The Effects Enhancers on the Kinetics of Transcription

Benjamin Clark
Master of Science

Experimental Medicine
McGill University
Montreal, Quebec, Canada

August 13, 2024

A thesis submitted to McGill University in partial
fulfillment of the requirements of the degree of
Master of Science

Abstract

At most loci across the eukaryotic genome, gene transcription has long been observed to be regulated in a punctuated, burst-like manner. Burst transcription, revealed by single molecule fluorescence techniques, is characterized by promoters being active for brief periods, capable of producing dozens to hundreds of nascent RNAs in a span of several minutes, followed by a long gap of transcriptional senescence ranging from minutes to several hours. Using a simple two-state model of transcription in tandem with smFISH or single-cell RNAseq, researchers can estimate both the frequency and size of these bursts in absolute numbers. A common observation at many isolated transgenes is that different enhancers impart different bursting frequencies and that both genomic distance and insulation by topologically-associated domain boundaries tend to negatively impact frequency. Furthermore, it has been observed that Mediator, an essential transcriptional co-factor, can initiate transcription when it forms condensates and enters proximity with a promoter and its associated enhancer. Mediator has long been understood to have multiple roles in transcription, both at enhancers by interacting with transcription factors and with the core promoter by facilitating RNA polymerase II activity. It is not well understood how Mediator nor how enhancers interacting with promoters shape transcriptional kinetics across the whole genome. In this project we investigate these relationships using two approaches. Firstly, we develop an experimental mouse model whereby Mediator can be partially depleted and used in conjunction with allele-specific, single-cell RNA sequencing to estimate whole genome bursting kinetics. Secondly, using public chromatin conformation and kinetics data in mouse embryonic fibroblasts we investigate the role of enhancer-promoter contacts in shaping kinetics across the genome. When examining multivalent promoter-enhancer interactions, enhancers appear to have a dual role in transcription. Burst frequency does not increase after a promoter interacts with more than two enhancers while burst size increases linearly with enhancer contacts, suggesting enhancers both help initiate transcription and maintain it via different models of cooperativity. Meanwhile, super-enhancers induce greater bursting frequencies at cognate promoters while burst size is increased at proximal and overlapping genes, suggesting chromatin looping is sufficient but not necessary for super-enhancer activity. Lastly, topologically-associated domain insulation and genomic distance of single-linked promoters do not appear to have any consequence on kinetics.

Abrégé

Dans la plupart des locus du génome eucaryote, on observe depuis longtemps que la transcription des gènes est régulée de manière ponctuelle et en rafale. La transcription en rafale, révélée par les techniques de fluorescence à molécule unique, est caractérisée par des promoteurs actifs pendant de brèves périodes, capables de produire des dizaines ou des centaines d'ARN naissants en l'espace de quelques minutes, suivies d'une longue période de sénescence transcriptionnelle allant de quelques minutes à plusieurs heures. En utilisant un modèle simple de transcription à deux états en tandem avec le smFISH ou le séquençage d'ARN à cellules uniques, les chercheurs peuvent estimer à la fois la fréquence et la taille de ces rafales en nombres absolus. Une observation commune à de nombreux transgènes isolés est que différents activateurs transmettent différentes fréquences de rafales et que la distance génomique et l'isolation par des limites de domaine topologiquement associées ont tendance à avoir un impact négatif sur la fréquence. En outre, il a été observé que Mediator, un cofacteur essentiel de la transcription, peut initier la transcription lorsqu'il forme des condensats et entre à proximité d'un promoteur et de l'enhancer qui lui est associé. On sait depuis longtemps que Mediator joue un rôle multiple dans la transcription, à la fois au niveau des enhancers en interagissant avec les facteurs de transcription et au niveau du promoteur proximal en facilitant l'activité de l'ARN polymérase II. On ne sait pas très bien comment Mediator ou les enhancers interagissant avec les promoteurs façonnent la cinétique de la transcription sur l'ensemble du génome. Dans ce projet, nous étudions ces relations selon deux approches. Tout d'abord, nous développons un modèle expérimental de souris dans lequel Mediator peut être partiellement supprimé et nous l'utilisons en conjonction avec le séquençage de l'ARN de cellules uniques spécifique d'un allèle pour estimer la cinétique de rafales transcriptionnelles à l'échelle du génome entier. Deuxièmement, en utilisant les données publiques de conformation de la chromatine et de et de cinétiques transcriptionnelles dans les fibroblastes embryonnaires de souris, nous étudions le rôle des contacts enhancer-promoteur dans le façonnement de la cinétique à travers le génome. En examinant les interactions promoteur-enhancer multivalentes, les enhancers semblent avoir un double rôle dans la transcription. La fréquence des rafales n'augmente pas lorsqu'un promoteur interagit avec plus de deux enhancers, alors que la taille des rafales augmente de façon linéaire avec les contacts entre promoteurs-enhancers, ce qui suggère que les enhancers contribuent à la fois à initier la transcription et à la maintenir par le biais de différents modèles de coopérativité. Parallèlement, les super-enhancers induisent des fréquences de rafales plus importantes au niveau des promoteurs apparentés, tandis que la taille des rafales augmente au niveau des gènes proximaux et chevauchants, ce qui suggère que le bouclage de la chromatine est suffisant, mais pas nécessaire, pour l'activité des super-enhancers. Enfin, l'isolation des domaines topologiquement associés et la distance génomique des promoteurs à lien unique ne semblent

