretable by investigating these organizational events.

1.7 3D Organization and Gene Expression

It is clear that the organization of genetic material inside the nucleus provides a function and informs on cellular status. We can currently observe this organization at two experimentally characterizable levels. At the first level, the positioning of chromosomes can be analyzed using FISH. While this technique does not inform directly on changes in gene expression, changes in chromosome

territory positioning occur reliably under repeated treatments and correlate to a change in cellular function and concomitant gene expression alteration. Several avenues have been pursued to understand and characterize chromosome territory positioning and organization, including the analysis of how sister chromatids 'face' each other during interphase (Sehgal et al., 2016). Also significantly, it was observed that the acrocentromeric chromosomes co-localize at the nucleolus where their Nucleolar Organizing Regions (NORs) organize and allow for the localized transcription of rRNA genes and subsequent ribosome biosynthesis. This example indicates that larger organization changes are functionally linked with highly controlled gene transcription.

To characterize the organization of the genome at the level of chromatin interactions and folding, several labs have employed the 3C technique or one of its variants (Barutcu et al., 2016; Dekker et al., 2002; Lando et al., 2018; Tolhuis et al., 2002). This technique probes interactions occurring between regions of chromatin and allows for the correlation of genome organization to gene expression at distinct and definable genomic loci. Specifically, it allows the examination of long-range interactions between two distinct genomic loci. For example, 3C can detect the interaction between a promoter and several putative enhancers regions, or an interaction occurring between two insulating CTCF sites. Additionally, 3C and its derivative techniques have been used to detect interactions occurring between active genes which are thought to be co-localizing during transcription, at sites suggested to be 'transcription factories' (Carter et al., 2008; Eskiw et al., 2010; Eskiw et al., 2008; Osborne et al., 2004; Schoenfelder et al., 2010; Sutherland and Bickmore, 2009; Xu and Cook, 2008).

Transcription factories are accumulations (40-320nm in size) of RNA polymerase II (4-30 RNA Pol II per factory) (Eskiw and Fraser, 2011; Eskiw et al., 2008; Iborra et al., 1996; Rieder et al., 2012). It has been well documented that there are distinct sites of transcription in the nucleus (Eskiw et al., 2008), containing active RNA Pol II as well as nascent RNA transcripts (Kimura et al., 2002). Contrary to the model of transcription whereby RNA Pol II locates to a gene to begin transcription, in the factory model

genes mobilize to these accumulations of RNA Pol II for transcription. It is worth noting that there are fewer of these distinct sites than genes being transcribed (Eskiw and Fraser, 2011). With fewer sites of transcription than genes expressed, seemingly genes must occupy the same factories – and not necessarily at random. Eskiw *et al* (2011) discovered ~60 transcription factories containing the Kruppel-like factor (KLF)-1 transcription factor in erythroid cells. When accounting for the genes regulated by KLF1, they estimated 15-25 genes were being transcribed per factory, albeit transiently.

Another factor to consider about transcription factories is the increase in efficiency of transcription, brought about by concentrating or specializing the necessary elements at distinct nuclear locations. Localizing transcription machinery and appropriate genes increases the frequency and speed of transcription (Sutherland and Bickmore, 2009). Furthermore, it has been observed that genes loop out from their chromosome territory and towards a transcription factory (Branco and Pombo, 2006; Osborne et al., 2007), and additionally that like-regulated genes tend to co-localize at the same transcription factory (Eskiw and Fraser, 2011). Genes localized on different chromosomes have been found producing mRNA at a single shared location (Schoenfelder et al., 2010). Organization of this type occurring in the nucleus is not unfounded. The nucleolus performs an extremely similar act, wherein it gathers genes of related function (rRNA) to a distinct nuclear location at a site containing concentrated RNA Polymerase I. Interestingly, the nucleolus has been suggested to be a prototypical transcription factory (Carter et al., 2002; Cope et al., 2010; Kimura et al., 2002). From this evidence, it is clear that the nucleus is capable of organizing transcriptional events to specific locations.

The organization of genes around a transcription factory is permissive to long-range interactions, where genes contained at the factory are folded in a way that appropriate enhancers are brought into close proximity to gene promoters – awaiting transcription factor and protein mediator binding. Some have argued that regions of chromatin may be structurally folded and ‘poised’ prior to

recruitment to transcription factories containing transcription factors that bridge long-range interactions (T_h2), while others subscribe that localization to a transcription factory and binding of transcription factors drives genome folding (Osborne et al., 2004; Spilianakis and Flavell, 2004; Spilianakis et al., 2005). It has been observed that a knockout of relevant transcription factors has caused organizational loss (Schoenfelder et al., 2010) and additionally that loss of transcription factors did not cause an organization loss (Spilianakis and Flavell, 2004). This contradiction represents a complexity in genome organization, asking the question of exactly what drives organization. Many factors (transcription factors, localization to a factory, RNA Polymerase II, mediator proteins, enhancers and promoters) are involved in genome organization, and complete mechanisms have not been fully explained yet. For example, genes reliant on a trans-enhancer require localization between the promoter and the trans-enhancer. This interaction may be facilitated by localization of both elements to a transcription factory, where the promoter-enhancer interaction is stabilized by transcription factors located there. Conversely, the enhancer-promoter interaction may be formed by transcription factors prior to localization at a transcription factory. Interestingly, the mechanism of transcription by RNA Polymerase II is also still under investigation.

The idea of RNA Polymerase II transcribing DNA by processing along the strand has been challenged (Iborra et al., 1996). Evidence has recently emerged that DNA is instead reeled through polymerases (Papantonis et al., 2010) (**Figure 1-7**). Through FISH assays, Papantonis and Cook identified that two genes*, Sterile Alpha Motif Domain Containing 4A* (*SAMD4a*) and *Tumor Necrosis Factor Alpha-Induced Protein 2* (*TNF-AIP2*), were transcribed in the same location of the nucleus and were both regulated by Tumor Necrosis Factor α (TNF-α) responsive promoters. Both genes were known to be rapidly induced by TNF-α. To challenge the notion that RNA Pol II tracks along templates as transcription occurs, the interaction of *TNFAIP2* to *SAMD4a* during transcription was traced using 3C. The *TNFAIP2* gene is 11 kb in size versus the much larger *SAMD4a* gene which is 221 kb and transcribes less than once

an hour upon induction. These genes reside 50 million base pairs apart on chromosome 14 and when not induced by TNF-α interactions between the two genes were not detectable. By tracing the interaction between the promoter of the *TNF-AIP2* gene versus the different sites along the length of the *SAMD4a* gene, they were able to show that as increasing amounts of time passed following induction the *TNFAIP2* promoter interacted further downstream of the *SAMD4a* promoter. This result suggested that the polymerases were immobile and were instead reeling the coding sequence of both genes through a shared transcription factory rather than tracing along the gene sequence (Papantonis et al., 2010).

Transcription factories display a level of gene and transcription factor specificity. To date several examples of like-regulated genes clustering at transcription factories have been observed. An example already mentioned includes the observation by Papantonis *et al* (2010)., wherein *TNFAIP2* and *SAMD4a*, genes both regulated by *NF-κB* (Nuclear Factor kappa-light-chain-enhancer of activated B cells), localize in a detectable and reproducible manner. Additionally, Osbourne *et al* (2007)., observed several immediate early (IE) genes co-localizing, also driving the rationale that the 3D organization of the genome during transcription was the cause of gene translocations between the *MYC* and *IgH* genes known to happen in Burkitt's Lymphoma (Osborne et al., 2007).

From the pooling of this evidence, there is a strong argument for the existence of transcription factories, and that their function is to concentrate appropriate transcriptional machinery and factors. During transcription, like-regulated (similarly regulated genes) are accumulated at these factories for efficient transcription where necessary factors are concentrated. Furthermore, an evolutionary precedent is set for this by the cluster of acrocentromeric chromosomes at the nucleolus, organized specifically for the biosynthesis of ribosomes from the multiple rRNA genes of different chromosomes (Floutsakou et al., 2013; Pederson, 2011; Pontvianne et al., 2016).

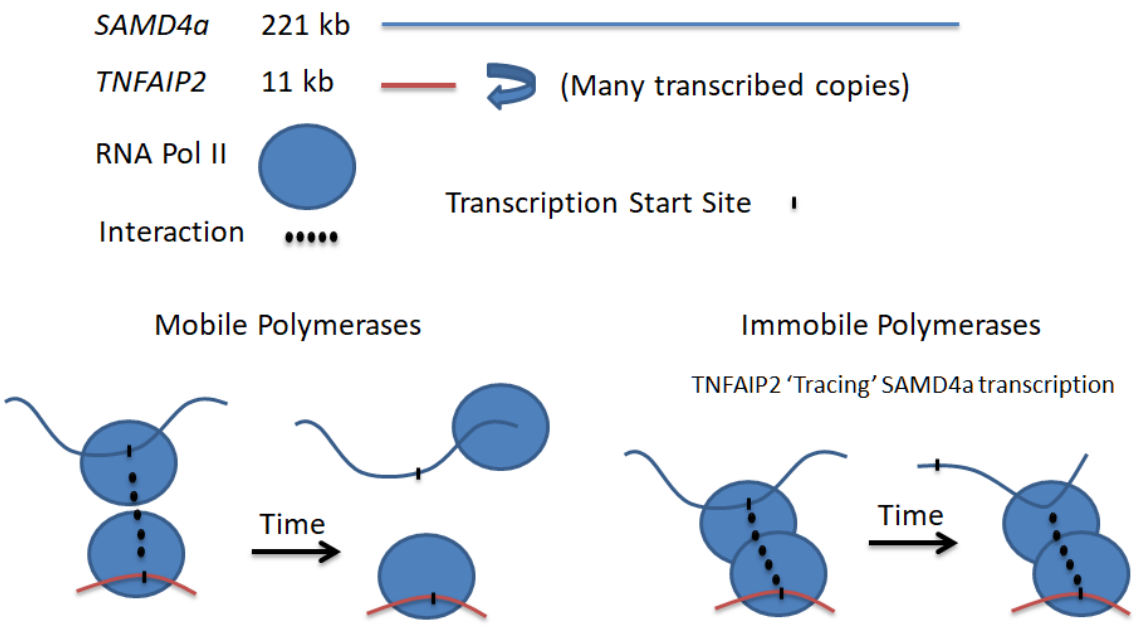


Figure 1-7: Mobile or Immobile Polymerase Investigations using TNF-α and SAMD4a. To understand whether RNA polymerases are mobile or immobile in the nucleus, Papantonis and Cook (2010)., investigated the interactions occurring between two genes known to be transcribed at the same time and in proximity to each other. By measuring the possible interactions occurring between these two genes, *Sterile Alpha Motif Domain Containing 4A* (*SAMD4a*) and *Tumor Necrosis Factor Alpha-Induced Protein 2* (*TNF-AIP2*), following activation of transcription by TNF-α. By tracing the transcription of the long *SAMD4a* gene (221 kb) using the nearby, quick and repeated transcription of TNFAIP2 (11 kb), Papantonis and Cook (2010)., provided evidence that RNA Polymerase II was immobile and suggested DNA was 'reeled' through.

1.7.1 Organization of the β-globin Locus

The investigations into β-globin locus represent one of the earliest examples of genome organization in a mammalian genome (Carter et al., 2002; Tolhuis et al., 2002). In their paper, Tolhuis *et al* (2002)., investigated the interaction between DNAse hypersensitivity sites (exposed DNA that are

digestible by DNAse I) occurring in the locus control region (LCR) and the head-to-tail cluster of *globin* genes in the globin locus. Resultantly, they discovered which sites of DNAse hypersensitivity (HS) had higher levels of interaction with the active globin gene, and observed a lack of interaction between these sites and the non-expressed *globin* genes in erythroid cells. It was also observed that the locus did not form long-range interactions to promote *globin* expression in cells that do not express these *globin* genes. These experiments displayed 'looping out' of inactive *globin* genes. They also demonstrated the preferential and detectable long-range interactions between distinct HS elements of the LCR and active *globin* genes. Furthermore, it was shown that topological organization is cell-type dependent, as the non-expressing brain cells had shown no significant long-range interactions in these instances.

These early investigations into the *globin* locus quickly expanded into other questions of topology, such as the function of CTCF in the globin locus (Splinter et al., 2006). CTCF insulates regions of the genome and prevents atypical interactions between enhancers and promoters. In the investigations by Splinter *et al* (2006)., CTCF binding sites were identified and were known to contact each other prior to *globin* expression in erythroid progenitor cells. It was discovered in the course of these experiments that a homozygous knockout of the CTCF protein in mice was embryonic lethal, displaying the significance of this protein, and a mutation of the CTCF binding sites resulted in a loss in long-range interactions between CTCF sites and of the regulatory interactions of the *globin* genes to the LCR. The authors suggest that CTCF is essential for establishing early organization of the locus, however that it is not essential as differentiation progresses. Further investigations into this region have highlighted specific histone modifications at the LCR, when interactions are formed between the *globin* locus and loci on other chromosomes (Miles et al., 2007). As inter-chromosomal interactions require spatiotemporal organization of chromatin, these investigations highlight the role of genome organization for gene expression (Pink et al., 2010). Overall the investigations into the *globin* locus

highlighted several aspects of genome organization, including the identification of specific regulatory elements in the LCR and the capacity of the genome to utilize the same enhancer on different genes as development progresses.

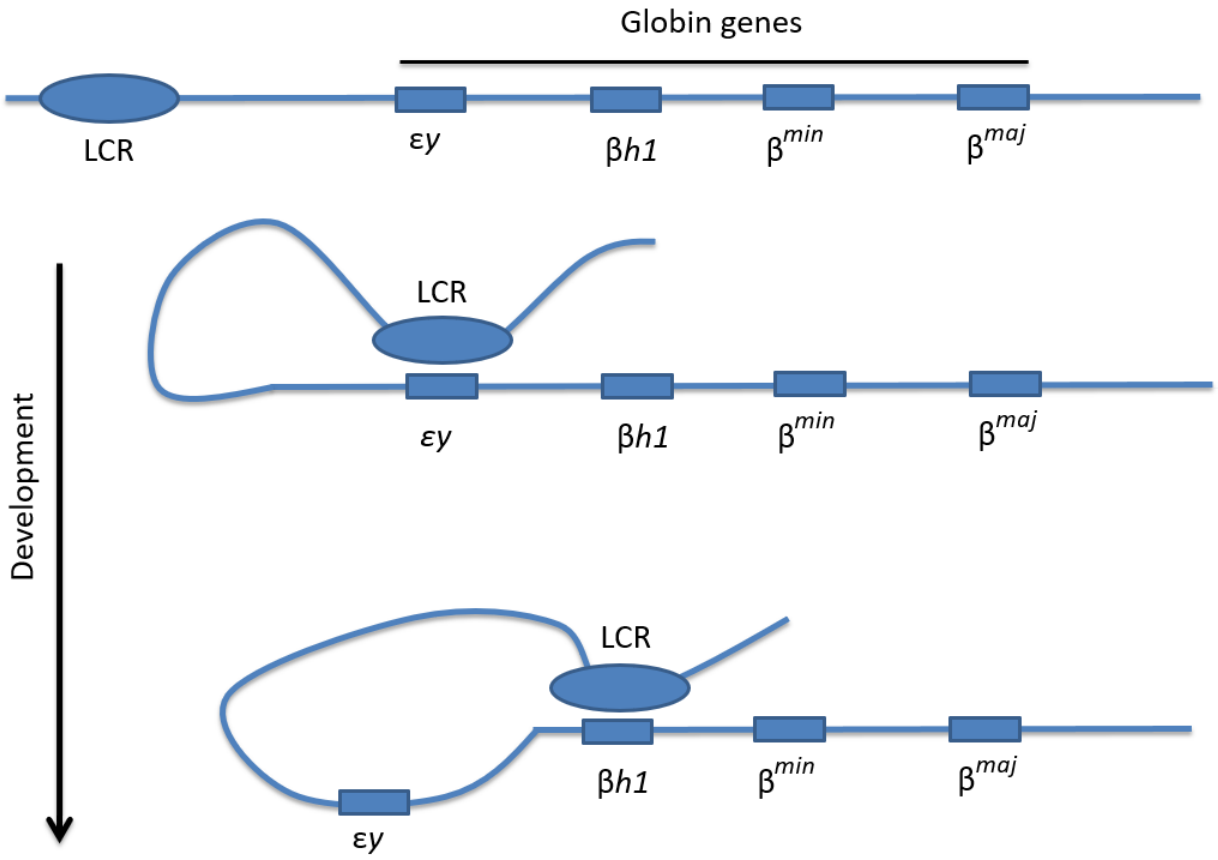


Figure 1-8: Organization of the β-globin locus in mice. Schematic illustrating the β-globin locus contained on chromosome 7 in mice. The expression of the contained globin genes is regulated by a nearby Locus Control Region (LCR). As development progresses the appropriate globin gene interacts with the HS2 site inside of the LCR to promote gene expression. Earlier genes are looped away from the LCR when no longer being transcribed and later expressing genes become bound to the LCR.

1.7.2 H Enhancer/ODR

Olfaction is a vital process for many eukaryotic species. The versatility in sense of smell stems from a genomic organization process, wherein the H enhancer selects from one of many *Odorant Receptor (OR)* genes to express as the olfactory neurons develop (Lomvardas et al., 2006). In mice, this requires the H enhancer (chromosome 14) to operate in *trans* to activate odorant receptor genes (1300 genes) located on other chromosomes, bypassing the odorant genes located downstream of itself. Upon selection of odorant gene, the second allele of the H enhancer is methylated and prevented from forming interactions with odorant genes. These observations that an enhancer can act in *trans* suggests that a level of chromosome territory organization must occur to be permissive for such interactions, and additionally that enhancers do not necessarily regulate nearby genes, but rather have the ability to bypass local genes to regulate genes on other chromosomes (Savarese and Grosschedl, 2006).

1.7.3 V(D)J Recombination

Variable (V), Diversity (D), and Joining (J) recombination requires genomic organization to successfully occur. During V(D)J recombination, T Cells and B Cells undergo somatic recombination to create a novel gene from coding sequences at different loci that they and their daughter progeny will express. The process of V(D)J recombination is sophisticated and requires precise three-dimensional organization of its genetic components to successfully occur, wherein distally located coding sequences are brought into close proximity and recombined into a novel product. Each of these genes (VDJ) contains a Recombination Signal Sequence (RSS) up-stream and downstream of the coding sequence that orchestrates several enzymes, notably Recombinase Activating Genes 1 and 2 (RAG1/2), to target the strand for nicking and subsequent recombination. Several labs have ascribed the CTCF sites in these regions as key parts of recombination process, due to their ability to organize the region by forming chromatin loops (Chaumeil and Skok, 2012; Frock et al., 2015; Guo et al., 2011; Jung et al., 2006; Lin et

al., 2015). Vitally, transcription is required for V(D)J recombination to occur. When transcription was inhibited this recombination did not occur (Bolland et al., 2007). As V(D)J recombination provides the ability for an adaptive immune system, it represents a significant example of how genome organization translates into a larger function.

1.8 Topologically Associating Domains

1.8.1 Overview

Topologically associating domains (TADs) represent a level of organization of the genome. These regions consist of sections of chromatin roughly 100 kb to 1 Mb in size. TADs cover a considerable portion of the mammalian genome and are estimated at 91% coverage in the mouse genome (Dixon et al., 2012). The borders of a TAD are defined by CTCF binding sites and housekeeping genes, as well as other elements, that isolate internal elements (enhancers, genes) from interacting with external features while simultaneously promoting internal interactions.

TADs are a conserved element of mammalian genomes and similar topological patterns have been observed between the *HoxA (*homeobox A cluster*)* locus in both mice and humans (Dixon et al., 2012). There is a high conservation of TAD boundary marks for when comparing differentiated tissues of the same species, as well as comparing between humans and mice. Conservation of such a feature indicates the importance of TADs and their role in gene regulation, as well as the organization of the genome on a whole.

The function of compartmentalization provided by TADs allows internal elements to operate independently of the local environment. Boundary elements of TADs have been ascribed roles in stopping the spread of heterochromatin (as reviewed by (Burgess-Beusse et al., 2002)). It has been observed that TADs are able to maintain distinct chromatin environments (euchromatin or

heterochromatin) from their neighbors (Pope et al., 2014). For example, a euchromatinized TAD may be positioned between two heretochromatinzed TADs. This chromatin state, which is maintained separately from TADs neighbours, delineates TADs as functional units of the genome with an independent and maintained environment.

Compartmentalization of TADs also provides roles during DNA repair. During double-stand break (DSB) repair, the histone variant responsible for marking damaged strands, H2AXγ, marks the TAD from the damaged site and up to the boundaries (Aymard et al., 2017; Marnef and Legube, 2017). In this 'intra-TAD model', H2AXγ integrates bidirectionally from the induced damage until reaching the TAD boundary. Following this marking by H2AXγ, the DSBs co-localize to distinct sites of the nucleus for repair. Alternatively, unrepairable DSBs may be localized to the periphery of the nucleus. The precision of H2AXγ integration further demonstrates the existence of TADs in the genome, as well as providing a specific example of a function not possible without their presence.

A growth issue in mice has been directly linked to aberrancies in TAD boundaries, observed in a boundary element deletion resulting in malformed digits (as reviewed in (Lupianez et al., 2015)). F-syndrome, the fusing of the index finger and thumb in humans, was attributed as the result of an inversion of a boundary element flanked by several enhancers – causing enhancers to act abnormally on another developmental gene *Wnt6* (*Wingless-type MMTV integration site family, member 6*) located in the neighboring TAD (Lupianez et al., 2015). Additionally, in humans, the deletion of a TAD boundary element results in nearby enhancers acting abnormally upon the *Lamin B1* (*LMNB1)* gene, inducing adult-onset-demyelinating leukemia (Giorgio et al., 2015). These examples display the importance of maintaining TAD structure and the controls TADs convey on compartmentalizing long-range interactions that (along with transcription factors) regulate gene expression.

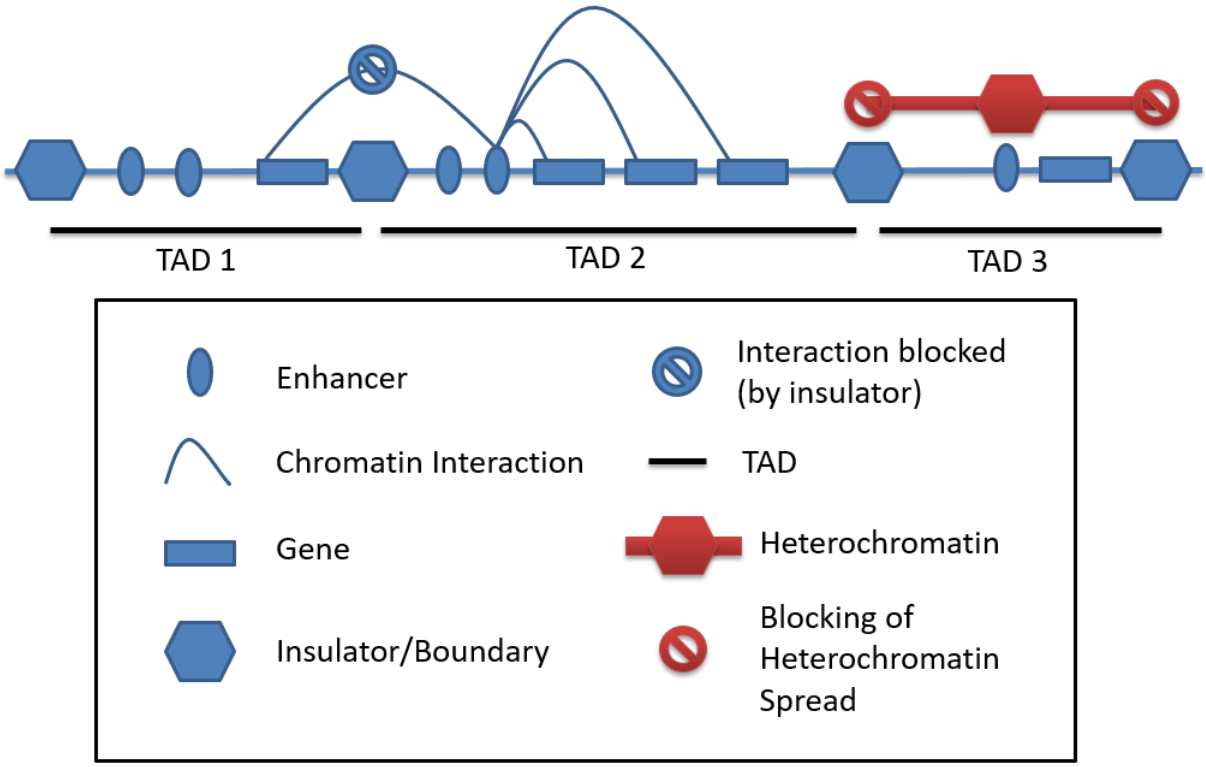


Figure 1-9: Topologically Associated Domains. Schematic of the organization of three TADs containing several features. TADs compartmentalize regions of the genome and facilitate internal interactions while blocking external interactions. TAD boundaries stop the spread of heterochromatin and maintain chromatin states unique from neighboring TADs.

