2019

Characterization Of Epigenetic Plasticity And Chromatin Dynamics In Cancer Cell Models

Diana Lea Gerrard
University of Vermont

CHARACTERIZATION OF EPIGENETIC PLASTICITY AND CHROMATIN DYNAMICS IN CANCER CELL MODELS

A Dissertation Presented

by

Diana Lea Gerrard

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfilment of the Requirements
For the Degree of Doctor of Philosophy
Specializing in Cellular, Molecular, and Biomedical Sciences

May, 2019

Defense Date: March 8, 2019
Dissertation Examination Committee:

Seth Frietze, Ph.D., Advisor
Sayyed Kaleem Zaidi, Ph.D., Chairperson
Frances Carr, Ph.D.
Jessica Heath, M.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

Cancer progression is driven by cumulative changes that promote and maintain the malignant phenotype. Epigenetic alterations are central to malignant transformation and to the development of therapy resistance. Changes in DNA methylation, histone acetylation and methylation, noncoding RNA expression and higher-order chromatin structures are epigenetic features of cancer, which are independent of changes in the DNA sequence. Despite the knowledge that these epigenetic alterations disrupt essential pathways that protect cells from uncontrolled growth, how these modifications collectively coordinate cancer gene expression programs remains poorly understood. In this dissertation, I utilize molecular and informatic approaches to define and characterize the genome-wide epigenetic patterns of two important human cancer cell models. I further explore the dynamic alterations of chromatin structure and its interplay with gene regulation in response to therapeutic agents.

In the first part of this dissertation, pancreatic ductal adenocarcinoma (PDAC) cell models were used to characterize genome-wide patterns of chromatin structure. The effects of histone acetyltransferase (HAT) inhibitors on chromatin structure patterns were investigated to understand how these potential therapeutics influence the epigenome and gene regulation. Accordingly, HAT inhibitors globally target histone modifications and also impacted specific gene pathways and regulatory domains such as super-enhancers. Overall, the results from this study uncover potential roles for specific epigenomic domains in PDAC cells and demonstrate epigenomic plasticity to HAT inhibitors.

In the second part of this dissertation, I investigate the dynamic changes of chromatin structure in response to estrogen signaling over a time-course using Estrogen Receptor (ER) positive breast cancer cell models. Accordingly, I generated genome-wide chromatin contact maps, ER, CTCF and regulatory histone modification profiles and compared and integrated these profiles to determine the temporal patterns of regulatory chromatin compartments. The results reveal that the majority of alterations occur in regions that correspond to active chromatin states, and that dynamic chromatin is linked to genes associated with specific cancer growth and metabolic signaling pathways. To distinguish ER-regulated processes in tamoxifen-sensitive and in tamoxifen-resistant (TAMR) cell models, we determined the corresponding chromatin and gene expression profiles using ER-positive TAMR cancer cell derivatives. Comparison of the patterns revealed characteristic features of estrogen responsiveness and show a global reprogramming of chromatin structure in breast cancer cells with acquired tamoxifen resistance.

Taken together, this dissertation reveals novel insight into dynamic epigenomic alterations that occur with extrinsic stimuli and provides insight into mechanisms underlying the therapeutic responses in cancer cells.

CITATIONS

Materials from this dissertation have been published in the following form:

Gerrard, D. L., Hawkinson, A., Sherman, T., Modahl, C. M., Hume, G., Campbell, C. L., . . . Frietze, S.. (2017). Transcriptomic Signatures of Tacaribe Virus-Infected Jamaican Fruit Bats. *mSphere*, *2*(5). doi:10.1128/mSphere.00245-17

Gerrard, D. L., Wang, Y., Gaddis, M., Zhou, Y., Wang, J., Witt, H., . . . Frietze, S. E.. (2018). Three-dimensional analysis reveals altered chromatin interaction by enhancer inhibitors harbors TCF7L2-regulated cancer gene signature. *J Cell Biochem*. doi:10.1002/jcb.27449

AND

Materials from this dissertation are "in press" for publication in the following form:

Gerrard, D. L., Zhou, Y., Wang, J., Li, T., Yang, Y., Fritz, A.J., . . . Frietze, S., Jin, V.X.. (2019). Temporal dynamic reorganization of 3D chromatin architecture in hormone induced breast cancer and endocrine resistance. Nature Communications.

AND

Materials from this dissertation are "in review" for publication in the following form:

Gerrard, D.L., Boyd, J.R., Stein, G.S., Jin, V.X., and Frietze, S.. (2019). Disruption of broad epigenetic domains in PDAC cells by HAT inhibitors. Scientific Reports.

DEDICATION

I dedicate this dissertation to my father, Master Chief Leslie James Gerrard (retired), for his never-ending support and encouragement; my sister, Dr. Crystal Lynn Gerrard, for her guidance through dedication and perseverance; and my mother, Florentina Cook, for her persistence in love and internal value for the educational opportunities available in the USA. You have each contributed a significant amount of yourselves to me, which has built me into who I am and have yet to become. Your never ending commitment to me and passion for what you love inspires me.

ACKNOWLEDGEMENTS

First, I must express my immense appreciation for my dissertation mentor, Dr. Seth Frietze, who has selflessly invested in my education while I aimed to establish myself and grow as a research scientist. Thank you, Seth, for always giving me opportunity and allowing me to advance the skills and expertise that I need and desire to build my future. You have been supportive and a great captain for my journey. I am forever grateful for the lessons you have given me within intellectual scientific creativity, technical skills and laboratory management. The mentorship opportunities you have trusted me with allowed me to foster the concept of being a collaborative scientist. I also want to extend my gratitude for the rest of the Frietze family. Having had roots in Vermont, you made sure that this Texan settled properly and have continued to support me throughout my time in the Frietze laboratory. I am so honored to have been trained and mentored by you, Seth, and feel pride for the opportunity to represent you in my future career.

I am extremely fortunate to have had Dr. Sayyed Kaleem Zaidi as a mentor throughout my PhD studies as the chair of my committee. Kaleem is the most holistic mentor I could have ever dreamed of having throughout these last five years and I am confident that he will continue to play an integral part of my future endeavors. Not only is he an inspirational scientist who provides critical evaluation and comprehensive perspectives on scientific concepts, but he also has inspired me to constantly challenge my perception of my life as a scientist within the bigger picture. Kaleem, I really want to express my sincere gratitude to you for continuously providing a listening ear and for always being

ready to follow that listening with VERY VALID ‘reality checks’. You are an extremely great mentor and I am forever grateful for you. Dr. Jessica Heath and Dr. Frances Carr are two Women in Science that I have been blessed to be mentored by. Dr. Heath’s immense dedication to her career as a Physician and a Research Scientist is truly inspiring. Thank you, Jess, for always igniting the confidence within me to strive for greatness. Dr. Carr has dedicated her life to several different facets of the scientific community and has thrived as an eminent researcher. Thank you Fran for always being receptive of ideas and for providing valuable feedback, encouragement and for continuously challenging me to consider my future.

I am grateful to have shared this process with two other graduate students in the Frietze lab, Michael Mariani and Princess Rodriguez. Thank you Mike for encouraging me daily and for helping in many Tron ways. My uttermost gratitude goes to Princess. You have been the best ‘ride or die’ and I am completely grateful for the support we have given each other through the last four years. The effort and the stamina needed to complete these projects was definitely inspired by your encouragement. We have done this together and I cannot wait to continue to encourage you throughout the rest of your PhD journey and celebrate with you at the end. As an honorary mention- Dr. Phyu Thwe has been the best graduate student mentor for me. I am forever indebted to you for your guidance!

Lastly, thank you to my partner Nathan Horsfall, who has been a part of this entire journey from applications to two cross-country moves. I am forever grateful for you.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

CHAPTER 1

Comprehensive Literature Review

1.1: Basic principles underlying genome organization

Research over many decades has enhanced our knowledge concerning eukaryotic genome organization. Genome organization refers to the structural orientation of deoxyribonucleic acid (DNA) and is highly controlled. The discovery of the structure of DNA in the 1950s [1] provided the platform for later investigations regarding its mechanisms within genetic inheritance, determination of cellular fates and contributions to disease phenotypes. Other studies in the 1950s-1960s elucidated significant aspects of DNA structure including the ratio of bases [2] and the determination of DNA as the 'backbone' of chromatin structure [1, 3]. After the discovery of its involvement in inheritance during the 19th century, its rediscovery inspired researchers to expand on this fundamental knowledge and to further investigate DNA's regulation and influences on cellular mechanisms. This became possible as new revelations discovered the genome to be densely organized inside of the approximate 10 μm nucleus [4]. In this section, I provide an overview of key concepts governing genome organization including DNA structure, epigenetics and gene expression.

1.1.1: The nucleosome as the fundamental subunit of DNA

Scientific investigations in the 1970s paved the way for succeeding chromatin research. In the early 1970s, chromatin fibers were discovered to be approximately 100 angstroms thick [5, 6] and in 1975 the structure of this fiber was confirmed after earlier propositions acknowledging that DNA was wrapped around eight histone proteins in repeating units, comprising chromatin [7]. These units, called nucleosomes, were first visualized by Oudet et al. with an electron microscope. They achieved these after

depleting lysine-rich histones, which allowed them to observe chromatin structure as a replicating unit of spherical particles [7]. The molecular structure of the nucleosome was subsequently solved [8, 9] and research began to reveal its significance in biological outcomes including transcription, DNA repair and cell cycle processes [10-12]. Elucidating the organization of the tightly packed chromatin within the nucleus provided the foundation needed to explore chromatin regulation in cellular mechanisms.

1.1.2: Gene expression

As the central dogma of molecular biology highlights, underlying information contained within DNA has the potential to be transferred to RNA followed by protein through successive steps (transcription and translation, respectively) [13]. The complexity of how these processes are achieved and the particular outcomes as a result, extends far beyond this basic fundamental principle. Underlying this complexity is temporal and spatial control of genes relative to the rest of the genome. This determines the accessibility of genes and the ability of regulatory protein complexes to bind those regions in order to undergo transcription. RNA polymerase proteins are the core of transcriptional machinery. There are many different types of RNA polymerases with distinct roles regarding the types of genes they are able to transcribe [14]. While RNA polymerases I and III transcribe ribosomal RNAs and transfer RNAs/small RNAs respectively [15-17], RNA polymerase II regulates protein coding genes and long non-coding RNA via sophisticated control steps assuring correct structure at the site of transcription and allows for correction of mismatched bases [18-20]. Additionally, RNA polymerase II coordinates with core general transcription factors, which acts to help RNA

polymerase II localize to transcriptional initiation regions known as promoters to carry out these activities [21].

A different type of transcription factor class contains proteins that are sequence specific and have the ability to bind to DNA binding domains such as the High Mobility Group-box (HMG) domain [22, 23]. This ubiquitous protein domain is found in many transcription factors including the Sry-type HMG box (SOX) family of transcription factors, which play transcriptional roles in development and differentiation [24]. Additionally, transcription factors have other effector domains allowing for various other functions; for example, the recruitment of other regulatory complexes [25]. The spatial and temporal control of the genome also contributes to the ability of these factors to recognize their DNA binding sites and carry out their given roles. To this end, there are far less transcription factors within the human genome than there are genes and so they function via combinatorial mechanisms and make up transcriptional regulatory networks [22]. The effector function of transcription factors can vary and dictate the regulatory outcome. For instance, they have the capability of activating or repressing gene expression programs through the recruitment of chromatin-modifying enzymes and other cofactors that influence DNA accessibility directly impacting transcription [26].

1.1.3: Foundational principles of epigenetics

Epigenetic mechanisms play key roles in chromatin accessibility. In regard to the term itself, 'epigenetics' was coined by C.H. Waddington around 1940-1950 [27]. While the field of genetics encompasses research elucidating gene-based mechanisms regarding heritability, epigenetics extends beyond this notion to not only include heritable changes

but rather concerns modifications occurring along the genome that do not include alterations to DNA sequences themselves. Consequences of research efforts over many years has revealed modifications influencing transcriptional control including DNA methylation, post translational histone modifications, chromatin architecture and non-coding RNAs to be among these epigenetic aspects [28-32]. DNA methylation was identified in the 1940s and refers to the transfer of methyl groups to CpG dinucleotides by DNA methyltransferase enzymes [33]. Mechanistically, this modification can result in gene silencing by blocking the binding of transcriptional regulators or by recruiting gene repressive complexes [34-37]. Histone modifications incorporate different edits to amino acid tails of histone proteins some of which include acetylation, methylation, ubiquitination and phosphorylation. These modifications can either be transferred on or removed and can further act as docking sites for chromatin regulatory proteins and transcriptional complexes modulating chromatin accessibility and gene regulation [38]. Importantly, epigenetic regulation contributes to the heterogeneity of cellular functions as a result of gene expression alterations, in part, due to these modifications.

1.2: Higher order chromatin organization

Aspects underlying chromatin architecture within the nucleus and its contributions to cellular phenotypes are still largely unknown. After the identification of 3-dimensional chromatin structure within the nucleus, researchers over the last several decades have desired to elucidate fundamental principles regarding the nuclear organization of chromatin. Moreover, questions within these explorations have included: (1) In the small nuclear environment, how is chromatin organized and why does it take on

this organization?; (2) What molecular consequences does chromatin organization play in driving differentiation and development?; (3) Where are the genome-wide chromosome contacts in different cell types and what implications do these have on gene expression programs?; and (4) What are the implications of chromatin organization on diseased outcomes? In this section, I highlight principles of chromatin architecture and its importance in normal cellular phenotypes as well as cancer.

1.2.1: Fundamental principles of chromatin architecture

Chromatin regulation is a dynamic process by which fluctuations between condensed and relaxed chromatin states have been recognized to impact various cellular processes. Among these mechanisms are DNA repair [39], transcription [40] and cell division [41]. Cells must maintain genomic structural integrity and functional identity throughout successive generations to prevent transformations into an aberrant phenotype [42, 43]. Unveiling 3-D chromosome organization within the nucleus is crucial for understanding direct structural mechanisms underlying gene regulation and other cell-type specification processes.

It was first proposed in 1885 by Carl Rabl that chromosomes exist in distinct chromosome territories, which was later supported with experimental findings conducted by Cremer et al. [44, 45]. How the chromosomes are oriented during interphase is important for transcriptional regulatory targets and therefore directly influences cellular phenotypic outcome. The radial positioning of chromosomes in the nucleus and the relationship between these territories and gene density have been determined in the interphase nucleus [46]. To this end, it is well accepted that these chromosome territories

are organized such that gene-dense chromosomes are positioned interiorly [45], allowing for long-range interactions between genomic regions and can help dictate the phenotypic outcome of the cell [47]. For example, a long range single nucleotide polymorphism region was shown to interact with the oncogenic gene c-MYC through long-range chromatin interaction in colorectal cancer cells, ultimately playing a role in the upregulation of this cancer driver [48].

Further elucidation of structures contributing to chromosome architecture and organization within the nucleus revealed additional components including: the nucleolus, transcriptional complexes, histone locus bodies, heterochromatin and euchromatin [49]. These discoveries were products of combinatorial efforts involving traditional microscopic techniques as well as newly developed molecular tools (discussed in Chapter 1.5). From here, the interest arose in focusing efforts to identify specific domains associated with distinct territories in the nucleus.

Among the different types of associating contacts identified were lamina associated domains, nucleolar associated domains and topologically associated domains [50]. While nucleolar and lamina associated domains involve the interaction of chromosomes with nuclear components, topological associated domains refer to DNA-DNA contacts. In addition to the identification of different interacting regions, genomic compartments were discovered. These spatially segregated compartments along the genome were defined at a 1 Mb resolution and are categorized into either compartment types A or B [47]. These compartments were coordinated with epigenetic and gene expression programs, allowing for further characterization of these regions. Type A compartments were found to be enriched in genes and associated with open chromatin

regions; therefore, compartment A is commonly referred to as an open compartment. In contrast, compartment B was determined to be more densely packed and is referred to as a closed compartment. Understanding these compartments have revealed functional relationships between chromatin structure and gene activity. The topologically associated domains mentioned previously were identified within these compartments [51]. Moreover, topologically associated domains are genomic regions with self-interacting chromatin regions. The frequency of interactions within 100 kb along the genome revealed dense regions of self-interactions. Observations of these topological associated domains also revealed regions along the chromatin where interaction density abruptly ends. These regions, termed boundary regions, separate topologically associated domains and contain high levels of the transcription factor CCCTC-binding factors (CTCF) [52] in addition to other chromatin remodeling proteins, condensin and cohesin [53].

Long-range chromatin interactions are involved in cis gene regulatory programs by facilitating enhancer-promoter interactions [54]. After the discovery of enhancer regions, it was postulated that enhancer regions interact with promoter regions to regulate gene expression. Their interaction is achieved through chromatin looping as these regulatory regions are far apart from each other on a one-dimensional level [55, 56]. For example, with the goal of defining mechanisms underlying a gene relevant to kidney cancer, Moisan et al., identified enhancer elements that were localized to the promoter region of the *PKD2* gene via CTCF stabilization mediated by chromatin looping in renal cancer cells [57].

Together the identification of compartments and topological associated domains have laid the groundwork needed for understanding spatial organization of chromosomes

and its regulatory link to functional biological outcomes. The architecture of chromosomes during interphase is important for mediating maintenance of cellular genomes and gene expression regulation. Current efforts seek to further elucidate mechanisms that these architectural elements play in normal and diseased processes.

1.2.2: Gene regulation mediated by higher order chromatin organization in development

While chromosome territories have been determined to display differences in cell types through repositioning [57, 58], topologically associated domains are largely conserved between species and have shown little variance during differentiation [59]. However, the interactions occurring with the topologically associated domains themselves can vary. Therefore, the separation of chromosomes into topologically associated domains provides a framework for distinct developmental-specific nuclear positioning. As cells differentiate, they respond to many different signals and rely on their underlying transcriptional machinery including protein complexes to respond to these cues. Among these responses are that of transcriptional regulatory control [60, 61]. The gene expression programs are controlled by many different proteins including transcription factors, co-regulatory proteins and chromatin regulators that bind to DNA at regulatory regions. As chromosomes are arranged in their territories during interphase, and within these territories, chromosomes contain regulatory interacting regions as discussed earlier. It is evident that organization of the genome is directly important for gene expression and therefore research has aimed to uncover their role in development through differentiation and cell-type specificity.

Among the hallmark studies investigating the role of chromatin organization and gene expression in regard to development include studies investigating the inactivation of the X chromosome [62-64]. A study seeking to characterize cis-regulatory mechanisms involved in X inactivation using female mouse embryonic fibroblasts found that while topologically associated domain organization remains largely the same for the X inactivated chromosome compared to the activated, there were differences in internal-topological associated domain interactions [63]. Additionally, when interrogating gene expression differences within these regions, they found a correlation between intra-topologically associated domain alterations and differential gene expression throughout differentiation. Lastly, they revealed that alterations of boundary regions resulted in the mis-regulation of long-range gene-networks. A more recent study by Dixon et al. mapped genome-wide chromatin interactions in H1 human embryonic stem cells and human embryonic stem cell-derived lineages and ultimately uncovered reorganization of chromatin architecture during stem cell differentiation [65]. This reorganization during lineage specification was observed through switches between chromosome compartment types (A and B) and changes in the frequency of local interactions within the topological associated domains. They associated these regions with gene expression and identified subsets of genes displaying alterations in gene expression patterns. Together these studies serve as examples to the groundbreaking efforts implicating differential 3D-genomic architecture within chromosome compartments and topologically associated domains, their influence on gene regulation and moreover the implications of this molecular network on cellular development.

1.2.3: Chromosome architecture in cancer programs

Some of the first hypotheses regarding chromosomal abnormalities within cancer were conceived in the late 1800s and early 1900s by the German scientists David Paul von Hansemann and Theodore Boveri who proposed that cancer is a result of aberrant chromosome regulation during the cell cycle [66,67]. For many years pathologists have used changes in nuclear structure as a diagnostic framework to detect cancer [68]. Expanding from fundamental observations of abnormal nuclear morphology in addition to the quantity of nuclei in cancer cells comes later observations, which have revealed changes in higher-order chromatin organization. Despite the advancements in our knowledge regarding abnormal gene expression programs within various cancers, the specific roles these changes in 3D-chromatin structure plays in cancer phenotypes remains to be further elucidated.

Changes in chromosome territories have been linked to differential gene expression programs in cancer [69, 70]. Marella et al. investigated chromatin territory organization in normal epithelial and breast cancer cell lines and observed alterations in chromosome territories through an increase in associations between chromosome 4 and chromosome 16 [71]. Another study investigating alterations in topologically associated domains in cancer revealed changes within interacting regions where they also identified differences in chromatin architecture surrounding genes that were differentially expressed within these altered long-range chromatin interactions [72]. While there are other similar observations in different cancers [73] and recent studies have produced large datasets for studying 3-D chromatin alterations in cancer, much still remains to be elucidated.

1.3: Post translational histone modifications

The structure of chromatin and gene expression has emerged as a key area of research for uncovering mechanisms underlying cellular phenotypes. Modifications of the core histone proteins influence gene expression programs in the way that some modifications can dynamically alter how open or closed the chromatin is, thereby affecting regulatory protein accessibility to genes during transcription. Research has identified modifications that can occur to the free amino acid tails of the core histone proteins at specific residues that are directly linked to a biological outcome; this is referred to as the histone code hypothesis [74]. Histone tail acetylation and methylation are the modifications that have been identified to primarily play the role in switching the dynamics of chromatin accessibility. To this end, therapeutic agents targeting epigenetic histone modifications have been developed to mitigate oncogenic development in the cancer phenotype.

1.3.1: The core histone proteins

Post translational histone modifications have emerged as important regulatory components of cellular processes influencing gene regulation. Histone proteins wrapped in DNA assemble into nucleosomes and are the foundational building blocks of eukaryotic chromatin. Early reports in the late 1880's suggest the discovery of histone proteins and work in the 1960s was the first to report a potential function for these proteins in controlling gene regulation [75]. Evolutionary processes have resulted in the four core histone proteins that make up the nucleosome and pack the DNA tightly. These four core proteins that make up the octamer are H2A, H2B, H3 and H4 [76]. Each histone

protein contains globular domains that modulate the interactions between the core histone proteins. Additionally, each protein contains amino tails that contain approximately 30 amino acids that extend away from the nucleosome structure and are subject to different modifications [77].

1.3.2: Modifications of the core histone proteins

Histone proteins have the potential to gain post translational covalent modifications on their extended free amino acid tails. While on an individual level, these modifications largely do not affect the structure of the nucleosomes themselves, higher order chromatin structure containing many nucleosomes can be impacted by these changes [78]. Modifications to the N-terminal tails of H2A, H2B, H3 and H4 were first reported in the 1963 by Phillips et al. [79]. The variety of modifications that can occur to the amino acid tails include methylation, acetylation, ubiquitylation, phosphorylation, ribosylation and sumolyation and each modification has the ability to translate into distinct biological outcomes [78]. The histone protein H3 is the most extensively modified histone protein and has been studied to examine epigenetic post translational histone modifications during normal cellular development as well as the progression and maintenance of diseased states [80]. In this review, I highlight histone acetylation and methylation of the core histone protein, H3.

1.3.3: Overview of histone acetylation and methylation

Histone acetylation and methylation was first described in 1964, when Allfrey et al. suggested its potential role in the regulation of RNA synthesis [81]. Notably, these

modifications have been extensively studied in regard to their contribution to closed and permissive chromatin states ultimately impacting nucleosome positioning and orientation of key regulatory sequences [82].

Histone acetylation is mediated by histone acetyltransferase (HAT) enzymes. There are different HATs containing specific subunits that dictate the distinct acetylation pattern it writes on the histone tail [83]. Human HATs can be grouped into five families, comprising of approximately thirty different HATs and while the sequences in HAT domains differ between families, they have high structural similarities within family members [84]. These enzymes rely on accessory proteins along the chromatin for localization to specific sites allowing the exchange for the acetyl group provided by Acetyl-CoA to modify the lysine [77]. This results in a more permissive, open chromatin state and allows the recruitment of proteins to these now accessible regions as the nucleosomes typically act as a barrier for RNA polymerase during this process [26].

The HAT ability of CBP/p300, one of the five HAT family members, was discovered in 1996 [85, 86]. These enzymes are the only HATs that are able to acetylate all of the core histone proteins [86]. In terms of a functional example of this HAT, CBP/p300 are recruited by the transcription factor DUX4 in myoblasts to globally mark the underlying genome with the H3K27ac modification [87]. This transcription factor is known to be involved in a form of muscular dystrophy; however, the mechanisms of action remain poorly defined. In this study, Choi et al., highlights a new epigenetic mechanism by which a HAT, recruited to the genome via a transcription factor, modifies histones resulting in open chromatin surrounding genes relevant to muscular dystrophy.

The dynamic exchange between HAT and histone deacetylase (HDAC) enzymes results in changes between chromatin states as HDACs remove acetyl groups typically resulting in a more closed chromatin arrangement [74]. In principle, the acetylation to the lysine group serves two overall purposes: (1) it neutralizes some of the positive charge of the histone protein and thus results in a weaker interaction of the DNA with the histone, thereby resulting in a euchromatin state [81] and (2) it acts as a docking site for the recruitment of other regulatory proteins that can come in and direct transcriptional changes.

In contrast to histone acetylation, histone methylation results in condensed chromatin. This modification to the histone tails can be acquired at the amino acids arginine or lysine; however, the role of arginine in chromatin dynamics and transcriptional regulation is not as well defined compared to those of lysine alterations [88]. Lysine has the ability to gain either one, two or three methyl groups which are mediated by the enzymes, histone methyltransferases (HMT), that recognize specific lysine residues. HMTs that mediate the acquisition of methyl groups of lysine residues are of two different broad classes, SET domain containing or non-SET domain containing, each with different catalytic mechanisms [89].

Histone methylation does not change the charge of the histone and so it has the ability to correlate with either transcriptional activation or transcriptional repression depending on the location of the methylation [77]. These methylation patterns are also dynamically regulated by demethylase enzymes and for the case of lysine demethylation, these are referred to as lysine demethylases. The methyl groups replace the hydrogen groups by exchanging with methyl groups provided by S-Adenosyl methionine to result

in either mono, di or tri methylation [90]. Similar to that of acetylation, regulatory proteins can recognize these methyl groups. Different chromodomains of HMTs prefer to bind to specific methylated lysine residues and go on to perform their specific effector function; this results in the outcome of either a more heterochromatin or euchromatin state [91].

Foundational studies characterized the fundamental core histone proteins and revealed their N-terminal tails are subject to different covalent post-translational histone modifications. Later research identified factors writing and removing these modifications from the tails, which was followed with advancements in our knowledge surrounding regulatory proteins reading these modifications and carrying out the gene regulatory programs influencing cellular phenotypes. The dynamic relationship within histone acetylation and histone methylation programs allows for the switch between open and closed chromatin states.

1.3.4: Histone acetylation and methylation in cancer

Cancer is a heterogeneous disease with underlying abnormalities in genomic and epigenomic profiles [92]. Aberrant regulation of posttranslational histone modifications is among these abnormalities, which has been shown to contribute to the cancer phenotype by altering accessibility of key target genes [93]. In fact, histone modifications have emerged as key regulators in this disease [94, 95]. This finding expanded on the prior determination of DNA methylation in cancer that revealed the interaction of HDACs with a chromatin-modifying factor that aids in methylating DNA within promoter regions of

tumor suppressor genes [96, 97]. Furthermore, the association of DNA methyltransferases (DNMTs) with HDACs was revealed [94, 95].

The dynamic mechanisms within histone acetylation and histone methylation have since been discovered to regulate known oncogenes and tumor suppressors ultimately influencing the progression and metastasis of cancers [98, 99]. These finding have inspired the expansion of the histone code hypothesis to include the regulatory roles of epigenetics in cancer transcriptional programs. Histone acetylation and methylation patterns are controlled by the interchange of their enzyme activities, which ultimately provide balance in a dynamic relationship between a euchromatin and heterochromatin state. Since these processes impact active and inactive gene states, continuing to elucidate these mechanisms in the development and maintenance in cancer cell models holds relevant purpose in better understanding the mechanisms underlying the disease and has also shown promise in therapeutic targets.

Histone acetylation landscapes that are altered in cancers can be attributed to the recruitment of HDACs to tumor suppressor genes that are important for silencing epigenetic programs. In contrast, increased HAT activity at oncogenes can contribute to growth and proliferation of cancer cells. For example, abnormal HDAC regulation has been identified in pancreatic cancer where HDAC1-3 and HDAC7 have been identified to be overexpressed [100-103]. The cellular consequence of these alterations is dependent on the specific enzyme involved. Notably, the aberrant expression of HDAC7 results in the ability to distinguish between advanced pancreatic cancer from earlier neoplasms [101] and HDAC1-3 enzymes have been shown to impact processes of p53, NFkB and p65 programs, which have been correlated with cancer progression including pancreatic

cancer [104, 105]. The HAT p300 has been shown to be influenced by the activity of nuclear factor of activated T cells (NFAT) and glioma-associated oncogene family zinc finger 3 (GLI3) in that these transcription factors recruit this active HAT and thereby mechanistically results in loosening of the chromatin and allows for the transcription of the *c-MYC* oncogene [106], a known regulator of many cancer phenotypes [107, 108].

1.4: Next generation sequencing methods for studying chromatin-based mechanisms

Early methods in determining DNA structure and the organization relied primarily on microscopic and biochemical techniques. These techniques have led us to groundbreaking findings that have since served as the foundation of normal and aberrant cell phenotypes. Modern technologies have allowed us to study the epigenome and transcriptome on a genome-wide scale. These technologies termed 'Next Generation Sequencing' technologies have revolutionized the way we study genomic based mechanisms and have notably provided great insight into cancer phenotypes. Here I describe methodologies used to study the various layers of genomic regulation described above including: (1) 3-dimensional chromosome architecture, (2) post translational histone modifications and (3) gene expression.

1.4.1: Methods for studying 3D chromosome organization

Research investigating nuclear organization initially relied on microscopic techniques. While on the fundamental level, light microscopy has allowed for the basic observation of nuclear morphology, advancements in microscopic methods have allowed for interrogation of nucleic acid sequences within the nucleus. Fluorescence in situ

hybridization is a microscopic technique developed in the 1980's that uses fluorescent probes designed to bind to a complementary sequence along the genome. This technique identified co-localization of distal genes during active transcription [109]. Fluorescence in situ hybridization and its derivatives holds power when investigating single cells or smaller cell populations. However, when inquiring about cell populations on a genome-wide level, this and other microscopic techniques fall short.

While microscopy still holds importance and value in genomic and molecular biology research, advancements in genome-wide methods have allowed for the investigation of entire genomes within single or whole cell populations. Modern methods for defining genome-wide landscapes of histone modifications uses Next Generation Sequencing technologies to expand characterization past single loci. A groundbreaking study in 2002 by Dekker et al. described the first chromosome conformation capture (3C) based assay, which allows for the association of genome-wide contacts [110]. Some of the derivatives of 3C technologies include 4C and 5C. While there are differences in the number of targeted genomic regions, the fundamental principle of 3C is shared across the other derivatives. In short, the chromatin is fixed using a fixative such as formaldehyde which results in covalent linkage between DNA and protein interactions [111]. The crosslinked chromatin is then digested typically using a 6 bp restriction endonuclease and is followed by chromatin dilution and ligation, resulting in the ligation of intramolecular fragments. The advancements of 3C into 4C (circularized chromosome conformation capture) and 5C (chromosome conformation capture carbon copy) builds on the limitations from the given predecessor. Whereas 3C interactions represent a 'one vs. one'

interaction, 4C allows for the detection of ‘one vs. all’ interactions and 5C expands to ‘many vs. many’.

The innovation of Next Generation Sequencing technologies permitted the development of methods determining all genome-wide interactions for any given loci. In 2009, Lieberman et al. developed the first among these methods named Hi-C [47]. This methodology does not require the identification of target loci and rather identifies all chromosome interactions. The principles of Hi-C are shared with that of 3C and its derivatives; however, modifications include filling the digested ends with a biotin and generation of a Next Generation Sequencing library following ligation and isolation of biotinylated fragments post dilution, ligation and linearization. Paired-end massive parallel sequencing is then performed on the fragments to amplify and identify the junctions. This method was modified in 2011 by Kalhor et al. with the aim of improving the signal to noise ratio [112]. This method biotinylates the proteins after crosslinking and solubilization of the chromatin. After enzyme digestion, the fragments are tethered to magnetic streptavidin beads and selected through magnetization allowing for reduction of inappropriate intermolecular interactions. The ends of the DNA are then filled in, biotinylated, purified, exonuclease treated, sheared into smaller fragments and pulled down by magnetic streptavidin beads. Lastly, the Next Generation Sequencing libraries are prepared and sequenced as described for Hi-C. This technique is referred to as Tethered Chromatin Conformation Capture. These technologies have advanced our understanding of chromosome contacts due to the high resolution they provide at any given interaction loci across the genome.

1.4.2: Methods for post translational histone modifications

Next Generation Sequencing has also led to the development of novel methods for identifying the genome-wide landscape of histone modifications. These technologies have afforded an advantage for genome-wide analyses compared to prior methods that limited investigations to specific loci. This method pairs the traditional chromatin immunoprecipitation (ChIP) technique with massive parallel sequencing and is termed ChIP-seq. ChIP allows for the investigation of protein-DNA interactions within the nucleus [113, 114] and when paired with NGS, results in genome-wide binding sites of the protein of interest [115]. To perform the ChIP, proteins are first crosslinked with formaldehyde and the cells are lysed to release the DNA from the nuclei [116]. Sonication is then used to fragment the DNA into approximately 500 bp fragments, which can then be used to isolate protein-DNA complexes. The complexes are immunoprecipitated with an antibody against the protein target of interest and isolated using protein A/G magnetic bead selection with a series of washes. The chromatin bound to the protein is then eluted, the cross-links are reversed and the proteins are proteinase digested. After the DNA is purified, enrichment can be detected using quantitative polymerase chain reaction (qPCR) with primers designed against known target regions for the protein of interest. For some studies this is the endpoint of the investigation; however, when inquiring about whole genome targets either microarrays or high-throughput sequencing can be used. Microarrays utilize a hybridization technique that allows for fluorescently labeled DNA fragments to hybridize to an array of genomic targets, when paired with ChIP this is referred to as ChIP-on-chip. In contrast, Next Generation Sequencing libraries can be built from the fragments and used in massive

parallel sequencing by synthesis to yield high throughput results of the proteins' targets along the entire genome. The enrichment of these targets are determined when comparing the ChIP-seq data for the protein of interest relative to an input control.

1.4.3: Evaluating gene expression

Many methods have been designed to quantify gene expression over the years. Quantitative reverse transcriptase PCR (qRT-PCR) can be used to determine expression of individual gene targets along the genome [117]. The development of two technologies, the microarray and RNA-sequencing, have advanced our capability of studying gene expression by allowing the investigation of thousands of genes. As mentioned above, microarray is a hybridization technique that utilizes a fluorescent dye for later quantification. The array is designed with fixed DNA probes against the RNAs of interest. The prepared cDNAs from the given experiment are passed over the microarray slide and any complementary sequences between the probes and cDNAs will have an affinity for hybridization. The fluorescent dyes are ultimately used for optimal measurements where relative amounts from each probe can be calculated, thereby correlating with expression quantification [118]. Next Generation Sequencing, described earlier, resulted in the development of RNA-sequencing. This technique was developed to investigate the quantity and presence of RNA at a given time. This method utilizes massive parallel sequencing to investigate the entire transcriptome.

1.5: Dynamics of chromatin structure

Chromatin accessibility is a dynamic process mediated by external stimuli and affects the capability of transcriptional regulators to bind their regulatory elements along the genome. The switch between open and closed chromatin states is referred to as chromatin-remodeling and is influenced by a variety of epigenetic mechanism. In addition to DNA methylation [29], posttranslational histone modifications [30] and non-coding RNAs [31], environmental stimuli can also influence chromatin organization and accessibility. For example, cigarette smoke has been shown to remodel chromatin by altering histone acetylation patterns resulting in increased gene expression of pro-inflammatory genes [119] and has recently been implicated in lung cancer phenotypes [120]. Moreover, chromatin-remodeling is a well-known facet underlying cancer phenotypes [121] and because of the reversibility of chromatin-remodeling, efforts have been made to target regulatory factors with the potential for therapeutic intervention. Therefore, elucidating chromatin-remodeling mechanisms underlying disease phenotypes is important for identifying agents capable of disease mitigation.

1.5.1: Chromatin-remodeling in response to hormone signaling

Nuclear receptors have been shown to mediate chromatin-remodeling by recruiting regulatory proteins that alter downstream transcriptional programs [122]. These nuclear receptors are transcription factors that are activated by their coordinated ligand and contribute to epigenetic and gene expression changes in normal and disease phenotypes.

A clear example of this relationship is that of hormones and their corresponding receptors. St. John et al. studied the hormone Vitamin D3 in the transition between osteoblasts and osteocytes [123]. Through the integration of genome-wide histone modifications and gene expression datasets in response to hormone-stimulation, they identified temporal changes in chromatin-remodeling around osteocyte-relevant genes. Since Vitamin D3 hormone is involved in bone remodeling, this provided mechanistic context for this hormone in chromatin-remodeling during osteoblast differentiation. Another study in mammary epithelium during different cellular states investigated genome-wide H3K27me3 changes in response to two hormones, prolactin and progestin [124]. Ultimately, they determined hormone-regulated histone methylation patterns to be mediated by the chromatin remodeling complex EZH2 in mammary cell lineage specification.

Aberrant nuclear hormone receptors are underlying phenotypes of several cancers and furthermore can be used to further classify cancer subtypes of breast cancers [125]. Among these prevalent cancers is the estrogen receptor (ER) positive subtype [126]. Other receptors linked to breast cancer are the progesterone (PR), androgen, thyroid and glucocorticoid receptors [127-129]. A study investigating chromatin-remodeling signatures in ER/PR positive breast cancer in response to temporal stimulation revealed structural transitions in topologically associated domains that coordinated with the hormones estrogen and progestin [130]. Furthermore, changes in genome-wide histone modification signatures were also observed that coordinated with hormone response and suggested unique and concordant hormone chromatin reorganization. Furthermore, changes in genome-wide compartments and histone modifications resulted in gene

expression differences in response to either hormone. In another study, alterations in long-range chromatin interactions and DNaseI hypersensitivity coordinated with changes in gene expression in response to glucocorticoid steroid hormone pulsations in a murine breast cancer model [131].

As nuclear receptors play a role in transcriptional processes and aberrant expression of nuclear receptors has been associated with disease phenotypes, elucidating mechanisms underlying their alterations in response to their given ligands is important. While there are many nuclear receptors whose downstream mechanisms result in distinct cellular phenotypes, among these are hormone nuclear receptors. Understanding the dynamic mechanisms within chromatin remodeling in regard to genome-wide chromatin architecture and histone modification alterations in cancer phenotypes will allow for elucidation of chromatin-mediated etiology of cancer and may provide insight into therapeutic targets. In Chapter 4 of this dissertation, I present high resolution chromatin dynamics in response to temporal stimulation of an estrogen derivative in ER positive breast cancer cell models.

1.5.2: Chromatin-remodeling in response to small molecules

Changes in chromatin structure are often mediated by enzymes that communicate toward downstream transcriptional programs as discussed earlier. As research has continuously supported epigenetic mechanisms governing chromatin-remodeling to be reversible, there has been increased interest in modulating these processes in human disease. To this end, small molecule inhibitors have emerged as promising for altering

mechanisms involving chromatin-remodeling and a subset of the selective inhibitors have advanced to clinical trials.

Some of the major overall classes of inhibitors that have been developed include: histone acetyltransferase inhibitors and histone deacetylase modulators [132], bromodomain inhibitors, DNA methyltransferase modulators and protein methyltransferase inhibitors [133]. HDAC inhibitors have been widely developed for cancer therapies [134]. Recently, an HDAC inhibitor was shown to increase expression of a tumor antigen and resulted in a decrease in immune suppressive cell types in non-small cell lung cancer cells [135].

1.6: Scope of Dissertation:

As discussed in the literature review portion of this chapter, aberrant chromatin based processes including higher order chromatin organization and post translational histone modifications are widely appreciated as contributing constituents to cancer development and maintenance. Elucidating these processes is important not only for gaining understanding regarding the molecular heterogeneity underlying this disease but targeting these mechanisms has also shown therapeutic promise. Here, I describe our work applying Next Generation Sequencing technologies to study functional genomics within the 3-dimensional and epigenetic landscape of two prevalent cancers.

Chapter 2 of this thesis embodies integrative genomic analyses revealing epigenetic plasticity in a pancreatic ductal adenocarcinoma cell-line model. The overall goal of this chapter was to reveal chromatin-based mechanisms in pancreatic ductal adenocarcinoma and to determine the effects of clinically relevant histone

acetyltransferase inhibitors on the cancer epigenome. This chapter builds on our study identifying the impact of ICG-001 and C646 on gene expression in pancreatic ductal adenocarcinoma and colorectal cancer cell-lines (highlighted in Appendix A). I define the 3-dimensional chromatin landscape of pancreatic ductal adenocarcinoma and characterizes the distribution of post translational histone modifications within these interacting regions and correlate these regions with gene regulatory programs. I then determine the effects of histone acetyltransferase inhibitors on chromatin loops and gene expression in the widely used pancreatic ductal adenocarcinoma cell-line, PANC1.

In Chapter 3, I characterize grade-specific broad H3K4me3 and H3K27ac regions in pancreatic ductal adenocarcinoma cell-lines and furthermore determine the influence of these histone acetyltransferase inhibitors on these broad epigenomic domains. Lastly, I propose genome-wide mechanisms mediated by these inhibitors within these regulatory regions.

In Chapter 4, I identified the dynamic reorganization of chromatin domains in estrogen receptor positive breast cancer cell models. The overall goal of this chapter was to elucidate chromatin-based mechanisms in response to temporal stimulation of the estrogen derivative, 17-β estradiol, in an estrogen receptor positive breast cancer cell-line and the tamoxifen resistant derivative of these cells. Furthermore, we aimed to establish the role of estrogen receptor alpha in mediating these dynamics and further linked these processes to gene expression.

This body of work focuses on the characterization of higher-order chromatin organization and histone-modification mechanisms of two deadly cancers. Furthermore,

we investigate the plasticity of these mechanisms in response to external stimuli to determine the dynamics of these mechanisms in regard to transcriptional outcomes.

References

1. Franklin, R.E. and R.G. Gosling, *Molecular configuration in sodium thymonucleate.* Nature, 1953. 171(4356): p. 740-1.
2. Chargaff, E., *Chemical specificity of nucleic acids and mechanism of their enzymatic degradation.* Experientia, 1950. 6(6): p. 201-9.
3. Gall, J.G., *Kinetics of deoxyribonuclease action on chromosomes.* Nature, 1963. 198: p. 36-8.
4. Lamond, A.I. and W.C. Earnshaw, *Structure and function in the nucleus.* Science, 1998. 280(5363): p. 547-53.
5. Ris, H. and D.F. Kubai, *Chromosome structure.* Annu Rev Genet, 1970. 4: p. 263-94.
6. Huberman, J.A., *Structure of chromosome fibers and chromosomes.* Annu Rev Biochem, 1973. 42: p. 355-78.
7. Oudet, P., M. Gross-Bellard, and P. Chambon, *Electron microscopic and biochemical evidence that chromatin structure is a repeating unit.* Cell, 1975. 4(4): p. 281-300.
8. Arents, G., et al., *The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix.* Proc Natl Acad Sci U S A, 1991. 88(22): p. 10148-52.
9. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. 389(6648): p. 251-60.

10. Workman, J.L. and R.E. Kingston, *Alteration of nucleosome structure as a mechanism of transcriptional regulation.* Annu Rev Biochem, 1998. 67: p. 545-79.

11. Tagami, H., et al., *Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis.* Cell, 2004. 116(1): p. 51-61.

12. Deniz, O., et al., *Nucleosome architecture throughout the cell cycle.* Sci Rep, 2016. 6: p. 19729.

13. Crick, F., *Central dogma of molecular biology.* Nature, 1970. 227(5258): p. 561-3.

14. Roeder, R.G. and W.J. Rutter, *Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms.* Nature, 1969. 224(5216): p. 234-7.

15. Russell, J. and J.C. Zomerdijk, *RNA-polymerase-I-directed rDNA transcription, life and works.* Trends Biochem Sci, 2005. 30(2): p. 87-96.

16. Weinmann, R. and R.G. Roeder, *Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes.* Proc Natl Acad Sci U S A, 1974. 71(5): p. 1790-4.

17. Borchert, G.M., W. Lanier, and B.L. Davidson, *RNA polymerase III transcribes human microRNAs.* Nat Struct Mol Biol, 2006. 13(12): p. 1097-101.

18. Reines, D., J.W. Conaway, and R.C. Conaway, *The RNA polymerase II general elongation factors.* Trends Biochem Sci, 1996. 21(9): p. 351-5.

19. Sydow, J.F., et al., *Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA.* Mol Cell, 2009. 34(6): p. 710-21.

20. Cheung, A.C. and P. Cramer, *Structural basis of RNA polymerase II backtracking, arrest and reactivation.* Nature, 2011. 471(7337): p. 249-53.

21. Matsui, T., et al., *Multiple factors required for accurate initiation of transcription by purified RNA polymerase II.* J Biol Chem, 1980. 255(24): p. 11992-6.

22. Vaquerizas, J.M., et al., *A census of human transcription factors: function, expression and evolution.* Nat Rev Genet, 2009. 10(4): p. 252-63.

23. Goodwin, G.H., C. Sanders, and E.W. Johns, *A new group of chromatin-associated proteins with a high content of acidic and basic amino acids.* Eur J Biochem, 1973. 38(1): p. 14-9.

24. Murugesapillai, D., et al., *Single-molecule studies of high-mobility group B architectural DNA bending proteins.* Biophys Rev, 2017. 9(1): p. 17-40.

25. Reiter, F., S. Wienerroither, and A. Stark, *Combinatorial function of transcription factors and cofactors.* Curr Opin Genet Dev, 2017. 43: p. 73-81.

26. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. 128(4): p. 707-19.

27. Van Speybroeck, L., *From epigenesis to epigenetics: the case of C. H. Waddington.* Ann N Y Acad Sci, 2002. 981: p. 61-81.

28. Martinowich, K., et al., *DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation.* Science, 2003. 302(5646): p. 890-3.

29. Maurano, M.T., et al., *Role of DNA Methylation in Modulating Transcription Factor Occupancy.* Cell Rep, 2015. 12(7): p. 1184-95.

30. Zhou, V.W., A. Goren, and B.E. Bernstein, *Charting histone modifications and the functional organization of mammalian genomes.* Nat Rev Genet, 2011. 12(1): p. 7-18.

31. Hirota, K., et al., *Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs.* Nature, 2008. 456(7218): p. 130-4.

32. Mercer, T.R. and J.S. Mattick, *Structure and function of long noncoding RNAs in epigenetic regulation.* Nat Struct Mol Biol, 2013. 20(3): p. 300-7.

33. Robertson, K.D., et al., *The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors.* Nucleic Acids Res, 1999. 27(11): p. 2291-8.

34. Cedar, H. and Y. Bergman, *Linking DNA methylation and histone modification: patterns and paradigms.* Nat Rev Genet, 2009. 10(5): p. 295-304.

35. Ng, H.H. and A. Bird, *DNA methylation and chromatin modification.* Curr Opin Genet Dev, 1999. 9(2): p. 158-63.

36. Bird, A.P. and A.P. Wolffe, *Methylation-induced repression--belts, braces, and chromatin.* Cell, 1999. 99(5): p. 451-4.

37. Ballestar, E. and A.P. Wolffe, *Methyl-CpG-binding proteins. Targeting specific gene repression.* Eur J Biochem, 2001. 268(1): p. 1-6.

38. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. 128(4): p. 693-705.

39. Misteli, T. and E. Soutoglou, *The emerging role of nuclear architecture in DNA repair and genome maintenance.* Nat Rev Mol Cell Biol, 2009. 10(4): p. 243-54.

40. Lai, F., et al., *Activating RNAs associate with Mediator to enhance chromatin architecture and transcription.* Nature, 2013. 494(7438): p. 497-501.

41. Kimmins, S. and P. Sassone-Corsi, *Chromatin remodelling and epigenetic features of germ cells.* Nature, 2005. 434(7033): p. 583-9.

42. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. 144(5): p. 646-74.

43. Hubner, M.R., M.A. Eckersley-Maslin, and D.L. Spector, *Chromatin organization and transcriptional regulation.* Curr Opin Genet Dev, 2013. 23(2): p. 89-95.

44. Cremer, T., et al., *Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments.* Hum Genet, 1982. 60(1): p. 46-56.

45. Cremer, M., et al., *Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells.* Chromosome Res, 2001. 9(7): p. 541-67.

46. Kreth, G., et al., *Radial arrangement of chromosome territories in human cell nuclei: a computer model approach based on gene density indicates a probabilistic global positioning code.* Biophys J, 2004. 86(5): p. 2803-12.

47. Lieberman-Aiden, E., et al., *Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science, 2009. 326(5950): p. 289-93.

48. Wright, J.B., S.J. Brown, and M.D. Cole, *Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells.* Mol Cell Biol, 2010. 30(6): p. 1411-20.

49. Cremer, T., et al., *Role of chromosome territories in the functional compartmentalization of the cell nucleus.* Cold Spring Harb Symp Quant Biol, 1993. 58: p. 777-92.

50. Ciabrelli, F. and G. Cavalli, *Chromatin-driven behavior of topologically associating domains.* J Mol Biol, 2015. 427(3): p. 608-25.

51. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. 485(7398): p. 376-80.

52. Tang, Z., et al., *CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription.* Cell, 2015. 163(7): p. 1611-27.

53. Hagstrom, K.A. and B.J. Meyer, *Condensin and cohesin: more than chromosome compactor and glue.* Nat Rev Genet, 2003. 4(7): p. 520-34.

54. Marsman, J. and J.A. Horsfield, *Long distance relationships: enhancer-promoter communication and dynamic gene transcription.* Biochim Biophys Acta, 2012. 1819(11-12): p. 1217-27.

55. Hatzis, P. and I. Talianidis, *Dynamics of enhancer-promoter communication during differentiation-induced gene activation.* Mol Cell, 2002. 10(6): p. 1467-77.

56. Whalen, S., R.M. Truty, and K.S. Pollard, *Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin.* Nat Genet, 2016. 48(5): p. 488-96.

57. Moisan, S., et al., *Novel long-range regulatory mechanisms controlling PKD2 gene expression.* BMC Genomics, 2018. 19(1): p. 515.

58. Tanabe, H., et al., *Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications.* Mutat Res, 2002. 504(1-2): p. 37-45.

59. Vietri Rudan, M., et al., *Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture.* Cell Rep, 2015. 10(8): p. 1297-309.

60. Li, G., et al., *Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation.* Cell, 2012. 148(1-2): p. 84-98.

61. Xiao, J., R. Jin, and D. Wagner, *Developmental transitions: integrating environmental cues with hormonal signaling in the chromatin landscape in plants.* Genome Biol, 2017. 18(1): p. 88.

62. Erwin, J.A. and J.T. Lee, *New twists in X-chromosome inactivation.* Curr Opin Cell Biol, 2008. 20(3): p. 349-55.

63. Nora, E.P., et al., *Spatial partitioning of the regulatory landscape of the X-inactivation centre.* Nature, 2012. 485(7398): p. 381-5.

64. Splinter, E., et al., *The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA.* Genes Dev, 2011. 25(13): p. 1371-83.

65. Dixon, J.R., et al., *Chromatin architecture reorganization during stem cell differentiation.* Nature, 2015. 518(7539): p. 331-6.

66. von Hansemann, D., *Ueber asymmetrische Zelltheilung in epithel Krebsen und deren biologische Bedeutung. Virchows Arch.* Path. Anat., 1890: p. 119, 299.

67. T, B., *Zur Frage der Entstehung maligner Tumoren.* G. Fischer, 1914.

68. Zink, D., A.H. Fischer, and J.A. Nickerson, *Nuclear structure in cancer cells.* Nat Rev Cancer, 2004. 4(9): p. 677-87.

69. Cremer, M., et al., *Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei.* J Cell Biol, 2003. 162(5): p. 809-20.

70. Parada, L.A., et al., *Conservation of relative chromosome positioning in normal and cancer cells.* Curr Biol, 2002. 12(19): p. 1692-7.

71. Marella, N.V., et al., *Cell type specific chromosome territory organization in the interphase nucleus of normal and cancer cells.* J Cell Physiol, 2009. 221(1): p. 130-8.

72. Taberlay, P.C., et al., *Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations.* Genome Res, 2016. 26(6): p. 719-31.

73. Lupianez, D.G., et al., *Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions.* Cell, 2015. 161(5): p. 1012-1025.

74. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. 403(6765): p. 41-5.

75. Johns, E.W. and S. Forrester, *Studies on nuclear proteins. The binding of extra acidic proteins to deoxyribonucleoprotein during the preparation of nuclear proteins.* Eur J Biochem, 1969. 8(4): p. 547-51.

76. Henikoff, S. and M.M. Smith, *Histone variants and epigenetics.* Cold Spring Harb Perspect Biol, 2015. 7(1): p. a019364.

77. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. 293(5532): p. 1074-80.

78. Khorasanizadeh, S., *The nucleosome: from genomic organization to genomic regulation.* Cell, 2004. 116(2): p. 259-72.

79. Phillips, D.M., *The presence of acetyl groups of histones.* Biochem J, 1963. 87: p. 258-63.

80. Xu, Y.M., J.Y. Du, and A.T. Lau, *Posttranslational modifications of human histone H3: an update.* Proteomics, 2014. 14(17-18): p. 2047-60.

81. Allfrey, V.G., R. Faulkner, and A.E. Mirsky, *Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis.* Proc Natl Acad Sci U S A, 1964. 51: p. 786-94.

82. al., K.T.K.e., *Histone Acetylation And Methylation.* Chromatin and Disease. Subcellular Biochemistry, 2007. 41.

83. Marmorstein, R., *Structure of histone acetyltransferases.* J Mol Biol, 2001. 311(3): p. 433-44.

84. McCullough, C.E. and R. Marmorstein, *Molecular Basis for Histone Acetyltransferase Regulation by Binding Partners, Associated Domains, and Autoacetylation.* ACS Chem Biol, 2016. 11(3): p. 632-42.

85. Bannister, A.J., et al., *Acetylation of importin-alpha nuclear import factors by CBP/p300.* Curr Biol, 2000. 10(8): p. 467-70.

86. Ogryzko, V.V., et al., *The transcriptional coactivators p300 and CBP are histone acetyltransferases.* Cell, 1996. 87(5): p. 953-9.

87. Choi, S.H., et al., *DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes.* Nucleic Acids Res, 2016. 44(11): p. 5161-73.

88. Litt, M., Y. Qiu, and S. Huang, *Histone arginine methylations: their roles in chromatin dynamics and transcriptional regulation.* Biosci Rep, 2009. 29(2): p. 131-41.

89. Dillon, S.C., et al., *The SET-domain protein superfamily: protein lysine methyltransferases.* Genome Biol, 2005. 6(8): p. 227.

90. Bannister, A.J., R. Schneider, and T. Kouzarides, *Histone methylation: dynamic or static?* Cell, 2002. 109(7): p. 801-6.

91. Heintzman, N.D., et al., *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome.* Nat Genet, 2007. 39(3): p. 311-8.

92. Litzenburger, U.M., et al., *Single-cell epigenomic variability reveals functional cancer heterogeneity.* Genome Biol, 2017. 18(1): p. 15.

93. Audia, J.E. and R.M. Campbell, *Histone Modifications and Cancer.* Cold Spring Harb Perspect Biol, 2016. 8(4): p. a019521.

94. Rountree, M.R., K.E. Bachman, and S.B. Baylin, *DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci.* Nat Genet, 2000. 25(3): p. 269-77.

95. Wade, P.A., et al., *Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation.* Nat Genet, 1999. 23(1): p. 62-6.

96. Keshet, I., J. Lieman-Hurwitz, and H. Cedar, *DNA methylation affects the formation of active chromatin.* Cell, 1986. 44(4): p. 535-43.

97. Nan, X., et al., *Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex.* Nature, 1998. 393(6683): p. 386-9.

98. Chi, P., C.D. Allis, and G.G. Wang, *Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers.* Nat Rev Cancer, 2010. 10(7): p. 457-69.

99. Khan, S.A., D. Reddy, and S. Gupta, *Global histone post-translational modifications and cancer: Biomarkers for diagnosis, prognosis and treatment?* World J Biol Chem, 2015. 6(4): p. 333-45.

100. Fritsche, P., et al., *HDAC2 mediates therapeutic resistance of pancreatic cancer cells via the BH3-only protein NOXA.* Gut, 2009. 58(10): p. 1399-409.

101. Ouaissi, M., et al., *High histone deacetylase 7 (HDAC7) expression is significantly associated with adenocarcinomas of the pancreas.* Ann Surg Oncol, 2008. 15(8): p. 2318-28.

102. Zhou, W., I.C. Liang, and N.S. Yee, *Histone deacetylase 1 is required for exocrine pancreatic epithelial proliferation in development and cancer.* Cancer Biol Ther, 2011. 11(7): p. 659-70.

103. Schneider, G., et al., *Acetylation as a transcriptional control mechanism-HDACs and HATs in pancreatic ductal adenocarcinoma.* J Gastrointest Cancer, 2011. 42(2): p. 85-92.

104. Wang, W., et al., *The nuclear factor-kappa B RelA transcription factor is constitutively activated in human pancreatic adenocarcinoma cells.* Clin Cancer Res, 1999. 5(1): p. 119-27.

105. Dolcet, X., et al., *NF-kB in development and progression of human cancer.* Virchows Arch, 2005. 446(5): p. 475-82.

106. Koenig, A., et al., *NFAT-induced histone acetylation relay switch promotes c-Myc-dependent growth in pancreatic cancer cells.* Gastroenterology, 2010. 138(3): p. 1189-99 e1-2.

107. Hessmann, E., et al., *MYC in pancreatic cancer: novel mechanistic insights and their translation into therapeutic strategies.* Oncogene, 2016. 35(13): p. 1609-18.

108. Schaub, F.X., et al., *Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas.* Cell Syst, 2018. 6(3): p. 282-300 e2.

109. Osborne, C.S., et al., *Active genes dynamically colocalize to shared sites of ongoing transcription.* Nat Genet, 2004. 36(10): p. 1065-71.

110. Dekker, J., et al., *Capturing chromosome conformation.* Science, 2002. 295(5558): p. 1306-11.

111. Hoffman, E.A., et al., *Formaldehyde crosslinking: a tool for the study of chromatin complexes.* J Biol Chem, 2015. 290(44): p. 26404-11.

112. Kalhor, R., et al., *Genome architectures revealed by tethered chromosome conformation capture and population-based modeling.* Nat Biotechnol, 2011. 30(1): p. 90-8.

113. Aparicio, O., J.V. Geisberg, and K. Struhl, *Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo.* Curr Protoc Cell Biol, 2004. Chapter 17: p. Unit 17 7.

114. Buck, M.J. and J.D. Lieb, *ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments.* Genomics, 2004. 83(3): p. 349-60.

115. Robertson, G., et al., *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.* Nat Methods, 2007. 4(8): p. 651-7.

116. O'Geen, H., S. Frietze, and P.J. Farnham, *Using ChIP-seq technology to identify targets of zinc finger transcription factors.* Methods Mol Biol, 2010. 649: p. 437-55.

117. Bustin, S.A., *Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays.* J Mol Endocrinol, 2000. 25(2): p. 169-93.

118. Allison, D.B., et al., *Microarray data analysis: from disarray to consolidation and consensus.* Nat Rev Genet, 2006. 7(1): p. 55-65.

119. Yang, S.R., et al., *Cigarette smoke induces proinflammatory cytokine release by activation of NF-kappaB and posttranslational modifications of histone deacetylase in macrophages.* Am J Physiol Lung Cell Mol Physiol, 2006. 291(1): p. L46-57.

120. O'Malley, M., et al., *Effects of cigarette smoking on metabolism and effectiveness of systemic therapy for lung cancer.* J Thorac Oncol, 2014. 9(7): p. 917-26.

121. Kumar, R., et al., *Epigenomic regulation of oncogenesis by chromatin remodeling.* Oncogene, 2016. 35(34): p. 4423-36.

122. Sever, R. and C.K. Glass, *Signaling by nuclear receptors.* Cold Spring Harb Perspect Biol, 2013. 5(3): p. a016709.

123. St John, H.C., et al., *The osteoblast to osteocyte transition: epigenetic changes and response to the vitamin D3 hormone.* Mol Endocrinol, 2014. 28(7): p. 1150-65.

124. Pal, B., et al., *Global changes in the mammary epigenome are induced by hormonal cues and coordinated by Ezh2.* Cell Rep, 2013. 3(2): p. 411-26.

125. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. 98(19): p. 10869-74.

126. Muscat, G.E., et al., *Research resource: nuclear receptors as transcriptome: discriminant and prognostic value in breast cancer.* Mol Endocrinol, 2013. 27(2): p. 350-65.

127. Barnard, M.E., C.E. Boeke, and R.M. Tamimi, *Established breast cancer risk factors and risk of intrinsic tumor subtypes.* Biochim Biophys Acta, 2015. 1856(1): p. 73-85.

128. Cochrane, D.R., et al., *Role of the androgen receptor in breast cancer and preclinical analysis of enzalutamide.* Breast Cancer Res, 2014. 16(1): p. R7.

129. Sorrentino, G., et al., *Glucocorticoid receptor signalling activates YAP in breast cancer.* Nat Commun, 2017. 8: p. 14073.

130. Le Dily, F., et al., *Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation.* Genes Dev, 2014. 28(19): p. 2151-62.

131. Stavreva, D.A., et al., *Dynamics of chromatin accessibility and long-range interactions in response to glucocorticoid pulsing.* Genome Res, 2015. 25(6): p. 845-57.

132. Finnin, M.S., et al., *Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors.* Nature, 1999. 401(6749): p. 188-93.

133. Kaniskan, H.U., K.D. Konze, and J. Jin, *Selective inhibitors of protein methyltransferases.* J Med Chem, 2015. 58(4): p. 1596-629.

134. West, A.C. and R.W. Johnstone, *New and emerging HDAC inhibitors for cancer treatment.* J Clin Invest, 2014. 124(1): p. 30-9.

135. Briere, D., et al., *The class I/IV HDAC inhibitor mocetinostat increases tumor antigen presentation, decreases immune suppressive cell types and augments checkpoint inhibitor therapy.* Cancer Immunol Immunother, 2018. 67(3): p. 381-392.

CHAPTER 2

Three-dimensional analysis reveals altered chromatin interaction by enhancer inhibitors harbors TCF7L2-regulated cancer gene signature

Abstract

Distal regulatory elements influence the activity of gene promoters through chromatin looping. Chromosome conformation capture (3C) methods permit identification of chromatin contacts across different regions of the genome. However, due to limitations in the resolution of these methods, the detection of functional chromatin interactions remains a challenge. In the current study, we employ an integrated approach to define and characterize the functional chromatin contacts of human pancreatic cancer cells. We applied Tethered Chromatin Capture (TCC) to define classes of chromatin domains on a genome-wide scale. We identified three types of structural domains (TAD, boundary and gap) and investigated the functional relationships of these domains with respect to chromatin state and gene expression. We uncovered six distinct sub-domains associated with epigenetic states. Interestingly, specific epigenetically active domains are sensitive to treatment with histone acetyltransferase (HAT) inhibitors and decrease in H3K27 acetylation levels. To examine whether the subdomains that change upon drug treatment are functionally linked to transcription factor regulation, we compared TCF7L2 chromatin binding and gene regulation to HAT inhibition. We identified a subset of coding RNA genes that together can stratify pancreatic cancer patients into distinct survival groups. Overall, this study describes a process to evaluate the functional features of chromosome architecture and reveals the impact of epigenetic inhibitors on chromosome architecture and identifies genes that may provide insight into disease outcome.

Introduction

The compartmentalization of the eukaryotic genome into highly organized chromatin domains is central to the regulation of gene expression and to cellular homeostasis (Dekker & Mirny, 2016). Until recently, the genome and its structural organization has largely been studied as a unidimensional entity where local chromatin structure is regulated by epigenetic mechanisms such as post-translational histone modifications, DNA methylation and chromatin-binding proteins. However, advances in genome-wide chromatin conformation capture (3C) methods have enabled the study of the three-dimensional (3D) organization of the genome. Studies employing various 3C-based methods, including 4C, 5C, ChIA-PET and Hi-C, have been developed to map long-range chromatin interactions, and have provided experimental evidence to explore the principles of 3D genomic architecture (Ramani, Shendure, & Duan, 2016). Collectively, these approaches support a model that interphase chromosomes occupy distinct chromosome territories and provide insight into how chromosomes fold within these territories (Duan et al., 2010; Heidari et al., 2014; Lieberman-Aiden et al., 2009; Sexton et al., 2012). However, the mechanisms that underlie the partitioning of the genome into these domains and their functional importance remains poorly defined.

Analysis of Hi-C data has revealed characteristic structural features of the genome, including chromatin compartments, topologically associated domains (TADs), and chromatin loops (Dixon et al., 2012; Lieberman-Aiden et al., 2009; Rao et al., 2014; van Steensel & Dekker, 2010). These distinctive higher order chromatin structures are believed to frame long-range enhancer-promoter interactions for epigenetic gene regulation (de Laat & Duboule, 2013; Dekker, Marti-Renom, & Mirny, 2013; G. Li et al.,

2012; Sanyal, Lajoie, Jain, & Dekker, 2012). However, large-scale structural studies generally provide little mechanistic detail regarding the functional relationships between higher order chromatin structure and cell-specific gene regulation. Recent computational and statistical approaches demonstrate that Hi-C data can be used to identify interacting genomic loci at a resolution of 8-20 kb (Ay, Bailey, & Noble, 2014; Jin et al., 2013; Lan et al., 2012), providing sufficient resolution to integrate higher order chromatin organization and gene expression data.

In cancer, altered regulation of epigenetic networks plays a central role in tumorigenesis and metastasis. While DNA methylation and histone modification patterns are frequently associated with both solid and hematological malignancies, it remains to be determined if 3D chromatin states are characteristic to specific cancer types and their gene expression programs (de Laat & Duboule, 2013; Dowen et al., 2014; Gondor & Ohlsson, 2009). The reversibility of histone modifications makes them an attractive target for cancer therapy and thus defining the epigenetic landscape of specific cancer types may provide important insight into the development of new therapeutic targets. Small molecule inhibitors that target histone modifying enzymes to disrupt the cancer cell epigenome are being developed for the treatment of cancer (Perri et al., 2017). In particular, histone acetyltransferases (HATs) are emerging targets in drug discovery with potential applications in cancer and other disease models (Wapenaar & Dekker, 2016). HATs catalyze the acetylation of lysine residues on histones during the epigenetic regulation of gene transcription (Grunstein, 1997). In addition to histones, HATs mediate the lysine acetylation of transcription factors, which is important for their function (Farnham, 2009; Singh et al., 2010; Vaquerizas et al., 2009). However, currently the role

that HATs, histone acetylation and HDACs play in regulating higher order chromatin structure remains unknown.

In this study, we investigate the relationship of higher order chromatin structure, histone modification and gene expression using the human pancreatic cancer cell line PANC1. We conducted Tethered Chromatin Capture (TCC), a modified Hi-C protocol (Kalhor et al., 2011), to identify and characterize chromosome interactions and domains in PANC1 cells. We integrated the interacting regions with chromatin state information (histone modifications, DNase hypersensitivity, and RNA Polymerase II binding) to uncover distinct types of subdomains associated with specific epigenetic states. We then determined the impact of two epigenetic inhibitors that target the histone acetyltransferases CBP (ICG-001) and EP300 (C646) on chromatin architecture (Bowers et al., 2010; Eguchi, Nguyen, Lee, & Kahn, 2005). Finally, we incorporated chromatin binding and gene expression data for the transcription factor TCF7L2 to examine the association of chromosome architecture and TF-mediated gene regulation. Overall, our analysis highlights (1) a process for evaluating chromosome architecture and epigenetic states (2) the impact of two histone acetyltransferase inhibitors on chromosome architecture and (3) chromatin architecture associated with TCF7L2-mediated gene regulation.

Results

Identification of chromosomal interacting regions in PANC1

We conducted our studies of higher order chromatin structure in the human pancreatic ductal adenocarcinoma (PDAC) cell-line PANC1, which is a model used for a variety of mechanistic and functional studies of pancreatic cancer. We identified the interacting regions of chromatin via tethered conformation capture (TCC) using 2 biological replicates (Kalhor et al., 2011). The TCC protocol decreases random intermolecular ligations between DNA fragments, particularly from interchromosomal interactions. We assessed the TCC data quality by comparison to available Encyclopedia of DNA Elements (ENCODE) HiC datasets for PANC1. **Figure 2.1-A** compares the genome-wide and chromosome 17 contacts for the TCC and HiC datasets binned at 1 Mb resolution, respectively, where the heatmap color indicates the contact frequency. Both interaction maps exhibit comparable patterns of the regional enrichment of long-range interactions. However, the TCC dataset has a notable depletion of interchromosomal interactions compared to the HiC dataset with a similar percentage of cis interactions greater than 20 kb (cis and trans interactions, respectively; **Fig. 2.1-B**). Pairwise comparison of TCC and Hi-C interaction matrices binned at 1 Mb have Pearson correlation coefficients greater than 0.9 (**Fig. 2.1-C**). The correlation of the TCC replicates for each chromosome binned at different resolutions (200 kb, 500 kb and 1 Mb) also correlate well, except for chromosome 9. Poor correlation for chromosome 9 has been found in other cell types (Rao et al., 2014). Further, the PANC1 TCC and HiC datasets have a comparable number of corresponding topological associated domains (TADs) and TAD boundaries (**Appendix B-1**). Overall, these results indicate a high

degree of similarity between the TCC and HiC datasets. Thus, TCC replicates were combined for downstream analyses of PANC1 chromatin structure.

Using the merged TCC replicates, we defined a total of 1,371 TADs, 709 boundary and 71 gap domains. Boundaries are interaction-sparse regions that lack inter-domain chromosomal interactions with neighboring TAD regions, whereas gaps are regions that lack interactions and are located between two identified domains. Gaps occurred in gene deserts or centromeres and few boundaries or gaps were found between two adjacent TADs. This is consistent with other studies showing that the genome is partitioned into Mb-sized local chromatin interaction domains (Dekker et al., 2013; Dixon et al., 2012; Rao et al., 2014). 60% of genes are contained within TADs whereas boundaries and gaps contain 38% and 2% of genes, respectively.

Classification of epigenetic marks

We next characterized the epigenetic states associated with the different types of PANC1 structural domains (TADs, boundaries, and gaps). We applied a Hidden Markov Model (HMM) to segment the genome based on combinatorial epigenetic states using histone ChIP-seq data (Bonneville & Jin, 2013). A12-state HMM with a 1 kb bin size and an optimized emission probability matrix using the best Bayesian Information Criterion (BIC) scores was used (**Fig. 2.2-A**). The resulting 12 epigenetic states are referred to as S1-S12, and can be categorized by regulatory potential by the emission probability values. In particular, values greater than 0.1 are considered valid marks for that state and values larger than 0.5 represent dominant marks. S1 and S7 are one-mark states enriched with the repressive mark H3K27me3, whereas S9 is a two-mark state having both

H3K27me3 and H3K9me3. Both S2 and S11 represent regions that are depleted of any epigenetic marks (emission probabilities less than 0.015), thus are termed depleted states.

We determined the proximity of the epigenetic states by evaluating their transition probabilities (**Fig. 2.2-B**). Three states (S1, S7, and S9) also have relatively high transition probabilities to each other, indicating a strong neighborhood of interspersed H3K36me3 and H3K9me3 repressive marks. The S6 state is a one-mark state enriched only with the repressive H3K9me3 and the S8 state enriched with both H3K36me3 and POLR2A. S3 and S4 are enriched with H3K27ac/H3K9ac/H3K4me1 and DNase/POLR2A. S10 is an intermediate active state, with a pattern similar to S3 and S4, but only enriched with H3K4me1. S5 and S12 are two mixture states showing enrichment of both active and repressed marks, as well as high POLR2A.

We categorized genomic regions into 8 different categories and determined the distribution of epigenetic states within each region (**Fig. 2.2-C**). Non-promoter regions, including 5' and 3' distal and gene body (intragenic) categories are enriched in repressive states (S1, S7 and S9). Active states are enriched in 5' TSS and 3' Proximal regions (states S4 and S5). **Figure 2.2-D** shows a region that contains a gap, boundary and TAD, with the corresponding epigenetic state. The bulk of gaps are S2 domains and are depleted of any epigenetic mark (**Fig. 2.2-E**). Interestingly, there are subgroups of boundaries and TADs that have varied patterns of histone modifications. We therefore further divided these into different categories; S1/S7 dominant (repressive marks), S4/S5/S10 dominant (active marks, near a TSS), or a mixture (mixed percentage of active and repressive states). This characterization indicates distinct epigenetic states are

physically connected and certain domains contain interspersed repressive epigenetic patterns (**Appendix B-2**).

Based on the association of epigenetic states, we classified adjacent or intra-domain states by defining subdomains. These combinations of states resulted in six subdomains referred to as SD1 to SD6 (**Fig. 2.2-F**). The subdomains are on average 60 kb. SD1 is a depleted subdomain (comprised of S2 and S11) and lacks marks, SD2 and SD3 are repressed subdomains (S6/S7 and S1/S7/S9, respectively) and SD4 is a gene body subdomain (S5/S8/S12). SD5 is an active enhancer subdomain (S4/S5/S10/S11 or only S10/S11), and SD6 is an active promoter subdomain (S3/S4/S5 states), which are centered by a promoter (S4) and extend up/downstream of the 5' TSS (S3).

Correlation of sub-domains and interacting peaks

To explore the relationships between chromosomal loops epigenetic states and domains, the loci of interacting chromatin regions were determined at a 10 kb resolution in 40 kb overlapping windows (interaction peaks (IPs). We identified 30,297 significant IPs (FDR <0.1 with a peak pair distance >20 kb). 90% of IPs are intra-domain interactions, whereby the two different loci are located within the same domain. Nearly 80% of the IPs are within a TAD, 19% of the IPs are in a boundary, and very few IPs are in gap regions. Since cancer cell-lines typically harbor chromosomal abnormalities, including chromosomal amplification, we investigated whether amplified regions contribute to the set of IPs. Only 0.72% of the IPs for PANC1 are in amplified regions of PANC1 cells, confirming that amplified regions are not enriched in the set of identified long-range interactions (Lan, Farnham, & Jin, 2012). Since the small number of IPs in the

amplified regions of PANC1 cells may play important roles in gene regulation, we included them in downstream analyses, which has been done previously (Fullwood et al., 2009).

The heatmap in **Figure 2.3-A** demonstrates that the specific subdomains of interacting loci tend to be the same on either end. For instance, an IP having SD4 at one end usually has a matched SD4 at the other end. This result is consistent with the hypothesis that the two ends of an IP are indeed physically close to or interacting with each other and thus have similar epigenetic states. We also found that many IPs have at least one locus in a depleted or repressed subdomain (SD1, SD2, or SD3) (summarized in **Fig. 2.3-A**, right panel).

We annotated the IPs according to gene regions and defined six promoter-centered and one non-promoter interaction groups (**Fig. 2.3-B**); Promoter (P; -5 to +1 kb of a TSS), Distal (D; 100 kb upstream or downstream of a TSS), and Far (F) regions (greater than 100 kb from a TSS). We defined promoters that interact with each other. Because the distance between two loci was calculated using the center of each locus (as opposed to the boundary), in certain cases, the two loci of one IP could actually overlap with each other. If this occurs, then it is possible that the same promoter is identified by both loci; these are designated PP1. There are also promoter-promoter (PP) IPs, where different promoters are at each loci (PP2). If one end of an IP is in the promoter region of one gene and the other end in the distal region of the same gene, this IP is categorized as PD1. PD2 is an IP that has one end in promoter region of Gene1 and the other end in distal region of Gene2. Similarly, if one end is in the promoter region of one gene and the other end in the far region of the same gene, this IP is categorized as PF1. PF2 is an IP

that has one end in the promoter region of Gene1 and the other end in far region of Gene2. If neither end is in the promoter region of any gene, that IP is classified as non-promoter-related and given the designation O-O (total number is 19,035).

In total, we obtained 11,262 IPs associated with at least one promoter, and thus referred them as P-centered IPs or looping. We found that approximately 26% of looping events represent promoter-promoter interactions with 951 PP1 and 2,059 PP2 (interactions between the promoters of different genes), respectively. Further, 24% promoter IPs (2,724) occur with distal regions of the same gene (PD1; 2,898 genes), whereas 27% of IPs (3,045) are between the promoter and distal regions of different genes (PD2). The expression level of genes in PANC1 cells linked to each IP categories was determined (**Fig. 2.3-C**) (Gaddis et al., 2015). Interestingly, promoter-centered loops either contain genes that are in the repressed states (low expression and in SD1-3) or genes that are in the active states (higher expression and in SD4-6). Genes in IPs corresponding to SD5 and SD6 are more highly expressed as compared to any other types of subdomains. Overall, these results indicate that connecting epigenetic states to topological structure can identify epigenetic subdomains that have distinct patterns of gene expression.

HAT inhibitors affect chromosomal organization in PANC1 cells

We previously reported the impact of the histone acetyltransferase (HAT) inhibitors ICG-001 and C646 on global gene expression in PANC1 cells (Gaddis et al., 2015). To examine whether PANC1 sub-domains are functionally linked to changes in gene expression, we treated cells with ICG-001 and C646 for 24 hours and performed

TCC on control and treated cells. The total number of TADs within each chromosome is equivalent between drug-treated and control-treated PANC1 cells (**Appendix B-3**). While a large proportion of the TADs, boundaries and gaps do not change with treatment (50%, 40%, and 80%), we identified some of the chromatin domains to be sensitive to HAT inhibitor treatment. We therefore classified the domain changes (**Fig. 2.4-A**). The most frequent type of change occurred within TADs, whereby treatment increases the TAD length by a maximum of 300 kb (**Fig. 2.4-B**; conserved-expand category, yellow bar). In contrast, a TAD in treated cells that overlaps with a TAD in untreated cells but the position shifts by more than 300 kb occurs much less frequently (**Fig. 2.4-B**; the shift category, blue bar). The boundaries were most sensitive to treatment and were more prone to change than either gap or TADs (**Fig. 2.4-B**, purple bar). Pearson correlation showed that domain type changes are associated with boundaries and represent the active subdomains (SD4-6; **Appendix B-4**). This suggests that changes of a domain type, especially the transition of boundary to TADs are linked to epigenetically active regions.

To examine how the HAT inhibition impacts histone acetylation within chromatin domains, we conducted ChIP-seq for H3K27ac in ICG001- and C646-treated PANC1 cells. We calculated the log2 Fold Change (log2FC) of normalized and averaged H3K27ac read signals in 100 bp bins in drug-treated versus untreated PANC1 cells for each domain and sub-domain. We observed a minor decrease of H3K27ac at altered categorical domains (Shift/Split/Type-Change) and a slight increase of H3K27ac in the No Change category (**Fig. 2.4-C**, left panel). While we did not observe major alterations in these domain categories, there were significant differences in the sub-domains.

Specifically, all of the active subdomains (SD4-6) showed loss of H3K27ac signal in the treated cells (**Fig. 2.4-C**, right panel). H3K27ac is enriched at promoter and at distal regions. We found that the loss of H3K27ac is more profound at the promoter active subdomain SD6 than at the enhancer active subdomain SD5, suggesting that these inhibitors may affect HAT activity at promoters more than at enhancers. To assess how IPs are altered in drug-treated cells, we performed an Interaction Peak (IP) analysis (described above) and identified 10,787 IPs in ICG001-treated and 13,773 IPs in C646-treated PANC1 cells. This represents an approximate 50% reduction in total IPs in treated PANC1 cells compared to the untreated control. Additionally, we identified that only approximately 50% of IPs in drug-treated cells were concordant with the control. Thus, treatment with ICG001 and C646 results in a decline in total IPs and also generation of new IPs (**Fig. 2.4-D**).

We previously identified 2,029 differentially expressed genes (DEGs) in ICG001-treated and 1,740 DEGs in C646-treated cells compared to control cells treated with DMSO (using a log2FC cutoff of 0.5 and a detection p-value < 0.05), with an overlap of 754 DEGs common to both drugs (Gaddis et al., 2015). We integrated expression data with domains and found that approximately 70% of the genes that respond to drug treatment are located in conserved domains. Strikingly, the subdomains SD5 and SD6 contain a large number of DEGs, regardless of the type of domain or domain change they are associated with (**Fig. 2.4-E**). After further associating DEGs with regions of differential H3K27ac enrichment and with looping events, we derived a list of 784 genes for ICG001-treated cells and 380 genes for C646-treated cells, for a combined total of 992 DEGs that have altered chromatin domains.