pas avoir de conséquences sur la cinétique.

Contributions

Dr.Francois Robert offered guidance and suggestions on all aspects of the research, as well as editorial work for all chapters. Dr.Martin Sauvageau provided materials and suggestions for the experiments in chapter two. Marie-Ève Wedge prepared the mouse embryonic fibroblasts used in chapter two. Additional training for experiments in chapter two provided by Élie Lambert. All experiments and data analysis was performed by the author unless stated otherwise.

Acknowledgements

I have deepest gratitude to the members of our lab that helped me grow as student and a scientist these past two years. I am indebted to Elie Lambert for helping me getting my feet wet with my first experiments and Christian Poitras for helping me get oriented with my bioinformatics work. I am thankful for Kavindu Puwakdandawa and Célia Jeronimo for their advice, encouragement and our many moments of commiseration over failed experiments.

I would like to extend my thanks to the members of the Sauvageau lab, David Ferland-McCollough and Jean-François Laurendeau for helping enter the world of mammalian cell culture, as well as Marie-Ève Wedge for preparing the cell line which became the center of my project.

I would also like to thank the members of my thesis committee for their enthusiastic support, Martin Sauvageau, Mohan Malleshaiah and Jun Ding. And lastly, a million thanks to my supervisor, Francois Robert for his endless patience and guidance over the years.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

Possibly one of the most fundamental questions in molecular biology is '*how are genes expressed?*' And while many technologies have emerged in last few decades that have become pillars of the discipline, these primarily address the relative abundance of transcripts between biological contexts. It remained an open question until this past decade as to what kind of regiment of expression gene promoters typically have; are they productive constitutively or periodically? Or some combination of both?

One of the first insights into these questions came from Miller Jr. and McKnight's seminal electron microscopy work on drosophila embryos [47]. Visible ribonucleoprotein fibres emerging laterally from the chromosome, believed to be nascent transcripts, appeared to be arranged in clusters at different positions along the gene body. The arrangements of clusters suggested a pattern of discontinuous expression whereby transcription occurs in short bursts, producing many RNAs at once. This notion will be revisited and formalized upon the arrival of newer fluorescent imaging techniques. Early experiments using flow cytometery in yeast demonstrated significant heterogeneity among cells harbouring fluorescent markers with identical promoters [22, 6]. Dubbed 'transcriptional noise', measurements of the coefficient of variation (CV^2) by Elowitz et al. between cells and alleles using two fluorescent reporters confirmed a pulsatile regime of gene expression [22].

Using single molecule fluorescence in-situ hybridization (smFISH), many groups followed the intuition of discontinuous expression by either measuring single cell RNA counts or by monitoring real-time transcription of a variety of genes [14, 60, 77, 61, 4, 98]. From these early reports, the emergence of a 'telegraph' or 'bursty' model of transcription was introduced to explain and estimate kinetics rates [4]. In this simplified model, gene promoters can only persist in either an ON or OFF state, whose rates of transitions between can be estimated from fixed mRNA counts among individual cells or by directly monitoring the production of fluorescently labelled mRNAs . In the next section we will focus on discrete estimators.