1.8.2 LADS, MADS, and NADs

Several other forms of domains exist, namely Lamina Associated Domains (LADs) sometimes also called Matrix Associated Domains (MADs), and Nucleolar Associated Domains (NADs). LADs are regions of the genome with low gene density and activity, spanning 10 kbp to 10 Mbp in size, that locate to the nuclear periphery and tether at the nuclear matrix (van Steensel and Belmont, 2017). By tethering chromosomes at the periphery of the nucleus, it is thought that LADs contribute to chromosome

topology and gene expression by providing an anchor point. A weakening of this binding caused by hypomethylation of a LAD has been observed in colon cancer (van Steensel and Belmont, 2017) – potentiating altered chromosome topology and subsequent mis-regulation of genes through a loss of an organizational element. NADs are regions of heterochromatin containing inactive genes as well as the rRNA genes that associate with the nucleolus (Pontvianne et al., 2016). LADs and NADs are further examples of genetically programed elements of nuclear organization and further signals that the organization of the genome is an important and conserved process.

1.9 Summary

RNA-Seq performed on proliferative, quiescent, and Rapamycin treated 2DD cells in our lab revealed a region of chromosome 4 containing several up-regulated genes at the 74 MB to 76 MB loci. As these growth conditions affect gene expression profiles, a question remained as to how the organization of the genome alters in these conditions, especially the up-regulation of the cluster of chemokine genes on chromosome 4, region 74 Mbp. Previous work in the Eskiw lab has displayed altered genome organization at the level of chromosome territories in these conditions. However, a deeper (gene-interaction) analysis has yet to be performed to show if alterations to organization are occurring at sub-chromosomal levels and are responsible to changes in gene expression profiles. The question is also raised for if chromosomal repositioning alters or facilitates sub-chromosomal organization, during our treatment conditions. To accomplish this level of analysis the technique 3C is necessary; however, its use has yet to be established in the Eskiw lab. To expand the capabilities of the Eskiw lab as well as investigate the region of interest on chromosome 4 and other potential targets this project will seek to employ and improve the 3C technique in accompaniment with traditional fluoresce *in situ* hybridization (FISH) microscopy to characterize changes in genome organization induced by

Rapamycin treatment.

1.10 Hypothesis

Based on the observations of the Eskiw laboratory and those from the literature, I hypothesize that changes in growth conditions, moving from proliferative growth to quiescence or under conditions of Rapamycin treatment, will result genome re-organization. Specifically, I hypothesize that chromosome 4, which contains a cluster of *chemokines* genes that dramatically change expression in response to Rapamycin, will change position within the nuclear volume. I further hypothesize that specific sub-regions of chromosome will undergo re-folding in response to these transcriptional changes resulting in new long range chromatin interaction forming and the disruption of others.

1.11 Objectives

To confirm these hypotheses, the following objectives were planned:

1) Identification of reliable internal controls for 3C.

2) To confirm that chromosomes are reorganizing and; therefore, there is a change in genome folding using chromosome painting by DNA FISH

3) Confirmation of a specific interaction computationally theorized to occur from RNA-Seq datasets, specifically a dataset found on Chromosome 4, region 74 Mb.

4) To utilize an optimized 3C protocol to identify specific changes in long range chromatin interactions occurring at other locations in the genome, in divergent growth conditions.

5) Perform Hi-C to model 3D spatial conformations in collaboration with Dr. Anthony Kusalik.

2.0 Materials and Methods

2.1 Cell Culture, Cell Counting

Cell Lines

Three cell lines were used in all outlined experiments and all were resuspended prior to reaching 70% confluency. Primary human foreskin fibroblasts (designated 2DD) were used during FISH and 3C experiments and were harvested between passage 14 and passage 17, with a confluency no greater than 70% (Bridger et al., 1993). NB1 hTERT were derived from a parental line (human fibroblasts) and immortalized by expression of human telomerase (Mehta et al., 2010). Human Umbilical Vein Endothelial Cells (HUVECs) (Lonza, USA, Cat #: CC-2517) were harvested between passage 4 and passage 7 and at a confluency no greater than 70%. NB1 hTERT and HUVEC lines were used in the initial optimization of the 3C protocol.

Cell Culture

NB1 hTERT and 2DD cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, L-glutamine & sodium pyruvate (Corning, USA, Ref: 10-013-CV) supplemented with 10% fetal bovine serum (FBS) (Gibco, Thermo Fisher Scientific, USA, Ref: 12483-020) and 1% penicillin-streptomycin (Hyclone, GE Healthcare Life Sciences USA, Ref: SH40003.01). HUVECs were cultured in EGM-2 SingleQuots (Lonza, USA, Ref: CC-4176) containing the full supplements provided and 1% penicillin-streptomycin. Cultured cells were resuspended prior to reaching 70% confluency using TrypLE^{TM} Express (Gibco, Life Technologies, USA, Ref: 12604-021). Cell counts were performed by hemocytometer and passaged cells were seeded at a density of 3000 cells/cm^2 (225,000 cells per T75 flask, 550,000 per 15 cm tissue culture plate). For FISH assays, cells were cultured in 6-well dishes containing 22mm X 22mm glass coverslips and removed with forceps prior to experimental processing.

Cell Treatments

Cultured cells were induced into quiescence by serum deprivation (0.5% FBS) in culture media. Rapamycin treatments occurred in standard culture media (DMEM, 10% FBS, 1% penicillin-streptomycin) at a concentration of 500 nM Rapamycin (Fischer BioReagents, Israel, Code: BP2963-1). Rapamycin was dissolved in DMSO (Fisher Chemical, USA, D128-4). Final Concentration of DMSO during Rapamycin treatment was less than 0.001%. Cell treatments occurred for 5 days, with media refreshment on the third day, prior to harvesting. Prior to harvesting for 3C, HUVECs were treated with 10 ng/ml Tumor Necrosis Factor- α (TNF-α) (Abcam, USA, Ref: ab9642) for 30 minutes.

2.2 Electrophoresis

2.2.1 Agarose Gels

To verify PCR products, a 2% agarose gel (Melford, USA, cat #: MB1202) 1x TAE (40 mM Tris, 20 mM acetic acid, 1 mM EDTA)) (Tris-Base: Fisher, USA, Cat #: BP154-1; EDTA: Boehringer Mannheim GmbH, Germany, Cat #: 808-288) containing ethidium bromide (Sigma Aldrich, USA, Cat #: E1510-10ML) was used. For evaluating the digestion and ligation of 3C samples a 1% agarose gel was used. Samples were diluted in 6X Orange G DNA loading buffer (Sigma Aldrich, USA, Cat #: 1936-15-8) and separated in 1x Tris-Acetate-EDTA (TAE) buffer at 100 volts (constant) for 30 minutes. 3C library digestion efficiency was analyzed by separating samples on a 1% agarose gel with otherwise identical conditions. Agarose gels were visualized on a Benchtop UV Transilluminator BioDoc-It^{TM} Imaging System.

2.2.2 Tris-Borate Gels

To verify the size of PCR products (50-250 base pairs) a 6% Tris-Borate-EDTA (TBE) (6% Bis-acrylamide, 0.1% Ammonium persulfate (Fisher, USA, Cat #: BP179-25, 0.01% TEMED, 0.5x TBE buffer) gel was separated at 200 volts (constant) for 20-25 minutes in TBE buffer (1x TBE 100 µM Tris-Base, 100 µM Boric Acid (EMD, USA, B10058), 1 µM EDTA). Samples were diluted with 10X TBE loading buffer (45 µM Tris-base, 45 µM Boric Acid, 50% glycerol (Fisher Chemical, USA, cat #: G33-1), 3.5 µM bromophenol blue (Fisher Biotech, USA, cat #: BP115-50), 4.5 µM xylene cyanol FF) (Sigma Aldritch, USA, Cat #:2650-17-1) prior to being separated. Following separation gels were immersed in water containing 0.001 M ethidium bromide for 10 minutes before being visualized on a Benchtop UV Transilluminator BioDoc-It™ Imaging System.

2.2.3 Sodium Borate Gels

To verify the size of PCR products (50-250 base pairs) a 6% Sodium-Borate (SB) (6% Bis-acrylamide, 0.1% Ammonium persulfate, 0.01% TEMED, 1x SB buffer) gel was separated at 250 volts (constant) for 12-16 minutes in 1x SB buffer. Samples were diluted with 10x TBE loading buffer (45 µM Tris-base, 45 µM Boric Acid, 50% glycerol (Fisher Chemical, USA, Cat #: G33-1), 3.5 µM bromophenol blue (Fisher Biotech, USA, Cat #: BP115-50), 4.5 µM xylene cyanol FF) prior to being separated. Following separation, gels were immersed in water containing 0.001 µM Ethidium Bromide for 10 minutes before being visualized on a Benchtop UV Transilluminator BioDoc-It™ Imaging System.

2.3 DNA Fluorescence *in situ* hybridization and Erosion Analysis

Fluorescence *in situ* Hybridization

Cultured cells were grown directly on glass coverslips, and washed with PBS, prior to fixation with ice-cold 3:1 (v/v) methanol-acetic acid (Methanol; BDH, Germany, cat #: BD1135-4LP, Acetic acid;

Fisher Chemical, USA, cat #: BP2401-212). Cells were then incubated for an hour at 4°C, and washed five times with 3:1 (v/v) methanol: acetic acid. Fixed cells were dried at 70°C for an hour, prior to a dehydration series of 5-minute ethanol washes (in order: 70%, 90%, 100% ethanol). Cells were then incubated at 70°C to complete the drying process. DNA of fixed cells was denatured by treating with 70% formamide (Alfa Aesar, USA, Cat #: A11076), and 2X saline sodium citrate (SSC) (Fisher, USA, Cat #: BP1325) for 2 minutes at 70°C. Cell then underwent a second dehydration series in ice-cold ethanol. Chromosome-specific DNA probes were generated in-house by amplifying chromosome specific sequences from flow-sorted chromosomes using degenerative oligonucleotide primers. In result, 300-500 nt (chromosome-specific) templates were generated by PCR with biotinylated-uracil residues incorporated for fluorescent signal amplification. For each coverslip to be labelled, 200-400ng of chromosome-specific DNA probes were ethanol precipitated with 3.5μg and 3μg herring sperm DNA and resuspended in 12.5 μl hybridization buffer (50% formamide, 4X SSC, 50mM sodium phosphate and 20% (w/v) dextran sulfate). Probes were denatured at 90°C for 30 min and renatured by incubation at 37°C for 20 mins before being added to fixed coverslips. Coverslips were sealed onto glass slides using rubber cement and incubated for 2 minutes at 60°C, prior to being placed in a humidity chamber for 12 hours at 37°C. Coverslips were then removed from glass slides while being washed in 50% formamide, 2X SSC pH 7.0, at 45°C. Coverslips were then washed several times with 0.5X SSC pH 7.0 at 60°C, then stored in PBS at room temperature. Directly labeled probes (Chromosome 4: Cytocell Aquarius, USA, cat #: LPP04G) were then mounted in Hoechst 33342 (Invitrogen, USA, Cat #: h3570). Coverslips with biotinylated probes were incubated with antibodies at room temperature for an hour, before being washed with 0.05% Triton X-100.The order of antibody addition was as follows: Streptavidin Cy3 (1:200; Vector laboratories, USA, cat #: Ba-0500), goat anti-rabbit biotin (1:200; Cedarlane, UK, Cat #: 111-065-003), Streptavidin Cy3. After treatment with antibodies, coverslips were washed with PBS and mounted in Hoechst 3342.

Fluorescence *in situ* Hybridization Image Capture and Image Processing

Images of nuclei were captured using a Nikon Y-IDP with an X-Cite fluorescence light source and a Nikon Digital Sight DS-U3 camera. Nuclei images were collected using three channels; Red - DOP Probes (Chromosomes X, 10, 18), Green - Direct-labelled probes (Chromosome 4), and Blue (Hoescht 33342). Resulting images were overlaid together using Adobe Photoshop (CS6) for use in Erosion Analysis.

Fluorescence in situ Hybridization - Erosion Analysis

Using Cell Image Analyzer software, images of nuclei were divided into five concentric rings of equal area, numbered one (periphery) to five (most internal) and the signal from each chromosome was measured and standardized to the Hoechst 33342 signal (stains DNA). The resulting data was plotted in bar graphs depicting standardized chromosome signal intensity per concentric ring or 'shell'. Signal intensity of each chromosome was normalized to the quantified background DNA in each shell as detected by Hoechst 33342. The resulting data were tested using a 2-tailed Student's T-test with unequal variance (Microsoft Excel) to demonstrate significance in the repositioning of chromosomes.

2.4 Chromosome Conformation Capture

Five-million cells were harvested for each replicate of 3C **(Figure 2-1: A)**, for each treatment condition. Cultured cells were harvested as follows: proliferative cells were harvested upon reaching 70% confluency, quiescent induced cells were harvested after 5 days of treatment with culture media containing only 0.5% FBS, Rapamycin treated cells were harvested after 5 days of treatment with 500 nM Rapamycin. Culture media was changed every 2-3 days to ensure that rapamycin was not depleted. Cells were suspended and fixed with 1% formaldehyde (Electron Microscopy Sciences, USA, Cat #:

15714) for 10 minutes at 37°C to capture three-dimensional chromatin contacts. Fixation was quenched with 2.0 M glycine (Fischer BioReagents, USA, Cat #: BP381-5) to a final glycine concentration of 125 mM and incubated at room temperature for 10 minutes. Cells were then washed in PBS to remove excess glycine. Cells were lysed for 10 mins at 65°C, in a solution containing 10 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1% SDS containing 1 X protease inhibitor cocktail. Following this incubation cells were chilled on ice for 5 min, before being diluted with a 1X *HindIII* restriction enzyme buffer (NEB Buffer 2.1). Samples were digested using 600 U of *HindIII* (NEB, USA, cat #: R0104M) in 1X 2.1 NEB Buffer (500 µl total volume) overnight at 37°C, 950 RPM. *HindIII* was heat inactivated by incubating samples at 80°C for 20 minutes. Samples were ligated with 1600 U of T4 DNA Ligase (NEB, USA, cat #: M2020L) in 1X T4 DNA Ligase buffer (1.0 ml) for four hours at 16°C, 950 RPM. Following ligation, samples were incubated overnight at 65°C with proteinase K (Invitrogen, USA, 100053) to remove cross-linked proteins. Samples were then treated with RNAse A for 1 hour at 37°C. Samples were extracted by equal volumes of 1:1 phenol:chloroform and 19:1 chloroform:isoamyl to remove proteins. DNA was precipitated using 1 µl glycogen (20 mg/ml – Roche. Cat #: 10-901-393-001), 10% volume NaOAc (Fisher chemical, USA, cat #: S208) and equal volume of isopropanol, and incubated for 20 minutes on ice. Following centrifugation, the resulting pellet was washed with 70% ethanol and air dried until resuspended in Tris-EDTA (TE) buffer (10 mM Tris-HCL pH 8.0, 0.1 mM EDTA).

Initial success of 3C assays were tested in the HUVECs (**Figure 2-1: B**) cell line recapitulating earlier work by Papantonis and Cook (Papantonis et al., 2010), wherein an interaction between *TNFAIP* and *SAMD4a* was detected by 3C. HUVEC cells were either untreated (Pro) or treated with TNF-α (10 ng/ml) for 30 mins (TNFα #1 and TNFα #2) prior to collection of cells.

To analyze the digestion and ligation quality of a 3C library, each sample is separated on a 0.8%-1.0% Agarose gel **(Figure 2-2)**. For each library a digestion efficiency (DE) and 3C library (3C) are separated and analyzed for adequate digestion in the DE sample, and adequate ligation in the 3C

sample. Adequate digestion is observed as a smear on the gel below the 10 kB shear limit of DNA. Adequate ligation of 3C libraries is observed in the formation of a '3C band' – a band occurring at a slighter higher height than the DNA shear limit band. 3C data was generated by qPCR amplification of primer pairs that capture potential interactions. The resulting data was analyzed for significance by 2-tailed Student's T-tests with unequal variance (Microsoft Excel).

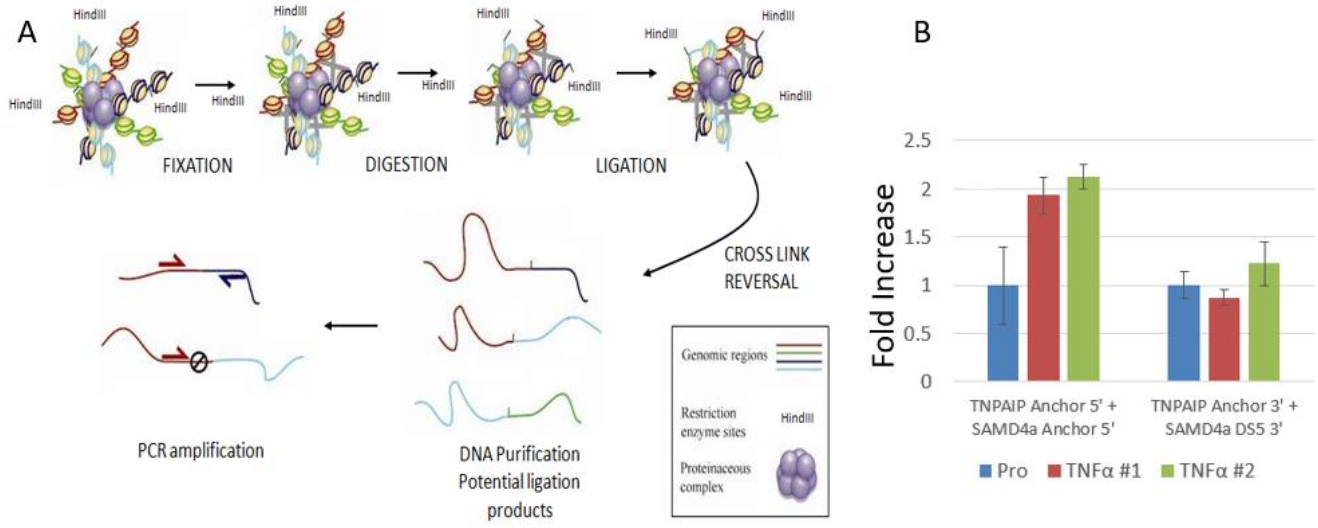


Figure 2-1: Chromosome Conformation Capture and Initial 3C assays in HUVECs. (A) Schematic of the 3C assay. Cells are fixed in formaldehyde prior to isolation of nuclei. Samples are then permeabilized and chromatin is digested using a restriction enzyme. Following digestion, ligation using T4 DNA ligase is performed in a large volume to favor intramolecular interactions. Resulting ligated products are purified and used in qPCR. **(B)** Initial 3C assays recapitulating interactions between *TNFAIP* and *SAMD4a* observed in earlier work by Papantonis and Cook (2010). HUVECs were either untreated (Pro) or treated with 10 ng/ml TNFα 30 mins prior to harvest. TNFAIP Anchor 5' and SAMD4a Anchor 5' detects an interaction occurring between the promoters of both (*TNFAIP2* and *SAMD4a*) genes. TNFAIP Anchor 3' and SAMD4a 3' detects an interaction occurring between the promoter of *TNFAIP2* and a region downstream of the promoter of *SAMD4a*. (Pro = Proliferative cells, TNFα #1 and TNFα #2 = Cells treated

with 10 ng/ml TNFα)

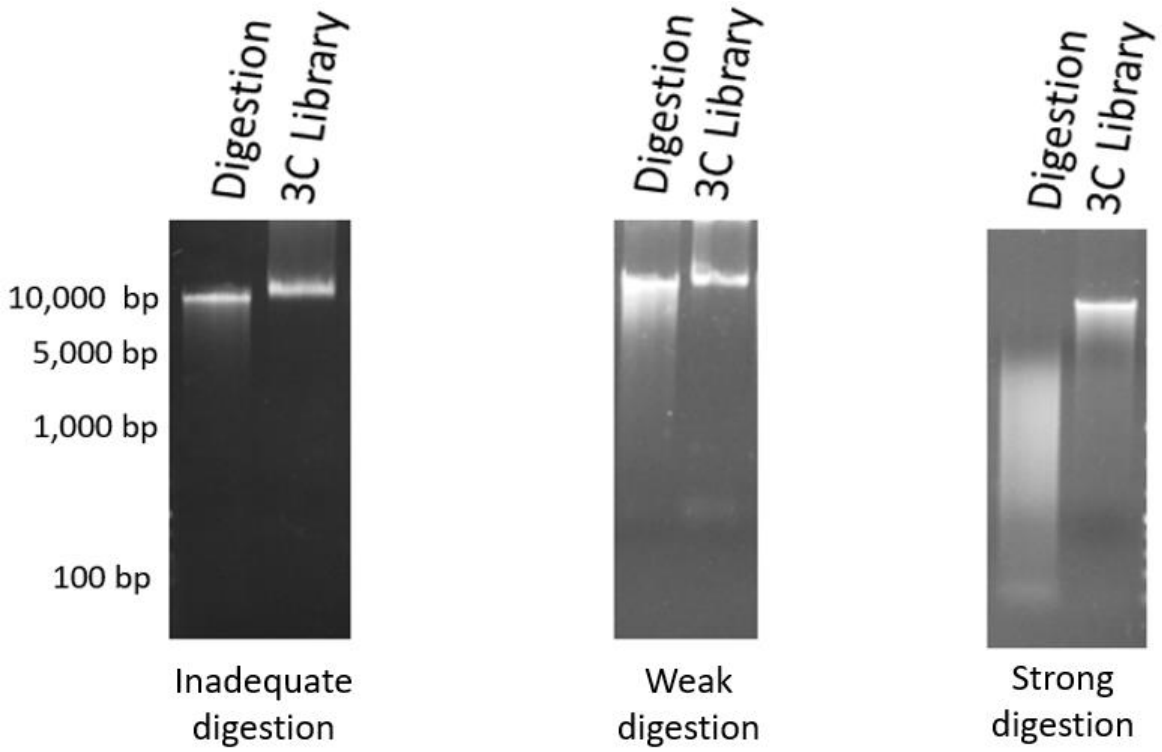
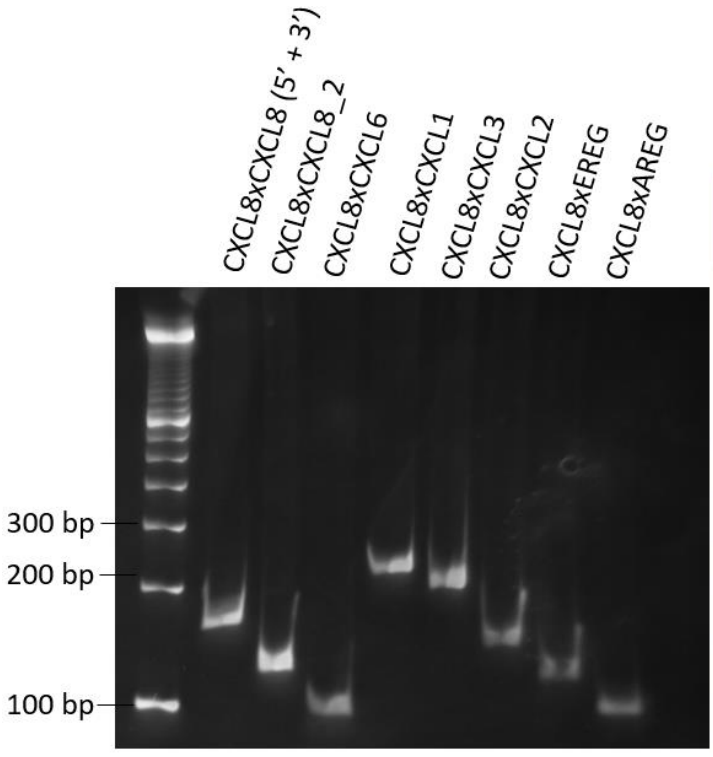


Figure 2-2: Analysis of 3C Digestion and Ligation. 3C samples were separated by electrophoresis and visualized using a 1.0% agarose gel by staining with ethidium bromide. Samples collected after digestion with *HindIII* were marked as 'Digestion' and completed 3C libraries are marked as '3C Library'. Digestion samples of good quality display a smear of DNA along the length of the agarose gel. Good quality 3C libraries display a concise band that separates at similar height to a 10,000 bp marker.

2.5 PCR and Synthetic Libraries

Synthetic libraries were generated to test the ability of 3C to detect products without preforming the entire protocol (**Figure 2-3**). Primer pairs for *HindIII* cut sites of interest were used to amplify genomic fragments using standard PCR reaction conditions (300 nM primers, 0.025 U/µl Taq

polymerase, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 50 ng/reaction genomic DNA) using 35 cycles of 95°C, 60°C and 72°C (PCR Mix; Thermo Scientific, USA, cat #: K0171). The appropriate primer pairs (matching the genomic location of interest) were used and equal volumes of reaction products were combined into a pooled volume. This mixture was digested with 500 U of *HindIII* enzyme for 2 hours at 37°C. Following digestion *HindIII* was heat inactivated by incubating at 65°C for 10 minutes. The resulting digested DNA was diluted into a solution containing a 1X buffer of T4 DNA ligase buffer and 400 U of T4 DNA Ligase. Samples were incubated at 16°C for 4 hours. Reaction mixtures were purified by phenol:chloroform:IAA extraction and precipitated using 10% volume NaOAc and 2.5x volume of ethanol and incubated for 2 hours at -20°C. To test if we could detect putative interactions, 10 ng of template were used in PCR reactions to determine if primers from different pairs (paralleling combinations used in 3C) could amplify products.