TCF7L2-regulated genes are involved in altered chromatin interactions

ICG001 and C646 inhibit the activity of CBP and P300 HATs and likely alter key signaling pathways. ICG001 was developed to be a specific inhibitor of the Wnt signaling pathway, which is important for developmental and disease processes (Eguchi et al., 2005; Emami et al., 2004). A key transcription factor involved in this pathway is TCF7L2, which recruits CBP/P300 to its target gene regulatory elements. Our previous study assessed the impact of TCF7L2 and HAT inhibitors in PANC1 cells; however, the relationship between these processes and chromatin interactions and their epigenetic states remains unknown. TCF7L2 has been linked to a variety of human diseases such as type II diabetes and cancer (Blahnik et al., 2010; Cauchi & Froguel, 2008). In a previous study exploring cell type-specific binding patterns of TCF7L2, we showed that the majority of TCF7L2 sites co-localize with H3K4me1 and H3K27ac (Frietze et al., 2012). Given the relationship between TCF7L2 and H3K27ac marked distal regulatory elements, we hypothesized that drug treatment would affect TCF7L2-associated chromatin loops in PANC1 cells. We therefore identified promoter-distal (PD) IPs that were bound by TCF7L2 in PANC1 cells that are no longer classified as IPs in the drug-treated cells. We isolated the genes associated with these IPs and compared them to genes differentially expressed upon drug treatment or upon TCF7L2 knockdown in PANC1 cells, which we identified in a previous study (**Fig. 2.5-A**) (Gaddis et al., 2015). We found that the highest fraction of these IPs were those containing interactions between promoter and distal regions of different genes (PD2-D). We derived a list of 39 genes that are differentially expressed in drug-treated PANC1 cells and are also regulated by TCF7L2. Pathway analysis using Gene Set Enrichment Analysis (GSEA) (**Fig. 2.5-B**)

(Subramanian et al., 2005) reveals enrichment in several cancer-related pathways, including Wnt signaling. We used SurvExpress (Aguirre-Gamboa et al., 2013) to determine if these genes can stratify survival risk of pancreatic cancer patients and found that this geneset predicts a significant survival correlation (**Fig. 2.5-C**, left panel, *p*-value 2.5e-07), with high-risk patients displaying a probability of an overall worse survival rate (Aguirre-Gamboa et al., 2013). Specifically, 25 of the candidate genes showed differential gene expression between the high- versus low-risk patient groups (**Fig. 2.5-C**, right panel). Thus, our results demonstrate that the HAT inhibitors not only alter chromatin interactions but also distinguish TCF7L2-regulated genes for potentially useful clinical signatures.

Discussion

Despite advances in 3C-based chromatin interaction mapping (Dekker, Rippe, Dekker, & Kleckner, 2002; Libbrecht et al., 2015), there is a lack of understanding of how nuclear architecture affects gene expression and cellular function. In particular, our knowledge of how the 3D chromatin architecture of cancer cells contributes to cancer cell-specific gene expression programs is limited. Due to the limitation of sequencing depth and the use of 6-mer cut sites of restriction enzymes, most studies of 3D chromatin architecture thus far have focused on characterizing very large 0.7-2Mb TADs. Although such studies provide important insights into chromosomal architecture (Deng et al., 2014; Lan, Farnham, et al., 2012), studies of large domains do not address the challenge of associating chromosomal interactions with transcriptional control at the individual gene level. Recent advances in both the experimental and computational aspects of

chromosomal interaction analyses now enable the exploration of the 3D chromatin architecture of the human genome at a much higher resolution than previously possible, allowing for the construction of a detailed genome-wide interaction map (Ay et al., 2014; Jin et al., 2013; Lan, Farnham, et al., 2012). A recent study used an in situ Hi-C protocol to achieve 1-5 kb resolution of interacting genomic segments and linked chromatin loops with promoters, enhancers, and CTCF sites (Rao et al., 2014); however, it did not address the relationship between gene loops and gene regulation.

In this study, we demonstrated that the method of TCC can partition the PANC1 genome into three types of structural domains termed gap, boundary and TAD. Our results are similar to previous studies of other cell types that used different experimental chromatin interaction methods (Dixon et al., 2012; Nora et al., 2012). Interestingly, we observed that both TAD and boundary domains (which are 1-5 Mb in length) were embedded with approximately 170,000 intra-domain chromatin interactions or interaction peaks. We found that these domains could be further categorized into six types of sub-domains, each with distinct epigenetic characteristics. We note that similar types of sub-compartments were defined in a previous study (Rao et al., 2014). However, there are notable differences between the method we present here and that of which was described in the previous study. The previous method divided each of two compartments with histone marks based on underlying interaction intensity and patterns. In contrast, we first used an unbiased training process in which we trained epigenetic states on the whole genome. We then associated the states with gene structure, expression, and other features resulting in the derivation of the six sub-domains. Using this approach, we found that promoter-centered looping genes within the three active subdomains (SD4-6) showed

much higher expression than those in the two repressed subdomains (SD2-3) (**Fig. 2.3-C**), suggesting these newly defined subdomains have functional distinctions.

We further examined the relationship of histone acetylation in chromatin architecture. Although two histone acetyltransferase inhibitors, ICG001 and C646, have been previously shown to alter gene expression in cancer cells (Emami et al., 2004; Gaddis et al., 2015; Oike et al., 2014), their impact on the 3D genome and epigenome structure has not been studied. Therefore, we conducted TCC and ChIP-seq of H3K27ac in ICG001- and C646-treated PANC1 cells. Interestingly, we uncovered five major types of domain changes that occur upon treatment of PANC1 cells with ICG001 or C646 (**Fig. 2.4-A**). In regards to Type-Change domains, we found that TADs are largely conserved and stable with drug treatment whereas boundary domains tend to switch to TADs. We also found that the drugs altered chromatin structures associated with positive regulatory elements. The H3K27ac enrichment is reduced predominantly within the active enhancer subdomains (SD6; **Fig. 2.4-C**) and the most significant gene expression changes occurred in the active-promoter subdomains (SD6; **Fig. 2.4-E**). We were able to link loops that are lost upon drug treatment with a list of 39 coding genes regulated by TCF7L2, a transcription factor important for developmental processes and implicated in human disease. This subset of genes is associated with cancer-related pathways and could separate pancreatic cancer patients into distinct survival groups.

In summary, we have developed a computational analytical approach for analysis of HiC/TCC data that can identify domains and subdomains and can classify chromatin looping events. Through the use of epigenetic inhibitors, our work also provides insights into the interdependence of 3D chromatin looping and transcriptional control. We

recognize that our current studies cannot determine if the enhancers that are affected by the epigenetic drugs are the same enhancers as identified by the chromosomal looping method. Future work using CRISPR/Cas9 to delete the TCF7L2-associated enhancers within the identified promoter-enhancer loops is needed to fully elucidate the mechanistic involvement of enhancer-mediated looping events in the regulation of drug-responsive genes. Nevertheless, our work provides genome-wide evidence that a strong association exists between a subset of enhancer-associated loops and enhancer-regulated genes.

Methods

Tethered Chromatin Capture (TCC)

TCC was performed as described (Kalhor et al., 2011). Briefly, approximately $5x10^7$ PANC1 cells were crosslinked with 1% formaldehyde for 10 minutes at room temperature, crosslinking was quenched with 0.125M glycine for 5 minutes at room temperature and cell pellets were collected and stored at -80°C. Nuclei were digested with 2000U HindIII prior to dilute solid-surface ligation reactions and TCC library preparation as described (Kalhor et al., 2011). For drug treatments, PANC1 cells were grown to 60% confluency before a 48-hour treatment with 10 μM ICG-001 (Tocris), 10 μM C646 (Sigma-Aldrich) or DMSO and fixed and harvested as described above.

Frequency contact matrix of TCC data

Paired raw reads of TCC data for the PANC1 cell line were aligned to the human reference genome (hg19) by BWA (H. Li & Durbin, 2009) with default parameters. Reads were trimmed by 5 bp until 25 bp and aligned iteratively. Multiple aligned reads

and reads with a MAPQ less than 30 were removed. After performing fragment filtering (such as removing self-circles, error-pairs, and PCR duplicate reads), the reads were binned into either 500 kb or 1 Mb size bins, where the sum of interaction pairs within the bins was used for all bin-bin interactions. The construction of a frequency contact matrix was done as described previously (Bau et al., 2011; Wang et al., 2013). Briefly, binned data was first subjected to normalization and transformation into Z-scores, the distribution of chromosomal interaction frequencies of both cell lines was then examined using a 500 kb or 1 Mb resolution for intra-chromosomal and a 1 Mb resolution for whole-genome contact matrices. More specifically, for every 500 kb or 1 Mb bin of chromosome regions, the number of interactions (*i.e.*, Z-scores not equal to zero) between each chromosome region and the rest of the chromosome regions was counted. The chromosomal interaction frequency of the region was then calculated as the counted number of interactions in the region divided by the total number of chromosome regions (e.g. with a 1 Mb resolution there are 3,029 chromosome regions in the human genome, and with a 500 kb resolution there are 498 chromosome regions in chr1). Z-scores of intra- or inter-chromosomal interaction matrices were then constructed as either a genome-wide contact heat map or a chromosome-specific intra-chromosomal contact heat map.

Topological domains of TCC data

A raw interaction contact matrix of each chromosome at 100 kb resolution was normalized using Hi-Corrector (W. Li, Gong, Li, Alber, & Zhou, 2015), which implements a set of scalable algorithms adapted from the original IC algorithm (Imakaev

et al., 2012) for parallel computing. Domains were detected using TopDom (Shin et al., 2016) based on the local minima of normalized contact matrix. For two consecutive local minima, if any bin does not show a significant difference between the contact frequencies of within interactions and between interactions, they are defined as being within a topological domain (TAD); otherwise, they are either a boundary or a gap. The boundary and gap regions represent TAD-free chromatin at the given sequencing resolution and current parameter settings. We note that a boundary does not refer to the left or right side of a TAD, but is a specific region that has low interactions within itself and also between neighboring regions. Thus, based on this definition, there is not always a boundary between two TADs. A Gap is a region depleted of interactions.

HMM epigenetic states and subdomains

Histone modification marks (H3K27ac, H3K4me1, H3K4me3, H3K36me3, H3K27me3, H3K9me3), RNA Polymerase II, DNase-seq and TCF7L2 datasets in PANC1 cell lines were obtained from the ENCODE Project (Consortium, 2012; Frietze et al., 2012).

The data was trained by a univariate first-order Hidden Markov Model (HMM) (Heinz et al., 2010) to identify combinatorial epigenetic states. For each bin on the genome, the reads of each epigenetic mark were evaluated to determine if that mark is enriched in that bin (1) or not (0). We then used this binarized information of all epigenetic marks to train the HMM model for the default 300 iterations. For each combination of bin size (1 kb) and number of states (8, 10, 12, and 20), 5 trainings of the HMM model were performed and the best model was selected based on the Bayesian

information criterion (BIC). The outputs of emission and transition matrices or states were visualized using the commercial MATLAB program. Consecutive bins of the same HMM states were merged into a single region, given that the bins were within the same domain defined by TopDom.

The emission probability of the HMM represents the distribution of the epigenetic marks in that particular bin, whereas the transition probability represents the possibility that a certain state should be assigned to a specific bin given the known state of the previous bin. If in a given state there were marks with an emission probability greater than 0.5, only these marks are considered as dominant marks for that state. For states without dominant marks, we used an emission probability cutoff of 0.1 for a mark to be considered as valid to identify a corresponding state.

The transition matrix indicates which states are frequently neighbors. In addition, states that have very low values in the emission matrix, such as S2 and S11 in **Figure 2.2-A**, may represent epigenetic mark-depleted states. Therefore, based on the transition and emission matrix, as well as the other genomic features, certain epigenetic states were merged together to a single region and biologically defined as a subdomain. These subdomains reflect the epigenetic modification context over a chromatin structural domain.

TCC data modeling using HOMER

We use HOMER (Heinz et al., 2010) to find significant interactions or interaction peaks (IPs) in our TCC data. HOMER can search for pairs of loci that have a greater

number of reads in interaction data than would be expected by chance using a background model. HOMER defines the expected number of reads as

$$\boldsymbol{e_{ij}} = \frac{\boldsymbol{f(i-j)(n_i^*)(n_j^*)}}{\boldsymbol{N^*}}$$

where f is the expected frequency of reads, N* is estimated total number of reads, and n* is the estimated total number of interaction reads at each region. HOMER uses the actual number of interaction reads at each region as the initial value and then iteratively calculates the expected number of reads using the above model until the error between expected and observed reads totals per region is near zero. We examined genomic regions at 40 kb resolution to find significant interactions, using a minimum distance of 10 kb to consider an interaction between regions. The peaks were then further filtered using an FDR cutoff of 0.1 and distance cutoff of 20 kb between loci centers.

In addition to the genomic location of the two interacting regions, HOMER also outputs a binomial p-value and FDR based on Benjamini correction. We further filtered the peaks using FDR<=0.1 and loci distance greater than 20 kb to isolate a more stringent set of IPs.

Associating chromatin interactions with epigenetic subdomains

For each IP, the subdomain overlaps (1 bp) with the two loci of the peaks were extracted. The average sizes of interaction peaks are longer than that of subdomains, so one peak locus may cover multiple subdomains. Changes in histone modifications in treated PANC1 cells were calculated by first extracting and averaging reads in the

subdomains in 100 bp bin, then dividing the averaged reads in drug-treated PANC1 cells by control PANC1 cells.

Associating chromatin interactions with annotated genes

For all annotated human RefSeq genes, we defined 5 genomic regions (**Fig. 2.3-B**) relevant to transcription start sites (TSS), which are Promoter (-5 kb to +1 kb), Distal (±100 kb), and Far (beyond 100 kb). Then, we defined the following seven categories of IPs: 1) PP1—any IPs with both ends within the same Promoter; 2) PP2-- any IPs having two ends located in two different Promoters; 3) PD1-- any IPs between a Promoter and a Distal region, with the closest TSS to the Distal region being the same gene as for the Promoter end; 4) PD2-- any IPs between a Promoter and a Distal region, with the closest TSS to the Distal region NOT being the same gene as for the Promoter end; 5). PF1-- any IPs between a Promoter and a Far region, with the closest TSS to the Far region being the same gene as for the Promoter end; 6) PF2-- any IPs between a Promoter and a Far region, with the closest TSS to the Far region NOT being the same gene as for the Promoter end; and 7) Other-- any IPs that do not involve a Promoter.

Integrating gene expression datasets

All of our expression datasets were obtained from our previous study (Gaddis et al., 2015). Further details for each component of our analysis are described. For assessing the differential expression analysis for IP alterations within untreated and HAT-inhibitor treated PANC1 cells, total RNA was collected for untreated PANC1 cells and cells treated with either epigenetic inhibitor for 96 hours. Total RNA was collected using

Trizol (Life Technologies). Ultimately, these RNAs were labeled, hybridized and analyzed with Illumina HT-12 v4 Expression BeadChips (catalog#: BD-103-0204) with the Direct Hybridization Assay and then scanned on an Illumina HiScan (catalog#: BD-103-0604). We analyzed the data as described. For this analysis we used a log2FC cutoff of 0.5 and *p*-value <0.05 for further analyses.

To incorporate TCF7L2 regulation we utilized our knockdown RNA-seq data. Total RNA after knockdown with 40nM siRNA targeting TCF7L2 or an siControl. We then performed RNA-sequencing on the polyA+ RNA selected True-Seq libraries using the Illumina HiSeq2000 platform and differential expression was determined as described (Gaddis et al., 2015). For this analysis we used a log2FC cutoff of 0.5 and *p*-value <0.05 for further analyses. Lastly, for determining the expression levels of PANC1 genes within promoter-centeric IPs we used our control dataset from the RNAseq experiment.

Ontology analyses were performed using GSEA (Subramanian et al., 2005) with default settings and survival analysis was done using SurvExpress (Aguirre-Gamboa et al., 2013).

Figures

Figure 2.1: Characteristics of interacting chromatin regions in PANC1 cells

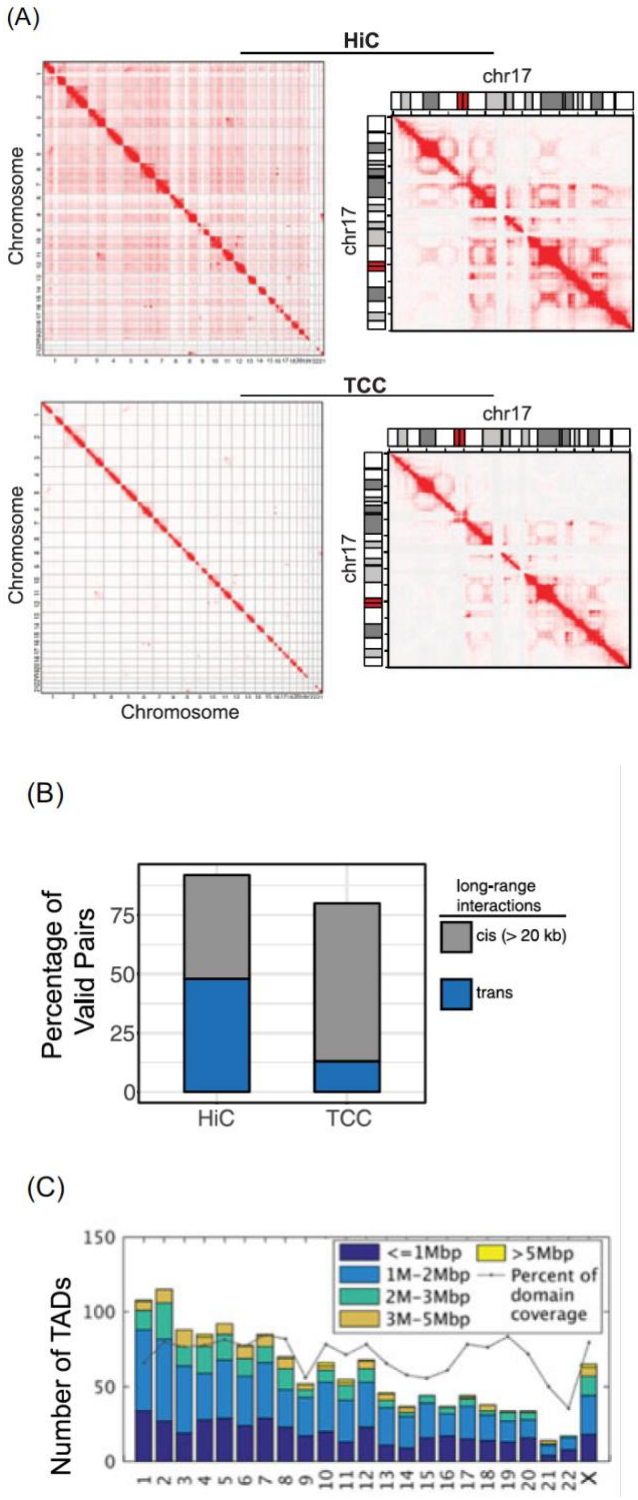


Figure 2.1: Characteristics of interacting chromatin regions in PANC1 cells. A, Genome-wide and chromosome 17 interaction matrices for PANC1 HiC (top) and TCC (bottom) datasets. The color intensity represents the normalized number of contacts between a pair of loci and the chromosome numbers are indicated on the outside of the matrix. B, The observed proportions of intra- and interchromosomal interactions in the valid HiC pairs using HiC or TCC (cis and trans, respectively). C, Histogram displaying the size distribution of TADs within each individual chromosome. TCC, tethered chromatin capture; TADs, topological associated domains

Figure 2.2. Classification of PANC1 domains with epigenetic marks

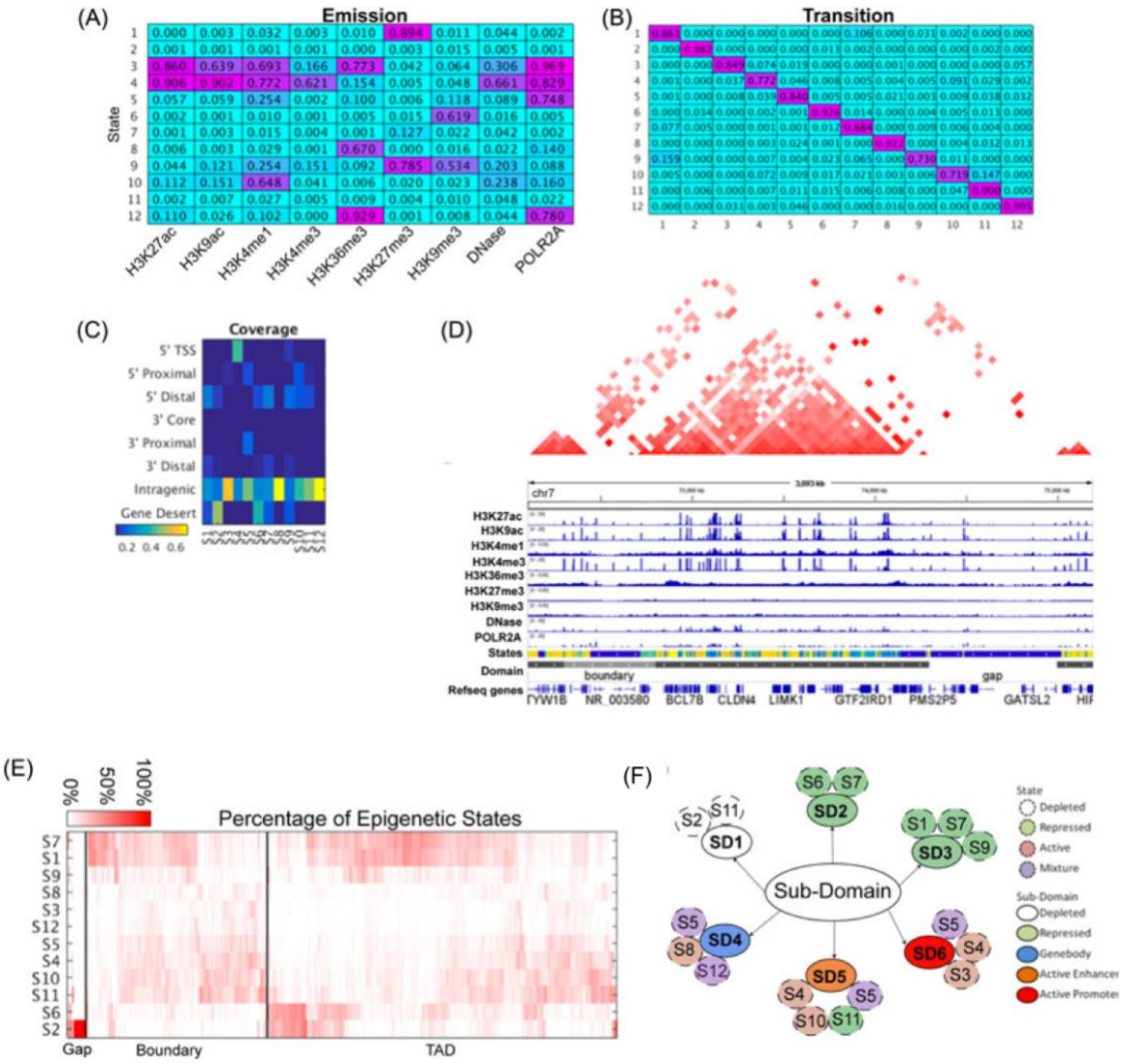
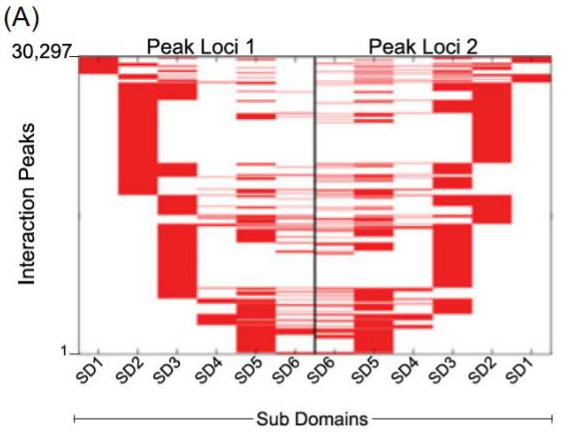
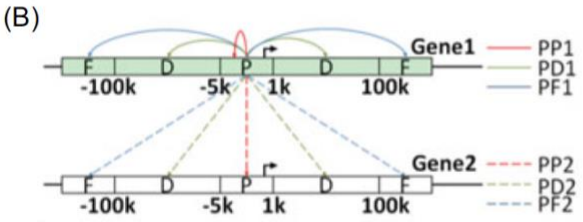


Figure 2.2. Classification of PANC1 domains with epigenetic marks. A, Emission probabilities of 12 epigenetic states trained by an HMM model on seven histone modifications, DNase and POLR2A. Marks containing emission probability values greater than 0.1 for a given state are considered to be valid and values greater than 0.5 are considered valid marks for that state and values larger than 0.5 represent dominant marks. B, Transition probabilities of the 12 epigenetic states mentioned in (A) with a high transition indicating a higher probability that a state is assigned to a given bin due to the state of the previous bin. C, Genome-wide location analysis of the 12 epigenetic states defined in (A). D, Illustration of one genomic region along chromosome 7 displaying a TAD, boundary and gap domain with the corresponding IGV snapshots of 7 histone modifications, DNase and POLR2A. E, Heatmap displaying clustering of the 12 epigenetic states within the corresponding domains. Each row corresponds to one domain and each column represents the percentage of each epigenetic state in each domain. Columns are clustered based on the TAD domains. F, We then categorized the 12 epigenetic states based on their regulatory potential, these new categorizations are referred to as "sub-domains." We identified six sub-domains (SD1-6). The epigenetic mark-depleted states S2 and S11 are merged into SD1, the interspersed S6-S7 transition regions are merged into a repressed SD2, the interspersed S1-S7-S9 regions are merged into a repressed SD3, and the regions having active states are merged into a genebody SD4 and two active enhancer/active promoter SD5 and SD6. TADs, topological associated domains; HMM, Hidden Markov Model.

Figure 2.3: The relationship between interaction peaks and sub-domains



Sub-Domain	States	Regulation
SD1	S2,S11	Depleted
SD2	S6,S7	Repressed
SD3	S1,S7,S9	Repressed
SD4	S5,S8,S12	Gene body
SD5	S4,S10,S11,S5	Active enhancer
SD6	S5,S4,S3	Active promoter



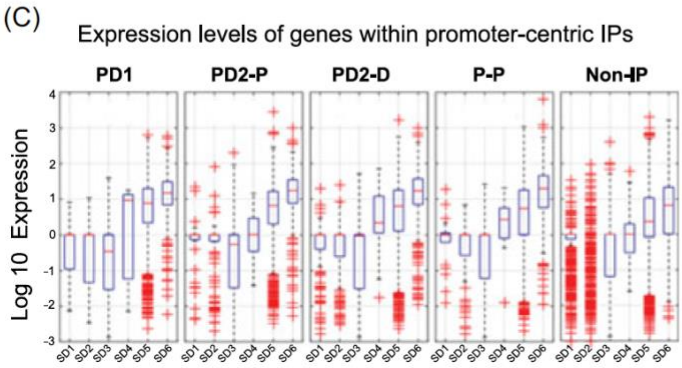


Figure 2.3: The relationship between interaction peaks and sub-domains. A, Subdomains at each locus of interaction peaks (IPs) (left). ‘Peak Loci 1′ and ‘Peak Loci 2′ represent the two ends of an IP. The table on the right summarizes the sub-domains identified in (Figure 2). B, Classification of promoter-centric IPs based on nearest genes. We defined a promoter region (P) to include 5 kb upstream to 1 kb downstream of a TSS, a distal region (D) as 100 kb upstream or downstream of a TSS, and any region beyond 100 kb from a TSS as a far (F) region. IPs for which the same gene is the nearest gene to both ends are defined as PP1, PD1, and PF1 whereas loops in which the nearest gene is different for each end are denoted as PP2, PD2, and PF2. C, Boxplots of expression for genes associated with promoter-centric IPs in PANC1 cells. For PD1 genes, only one gene is involved and that data is plotted in the PD1 panel, for PD2 genes, the expression of the gene at the promoter end is plotted in the PD2-P panel and the expression of the gene at the other end is plotted in the PD2-D panel. P-P is PP1 and PP2 combined. The expression of all other genes that are not involved in IPs in PANC1 cells are plotted in the Non-IP panel. The genes are grouped by the type of subdomain where the promoter is located. PP, promoter-promoter; TSS, transcription start sites.

Figure 2.4: Effects of histone acetyltransferase inhibitors on chromatin loops and gene expression in PANC1 cells

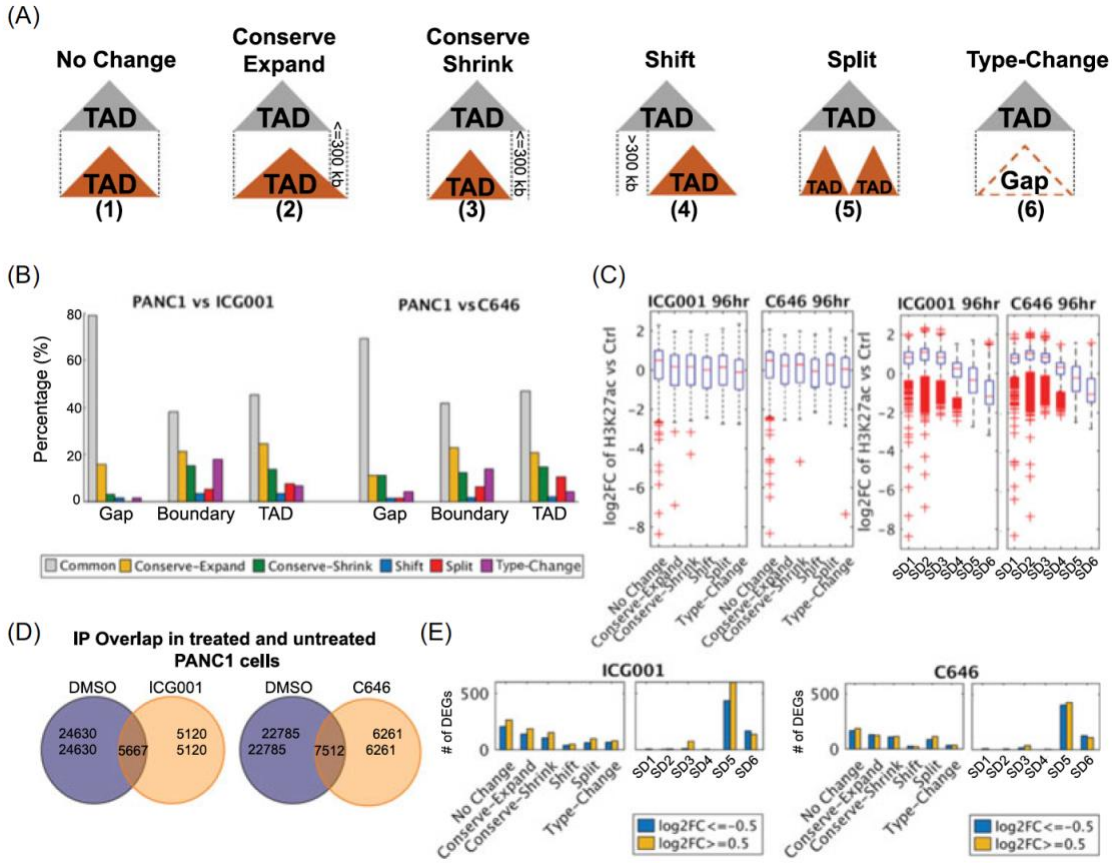


Figure 2.4: Effects of histone acetyltransferase inhibitors on chromatin loops and gene expression in PANC1 cells. A, Diagram of observed types of domain alterations after HAT inhibitor treatment. (1) No change – regions which match exactly between control and treated cells, (2) conserve-expand – regions identified as a TAD in both control and treated cells, with the length of the TAD increasing by at most 300 kb in treated cells, (3) conserve-shrink – regions identified as a TAD in both control and treated cells, with the length of the TAD decreasing by at most 300 kb in treated cells, (4) shift – a region identified as a TAD in treated cells that overlaps with a TAD in control cells, with the position shifting by more than 300 kb, (5) split – a region identified as one TAD in control cells but covers multiple TADs in treated cells, and (6) type-change – a region identified as a TAD in control cells but has switched to a gap or boundary in treated cells. B, Percentage of domain changes after treatment for 96 hours with ICG001 (left) and C646 (right). C, Log2 fold change of H3K27ac levels after treatment, separated by type of domain changes described in (A) (left panels) and type of sub-domains (right panels). D, Overlap between IPs identified in untreated and ICG001 treated (left) or C646 treated (right) PANC1. E, Number of differentially expressed genes after ICG001 (left) or C646 (right) treatment, separated by type of domain change or sub-domain. HAT, histone acetyltransferase; TADs, topological associated domains.

Figure 2.5: Effects of histone acetyltransferase inhibitors on TCF7L2-mediated looping in PANC1 cells

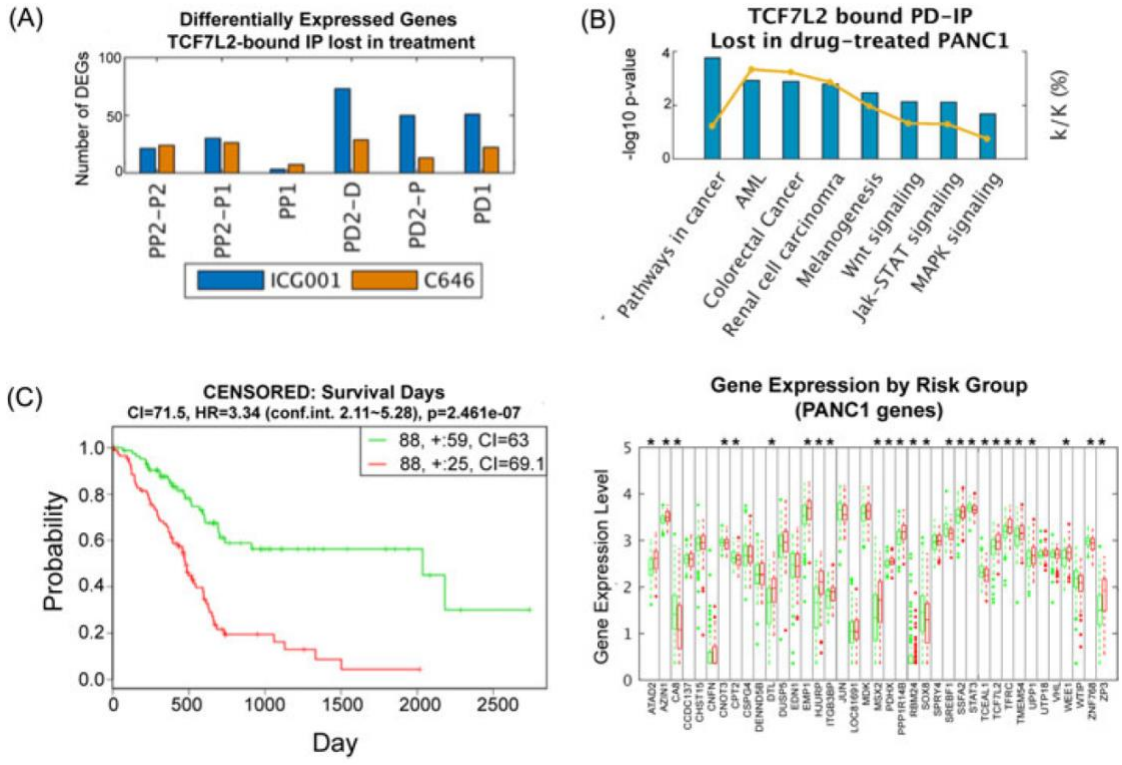


Figure 2.5: Effects of histone acetyltransferase inhibitors on TCF7L2-mediated looping in PANC1 cells. A, Number of differentially expressed genes within promoter-centric IPs that were bound by TCF7L2 in untreated PANC1 cells that are no longer classified as IPs in the drug treated cells. These differentially expressed genes were altered in siTCF7L2 knockdown cells as well as drug treated cells. PP2-P1 and PP2-P2 are genes of which the promoters are associated with a PP2 IP, PP1 is the gene of which the promoter is associated with a PP1 IP, PP2-D is the distal gene, while PP2-P is the promoter gene that are associated with a PD2 IP and PD1 is the gene that is associated with PD1 IP. B, KEGG pathway analysis of the genes (n = 39) that are associated with promoter-distal interactions and that are differentially expressed in drug treated PANC1 cells that are regulated by TCF7L2. C, Survival analysis of 176 TCGA pancreatic adenocarcinoma patients of the 39 genes identified in (A). The red line is the survival of the High Risk group, and green line is the survival of the low risk group patients. "+" in the legend stands for the censored patients in each risk group (left). Boxplots displaying the expression of the 39 genes in the two risk groups (right). CI, confidence interval; HR, hazard ratio; P, P value

References:

1. Aguirre-Gamboa, R., Gomez-Rueda, H., Martinez-Ledesma, E., Martinez-Torteya, A., Chacolla-Huaringa, R., Rodriguez-Barrientos, A., . . . Trevino, V. (2013). SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis. *PLoS One, 8*(9), e74250. doi:10.1371/journal.pone.0074250
2. Ay, F., Bailey, T. L., & Noble, W. S. (2014). Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res, 24*(6), 999-1011. doi:10.1101/gr.160374.113
3. Bau, D., Sanyal, A., Lajoie, B. R., Capriotti, E., Byron, M., Lawrence, J. B., . . . Marti-Renom, M. A. (2011). The three-dimensional folding of the alpha-globin gene domain reveals formation of chromatin globules. *Nat Struct Mol Biol, 18*(1), 107-114. doi:10.1038/nsmb.1936
4. Blahnik, K. R., Dou, L., O'Geen, H., McPhillips, T., Xu, X., Cao, A. R., . . . Farnham, P. J. (2010). Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data. *Nucleic Acids Res, 38*(3), e13. doi:10.1093/nar/gkp1012
5. Bonneville, R., & Jin, V. X. (2013). A hidden Markov model to identify combinatorial epigenetic regulation patterns for estrogen receptor alpha target genes. *Bioinformatics, 29*(1), 22-28. doi:10.1093/bioinformatics/bts639
6. Bowers, E. M., Yan, G., Mukherjee, C., Orry, A., Wang, L., Holbert, M. A., . . . Cole, P. A. (2010). Virtual ligand screening of the p300/CBP histone

acetyltransferase: identification of a selective small molecule inhibitor. *Chem Biol, 17*(5), 471-482. doi:10.1016/j.chembiol.2010.03.006

7. Cauchi, S., & Froguel, P. (2008). TCF7L2 genetic defect and type 2 diabetes. *Curr Diab Rep, 8*(2), 149-155.

8. Consortium, E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*(7414), 57-74. doi:10.1038/nature11247

9. de Laat, W., & Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature, 502*(7472), 499-506. doi:10.1038/nature12753

10. Dekker, J., Marti-Renom, M. A., & Mirny, L. A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet, 14*(6), 390-403. doi:10.1038/nrg3454

11. Dekker, J., & Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. *Cell, 164*(6), 1110-1121. doi:10.1016/j.cell.2016.02.007

12. Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science, 295*(5558), 1306-1311. doi:10.1126/science.1067799

13. Deng, W., Rupon, J. W., Krivega, I., Breda, L., Motta, I., Jahn, K. S., . . . Blobel, G. A. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell, 158*(4), 849-860. doi:10.1016/j.cell.2014.05.050

14. Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., . . . Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature, 485*(7398), 376-380. doi:10.1038/nature11082

15. Dowen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., . . . Young, R. A. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell, 159*(2), 374-387. doi:10.1016/j.cell.2014.09.030

16. Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., . . . Noble, W. S. (2010). A three-dimensional model of the yeast genome. *Nature, 465*(7296), 363-367. doi:10.1038/nature08973

17. Eguchi, M., Nguyen, C., Lee, S. C., & Kahn, M. (2005). ICG-001, a novel small molecule regulator of TCF/beta-catenin transcription. *Med Chem, 1*(5), 467-472.

18. Emami, K. H., Nguyen, C., Ma, H., Kim, D. H., Jeong, K. W., Eguchi, M., . . . Kahn, M. (2004). A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. *Proc Natl Acad Sci U S A, 101*(34), 12682-12687. doi:10.1073/pnas.0404875101

19. Farnham, P. J. (2009). Insights from genomic profiling of transcription factors. *Nat Rev Genet, 10*(9), 605-616. doi:10.1038/nrg2636

20. Frietze, S., Wang, R., Yao, L., Tak, Y. G., Ye, Z., Gaddis, M., . . . Jin, V. X. (2012). Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. *Genome Biol, 13*(9), R52. doi:10.1186/gb-2012-13-9-r52

21. Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., . . . Ruan, Y. (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature, 462*(7269), 58-64. doi:10.1038/nature08497

22. Gaddis, M., Gerrard, D., Frietze, S., & Farnham, P. J. (2015). Altering cancer transcriptomes using epigenomic inhibitors. *Epigenetics Chromatin, 8*, 9. doi:10.1186/1756-8935-8-9

23. Gondor, A., & Ohlsson, R. (2009). Chromosome crosstalk in three dimensions. *Nature, 461*(7261), 212-217. doi:10.1038/nature08453

24. Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature, 389*(6649), 349-352. doi:10.1038/38664

25. Heidari, N., Phanstiel, D. H., He, C., Grubert, F., Jahanbani, F., Kasowski, M., . . . Snyder, M. P. (2014). Genome-wide map of regulatory interactions in the human genome. *Genome Res, 24*(12), 1905-1917. doi:10.1101/gr.176586.114

26. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell, 38*(4), 576-589. doi:10.1016/j.molcel.2010.05.004

27. Imakaev, M., Fudenberg, G., McCord, R. P., Naumova, N., Goloborodko, A., Lajoie, B. R., . . . Mirny, L. A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods, 9*(10), 999-1003. doi:10.1038/nmeth.2148

28. Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., . . . Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature, 503*(7475), 290-294. doi:10.1038/nature12644

29. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., & Chen, L. (2011). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat Biotechnol, 30*(1), 90-98. doi:10.1038/nbt.2057

30. Lan, X., Farnham, P. J., & Jin, V. X. (2012). Uncovering transcription factor modules using one- and three-dimensional analyses. *J Biol Chem, 287*(37), 30914-30921. doi:10.1074/jbc.R111.309229

31. Lan, X., Witt, H., Katsumura, K., Ye, Z., Wang, Q., Bresnick, E. H., . . . Jin, V. X. (2012). Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages. *Nucleic Acids Res, 40*(16), 7690-7704. doi:10.1093/nar/gks501

32. Li, G., Ruan, X., Auerbach, R. K., Sandhu, K. S., Zheng, M., Wang, P., . . . Ruan, Y. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell, 148*(1-2), 84-98. doi:10.1016/j.cell.2011.12.014

33. Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics, 25*(14), 1754-1760. doi:10.1093/bioinformatics/btp324

34. Li, W., Gong, K., Li, Q., Alber, F., & Zhou, X. J. (2015). Hi-Corrector: a fast, scalable and memory-efficient package for normalizing large-scale Hi-C data. *Bioinformatics, 31*(6), 960-962. doi:10.1093/bioinformatics/btu747

35. Libbrecht, M. W., Ay, F., Hoffman, M. M., Gilbert, D. M., Bilmes, J. A., & Noble, W. S. (2015). Joint annotation of chromatin state and chromatin conformation reveals relationships among domain types and identifies domains of

cell-type-specific expression. *Genome Res, 25*(4), 544-557. doi:10.1101/gr.184341.114

36. Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., . . . Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science, 326*(5950), 289-293. doi:10.1126/science.1181369

37. Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., . . . Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature, 485*(7398), 381-385. doi:10.1038/nature11049

38. Oike, T., Komachi, M., Ogiwara, H., Amornwichet, N., Saitoh, Y., Torikai, K., . . . Kohno, T. (2014). C646, a selective small molecule inhibitor of histone acetyltransferase p300, radiosensitizes lung cancer cells by enhancing mitotic catastrophe. *Radiother Oncol, 111*(2), 222-227. doi:10.1016/j.radonc.2014.03.015

39. Perri, F., Longo, F., Giuliano, M., Sabbatino, F., Favia, G., Ionna, F., . . . Pisconti, S. (2017). Epigenetic control of gene expression: Potential implications for cancer treatment. *Crit Rev Oncol Hematol, 111*, 166-172. doi:10.1016/j.critrevonc.2017.01.020

40. Ramani, V., Shendure, J., & Duan, Z. (2016). Understanding Spatial Genome Organization: Methods and Insights. *Genomics Proteomics Bioinformatics, 14*(1), 7-20. doi:10.1016/j.gpb.2016.01.002

41. Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., . . . Aiden, E. L. (2014). A 3D map of the human genome at

kilobase resolution reveals principles of chromatin looping. *Cell, 159*(7), 1665-1680. doi:10.1016/j.cell.2014.11.021

42. Sanyal, A., Lajoie, B. R., Jain, G., & Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature, 489*(7414), 109-113. doi:10.1038/nature11279

43. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., . . . Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell, 148*(3), 458-472. doi:10.1016/j.cell.2012.01.010

44. Shin, H., Shi, Y., Dai, C., Tjong, H., Gong, K., Alber, F., & Zhou, X. J. (2016). TopDom: an efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res, 44*(7), e70. doi:10.1093/nar/gkv1505

45. Singh, B. N., Zhang, G., Hwa, Y. L., Li, J., Dowdy, S. C., & Jiang, S. W. (2010). Nonhistone protein acetylation as cancer therapy targets. *Expert Rev Anticancer Ther, 10*(6), 935-954. doi:10.1586/era.10.62

46. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., . . . Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A, 102*(43), 15545-15550. doi:10.1073/pnas.0506580102

47. van Steensel, B., & Dekker, J. (2010). Genomics tools for unraveling chromosome architecture. *Nat Biotechnol, 28*(10), 1089-1095. doi:10.1038/nbt.1680

48. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. *Nat Rev Genet, 10*(4), 252-263. doi:10.1038/nrg2538

49. Wang, J., Lan, X., Hsu, P. Y., Hsu, H. K., Huang, K., Parvin, J., . . . Jin, V. X. (2013). Genome-wide analysis uncovers high frequency, strong differential chromosomal interactions and their associated epigenetic patterns in E2-mediated gene regulation. *BMC Genomics, 14*, 70. doi:10.1186/1471-2164-14-70

50. Wapenaar, H., & Dekker, F. J. (2016). Histone acetyltransferases: challenges in targeting bi-substrate enzymes. *Clin Epigenetics, 8*, 59. doi:10.1186/s13148-016-0225-2

CHAPTER 3

Disruption of broad epigenetic domains in PDAC cells by HAT inhibitors

Abstract

The spreading of epigenetic domains has emerged as a distinguishing epigenomic phenotype for diverse cell types. In particular, clusters of H3K27ac- and H3K4me3-marked elements, referred to as super-enhancers and broad H3K4me3 domains, respectively, have been linked to cell identity and disease states. Here, we characterized the broad domains from different pancreatic ductal adenocarcinoma (PDAC) cell lines that represent distinct histological grades. We find that distinct PDAC grades exhibit characteristic broad epigenetic features that are predictive of patient prognosis and provide insight into pancreatic cancer cell identity. In particular, we find that genes marked by overlapping Low-Grade broad domains correspond to epithelial phenotype and hold potential as a marker for patient stratification. We further used ChIP-seq to compare the effects of histone acetyltransferase (HAT) inhibitors to detect global changes in histone acetylation and methylation levels at broad domains. HAT inhibitors treatment influence subclasses of broad domains in pancreatic cancer cells, which are potentially reflective of therapeutic responses. Thus revealing imperative insight into nuclear signals. The results reveal potential roles for broad domains in cells from distinct PDAC grades and demonstrates domain-specific responses to epigenetic inhibition.

Background

Cancer is a complex disease arising from both genetic and epigenetic alterations that impact changes in gene expression to drive and maintain the malignant phenotype. In recent years, epigenomic profiling has revealed that cancer progression involves a global reprogramming of networks of functional DNA regulatory elements including enhancers [1]. Enhancers are cis-acting elements that positively control the transcription of target genes and play central roles in regulating cell-type or tissue-type specific genes during development and differentiation [2]. Enhancer sequences are comprised of DNA sequence motifs that allow transcription factors to bind in a sequence-specific manner, and to recruit various histone writers to regulate transcriptional regulation. Recently, clusters of enhancer elements, referred to as super-enhancers have been linked to cell identity and disease states [3-7]. In addition, regions with widespread H3K4me3 modification called broad H3K4me3 domains have also emerged as important domains linked to the expression of tumor suppressor and cell identity genes [8, 9]. Understanding the functional roles of these epigenomic domains in different cancer types has the potential to uncover new strategies for the development of new cancer therapies [10].

Pancreatic ductal adenocarcinoma (PDAC) is the most common form of pancreatic cancer and ranks as one of the deadliest diseases with a five-year survival rate of less than 5% [11, 12]. PDAC is associated with a number of genetic and epigenetic alterations, leading to the activation of growth promoting and cell survival pathways and the inactivation of apoptotic and tumor suppressor pathways [13]. Recent reports have demonstrated the PDAC enhancer landscape and have classified enhancers associated

with PDAC progression [14-16]. However, the genes regulated by broad epigenomic domains in PDAC cells remains unanswered.

To increase our understanding of broad epigenomic domains and their association with PDAC gene regulation in cancer progression, we classified super-enhancer and broad H3K4me3 domains in human PDAC cell lines. We specifically defined groups of epigenomic domains that correspond to distinct histological grades and compared their enriched pathways and linked gene expression levels. We show that broad domains correlate with clinical features and hold potential as markers for patient stratification. As epigenetic inhibitors are promising avenues for cancer treatment, we also explored the ability of these compounds to target PDAC epigenomic domains.

Results

Classifying the broad domains of different PDAC cell lines

We analyzed ChIP-seq data from different human PDAC cell lines to identify super-enhancer and broad H3K4me3 domains, respectively [14]. This data was derived from a panel of human PDAC cell lines that are representative of both Low and High PDAC tumor grades, based on genotypic and phenotypic characteristics [17-19]. For example, the 'High-Grade' PANC1, MiaPaCa2 and PT45P1 cell lines all express mesenchymal genes [14, 19], show mesenchymal spindle-shaped cell morphology [20], and are considered to be poorly differentiated [17, 21, 22]. In contrast, the 'Low-Grade' PDAC cell lines CAPAN1, CAPAN2, CFPAC1 and HFPAC1 display epithelial-like features and are considered to be well-differentiated [23-26]. To determine the super-enhancer domains using these different PDAC datasets, we essentially followed the same procedures used in

Hnisz *et al*. [3] with some slight modifications. Briefly, H3K27ac peaks were called from ChIP-seq data against input and enriched peaks that clustered within 5 kb were stitched together. These stitched regions were then ranked to determine super-enhancers. Broad H3K4me3 domains were determined from enriched peaks identified from ChIP-seq, where the top 5% of peaks based on domain size were used to call broad H3K4me3 domains [27]. In total, we identified between 457 to 1,346 super-enhancers and 1,214 to 2,559 broad H3K4me3 domains in seven different PDAC cell lines that correspond to Low- and High-Grade groups, respectively (**Figure 3.1-A**). We observed that many genes were differentially marked by broad domains according to the assigned PDAC Grade group. As an example, the *VIM* gene encoding VIMENTIN, which is central to metastasis and is highly expressed in poorly differentiated High-Grade cells, is bound by both types of broad epigenetic domains only in High-Grade cells (**Figure 3.1-B**). We inspected the profiles of other regulatory histone modifications within the broad epigenetic domains using ENCODE data from PANC1 cells [28, 29]. The PANC1 super-enhancers exhibit higher H3K4me3 and H3K4me1 signal compared to typical enhancers (**Appendix C-1**). Both typical and broad H3K4me3 domains display low H3K4me1 enrichment, whereas broad H3K4me3 domains display higher H3K27ac than typical H3K4me3 regions (**Appendix C-2**). Comparison of the ChIP-seq signal of the different domains between cell lines reveals that the majority of super-enhancers and broad H3K4me3 domains are uniquely enriched in a given cell line (**Figure 3.1-C and D**).

We hypothesized that cells in separate differentiation states exhibit characteristic broad epigenetic patterns. We therefore compared the regions between different PDAC cell lines to define Grade-specific broad domains. Altogether, 38 super-enhancers were

common to all PDAC cell lines, and 61 and 224 super-enhancers were unique to High-Grade and unique to Low-Grade groups (HGU and LGU, respectively) (**Figure 3.2-A**). Similarly, we identified 228 common and 177 HGU and 302 LGU broad H3K4me3 domains (**Figure 3.2-B**). We further compared the overlap of both types of broad domains for the different PDAC groups by clustering these regions. There were 87 overlapping super-enhancer and broad H3K4me3 domains for LGU, compared to the 34 overlapping HGU (**Figure 3.2-C**). In general, Low-Grade PDAC cells had an increased number of super-enhancer domains compared to High-Grade PDAC cells, whereas both groups have a similar number of broad H3K4me3 domains. Overall, this analysis revealed that distinctive PDAC Grades exhibit characteristic broad epigenomic domains.

Broad epigenomic domains mark distinctive PDAC pathways

Prior studies have demonstrated that broad domains are associated with developmental and cell identity genes and broad H3K4me3 domains in particular have been shown to mark tumor suppressor genes [8, 9, 30]. To explore the gene pathways associated with super-enhancers and broad H3K4me3 domains, we annotated genes marked by each type of domain and determined their functional classifications (**Appendix C-3** and **Figure 3.2-D**). Interestingly, pathway enrichment analysis showed that the genes marked by overlapping domains are involved with a variety of signaling pathways that were either specific to LGU, HGU, or common to all PDAC cells. For example, pathways that enriched common to all PDAC cells included TGFB, microRNAs in cancer and cell cycle. Pathways specific to HGU included VEGF and Ras signaling pathways, whereas pathways specific to LGU were tight junction and Hippo signaling (**Figure 3.2-D**).

Broad regions predict poorer survival in PDAC patients

Super-enhancers and broad H3K4me3 domains have been linked to increased gene expression in a variety of tissue types [31, 32]. We therefore inspected the relative expression levels of genes marked by different broad domains across PDAC Grades using available RNA-seq data from the corresponding cell lines [14]. As expected, both super-enhancer and broad H3K4me3 domains unique to each Grade group (LGU and HGU) showed appreciably elevated expression in the corresponding group compared to the contrasting group (**Figure 3.2-E, Appendix C-4**). For example, genes marked by HGU super-enhancers had significantly higher expression levels in High-Grade cells compared to Low-Grade cells.

We next explored the clinical association of the gene expression for genes uniquely marked by different broad domains relative to patients' overall survival using the TCGA PDAC dataset [33]. Kaplan-Meier survival analysis showed that gene expression marked by HGU super-enhancers as well as HGU broad H3K4me3 domains are strong predictors of poor survival (**Figure 3.2-F**). In contrast, gene expression linked to LGU domains do not predict a poorer survival rate. However, the expression levels of genes bound by both LGU super-enhancers and broad H3K4me3 domains are significantly associated with a worse overall survival in PDAC patients (**Figure 3.2-F**). Overall, these results reveal that genes uniquely bound by different domains are predictive of PDAC patient outcome.

Histone acetyltransferase inhibitors alter global H3K27ac and H3K4me3 levels

Epigenetic modulation via small molecule inhibitors has been proposed as an approach for treating various malignancies, including pancreatic cancer [34, 35]. We have previously shown that the two HAT inhibitors ICG-001 and C646 differentially impair the global gene expression levels in human pancreatic and colorectal cancer cell lines [36]. However, the impact of HAT inhibitor treatment on histone acetylation remains unknown. To determine the effect of HAT inhibitors on genome-wide H3K27ac enrichment, we treated PANC1 cells with ICG-001, C646 or a vehicle control and performed H3K27ac ChIP-seq, each with biological replicates. As expected, HAT inhibitor treatment caused significant changes in genome-wide H3K27ac patterns compared to vehicle treatment (**Figure 3.3-A**). Differential H3K27ac analysis showed that 4,675 and 5,362 regions with reduced H3K27ac levels in ICG-001 and C646 treated cells, respectively (FDR <0.1). There were also 2,391 and 4,383 sites with elevated H3K27ac levels in ICG-001 or C646 treated cells, respectively (**Figure 3.3-B**). Interestingly, for either treatment the majority of higher H3K27ac enrichment clustered within the gene body (intragenic regions), whereas the bulk of reduced H3K27ac enrichment corresponded to distal intergenic regions (**Figure 3.3-C**). Examples of genes that display altered H3K27ac patterns for each treatment are shown in **Figure 3.3-D**. Pathway enrichment analysis indicated that HAT inhibitor treatments influence H3K27ac enrichment at genes that map to a variety of pathways (**Figure 3.3-E**). For example, both treatments decreased H3K27ac levels at genes that belong to pancreatic, glioma, breast and gastric cancers, whereas C646 decreased the H3K27ac at HIF-1 and phosphatidylinositol signaling gene pathways.

We further investigated the impact of ICG-001 treatment on global H3K4me3 levels in PANC1 cells. ChIP-seq for H3K4me3 showed global H3K4me3 enrichment alterations in ICG-001 treated PANC1 cells. In total there were 6,847 increased and 3,219 decreased regions (**Appendix C-5**). *PPP2R2C*, a tumor suppressor gene [37], exhibited elevated H3K4me3 signal with ICG-001 treatment compared to control. In contrast *NKIRAS1*, encoding a RAS-like protein exhibited decreased H3K4me3 signal with ICG-001 treatment (**Appendix C-5**).

HAT inhibitors alter PDAC broad epigenomic domains

We next determined the impact of HAT inhibitor treatment on H3K27ac signal at super-enhancers and found that both treatments specifically impact H3K27ac enrichment levels at many super-enhancers (**Figure 3.4-A**). In total, there were 136 and 128 super-enhancers with reduced H3K27ac levels, whereas 121 and 117 super-enhancers showed significant gains in H3K27ac levels after ICG-001 and C646 treatments, respectively (FDR < 0.1). Notably, the majority of super-enhancers with decreased H3K27ac are the same between either treatment (~90%), whereas only ~50% of the domains with increased H3K27ac are the same between either treatment (**Figure 3.4-B**). Increases in H3K27ac signal at super-enhancers was associated with elevated gene expression levels and similarly decreased enrichment at super-enhancers corresponded to a reduced gene expression levels, but only with ICG-001 treatment (**Figure 3.4-C**). Pathway enrichment analysis showed that different pathways are linked to the super-enhancers targeted by HAT inhibitors (**Figure 3.4-D**). ICG-001 treatment targets super-enhancers that correspond to genes involved with AGE-RAGE signaling complications in diabetes. In addition to super-

enhancers, ICG-001 treatment resulted in 4 decreased and 113 increased broad H3K4me3 domains. A comparison of both domains targeted by ICG-001 treatment reveals an increase of H3K27ac and H3K4me3 enrichment at 24 distinct genomic regions (**Appendix C-5).** Overall, these results delineate the broad epigenomic domains that are sensitive to HAT inhibitor treatment in PANC1 cells.

HAT inhibitor treatment targets broad domains that are enriched at TAD boundary regions

Recently, super-enhancer domains that overlap with broad H3K4me3 domains were shown to be linked to higher-order chromatin interactions, signifying a unique spatial organization of chromatin around cell-specific epigenetic domains [8, 38]. We therefore examined the relationship between higher-order chromatin organization and broad epigenetic domains in PANC1 cells. Genome-wide chromatin contacts were determined by analyzing tethered-chromatin conformation capture (TCC) data in PANC1 [39]. Chromatin contacts were partitioned into topologically associated domains (TADs) and TAD boundary regions using a resolution of 40 kb (**Figure 3.5-A** and **Appendix C-6**). We examined TADs in relation to CTCF and PANC1 broad domains. As expected, CTCF was significantly enriched at TAD boundaries (**Figure 3.5-B**). Similarly, broad H3K4me3 domains were significantly linked to TAD boundaries, however super-enhancers were not found to be enriched at TAD boundaries (**Figures 3.5-C and 3.5-D**). To study if HAT inhibitor-sensitive domains are linked to higher-order chromatin structures, we investigated the significance of association of both types of broad domains that are impacted by ICG-001 treatment with TAD boundaries. We found that the domains with increased H3K4me3 and H3K27ac enrichment were significantly associated with TAD

boundaries, whereas domains with decreased enrichment were not associated with TAD boundaries (**Appendix C-7**). Thus, these results show that broad domains that gain enrichment of either active histone modification are linked to TAD boundaries. Overall, this analysis demonstrates a global impact of drug treatment on the epigenome and shows that certain classes of broad domains within TAD boundaries are sensitive to epigenetic inhibitors.

Discussion

An improved understanding of PDAC tumor biology and tumor grading should leverage available therapies and data. Here we extend prior data that indicated HAT inhibitors elicit distinctive effects on cancer cell transcriptomes by exploring the susceptibility of broad domains to HAT inhibitor treatment [36]. While drugs that target epigenetic mediators are currently in development [40, 41], the downstream effects on the epigenome of existing drugs has not yet been thoroughly examined. In particular, epigenome-wide studies of their effects remains largely undetermined. In this study, we utilize ChIP-seq to compare the effects of treatment with C646, which is a competitive inhibitor of both p300 and CBP [42] to the effects of ICG-001, which prevents CBP interaction with the co-activator β-catenin [43, 44]. Indeed, following treatment with either drug we detect global changes in histone acetylation levels. Our results suggest that, in general, these two drugs have similar effects on the epigenome of PDAC cells; however, we were able to identify cell-specific and drug-specific responses after treatment.

We observed dramatic effects on the epigenome upon treatment with either ICG-001 or C646, with hundreds of regions showing differential enrichment of H3K27ac or

H3K4me3. Interestingly, both drugs targeted similar super-enhancers, causing a reduction in histone acetylation levels near genes involved in pancreatic cancer and other solid cancers (**Figure 3.4**). Since it is of current interest to target super-enhancers, we find that both super-enhancers and broad H3K4me3 domains are sensitive to epigenetic modulation. Thus, our results provide insight into the plasticity of these domains in response to epigenetic modulation. Future work could tailor these therapeutics to target such domains; thereby impacting specific cellular pathways involved with PDAC tumorigenesis.

The extension of epigenetic regulatory domains has emerged as a diagnostic marker that can serve to distinguish cancer cell identity and disease state. Accordingly, we characterized the broad domains in several different cell lines that represent distinct PDAC histological grades. By clustering the domains from 7 different PDAC cell lines into High- and Low-Grade groups, we find that different PDAC grades exhibit characteristic epigenetic features that are predictive of PDAC prognosis and provide insight into pancreatic cancer cell identity. Of particular interest are the genes marked by both super-enhancer and broad H3K4me3 domains in Low-Grade groups. Low-Grade groups demonstrate an enrichment for a greater number of unique gene pathways, which include several pathways significant to PDAC progression. Such pathways include tight junction, glycerophospholipid and Rap1 signaling pathways. We also provide evidence that genes marked by overlapping Low-Grade broad domains correspond to epithelial phenotype and hold potential as a marker for patient stratification (**Figure 3.2**). Thus, different PDAC grades exhibit characteristic pathways marked by broad epigenomic domains that provide insight into pancreatic cancer cell identity in the context of PDAC progression.

Pancreatic Ductal Adenocarcinoma (PDAC) broad domains span numerous distinctive loci including the HOX, SMAD, and FOX family of genes, proteins that have known roles in cell-type specific functions and are known factors in PDAC tumor cell biology [45-47]. After annotating the genes marked by different broad domains, we identified known PDAC signaling pathways including the TGFβ and MAPK pathways, which are downstream effectors of oncogenic KRAS. Oncogenic KRAS is an established driver of pancreatic cancer and several pathways that are known downstream effectors of KRAS signaling and play central roles in PDAC cancer cell growth and survival [48-50], were found to be marked by broad domains in domains common to all PDAC cells. Super-enhancers common to all PDAC cell lines were significantly enriched with a variety of cancer signaling pathways including focal adhesion, PI3K-AKT, microRNAs and Hippo signaling. LGU super-enhancers were uniquely associated with several pathways that include tight junction, Rap1 signaling and glycerophospholipid metabolism. Aberrant lipid synthesis and the reprogramming of lipid metabolism has been associated with the development and progression of pancreatic cancer [51] and several phospholipids have been identified as potential biomarkers in different types of pancreatic cancers [52, 53]. MAPK signaling was the singular KEGG pathway enriched in HGU super-enhancers (**Appendix C-3**). Similarly, the broad H3K4me3 domains common to all PDAC groups were associated with distinctive pathways including transcription corepressor, protein kinase, cadherin binding and RNA binding pathways (**Appendix C-3**). Several enriched pathways linked to LGU broad H3K4me3 domains include SMAD, protein kinase C, TGF beta, and beta-catenin pathways, whereas the HGU broad H3K4me3 domains solely mark genes enriched in transcriptional corepressor pathways. Examples of broad PDAC

epigenomic domains encompassing disease-associated genes include *SMADs* and *FOXC2* for super-enhancer regions and *MYC* and *CCND1* for broad H3K4me3 domains (**Appendix C-1 and 2**).

Interestingly, our analysis indicates that Low-Grade unique broad H3K4me3 domains are enriched for TGFβ signaling pathways. TGFβ acts as a tumor suppressor with growth-inhibitory activity in epithelial cells during early pancreatic tumorigenesis. However, TGFβ appears to promote tumor progression in advanced disease [54]. We also found broad epigenomic domains mark several others pathways with less well-characterized roles in PDAC tumor biology, including microRNAs and proteoglycans in cancer.

The histone modifications H3K4me3 has been widely recognized as a mark of active promoter regions [55]. Recent studies have correlated broad H3K4me3 domains with enhancer activity at tumor suppressor genes in normal and cancer cells to provide mutation-independent insight into tumor suppressor pathways of disease states [9]. Here, we found broad H3K4me3 domains span a number of genes including *HOX*, *MYC* and *CCND1* genes. As super-enhancers and broad H3K4me3 domains have been shown to function coordinately through chromatin interactions [8, 38], we identified regions containing both domains in both High- and Low-Grade cells. As mentioned previously, the expression of genes marked by both domains is significantly associated with poor prognosis in pancreatic cancer patients.

In summary, our data provides new perspective on the effect of HAT inhibitors on the epigenome and provides knowledge of the broad domains unique to different histological grades of pancreatic cancer. Our data show that epigenomic domains that

correlate with clinical features, that they are plastic and hold potential as markers for patient stratification.

Methods

Cell culture and epigenetic inhibitor experiments

The human cell line PANC1 (ATCC #CRL-1469) was obtained from the American Type Culture Collection. The cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. We obtained ICG-001 from Michael Kahn (University of Southern California) and C646 from VWR (catalog# 102516-240). Cells were grown to 70% confluency followed by treatment with 10μM ICG-001 or 10μM C646 and were collected after 12hrs.

ChIP-sequencing

After 12-hour incubation with either ICG-001 or C646, cells were crosslinked with 1% formaldehyde (Thermo Scientific #28908) for 10 minutes and quenched with 0.125M Glycine. The ChIP-seq experiments were further performed as described by O'Geen et al. [56] and the antibodies used for the given targets were as follows: H3K27ac (Abcam, Cambridge, MA, USA; Ab4729 lot#GR16377-1) and H3K4me3 (Abcam, Cambridge, MA, USA; Ab8580). We performed duplicate ChIP-seq experiments for each histone. For each histone ChIP-seq assay, 10μg of chromatin was incubated with (2.5-5μg) of antibody. To confirm enrichment of target sequences, we performed qPCR in ChIP versus input samples. DNA was quantified using Qubit (Invitrogen) and libraries were prepared using the

NEBNext ChIP-seq Illumina Sequencing library preparation kit (New England Biolabs, Ipswich, MA, USA).

Tethered Chromatin Capture (TCC)

Tethered chromatin capture (TCC) was performed as detailed by Kalhor et al. [57]. Briefly, approximately 5 x 10^7 PANC1 cells were crosslinked as described above for ChIP-seq experiments and cell pellets were collected and stored at -80C. Nuclei were digested with 2000U HindIII prior to dilute solid-surface ligation reactions and TCC library preparation was performed as described [57].

ChIP-sequencing analysis

For all datasets raw sequencing reads were aligned to the human reference genome hg19 using bowtie2 with default parameters [58]. ENCODE ChIP-seq datasets were obtained from GSE31755. We determined binding sites of each ChIP-seq experiment using MACS2 with default parameters with the exception of using the flag '-broad' for determining broad H3K4me3 binding sites [59]. Super-enhancer regions were identified over typical-enhancer regions using the Ranked Ordering of Super Enhancer (ROSE) tool [5, 60]. Briefly, peaks were called from the H3K27ac ChIP-seq data and stitched together in 12.5 kb windows, which were further used to identify super-enhancers (ranked cutoff score of 19701.68). Broad H3K4me3 regions were obtained via filtering for the top 5% of peaks (largest by domain size). Enrichment of signal within regions was plotted with the functions 'plotProfile' and 'plotHeatmap' within deepTools [61]. Overlapping binding regions were determined using peak interesectR. To determine peak locations relative to

gene regions, we utilized the 'annotatePeaks.pl' function within HOMER [62]. Visualization snapshots of ChIP-seq regions were obtained by building a signal track via the 'bdgcompare' utility in MACS [59]. Bigwigs were obtained by 'bedGraphToBigWig' via UCSC tools [63] and further visualized using Integrated Genomic Viewer [64, 65].

Coordination of Grade-Specific Broad Domains

H3K4me3, H3K27ac and control ChIP-seq datasets for PDAC cell lines were obtained from GSE64557 and processed as described above. We utilized the Bioconductor package 'seqsetvis' [66] to visualize the distribution and overlap of genomic regions.

Differential Binding Analysis of ChIP-seq datasets

To determine differential typical and broad H3K27ac and H3K4me3 regions we used the DIffBind R package [67, 68] with an FDR cutoff of <0.1.

Pathway Enrichment Analysis

Annotation of enriched gene pathways was performed using either ClusterProfiler or the Stanford GREAT tool [69, 70].

Integration of Gene Expression Datasets

For integrating High- and Low-Grade gene expression for the different cell lines, we retrieved the corresponding RNA-sequencing expression datasets (GSE64558). We mapped these datasets to the human genome reference, hg19 and gene counts were

nomalized using DESeq2 [71]. The mean of the normalized counts between replicates of each cell line was used to generate expression heatmaps. ICG-001 and C646 Illumina beadchip expression datasets were obtained from our previous study (GSE64038). We overlapped these differentially expressed genes within treatments with our genes annotated within drug altered broad regions to demonstrate expression changes at these genes.

Topological domains from TCC data

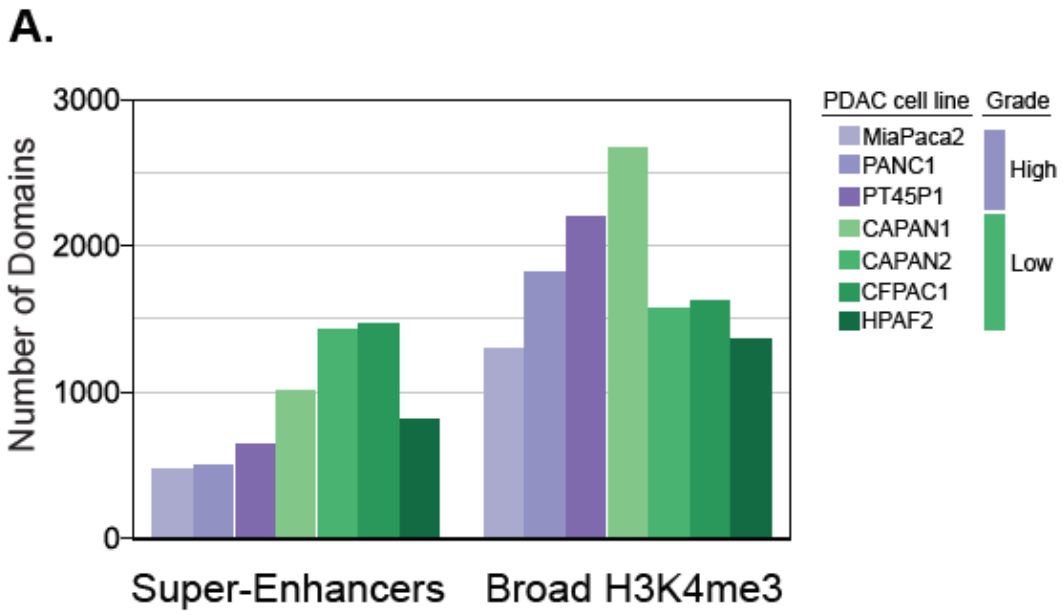
Paired raw reads of TCC data for the PANC1 cell line was processed using the HiC-Pro pipeline [72]. Briefly, these reads were aligned to the human reference genome hg19 using bowtie2 with default parameters [58]. 25bp of the reads were trimmed and the reads were then aligned iteratively. Reads with a MAPQ score less than 30 were removed and the fragments were filtered for self-ligated fragments, duplicated reads from PCR and error-pairs. Domains were detected using TopDom based on the local minima of normalized contact matrix [73, 74]. To visualize the relationship between broad domains and TADs, we used HiCPlotter [75].

Feature Enrichment Analysis within TAD Boundary Regions

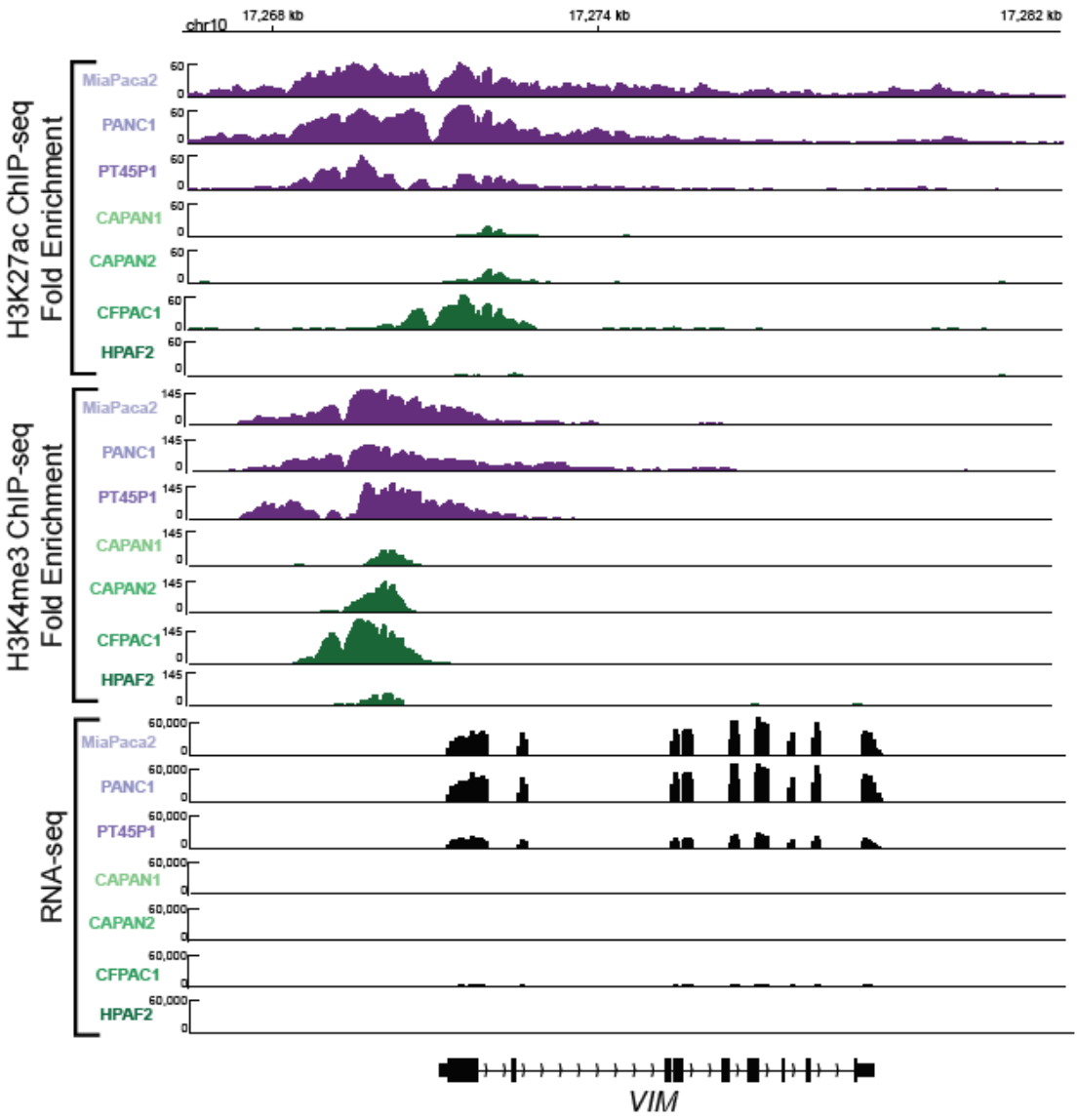
To determine enrichment of broad domains within +/- 20 kb of the Boundary regions, we iteratively determined the expected distribution either CTCF, broad H3K4me3 or super-enhancer regions within our TAD boundary domains within 8,354 randomly selected bins and iteratively repeated this 1,000 times. We then calculated our observed estimation of the given regions (either for CTCF, broad H3K4me3 or super-enhancers) within TAD Boundaries.

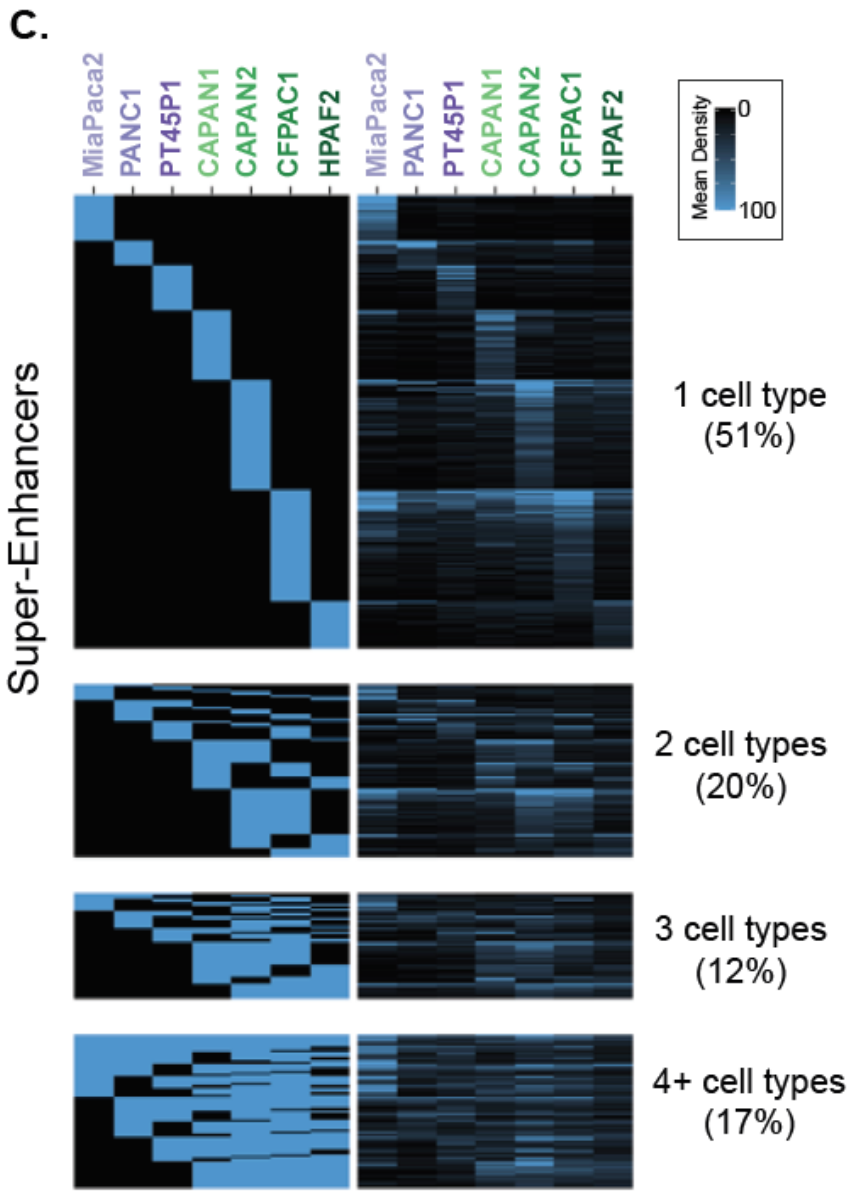
Figures

Figure 3.1: Determination of super-enhancers and broad H3K4me3 domains in PDAC cell lines that correspond to different histological grades



B.


chr10
17,268 kb
17,274 kb
17,282 kb
H3K27ac ChIP-seq Fold Enrichment
MiaPaca2
PANC1
PT45P1
CAPAN1
CAPAN2
CFPAC1
HPAF2
60
0
H3K4me3 ChIP-seq Fold Enrichment
145
RNA-seq
60,000
VIM


C.
MiaPaca2
PANC1
PT45P1
CAPAN1
CAPAN2
CFPAC1
HPAF2
MiaPaca2
PANC1
PT45P1
CAPAN1
CAPAN2
CFPAC1
HPAF2
Mean Density
0
100
Super-Enhancers
1 cell type
(51%)
2 cell types
(20%)
3 cell types
(12%)
4+ cell types
(17%)

D.