1.1.1 Burst Transcription Estimators From Discrete Data

There are many methods for estimating kinetic parameters, the most common being a frequentist approach such as the maximum-likelihood estimator (MLE) on a assumed probability density function (PDF). More precisely, the PDF can be defined as either a Markovian life and death process [55], or more recently by a beta-poisson distribution [85, 90]. The former has the advantage of being able to approximate parameters from common distribution measures which has been used to great effect in smFISH studies[4, 17, 50, 97, 77]. Recent advances in single-cell RNA sequencing (scRNA-seq) has enabled researchers estimate whole-transcriptome bursting kinetics. The beta-poisson model of expression supports a long-tail distribution of transcripts and is more robust to drop-outs and low integer counts typical of scRNA-seq data [85, 90, 36]. Using this model, the fraction of time a gene is active is defined as [90, 38]:

$$\bar{p} = \frac{Kon}{Kon + Koff}$$

The individual parameters can be contextualized as such:

$$x|K_{syn}, p \sim Poisson(K_{syn}p)$$

$$p|K_{on}, K_{off} \sim Beta(K_{on}, K_{off})$$

Where *Ksyn* is the rate of nascent transcription, and *Kon* and *Koff* are the rates of transition between active and inactive states [38, 85, 90]. *Kon* is typically characterized as *burst frequency* while *burst size* can be estimated by $\frac{K_{syn}}{K_{off}}$. The analytical solution for the PDF for its use with MLE has been provided by Vu et al. with extended parameters to accommodate non-discrete values [85]. Alternatively, a hierarchical bayesian approach can be used as demonstrated by Kim and Marioni [38].

The MLE approach has been extended by Larsson et al. for its use allele specific scRNA-seq [16, 40, 41]. By using a hybrid mouse strain and full length transcript sequencing platform (SmartSeq), reads containing individual single nucleotide polymorphisms (SNPs) can attributed to their allele of origin. From this it was found nearly 12 to 24% of autosomal genes in mouse embryonic stem cells (mESCs) were monoallelic in expression [16]. For Larsson et al., nearly 300 genes out of the 5 thousand captured in mouse embryonic fibroblasts (MEFs) differed significantly for either kinetic parameter [40]. Allelic imbalance in expression could mostly be explained by independent differences in burst size and frequency [41]. From this we can infer the importance of allele specificity when estimating the nuances of expression.

1.1.2 Timescales of Expression at the Core Promoter

In order to understand transcription, and further its kinetics, it is important to identify the disparate events which compose it. It is also important to note that the two-state model of expression is a phenomenological model; it makes no assumptions of the underlying molecular mechanisms. It is the researcher's game to perturb biological contexts and theorize on its kinetic outcomes.

Perhaps the most obvious place to interrogate is the formation of the pre-initiation complex (PIC). The PIC is composed of RNA polymerase II (PolII), which is itself made up of

12 subunits, and five general transcription factors: TFIIB, TFIID, TFIIE, TFIIF, TFIIH [65]. TFIID itself is composed of the TATA binding protein (TBP) and TBP associated factors (TAFs). For transcription to occur, all of the above components most assemble at the core promoter, proposed to occur in a classical stepwise manner [65]. Mediator, a large multi-subunit cofactor, stabilizes the formation of the PIC and facilitates for the release of PolII into productive elongation [3] . The presence of a re-initiation complex has been proposed whereby elements of the PIC remains after the polymerase is released, allowing for immediate reintegration of a new polymerase in waiting [96]. More recently, Tantale et al. leveraged the use of sensitive, fluorescent live-cell nascent mRNA imaging and proposed a 'polymerase convoy' model of expression at an HIV promoter in fly cells [79]. Under a such a model, tens of polymerases fire from the promoter and travel the length of the gene body equidistant from each other. Such a model gives a mechanistic basis for transcriptional bursting in highly active genes whereby multiple convoys could compose an active ON period at the promoter. Following the intuition that the TBP remains in the reinitiation complex, Tantale et al. mutated the HIV TATA box. Interestingly, the mutations had no effect on the average number of polymerases acting in a convoy but increased the probability of an OFF state occurring during the eight hour recording period [79]. These findings suggest that the binding of the TBP and stability at the TATA box mark the temporal boundaries of the promoter state. Contradicting this, a single molecule tracking (SMT) study in human breast cancer cells showed that, while much longer than other elements of the PIC, the TBP had an average global residence time of 9.6s, falling far below the average ON time of ten minutes [54]. However, the study also observed both a global increase in burst duration and a doubling TBP residence times when the oncogene MYC was overexpressed, suggesting a functional link. The mere presence of a TATA box itself has been repeatedly shown to be associated with greater burst sizes, both from in vivo imaging data and from fixed single-cell parameter estimates [40, 56, 53]. From this perspective it seems unlikely that a re-initiation complex ought to be present for the entirety of a burst, but undoubtedly it facilitates its duration and size in some manner related to TBP binding dynamics that are not yet fully realized.