Interaction	Observed Size	Expected Size
CXCL8xCXCL8 (5'+3')	180	158
CXCL8xCXCL8 _2	140	140
CXCL8xCXCL6	110	111
CXCL8xCXCL1	240	214
CXCL8xCXCL3	220	197
CXCL8xCXCL2	160	145
CXCL8xEREG	140	122
CXCL8xAREG	100	114

Figure 2-3: Construction of Synthetic Libraries. To confirm if desired 3C products could be produced by qPCR, a synthetic library was produced using PCR to amplify genomic regions for targets of interest (*CXCL* locus, chromosome 4). A 1 kB DNA ladder (Invitrogen, USA, Cat #: 10488090) was used for comparison. Amplified genomic regions were digested using *HindIII*, prior to ligation with T4 DNA ligase. This digestion and ligated yielded a synthetic library containing possible 3C products. The synthetic library was then tested against combinations of primers used for 3C for the *CXCL* locus on chromosome 4. Resulting PCR products were separated on a 1% agarose gel and sized. Correct product formation was evaluated by the resulting size of products by comparing expected and observed sizes.

2.6 Quantitative PCR

qPCR reactions were performed in triplicate on the Rotor-Gene Q (Qiagen) qPCR machine using SYBR® Select Master Mix (Applied Biosystems, USA, Ref: 4472908), template DNA, and 300 nM of both forward and reverse primers in 10 µl reactions. Reactions occurred for 40 cycles under the following profile conditions: 95^{o}C for 5 minutes, 95^{o}C for 20 seconds followed by 60^{o}C for 30 second (repeated for 40 cycles), then held at 50^{o}C for 30 seconds prior to Melt Analysis by the software. Results were analyzed using the Rotor Gene Q series software. PCR products were analyzed by melt peaks produced by the software and size identification by separation on a 6% Tris-Borate-EDTA (TBE) or 2% Agarose gel. Following slope correction CT values of products were measured using a standardized threshold of 0.01, ignoring the first 15 cycles.

2.7 3C Primer Design

In designing of 3C experiments primer pairs were chosen that annealed upstream (5') of a restriction site and downstream (3') of the restriction site. In **Table 2-1** below these are referred to as

forward (5') and reverse (3'). Downstream (3' or reverse) primers capture the restriction fragment of interest that is used to detect occurrence of interactions. Primers designed against *EFEMP2* were utilized to quantify the DNA and normalize loading of 3C library during qPCR assays.

Primer Name	Forward	Reverse
CXCL8	TTTGGACTTAGACTTTATGCCTGAC	ACCCAGTTTTCCTTGGGGTC
Range Finding Site 1	GCCACAAAGACATCCTTAGAATTTAG	ATTCTGACAAATTGAACATATGCC
Range Finding Site 2	GAGCTTCTGGAACCTTACAACC	GGGGGCAAGTACAGCTAATAGTC
Range Finding Site 3	CAAAAGGCTGTCTAAGGCAGAG	TACTGGCAGGGGCAAACT
CXCL6	CTTCTTTCCACACTGCCCCC	CCGACTGGCAGGAAACTCAA
CXCL1	CACACTCAAGAATGGGCGGA	GGGGCACATGCCAGTATTTC
CXCL3	CCTAGAAAGCTGCTGTTCTCTT	TCCAAGCAACTTCAGAGTGAC
CXCL2	GAAGACTTCTCCTAAGTGATGCTC	GCACCCATTTTCTCATTACAG
EREG_1	ACGCACCCTAGGCCAATAAG	ACCTACCGTGGGGGATAATG
EREG_2	TTTCCTGAGCTAAATGCTTCATTG	TCTTGAGCTATGACCTGGGA
AREG	TGACAAGGCGTGTCTTCTTGA	ACGTGCGGTGTGTTTTAAGC
GAPDH P	AGGGTGCAGCTGAGCTAGG	AACACCCCCAGTCATACGAA
GAPDH E	TGTTTGCCAATCTCCTTGTTT	TCTCTGCTGCATGAAACTGG
LIF Anchor 1	CGCTCCCTGCAGCAGGACAA	TACCCCTGGGATACTGACAGGAGA
LIF Anchor 2	GCCCTACCCACCTCCACAAA	AGGAGGTCACTTGGCATTCAGG
LIF Downstream Site 4	CCGGACTTTCCCCGGTTTT	GTGGGCAGTGAGGGAAGTGAG
LIF Downstream Site 3	AGCCATAGCATGGGCTGGC	CCTTGGGATGGGGTTTTACATATGGGATA
LIF Downstream Site 2	CCCACCTCATGGCATTTGCAGA	CACAGGCACAAACTGGGGCT
LIF Downstream Site 1	ACCCAGGACTTAGCTGTCATTCTCT	AGTCTCCAGGAGGCTGGCTC
LIF Upstream Site 1	GATGTGCCCATGTCCCCCA	TGGATGTCAGTGTCCTAATGGAGGC
LIF Upstream Site 2	CCCAGTAGTATGTGGACTAGTGGCC	CCCGGCCCCAAAGAACCTTTT
LIF Upstream Site 3	GCCTGGTGGCATTGAAACTTCC	CCACTGCAGGCCTTAGTTTCCC
LIF Upstream Site 4	CTTTTGGCTCAGCCACTGCTGG	CCTGGTATGGGAGAATTGTATGGCCTC
LIF Upstream Site 5	GGAGACTGCAAAAAGACCCCAGG	CCCCTTTCTCAGCCTAGCTACTTCC
LIF Upstream Site 6	TCATCAGTGATGGTCATTCGCGTG	TGCTAAGAGATTACAAGACATGTTGCAACCA
KRTAP1-1	GCATCAGCAAAAATAGACAAGTG	TCACTCAGTTTAGTGTTTTATTGGC
KRTAP1-5	GGTTTGATTTCCTCATCTACAAAG	GAACCATCCTACCCACCTCC
KRTAP2-1	GGCAATATCTTCCTGATTAGTTTCT	TTGTATGCTTTGAATACTAGCCAAC
KRTAP2-3	AAATTAAAAAGACTAAAGACCCCTG	CCTAGCATGTGATAGAAAAGAGATTTC
KRTAP2-4	GTCAGTTTACTACCAGTGGTCATTC	CCTTATCCATAGAAATCTCATCTATTCA
KRTAP4-12	CCAAGGTTAAAGAGAAAGTAAGCC	GAATAAATGCCATAGCGCAA
KRTAP4-6	TCTATCAGTTACTTTGTCACAACTGG	GGTCACTTGAAAAAGTAAAGGACAC
KRTAP3-1	CATTCTGTTAATCTCTAGACTCTAGCATAG	GCTCTCATAGTGGAATTTATCGAA
KRTAP4-11	GATTTGAATCAAGATGGAATGC	CTTTCTCCCTGTTCTCTCTTTGT
KRTAP4-4	GCTTATCAAATCCAAGAGTTTTTTG	CTAGGAAATACTCTTCTAGACCTTGATCTA
KRTAP4-5	GAGTCTTACTAGGTTTGTTGAGTTTCC	CCAGAAAGGCAACAATAGATGCCTGTTTTG
KRTAP4-7	GGGTTTTGACAAATGAGTAGGAG	CAAAATGCTTGTCCAAGATTAAAG

Table 2-1: 3C Primers. Primers designed for use in Chromosome Conformation Capture Assays (3C). For the detection of interactions only reverse primers were used. For construction of the synthetic libraries to both 5' and 3' primers were used in amplifying the initial genomic fragments prior to digestion and ligation, where final products were detected by combinations of 3' primers. All primers are given in the 5'-3' direction.

2.8 3C Controls

To quantitatively measure the efficiency of library formation, an interaction between two neighboring restriction fragments was detected and used to normalize 3C data. Neighbouring restriction fragments ligate together at high frequency, due to their immediate linear proximity and likelihood of being cross-linked during fixation with formaldehyde. An interaction between two neighboring restriction fragments at the *LIF* gene (known as *LIF* A1xA2) was measured by qPCR and used to normalize data collected. Alternately, another set of primers was designed to detect an interaction between neighbouring restriction fragments occurring at the *Glyceradeldehyde 3-phosphate dehydrogenase* (*GAPDH)* gene. These control interactions measure an 'inversion' in the linear sequence of DNA, and can only be detected in samples that have successfully undergone both digestion (by *HindIII*) and ligation (by T4 DNA ligase) to create a 3C library.

Linear distance from the anchor fragment plays a significant role in interaction strength at close distances (<25,000 bp genomic distance). To ensure our libraries were performing adequately, a set of 'Range-finding site' (RFS) primers were designed at sites near the *CXCL8* gene on chromosome 4. These primers captured restriction fragments occurring at distances of ~5000, ~15,000, and ~25,000 bp away from the *CXCL8* gene. Using qPCR, we amplified the library using combinations of the *CXCL8* anchor

primer and these RFS primers. We observed declining interaction frequencies as a result of increased distance and concluded our 3C libraries we're performing in a predictable and appropriate manner.

2.9 Graphing of 3C Data

From Ct to Relative Value

Cycle threshold (Ct) values are standardized by alignment of a neighboring interaction at the *LIF* gene (*LIF* A1xA2). This novel ligation product is formed by the inversion of one of the restriction fragments of the neighboring pair and is used to estimate the efficiency of 3C library synthesis. To standardize differences in the formation efficiencies of 3C libraries these interactions are transformed by adding or removing CT value to all interactions in that biological replicate, normalizing the interaction profile under the assumption the only difference between *LIF* A1xA2 interaction frequency is due to library formation efficiency. Once the data has been transformed it can be graphed in a straight line scatter plot.

For Example:

3C Sample	non-normalized CT (LIF A1xA2)	normalized CT (value = 25)
Proliferative	26.72	25.00 (-1.72 CT to all other CT)
Quiescence	24.84	25.00 (+0.16 CT to all other CT)
Rapamycin Treated	26.34	25.00 (-1.34 CT to all other CT)

Interaction Frequency Calculation

To measure interaction frequency the normalized Ct value undergoes the following transformation, where 40 is the maximum Ct value possible by qPCR as it represents 40 cycles of amplification:

Interaction Frequency = 2^(40-3C Sample CT)

Ex. **Interaction Frequency** = 2^(40-25) = 32768

Calculation of Error

Error is calculated from the standard error of the mean (SEM) between the number of biological replicates for each treatment.

(Standard Deviation between biologicals) / Sqrt (# of biological replicates)

To use this error in the profile the result must undergo a transformation/interpretation of the appropriate range of error.

Maximum Error = Interaction Frequency + Error

Minimum Error = Interaction Frequency – Error

Interaction Frequency Error = (Maximum Error - Minimum Error)/2

Graphing 3C Data

Once the data has been transformed it can be plotted in a straight-line scatter. The data is plotted with the Y-axis set as the interaction strength of the assayed interaction, and the X-axis as the distance of assayed interaction from the anchor.

2.10 3C Product Size Confirmation

Correct product formation in 3C-qPCR analysis was confirmed by electrophoresis to separate 3C qPCR products and sizing the resulting bands on either a 6% TBE or a 6% SB gel. As our 3C technique and analysis was developed, we transitioned from analyzing the size of PCR fragments on a 2% agarose gel,

to either a 6% TBE gel (**Figure 2-4**) or 6% SB gel (**Figure 2-5**). These gels provided much higher resolution than the 2% agarose gel and made confirming correct product size easier. The observed size of 3C qPCR products was compared to the theoretical size of the fragment, calculated by combining the linear distances from the beginning of each primer sequence to the restriction site (the 'amplicon contribution') and adding six base pairs to accommodate for the site of the restriction site. By combining the qPCR melt-curve data to this analysis, we were able to confirm that the correct and unique product was being formed.

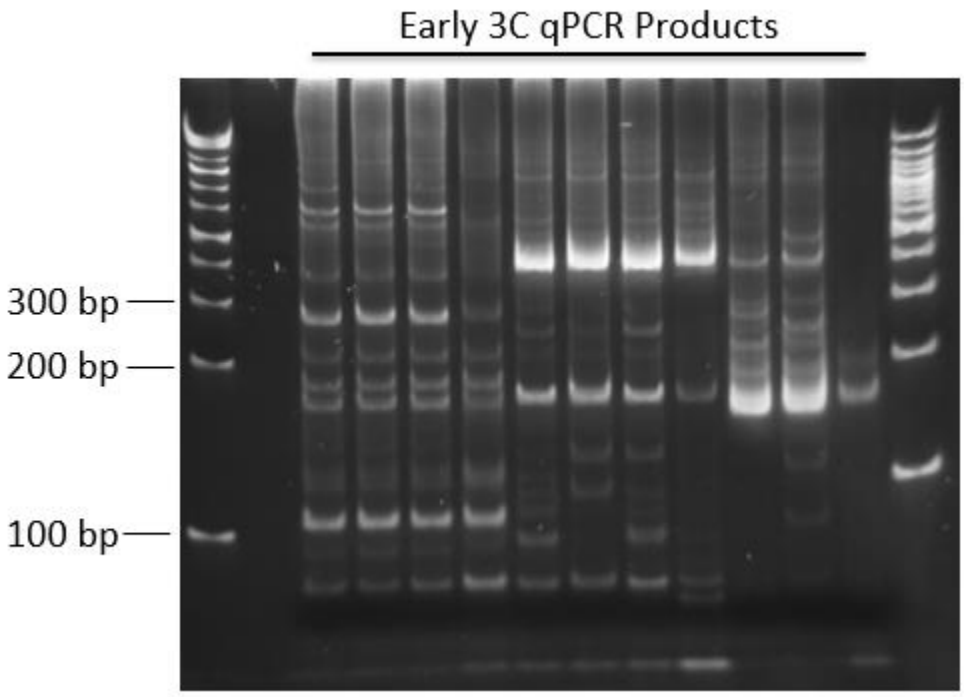


Figure 2-4: Sizing of PCR Products by Tris-Borate Gel. Separation of several 3C early qPCR products on a 6% TBE gel. A 1 kB DNA ladder (Invitrogen, USA, Cat #: 10488090) was used to estimate product sizes. TBE Gels provide much higher band resolution compared to agarose gels, even between products of similar sizes. Many 3C primers were designed to amplify final products ranging 100-300 bp and TBE gels effectively resolved this size range.

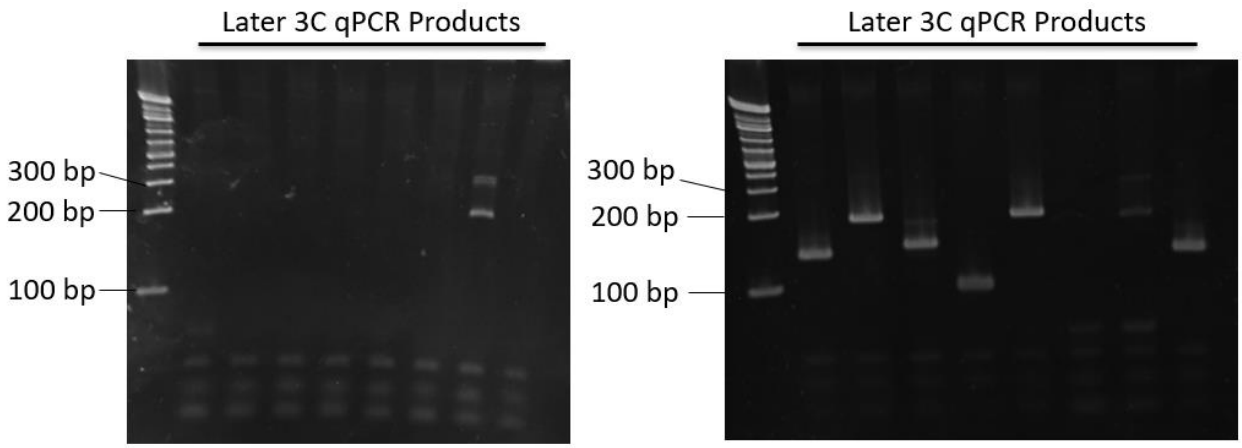


Figure 2-5: Sizing of PCR Products by Sodium Borate Gel. Separation of several later 3C qPCR products on a 6% sodium borate (SB) gel. A 1 kB DNA ladder (Invitrogen, USA, Cat #: 10488090) was used to estimate product sizes. 6% SB gels provided even better band resolution than 6% TBE gels and were also able to separate of low (100-250 bp) size products. These gels could also be electrophoretically separated in less than 10 minutes.

2.11 3C Optimization

Our 3C protocol has been optimized at several steps of the protocol.

Cell Harvesting

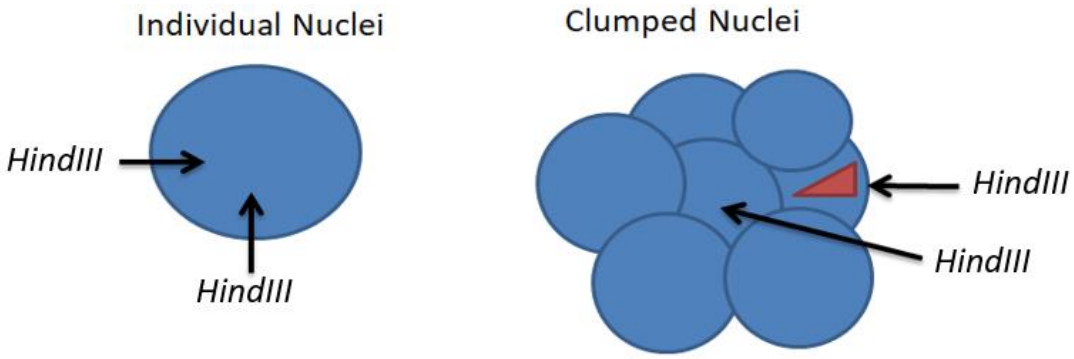
Traditionally the 3C protocol has been performed by initially re-suspending cultured cells with Trypsin or Tryple Express™ prior to fixation with formaldehyde. In the course of this project we were made aware that genome organization at the chromosome level is capable of altering within fifteen minutes in response to stimuli (Mehta *et al*., 2010). As resuspending cultured cells can require as much as fifteen minutes to perform, especially if many samples are being prepared at once, a concern was raised regarding the re-organization of the genome in response to this resuspension. To try to best

preserve the three-dimensional organization of the genome we sought an alternative method of fixing cultured cells with formaldehyde. Similar to ChIP and even a recent 3C-ChIP protocol (Tiwari and Baylin, 2009), we fixed cells in culture without resuspension, instead replacing culture media with serum-free media containing 1% formaldehyde. In our initial attempt we discovered cells would not resuspend after the addition of trypsin, as expected, and would require the use of cell culture tissue scrapers to liberate cells from the plate. This avenue produced three immediate problems. Cells became quite difficult to count following scraping and many would be destroyed by the process causing us to lose information on exactly how many cells we began with. The collection of cells was less efficient as many were unrecoverable from the surface of the plate even following repeated scraping and washing with PBS. The resulting cells we did successfully collect endured over 200 strokes with a glass Dounce homogenizer in a cold cell lysis buffer. This result lead to a search for alternate methods to isolate and permeabilize cell nuclei (discussed in Nuclei Isolation). We were able to create successful 3C libraries from this process; however, they suffered from an immediate large loss of cell number which limited the production of 3C products and subsequent breadth of analysis. In the end we decided on fixing the cells immediately upon resuspension, skipping a five-minute centrifugation step and adding media containing formaldehyde directly to the resuspension volume. Resultantly we reduced the time between culture and fixation to 7-9 minutes from ~15 minutes.

Nuclei Isolation and Digestion Efficiency

Following concerns of nuclei clumping prior to digestion with *HindIII*, we examined samples under a light microscope and observed variable levels of clumping/coalescence. To some degree the clumping could be reduced, albeit not entirely, by repeated vortexing and pipetting. However, as this was time-intensive for even one sample and we were often working with upwards of six samples, it was desirable to find other means to individualize nuclei. Individualization of nuclei is an important step for

digestion as clumping occludes nuclei or limits surface area (**Figure 2-6**), reducing the total amount of genomic DNA available for digestion by *HindIII* in solution and subsequent ligation. Differential surface area also biases digestion, as ease of access into genomic DNA will facilitate a gradient wherein the nearest-to-surface DNA will be digested more thoroughly and more internalized/occluded DNA will be less digested proportional to increased distance from surface. Undigested DNA caused by either clumping and complete occlusion of internalized cells or cells whose surface area is unequal due to clumping which may cause DNA to remain undigested may look similar to genomic DNA when separated during our analysis by separation of samples by 1.0% agarose gel.



Surface Area Exposure	100% (Equal Area)	??% (Variable Area)
Resulting Digestion	Adequate	Insufficient and possibly biased

Figure 2-6: Clumping of nuclei reduces digestion efficiency. Illustration depicting the access of the restriction enzyme *HindIII* to formaldehyde fixed nuclei during the 3C protocol. Height of the red triangle in the second figure represents a biased gradient of digestion anticipated across the nuclear volume (in a cell with limited surface area due to clumping). Cells at the center of clumps may not be accessible to *HindIII* (or other) restriction enzymes.

To remediate the clumping of isolated nuclei we attempted two changes to the protocol. One change was the removal of sodium dodecyl sulphate (SDS) during resuspension, inspired by the observation that clumping significantly worsens when the cells were resuspended in restriction enzyme digestion buffer containing SDS. Our suspicion was that SDS was increasing the 'stickiness' of the nuclei and worsening the clumping (SDS is added to solubilize the membrane). Additionally, we also suspected that the addition of BSA to coat the nuclei would reduce clumping. As can be observed in **Figure 2-7** both early attempts to reduce clumping caused the opposite effect. Through experience we eventually discovered that reducing the speed of the centrifuge during resuspension (a cleanup step) and diluting fewer cells in the same volume resulted in less clumping. By evaluating 3C libraries that had digested well (>70% digestion) we observed that fewer cells at time of harvest correlated with better digestion. These samples were initially thought to be more successful because the restriction enzyme had less material to digest; however, it has since been observed that individualization of even a large number of initial nuclei in a sample can result in greater than 70% digestion of a target site.

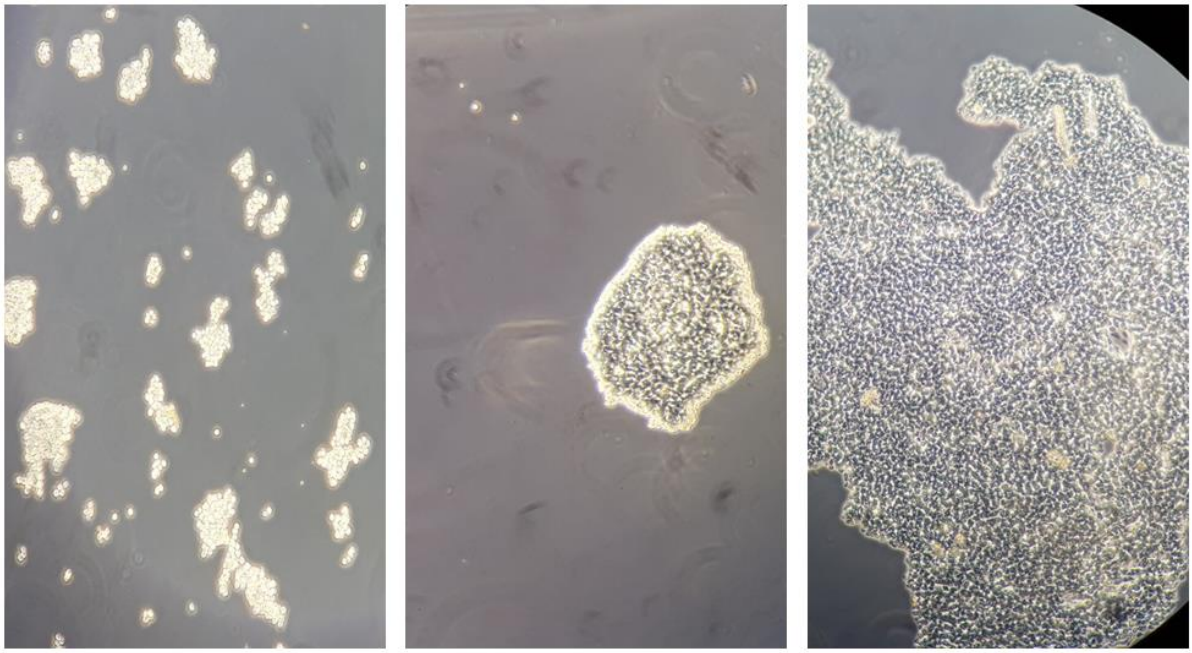


Figure 2-7: Modifications to the 3C protocol to mitigate the clumping of nuclei. During troubleshooting of the 3C protocol, several methods to reduce clumping were performed. These alterations occur just prior to the digestion step with *HindIII* and involved either no alteration to the protocol (Left), the removal of SDS (Middle), and the addition of BSA to coat nuclei and reduce cohesion (Right).