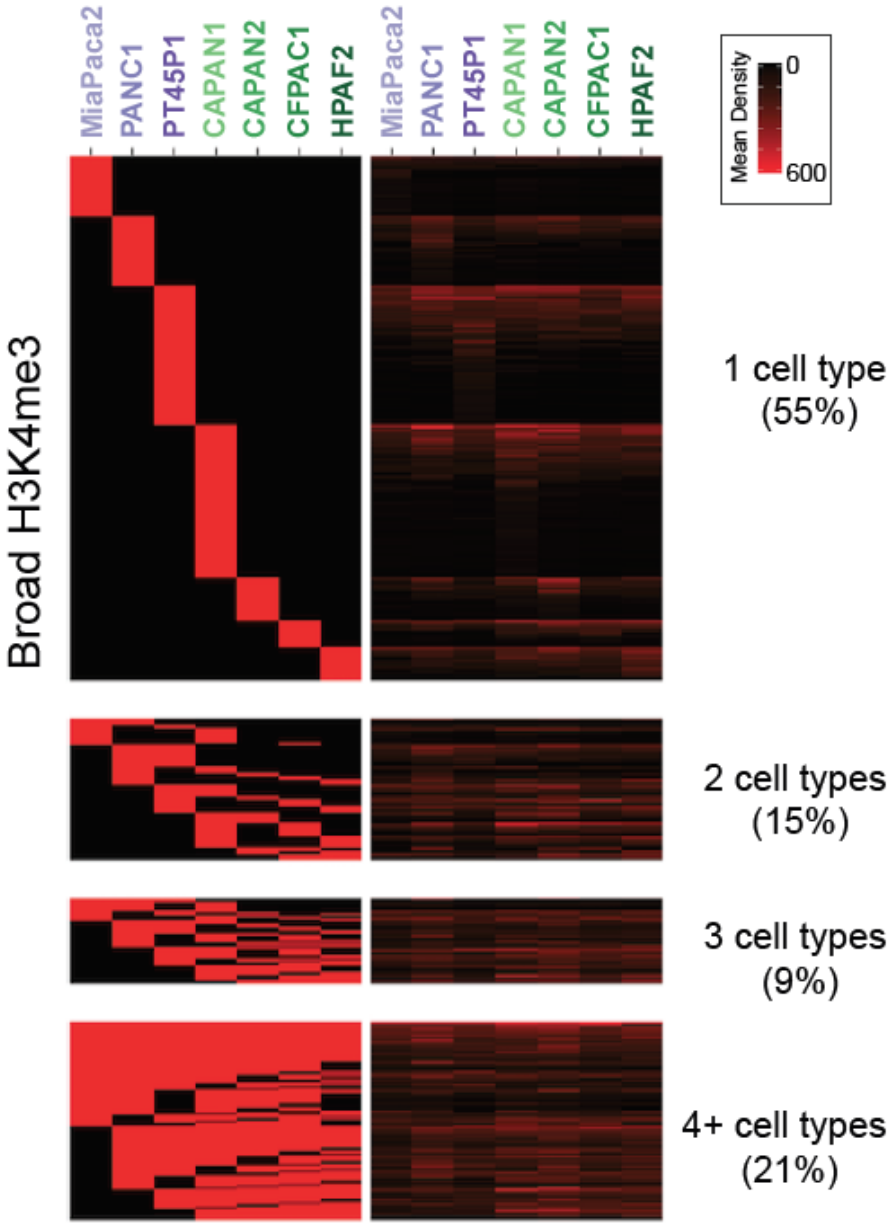
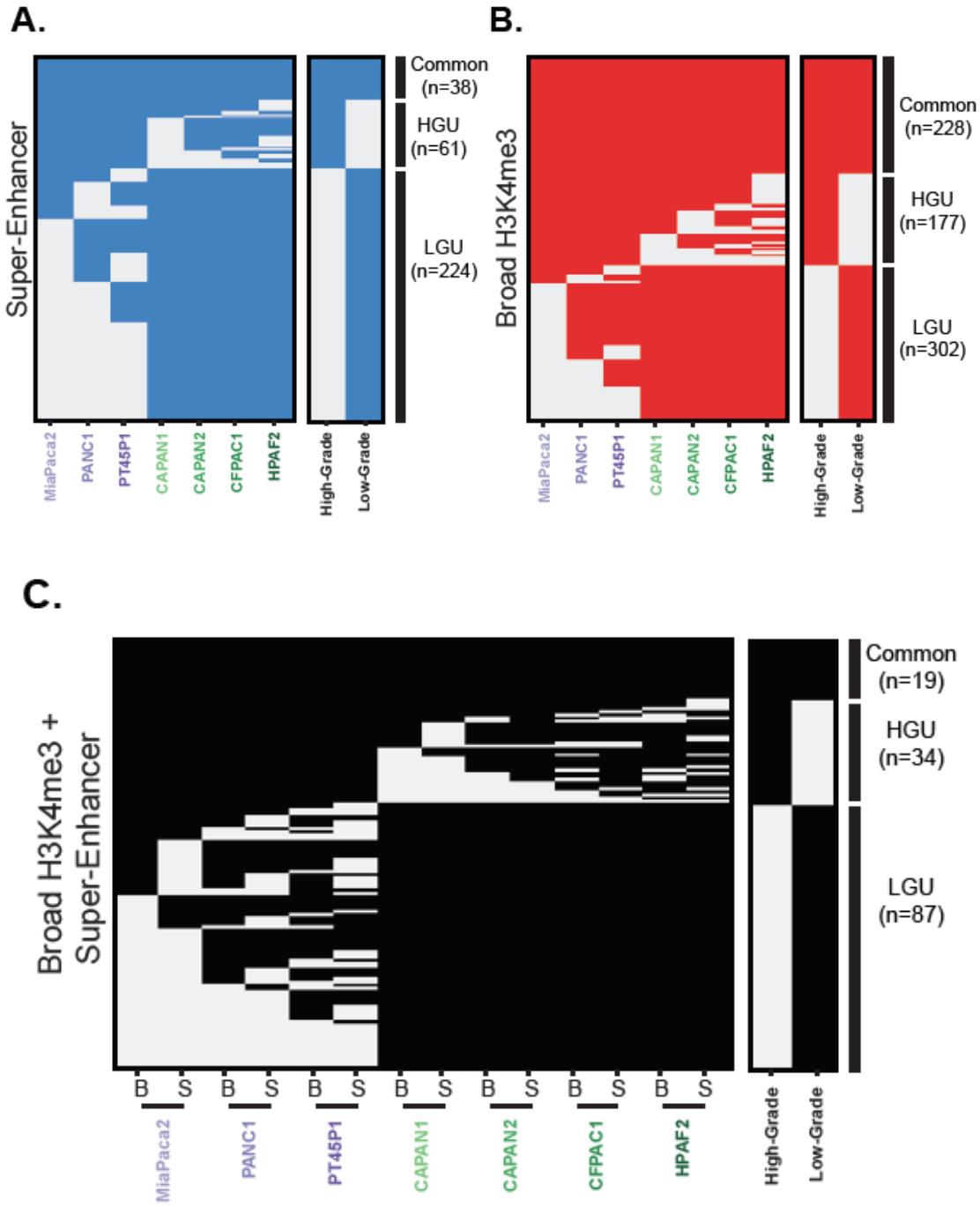
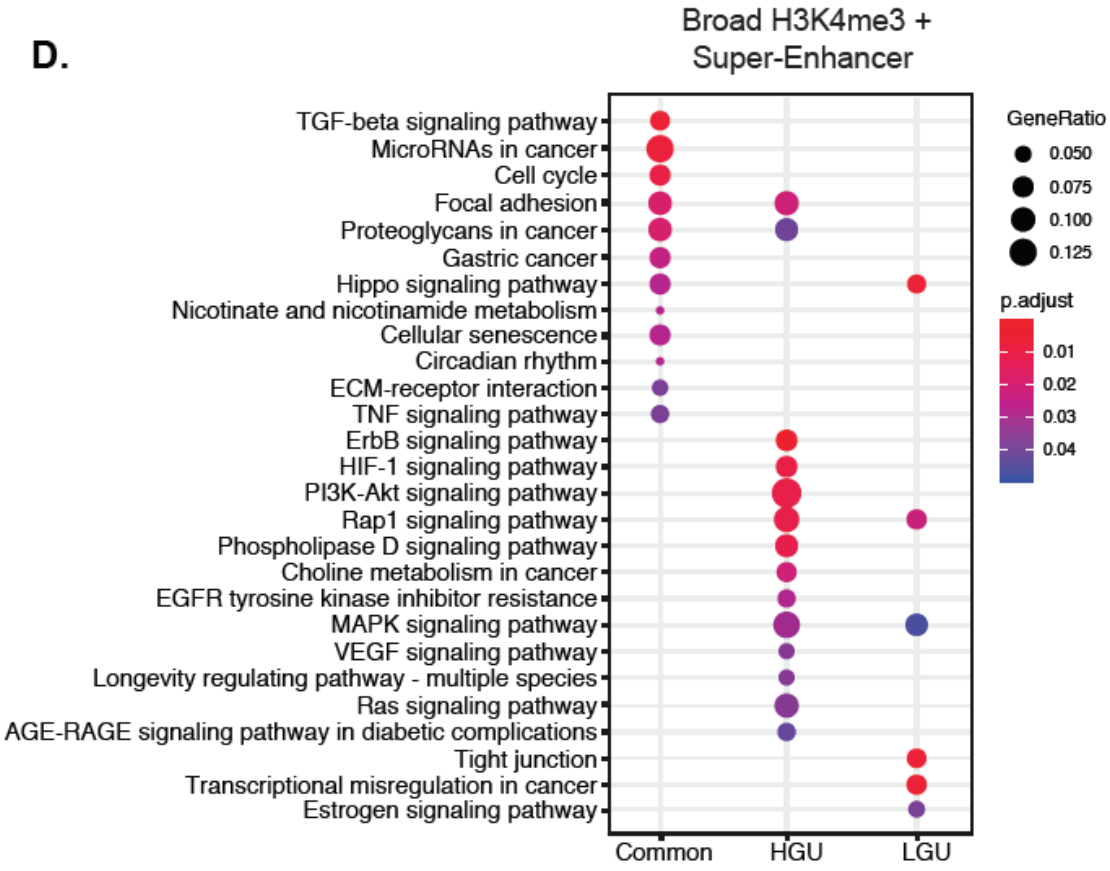

MiaPaca2
PANC1
PT45P1
CAPAN1
CAPAN2
CFPAC1
HPAF2
MiaPaca2
PANC1
PT45P1
CAPAN1
CAPAN2
CFPAC1
HPAF2
Mean Density
0
600
Broad H3K4me3
1 cell type
(55%)
2 cell types
(15%)
3 cell types
(9%)
4+ cell types
(21%)

Figure 3.1. Determination of super-enhancers and broad H3K4me3 domains in PDAC cell lines that correspond to different histological grades. A. The total number of super-enhancers and broad H3K4me3 domains found in different human PDAC cell lines that represent either High-Grade (purple) or Low-Grade (green) PDAC groups. B. Genome browser representation of the H3K27ac and H3K4me3 ChIP-seq signal, as well as the RNA-seq signal over an approximate 14 kb region surrounding the *VIM* locus. C. Heatmap displaying the classification of super-enhancer domains across 7 human PDAC cell lines. The color scale reflects the density of H3K27ac signal at super-enhancer regions. D. Heatmap displaying the classification of broad H3K4me3 domains across 7 human PDAC cell lines. The color scale reflects the density of H3K4me3 signal at broad H3K4me3 domains.

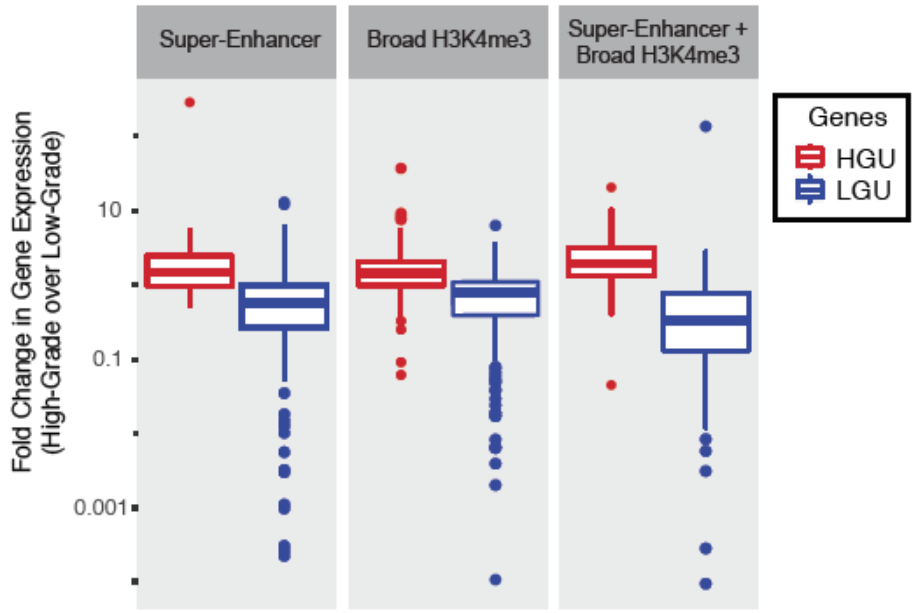
Figure 3.2: Broad domains mark distinctive pathways and are predictive of poorer patient survival



D.


Broad H3K4me3 +
Super-Enhancer
TGF-beta signaling pathway
MicroRNAs in cancer
Cell cycle
Focal adhesion
Proteoglycans in cancer
Gastric cancer
Hippo signaling pathway
Nicotinate and nicotinamide metabolism
Cellular senescence
Circadian rhythm
ECM-receptor interaction
TNF signaling pathway
ErbB signaling pathway
HIF-1 signaling pathway
PI3K-Akt signaling pathway
Rap1 signaling pathway
Phospholipase D signaling pathway
Choline metabolism in cancer
EGFR tyrosine kinase inhibitor resistance
MAPK signaling pathway
VEGF signaling pathway
Longevity regulating pathway - multiple species
Ras signaling pathway
AGE-RAGE signaling pathway in diabetic complications
Tight junction
Transcriptional misregulation in cancer
Estrogen signaling pathway
Common
HGU
LGU
GeneRatio
0.050
0.075
0.100
0.125
p.adjust
0.01
0.02
0.03
0.04

E.


Super-Enhancer
Broad H3K4me3
Super-Enhancer +
Broad H3K4me3
Fold Change in Gene Expression
(High-Grade over Low-Grade)
10
0.1
0.001
Genes
HGU
LGU

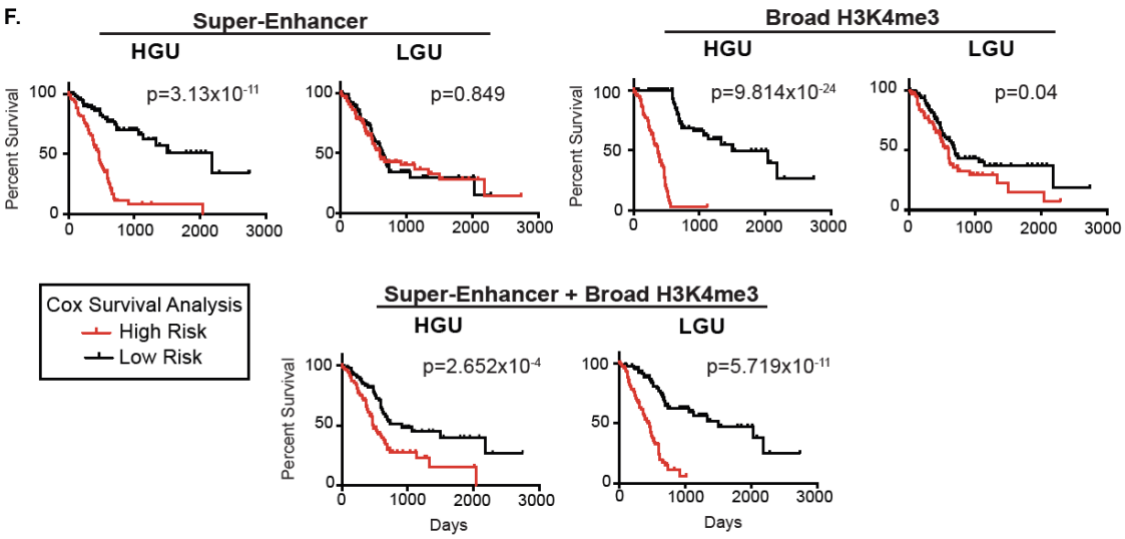
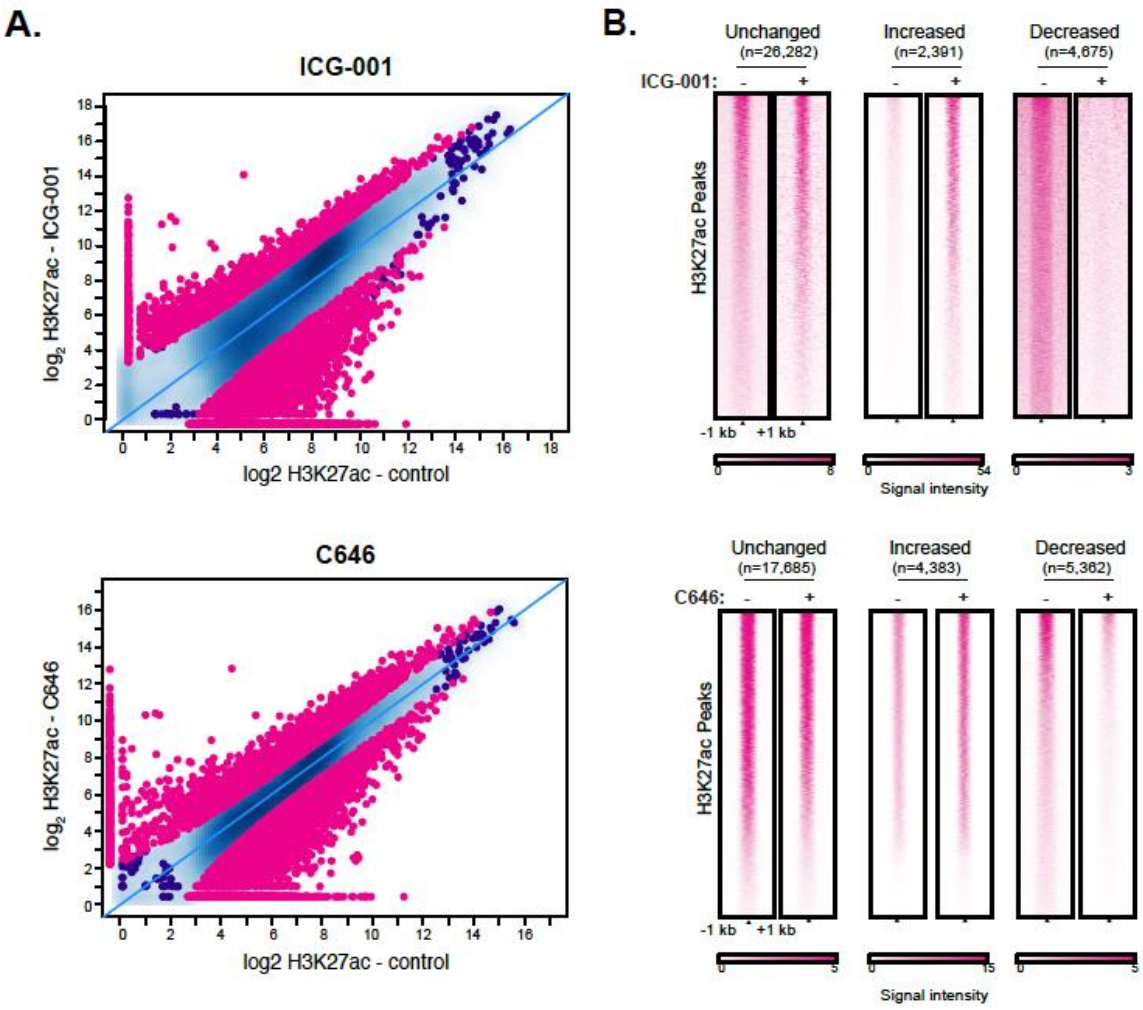
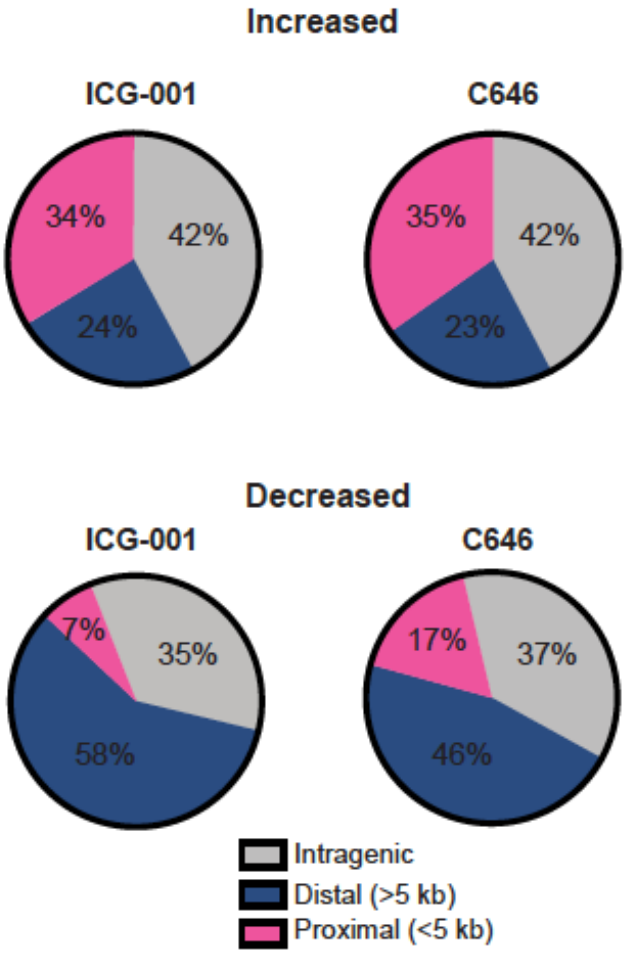

F.
Super-Enhancer
HGU
LGU
Broad H3K4me3
HGU
LGU
Percent Survival
p=3.13x10-11
p=0.849
p=9.814x10-24
p=0.04
Cox Survival Analysis
High Risk
Low Risk
Super-Enhancer + Broad H3K4me3
HGU
LGU
p=2.652x10-4
p=5.719x10-11
Days

Figure 3.2: Broad domains mark distinctive pathways and are predictive of poorer PDAC patient survival. A. Clustering of genomic regions encompassing super-enhancers across seven human PDAC cell lines to define common, High- and Low-Grade unique super-enhancers. B. A similar analysis was performed on broad H3K4me3 domains to define common, High- and Low-Grade unique (HGU and LGU, respectively) broad H3K4me3 domains. C. A similar analysis was performed to look at overlapping domains, where 'B' represents Broad H3K4me3 domains and 'S' represents super-enhancers. D. Gene Ontology pathway enrichment profiles of genes marked by both domains that are common, HGU or LGU domains. E. Comparison of expression levels of genes marked by the indicated domains in High-Grade and Low-Grade PDAC cells. Data is derived from mean normalized expression counts of genes. F. Kaplan-Meier survival analysis of high- and low-risk groups (red and black, respectively) for genes marked High-Grade unique (HGU) and Low-Grade unique (LGU) domains for the indicated domain type.

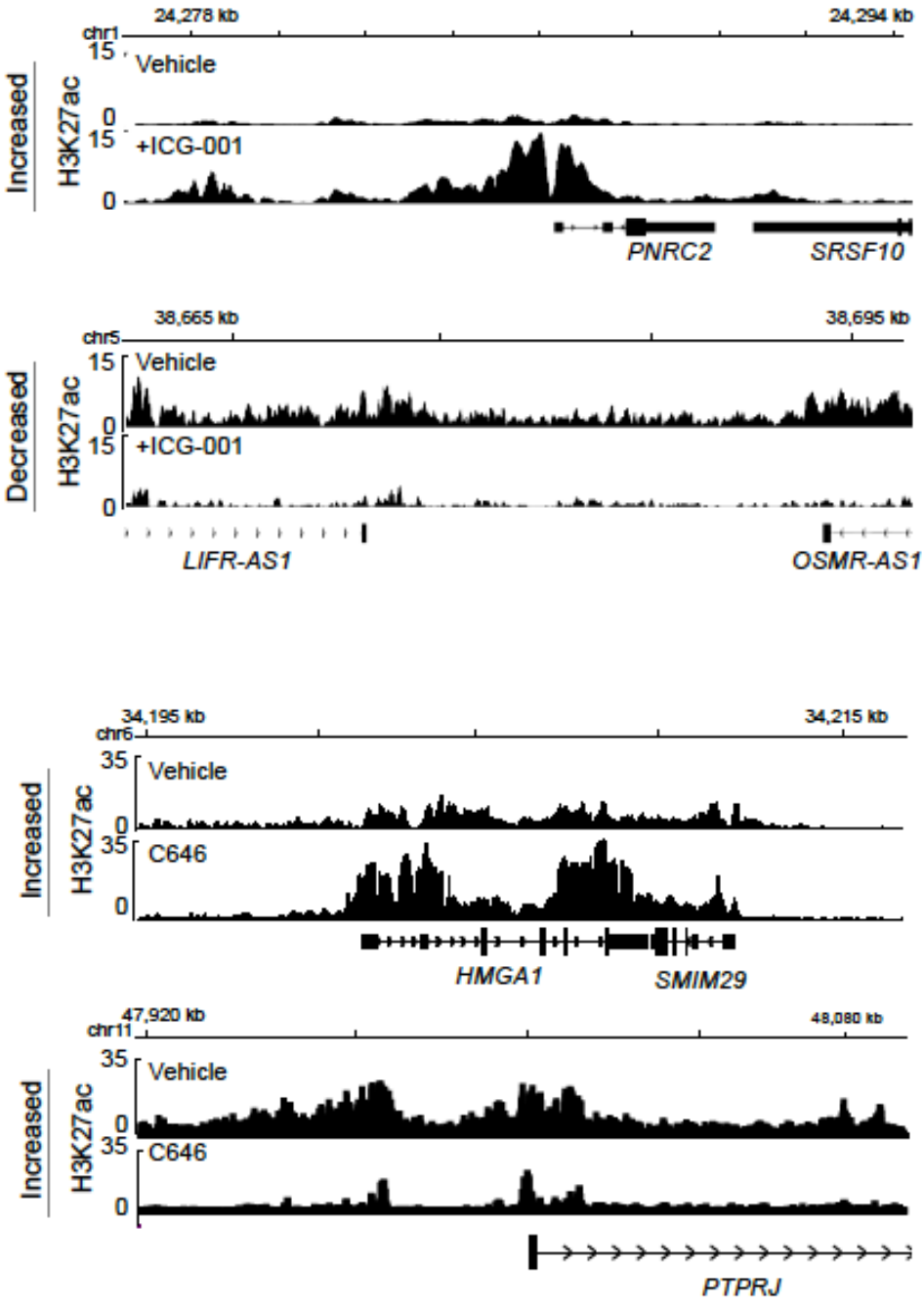
Figure 3.3: Inhibitors of histone acetyltransferases impact global H3K27ac levels



C.


Increased
ICG-001
C646
34%
42%
24%
35%
42%
23%
Decreased
ICG-001
C646
7%
35%
58%
17%
37%
46%
Intragenic
Distal (>5 kb)
Proximal (<5 kb)

D.


24,278 kb
24,294 kb
chr1
Increased
H3K27ac
15
Vehicle
0
15
+ICG-001
0
PNRC2
SRSF10
38,665 kb
38,695 kb
chr5
Decreased
H3K27ac
15
Vehicle
0
15
+ICG-001
0
LIFR-AS1
OSMR-AS1
34,195 kb
34,215 kb
chr6
Increased
H3K27ac
35
Vehicle
0
35
C646
0
HMGA1
SMIM29
47,920 kb
48,080 kb
chr11
Increased
H3K27ac
35
Vehicle
0
35
C646
0
PTPRJ

E.

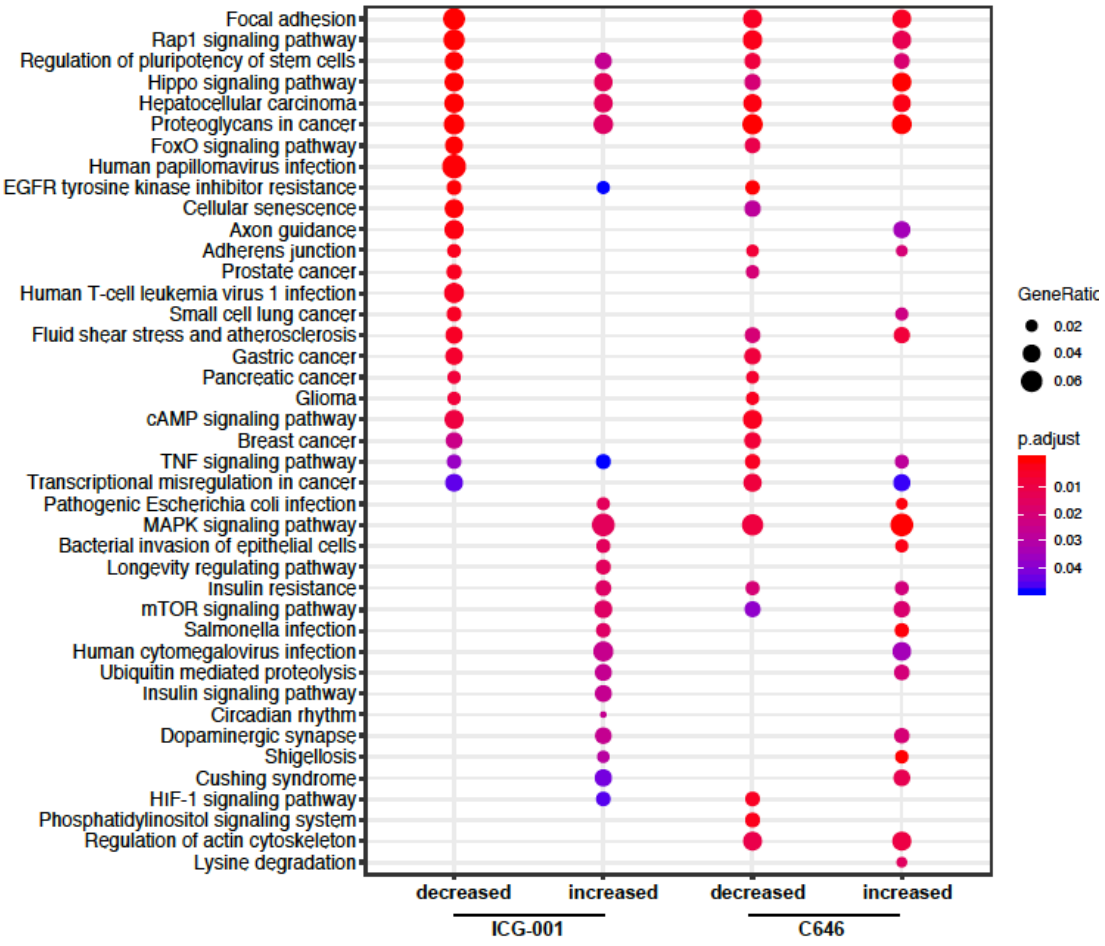
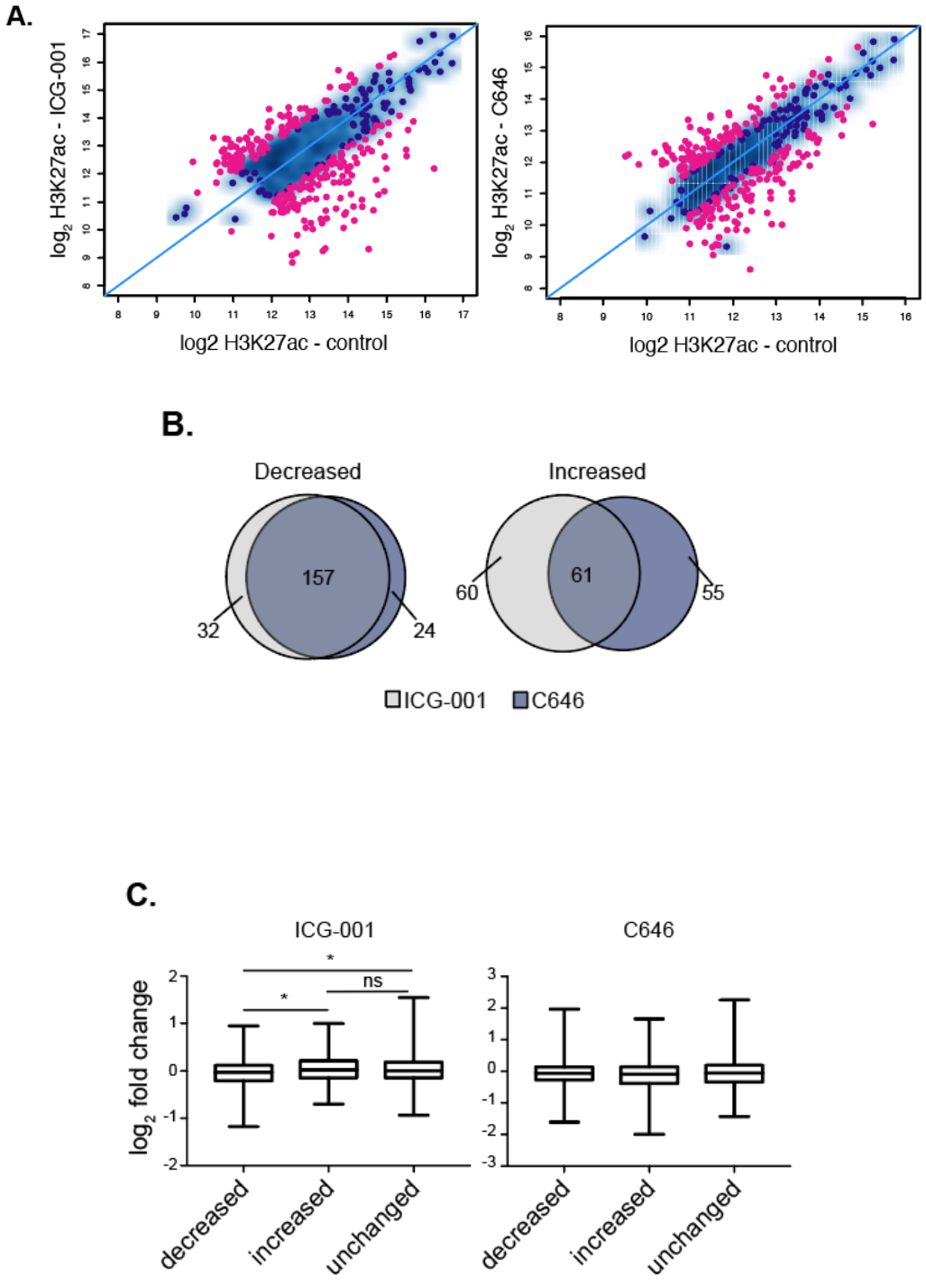

Focal adhesion
Rap1 signaling pathway
Regulation of pluripotency of stem cells
Hippo signaling pathway
Hepatocellular carcinoma
Proteoglycans in cancer
FoxO signaling pathway
Human papillomavirus infection
EGFR tyrosine kinase inhibitor resistance
Cellular senescence
Axon guidance
Adherens junction
Prostate cancer
Human T-cell leukemia virus 1 infection
Small cell lung cancer
Fluid shear stress and atherosclerosis
Gastric cancer
Pancreatic cancer
Glioma
cAMP signaling pathway
Breast cancer
TNF signaling pathway
Transcriptional misregulation in cancer
Pathogenic Escherichia coli infection
MAPK signaling pathway
Bacterial invasion of epithelial cells
Longevity regulating pathway
Insulin resistance
mTOR signaling pathway
Salmonella infection
Human cytomegalovirus infection
Ubiquitin mediated proteolysis
Insulin signaling pathway
Circadian rhythm
Dopaminergic synapse
Shigellosis
Cushing syndrome
HIF-1 signaling pathway
Phosphatidylinositol signaling system
Regulation of actin cytoskeleton
Lysine degradation
decreased
increased
decreased
increased
ICG-001
C646
GeneRatic
0.02
0.04
0.06
p.adjust
0.01
0.02
0.03
0.04

Figure 3.3: Inhibitors of histone acetyltransferases impact global H3K27ac levels. A. Differential H3K27ac enrichment analysis reveals significantly altered genome-wide H3K27ac sites in response to ICG-001 (Top) or C646 (bottom) treatment (FDR <0.1), signal is represented as log2 normalized read count for the indicated condition. B. Signal heatmaps representing the H3K27ac within altered regions after ICG-001 or C646 treatment identified from the differential analysis. C. Location analysis of increased or decreased H3K27ac signal after treatment relative to gene regions. D. Example of genes with increased or decreased H3K27ac signal after ICG-001 treatment, signal is represented as fold enrichment over input. E. KEGG pathway enrichment analysis of genes within altered regions after HAT inhibitor treatment.

Figure 3.4: HAT inhibitors influence the acetylation levels at super-enhancers



D.

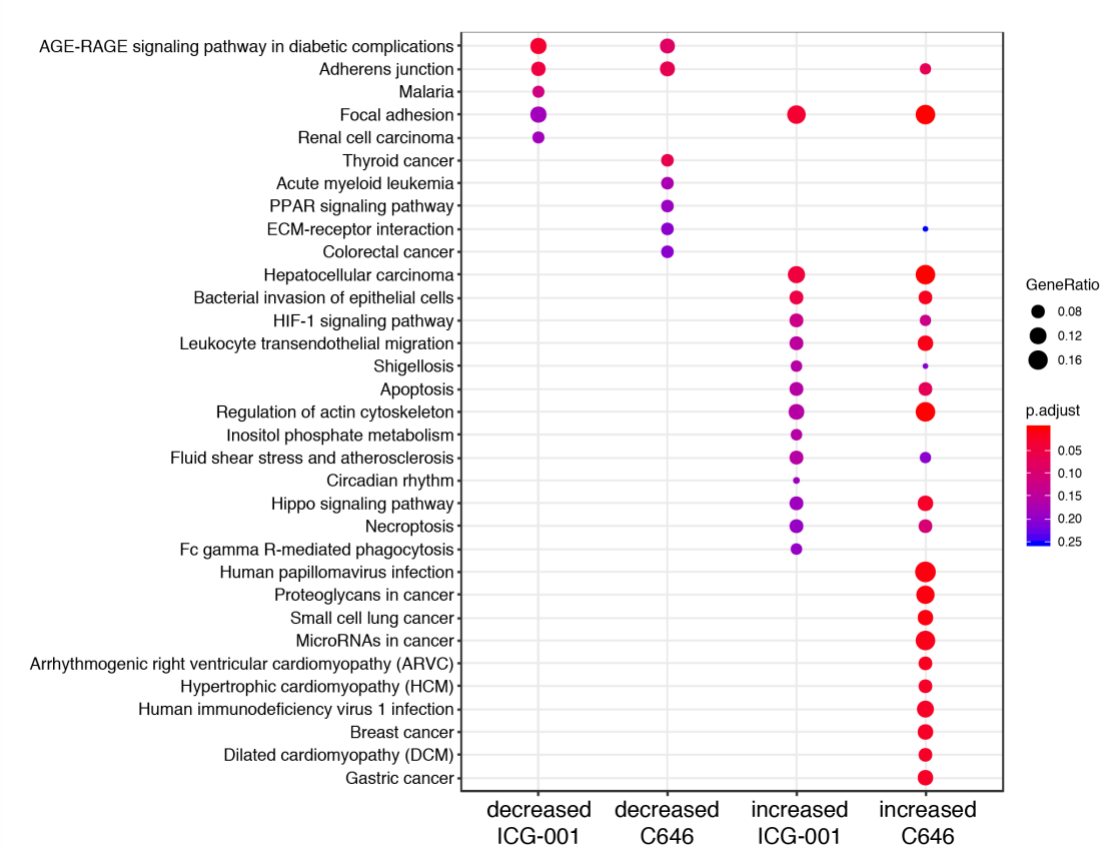
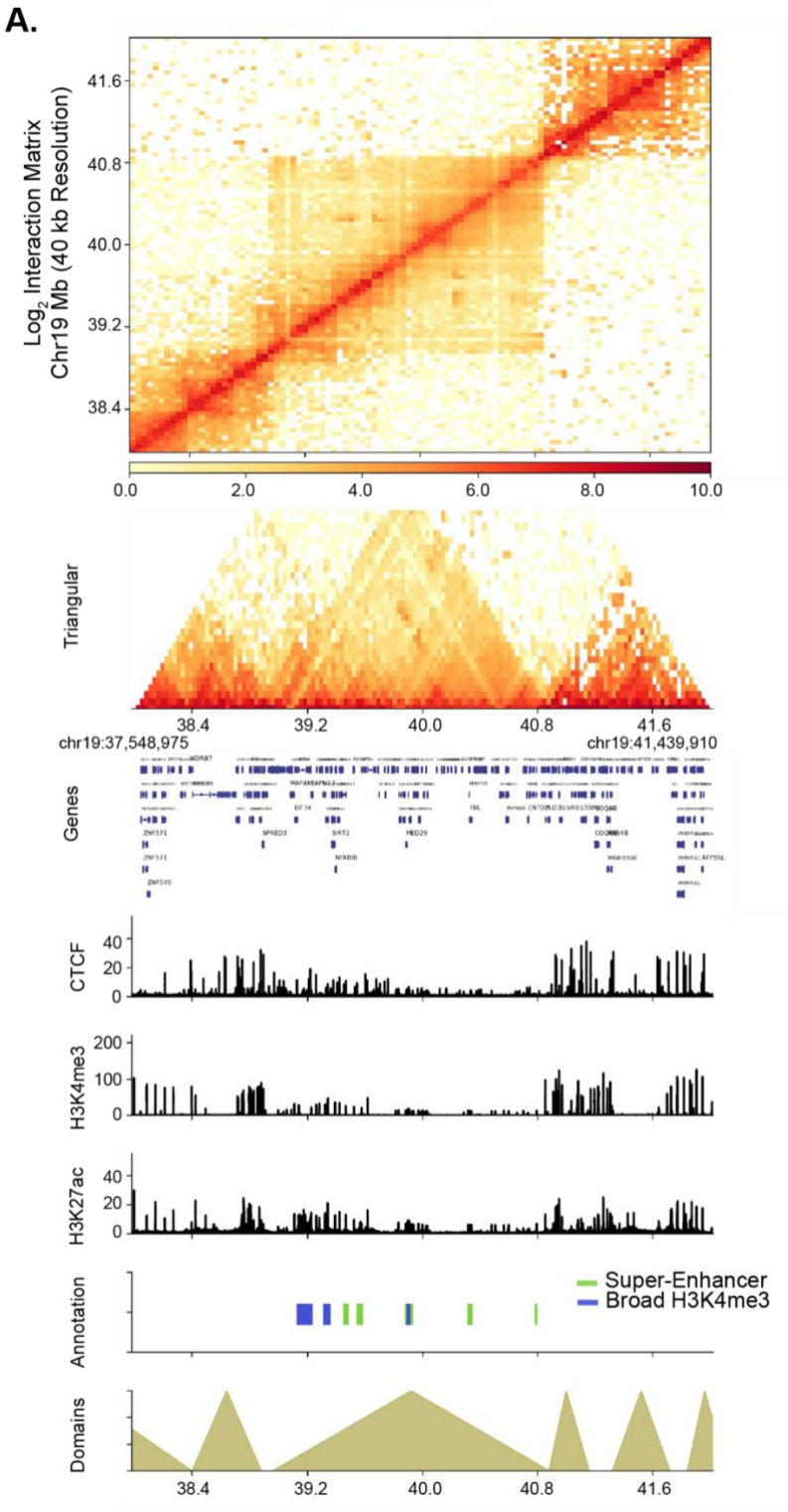

AGE-RAGE signaling pathway in diabetic complications
Adherens junction
Malaria
Focal adhesion
Renal cell carcinoma
Thyroid cancer
Acute myeloid leukemia
PPAR signaling pathway
ECM-receptor interaction
Colorectal cancer
Hepatocellular carcinoma
Bacterial invasion of epithelial cells
HIF-1 signaling pathway
Leukocyte transendothelial migration
Shigellosis
Apoptosis
Regulation of actin cytoskeleton
Inositol phosphate metabolism
Fluid shear stress and atherosclerosis
Circadian rhythm
Hippo signaling pathway
Necroptosis
Fc gamma R-mediated phagocytosis
Human papillomavirus infection
Proteoglycans in cancer
Small cell lung cancer
MicroRNAs in cancer
Arrhythmogenic right ventricular cardiomyopathy (ARVC)
Hypertrophic cardiomyopathy (HCM)
Human immunodeficiency virus 1 infection
Breast cancer
Dilated cardiomyopathy (DCM)
Gastric cancer
decreased ICG-001
decreased C646
increased ICG-001
increased C646
GeneRatio
0.08
0.12
0.16
p.adjust
0.05
0.10
0.15
0.20
0.25

Figure 3.4. HAT inhibitors influence the acetylation levels at super-enhancers.

A. Scatterplot displaying the differential enrichment of H3K27ac (log2 ChIP-seq read count) for ICG-001 (left) and C646 (right) compared to control, the blue diagonal line separates those of increasing or decreasing signal with the colored dots corresponding to regions with significant changes in treatment compared to control (FDR <0.1). B. Overlap analysis of regions comparing the increased or decreased H3K27ac regions after ICG-001 and C646 treatment. C. Boxplots displaying log2 fold change of genes (treatment vs control) within the given differential domains identified in A. D. Pathway enrichment analysis of genes within altered super-enhancers after ICG-001 or C646 treatment.

Figure 3.5: Broad domains are linked to topological associated domain boundaries



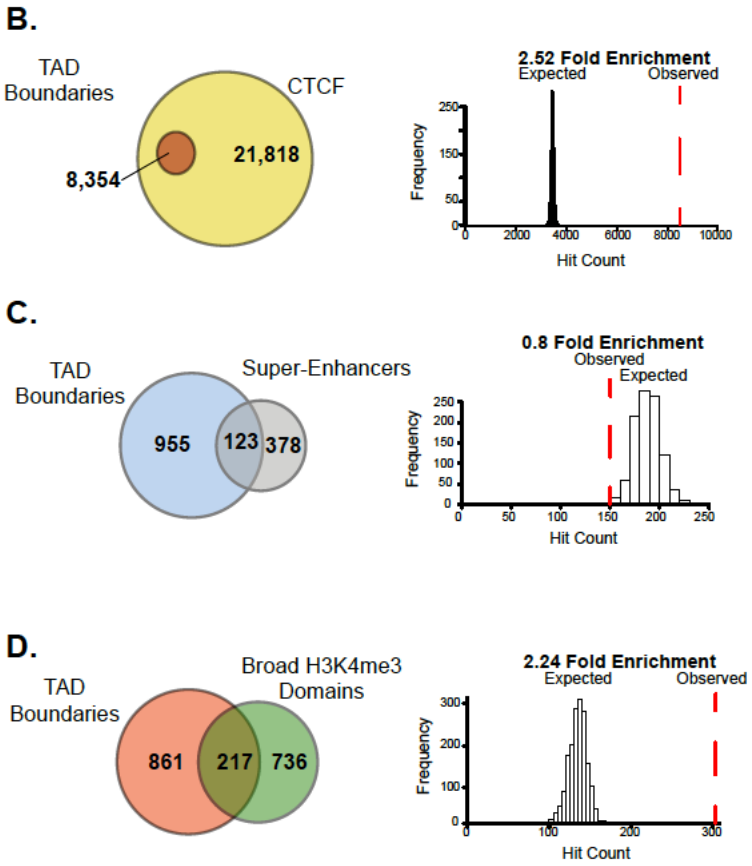

B.
TAD Boundaries
CTCF
8,354
21,818
2.52 Fold Enrichment
Expected
Observed
Frequency
Hit Count
C.
TAD Boundaries
Super-Enhancers
955
123
378
0.8 Fold Enrichment
Observed
Expected
Frequency
Hit Count
D.
TAD Boundaries
Broad H3K4me3 Domains
861
217
736
2.24 Fold Enrichment
Expected
Observed
Frequency
Hit Count

Figure 3.5. Broad domains are linked to topological associated domain boundaries. A. Chromatin interaction matrix at 40 kb resolution showing broad domains contained within TAD regions. Statistical associations of TAD boundaries were performed for B. CTCF, C. super-enhancers and D. Broad H3K4me3 domains.

References

1. Feinberg, A.P., M.A. Koldobskiy, and A. Gondor, Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet, 2016. 17(5): p. 284-99.
2. Heinz, S., et al., The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol, 2015. 16(3): p. 144-54.
3. Hnisz, D., et al., Super-enhancers in the control of cell identity and disease. Cell, 2013. 155(4): p. 934-47.
4. Parker, S.C., et al., Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A, 2013. 110(44): p. 17921-6.
5. Whyte, W.A., et al., Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell, 2013. 153(2): p. 307-19.
6. Adam, R.C., et al., Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature, 2015. 521(7552): p. 366-70.
7. Ding, J., et al., Tex10 Coordinates Epigenetic Control of Super-Enhancer Activity in Pluripotency and Reprogramming. Cell Stem Cell, 2015. 16(6): p. 653-68.
8. Cao, F., et al., Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. Sci Rep, 2017. 7(1): p. 2186.
9. Chen, K., et al., Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet, 2015. 47(10): p. 1149-57.

10. Mack, S.C., et al., Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling. Nature, 2018. 553(7686): p. 101-105.
11. Siegel, R.L., K.D. Miller, and A. Jemal, Cancer Statistics, 2017. CA Cancer J Clin, 2017. 67(1): p. 7-30.
12. Ying, H., et al., Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev, 2016. 30(4): p. 355-85.
13. Jones, S., et al., Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science, 2008. 321(5897): p. 1801-6.
14. Diaferia, G.R., et al., Dissection of transcriptional and cis-regulatory control of differentiation in human pancreatic cancer. EMBO J, 2016. 35(6): p. 595-617.
15. Roe, J.S., et al., Enhancer Reprogramming Promotes Pancreatic Cancer Metastasis. Cell, 2017. 170(5): p. 875-888 e20.
16. Mostoslavsky, R. and N. Bardeesy, Reprogramming Enhancers to Drive Metastasis. Cell, 2017. 170(5): p. 823-825.
17. Sipos, B., et al., A comprehensive characterization of pancreatic ductal carcinoma cell lines: towards the establishment of an in vitro research platform. Virchows Arch, 2003. 442(5): p. 444-52.
18. Deer, E.L., et al., Phenotype and genotype of pancreatic cancer cell lines. Pancreas, 2010. 39(4): p. 425-35.
19. Collisson, E.A., et al., Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med, 2011. 17(4): p. 500-3.

20. Dhayat, S.A., et al., Clinical Impact of Epithelial-to-Mesenchymal Transition Regulating MicroRNAs in Pancreatic Ductal Adenocarcinoma. Cancers (Basel), 2018. 10(9).

21. Lieber, M., et al., Establishment of a continuous tumor-cell line (panc-1) from a human carcinoma of the exocrine pancreas. Int J Cancer, 1975. 15(5): p. 741-7.

22. Yunis, A.A., G.K. Arimura, and D.J. Russin, Human pancreatic carcinoma (MIA PaCa-2) in continuous culture: sensitivity to asparaginase. Int J Cancer, 1977. 19(1): p. 128-35.

23. Kyriazis, A.P., et al., Human pancreatic adenocarcinoma line Capan-1 in tissue culture and the nude mouse: morphologic, biologic, and biochemical characteristics. Am J Pathol, 1982. 106(2): p. 250-60.

24. Kyriazis, A.A., et al., Morphological, biological, biochemical, and karyotypic characteristics of human pancreatic ductal adenocarcinoma Capan-2 in tissue culture and the nude mouse. Cancer Res, 1986. 46(11): p. 5810-5.

25. Schoumacher, R.A., et al., A cystic fibrosis pancreatic adenocarcinoma cell line. Proc Natl Acad Sci U S A, 1990. 87(10): p. 4012-6.

26. Gower, W.R., Jr., et al., HPAC, a new human glucocorticoid-sensitive pancreatic ductal adenocarcinoma cell line. In Vitro Cell Dev Biol Anim, 1994. 30A(3): p. 151-61.

27. Dincer, A., et al., Deciphering H3K4me3 broad domains associated with gene-regulatory networks and conserved epigenomic landscapes in the human brain. Transl Psychiatry, 2015. 5: p. e679.

28. Gerstein, M.B., et al., Architecture of the human regulatory network derived from ENCODE data. Nature, 2012. 489(7414): p. 91-100.

29. Frietze, S., et al., Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. Genome Biol, 2012. 13(9): p. R52.

30. Benayoun, B.A., et al., H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell, 2014. 158(3): p. 673-88.

31. Ucar, D. and D. Bayarsaihan, Cell-specific gene promoters are marked by broader spans of H3K4me3 and are associated with robust gene expression patterns. Epigenomics, 2015. 7(2): p. 129-31.

32. Khan, A., A. Mathelier, and X. Zhang, Super-enhancers are transcriptionally more active and cell type-specific than stretch enhancers. Epigenetics, 2018. 13(9): p. 910-922.

33. Cancer Genome Atlas Research Network. Electronic address, a.a.d.h.e. and N. Cancer Genome Atlas Research, Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell, 2017. 32(2): p. 185-203 e13.

34. Herbertz, S., et al., Clinical development of galunisertib (LY2157299 monohydrate), a small molecule inhibitor of transforming growth factor-beta signaling pathway. Drug Des Devel Ther, 2015. 9: p. 4479-99.

35. Abulwerdi, F., et al., A novel small-molecule inhibitor of mcl-1 blocks pancreatic cancer growth in vitro and in vivo. Mol Cancer Ther, 2014. 13(3): p. 565-75.

36. Gaddis, M., et al., Altering cancer transcriptomes using epigenomic inhibitors. Epigenetics Chromatin, 2015. 8: p. 9.

37. Wu, A.H., et al., MiR-572 prompted cell proliferation of human ovarian cancer cells by suppressing PPP2R2C expression. Biomed Pharmacother, 2016. 77: p. 92-7.

38. Huang, J., et al., Dissecting super-enhancer hierarchy based on chromatin interactions. Nat Commun, 2018. 9(1): p. 943.

39. Gerrard, D.L., et al., Three-dimensional analysis reveals altered chromatin interaction by enhancer inhibitors harbors TCF7L2-regulated cancer gene signature. J Cell Biochem, 2019. 120(3): p. 3056-3070.

40. Jones, P.A., J.P. Issa, and S. Baylin, Targeting the cancer epigenome for therapy. Nat Rev Genet, 2016. 17(10): p. 630-41.

41. Morera, L., M. Lubbert, and M. Jung, Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. Clin Epigenetics, 2016. 8: p. 57.

42. Bowers, E.M., et al., Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. Chem Biol, 2010. 17(5): p. 471-82.

43. Eguchi, M., et al., ICG-001, a novel small molecule regulator of TCF/beta-catenin transcription. Med Chem, 2005. 1(5): p. 467-72.

44. Emami, K.H., et al., A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. Proc Natl Acad Sci U S A, 2004. 101(34): p. 12682-7.

45. Krumlauf, R., Hox genes in vertebrate development. Cell, 1994. 78(2): p. 191-201.

46. Korkut, A., et al., A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF-beta Superfamily. Cell Syst, 2018. 7(4): p. 422-437 e7.

47. Riggins, G.J., et al., Frequency of Smad gene mutations in human cancers. Cancer Res, 1997. 57(13): p. 2578-80.

48. Murthy, D., K.S. Attri, and P.K. Singh, Phosphoinositide 3-Kinase Signaling Pathway in Pancreatic Ductal Adenocarcinoma Progression, Pathogenesis, and Therapeutics. Front Physiol, 2018. 9: p. 335.

49. Rozengurt, E., J. Sinnett-Smith, and G. Eibl, Yes-associated protein (YAP) in pancreatic cancer: at the epicenter of a targetable signaling network associated with patient survival. Signal Transduct Target Ther, 2018. 3: p. 11.

50. Jiang, H., et al., Targeting focal adhesion kinase renders pancreatic cancers responsive to checkpoint immunotherapy. Nat Med, 2016. 22(8): p. 851-60.

51. Sunami, Y., A. Rebelo, and J. Kleeff, Lipid Metabolism and Lipid Droplets in Pancreatic Cancer and Stellate Cells. Cancers (Basel), 2017. 10(1).

52. Di Gangi, I.M., et al., Metabolomic profile in pancreatic cancer patients: a consensus-based approach to identify highly discriminating metabolites. Oncotarget, 2016. 7(5): p. 5815-29.

53. Fang, F., et al., Discrimination of metabolic profiles of pancreatic cancer from chronic pancreatitis by high-resolution magic angle spinning 1H nuclear magnetic resonance and principal components analysis. Cancer Sci, 2007. 98(11): p. 1678-82.

54. Principe, D.R., et al., TGFbeta Signaling in the Pancreatic Tumor Microenvironment Promotes Fibrosis and Immune Evasion to Facilitate Tumorigenesis. Cancer Res, 2016. 76(9): p. 2525-39.

55. Heintzman, N.D., et al., Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet, 2007. 39(3): p. 311-8.

56. O'Geen, H., S. Frietze, and P.J. Farnham, Using ChIP-seq technology to identify targets of zinc finger transcription factors. Methods Mol Biol, 2010. 649: p. 437-55.

57. Kalhor, R., et al., Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat Biotechnol, 2011. 30(1): p. 90-8.

58. Langmead, B. and S.L. Salzberg, Fast gapped-read alignment with Bowtie 2. Nat Methods, 2012. 9(4): p. 357-9.

59. Zhang, Y., et al., Model-based analysis of ChIP-Seq (MACS). Genome Biol, 2008. 9(9): p. R137.

60. Loven, J., et al., Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell, 2013. 153(2): p. 320-34.

61. Ramirez, F., et al., deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res, 2016. 44(W1): p. W160-5.

62. Heinz, S., et al., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell, 2010. 38(4): p. 576-89.

63. Kent, W.J., et al., BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics, 2010. 26(17): p. 2204-7.

64. Robinson, J.T., et al., Integrative genomics viewer. Nat Biotechnol, 2011. 29(1): p. 24-6.

65. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform, 2013. 14(2): p. 178-92.

66. Boyd J (2018). seqsetvis: Set Based Visualizations for Next-Gen Sequencing Data. R package version 1.2.0.

67. Stark R and Brown G (2011). DiffBind: differential binding analysis of ChIP-seq peak datasets.

68. Ross-Innes, C.S., et al., Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature, 2012. 481(7381): p. 389-93.

69. McLean, C.Y., et al., GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol, 2010. 28(5): p. 495-501.

70. Hiller, M., et al., Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: application to zebrafish. Nucleic Acids Res, 2013. 41(15): p. e151.

71. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 2014. 15(12): p. 550.

72. Servant, N., et al., HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol, 2015. 16: p. 259.

73. Sexton, T., et al., Three-dimensional folding and functional organization principles of the Drosophila genome. Cell, 2012. 148(3): p. 458-72.

74. Shin, H., et al., TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res, 2016. 44(7): p. e70.

75. Akdemir, K.C. and L. Chin, HiCPlotter integrates genomic data with interaction matrices. Genome Biol, 2015. 16: p. 198.

CHAPTER 4

Temporal dynamic reorganization of 3D chromatin architecture in hormone-induced breast cancer and endocrine resistance

Abstract

Recent studies have demonstrated that chromatin architecture is linked to the progression of cancers. However, the roles of 3D structure and its dynamics in hormone-dependent breast cancer and endocrine resistance are largely unknown. Here we report the dynamics of 3D chromatin structure across a time course of estradiol (E2) stimulation in human estrogen receptor α (ERα) positive breast cancer cells. We identified subsets of temporally highly dynamic compartments predominantly associated with active open chromatin and found that these highly dynamic compartments showed higher alteration in tamoxifen resistant breast cancer cells. Remarkably, these compartments are characterized by active chromatin states, and enhanced ERα binding but decreased transcription factor CCCTC-binding factor (CTCF) binding. Lastly, we identified a set of ERα-bound promoter-enhancer looping genes enclosed within altered domains that are enriched with cancer invasion, aggressiveness or metabolism signaling pathways. This large-scale analysis expands our understanding of higher order temporal chromatin reorganization underlying hormone-dependent breast cancer.

Introduction

Numerous efforts have been devoted to revealing the basic principle of three dimensional (3D) chromatin architecture and genome organization inside the cell nucleus of various mammalian genomes[1-8]. One prominent structural feature of genome organization is the formation of various types of chromosomal domains[9] defined as spatial compartments[1,10], topologically association domains (TAD)[3] or lamina-associated domains (LAD)[11]. The discrete TADs ranging from several hundreds of kilobases (kb) to several megabases (Mb) are usually stable in diverse cell types and are highly conserved across different mammalian species, suggesting that they are inherent and important functional units of mammalian genomes[12,13]. By contrast, spatial compartments comprised of two types, compartment A or B, form an alternating pattern of active and inactive domains along chromosomes. Their sizes usually range around 5 Mb and are characterized by genomic features associated with transcriptional activity, such as chromatin accessibility, active or repressive histone marks, gene density, GC content and repetitive regions[14,15]. Furthermore, A and B compartments show tissue- or cell-type specificity that are correlated with cell-type specific gene expression patterns[16,17]. However, a recent study found A or B compartments may be much smaller in size at a couple of hundred kb when using improved Hi-C protocols in higher resolution maps[18]. These maps are similar in size to the topologically constrained domains[19]. It is also increasingly recognized that spatial compartments and TADs are fundamentally two independent chromosomal organization modes[20,21], thus disputing the common notion of a hierarchical folding principle that TADs are the building blocks of larger compartment domains.

Recent efforts have focused on understanding the relationship between higher-order structures and human development and disease.[22-24]. For instance, new studies demonstrated that the reprogramming of high-order structures of both the paternal and maternal genomes gradually occurs during early mammalian development[23,24]. Another study showed that disorganization of prostate cancer 3D genome architecture occurs coincident with long-range epigenetically activated or silenced regions, which coordinated with gene transcription[25]. Despite the advances in our knowledge of 3D genome regulation, several critical questions remain to be answered in the field. For example: (1) how stable or dynamic are chromosome domains upon signaling stimuli as cells respond to external cues? (2) To what extent do these changes affect establishing or re-establishing the compartmentalized architecture? (3) What degree of impact do the master or key transcription factors in a particular cell system have on chromatin reorganization? (4) What are the roles of chromatin architecture in governing the progression of human diseases, such as cancers?

Estrogen (E2) signaling plays a crucial role in driving estrogen receptor α positive (ERα+) breast cancer cell growth and proliferation[26,27]. The cellular response to E2 induction is characterized by timed and coordinated transcriptional regulation primarily mediated by ERα. Thus, it has been frequently used as a model system to illustrate the mechanisms underlying transcriptional controls in cancer development and progression as well as in fundamental biological process[28-32]. Using genome-wide approaches, we and others demonstrated very little coincidence between ERα targeted genes in breast cancer cells versus acquired endocrine resistant breast cancer cells indicating distinct transcriptional regulatory mechanisms underlying endocrine resistance[33-37]. In a recent

study, we used a 3C-based high throughput protocol to identify two densely mapped distant estrogen response element (DERE) regions which were frequently amplified in ERα+ breast cancer[38,39]. Interestingly, these aberrantly amplified DEREs deregulated target gene expression linked to cancer development and tamoxifen resistance. However, the roles of 3D structure and its dynamics in hormone-dependent breast cancer and endocrine resistance are largely unknown.

To establish a basis for data-driven learning and modeling of the temporal dynamics and 3D chromatin reorganization, we applied tethered chromatin conformation (TCC), a modified Hi-C protocol[40] for high depth sequencing. We performed TCC in a time-series of E2-induction in the ERα+ breast cancer cell line, MCF7 as well as the tamoxifen resistant MCF7 (TamR) cell line. Here, we present a time-series of genome-wide maps of chromatin contacts, identify the temporal dynamic patterns of chromatin compartments, compare the patterns between MCF7 and TamR cells and examine the enrichment of ERα and CTCF binding as well as five active and repressive histone marks in the patterns. We further identify ERα-bound promoter-enhancer (ERα-PE) looping genes enclosed within TamR altered dynamic compartments. This large 3D-scale of chromatin data provides a rich resource for studying the basic characteristics of hormone-dependent breast cancers and provides further insight into the mechanisms of tamoxifen resistance.

Results

Re-compartmentalization of chromatin at early E2 treatment

Despite many studies demonstrating that E2 induces the highest levels of ERα binding and gene activity around 45min to 1hr[28,30], little has been done to comprehensively characterize the changes of chromatin architecture in MCF7 cells in a genome-wide manner. In this study, we conducted TCC analysis in hormone-starved MCF7 cells (T0) and compared this to MCF7 cells treated with 1hr of E2 (T1). The Pearson correlation of chromatin interactions showed that biological replicates are largely correlated for each treatment at different resolutions (**Figure 4.1-A**), illustrating good quality TCC data. We thus combined two replicates at either time point to identify chromosome compartments with HiCLib[10] (**Figure 4.1-B**). With a sequencing depth of around 200 million (75 million uniquely mapped) paired-end reads for each data set, we expect the resolution of compartments around 40-50kb[41]. Surprisingly, the genomic size of a majority of compartments is either smaller than 1Mb or between 1-2Mb, and very few are larger than 5Mb (**Figure 4.1-C and D**). Our data seems to contradict earlier studies that determine compartment sizes to be larger than 2Mb[1,10]; however, it is consistent with newly reported studies using improved Hi-C protocols in conjuction with a higher sequencing depth[18]. We found a similar number of compartments, 2,067 and 2,039, in the untreated and E2-treated cells respectively, where approximately half consisted of compartment A (active chromosome domains) and half were compartment B (inactive chromosome domains) at each time point (**Figure 4.1-E**). We then compared the compartments between untreated and E2-treated cells, and found that the type of compartments drastically changes following E2 treatment. The number of common or

conserved compartments between treatments increases from 28%, 55%, 74% to 78% using different bin sizes of 100kb, 250kb, 500kb and 1Mb, respectively **(Figure 4.1-F**). Our results at lower resolution (500kb or 1Mb) are in-line with many other studies that report approximately 80% conserved domains among different conditions or cell types[3,12]. However, at a higher resolution of 100kb, we identified 576 Common compartments between untreated and E2-treated cells and 1,463 Transit compartments sensitive to E2 treatment that shifted in size or flipped between A/B compartments. Of Transit compartments, 743 (51%) of them shift only 100kb and 247 (17%) shift more than 400kb while 116 (8%) flipped between A/B compartments **(Figure 4.1-G)**. These data demonstrate a re-compartmentalization of higher-order chromatin domains following 1 hour of E2 treatment.

Temporal dynamic chromatin along prolonged E2 treatments

Several studies found that transcriptional responses to longer E2 treatments were dramatically different than responses to shorter treatments[30,31]. In order to understand the dynamics of chromatin structure in longer E2 treatment periods in MCF7 cells, we further conducted TCC analysis in three more time points, 4hr (T4), 16hr (T16) and 24hr (T24), each with biological replicates. In order to capture the dynamic patterns following a prolonged E2 treatment, we determined compartments at 100kb and compared the compartments among the five time points. As expected, the number of compartments was very similar among the five time points. We used the Common and Transit compartments obtained from the comparison of T1 *vs.* T0 and compared them to T4, T16 and T24 respectively. We also separated the comparison of active open chromatin compartment A

(**Figure 4.2-A**, upper panel) from inactive closed chromatin compartment B (**Figure 4.2-A**, lower panel). When re-examining these sets of Common or Transit compartments, we identified 15 patterns of changed compartments from 16 sets (**Appendix D-1**) and 9 additional patterns from the last set (labeled X in **Appendix D-1**) based on the converted bins. We were able to categorize these 24 patterns of chromatin into six types of temporally dynamic re-compartmentalization (TDRC): Highly Common Compartments (HCC, patterns 1-4 with a FDR of 0.268), Early Transit Compartments (ETC, patterns 5-8 with a FDR of 0.230), Late Transit Compartments (LTC, patterns 9-12 with a FDR of 0.192), Lowly Dynamic Compartments (LDC, patterns 13-16 with a FDR of 0.178), Moderately Dynamic Compartments (MDC, patterns 17-20 with a FDR of 0.161), and Highly Dynamic Compartments (HDC, patterns 21-24 with a FDR of 0.201) (**Figure 4.2-B**). There is also a statistically significant difference ($p=7.6x10^{-11}$, Wilcoxon rank-sum test) between highly dynamic compartments (MDC and HDC) and lowly dynamic compartments (HCC, ETC, LTC and LDC) (**Figure 4.2-C**). Interestingly, the MDC and HDC are predominantly composed of compartment A mainly from Transit to Transit compartments along the time courses of E2 treatment (**Figure 4.2-D** and **E**), while the LDC has the most changed compartment B. This data suggests that active chromatin domains are more susceptible to change in response to E2 stimulation over time.

Altered chromatin compartmentalization in resistant cells

Increasing evidence suggests that ERα-mediated gene deregulation or epigenetic alterations may be key mechanisms underlying acquired tamoxifen resistant breast cancer[42,43]. However, the knowledge of endocrine resistant associated 3D regulation is

still limited. To delineate the altered 3D architectures, we further conducted TCC analysis in a tamoxifen resistant MCF7 cell line, MCF7-TamR[44]. At a resolution of 100kb, we identified 2,103 compartments including both A and B (**Figure 4.3-A**). We further defined a compartment to be a tamoxifen resistant altered compartment (TRAC) if it was a Transit compartment and if there was at least one converted bin between TDRC and a TamR compartment. As such, we obtained six corresponding types of TRACs: TA-HCC (FDR of 0.154), TA-ETC (FDR of 0.250), TA-LTC (FDR of 0.154), TA-LDC (FDR of 0.165), TA-MDC (FDR of 0.139) and TA-HDC (FDR of 0.226). Patterns 17, 19, 21, 23, and 24 in TA-MDC and TA-HDC types showed higher alteration than other patterns, suggesting that the higher dynamics of the compartments in E2-induced MCF7 cells, the stronger alterations of the compartments in TamR cells (**Figure 4.3-B**). Further, we observed that the average size of TA-HDC and TA-ETC types were longer than those in unaltered compartments, while the size of TamR unaltered compartments (TRUCs) were longer than TA-HCC (**Figure 4.3-C**). Att a bird's-eye-view, we identified three interesting types of TRACs: Shrunk, Expanded and Flipped (**Figure 4.3-D-F**). Our data suggest that a group of genes within the same domain may be concordantly regulated during acquired tamoxifen resistance.

Epigenetic states in dynamic re-compartmentalization

Epigenetic marks have been shown to classify genomic compartments and chromosomal domains into subcompartments or subdomains in diverse cell types[18]. There is little known about the structural roles of one-dimensional (1D) epigenetic states in E2-induced 3D chromatin structure. We performed ChIP-seq of three active marks,

H3K27ac, H3K4me1, and H3K4me3, and two repressive marks, H3K9me3 and H3K27me3 in a time course of E2 treatment in MCF7 cells as well as in asynchronous TamR cells, each with biological replicates. We first trained a total of 30 histone modification data by ChromHMM at various parameters[45], and obtained ten HMM states (**Figure 4.4-A** and **B**). Through interpreting both HMM emission and transition probabilities, we inferred seven biologically meaningful epigenetic states, including two active states (S9 and S10), two bivalent states (S3 and S8) and three repressive states (S1, S4 and S6) (**Figure 4.4-C**). We were also able to map these states back into each of six types of E2-induced dynamic compartments and TamR altered compartments. Overall, we observed that more active states were distributed in compartment A and more repressive states in compartment B, while three types of dynamic changed compartments, *i.e.*, LDC, MDC and HDC, have a higher percentage of active states than HCC does (**Figure 4.4-D-G**). Surprisingly, S1 showed a high percentage of distribution in compartment A despite that it is a repressive state.

ERα and CTCF binding in dynamic re-compartmentalization

Since ERα is a master transcription factor mainly in response to E2 stimulation in MCF7 cells and CTCF is a chromatin organizer known to regulate the 3D architecture, we wanted to understand their regulatory roles in mediating these 3D structural dynamics. We performed ChIP-seq of ERα and CTCF at five time points of E2-stimulated MCF7 cells and in TamR cells, each with biological replicates. We used MACS[46] to call ERα binding sites (peaks) in each of the 12 data sets and obtained 7,553 peaks in untreated MCF7 cells (T0), and between 13,000-20,000 ERα peaks in E2-treated

MCF7 cells at four time points and untreated TamR cells respectively (**Figure 4.5-A**). Interestingly, we found Patterns 18 and 20 of MDC and Pattern 24 of HDC had the highest number of ERα binding sites in compartment A but not in compartment B in E2-induced MCF7 cells as well as in TamR cells (**Figure 4.5-B**), illustrating that there are more ERα binding sites in higher dynamic active chromatin.

Furthermore, we identified approximately 50,000 CTCF peaks in each of the five time points of E2-induced MCF7 cells and in TamR cells (**Figure 4.5-C**). On the switched domain boundary between two compartments, we observed generally lower averages of CTCF binding sites in three types of TDRCs (LDC, MDC, HDC), and three types of TRACs (TA-LDC, TA-MDC, TA-HDC) than other three types (**Figure 4.5-D**). When testing the correlation of ERα binding within the compartments *vs* CTCF binding on the boundary regions, Pattern 24 of HDC or TA-HDC was the only pattern having more than 80% ERα peaks and less than 0.4 CTCF peaks per compartment in E2-induced MCF7 cells (**Figure 4.5-E**—left panel) or in TamR cells (**Figure 4.5-E**—right panel). Collectively, our results indicate a reciprocal relationship to ERα binding and CTCF binding at highly dynamic changed compartments during the temporal response to E2 stimulation which is also observed in cells with acquired tamoxifen resistance.

Differentially expressed genes and putative loops

We next examined the gene expression and loops within these TDRCs and TRACs. We utilized a publicly available time-series RNA-seq dataset[47], profiled at 10 time points of E2-treated MCF7 cells, to identify differentially expressed genes. By picking 5 time points close to this study (T0 = 0 min, T1 = 40 min, T4 = 160 min, T16 =

640 min, T24 = 1280 min), we identified 4,106 dynamic differentially expressed genes (DDEGs) across all time points. As expected, a majority of such genes were located in dynamically changed compartment A with most in the MDC and HDC types (**Figure 4.6-A**). There seems to be minimal difference in the average of gene expression levels within each of the six types of TDRCs. However, the variance of DDEGs in LDC, MDC and HDC is higher than HCC, ETC and LTC (**Figure 4.6-B**). GSEA analysis[48] showed ribosome, tight junction, endocytosis, lysosome, cell cycle, WNT signaling pathway, insulin signaling pathway, focal adhesion, and MAPK signaling pathway were among the top functional categories for the 1,396 DDEGs in MDC and HDC types (**Figure 4.6-C**).

We further performed RNA-seq in parental MCF7 and TamR cells, each with three biological replicates and identified a total of 2,097 TamR-specific differentially expressed genes (TDEGs). More than half of them (1,188) were in the combined TA-M&HDC types (**Figure 4.6-D**). We then identified 42,390 TamR-specific significant interaction pairs or putative loops from TamR TCC data by HOMER[49] and using our T0 TCC dataset as the contrast. 3C-qPCR validations further confirmed the differential looping intensity of seven randomly selected pairs between parental MCF7 and TamR cells (**Appendix D-2**). Of the 42,390 identified loops, 16,807 were overlapped with a promoter (-5kb/+1kb around the 5'TSS), 9,638 of them had either H3K27ac or H3K4me1 peaks in the distal loci, 4,122 of them had at least one ERα binding site at either loci of the loop and were thus considered as ERα regulated promoter-enhancer (ERα-PE) loops (**Figure 4.6-E**). Finally, 396 TDEGs within 599 ERα-PE loops in the combined TA-M&HDC types were identified as ERα dysregulated dynamic looping genes in resistant cells. Functional annotation and gene pathway analysis with GSEA showed these genes

were mapped to nine KEGG signaling pathways related to cancer invasion and aggressiveness, as well as glycolysis and metabolism (**Figure 4.6-F**). Taken together, our results demonstrate that these ERα-associated dynamically reorganized active domains regulating gene looping events may result in higher susceptibility to alterations in tamoxifen resistant cells. This prompts us to speculate that these genome domains and looping genes may be responsible for driving the acquired tamoxifen resistance.

Discussion

Despite the increasing developments of various 3C-derived high throughput sequencing techniques in which it advances our understanding of the principles of 3D genome architecture, several important questions remain to be answered in the field. One of the many aspects is to elucidate how stable or dynamic chromosome domains are in response to signaling stimuli and to what extent these changes affect establishing or re-establishing the compartmentalized architecture. Our main goal of this study is to establish a basis for data-driven modeling of temporal dynamics and 3D chromatin reorganization given that such studies are very limited. While mega-sized TADs are conserved among different cell types and mammalian species[3,5,12], 100-500kb size of subTADs or compartments are considered to be dynamic where the boundaries are non-conserved[18,50]. Though our data showed the total number of compartment domains are quite similar among different time points, the changes in size (of at least 100kb) of compartments are very pronounced, particularly in these E2-induced highly dynamic compartments. With a very loose definition of dynamic changes requiring a minimum of 100kb, we were able to unveil 24 temporal dynamic patterns upon E2-induction which

were further categorized into six major types. Indeed, the MDCs and HDCs were not only predominately of active compartment A but also contained dynamic differentially expressed genes enriched with biological process terms ribosome, tight junction, cell cycle and others (**Figure 4.6-C**). Many of them characterize known effects of estrogen on the MCF7 cell phenotype[51]. In contrast, there were no significant differences between E2-induced early and late changed compartments in which both types were comprised of very few compartments. Our data implied that these moderate to high dynamic compartments may play an essential role in governing hormone-mediated luminal breast cancer development.

Many studies including ours have demonstrated that E2 instructed dynamic transcriptional programs rewired or altered transcription regulatory networks in tamoxifen resistant breast cancer cells[36,52,53]; however, very few focused on examining 3D regulatory roles in tamoxifen resistance. Our previous studies utilized a 3C-seq technique to identify two densely mapped DERE regions located on chromosomes 17q23 and 20q13 frequently amplified in MCF7 cells and found their aberrantly amplified DEREs deregulated target genes were potentially linked to cancer development and tamoxifen resistance[38]. However, this study containd many limitations including the technique itself, smaller data volumes and fewer computational tools available for a thorough analysis. Our current work has significantly improved in the following aspects: 1) we generated high quality TCC datasets with increased sequencing depth allowing for the detection at 40kb resolution; 2) we produced TCC datasets in an E2-induced time series in MCF7 cells and then compared it to TamR cells; and 3) we utilized many state-of-art computational tools to process TCC and ChIP-seq datasets. Remarkably, our

integrative analyses uncovered many temporal dynamic patterns characterizing the 3D chromatin reorganization upon E2-induction. Interestingly, two types of temporal dynamic re-compartmentalization (TDRC), *i.e.*, moderately and highly dynamic compartments (MDCs and HDCs), showed higher alteration in TamR cells (**Figure 4.3-B**). Furthermore, the looping gene signatures enclosed in these two altered dynamic domains were highly enriched with GO terms cancer invasion and aggressiveness or metabolism. All of these biological processes captured the nature of acquired resistant breast cancer cells[54].

Our definition of 24 patterns or six types of dynamic changed compartments were based upon how re-compartmentalization in MCF7 cells respond to E2 induction. Interestingly, the resulted six types are identical to the analysis based on a mathematical calculation resulting in a total of 256 combinations of E2-induced time-dependent compartments (**Appendix D-3**) when using T0 as a contrast. Furthermore, we observed the same trends of E2-induced time-dependent compartments in both MCF7 and T47D cell lines, where the major trend are miscellaneous dynamic compartments in both MCF7 and T47D cell lines. Moreover, the altered compartments of both MCF7-TamR and T47D-TamR have higher percentage of miscellaneous compartments. Our results suggest that our analytical strategy and observations are generalizable in various cell lines.

Our findings further illustrated an anti-correlative trend of binding enrichments between intradomain ERα sites and boundary CTCF sites (**Figure 4.4-D** and **E**). Interestingly, the average of CTCF sites is generally lower regardless of its distance from the boundary in these highly dynamic changed domains. Although our result is not so surprising, it nevertheless implicates an underlying molecular event that ERα regulated

high re-compartmentalization may be required to loosen CTCF insulator at the domain boundary[4,55]. Furthermore, our data may support the higher-order chromatin architectural role of ERα signaling in mediating hormonal activity, expanding our current understanding of the molecular mechanism underlying the E2-induced breast carcinogenesis through ERα regulation.

Collectively, our observations suggest a possible molecular mechanistic model (**Figure 4.7-A-D**). A constitutive estrogen stimulation in breast cancer cells enhances stronger ERα activity and further recruits its distal regulatory machinery including different co-regulators, mediators, cohesions and chromatin remodelers, and then mobilizes highly dynamic gene looping events which essentially expand to render a 3D genome re-compartmentalization accompanying with lower CTCF binding at the compartment boundary. In contrast, in breast cancer cells with the acquired resistance, increased crosstalk between ERα and other signal transduction pathways such as EGFR/HER2[56], IGF-IR[53], and AKT/PTEN[57] or altered expression of some key co-regulators particularly reshuffle these highly dynamic gene looping resulting in altered chromatin reorganization. To substantiate this model, we propose further functional or mechanistic experiments in our follow up studies: 1) establishing genome-edited TamR sublines by editing out a handful ERα sites selected from 599 gene loops using the CRISPR/Cas9 technique[58]; in these sublines, examining the changes of ERα regulated looping using ChIP-3C-qPCR and determining if resistant cells are re-sensitized; and 2) establishing an *in vivo* model of TamR xenografts[44]; in this model, examining selected ERα regulated loops using ChIP-3C-qPCR by comparing untreated vs. treated Gefitinib,

an inhibitor to EGFR/HER2. With these results, we might expect to elucidate the detailed 3D ERα regulatory role in mediating tamoxifen resistance.

In summary, the high quality and large 3D chromatin data along with many ChIP-seq and RNA-seq data provided a comprehensive resource for understanding how estrogen exposure drives genome-wide 3D chromatin reorganization in ERα positive breast cancer cells as well as how their alterations occur in hormone resistant cells. Our integrative analysis reveals temporal dynamic patterning and 3D chromatin reorganization of the breast cancer genome that occurs in response to E2 stimulation over time. Our work may give further insight into the effective treatment strategies to overcome tamoxifen resistance and discovery of novel epigenetic therapeutic targets.