Broadly speaking, most common elements of the core promoter have been associated with greater burst sizes. This is exemplified by Larsson et al.'s scRNA-seq study in mouse embryonic fibroblasts which proposed a linear model whereby the promoter elements, TATA box and initiator, had an additive effect on burst size but not burst frequency [40]. Suter et al. also found greater burst intensities when promoters contained multiple copies of the CCAAT box, a regulatory region known for recruiting the NF-Y transcription factor [77]. Supporting this, Ochiai et al. in their global kinetics estimates in mESCs found that the ChIP normalized read counts of many elements of the PIC and proteins associated with PolII elongation and initiation are positively correlated with burst size when bound to promoters [53]. These elements include: the TBP, the proximal pause release factor CDK9 and members of the super elongation complex AFF4, ELL2 and BRD4 [53]. From this accumulated evidence we can generally conclude that the core promoter facilitates longer and more productive bursts of transcription.

1.1.3 Faster Action at a Distance: Enhancer Dynamics and Burst Frequency

At the majority of endogenous genes, K_{on} values are much smaller than K_{off} such that dozens to hundreds of transcripts are produced in discrete moments followed by long periods of senescence [4, 40, 77]. From another perspective it can be said that transitions into the ON state of a promoter is less favoured than into its OFF state. What controls the transition between these states? More precisely, what is the genomic basis for the infrequent transition between inactive and active states? A cursory survey of the literature will identify enhancers as the main driver of burst frequency [5, 37, 25, 40, 99, 53]. Enhancers are typically defined in as distal regulatory sequences responsible for recruiting transcription factors (TFs) which help activate gene expression at their associated promoter [71]. The classical model has enhancers looping to the promoter via the interaction between the DNA bound transcription factor and a cofactor, such as Mediator or BRD4, which in turn associates with elements of the PIC [62]. This model has been challenged by several contradicting results, particularly in the way enhancer proximity and chromatin folding relate to bursting kinetics[93, 46]. Enhancer models are multiplicitous and beyond the scope of this review, rather here we present the evidence by which enhancers and burst frequency are related and how that might challenge our notions of gene expression.

Perhaps the most direct evidence of a functional link, Bartman et al. used a enhancer-promoter forced looping model at the β-globin locus in GATA-1 deficient mESCs [5]. By approximating burst frequency by counting the number of active alleles using smFISH in a population of cells, dubbed *burst fraction*, Bartman et al. found that when the primary enhancer for β-globin was anchored to the promoter burst fraction was elevated in isolation of burst size. This evidence of kinetic dissociation was shared by scRNAseq studies using hybrid MEFs. Enhancers harbouring single nucleotide polymorphisms have a tendency to have significantly different burst frequencies across alleles at their target promoters [40]. In the same study, a SOX2 enhancer deletion led to a loss in K_{on} values, but not burst size [40]. H3K27ac, a chromatin mark for active enhancers and promoters, can be correlated to burst frequency when found in enhancers [40, 53].

Some of the most revealing insights about the enhancer dynamics is a result live imaging studies in the fly embryo [25]. Using the MS2 system where the production of nascent transcripts can be measured in real time via fluorescent proteins targeting the first intron, the researchers examined the expression of the *yellow* reporter under the influence of different endogenous enhancers. What resulted was a differential number of spikes of transcription, but the average amplitude of each peak remained the same [25]. Enhancers were differentially active depending on the nuclei position in the embryo and their total output was mostly attributed to changing burst frequencies. Remarkably, two promoters were able to share an enhancer by evidence of an synchronized pattern of expression, disrupting the classical model of an enhancer entering close proximity to a single promoter at a time [25]. And while these results do not exclude the possibility of a contact model of transcription, the functional distance required for enhancers alleviate the thermodynamic burden of initiating transcription seems much more flexible than the assumed tens of nanometers which comprise a TF-Mediator-PolII complex [63]. Many imaging studies seem to reinforce such a activity at a distance: the 3D distance required between enhancers and promoters for an ON state

can often be in the range of hundreds of nanometers [12, 19]. How cells manage to bridge this physical gap will be discussed in the following section.