Improved Ligation Efficiencies

The ligation step is fundamental to the success of creating 3C libraries, but its reproducibility is limited by the level of digestion that occurred to create the necessary sticky/overhang fragments. If digestion was complete, it would generate ~900,000 restriction fragments per genome, assuming a 6-cutter (AAGCTT) restriction enzyme such as *HindIII* was used. As T4 DNA Ligase is dependent on ATP for activity, using one ATP per ligation event, between ~450,000 (one ligation event per pair) to ~900,000 (all fragments ligated into one strand) molecules of ATP will be consumed. Per million genomes this number rises to 4.5 x 10^11 to 9.0 x 10^11 molecules of ATP, or 7.5 E-13 to 1.5 E-12 Mol of ATP. In considering these requirements against the ATP provided by the T4 DNA Ligase Buffer (1x) – See **Table 2-2** – it is clear that there is an extreme excess of ATP available in solution even at substantially smaller solution volumes. However, several protocols (Hagege et al., 2007; Naumova et al., 2012), either add excess ATP to these reactions or delay adding the buffer until just prior to the ligation event. At early stages of the developing our protocol we observed some well-digested 3C libraries did not ligate all restriction products (**Figure 2-8**).

T4 DNA Ligase (1x) Buffer	Total ATP (mols)	Required ATP for 10 Million Cells (Mol)	Excess ATP	Percent ATP Consumed
7 ml Solution	7.00E-06	1.49502E-11	7.00E-06	0.000%
1 ml Solution	1.00E-06	1.50E-11	1.00E-06	0.001%

0.5 ml Solution	5.00E-07	1.50E-11	5.00E-07	0.003%

Table 2-2: Theoretical Consumption of ATP during the ligation step of 3C. This table describes 3 total volumes ligation may be performed in, 7 ml, 1 ml, and 0.5 ml. For each total volume the total mols of ATP has been calculated as well as the required ATP by T4 DNA Ligase to complete ligation of a 3C library, assuming 100% digestion efficiency. The two rightmost columns describe the excess ATP available in solution and the percent total ATP consumed of the solution to complete ligation.

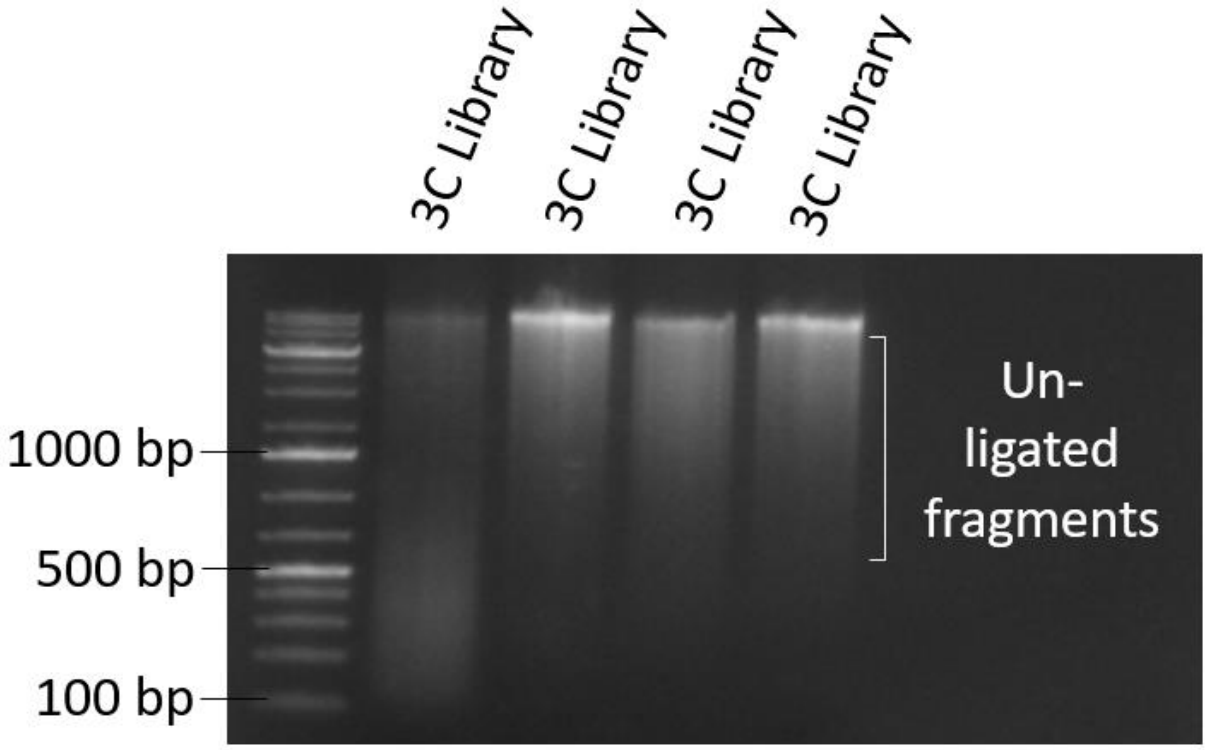


Figure 2-8: Incomplete ligation of 3C libraries. Separation of several 3C library samples on a 1.0% agarose gel alongside a 1 kB DNA ladder (Invitrogen, USA, Cat #: 10488090). Some early well-digested (> 80%) 3C Libraries did not ligate all restriction fragments and present a smear below the '3C band' occurring at roughly 10 Kbp.

Different approaches by other labs (Hagege et al., 2007; Naumova et al., 2012) suggest adding the T4 DNA Ligase buffer immediately prior to ligation, avoiding its addition prior to a one-hour 37°C

incubation in Triton-X that may have degraded ATP. Though simple, this alteration has resulted in libraries produced after the change to having no smear of un-ligated products below the "3C Band" (See Agarose Gel Controls). While the exact rationale for the success of this alteration is not clear, several labs likely experienced incomplete ligations, despite the theoretical extreme excess of ATP. Storage conditions provided by Sigma-Aldrich (Cat #: 51963-61-2) suggest that aqueous ATP rapidly breaks down over the course of one-week when refrigerated, or several months when frozen. Additionally, as the ligation step of 3C occurs prior to phenol/chloroform extraction, proteinaceous components and debris are still present in solution and may consume or sequester ATP. Considering that ATP is a major energy currency in biology, is highly unstable, and is shipped as a desiccated powder rather than an aqueous solution, the conclusion is possible that early 3C ligation reactions suffered from rapid ATP degradation during the hour long incubation at 37°C prior to ligation, possibly amplified by cell debris.

Improved DNA Quantification Accuracy

At early stages of this project our 3C libraries suffered from very large (2.1 mg/µl) and inaccurate DNA reads following ethanol precipitation (**Figure 2-9**). Our DNA quantification was performed by spectrophotometry (Nanodrop™) (example of an early examination is in **Figure 2-9)**. Of note is the distinct 'bumpy' curve produced during spectrophotometry. This problem, caused by a contamination of dithiolthreitol (DTT) which is present in the T4 DNA Ligase Buffer, was originally reduced by column purification. However, column purification introduced another step that resulted in the loss of total DNA collected at the end of the experiment. As 3C libraries can be quickly consumed by analysis, with each replicate of a triplicate for qPCR requiring 100 ng of template, investigation are quite limited by the amount of DNA available per biological replicate. During a review of the literature (Hagege et al., 2007) we discovered that a simple dilution of the ligation mix (increasing total volume by ~40%) with water prior to phenol/chloroform extraction and ethanol precipitation prevented DTT from

contaminating the precipitated 3C sample. Further improvements to the protocol reduced the total DTT in solution by reducing the total ligation volume. In practice, these alterations have severely reduced (possibly eliminated) DTT contamination and entirely removed the need of column purification. Accuracy of quantification by spectrophotometry has been verified by loading equal (by spectrophotometric read) amounts of a template on an agarose gel (**Figure 2-10**), as well as by qPCR against primers against a genomic DNA site at the *EFEMP2* (*EGF-containing fibulin-like extracellular matrix protein 2*) gene. Primers designed against the *EFEMP2* gene do not contain a restriction site for *HindIII* and 'count' the number of genomes present in a DNA sample. By amplifying *EFEMP2* from a known concentration of genomic DNA using qPCR, a comparison can be made between this known concentration and the unknown concentration of another sample also amplifying *EFEMP2*.

Isopropanol Precipitation supplants Ethanol Precipitation

Early 3C libraries also suffered from contamination from the buffers used during restriction digestion and ligation, from cell debris, and from the salt used during the precipitation of DNA. High salt concentration and other persistent contaminants negatively alter the accuracy of quantification by Nanodrop spectrometer and required us to measure template concentration by using *EFEMP2* (mentioned earlier). Initially the DNA precipitation step was performed in a large volume (>35 mL, containing final concentration of 70% ethanol and 150 mM sodium acetate). This precipitation commonly required several days to complete (stored at RT or -20°C) and originally contained large amounts of DTT and precipitated salt (noted by the appearance of a strong white pellet). Additionally, the accompanying digestion efficiency sample (collected as 10% of the total volume of sample following digestion) often yielded weak or absent pellets compared to expectations based on 3C library pellets. In suspecting that the large volume, contaminants, and low DNA content of DE samples were problematic we supplanted the ethanol precipitation. We introduced an isopropanol precipitation (4°C, 20 minutes,

1:1 isopropanol, 10% (v/v) 3M NaOAc) aided by 1 ul of 20 mg/ml glycogen). This change increased the yield and speed of precipitation for 3C samples; the yield, speed of precipitation, and reliability of visible pellet formation for DE samples.

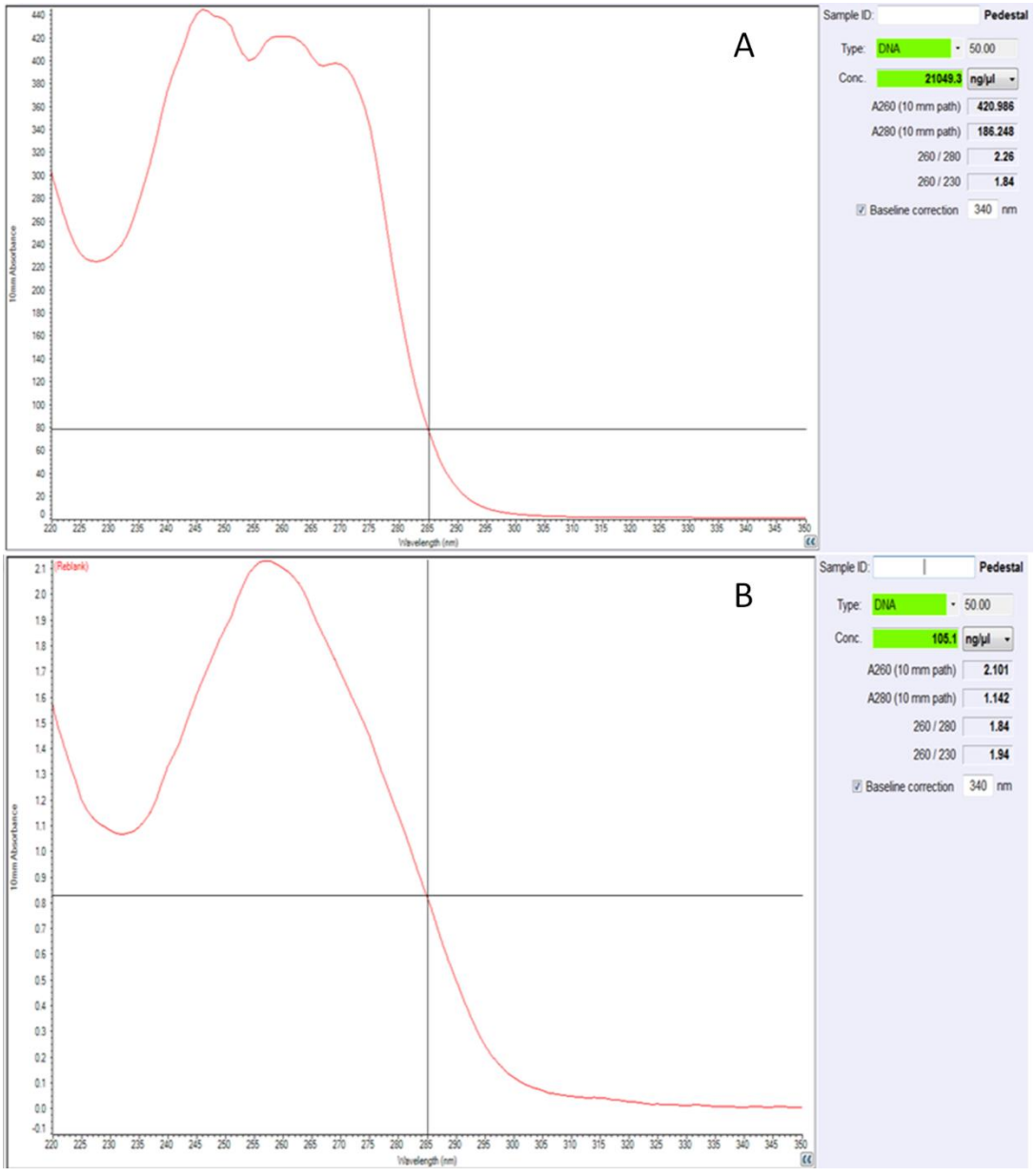


Figure 2-9: Computational output of 3C samples by Thermo Scientific Nanodrop 2000. 3C libraries were quantified using a Thermo Scientific Nandrop 2000. Early 3C samples suffered from contamination with dithiolthreitol (DTT) that caused an erroneous and inflated estimation of DNA. Spectra of a DTT-

contaminated 3C sample (A) and a DTT-free 3C sample (B) are shown above, as measured by a Thermo Scientific Nanodrop 2000.

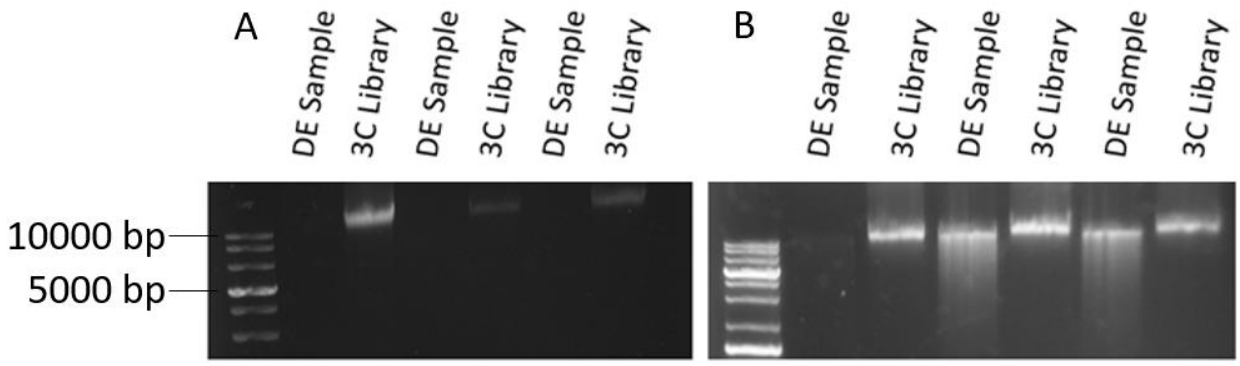


Figure 2-10: Modification to DNA precipitation increases the accuracy of DNA quantification. Separation of 3C libraries on a 1.0% agarose gel alongside a 1 kB DNA ladder (Invitrogen, USA, Cat #: 10488090). In each lane, 300 ng of sample was loaded as measured by a Thermo Scientific Nanodrop 2000. **(A)** Early 3C samples suffered from inaccuracy in quantification. **(B)** Later 3C samples benefitted from increased accuracy in quantification due to higher purification and reduction of DTT contamination.

3.0 Results: Analysis of chromosome territory positioning in response to quiescence induction or Rapamycin treatment.

Genome organization in the 3D nuclear volume can alter in response to stimuli. In earlier assays performed in the Eskiw laboratory (Gillespie et al., 2015) and others (Mehta et al., 2010) a reorganization of chromosome territory position has been noted for both chromosomes 10 and 18 following induction of quiescence and treatment with 500 nM Rapamycin using FISH. These alterations occurred alongside changes to gene expression as measured by RNA-Seq (**Table 3-1**). For the purpose of this thesis, these experiments have been recapitulated and expanded upon by the addition of chromosome 4, which contains a cluster of *chemokine* genes up-regulated by treatment with 500 nM Rapamycin. Considering the up-regulation of this cluster of genes we anticipated that chromosome 4 will reposition in the nuclear volume in response to treatment with Rapamycin in our 2DD human foreskin fibroblast cell line.

Identical to work previously performed, cells were cultured in proliferative, quiescent (0.5% FBS), or Rapamycin-treated (500 nM) conditions and subjected to analysis by FISH and subsequent erosion analysis. Specifically, in FISH chromosomes are ‘painted' using probes that hybridize to strands of a target chromosome in the presence of a Hoechst 33342 (H33342) stain that stains all chromatin non-specifically. By quantifying the signal produced from each chromosome and standardizing these values against the non-specific DNA signal produced by H33342 staining, the exact positioning of each chromosome can be characterized. In this characterization known as Erosion Analysis (**Figure 3-1**) each nucleus is divided into five concentric rings of equal area and the chromosome occupancy per ‘shell’ or ring is given a numerical value. By comparing these profiles in each treatment, the repositioning of chromosome territories can be measured. Chromosome X was used as a control.

Repositioning of a chromosome is suggestive of altered gene expression. It has been previously observed that positioning towards the nuclear interior relates to increased expression of contained

genes and positioning near the exterior related to decreased expression of contained genes. Chromosome positioning; however, has other features to consider such as which chromosomes are neighbours. By repositioning two chromosomes closer together in the nuclear volume, *trans*-interactions that were previously impossible can now occur such as between an enhancer and a promoter. Furthermore, relocating a chromosome may allow access to a specialized transcription factory previously denied by positioning, or may facilitate the rapid expulsion of newly transcribed mRNA from the nucleus by occupying a more peripheral position. Chromosome 4 contains a cluster of chemokine genes up-regulated in response to Rapamycin. How this chromosome alters position in response to this up-regulation may inform on the underlying processes organizing chromosome territories. Alternately, repositioning of this chromosome may be responsible for the up-regulation, by allowing contained genes to interact with nuclear elements previously denied, such as transcription factories or other chromosomes. Chromosome 4 contains a cluster of *CXCL* genes that are up-regulated in response to Rapamycin. This 'chemokine cluster' serves as a good target to investigate, to characterize if repositioning of the chromosome is necessary for increased expression. In the following experiments we measured the positioning of chromosome 4, 10, 18, and X.

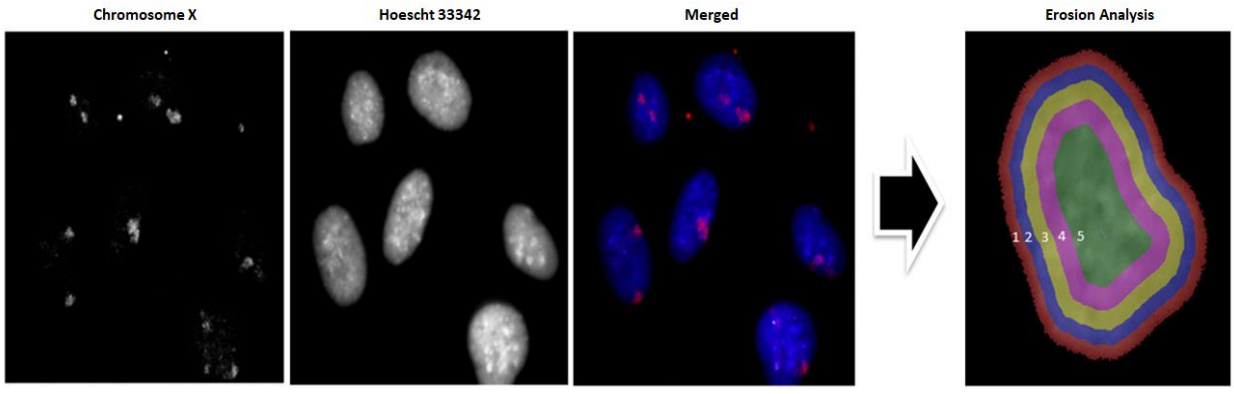


Figure 3-1: **Schematic of fluorescence in situ hybridization image collation and erosion analysis.** Images were captured using Fluorescence microscopy for chromosomes of interest (4, 10, 18, X) and are

computationally merged prior to Erosion analysis. H33342 dye is used to stain the genome non-specifically. Chromosome-specific probes containing fluorophores are used to identify targeted chromosomes. During Erosion analysis the positioning of chromosomes is quantified in each of five concentric rings of equal area.

Gene	Also known as	Full Name	Chromosome Location	Function	Regulation in Quiescence Induced RNA-Seq Dataset (Fold Change)	Regulation in Rapamycin treated RNA-Seq Dataset (Fold Change)
CXCL8	*IL8*	C-X-C chemokine motif ligand 8	Chr 4 – 74 Mbp	Angiogenesis, chemotaxis, innate immune response, growth, metalloprotease activation	0.406	28.5
CXCL6	*GCP-2*	C-X-C chemokine motif ligand 6	Chr 4 – 74 Mbp	Growth, anti-bacterial, inflammation	0.655	2.6
CXCL1	*GRO1, GROα*	C-X-C chemokine motif ligand 1	Chr 4 – 74 Mbp	Growth, angiogenesis, cell motility, inflammation, wound healing	0.425	6.4
CXCL3	*GRO3, GROγ*	C-X-C chemokine motif ligand 3	Chr 4 – 74 Mbp	Same as CXCL1	0.193	6.3
CXCL2	*GRO2, GROβ*	C-X-C chemokine motif ligand 2	Chr 4 – 74 Mbp	Same as CXCL1, elevated in many cancers.	0.397	3.2

Table 3-1: Gene from chromosome 4 (74-75 Mbp) are differentially regulated in response to quiescence and Rapamycin. Summarized changes in regulation of multiple genes on chromosome 4 from an RNA-Seq data (Gillespie *et al*, 2015) (GEO Accession: GSE65145). Each treatment condition (proliferative, quiescent, Rapamycin) had two biological replicates. Fold change is measured in comparison to proliferative samples. Samples were either induced into quiescence by reducing media serum to 0.5% or by treating cells with 500 nM Rapamycin for 5 days. Contained genes occur in a cluster of Chemokine C-X-C Motif Ligand genes on chromosome 4 region 74 Mbp (*CXCL8*, *CXCL6*, *CXCL1*, *CXCL3*, and *CXCL2*) that were discovered up-regulated at this location during treatment with 500 nM Rapamycin in the 2DD normal human dermal fibroblast cell line.

3.1 Chromosomes 10 and 18 reposition in response to quiescence induction and Rapamycin treatment

In agreement with earlier results (Gillespie et al., 2015; Mehta et al., 2010), I observed a repositioning of chromosomes 10 and 18 in both quiescence and Rapamycin-treated conditions. In proliferative cells, chromosome 10 occupies a central position in the nucleus (shells 4 and 5, the most internal shells) (**Table 3-2**). In Rapamycin-treated cells a repositioning of chromosome 10 to shells 1 and 2 (the most external shells) was observed. Similarly, this repositioning of chromosome 10 from shells 4 and 5, to shells 1 and 2 was also observed in quiescent cells. Chromosome 18, which occupies a peripheral position (shells 1 and 2) in proliferative cells, also underwent a significant nuclear repositioning. For both quiescence and Rapamycin-treated, chromosome 18 repositioned to the innermost shells (4 and 5). In all three treatments chromosome X displayed little or no movement and serves as a control for these experiments.

These results validate earlier findings and confirm our three growth conditions are altering genome organization in a predictable and repeatable manner. While the significance and exact effect upon gene expression from this repositioning is not clear, the genome is re-organizing in response to both of our treatments compared to proliferative cells. Furthermore, these results display the capability of the genome to significantly alter the position of chromosomes in response to stimuli. As this repositioning is reproducible and has been previously observed, it suggests it provides a function to the cell. However, our current investigations are focused on the organization and repositioning of chromosome 4.