Methods

Cell lines and reagents

The human parental MCF7, T47D and Tamoxifen Resistant (TamR) cell lines were derived from Osborne et al. 1994[44]. MCF7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine and 1% penicillin/streptomycin (pen/strep) until 90% confluent. For temporal estrogen responsiveness, MCF7 cells were hormone-starved for 72hrs followed by the addition of 100nM β-Estradiol (MP Biomedicals, Inc.) at 1hr, 4hr, 16hr and 24hr. To hormone starve MCF7 or T47D cells, these cells were grown to 80% confluency as described above. Once the desired confluency was reached, the cells were washed one time with Phosphate Buffered Saline and the media was replaced with phenol-red free DMEM supplemented with 5% charcoal-stripped FBS, 2mM (L-glutamine) and 1%

(pen/strep). For the 0hr time point, cells were immediately crosslinked following 72hrs of hormone starvation.

TamR cells were cultured in phenol red free DMEM supplemented with 10% charcoal-stripped FBS, 2mM L-glutamine, 1% pen/strep, and 100nM Tamoxifen (Sigma-Aldrich). Tamoxifen was replenished every 48hrs and cells were crosslinked at 90% confluency.

Tethered chromatin capture (TCC)

TCC was performed as Kalhor et al. described[40]. Approximately 50 million MCF7 or T47D cells (either parental hormone starved with temporal addition of β-Estradiol (E2) or TamR cells as described above) were crosslinked with 1% formaldehyde for 10 minutes at room temperature, crosslinking was quenched with 0.125 M glycine for 5 minutes at room temperature and cell pellets were collected and stored at -80ºC. The crosslinked cells were lysed with 550µL of Cell Lysis Buffer (10mM HEPES pH 8.0, 10mM NaCl, 0.2% Igepal CA-630, containing 1X protease inhibitor cocktail (Thermo Scientific #88665) and 1X PMSF (Acros Organics #215740050). The cells were homogenized with a dounce homogenizer for 20 strokes with pestle A after incubation on ice for 15 minutes. The lysate was then centrifuged at 2,500 rcf for 5 minutes at room temperature. The supernatant was discarded and the remaining pellet containing the nuclei was washed twice with ice-cold wash buffer #1 (50mM Tris-HCl pH8, 50mM NaCl, 1mM EDTA) and resuspended in 250µL of wash buffer #1. The chromatin was then solubilized by the addition of 95µL of 2% SDS followed by an incubation at 65ºC for 10 minutes. The cysteine residues were biotinylated by the addition of 105µL of

25mM EZlink Iodoacetyl-PEG2-Biotin (IPB) (Thermo Scientific #21334) and incubated at room temperature for 1 hour while rocking. The SDS was neutralized with 1300µL of 1X NEBuffer 2 (NEB #B7002S) on ice for 5 minutes, followed by the addition of 225µL of 10% Triton X-100, which was then incubated on ice for 10 minutes followed by a final incubation at 37ºC for 10 minutes. DNA was digested overnight at 37ºC with the following reagents: 100µL of 10X NEBuffer 2, 5µL of 1M DTT, 430µL of water, and 2000U of HindIII (100U μL^{-1}; NEB #R0104M). To remove remaining IPB, the samples were then dialyzed for 4 hours at room temperature using a Slide-A-Lyzer Dialysis Cassette with a 20kD cutoff (Thermo Scientific #87735) in 1L of dialysis buffer (10mM Tris-HCl, pH 8.0 and 1mM EDTA). The biotinylated chromatin was then tethered using 400µL of MyOne Streptavidin T1 beads (Invitrogen #65601) after washing the beads three times with PBST (1X PBS containing 0.01% Tween20) and re-suspending in 2mL of PBST. 400µL of washed Streptavidin T1 beads was then added into each of five equal aliquots of dialyzed sample. Binding occurred at room temperature for 30 minutes followed by the addition of 150µL of 25mM IPB neutralized with 25mM 2-mercaptoethanol, which was then incubated at room temperature for 15 minutes. Non-crosslinked DNA and non-biotinylated chromatin was removed by washing the beads once with 600µL PBST followed by one wash with 600µL wash buffer #2 (10mM Tris-HCl, pH 8.0, 50mM NaCl, 0.4% Triton X-100). To wash the beads, we utilized a magnetic rack and ensured beads bound to the magnet before aspirating the buffer out. The beads were then resuspended in 100µL of wash buffer #2. The 5’ overhangs were filled with 63µL water, 1µL 1M MgCl2, 10µL 10X NEBuffer 2, 0.7µL 10mM dATP (NEB #N0440S), 0.7µL 10mM dTTP (NEB #N0443S), 0.7µL 10mM 2’-

Deoxyguanosine-5'-O(1-thiotriphosphate) sodium salt (dGTPαS) (Biolog Life Science Institute #D031-05), 15µL of 0.4mM Biotin-14-dCTP (Invitrogen #19518-018), 4µL of 10% Triton X-100, and 25U Klenow-large fragment (NEB #M0210L) and rocked at room temperature for 40 minutes. 5µL of 0.5M EDTA was added to stop the reaction and the beads were washed twice with wash buffer #3 (50mM Tris-HCl pH7.4, 0.4% Triton X-100, 0.1mM EDTA) and resuspended in 500µL of wash buffer #3. The crosslinks were reversed with 400µL of extraction buffer (50mM Tris-HCl pH8, 0.2% SDS, 1mM EDTA, 100mM NaCl) followed by the addition of 400µg of proteinase K (NEB #P8107S) and incubation for two hours at 65°C. The initial conformation capture library (the supernatant) was extracted twice with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1 v/v) and once with an equal volume of chloroform. NaCl was then added to a final concentration of 20mM and glycogen to 2ug μL^{-1} followed by precipitation of the DNA with the addition of 900µL of ethanol (200 proof) and incubation at -20°C overnight. The DNA was pelleted via centrifugation at 20,000 rcf at 4°C for 20 minutes. The pellet was then immersed in 500µL of 80% ethanol and centrifuged at 20,000 rcf for 10 minutes. The ethanol was removed and the pellet was air dried until approximately 90% dry and resuspended in 20µL of 10mM Tris-HCl pH8. The five aliquots were combined and the RNA was removed via RNAseA digestion (10µg RNAseA) for 30 minutes at 37°C. The DNA was purified using the Invitrogen Purelink Quick PCR purification kit (Invitrogen #K310001). Biotin from non-ligated DNA was removed from 5µg of purified DNA using 300U EXOIII (NEB #M0206S), adjusting the total volume to 90µL with 10X NEBuffer 1 (NEB #B7001S). This reaction was incubated at 37°C for 1 hour. The reaction was stopped with 2µL of 0.5M EDTA and

2µL of 5M NaCl followed by incubation at 70°C for 20 minutes. A Covaris Focused-ultrasonicator (Covaris S220) was used to shear the DNA, with a duty factor of 5%, peak power of 175W, and 200 cycles per burst. Each sample was sonicated for 180 seconds and purified using the Purelink Quick PCR purification kit and eluted in 50µL of elution buffer. Libraries were generated with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB #E7645L). First, end-repair was performed after sonication. 1µg of DNA was used and the total volume of sample was brought up to 50µL with 0.1X TE. The end repair was carried out as outlined in the manufacturer's protocol. After end repair, the biotinylated DNA was pulled down using 10µL of MyOne Streptavidin C1 beads (Invitrogen #65001). The beads were first washed twice with 500µL of 1X Binding and Wash buffer (for 2X Binding and Wash buffer: 10mM Tris-HCl pH7.5, 1mM EDTA, and 2M NaCl) and resuspended in 2X Binding and Wash buffer, which was then added to the end-repaired DNA. Lo-bind tubes (Eppendorf #022431021) were used to prevent sticking of beads to the sides of the tubes. The samples were rocked for 30 minutes at room temperature. The beads were washed one time with 1X Binding and Wash buffer containing 0.1% Triton-X100 followed by one wash with 10mM Tris-HCl, pH8 and the beads were collected in 60µL of 10mM Tris-HCl, pH8. Next, adaptor ligation was performed as described in the NEBNext Ultra II DNA Library Prep protocol, following the instructions for 1µg of input DNA. After ligation of Illumina adaptors on the beads, the beads were washed twice with 1X Binding and Wash buffer and twice with 0.1X TE. The beads were resuspended in 30µL of 10mM Tris-HCl pH8. 15µL of the beads containing adaptor-ligated DNA was transferred to a new tube and we continued on to PCR enrichment of adaptor-ligated DNA on the beads. The remaining 15µL was saved

and stored it at 4ºC. The PCR was carried out as outlined in the NEBNext Ultra II protocol, NEBNext Multiplex oligos for Illumina (NEB #E7335S and #E7500S) were used for the individual barcodes and the enrichment was performed using 10 cycles. 45µL of the supernatant containing the PCR products was transferred to a new tube and were cleaned using a 0.8X bead cleanup of the PCR reaction with Agencourt AMPure XP beads (Beckman Coulter #A63881). The AMPure XP beads were brought to room temperature and resuspended. 36µL of the resuspended beads were added to the libraries and mixed by pipetting. The beads were incubated at room temperature for 5 minutes and the supernatant was discarded. The beads were washed twice while on the magnet with 200µL of 80% 200 proof ethanol. After air-drying the beads, the library was eluted off the beads with 23µL of 0.1X TE and transferred to a new tube. The final library was quantified using a Qubit fluorometer (Applied Biosystems) and analyzed using a Bioanalyzer (Agilent Technologies).

Chromatin immunoprecipitation sequencing (ChIP-seq)

The antibodies used for ChIP-seq were: H3K27ac (Abcam, Cambridge, MA, USA; Ab4729 lot #GR238071-1), H3K27me3 (Abcam, Cambridge, MA, USA; Ab6002 lot #GR137554-5), H3K4me3 (Abcam, Cambridge, MA, USA; Ab8580 lot #GR240214-1), H3K4me1 (Abcam, Cambridge, MA, USA; Ab8895 lot #GR114265-2), H3K9me3 (Abcam, Cambridge, MA, USA; Ab8898 lot #GR216368-1), ERα (Santa Cruz Biotechnology, Santa Cruz, CA, USA; sc-543X lot #J0313) and CTCF (Cell Signaling Technology, Danvers, MA, USA; D31H2 lot#1). We performed duplicate ChIP-seq experiments for each histone or factor using chromatin collected on different cell culture

dates. For each histone ChIP-seq assay, 10µg of chromatin was incubated with (2.5-5µg) of antibody. 150µg of chromatin was used for CTCF ChIP-seq (with 20µL of antibody) and 250µg of chromatin was used for ERα ChIP-seq (with 12 µg of antibody). ChIP-seq samples were prepared as O'Geen et al. described[59] with minor adjustments. The cells were crosslinked as described above for TCC experiments. The crosslinked cell pellets were washed twice with ice-cold 1X PBS and stored at -80°C until sonication. Crosslinked cell pellets were thawed on ice and resuspended in 1mL ice-cold cell lysis buffer (5mM PIPES pH8, 85mM KCl, Igepal 10µL mL^{-1}) containing 1X protease inhibitor cocktail and 1X PMSF. After incubation on ice for 15 minutes the samples were then homogenized using a 2mL dounce homogenizer fitted with pestle 'B', using 20 strokes. The samples were then centrifuged at 430 rcf for 5 minutes at 4°C. The supernatant was removed and the pelleted nuclei were lysed with 1mL ice-cold nuclei lysis buffer (50mM Tris-HCl pH8.1, 10mM EDTA, 1% SDS) containing protease inhibitors (1X protease inhibitor cocktail and 1X PMSF). The nuclei were lysed while incubating on ice for 30 minutes. Sonication was performed for 12 minutes using a Covaris Focused-ultrasonicator (Covaris S220) with a peak power of 140W, duty factor of 10%, and 200 cycles per burst. The sonicated material was then centrifuged at 20,000 rcf for 15 minutes at 4°C and transferred to a new tube. To quantify the chromatin, 20µL of the sonicated chromatin was added to 80µL of ChIP elution buffer (50mM $NaHCO_3$ and 1%SDS) followed by the addition of 12µL of 5M NaCl. The samples were boiled at 97°C for 15 minutes and 10µg of RNAseA was added to the tubes once the sample was cooled to room temperature. The sample was incubated at 37°C for 10 minutes to allow for RNA digestion. The reverse-crosslinked chromatin was then purified using the

Purelink Quick PCR purification kit and eluted in 20µL of nuclease free water. After quantification via a nanodrop (Thermo Scientific) the total chromatin yield in our sonicated material was calculated. To visualize the fragment sizes of the sonicated chromatin, we ran 1µg of purified chromatin on a 1.5% agarose gel. If the chromatin fragments were concentrated around the 300-500bp range, we continued onto immunoprecipitation. If under-sonicated, additional sonication was performed as needed. 500ng of purified chromatin sample was saved as our input samples, these samples were brought to a total volume of 150µL with ChIP elution buffer and stored at -20°C. ChIP for each target was carried out using the quantities of chromatin and antibody above-mentioned. The chromatin for each target was diluted with 5 times the volume of ice-cold 1X IP dilution buffer (50mM Tris pH7.4, 150mM NaCl, 1% Igepal (v/v), 0.25% Deoxycholic acid, 1mM EDTA pH 8.0) containing protease inhibitors. The appropriate amount of antibody for each reaction was added and rotated overnight at 4°C. The antibody/chromatin complexes were captured by the addition of 150µL of protein A/G beads (Pierce #88803), which were first washed twice with 1X IP dilution buffer, for the transcription factor ChIPs and 15µL of protein A/G beads for the histone ChIPs. These complexes were rotated at 4°C for 2 hours. Following incubation, the beads were captured using a magnetic rack and washed twice with IP wash buffer #1 (50mM Tris-HCl pH 7.4, 150mM NaCl, 1% Igepal (v/v) 0.25% Deoxycholic acid and 1mM EDTA, pH8). The beads were resuspended in the wash buffer for each wash and the supernatant was removed between each wash. The beads were washed three times with IP wash buffer #2 (100mM Tris-Cl pH9, 500mM LiCl, 1% Igepal, and 1% Deoxycholic acid). The beads were transferred to a new tube on the third wash. The complexes were then

eluted off of the beads by the addition of 75µL of ChIP elution buffer while vortexing at room-temperature for 30 minutes. The supernatant was transferred to a new tube and the elution step was repeated. The ChIP input samples were thawed and 20µL of 5M NaCl was added to the 150µL of final eluted complexes and to the input samples. Crosslinks were reversed overnight at 65ºC and the ChIPs were purified using the Purelink Quick PCR purification kit and the samples were eluted in 35µL of elution buffer. We performed qPCR against targets enriched for each of the ChIPs. The ChIPs were diluted 1:5 and the input samples were diluted to 1ng μL^{-1}. 2µL of DNA was used for each PCR and 1ng was used for the input sample. Primers against GAPDH were positive for CTCF and H3K4me3; STX16 for CTCF; GREB1 for H3K4me1 and H3K27ac; TFF1 for H3K4me1, H3K4me3 and H3K27ac; ZNF180 and ZNF333 for H3K9me3; HOXB2 for H3K27me3; and HES3 for H3K27me3. ZNF333 and ZNF180 were negative targets for CTCF, ER-α, H3K4me1, H3K4me3, H3K27ac, and H3K27me3. TFF1 and SHISA5 were used as negative targets for H3K9me3.

ChIP-seq libraries were generated using the NEBNext ChIP-seq Library Prep Master Mix Set for Illumina (NEB#E6240L) as per manufacturer's protocol with size selection for the insert size of 300bp. Half of adaptor-ligated DNA was saved at 4ºC before PCR enrichment of adaptor ligated DNA. PCR enrichment was done using 10 cycles and cleaned with AMPure XP beads at 0.9X as outlined in the protocol. The final library was eluted off of the beads using 30µL of 0.1X TE and the quality was analyzed with a bioanalyzer (Agilent Technologies).

Chromosome conformation capture coupled with qPCR (3C-qPCR)

Experiments using 3C-qPCR experiment were conducted as Hagège et al. described[60]. Ten million cells (MCF7 or TamR) were harvested and then fixed with 1% formaldehyde for 10 min at room temperature followed by 0.2M glycine to quench the reaction. Cells were lysed with 0.2% Igepal CA630 for 1hr on ice, then the pelleted nuclei were solubilized with 0.3% sodium dodecyl sulfate (SDS) for 1hr at 37°C and diluted with 2% Triton X-100 for 1 h at 37°C. The genomic DNA was digested with 400 U HindIII overnight at 37 °C and the digestion was stopped with 1.6% SDS for 20 min at 65°C. The digested nuclei were diluted with 1:1 volume of ligation buffer and then ligated with 100U T4 DNA ligase. The ligated DNA was de-crosslinked with 300µg proteinase K overnight at 65°C and purified by phenol–chloroform extraction. The 3C template was dissolved in 10mM Tris-HCl and analyzed with the quantitative PCR.

Parental and TamR RNA sequencing (RNA-seq) sample preparation and processing

Total RNA were extracted using the ZYMO Research Quick-RNA MiniPrep kit. Ten million MCF7 or MCF7-TamR cells were lysed in RNA Lysis Buffer followed by removing the majority of gDNA with a Spin-Away Filter. The mixture of RNA and ethanol were then loaded onto Zymo-Spin IIICG Column. Trace DNA was removed by DNase I on the column followed by washing twice with RNA Wash Buffer. The total RNA was eluted with 50 µl DNase/RNase-Free Water. RNA-seq libraries were prepared with Illumina TruSeq stranded mRNA kit. 4µg of total RNA of either parental MCF7 or MCF7-TamR cells was incubated with RNA purification beads and then washed with bead washing buffer. The mRNA was eluted with elution buffer and then reverse

transcribed with Superscript III reverse transcriptase. The first strand cDNA was synthesized with first strand synthesis act D mix and the second strand cDNA was synthesized with second strand marking master mix. After cDNA was synthesized, a single adenylate was added to the 3' end with A-tailing mix and adapters were ligated with ligation mix. DNA fragments were enriched with PCR master mix and then purified to build the DNA library. The library was sequenced with Illumina HiSeq 2000. The differentially expressed genes were identified with CuffDiff[61]. The 50bp single end sequencing reads were aligned with the Tophat module, and then transcripts were assembled with Cufflinks. The transcript assemblies were compared to annotation using the Cuffcompare and two or more transcript assemblies were merged with Cuffmerge. Lastly, the differentially expressed genes and transcripts were found with Cuffdiff.

Identification of compartment patterns and types

All TCC data were analyzed with HiCLib python package[10] to identify chromatin compartment A or B. Paired-end reads of TCC data were iteratively aligned to the human reference genome (hg19) by bowtie2[62] with the minimal sequencing length of 20bp and the length step of 5bp in the module of HiCLib mapping. The following reads were removed from the dataset in the HiCLib HiCdataset object: beginning with the 5bp range from the restriction enzyme cut site; the duplicate molecules; the fragment pairs separated by less than 2 restriction sites within the same chromosome; extremely large restriction fragments (more than 10,000bp) and extremely small restriction fragments (smaller than 100 bp); both ends of the pairs starting at exactly the same positions; the top 0.5% most frequently identified restriction fragments. At this stage the self-circles, dangling ends

and PCR duplicate reads were removed and a maximum molecule length of 500bp was specified at the initiation of the object. The correlation of two replicates was computed as the following: the counts of mapped reads pairs were accumulated at the 1Mb bin, and then the correlation of these counts for each chromosome were calculated separately. After filtering the reads, the frequency contact matrices were constructed at a bin size of 100kb with the HiCLib fragmentHiC module. The contacts between loci located within the same bin were then removed from the raw heatmap. The bins with less than half of a bin sequenced and the 1% of regions with low coverage were also removed. The top 0.05% of interchromosomal counts as the possible PCR blowouts were truncated followed by iterative correction to get the ICE heatmap using the HiCLib binnedData module. All bins of the ICE heatmap on a diagonal were removed with the HiCLib binnedData module. The bins with less than half of a bin sequenced were also removed. All cis contacts were set to zero to obtain only the trans contacts. The cis contacts were forged in an interative way. After removing the bins with zero counts the eigenvector expansion was performed with the HiCLib binnedData module to get the first eigenvectors of compartments. The continuous genomic regions of positive first eigenvectors were defined as compartment A (active chromatin), and the continuous genomic regions of negative first eigenvectors were defined as compartment B (inactive chromatin) individually at the 100kb scale. The compartments of five time points (T0, T1, T4, T16, T24) were compared to identify the dynamic patterns (**Appendix D-1**). First, two kinds of compartments: T0 vs. T1 Common and T0 vs. T1 Transit were identified by comparing compartments of T0/T1. The Common compartments are the overlapping compartments and the Transit compartments are differential compartments,

which will be used in the following steps as well. Next, T0 vs. T1 Common and T0 vs. T1 transit were compared with T4, T16, T24 independently to generate the (a) T0 vs. T1 Common vs. T4/T16/T24 Common, (b) T0 vs. T1 Common vs T4/T16/T24 Transit, (c) T0 vs. T1 transit vs T4/T16/T24 Common, (d) T0 vs. T1 transit vs T4/T16/T24 Transit. Thirdly the pattern 1-15 were produced by comparing the various time points (T4, T16 and T24) of last step subsets a, b, c, and d, which are vs. T4, vs. T16, vs. T24. The rest subsets were divided into pattern 16-24 according to the numbers of converted bins. Finally, 24 patterns were identified from the intersection and difference among subsets a, b, c, and d. According to their biological meanings, these patterns were able to categorize into six types of dynamic changed compartments (DCCs): HCC (patterns 1-4), ETC (patterns 5-8), LTC (patterns 9-12), LDC (patterns 13-16), MDC (patterns 17-20), HDC (patterns 21-24).

Computation of differential compartments

The variance of first eigenvector values of compartments identified in T0, T1, T4, T16, T24 was computed at the 100kb scale. The difference of HCC/ETC/LTC/LDC with MDC/HDC was determined by two-sided Wilcoxon rank-sum test for their averaged variance of first eigenvector values. The estimate of False Discover Rate (FDR) of the differential compartment between any two compartments was conducted by a permutation-based test. In brief, the difference in Means of Eigenvector Values of each of two compartments: Compartment 1 and Compartment 2, was first calculated as the Observed Value. The Eigenvector Values of two compartments were then pooled together and randomly selected one half as randomized Compartment 1 and the other half

as randomized Compartment 2. This was done for a total of 1,000 rounds of permutation. In each round, the difference in means of Eigenvector Values of each of two randomized compartments was calculated as a Permutated Value. All permutated values were combined into a null distribution. The FDR is estimated based on how many permutated values are above the observed value and the permutated null.

Epigenetic states

ChIP-seq of H3K4me3, H3K27ac, H3K4me1, H3K27me3, and H3K9me3 data sets from five time points of E2-induced MCF7 cells and TamR cells were aligned to the human reference genome hg19. We then utilized the Java program ChromHMM v1.17[45] to characterize chromatin epigenetic states by integrating the histone modification ChIP-seq datasets to identify de novo major re-occuring combinatorial and spatial patterns based on a multivariate Hidden Markov Model. The results of the model were then used to systematically annotate genome-wide maps of chromatin states. After the ChIP-seq data were mapped to the human genome, the BinarizeBam module was used to binarize uniquely mapped reads into 1kb bins for model learning. The binarized data were then trained with LearnModel and ten epigenetic states were finally identified with a minimum *p*-value after averaging five training rounds. The emission and transition matrices were visualized using R. The ten ChromHMM states were classified into three kinds of epigenetic states according to the combination of histone marks with the *p*-value cutoff of 0.3 for emission matrices. The active states were defined by H3K4me3, H3K27ac and H3K4me1, repressive states defined by H3K27me3, H3K9me3 and bivalent states including both active and repressive states.

Identification of ERα and CTCF binding sites

ChIP-seq data for ERα and CTCF were aligned to the human reference genome, hg19. We then determined binding sites by peaks calling with MACS v1.4.2[46]. The identified peaks were coordinated to the various compartments and boundary regions to obtain the coincident sites. The peak summits generated by MACS were defined as binding sites for the subsequent analyses.

Differential binding analysis of ERα

We identified ERα differential peaks (DPs) with the R Bioconductor package DiffBind v2.6.6[63,64] using TamR vs. T0/T1/T4/T16/T24 as the contrast. Within the anlaysis, peaks were first enriched for genomic loci from ChIP-seq data and then read by DiffBind. Next, overlaps of peaks were examined to determine how well similar samples cluster together with the function dba.count. Then, overlapped reads in each interval for each unique sample were counted with the function dba.contrast. Lastly, a contrast was established and then the core analysis of DiffBind was executed by default using DESeq2[65] with the function dba.analyze. Finally, the results were reported and plotted with the function dba.report.

Time-series RNA-seq data analysis

Time-series RNA-seq data of E2-treated MCF7 cells[47] were acquired from GSE62789. Five time points close to this study (T0 = 0 min, T1 = 40 min, T4 = 160 min, T16 = 640 min, T24 = 1280 min) were selected and mapped to the human genome using Tophat and gene expression was analyzed with Cuffdiff[61]. After gene expression values

were normalized with log2 transformation, the variances were calculated between the five time points. The normalized gene expression values of genes located at each compartment types in five time points were visualized with the R program.

Significant Interaction Loops

The uniquely mapped paired-end reads were inputted into HOMER v4.7[49] to generate the significant interaction loops. In brief, HOMER was originally developed for a de novo motif discovery program and is now able to identify significant loops. The HOMER analyzeHiC module was used to make interaction matrices, normalize interaction counts and identify the significant interaction loops. The loops were further filtered with a LogP cutoff of -6 and distance cutoff 20kb between loci centers. Tamoxifen resistant differential loops of MCF7-TamR were obtained via the analyeHiC module of HOMER using MCF7 T0 as the contrast with the cutoff of FDR<=0.1 and the distance of loci pair center at 40kb to 5Mb.

Enrichment of KEGG pathway

Differentially expressed genes in various compartments were analyzed using Gene Set Enrichment Analysis (GSEA) v3.0[48]. Kyoto Encyclopedia of Genes and Genomes (KEGG) were selected as the geneset database. Gene ranking was determined by the ratio of log2 fold change to p-value of differential expression.

Data availability

Raw and processed TCC, ChIP-seq and RNA-seq data for MCF7 and TamR cells is deposited in GEO under the accession number GSE108787, and raw and processed TCC data for T47D and TamR cells is deposited in GEO under accession number GSE119890. The RNA-seq data of E2-treatment time series MCF7 is available at GEO accession number GSE62789.

Figures

Figure 4.1: Identification of E2-induced compartments in MCF7 cells at T1 versus T0

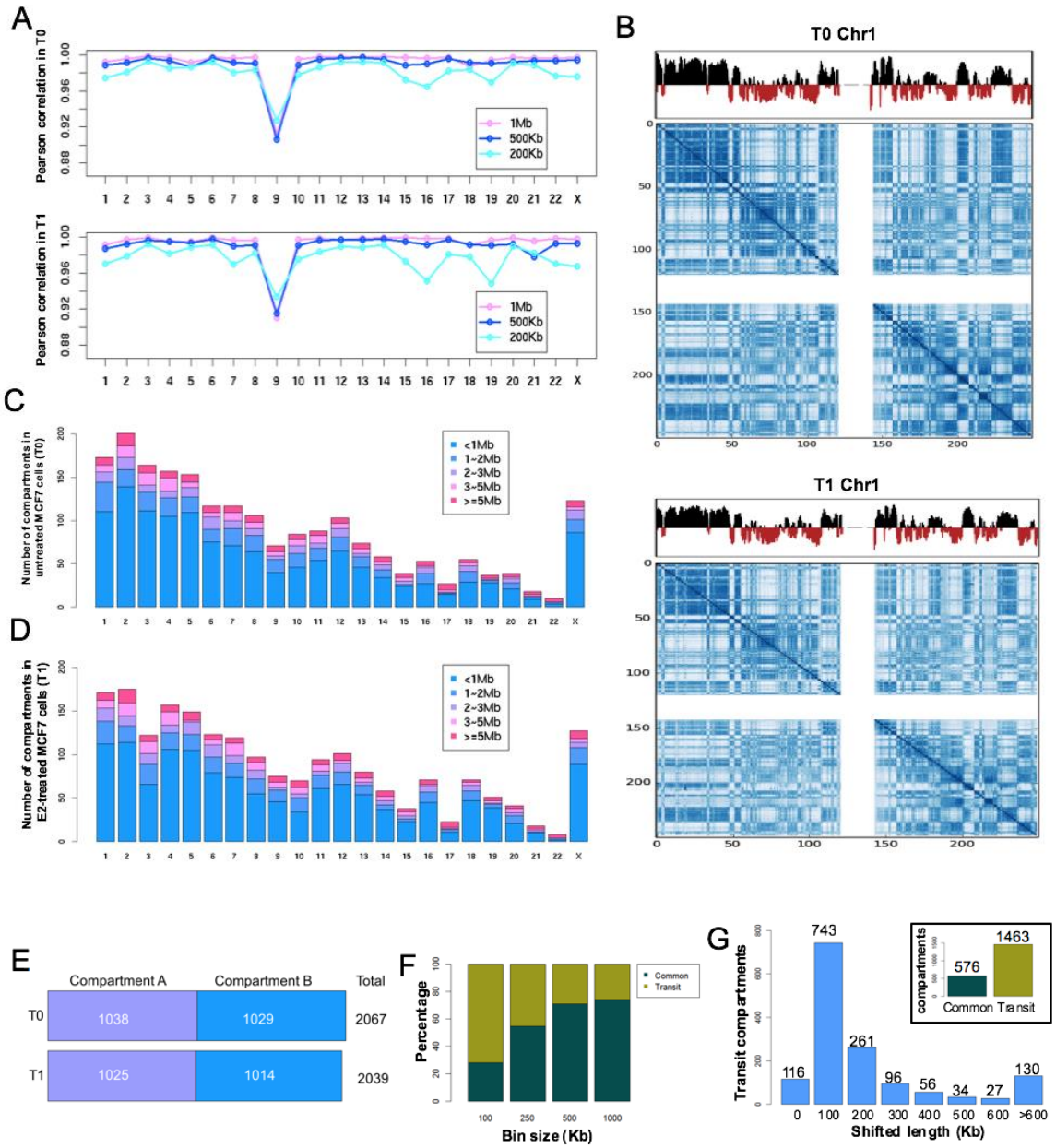
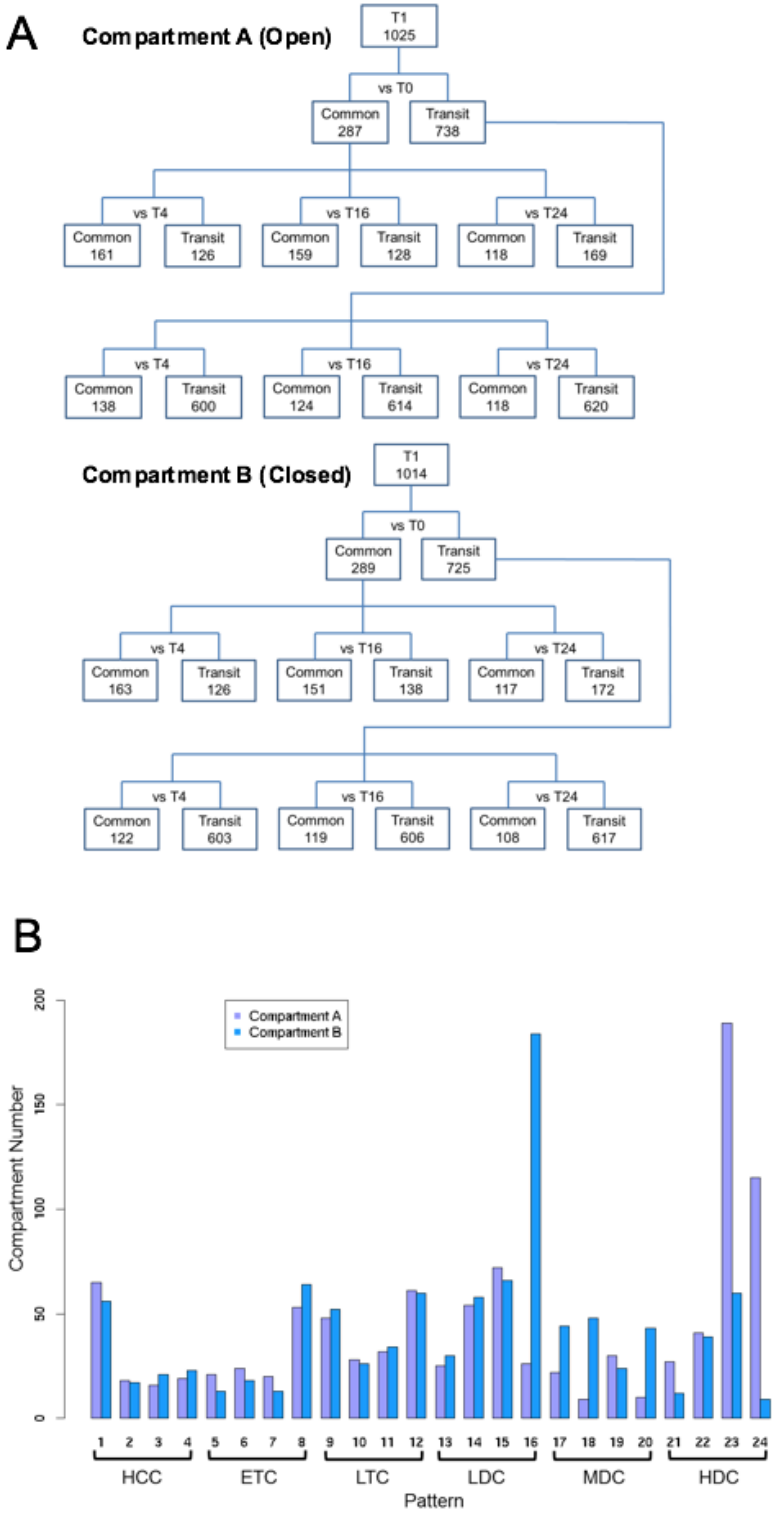


Figure 4.1: Identification of E2-induced compartments in MCF7 cells at T1 versus T0. (A) Pearson correlation of two biological replicates with the bin size of 1Mb, 500kb and 200kb on individual chromosomes at T0 (Top) and T1 (Bottom). (B) Contact matrices of compartment A (Black) and B (Red) at T0 (Top) and T1 (Bottom) respectively. (C,D) Histograms of compartments with different sizes at T0 and T1 respectively. (E) Distribution of compartments A and B in T0 and T1 respectively. (F) The percentage of common and transit compartments at T1 vs T0 with various bin sizes. (G) The number of compartments of T1 vs. T0 with various shifted length when bin size is 100kb.

Figure 4.2: Defining E2-induced temporal dynamic re-compartmentalization (TDRC) in MCF7 cells



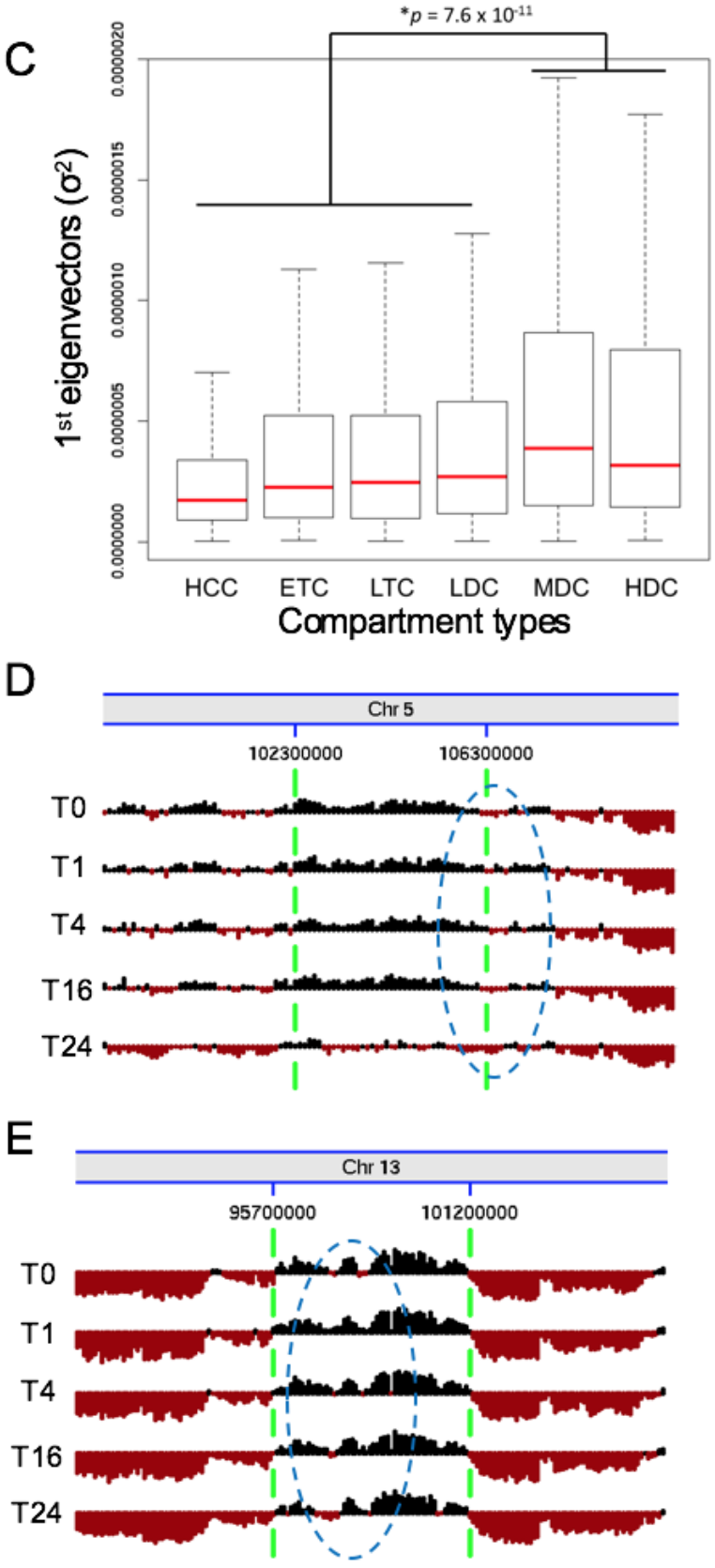

C
*p = 7.6 x 10-11
1st eigenvectors (σ2)
0.0000000
0.0000005
0.0000010
0.0000015
0.0000020
HCC
ETC
LTC
LDC
MDC
HDC
Compartment types
D
Chr 5
102300000
106300000
T0
T1
T4
T16
T24
E
Chr 13
95700000
101200000
T0
T1
T4
T16
T24

Figure 4.2: Defining E2-induced temporal dynamic re-compartmentalization (TDRC) in MCF7 cells. (A) An instructive tree displaying the comparisons of compartments A (Top) and B (Bottom) at E2-induced five time-points. T0: control, T1: 1 hour; T4: 4 hours; T16: 16 hours; and T24: 24 hours. (B) Number of compartments at each of the 24 individual patterns categorized into six types of altered compartments: HCC including patterns 1-4; ETC including patterns 5-8; LTC including patterns 9-12; LDC including patterns 13-16; MDC including patterns 17-20; HDC including patterns 21-24. (C) The variance of first eigenvectors of compartment types in T0, T1, T4, T16, T24. The *p* value was determined by the Wilcoxon rank-sum test. (D,E) Snapshots displaying examples of compartment changes, MDC (D) and HDC (E) along five time-points of E2-induction. The compartment of interest is indicated between the green dashed lines. The blue dashed ovals highlight the dynamic region.

Figure 4.3: Identification of altered compartments in TamR cells

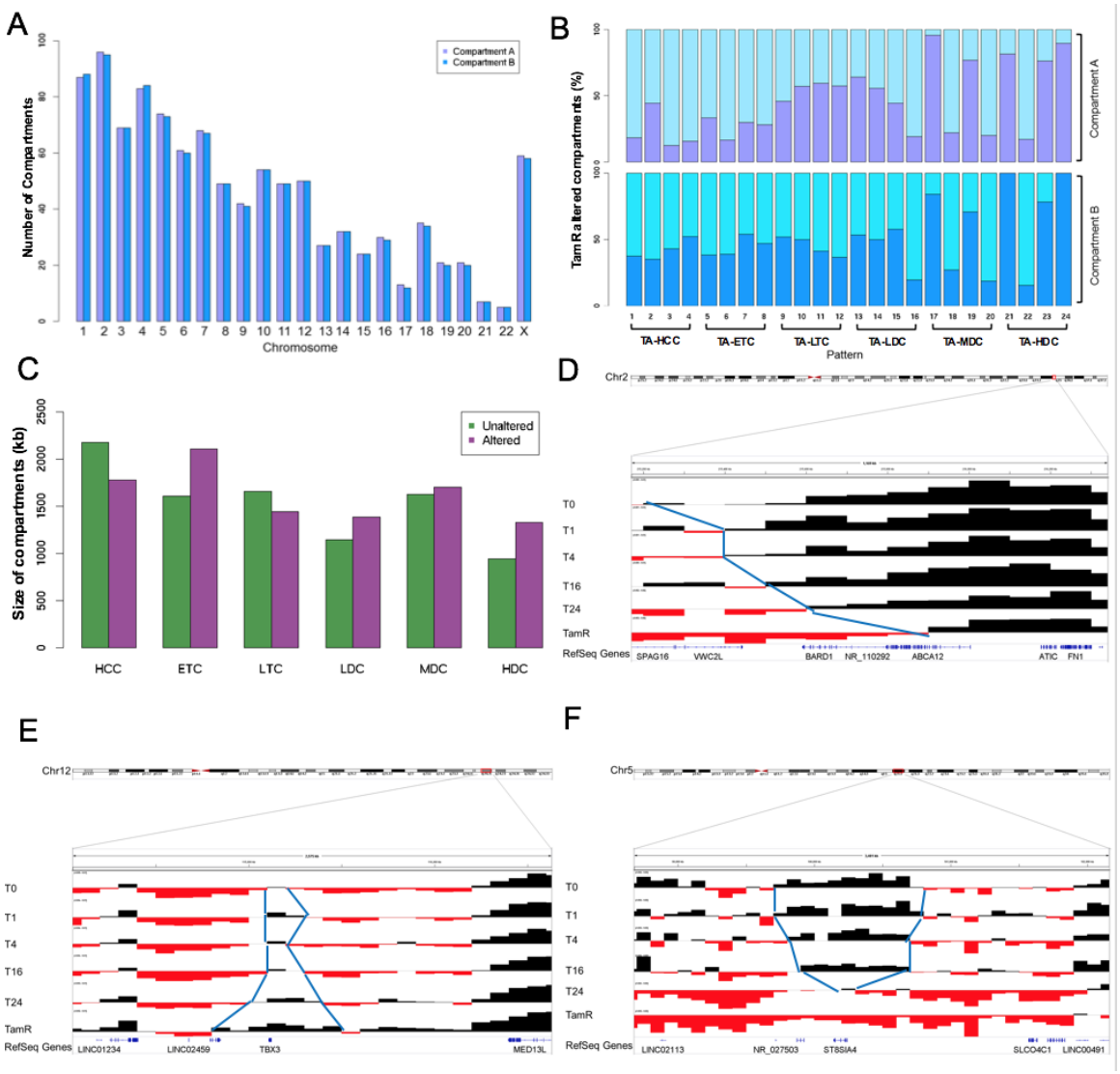
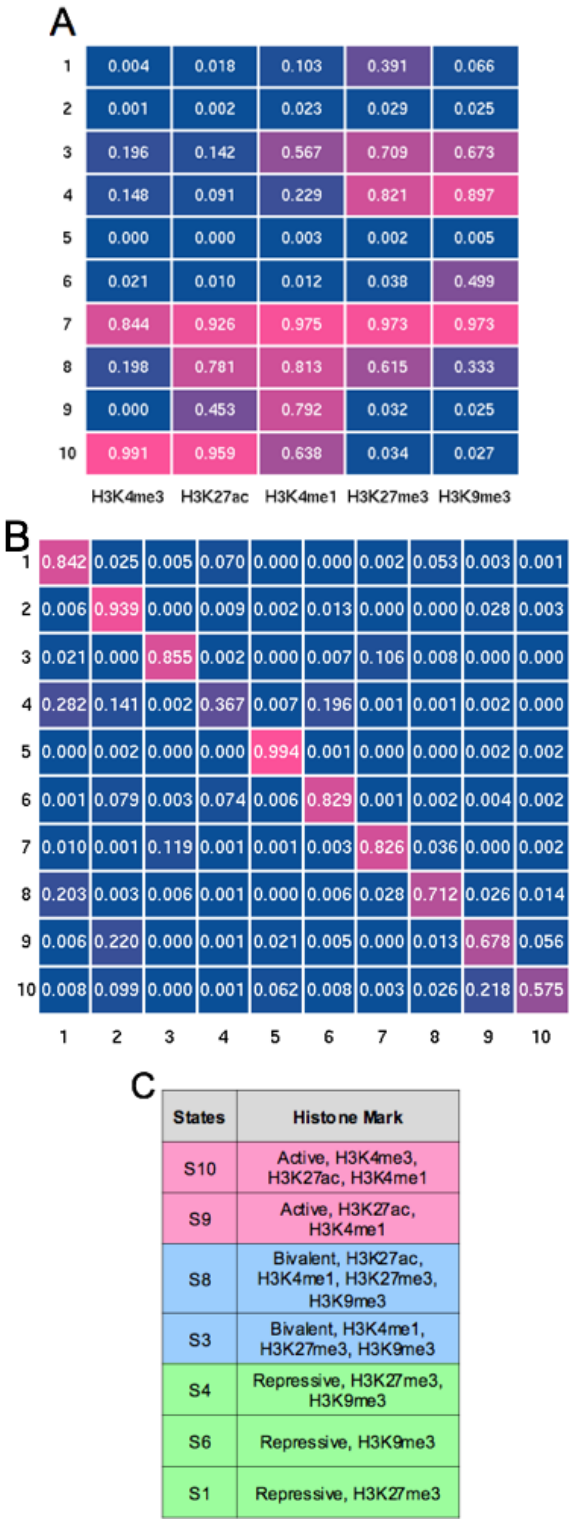


Figure 4.3: Identification of altered compartments in TamR cells. (A) The number of tamoxifen resistant compartments along chromosomes in MCF7 TamR cells. (B) The percentage of tamoxifen resistant altered compartments (TRACs), A (Top panel) and B (Bottom panel) in TamR cells. The x axis displays the identified compartments after comparison with the 24 states identified in responsive cells. (C) Genomic size of six types of TRACs and tamoxifen resistant unaltered compartments (TRUCs). (D-F) UCSC genome browser snapshots of TRACs and the enclosed genes within their loci. Dark: Compartment A; Gray: Compartment B. (D) contracted compartment. Blue lines represent the compartment boundary. (E) Expanded compartment. (F) Flipped compartment.

Figure 4.4: Epigenetic modifications within E2-induced TDRCs and TRACs



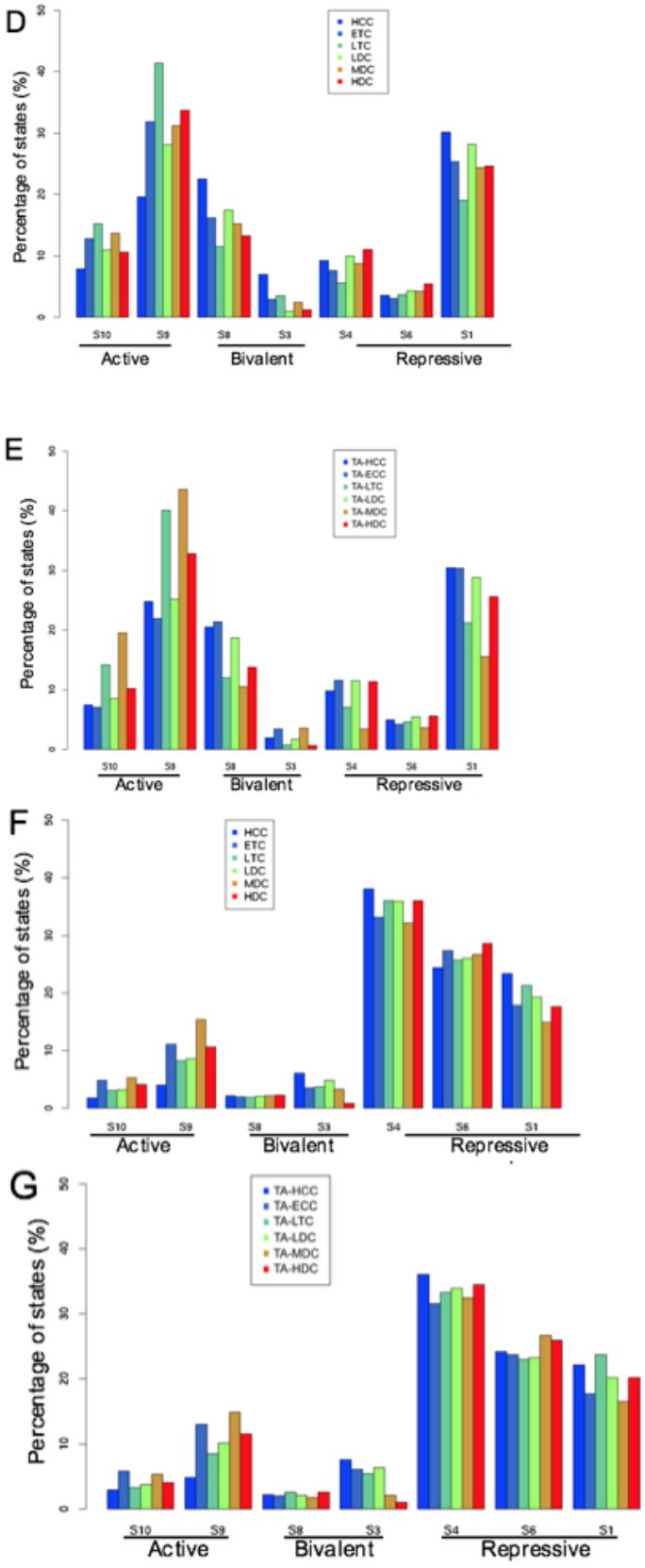
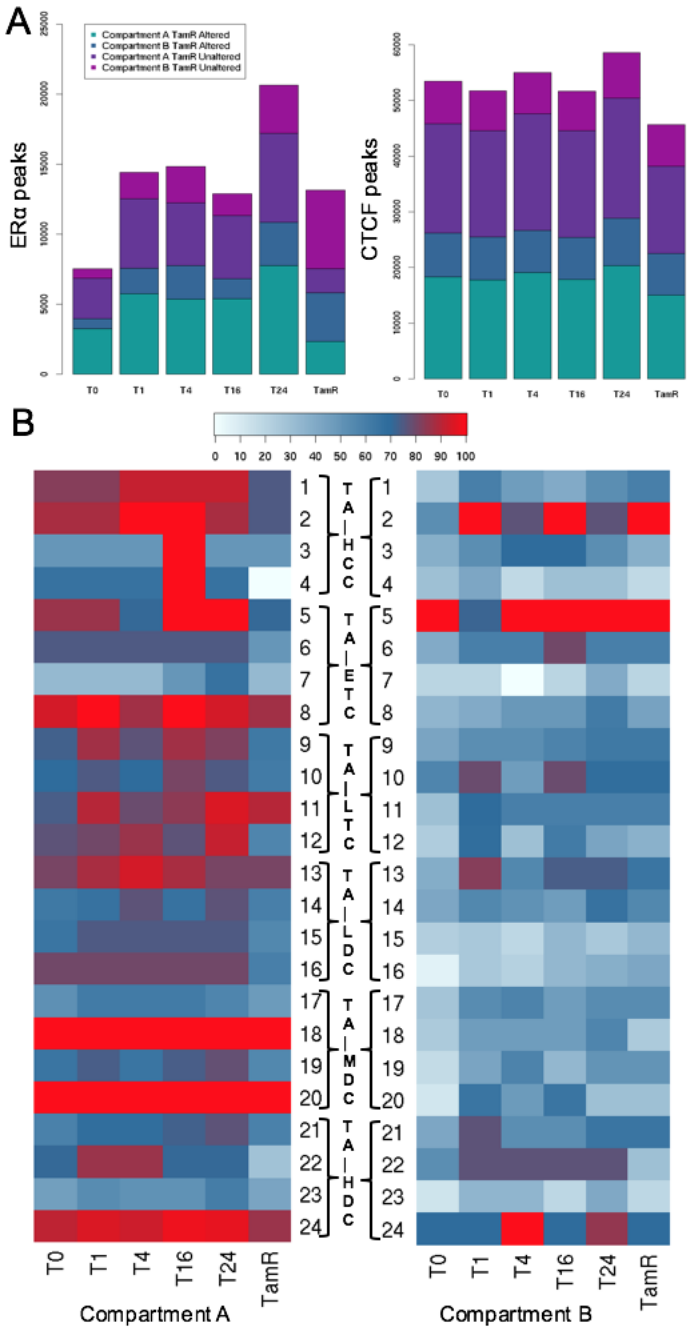

D
HCC
ETC
LTC
LDC
MDC
HDC
Percentage of states (%)
S10
S9
S8
S3
S4
S6
S1
Active
Bivalent
Repressive
E
TA-HCC
TA-ECC
TA-LTC
TA-LDC
TA-MDC
TA-HDC
Percentage of states (%)
S10
S9
S8
S3
S4
S6
S1
Active
Bivalent
Repressive
F
HCC
ETC
LTC
LDC
MDC
HDC
Percentage of states (%)
S10
S9
S8
S3
S4
S6
S1
Active
Bivalent
Repressive
G
TA-HCC
TA-ECC
TA-LTC
TA-LDC
TA-MDC
TA-HDC
Percentage of states (%)
S10
S9
S8
S3
S4
S6
S1
Active
Bivalent
Repressive

Figure 4.4: Epigenetic modifications on E2-induced TDRCs and TRACs. (A) Emission probabilities of ten epigenetic states trained by a HMM model on five histones. The enrichment of the corresponding mark is indicated by higher values and corresponds with a darker pink color. (B) Transition probabilities of ten epigenetic states trained by a HMM model on five histones. The enrichment of the corresponding mark is indicated by higher values and with a darker pink color. (C) The summary of the corresponding histone marks and each of the defined epigenetic states. (D) The percentage of epigenetic states on dynamic E2-induced compartment A. (E) The percentage of epigenetic states on altered TamR compartment A. (F) The percentage of epigenetic states on dynamic E2-induced compartment B. (G) The percentage of epigenetic states on altered TamR compartment B.

Figure 4.5: A distribution of ERα and CTCF peaks in E2-induced TDRCs and TRACs



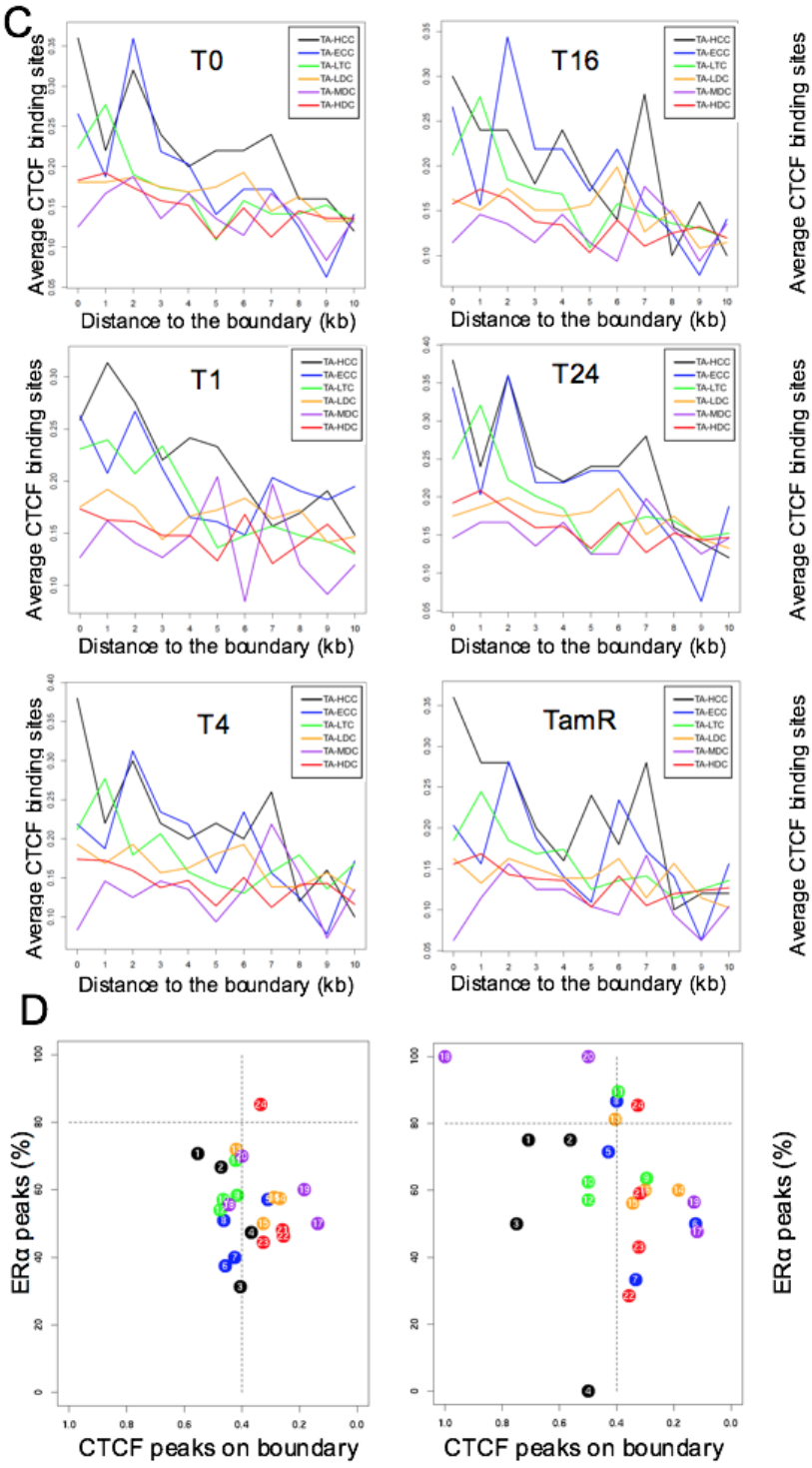
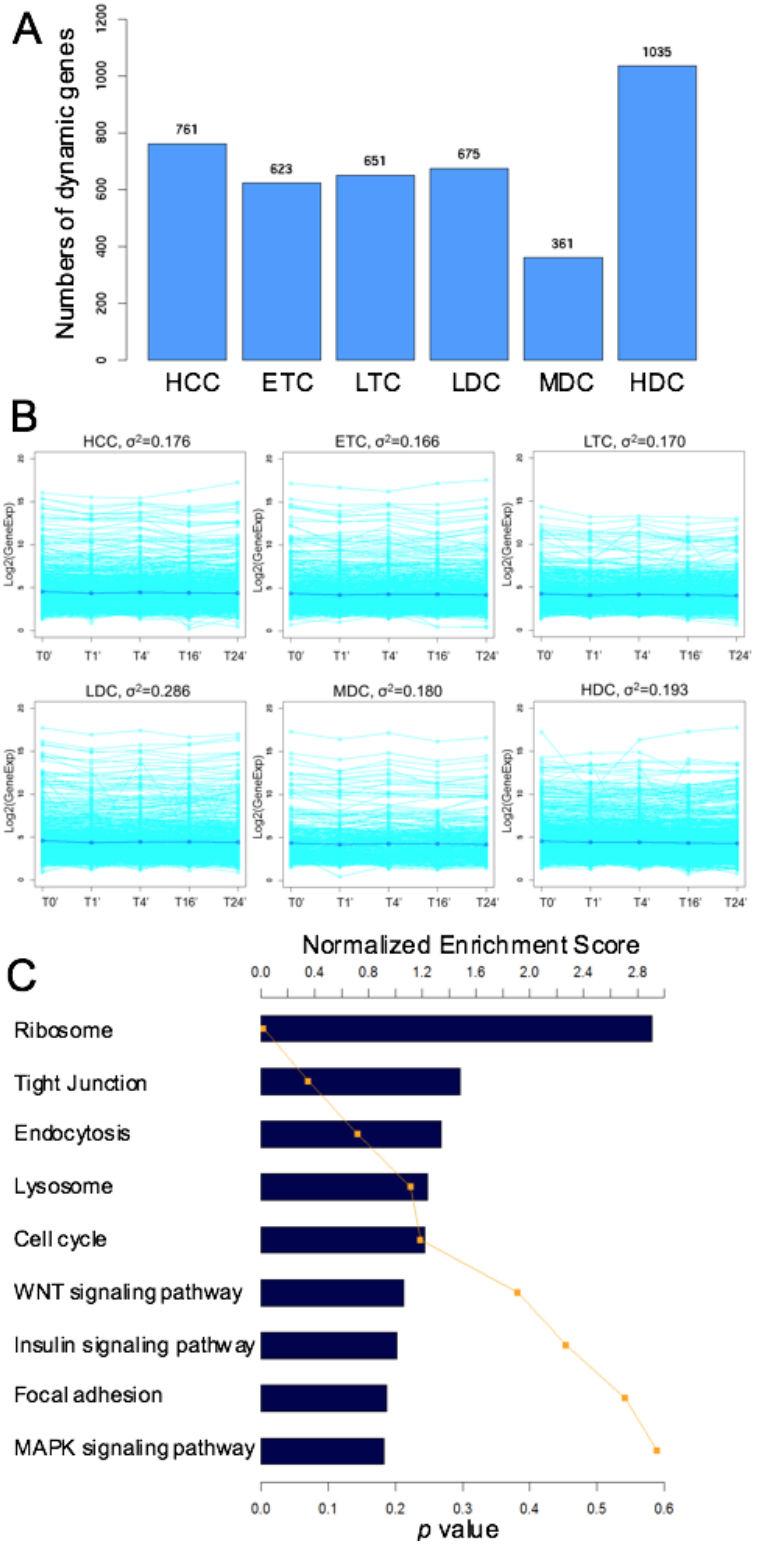

C
T0
T16
T1
T24
T4
TamR
TA-HCC
TA-ECC
TA-LTC
TA-LDC
TA-MDC
TA-HDC
Average CTCF binding sites
Distance to the boundary (kb)
D
ERα peaks (%)
CTCF peaks on boundary

Figure 4.5: A distribution of ERα and CTCF peaks in E2-induced TDRCs and TRACs. (A) The number of ERα (Left) and CTCF (Right) peaks within altered and unaltered TamR compartments. (B) Heatmap showing the percentage of altered TamR compartments A (Left) and B (Right) with ERα peaks. (C) The average number of CTCF peaks on altered TamR compartment boundary in E2-induced five time-points and TamR MCF7 cell lines. (D) The distribution of each of the 24 patterns (circled) of ERα peaks within compartments and CTCF peaks on compartment boundary.

Figure 4.6. Gene expression and looping in E2-induced TDRCs and TRACs



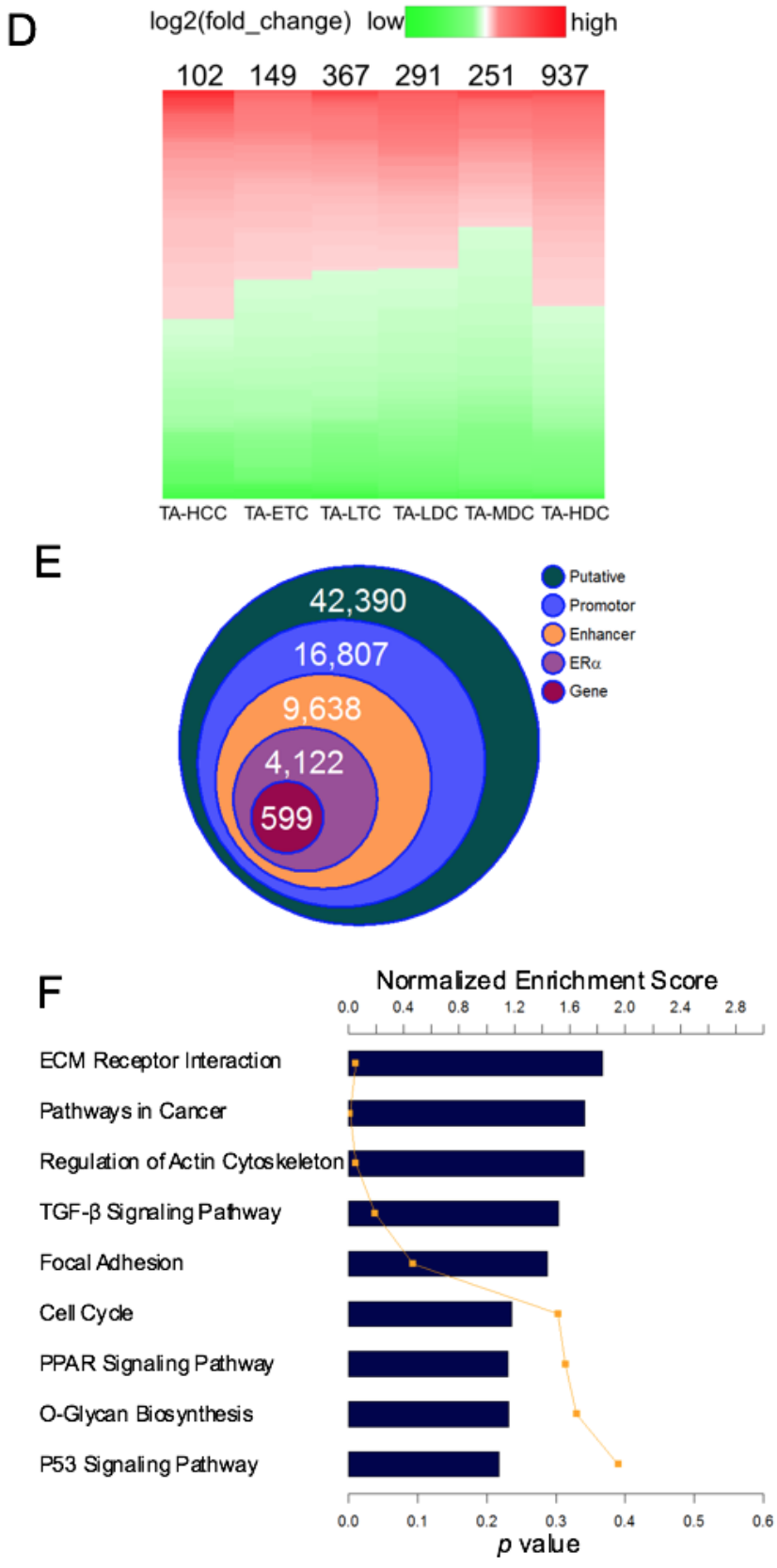

D
log2(fold_change) low high
102 149 367 291 251 937
TA-HCC TA-ETC TA-LTC TA-LDC TA-MDC TA-HDC
E
42,390
16,807
9,638
4,122
599
Putative
Promotor
Enhancer
ERα
Gene
F
Normalized Enrichment Score
0.0 0.4 0.8 1.2 1.6 2.0 2.4 2.8
ECM Receptor Interaction
Pathways in Cancer
Regulation of Actin Cytoskeleton
TGF-β Signaling Pathway
Focal Adhesion
Cell Cycle
PPAR Signaling Pathway
O-Glycan Biosynthesis
P53 Signaling Pathway
0.0 0.1 0.2 0.3 0.4 0.5 0.6
p value

Figure 4.6. Gene expression and looping in E2-induced TDRCs and TRACs. (A) The number of dynamic DEGs located in each of six types of TDRCs-A. (B) A plot of gene expression values along five time-points in each of the six types of TDRCs-A. Each green line represents one gene, and blue line represents the average gene expression value. (C) The enrichment of KEGG pathways of 1,396 genes in MDC/HDCs-A. (D) The heatmap of DEGs of TamR versus MCF7 in each of six types of TRACs-A. (E) The number of loops defined by HOMER. Putative: putative loops identified by HOMER. Promoter: one locus of putative loops within -5kb/+1kb of TSS. Enhancer: one locus of putative loops within -5kb/+1kb of TSS and the other with either H3K27ac/H3K4me1 peaks. ERα: at least one locus of loops with an ERα peak located within a promoter-enhancer region. Gene: genes associated with ERα-PE loops showing differentially expressed genes at TA-MDCs-A and TA-HDCs-A. (F) The enrichment of KEGG pathways of 396 genes associated 599 ERα-PE loops.

Figure 4.7: A proposed model for dynamic 3D chromatin architecture

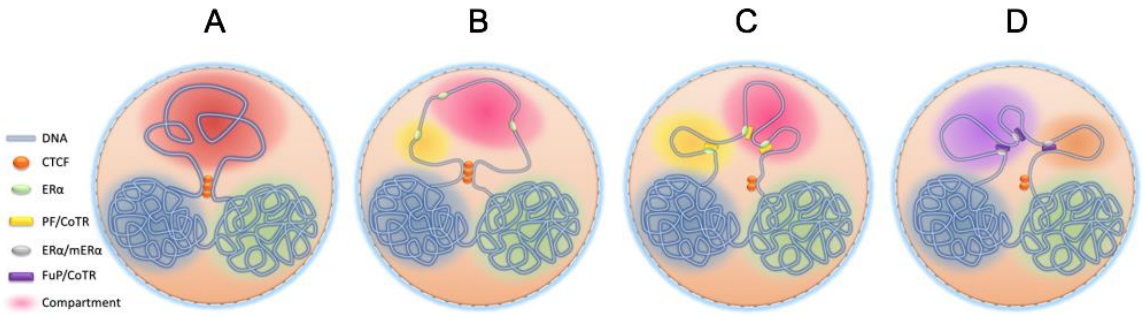


Figure 4.7. A proposed model for dynamic 3D chromatin architecture. (A-C). Constitutive estrogen stimulation in breast cancer cells enhances stronger ERα activity and further recruits its distal regulatory machinery and then mobilizes highly dynamic gene looping which essentially expand to render a 3D genome re-compartmentalization meanwhile force CTCF eviction resulting in reduced insulation activity at the compartment boundary. (D). In acquired resistant breast cancer cells, increased crosstalk between ERα and other signal transduction pathways such as EGFR/HER2, IGF-IR, and AKT/PTEN or altered expression of some key co-regulators particularly reshuffle these highly dynamic gene looping resulting in altered chromatin reorganization.

References

1. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293 (2009).
2. Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y.J., Lee, C., Shendure, J., Fields, S., Blau, C.A., Noble, W.S. A three-dimensional model of the yeast genome. *Nature* 465, 363–367 (2011).
3. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380 (2012).
4. Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., Bruneau, B.G. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 169, 930-944.e22. (2017).
5. Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., Cavalli, G. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* 148, 458–472 (2012).
6. Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., and Young, R.A. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159, 374–387 (2014).

7. Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Scho¨ pflin, R., Kraft, K., Kempfer, R., Jerkovic, I., Chan, W.-L., et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* 538, 265–269 (2016).

8. Ji, X., Dadon, D.B., Powell, B.E., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., Hnisz, D., Pegoraro, G., Lee, T.I., et al. 3D Chromosome regulatory landscape of human pluripotent cells. *Cell Stem Cell* 18, 262–275 (2016).

9. Bonev, B., and Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* 17, 661–678 (2016).

10. Imakaev, M., Fudenberg, G., McCord, R.P., Naumova, N., Goloborodko, A., Lajoie, B.R., Dekker, J., and Mirny, L.A. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* 9, 999–1003 (2012).

11. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., and van Steensel, B. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951 (2008).

12. Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385 (2012).

13. Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., Hadjur, S. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10, 1297–1309 (2015).

14. Naumova, N. and Dekker, J. Integrating one-dimensional and three-dimensional maps of genomes. *Journal of Cell Science* 123, 1979-1988 (2010).

15. Hou, C., Li, L., Qin, Z. S. & Corces, V. G. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol. Cell *48*, 471–484 (2012).

16. 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. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* 42, 53-61 (2010).

17. Dekker, J., Marti-Renom, M.A. and Mirny, L.A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Review Genetics* 14, 390-403 (2013).

18. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680 (2014).

19. Zehnbauer, B.A., and Vogelstein, B. Supercoiled loops and the organization of replication and transcription in eukaryotes. *BioEssays* 2, 52–54 (1985).

20. Flyamer, I.M., Gassler, J., Imakaev, M., Brandão, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., Tachibana-Konwalski, K. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* 544, 110-114 (2017).

21. Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C., Mirny, L., et al. Two independent modes of chromosome organization are revealed by cohesin removal. *Nature* 551, 51-56 (2017).

22. Barutcu, A.R., Lajoie, B.R., Fritz, A.J., McCord, R.P., Nickerson, J.A., van Wijnen, A.J., Lian, J.B., Stein, J.L., Dekker, J., Stein, G.S., et al. SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells. *Genome Res*. 26, 1188-201 (2016).

23. Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., Gao, L., et al. 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* 170, 367-381.e20 (2017).

24. Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y., et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* 547, 232-235 (2017).

25. Taberlay, P.C., Achinger-Kawecka, J., Lun, A.T., Buske, F.A., Sabir, K., Gould, C.M., Zotenko, E., Bert, S.A., Giles, K.A., Bauer, D.C., et al. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. *Genome Res*. 26, 719-31 (2016).

26. Heldring, N., Pike, A., Andersson, S., Matthews, J., Cheng, G., Hartman, J., Tujague, M., Strom, A., Treuter, E., Warner, M., Gustafsson, J.A. Estrogen receptors: how do they signal and what are their targets. *Physiol. Rev*. 87, 905–931 (2007).

27. Jordan, V.C. A century of deciphering the control mechanisms of sex steroid action in breast and prostate cancer: the origins of targeted therapy and chemoprevention. *Cancer Res*. 69, 1243–1254 (2009).

28. Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., et al. Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet*. 38, 1289-97 (2006).

29. Lin, C.Y., Vega, V.B., Thomsen, J.S., Zhang, T., Kong, S.L., Xie, M., Chiu, K.P., Lipovich, L., Barnett, D.H., Stossi, F., et al. Whole-genome cartography of estrogen receptor alpha binding sites. *PLoS Genet*. 3, e87 (2007).

30. Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 462, 58-64 (2009).

31. Cicatiello, L., Mutarelli, M., Grober, O.M., Paris, O., Ferraro, L., Ravo, M., Tarallo, R., Luo, S., Schroth, G.P., Seifert, M., Zinser, C., et al. Estrogen receptor alpha controls a gene network in luminal-like breast cancer cells comprising multiple transcription factors and microRNAs. *Am. J. Pathol*. 176, 2113-30 (2010).

32. Wang, J., Lan, X., Hsu, P.Y., Hsu, H.K., Huang, K., Parvin, J., Huang, T.H., Jin, V.X. Genome-wide analysis uncovers high frequency, strong differential chromosomal interactions and their associated epigenetic patterns in E2-mediated gene regulation. *BMC Genomics* 14, 70 (2013).

33. Fan, M., Yan, P.S., Hartman-Frey, C., Chen, L., Paik, H., Oyer, S.L., Salisbury, J.D., Cheng, A.S., Li, L., Abbosh, P.H., et al. Diverse gene expression and DNA

methylation profiles correlate with differential adaptation of breast cancer cells to the antiestrogens tamoxifen and fulvestrant. *Cancer Res.* 66, 11954-11966 (2006).

34. Frasor, J., Chang, E.C., Komm, B., Lin, C.Y., Vega, V.B., Liu, E.T., Miller, L.D., Smeds, J., Bergh, J., Katzenellenbogen, B.S. Gene expression preferentially regulated by tamoxifen in breast cancer cells and correlations with clinical outcome. *Cancer Res*. 66, 7334-40 (2006).

35. Welboren, W.J., van Driel, M.A., Janssen-Megens, E.M., van Heeringen, S.J., Sweep, F.C.G.J., Span, P.N., Stunnenberg, H.G. ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* 28, 1418-1428 (2009).

36. Gu, F., Hsu, H.K., Hsu, P.Y., Wu, J., Ma, Y., Parvin, J., Huang, T.H., Jin, V.X. Inference of hierarchical regulatory network of estrogen-dependent breast cancer through ChIP-based data. *BMC Syst. Biol.* 4, 170 (2010).

37. Magnani, L., Stoeck, A., Zhang, X., Lánczky, A., Mirabella, A.C., Wang, T.L., Gyorffy, B., Lupien, M. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proc. Natl. Acad. Sci. USA*. 110, E1490-9 (2013).

38. Hsu, P.Y., Hsu, H.K., Lan, X., Juan, L., Yan, P.S., Labanowska, J., Heerema, N., Hsiao, T.H., Chiu, Y.C., Chen, Y., et al. Amplification of distant estrogen response elements deregulates target genes associated with tamoxifen resistance in breast cancer. *Cancer Cell* 24, 197-212 (2013).

39. Mourad, R., Hsu, P.Y., Juan, L., Shen, C., Koneru, P., Lin, H., Liu, Y., Nephew, K., Huang, T.H., Li, L. Estrogen induces global reorganization of chromatin structure in human breast cancer cells. *PLoS One* 9, e113354 (2014).

40. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., Chen, L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* 30, 90-8 (2011).

41. Ay, F., Noble, W.S. Analysis methods for studying the 3D architecture of the genome. *Genome Biol.* 16, 183 (2015).

42. Hurtado, A., Holmes, K.A., Geistlinger, T.R., Hutcheson, I.R., Nicholson, R.I., Brown, M., Jiang, J., Howat, W.J., Ali, S., Carroll, J.S. Regulation of ERBB2 by oestrogen receptor-PAX2 determines response to tamoxifen. *Nature* 456, 663-666 (2008).

43. Fu X., Jeselsohn, R., Pereira, R., Hollingsworth, E.F., Creighton, C.J., Li, F., Shea, M., Nardone, A., De Angelis, C., Heiser, L.M., et al. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. *Proc. Natl. Acad. Sci. USA.* 113, E6600-E6609 (2016).

44. Osborne, C.K., Hobbs, K., Clark, G.M. Effect of estrogens and antiestrogens on growth of human breast cancer cells in athymic nude mice. *Cancer Res*. 45, 584–90 (1985).

45. Ernst, J., Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* 9, 215-6 (2012).

46. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137 (2008).

47. Honkela, A., Peltonen, J., Topa, H., Charapitsa, I., Matarese, F., Grote, K., Stunnenberg, H.G., Reid, G., Lawrence, N.D., Rattray, M. Genome-wide modeling of transcription kinetics reveals patterns of RNA production delays. *Proc. Natl. Acad. Sci. USA.* 112, 13115-20 (2015).

48. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA.* 102, 15545-50 (2005).

49. Heinz, S., Benner, C., Spann, N., Bertolino, E., et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576-589 (2010).

50. Yu, W., He, B., Tan, K. Identifying topologically associating domains and subdomains by Gaussian Mixture model and Proportion test. *Nature Communications* 8, 535 (2017).

51. Frasor, J., Danes, J.M., Komm, B., Chang, K.C., Lyttle, C.R., Katzenellenbogen, B.S. Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology* 144, 4562-74 (2003).

52. Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D., Carroll J.S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nature Genetics* 43, 27-33 (2011).

53. Browne, B.C., Hochgräfe, F., Wu, J., Millar, E.K., Barraclough, J., Stone, A., McCloy, R.A., Lee, C.S., Roberts, C., Ali, N.A., et al. Global characterization of signalling networks associated with tamoxifen resistance in breast cancer. *FEBS J*. 280, 5237-57 (2013).

54. Osborne, C.K., Jarman, M., McCague, R., Coronado, E.B., Hilsenbeck, S.G., Wakeling, A.E. The importance of tamoxifen metabolism in tamoxifen-stimulated breast tumor growth. *Cancer Chemother. Pharmacol.* 34, 89–95 (1994).

55. Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van IJcken, W.F., et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. USA*. 111, 996-1001 (2014).

56. Massarweh, S., Osborne, C.K., Creighton, C.J., Qin, L., Tsimelzon, A., Huang, S., Weiss, H., Rimawi, M., Schiff, R. Tamoxifen resistance in breast tumors is driven by growth factor receptor signaling with repression of classic estrogen receptor genomic function. *Cancer Res*. 68, 826-833 (2008).

57. Shoman, N., Klassen, S., McFadden, A., Bickis, M.G., Torlakovic, E. and Chibbar, R. Reduced PTEN expression predicts relapse in patients with breast carcinoma treated by tamoxifen. *Mod. Pathol*. 18, 250–259 (2005).

58. Hsu, P.Y., Hsu, H.K., Hsiao, T.H., Ye, Z., Wang, E., Profit, A.L., Jatoi, I., Chen, Y., Kirma, N.B., Jin, V.X., et al. Spatiotemporal control of estrogen-responsive transcription in ERα-positive breast cancer cells. *Oncogene* 35, 2379-89 (2016).

59. O'Geen, H., Frietze, S., Farnham, P.J. Using ChIP-seq technology to identify targets of zinc finger transcription factors. *Methods Mol. Biol*. 649, 437-55 (2010).

60. Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., Forné, T. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc*. 2, 1722-33 (2007).

61. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* 7, 562–578 (2012).

62. Langmead, B., Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357-359 (2012).

63. Stark R, Brown G. DiffBind: differential binding analysis of ChIP-Seq peak data. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf. (2011)

64. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S, Palmieri C, Caldas C, Carroll JS. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 481, 389-93 (2012).

65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15, 550 (2014).

66. Lanczky A, Nagy A, Bottai G, Munkacsy G, Paladini L, Szabo A, Santarpia L, Gyorffy B. miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2,178 breast cancer patients. *Breast Cancer Res. Treat*. 160(3):439-446 (2016).