1.1.4 The Missing Middle: Mediator and its Kinetic Contexts

Mediator is a large multi-subunit protein complex that is functionally conserved in eukaryotes. It has a diverse and multifarious role in facilitating transcription [3, 63, 39]. Canonically it can be found in either two compositions: Mediator alone and when it is bound to the CDK8-kinase-module (CKM) to form the cyclin-dependent kinase Mediator complex . CKM is composed of four subunits and is functionally distinct from Mediator proper. It can reversibly bind with Mediator whose biochemical implications will be discussed later.

At its core, Mediator can be thought of as a nexus of transcriptional regulation whose primary means of control is to facilitate the formation of the PIC and the release of PolII from the promoter. Mediator, with the help of TFIID, positions the XPB to enable ATP-dependent promoter melting [63]. Concurrently, Mediator and TFIID captures PolII, orients the CDK-activating component of TFIIH (CAK) to enable the extensive phosphorylation of the c-terminal domain (CTD) of PolII [63]. Upon several rounds of phosphorylation at the serine-rich CTD, PolII is released from the crowded PIC and allowed to enter productive elongation[63].

At most genes PolII stops transcription at roughly 50 base-pairs from the transcription start site with a highly variable duration per loci [45]. This 'paused' polymerase state has been shown to be important for RNA stability and regulation at large [45]. There is some limited evidence that the Mediator-CKM interacts with elements from the super elongation complex (SEC) which facilitates pause release [3, 27]. This same complex effectively inhibits PIC formation by sequestering Mediator from PolII via competitive binding [63].

Mediator and CKM activity are not exclusive to the core promoter. Mediator has famously been implicated in many different functions at enhancers [63]. Med1 for instance, a member with many known transcription factor interactors, has often been hailed as a common ChIP-Seq signal for enhancer and super-enhancer activity, as is Med12 for the CKM [89, 48] . Many other Mediator sub-units have been shown to directly bind with many different transcription factors [63]. In addition to recruiting Mediator to enhancers, it has also been shown to change conformation when bound to transcription factors which helps stabilize the PIC [10, 87, 91]. Mediator has often been described as a functional bridge rather than a physical one; an intermediary protein linking the information dense, TF bound enhancers to promoters. This model is increasingly relevant given new chromatin capture methods which show a lack of a significant difference in chromatin interaction when the complex is depleted [21]. The exact mechanism by which Mediator forms such a bridge, one which does not require enhancer-promoters to be in close proximity, is unclear. However, new studies investigating its capacity to form protein condensates appears to be the most plausible model for enhancers operating at a distance.

Protein condensates can be thought of as enriched local concentrations of a population of proteins which selectively associates with other compositionally similar factors via weak, multivalent bonds [81, 86]. What can result is the formation of droplets formed by liquid-liquid phase separation (LLPS), this phenomenon is typically been associated with membraneless organelles, such as paraspeckles and nucleoli. These weak interactions are often attributed to the presence of intrinsically disordered regions (IDRs), although many other protein or

protein-RNA interactions can exhibit liquid-like behavior [81, 86]. Mediator, as does PolII, possess IDRs in many of their subunits and have been to shown to form droplets *in-vitro* and *in-vivo* in mESCs by co-localizing near genes activated by super-enhancers [7, 13, 19, 37]. Crucially, the size of these condensates appear to be between 200-300nm in diameter, bridging the gap previously observed in earlier enhancer-promoter imaging studies [13, 19, 37]. From this, a condensate model of expression has been proposed whereby condensates become repositories for PIC components, enabling the rapid assembly and loading of polymerases to form convoys. Bursts of transcription could occur via a cascade of assembling condensates, whereby co-factor droplets of dozens Mediator or BRD4 molecules co-localize with bound transcription factors or PolII at enhancers or the promoter respectively [33, 93]. In this model, Mediator not only stabilizes the PIC and performs post-translational modifications, but also delivers general transcription factors (GTFs), potentially accompanied by chromatin remodelers and pause release factors, by co-condensing with Mediator droplets. Many recent imaging studies seem corroborate some elements of this model, such as the observation of burst induction co-occurring when Mediator condensates localizes with the Sox2 super-enhancer and the promoter in mESCs [19] . Mediator condensates have been shown to co-localize with both TFs like Oct4 and estrogen receptor (ER), and with PolII in *in-vitro* [8, 30]. Acute Mediator abrogation has been shown to lead to a loss of PolII clusters previously co-localized with Mediator clusters *in-vivo* [35]. CTD phosphorylation has even been shown to enable condensate switching at PolII foci from Mediator to splicing factors [30]. Furthermore, it was shown that when Med25 was tethered to a Gal4 UAS in the developing fly embryo, more nuclei became active and produced more transcripts [37]. Larger Med25 labeled condensates, as opposed to smaller ones, seem to be able to activate transcription at two fluorescently labelled transgenes at once when all three co-localize briefly [37].