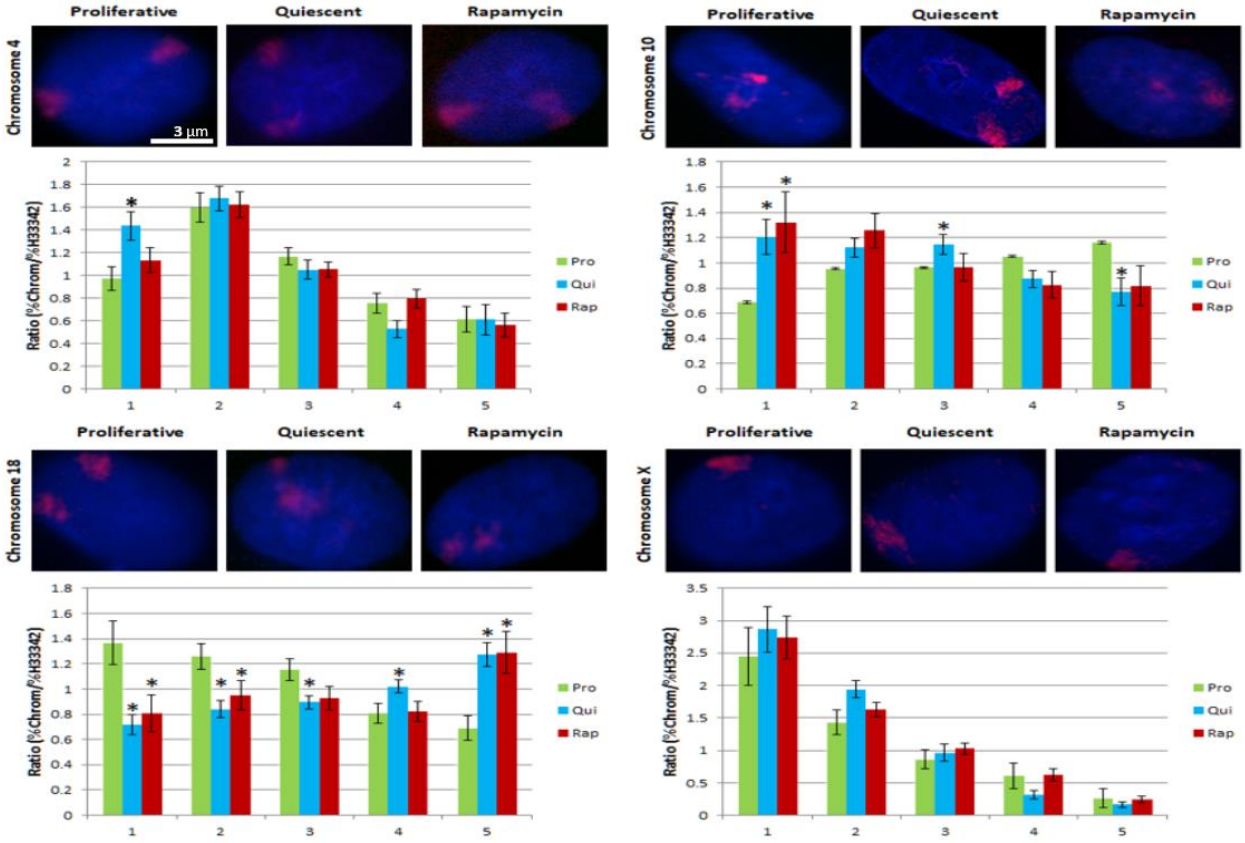


Figure 3-2: **Positioning of several chromosome territories during proliferative, quiescent, or Rapamycin-treated conditions.** Bar graphs depicting the results of erosion analysis on images collected by fluorescence microscopy. During erosion analysis the nucleus is divided into five concentric rings of equal area, with shell 1 being the most external and shell 5 being the most internal. Height of bars is determined by the ratio of specific chromosome painting divided by strength of Hoechst 33342 staining for the relevant shell. Below each graph are representative images of cells from each treatment. Chromosomes of interest are stained red and DNA is non-specifically stained blue with DAPI. (* = $P < 0.05$.) Pro = proliferative, Qui = quiescence cells, Rap = Rapamycin-treated cells. Representative images for each treated and chromosome were collected from at least 30 nuclei images.

Chromosome	Proliferative	Quiescence	Rapamycin
4	Peripheral (Shell 2/3)	Peripheral (Shell 1/2)	Peripheral (Shell 1/2)
10	Central (Shell 4/5)	Peripheral (Shell 1/2)	Peripheral (Shell 1/2)
18	Peripheral (Shell 1/2)	Central (Shell 4/5)	Central (Shell 4/5)
X	Peripheral (Shell 1/2)	Peripheral (Shell 1/2)	Peripheral (Shell 1/2)

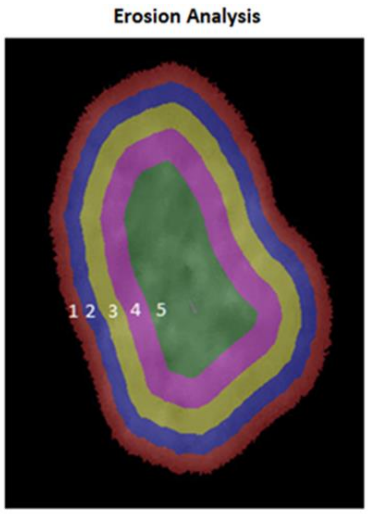


Table 3-2: Summarized results of chromosome territory positioning during proliferative, quiescent, and Rapamycin-treated conditions. Left) Summarized results following quantification of chromosome positioning by Erosion Analysis. Numerical data of chromosome shell occupancy has been simplified to a positioning (Central or peripheral) and the two shells with the highest chromosome signal relative to H33342 ratios. Right) Image depicting the numbering system for Erosion Analysis. Shell 1 represents the most peripheral shell while shell 5 is the most central. (Erosion Analysis image credit: Zoe Gillespie).

3.2 Chromosome 4 repositions in response to quiescence induction

Chromosome 4, which contains our cluster of *chemokine* genes of interest at 74 Mbp, has yet to be characterized by FISH in all three treatment conditions. We observed an up-regulation of these chemokine genes on chromosome 4 in our RNA-Seq data set, from this we anticipated a repositioning of chromosome 4 in the nuclear volume in response to increased expression of these genes.

We performed analysis by DNA-FISH on chromosome 4 **(Figure 3-2, Table 3-2)**, and characterized chromosome 4 position in shells 2 and 3 in proliferative cells. When treated with Rapamycin or induced into quiescence, there was a shift in positioning to a more peripheral position. In both quiescence and Rapamycin-treatment the highest occupancy shifted from shells 2 and 3 to shells 1

and 2. During Rapamycin treatment this test is non-significant ($P > 0.05$); however, in quiescent cells the movement towards the periphery is more pronounced and is significant ($P < 0.05$).

3.3 Discussion

Overall these findings on the repositioning of chromosome 18 and 10 during Rapamycin treatment and quiescence induction indicate treatments were successful in reproducing earlier results (Gillespie et al., 2015; Mehta et al., 2010), confirming that the genome is capable of re-organizing in a predictable and repeatable manner, and that our treatments are behaving as seen previously. These examinations by DNA-FISH were expanded upon to include chromosome 4, which contains our cluster of chemokine genes of interest at region 74 Mbp. In proliferative cells chromosome 4 occupies a position near the periphery. However, during treatment this positioning only slightly alters in quiescent and Rapamycin-treatment to an even more peripheral position. Though subtle, this chromosome territory repositioning was significant ($P<0.05$) during quiescence induction; however, the effect non-significant in Rapamycin-treated cells.

Interestingly, there was not a significant repositioning during the up-regulation of genes by Rapamycin-treatment. Furthermore, we did observe a repositioning of chromosome 4 to the periphery during quiescence, which displays lower expression of the *chemokine* genes of interest (*CXCL8*, *CXCL6*, *CXCL1*, *CXCL3*, *CXCL2*). Despite the up-regulation of these *CXCL* genes during Rapamycin treatment we did not observe significant repositioning of the chromosome. This may indicate that the up-regulation of a single cluster of genes is not enough to re-position the whole chromosome, and rather a more subtle effect may be occurring at sub-chromosomal levels.

Previous analyses indicate a correlation repositioning of chromosomes towards the nuclear periphery with decreased transcription of contained genes, though this relationship is incomplete and warrants further investigation. In the context of observed changes for quiescent cells, which have

reduced gene expression of the chemokine cluster, chromosome 4 did move to a more peripheral position. However, we did not observe a repositioning to the interior during treatment with Rapamycin that upregulated expression of several genes on chromosome 4. It is possible that many factors are at play, such as the expression of other genes on the chromosome, which may also influence the chromosome to a more internal or peripheral position.

The nucleus is a highly dynamic and crowded organelle. Many factors orchestrate this organization and perhaps the changes to three-dimensional organization of chromosome 4 are occurring more subtly – at the 'sub-chromosomal' level. At this sub-chromosomal level, alterations in three-dimensional organization manifest as enhancer-promoter interactions, or as interactions occurring between genes caught in the proximity of the same transcription factory. These subtle changes would not be measurable by DNA-FISH, due to optical microscopy limitations. Investigation at this level requires the use of the 3C technique, wherein local folding and interactions of a DNA strand can be characterized.

4.0 Results: 3C Analysis of *CXCL* Locus on Chromosome 4

Previously, analysis of RNA-Seq data of Rapamycin-treated primary human foreskin fibroblasts (2DD) identified a region on chromosome 4 (region 74 Mbp) which contained many *chemokine* genes that were Up-regulated. These genes are members of the chemokine C-X-C motif ligand (CXCL) family and include *CXCL8*, *CXCL6*, *CXCL1*, *CXCL3*, *CXCL2*. These genes are contained in a putative TAD spanning ~360,000 base pairs on chromosome 4. Many of these genes have roles in regulating growth and proliferation and have been ascribed roles in many cancers (commonly referred to as Growth Related Oncogene (GRO) α, β, or γ in cancer) (Dehghani et al., 2005; Kim et al., 2015). These *CXCL* genes were discovered in tight sequence proximity on chromosome 4 that changed expression in response to treatment with 500 nM Rapamycin treatment, based on RNA-seq data (See **Table 3-1**) (Gillespie et al., 2015). As the expression of these *CXCL* genes all increased following Rapamycin-treatment, we suspected these genes would coalesce in three-dimensional space during transcription at a transcription factory, recruited by the transcription factor Signal Transducer and Activator of Transcription 5A/B (STAT5A/B). Earlier work from the Eskiw lab indicated that STAT5A/B mediates changes in transcription profile of cells treated with Rapamycin (Gillespie et al., 2015).

The observation that a cluster of chemokine genes was up-regulated by Rapamycin treatment, provided a candidate region of the genome to test if the up-regulation of genes equated to re-organization of sub-chromosomal domains. Furthermore, with the idea of the transcription factory in mind, we hypothesized that this cluster of genes likely shared a single transcription factory, or a cluster of factories containing similar transcriptional machinery, and may come together (more frequently) in three-dimensional space during up-regulation. Lastly, this cluster of *chemokine* genes on chromosome 4 also provided an example of a possible TAD, which is a cluster of like-regulated genes with boundaries that isolate its interaction with external elements while facilitating internal interactions. To test these

hypotheses, we performed 3C on a primary normal human foreskin fibroblast cell line (2DD) in three conditions: proliferative (10% FBS), quiescent (0.5% FBS), and Rapamycin-treated (500 nM).

In the 3C assay, cultured cells are incubated with formaldehyde and the resulting fixed nuclei are isolated for downstream steps. These nuclei are then incubated with a restriction enzyme (*HindIII*) that digests the genome at known sequences (AAGCTT). Following digestion formaldehyde linked restriction fragments are then ligated with T4 DNA ligase to produce a '3C Library'. These libraries contain ligated fragments that represent three-dimensional interactions occurring in the genome and are detectable by qPCR. By designing primers to flank restriction sites in our region of interest (for example, chromosome 4, region 74 Mbp), combinations of these primers can be used to characterize the local folding environment. This interaction is measured as 'interaction frequency'. To measure this qPCR is performed to measure the presence or absence of 3C products. These values are then normalized to an interaction that does not change during treatment. For these assays, 3C data was normalized to a pair of neighboring restriction fragments at the *LIF* gene. When no interactions are occurring, interaction frequency decreases as distance between two restriction fragment increases. Neighboring restriction fragments provide an accurate way of normalizing data due to their high occurrence rate due to minimal linear distance and can only be detected when a 3C product has been successfully produced. This is accomplished by using two 3' or reverse primers that cannot amplify genomic DNA that has not been digested, ligated, and inverted. Normalized interaction frequencies are measured in fold change as compared to the minimal level of detectable product ('background noise').

To characterize a region of interest, such as the chemokine cluster on chromosome 4, an anchor fragment must be selected. This anchor fragment is a constant fragment that is used in combinations of other nearby sites. For example, analysis of the chemokine cluster used *CXCL8* as an anchor which was used in combination with all sites of interest in the region (other *CXCL* genes). By measuring these interaction frequencies, a 3C profile can be generated based on the interactions occurring between the

anchor (*CXCL8*) and sites of interest, and summarily plotted. For the chemokine cluster on chromosome 4, region 74 Mbp, this was performed using both *CXCL8* and *CXCL1* as anchor fragments.

4.1 *CXCL8* anchored interactions in the chemokine locus on chromosome 4.

Previously, we identified up-regulation of five *chemokine* genes in response to treatment with 500 nM Rapamycin. We further discovered that five of these genes (*CXCL8, CXCL6, CXCL1, CXCL3, CXCL2*) are clustered on chromosome 4, region 74 Mbp in a putative TAD. During treatment with Rapamycin these genes were all up-regulated (*CXCL8*: 28.5-fold, *CXCL6*: 2.6-fold, *CXCL1*: 6.4-fold, *CXCL3*: 6.3-fold, *CXCL2*: 3.2-fold) and during quiescence they were all discovered to be down-regulated (*CXCL8*: 0.41-fold, *CXCL6*: 0.66-fold, *CXCL1*: 0.43-fold, *CXCL3*: 0.19-fold, *CXCL2*: 0.40-fold). From this analysis, it was identified that these genes were co-regulated by common transcription factors (including STAT5A/B).

Our hypothesis for this cluster of *CXCL* genes had two elements; we suspected these genes are like-regulated and may be coming together in three-dimensional space for transcription, and that this would occur at a shared transcription factory. We suspected that Rapamycin-treatment was driving the entire cluster to engage more actively with a transcription factory, and during quiescence the cluster shifted away from a transcription factory to decrease transcription. During transcription at a transcription factory, interactions occurring between co-transcribed genes would be detectable. Conversely, genes being transcribed less (such as during quiescence) would display decreased or non-detectable levels of interaction frequencies. By analyzing the interaction frequencies between the genes, we would be able to characterize if the genes are interacting more during transcription, and if they are coming together at a single shared site for transcription.

My initial analysis used *CXCL8* as an anchor, as it was the most highly up-regulated gene in the cluster. We measured *CXCL8* interaction frequencies with several nearby downstream sites, including a

set of control sites and the nearby *CXCL* genes. These sites include 'Range-finding sites' (RFS) to measure the decay of interaction frequency as linear distance increases (RFS1: ~5000 bp, RFS2: ~15,000 bp, RFS3: ~25,000 bp), and the nearby *CXCL* genes (*CXCL6*: 96,600 bp, *CXCL1*: 128,800 bp, *CXCL3*: 297,700 bp, *CXCL2*: 360,000 bp).

To verify the success of our 3C libraries, we first tested the anticipated decrease of interaction frequency as linear distance increased from the *CXCL8* restriction fragment (**Figure 4-1, Table 4-1**) to nearby control sites. We tested this anticipated decrease interaction frequency using 3 sites, RFS1, RFS2, RFS3. As expected, we observed a strong interaction frequency between *CXCL8* and RFS1 (~5000 bp distance) in all three treatment conditions as measured in arbitrary interaction strength units above background (Pro: 741 AU, Qui: 828 AU, Rap: 755 AU). The next site (RFS2, ~15,000 bp) displayed significantly decreased interaction frequencies in all conditions (Pro: 98.0 AU, Qui: 188 AU, Rap: 212 AU) representing a 7.6- fold reduction in proliferative, 4.4-fold reduction in quiescence and 3.6-fold reduction in Rapamycin treated samples. For RFS3 (~25,000 bp distance), we observed proliferative cells had an AU value of 65.6, quiescent cells had 270 AU and Rapamycin treated samples had an interaction strength of 390 AU. We anticipated that since these were non-coding regions that there would be a decrease in all interactions with *CXCL8* as distance from the *CXCL8* anchor increased. This was the case for proliferative cells at 25 Kb distance (exhibiting a 0.09-fold interaction frequency compared to the 5 Kb distance); however, there was an unexpected increase in interaction strength between these sites in both quiescent cells (4.1-fold) and Rapamycin-treated cells (6.0-fold) samples as compared to the 15 Kb distance. Interestingly, this increase in interaction frequency between *CXCL8* and RFS3 in quiescent and Rapamycin-treated cells may be a functional interaction with a currently unknown element (e.g. Enhancer).

Next we measured the interaction frequency between *CXCL8* and the downstream *CXCL* genes (in linear order: *CXCL6*, *CXCL1*, *CXCL3*, and *CXCL2*). Interaction frequency between *CXCL8* and *CXCL6*

(96,600 bp) was similar in quiescence (0.91- fold, relative to proliferative) and decreased Rapamycin-treatment (0.48-fold, compared to proliferative) (Pro: 316 AU, Qui: 289 AU, Rap: 150 AU, compared to background). Interestingly, these interactions occurred with higher frequencies than interactions occurring between *CXCL8* and the nearby (~15,000 bp) RFS2 site, despite being positioned nearly 100,000 bp apart indicating that this may be a functional interaction based on folding and not proximity of these regions on the linear genome. The interaction frequency between *CXCL8* and *CXCL1* (128,800 bp distance) was similar for proliferative and Rapamycin treated samples (0.89-fold, relative to proliferative); however, it dropped significantly ($P < 0.10$, Student's 2-tailed T-test) in quiescent cells (0.25-fold, relative to proliferative) (Pro: 347 AU, Qui: 88.4 AU, Rap: 310 AU, compared to background). For the interaction frequency between *CXCL8* and *CXCL3* (297,700 bp), we discovered similar and comparatively weak interaction frequencies for all three treatments (Pro: 32.6-fold, Qui: 33.7-fold, Rap: 26.5-fold, compared to background) and was unchanged between the treatments (Qui 1.0-fold change and Rap 0.8-fold change compared to proliferative). Lastly, we identified a significant ($P < 0.05$, Student's two-tailed t-test) increase in interaction frequency between *CXCL8* and *CXCL2* (360,000 bp distance) in quiescence cells (27.3-fold, relative to proliferative) compared to proliferative and Rapamycin-treated (Pro: 12.3-fold, Qui: 335-fold, Rap: 22.3-fold, compared to background).

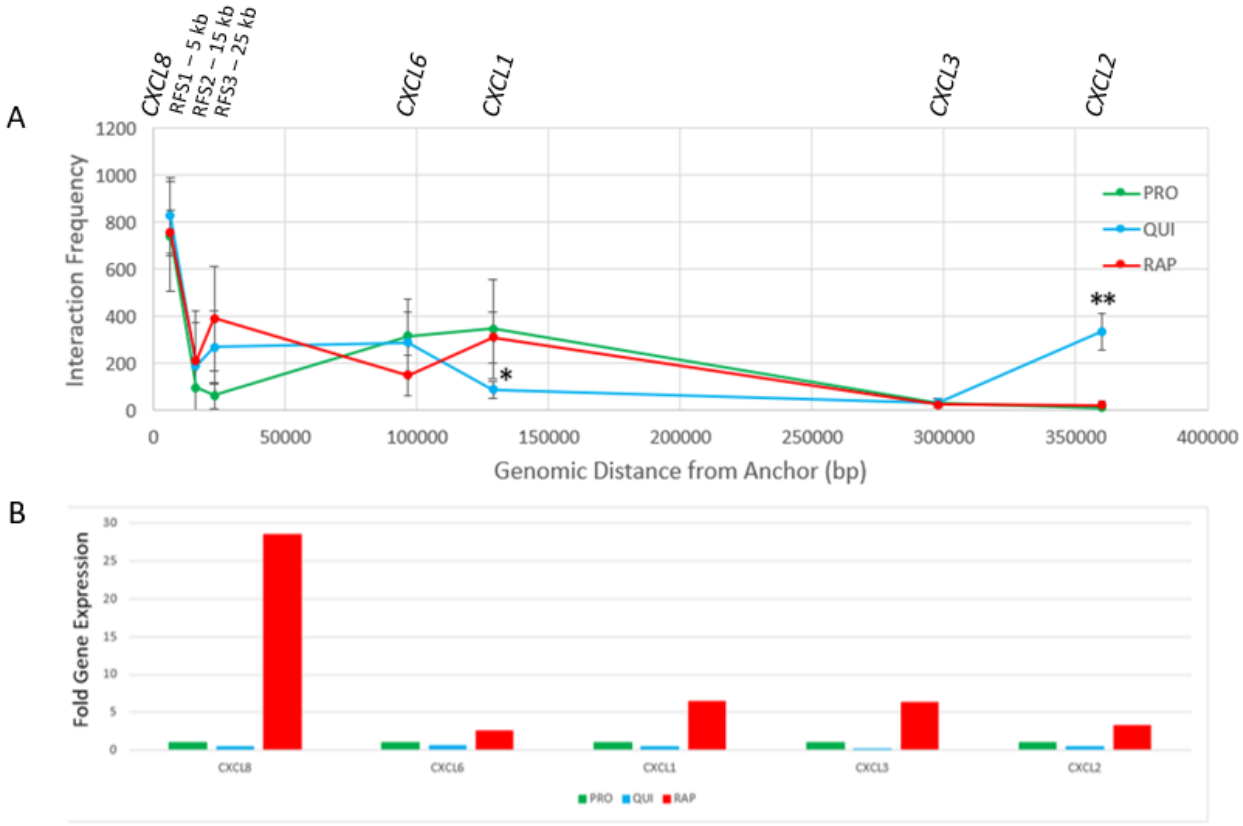


Figure 4-1: **3C profile of the *CXCL* locus on chromosome 4, region 74 Mbp with *CXCL8* as the anchor fragment. (A)** Graphed 3C profile of interactions occurring between the *CXCL8* gene and seven downstream sites (marked above graph). Interaction frequency is measured by fold change in comparison to detected products that have been normalized to a control interaction (between neighboring restriction fragments at the *LIF* gene) against background. Genomic distance is measured as the distance between the tested site and the anchor (at the vertical axis). **(B)** Relative expression of several *CXCL* genes during three different treatment conditions as determined by RNA-Sequencing (Gillespie *et al*, 2015) (GEO Accession: GSE65145). Height of bars indicates fold change in gene expression compared to proliferative values. Pro = proliferative (10% FBS), Qui = quiescence (0.5% FBS), Rap = Rapamycin Treated (500 nM). (* = $P < 0.10$, ** = $P < 0.05$, Student's two-tailed t-test). (n = 5 biological replicates for each treatment).

Interaction (*CXCL8* anchor)	Distance from anchor (base pairs)	Proliferative (fold change versus proliferative)	Quiescence (fold change versus proliferative)	Rapamycin-treated (fold change versus proliferative)
CXCL8 x RFS1	6400	1.00	1.12	1.02
CXCL8 x RFS2	16000	1.00	1.93	2.17
CXCL8 x RFS3	23000	1.00	4.12	5.96
CXCL8 x *CXCL6*	96600	1.00	0.91	0.48
CXCL8 x *CXCL1*	128800	1.00	0.25*	0.89
CXCL8 x *CXCL3*	297700	1.00	1.03	0.81
CXCL8 x *CXCL2*	360000	1.00	27.31**	1.82

Table 4-1: CXCL8 anchored 3C investigations of chromosome 4 (74-75 Mbp) relative to proliferative samples. Table depicting interaction frequency relative to proliferative samples for examined 3C interactions on chromosome 4, region 74 Mbp between five chemokine *C-X-C motif ligand* (*CXCL*) genes and 'Range-finding sites' (RFS) located near *CXCL8*. Interactions are defined by a pair of primers with one primer chosen as a constant or 'anchor' primer for the assays and another as a variable primer. Distance is measured from the site of the anchor primer to the site of the variable primer. Fold change is calculated as a ratio of interaction frequency compared to proliferative samples using the same primer pair. (** = $P < 0.05$, * = $P < 0.10$. Student's 2-tailed T-test.)

4.2 *CXCL1* anchored interactions in the chemokines locus on chromosome 4, region 74 Mbp.

To further characterize the organization of the chemokine cluster on chromosome 4, we chose *CXCL1* as a second anchor to probe the region (**Figure 4-2**). *CXCL1* was chosen due to it having the second highest up-regulation during treatment with Rapamycin (6.4 fold) and that it resides in a more

central position in the chemokine cluster than *CXCL8*. Additionally, by incorporating a second anchor site in our analysis to identify potential biases caused by using only a single anchor site. As discussed previously, interaction frequency between *CXCL1* and *CXCL8* (128,800 bp distance) decreases significantly in quiescent cells (0.25-fold, relative to proliferative, $P < 0.10$) compared to proliferative and Rapamycin-treated cells (Pro: 347 AU, Qui: 88.4 AU, Rap: 310 AU compared to background). For interaction frequency occurring between *CXCL1* and *CXCL6* (32,200 distance, the nearest *CXCL* gene), we observed a decrease in interaction frequency in quiescent cells (0.83-fold, relative to proliferative) and an increase in interaction frequency with Rapamycin-treated cells (1.22-fold, relative to proliferative) (Pro: 673 AU, Qui: 558 AU, Rap: 819 AU, compared to background). The interactions between *CXCL1* and *CXCL6* (32,200 bp) occur at similar frequencies to *CXCL8* and RFS1 at only ~5,000 bp distance. Interaction frequencies between *CXCL1* and *CXCL3* (168,000 bp distance) increased during quiescence (1.34-fold, relative to proliferative) and Rapamycin treatment (1.39-fold, relative to proliferative) compared to proliferative cells (Pro: 109 AU, Qui: 146 AU, Rap: 152 AU compared to background). Lastly, *CXCL1* and *CXCL2* interaction frequency increases during Rapamycin-treatment (2.21-fold, relative to proliferative) and is similar to proliferative cells during quiescence (1.09-fold, relative to proliferative) (Pro: 120 AU, Qui: 131 AU, Rap: 267 AU, compared to background).