CHAPTER 5

Discussion and Concluding Remarks

Chapter 5.1: Summary

Chromatin based-mechanisms involved in the establishment and maintenance of cellular phenotypes are mediated by higher-order chromatin organization and post translational histone modifications. These regulatory processes coordinate the recruitment of protein complexes to specific genomic targets that result in gene expression alterations. This dissertation seeks to employ integrative high-throughput methodologies to investigate dynamic epigenetic changes in cancer cell models. Presented here are new contributions to the field of cancer genomics where I investigate the effects of external stimuli on higher-order chromatin structure, post translational histone modifications and gene expression. In Chapters 2 and 3 I reveal dynamic epigenetic mechanisms in pancreatic cancer cell models in response to epigenetic inhibitors. In Chapter 4, I present my work in ER positive breast cancer cells where plasticity was observed during temporal estrogen stimulation in cell models sensitive or resistant to tamoxifen. Taken together, this dissertation reveals novel insight into dynamic epigenomic alterations that occur with external stimuli and provides insight into mechanisms underlying the therapeutic responses in cancer cells.

Chapter 5.2: Implications of epigenetic plasticity in pancreatic cancer cell models

5.2.1: Overall outlook and contributions to the field

The work presented in Chapters 2 and 3 ultimately present an analysis framework for investigating genome-wide histone modification landscapes in relation to 3D-chromatin topology. Additionally, I provide characterization of broad epigenetic domains

corresponding to histological grade and give insight into genome-wide domain targets of epigenetic inhibitors. Lastly, this work provides datasets in pancreatic cancer cell-lines that are now available for public use for studies desiring to investigate: (1) chromatin contacts and (2) epigenomic (histone modifications and 3D chromatin architecture) and transcriptomic alterations in response to ICG001 and C646. In section 5.2.4, I propose future directions that can build on the studies presented here.

5.2.2: Defining chromatin contacts and chromatin states in PDAC cells

The goal underlying the first part of this dissertation seeks to characterize genome-wide patterns of chromatin structure in pancreatic cancer cell-lines and to investigate the effects of two epigenomic inhibitors on these regions (Appendix A, Chapters 2 and 3). This work provides an important basis for future studies regarding epigenomic regulation and the effects of epigenomic inhibitors on the cancer genome.

In Chapter 2 of this dissertation, I present the genome-wide interactions within a widely used pancreatic cancer cell-line and functionally characterize these regions by: (1) defining genome-wide chromatin contacts; (2) classifying the surrounding domains in relation to histone modifications; (3) determining how histone acetyltransferase inhibitors impact chromosomal organization; and (4) investigating these interactions in regard to chromatin regulation mediated by transcription factor binding. This chapter demonstrates a computational analytical approach to investigating higher-order chromatin regulation data relating to histone modifications and gene expression to classify chromatin looping events. For the first time in the literature, we present a genome-wide view of chromatin domains in the extensively studied pancreatic cancer cell-line, PANC1. After classifying

these domains with epigenetic histone modifications, annotating these defined states to gene regions and correlating PANC1 gene expression with these classified domains, we found that domains coordinating with active epigenetic states contained increased gene expression patterns. Repressed epigenetic states corresponded with decreased expression, providing confidence in our domain classification pipeline as these epigenetic states are correlative of gene expression [1]. Modifications such as H3K4me1 are known to play a role in gene repression [2] but when co-marked with H3K27ac this signature typically corresponds to active enhancers [3]. While there are general trends between histone modifications and chromatin state, it is imperative to further investigate chromatin machinery within these mechanisms to gain further insight about underlying regulation.

In the next part of Chapter 2, I investigated changes in interaction domain loci in response to two epigenetic inhibitors and associate these altered regions with gene expression. In our initial study (Appendix A), we investigated the effect of two histone acetyltransferase (HAT) inhibitors on the pancreatic cancer transcriptome and found shared and independent targets of ICG-001 and C646. Recent reports have identified C646 to help sensitize pancreatic cancer cell-lines to the canonical therapeutic agent used to treat pancreatic cancer [4] and has shown efficacy in other cancers [5, 6]. Additionally, the efficacy of ICG-001 has shown similar effect [7, 8] and has furthermore been promoted its advancement to clinical trials (NCT01764477; NCT01606579; NCT01302405). With the potential in reverting the cancer phenotype, little is known about the impact of these inhibitors genome-wide. Our study highlighted in Appendix A was the first study to characterize the effects of these inhibitors on the cancer transcriptome. Chapter 2 builds on this study by investigating potential chromatin-based

mechanisms that may be influencing the transcriptional targets of these inhibitors. While changes in 3D chromatin topology has been observed in response to temperature stress [9], the impact of epigenetic inhibitors on genome-wide contacts was not known. Our characterization identified shifts in domains that were correlated with transcription factor-mediated regulation using TCF7L2. Ultimately, I suggest a potential mechanism by which transcription factor associated chromosome interactions containing an active epigenetic signature surrounding these regions are altered by epigenomic inhibition.

5.2.3: HAT inhibitors on broad epigenomic domains in PDAC cells

In Chapter 3, I further investigate the active histone modification landscape by defining typical and broad H3K4me3 and H3K27ac domains in pancreatic cancer cell models corresponding to histological grade. I then determine the impact of histone acetyltransferase inhibitors on these histone modification domains in the high-grade PDAC cell-line, PANC1, the model in which we explore their impact on the transcriptome (Appendix A) and chromatin architecture (Chapter 2). Broad epigenomic domains have recently been linked to cancer-related genes programs [10, 11]. Our comprehensive analysis of super-enhancers and broad H3K4me3 domains in high- and low-grade pancreatic ductal adenocarcinoma cell-lines revealed, for the first time, that both of genome wide landscapes of these broad epigenomic domains have the ability to correspond to distinct high and low histological grades. When correlating annotated genes within the broad domains with gene expression in the given cell-lines, we find that gene expression correlates with predicted patient prognosis. We observed global

acetylation and methylation changes after histone acetyltransferase inhibitor treatment and furthermore, alteration in broad domains.

5.2.4: Future Directions

Cell-lines are important tools for investigating multiple facets of cellular mechanisms. However; performing these analyses in low-grade and additional high-grade samples would provide additional insight into the aggressive phenotype of PDAC and could also allow us to investigate potential mechanisms underlying the different cellular phenotype of the individual cell-lines. Recent reports have characterized the PDAC transcriptome, genome and epigenome in primary and metastatic PDAC samples as well as patient derived xenografts of PDAC of the classical and basal subtypes [12, 13]. Integrating the domain characterizations and gene targets from our studies could narrow down specific targets and candidates for further investigations. We examine the role of TCF7L2 and for the first time, correlate this transcription factor with chromatin architecture. However, its direct involvement in the aggressive phenotype in PDAC is not well understood. TCF7L2 is a downstream transcription factor of this pathway, there are other TCF/LEF family members that could be of importance. Furthermore, future studies could incorporate additional transcription factor regulation to determine other regulatory mechanisms and moreover the biology of PDAC. These could functionally be tested through perturbation experiments of the given transcription factor or chromatin regulator and the described experiments could be repeated and integrated with the current datasets to infer direct regulatory roles. Future studies modulating target regions within specific chromatin states would provide further insight into the functional significance of our

associations. For example, the advancements in CRISPR/Cas9 technologies provides tools for which we can edit the genome in order to explore molecular consequences. A future study could utilize this technology to create a double strand break at a super-enhancer or broad H3K4me3 region containing a gene potentially contributing to the PDAC phenotype using a high- and low-grade cell model to determine the role these epigenetic regions are playing on gene expression of the gene of interest. Additionally, chromatin architecture data could be used to determine how the absence of this region alters chromatin interactions nearby and further phenotypic studies could be carried out to determine if targeting the broad region of interest provides clinical potential in reverting the cancer phenotype. Ultimately, exploring the relevance would be best determined by using human PDAC samples to interrogate these processes. In the realm of the genome-wide effects of epigenetic inhibitors on chromatin regulation, additional cell models should be used to gain insight into more generalized and cell-type specific effects of these inhibitors genome-wide.

Chapter 5.3: Implications of chromatin dynamics in breast cancer cell models

5.3.1: Overall outlook and contributions to the field

The work presented in Chapter 4 of this dissertation provides a report of the dynamics of 3D chromatin structure across a time course of estradiol (E2) stimulation in human estrogen receptor α (ERα) positive breast cancer cells in tamoxifen sensitive and resistant cells. In addition to these new revelations regarding chromatin compartment regulated through temporal estradiol stimulation, we proved new high-throughput datasets available to the general public for further investigations. These datasets include:

(1) genome-wide 3D architecture, (2) ChIP-sequencing and (3) RNA-sequencing throughout temporal estradiol stimulation in tamoxifen sensitive and resistant cells.

5.3.2: Dynamics during temporal estradiol stimulation

In Chapter 4, we present our study in estrogen receptor positive (ER+) breast cancer, where we investigate the dynamics of 3D chromatin reorganization. Our integrative genomic analyses using ER-α positive breast cell models provided new insight into epigenomic regulation of ER+ breast cancer mediated by ER-α. Also, we investigate epigenetic alterations in tamoxifen resistant derivatives. We observed genome-wide changes in chromatin compartments (open or closed) in breast cancer cells hormone starved followed by temporal stimulation of estradiol, an estrogen derivative. Moreover, we characterized the changes over five estradiol induction periods with respect to histone modifications and gene expression. Our results revealed coordinated changes in chromatin compartments in tamoxifen resistant cells, where frequently changed compartments were associated with increased ER-α activity. We identify ERα-bound promoter-enhancer loops within altered compartments that are linked to TamR differential gene expression. Ultimately, this large 3D-scale chromatin data provides a rich resource for studying the basic characteristics of hormone-dependent breast cancers and provides further insight into the mechanisms of tamoxifen resistance.

5.3.3: Future Directions

When identifying genome-wide compartmentalization changes, we used two different ER+ breast cancer cell-lines, MCF7 and T47D. Our goal in this Chapter was to

characterize these compartments and associate epigenetics and gene expression programs with these compartments using the MCF7 cell line; however, since these two breast cancer cell-lines are of the same classification and contain the same estrogen, progesterone and Her2 expression [14] we utilized T47D to confirm that the identification of estradiol-induced compartments were consistent between both cell-lines. Future studies using more cell-lines with diverse subtypes will provide additional insight into the biology of re-compartmentalization in breast cancer models. Additionally, carrying these investigations further into human primary samples with matched clinical controls will provide more clinical perspective into these dynamic mechanisms in tamoxifen resistance. Future studies incorporating high-throughput proteomics data for ER-α with our datasets can isolate specific candidates for additional functional studies regarding ER-α mediated chromatin regulation. Additionally, disrupting an ER-α region within the highly dynamic compartments will allow us to identify the potential implications of ER-α within these regulatory regions that are highly altered in response to estradiol and also largely changed in response to tamoxifen.

Chapter 5.4: Concluding remarks

Chromatin-based mechanisms including higher-order chromatin organization and histone modifications have been identified to alter chromatin accessibility and directly influence transcriptional programs and furthermore biological outcomes. This dissertation seeks to characterize these domains in cell models for pancreatic and breast cancer and to ultimately investigate epigenetic plasticity in response to different conditions. We observed dynamic changes in chromatin structure in pancreatic cancer in response to

epigenetic inhibitors and in breast cancer in response to estrogen stimulation. Future work exploring the molecular consequences of these genome-wide alterations are needed and furthermore, studies in primary patient samples will provide direct insight into clinical importance and can help narrow down directed targets for phenotypic studies. Additionally, 3D-FISH experiments can be designed in these models to confirm our 3D-chromatin findings. This dissertation ultimately provides insight into dynamic epigenetic processes. Importantly, this work contributes large genomic datasets that are publically available for the scientific community, which provides opportunities for new research projects. These datasets can be combined with other publically available data to investigate SNPS, other epigenomic modifications and additional DNA binding factors to conduct further meta-analyses for mechanistic perspectives.

References:

1. Nebbioso, A., et al., *Cancer epigenetics: Moving forward.* PLoS Genet, 2018. 14(6): p. e1007362.
2. Cheng, J., et al., *A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers.* Mol Cell, 2014. 53(6): p. 979-92.
3. Andersson, R., et al., *An atlas of active enhancers across human cell types and tissues.* Nature, 2014. 507(7493): p. 455-461.
4. Ono, H., M.D. Basson, and H. Ito, *P300 inhibition enhances gemcitabine-induced apoptosis of pancreatic cancer.* Oncotarget, 2016. 7(32): p. 51301-51310.
5. Oike, T., et al., *C646, a selective small molecule inhibitor of histone acetyltransferase p300, radiosensitizes lung cancer cells by enhancing mitotic catastrophe.* Radiother Oncol, 2014. 111(2): p. 222-7.
6. Gao, X.N., et al., *A histone acetyltransferase p300 inhibitor C646 induces cell cycle arrest and apoptosis selectively in AML1-ETO-positive AML cells.* PLoS One, 2013. 8(2): p. e55481.
7. Liu, Y., et al., *ICG-001 suppresses growth of gastric cancer cells and reduces chemoresistance of cancer stem cell-like population.* J Exp Clin Cancer Res, 2017. 36(1): p. 125.
8. Zhang, X., et al., *Wnt blockers inhibit the proliferation of lung cancer stem cells.* Drug Des Devel Ther, 2015. 9: p. 2399-407.
9. Li, L., et al., *Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing.* Mol Cell, 2015. 58(2): p. 216-31.

10. Hnisz, D., et al., *Super-enhancers in the control of cell identity and disease.* Cell, 2013. 155(4): p. 934-47.

11. Chen, K., et al., *Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes.* Nat Genet, 2015. 47(10): p. 1149-57.

12. Connor, A.A., et al., *Integration of Genomic and Transcriptional Features in Pancreatic Cancer Reveals Increased Cell Cycle Progression in Metastases.* Cancer Cell, 2019. 35(2): p. 267-282 e7.

13. Lomberk, G., et al., *Distinct epigenetic landscapes underlie the pathobiology of pancreatic cancer subtypes.* Nat Commun, 2018. 9(1): p. 1978.

14. Neve, R.M., et al., *A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes.* Cancer Cell, 2006. 10(6): p. 515-27.

COMPREHENSIVE BIBLIOGRAPHY

1. Boyd J (2018). seqsetvis: Set Based Visualizations for Next-Gen Sequencing Data. R package version 1.2.0.
2. Stark R and Brown G (2011). DiffBind: differential binding analysis of ChIP-seq peak datasets.
3. Abulwerdi F, Liao C, Liu M, Azmi AS, Aboukameel A, Mady AS, Gulappa T, Cierpicki T, Owens S, Zhang T, Sun D, Stuckey JA, Mohammad RM, Nikolovska-Coleska Z. A novel small-molecule inhibitor of mcl-1 blocks pancreatic cancer growth in vitro and in vivo. Mol Cancer Ther. 2014;13(3):565-75. doi: 10.1158/1535-7163.MCT-12-0767. PubMed PMID: 24019208; PMCID: PMC4174574.
4. Adam RC, Yang H, Rockowitz S, Larsen SB, Nikolova M, Oristian DS, Polak L, Kadaja M, Asare A, Zheng D, Fuchs E. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature. 2015;521(7552):366-70. doi: 10.1038/nature14289. PubMed PMID: 25799994; PMCID: PMC4482136.
5. Aguilar-Setien A, Loza-Rubio E, Salas-Rojas M, Brisseau N, Cliquet F, Pastoret PP, Rojas-Dotor S, Tesoro E, Kretschmer R. Salivary excretion of rabies virus by healthy vampire bats. Epidemiol Infect. 2005;133(3):517-22. PubMed PMID: 15966107; PMCID: 2870282.
6. Aguirre-Gamboa R, Gomez-Rueda H, Martinez-Ledesma E, Martinez-Torteya A, Chacolla-Huaringa R, Rodriguez-Barrientos A, Tamez-Pena JG, Trevino V. SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis. PLoS One. 2013;8(9):e74250. doi: 10.1371/journal.pone.0074250. PubMed PMID: 24066126; PMCID: PMC3774754.
7. Akdemir KC, Chin L. HiCPlotter integrates genomic data with interaction matrices. Genome Biol. 2015;16:198. doi: 10.1186/s13059-015-0767-1. PubMed PMID: 26392354; PMCID: PMC4576377.
8. al. KTKe. Histone Acetylation And Methylation. Chromatin and Disease Subcellular Biochemistry. 2007;41.
9. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A. 1964;51:786-94. PubMed PMID: 14172992; PMCID: PMC300163.
10. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet. 2006;7(1):55-65. doi: 10.1038/nrg1749. PubMed PMID: 16369572.
11. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9. doi: 10.1093/bioinformatics/btu638. PubMed PMID: 25260700; PMCID: PMC4287950.
12. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfischer L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, Jorgensen M, Andersen PR, Bertin N, Rackham O, Burroughs AM, Baillie JK, Ishizu Y, Shimizu Y, Furuhata E, Maeda S, Negishi Y, Mungall CJ, Meehan TF, Lassmann

T, Itoh M, Kawaji H, Kondo N, Kawai J, Lennartsson A, Daub CO, Heutink P, Hume DA, Jensen TH, Suzuki H, Hayashizaki Y, Muller F, Forrest ARR, Carninci P, Rehli M, Sandelin A. An atlas of active enhancers across human cell types and tissues. Nature. 2014;507(7493):455-61. doi: 10.1038/nature12787. PubMed PMID: 24670763; PMCID: PMC5215096.

13. Andoniou CE, Andrews DM, Degli-Esposti MA. Natural killer cells in viral infection: more than just killers. Immunol Rev. 2006;214:239-50. doi: 10.1111/j.1600-065X.2006.00465.x. PubMed PMID: 17100889.
14. Aparicio O, Geisberg JV, Struhl K. Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. Curr Protoc Cell Biol. 2004;Chapter 17:Unit 17 7. doi: 10.1002/0471143030.cb1707s23. PubMed PMID: 18228445.
15. Arensman MD, Telesca D, Lay AR, Kershaw KM, Wu N, Donahue TR, Dawson DW. The CREB-binding protein inhibitor ICG-001 suppresses pancreatic cancer growth. Mol Cancer Ther. 2014;13(10):2303-14. doi: 10.1158/1535-7163.MCT-13-1005. PubMed PMID: 25082960; PMCID: PMC4188417.
16. Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A. 1991;88(22):10148-52. PubMed PMID: 1946434; PMCID: PMC52885.
17. Atkinson SR, Marguerat S, Bahler J. Exploring long non-coding RNAs through sequencing. Semin Cell Dev Biol. 2012;23(2):200-5. doi: 10.1016/j.semcdb.2011.12.003. PubMed PMID: 22202731.
18. Au WC, Yeow WS, Pitha PM. Analysis of functional domains of interferon regulatory factor 7 and its association with IRF-3. Virology. 2001;280(2):273-82. doi: 10.1006/viro.2000.0782. PubMed PMID: 11162841.
19. Audia JE, Campbell RM. Histone Modifications and Cancer. Cold Spring Harb Perspect Biol. 2016;8(4):a019521. doi: 10.1101/cshperspect.a019521. PubMed PMID: 27037415; PMCID: PMC4817802.
20. Ay F, Bailey TL, Noble WS. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res. 2014;24(6):999-1011. doi: 10.1101/gr.160374.113. PubMed PMID: 24501021; PMCID: PMC4032863.
21. Ay F, Noble WS. Analysis methods for studying the 3D architecture of the genome. Genome Biol. 2015;16:183. doi: 10.1186/s13059-015-0745-7. PubMed PMID: 26328929; PMCID: PMC4556012.
22. Ballestar E, Wolffe AP. Methyl-CpG-binding proteins. Targeting specific gene repression. Eur J Biochem. 2001;268(1):1-6. PubMed PMID: 11121095.
23. Bannister AJ, Miska EA, Gorlich D, Kouzarides T. Acetylation of importin-alpha nuclear import factors by CBP/p300. Curr Biol. 2000;10(8):467-70. PubMed PMID: 10801418.
24. Bannister AJ, Schneider R, Kouzarides T. Histone methylation: dynamic or static? Cell. 2002;109(7):801-6. PubMed PMID: 12110177.
25. Barnard ME, Boeke CE, Tamimi RM. Established breast cancer risk factors and risk of intrinsic tumor subtypes. Biochim Biophys Acta. 2015;1856(1):73-85. doi: 10.1016/j.bbcan.2015.06.002. PubMed PMID: 26071880.

26. Barutcu AR, Lajoie BR, Fritz AJ, McCord RP, Nickerson JA, van Wijnen AJ, Lian JB, Stein JL, Dekker J, Stein GS, Imbalzano AN. SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells. Genome Res. 2016;26(9):1188-201. doi: 10.1101/gr.201624.115. PubMed PMID: 27435934; PMCID: PMC5052043.
27. Bau D, Sanyal A, Lajoie BR, Capriotti E, Byron M, Lawrence JB, Dekker J, Marti-Renom MA. The three-dimensional folding of the alpha-globin gene domain reveals formation of chromatin globules. Nat Struct Mol Biol. 2011;18(1):107-14. doi: 10.1038/nsmb.1936. PubMed PMID: 21131981; PMCID: PMC3056208.
28. Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, Devarajan K, Daugherty AC, Kundaje AB, Mancini E, Hitz BC, Gupta R, Rando TA, Baker JC, Snyder MP, Cherry JM, Brunet A. H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell. 2014;158(3):673-88. doi: 10.1016/j.cell.2014.06.027. PubMed PMID: 25083876; PMCID: PMC4137894.
29. Bird AP, Wolffe AP. Methylation-induced repression--belts, braces, and chromatin. Cell. 1999;99(5):451-4. PubMed PMID: 10589672.
30. Blahnik KR, Dou L, O'Geen H, McPhillips T, Xu X, Cao AR, Iyengar S, Nicolet CM, Ludascher B, Korf I, Farnham PJ. Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data. Nucleic Acids Res. 2010;38(3):e13. doi: 10.1093/nar/gkp1012. PubMed PMID: 19906703; PMCID: PMC2817454.
31. Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, Coleman JT, Darling SR, Gargas A, Niver R, Okoniewski JC, Rudd RJ, Stone WB. Bat white-nose syndrome: an emerging fungal pathogen? Science. 2009;323(5911):227. Epub 2008/11/01. doi: 1163874 [pii] 10.1126/science.1163874. PubMed PMID: 18974316.
32. Bonev B, Cavalli G. Organization and function of the 3D genome. Nat Rev Genet. 2016;17(11):661-78. doi: 10.1038/nrg.2016.112. PubMed PMID: 27739532.
33. Bonneville R, Jin VX. A hidden Markov model to identify combinatorial epigenetic regulation patterns for estrogen receptor alpha target genes. Bioinformatics. 2013;29(1):22-8. doi: 10.1093/bioinformatics/bts639. PubMed PMID: 23104890.
34. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol. 2006;13(12):1097-101. doi: 10.1038/nsmb1167. PubMed PMID: 17099701.
35. Bowen MD, Peters CJ, Nichol ST. The phylogeny of New World (Tacaribe complex) arenaviruses. Virology. 1996;219(1):285-90. PubMed PMID: 8623541.
36. Bowen MD, Peters CJ, Nichol ST. Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. Mol Phylogenet Evol. 1997;8(3):301-16. doi: 10.1006/mpev.1997.0436. PubMed PMID: 9417890.
37. Bowers EM, Yan G, Mukherjee C, Orry A, Wang L, Holbert MA, Crump NT, Hazzalin CA, Liszczak G, Yuan H, Larocca C, Saldanha SA, Abagyan R, Sun Y,

Meyers DJ, Marmorstein R, Mahadevan LC, Alani RM, Cole PA. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. Chem Biol. 2010;17(5):471-82. doi: 10.1016/j.chembiol.2010.03.006. PubMed PMID: 20534345; PMCID: PMC2884008.

38. Brand JM, Frohn C, Cziupka K, Brockmann C, Kirchner H, Luhm J. Prolactin triggers pro-inflammatory immune responses in peripheral immune cells. Eur Cytokine Netw. 2004;15(2):99-104. PubMed PMID: 15319167.
39. Briere D, Sudhakar N, Woods DM, Hallin J, Engstrom LD, Aranda R, Chiang H, Sodre AL, Olson P, Weber JS, Christensen JG. The class I/IV HDAC inhibitor mocetinostat increases tumor antigen presentation, decreases immune suppressive cell types and augments checkpoint inhibitor therapy. Cancer Immunol Immunother. 2018;67(3):381-92. doi: 10.1007/s00262-017-2091-y. PubMed PMID: 29124315.
40. Brown MS, Goldstein JL. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell. 1997;89(3):331-40. PubMed PMID: 9150132.
41. Browne BC, Hochgrafe F, Wu J, Millar EK, Barraclough J, Stone A, McCloy RA, Lee CS, Roberts C, Ali NA, Boulghourjian A, Schmich F, Linding R, Farrow L, Gee JM, Nicholson RI, O'Toole SA, Sutherland RL, Musgrove EA, Butt AJ, Daly RJ. Global characterization of signalling networks associated with tamoxifen resistance in breast cancer. FEBS J. 2013;280(21):5237-57. doi: 10.1111/febs.12441. PubMed PMID: 23876235.
42. Buck MJ, Lieb JD. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. Genomics. 2004;83(3):349-60. PubMed PMID: 14986705.
43. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. J Mol Endocrinol. 2000;25(2):169-93. PubMed PMID: 11013345.
44. Cairns BR. Emerging roles for chromatin remodeling in cancer biology. Trends Cell Biol. 2001;11(11):S15-21. PubMed PMID: 11684437.
45. Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. Bats: important reservoir hosts of emerging viruses. Clin Microbiol Rev. 2006;19(3):531-45. PubMed PMID: 16847084.
46. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell. 2013;49(5):825-37. doi: 10.1016/j.molcel.2013.01.038. PubMed PMID: 23473601; PMCID: PMC3857148.
47. Campbell CAR. Bats, mosquitoes, and dollars. Boston, Mass.: The Stratford Co.
1. ; 1925.
48. Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell. 2017;32(2):185-203 e13. doi: 10.1016/j.ccell.2017.07.007. PubMed PMID: 28810144; PMCID: PMC5964983.
49. Cao F, Fang Y, Tan HK, Goh Y, Choy JYH, Koh BTH, Hao Tan J, Bertin N, Ramadass A, Hunter E, Green J, Salter M, Akoulitchev A, Wang W, Chng WJ,

Tenen DG, Fullwood MJ. Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. Sci Rep. 2017;7(1):2186. doi: 10.1038/s41598-017-02257-3. PubMed PMID: 28526829; PMCID: PMC5438348.

50. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M. Genome-wide analysis of estrogen receptor binding sites. Nat Genet. 2006;38(11):1289-97. doi: 10.1038/ng1901. PubMed PMID: 17013392.
51. Cauchi S, Froguel P. TCF7L2 genetic defect and type 2 diabetes. Curr Diab Rep. 2008;8(2):149-55. PubMed PMID: 18445358.
52. Cawston TE, Young DA. Proteinases involved in matrix turnover during cartilage and bone breakdown. Cell Tissue Res. 2010;339(1):221-35. doi: 10.1007/s00441-009-0887-6. PubMed PMID: 19915869.
53. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009;10(5):295-304. doi: 10.1038/nrg2540. PubMed PMID: 19308066.
54. Chargaff E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. Experientia. 1950;6(6):201-9. PubMed PMID: 15421335.
55. Chen D, Niu M, Jiao X, Zhang K, Liang J, Zhang D. Inhibition of AKT2 enhances sensitivity to gemcitabine via regulating PUMA and NF-kappaB signaling pathway in human pancreatic ductal adenocarcinoma. Int J Mol Sci. 2012;13(1):1186-208. doi: 10.3390/ijms13011186. PubMed PMID: 22312312; PMCID: PMC3269746.
56. Chen K, Chen Z, Wu D, Zhang L, Lin X, Su J, Rodriguez B, Xi Y, Xia Z, Chen X, Shi X, Wang Q, Li W. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet. 2015;47(10):1149-57. doi: 10.1038/ng.3385. PubMed PMID: 26301496; PMCID: PMC4780747.
57. Cheng J, Blum R, Bowman C, Hu D, Shilatifard A, Shen S, Dynlacht BD. A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. Mol Cell. 2014;53(6):979-92. doi: 10.1016/j.molcel.2014.02.032. PubMed PMID: 24656132; PMCID: PMC4031464.
58. Cheung AC, Cramer P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature. 2011;471(7337):249-53. doi: 10.1038/nature09785. PubMed PMID: 21346759.
59. Chi P, Allis CD, Wang GG. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nat Rev Cancer. 2010;10(7):457-69. doi: 10.1038/nrc2876. PubMed PMID: 20574448; PMCID: PMC3262678.
60. Choi SH, Gearhart MD, Cui Z, Bosnakovski D, Kim M, Schennum N, Kyba M. DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. Nucleic Acids Res. 2016;44(11):5161-73. doi: 10.1093/nar/gkw141. PubMed PMID: 26951377; PMCID: PMC4914088.
61. Ciabrelli F, Cavalli G. Chromatin-driven behavior of topologically associating

domains. J Mol Biol. 2015;427(3):608-25. doi: 10.1016/j.jmb.2014.09.013. PubMed PMID: 25280896.
62. Cicatiello L, Mutarelli M, Grober OM, Paris O, Ferraro L, Ravo M, Tarallo R, Luo S, Schroth GP, Seifert M, Zinser C, Chiusano ML, Traini A, De Bortoli M, Weisz A. Estrogen receptor alpha controls a gene network in luminal-like breast cancer cells comprising multiple transcription factors and microRNAs. Am J Pathol. 2010;176(5):2113-30. doi: 10.2353/ajpath.2010.090837. PubMed PMID: 20348243; PMCID: PMC2861078.
63. Ciliberto D, Botta C, Correale P, Rossi M, Caraglia M, Tassone P, Tagliaferri P. Role of gemcitabine-based combination therapy in the management of advanced pancreatic cancer: a meta-analysis of randomised trials. Eur J Cancer. 2013;49(3):593-603. doi: 10.1016/j.ejca.2012.08.019. PubMed PMID: 22989511.
64. Clark DW, Mitra A, Fillmore RA, Jiang WG, Samant RS, Fodstad O, Shevde LA. NUPR1 interacts with p53, transcriptionally regulates p21 and rescues breast epithelial cells from doxorubicin-induced genotoxic stress. Curr Cancer Drug Targets. 2008;8(5):421-30. PubMed PMID: 18690848.
65. Clevers H. Wnt/beta-catenin signaling in development and disease. Cell. 2006;127(3):469-80. doi: 10.1016/j.cell.2006.10.018. PubMed PMID: 17081971.
66. Cochrane DR, Bernales S, Jacobsen BM, Cittelly DM, Howe EN, D'Amato NC, Spoelstra NS, Edgerton SM, Jean A, Guerrero J, Gomez F, Medicherla S, Alfaro IE, McCullagh E, Jedlicka P, Torkko KC, Thor AD, Elias AD, Protter AA, Richer JK. Role of the androgen receptor in breast cancer and preclinical analysis of enzalutamide. Breast Cancer Res. 2014;16(1):R7. doi: 10.1186/bcr3599. PubMed PMID: 24451109; PMCID: PMC3978822.
67. Cogswell-Hawkinson A, Bowen R, James S, Gardiner D, Calisher CH, Adams R, Schountz T. Tacaribe virus causes fatal infection of an ostensible reservoir host, the Jamaican fruit bat. J Virol. 2012;86(10):5791-9. doi: 10.1128/JVI.00201-12. PubMed PMID: 22379103; PMCID: 3347293.
68. Cogswell-Hawkinson AC, Mitchell E. McGlaughlin, Charles H. Calisher, Rick Adams, and Tony Schountz. Molecular and Phylogenetic Characterization of Cytokine Genes from Seba's Short-tailed Fruit Bat (Carollia perspicillata). Open Immunology Journal. 2011;4:31-9.
69. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, Cooc J, Weinkle J, Kim GE, Jakkula L, Feiler HS, Ko AH, Olshen AB, Danenberg KL, Tempero MA, Spellman PT, Hanahan D, Gray JW. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med. 2011;17(4):500-3. doi: 10.1038/nm.2344. PubMed PMID: 21460848; PMCID: PMC3755490.
70. Conesa A, Gotz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics. 2008;2008:619832. doi: 10.1155/2008/619832. PubMed PMID: 18483572; PMCID: 2375974.
71. Connor AA, Denroche RE, Jang GH, Lemire M, Zhang A, Chan-Seng-Yue M, Wilson G, Grant RC, Merico D, Lungu I, Bartlett JMS, Chadwick D, Liang SB, Eagles J, Mbabaali F, Miller JK, Krzyzanowski P, Armstrong H, Luo X, Jorgensen LGT, Romero JM, Bavi P, Fischer SE, Serra S, Hafezi-Bakhtiari S,

Caglar D, Roehrl MHA, Cleary S, Hollingsworth MA, Petersen GM, Thayer S, Law CHL, Nanji S, Golan T, Smith AL, Borgida A, Dodd A, Hedley D, Wouters BG, O'Kane GM, Wilson JM, Zogopoulos G, Notta F, Knox JJ, Gallinger S. Integration of Genomic and Transcriptional Features in Pancreatic Cancer Reveals Increased Cell Cycle Progression in Metastases. Cancer Cell. 2019;35(2):267-82 e7. doi: 10.1016/j.ccell.2018.12.010. PubMed PMID: 30686769.

72. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74. doi: 10.1038/nature11247. PubMed PMID: 22955616; PMCID: PMC3439153.
73. Cremer M, Kupper K, Wagler B, Wizelman L, von Hase J, Weiland Y, Kreja L, Diebold J, Speicher MR, Cremer T. Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. J Cell Biol. 2003;162(5):809-20. doi: 10.1083/jcb.200304096. PubMed PMID: 12952935; PMCID: PMC2172812.
74. Cremer M, von Hase J, Volm T, Brero A, Kreth G, Walter J, Fischer C, Solovei I, Cremer C, Cremer T. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. Chromosome Res. 2001;9(7):541-67. PubMed PMID: 11721953.
75. Cremer T, Cremer C, Baumann H, Luedtke EK, Sperling K, Teuber V, Zorn C. Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. Hum Genet. 1982;60(1):46-56. PubMed PMID: 7076247.
76. Cremer T, Kurz A, Zirbel R, Dietzel S, Rinke B, Schrock E, Speicher MR, Mathieu U, Jauch A, Emmerich P, Scherthan H, Ried T, Cremer C, Lichter P. Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harb Symp Quant Biol. 1993;58:777-92. PubMed PMID: 7525149.
77. Crick F. Central dogma of molecular biology. Nature. 1970;227(5258):561-3. PubMed PMID: 4913914.
78. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalatskaya I, Mahajan S, May B, Ndegwa N, Schmidt E, Shamovsky V, Yung C, Birney E, Hermjakob H, D'Eustachio P, Stein L. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res. 2011;39(Database issue):D691-7. doi: 10.1093/nar/gkq1018. PubMed PMID: 21067998; PMCID: 3013646.
79. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. Nature. 2013;502(7472):499-506. doi: 10.1038/nature12753. PubMed PMID: 24153303.
80. Deer EL, Gonzalez-Hernandez J, Coursen JD, Shea JE, Ngatia J, Scaife CL, Firpo MA, Mulvihill SJ. Phenotype and genotype of pancreatic cancer cell lines. Pancreas. 2010;39(4):425-35. doi: 10.1097/MPA.0b013e3181c15963. PubMed PMID: 20418756; PMCID: PMC2860631.
81. Dekker FJ, Haisma HJ. Histone acetyl transferases as emerging drug targets. Drug Discov Today. 2009;14(19-20):942-8. doi: 10.1016/j.drudis.2009.06.008.

PubMed PMID: 19577000.
82. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet. 2013;14(6):390-403. doi: 10.1038/nrg3454. PubMed PMID: 23657480; PMCID: PMC3874835.
83. Dekker J, Mirny L. The 3D Genome as Moderator of Chromosomal Communication. Cell. 2016;164(6):1110-21. doi: 10.1016/j.cell.2016.02.007. PubMed PMID: 26967279; PMCID: PMC4788811.
84. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002;295(5558):1306-11. doi: 10.1126/science.1067799. PubMed PMID: 11847345.
85. Deng W, Rupon JW, Krivega I, Breda L, Motta I, Jahn KS, Reik A, Gregory PD, Rivella S, Dean A, Blobel GA. Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell. 2014;158(4):849-60. doi: 10.1016/j.cell.2014.05.050. PubMed PMID: 25126789; PMCID: PMC4134511.
86. Deniz O, Flores O, Aldea M, Soler-Lopez M, Orozco M. Nucleosome architecture throughout the cell cycle. Sci Rep. 2016;6:19729. doi: 10.1038/srep19729. PubMed PMID: 26818620; PMCID: PMC4730144.
87. Deutsch VR, Tomer A. Megakaryocyte development and platelet production. Br J Haematol. 2006;134(5):453-66. doi: 10.1111/j.1365-2141.2006.06215.x. PubMed PMID: 16856888.
88. Dhayat SA, Traeger MM, Rehkaemper J, Stroese AJ, Steinestel K, Wardelmann E, Kabar I, Senninger N. Clinical Impact of Epithelial-to-Mesenchymal Transition Regulating MicroRNAs in Pancreatic Ductal Adenocarcinoma. Cancers (Basel). 2018;10(9). doi: 10.3390/cancers10090328. PubMed PMID: 30217058; PMCID: PMC6162771.
89. Di Gangi IM, Mazza T, Fontana A, Copetti M, Fusilli C, Ippolito A, Mattivi F, Latiano A, Andriulli A, Vrhovsek U, Pazienza V. Metabolomic profile in pancreatic cancer patients: a consensus-based approach to identify highly discriminating metabolites. Oncotarget. 2016;7(5):5815-29. doi: 10.18632/oncotarget.6808. PubMed PMID: 26735340; PMCID: PMC4868723.
90. Diaferia GR, Balestrieri C, Prosperini E, Nicoli P, Spaggiari P, Zerbi A, Natoli G. Dissection of transcriptional and cis-regulatory control of differentiation in human pancreatic cancer. EMBO J. 2016;35(6):595-617. doi: 10.15252/embj.201592404. PubMed PMID: 26769127; PMCID: PMC4801945.
91. Dillon SC, Zhang X, Trievel RC, Cheng X. The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. 2005;6(8):227. doi: 10.1186/gb-2005-6-8-227. PubMed PMID: 16086857; PMCID: PMC1273623.
92. Dincer A, Gavin DP, Xu K, Zhang B, Dudley JT, Schadt EE, Akbarian S. Deciphering H3K4me3 broad domains associated with gene-regulatory networks and conserved epigenomic landscapes in the human brain. Transl Psychiatry. 2015;5:e679. doi: 10.1038/tp.2015.169. PubMed PMID: 26575220; PMCID: PMC5068762.
93. Ding J, Huang X, Shao N, Zhou H, Lee DF, Faiola F, Fidalgo M, Guallar D, Saunders A, Shliaha PV, Wang H, Waghray A, Papatsenko D, Sanchez-Priego C,

Li D, Yuan Y, Lemischka IR, Shen L, Kelley K, Deng H, Shen X, Wang J. Tex10 Coordinates Epigenetic Control of Super-Enhancer Activity in Pluripotency and Reprogramming. Cell Stem Cell. 2015;16(6):653-68. doi: 10.1016/j.stem.2015.04.001. PubMed PMID: 25936917; PMCID: PMC4458159.

94. DiNorcia J, Lee MK, Moroziewicz DN, Winner M, Suman P, Bao F, Remotti HE, Zou YS, Yan SF, Qiu W, Su GH, Schmidt AM, Allendorf JD. RAGE gene deletion inhibits the development and progression of ductal neoplasia and prolongs survival in a murine model of pancreatic cancer. J Gastrointest Surg. 2012;16(1):104-12; discussion 12. doi: 10.1007/s11605-011-1754-9. PubMed PMID: 22052106; PMCID: PMC4049447.
95. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenkov VV, Ecker JR, Thomson JA, Ren B. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015;518(7539):331-6. doi: 10.1038/nature14222. PubMed PMID: 25693564; PMCID: PMC4515363.
96. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485(7398):376-80. doi: 10.1038/nature11082. PubMed PMID: 22495300; PMCID: PMC3356448.
97. Djavani M, Crasta OR, Zhang Y, Zapata JC, Sobral B, Lechner MG, Bryant J, Davis H, Salvato MS. Gene expression in primate liver during viral hemorrhagic fever. Virol J. 2009;6:20. doi: 10.1186/1743-422X-6-20. PubMed PMID: 19216742; PMCID: PMC2657139.
98. Dolcet X, Llobet D, Pallares J, Matias-Guiu X. NF-kB in development and progression of human cancer. Virchows Arch. 2005;446(5):475-82. doi: 10.1007/s00428-005-1264-9. PubMed PMID: 15856292.
99. Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schujiers J, Lee TI, Zhao K, Young RA. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell. 2014;159(2):374-87. doi: 10.1016/j.cell.2014.09.030. PubMed PMID: 25303531; PMCID: PMC4197132.
100. Downs WG, C.R. Anderson, L. Spence, T.H.G. Aitken, and A.H. Greenhall. Tacaribe Virus, a New Agent Isolated from Artibeus Bats and Mosquitos in Trinidad, West Indies. American Journal of Tropical Medicine and Hygiene. 1963;12:640-6.
101. Doyle S, Vaidya S, O'Connell R, Dadgostar H, Dempsey P, Wu T, Rao G, Sun R, Haberland M, Modlin R, Cheng G. IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity. 2002;17(3):251-63. PubMed PMID: 12354379.
102. Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, Ma J, Zhang X, Zhang K, Wang Y, Zhang MQ, Gao J, Dixon JR, Wang X, Zeng J, Xie W. Allelic reprogramming of 3D chromatin architecture during early mammalian development. Nature. 2017;547(7662):232-5. doi: 10.1038/nature23263. PubMed PMID: 28703188.
103. Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, Lee C, Shendure J,

Fields S, Blau CA, Noble WS. A three-dimensional model of the yeast genome. Nature. 2010;465(7296):363-7. doi: 10.1038/nature08973. PubMed PMID: 20436457; PMCID: PMC2874121.

104. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst. 2016;3(1):95-8. doi: 10.1016/j.cels.2016.07.002. PubMed PMID: 27467249; PMCID: PMC5846465.
105. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792-7. doi: 10.1093/nar/gkh340. PubMed PMID: 15034147; PMCID: PMC390337.
106. Eguchi M, Nguyen C, Lee SC, Kahn M. ICG-001, a novel small molecule regulator of TCF/beta-catenin transcription. Med Chem. 2005;1(5):467-72. PubMed PMID: 16787331.
107. Emami KH, Nguyen C, Ma H, Kim DH, Jeong KW, Eguchi M, Moon RT, Teo JL, Kim HY, Moon SH, Ha JR, Kahn M. A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. Proc Natl Acad Sci U S A. 2004;101(34):12682-7. doi: 10.1073/pnas.0404875101. PubMed PMID: 15314234; PMCID: PMC515116.
108. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9(3):215-6. doi: 10.1038/nmeth.1906. PubMed PMID: 22373907; PMCID: PMC3577932.
109. Erwin JA, Lee JT. New twists in X-chromosome inactivation. Curr Opin Cell Biol. 2008;20(3):349-55. doi: 10.1016/j.ceb.2008.04.007. PubMed PMID: 18508252; PMCID: PMC2491718.
110. Fan M, Yan PS, Hartman-Frey C, Chen L, Paik H, Oyer SL, Salisbury JD, Cheng AS, Li L, Abbosh PH, Huang TH, Nephew KP. Diverse gene expression and DNA methylation profiles correlate with differential adaptation of breast cancer cells to the antiestrogens tamoxifen and fulvestrant. Cancer Res. 2006;66(24):11954-66. doi: 10.1158/0008-5472.CAN-06-1666. PubMed PMID: 17178894.
111. Fang F, He X, Deng H, Chen Q, Lu J, Spraul M, Yu Y. Discrimination of metabolic profiles of pancreatic cancer from chronic pancreatitis by high-resolution magic angle spinning 1H nuclear magnetic resonance and principal components analysis. Cancer Sci. 2007;98(11):1678-82. doi: 10.1111/j.1349-7006.2007.00589.x. PubMed PMID: 17727683.
112. Farnham PJ. Insights from genomic profiling of transcription factors. Nat Rev Genet. 2009;10(9):605-16. doi: 10.1038/nrg2636. PubMed PMID: 19668247; PMCID: PMC2846386.
113. Feinberg AP, Koldobskiy MA, Gondor A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet. 2016;17(5):284-99. doi: 10.1038/nrg.2016.13. PubMed PMID: 26972587; PMCID: PMC4888057.
114. Finnin MS, Donigian JR, Cohen A, Richon VM, Rifkind RA, Marks PA, Breslow R, Pavletich NP. Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature. 1999;401(6749):188-93. doi: 10.1038/43710. PubMed PMID: 10490031.

115. Flyamer IM, Gassler J, Imakaev M, Brandao HB, Ulianov SV, Abdennur N, Razin SV, Mirny LA, Tachibana-Konwalski K. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature. 2017;544(7648):110-4. doi: 10.1038/nature21711. PubMed PMID: 28355183; PMCID: PMC5639698.
116. Foley P, Bunyan D, Stratton J, Dillon M, Lynch SA. Further case of Rubinstein-Taybi syndrome due to a deletion in EP300. Am J Med Genet A. 2009;149A(5):997-1000. doi: 10.1002/ajmg.a.32771. PubMed PMID: 19353645.
117. Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schopflin R, Kraft K, Kempfer R, Jerkovic I, Chan WL, Spielmann M, Timmermann B, Wittler L, Kurth I, Cambiaso P, Zuffardi O, Houge G, Lambie L, Brancati F, Pombo A, Vingron M, Spitz F, Mundlos S. Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature. 2016;538(7624):265-9. doi: 10.1038/nature19800. PubMed PMID: 27706140.
118. Franklin RE, Gosling RG. Molecular configuration in sodium thymonucleate. Nature. 1953;171(4356):740-1. PubMed PMID: 13054694.
119. Frasor J, Chang EC, Komm B, Lin CY, Vega VB, Liu ET, Miller LD, Smeds J, Bergh J, Katzenellenbogen BS. Gene expression preferentially regulated by tamoxifen in breast cancer cells and correlations with clinical outcome. Cancer Res. 2006;66(14):7334-40. doi: 10.1158/0008-5472.CAN-05-4269. PubMed PMID: 16849584.
120. Frasor J, Danes JM, Komm B, Chang KC, Lyttle CR, Katzenellenbogen BS. Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. Endocrinology. 2003;144(10):4562-74. doi: 10.1210/en.2003-0567. PubMed PMID: 12959972.
121. Frick WF, Pollock JF, Hicks AC, Langwig KE, Reynolds DS, Turner GG, Butchkoski CM, Kunz TH. An emerging disease causes regional population collapse of a common North American bat species. Science. 2010;329(5992):679-82. Epub 2010/08/07. doi: 329/5992/679 [pii]
2. 10.1126/science.1188594. PubMed PMID: 20689016.
122. Fridman WH. Fc receptors and immunoglobulin binding factors. FASEB J. 1991;5(12):2684-90. PubMed PMID: 1916092.
123. Frietze S, Wang R, Yao L, Tak YG, Ye Z, Gaddis M, Witt H, Farnham PJ, Jin VX. Cell type-specific binding patterns reveal that TCF7L2 can be tethered to the genome by association with GATA3. Genome Biol. 2012;13(9):R52. doi: 10.1186/gb-2012-13-9-r52. PubMed PMID: 22951069; PMCID: PMC3491396.
124. Fritsche P, Seidler B, Schuler S, Schnieke A, Gottlicher M, Schmid RM, Saur D, Schneider G. HDAC2 mediates therapeutic resistance of pancreatic cancer cells via the BH3-only protein NOXA. Gut. 2009;58(10):1399-409. doi: 10.1136/gut.2009.180711. PubMed PMID: 19528037.
125. Fu X, Jeselsohn R, Pereira R, Hollingsworth EF, Creighton CJ, Li F, Shea M, Nardone A, De Angelis C, Heiser LM, Anur P, Wang N, Grasso CS, Spellman PT, Griffith OL, Tsimelzon A, Gutierrez C, Huang S, Edwards DP, Trivedi MV, Rimawi MF, Lopez-Terrada D, Hilsenbeck SG, Gray JW, Brown M, Osborne

CK, Schiff R. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. Proc Natl Acad Sci U S A. 2016;113(43):E6600-E9. doi: 10.1073/pnas.1612835113. PubMed PMID: 27791031; PMCID: PMC5087040.

126. Fulhorst CF, Bowen MD, Salas RA, Duno G, Utrera A, Ksiazek TG, De Manzione NM, De Miller E, Vasquez C, Peters CJ, Tesh RB. Natural rodent host associations of Guanarito and pirital viruses (Family Arenaviridae) in central Venezuela. Am J Trop Med Hyg. 1999;61(2):325-30. Epub 1999/08/27. PubMed PMID: 10463688.
127. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature. 2009;462(7269):58-64. doi: 10.1038/nature08497. PubMed PMID: 19890323; PMCID: PMC2774924.
128. Gabitova L, Gorin A, Astsaturov I. Molecular pathways: sterols and receptor signaling in cancer. Clin Cancer Res. 2014;20(1):28-34. doi: 10.1158/1078-0432.CCR-13-0122. PubMed PMID: 24158702; PMCID: PMC3859141.
129. Gaddis M, Gerrard D, Frietze S, Farnham PJ. Altering cancer transcriptomes using epigenomic inhibitors. Epigenetics Chromatin. 2015;8:9. doi: 10.1186/1756-8935-8-9. PubMed PMID: 26191083; PMCID: PMC4506402.
130. Gall JG. Kinetics of deoxyribonuclease action on chromosomes. Nature. 1963;198:36-8. PubMed PMID: 14010299.
131. Gao XN, Lin J, Ning QY, Gao L, Yao YS, Zhou JH, Li YH, Wang LL, Yu L. A histone acetyltransferase p300 inhibitor C646 induces cell cycle arrest and apoptosis selectively in AML1-ETO-positive AML cells. PLoS One. 2013;8(2):e55481. doi: 10.1371/journal.pone.0055481. PubMed PMID: 23390536; PMCID: PMC3563640.
132. Gargas A, M.T. Trest, M. Christensen, T.J. Volk, D. S. Blehert. Geomyces destructans sp. nov. associated with bat white-nose syndrome. Mycotaxon. 2009;108:147-54.
133. Gerrard DL, Wang Y, Gaddis M, Zhou Y, Wang J, Witt H, Lin S, Farnham PJ, Jin VX, Frietze SE. Three-dimensional analysis reveals altered chromatin interaction by enhancer inhibitors harbors TCF7L2-regulated cancer gene signature. J Cell Biochem. 2019;120(3):3056-70. doi: 10.1002/jcb.27449. PubMed PMID: 30548288.
134. Gerstein MB, Kundaje A, Hariharan M, Landt SG, Yan KK, Cheng C, Mu XJ, Khurana E, Rozowsky J, Alexander R, Min R, Alves P, Abyzov A, Addleman N, Bhardwaj N, Boyle AP, Cayting P, Charos A, Chen DZ, Cheng Y, Clarke D, Eastman C, Euskirchen G, Frietze S, Fu Y, Gertz J, Grubert F, Harmanci A, Jain P, Kasowski M, Lacroute P, Leng JJ, Lian J, Monahan H, O'Geen H, Ouyang Z, Partridge EC, Patacsil D, Pauli F, Raha D, Ramirez L, Reddy TE, Reed B, Shi M, Slifer T, Wang J, Wu L, Yang X, Yip KY, Zilberman-Schapira G, Batzoglou S,

Sidow A, Farnham PJ, Myers RM, Weissman SM, Snyder M. Architecture of the human regulatory network derived from ENCODE data. Nature. 2012;489(7414):91-100. doi: 10.1038/nature11245. PubMed PMID: 22955619; PMCID: PMC4154057.

135. Geutjes EJ, Bajpe PK, Bernards R. Targeting the epigenome for treatment of cancer. Oncogene. 2012;31(34):3827-44. doi: 10.1038/onc.2011.552. PubMed PMID: 22139071.
136. GF Gunnell NS. Fossil evidence and the Origin of Bats. Journal of Mammalian Evolution. 2005;12:209-46.
137. Giles RH, Peters DJ, Breuning MH. Conjunction dysfunction: CBP/p300 in human disease. Trends Genet. 1998;14(5):178-83. PubMed PMID: 9613201.
138. Giroux V, Malicet C, Barthet M, Gironella M, Archange C, Dagorn JC, Vasseur S, Iovanna JL. p8 is a new target of gemcitabine in pancreatic cancer cells. Clin Cancer Res. 2006;12(1):235-41. doi: 10.1158/1078-0432.CCR-05-1700. PubMed PMID: 16397047.
139. Gondor A, Ohlsson R. Chromosome crosstalk in three dimensions. Nature. 2009;461(7261):212-7. doi: 10.1038/nature08453. PubMed PMID: 19741702.
140. Goodwin GH, Sanders C, Johns EW. A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. Eur J Biochem. 1973;38(1):14-9. PubMed PMID: 4774120.
141. Gorin A, Gabitova L, Astsaturov I. Regulation of cholesterol biosynthesis and cancer signaling. Curr Opin Pharmacol. 2012;12(6):710-6. doi: 10.1016/j.coph.2012.06.011. PubMed PMID: 22824431; PMCID: PMC3504641.
142. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 2008;36(10):3420-35. doi: 10.1093/nar/gkn176. PubMed PMID: 18445632; PMCID: 2425479.
143. Goubau D, Deddouche S, Reis e Sousa C. Cytosolic sensing of viruses. Immunity. 2013;38(5):855-69. doi: 10.1016/j.immuni.2013.05.007. PubMed PMID: 23706667.
144. Gower WR, Jr., Risch RM, Godellas CV, Fabri PJ. HPAC, a new human glucocorticoid-sensitive pancreatic ductal adenocarcinoma cell line. In Vitro Cell Dev Biol Anim. 1994;30A(3):151-61. PubMed PMID: 25939163.
145. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644-52. doi: 10.1038/nbt.1883. PubMed PMID: 21572440; PMCID: 3571712.
146. Grunstein M. Histone acetylation in chromatin structure and transcription. Nature. 1997;389(6649):349-52. doi: 10.1038/38664. PubMed PMID: 9311776.
147. Gu F, Hsu HK, Hsu PY, Wu J, Ma Y, Parvin J, Huang TH, Jin VX. Inference of hierarchical regulatory network of estrogen-dependent breast cancer through ChIP-based data. BMC Syst Biol. 2010;4:170. doi: 10.1186/1752-0509-4-170. PubMed PMID: 21167036; PMCID: PMC3012048.

148. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, van Steensel B. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature. 2008;453(7197):948-51. doi: 10.1038/nature06947. PubMed PMID: 18463634.
149. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8(8):1494-512. doi: 10.1038/nprot.2013.084. PubMed PMID: 23845962; PMCID: 3875132.
150. Hagege H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, de Laat W, Forne T. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc. 2007;2(7):1722-33. doi: 10.1038/nprot.2007.243. PubMed PMID: 17641637.
151. Hagstrom KA, Meyer BJ. Condensin and cohesin: more than chromosome compactor and glue. Nat Rev Genet. 2003;4(7):520-34. doi: 10.1038/nrg1110. PubMed PMID: 12838344.
152. Hamidi T, Algul H, Cano CE, Sandi MJ, Molejon MI, Riemann M, Calvo EL, Lomberk G, Dagorn JC, Weih F, Urrutia R, Schmid RM, Iovanna JL. Nuclear protein 1 promotes pancreatic cancer development and protects cells from stress by inhibiting apoptosis. J Clin Invest. 2012;122(6):2092-103. doi: 10.1172/JCI60144. PubMed PMID: 22565310; PMCID: PMC3366404.
153. Hanada T, Yoshimura A. Regulation of cytokine signaling and inflammation. Cytokine Growth Factor Rev. 2002;13(4-5):413-21. PubMed PMID: 12220554.
154. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74. doi: 10.1016/j.cell.2011.02.013. PubMed PMID: 21376230.
155. Hatzis P, Talianidis I. Dynamics of enhancer-promoter communication during differentiation-induced gene activation. Mol Cell. 2002;10(6):1467-77. PubMed PMID: 12504020.
156. Heidari N, Phanstiel DH, He C, Grubert F, Jahanbani F, Kasowski M, Zhang MQ, Snyder MP. Genome-wide map of regulatory interactions in the human genome. Genome Res. 2014;24(12):1905-17. doi: 10.1101/gr.176586.114. PubMed PMID: 25228660; PMCID: PMC4248309.
157. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007;39(3):311-8. doi: 10.1038/ng1966. PubMed PMID: 17277777.
158. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89. doi: 10.1016/j.molcel.2010.05.004. PubMed PMID: 20513432; PMCID: PMC2898526.
159. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell

type-specific enhancers. Nat Rev Mol Cell Biol. 2015;16(3):144-54. doi: 10.1038/nrm3949. PubMed PMID: 25650801; PMCID: PMC4517609.
160. Heldring N, Pike A, Andersson S, Matthews J, Cheng G, Hartman J, Tujague M, Strom A, Treuter E, Warner M, Gustafsson JA. Estrogen receptors: how do they signal and what are their targets. Physiol Rev. 2007;87(3):905-31. doi: 10.1152/physrev.00026.2006. PubMed PMID: 17615392.
161. Helin K, Dhanak D. Chromatin proteins and modifications as drug targets. Nature. 2013;502(7472):480-8. doi: 10.1038/nature12751. PubMed PMID: 24153301.
162. Henikoff S, Smith MM. Histone variants and epigenetics. Cold Spring Harb Perspect Biol. 2015;7(1):a019364. doi: 10.1101/cshperspect.a019364. PubMed PMID: 25561719; PMCID: PMC4292162.
163. Herbertz S, Sawyer JS, Stauber AJ, Gueorguieva I, Driscoll KE, Estrem ST, Cleverly AL, Desaiah D, Guba SC, Benhadji KA, Slapak CA, Lahn MM. Clinical development of galunisertib (LY2157299 monohydrate), a small molecule inhibitor of transforming growth factor-beta signaling pathway. Drug Des Devel Ther. 2015;9:4479-99. doi: 10.2147/DDDT.S86621. PubMed PMID: 26309397; PMCID: PMC4539082.
164. Hessmann E, Schneider G, Ellenrieder V, Siveke JT. MYC in pancreatic cancer: novel mechanistic insights and their translation into therapeutic strategies. Oncogene. 2016;35(13):1609-18. doi: 10.1038/onc.2015.216. PubMed PMID: 26119937.
165. Hiller M, Agarwal S, Notwell JH, Parikh R, Guturu H, Wenger AM, Bejerano G. Computational methods to detect conserved non-genic elements in phylogenetically isolated genomes: application to zebrafish. Nucleic Acids Res. 2013;41(15):e151. doi: 10.1093/nar/gkt557. PubMed PMID: 23814184; PMCID: PMC3753653.
166. Hirota K, Miyoshi T, Kugou K, Hoffman CS, Shibata T, Ohta K. Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature. 2008;456(7218):130-4. doi: 10.1038/nature07348. PubMed PMID: 18820678.
167. Hirsch RL. The complement system: its importance in the host response to viral infection. Microbiol Rev. 1982;46(1):71-85. PubMed PMID: 7045625; PMCID: PMC373211.
168. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. Super-enhancers in the control of cell identity and disease. Cell. 2013;155(4):934-47. doi: 10.1016/j.cell.2013.09.053. PubMed PMID: 24119843; PMCID: PMC3841062.
169. Hoffman EA, Frey BL, Smith LM, Auble DT. Formaldehyde crosslinking: a tool for the study of chromatin complexes. J Biol Chem. 2015;290(44):26404-11. doi: 10.1074/jbc.R115.651679. PubMed PMID: 26354429; PMCID: PMC4646298.
170. Honkela A, Peltonen J, Topa H, Charapitsa I, Matarese F, Grote K, Stunnenberg HG, Reid G, Lawrence ND, Rattray M. Genome-wide modeling of transcription kinetics reveals patterns of RNA production delays. Proc Natl Acad Sci U S A. 2015;112(42):13115-20. doi: 10.1073/pnas.1420404112. PubMed PMID:

26438844; PMCID: PMC4620908.
171. Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol Cell. 2012;48(3):471-84. doi: 10.1016/j.molcel.2012.08.031. PubMed PMID: 23041285; PMCID: PMC3496039.
172. Hsu PY, Hsu HK, Hsiao TH, Ye Z, Wang E, Profit AL, Jatoi I, Chen Y, Kirma NB, Jin VX, Sharp ZD, Huang TH. Spatiotemporal control of estrogen-responsive transcription in ERalpha-positive breast cancer cells. Oncogene. 2016;35(18):2379-89. doi: 10.1038/onc.2015.298. PubMed PMID: 26300005; PMCID: PMC4865474.
173. Hsu PY, Hsu HK, Lan X, Juan L, Yan PS, Labanowska J, Heerema N, Hsiao TH, Chiu YC, Chen Y, Liu Y, Li L, Li R, Thompson IM, Nephew KP, Sharp ZD, Kirma NB, Jin VX, Huang TH. Amplification of distant estrogen response elements deregulates target genes associated with tamoxifen resistance in breast cancer. Cancer Cell. 2013;24(2):197-212. doi: 10.1016/j.ccr.2013.07.007. PubMed PMID: 23948299; PMCID: PMC3890247.
174. Huang J, Li K, Cai W, Liu X, Zhang Y, Orkin SH, Xu J, Yuan GC. Dissecting super-enhancer hierarchy based on chromatin interactions. Nat Commun. 2018;9(1):943. doi: 10.1038/s41467-018-03279-9. PubMed PMID: 29507293; PMCID: PMC5838163.
175. Huberman JA. Structure of chromosome fibers and chromosomes. Annu Rev Biochem. 1973;42:355-78. doi: 10.1146/annurev.bi.42.070173.002035. PubMed PMID: 4581228.
176. Hubner MR, Eckersley-Maslin MA, Spector DL. Chromatin organization and transcriptional regulation. Curr Opin Genet Dev. 2013;23(2):89-95. doi: 10.1016/j.gde.2012.11.006. PubMed PMID: 23270812; PMCID: PMC3612554.
177. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001;17(8):754-5. PubMed PMID: 11524383.
178. Hung HL, Kim AY, Hong W, Rakowski C, Blobel GA. Stimulation of NF-E2 DNA binding by CREB-binding protein (CBP)-mediated acetylation. J Biol Chem. 2001;276(14):10715-21. doi: 10.1074/jbc.M007846200. PubMed PMID: 11154691.
179. Hurtado A, Holmes KA, Geistlinger TR, Hutcheson IR, Nicholson RI, Brown M, Jiang J, Howat WJ, Ali S, Carroll JS. Regulation of ERBB2 by oestrogen receptor-PAX2 determines response to tamoxifen. Nature. 2008;456(7222):663-6. doi: 10.1038/nature07483. PubMed PMID: 19005469; PMCID: PMC2920208.
180. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat Genet. 2011;43(1):27-33. doi: 10.1038/ng.730. PubMed PMID: 21151129; PMCID: PMC3024537.
181. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods. 2012;9(10):999-1003. doi: 10.1038/nmeth.2148. PubMed PMID: 22941365; PMCID: PMC3816492.
182. IUCN. IUCN Red List Version 2015.1, Table 3a- Status Category summary by

major taxonomic group (animals): Conservation International, Arizona State University, Texas A&M University, University of Rome, University of Virginia, Zoological Society London. ; 2015 [cited 2015 June 11, 2015]. Available from: www.iucnredlist.org/mammals.
183. IUCN. IUCN Red List Version 2015.1, Table 4a- Number of species in each IUCN Red List Category in each major animal taxonomic group (class, order): Conservation International, Arizona State University, Texas A&M University, University of Rome, University of Virginia, Zoological Society London. ; 2015 [cited 2015 June 11, 2015]. Available from: www.iucnredlist.org/mammals.
184. Jayme SI, Field HE, de Jong C, Olival KJ, Marsh G, Tagtag AM, Hughes T, Bucad AC, Barr J, Azul RR, Retes LM, Foord A, Yu M, Cruz MS, Santos IJ, Lim TM, Benigno CC, Epstein JH, Wang LF, Daszak P, Newman SH. Molecular evidence of Ebola Reston virus infection in Philippine bats. Virol J. 2015;12:107. doi: 10.1186/s12985-015-0331-3. PubMed PMID: 26184657; PMCID: 4504098.
185. Jenuwein T, Allis CD. Translating the histone code. Science. 2001;293(5532):1074-80. doi: 10.1126/science.1063127. PubMed PMID: 11498575.
186. Ji X, Dadon DB, Powell BE, Fan ZP, Borges-Rivera D, Shachar S, Weintraub AS, Hnisz D, Pegoraro G, Lee TI, Misteli T, Jaenisch R, Young RA. 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. Cell Stem Cell. 2016;18(2):262-75. doi: 10.1016/j.stem.2015.11.007. PubMed PMID: 26686465; PMCID: PMC4848748.
187. Jia J, Parikh H, Xiao W, Hoskins JW, Pflicke H, Liu X, Collins I, Zhou W, Wang Z, Powell J, Thorgeirsson SS, Rudloff U, Petersen GM, Amundadottir LT. An integrated transcriptome and epigenome analysis identifies a novel candidate gene for pancreatic cancer. BMC Med Genomics. 2013;6:33. doi: 10.1186/1755-8794-6-33. PubMed PMID: 24053169; PMCID: PMC3849454.
188. Jiang H, Hegde S, Knolhoff BL, Zhu Y, Herndon JM, Meyer MA, Nywening TM, Hawkins WG, Shapiro IM, Weaver DT, Pachter JA, Wang-Gillam A, DeNardo DG. Targeting focal adhesion kinase renders pancreatic cancers responsive to checkpoint immunotherapy. Nat Med. 2016;22(8):851-60. doi: 10.1038/nm.4123. PubMed PMID: 27376576; PMCID: PMC4935930.
189. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen CA, Schmitt AD, Espinoza CA, Ren B. A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature. 2013;503(7475):290-4. doi: 10.1038/nature12644. PubMed PMID: 24141950; PMCID: PMC3838900.
190. Johns EW, Forrester S. Studies on nuclear proteins. The binding of extra acidic proteins to deoxyribonucleoprotein during the preparation of nuclear proteins. Eur J Biochem. 1969;8(4):547-51. PubMed PMID: 5796141.
191. Jones PA, Issa JP, Baylin S. Targeting the cancer epigenome for therapy. Nat Rev Genet. 2016;17(10):630-41. doi: 10.1038/nrg.2016.93. PubMed PMID: 27629931.
192. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Kamiyama H, Jimeno A, Hong SM, Fu B, Lin MT, Calhoun ES, Kamiyama M, Walter K, Nikolskaya T, Nikolsky Y, Hartigan J, Smith DR,

Hidalgo M, Leach SD, Klein AP, Jaffee EM, Goggins M, Maitra A, Iacobuzio-Donahue C, Eshleman JR, Kern SE, Hruban RH, Karchin R, Papadopoulos N, Parmigiani G, Vogelstein B, Velculescu VE, Kinzler KW. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008;321(5897):1801-6. doi: 10.1126/science.1164368. PubMed PMID: 18772397; PMCID: PMC2848990.

193. Jordan VC. A century of deciphering the control mechanisms of sex steroid action in breast and prostate cancer: the origins of targeted therapy and chemoprevention. Cancer Res. 2009;69(4):1243-54. doi: 10.1158/0008-5472.CAN-09-0029. PubMed PMID: 19208829.
194. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat Biotechnol. 2011;30(1):90-8. doi: 10.1038/nbt.2057. PubMed PMID: 22198700; PMCID: PMC3782096.
195. Kalkhoven E. CBP and p300: HATs for different occasions. Biochem Pharmacol. 2004;68(6):1145-55. doi: 10.1016/j.bcp.2004.03.045. PubMed PMID: 15313412.
196. Kang R, Tang D, Schapiro NE, Loux T, Livesey KM, Billiar TR, Wang H, Van Houten B, Lotze MT, Zeh HJ. The HMGB1/RAGE inflammatory pathway promotes pancreatic tumor growth by regulating mitochondrial bioenergetics. Oncogene. 2014;33(5):567-77. doi: 10.1038/onc.2012.631. PubMed PMID: 23318458; PMCID: PMC3795800.
197. Kaniskan HU, Konze KD, Jin J. Selective inhibitors of protein methyltransferases. J Med Chem. 2015;58(4):1596-629. doi: 10.1021/jm501234a. PubMed PMID: 25406853; PMCID: PMC4345896.
198. Kawai T, Takeuchi O, Fujita T, Inoue J, Muhlradt PF, Sato S, Hoshino K, Akira S. Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes. J Immunol. 2001;167(10):5887-94. PubMed PMID: 11698465.
199. Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, Chen H, Du Z, Xie W, Xu X, Huang X, Liu J. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. Cell. 2017;170(2):367-81 e20. doi: 10.1016/j.cell.2017.06.029. PubMed PMID: 28709003.
200. Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics. 2010;26(17):2204-7. doi: 10.1093/bioinformatics/btq351. PubMed PMID: 20639541; PMCID: PMC2922891.
201. Keshet I, Lieman-Hurwitz J, Cedar H. DNA methylation affects the formation of active chromatin. Cell. 1986;44(4):535-43. PubMed PMID: 3456276.
202. Ketscher L, Hannss R, Morales DJ, Basters A, Guerra S, Goldmann T, Hausmann A, Prinz M, Naumann R, Pekosz A, Utermohlen O, Lenschow DJ, Knobeloch KP. Selective inactivation of USP18 isopeptidase activity in vivo enhances ISG15 conjugation and viral resistance. Proc Natl Acad Sci U S A. 2015;112(5):1577-82. doi: 10.1073/pnas.1412881112. PubMed PMID: 25605921; PMCID: 4321242.

203. Khan A, Mathelier A, Zhang X. Super-enhancers are transcriptionally more active and cell type-specific than stretch enhancers. Epigenetics. 2018;13(9):910-22. doi: 10.1080/15592294.2018.1514231. PubMed PMID: 30169995; PMCID: PMC6284781.
204. Khan SA, Reddy D, Gupta S. Global histone post-translational modifications and cancer: Biomarkers for diagnosis, prognosis and treatment? World J Biol Chem. 2015;6(4):333-45. doi: 10.4331/wjbc.v6.i4.333. PubMed PMID: 26629316; PMCID: PMC4657128.
205. Khorasanizadeh S. The nucleosome: from genomic organization to genomic regulation. Cell. 2004;116(2):259-72. PubMed PMID: 14744436.
206. Kimmins S, Sassone-Corsi P. Chromatin remodelling and epigenetic features of germ cells. Nature. 2005;434(7033):583-9. doi: 10.1038/nature03368. PubMed PMID: 15800613.
207. Koenig A, Linhart T, Schlengemann K, Reutlinger K, Wegele J, Adler G, Singh G, Hofmann L, Kunsch S, Buch T, Schafer E, Gress TM, Fernandez-Zapico ME, Ellenrieder V. NFAT-induced histone acetylation relay switch promotes c-Myc-dependent growth in pancreatic cancer cells. Gastroenterology. 2010;138(3):1189-99 e1-2. doi: 10.1053/j.gastro.2009.10.045. PubMed PMID: 19900447; PMCID: PMC2895621.
208. Korkut A, Zaidi S, Kanchi RS, Rao S, Gough NR, Schultz A, Li X, Lorenzi PL, Berger AC, Robertson G, Kwong LN, Datto M, Roszik J, Ling S, Ravikumar V, Manyam G, Rao A, Shelley S, Liu Y, Ju Z, Hansel D, de Velasco G, Pennathur A, Andersen JB, O'Rourke CJ, Ohshiro K, Jogunoori W, Nguyen BN, Li S, Osmanbeyoglu HU, Ajani JA, Mani SA, Houseman A, Wiznerowicz M, Chen J, Gu S, Ma W, Zhang J, Tong P, Cherniack AD, Deng C, Resar L, Cancer Genome Atlas Research N, Weinstein JN, Mishra L, Akbani R. A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF-beta Superfamily. Cell Syst. 2018;7(4):422-37 e7. doi: 10.1016/j.cels.2018.08.010. PubMed PMID: 30268436.
209. Koster R, di Pietro A, Timmer-Bosscha H, Gibcus JH, van den Berg A, Suurmeijer AJ, Bischoff R, Gietema JA, de Jong S. Cytoplasmic p21 expression levels determine cisplatin resistance in human testicular cancer. J Clin Invest. 2010;120(10):3594-605. doi: 10.1172/JCI41939. PubMed PMID: 20811155; PMCID: PMC2947220.
210. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693-705. doi: 10.1016/j.cell.2007.02.005. PubMed PMID: 17320507.
211. Kreth G, Finsterle J, von Hase J, Cremer M, Cremer C. Radial arrangement of chromosome territories in human cell nuclei: a computer model approach based on gene density indicates a probabilistic global positioning code. Biophys J. 2004;86(5):2803-12. doi: 10.1016/S0006-3495(04)74333-7. PubMed PMID: 15111398; PMCID: PMC1304150.
212. Krumlauf R. Hox genes in vertebrate development. Cell. 1994;78(2):191-201. PubMed PMID: 7913880.
213. Kumar R, Li DQ, Muller S, Knapp S. Epigenomic regulation of oncogenesis by

chromatin remodeling. Oncogene. 2016;35(34):4423-36. doi: 10.1038/onc.2015.513. PubMed PMID: 26804164.
214. Kuno G. Persistence of arboviruses and antiviral antibodies in vertebrate hosts: its occurrence and impacts. Rev Med Virol. 2001;11(3):165-90. PubMed PMID: 11376480.
215. Kunz TH, and Linda F. Lumsden. Bat Ecology. Kunz TH, and M. Brock Fenton, editor. Chicago: The University of Chicago Press; 2003. 3-89 p.
216. Kyriazis AA, Kyriazis AP, Sternberg CN, Sloane NH, Loveless JD. Morphological, biological, biochemical, and karyotypic characteristics of human pancreatic ductal adenocarcinoma Capan-2 in tissue culture and the nude mouse. Cancer Res. 1986;46(11):5810-5. PubMed PMID: 3019537.
217. Kyriazis AP, Kyriazis AA, Scarpelli DG, Fogh J, Rao MS, Lepera R. Human pancreatic adenocarcinoma line Capan-1 in tissue culture and the nude mouse: morphologic, biologic, and biochemical characteristics. Am J Pathol. 1982;106(2):250-60. PubMed PMID: 6278935; PMCID: PMC1916189.
218. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature. 2013;494(7438):497-501. doi: 10.1038/nature11884. PubMed PMID: 23417068; PMCID: PMC4109059.
219. Lam QL, Lu L. Role of leptin in immunity. Cell Mol Immunol. 2007;4(1):1-13. PubMed PMID: 17349207.
220. Lamond AI, Earnshaw WC. Structure and function in the nucleus. Science. 1998;280(5363):547-53. PubMed PMID: 9554838.
221. Lan X, Farnham PJ, Jin VX. Uncovering transcription factor modules using one- and three-dimensional analyses. J Biol Chem. 2012;287(37):30914-21. doi: 10.1074/jbc.R111.309229. PubMed PMID: 22952238; PMCID: PMC3438924.
222. Lan X, Witt H, Katsumura K, Ye Z, Wang Q, Bresnick EH, Farnham PJ, Jin VX. Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages. Nucleic Acids Res. 2012;40(16):7690-704. doi: 10.1093/nar/gks501. PubMed PMID: 22675074; PMCID: PMC3439894.
223. Lanczky A, Nagy A, Bottai G, Munkacsy G, Szabo A, Santarpia L, Gyorffy B. miRpower: a web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. Breast Cancer Res Treat. 2016;160(3):439-46. doi: 10.1007/s10549-016-4013-7. PubMed PMID: 27744485.
224. Lanfear R, Calcott B, Ho SY, Guindon S. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 2012;29(6):1695-701. doi: 10.1093/molbev/mss020. PubMed PMID: 22319168.
225. Langevin C, van der Aa LM, Houel A, Torhy C, Briolat V, Lunazzi A, Harmache A, Bremont M, Levraud JP, Boudinot P. Zebrafish ISG15 exerts a strong antiviral activity against RNA and DNA viruses and regulates the interferon response. J Virol. 2013;87(18):10025-36. doi: 10.1128/JVI.01294-12. PubMed PMID: 23824820; PMCID: 3753986.
226. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat

Methods. 2012;9(4):357-9. doi: 10.1038/nmeth.1923. PubMed PMID: 22388286; PMCID: PMC3322381.
227. Lanier LL. DAP10- and DAP12-associated receptors in innate immunity. Immunol Rev. 2009;227(1):150-60. doi: 10.1111/j.1600-065X.2008.00720.x. PubMed PMID: 19120482; PMCID: 2794881.
228. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH, Pugh TJ, Stransky N, Helman E, Kim J, Sougnez C, Ambrogio L, Nickerson E, Shefler E, Cortes ML, Auclair D, Saksena G, Voet D, Noble M, DiCara D, Lin P, Lichtenstein L, Heiman DI, Fennell T, Imielinski M, Hernandez B, Hodis E, Baca S, Dulak AM, Lohr J, Landau DA, Wu CJ, Melendez-Zajgla J, Hidalgo-Miranda A, Koren A, McCarroll SA, Mora J, Crompton B, Onofrio R, Parkin M, Winckler W, Ardlie K, Gabriel SB, Roberts CWM, Biegel JA, Stegmaier K, Bass AJ, Garraway LA, Meyerson M, Golub TR, Gordenin DA, Sunyaev S, Lander ES, Getz G. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013;499(7457):214-8. doi: 10.1038/nature12213. PubMed PMID: 23770567; PMCID: PMC3919509.
229. Le Dily F, Bau D, Pohl A, Vicent GP, Serra F, Soronellas D, Castellano G, Wright RH, Ballare C, Filion G, Marti-Renom MA, Beato M. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes Dev. 2014;28(19):2151-62. doi: 10.1101/gad.241422.114. PubMed PMID: 25274727; PMCID: PMC4180976.
230. Le Y, Zhou Y, Iribarren P, Wang J. Chemokines and chemokine receptors: their manifold roles in homeostasis and disease. Cell Mol Immunol. 2004;1(2):95-104. PubMed PMID: 16212895.
231. Lenschow DJ, Walunas TL, Bluestone JA. CD28/B7 system of T cell costimulation. Annu Rev Immunol. 1996;14:233-58. doi: 10.1146/annurev.immunol.14.1.233. PubMed PMID: 8717514.
232. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell. 2007;128(4):707-19. doi: 10.1016/j.cell.2007.01.015. PubMed PMID: 17320508.
233. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, Sim HS, Peh SQ, Mulawadi FH, Ong CT, Orlov YL, Hong S, Zhang Z, Landt S, Raha D, Euskirchen G, Wei CL, Ge W, Wang H, Davis C, Fisher-Aylor KI, Mortazavi A, Gerstein M, Gingeras T, Wold B, Sun Y, Fullwood MJ, Cheung E, Liu E, Sung WK, Snyder M, Ruan Y. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 2012;148(1-2):84-98. doi: 10.1016/j.cell.2011.12.014. PubMed PMID: 22265404; PMCID: PMC3339270.
234. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60. doi: 10.1093/bioinformatics/btp324. PubMed PMID: 19451168; PMCID: PMC2705234.
235. Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong CT, Cubenas-Potts C, Hu M, Lei EP, Bosco G, Qin ZS, Corces VG. Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing.

Mol Cell. 2015;58(2):216-31. doi: 10.1016/j.molcel.2015.02.023. PubMed PMID: 25818644; PMCID: PMC4402144.
236. Li W, Gong K, Li Q, Alber F, Zhou XJ. Hi-Corrector: a fast, scalable and memory-efficient package for normalizing large-scale Hi-C data. Bioinformatics. 2015;31(6):960-2. doi: 10.1093/bioinformatics/btu747. PubMed PMID: 25391400; PMCID: PMC4380031.
237. Libbrecht MW, Ay F, Hoffman MM, Gilbert DM, Bilmes JA, Noble WS. Joint annotation of chromatin state and chromatin conformation reveals relationships among domain types and identifies domains of cell-type-specific expression. Genome Res. 2015;25(4):544-57. doi: 10.1101/gr.184341.114. PubMed PMID: 25677182; PMCID: PMC4381526.
238. Lieber M, Mazzetta J, Nelson-Rees W, Kaplan M, Todaro G. Establishment of a continuous tumor-cell line (panc-1) from a human carcinoma of the exocrine pancreas. Int J Cancer. 1975;15(5):741-7. PubMed PMID: 1140870.
239. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289-93. doi: 10.1126/science.1181369. PubMed PMID: 19815776; PMCID: PMC2858594.
240. Lin CY, Vega VB, Thomsen JS, Zhang T, Kong SL, Xie M, Chiu KP, Lipovich L, Barnett DH, Stossi F, Yeo A, George J, Kuznetsov VA, Lee YK, Charn TH, Palanisamy N, Miller LD, Cheung E, Katzenellenbogen BS, Ruan Y, Bourque G, Wei CL, Liu ET. Whole-genome cartography of estrogen receptor alpha binding sites. PLoS Genet. 2007;3(6):e87. doi: 10.1371/journal.pgen.0030087. PubMed PMID: 17542648; PMCID: PMC1885282.
241. Lin JY, Hu GB, Liu DH, Li S, Liu QM, Zhang SC. Molecular cloning and expression analysis of interferon stimulated gene 15 (ISG15) in turbot, Scophthalmus maximus. Fish Shellfish Immunol. 2015;45(2):895-900. doi: 10.1016/j.fsi.2015.05.050. PubMed PMID: 26095010.
242. Litt M, Qiu Y, Huang S. Histone arginine methylations: their roles in chromatin dynamics and transcriptional regulation. Biosci Rep. 2009;29(2):131-41. doi: 10.1042/BSR20080176. PubMed PMID: 19220199; PMCID: PMC5433800.
243. Litzenburger UM, Buenrostro JD, Wu B, Shen Y, Sheffield NC, Kathiria A, Greenleaf WJ, Chang HY. Single-cell epigenomic variability reveals functional cancer heterogeneity. Genome Biol. 2017;18(1):15. doi: 10.1186/s13059-016-1133-7. PubMed PMID: 28118844; PMCID: PMC5259890.
244. Liu Y, Chen H, Zheng P, Zheng Y, Luo Q, Xie G, Ma Y, Shen L. ICG-001 suppresses growth of gastric cancer cells and reduces chemoresistance of cancer stem cell-like population. J Exp Clin Cancer Res. 2017;36(1):125. doi: 10.1186/s13046-017-0595-0. PubMed PMID: 28893318; PMCID: PMC5594604.
245. Logsdon CD, Fuentes MK, Huang EH, Arumugam T. RAGE and RAGE ligands in cancer. Curr Mol Med. 2007;7(8):777-89. PubMed PMID: 18331236.
246. Lomberk G, Blum Y, Nicolle R, Nair A, Gaonkar KS, Marisa L, Mathison A, Sun Z, Yan H, Elarouci N, Armenoult L, Ayadi M, Ordog T, Lee JH, Oliver G, Klee

E, Moutardier V, Gayet O, Bian B, Duconseil P, Gilabert M, Bigonnet M, Garcia S, Turrini O, Delpero JR, Giovannini M, Grandval P, Gasmi M, Secq V, De Reynies A, Dusetti N, Iovanna J, Urrutia R. Distinct epigenetic landscapes underlie the pathobiology of pancreatic cancer subtypes. Nat Commun. 2018;9(1):1978. doi: 10.1038/s41467-018-04383-6. PubMed PMID: 29773832; PMCID: PMC5958058.

247. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PMCID: PMC4302049.

248. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;153(2):320-34. doi: 10.1016/j.cell.2013.03.036. PubMed PMID: 23582323; PMCID: PMC3760967.

249. Lu P, Takai K, Weaver VM, Werb Z. Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harb Perspect Biol. 2011;3(12). doi: 10.1101/cshperspect.a005058. PubMed PMID: 21917992; PMCID: 3225943.

250. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389(6648):251-60. doi: 10.1038/38444. PubMed PMID: 9305837.

251. Luis AD, Hayman DT, O'Shea TJ, Cryan PM, Gilbert AT, Pulliam JR, Mills JN, Timonin ME, Willis CK, Cunningham AA, Fooks AR, Rupprecht CE, Wood JL, Webb CT. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proc Biol Sci. 2013;280(1756):20122753. doi: 10.1098/rspb.2012.2753. PubMed PMID: 23378666; PMCID: 3574368.

252. Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, Santos-Simarro F, Gilbert-Dussardier B, Wittler L, Borschiwer M, Haas SA, Osterwalder M, Franke M, Timmermann B, Hecht J, Spielmann M, Visel A, Mundlos S. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell. 2015;161(5):1012-25. doi: 10.1016/j.cell.2015.04.004. PubMed PMID: 25959774; PMCID: PMC4791538.

253. Luu HH, Zhang R, Haydon RC, Rayburn E, Kang Q, Si W, Park JK, Wang H, Peng Y, Jiang W, He TC. Wnt/beta-catenin signaling pathway as a novel cancer drug target. Curr Cancer Drug Targets. 2004;4(8):653-71. PubMed PMID: 15578921.

254. Ma H, Nguyen C, Lee KS, Kahn M. Differential roles for the coactivators CBP and p300 on TCF/beta-catenin-mediated survivin gene expression. Oncogene. 2005;24(22):3619-31. doi: 10.1038/sj.onc.1208433. PubMed PMID: 15782138.

255. Mack SC, Pajtler KW, Chavez L, Okonechnikov K, Bertrand KC, Wang X, Erkek S, Federation A, Song A, Lee C, Wang X, McDonald L, Morrow JJ, Saiakhova A, Sin-Chan P, Wu Q, Michaelraj KA, Miller TE, Hubert CG, Ryzhova M, Garzia L, Donovan L, Dombrowski S, Factor DC, Luu B, Valentim CLL, Gimple RC, Morton A, Kim L, Prager BC, Lee JJY, Wu X, Zuccaro J, Thompson Y, Holgado BL, Reimand J, Ke SQ, Tropper A, Lai S, Vijayarajah S, Doan S, Mahadev V,

Minan AF, Grobner SN, Lienhard M, Zapatka M, Huang Z, Aldape KD, Carcaboso AM, Houghton PJ, Keir ST, Milde T, Witt H, Li Y, Li CJ, Bian XW, Jones DTW, Scott I, Singh SK, Huang A, Dirks PB, Bouffet E, Bradner JE, Ramaswamy V, Jabado N, Rutka JT, Northcott PA, Lupien M, Lichter P, Korshunov A, Scacheri PC, Pfister SM, Kool M, Taylor MD, Rich JN. Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling. Nature. 2018;553(7686):101-5. doi: 10.1038/nature25169. PubMed PMID: 29258295; PMCID: PMC5993422.

256. Magnani L, Stoeck A, Zhang X, Lanczky A, Mirabella AC, Wang TL, Gyorffy B, Lupien M. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. Proc Natl Acad Sci U S A. 2013;110(16):E1490-9. doi: 10.1073/pnas.1219992110. PubMed PMID: 23576735; PMCID: PMC3631697.

257. Marella NV, Bhattacharya S, Mukherjee L, Xu J, Berezney R. Cell type specific chromosome territory organization in the interphase nucleus of normal and cancer cells. J Cell Physiol. 2009;221(1):130-8. doi: 10.1002/jcp.21836. PubMed PMID: 19496171.

258. Marmorstein R. Structure of histone acetyltransferases. J Mol Biol. 2001;311(3):433-44. doi: 10.1006/jmbi.2001.4859. PubMed PMID: 11492997.

259. Marsman J, Horsfield JA. Long distance relationships: enhancer-promoter communication and dynamic gene transcription. Biochim Biophys Acta. 2012;1819(11-12):1217-27. doi: 10.1016/j.bbagrm.2012.10.008. PubMed PMID: 23124110.

260. Martinez-Sobrido L, Giannakas P, Cubitt B, Garcia-Sastre A, de la Torre JC. Differential inhibition of type I interferon induction by arenavirus nucleoproteins. J Virol. 2007;81(22):12696-703. doi: 10.1128/JVI.00882-07. PubMed PMID: 17804508; PMCID: PMC2168988.

261. Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. Science. 2003;302(5646):890-3. doi: 10.1126/science.1090842. PubMed PMID: 14593184.

262. Marwick JA, Kirkham PA, Stevenson CS, Danahay H, Giddings J, Butler K, Donaldson K, Macnee W, Rahman I. Cigarette smoke alters chromatin remodeling and induces proinflammatory genes in rat lungs. Am J Respir Cell Mol Biol. 2004;31(6):633-42. doi: 10.1165/rcmb.2004-0006OC. PubMed PMID: 15333327.

263. Massarweh S, Osborne CK, Creighton CJ, Qin L, Tsimelzon A, Huang S, Weiss H, Rimawi M, Schiff R. Tamoxifen resistance in breast tumors is driven by growth factor receptor signaling with repression of classic estrogen receptor genomic function. Cancer Res. 2008;68(3):826-33. doi: 10.1158/0008-5472.CAN-07-2707. PubMed PMID: 18245484.

264. Matsui T, Segall J, Weil PA, Roeder RG. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. J Biol Chem. 1980;255(24):11992-6. PubMed PMID: 7440580.

265. Maurano MT, Wang H, John S, Shafer A, Canfield T, Lee K,

Stamatoyannopoulos JA. Role of DNA Methylation in Modulating Transcription Factor Occupancy. Cell Rep. 2015;12(7):1184-95. doi: 10.1016/j.celrep.2015.07.024. PubMed PMID: 26257180.
266. McCullough CE, Marmorstein R. Molecular Basis for Histone Acetyltransferase Regulation by Binding Partners, Associated Domains, and Autoacetylation. ACS Chem Biol. 2016;11(3):632-42. doi: 10.1021/acschembio.5b00841. PubMed PMID: 26555232; PMCID: PMC5067162.
267. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010;28(5):495-501. doi: 10.1038/nbt.1630. PubMed PMID: 20436461; PMCID: PMC4840234.
268. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature. 2013;501(7467):328-37. doi: 10.1038/nature12624. PubMed PMID: 24048065; PMCID: PMC4521623.
269. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol. 2013;20(3):300-7. doi: 10.1038/nsmb.2480. PubMed PMID: 23463315.
270. Miki T, Yasuda SY, Kahn M. Wnt/beta-catenin signaling in embryonic stem cell self-renewal and somatic cell reprogramming. Stem Cell Rev. 2011;7(4):836-46. doi: 10.1007/s12015-011-9275-1. PubMed PMID: 21603945.
271. Milazzo ML, Barragan-Gomez A, Hanson JD, Estrada-Franco JG, Arellano E, Gonzalez-Cozatl FX, Fernandez-Salas I, Ramirez-Aguilar F, Rogers DS, Bradley RD, Fulhorst CF. Antibodies to Tacaribe serocomplex viruses (family Arenaviridae, genus Arenavirus) in cricetid rodents from new Mexico, Texas, and Mexico. Vector Borne Zoonotic Dis. 2010;10(6):629-37. Epub 2010/08/28. doi: 10.1089/vbz.2009.0206. PubMed PMID: 20795917.
272. Misteli T, Soutoglou E. The emerging role of nuclear architecture in DNA repair and genome maintenance. Nat Rev Mol Cell Biol. 2009;10(4):243-54. doi: 10.1038/nrm2651. PubMed PMID: 19277046; PMCID: PMC3478884.
273. Miyamoto MM. A congruence study of molecular and morphological data for Eutherian mammals. Mol Phylogenet Evol. 1996;6(3):373-90. doi: DOI 10.1006/mpev.1996.0087. PubMed PMID: WOS:A1996WB05700005.
274. Moisan S, Levon S, Cornec-Le Gall E, Le Meur Y, Audrezet MP, Dostie J, Ferec C. Novel long-range regulatory mechanisms controlling PKD2 gene expression. BMC Genomics. 2018;19(1):515. doi: 10.1186/s12864-018-4892-6. PubMed PMID: 29986647; PMCID: PMC6038307.
275. Morera L, Lubbert M, Jung M. Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. Clin Epigenetics. 2016;8:57. doi: 10.1186/s13148-016-0223-4. PubMed PMID: 27222667; PMCID: PMC4877953.
276. Morris JPt, Wang SC, Hebrok M. KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. Nat Rev Cancer. 2010;10(10):683-95. doi: 10.1038/nrc2899. PubMed PMID: 20814421; PMCID: PMC4085546.
277. Mostoslavsky R, Bardeesy N. Reprogramming Enhancers to Drive Metastasis.

Cell. 2017;170(5):823-5. doi: 10.1016/j.cell.2017.08.010. PubMed PMID: 28841414.
278. Mourad R, Hsu PY, Juan L, Shen C, Koneru P, Lin H, Liu Y, Nephew K, Huang TH, Li L. Estrogen induces global reorganization of chromatin structure in human breast cancer cells. PLoS One. 2014;9(12):e113354. doi: 10.1371/journal.pone.0113354. PubMed PMID: 25470140; PMCID: PMC4255042.
279. Mu GG, Zhang LL, Li HY, Liao Y, Yu HG. Thymoquinone Pretreatment Overcomes the Insensitivity and Potentiates the Antitumor Effect of Gemcitabine Through Abrogation of Notch1, PI3K/Akt/mTOR Regulated Signaling Pathways in Pancreatic Cancer. Dig Dis Sci. 2015;60(4):1067-80. doi: 10.1007/s10620-014-3394-x. PubMed PMID: 25344906.
280. Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. Molecular phylogenetics and the origins of placental mammals. Nature. 2001;409(6820):614-8. doi: 10.1038/35054550. PubMed PMID: 11214319.
281. Murthy D, Attri KS, Singh PK. Phosphoinositide 3-Kinase Signaling Pathway in Pancreatic Ductal Adenocarcinoma Progression, Pathogenesis, and Therapeutics. Front Physiol. 2018;9:335. doi: 10.3389/fphys.2018.00335. PubMed PMID: 29670543; PMCID: PMC5893816.
282. Murugesapillai D, McCauley MJ, Maher LJ, 3rd, Williams MC. Single-molecule studies of high-mobility group B architectural DNA bending proteins. Biophys Rev. 2017;9(1):17-40. doi: 10.1007/s12551-016-0236-4. PubMed PMID: 28303166; PMCID: PMC5331113.
283. Muscat GE, Eriksson NA, Byth K, Loi S, Graham D, Jindal S, Davis MJ, Clyne C, Funder JW, Simpson ER, Ragan MA, Kuczek E, Fuller PJ, Tilley WD, Leedman PJ, Clarke CL. Research resource: nuclear receptors as transcriptome: discriminant and prognostic value in breast cancer. Mol Endocrinol. 2013;27(2):350-65. doi: 10.1210/me.2012-1265. PubMed PMID: 23292282; PMCID: PMC5417325.
284. Nakamoto M, Matsuyama A, Shiba E, Shibuya R, Kasai T, Yamaguchi K, Hisaoka M. Prognostic significance of WNT signaling in pancreatic ductal adenocarcinoma. Virchows Arch. 2014;465(4):401-8. doi: 10.1007/s00428-014-1642-2. PubMed PMID: 25146168.
285. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature. 1998;393(6683):386-9. doi: 10.1038/30764. PubMed PMID: 9620804.
286. Naumova N, Dekker J. Integrating one-dimensional and three-dimensional maps of genomes. J Cell Sci. 2010;123(Pt 12):1979-88. doi: 10.1242/jcs.051631. PubMed PMID: 20519580; PMCID: PMC2880011.
287. Navarro L, David M. p38-dependent activation of interferon regulatory factor 3 by lipopolysaccharide. J Biol Chem. 1999;274(50):35535-8. PubMed PMID: 10585427.
288. Nebbioso A, Tambaro FP, Dell'Aversana C, Altucci L. Cancer epigenetics: Moving forward. PLoS Genet. 2018;14(6):e1007362. doi:

10.1371/journal.pgen.1007362. PubMed PMID: 29879107; PMCID: PMC5991666.
289. Neefjes J, Jongsma ML, Paul P, Bakke O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol. 2011;11(12):823-36. doi: 10.1038/nri3084. PubMed PMID: 22076556.
290. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell. 2006;10(6):515-27. doi: 10.1016/j.ccr.2006.10.008. PubMed PMID: 17157791; PMCID: PMC2730521.
291. Ng HH, Bird A. DNA methylation and chromatin modification. Curr Opin Genet Dev. 1999;9(2):158-63. PubMed PMID: 10322130.
292. Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell. 2017;169(5):930-44 e22. doi: 10.1016/j.cell.2017.05.004. PubMed PMID: 28525758; PMCID: PMC5538188.
293. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Bluthgen N, Dekker J, Heard E. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012;485(7398):381-5. doi: 10.1038/nature11049. PubMed PMID: 22495304; PMCID: PMC3555144.
294. O'Geen H, Frietze S, Farnham PJ. Using ChIP-seq technology to identify targets of zinc finger transcription factors. Methods Mol Biol. 2010;649:437-55. doi: 10.1007/978-1-60761-753-2_27. PubMed PMID: 20680851; PMCID: PMC4151297.
295. O'Malley M, King AN, Conte M, Ellingrod VL, Ramnath N. Effects of cigarette smoking on metabolism and effectiveness of systemic therapy for lung cancer. J Thorac Oncol. 2014;9(7):917-26. doi: 10.1097/JTO.0000000000000191. PubMed PMID: 24926542.
296. O'Shea TJ, Cryan PM, Cunningham AA, Fooks AR, Hayman DT, Luis AD, Peel AJ, Plowright RK, Wood JL. Bat flight and zoonotic viruses. Emerg Infect Dis. 2014;20(5):741-5. doi: 10.3201/eid2005.130539. PubMed PMID: 24750692; PMCID: 4012789.
297. Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell. 1996;87(5):953-9. PubMed PMID: 8945521.
298. Oike T, Komachi M, Ogiwara H, Amornwichet N, Saitoh Y, Torikai K, Kubo N, Nakano T, Kohno T. C646, a selective small molecule inhibitor of histone acetyltransferase p300, radiosensitizes lung cancer cells by enhancing mitotic catastrophe. Radiother Oncol. 2014;111(2):222-7. doi: 10.1016/j.radonc.2014.03.015. PubMed PMID: 24746574.
299. Oliner JD, Andresen JM, Hansen SK, Zhou S, Tjian R. SREBP transcriptional

activity is mediated through an interaction with the CREB-binding protein. Genes Dev. 1996;10(22):2903-11. PubMed PMID: 8918891.
300. Omatsu T, Watanabe S, Akashi H, Yoshikawa Y. Biological characters of bats in relation to natural reservoir of emerging viruses. Comp Immunol Microbiol Infect Dis. 2007;30(5-6):357-74. Epub 2007/08/21. doi: S0147-9571(07)00048-3 [pii]
3. 10.1016/j.cimid.2007.05.006. PubMed PMID: 17706776.
301. Ono H, Basson MD, Ito H. P300 inhibition enhances gemcitabine-induced apoptosis of pancreatic cancer. Oncotarget. 2016;7(32):51301-10. doi: 10.18632/oncotarget.10117. PubMed PMID: 27322077; PMCID: PMC5239476.
302. Ono K, Demchak B, Ideker T. Cytoscape tools for the web age: D3.js and Cytoscape.js exporters. F1000Res. 2014;3:143. doi: 10.12688/f1000research.4510.2. PubMed PMID: 25520778.
303. Osborne CK, Hobbs K, Clark GM. Effect of estrogens and antiestrogens on growth of human breast cancer cells in athymic nude mice. Cancer Res. 1985;45(2):584-90. PubMed PMID: 3967234.
304. Osborne CK, Jarman M, McCague R, Coronado EB, Hilsenbeck SG, Wakeling AE. The importance of tamoxifen metabolism in tamoxifen-stimulated breast tumor growth. Cancer Chemother Pharmacol. 1994;34(2):89-95. PubMed PMID: 8194171.
305. Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P. Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet. 2004;36(10):1065-71. doi: 10.1038/ng1423. PubMed PMID: 15361872.
306. Ouaissi M, Sielezneff I, Silvestre R, Sastre B, Bernard JP, Lafontaine JS, Payan MJ, Dahan L, Pirro N, Seitz JF, Mas E, Lombardo D, Ouaissi A. High histone deacetylase 7 (HDAC7) expression is significantly associated with adenocarcinomas of the pancreas. Ann Surg Oncol. 2008;15(8):2318-28. doi: 10.1245/s10434-008-9940-z. PubMed PMID: 18506539.
307. Oudet P, Gross-Bellard M, Chambon P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell. 1975;4(4):281-300. PubMed PMID: 1122558.
308. Pal B, Bouras T, Shi W, Vaillant F, Sheridan JM, Fu N, Breslin K, Jiang K, Ritchie ME, Young M, Lindeman GJ, Smyth GK, Visvader JE. Global changes in the mammary epigenome are induced by hormonal cues and coordinated by Ezh2. Cell Rep. 2013;3(2):411-26. doi: 10.1016/j.celrep.2012.12.020. PubMed PMID: 23375371.
309. Papenfuss AT, Baker ML, Feng ZP, Tachedjian M, Crameri G, Cowled C, Ng J, Janardhana V, Field HE, Wang LF. The immune gene repertoire of an important viral reservoir, the Australian black flying fox. BMC Genomics. 2012;13:261. doi: 10.1186/1471-2164-13-261. PubMed PMID: 22716473; PMCID: 3436859.
310. Parada LA, McQueen PG, Misteli T. Tissue-specific spatial organization of genomes. Genome Biol. 2004;5(7):R44. doi: 10.1186/gb-2004-5-7-r44. PubMed PMID: 15239829; PMCID: PMC463291.
311. Parada LA, McQueen PG, Munson PJ, Misteli T. Conservation of relative chromosome positioning in normal and cancer cells. Curr Biol.

2002;12(19):1692-7. PubMed PMID: 12361574.
312. Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N, Program NCS, Black BL, Visel A, Pennacchio LA, Collins FS, National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program A, Authors NCSP. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A. 2013;110(44):17921-6. doi: 10.1073/pnas.1317023110. PubMed PMID: 24127591; PMCID: PMC3816444.
313. Pawan JL. Rabies in the vampire bat of Trinidad, with special reference to the clinical course and the latency of infection. Caribb Med J. 1959;21:137-56. PubMed PMID: 14431118.
314. Perri F, Longo F, Giuliano M, Sabbatino F, Favia G, Ionna F, Addeo R, Della Vittoria Scarpati G, Di Lorenzo G, Pisconti S. Epigenetic control of gene expression: Potential implications for cancer treatment. Crit Rev Oncol Hematol. 2017;111:166-72. doi: 10.1016/j.critrevonc.2017.01.020. PubMed PMID: 28259291.
315. Pettigrew JD, Jamieson BGM, Robson SK, Hall LS, Mcanally KI, Cooper HM. Phylogenetic Relations between Microbats, Megabats and Primates (Mammalia, Chiroptera and Primates). Philos T Roy Soc B. 1989;325(1229):489-559. doi: DOI 10.1098/rstb.1989.0102. PubMed PMID: WOS:A1989CD42900001.
316. Phillips DM. The presence of acetyl groups of histones. Biochem J. 1963;87:258-63. PubMed PMID: 13943142; PMCID: PMC1201885.
317. Polakis P. Wnt signaling in cancer. Cold Spring Harb Perspect Biol. 2012;4(5). doi: 10.1101/cshperspect.a008052. PubMed PMID: 22438566; PMCID: PMC3331705.
318. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol. 2010;28(10):1057-68. doi: 10.1038/nbt.1685. PubMed PMID: 20944598.
319. Principe DR, DeCant B, Mascarinas E, Wayne EA, Diaz AM, Akagi N, Hwang R, Pasche B, Dawson DW, Fang D, Bentrem DJ, Munshi HG, Jung B, Grippo PJ. TGFbeta Signaling in the Pancreatic Tumor Microenvironment Promotes Fibrosis and Immune Evasion to Facilitate Tumorigenesis. Cancer Res. 2016;76(9):2525-39. doi: 10.1158/0008-5472.CAN-15-1293. PubMed PMID: 26980767; PMCID: PMC4873388.
320. Pythoud C, Rodrigo WW, Pasqual G, Rothenberger S, Martinez-Sobrido L, de la Torre JC, Kunz S. Arenavirus nucleoprotein targets interferon regulatory factor-activating kinase IKKepsilon. J Virol. 2012;86(15):7728-38. doi: 10.1128/JVI.00187-12. PubMed PMID: 22532683; PMCID: PMC3421673.
321. Raghavan M, Bjorkman PJ. Fc receptors and their interactions with immunoglobulins. Annu Rev Cell Dev Biol. 1996;12:181-220. doi: 10.1146/annurev.cellbio.12.1.181. PubMed PMID: 8970726.
322. Ramani V, Shendure J, Duan Z. Understanding Spatial Genome Organization: Methods and Insights. Genomics Proteomics Bioinformatics. 2016;14(1):7-20. doi: 10.1016/j.gpb.2016.01.002. PubMed PMID: 26876719; PMCID: PMC4792841.

323. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160-5. doi: 10.1093/nar/gkw257. PubMed PMID: 27079975; PMCID: PMC4987876.
324. Ramos YF, Hestand MS, Verlaan M, Krabbendam E, Ariyurek Y, van Galen M, van Dam H, van Ommen GJ, den Dunnen JT, Zantema A, t Hoen PA. Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. Nucleic Acids Res. 2010;38(16):5396-408. doi: 10.1093/nar/gkq184. PubMed PMID: 20435671; PMCID: PMC2938195.
325. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159(7):1665-80. doi: 10.1016/j.cell.2014.11.021. PubMed PMID: 25497547; PMCID: PMC5635824.
326. Rebel VI, Kung AL, Tanner EA, Yang H, Bronson RT, Livingston DM. Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. Proc Natl Acad Sci U S A. 2002;99(23):14789-94. doi: 10.1073/pnas.232568499. PubMed PMID: 12397173; PMCID: PMC137497.
327. Reines D, Conaway JW, Conaway RC. The RNA polymerase II general elongation factors. Trends Biochem Sci. 1996;21(9):351-5. PubMed PMID: 8870500; PMCID: PMC3374595.
328. Reiter F, Wienerroither S, Stark A. Combinatorial function of transcription factors and cofactors. Curr Opin Genet Dev. 2017;43:73-81. doi: 10.1016/j.gde.2016.12.007. PubMed PMID: 28110180.
329. Riggins GJ, Kinzler KW, Vogelstein B, Thiagalingam S. Frequency of Smad gene mutations in human cancers. Cancer Res. 1997;57(13):2578-80. PubMed PMID: 9205057.
330. Ris H, Kubai DF. Chromosome structure. Annu Rev Genet. 1970;4:263-94. doi: 10.1146/annurev.ge.04.120170.001403. PubMed PMID: 4950239.
331. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods. 2007;4(8):651-7. doi: 10.1038/nmeth1068. PubMed PMID: 17558387.
332. Robertson KD, Uzvolgyi E, Liang G, Talmadge C, Sumegi J, Gonzales FA, Jones PA. The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. Nucleic Acids Res. 1999;27(11):2291-8. PubMed PMID: 10325416; PMCID: PMC148793.
333. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-6. doi: 10.1038/nbt.1754. PubMed PMID: 21221095; PMCID: PMC3346182.
334. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40. doi: 10.1093/bioinformatics/btp616. PubMed PMID: 19910308; PMCID: PMC2796818.

335. Roe JS, Hwang CI, Somerville TDD, Milazzo JP, Lee EJ, Da Silva B, Maiorino L, Tiriac H, Young CM, Miyabayashi K, Filippini D, Creighton B, Burkhart RA, Buscaglia JM, Kim EJ, Grem JL, Lazenby AJ, Grunkemeyer JA, Hollingsworth MA, Grandgenett PM, Egeblad M, Park Y, Tuveson DA, Vakoc CR. Enhancer Reprogramming Promotes Pancreatic Cancer Metastasis. Cell. 2017;170(5):875-88 e20. doi: 10.1016/j.cell.2017.07.007. PubMed PMID: 28757253; PMCID: PMC5726277.
336. Roeder RG, Rutter WJ. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature. 1969;224(5216):234-7. PubMed PMID: 5344598.
337. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin SF, Palmieri C, Caldas C, Carroll JS. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature. 2012;481(7381):389-93. doi: 10.1038/nature10730. PubMed PMID: 22217937; PMCID: PMC3272464.
338. Rountree MR, Bachman KE, Baylin SB. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet. 2000;25(3):269-77. doi: 10.1038/77023. PubMed PMID: 10888872.
339. Rozengurt E, Sinnett-Smith J, Eibl G. Yes-associated protein (YAP) in pancreatic cancer: at the epicenter of a targetable signaling network associated with patient survival. Signal Transduct Target Ther. 2018;3:11. doi: 10.1038/s41392-017-0005-2. PubMed PMID: 29682330; PMCID: PMC5908807.
340. Russell J, Zomerdijk JC. RNA-polymerase-I-directed rDNA transcription, life and works. Trends Biochem Sci. 2005;30(2):87-96. doi: 10.1016/j.tibs.2004.12.008. PubMed PMID: 15691654; PMCID: PMC3858833.
341. Ryu D, Joung JG, Kim NK, Kim KT, Park WY. Deciphering intratumor heterogeneity using cancer genome analysis. Hum Genet. 2016;135(6):635-42. doi: 10.1007/s00439-016-1670-x. PubMed PMID: 27126234.
342. Sandi MJ, Hamidi T, Malicet C, Cano C, Loncle C, Pierres A, Dagorn JC, Iovanna JL. p8 expression controls pancreatic cancer cell migration, invasion, adhesion, and tumorigenesis. J Cell Physiol. 2011;226(12):3442-51. doi: 10.1002/jcp.22702. PubMed PMID: 21344397.
343. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature. 2012;489(7414):109-13. doi: 10.1038/nature11279. PubMed PMID: 22955621; PMCID: PMC3555147.
344. Sato M, Suemori H, Hata N, Asagiri M, Ogasawara K, Nakao K, Nakaya T, Katsuki M, Noguchi S, Tanaka N, Taniguchi T. Distinct and essential roles of transcription factors IRF-3 and IRF-7 in response to viruses for IFN-alpha/beta gene induction. Immunity. 2000;13(4):539-48. PubMed PMID: 11070172.
345. Sayler KA, Barbet AF, Chamberlain C, Clapp WL, Alleman R, Loeb JC, Lednicky JA. Isolation of Tacaribe virus, a Caribbean arenavirus, from host-seeking Amblyomma americanum ticks in Florida. PLoS One. 2014;9(12):e115769. doi: 10.1371/journal.pone.0115769. PubMed PMID: 25536075; PMCID: 4275251.
346. Schaub FX, Dhankani V, Berger AC, Trivedi M, Richardson AB, Shaw R, Zhao W, Zhang X, Ventura A, Liu Y, Ayer DE, Hurlin PJ, Cherniack AD, Eisenman

RN, Bernard B, Grandori C, Cancer Genome Atlas N. Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. Cell Syst. 2018;6(3):282-300 e2. doi: 10.1016/j.cels.2018.03.003. PubMed PMID: 29596783; PMCID: PMC5892207.

347. Schneider G, Kramer OH, Schmid RM, Saur D. Acetylation as a transcriptional control mechanism-HDACs and HATs in pancreatic ductal adenocarcinoma. J Gastrointest Cancer. 2011;42(2):85-92. doi: 10.1007/s12029-011-9257-1. PubMed PMID: 21271301.
348. Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, Kurukuti S, Mitchell JA, Umlauf D, Dimitrova DS, Eskiw CH, Luo Y, Wei CL, Ruan Y, Bieker JJ, Fraser P. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet. 2010;42(1):53-61. doi: 10.1038/ng.496. PubMed PMID: 20010836; PMCID: PMC3237402.
349. Schoumacher RA, Ram J, Iannuzzi MC, Bradbury NA, Wallace RW, Hon CT, Kelly DR, Schmid SM, Gelder FB, Rado TA, et al. A cystic fibrosis pancreatic adenocarcinoma cell line. Proc Natl Acad Sci U S A. 1990;87(10):4012-6. PubMed PMID: 1692630; PMCID: PMC54034.
350. Schountz T. Immunology of bats and their viruses: challenges and opportunities. Viruses. 2014;6(12):4880-901. doi: 10.3390/v6124880. PubMed PMID: 25494448; PMCID: 4276934.
351. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, C HH, Mirny L, Spitz F. Two independent modes of chromatin organization revealed by cohesin removal. Nature. 2017;551(7678):51-6. doi: 10.1038/nature24281. PubMed PMID: 29094699; PMCID: PMC5687303.
352. Seppola M, Stenvik J, Steiro K, Solstad T, Robertsen B, Jensen I. Sequence and expression analysis of an interferon stimulated gene (ISG15) from Atlantic cod (Gadus morhua L.). Dev Comp Immunol. 2007;31(2):156-71. doi: 10.1016/j.dci.2006.05.009. PubMed PMID: 16824598.
353. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 2015;16:259. doi: 10.1186/s13059-015-0831-x. PubMed PMID: 26619908; PMCID: PMC4665391.
354. Sever R, Glass CK. Signaling by nuclear receptors. Cold Spring Harb Perspect Biol. 2013;5(3):a016709. doi: 10.1101/cshperspect.a016709. PubMed PMID: 23457262; PMCID: PMC3578364.
355. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell. 2012;148(3):458-72. doi: 10.1016/j.cell.2012.01.010. PubMed PMID: 22265598.
356. Shaw TI, Srivastava A, Chou WC, Liu L, Hawkinson A, Glenn TC, Adams R, Schountz T. Transcriptome sequencing and annotation for the Jamaican fruit bat (Artibeus jamaicensis). PLoS One. 2012;7(11):e48472. doi: 10.1371/journal.pone.0048472. PubMed PMID: 23166587; PMCID: 3499531.
357. Shimano H. Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. Prog Lipid Res. 2001;40(6):439-52. PubMed

PMID: 11591434.
358. Shin H, Shi Y, Dai C, Tjong H, Gong K, Alber F, Zhou XJ. TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res. 2016;44(7):e70. doi: 10.1093/nar/gkv1505. PubMed PMID: 26704975; PMCID: PMC4838359.
359. Shitashige M, Hirohashi S, Yamada T. Wnt signaling inside the nucleus. Cancer Sci. 2008;99(4):631-7. doi: 10.1111/j.1349-7006.2007.00716.x. PubMed PMID: 18177486.
360. Shoman N, Klassen S, McFadden A, Bickis MG, Torlakovic E, Chibbar R. Reduced PTEN expression predicts relapse in patients with breast carcinoma treated by tamoxifen. Mod Pathol. 2005;18(2):250-9. doi: 10.1038/modpathol.3800296. PubMed PMID: 15475931.
361. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67(1):7-30. doi: 10.3322/caac.21387. PubMed PMID: 28055103.
362. Singh BN, Zhang G, Hwa YL, Li J, Dowdy SC, Jiang SW. Nonhistone protein acetylation as cancer therapy targets. Expert Rev Anticancer Ther. 2010;10(6):935-54. doi: 10.1586/era.10.62. PubMed PMID: 20553216; PMCID: PMC3273412.
363. Sipos B, Moser S, Kalthoff H, Torok V, Lohr M, Kloppel G. A comprehensive characterization of pancreatic ductal carcinoma cell lines: towards the establishment of an in vitro research platform. Virchows Arch. 2003;442(5):444-52. doi: 10.1007/s00428-003-0784-4. PubMed PMID: 12692724.
364. Solis M, Goubau D, Romieu-Mourez R, Genin P, Civas A, Hiscott J. Distinct functions of IRF-3 and IRF-7 in IFN-alpha gene regulation and control of anti-tumor activity in primary macrophages. Biochem Pharmacol. 2006;72(11):1469-76. doi: 10.1016/j.bcp.2006.06.002. PubMed PMID: 16846591.
365. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869-74. doi: 10.1073/pnas.191367098. PubMed PMID: 11553815; PMCID: PMC58566.
366. Sorrentino G, Ruggeri N, Zannini A, Ingallina E, Bertolio R, Marotta C, Neri C, Cappuzzello E, Forcato M, Rosato A, Mano M, Bicciato S, Del Sal G. Glucocorticoid receptor signalling activates YAP in breast cancer. Nat Commun. 2017;8:14073. doi: 10.1038/ncomms14073. PubMed PMID: 28102225; PMCID: PMC5253666.
367. Splinter E, de Wit E, Nora EP, Klous P, van de Werken HJ, Zhu Y, Kaaij LJ, van Ijcken W, Gribnau J, Heard E, de Laat W. The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes Dev. 2011;25(13):1371-83. doi: 10.1101/gad.633311. PubMed PMID: 21690198; PMCID: PMC3134081.
368. St John HC, Bishop KA, Meyer MB, Benkusky NA, Leng N, Kendziorski C, Bonewald LF, Pike JW. The osteoblast to osteocyte transition: epigenetic changes and response to the vitamin D3 hormone. Mol Endocrinol. 2014;28(7):1150-65.

doi: 10.1210/me.2014-1091. PubMed PMID: 24877565; PMCID: PMC5414828.
369. Stavreva DA, Coulon A, Baek S, Sung MH, John S, Stixova L, Tesikova M, Hakim O, Miranda T, Hawkins M, Stamatoyannopoulos JA, Chow CC, Hager GL. Dynamics of chromatin accessibility and long-range interactions in response to glucocorticoid pulsing. Genome Res. 2015;25(6):845-57. doi: 10.1101/gr.184168.114. PubMed PMID: 25677181; PMCID: PMC4448681.
370. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000;403(6765):41-5. doi: 10.1038/47412. PubMed PMID: 10638745.
371. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50. doi: 10.1073/pnas.0506580102. PubMed PMID: 16199517; PMCID: PMC1239896.
372. Sulkin SE, Allen R. Virus infections in bats. Monogr Virol. 1974;8(0):1-103. PubMed PMID: 4367453.
373. Sunami Y, Rebelo A, Kleeff J. Lipid Metabolism and Lipid Droplets in Pancreatic Cancer and Stellate Cells. Cancers (Basel). 2017;10(1). doi: 10.3390/cancers10010003. PubMed PMID: 29295482; PMCID: PMC5789353.
374. Suzuki S, Suzuki Y, Yamamoto N, Matsumoto Y, Shirai A, Okubo T. Influenza A virus infection increases IgE production and airway responsiveness in aerosolized antigen-exposed mice. J Allergy Clin Immunol. 1998;102(5):732-40. PubMed PMID: 9819289.
375. Sydow JF, Brueckner F, Cheung AC, Damsma GE, Dengl S, Lehmann E, Vassylyev D, Cramer P. Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. Mol Cell. 2009;34(6):710-21. doi: 10.1016/j.molcel.2009.06.002. PubMed PMID: 19560423.
376. T B. Zur Frage der Entstehung maligner Tumoren. G Fischer. 1914.
377. Taberlay PC, Achinger-Kawecka J, Lun AT, Buske FA, Sabir K, Gould CM, Zotenko E, Bert SA, Giles KA, Bauer DC, Smyth GK, Stirzaker C, O'Donoghue SI, Clark SJ. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. Genome Res. 2016;26(6):719-31. doi: 10.1101/gr.201517.115. PubMed PMID: 27053337; PMCID: PMC4889976.
378. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell. 2004;116(1):51-61. PubMed PMID: 14718166.
379. Takahashi-Yanaga F, Kahn M. Targeting Wnt signaling: can we safely eradicate cancer stem cells? Clin Cancer Res. 2010;16(12):3153-62. doi: 10.1158/1078-0432.CCR-09-2943. PubMed PMID: 20530697.
380. Tanabe H, Habermann FA, Solovei I, Cremer M, Cremer T. Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. Mutat Res. 2002;504(1-2):37-45. PubMed PMID: 12106644.
381. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A,

Wlodarczyk J, Ruszczycki B, Michalski P, Piecuch E, Wang P, Wang D, Tian SZ, Penrad-Mobayed M, Sachs LM, Ruan X, Wei CL, Liu ET, Wilczynski GM, Plewczynski D, Li G, Ruan Y. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell. 2015;163(7):1611-27. doi: 10.1016/j.cell.2015.11.024. PubMed PMID: 26686651; PMCID: PMC4734140.

382. Teo JL, Kahn M. The Wnt signaling pathway in cellular proliferation and differentiation: A tale of two coactivators. Adv Drug Deliv Rev. 2010;62(12):1149-55. doi: 10.1016/j.addr.2010.09.012. PubMed PMID: 20920541.

383. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14(2):178-92. doi: 10.1093/bib/bbs017. PubMed PMID: 22517427; PMCID: PMC3603213.

384. Tian W, Han X, Yan M, Xu Y, Duggineni S, Lin N, Luo G, Li YM, Han X, Huang Z, An J. Structure-based discovery of a novel inhibitor targeting the beta-catenin/Tcf4 interaction. Biochemistry. 2012;51(2):724-31. doi: 10.1021/bi201428h. PubMed PMID: 22224445.

385. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7(3):562-78. doi: 10.1038/nprot.2012.016. PubMed PMID: 22383036; PMCID: PMC3334321.

386. Turmelle AS, Jackson FR, Green D, McCracken GF, Rupprecht CE. Host immunity to repeated rabies virus infection in big brown bats. J Gen Virol. 2010;91(Pt 9):2360-6. Epub 2010/06/04. doi: vir.0.020073-0 [pii]

4. 10.1099/vir.0.020073-0. PubMed PMID: 20519458.

387. Ucar D, Bayarsaihan D. Cell-specific gene promoters are marked by broader spans of H3K4me3 and are associated with robust gene expression patterns. Epigenomics. 2015;7(2):129-31. doi: 10.2217/epi.14.87. PubMed PMID: 25942528.

388. Ugai H, Uchida K, Kawasaki H, Yokoyama KK. The coactivators p300 and CBP have different functions during the differentiation of F9 cells. J Mol Med (Berl). 1999;77(6):481-94. PubMed PMID: 10475063.

389. Unkeless JC, Scigliano E, Freedman VH. Structure and function of human and murine receptors for IgG. Annu Rev Immunol. 1988;6:251-81. doi: 10.1146/annurev.iy.06.040188.001343. PubMed PMID: 2968084.

390. Van Speybroeck L. From epigenesis to epigenetics: the case of C. H. Waddington. Ann N Y Acad Sci. 2002;981:61-81. PubMed PMID: 12547674.

391. van Steensel B, Dekker J. Genomics tools for unraveling chromosome architecture. Nat Biotechnol. 2010;28(10):1089-95. doi: 10.1038/nbt.1680. PubMed PMID: 20944601; PMCID: PMC3023824.

392. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. Nat Rev Genet. 2009;10(4):252-63. doi: 10.1038/nrg2538. PubMed PMID: 19274049.

393. Varet H, Brillet-Gueguen L, Coppee JY, Dillies MA. SARTools: A DESeq2- and

EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. PLoS One. 2016;11(6):e0157022. doi: 10.1371/journal.pone.0157022. PubMed PMID: 27280887; PMCID: PMC4900645.
394. Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep. 2015;10(8):1297-309. doi: 10.1016/j.celrep.2015.02.004. PubMed PMID: 25732821; PMCID: PMC4542312.
395. Vincent AJ, Ren S, Harris LG, Devine DJ, Samant RS, Fodstad O, Shevde LA. Cytoplasmic translocation of p21 mediates NUPR1-induced chemoresistance: NUPR1 and p21 in chemoresistance. FEBS Lett. 2012;586(19):3429-34. doi: 10.1016/j.febslet.2012.07.063. PubMed PMID: 22858377.
396. von Hansemann D. Ueber asymmetrische Zelltheilung in epithel Krebsen und deren biologische Bedeutung. Virchows Arch. Path Anat. 1890:119, 299.
397. Wade PA, Gegonne A, Jones PL, Ballestar E, Aubry F, Wolffe AP. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat Genet. 1999;23(1):62-6. doi: 10.1038/12664. PubMed PMID: 10471500.
398. Wall I, Schmidt-Wolf IG. Effect of Wnt inhibitors in pancreatic cancer. Anticancer Res. 2014;34(10):5375-80. PubMed PMID: 25275031.
399. Wang J, Lan X, Hsu PY, Hsu HK, Huang K, Parvin J, Huang TH, Jin VX. Genome-wide analysis uncovers high frequency, strong differential chromosomal interactions and their associated epigenetic patterns in E2-mediated gene regulation. BMC Genomics. 2013;14:70. doi: 10.1186/1471-2164-14-70. PubMed PMID: 23368971; PMCID: PMC3599885.
400. Wang L, Tang Y, Cole PA, Marmorstein R. Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. Curr Opin Struct Biol. 2008;18(6):741-7. doi: 10.1016/j.sbi.2008.09.004. PubMed PMID: 18845255; PMCID: PMC2643075.
401. Wang L-F. Bats and Viruses: a Breif Review. Virologica Sinica. 2009;24(2):93-9.
402. Wang LF, Walker PJ, Poon LL. Mass extinctions, biodiversity and mitochondrial function: are bats 'special' as reservoirs for emerging viruses? Curr Opin Virol. 2011;1(6):649-57. doi: 10.1016/j.coviro.2011.10.013. PubMed PMID: 22440923.
403. Wang W, Abbruzzese JL, Evans DB, Larry L, Cleary KR, Chiao PJ. The nuclear factor-kappa B RelA transcription factor is constitutively activated in human pancreatic adenocarcinoma cells. Clin Cancer Res. 1999;5(1):119-27. PubMed PMID: 9918209.
404. Wapenaar H, Dekker FJ. Histone acetyltransferases: challenges in targeting bi-substrate enzymes. Clin Epigenetics. 2016;8:59. doi: 10.1186/s13148-016-0225-2. PubMed PMID: 27231488; PMCID: PMC4881052.
405. Weinmann R, Roeder RG. Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. Proc Natl Acad Sci U S A. 1974;71(5):1790-4. PubMed PMID: 4525293; PMCID: PMC388326.
406. Weiss RH. p21Waf1/Cip1 as a therapeutic target in breast and other cancers.

Cancer Cell. 2003;4(6):425-9. PubMed PMID: 14706334.
407. Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG. ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J. 2009;28(10):1418-28. doi: 10.1038/emboj.2009.88. PubMed PMID: 19339991; PMCID: PMC2688537.
408. West AC, Johnstone RW. New and emerging HDAC inhibitors for cancer treatment. J Clin Invest. 2014;124(1):30-9. doi: 10.1172/JCI69738. PubMed PMID: 24382387; PMCID: PMC3871231.
409. Whalen S, Truty RM, Pollard KS. Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. Nat Genet. 2016;48(5):488-96. doi: 10.1038/ng.3539. PubMed PMID: 27064255; PMCID: PMC4910881.
410. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153(2):307-19. doi: 10.1016/j.cell.2013.03.035. PubMed PMID: 23582322; PMCID: PMC3653129.
411. Wibbelt G, Moore MS, Schountz T, Voigt CC. Emerging diseases in Chiroptera: why bats? Biol Lett. 2010;6(4):438-40. Epub 2010/04/30. doi: rsbl.2010.0267 [pii]
5. 10.1098/rsbl.2010.0267. PubMed PMID: 20427329.
412. Workman JL, Kingston RE. Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annu Rev Biochem. 1998;67:545-79. doi: 10.1146/annurev.biochem.67.1.545. PubMed PMID: 9759497.
413. Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Mol Cell Biol. 2010;30(6):1411-20. doi: 10.1128/MCB.01384-09. PubMed PMID: 20065031; PMCID: PMC2832500.
414. Wu AH, Huang YL, Zhang LZ, Tian G, Liao QZ, Chen SL. MiR-572 prompted cell proliferation of human ovarian cancer cells by suppressing PPP2R2C expression. Biomed Pharmacother. 2016;77:92-7. doi: 10.1016/j.biopha.2015.12.005. PubMed PMID: 26796271.
415. Xiao J, Jin R, Wagner D. Developmental transitions: integrating environmental cues with hormonal signaling in the chromatin landscape in plants. Genome Biol. 2017;18(1):88. doi: 10.1186/s13059-017-1228-9. PubMed PMID: 28490341; PMCID: PMC5425979.
416. Xiao X, Song BL. SREBP: a novel therapeutic target. Acta Biochim Biophys Sin (Shanghai). 2013;45(1):2-10. doi: 10.1093/abbs/gms112. PubMed PMID: 23257291.
417. Xu W, Wang Z, Zhang W, Qian K, Li H, Kong D, Li Y, Tang Y. Mutated K-ras activates CDK8 to stimulate the epithelial-to-mesenchymal transition in pancreatic cancer in part via the Wnt/beta-catenin signaling pathway. Cancer Lett. 2015;356(2 Pt B):613-27. doi: 10.1016/j.canlet.2014.10.008. PubMed PMID: 25305448.
418. Xu YM, Du JY, Lau AT. Posttranslational modifications of human histone H3: an update. Proteomics. 2014;14(17-18):2047-60. doi: 10.1002/pmic.201300435. PubMed PMID: 25044606.

419. Yang SR, Chida AS, Bauter MR, Shafiq N, Seweryniak K, Maggirwar SB, Kilty I, Rahman I. Cigarette smoke induces proinflammatory cytokine release by activation of NF-kappaB and posttranslational modifications of histone deacetylase in macrophages. Am J Physiol Lung Cell Mol Physiol. 2006;291(1):L46-57. doi: 10.1152/ajplung.00241.2005. PubMed PMID: 16473865.
420. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586-91. doi: 10.1093/molbev/msm088. PubMed PMID: 17483113.
421. Ying H, Dey P, Yao W, Kimmelman AC, Draetta GF, Maitra A, DePinho RA. Genetics and biology of pancreatic ductal adenocarcinoma. Genes Dev. 2016;30(4):355-85. doi: 10.1101/gad.275776.115. PubMed PMID: 26883357; PMCID: PMC4762423.
422. Yu W, He B, Tan K. Identifying topologically associating domains and subdomains by Gaussian Mixture model And Proportion test. Nat Commun. 2017;8(1):535. doi: 10.1038/s41467-017-00478-8. PubMed PMID: 28912419; PMCID: PMC5599511.
423. Yunis AA, Arimura GK, Russin DJ. Human pancreatic carcinoma (MIA PaCa-2) in continuous culture: sensitivity to asparaginase. Int J Cancer. 1977;19(1):128-35. PubMed PMID: 832918.
424. Zavoral M, Minarikova P, Zavada F, Salek C, Minarik M. Molecular biology of pancreatic cancer. World J Gastroenterol. 2011;17(24):2897-908. doi: 10.3748/wjg.v17.i24.2897. PubMed PMID: 21734801; PMCID: PMC3129504.
425. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, Tachedjian M, Zhu Y, Zhou P, Jiang X, Ng J, Yang L, Wu L, Xiao J, Feng Y, Chen Y, Sun X, Zhang Y, Marsh GA, Crameri G, Broder CC, Frey KG, Wang LF, Wang J. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science. 2013;339(6118):456-60. doi: 10.1126/science.1230835. PubMed PMID: 23258410.
426. Zhang X, Lou Y, Zheng X, Wang H, Sun J, Dong Q, Han B. Wnt blockers inhibit the proliferation of lung cancer stem cells. Drug Des Devel Ther. 2015;9:2399-407. doi: 10.2147/DDDT.S76602. PubMed PMID: 25960639; PMCID: PMC4423515.
427. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137. doi: 10.1186/gb-2008-9-9-r137. PubMed PMID: 18798982; PMCID: PMC2592715.
428. Zhang Y, Morris JPt, Yan W, Schofield HK, Gurney A, Simeone DM, Millar SE, Hoey T, Hebrok M, Pasca di Magliano M. Canonical wnt signaling is required for pancreatic carcinogenesis. Cancer Res. 2013;73(15):4909-22. doi: 10.1158/0008-5472.CAN-12-4384. PubMed PMID: 23761328; PMCID: PMC3763696.
429. Zhao C, Denison C, Huibregtse JM, Gygi S, Krug RM. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. Proc Natl Acad Sci U S A. 2005;102(29):10200-5. doi: 10.1073/pnas.0504754102. PubMed PMID:

16009940; PMCID: 1177427.
430. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet. 2011;12(1):7-18. doi: 10.1038/nrg2905. PubMed PMID: 21116306.
431. Zhou W, Liang IC, Yee NS. Histone deacetylase 1 is required for exocrine pancreatic epithelial proliferation in development and cancer. Cancer Biol Ther. 2011;11(7):659-70. PubMed PMID: 21301206; PMCID: PMC3084970.
432. Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. Nat Rev Cancer. 2004;4(9):677-87. doi: 10.1038/nrc1430. PubMed PMID: 15343274.
433. Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, van de Corput MP, van de Werken HJ, Knoch TA, van IWF, Grosveld FG, Ren B, Wendt KS. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc Natl Acad Sci U S A. 2014;111(3):996-1001. doi: 10.1073/pnas.1317788111. PubMed PMID: 24335803; PMCID: PMC3903193.

Appendix A

Altering cancer transcriptomes using epigenomic inhibitors

Published in the following citation: Gaddis, M., Gerrard, D., Frietze, S., & Farnham, P. J. (2015). Altering cancer transcriptomes using epigenomic inhibitors. *Epigenetics Chromatin, 8*, 9. doi:10.1186/1756-8935-8-9

Abstract

Background: Due to the hyper-activation of WNT signaling in a variety of cancer types, there has been a strong drive to develop pathway-specific inhibitors with the eventual goal of providing a chemotherapeutic antagonist of WNT signaling to cancer patients. A new category of drugs, called epigenetic inhibitors, are being developed that hold high promise for inhibition of the WNT pathway. The canonical WNT signaling pathway initiates when WNT ligands bind to receptors, causing the nuclear localization of the co-activator β-catenin (CTNNB1), which leads to an association of β-catenin with a member of the TCF transcription factor family at regulatory regions of WNT-responsive genes. The TCF/β-catenin complex then recruits CBP (CREBBP) or p300 (EP300), leading to histone acetylation and gene activation. A current model in the field is that CBP-driven expression of WNT target genes supports proliferation whereas p300-driven expression of WNT target genes supports differentiation. The small molecule inhibitor ICG-001 binds to CBP, but not to p300, and competitively inhibits the interaction of CBP with β-catenin. Upon treatment of cancer cells, this should reduce expression of CBP-regulated transcription, leading to reduced tumorigenicity and enhanced differentiation.

Results: We have compared the genome-wide effects on the transcriptome after treatment with ICG-001 (the specific CBP inhibitor) versus C646, a compound that competes with acetyl-coA for the Lys-coA binding pocket of both CBP and p300. We found that both drugs cause large-scale changes in the transcriptome of HCT116 colon cancer cells and PANC1 pancreatic cancer cells and reverse some tumor-specific changes in gene expression. Interestingly, although the epigenetic inhibitors affect cell cycle pathways in both the colon and pancreatic cancer cell lines, the WNT signaling pathway was affected

only in the colon cancer cells. Notably, WNT target genes were similarly downregulated after treatment of HCT116 with C646 as with ICG-001.

Conclusion: Our results suggest that treatment with a general HAT inhibitor causes similar effects on the transcriptome as does treatment with a CBP-specific inhibitor and that epigenetic inhibition affects the WNT pathway in HCT116 cells and the cholesterol biosynthesis pathway in PANC1 cells.

Background

Due to the hyper-activation of WNT signaling in a variety of cancer types *(1,2)*, there has been a strong drive to develop antagonists of WNT signaling for cancer treatment. Standard inhibitors of the WNT signaling pathway include biologic inhibitors, such as small interfering RNAs, antibodies, and recombinant proteins, and chemical inhibitors, such as NSAIDs, vitamins, and polyphenols, that have fairly general (or unknown) targets *(1,3,4)*. However, a new category of drugs to target the WNT pathway is being developed that holds high promise as chemotherapeutics. These drugs, called epigenetic inhibitors, function to modify chromatin structure. Chromatin is composed of nucleosomes, which are comprised of 146 bp of DNA wrapped around eight core histone proteins (two copies each of H2A, H2B, H3, and H4). The N terminal tails of the core histones that constitute the nucleosome are subject to various different types of modifications that can influence chromatin structure and either enhance or inhibit the ability of transcription factors to bind to and regulate their target genes. The pattern of histone modifications throughout the genome, in combination with the pattern of DNA methylation, is called the epigenome. Recent studies have revealed that different histone

modifications are associated with active vs. silenced chromatin, that different cell types show different epigenomic patterns of silenced vs. active chromatin, and that changes in chromatin structure can have a dramatic effect on cell proliferation, differentiation, and survival. One widely studied histone modification is acetylation; histone acetylation is a critical regulatory mechanism of gene expression and plays an important role in gene expression. In fact, acetylation of histone H3 on lysine 27 is the epigenetic modification that most precisely identifies distal regulatory regions that serve as active enhancers *(5)*. Because cancer genomes show changes in histone acetylation patterns, there is great interest in the use of acetylation inhibitors that inhibit signaling pathways linked to human cancers for epigenetic therapy *(6)*.

Drugs that inhibit acetylation are particularly relevant for inhibition of the WNT pathway. The canonical WNT signaling pathway initiates when WNT ligands bind to receptors, resulting in the nuclear localization of the co-activator β-catenin (CTNNB1), which leads to an association of β-catenin with a member of the TCF/LEF transcription factor family at regulatory regions of WNT responsive genes *(7,8)*. The TCF/β-catenin complex can interact with co-activators such as CBP (CREBBP) and p300 (EP300) which function in part through the acetylation of histone H3 on lysine 27 *(5)*. Thus, it has been proposed that the initiation of the WNT signaling pathway ultimately ends with histone acetylation and a relaxing of the chromatin structure, a process necessary for gene activation. The small molecule inhibitor ICG-001 binds to CBP and competitively inhibits the interaction of CBP with β-catenin *(9,10)*, with the expected result of loss of active histone at promoters and enhancers regulated by TCF/β-catenin/CBP complexes (**Appendix A-1A**). Importantly, ICG-001 does not bind to the highly related histone

acetyltransferase (HAT) p300 and should not affect the activity of promoters or enhancers bound by TCF/β-catenin/p300 complexes. Thus, ICG-001 is thought to specifically decrease the expression of only the subset of WNT target genes regulated by β-catenin/CBP interactions. These proposed effects of ICG-001 are in contrast to those of C646 an inhibitor that competes with acetyl-coA for the Lys-coA binding pocket of p300 (**Appendix A-1B**). C646 is very selective for p300 versus six other unrelated histone acetyltransferases *(11)*. Although no direct comparisons have been performed, due to the mode of action of C646 and because the HAT domains of p300 and CBP have greater than 90% similarity, it has been proposed that C646 is a general inhibitor for both CBP and p300 *(11)*. Of importance for the role of ICG-001 as a chemotherapeutic drug, studies suggest that CBP-driven transcription helps to maintain pluripotency whereas p300-driven transcription pushes cells toward a differentiated state *(3,12-15)*; examples of genes thought to be regulated by CBP vs. p300 are shown in **Appendix A-1C**. However, the hypothesis that ICG-001 specifically downregulates only the subset of WNT target genes involved in proliferation (such as BIRC5 and CCND1) has not been tested on a genome-wide scale. Because a derivative of ICG-001 called PRI-724 is now in clinical trials (NCT01302405 and NCT01606579), it is critical to have a thorough understanding of the specificity and effectiveness of this drug. Therefore, we have compared the genome-wide effects on the transcriptome of ICG-001 versus C646 in two cancer cell lines that constitutively express the components of the transcription complex that mediates WNT signaling (**Appendix A-1D**).

Results

ICG-001 and C646 have similar effects on the transcriptome of HCT116 colon cancer cells.

Constitutive activation of WNT target genes via a TCF/β-catenin/CBP complex is thought to be a major driver of colorectal cancer. Therefore, it has been proposed that treatment of colon cancer cells with ICG-001 should specifically inhibit the WNT pathway (by preventing recruitment of the co-activator CBP to TCF/β-catenin target genes) and reduce the tumorigenicity of the cells. In support of this hypothesis, Emami et al. (*10*) have shown that ICG-001 reduces growth of colon carcinoma cells in culture and reduces the formation of colon and small intestinal polyps in a mouse model system. As noted above, CBP is highly related to another HAT called p300 and many studies have shown similar functions for p300 and CBP *(16)*. In fact, a ChIP-seq analysis of p300 and CBP in T98G glioblastoma cells immediately after release from serum starvation arrest showed that almost all of the CBP genomic binding sites were also bound by p300 *(17)*. However, under the tested conditions, a small set of genomic sites were preferentially bound by either CBP or p300, suggesting that there might be some specificity in their action. It is also possible that cell type plays a critical role in specifying CBP vs. p300 contributions to regulating the transcriptome. For example, approximately 50% of Rubinstein-Taybi syndrome patients have mutations in CBP but only 3% of patients have mutations in p300 *(18)*. Of course, functional specificity can also occur post-DNA binding because the two HATs only share extensive, but not complete, homology. If, for example, CBP and p300 recruit different interaction partners they could have opposite effects on transcription at a given promoter. In support of this hypothesis, Ma et al. have

shown that both CBP and p300 can bind to the BIRC5 promoter but they have opposite effects on transcription *(19)*.

To determine if the effects on the transcriptome after specifically inhibiting CBP are different than the effects after inhibiting both CBP and p300, we treated HCT116 colon cancer cells with 0.05% DMSO, 10 μM ICG-001, or 10 μM C646 for 12 and 96 h. Samples were prepared in replicate and Illumina HumanHT-12 v4 expression arrays were used to detect changes in gene expression (**Appendix A-2**). Genes having a detection P value less than 0.01 in any of the control or treated cell populations were selected for further analysis; this constituted a total of 15,092 genes from HCT116 cells, of which 3,689 showed differential expression in drug-treated cells (differential expression P value less than 0.05). After selecting the significant differentially expressed genes, the expression fold change was calculated for each gene and Euclidean distance was used for K-means clustering of expression fold change (**Appendix A-3**). We found that, contrary to our initial expectations, a very similar response was observed for both drugs. Genes that were downregulated by both drugs were involved in the cell cycle and WNT signaling (**Appendix A-3**). However, some genes did show drug-specific changes in HCT116 cells. According to the mechanism of action of each drug, genes with decreased levels of expression only after treatment with ICG-001 should be regulated by CBP but not by p300, whereas genes with decreased levels of expression only after treatment with C646 but not with ICG-001 should be regulated by p300 but not by CBP. A gene ontology analysis of the approximately 400 genes affected only by ICG-001 revealed a strong enrichment for genes controlling the cell cycle whereas the approximately 500 genes only affected by C646 were not related to cell proliferation. Thus, in HCT116 cells,

both drugs have a broad effect on gene regulation that includes downregulation of genes involved in proliferation control. However, treatment of colorectal cancer cells with ICG-001 alters the expression of a greater number of cell cycle-regulated genes than does treatment with C646.

*ICG-001 and C646 have similar effects on the transcriptome of PANC1 cells***.**

As noted above, the WNT/TCF/β-catenin/CBP path- way has been proposed to be a major positive regulator of proliferation of colon cancer cells. Perhaps β-catenin/ CBP complexes play a prominent role in WNT-mediated gene expression in HCT116 cells (with little contribution by β-catenin/p300 complexes), explaining why the effects of ICG-001 were so widespread and why treatment with the two drugs elicited similar responses. To determine if ICG-001 has a similar widespread effect on other cancer cells, we also examined pancreatic cancer cells. Pancreatic ductal adenocarcinoma, the most common form of pancreatic cancer, displays activation of the WNT/β-catenin pathway *(20-25)* and is therefore predicted to respond to treatment with ICG-001. We treated PANC1 cells with ICG-001 or C646 and analyzed gene expression. Again, we found that ICG-001 and C646 have similar effects on PANC1 cells (**Appendix A-4**), with genes involved in cell cycle regulation being down-regulated by both drugs. However, in this case, cell proliferation-related genes were not enriched categories in gene sets downregulated specifically by either ICG-001 or C646. Interestingly, in PANC1 cells, the cholesterol biosynthesis pathway was highly enriched for genes specifically downregulated by ICG-001, suggesting that perhaps genes involved in cholesterol biosynthesis are specific CBP, but not p300, target genes. In contrast, p300-specific genes

(identified as those responsive only to C646) appear to be involved in various types of signaling pathways, including PI3K/AKT signaling which is linked to cell survival. To determine if gene responses to the drugs were cell type-specific, we compared the genes whose expression was altered by ICG-001 or C646 in both HCT116 and PANC1 cells (a total of 6,732 genes). Genes that were significantly detected in HCT116 or in PANC1 cells (*P* value <0.01) and which had a differential *P* value <0.05 and a fold change greater than 1.2 (5,182 genes) were compared using hierarchical clustering with Euclidean distance and average linkage measures (**Appendix A-5**). We found that although some genes were altered in a cell type-specific manner, most genes were similarly affected in both cell types. A gene ontology analysis revealed that the top two categories of genes downregulated by ICG-001 or C646 in both HCT116 and PANC1 cells were oxidative phosphorylation and mitochondrial dysfunction. Genes that were commonly upregulated by the drugs in both cell types are involved in pathways such as death receptor signaling and INOS signaling.

Effectiveness of the epigenetic inhibitors in reverting a tumor cell phenotype.

The ultimate goal of epigenetic therapy is to revert the transcriptome from a tumor-specific pattern of gene expression back to the expression patterns seen in nor- mal cells. To determine the extent to which the epigenetic inhibitors ICG-001 and C646 are effective in this goal, we obtained RNA-seq expression data for 41 normal and 274 tumor colon cells from the TCGA Consortium. Using this data, we identified 16,416 genes that were expressed in either normal or colon samples, using log2 (RSEM + 1) >2. Of these, 11,824 genes were differentially expressed (adjusted differential *P* value <0.001) in the

tumor samples as compared to the normal tissues. To determine if the drugs were effective in reverting the expression of these genes back to normal levels, we compared the set of genes deregulated in the tumors with the set of genes responsive to the drug treatments, identifying a set of 2,028 common genes. If the drugs are having an anti-tumor effect, then genes that are up- regulated in tumors should be downregulated by the drugs and genes that are downregulated in tumors should be up- regulated by the drug. Using a log2(RSEM + 1) cutoff of 2, we identified 2,029 genes that showed expression changes (adjusted P value <0.05) in colon tumor cells, as compared to the normal tissues. An analysis of these expression patterns (**Appendix A-6**) shows that many genes had expression changes in the correct direction as a result of treatment with at least one drug (that is, a gene that is upregulated in tumors was down- regulated by a drug or a gene that is downregulated in tumors was upregulated by a drug). Analysis of four normal and 125 pancreatic tumor samples revealed a much smaller set of genes showing expression changes in tumors. Using a log2(RSEM + 1) cutoff of 2, we identified only 167 genes that showed expression changes (adjusted P value <0.05) in pancreatic tumor cells, as compared to the normal tissues. It is unclear as to whether the small number of differentially expressed genes in the pancreatic tumors as compared to the colon tumors is due to real differences in cancer phenotypes, to the small number of normal pancreatic samples, or other possibilities such as tumor heterogeneity. To increase the number of analyzed genes, we also obtained a list of 596 genes that are differentially expressed in normal hTERT-HPNE pancreatic cells as compared to PANC1 cells *(26)*. We examined the responses of the 167 genes that are differentially regulated in normal pancreatic tissue vs. tumors and the 596 genes that are differentially regulated in normal HPNE cells

grown in culture vs. PANC1 cells to drug treatment. We found that many of the genes whose expression is deregulated in pancreatic tumors or PANC1 cells showed appropriate responses to at least one drug (that is, genes upregulated in tumors or PANC1 were downregulated by the drugs and genes downregulated in tumors or PANC1 were upregulated by the drugs) (**Appendix A-6**). Thus, treatment with the epigenetic inhibitors is effective in reverting some of the tumor-specific transcriptome to a normal pattern.

Direct targeting of a component of the transcription complex that mediates WNT signaling.

As described above, ICG-001 was developed to be a specific inhibitor of the WNT pathway. We therefore directly analyzed the WNT pathway using a list of genes previously implicated as components of this pathway (http://www.stanford.edu/group/nusselab/cgi-bin/wnt/). We found that a subset of these proposed WNT target genes were expressed in HCT116 and/or PANC1 cells and were significantly affected by treatment with ICG-001 or C646 (**Appendix A-7**). The overall trend of the effects of ICG-001 and C646 on WNT targets was similar in a given cell line. However, the WNT pathway-related genes responded quite differently to the epigenetic inhibitors in the different cell lines. In general, the response of the genes listed in **Appendix A-1C** was more similar to what was predicted when HCT116 cells were treated with the epigenetic inhibitors than when PANC1 cells were treated with the drugs. For example, expression of the transcription factor *JUN* (which is involved in specifying differentiated phenotypes) is increased by both drugs in HCT116 but is decreased by both drugs in

PANC1. Conversely, the expression of *MYC*, a transcription factor involved in cell proliferation, is reduced by both drugs in HCT116 but is increased by both drugs in PANC1 cells. The gene ontology results suggest that ICG-001 and C646 affect the WNT pathway in HCT116 cells but not in PANC1 cells. Of course, it is also possible that different downstream target genes mediate the WNT pathway in pancreatic cancer cells as compared to colon cancer cells. The HATs CBP and p300 are brought to genomic regulatory elements by the DNA binding protein TCF7L2 via interaction with the bridging protein β-catenin. If ICG-001 and C646, which block the recruitment or function of the HAT activity of the co-activators CBP and p300, are specific inhibitors of the WNT signaling pathway in PANC1 cells, then targeting TCF7L2 should result in similar effects on the transcriptome as does drug treatment. In contrast, if the epigenetic inhibitors are in fact targeting a different pathway in PANC1 cells, then genes affected by reduction of TCF7L2 should be different than the set of genes affected by the drugs. To identify genes affected by direct targeting of a component of the transcriptional complex implicated in WNT regulation, we used siRNAs to knockdown *TCF7L2* in PANC1 cells. Cells were treated with control siRNAs or siRNAs specific for *TCF7L2* and RNA was analyzed by RNA-seq. We analyzed the top 1,000 genes that were affected by knockdown of *TCF7L2* and the top 1,000 genes affected by treatment with ICG-001 (**Appendix A-8**). Interestingly, there were very few genes affected by reduction of *TCF7L2* that were also affected by ICG-001. Specifically, the WNT pathway was identified in the set of genes affected upon reduction of *TCF7L2* but not by treatment with ICG-001 (**Appendix A-4**). These results suggest that in PANC1 cells co-activators other than CBP cooperate with TCF7L2 to regulate gene expression and support the

hypothesis that the anti-proliferative effects of ICG-001 in PANC1 cells are not due to inhibition of the WNT pathway.

Discussion

Recent studies have shown large changes in the epigenomic patterns in normal vs. cancer cells, suggesting that epigenetic therapy may be commonly applicable to treatments of various cancers. Drugs that target epigenetic regulators are being developed *(27-29)*, some of which are moving into clinical trials. However, the specificity of action of many of these drugs has not yet been thoroughly examined. In particular, genome-wide analyses of their effects have not been determined. In our study, we compare the effects of treatment with C646, which is thought to compete with acetyl-coA for the Lys-coA binding pocket of both p300 and CBP *(11)* to the effects of ICG-001, which specifically binds to CBP and prevents its interaction with the co-activator β-catenin. Theoretically, ICG-001 is expected to be of higher specificity than C646 because it should only affect β-catenin/ CBP-driven transcription whereas C646 should affect all genes regulated by either CBP or p300, regardless of whether β-catenin is involved. However, it is possible that ICG-001 has broader effects than anticipated if the drug affects the ability of CBP to interact with other as- of-yet unknown co-activators. In addition, we note that CBP and p300 can acetylate non-histone proteins *(30)*; thus, both compounds could also have effects on non- chromatin bound proteins. Although we initially expected cells to respond differently to C646 and ICG-001, our results suggest that generally these two drugs have similar effects on the transcriptome of tumor cells.

However, we did identify some cell-specific and drug-specific responses after epigenetic inhibition.

We observed dramatic effects on the transcriptome upon treatment of HCT116 colon cancer cells with either ICG-001 or C646, with thousands of genes showing differential expression. Interestingly, the responses to the two drugs were quite similar overall, with both drugs causing a reduction in certain genes involved in the WNT pathway. Because ICG-001 affects only CBP-driven transcription and not p300-driven transcription, these results suggest that perhaps the majority of the WNT-related active regulatory elements in HCT116 cells are bound by β-catenin/CBP complexes. We did identify a set of approximately 500 genes whose expression was down- regulated by ICG-001 and not by C646 (these are potential CBP-specific target genes) and a set of approximately 500 genes whose expression was downregulated by C646 but not by ICG-001 (these are potential p300-specific target genes). These results are similar to a previous study of CBP and p300 in T98G glioblastoma cells that found that the two factors bound mainly to the same sites but that some specific binding sites could be identified (*17*). Interestingly, the genes specifically responsive to ICG-001 but not to C646 in HCT116 cells showed enrichment for cell proliferation-related gene ontology categories. Taken together, these results suggest that thousands of genes are regulated both by p300 and CBP (many of which are involved in cell proliferation) and that CBP-specific genes may also include additional genes that regulate cell proliferation whereas p300-specific genes are involved in other processes. In general, our results in HCT116 cells support the current model implicating WNT-mediated cell signaling as a critical regulator of cancer cell proliferation.

Although the WNT pathway has been implicated in the development of pancreatic cancer, the studies are not as extensive as those related to WNT's role in colon cancer *(20-25)*. We show that, in general, the effects of ICG-001 and C646 on the transcriptome of PANC1 cells are similar to those observed upon treatment of HCT116 cells. For example, a set of genes involved in cell proliferation show reduced expression upon treatment of PANC1 with either ICG-001 or C646. However, we did observe several differences in the response of PANC1 cells to the epigenetic inhibitors, as compared to HCT116 cells. First, we found that many of the enriched gene categories that responded specifically to ICG-001 treatment of PANC1 cells are involved in cholesterol biosynthesis. Interestingly, many cancers have a high dependency on accelerated biogenesis and uptake of lipids and cholesterol and inhibition of these pathways has been proposed to be a therapeutic opportunity for metabolic targeting of cancer growth *(31,32)*. Cholesterol homeostasis in mammalian cells is maintained in part by a basic-helix-loop-helix family of transcription factors called the sterol regulatory element binding proteins (SREBPs) *(33,34)*. The SREBP family members activate a number of target genes involved in cholesterol and fatty acid metabolism through binding to sterol regulatory elements in the promoters of target genes. In fact, SREBP transcription factors have been suggested to be novel therapeutic targets *(35)*. Interestingly, SREBP proteins require interaction with CBP to mediate transcriptional activation *(36)*. Thus, the treatment of PANC1 cells with ICG-001 likely disrupts a functional interaction between CBP and a SREBP family member, causing downregulation of genes involved in the cholesterol biosynthetic pathway (**Appendix A-9**). Second, in PANC1 cells the WNT pathway was not enriched in downregulated genes after treatment with either drug and

several critical WNT target genes showed unexpected transcriptional responses. Notably, expression of *JUN* (which promotes differentiation) was predicted to be increased upon treatment but in PANC1 cells *JUN* expression was decreased (*JUN* did show the expected response in HCT116 cells). Similarly, expression of *MYC* (which promotes proliferation) was predicted to be de- creased upon treatment but in PANC1 cells *MYC* expression was increased (*MYC* did show the expected response in HCT116 cells). The transcriptional response of the *MYC* gene was particularly surprising because it is considered to be a critical mediator of WNT signaling. Upregulation of *MYC* in PANC1 suggests that the drugs do not inhibit the WNT pathway in these cells. This hypothesis is supported by our finding that in PANC1 cells knockdown of TCF7L2, the transcription factor that brings β-catenin and CBP to regulatory elements to regulate WNT-responsive genes, does not affect expression of the same genes as are affected by treatment with ICG-001. While our work was in progress, another group reported treatment of pancreatic cancer cells with ICG- 001 *(37)*. They showed that treatment of PANC1 cells with 10uM ICG-001 was effective at reducing cell proliferation in culture and reducing colony formation in soft agar. Al- though global effects on the PANC1 transcriptome were not examined in that study, the noted effects on proliferation are consistent with our finding that cell cycle-related genes are downregulated in response to ICG-001 and C646. That study did, however, perform microarray expression analysis after treatment of a different pancreatic cancer cell line (AsPC-1) with ICG-001 and found that 569 transcripts were upregulated and 150 transcripts were downregulated. Because only 117 of the 719 drug-responsive genes were altered in β-catenin knockdown cells, they concluded that ICG-

001 had a broader effect than simply as a disrupter of WNT/β-catenin signaling in AsPC-1 cells.

As noted above, epigenetic inhibitors are considered promising new drugs for cancer treatment. One current clinical trial employs PRI-724, a derivative of ICG-001, in combination with gemcitabine in patients with advanced or metastatic pancreatic adenocarcinoma (NCT01764477). Gemcitabine is considered a first-line treatment for pancreatic adenocarcinoma but has poor overall efficacy because pancreatic cancer cells develop resistance to the drug *(38)*. While investigating the pathways that lead to drug resistance, the transcriptional regulator NUPR1 (also known as anti-apoptotic protein p8 or Candidate of Metastasis-1) was identified as being involved in the acquisition of gemcitabine resistance by pancreatic cancer cells *(39)*. NUPR1 normally functions as a stress response gene in the pancreas, but it has been shown to contribute to metastasis, anti-apoptotic activity and pancreatic cancer development *(40,41)*. Interestingly, our genome-wide analyses identified *NUPR1* as one of the top upregulated genes after treatment of PANC1 cells with ICG-001. The upregulation of *NUPR1* by ICG-001 may explain why ICG-001 plus gemcitabine did not increase overall lifespan in an *in vivo* pancreatic cancer cell xenograft model *(37)*. Although the mechanism by which NUPR1 promotes oncogenesis and/or drug resistance in pancreatic cells is not yet known, NUPR1 has been shown to form a complex with p300 and TP53 to upregulate and promote cytoplasmic translocation of CDKN1A (p21) in breast cancer cells *(42)*. Although nuclear p21 is a negative regulator of cell cycle progression, studies have associated cytoplasmic p21 with drug resistance and oncogenic activity in breast and testicular cancer *(43-45)*. Vincent *et al. (44)* showed that treatment of NUPR1-expressing cells with

PI3K-AKT inhibitors could reverse cytoplasmic p21 localization and re-sensitize cells to doxorubicin. Importantly, studies have also shown that inhibition of the PI3K-AKT pathway in pancreatic cancer helps re-sensitize cells to gemcitabine *(46,47)*. Thus, adding a PI3K-AKT inhibitor to the combined usage of ICG-001 plus gemcitabine may be the most effective treatment combination. However, it should also be noted that C646 caused only a modest increase in *NUPR1* in PANC1 cells and that C646, but not ICG- 001, specifically inhibited the PI3K-AKT pathway (see Figure 4). Taken together, these results suggest that per- haps C646 plus gemcitabine would be more effective than ICG-001 plus gemcitabine in the treatment of pancreatic cancer.

Conclusions

We have compared the genome-wide effects on the transcriptome of ICG-001 (a specific CBP inhibitor) versus C646 (a compound that competes with acetyl-coA for the Lys-coA binding pocket of both CBP and p300). We found that ICG-001 has a similar broad specificity as C646 in HCT116 colon cancer, with both drugs decreasing the expression of cell cycle-related and WNT pathway genes. In contrast, ICG-001 and C646 affect cell cycle- related genes but do not result in appropriate responses of critical WNT target genes in PANC1 cancer cells. The effects of ICG-001 on PANC1 cells and comparison to gene expression patterns in TCF7L2 knockdown cells suggests that ICG-001 inhibits proliferation of pancreatic cancer cells via a mechanism different than the WNT pathway. Gene ontology analyses point toward disruption of SREBP-CBP functional interactions as a possible cause of the anti-proliferative function of ICG-001 in pancreatic cancer cells. Importantly, both epigenetic inhibitors are effective at reversing

some tumor-specific changes in gene expression that are observed in colon or pancreatic tumor cells.

Methods

Cell growth conditions

The human cell lines HCT116 (ATCC #CCL-247) and PANC1 (ATCC #CRL-1469) were obtained from the American Type Culture Collection. HCT116 and PANC1 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. Michael Kahn (University of Southern California) provided ICG-001 and C646 was obtained from VWR (catalog # 102516–240). Cells were treated with 10 μM ICG-001, 10 μM C646, or 0.05% DMSO and collected after 12 or 96 h. Cells for the 12-h treatments were grown to 70% confluency before addition of the drugs or DMSO. Cells for the 96-h treatments were grown at 40% to 50% confluency before addition of the drugs or DMSO and were passaged before they could reach 90% confluency. New media and drugs were added every 24 h. After treatment, gene expression was analyzed using Illumina BeadChips.

Microarray RNA expression

Total RNA was collected using Trizol according to the manufacturer's instructions (Life Technologies). To confirm RNA samples were not degraded, RNA quality was checked with the Experion StdSens kit (Bio-Rad) prior to amplification and labeling. The Illumina TotalPrep RNA Amplification Kit (Life Technologies catalog # AMIL1791) was used according to the manufacturer's instructions to amplify and label

RNA samples for Illumina array hybridization. Labeled RNAs were analyzed with Illumina HT-12 v4 Expression BeadChips (Catalog #: BD-103-0204) with the Direct Hybridization Assay and then scanned on an Illumina HiScan (catalog # BD-103- 0604). The data were analyzed and exported from Illumina's GenomeStudio software using quantile normalization with- out background subtraction. Each drug/DMSO treatment and time point was performed using two independent bio- logical replicates. The correlation between replicates was calculated to ensure that the data were reproducible, replicate samples were averaged together and genes with a detection *P* value <0.01 were considered for further analysis. Differential expression analysis was performed using Illumina's custom differential expression error model, which assumes a normal distribution of the target signal intensity and takes into account biological variation, non-specific biological variation, and technical error. For more detail on Illumina's custom error model, see GenomeStudio Gene Expression Module v1.0 User Guide (pages 103 and 104). Genes with a differential expression *P* value <0.05 were considered to be significantly differentially expressed.

TCF7L2 knockdown

TCF7L2 knockdown was performed in triplicate by siRNA transfection. Transfections were performed using Lipofectamine RNAiMax (Life Technologies) according to manufacturer's instructions. A final concentration of 40nM siRNAs targeting either *TCF7L2* (catalog # 4392420, Life Technologies) or a non-specific negative control siRNA (catalog # AM4611, Life Technologies) were used using reduced serum OptiMEM media (Life Technologies). Media was changed 12 h post transfection

and total RNA was collected 48 h post transfection using Trizol according to manufacturer's instructions (Life Technologies). Knockdown efficiency was detected using RT-qPCR and then samples were analyzed by RNA-seq.

RNA-Seq

Total RNA was used for polyA+ RNA selection using oligo-dT beads and subjected to library construction by True-Seq library preparation kits (Illumina), followed by Illumina HiSeq2000 sequencing. The RNA-seq reads were aligned to the human genome hg19 using Bowtie2 with ultrasensitive parameters. The RNA-seq reads were counted over gene exons using HTSeq *(48)*. EdgeR was used for statistical analyses of siControl and siTCF7L2 samples, and a fold change of 2 was used to call the differentially expressed genes *(49)*.