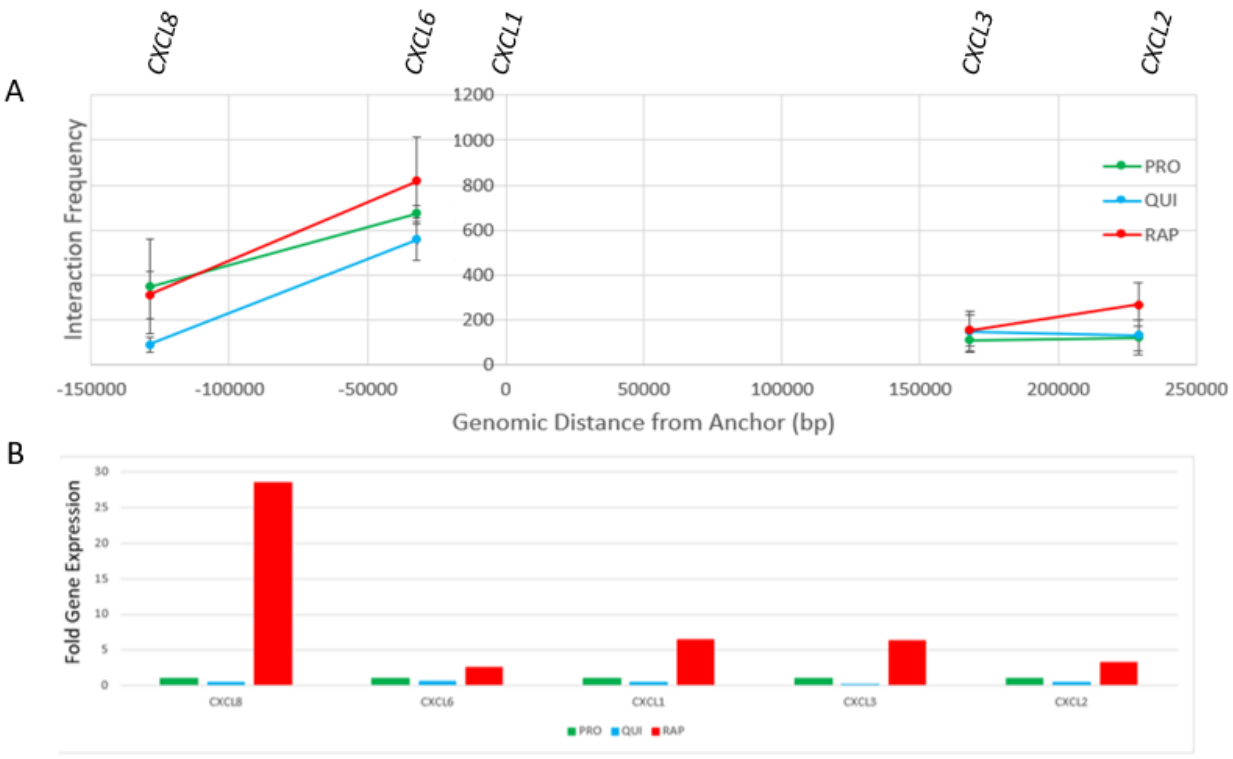


Figure 4-2: **3C profile of the *CXCL* locus on chromosome 4, region 74 Mbp with *CXCL1* as the anchor fragment. (A)** Graphed 3C profile of interactions occurring between the *CXCL1* gene and other *CXCL* genes nearby (*CXCL8*, *CXCL6*, *CXCL3*, *CXCL2*). Interaction Frequency is measured by fold comparison of detected products that have been normalized to a control interaction against background. Genomic distance is measured as the distance between the tested site and the anchor position (at '0' or the vertical axis). **(B)** Relative expression of several *CXCL* genes during three different treatment conditions as determined by RNA-Sequencing (Gillespie *et al*, 2015) (GEO Accession: GSE65145). Height of bars indicates fold change in gene expression compared to proliferative values. Pro = proliferative (10% FBS), Qui = quiescence (0.5% FBS), Rap = Rapamycin Treated (500 nM). (* = $P < 0.10$, ** = $P < 0.05$, Student's two-tailed t-test). (n = 5 biological replicates for each treatment).

Interaction (*CXCL1* anchor)	Distance from anchor (base pairs)	Proliferative (fold change versus proliferative)	Quiescence (fold change versus proliferative)	Rapamycin-treated (fold change versus proliferative)
CXCL1 x *CXCL8*	-128800	1.00	0.25*	0.89
CXCL1 x *CXCL6*	-32200	1.00	0.83	1.22
CXCL1 x *CXCL3*	168000	1.00	1.34	1.39
CXCL1 x *CXCL2*	229000	1.00	1.09	2.21

Table 4-2: CXCL1 anchored 3C investigations of chromosome 4 (74-75 Mbp) relative to proliferative samples. Table depicting interaction frequency relative to proliferative samples for examined 3C interactions on chromosome 4, region 74 Mbp between several chemokine *C-X-C motif ligand* (*CXCL*) genes. Interactions are defined by a pair of primers with one primer chosen as a constant or 'anchor' primer for the assays and another as a variable primer. Distance is measured from the site of the anchor primer to the site of the variable primer. Negative distances indicate the variable primer is upstream of the anchor primer. Positive distances indicate the variable primer is downstream of the anchor primer. Fold change is calculated as a ratio of interaction frequency compared to proliferative samples using the same primer pair. (** = $P < 0.05$, * = $P < 0.10$. Student's 2-tailed T-test.)

We investigated if interactions could be detected between the *CXCL2* gene (occurring in one putative TAD) and the nearby *epiregulin* (*EREG*) gene occurring in a second, neighboring, putative TAD **(Table 4-3)**. If an interaction was detectable between *AREG* or *EREG* and the *CXCL2* gene, it would suggest they may be occupying a single TAD rather than two TADs. We detected interaction frequencies occurring between *EREG* and *AREG* which supports the idea that *EREG* and *AREG* occupy the same TAD (Pro: 1.00-fold, Qui: 1.94-fold, Rap: 1.89-fold, relative to proliferative samples) (Pro: 3502-AU, Qui: 6780 AU, Rap: 6623 AU as compared to background). However, we were unable to detect interactions occurring between the *CXCL2* gene and AREG or EREG, suggesting *EREG* and *AREG* occupy a different

TAD than the *CXCL* genes of interest.

Interaction	Distance apart (base pairs)	Proliferative (fold change versus proliferative)	Quiescence (fold change versus proliferative)	Rapamycin-treated (fold change versus proliferative)
EREG X CXCL2	-280,000	N/A	N/A	N/A
EREG x AREG	-72,000	1.00	1.94	1.89

Table 4-3: **EREG anchored 3C investigations of chromosome 4 (74-75 Mbp) relative to proliferative samples.** Table depicting interaction frequency relative to proliferative samples for examined 3C interactions on chromosome 4, region 74 Mbp between the *EREG*, *AREG*, and *CXCL2* genes. Interactions are defined by a pair of primers with one primer chosen as a constant or 'anchor' primer for the assays and another as a variable primer. Distance is measured from the site of the anchor primer to the site of the variable primer. Negative distances indicate the variable primer is upstream of the anchor primer. Positive distances indicate the variable primer is downstream of the anchor primer. Fold change is calculated as a ratio of interaction frequency compared to proliferative samples using the same primer pair.

4.3 Discussion

Our analysis of the cluster of *chemokine* genes on chromosome 4, region 74 Mb, revealed several aspects of the local organization of this putative TAD. The initial analysis of interaction decay, using sites located at 5,000, 15,000, and 25,000 bp away from the *CXCL8*, displayed high interaction frequency at 5,000 bp (Average: 775-AU), lower interaction frequency at 15,000 bp (Average: 166-AU), and even lower interaction frequency at 25,000 bp in proliferative cells (Pro: 66-AU). Interestingly,

quiescent and Rapamycin-treated cells displayed increased interaction frequency (Qui: 4.12-fold, Rap: 5.96-fold relative to proliferative) between *CXCL8* and this site 25,000 bp away (Qui: 270 AU, Rap: 390 AU compared to background), and were forming functional but currently uncharacterized interactions.

For the *CXCL* genes of interest on chromosome 4, a complicated folding pattern of these region was identified. *CXCL8* interacted more strongly with *CXCL2* following quiescence induction; however, quiescent treatment decreased the interaction strength with *CXCL1*. This loss of interaction strength may be indicative of a gain of interaction with *CXCL8* with other segments of this regions (such as *CXCL2*) or *CXCL1* favoring other interacting partners. Rapamycin appeared to have a slight change in interaction strength with *CXCL2* but did not meet our criteria for significance. These observations suggest that this region of the genome might be already folded for the expression of these genes in response to rapamycin treatment, and that little re-organization is required for dramatic changes in expression. The significance of the relationship between *CXCL8* with *CXCL1* and *CXCL2* in response to quiescence remains to be elucidated. However, we did provide evidence that the *CXCL* genes are clustered in a TAD separate from the neighboring *EREG* and *AREG* genes.

While many interactions occur at high frequency during all growth conditions, our analyses revealed several aspects of the genes contained inside this putative TAD on chromosome 4, located between 74 and 75 Mbp. On average, we observed a higher interaction frequency for interactions in either proliferative or Rapamycin-treated cells, as compared to quiescent cells. Interestingly, proliferative and Rapamycin-treated cells displayed a similar organization, perhaps due to both growth conditions having above a certain level of transcription. In contrast, the lowered transcription during quiescence may be weakening these interaction frequencies. In the case of the increased interaction frequency between *CXCL8* and *CXCL2* in quiescent cells, this may be the result a different interaction forming. Each growth condition also displayed unique aspects. Both quiescent and Rapamycin-treated cells displayed increased interaction frequency between *CXCL8* and the RFS3 site ~25,000 bp upstream

of the *CXCL8* gene compared to proliferative cells. Quiescent cells uniquely featured a detectable interaction between *CXCL8* and *CXCL2* (360,000 bp apart) that was absent in both proliferative and Rapamycin-treated cells and had a significantly weakened interaction between *CXCL8* and *CXCL1* (128,800 bp). From these results it seems the organization of this region on chromosome 4 may alter during different growth conditions; however, some of the changes are subtle and are not simply the loss or gain of interactions.

5.0 Results: 3C Analysis of other loci

Following our analysis of the region on chromosome 4 containing five *chemokine* genes, we investigated two other regions of the genome encoding genes with differential expression in response to Rapamycin treatment or quiescence induction to determine if they also re-organized. The first region contained the *Leukemia inhibitory factor (LIF) gene* on chromosome 22, and the second, a cluster of *KRTAP* genes on chromosome 17, region 41 Mb. These regions were selected based on our previous observations that *LIF* was up-regulated 86-fold during treatment with 500 nM Rapamycin (Gillespie et al 2015) and upon examination of the region using the UCSC Genome Browser we discovered several nearby sites featuring histone 3 lysine 27 acetylation (H3K27Ac) marks that denote possible enhancers for the gene, as well we observed possible 'eRNA' (enhancer RNA) being produced at a site near the *LIF* gene in our RNA-Seq data (**Figure 5-1**). eRNA is produced from enhancers caught by transcriptional machinery and appear as short RNA fragments that match intergenic sequences (Rothschild and Basu, 2017). Considering the up-regulation of *LIF* during treatment with Rapamycin, this was an ideal opportunity to identify the location of this enhancer and confirm if the possible eRNA signal matches the location of this enhancer, and to observe the possible interaction between the promoter of *LIF* and its nearby enhancer during up-regulation. To discover the location of the LIF enhancer, LIF was used as the anchor and primers were designed against ten nearby sites that captured possible enhancers. These sites were selected by the presence of H3K27Ac (enhancer) marks in the UCSC Genome Browser (for various cell lines).

A region was discovered, containing multiple *KRTAP* genes that were up-regulated during Rapamycin treatment. Using computational tools that identified groups of genes responding to stimuli within a specific window distance, this region was identified on chromosome 17. Like our earlier analysis of the *chemokine* genes on chromosome 4, we hypothesized that co-transcription and localization of these genes to a transcription factory would be observable during treatment with Rapamycin. To

analyze the *KRTAP* region on chromosome 17, we utilized two genes as anchors, *KRTAP2-3* and *KRTAP4-5*.

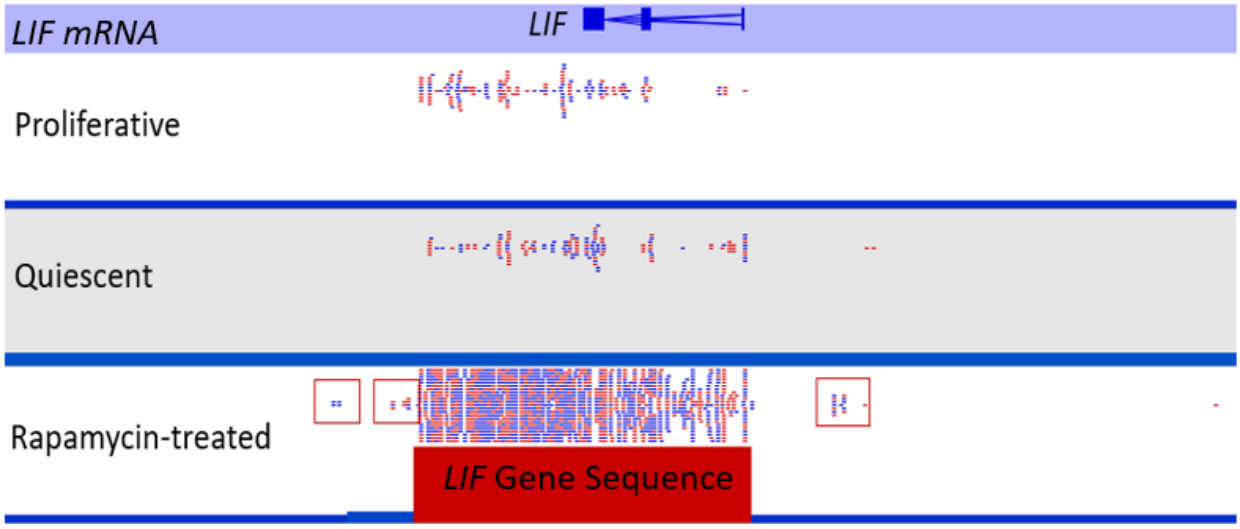


Figure 5-1: Screen capture of the Seq-Monk software used to analyze RNA-Seq data produced for proliferative, quiescent, and Rapamycin-treated 2DD cells (Gillespie *et al*, 2015) (GEO Accession: GSE65145). The image captures a section of chromosome 22, between 30,630,000 and 30,650,000 bp. Sequence read alignments are displayed as red and blue markings. The *LIF* gene is marked as the red box. The three treatment conditions sequence reads are contained in the rows marked 'Proliferative', 'Quiescent', and 'Rapamycin-treated'. Hollow red boxes have been placed over potential eRNA reads captured by RNA-sequencing in Rapamycin-treated samples.

5.1 LIF interacts with high frequency with two nearby putative enhancers

LIF provided an ideal candidate to search for a nearby enhancer as it is highly up-regulated during treatment with Rapamycin and featured nearby eRNA transcripts in our RNA-Seq data. I hypothesized that the Rapamycin mediated up-regulation was due to an interaction forming between

LIF and an appropriate enhancer. Using 3C, we discovered that three sites in the *LIF* locus displayed high interaction frequency compared to the rest of the locus (**Figure 5-2**). These sites occurred at 43,400 bp upstream and at 5,800 bp and 7,900 bp downstream of the gene. The downstream site 5,800 bp away decreases interaction frequency in both quiescent (0.10-fold, relative to proliferative) and Rapamycin-treated cells (0.12-fold, relative to proliferative) (Pro: 133 AU, Qui: 12.8 AU, Rap: 16.2 AU compared to background). Likewise, the downstream site 7,900 bp decreases interaction frequency in both quiescent (0.49-fold, relative to proliferative) and Rapamycin-treated cells (0.59-fold, relative to proliferative) (Pro: 133 AU, Qui: 12.8 AU, Rap: 16.2 AU). For the interaction between *LIF* and a site 43,400 bp upstream, the highest interaction was detected in proliferative cells, a weaker interaction frequency in Rapamycin-treated cells (0.67-fold, relative to proliferative), and a lower interaction frequency in quiescent cells (0.45-fold, relative to proliferative) (Pro: 246 AU, Qui: 110 AU, Rap: 164 AU compared to background).

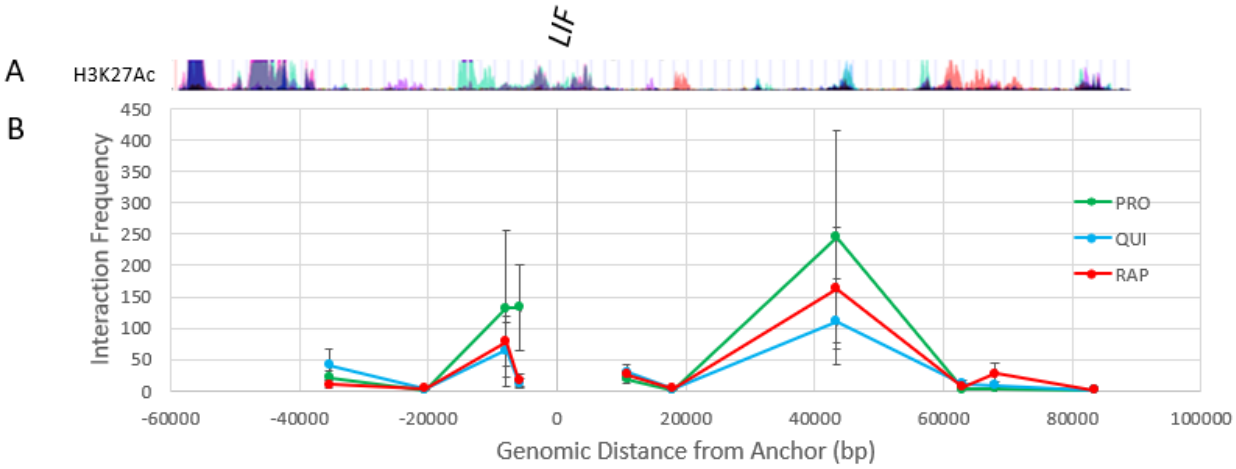


Figure 5-2: 3C profile of the *LIF* locus on chromosome 22 with *LIF*. (A) Screen-capture of UCSC Genome Browser featuring the H3K27Ac profile of all available cell types that was used when designing primers against putative enhancers. **(B)** Graphed 3C profile of interactions occurring between the *LIF* gene and several downstream and upstream putative enhancer sites. Interaction Frequency is measured by fold change comparison of detected products that have been normalized to a control interaction against background. Genomic distance is measured as the distance of the tested site versus the position of the

anchor (at the intercept or the vertical axis). Pro = proliferative (10% FBS), Qui = quiescence (0.5% FBS), Rap = Rapamycin Treated (500 nM). (* = P < 0.10, ** = P < 0.05, Student's two-tailed t-test). (n = 5 for each treatment).

Interaction (*LIF* anchor)	Distance from anchor (base pairs)	Proliferative (fold change versus proliferative)	Quiescence (fold change versus proliferative)	Rapamycin-treated (fold change versus proliferative)
LIF A1 x DS4	-35400	1.00	1.97	0.49
LIF A1 x DS3	-20500	1.00	1.55	1.99
LIF A1 x DS2	-7900	1.00	0.49	0.59
LIF A1 x DS1	-5800	1.00	0.10	0.12
LIF A1 x US2	10800	1.00	1.64	1.44
LIF A1 x US1	18000	1.00	2.05	3.42
LIF A1 x US3	43400	1.00	0.45	0.67
LIF A1 x US4	63000	1.00	3.55	2.06
LIF A1 x US5	68000	1.00	2.52	8.16
LIF A1 x US6	83400	1.00	1.00	1.00

Table 5-1: LIF anchored 3C investigations of chromosome 17 relative to proliferative samples. Table depicting interaction frequency relative to proliferative samples for examined 3C interactions on chromosome 22, between the *Leukemia inhibitory factor* (*LIF*) gene and several upstream (US) or downstream (DS) putative enhancer sites. Interactions are defined by a pair of primers with one primer chosen as a constant or 'anchor' primer for the assays and another as a variable primer. Distance is measured from the site of the anchor primer to the site of the variable primer. Negative distances indicate the variable primer is upstream of the anchor primer. Positive distances indicate the variable primer is downstream of the anchor primer. Fold change is calculated as a ratio of interaction frequency

compared to proliferative samples using the same primer pair. (** = P < 0.05, * = P < 0.10. Student's 2-tailed T-test).

5.2 KRTAP2-3 anchored investigations of the *KRTAP* locus on chromosome 17.

We identified a cluster of k*eratin associated proteins (KRTAP)* on chromosome 17, region 41 Mbp that was differentially regulated in response to quiescence-induction and Rapamycin-treatment, and; therefore, a candidate region for investigating sub-chromosomal re-organization in response to gene regulation. In response to Rapamycin treatment many of these genes are up-regulated (*KRTAP3-1, KRTAP4-7, KRTAP4-11, KRTAP4-12, KRTAP4-6, KRTAP4-5, KRTAP4-4)*, and during quiescence many of these genes are down-regulated (*KRTAP1-5*, *KRTAP1-1*, *KRTAP2-3, KRTAP2-4, KRTAP4-12*). Similar to our investigation of the cluster of *chemokine* genes on chromosome 4, we investigated if these genes were coming together for transcription during gene up-regulation, and further if they were disassociating during down-regulation. *KRTAP2-3* was selected as the anchor, which is down-regulated during quiescence.

Our 3C analysis (**Figure 5-3**) revealed that *KRTAP2-3* maintains several high frequency interactions with upstream genes of the *KRTAP* cluster on chromosome 17, as well as several weaker interactions with downstream genes. *KRTAP2-3* displayed a high interaction frequency with *KRTAP3-1* (50,100 bp away) in proliferative cells; however, this interaction frequency significantly declined (P < 0.10, student's two-tailed t-test) for both quiescent (0.20-fold, relative to proliferative) and Rapamycin-treated cells (0.19-fold, relative to proliferative) (Pro: 2786 AU, Qui: 558 AU, Rap: 542 AU compared to background). *KRTAP2-3* also displayed high interaction frequency with *KRTAP1-5* (39,500 bp away) in proliferative cells, and again displayed decreased interaction frequency for both quiescence (0.30-fold, relative to proliferative) and Rapamycin treated cells (0.29, relative to proliferative) (Pro: 2281 AU, Qui: 681 AU, Rap: 661 AU, compared to background). *KRTAP2-3* and *KRTAP1-1* (27,100 bp apart) displayed

the highest interaction frequency in Rapamycin-treated cells, weaker for proliferative, and the lowest for quiescence cells (Pro: 3761 AU, Qui: 2906 AU, Rap: 4535 AU compared to background). Upstream of *KRTAP2-3*, interaction frequencies were detected for several genes. *KRTAP2-3* displayed a unique interaction with *KRTAP2-4* (6,500 bp away) only in quiescent and Rapamycin-treated cells (Pro: 1 AU, Qui: 403 AU, Rap: 418 AU compared to background). *KRTAP2-3* also displayed detectable interactions with the distal *KRTAP4-5* (87,800 bp away) with the highest interaction frequency in proliferative cells and lower interaction frequencies in quiescence (0.29-fold, relative to proliferative) and Rapamycin-treated cells (0.41-fold, relative to proliferative) (Pro: 272 AU, Qui: 80.0 AU, Rap: 112 AU compared to background). *KRTAP2-3* interacted with *KRTAP4-4* (97,200 bp away) with the highest interaction frequency again in proliferative cells and lower interaction frequencies during quiescence (0.20-fold, relative to proliferative) and Rapamycin-treatment (0.16-fold, relative to proliferative) (Pro: 329 AU, Qui: 66.1 AU, Rap: 53.0 AU, compared to background).