Ingenuity pathway analysis

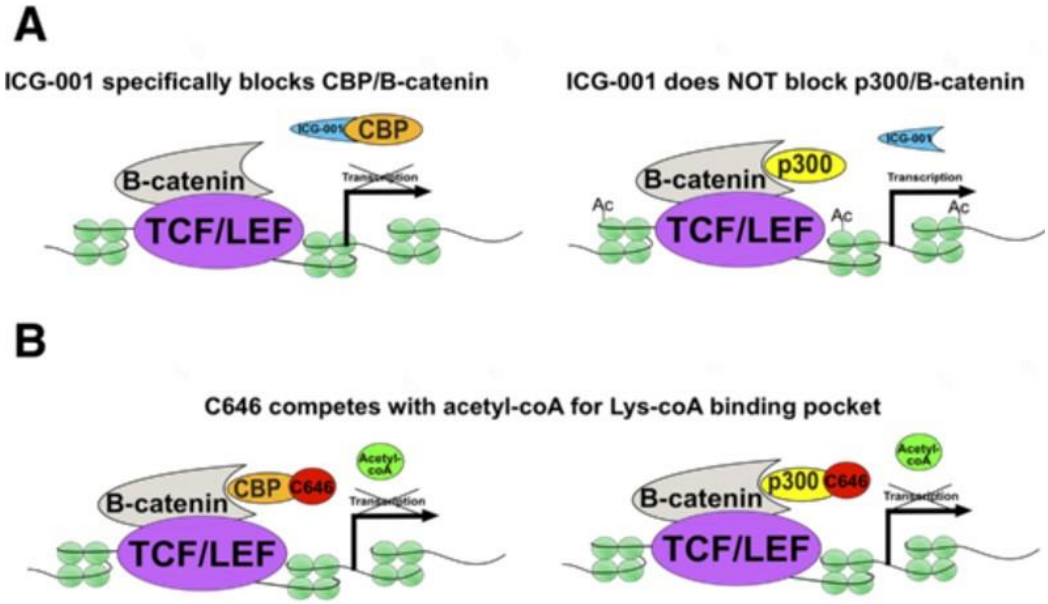
Gene network diagrams in **Appendix A-9** were created through use of IPA. The expression data were analyzed through the use of QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). For each subset of genes, a core analysis was run with parameters set to consider only direct relationships and relationships between molecules that have been experimentally observed. The reference gene set used for *P* value calculations was the Ingenuity Knowledge Base (genes only).

Data access

Expression array analyses for control and treated cells and RNA-seq datasets for *TCF7L2* knockdown experiments have been deposited in GEO (GSE64039 and GSE63776). The TCGA RNA-seq can be downloaded at https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp.

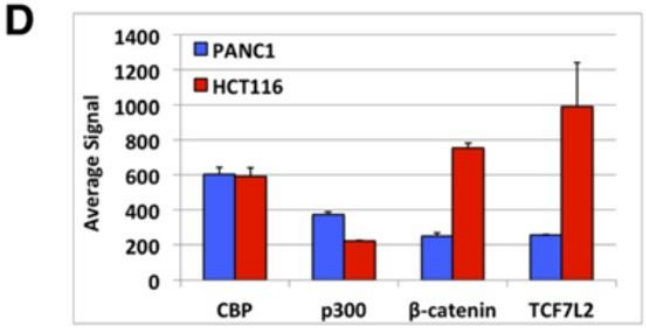
Figures

Appendix A-1: Targeting the WNT pathway using epigenetic inhibitors



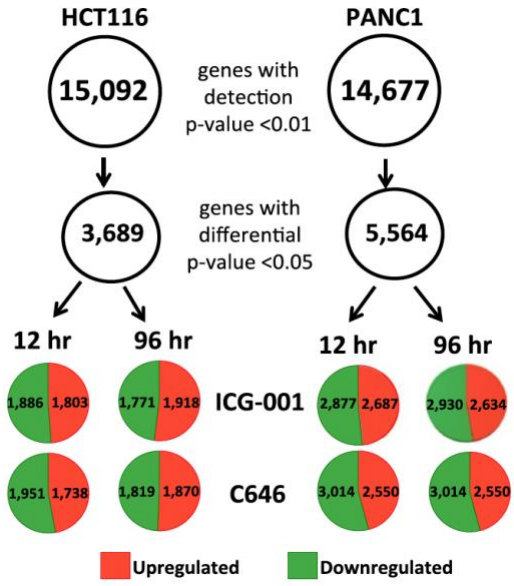
C

Co-regulator	Down-regulated genes	Up-regulated genes	Outcome
p300	CCND1, POU5F1, BIRC5,S100A4, AXIN2	JUN, FOSL1	Differentiation
CBP	JUN, FOSL1	CCND1, POU5F1, BIRC5,S100A4, AXIN2	Stem cell maintenance, proliferation



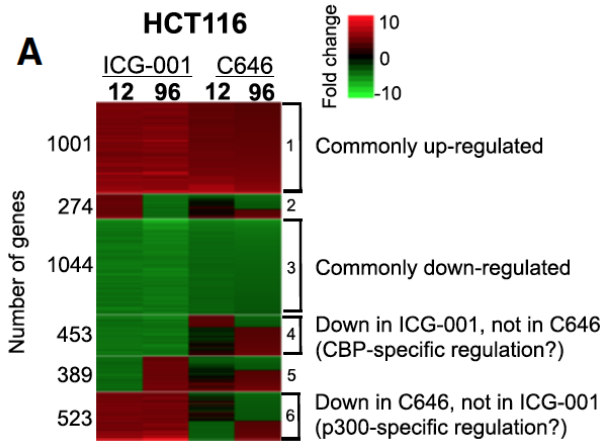
Appendix A-1. Targeting the WNT pathway using epigenetic inhibitors. WNT signaling culminates when, upon recruitment of β-catenin/CBP or β-catenin/p300 complexes to the DNA via a TCF/LEF family member, CBP and p300 activate transcription by acetylating histone H3. (A) Treatment with ICG-001 disrupts the interaction of CBP with β-catenin, blocking CBP-driven, but not p300-driven transcription. (B) In contrast to the effects of ICG-001, C646 competes with acetyl-coA for the Lys-coA binding pocket of both CBP and p300, preventing HAT activity of both complexes. (C) Examples of predicted gene expression differences mediated by β-catenin/CBP vs. β-catenin/p300 complexes *(15)*. (D) RNA levels in HCT116 and PANC1 cells of the various components of the WNT signaling model.

Appendix A-2. The effects of epigenetic inhibitors on the transcriptome of HCT116 and PANC1 cells



Appendix A-2. The effects of epigenetic inhibitors on the transcriptome of HCT116 and PANC1 cells. HCT116 colon cancer cells and PANC1 pancreatic adenocarcinoma cells were treated in duplicate with DMSO or 10 μM ICG-001 or C646 for 12 or 96 h (12 samples per cell line). Cells were harvested and RNA was analyzed using Illumina HumanHT-12 v4 expression arrays. Any gene having a detection *P* value <0.01 in any of the samples was selected for differential gene analysis; genes having a differential *P* value <0.05 were further analyzed. The number of upregulated (red) and downregulated (green) genes under each condition for each cell line is shown.

Appendix A-3. Effects of epigenetic inhibitors on gene expression in HCT116 cells



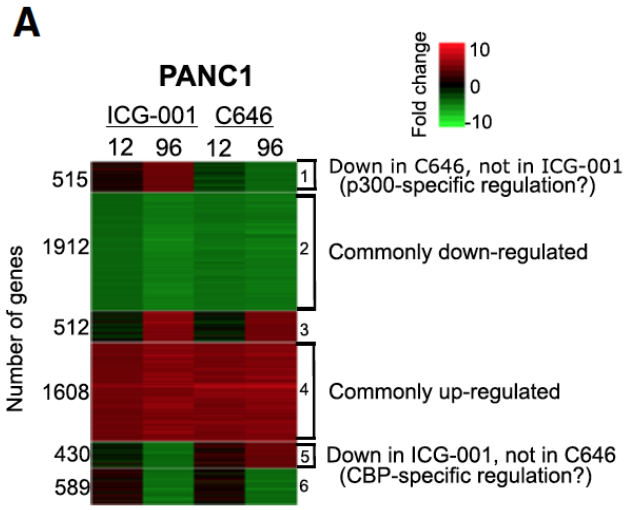
B

Up in ICG-001 and C646		Down in ICG-001 and C646		Down in ICG-001 only		Down in C646 only	
Pathway	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**
Death Receptor Signaling	4.0E-05	Mitochondrial Dysfunction	4.0E-11	Cell Cycle: G2/M DNA Damage Checkpoint	1.9E-08	Unfolded protein response	1.1E-06
TWEAK Signaling	2.3E-04	Oxidative Phosphorylation	3.2E-09	Cholesterol Biosynthesis	6.2E-07	Tetrahydrofolate Salvage	7.8E-05
tRNA Charging	5.6E-04	EIF2 Signaling	3.6E-07	Mitotic Roles of Polo-Like Kinase	3.6E-06	Adipogenesis Pathway	2.0E-04
Protein Ubiquitination	6.9E-04	Cholesterol Biosynthesis	5.4E-05	Cell Cycle Control of Chromosomal Replication	8.5E-06	Glucocorticoid Receptor Signaling	2.6E-04
Hypoxia Signaling	7.1E-04	eIF4 and p70S6K Signaling	1.0E-04	Geranylgeranyldiphosphate Biosynthesis	1.1E-05	Histidine Degradation	4.1E-04
Heparan Sulfate Biosynthesis	1.1E-03	Antiproliferative Role of TOB in T Cell Signaling	3.0E-04	Mevalonate Pathway	7.4E-05	Folate Transformation	6.0E-04
NRF2-mediated Oxidative Stress	3.5E-03	Wnt/β-catenin Signaling	5.4E-04	DNA damage-induced 14-3-3 Signaling	3.7E-04	Endoplasmic Reticulum Stress Pathway	7.2E-04
Chondroitin Sulfate Biosynthesis	3.6E-03	Cyclins and Cell Cycle Regulation	5.8E-04	p53 Signaling	4.9E-04	Death Receptor Signaling	2.5E-03
Dermatan Sulfate Biosynthesis	4.5E-03	BER Pathway	7.2E-04	CHK Proteins in Cell Cycle Checkpoint Control	5.4E-04	Assembly of RNA Polymerase II Complex	3.2E-03
Induction of Apoptosis	7.2E-03	RAR Activation	8.3E-04	ATM Signaling	7.8E-04	JAK/Stat Signaling	3.2E-03

Appendix A-3. Effects of epigenetic inhibitors on gene expression in HCT116 cells. (A) Genes differentially expressed after treatment of HCT116 cells with ICG-001 or C646 (see Figure 2) were analyzed using Euclidean distance and K-means clustering of expression fold change. (B) Gene ontology analyses are shown for the genes commonly up- and downregulated by both drugs and for the genes that are downregulated only by one of the drugs in HCT116 cells. Terms related to the cell cycle are shown in red and terms related to WNT signaling are shown in blue. The numbers 1 to 6 in the brackets in panel A refer to different clusters that were used in the gene ontology analyses shown in panel B.

Appendix A-4. Effects of epigenetic inhibitors on gene expression in PANC1 cells

A

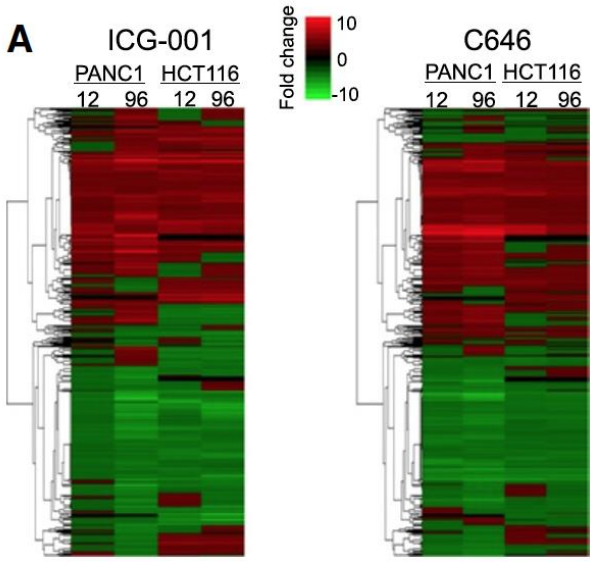


B

Up in ICG-001 and C646		Down in ICG-001 and C646		Down in ICG-001 only		Down in C646 only	
Pathway	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**
Death Receptor Signaling	1.0E-04	EIF2 Signaling	3.2E-15	Cholesterol Biosynthesis	7.9E-16	EIF2 Signaling	8.1E-08
EIF2 Signaling	7.2E-04	EIF4 and p70S6K Signaling	1.3E-08	Cholesterol Biosynthesis I	1.0E-11	PI3K/AKT Signaling	3.5E-06
MethylmalonylPathway	8.1E-04	RAN Signaling	4.7E-08	Cholesterol Biosynthesis II	1.0E-11	RANK Signaling	5.9E-06
iNOS Signaling	1.0E-03	mTOR Signaling	6.6E-08	Cholesterol Biosynthesis III	1.0E-11	EIF4 and p70S6K Signaling	2.1E-05
Interferon Signaling	1.1E-03	Cyclins and Cell Cycle Regulation	4.7E-07	Geranylgeranyldi-phosphate Biosynthesis	1.2E-05	IL-8 Signaling	4.5E-05
Prostate Cancer Signaling	1.1E-03	Mitotic Roles of Polo-Like Kinase	6.5E-07	TCA Cycle	4.5E-05	CD27 Signaling	5.0E-05
Activation of IRF	1.3E-03	Androgen Signaling	1.0E-06	Mevalonate Pathway	8.5E-05	PTEN Signaling	7.8E-05
Estrogen Receptor Signaling	1.5E-03	Remodeling of Epithelial Adherens Junctions	1.0E-06	Mitochondrial Dysfunction	1.0E-03	IL-17A Signaling	1.9E-04
2-oxobutanoate Degradation	1.9E-03	Cell Cycle: G2/M DNA Damage Checkpoint Regulation	1.1E-06	Aldosterone Signaling	1.6E-03	Molecular Mechanisms of Cancer	2.0E-04
PDGF Signaling	2.0E-03	Mitochondrial Dysfunction	1.8E-06	Gap Junction Signaling	1.9E-03	CD40 Signaling	2.1E-04

Appendix A-4. Effects of epigenetic inhibitors on gene expression in PANC1 cells. (A) Genes differentially expressed after treatment of PANC1 cells with ICG-001 or C646 (see Figure 2) were analyzed using Euclidean distance and K-means clustering of expression fold change. (B) Gene ontology analyses are shown for the genes commonly up- and downregulated by both drugs and for the genes that are downregulated only by one of the drugs in PANC1 cells. Terms related to the cell cycle are shown in red. The numbers 1 to 6 in the brackets in panel A refer to different clusters that were used in the gene ontology analyses shown in Panel B.

Appendix A-5. ICG-001 and C646 affect many of the same genes in HCT116 and PANC1 cells

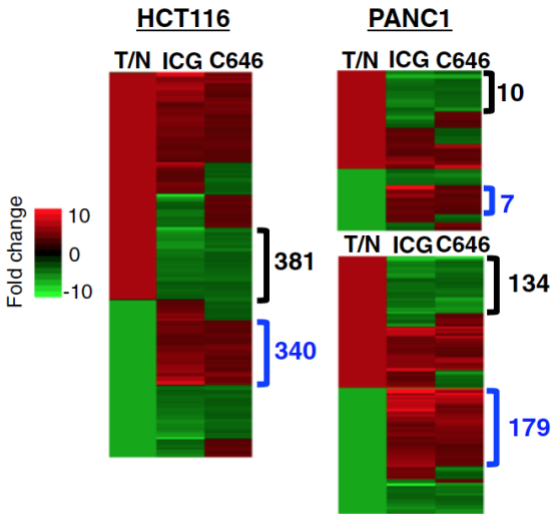


B

Downregulated by ICG in both HCT116 and PANC1		Upregulated by ICG in HCT116 and PANC1		Downregulated by C646 in both HCT116 and PANC1		Upregulated by C646 in both both HCT116 and PANC1	
Pathway	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**
Oxidative Phosphorylation	3.9E-09	Death Receptor Signaling	1.5E-04	Mitochondrial Dysfunction	1.0E-08	BRCA1 and DNA Damage Response	2.1E-04
Mitochondrial Dysfunction	4.8E-09	Retinoic acid Mediated Apoptosis Signaling	8.1E-04	Oxidative Phosphorylation	1.1E-07	PI3K Signaling	9.1E-04
EIF2 Signaling	1.0E-06	ATM Signaling	6.2E-03	Androgen Signaling	4.2E-06	Death Receptor Signaling	3.8E-03
Androgen Signaling	1.9E-05	iNOS Signaling	7.6E-03	Cyclins and Cell Cycle Regulation	1.1E-05	D-myo-inositol (1,4,5)-trisphosphate Metabolism	8.3E-03
Mitotic Roles of Polo-Like Kinase	7.8E-05	TWEAK Signaling	7.9E-03	EIF2 Signaling	1.7E-05	D-myo-inositol (1,4,5)-trisphosphate Degradation	1.0E-02
Remodeling of Epithelial Adherens Junctions	1.0E-04	Activation of IRF by Cytosolic Pattern Recognition Receptors	9.5E-03	Cell Cycle: G1/S Checkpoint Regulation	1.4E-04	iNOS Signaling	1.1E-02
BER Pathway	1.2E-04	Assembly of RNA Polymerase II Complex	1.0E-02	Cell Cycle Regulation by BTG Family Proteins	1.4E-04	Circadian Rhythm Signaling	1.2E-02
Epithelial Adherens Junction Signaling	1.4E-04	p38 MAPK Signaling	1.2E-02	Protein Ubiquitination Pathway	1.5E-04	Urate Biosynthesis/ Inosine 5'-phosphate Degradation	1.2E-02
Assembly of RNA Polymerase II Complex	1.9E-04	Heparan Sulfate Biosynthesis	1.4E-02	HIPPO Signaling	1.6E-04	Aryl Hydrocarbon Receptor Signaling	1.5E-02
Telomerase Signaling	2.2E-04	Chondroitin Sulfate Biosynthesis	1.4E-02	Telomerase Signaling	1.6E-04	iCOS-iCOSL Signaling	1.5E-02
Estrogen Receptor Signaling	2.5E-04	Dermatan Sulfate Biosynthesis	1.5E-02	Estrogen Receptor Signaling	2.2E-04	Heparan Sulfate Biosynthesis	2.0E-02

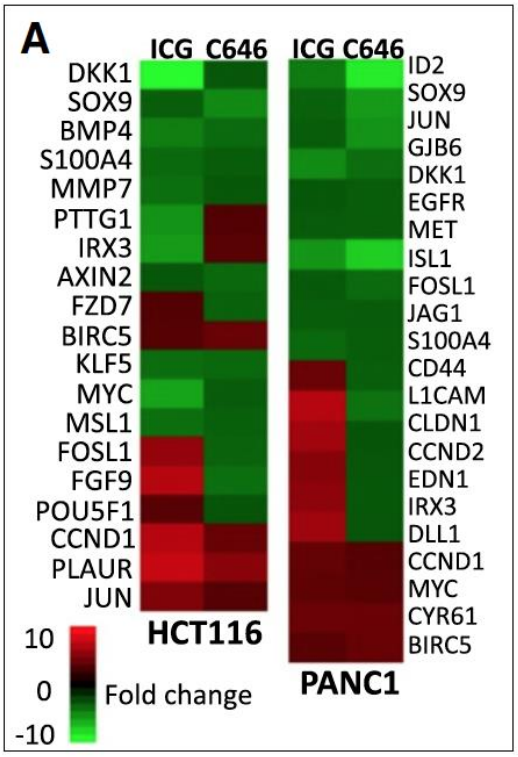
Appendix A-5. ICG-001 and C646 affect many of the same genes in HCT116 and PANC1 cells. (A) Genes that were significantly detected in both HCT116 or in PANC1 cells (P value <0.01) were analyzed for expression differences caused by drug treatment. All genes having a differential P value <0.05 and a fold change greater than 1.2 were analyzed using Euclidean distance and hierarchical clustering. (B) Gene ontology analyses are shown for the genes commonly up- or downregulated in HCT116 and PANC1 cells by the drugs. Terms related to the cell cycle are shown in red.

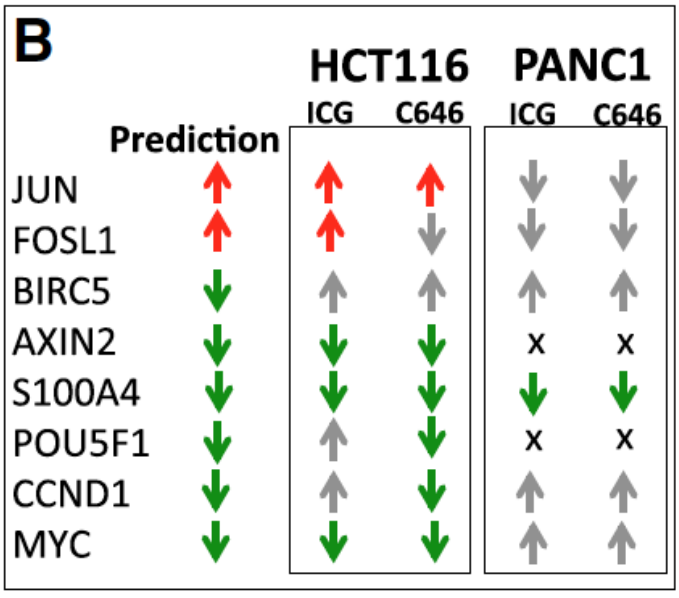
Appendix A-6. Epigenetic inhibitors can partially restore a normal expression pattern to tumor cells



Appendix A-6. Epigenetic inhibitors can partially restore a normal expression pattern to tumor cells. Genes that showed tumor-specific changes in expression in TCGA colon RNA-seq samples (left), TCGA pancreatic RNA-seq samples (right, top), plus differentially expressed genes identified by comparison of normal to tumor pancreatic cell lines (right bottom) were analyzed for responses to drug treatments. In the T/N columns, green indicates that the gene was downregulated in the tumor cells whereas red indicates the gene was upregulated in the tumor cells. The blue brackets indicate genes that were downregulated in the tumor cells and upregulated by the drugs (resulting in an expression level closer to that in normal cells) whereas the black brackets represent the genes that were upregulated in the tumor cells and downregulated by the drugs (resulting in an expression level closer to that in normal cells). The color scale indicates the fold change of gene expression in HCT116 or PANC1 cells after treatment with ICG-001 (ICG) or C646.

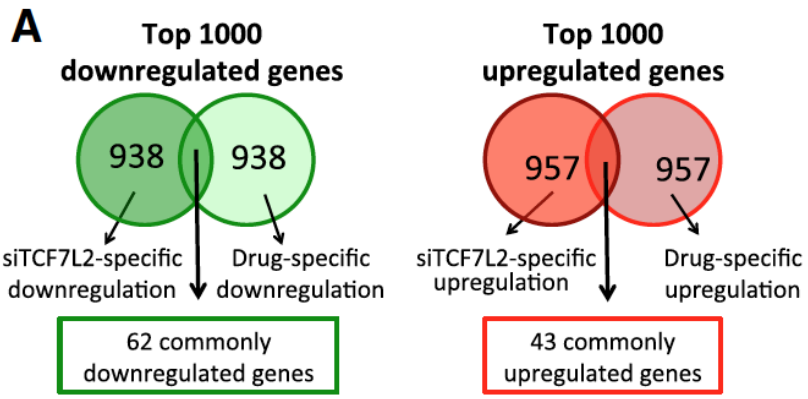
Appendix A-7. Effects of drug treatments on WNT pathway genes





Appendix A-7. Effects of drug treatments on WNT pathway genes. (A) Shown are the expression changes in previously identified WNT pathway genes (http://www.stanford.edu/group/nusselab/cgi-bin/wnt/) that have a detection P value <0.01 and a differential P value <0.05 after 96 h of treatment of HCT116 or PANC1 cells with either ICG-001 (ICG) or C646. (B) Shown are the predicted results (based on the model shown in Figure 1) and the actual responses to the drugs after treatment of HCT116 or PANC1 cells for a set of WNT target genes. In the prediction column, a red arrow indicates that the gene should have been upregulated by ICG-001and the green arrow indicates that the gene should have been downregulated by ICG-001, according to the model. For each cell type, the actual response is shown for both drugs: a red arrow indicates that expression was increased as predicted by the model, a green arrow indicates expression was decreased as predicted by the model, a gray arrow indicates that the expression pattern upon treatment did not correspond to the prediction, and an x indicates that the gene was not expressed in that cell line.

Appendix A-8. In PANC1 cells, treatment with ICG-001 does not affect the same genes as does reduction in levels of TCF7L2

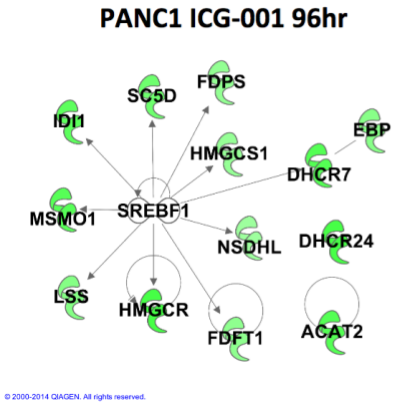


B

Down in siTCF7L2 and ICG-001		Down in siTCF7L2 only		Up in siTCF7L2 only		Up in siTCF7L2 and ICG-001	
Pathway	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**	**Pathway**	**p-val**
Estrogen Receptor Signaling	5.1E-04	Cyclins and Cell Cycle Regulation	5.5E-05	Colorectal Cancer Metastasis Signaling	2.6E-06	1D-myo-inositol Hexakisphosphate Biosynthesis	5.6E-04
Ephrin Receptor Signaling	1.7E-03	Tight Junction Signaling	1.2E-04	Hepatic Fibrosis/ Stellate Cell Activation	1.0E-05	D-myo-inositol (1,3,4)-trisphosphate Biosynthesis	5.6E-04
G Beta Gamma Signaling	2.1E-03	Cell Cycle: G1/S Checkpoint Regulation	2.0E-04	cAMP-mediated signaling	1.9E-05	D-myo-inositol (1,4,5)-trisphosphate Metabolism	9.1E-04
Virus Entry via Endocytic Pathways	2.2E-03	ERK/MAPK Signaling	1.3E-03	Dopamine-DARPP32 in cAMP Signaling	5.1E-05	Leucine Degradation	1.7E-02
Breast Cancer Regulation by Stathmin1	2.3E-03	CHK Proteins in Cell Cycle Checkpoint Control	1.4E-03	Thrombin Signaling	6.0E-05	Isoleucine Degradation	2.6E-02
SAPK/JNK Signaling	2.6E-03	WNT/Ca+ Pathway	1.5E-03	NFAT in Cardiac Hypertrophy	2.1E-04	nNOS Signaling in Skeletal Muscle Cells	2.8E-02
fMLP Signaling in Neutrophils	3.9E-03	Glutamate Biosynthesis	1.6E-03	G-Protein Coupled Receptor Signaling	2.3E-04	Epithelial Adherens Junction Signaling	3.0E-02
Tissue Factor in Cancer	4.1E-03	Glutamate Degradation	1.6E-03	Protein Kinase A Signaling	4.5E-04	Valine Degradation	3.3E-02
Androgen Signaling	4.2E-03	IGF-1 Signaling	1.6E-03	Synaptic Long Term Potentiation	5.0E-04	D-myo-inositol (1,4,5)-trisphosphate Degradation	3.3E-02
Hereditary Breast Cancer	4.6E-03	Circadian Rhythm Signaling	1.7E-03	Gαi Signaling	5.4E-04	IL-22 Signaling	4.4E-02

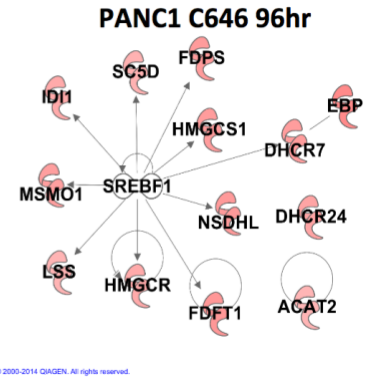
Appendix A-8. In PANC1 cells, treatment with ICG-001 does not affect the same genes as does reduction in levels of TCF7L2. (A) PANC1 cells were treated with siRNAs to TCF7L2 and RNA-seq was performed. The top 1,000 differentially expressed genes after knockdown of TCF7L2 were compared to the top 1,000 genes identified to be responsive to ICG-001 in PANC1 cells. (B) Gene ontology analyses are shown for the genes commonly up- and downregulated by knockdown of TCF7L2 and treatment with ICG-001 and genes that are only affected by knockdown of TCF7L2. Terms related to the cell cycle are shown in red and terms related to the WNT pathway are shown in blue.

Appendix A-9: ICG-001 negatively regulates the cholesterol biosynthesis network



Fold change

Symbol	ICG-001 96hr	C646 96hr
ACAT2	-2.0	1.0
DHCR24	-2.1	1.3
DHCR7	-2.0	1.4
EBP	-1.4	1.6
FDFT1	-1.3	1.5
FDPS	-1.3	1.4
HMGCR	-2.1	1.2
HMGCS1	-1.2	1.4
IDI1	-1.9	1.2
LSS	-1.4	1.1
MSMO1	-1.9	1.3
NSDHL	-1.1	1.4



Appendix A-9: ICG-001 negatively regulates the cholesterol biosynthesis network. IPA was used to show the rela0onships between SREBF1 and other genes involved in cholesterol biosynthesis that are affected by treatment of PANC1 cells with ICG-001 or C646. The arrows indicate direction interactions between the SREBF1 transcription factor and the other genes. Each of the indicated genes was down--regulated (indicated by the green color) by ICG-001 but up-regulated or unaffected by C646 (indicated by the red color); fold change for each gene is shown in the table.

References

1. Luu HH, Zhang R, Haydon RC, Rayburn E, Kang Q, Si W, et al. Wnt/beta-catenin signaling pathway as a novel cancer drug target. Curr Cancer Drug Targets. 2004;4:653–71.
2. Polakis P. Wnt signaling in cancer. Cold Spring Harb Perspect Biol. 2012;4:a008052.
3. Takahashi-Yanaga F, Kahn M. Targeting Wnt signaling: can we safely eradicate cancer stem cells? Clin Cancer Res. 2010;16:3153–62.
4. Tian W, Han X, Yan M, Xu Y, Duggineni S, Lin N, et al. Structure-based discovery of a novel inhibitor targeting the beta-Catenin/Tcf4 interaction. Biochemistry. 2012;51:724–31.
5. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell. 2013;49:825–37.
6. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol. 2010;28:1057–68.
7. Clevers H. Wnt/beta-catenin signaling in development and disease. Cell. 2006;127:469–80.
8. Shitashige M, Hirohashi S, Yamada T. Wnt signaling inside the nucleus. Cancer Sci. 2008;99:631–7.
9. Eguchi M, Nguyen C, Lee SC, Kahn M. ICG-001, a novel small molecule regulator of TCF/beta-catenin transcription. Med Chem. 2005;1:467–72.

10. Emami KH, Nguyen C, Ma H, Kim DH, Jeong KW, Eguchi M, et al. A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. Proc Natl Acad Sci U S A. 2004;101:12682–7.

11. Bowers EM, Yan G, Mukherjee C, Orry A, Wang L, Holbert MA, et al. Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. Chem Biol. 2010;17:471–82.

12. Miki T, Yasuda SY, Kahn M. Wnt/beta-catenin signaling in embryonic stem cell self-renewal and somatic cell reprogramming. Stem Cell Rev. 2011;7:836–46.

13. Rebel VI, Kung AL, Tanner EA, Yang H, Bronson RT, Livingston DM. Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. Proc Natl Acad Sci U S A. 2002;99:14789–94.

14. Ugai H, Uchida K, Kawasaki H, Yokoyama KK. The coactivators p300 and CBP have different functions during the differentiation of F9 cells. J Mol Med (Berl). 1999;77:481–94.

15. Teo JL, Kahn M. The Wnt signaling pathway in cellular proliferation and differentiation: a tale of two coactivators. Adv Drug Deliv Rev. 2010;62:1149–55.

16. Kalkhoven E. CBP and p300: HATs for different occasions. Biochem Pharmacol. 2004;68:1145–55.

17. Ramos YF, Hestand MS, Verlaan M, Krabbendam E, Ariyurek Y, van Galen M, et al. Genome-wide assessment of differential roles for p300 and CBP in transcription regulation. Nucleic Acids Res. 2010;38:5396–408.

18. Foley P, Bunyan D, Stratton J, Dillon M, Lynch SA. Further case of Rubinstein-Taybi syndrome due to a deletion in EP300. Am J Med Genet A. 2009;149A:997–1000.

19. Ma H, Nguyen C, Lee KS, Kahn M. Differential roles for the coactivators CBP and p300 on TCF/beta-catenin-mediated survivin gene expression. Oncogene. 2005;24:3619–31.

20. Morris JP, Wang SC, Hebrok M. KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. Nat Rev Cancer. 2010;10:683–95.

21. Zavoral M, Minarikova P, Zavada F, Salek C, Minarik M. Molecular biology of pancreatic cancer. World J Gastroenterol. 2011;17:2897–908.

22. Zhang Y, Morris JP, Yan W, Schofield HK, Gurney A, Simeone DM, et al. Canonical wnt signaling is required for pancreatic carcinogenesis. Cancer Res. 2013;73:4909–22.

23. Nakamoto M, Matsuyama A, Shiba E, Shibuya R, Kasai T, Yamaguchi K, et al. Prognostic significance of WNT signaling in pancreatic ductal adenocarcinoma. Virchows Arch. 2014;465:401–8.

24. Wall I, Schmidt-Wolf IG. Effect of Wnt inhibitors in pancreatic cancer. Anticancer Res. 2014;34:5375–80.

25. Xu W, Wang Z, Zhang W, Qian K, Li H, Kong D, et al. Mutated K-ras activates CDK8 to stimulate the epithelial-to-mesenchymal transition in pancreatic cancer in part via the Wnt/beta-catenin signaling pathway. Cancer Lett. 2015;356:613–27.

26. Jia J, Parikh H, Xiao W, Hoskins JW, Pflicke H, Liu X, et al. An integrated transcriptome and epigenome analysis identifies a novel candidate gene for pancreatic cancer. BMC Med Genomics. 2013;6:33.

27. Dekker FJ, Haisma HJ. Histone acetyl transferases as emerging drug targets. Drug Discov Today. 2009;14:942–8.

28. Geutjes EJ, Bajpe PK, Bernards R. Targeting the epigenome for treatment of cancer. Oncogene. 2012;31:3827–44.

29. Helin K, Dhanak D. Chromatin proteins and modifications as drug targets. Nature. 2013;502:480–8.

30. Wang L, Tang Y, Cole PA, Marmorstein R. Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. Curr Opin Struct Biol. 2008;18:741–7.

31. Gorin A, Gabitova L, Astsaturov I. Regulation of cholesterol biosynthesis and cancer signaling. Curr Opin Pharmacol. 2012;12:710–6.

32. Gabitova L, Gorin A, Astsaturov I. Molecular pathways: sterols and receptor signaling in cancer. Clin Cancer Res. 2014;20:28–34.

33. Brown MS, Goldstein JL. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell. 1997;89:331–40.

34. Shimano H. Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. Prog Lipid Res. 2001;40:439–52.

35. Xiao X, Song BL. SREBP: a novel therapeutic target. Acta Biochim Biophys Sin (Shanghai). 2013;45:2–10.

36. Oliner JD, Andresen JM, Hansen SK, Zhou S, Tjian R. SREBP transcriptional activity is mediated through an interaction with the CREB-binding protein. Genes Dev. 1996;10:2903–11.

37. Arensman MD, Telesca D, Lay AR, Kershaw KM, Wu N, Donahue TR, et al. The CREB-binding protein inhibitor ICG-001 suppresses pancreatic cancer growth. Mol Cancer Ther. 2014;13:2303–14.

38. Ciliberto D, Botta C, Correale P, Rossi M, Caraglia M, Tassone P, et al. Role of gemcitabine-based combination therapy in the management of advanced pancreatic cancer: a meta-analysis of randomised trials. Eur J Cancer. 2013;49:593–603.

39. Giroux V, Malicet C, Barthet M, Gironella M, Archange C, Dagorn JC, et al. p8 is a new target of gemcitabine in pancreatic cancer cells. Clin Cancer Res. 2006;12:235–41.

40. Hamidi T, Algul H, Cano CE, Sandi MJ, Molejon MI, Riemann M, et al. Nuclear protein 1 promotes pancreatic cancer development and protects cells from stress by inhibiting apoptosis. J Clin Invest. 2012;122:2092–103.

41. Sandi MJ, Hamidi T, Malicet C, Cano C, Loncle C, Pierres A, et al. p8 expression controls pancreatic cancer cell migration, invasion, adhesion, and tumorigenesis. J Cell Physiol. 2011;226:3442–51.

42. Clark DW, Mitra A, Fillmore RA, Jiang WG, Samant RS, Fodstad O, et al. NUPR1 interacts with p53, transcriptionally regulates p21 and rescues breast epithelial cells from doxorubicin-induced genotoxic stress. Curr Cancer Drug Targets. 2008;8:421–30.

43. Koster R, di Pietro A, Timmer-Bosscha H, Gibcus JH, van den Berg A, Suurmeijer AJ, et al. Cytoplasmic p21 expression levels determine cisplatin resistance in human testicular cancer. J Clin Invest. 2010;120:3594–605.

44. Vincent AJ, Ren S, Harris LG, Devine DJ, Samant RS, Fodstad O, et al. Cytoplasmic translocation of p21 mediates NUPR1-induced chemoresistance: NUPR1 and p21 in chemoresistance. FEBS Lett. 2012;586:3429–34.

45. Weiss RH. p21Waf1/Cip1 as a therapeutic target in breast and other cancers. Cancer Cell. 2003;4:425–9.

46. Chen D, Niu M, Jiao X, Zhang K, Liang J, Zhang D. Inhibition of AKT2 enhances sensitivity to gemcitabine via regulating PUMA and NF-kappaB signaling pathway in human pancreatic ductal adenocarcinoma. Int J Mol Sci. 2012;13:1186–208.

47. Mu GG, Zhang LL, Li HY, Liao Y, Yu HG. Thymoquinone pretreatment overcomes the insensitivity and potentiates the antitumor effect of gemcitabine through abrogation of notch1, PI3K/Akt/mTOR regulated signaling pathways in pancreatic cancer. Dig Dis Sci. 2014. doi:10.1007/s10620-014-3394-x.

48. Anders S, Pyl PT, Huber W. HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9.

49. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.

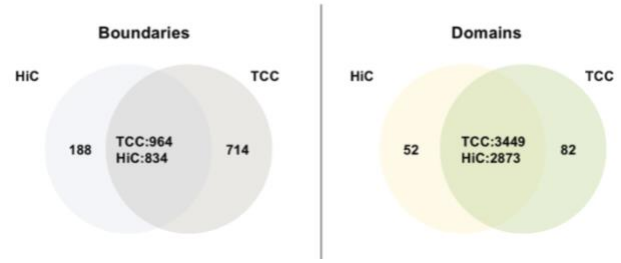
Appendix B

Supplemental Materials for Chapter 2

Appendix B-1: Comparison of TCC and HiC TADs and boundaries

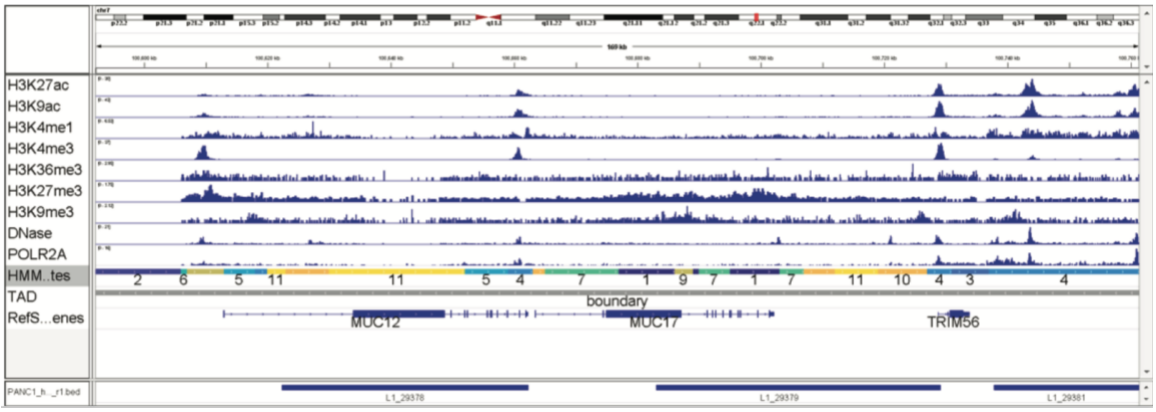
TADs identified TopDom

	HiC_merged	TCC_merged
# of Domains hg19 + 20KB	2,925	3,531
# of Boundaries hg19 + 20KB	1,022	1,678



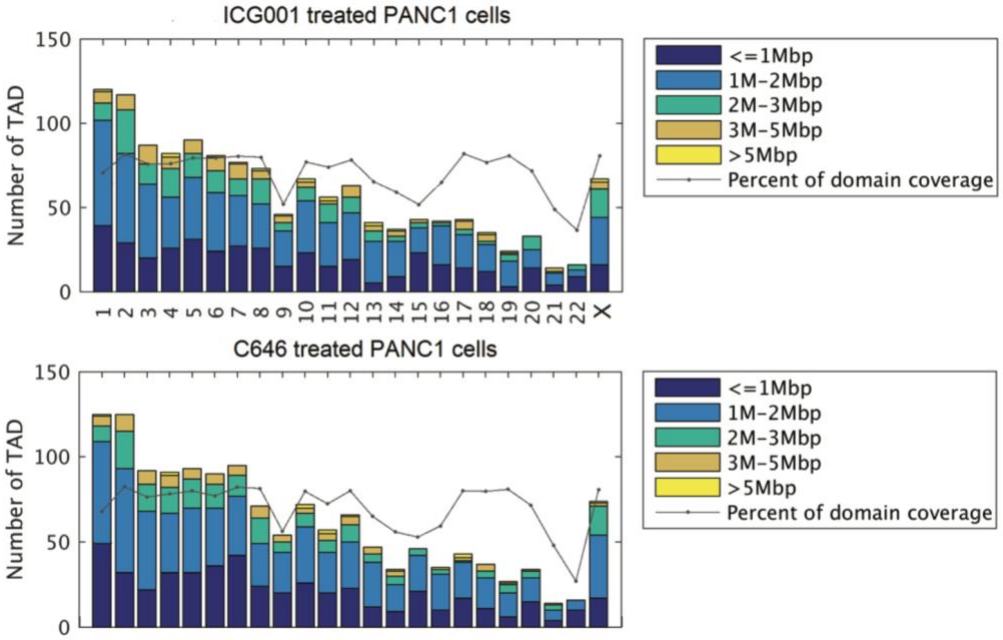
Appendix B-1: Comparison of TCC and HiC TADs and boundaries. A comparison of the number of topological associated domains and TAD boundaries between PANC1 TCC and HiC datasets.

Appendix B-2: Visualization of Repeated HMM state S1/S7/S9



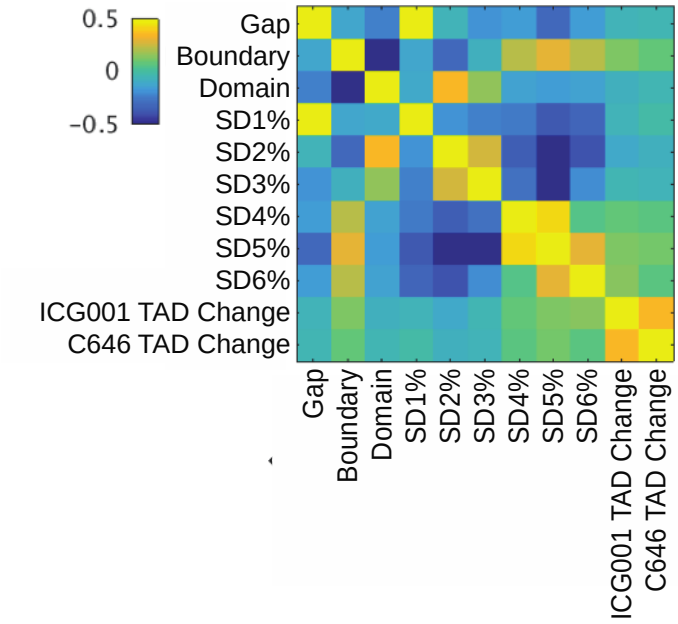
Appendix B-2: Visualization of Repeated HMM state S1/S7/S9. IGV snapshot of S1/S7/S9 state.

Appendix B-3: Length distribution of topological domains in drug treated PANC1



Appendix B-3: Length distribution of topological domains in drug-treated PANC1. Distribution of TAD length by chromosome for ICG001 (A) and C646 (B) treated PANC1 cells.

Appendix B-4: Pearson correlation between sub-domain and change of sub-domain

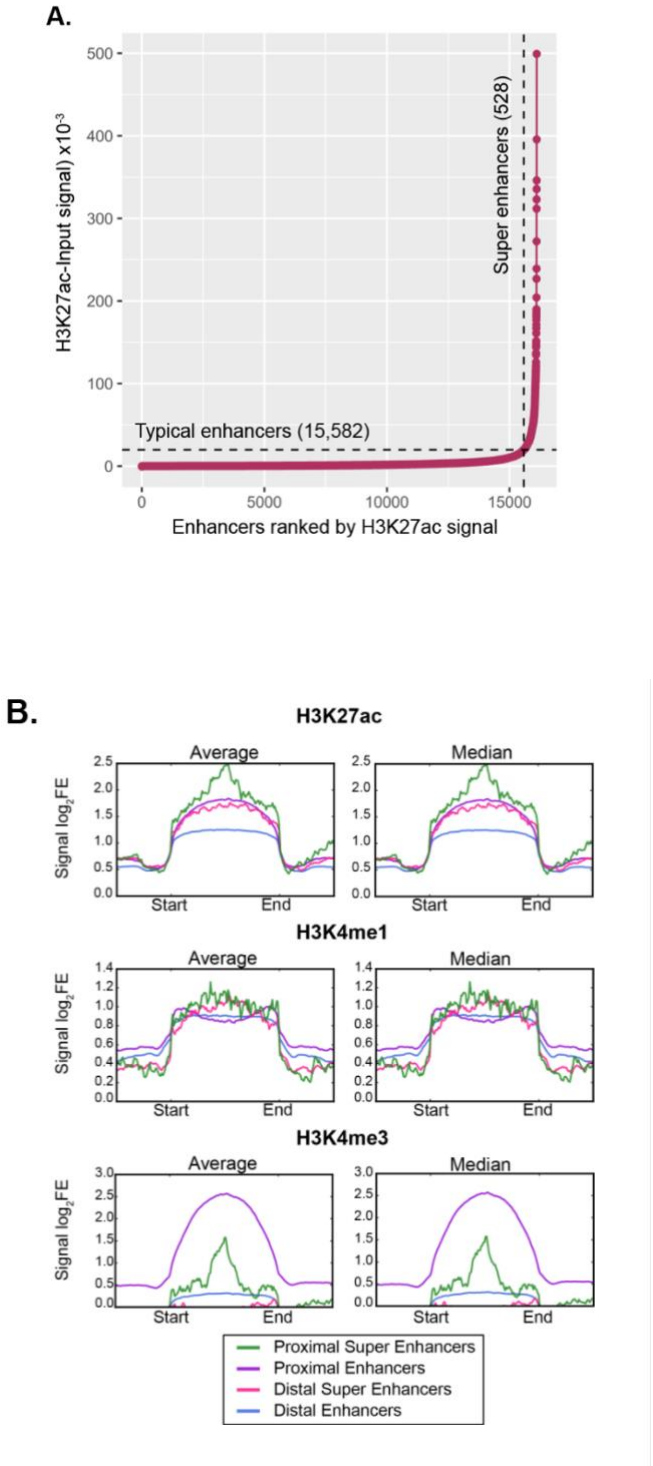


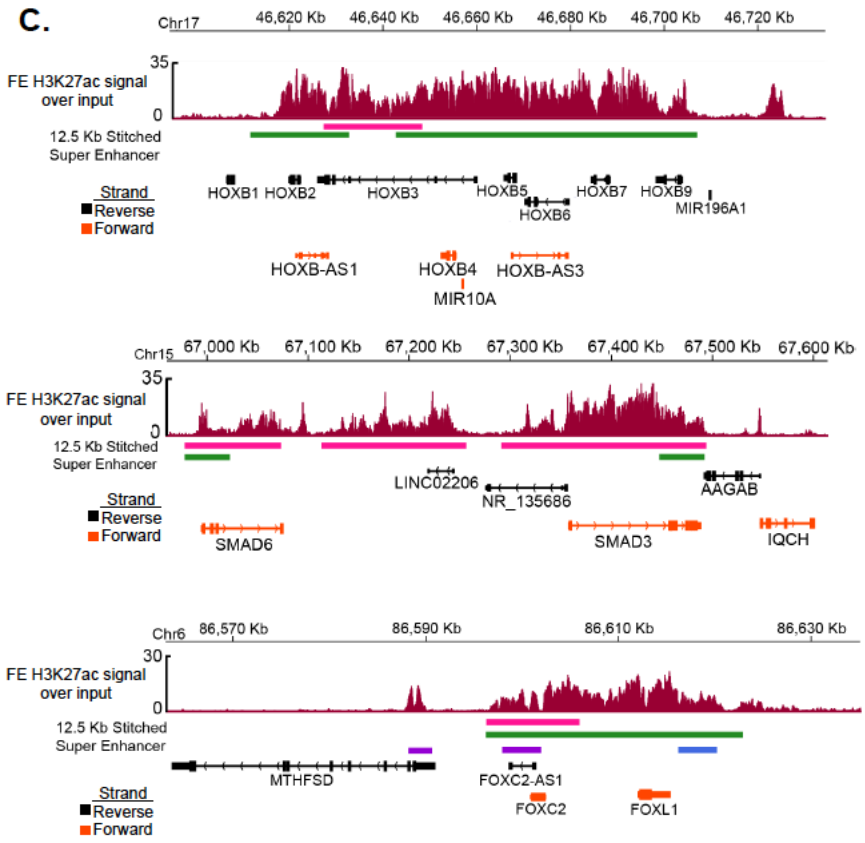
Appendix B-4: Pearson correlation between sub-domain and change of sub-domain. Heatmap displaying the Pearson correlation scores for different domains and domain changes.

Appendix C

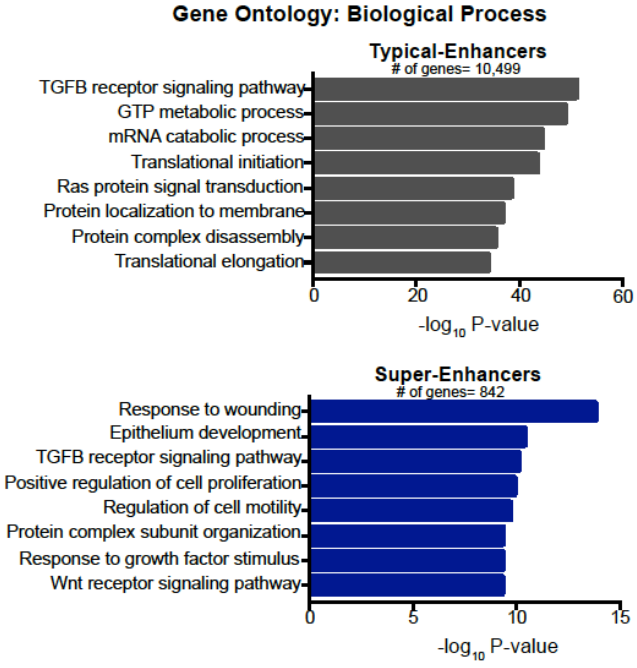
Supplemental Materials for Chapter 3

Appendix C-1: Identification of super-enhancers in PANC1 cells



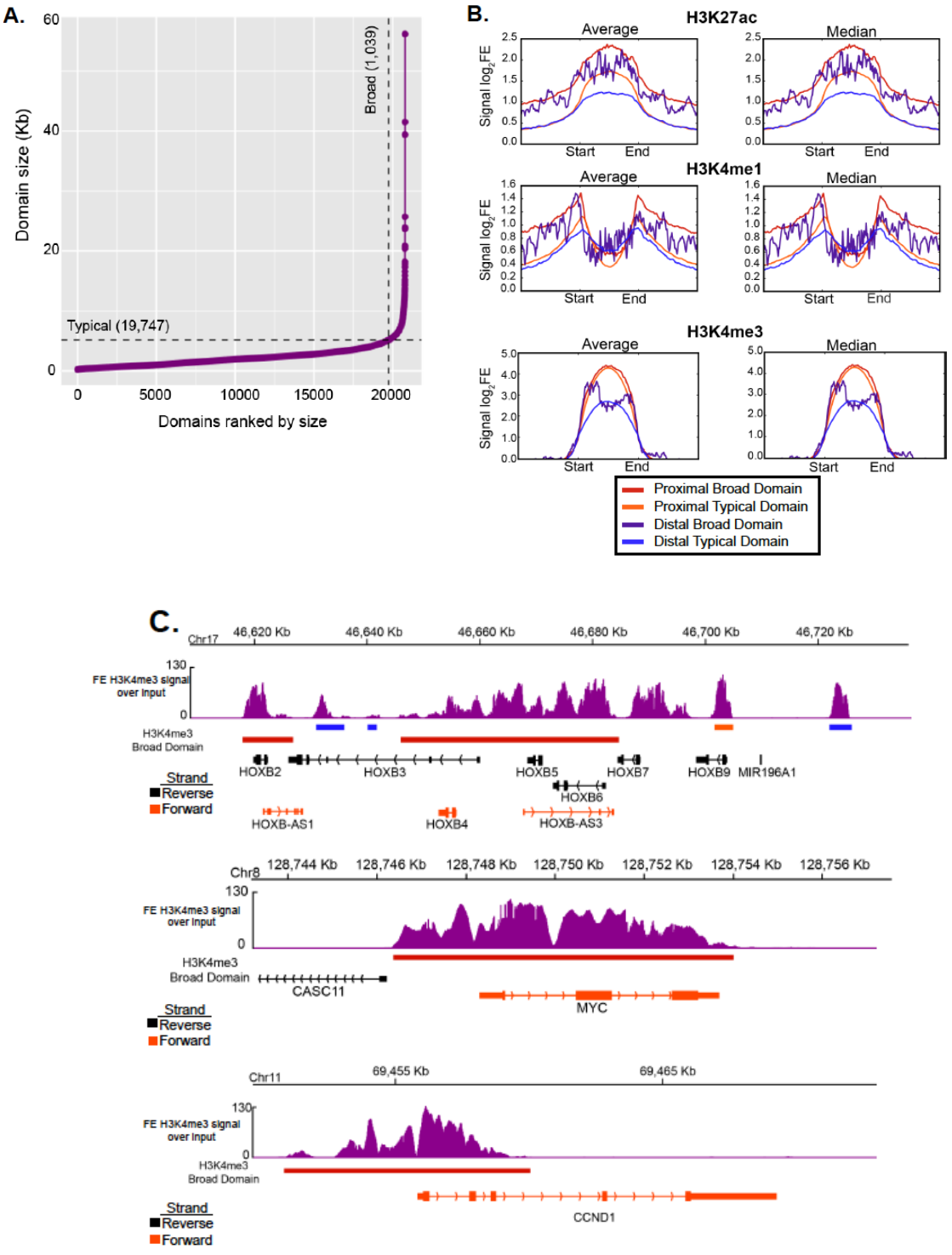

C.
Chr17
46,620 Kb
46,640 Kb
46,660 Kb
46,680 Kb
46,700 Kb
46,720 Kb
FE H3K27ac signal over input
35
0
12.5 Kb Stitched Super Enhancer
Strand
Reverse
Forward
HOXB1
HOXB2
HOXB3
HOXB5
HOXB6
HOXB7
HOXB9
MIR196A1
HOXB-AS1
HOXB4
MIR10A
HOXB-AS3
Chr15
67,000 Kb
67,100 Kb
67,200 Kb
67,300 Kb
67,400 Kb
67,500 Kb
67,600 Kb
LINC02206
NR_135686
AAGAB
SMAD6
SMAD3
IQCH
Chr6
86,570 Kb
86,590 Kb
86,610 Kb
86,630 Kb
30
MTHFSD
FOXC2-AS1
FOXC2
FOXL1

D.


Gene Ontology: Biological Process
Typical-Enhancers
# of genes= 10,499
TGFB receptor signaling pathway
GTP metabolic process
mRNA catabolic process
Translational initiation
Ras protein signal transduction
Protein localization to membrane
Protein complex disassembly
Translational elongation
0
20
40
60
-log10 P-value
Super-Enhancers
# of genes= 842
Response to wounding
Epithelium development
TGFB receptor signaling pathway
Positive regulation of cell proliferation
Regulation of cell motility
Protein complex subunit organization
Response to growth factor stimulus
Wnt receptor signaling pathway
0
5
10
15
-log10 P-value

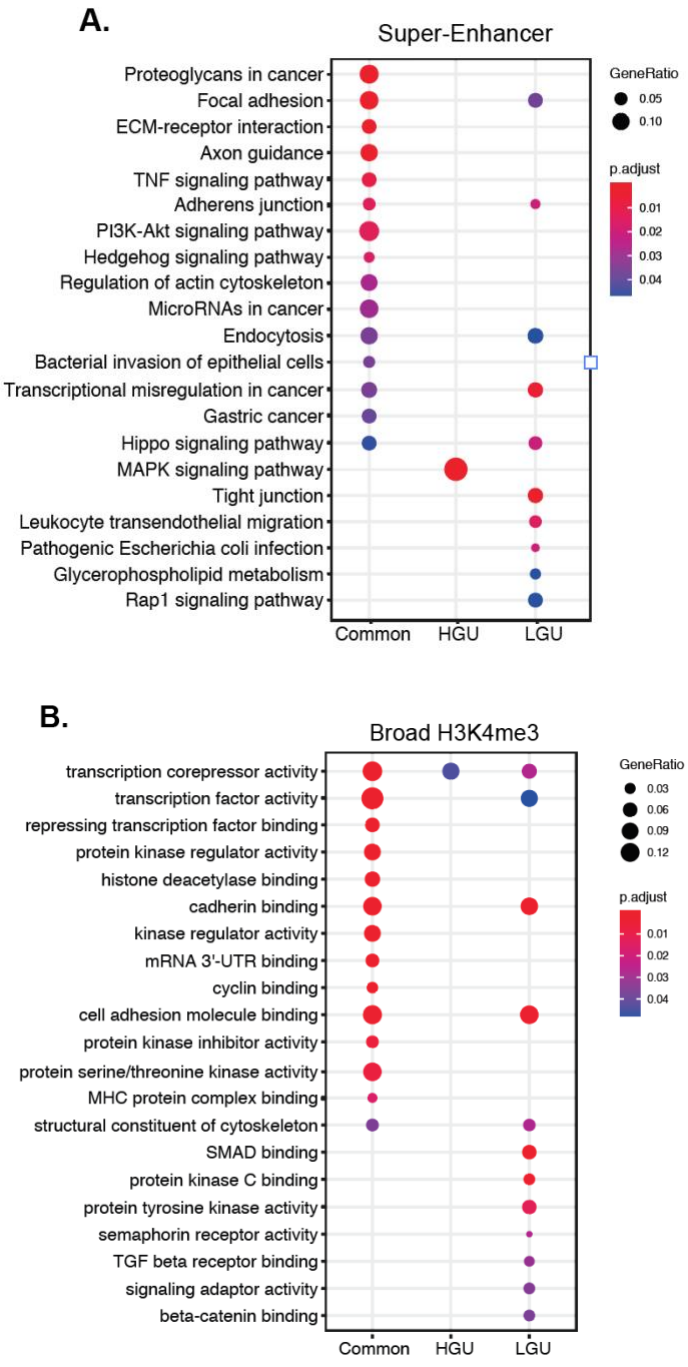
Appendix C-1. Identification of super-enhancers in PANC1 cells. A. Super-enhancers were identified using Ranked Ordering of Super-Enhancers (ROSE) [5, 60], where H3K27ac signal relative to input is ranked and visualized. The dashed line distinguishes between typical-enhancers and super-enhancers. B. We visualized the signal of core histone modifications including: H3K27ac, H3K4me1 and H3K4me3 within typical- or super-enhancers relative to gene regions (proximal <5 kb and distal >5 kb from the TSS). The signal is represented as log2 fold enrichment over input signal. C. Genome snapshots of representative gene regions marked by super-enhancers in PANC1. D. Gene ontology analysis for 'Biological Process' was performed for typical- and super-enhancers.

Appendix C-2: Identifying Broad H3K4me3 domains in PANC1 cells



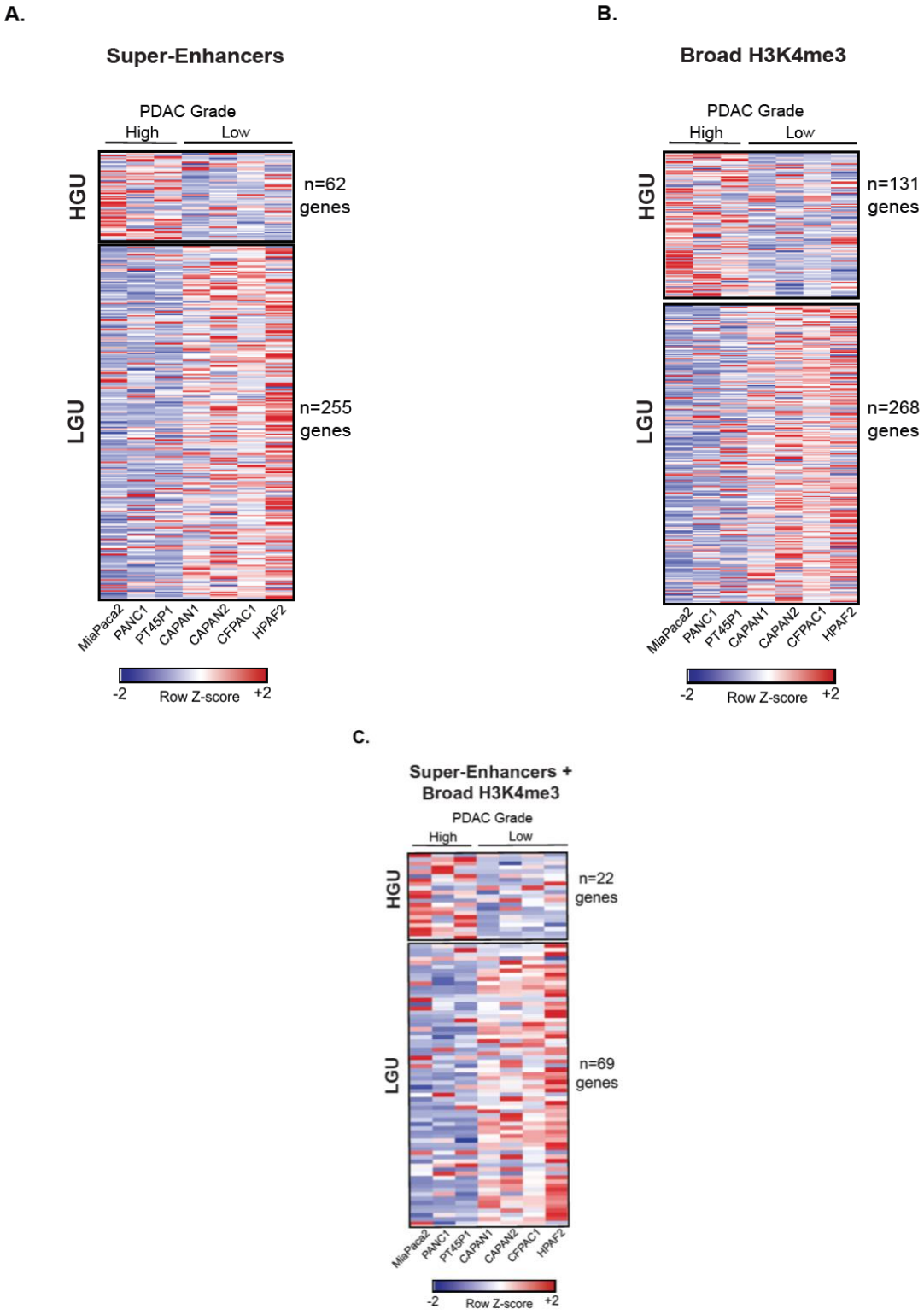
Appendix C-2. Identifying Broad H3K4me3 domains in PANC1 cells. A. Broad H3K4me3 domains were determined by MACS2 with the –broad flag activated. We then ranked the size of the domains and visualized them. The dashed line represents the cutoff of the top 5% of H3K4me3, which are defined as broad H3K4me3 domains. B. We then investigated the signal of core histone modifications including: H3K27ac, H3K4me1 and H3K4me3 within typical- and broad- H3K4me3 regions proximal (<5 kb) or distal (>5 kb) to TSS. C. Representative genome snapshots of regions containing broad H3K4me3 marks.

Appendix C-3: Pathway analysis of super-enhancers and broad H3K4me3 domains in different PDAC grade groups



Appendix C-3. Pathway analysis of super-enhancers and broad H3K4me3 domains in different PDAC grade groups. Gene ontology for PDAC grade groups within broad regions.

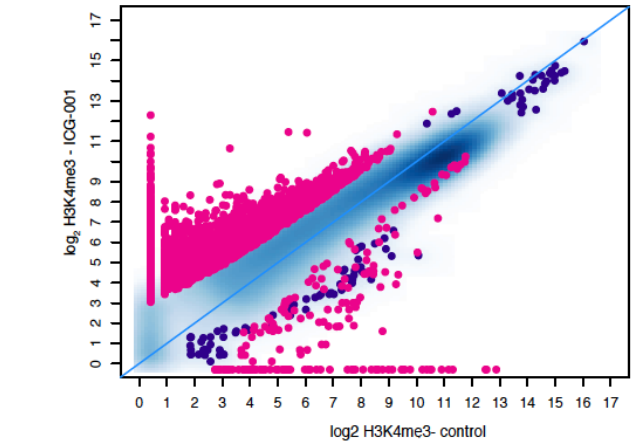
Appendix C-4: Gene expression relative to broad domains

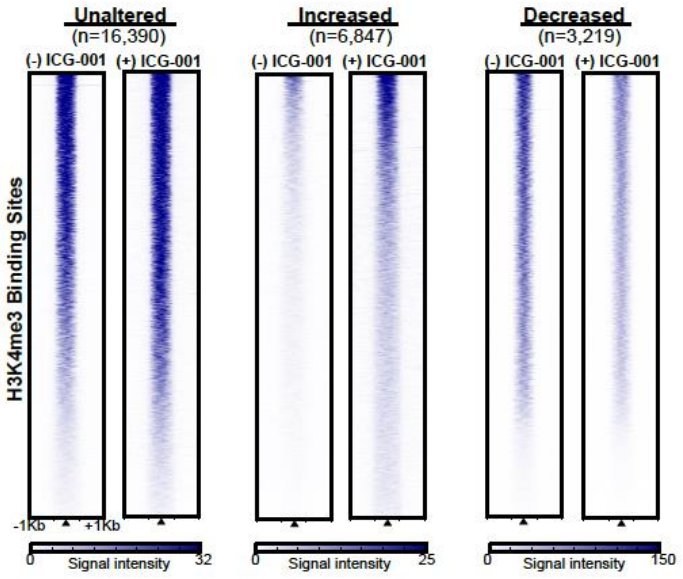


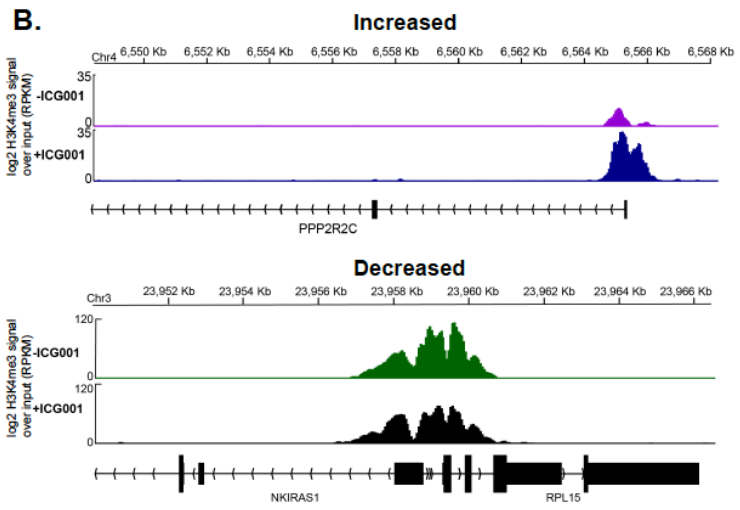
Appendix C-4. Gene expression relative to broad domains. Heatmaps showing the relative expression levels of genes marked by A. super-enhancers, B. broad H3K4me3, or C. both broad domains across 7 human PDAC cell lines. The number of genes belonging to each domain type is indicated. HGU corresponds to High-Grade Unique and LGU corresponds to Low-Grade Unique.

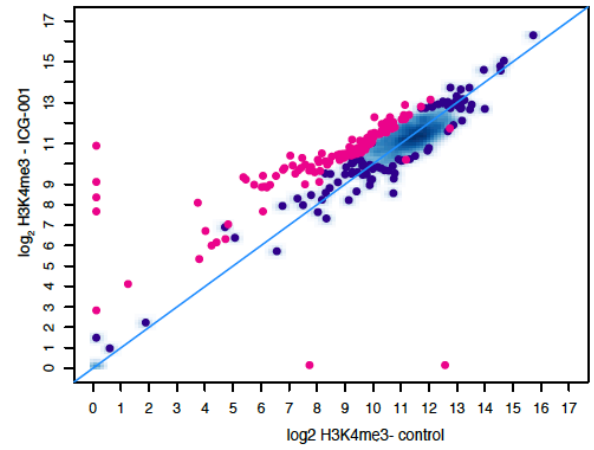
Appendix C-5: The impact of ICG-001 treatment on broad H3K4me3 domains

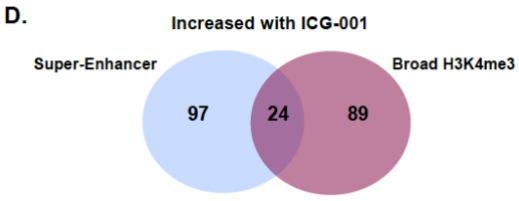
A.





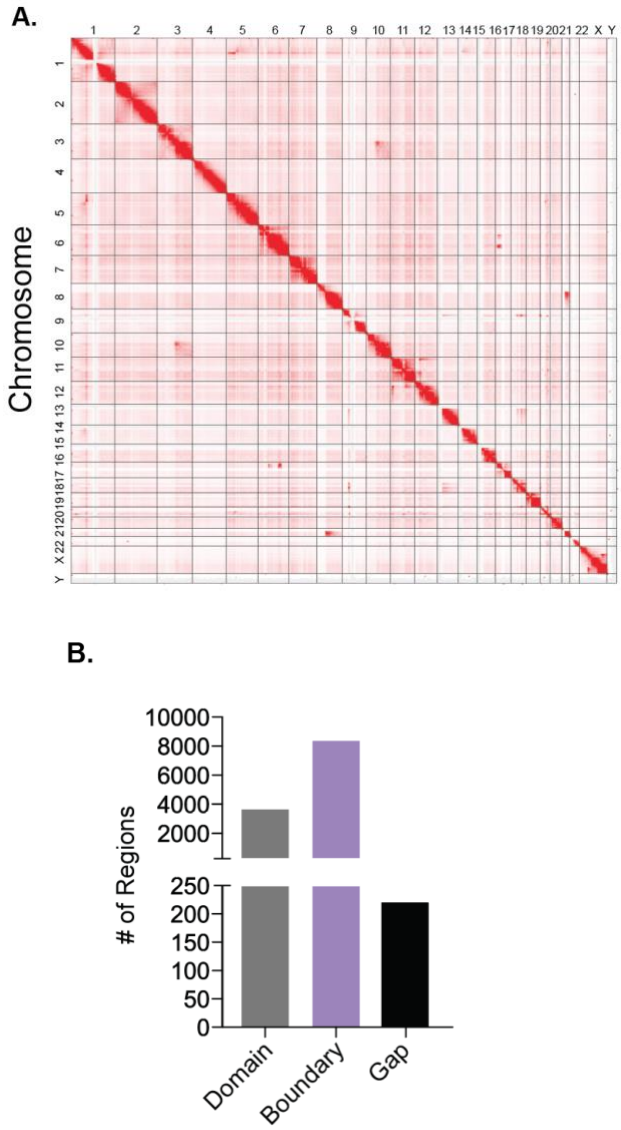

B.
Increased
Chr4
6,550 Kb
6,552 Kb
6,554 Kb
6,556 Kb
6,558 Kb
6,560 Kb
6,562 Kb
6,564 Kb
6,566 Kb
6,568 Kb
log2 H3K4me3 signal over input (RPKM)
-ICG001
+ICG001
35
0
PPP2R2C
Decreased
Chr3
23,952 Kb
23,954 Kb
23,956 Kb
23,958 Kb
23,960 Kb
23,962 Kb
23,964 Kb
23,966 Kb
120
NKIRAS1
RPL15


C.
log2 H3K4me3 - ICG-001
log2 H3K4me3- control


D.
Increased with ICG-001
Super-Enhancer
Broad H3K4me3
97
24
89

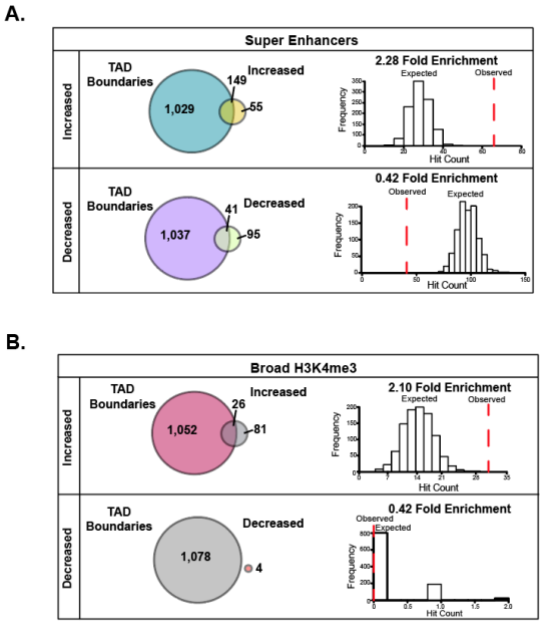
Appendix C-5. The impact of ICG-001 treatment on broad H3K4me3 domains. A. Differential analysis of H3K4me3 genome-wide enrichment in PANC1 cells treated with ICG-001. Significantly altered regions are indicated by the colored dots (FDR <0.1). B. Genome browser snapshots of altered H3K4me3 regions. C. Differential binding analysis of broad H3K4me3 regions. D. Venn diagrams showing overlap of super-enhancers and broad H3K4me3 regions with increased histone modification levels.

Appendix C-6: Chromatin-interacting domains in PANC1 cells



Appendix C-6: Chromatin-interacting domains in PANC1 cells. A. Genome-wide interaction matrix of chromosome contacts identified from TCC. B. Distribution of identified domain types using TopDom (40 kb resolution).

Appendix C-7: Domains with increased ChIP-seq signal after HAT inhibitor treatment are enriched in boundary regions

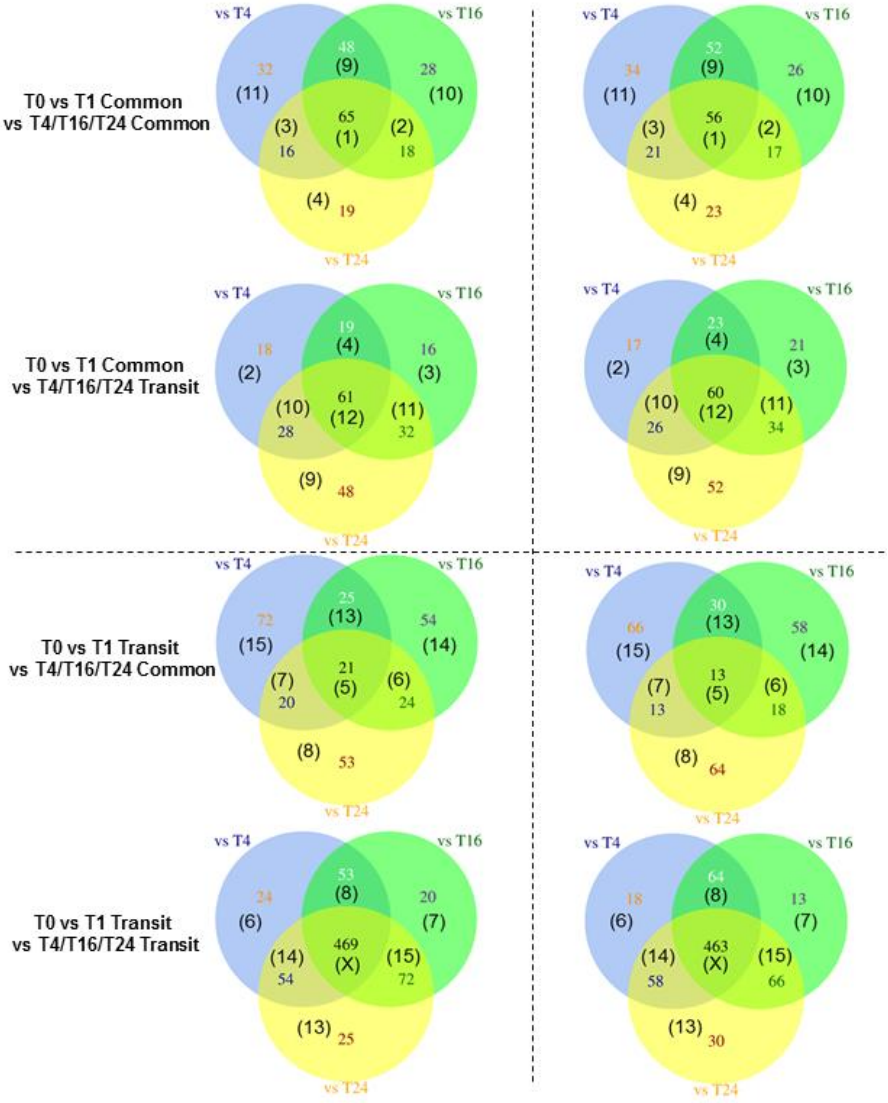


Appendix C-7: Domains with increased ChIP-seq signal after HAT inhibitor treatment are enriched in boundary regions. Feature enrichment analysis of super-enhancers (A) and broad H3K4me3 domains (B) increasing or decreasing in signal.

Appendix D

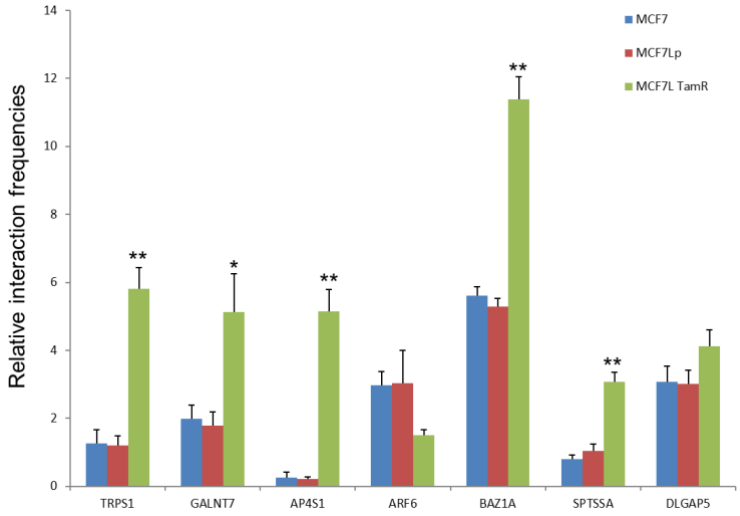
Supplemental Figures for Chapter 4

Appendix D-1: Identification of 24 patterns of dynamic compartments



Appendix D-1. Identification of 24 patterns of dynamic compartments. The continuous genomic regions of positive first eigenvector were defined as compartment A (open chromatin) and the coutinuous genomic regions of negative first eigenvector was defined as compartment B (close chromatin). The time series compartments (A or B individually) were compared as follows: First, two kinds of compartments: T0 vs T1 Common and T0 vs T1 Transit were identified by comparing compartments T0/T1. The "Common" compartments are the overlapping compartments and the "Transit" compartments are differential compartments, which were used in the following steps as well. Next, the common compartments identified from T0 vs T1 and Transit compartments identified from T0 vs T1 were compared with T4, T16, T24 independently to generate the (a) T0 vs T1 Common vs T4/T16/T24 Common, (b) T0 vs T1 Common vs T4/T16/T24 Transit, (c) T0 vs T1 Transit vs T4/T16/T24 Common, (d) T0 vs T1 Transit vs T4/T16/T24 Transit.Lastly, the patterns denoted as 1-15 were produced by comparing the various time points (T4, T16 and T24) of subsets (a, b, c, and d) identified from the previous step, which we refer to as "vs T4", "vs T16", "vs T24" as shown in the Venn diagram. The rest of the subsets (X shown in Venn diagram) were divided into patterns 16-24 according to the numbers of converted bins (Suppl. Figure S15). Left column: compartment A, right column: compartment B. In the Venn diagram, numbers without parentheses are the numbers of compartments, numbers with parenthases are the patterns.