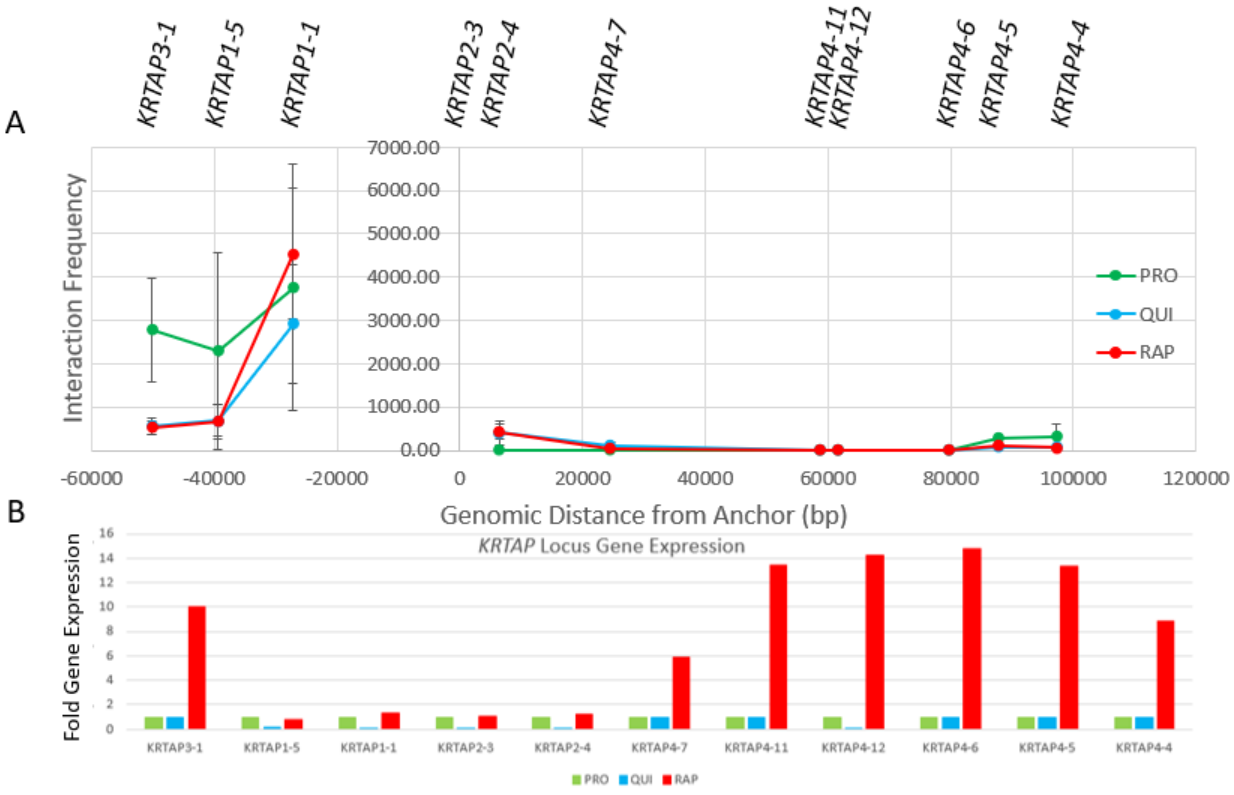


Figure 5-3: 3C profile of the *KRTAP* locus on chromosome 17 with *KRTAP2-3* as the anchor fragment. **(A)** Graphed 3C profile of interactions occurring between the *KRTAP2-3* (green box above graph) gene and several other *KRTAP* genes (blue boxes above graph). Interaction Frequency is measured by fold comparison of detected products that have been normalized to a control interaction against background. Genomic distance is measured as the distance of the tested site versus the position of the anchor (at '0' or the vertical axis). **(B)** Relative expression of several *KRTAP* genes during three different treatment conditions as determined by RNA-Sequencing data (Gillespie *et al*, 2015) (GEO Accession: GSE65145). Y-axis indicates fold change in gene expression compared to proliferative values. Pro = proliferative (10% FBS), Qui = quiescence (0.5% FBS), Rap = Rapamycin Treated (500 nM). (* = P < 0.10, ** = P < 0.05, Student's two-tailed t-test). (n = 5 for each treatment).

Interaction (KRTAP2-3 anchor)	Distance from anchor (base pairs)	Proliferative (fold change versus proliferative)	Quiescence (fold change versus proliferative)	Rapamycin-treated (fold change versus proliferative)
KRTAP2-3 x *KRTAP3-1*	-50125	1.00	0.20*	0.19*
KRTAP2-3 x *KRTAP1-5*	-39352	1.00	0.30	0.29
KRTAP2-3 x *KRTAP1-1*	-27145	1.00	0.77	1.21
KRTAP2-3 x *KRTAP2-4*	6467	1.00	559.91	581.40*
KRTAP2-3 x *KRTAP4-7*	24503	1.00	110.85	50.99
KRTAP2-3 x *KRTAP4-11*	58770	1.00	1.00	1.00
KRTAP2-3 x *KRTAP4-12*	61603	1.00	1.00	1.00
KRTAP2-3 x *KRTAP4-6*	79650	1.00	1.00	1.00
KRTAP2-3 x *KRTAP4-5*	87816	1.00	0.29	0.41
KRTAP2-3 x *KRTAP4-4*	97249	1.00	0.20	0.16

Table 5-2: KRTAP2-3 anchored 3C investigations of chromosome 17 relative to proliferative samples. Table depicting interaction frequency relative to proliferative samples for examined 3C interactions on chromosome 17 between *keratin associated protein* (*KRTAP*) genes. Interactions are defined by a pair of

primers with one primer chosen as a constant or 'anchor' primer for the assays and another as a variable primer. Distance is measured from the site of the anchor primer to the site of the variable primer. Negative distances indicate the variable primer is upstream of the anchor primer. Positive distances indicate the variable primer is downstream of the anchor primer. Fold change is calculated as a ratio of interaction frequency compared to proliferative samples using the same primer pair. (** = $P < 0.05$, * = $P < 0.10$. Student's 2-tailed T-test).

5.3 KRTAP4-5 anchored investigations of the *KRTAP* locus on chromosome 17

We chose a second anchor to further investigate the *KRTAP* locus on chromosome 17, region 41 Mbp. *KRTAP4-5* was selected as an anchor due to its up-regulation during Rapamycin treatment (**Figure 5-4, Table 5-3**). By using a second anchor, we allow the identification of anchor bias and a characterization of interactions from a different viewpoint that will further instruct on local interactions. *KRTAP4-5* displayed high interaction frequencies with *KRTAP3-1* (137,900 bp away) with the highest frequency occurring in proliferative cells, weaker frequency in Rapamycin-treated cells (0.26-fold, relative to proliferative), and the lowest interaction frequency in quiescent cells (0.26-fold, relative to proliferative) (Pro: 647 AU, Qui: 268 AU, Rap: 166 AU, compared to background). *KRTAP4-5* displayed the highest interaction frequencies with *KRTAP1-1* (115,000 bp away). KRTAP4-5 to KRTAP1-1 interaction frequency was highest in proliferative, lower in quiescent cells (0.73-fold, relative to proliferative), and the lowest in Rapamycin-treated cells (0.56-fold, relative to proliferative) (Pro: 1719 AU, Qui: 1560 AU, Rap: 956 AU compared to background). *KRTAP4-5* also displayed interaction frequencies with *KRTAP2-3* (87,800 bp) (discussed previously) and *KRTAP4-4* (8,200 bp away) (Pro: 144 AU, Qui: 129 AU, Rap: 88.4 AU compared to background).

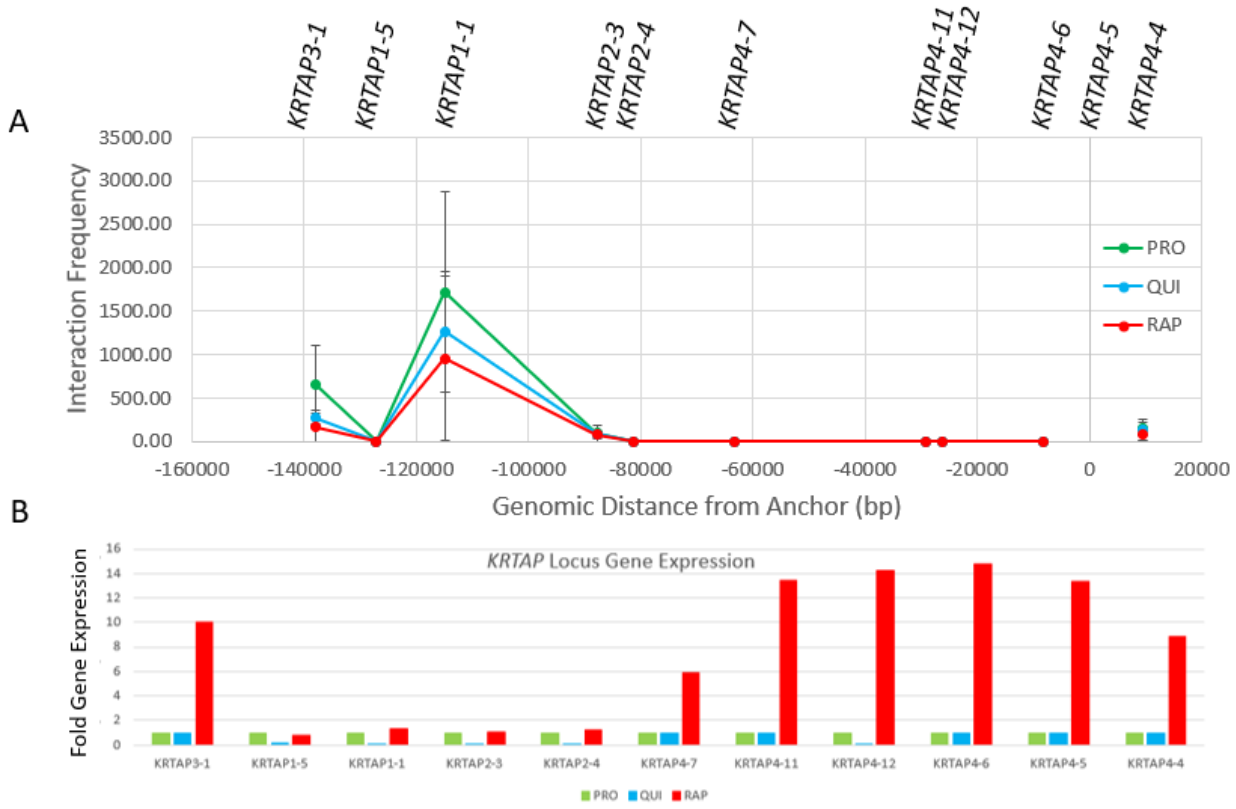


Figure 5-4: 3C profile of the *KRTAP* locus on chromosome 17 with *KRTAP4-5* as the anchor fragment. **(A)** Graphed 3C profile of interactions occurring between the *KRTAP4-5* (green box above graph) gene and several other *KRTAP* genes (blue boxes above graph). Interaction Frequency is measured by fold comparison of detected products that have been normalized to a control interaction against background. Genomic distance is measured as the distance of the tested site versus the position of the anchor (at '0' or the vertical axis). **(B)** Relative expression of several *KRTAP* genes during three different treatment conditions as determined by RNA-Seq (Gillespie *et al*, 2015) (GEO Accession: GSE65145). Height of bars indicates fold change in gene expression compared to proliferative values. Pro = proliferative (10% FBS), Qui = quiescence (0.5% FBS), Rap = Rapamycin Treated (500 nM).

Interaction (*KRTAP4-5* anchor)	Distance from anchor (base pairs)	Proliferative (fold change versus proliferative)	Quiescence (fold change versus proliferative)	Rapamycin-treated (fold change versus proliferative)
KRTAP4-5 x *KRTAP3-1*	-137941	1.00	0.41	0.26
KRTAP4-5 x *KRTAP1-5*	-127168	1.00	1.00	1.00
KRTAP4-5 x *KRTAP1-1*	-114961	1.00	0.73	0.56
KRTAP4-5 x *KRTAP2-3*	-87816	1.00	0.87	0.82
KRTAP4-5 x *KRTAP2-4*	-81349	1.00	1.00	1.00
KRTAP4-5 x *KRTAP4-7*	-63313	1.00	1.00	1.00
KRTAP4-5 x *KRTAP4-11*	-29046	1.00	1.00	1.00
KRTAP4-5 x *KRTAP4-12*	-26213	1.00	1.00	1.00
KRTAP4-5 x *KRTAP4-6*	-8166	1.00	1.00	1.00
KRTAP4-5 x *KRTAP4-4*	9433	1.00	0.89	0.61

Table 5-3: KRTAP4-5 anchored 3C investigations of chromosome 17 relative to proliferative samples. Table depicting interaction frequency relative to proliferative samples for examined 3C interactions on chromosome 17 between *keratin associated protein* (*KRTAP*) genes. Interactions are defined by a pair of primers with one primer chosen as a constant or 'anchor' primer for the assays and another as a variable primer. Distance is measured from the site of the anchor primer to the site of the variable primer. Negative distances indicate the variable primer is upstream of the anchor primer. Positive distances indicate the variable primer is downstream of the anchor primer. Fold change is calculated as a ratio of interaction frequency compared to proliferative samples using the same primer pair.

5.4 Discussion

In our analysis of gene architecture near the *LIF* gene, during quiescence we identified an expected decrease in interaction frequency with nearby sites compared to proliferative. However, we

also observed an unexpected decrease in interaction frequency during Rapamycin treatment when the gene is up-regulated 86-fold and expected to increase interaction frequency with these putative enhancers. While the explanation for this decrease in interaction frequency between *LIF* and nearby putative enhancers during Rapamycin treatment is unexpected, it is possible that these interactions are occurring at lower frequency due to the formation of new, as-of-yet unidentified interactions. For practicality, we only searched a limited distance upstream and downstream of the *LIF* gene for a possible enhancer. Certainly, more work must be performed on the region to characterize the functional implications of these changes, such as characterizing its relationship with other putative enhancers (H3K27Ac marks) in the region and discovering the protein factors responsible for mediating looping between *LIF* and regulatory elements. One relationship that can be drawn from this investigation is that during decreased nutrient available and growth signaling (quiescence) and perceived decreased nutrient availability (through inhibition of mTORC1 by Rapamycin), interaction frequency between LIF and these targeted putative enhancer sites decrease compared to proliferative cells.

In the *KRTAP* locus on chromosome 17, we discovered that proliferative samples displayed the highest interaction frequencies despite Rapamycin-treated samples having the highest expression of the contained *KRTAP* genes. Interestingly, we discovered similar interaction profiles when comparing quiescent and Rapamycin-treated samples, despite the opposite levels of transcription of the contained genes. Some portions of the *KRTAP* locus produced no interactions with either anchor despite up-regulation of genes during Rapamycin treatment (*KRTAP4-7*, *KRTAP4-11*, *KRTAP4-12*, *KRTAP4-6*, *KRTAP4-4*). This could be due to several reasons, namely that either interactions are not occurring between our anchors (*KRTAP2-3*, *KRTAP4-5*) and those genes, or that the difficult nature of primer design for the region precluded successful investigation of interactions. Upon further examination of this region using Hi-C, these concerns of primer design would be alleviated by a sequencing-based approach rather than a PCR-based method.

Similar to results with the *LIF* locus, interaction frequencies of the *KRTAP* locus tended to decrease for both quiescent and Rapamycin-treated samples compared to proliferative samples. As mentioned previously, this decrease of interaction frequency may be the result of decreased nutrient and growth signals or the perceived decrease in nutrients due to mTORC1 inhibition. This is not unreasonable when considering that resources would be either limited, or perceived as limited, for the cell to accomplish normal function, and may result in decreasing interaction frequency to accommodate energy contexts. However, this fails to explain why several genes are still capable of considerable up-regulation during treatment with Rapamycin, when actual energy is available, though downstream signaling of mTORC1 is inhibited.

6.0 Conclusions

Our experiments sought to elaborate the relationship between growth conditions and nutrient sensing to alterations in genome organization, focusing on genes up-regulated by treatment with 500 nM Rapamycin. By comparing genome organization between three different growth conditions; proliferative, quiescence induction (serum reduction), and Rapamycin-treatment (500 nM), we sought to understand if reorganization of the genome would occur at the level of chromosome territories as well as at the sub-chromosomal level.

Our analysis of chromosome territory positioning by FISH, in response to our three conditions, yielded insights into the complicated relationship between genome organization and gene expression in response. To our surprise, and despite significant changes of the gene expression profile induced by quiescence induction (0.5% FBS) and Rapamycin-treatment (500 nM), we observed only minimal changes in the positioning of chromosome 4 which contained many genes with altered regulation. As expected with previous results, we recapitulated earlier observations that chromosome 18 and 10 reposition in response to treatment with Rapamycin and induction into quiescence by serum starvation.

During our analysis of sub-chromosomal organization using 3C against the *CXCL* genes of interest, we discovered unique interactions when comparing cells under proliferative status versus quiescent induction or Rapamycin treatment. Previously, quiescence and Rapamycin treatment were thought to parallel each other, as they both adopted a flattened cellular morphology with decreased cellular proliferation during treatment. Work performed previously was the first to demonstrate these two treatments, though appearing similar at first, are quite distinct. The findings presented in this thesis agree with these earlier findings and further elaborate on the subtle organizational changes that occur concomitant with changes to gene expression. Significantly, our investigation of the *CXCL* locus on chromosome 4, region 74 Mbp, revealed a complex relationship between transcription and interaction frequency of the contained genes. For example, despite *CXCL8* and *CXCL2* being down-regulated during

quiescence, *CXCL8* and *CXCL2* displayed an increased interaction frequency compared to other treatments. Also interesting, both quiescence-induced and Rapamycin-treated cells displayed increased interaction frequency between *CXCL8* and a site occurring 25,000 base pairs downstream of the gene, despite opposite levels of regulation for the gene (Pro: 1-fold, Qui: 4.12-fold, Rap: 5.96-fold, relative to the proliferative state). We also observed that many genes maintained interactions during altered gene expression, such as between *CXCL8* and *CXCL1,* however proliferative and Rapamycin-treated cells often displayed the highest interaction frequency between genes – to some degree supporting the notion that increased transcription correlates with increased interaction frequency.

Outside of chromosome 4 we also investigated two cases of interest: the *LIF* gene on chromosome 22 and another cluster of genes (*KRTAP*) contained on chromosome 17, region 41 Mbp. Analysis of the *LIF* locus, in search for its enhancer(s), revealed high frequency interactions occurring between the gene and three nearby sites. These sites, occurring at ~5000 and ~8000 base pairs downstream and ~40,000 base pairs upstream of the gene, were both observed with the highest interaction frequency in proliferative cells, and at high frequency in both quiescent and Rapamycin-treated cells. Interestingly, quiescent cells also displayed an increased interaction frequency with a site ~35,000 bp downstream of the *LIF* gene. These results are complicated by the observation of lower interaction frequency between *LIF* and these putative enhancers during Rapamycin-treatment as compared to proliferative cells. As anticipated, if these sites are enhancers of the LIF gene, an increase in interaction frequency would be observed. Alternately, it is possible that these sites are alternate enhancers to a more effective enhancer, or that during up-regulation these interactions are weakened to favor another interaction. Currently the exact effect of these interactions are unclear, however they were observed occurring in all three conditions and most highly in proliferative cells.

During our analysis of the *KRTAP* locus on chromosome 17, containing eleven *KRTAP* genes, we discovered differential organization of the region when comparing among treatment conditions.

Proliferative cells again represented the highest frequency of interactions, as seen in the interactions occurring between *KRTAP2-3* and *KRTAP1-1, KRTAP2-3* and *KRTAP1-5,* and *KRTAP2-3* and *KRTAP3-1*. The interactions between *KRTAP2-3* and *KRTAP1-5*, or *KRTAP2-3* and *KRTAP3-1,* were observed occurring at lower frequencies during quiescence and Rapamycin-treatment. When analyzing the region from a different viewpoint, *KRTAP4-5*, we observed a high interaction frequency between *KRTAP4-5* and *KRTAP1-1* at nearly 110,000 base pairs away, as well as an interaction occurring between *KRTAP4-5* and *KRTAP3-1* at nearly 140,000 base pairs away. Again, we observed the highest interaction frequency occurring in proliferative cells compared to the other two treatments. This is interesting due to Rapamycin-treated cells and quiescent cells having opposite levels of expression for many of these genes, but similarly decreased interaction frequency profiles for the *KRTAP* locus.

These findings indicate a complexity connecting genome organization and gene expression, in response to our proliferative, quiescent, and Rapamycin-treated conditions. Several possibilities exist to explain the trends observed in the presented data.

Firstly, it is unclear if the organization of the *CXCL* locus and *KRTAP* locus is centered on a transcription factory. During all three growth conditions, we observed interactions occurring between many of the *CXCL* genes, with some changes occurring during quiescent treatment. The *KRTAP* locus did display increased interaction frequency between *KRTAP2-3* and *KRTAP3-1* in proliferative cells, however transcription of the *KRTAP* genes was highest during treatment with rapamycin. Overall, the anticipated relationship between increased transcription during Rapamycin treatment, and increased interaction between the genes inside our loci of interest (*CXCL* and *KRTAP* loci), was not observed.

Secondly, as none of the genes (*CXCL*, *KRTAP*, or *LIF*) are ever completely silenced, it is possible that even low levels of transcription result in a transcriptionally permissive arrangement of the regions. Possibly, the regions are poised for transcription and interactions are maintained while awaiting transcriptional activation.

Lastly, it seems possible that higher access to nutrients and other cellular resources may facilitate increased interaction frequency. In general, proliferative cells often displayed the highest interaction frequency between two sites, whereas quiescent and Rapamycin-treated cells displayed lower interaction frequencies. Quiescent and Rapamycin-treated cells both have reduced use of the mTORC1 signalling pathway, due to either less nutrients/resources available or inhibition of the pathway by Rapamycin. This may result in a more specialized interactions forming, compared to the proliferative state which may adopt a more elastic or promiscuous state made possible by (perceived) higher resource availability and the wide-spread transcription promotion of many genes by the mTORC1 pathway.

In summary, while we observed altered interactions in our treatments, a direct relationship between increased transcription and increased interaction, or decreased transcription and decreased interaction, was not clear. Possibly, there are interactions altering in response to our treatments which were not captured by our investigations. It does remain clear, with the repositioning of chromosome 10 and 18, as well as our observed interaction changes, that the genome is reorganizing in response to these quiescence and Rapamycin treatment.

How these findings relate to the ageing process or genome organization as a whole are complicated. Caloric restriction and Rapamycin-treatment have been observed extending the health span of several model organisms (Bjedov et al., 2010; Ehninger et al., 2014; Harrison et al., 2009) and here we have shown that genome organization varies between our three treatment conditions (proliferative, quiescence, and Rapamycin-treated).

7.0 Future Direction

Future investigations stemming from this project would naturally involve the use of two high-throughput techniques to characterize genome organization in healthy cells, aged cells, and cells being treated with Rapamycin to reverse ageing. Firstly, analysis of whole genome organization by Hi-C would allow the characterization of lost and gained interactions as a result of treatment with the putative anti-ageing drug Rapamycin, and perhaps give insight into regulatory mechanisms beneficial to mitigating the aging process, and mechanisms that worsen or incur ageing. Secondly, use of ChIP-Seq to characterize histone marks (such as H3K27Ac, a mark of enhancers) or transcription factors of interest (such as STAT5A/B – potential activator of many of the up-regulated *CXCL* genes) could elaborate on the mechanisms facilitating genome organization and reorganization. In characterizing these realms thoroughly both the alterations to organization and the effectors or organization could be identified. By understanding the alterations acquired in the ageing process, and the mitigating changes induced by Rapamycin treatment or other drugs, ageing mechanisms can be characterized. This could then be expanded into understanding how nutrition and nutrient sensing affect health and life span and how this can be used practically.

Characterizing genome organization in the context of ageing is fundamental to the process of combating it. In cells of patients with HGPS, a malformed nucleus is observed caused by inappropriate function of the LMNA/C protein, with clear consequences on nuclear organization and organismal health such as reduced DNA repair. As the symptoms of HGPS are highly similar to the normal ageing process, albeit at a much faster rate (HGPS sufferers often only live into mid-teens), it is readily apparent that genome organization has links to ageing phenotypes. This disorganization is caused by the accumulation of an aberrant LMNA/C protein inside the nucleus. Hope for treatment of HGPS lies in inducing the increased recycling of these gradually accumulating proteins that may allow the nucleus to return to a healthy phenotype with appropriate organization. Currently, investigations are underway in the Eskiw

laboratory to induce this recycling by restricting essential amino acids and inducing cells to recycle the aberrant and accumulated LMNA/C proteins.

Following the characterization of alterations to genome topology in response to Rapamycin treatment or dietary changes in normal healthy cells, a 'healthy and young' model of genome organization could be established. As cells age (and potentially devolve into a cancerous lineage), gene expression can alter inappropriately in response to genome changes, such as mutations, translations, deletions, or alterations to chromatin and epigenetics marks. In characterizing either young phenotype, or an anti-aging phenotype, and aged phenotypes, the alterations that occur in older and degrading cells may become readily apparent and emerge as targets to mitigate aging. By identifying and remediating these age-based alterations to the genome and genome topology, cellular health and age phenotype may be reversible to a normal status through the tailored use of drugs, diet, and gene editing in cellular models.

Once these anti-aging mechanisms are established in a cell model, they could be elevated for use in mice or another model organism. Whole organismal models provide significant boundaries compared to cellular models due to the complexity, variety, and accessibility to various tissues. A significant benefit of mediating ageing phenotypes through diet, is the effectiveness and tissue reach of nutrients. While drugs are limited by bioavailability, all cells require a certain upkeep of nutrients and actively ensure they can be provided for the cell. Nutrition provides a 'highly penetrative' approach to tackling ageing, since cells actively absorb nutrients. Diet can also be highly controlled by mediating ingestion of targeted amino acids, or other substances such as polyphenols. Perhaps by characterizing the genome topology changes of ageing, and understanding how nutrition and nutritional response drive topology changes, moderate and practical changes to diet can extend health and lifespan as well as combat other disease states.

8.0 References

Abraham, R.T., and Wiederrecht, G.J. (1996). Immunopharamcology of Rapamycin. Annu Rev Immunol *14*, 483-510.

Akeno, N., Miller, A.L., Ma, X., and Wikenheiser-Brokamp, K.A. (2015). p53 suppresses carcinoma progression by inhibiting mTOR pathway activation. Oncogene *34*, 589-599.

Albiez, H., Cremer, M., Tiberi, C., Vecchio, L., Schermelleh, L., Dittrich, S., Kupper, K., Joffe, B., Thormeyer, T., von Hase, J.*, et al.* (2006). Chromatin domains and the interchromatin compartment form structurally defined and functionally interacting nuclear networks. Chromosome Res *14*, 707-733.

Allen, B.L., and Taatjes, D.J. (2015). The Mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol *16*, 155-166.

Arlia-Ciommo A., Leonov, A., Mohammad, K., Beach, A., Richard, V., Bourque, S., Burstein, M., Goldberg, A., Kyryakov, P., Gomez-Perez, A., Koupaki O., Titorenko, V. (2018). Mechanisms through which lithocholic acid delays yeast chronological aging under caloric restriction conditions. Oncotarget *89*, 34945-34971.

Aymard, F., Aguirrebengoa, M., Guillou, E., Javierre, B.M., Bugler, B., Arnould, C., Rocher, V., Iacovoni, J.S., Biernacka, A., Skrzypczak, M.*, et al.* (2017). Genome-wide mapping of long-range

contacts unveils clustering of DNA double-strand breaks at damaged active genes. Nat Struct Mol Biol *24*, 353-361.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Barkess, G., and West, A.G. (2012). Chromatin insulator elements: establishing barriers to set heterochromatin boundaries. Epigenomics *4*, 67-80.

Barolo, S. (2012). Shadow enhancers: frequently asked questions about distributed cis-regulatory information and enhancer redundancy. Bioessays *34*, 135-141.

Barutcu, A.R., Fritz, A.J., Zaidi, S.K., van Wijnen, A.J., Lian, J.B., Stein, J.L., Nickerson, J.A., Imbalzano, A.N., and Stein, G.S. (2016). C-ing the Genome: A Compendium of Chromosome Conformation Capture Methods to Study Higher-Order Chromatin Organization. J Cell Physiol *231*, 31-35.

Bauer, N.C., Doetsch, P.W., and Corbett, A.H. (2015). Mechanisms Regulating Protein Localization. Traffic *16*, 1039-1061.

Belyaeva, A., Venkatachalapathy, S., Nagarajan, M., Shivashankar, G.V., and Uhler, C. (2017). Network analysis identifies chromosome intermingling regions as regulatory hotspots for transcription. Proc Natl Acad Sci U S A, 1-6.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat Rev Genet *15*, 163-175.

Bjedov, I., Toivonen, J.M., Kerr, F., Slack, C., Jacobson, J., Foley, A., and Partridge, L. (2010). Mechanisms of life span extension by rapamycin in the fruit fly Drosophila melanogaster. Cell Metab *11*, 35-46.

Blackwood, E.M., and Kadonaga, J.T. (1998). Going the Distance: A Current View of Enhancer Action. Science *281*, 60-65.

Blagosklonny, M. (2010). Calorie Restriction: decelerating mTORC-driven aging from cells to organisms (including humans). Cell Cycle *4*, 683-688.

Bolland, D.J., Wood, A.L., Afshar, R., Featherstone, K., Oltz, E.M., and Corcoran, A.E. (2007). Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. Mol Cell Biol *27*, 5523-5533.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R., *et al.* (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol *3*, 1-17.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol *4*, 1-9.

Bridger, J., and Bickmore, W. (1998). Putting the genome on the map. Elsevier *14*, 403-410.

Bridger, J.M., Kill, I.R., O'Farrell, M., and Hutchison, C.J. (1993). Internal lamin structures within G1 nuclei of human dermal fibroblast. Journal of Cell Science *104*, 297-306.

Brookes, E., and Pombo, A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. EMBO Rep *10*, 1213-1219.

Brugarolas, J., Lei, K., Hurley, R.L., Manning, B.D., Reiling, J.H., Hafen, E., Witters, L.A., Ellisen, L.W., and Kaelin, W.G., Jr. (2004). Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. Genes Dev *18*, 2893-2904.

Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A., and Felsenfeld, G. (2002). The insulation of genes from external enhancers and silencing chromatin. Proc Natl Acad Sci U S A *99 Suppl 4*, 16433-16437.

Cannavo, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Korbel, J.O., and Furlong, E.E. (2016). Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Curr Biol *26*, 38-51.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.

Carter, D.R., Eskiw, C., and Cook, P.R. (2008). Transcription factories. Biochem Soc Trans *36*, 585-589.

Chantranupong, L., Scaria, S.M., Saxton, R.A., Gygi, M.P., Shen, K., Wyant, G.A., Wang, T., Harper, J.W., Gygi, S.P., and Sabatini, D.M. (2016). The CASTOR Proteins Are Arginine Sensors for the mTORC1 Pathway. Cell *165*, 153-164.

Chaumeil, J., and Skok, J.A. (2012). The role of CTCF in regulating V(D)J recombination. Curr Opin Immunol *24*, 153-159.

Cheng, M.K., and Disteche, C.M. (2004). Silence of the fathers: early X inactivation. Bioessays *26*, 821-824.

Clamp, M., Fry, B., Kamal, M., Xie, X., Cuff, J., Lin, M.F., Kellis, M., Lindblad-Toh, K., and Lander, E.S. (2007). Distinguishing protein-coding and noncoding genes in the human genome. Proc Natl Acad Sci U S A *104*, 19428-19433.

Colman, R., Anderson, R., Johnson, S., Kastman, K., Beasley, T., Allison, D., Cruzen, C., Summons, H., Kemnitz, J., Weindruch, R. (2009) Caloric restriction delays disease onset and mortality in rhesus monkeys. Science *5937*, 201-204.

Cope, N.F., Fraser, P., and Eskiw, C.H. (2010). The yin and yang of chromatin spatial organization. Genome Biol *11*, 204-212.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J.*, et al.* (2008). Structure of eukaryotic RNA polymerases. Annu Rev Biophys *37*, 337-352.

Cremer, M., Kupper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., and Cremer, T. (2003). Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. J Cell Biol *162*, 809-820.

Cremer, T., and Cremer, C. (2001). Chromosome Territories, nuclear architecture and gene regulation in mammalian cells. Nature (Genetics) *2*, 292-302.

Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harb Perspect Biol *2*, 1-23.

Cremer, T., Cremer, M., and Cremer, C. (2018). The 4D Nucleome: Genome Compartmentalization in an Evolutionary Context. Biochemistry (Mosc) *83*, 313-325.

Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I., and Fakan, S. (2006). Chromosome territories--a functional nuclear landscape. Curr Opin Cell Biol *18*, 307-316.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W. (1999). Differences in the Localization and Morphology of Chromosomes in the Human Nucleus. Journal of Cell Biology *145*, 1119-1131.

Dechat, T., Adam, S.A., and Goldman, R.D. (2009). Nuclear lamins and chromatin: when structure meets function. Adv Enzyme Regul *49*, 157-166.

Dehghani, H., Dellaire, G., and Bazett-Jones, D.P. (2005). Organization of chromatin in the interphase mammalian cell. Micron *36*, 95-108.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science *295*, 1306-1313.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Doerks, T., Copley, R.R., Schultz, J., Ponting, C.P., and Bork, P. (2002). Systematic identification of novel protein domain families associated with nuclear functions. Genome Res *12*, 47-56.

Dumont, F.J.S., Q. (1995). Mechanism of action of the immunosuppressant Rapamycin. Elsevier *58*, 373-395.

Dundr, M. (2012). Nuclear bodies: multifunctional companions of the genome. Curr Opin Cell Biol *24*, 415-422.

Dunn, K.L., and Davie, J.R. (2003). The many roles of the transcriptional regulator CTCF. Biochem Cell Biol *81*, 161-167.

Ehninger, D., Neff, F., and Xie, K. (2014). Longevity, aging and rapamycin. Cell Mol Life Sci *71*, 4325-4346.

Eskiw, C.H., Cope, N.F., Clay, I., Schoenfelder, S., Nagano, T., and Fraser, P. (2010). Transcription Factories and Nuclear Organization of the Genome. Cold Spring Harbor Laboratory Press *LXXW*, 501-508.

Eskiw, C.H., and Fraser, P. (2011). Ultrastructural study of transcription factories in mouse erythroblasts. J Cell Sci *124*, 3676-3683.

Eskiw, C.H., Rapp, A., Carter, D.R., and Cook, P.R. (2008). RNA polymerase II activity is located on the surface of protein-rich transcription factories. J Cell Sci *121*, 1999-2007.

Feng, Z., Zhang, H., Levine, A.J., and Jin, S. (2005). The coordinate regulation of the p53 and mTOR pathways in cells. Proc Natl Acad Sci U S A *102*, 8204-8209.

Floutsakou, I., Agrawal, S., Nguyen, T.T., Seoighe, C., Ganley, A.R., and McStay, B. (2013). The shared genomic architecture of human nucleolar organizer regions. Genome Res *23*, 2003-2012.

Fontana, L., Partridge, L. (2015) Promoting health and longevity through diet: from model organisms to humans. Cell *161*, 106-118.

Frock, R.L., Hu, J., and Alt, F.W. (2015). Mechanisms of Recurrent Chromosomal Translocations. 27-51.

Gillespie, Z.E., MacKay, K., Sander, M., Trost, B., Dawicki, W., Wickramarathna, A., Gordon, J., Eramian, M., Kill, I.R., Bridger, J.M.*, et al.* (2015). Rapamycin reduces fibroblast proliferation without causing quiescence and induces STAT5A/B-mediated cytokine production. Nucleus *6*, 490-506.

Gilmore, A., Redman, L. (2018) Calorie restriction for human aging: is there a potential benefit for cancer? Mol Cell Oncol *5*, e1481811.

Giorgio, E., Robyr, D., Spielmann, M., Ferrero, E., Di Gregorio, E., Imperiale, D., Vaula, G., Stamoulis, G., Santoni, F., Atzori, C.*, et al.* (2015). A large genomic deletion leads to enhancer adoption by the lamin B1 gene: a second path to autosomal dominant adult-onset demyelinating leukodystrophy (ADLD). Hum Mol Genet *24*, 3143-3154.

Goodrich, J.A., and Tjian, R. (1994). Transcription Factors IIE and IIH and ATP Hydrolysis Direct Promoter Clearance by RNA Polymerase II. Cell *17*, 145-156.

Graziani, E.I. (2009). Recent advances in the chemistry, biosynthesis and pharmacology of rapamycin analogs. Nat Prod Rep *26*, 602-609.

Guo, C., Yoon, H.S., Franklin, A., Jain, S., Ebert, A., Cheng, H.L., Hansen, E., Despo, O., Bossen, C., Vettermann, C*., et al.* (2011). CTCF-binding elements mediate control of V(D)J recombination. Nature *477*, 424-430.

Gwinn, D.M., Shackelford, D.B., Egan, D.F., Mihaylova, M.M., Mery, A., Vasquez, D.S., Turk, B.E., and Shaw, R.J. (2008). AMPK phosphorylation of raptor mediates a metabolic checkpoint. Mol Cell *30*, 214-226.

Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forne, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc *2*, 1722-1733.

Hancock, R. (2014). Structures and functions in the crowded nucleus: new biophysical insights. Frontiers in Physics *2*, 1-17.

Harrison, D.E., Strong, R., Sharp, Z.D., Nelson, J.F., Astle, C.M., Flurkey, K., Nadon, N.L., Wilkinson, J.E., Frenkel, K., Carter, C.S*., et al.* (2009). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature *460*, 392-395.

Hasty, P., Sharp, Z.D., Curiel, T.J., and Campisi, J. (2013). mTORC1 and p53: clash of the gods? Cell Cycle *12*, 20-25.

Herold, M., Bartkuhn, M., and Renkawitz, R. (2012). CTCF: insights into insulator function during development. Development *139*, 1045-1057.

Holwerda, S.J., and de Laat, W. (2013). CTCF: the protein, the binding partners, the binding sites and their chromatin loops. Philos Trans R Soc Lond B Biol Sci *368*, 1-8.

Holz, M.K., and Blenis, J. (2005). Identification of S6 kinase 1 as a novel mammalian target of rapamycin (mTOR)-phosphorylating kinase. J Biol Chem *280*, 26089-26093.

Horton, W., Miyashita, T., Kohno, K., Hassel, J.R., and Yamada, Y. (1987). Identification of a phenotype-specific enhancer in the first intron of the rat collagen II gene. Proc. Nati. Acad. Sci. USA *84*, 8864-8868.

Huang, J., and Manning, B.D. (2008). The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. Biochem J *412*, 179-190.

Iborra, F., Pombo, A., Jackson, D., and Cook, P. (1996). Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. Journal of Cell Science *109*, 1427-1437.

Inoki, K., Ouyang, H., Zhu, T., Lindvall, C., Wang, Y., Zhang, X., Yang, Q., Bennett, C., Harada, Y., Stankunas, K.*, et al.* (2006). TSC2 integrates Wnt and energy signals via a coordinated phosphorylation by AMPK and GSK3 to regulate cell growth. Cell *126*, 955-968.

Inukai, S., Kock, K.H., and Bulyk, M.L. (2017). Transcription factor-DNA binding: beyond binding site motifs. Curr Opin Genet Dev *43*, 110-119.

Jewell, J.L., Kim, Y.C., Russell, R.C., Yu, F.X., Park, H.W., Plouffe, S.W., Tagliabracci, V.S., and Guan, K.L. (2015). Metabolism. Differential regulation of mTORC1 by leucine and glutamine. Science *347*, 194-198.

Jewell, J.L., Russell, R.C., and Guan, K.L. (2013). Amino acid signalling upstream of mTOR. Nat Rev Mol Cell Biol *14*, 133-139.

Jung, C.H., Ro, S.H., Cao, J., Otto, N.M., and Kim, D.H. (2010). mTOR regulation of autophagy. FEBS Lett *584*, 1287-1295.

Jung, D., Giallourakis, C., Mostoslavsky, R., and Alt, F.W. (2006). Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annu Rev Immunol *24*, 541-570.

Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol *339*, 225-229.

Kapoor, P., Chen, M., Winkler, D.D., Luger, K., and Shen, X. (2013). Evidence for monomeric actin function in INO80 chromatin remodeling. Nat Struct Mol Biol *20*, 426-432.

Katahira, J. (2015). Nuclear export of messenger RNA. Genes (Basel) *6*, 163-184.

Kim, M.S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S.*, et al.* (2014). A draft map of the human proteome. Nature *509*, 575-581.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231-1245.

Kim, T.K., Hemberg, M., and Gray, J.M. (2015). Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. Cold Spring Harb Perspect Biol *7*, 1-3.

Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. J Cell Biol *159*, 777-782.

Kitada, M., Koya, D. (2013). The use of calorie restriction mimetics to study aging. Methods Mol Biol *1048*, 95-107.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kramer, J.M. (2016). Regulation of cell differentiation and function by the euchromatin histone methyltranserfases G9a and GLP. Biochem Cell Biol *94*, 26-32.

Kugel, J.F.G., J. A. (2017). Finding the start site: redefining the human initiator element. Genes Dev *31*, 6-11.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650-665.

Lando, D., Basu, S., Stevens, T.J., Riddell, A., Wohlfahrt, K.J., Cao, Y., Boucher, W., Leeb, M., Atkinson, L.P., Lee, S.F*., et al.* (2018). Combining fluorescence imaging with Hi-C to study 3D genome architecture of the same single cell. Nat Protoc *13*, 1034-1061.

Li, J., Kim, S.G., and Blenis, J. (2014). Rapamycin: one drug, many effects. Cell Metab *19*, 373-379.

Lieberman-Aiden, E., (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326,* 289-293.

Lin, S.G., Guo, C., Su, A., Zhang, Y., and Alt, F.W. (2015). CTCF-binding elements 1 and 2 in the Igh intergenic control region cooperatively regulate V(D)J recombination. Proc Natl Acad Sci U S A *112*, 1815-1820.

Liu, B., Wang, J., Chan, K.M., Tjia, W.M., Deng, W., Guan, X., Huang, J.D., Li, K.M., Chau, P.Y., Chen, D.J*., et al.* (2005). Genomic instability in laminopathy-based premature aging. Nat Med *11*, 780-785.

Loewith, R., and Hall, M.N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. Genetics *189*, 1177-1201.

Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R*., et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012-1025.

Lupianez, D.G., Spielmann, M., and Mundlos, S. (2016). Breaking TADs: How Alterations of Chromatin Domains Result in Disease. Trends Genet *32*, 225-237.

Magnuson, B., Ekim, B., and Fingar, D.C. (2012). Regulation and function of ribosomal protein S6 kinase (S6K) within mTOR signalling networks. Biochem J *441*, 1-21.

Mahy, N.L., Perry, P.E., Gilchrist, S., Baldock, R.A., and Bickmore, W.A. (2002). Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. J Cell Biol *157*, 579-589.

Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nat Rev Genet *11*, 285-296.

Marnef, A., and Legube, G. (2017). Organizing DNA repair in the nucleus: DSBs hit the road. Curr Opin Cell Biol *46*, 1-8.

Martina, J.A., Chen, Y., Gucek, M., and Puertollano, R. (2012). MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB. Autophagy *8*, 903-914.

Matoulkova, E., Michalova, E., Vojtesek, B., and Hrstka, R. (2012). The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. RNA Biol *9*, 563-576.

Mehta, I.S., Amira, M., Harvey, A.J., and Bridger, J.M. (2010). Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. Genome Biol *11*, 1-17.

Melnik, S., Deng, B., Papantonis, A., Baboo, S., Carr, I.M., and Cook, P.R. (2011). The proteomes of transcription factories containing RNA polymerases I, II or III. Nat Methods *8*, 963-968.

Miles, J., Mitchell, J.A., Chakalova, L., Goyenechea, B., Osborne, C.S., O'Neill, L., Tanimoto, K., Engel, J.D., and Fraser, P. (2007). Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus. PLoS One *2*, 630-642.

Nativio, R., Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.M., and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS Genet *5*, 1-15.

Naumova, N., Smith, E.M., Zhan, Y., and Dekker, J. (2012). Analysis of long-range chromatin interactions using Chromosome Conformation Capture. Methods *58*, 192-203.

Noda, T. (2017). Regulation of Autophagy through TORC1 and mTORC1. Biomolecules *7*.

Oh, W.J., and Jacinto, E. (2011). mTOR complex 2 signaling and functions. Cell Cycle *10*, 2305-2316.

Ong, C.T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet *15*, 234-246.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W.*, et al.* (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065-1071.

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., and Fraser, P. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol *5*, 1-10.

Palazzo, A.G., R (2014). The Case for Junk DNA. Plos Genet *10*, 1-8.

Papantonis, A., Larkin, J.D., Wada, Y., Ohta, Y., Ihara, S., Kodama, T., and Cook, P.R. (2010). Active RNA polymerases: mobile or immobile molecular machines? PLoS Biol *8*, 1-10.

Pederson, T. (2011). The nucleolus. Cold Spring Harb Perspect Biol *3*, 1-15.

Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. Nat Rev Genet *14*, 288-295.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y.*, et al.* (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281-1295.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Pink, R.C., Eskiw, C.H., Caley, D.P., and Carter, D.R. (2010). Analysis of beta-globin chromatin micro-environment using a novel 3C variant, 4Cv. PLoS One *5*, 1-8.

Pontvianne, F., Carpentier, M.C., Durut, N., Pavlistova, V., Jaske, K., Schorova, S., Parrinello, H., Rohmer, M., Pikaard, C.S., Fojtova, M.*, et al.* (2016). Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the A. thaliana Genome. Cell Rep *16*, 1574-1587.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K.*, et al.* (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402-405.

Raab, J.R., and Kamakaka, R.T. (2010). Insulators and promoters: closer than we think. Nat Rev Genet *11*, 439-446.

Redwood, A.B., Perkins, S.M., Vanderwaal, R.P., Feng, Z., Biehl, K.J., Gonzalez-Suarez, I., Morgado-Palacin, L., Shi, W., Sage, J., Roti-Roti, J.L.*, et al.* (2011). A dual role for A-type lamins in DNA double-strand break repair. Cell Cycle *10*, 2549-2560.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev *23*, 1247-1269.

Rieder, D., Trajanoski, Z., and McNally, J.G. (2012). Transcription factories. Front Genet *3*, 221.

Rothschild, G., and Basu, U. (2017). Lingering Questions about Enhancer RNA and Enhancer Transcription-Coupled Genomic Instability. Trends Genet *33*, 143-154.

Rubio, E.D., Reiss, D.J., Welcsh, P.L., Disteche, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A *105*, 8309-8314.

Sarbassov, D.D., Ali, S.M., Sengupta, S., Sheen, J.H., Hsu, P.P., Bagley, A.F., Markhard, A.L., and Sabatini, D.M. (2006). Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. Mol Cell *22*, 159-168.

Sarvaiya, P., J., Guo, D., Ulasov, I., Gabikian, P., and Lesniak, M. Chemokines in tumor progression and metastasis. Oncotarget *4*, 2171-2186.

Savarese, F., and Grosschedl, R. (2006). Blurring cis and trans in gene regulation. Cell *126*, 248-250.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A *103*, 1412-1417.

Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell *168*, 960-976.

Schindler, C., Levy, D.E., and Decker, T. (2007). JAK-STAT signaling: from interferons to cytokines. J Biol Chem *282*, 20059-20063.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S.*, et al.* (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet *42*, 53-61.

Sehgal, N., Baker, H., Vezina, C. (1975). Rapamycin (AY-22,989), a new fungal antibiotic. II. Fermentation, isolation and characterization. Journal of Antibiotics (Tokyo) *10*, 727-732.

Sehgal, N., Fritz, A.J., Vecerova, J., Ding, H., Chen, Z., Stojkovic, B., Bhattacharya, S., Xu, J., and Berezney, R. (2016). Large-scale probabilistic 3D organization of human chromosome territories. Hum Mol Genet *25*, 419-436.

Shandilya, J., and Roberts, S.G. (2012). The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim Biophys Acta *1819*, 391-400.

Shchuka, V.M., Malek-Gilani, N., Singh, G., Langroudi, L., Dhaliwal, N.K., Moorthy, S.D., Davidson, S., Macpherson, N.N., and Mitchell, J.A. (2015). Chromatin Dynamics in Lineage Commitment and Cellular Reprogramming. Genes (Basel) *6*, 641-661.

Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nat Rev Mol Cell Biol *18*, 655-670.

Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. Curr Opin Cell Biol *21*, 344-351.

Smale, S.T., and Baltimore, D. (1989). The initiator as a transcription control element. Cell Press *57*, 103-113.

Sohal, R., Forster, M. (2014). Caloric restriction and the aging process: a critique. Free Radic Biol Med *73*, 366-382.

Solovei, I., Kreysing, M., Lanctot, C., Kosem, S., Peichl, L., Cremer, T., Guck, J., and Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell *137*, 356-368.

Spann, T.P., Goldman, A.E., Wang, C., Huang, S., and Goldman, R.D. (2002). Alteration of nuclear lamin organization inhibits RNA polymerase II-dependent transcription. J Cell Biol *156*, 603-608.

Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. Cold Spring Harb Perspect Biol *3*.

Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat Immunol *5*, 1017-1027.

Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature *435*, 637-645.

Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat Rev Genet *13*, 613-626.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev *20*, 2349-2354.

Sutherland, H., and Bickmore, W.A. (2009). Transcription factories: gene expression in unions? Nat Rev Genet *10*, 457-466.

Tiwari, V.K., and Baylin, S.B. (2009). Combined 3C-ChIP-cloning (6C) assay: a tool to unravel protein-mediated genome architecture. Cold Spring Harb Protoc *2009*, 1-10.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., and Laat, W.d. (2002). Looping and Interaction between Hypersensitive Sites in the Active B-globin Locus. Mol Cell Biol *10*, 1453-1465.

Travers, A., and Muskhelishvili, G. (2015). DNA structure and function. FEBS J *282*, 2279-2295.

Ulianov, S.V., Khrameeva, E.E., Gavrilov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M.V., Chertovich, A.*, et al.* (2016). Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. Genome Res *26*, 70-84.

van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780-791.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., and Luscombe, N.M. (2009). A census of human transcription factors: function, expression and evolution. Nat Rev Genet *10*, 252-263.

Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. Nat Rev Mol Cell Biol *16*, 178-189.

Villicaña, C., Cruz, G., and Zurita, M. (2014). The basal transcription machinery as a target for cancer therapy. BMC Cancer Cell *14*, 1-15.

Vogelmann, J., Valeri, A., Guillou, E., Cuvier, O., and Nollmann, M. (2011). Roles of chromatin insulator proteins in higher-order chromatin organization and transcription regulation. Nucleus *2*, 358-369.

Watson, J.D., and Crick, F.H. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribonucleic Acids. Nature, 737-738.

Wolfson, R.L., Chantranupong, L., Saxton, R.A., Shen, K., Scaria, S.M., Cantor, J.R., and Sabatini, D.M. (2016). METABOLISM Sestrin2 is a leucine sensor for the mTORC1 pathway. Science *351*, 43-48.

Wraught, B., and Gingras, A.C. (1999). eILF4 activity is regulated at multiple levels. Journal of Biochemistry and Cell Biology *31*, 43-57.

Xu, M., and Cook, P.R. (2008). Similar active genes cluster in specialized transcription factories. J Cell Biol *181*, 615-623.

Yanga, C., Bolotina, E., Jiang, T., Sladek, F.M., and Martinez, E. (2007). Prevalence of the Initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene 53-65.

Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature *456*, 125-129.

Zoncu, R., Efeyan, A., and Sabatini, D.M. (2011). mTOR: from growth signal integration to cancer, diabetes and ageing. Nat Rev Mol Cell Biol *12*, 21-35.

Zorn, C., Cremer, T., Cremer, C., and Zimmer, J.r. (1976). Laser UV microirradiation of interphase nuclei and post-treatment with caffeine. Human Genetics *35*, 83-89.