Appendix D-2. Validation of 3C-qPCR for HOMER loops



MCF7: MCF7 cell lines

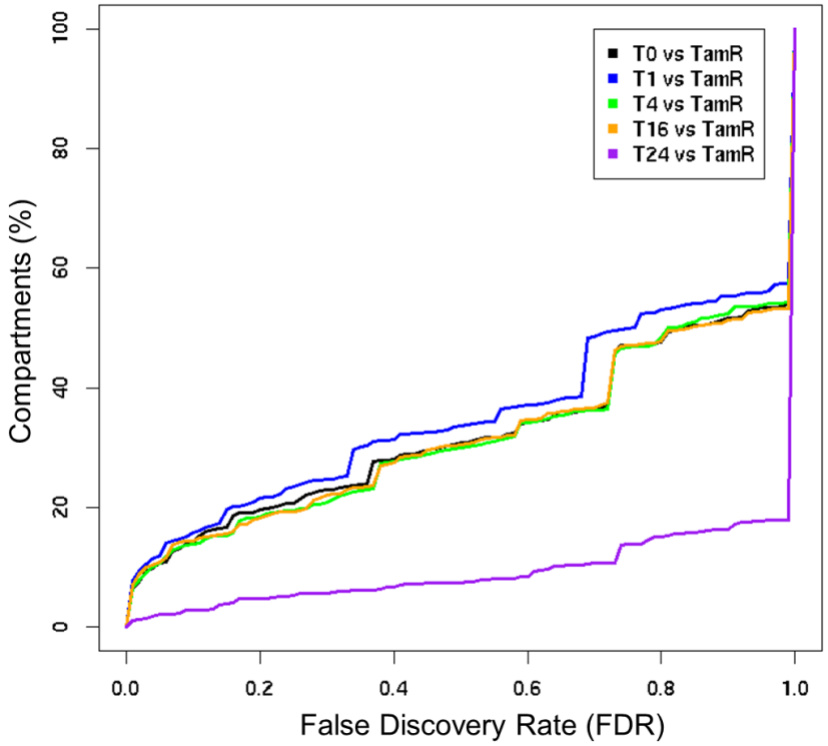
MCF7Lp: MCF7L parental cells

MCF7L TamR: MCF7L Tamoxifen resistant cells

*: $p < 0.05$; **: $p < 0.01$, t-test

Appendix D-2. Validation of 3C-qPCR for HOMER loops. HOMER interaction loops are named as genes located at the interaction loops. Three 3C-qPCRs were performed for each loop. The MCF7L TamR group was compared with MCF7 or MCF7Lp by ANOVA analysis.

Appendix D-3. False Discovery Rate (FDR) and percentage of compartments



Appendix D-3. For each 100 kb bin, the *p* values of first eigenvectors of compartments of T0 vs TamR, T1 vs TamR, T4 vs TamR, T16 vs TamR, T24 vs TamR were computed by Wilcoxon rank-sum test. Then False Discovery Rate (FDR) was determined by the adjustment methods of the Benjamini–Hochberg procedure.

Appendix E

Transcriptomic Signatures of Tacaribe Virus-Infected Jamaican Fruit Bats

Abstract

Tacaribe virus (TCRV) is a mammalian arenavirus that was first isolated from artibeus bats in the 1950s. Subsequent experimental infection of Jamaican fruit bats (Artibeus jamaicensis) caused a disease similar to that of naturally infected bats. Although substantial attention has focused on bats as reservoir hosts of viruses that cause human disease, little is known about the interactions between bats and their pathogens. We performed a transcriptome-wide study to illuminate the response of Jamaican fruit bats experimentally infected with TCRV. Differential gene expression analysis of multiple tissues revealed global and organ-specific responses associated with innate antiviral responses, including interferon alpha/beta and Toll-like receptor signaling, activation of complement cascades, and cytokine signaling, among others. Genes encoding proteins involved in adaptive immune responses, such as gamma interferon signaling and co-stimulation of T cells by the CD28 family, were also altered in response to TCRV infection. Immunoglobulin gene expression was also elevated in the spleens of infected bats, including IgG, IgA, and IgE isotypes. These results indicate an active innate and adaptive immune response to TCRV infection occurred but did not prevent fatal disease. This de novo assembly provides a high-throughput data set of the Jamaican fruit bat and its host response to TCRV infection, which remains a valuable tool to understand the molecular signatures involved in antiviral responses in bats.

Importance

As reservoir hosts of viruses associated with human disease, little is known about the interactions between bats and viruses. Using Jamaican fruit bats infected with

Tacaribe virus (TCRV) as a model, we characterized the gene expression responses to infection in different tissues and identified pathways involved with the response to infection. This report is the most detailed gene discovery work in the species to date and the first to describe immune gene expression responses in bats during a pathogenic viral infection.

Introduction

Bats are a phylogenetically and geographically diverse group of mammals, with about 1,150 species (1, 2). Certain bat species have been identified as reservoir hosts of zoonotic viruses associated with significant human morbidity and mortality, including rabies virus and other lyssaviruses, Marburg virus, Nipah virus, and Hendra virus (3). They also are suspected reservoirs of other viruses, such as the ebolaviruses, and Middle East respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS) coronaviruses (CoVs) (4–6). Each of these viruses can cause severe disease in humans but are not known to cause disease in their reservoir hosts (3, 7). Although nearly 200 viruses have been associated with bats, there are likely many more (8). As non-model organisms, virtually nothing is known about bat immune responses. Although bats appear to have small genomes relative to other mammals (9), genomic analyses suggest that bats share most features of other mammals (8, 10–12).

Despite serving as reservoir hosts of several zoonotic viruses, some bats are also susceptible to infectious diseases. White nose syndrome, which has caused the deaths of millions of bats in North America, is a fungal disease threatening some species with extinction (13–16). Bats can shed rabies virus and other lyssaviruses for prolonged

periods, but the infection is always fatal (3, 17–21). Because bats are important members of their ecosystems, a better understanding of the immune responses and subsequent pathogenesis to infectious agents is essential. To this end, we developed a laboratory model for the study of infection of Jamaican fruit bats (Artibeus jamaicensis) by a natural bat pathogen, Tacaribe virus (TCRV) (11, 21, 22).

TCRV is a mammarenavirus first isolated from two species of diseased artibeus bats in the late 1950s near Port-of-Spain, Trinidad, and is most closely related to Junïn and Machupo viruses, which cause Argentine and Bolivian hemorrhagic fevers, respectively (23–25). Each arenavirus is associated with a specific host species, and the distribution of the host therefore dictates the distribution of the virus. All known reservoir hosts of mammarenaviruses are rodents; however, the reservoir host of TCRV remains unclear. It was suspected that artibeus bats were reservoirs of TCRV given its original isolation from multiple artibeus bats and the inability to detect it in other mammals (25–27). Interestingly, TCRV was isolated from lone star ticks collected in Florida in 2012 (28). The tick-derived isolate was nearly identical to the TCRV isolate from Trinidad (TRVL-11573), with 99.6% nucleotide identity across its genome (28). Recent studies by our group found that TCRV causes fatal disease or is cleared without pathology in Jamaican fruit bats, features that are inconsistent for a reservoir host (22). In many of these bats, substantial neutrophil and lymphocytic infiltration into tissues occurred, which suggests a role for these cells in the host response to TCRV (22).

The present study was designed to characterize the transcriptional responses of bats with TCRV disease. Accordingly, we performed RNA sequencing of spleens and

liver and kidney samples from experimentally infected bats and generated a broad bat transcriptome rich in annotated genes. These target tissues were chosen because they represent the organs with the most significant pathology in our previous report (8). This report is the most comprehensive gene discovery work in the species to date and the first to describe immune gene expression responses in bats during an arenavirus infection.

Results

High-quality de novo assembly and annotation of the Jamaican fruit bat transcriptome

We previously reported a high mortality rate in Jamaican fruit bats experimentally infected with TCRV, in which high-dose inoculations (106 50% tissue culture infective doses [TCID50]) caused significant and fatal disease as early as 10 days post infection (22). Histopathologic findings revealed multiple organ involvement in TCRV disease, including acute neutrophilic splenitis and white pulp hyperplasia, as well as plasmacytic and histiocytic splenitis. To profile the host pathogenic transcriptional response, we generated stranded poly(A) Illumina RNA-Seq (transcriptome sequencing) libraries using RNA extracted from the organs of experimentally infected bats. For this analysis, we harvested the livers, kidneys, and spleens from 2 control bats (Dulbecco's phosphate-buffered saline [DPBS] treated) and 2 TCRV-infected bats with fatal disease (**Appendix E-1**). Our previous analysis indicated TCRV RNA was present in each of these tissues at time of collection (22). A total of 12 pooled samples were sequenced, generating 693,106,150 raw 100-bp paired-end reads. After demultiplexing, trimming of poor-quality reads and adapter sequences, and removing duplicate reads, 691,108,820

nonredundant reads per sample were used for the transcriptome assembly. De novo assembly of the global transcriptome was performed using Trinity, resulting in 349,855 assembled transcripts of greater than or equal to 300 bp (mean length of 997 bp) with an N50 of 3,419 bases that were clustered into 175,144 nonredundant clustered transcripts (unigenes) (**Appendix E-2A**) (29). Inspection of these unigenes identified from the combined transcriptome showed that 35% of the contigs (12,600) are expressed in each of the three different tissues (fragments per kilobase per million [FPKM]), whereas the expression of many tissue-specific contigs was identified in the spleen, liver, and kidney (**Appendix E-2B**).

The combined Jamaican fruit bat transcriptome was systematically annotated using the Trinotate pipeline, a software suite that automates the functional annotation of the assembled contigs (30). The annotation report for the combined assembly from the Trinotate pipeline represents the predicted coding sequences of Jamaican fruit bat genes and the results of homology searches against the databases listed in Data Set S1. Among the 227,656 transcripts containing complete open reading frame (ORF) sequences, 124,204 non-redundant ORFs (54%) were associated with high-confidence coding predictions, BLAST homology and PFAM domain content. We compared this combined Jamaican fruit bat transcriptome assembly to those of other mammals through BLASTX analysis. The bat Brandt's myotis (*Myotis brandtii*) had the highest number of related sequences (8,060 similar sequences). Among other mammals were the big brown bat (*Eptesicus fuscus*) and the black flying fox (*Pteropus alecto*) (with 7,947 and 6,955 similar sequences, respectively) (**Appendix E-2C**).

Differential gene expression following TCRV infection

To investigate the molecular response of bats to TCRV infection, differential gene expression analysis was performed. We used a pairwise comparison of TCRV-infected samples against the corresponding controls and found that the expression levels of hundreds of different genes were altered during TCRV infection (**Appendix E-3A**). The spleen had the largest number of differentially expressed genes (DEGs); among these 1,912 DEGs, 1,187 were upregulated and 725 were downregulated following infection (**Appendix E-3C**; false discovery rate [FDR], =0.01; log2 fold change =2). We also determined that the kidney and liver each had a greater number of upregulated genes (251 and 188, respectively) compared to the number of downregulated genes in these tissues following TCRV infection (123 and 72, respectively). A comparison of all TCRV-infected tissues against all of the uninfected controls revealed 62 upregulated and 16 downregulated genes (**Appendix E-3B and C**).

Immune gene expression profile in response to TCRV infection

To gain specific insight into the immune-related gene expression altered in response to TCRV infection, we utilized the ImmPort database to identify those TCRV-altered genes that relate to immune-system functions (31). Approximately 23% of the 4,723 genes available in the database corresponded to the differentially expressed genes annotated in our analysis. The coordinating transcript expression values of these identified immune genes were used to evaluate the relationship between the specific uninfected and infected tissues (**Appendix E-4A**). While all three tissue types studied had unique expression profiles in the infected samples, we further analyzed the

transcripts contained in cluster 3, which represent sequences with overall shared expression patterns and found that these corresponding genes map to pathways identified to be affected in response to viral infection (**Appendix E-4B**). Notably, with the use of Ingenuity Pathway Analysis (IPA [Qiagen]), we identified the interferon (IFN) signaling pathway to be among the top pathways altered upon TCRV infection. IFNs are a family of cytokines secreted by host cells in response to viruses and other pathogens to confer antiviral states upon uninfected neighboring cells in an effort to prevent spread of infection (32). Given that the IFN response has been explored in bats in regard to pathogen-host response (33), we then further examined the relationship between these factors within the spleen, kidney, and liver in response to TCRV infection and found that while most of the identified IFN pathway-related genes were upregulated, all of the factors identified in this pathway had statistically significant differential expression (log2 fold change =2; FDR=0.01) in the spleen (**Appendix E-4C**). We validated differential expression of select immune genes via reverse transcription-quantitative PCR (RT-qPCR) and confirmed upregulation of ISG15 and IRF7 in the spleen and kidney tissues and downregulation of HLA-DRA in the kidney.

In addition to the IFN-signaling pathway, we identified signaling pathways for Toll-like receptors (TLRs) and interleukin-6 (IL-6) (among other cytokines), as well as pathways for T-helper cell differentiation and the Th1 pathway (**Appendix E-4B**). Further analysis of all DEGs via the Reactome plugin (Cytoscape) identified additional key pathways involved in the immune response. Specifically, we identified increased transcript levels of several cytokine genes (IL6, IL8, IL1A, IL1B, and IFNG) and chemokine genes (CXCL1, CXCL2, CXCL3, CXCL5, and CXCL6). To highlight

markers associated with circulating immune cells, we focused on those DEGs that were common to two or more tissues, and the data are consistent with increased infiltration of neutrophils into the infected tissues. In kidneys, neutrophil infiltration can cause hyper-inflammation and kidney damage (34). This is further supported by the presence of enriched expression levels of neutrophil gelatinase-associated lipocalin (NGALl) in all three tissues, which is a biomarker for renal damage in humans (35).

Transcripts for IgG, IgM, IgA, and IgE were identified in the spleen data, and the level of each was significantly elevated in the infected bats. Six transcripts of IgG heavy chains were identified, including 5 with complete and distinct V regions. The six IgG constant regions were identical, other than one that contained a Thr in place of an Ala, which could represent an allele or a sequencing error. The hinge regions, which are frequently different between IgG subclasses within a species, were identical in all 6 transcripts. These two features suggest that Jamaican fruit bats have a single IgG isotype. The 5 V regions contain the canonical mammalian Ig sequences, including 4 framework regions (FR) and 3 complementarity-determining regions (CDRs) (**Appendix E-5**). Three distinct CDR3 sequences were present in these 5 transcripts. Two had 2 tyrosine residues, whereas the three that were identical had 6, substantially more than what has been reported in the CDR3s of other bat species (36, 37). Sequences for T-cell receptor alpha constant region domains were present in the assembly, although none had complete V regions. Expression of TCR-α was identified in all uninfected and infected tissues, TCR-β in all spleen and liver tissues, TCR-γ in all spleen and liver tissues, and TCR-δ in all spleen and uninfected liver tissues.

Gene ontology of annotated differentially expressed genes in TCRV-infected tissues

To characterize the overall transcriptome in response to TCRV infection, we performed an unbiased evaluation of the top 10 Reactome pathways (ranked by P value) associated with DEGs in various organs. In all three tissues, genes controlling cell cycle progression were elevated, including many associated with hypoxia, cell stress, senescence, and chromatin organization.

Spleen differential expression analysis indicated that immune system pathways were significantly elevated, including type I and II IFN signaling, antiviral IFN-stimulated genes (ISGs), interleukin signaling, and T and B cell activation pathways. Interestingly, genes involved in the complement cascade were repressed, including the genes for phosphatidylinositol 3-kinase and complement receptor 1. SH2B1, a gene encoding an important signal transduction adaptor in several pathways, including JAK, prolactin, platelet-derived growth factor, and nerve growth factor signaling, was also significantly downregulated in the spleen (38). We also identified repression of pro-apoptotic genes BMP and PMAIP1 as well as repression of genes involved in calcium mobilization following TCRV infection (39).

In the liver, Reactome analysis revealed strong immune activation signatures, including T-cell receptor and CD28 costimulatory signaling. No evidence of B-cell or NK cell activities was present. TLR and RIG-I/MDA5 signaling for type I IFN responses was also elevated, despite no evidence of differentially expressed type I IFN genes. Unlike the spleen, complement pathways were also enriched. The IFN-γ signaling pathway was also identified, although IFNG itself was not differentially expressed. Despite these findings, further indications of apoptotic activation were not differentially

expressed in the livers. As in the spleen, many metabolic genes were elevated, whereas genes involved in calcium mobilization were repressed. In the kidneys, Reactome analysis suggested that platelet calcium-associated degranulation may occur; the genes F13A1 and TMSB4X were elevated along with other genes involved in calcium mobilization. As in the other tissues, evidence of innate antiviral responses was present, including TLR signaling, RIG-I/MDA5 activity, and type I IFN signaling genes. Interleukin and IFN-γ signaling were also among the pathways characterized secondary to gene upregulation in the kidneys.

Discussion

Our previous work demonstrated that TCRV is pathogenic to Jamaican fruit bats and has allowed us to perform the most complete examination to date of a pathogenic virus infection in a bat species (22). Despite their importance to human health as reservoirs of emerging viruses, the characterization of infections in bats at the cellular and molecular levels has been limited relative to other model organisms, such as rodents. Fortunately, the emerging advantage of next-generation sequencing technologies has been fundamental to our understanding of disease responses; however, minimal reference data sets are available for bats. To this end, our group was among the first to perform next-generation sequencing on bats with a small-scale Illumina sequencing of kidney and lung tissues in a single library from the Jamaican fruit bat (8,11). Furthermore, in the present study we generated a high-quality transcriptomic data set for the Jamaican fruit bat and comprehensively profiled the altered immune genes in response to TCRV infection.

To gain insight into the pathogenic infection of Jamaican fruit bats, we performed high-throughput RNA sequencing of TCRV-infected spleen, liver, and kidney tissues and corresponding sham-inoculated controls. We produced high-quality non-redundant reads, and our Trinity de novo assembly resulted in 349,855 transcripts, which were further assembled into 124,204 contigs. The number of non-redundant contigs we identified is similar to those from other transcriptome assemblies reported for the black flying fox (126,378) (10), Rickett's big-footed bat (104,987), and the greater short-nosed fruit bat (171,394) (40).

We employed a pairwise comparison of all infected tissues versus sham-inoculated controls to identify altered gene expression levels upon pathogenic TCRV infection. We utilized a log2 fold change cutoff of >2 with an FDR of <0.01. We chose a stringent cutoff because of our small sample size (n = 2) for each tissue type under each condition. This revealed approximately 25% more genes upregulated than downregulated. The spleen is instrumental in systemic and local immune responses and has been used to study viral responses in many organisms, including bats (3, 41). We observed the greatest number of differentially expressed genes in spleen tissues compared to the liver and kidneys. Further analysis revealed that the majority of these differentially expressed genes identified in the spleen belonged to immune-related pathways.

Ingenuity Pathway Analysis identified enrichment of the helper T cell differentiation and Th1 pathway (**Appendix E-4B**) genes IFNG, IFNGR2, IL12RB2, IL6ST, SOCS1, and SOCS2, supporting a role for mobilization of a Th1 response. Despite this, levels of helper T cell genes, such as CD4 or T-cell receptor (TCR) genes, were not statistically different in infected bats. CD4 sequences were not in the assembly,

suggesting the CD4 level was below the threshold of depth of RNA-Seq. TCR-α, TCR-β, TCR-γ, and TCR-δ sequences were present in the assembly, and they appear to share features found in TCRs of other species. There was insufficient sequence data to evaluate TCR variable, diversity, or joining segments for T-cell receptors. Further studies using next-generation repertoire sequencing will be required to fully examine the TCR loci. Unfortunately, without monoclonal antibodies to identify CD4+ or CD8- cells by flow cytometry, it is difficult to determine whether T cells are expanding in response to infection. Other indicators of T-cell activation include the elevated expression of granzyme A and B genes (GZMA, GZMB), IL-12 and CCL5 (RANTES), and the activated T-cell chemotactic factor gene CXCL1 in the spleens of infected bats.

Transcripts for IgG, IgM, IgA, and, interestingly, IgE were significantly higher in the infected bats. IgE is not typically associated with viral infections, but has been associated with anaphylaxis after influenza vaccinations (42–44). No transcripts for IgD were present in the transcriptome, similar to what has been observed for other microbats (45). Alignments of the 6 IgG transcripts were identical, except for one transcript that had a Thr instead of Ala at position 395, which likely represents an allele or sequencing error. Only one IgG transcript has been found in Seba's fruit bat (Carollia perspicillata) (45); thus, it is not unexpected that Jamaican fruit bats may only have a single IgG isotype. The Jamaican fruit bat IgG shares 94% identity and 96% similarity with the Seba's fruit bat IgG constant region. The hinge regions of all IgG transcripts were also identical and distinct from those of Seba's fruit bat IgG. Hinge regions are generally considered hallmark indicators of IgG subclasses (46). It is possible that Jamaican fruit bats have IgG subclasses but without a genome or transcriptome profiling of Ig

transcripts this question could be difficult to address.

The heavy-chain variable regions of the 5 Jamaican fruit bat transcripts showed many differences, suggesting they represent distinct segments and multiple V region gene families. The limited number of V regions makes it difficult to assign Jamaican fruit bat sequences to gene families. We are unable to estimate the number of V, D, or J segments with the transcriptome data; however, bats appear to have much larger numbers of these segments than most mammals (36, 37). It is noteworthy that the three CDR3 regions have more tyrosine residues than are found in most other bat species immunoglobulins. The presence of tyrosines is thought to contribute to antibody interactions with a spectrum of epitopes (47–49), and the lack of these in bat antibodies has been postulated to account for why bats have generally poorer responses to infectious agents (50, 51).

Only a single variable region light-chain sequence was significantly elevated in the infected bats, which had most similarity to the IgLV7 variable gene family. Studies of big brown bats (Eptisicus fuscus) suggest they express predominantly, if not exclusively, λ light chains; thus our findings are similar (52). Considering that a single light chain was elevated in infected bats, it may be possible to clone this cDNA and co-express it with each of the 5 heavy-chain sequences described herein to determine if the antibodies are reactive to TCRV antigens.

We also detected elevated expression of polymeric immunoglobulin receptor (PIGR), which exports IgA antibodies across the epithelium into mucosa (53), in the spleens of infected bats. Considering the presence of TCRV in oral and rectal swabs and

in the lungs (22), it is likely that virus-specific IgA is present at these sites. The development of antibodies to artibeus IgA will be necessary to verify this.

The principal gene for somatic hypermutation (SHM) that leads to affinity maturation is activation-induced cytidine deaminase (AID) (54); however, despite its presence in all four bats in this study, its expression was not significantly elevated in the spleens. Other genes involved in SHM (54) were elevated, including those coding for DNA polymerase θ (POLQ), polymerase (POLN), and replication protein A (RPA). The level of APOBEC3, coding for another RNA-editing enzyme with lower SHM activity (55), was not elevated. Examination of little brown bat (Myotis lucifugus) antibody cDNA sequences suggests bats do not use SHM to a great extent (36), and our findings are congruent with this observation. In our previous work with TCRV and MERS coronavirus (CoV) infection of Jamaican fruit bats (22, 56), antibody responses were poor, suggesting that affinity maturation is limited in bats.

Global differential expression evaluation of TCRV-infected tissues revealed alterations in calcium mobilization, a characteristic mechanism of host response to infection by viruses, including arenaviruses (57). Additionally, our analysis revealed few indications of NK cell activation and minimal expression of genes that are associated with T-cell exhaustion (i.e., Ly6e and Fcgr3). It is noteworthy that bats appear to be missing many NK cell-associated genes (10, 58, 59); thus it may be that the functions of bat NK cells are substantially different from those of human or mouse NK cells. We detected increased IFNG, GZMA, and GZMB expression in the spleen infected tissues, and while these proteins are produced by both NK cells and T cells, we believe their presence correlates more strongly with a T-cell origin due to the increased number of T-

cell associated genes upregulated relative to NK cell genes. Moreover, the bats in this report were euthanized on days 10 and 11, a time point at which T-cell activation should be occurring. Thus, T-cell exhaustion, a feature of some lymphocytic choriomeningitis virus (LCMV) isolates, is likely not occurring in bats infected with TCRV (60, 61).

We identified several genes associated with neutrophil activation. These results areconsistent with our previous histopathological findings in this species, where we noted neutrophilic infiltration that was likely a result of proliferating lymphocytes (22). Additionally, our results are also consistent with a recent Lassa virus isolate from Mali that similarly induces neutrophil infiltration in nonhuman primates (62). The abundant expression of NGAL may provide a diagnostic tool; its protein, neutrophil gelatinase associated lipocalin, is secreted in the urine, which is detectable with commercially available diagnostic kits (e.g., Pacific Biomarkers, Seattle, WA).

A recent study looked at differential gene expression in an embryonic cell line from Egyptian fruit bats infected with Marburg virus (63). In contrast to the observed host responses in the Egyptian fruit bat cells, we identified the JAK/STAT signaling pathway as one of the immune-related pathways upregulated in response to pathogenic TCRV infection, suggesting a contributory role for this pathway in pathogenesis. Additionally, a study exploring the innate immune response to Newcastle disease virus in large flying fox cells, a newly characterized a subset of antiviral factors was found (64). Among these factors was the CHAC1 gene, which we identified to be 4-fold upregulated in spleen and kidney tissues. Together, this evidence along with our previous pathogenicity studies shows that a typical antiviral response occurs to TCRV in Jamaican fruit bats.

We focused additional analyses on immune genes with similar expression in all tissues (cluster 3 [**Appendix E-4A**]). Among the top pathways identified was the IFN signaling pathway. The signaling factors in this pathway exert their antiviral activities through the induction of other antiviral proteins (32). The IFN response has been explored in bat cells (33), and in all bat species examined, the type I IFN locus has undergone substantial contraction, with only three functional IFN-α genes but with constitutive IFN-α expression in at least one species (65). Specifically, recent discoveries have revealed enhanced IFN signaling in antiviral immunity and have identified its involvement in arenavirus response mechanisms. We therefore furthered our analysis regarding these pathways (51, 66). We found that most IFN signaling genes identified in this subset were upregulated in all tissues; however, 5 of these genes had no significant differential expression identified in the kidney and 2 had none in the liver. Notably, of the differentially expressed factors, the IFNAR1 gene was downregulated in the spleen and BCL2 was downregulated in the kidney. Apoptotic pathways play a critical role as defense mechanisms for a host when infected by a viral pathogen; BCL2 encodes an anti-apoptotic protein that is known to be involved in a typical antiviral response (67), and the observed downregulation of BCL2 in the kidneys upon TCRV infection suggests promotion of apoptotic pathways stimulated by IFN signaling in response to infection. In contrast, BCL2 was determined to be upregulated in spleen and liver tissues. Additionally, another anti-apoptotic factor gene, Mcl-1, was also upregulated in these tissues. Recent work with mice infected with LCMV, as well as other studies, has demonstrated the involvement of these factors in promoting naive T-cell survival and memory T cell activation (68, 69). Together, these results support congruency of our

annotated transcriptome given what is known about the coordination of immune genes altered in response to viral infection as well as the identification of genes specific to the antiviral response in bats (61).

As might be expected during an acute antiviral response, IFN-stimulated gene 15 (ISG15) was elevated in infected tissues. Reactome pathway analysis identified ISG15 in several immune pathways, including the innate immune response, cytokine signaling, IFN-α/β signaling, and RIG-I/MDA5-mediated induction of IFN-α/β pathways, which has also been previously identified in a bat pathogenic viral response (64). ISG15 is an important gene in the innate immune response, particularly the type I IFN antiviral response; however, the ISG15-encoded protein has recently been demonstrated to have additional functions as a ubiquitin-like modifier that covalently conjugates to other cellular proteins to form an "ISGylated" complex (70). Various roles of ISG15 have been identified in immune responses; when secreted extracellularly, ISG15 can act to drive expression of IFN-γ, which was elevated in the spleen. Alternatively, intracellular expression can modulate type I IFN signaling (71).

Although IFN-α, IFNB, IFNL, and IFNG transcripts were present in all of the tissues analyzed, the only differentially expressed transcript was IFNG in the spleen. In contrast, indications of downstream signaling initiated by IFN type I and type II were present, suggesting either transcript turnover prior to the time of sample collections or the potential for alternative routes of pathway activation. Previous work examining in vitro infection of the black flying fox with Tioman virus suggests a prominent role for IFN-γ (72). To this end, there are potential differences between bat species in terms of their responses to viruses that may account for apathogenic infections (e.g., reservoir hosts) or

disease. Future work with cell culture from our model system may help to clarify these points. Furthermore, the DEGs involved in IFN-α/β signaling suggest that a typical antiviral innate immune response occurred in the bats. Within the spleen, expression of 36% of the genes was upregulated in the IFN-α/β signaling pathway, whereas the kidneys and livers had approximately 16% and 34% elevated expression of these same genes, respectively. This indicates a more robust type I IFN response in the spleen.

Together, de novo transcriptome analysis of our high-throughput RNA-Seq data from Jamaican fruit bats infected with TCRV provides a high-quality data set and also a comprehensive gene expression analysis of immune gene expression responses in bats during a pathogenic infection. This data set will provide a strong basis for additional analyses. Further investigation of our identified pathways in vitro and in vivo will significantly contribute to our understanding of pathogenic viral infections in bats. Moreover, the data here will facilitate future experimental studies of artibeus bats and their cells, which have been used as models for MERS CoV and Zaire Ebola virus and which are suspected reservoirs of the recently discovered bat influenza viruses (56,73, 74).

Conclusion

This study provides a comprehensive analysis of the transcriptional landscape of Jamaican fruit bats during infection with Tacaribe virus. This natural pathogen of artibeus bats causes high-mortality disease with similar clinical manifestations to the South American hemorrhagic fevers and Lassa fever. In summary, this analysis identified the global response to TCRV infection. Our results suggest diverse immune responses,

including alterations in neutrophil activation, interferon signaling, markers for lymphocytes, and antibodies. We found substantial signatures of neutrophil activation in the spleen, kidney, and liver of bats with fatal disease. The innate and adaptive immune response appeared to be functional and typical of the canonical antiviral response. Many activation markers of T and B lymphocytes were also found; however, few indications of NK cell activity or T-cell exhaustion were apparent. IgG, IgM, IgA, and IgE sequences were abundantly expressed in the spleens of infected bats, and five immunoglobulin heavy-chain V segments were identified. Despite the clear evidence of antibody synthesis during infection, AID expression was not elevated, suggesting somatic hypermutation and affinity maturation were absent or minimal. Analysis of immunoglobulin heavy-chain and TCR V regions suggests that Jamaican fruit bats have canonical immunoglobulin and TCR genes found in most mammals. Moreover, the species appears to have a single IgG subclass. These results are the most extensive gene discovery work completed in Jamaican fruit bats to date and the first to describe differential immune gene expression in bats during a pathogenic virus infection.

Materials and Methods

Experimental TCRV infection in bats, sample collection, and RNA extraction

Experimental infections of Jamaican fruit bats were previously reported (22). Briefly, two Jamaican fruit bats were inoculated with 100 µl of sterile Dulbecco's phosphate-buffered saline (DPBS) as negative controls (bat IDs 688 and 689), and two bats were inoculated with 100 µl containing 106 TCID50 TCRV (bat IDs 714 and 729). Negative control bats were euthanized at the end of the experimental period (45 days),

whereas TCRV-infected bats were euthanized as they became moribund (days 11 and 18, respectively). Necropsies were performed directly following euthanasia, and organs were harvested and collected in RNAlater stabilization reagent (Qiagen). RNA was extracted from flash-frozen tissues by homogenization with a Mini Bead Beater (BioSpec Products, Inc.), using QiaShredder columns with the RNeasy kit (Qiagen).

RNA-Seq

Stranded Illumina libraries for each tissue were prepared from total RNA using the NEB Ultra Directional RNA library prep kit with poly(A) selection. Sequencing (paired-end 100 bp) was performed on the Illumina HiSeq-2000 platform at the UC Denver Genomics core.

Read processing and assembly

For transcriptome assembly, raw reads were filtered for adapter sequences and low-quality reads, and assembly was performed using Trinity (30) with the following parameters: —min_contig_length 300 —min_glue 3—min_kmer_cov 2. Resulting contigs were processed for read alignment and abundance estimation with Bowtie and RSEM (75, 76). Differential expression was performed using the edgeR package within the Trinity differential analysis pipeline using default parameters (77). A pairwise comparison was made between TCRV-infected samples and control uninfected samples. Genes were considered differentially expressed with an FDR of <0.01 and a log2 fold change > 2.

Gene ontology and pathway analysis

BLAST alignments and functional annotations were performed using Blast2GO Pro or Ingenuity Pathway Analysis (78, 79). Direct pathway analysis for immune-related genes was performed using the gene list from the ImmPort database (31).

Immunoglobulin sequence analysis

Contigs for immunoglobulins were translated using the default translation table of MacVector software. MUSCLE alignments were made to identify leader, framework regions, and complementarity-determining regions of the V segments using a black flying fox sequence as a reference (NCBI GenBank accession no. ADD71702.1) (50). Heavy chains and hinge regions were identified by BLAST against other Chiroptera.

RT-qPCR validation of RNA-Seq data

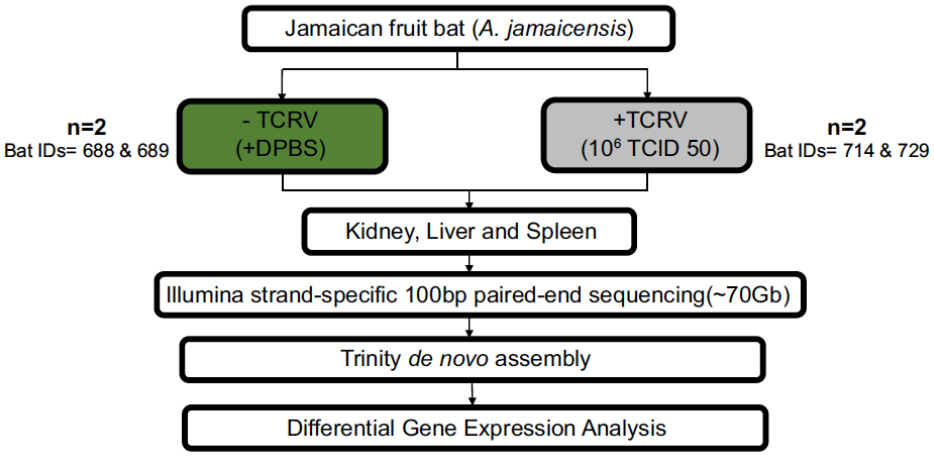
The experimental primer sequences used in RT-qPCR analysis are listed in Fig. S2. cDNA was generated using SuperScript III reverse transcriptase (Thermo Scientific) and SYBR Select master mix for CFX (Applied Biosystems). The same two Jamaican fruit bats that were used for uninfected samples in RNA-Seq were also used as uninfected samples for RT-qPCR.

Accession number(s)

Raw reads have been deposited into GenBank under GenBank accession no. GSE75771.

Figures

Appendix E-1: Transcriptomic analysis of Jamaican fruit bats infected with Tacaribe virus (TCRV)



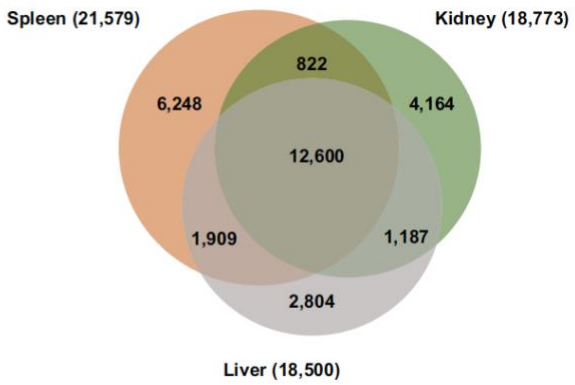
Appendix E-1: Transcriptomic analysis of Jamaican fruit bats infected with Tacaribe virus (TCRV). Jamaican fruit bats were inoculated with either TCRV or DPBS (n=2 for each condition). De novo assembly of the Jamaican fruit bat transcriptome was performed using RNA-Seq data from kidney, liver, and spleen tissues. Differentially expressed genes were then identified in the uninfected and infected tissues using edgeR.

Appendix E-2: De novo assembly of the Jamaican fruit bat transcriptome

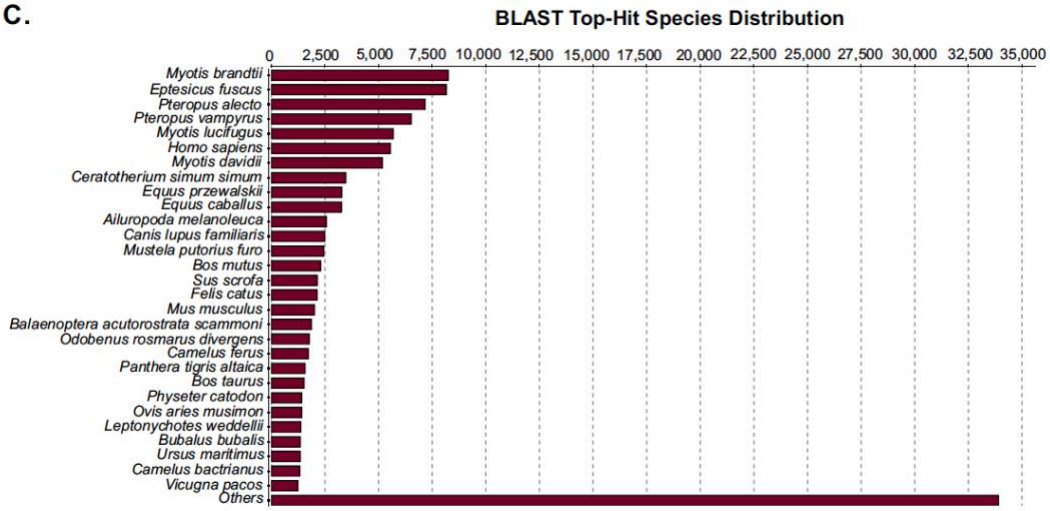
A.

Assembly Metric	Value
Total Trinity 'genes'	175,144
Total Trinity transcripts	349,855
Percent GC	50.96
Longest contig (bp)	27,622
# of contigs >20kb	100
# of contigs >10kb	4,118
# of contigs >5kb	36,613
Transcript contig N50	3,419
Median contig length (bp)	997
Average contig length (bp)	1137.44
Total assembeled bases	644,933,364

B.

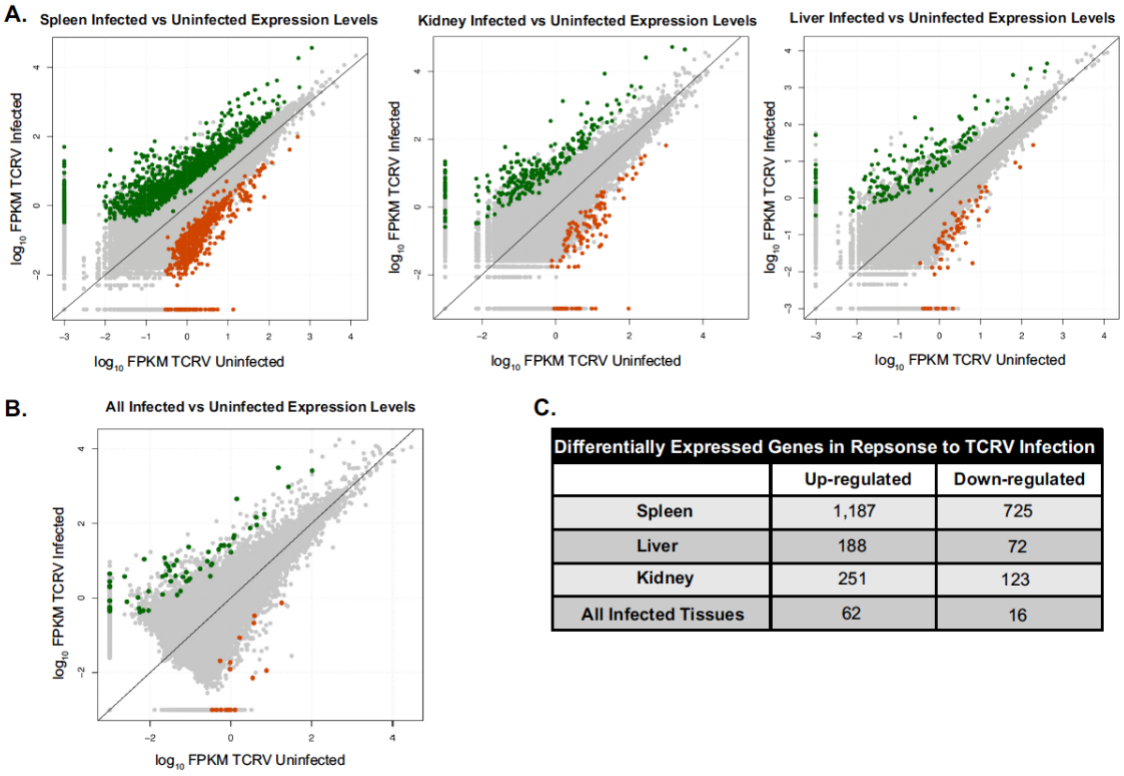


C.



Appendix E-2: De novo assembly of the Jamaican fruit bat transcriptome. (A) Trinity assembly resulted in the construction of the de novo-assembled Jamaican fruit bat transcriptome with 644,933,364 assembled bases. (B) Examination of the identified contigs from the transcriptome assembly showed that 35% (12,600) are expressed in the spleen, kidney, and liver tissues (FRKM, greater than 1). (C) We compared this transcriptome assembly to those of other mammals through BLASTX analysis and identified transcripts similar to those present in other bat species.

Appendix E-3: Differential gene expression analysis following TCRV infection in Jamaican fruit bats

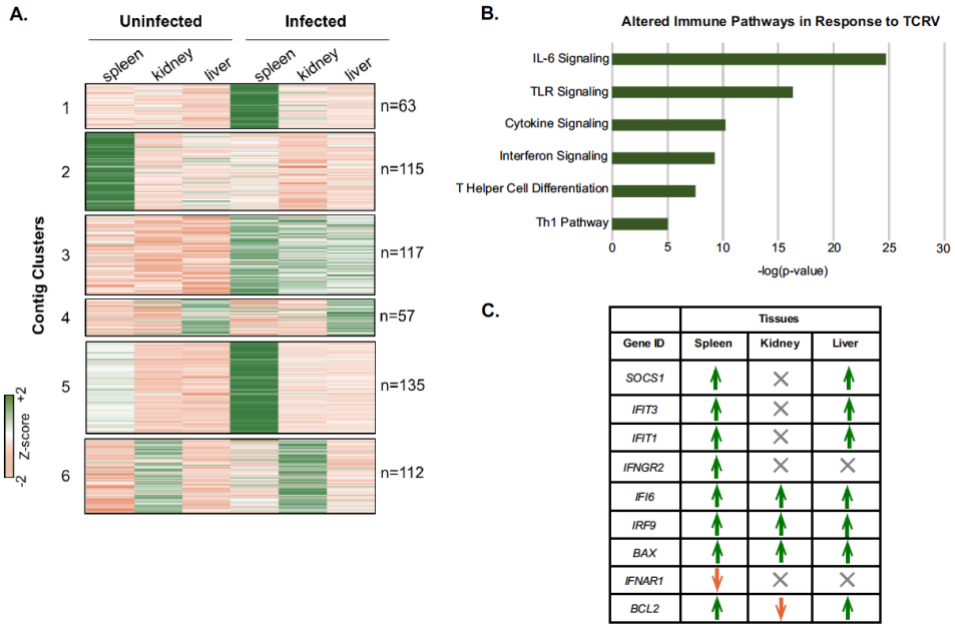


C.

Differentially Expressed Genes in Repsonse to TCRV Infection		
	Up-regulated	Down-regulated
Spleen	1,187	725
Liver	188	72
Kidney	251	123
All Infected Tissues	62	16

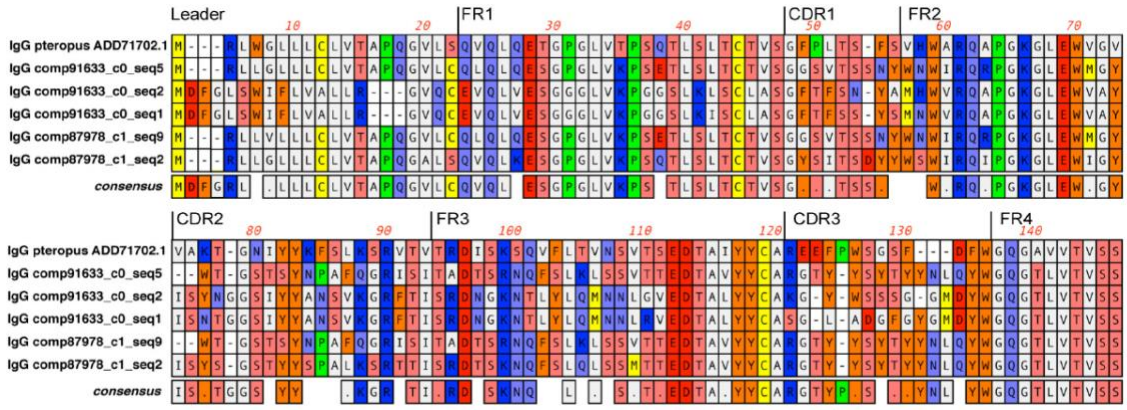
Appendix E-3: Differential gene expression analysis following TCRV infection in Jamaican fruit bats. We used a pairwise comparison of TCRV-infected samples against the corresponding control uninfected samples and found that the expression levels of hundreds of genes were altered with TCRV infection in the different tissues. (A) Differential expression analysis revealed upregulated genes (green) and downregulated genes (orange) defined by edgeR (log2 fold change > 2 and FDR < 0.01). (B) Inspection of altered genes in all infected tissues versus control tissues showed fewer changed genes common to all tissues. (C) Quantification of differentially expressed genes from panels A and B.

Appendix E-4: Immune-specific expression analysis of TCRV-infected Artibeus jamaicensis bats



Appendix E-4: Immune-specific expression analysis of TCRV-infected Artibeus jamaicensis bats. (A) Immport database immune-related genes and their expression values (FPKM) were clustered (k-means = 6) to investigate the relationship between uninfected and infected tissues. (B) We performed Ingenuity Pathway Analysis to characterize the specific immune pathways for those genes identified in cluster 3 (n = 117) from panel A. (C) Interferon signaling was among the top pathways identified to be altered after TCRV infection. We then identified these specific genes involved in interferon signaling and explored their alterations in the different tissues in response to TCRV infection. Green corresponds to upregulated and orange to downregulated (FDR, <0.01; log2 fold change, >2); gray indicates no significant differential expression.

Appendix E-5: Amino acid alignment of Jamaican fruit bat IgG V regions



Appendix E-5: Amino acid alignment of Jamaican fruit bat IgG V regions. Five V region transcripts were identified in the Jamaican fruit bat spleen transcriptomes and aligned with a V region of an annotated black flying fox V region (50). Framework regions (FR1 to FR4) and complementarity-determining regions (CDR1 to CDR3) were identified and exhibited differences between each transcript.

References

1. IUCN. 2015. IUCN Red List Version 2015.1, Table 3a- Status Category summary by major taxonomic group (animals), *on* Conservation International, Arizona State University, Texas A&M University, University of Rome, University of Virginia, Zoological Society London. . http://www.iucnredlist.org/mammals. Accessed June 11, 2015.
2. IUCN. 2015. IUCN Red List Version 2015.1, Table 4a- Number of species in each IUCN Red List Category in each major animal taxonomic group (class, order), *on* Conservation International, Arizona State University, Texas A&M University, University of Rome, University of Virginia, Zoological Society London. . http://www.iucnredlist.org/mammals. Accessed June 11, 2015.
3. Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. 2006. Bats: important reservoir hosts of emerging viruses. Clin Microbiol Rev 19:531-545.
4. Campbell CAR. 1925. Bats, mosquitoes, and dollars. The Stratford Co., Boston, Mass.
5. Kuno G. 2001. Persistence of arboviruses and antiviral antibodies in vertebrate hosts: its occurrence and impacts. Rev Med Virol 11:165-190.
6. Jayme SI, Field HE, de Jong C, Olival KJ, Marsh G, Tagtag AM, Hughes T, Bucad AC, Barr J, Azul RR, Retes LM, Foord A, Yu M, Cruz MS, Santos IJ, Lim TM, Benigno CC, Epstein JH, Wang LF, Daszak P, Newman SH. 2015. Molecular evidence of Ebola Reston virus infection in Philippine bats. Virol J 12:107.

7. Lau SK, Woo PC, Li KS, Huang Y, Tsoi HW, Wong BH, Wong SS, Leung SY, Chan KH, Yuen KY. 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci U S A 102:14040-14045.

8. Memish ZA, Mishra N, Olival KJ, Fagbo SF, Kapoor V, Epstein JH, Alhakeem R, Durosinloun A, Al Asmari M, Islam A, Kapoor A, Briese T, Daszak P, Al Rabeeah AA, Lipkin WI. 2013. Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. Emerg Infect Dis 19:1819-1823.

9. Kunz TH, and Linda F. Lumsden. 2003. Bat Ecology. The University of Chicago Press, Chicago.

10. Shaw TI, Srivastava A, Chou WC, Liu L, Hawkinson A, Glenn TC, Adams R, Schountz T. 2012. Transcriptome sequencing and annotation for the Jamaican fruit bat (Artibeus jamaicensis). PLoS One 7:e48472.

11. Kapusta A, Suh A, Feschotte C. 2017. Dynamics of genome size evolution in birds and mammals. Proc Natl Acad Sci U S A 114:E1460-E1469.

12. Papenfuss AT, Baker ML, Feng ZP, Tachedjian M, Crameri G, Cowled C, Ng J, Janardhana V, Field HE, Wang LF. 2012. The immune gene repertoire of an important viral reservoir, the Australian black flying fox. BMC Genomics 13:261.

13. Wang L-F. 2009. Bats and Viruses: a Breif Review. Virologica Sinica 24:93-99.

14. Wibbelt G, Moore MS, Schountz T, Voigt CC. 2010. Emerging diseases in Chiroptera: why bats? Biol Lett 6:438-440.

15. Cogswell-Hawkinson AC, Mitchell E. McGlaughlin, Charles H. Calisher, Rick Adams, and Tony Schountz. 2011. Molecular and Phylogenetic Characterization

of Cytokine Genes from Seba's Short-tailed Fruit Bat (*Carollia perspicillata*). Open Immunology Journal 4:31-39.

16. O'Shea TJ, Cryan PM, Cunningham AA, Fooks AR, Hayman DT, Luis AD, Peel AJ, Plowright RK, Wood JL. 2014. Bat flight and zoonotic viruses. Emerg Infect Dis 20:741-745.

17. Omatsu T, Watanabe S, Akashi H, Yoshikawa Y. 2007. Biological characters of bats in relation to natural reservoir of emerging viruses. Comp Immunol Microbiol Infect Dis 30:357-374.

18. Luis AD, Hayman DT, O'Shea TJ, Cryan PM, Gilbert AT, Pulliam JR, Mills JN, Timonin ME, Willis CK, Cunningham AA, Fooks AR, Rupprecht CE, Wood JL, Webb CT. 2013. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? Proc Biol Sci 280:20122753.

19. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, Tachedjian M, Zhu Y, Zhou P, Jiang X, Ng J, Yang L, Wu L, Xiao J, Feng Y, Chen Y, Sun X, Zhang Y, Marsh GA, Crameri G, Broder CC, Frey KG, Wang LF, Wang J. 2013. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science 339:456-460.

20. Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, Coleman JT, Darling SR, Gargas A, Niver R, Okoniewski JC, Rudd RJ, Stone WB. 2009. Bat white-nose syndrome: an emerging fungal pathogen? Science 323:227.

21. Frick WF, Pollock JF, Hicks AC, Langwig KE, Reynolds DS, Turner GG, Butchkoski CM, Kunz TH. 2010. An emerging disease causes regional population collapse of a common North American bat species. Science 329:679-682.

22. Gargas A, M.T. Trest, M. Christensen, T.J. Volk, D. S. Blehert. 2009. *Geomyces destructans* sp. nov. associated with bat white-nose syndrome. Mycotaxon 108:147-154.

23. Schountz T. 2014. Immunology of bats and their viruses: challenges and opportunities. Viruses 6:4880-4901.

24. Aguilar-Setien A, Loza-Rubio E, Salas-Rojas M, Brisseau N, Cliquet F, Pastoret PP, Rojas-Dotor S, Tesoro E, Kretschmer R. 2005. Salivary excretion of rabies virus by healthy vampire bats. Epidemiol Infect 133:517-522.

25. Pawan JL. 1959. Rabies in the vampire bat of Trinidad, with special reference to the clinical course and the latency of infection. Caribb Med J 21:137-156.

26. Sulkin SE, Allen R. 1974. Virus infections in bats. Monogr Virol 8:1-103.

27. Turmelle AS, Jackson FR, Green D, McCracken GF, Rupprecht CE. 2010. Host immunity to repeated rabies virus infection in big brown bats. J Gen Virol 91:2360-2366.

28. Wang LF, Walker PJ, Poon LL. 2011. Mass extinctions, biodiversity and mitochondrial function: are bats 'special' as reservoirs for emerging viruses? Curr Opin Virol 1:649-657.

29. Cogswell-Hawkinson A, Bowen R, James S, Gardiner D, Calisher CH, Adams R, Schountz T. 2012. Tacaribe virus causes fatal infection of an ostensible reservoir host, the Jamaican fruit bat. J Virol 86:5791-5799.

30. Bowen MD, Peters CJ, Nichol ST. 1996. The phylogeny of New World (Tacaribe complex) arenaviruses. Virology 219:285-290.

31. Bowen MD, Peters CJ, Nichol ST. 1997. Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. Mol Phylogenet Evol 8:301-316.

32. Downs WG, C.R. Anderson, L. Spence, T.H.G. Aitken, and A.H. Greenhall. 1963. Tacaribe Virus, a New Agent Isolated from Artibeus Bats and Mosquitos in Trinidad, West Indies. American Journal of Tropical Medicine and Hygiene 12:640-646.

33. Fulhorst CF, Bowen MD, Salas RA, Duno G, Utrera A, Ksiazek TG, De Manzione NM, De Miller E, Vasquez C, Peters CJ, Tesh RB. 1999. Natural rodent host associations of Guanarito and pirital viruses (Family *Arenaviridae*) in central Venezuela. Am J Trop Med Hyg 61:325-330.

34. Milazzo ML, Barragan-Gomez A, Hanson JD, Estrada-Franco JG, Arellano E, Gonzalez-Cozatl FX, Fernandez-Salas I, Ramirez-Aguilar F, Rogers DS, Bradley RD, Fulhorst CF. 2010. Antibodies to Tacaribe serocomplex viruses (family *Arenaviridae*, genus *Arenavirus*) in cricetid rodents from new Mexico, Texas, and Mexico. Vector Borne Zoonotic Dis 10:629-637.

35. Sayler KA, Barbet AF, Chamberlain C, Clapp WL, Alleman R, Loeb JC, Lednicky JA. 2014. Isolation of Tacaribe virus, a Caribbean arenavirus, from host-seeking *Amblyomma americanum* ticks in Florida. PLoS One 9:e115769.

36. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A,

Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644-652.

37. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494-1512.

38. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, Berger P, Desborough V, Smith T, Campbell J, Thomson E, Monteiro R, Guimaraes P, Walters B, Wiser J, Butte AJ. 2014. ImmPort: disseminating data to the public for the future of immunology. Immunol Res 58:234-239.

39. de Weerd NA, Samarajiwa SA, Hertzog PJ. 2007. Type I interferon receptors: biochemistry and biological functions. J Biol Chem 282:20053-20057.

40. Kuzmin IV, Schwarz TM, Ilinykh PA, Jordan I, Ksiazek TG, Sachidanandam R, Basler CF, Bukreyev A. 2017. Innate Immune Responses of Bat and Human Cells to Filoviruses: Commonalities and Distinctions. J Virol 91.

41. Majer O, Bourgeois C, Zwolanek F, Lassnig C, Kerjaschki D, Mack M, Muller M, Kuchler K. 2012. Type I interferons promote fatal immunopathology by regulating inflammatory monocytes and neutrophils during *Candida* infections. PLoS Pathog 8:e1002811.

42. Martensson J, Bellomo R. 2014. The rise and fall of NGAL in acute kidney injury. Blood Purif 37:304-310.

43. Bratsch S, Wertz N, Chaloner K, Kunz TH, Butler JE. 2011. The little brown bat, *M. lucifugus*, displays a highly diverse V H, D H and J H repertoire but little evidence of somatic hypermutation. Dev Comp Immunol 35:421-430.

44. Larsen PA, Marchan-Rivadeneira MR, Baker RJ. 2010. Natural hybridization generates mammalian lineage with species characteristics. Proc Natl Acad Sci U S A 107:11447-11452.

45. Maures TJ, Kurzer JH, Carter-Su C. 2007. SH2B1 (SH2-B) and JAK2: a multifunctional adaptor protein and kinase made for each other. Trends Endocrinol Metab 18:38-45.

46. Pecina-Slaus N. 2010. Wnt signal transduction pathway and apoptosis: a review. Cancer Cell Int 10:22.

47. Dong D, Lei M, Liu Y, Zhang S. 2013. Comparative inner ear transcriptome analysis between the Rickett's big-footed bats (*Myotis ricketti*) and the greater short-nosed fruit bats (*Cynopterus sphinx*). BMC Genomics 14:916.

48. Jones ME, Schuh AJ, Amman BR, Sealy TK, Zaki SR, Nichol ST, Towner JS. 2015. Experimental Inoculation of Egyptian Rousette Bats (*Rousettus aegyptiacus*) with Viruses of the Ebolavirus and Marburgvirus Genera. Viruses 7:3420-3442.

49. Nagao M, Fujisawa T, Ihara T, Kino Y. 2016. Highly increased levels of IgE antibodies to vaccine components in children with influenza vaccine-associated anaphylaxis. J Allergy Clin Immunol 137:861-867.

50. Nakayama T, Kumagai T, Nishimura N, Ozaki T, Okafuji T, Suzuki E, Miyata A, Okada K, Ihara T. 2015. Seasonal split influenza vaccine induced IgE sensitization against influenza vaccine. Vaccine 33:6099-6105.

51. Smith-Norowitz TA, Kusonruksa M, Wong D, Norowitz MM, Joks R, Durkin HG, Bluth MH. 2012. Long-term persistence of IgE anti-influenza A HIN1 virus antibodies in serum of children and adults following influenza A vaccination with subsequent H1N1 infection: a case study. J Inflamm Res 5:111-116.

52. Butler JE, Wertz N, Zhao Y, Zhang S, Bao Y, Bratsch S, Kunz TH, Whitaker JO, Jr., Schountz T. 2011. The two suborders of chiropterans have the canonical heavy-chain immunoglobulin (Ig) gene repertoire of eutherian mammals. Dev Comp Immunol 35:273-284.

53. Butler JE, Wertz N, Deschacht N, Kacskovics I. 2009. Porcine IgG: structure, genetics, and evolution. Immunogenetics 61:209-230.

54. Berens SJ, Wylie DE, Lopez OJ. 1997. Use of a single VH family and long CDR3s in the variable region of cattle Ig heavy chains. Int Immunol 9:189-199.

55. Golub R, Fellah JS, Charlemagne J. 1997. Structure and diversity of the heavy chain VDJ junctions in the developing Mexican axolotl. Immunogenetics 46:402-409.

56. Reynaud CA, Dahan A, Anquez V, Weill JC. 1989. Somatic hyperconversion diversifies the single Vh gene of the chicken with a high incidence in the D region. Cell 59:171-183.

57. Baker ML, Tachedjian M, Wang LF. 2010. Immunoglobulin heavy chain diversity in Pteropid bats: evidence for a diverse and highly specific antigen binding repertoire. Immunogenetics 62:173-184.

58. Brook CE, Dobson AP. 2015. Bats as 'special' reservoirs for emerging zoonotic pathogens. Trends Microbiol 23:172-180.

59. Lee WT, Jones DD, Yates JL, Winslow GM, Davis AD, Rudd RJ, Barron CT, Cowan C. 2016. Identification of secreted and membrane-bound bat immunoglobulin using a Microchiropteran-specific mouse monoclonal antibody. Dev Comp Immunol 65:114-123.

60. Shimada S, Kawaguchi-Miyashita M, Kushiro A, Sato T, Nanno M, Sako T, Matsuoka Y, Sudo K, Tagawa Y, Iwakura Y, Ohwaki M. 1999. Generation of polymeric immunoglobulin receptor-deficient mouse with marked reduction of secretory IgA. J Immunol 163:5367-5373.

61. Teng G, Papavasiliou FN. 2007. Immunoglobulin somatic hypermutation. Annu Rev Genet 41:107-120.

62. Halemano K, Guo K, Heilman KJ, Barrett BS, Smith DS, Hasenkrug KJ, Santiago ML. 2014. Immunoglobulin somatic hypermutation by APOBEC3/Rfv3 during retroviral infection. Proc Natl Acad Sci U S A 111:7759-7764.

63. Munster VJ, Adney DR, van Doremalen N, Brown VR, Miazgowicz KL, Milne-Price S, Bushmaker T, Rosenke R, Scott D, Hawkinson A, de Wit E, Schountz T, Bowen RA. 2016. Replication and shedding of MERS-CoV in Jamaican fruit bats (*Artibeus jamaicensis*). Sci Rep 6:21878.

64. Han Z, Madara JJ, Herbert A, Prugar LI, Ruthel G, Lu J, Liu Y, Liu W, Liu X, Wrobel JE, Reitz AB, Dye JM, Harty RN, Freedman BD. 2015. Calcium Regulation of Hemorrhagic Fever Virus Budding: Mechanistic Implications for Host-Oriented Therapeutic Intervention. PLoS Pathog 11:e1005220.

65. Lee AK, Kulcsar KA, Elliott O, Khiabanian H, Nagle ER, Jones ME, Amman BR, Sanchez-Lockhart M, Towner JS, Palacios G, Rabadan R. 2015. De novo transcriptome reconstruction and annotation of the Egyptian rousette bat. BMC Genomics 16:1033.

66. Lee MS, Park CH, Jeong YH, Kim YJ, Ha SJ. 2013. Negative regulation of type I IFN expression by OASL1 permits chronic viral infection and $CD8^+$ T-cell exhaustion. PLoS Pathog 9:e1003478.

67. Penaloza-MacMaster P, Kamphorst AO, Wieland A, Araki K, Iyer SS, West EE, O'Mara L, Yang S, Konieczny BT, Sharpe AH, Freeman GJ, Rudensky AY, Ahmed R. 2014. Interplay between regulatory T cells and PD-1 in modulating T cell exhaustion and viral control during chronic LCMV infection. J Exp Med 211:1905-1918.

68. Safronetz D, Strong JE, Feldmann F, Haddock E, Sogoba N, Brining D, Geisbert TW, Scott DP, Feldmann H. 2013. A recently isolated Lassa virus from Mali demonstrates atypical clinical disease manifestations and decreased virulence in cynomolgus macaques. J Infect Dis 207:1316-1327.

69. Holzer M, Krahling V, Amman F, Barth E, Bernhart SH, Carmelo VA, Collatz M, Doose G, Eggenhofer F, Ewald J, Fallmann J, Feldhahn LM, Fricke M, Gebauer J, Gruber AJ, Hufsky F, Indrischek H, Kanton S, Linde J, Mostajo N,

Ochsenreiter R, Riege K, Rivarola-Duarte L, Sahyoun AH, Saunders SJ, Seemann SE, Tanzer A, Vogel B, Wehner S, Wolfinger MT, Backofen R, Gorodkin J, Grosse I, Hofacker I, Hoffmann S, Kaleta C, Stadler PF, Becker S, Marz M. 2016. Differential transcriptional responses to Ebola and Marburg virus infection in bat and human cells. Sci Rep 6:34589.

70. Glennon NB, Jabado O, Lo MK, Shaw ML. 2015. Transcriptome Profiling of the Virus-Induced Innate Immune Response in *Pteropus vampyrus* and Its Attenuation by Nipah Virus Interferon Antagonist Functions. J Virol 89:7550-7566.

71. Zhou P, Tachedjian M, Wynne JW, Boyd V, Cui J, Smith I, Cowled C, Ng JH, Mok L, Michalski WP, Mendenhall IH, Tachedjian G, Wang LF, Baker ML. 2016. Contraction of the type I IFN locus and unusual constitutive expression of IFN-alpha in bats. Proc Natl Acad Sci U S A 113:2696-2701.

72. Koma T, Huang C, Kolokoltsova OA, Brasier AR, Paessler S. 2013. Innate immune response to arenaviral infection: a focus on the highly pathogenic New World hemorrhagic arenaviruses. J Mol Biol 425:4893-4903.

73. Hardwick JM, Bellows DS. 2003. Viral versus cellular BCL-2 proteins. Cell Death Differ 10 Suppl 1:S68-76.

74. Sochalska M, Tuzlak S, Egle A, Villunger A. 2015. Lessons from gain- and loss-of-function models of pro-survival Bcl2 family proteins: implications for targeted therapy. FEBS J 282:834-849.

75. Tripathi P, Koss B, Opferman JT, Hildeman DA. 2013. Mcl-1 antagonizes Bax/Bak to promote effector $CD4^+$ and $CD8^+$ T-cell responses. Cell Death Differ 20:998-1007.

76. Zhao C, Denison C, Huibregtse JM, Gygi S, Krug RM. 2005. Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. Proc Natl Acad Sci U S A 102:10200-10205.

77. Zhang X, Bogunovic D, Payelle-Brogard B, Francois-Newton V, Speer SD, Yuan C, Volpi S, Li Z, Sanal O, Mansouri D, Tezcan I, Rice GI, Chen C, Mansouri N, Mahdaviani SA, Itan Y, Boisson B, Okada S, Zeng L, Wang X, Jiang H, Liu W, Han T, Liu D, Ma T, Wang B, Liu M, Liu JY, Wang QK, Yalnizoglu D, Radoshevich L, Uze G, Gros P, Rozenberg F, Zhang SY, Jouanguy E, Bustamante J, Garcia-Sastre A, Abel L, Lebon P, Notarangelo LD, Crow YJ, Boisson-Dupuis S, Casanova JL, Pellegrini S. 2015. Human intracellular ISG15 prevents interferon-alpha/beta over-amplification and auto-inflammation. Nature 517:89-93.

78. Zhou P, Cowled C, Todd S, Crameri G, Virtue ER, Marsh GA, Klein R, Shi Z, Wang LF, Baker ML. 2011. Type III IFNs in pteropid bats: differential expression patterns provide evidence for distinct roles in antiviral immunity. J Immunol 186:3138-3147.

79. Miller MR, McMinn RJ, Misra V, Schountz T, Muller MA, Kurth A, Munster VJ. 2016. Broad and Temperature Independent Replication Potential of Filoviruses on

Cells Derived From Old and New World Bat Species. J Infect Dis 214:S297-S302.

80. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, Chen LM, Johnson A, Tao Y, Dreyfus C, Yu W, McBride R, Carney PJ, Gilbert AT, Chang J, Guo Z, Davis CT, Paulson JC, Stevens J, Rupprecht CE, Holmes EC, Wilson IA, Donis RO. 2013. New world bats harbor diverse influenza A viruses. PLoS Pathog 9:e1003657.

81. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

82. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323.

83. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

84. Conesa A, Gotz S. 2008. Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008:619832.

85. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420-3435.

86. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalatskaya I, Mahajan S, May B, Ndegwa N, Schmidt E, Shamovsky V, Yung C, Birney E, Hermjakob H, D'Eustachio P, Stein

L. 2011. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 39:D691-697.

87. Ono K, Demchak B, Ideker T. 2014. Cytoscape tools for the web age: D3.js and Cytoscape.js exporters. F1000Res 3:143.

Appendix F

Data access for publically available datasets

The following tables contain data access information for publically available datasets discussed in this dissertation.

Table 1: Data accession corresponding to publically available datasets (abbreviations: Ch=Chapter; Rep=Replicate)

Ch	Type	Cell line	Sample	Target	Rep	Database	Accession
2	ChIP-seq	PANC1	parental	TCF7L2	1	ENCODE	ENCLB828OGC
2	ChIP-seq	PANC1	parental	TCF7L2	2	ENCODE	ENCLB718WRS
2,3	ChIP-seq	PANC1	parental	H3K27ac	1	ENCODE	ENCLB737KTI
2,3	ChIP-seq	PANC1	parental	H3K27ac	2	ENCODE	ENCLB797ODR
2,3	ChIP-seq	PANC1	parental	H3K4me1	1	ENCODE	ENCLB482XLE
2,3	ChIP-seq	PANC1	parental	H3K4me1	2	ENCODE	ENCLB346ZDV
2,3	ChIP-seq	PANC1	parental	H3K4me3	1	ENCODE	ENCLB792WMR
2,3	ChIP-seq	PANC1	parental	H3K4me3	2	ENCODE	ENCLB135CDR
2	ChIP-seq	PANC1	parental	H3K36me3	1	ENCODE	ENCLB555ABM
2	ChIP-seq	PANC1	parental	H3K36me3	2	ENCODE	ENCLB555ABN

2	ChIP-seq	PANC1	parental	H3K27me3	1	ENCODE	ENCLB555ABQ
2	ChIP-seq	PANC1	parental	H3K27me3	2	ENCODE	ENCLB555ABR
2	ChIP-seq	PANC1	parental	H3K9me3	1	ENCODE	ENCLB555ABO
2	ChIP-seq	PANC1	parental	H3K9me3	2	ENCODE	ENCLB555ABP
2,3	ChIP-seq	PANC1	parental	Input	1,2	ENCODE	ENCLB499BUK
3	ChIP-seq	PANC1	parental		1	ENCODE	ENCLB941MSL
3	ChIP-seq	PANC1	parental		2	ENCODE	ENCLB342MOC
3	ChIP-seq	CAPAN1	parental	H3K27ac	1,2	GEO	GSM1574235
3	ChIP-seq	CAPAN2	parental	H3K27ac	1,2	GEO	GSM1574236
3	ChIP-seq	CFPAC1	parental	H3K27ac	1,2	GEO	GSM1574237
3	ChIP-seq	HPAF2	parental	H3K27ac	1,2	GEO	GSM1574238
3	ChIP-seq	MiaPaca2	parental	H3K27ac	1,2	GEO	GSM1574239
3	ChIP-seq	PANC1	parental	H3K27ac	1,2	GEO	GSM1574240

3	ChIP-seq	PT45P1	parental	H3K27ac	1,2	GEO	GSM1574241
3	ChIP-seq	CAPAN1	parental	H3K4me3	1,2	GEO	GSM1574256
3	ChIP-seq	CAPAN2	parental	H3K4me3	1,2	GEO	GSM1574258
3	ChIP-seq	CFPAC1	parental	H3K4me3	1,2	GEO	GSM1574259
3	ChIP-seq	HPAF2	parental	H3K4me3	1,2	GEO	GSM1574260
3	ChIP-seq	MiaPaca2	parental	H3K4me3	1,2	GEO	GSM1574261
3	ChIP-seq	PANC1	parental	H3K4me3	1,2	GEO	GSM1574262
3	ChIP-seq	PT45P1	parental	H3K4me3	1,2	GEO	GSM1574263
3	ChIP-seq	CAPAN1	parental	Input	1,2	GEO	GSM1574271
3	ChIP-seq	MiaPaca2	parental	Input	1,2	GEO	GSM1574272
4	ChIP-seq	MCF7	parental/0hr	Input	1,2	GEO	GSM2913215
4	ChIP-seq	MCF7	parental/0hr	H3K4me3	1,2	GEO	GSM2913216
4	ChIP-seq	MCF7	parental/0hr	H3K27ac	1,2	GEO	GSM2913217

4	ChIP-seq	MCF7	parental/0hr	H3K4me1	1,2	GEO	GSM2913218
4	ChIP-seq	MCF7	parental/0hr	H3K27me3	1,2	GEO	GSM2913219
4	ChIP-seq	MCF7	parental/0hr	H3K9me3	1,2	GEO	GSM2913220
4	ChIP-seq	MCF7	parental/0hr	ER-alpha	1,2	GEO	GSM2913221
4	ChIP-seq	MCF7	parental/0hr	CTCF	1,2	GEO	GSM2913222
4	ChIP-seq	MCF7	100nM Estradiol/1hr	Input	1,2	GEO	GSM2913223
4	ChIP-seq	MCF7	100nM Estradiol/1hr	H3K4me3	1,2	GEO	GSM2913224
4	ChIP-seq	MCF7	100nM Estradiol/1hr	H3K27ac	1,2	GEO	GSM2913225
4	ChIP-seq	MCF7	100nM Estradiol/1hr	H3K4me1	1,2	GEO	GSM2913226
4	ChIP-seq	MCF7	100nM Estradiol/1hr	H3K27me3	1,2	GEO	GSM2913227
4	ChIP-seq	MCF7	100nM Estradiol/1hr	H3K9me3	1,2	GEO	GSM2913228
4	ChIP-seq	MCF7	100nM Estradiol/1hr	ER-alpha	1,2	GEO	GSM2913229
4	ChIP-seq	MCF7	100nM Estradiol/1hr	CTCF	1,2	GEO	GSM2913230

4	ChIP-seq	MCF7	100nM Estradiol/4hr	Input	1,2	GEO	GSM2913231
4	ChIP-seq	MCF7	100nM Estradiol/4hr	H3K4me3	1,2	GEO	GSM2913232
4	ChIP-seq	MCF7	100nM Estradiol/4hr	H3K27ac	1,2	GEO	GSM2913233
4	ChIP-seq	MCF7	100nM Estradiol/4hr	H3K4me1	1,2	GEO	GSM2913234
4	ChIP-seq	MCF7	100nM Estradiol/4hr	H3K27me3	1,2	GEO	GSM2913235
4	ChIP-seq	MCF7	100nM Estradiol/4hr	H3K9me3	1,2	GEO	GSM2913236
4	ChIP-seq	MCF7	100nM Estradiol/4hr	ER-alpha	1,2	GEO	GSM2913237
4	ChIP-seq	MCF7	100nM Estradiol/4hr	CTCF	1,2	GEO	GSM2913238
4	ChIP-seq	MCF7	100nM Estradiol/16hr	Input	1,2	GEO	GSM2913239
4	ChIP-seq	MCF7	100nM Estradiol/16hr	H3K4me3	1,2	GEO	GSM2913240
4	ChIP-seq	MCF7	100nM Estradiol/16hr	H3K27ac	1,2	GEO	GSM2913241
4	ChIP-seq	MCF7	100nM Estradiol/16hr	H3K4me1	1,2	GEO	GSM2913242
4	ChIP-seq	MCF7	100nM Estradiol/16hr	H3K27me3	1,2	GEO	GSM2913243

4	ChIP-seq	MCF7	100nM Estradiol/16hr	H3K9me3	1,2	GEO	GSM2913244
4	ChIP-seq	MCF7	100nM Estradiol/16hr	ER-alpha	1,2	GEO	GSM2913245
4	ChIP-seq	MCF7	100nM Estradiol/16hr	CTCF	1,2	GEO	GSM2913246
4	ChIP-seq	MCF7	100nM Estradiol/24hr	Input	1,2	GEO	GSM2913247
4	ChIP-seq	MCF7	100nM Estradiol/24hr	H3K4me3	1,2	GEO	GSM2913248
4	ChIP-seq	MCF7	100nM Estradiol/24hr	H3K27ac	1,2	GEO	GSM2913249
4	ChIP-seq	MCF7	100nM Estradiol/24hr	H3K4me1	1,2	GEO	GSM2913250
4	ChIP-seq	MCF7	100nM Estradiol/24hr	H3K27me3	1,2	GEO	GSM2913251
4	ChIP-seq	MCF7	100nM Estradiol/24hr	H3K9me3	1,2	GEO	GSM2913252
4	ChIP-seq	MCF7	100nM Estradiol/24hr	ER-alpha	1,2	GEO	GSM2913253
4	ChIP-seq	MCF7	100nM Estradiol/24hr	CTCF	1,2	GEO	GSM2913254
2	DNase-seq	PANC1	parental	DNase	1	ENCODE	ENCLB574ZZZ
2	DNase-seq	PANC1	parental	DNase	2	ENCODE	ENCLB573ZZZ
A	expression	PANC1	10μM ICG-001	12hr	1,2,3	GEO	GSM1563237

	beadchip						
2,A	expression beadchip	PANC1	10μM ICG-001	96hr	1,2	GEO	GSM1563238
A	expression beadchip	PANC1	10μM C646	12hr	1,2	GEO	GSM1563239
2,A	expression beadchip	PANC1	10μM C646	96hr	1,2	GEO	GSM1563240
A	expression beadchip	PANC1	0.05% DMSO	12hr	1,2	GEO	GSM1563235
2,A	expression beadchip	PANC1	0.05% DMSO	96hr	1,2	GEO	GSM1563236
2	HiC	PANC1	parental		1	GEO	GSM2827313
2	HiC	PANC1	parental		2	GEO	GSM2827314
2,A	RNA-seq	PANC1	siTCF7L2	48hr	1	GEO	GSM1556985
2,A	RNA-seq	PANC1	siTCF7L2	48hr	2	GEO	GSM1556986
2,A	RNA-seq	PANC1	siTCF7L2	48hr	3	GEO	GSM1556987
2,A	RNA-seq	PANC1	siControl	48hr	1	GEO	GSM1556982
2,A	RNA-seq	PANC1	siControl	48hr	2	GEO	GSM1556983
2,A	RNA-seq	PANC1	siControl	48hr	3	GEO	GSM1556984
3	RNA-seq	CAPAN1	parental		1	GEO	GSM1574297
3	RNA-seq	CAPAN1	parental		2	GEO	GSM1574298
3	RNA-seq	CAPAN2	parental		1	GEO	GSM1574299
3	RNA-seq	CAPAN2	parental		2	GEO	GSM1574300
3	RNA-seq	CFPAC1	parental		1	GEO	GSM1574301

3	RNA-seq	CFPAC1	parental		2	GEO	GSM1574302
3	RNA-seq	HPAF2	parental		1	GEO	GSM1574303
3	RNA-seq	HPAF2	parental		2	GEO	GSM1574304
3	RNA-seq	MiaPaca 2	parental		1	GEO	GSM1574305
3	RNA-seq	MiaPaca 2	parental		2	GEO	GSM1574306
3	RNA-seq	PANC1	parental		1	GEO	GSM1574307
3	RNA-seq	PANC1	parental		2	GEO	GSM1574308
3	RNA-seq	PT45P1	parental		1	GEO	GSM1574309
3	RNA-seq	PT45P1	parental		2	GEO	GSM1574310
4	RNA-seq	MCF7	Parental		1	GEO	GSM2913255
4	RNA-seq	MCF7	Parental		2	GEO	GSM2913256
4	RNA-seq	MCF7	Parental		3	GEO	GSM2913257
4	RNA-seq	MCF7	TamR		1	GEO	GSM2913267
4	RNA-seq	MCF7	TamR		2	GEO	GSM2913268
4	RNA-seq	MCF7	TamR		3	GEO	GSM2913269
2,3	TCC	PANC1	parental		1,2	GEO	GSM1684570
4	TCC	MCF7	parental	0hr	1	GEO	GSM2913210
4	TCC	MCF7	100nM Estradiol	1hr	1	GEO	GSM2913211
4	TCC	MCF7	100nM Estradiol	4hr	1	GEO	GSM2913212
4	TCC	MCF7	100nM Estradiol	16hr	1	GEO	GSM2913213
4	TCC	MCF7	100nM Estradiol	24hr	1	GEO	GSM2913214
4	TCC	MCF7	TamR		1	GEO	GSM2913258
4	TCC	MCF7	TamR	Input	1,2	GEO	GSM2913259
4	TCC	MCF7	TamR	H3K4me3	1,2	GEO	GSM2913260
4	TCC	MCF7	TamR	H3K27ac	1,2	GEO	GSM2913261
4	TCC	MCF7	TamR	H3K4me1	1,2	GEO	GSM2913262
4	TCC	MCF7	TamR	H3K27me3	1,2	GEO	GSM2913263

4	TCC	MCF7	TamR	H3K9me3	1,2	GEO	GSM2913264
4	TCC	MCF7	TamR	ER-alpha	1,2	GEO	GSM2913265
4	TCC	MCF7	TamR	CTCF	1,2	GEO	GSM2913266
4	TCC	T47D	parental	0hr	1,2	GEO	GSM3386607
4	TCC	T47D	100nM Estradiol	1hr	1,2	GEO	GSM3386608
4	TCC	T47D	100nM Estradiol	4hr	1,2	GEO	GSM3386609
4	TCC	T47D	100nM Estradiol	16hr	1,2	GEO	GSM3386610
4	TCC	T47D	100nM Estradiol	24hr	1,2	GEO	GSM3386611
4	TCC	T47D	TamR		1,2	GEO	GSM3386612
E	RNA-seq	Jamaican fruit bat	Spleen	Control	1	GEO	GSM1967338
E	RNA-seq	Jamaican fruit bat	Spleen	Control	2	GEO	GSM1967339
E	RNA-seq	Jamaican fruit bat	Spleen	Infected	1	GEO	GSM1967340
E	RNA-seq	Jamaican fruit bat	Spleen	Infected	2	GEO	GSM1967341
E	RNA-seq	Jamaican fruit bat	Kidney	Control	1	GEO	GSM1967342
E	RNA-seq	Jamaican fruit bat	Kidney	Control	2	GEO	GSM1967343
E	RNA-seq	Jamaican fruit bat	Kidney	Infected	1	GEO	GSM1967344
E	RNA-seq	Jamaican fruit bat	Kidney	Infected	2	GEO	GSM1967345
E	RNA-seq	Jamaican fruit bat	Liver	Control	1	GEO	GSM1967346
E	RNA-seq	Jamaican fruit bat	Liver	Control	2	GEO	GSM1967347
E	RNA-seq	Jamaican fruit bat	Liver	Infected	1	GEO	GSM1967348
E	RNA-seq	Jamaican Fruit bat	Liver	Infected	2	GEO	GSM1967349