It is interesting to note that in all *in vivo* observations, Mediator condensates only reside at the enhancer-promoter foci for seconds before dissociating away [13, 19]. This is especially clear when, at Sox2, PolII labeled condensates appears to increase with fluorescence intensity over time while Mediator does not [19]. This observation falls in line with enhancer-promoter dynamics; enhancers have been shown to only interact with cognate promoters transiently, possibly being evicted by the accumulation of RNA [13, 57]. From these lines of evidence Mediator appears to initiate transcription by facilitating enhancer-TF activity at a distance, at least in its condensate form.

If we return to the polymerase convoy model, however, Mediator appears to do the opposite: a Med11 knockdown using siRNAs had no effect on the probability of the HIV transgene to enter an ON state [79]. Meanwhile the average number of PolII in each convoy reduced by a half, leading to a total reduction of transcription [79]. The researchers propose a model by which Mediator operates at fine timescales, motivating the re-initiation of PolII at an already active promoter. Med11 has been shown directly interact with TFIIH and facilitating its activity, therefore it follows that its depletion might inhibit CTD phosphorylation and PolII elongation [23]. While informative, this model stems from the observations of one HIV transgene and the depletion of a single sub-unit. It is very likely that non-canonical complexes may still form and thus phase-separate without Med11. Regardless, imaging analysis of condensates suggest is that Mediator may have a more context specific role in transcription, particularly dependant on the enhancers that surround their cognate promoters. How the depletion of the full Mediator complex might effect the kinetics of transcription, is still

lacking a comprehensive answer.

1.2 Project Aims

Given the sparsity of the data and the lack of reproduction, our understanding of how Mediator is implicated in bursting kinetics, particularly how they relate to enhancers, is still underdeveloped. In this study we aim to strengthen our understanding of Mediator and enhancers' role in regulating gene expression in the context of transcriptional kinetics. This will be approached in two parts. The first is to develop a mouse model whereby Mediator is to be depleted using an inducible degradation system. Upon the depletion of Mediator, single cell RNA sequencing can be performed whereby whole genome kinetics can be estimated in perturbed and wild-type cells using the discrete estimators described previously. Secondly, using public scRNAseq, ChIP-seq and H3K27ac chromatin capture data, globally re-contextualize enhancers' role in regulating bursting kinetics. This will be done by identifying enhancer-promoter (EP) contacts, enhancer hubs and super-enhancers in murine cells and comparing the kinetic parameters, namely burst size and burst frequency, with the rest of the genome. We also explore the impact topologically associated domains have on these kinetic parameters. To the author's knowledge, no whole-genome analysis of EPs or Mediator abrogation have been performed within the context of two-state bursting kinetics. This study aims to develop a model whereby the two are plausibly integrated.

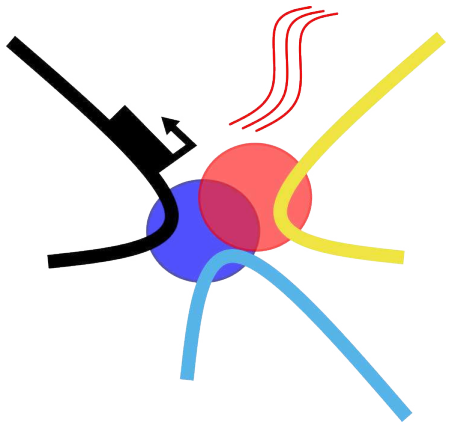

Figure 1.1: Illustration of enhancer-promoter locus in active transcription.

2

Design and Implementation of Auxin Inducible Degradation of Mediator

List of Figures

List of Tables

2.1 Introduction

One of the most essential tools for probing genes and their protein products is targeted knockouts using CRISPR-Cas9 or TALEN zinc-finger genome editing. By adding deleterious point mutations or truncations to open reading frames, researchers are able to measure the effects using transcriptomic or proteomic approaches. However, there are several caveats with gene knockouts which limit its potential utility. Firstly, removing a gene entirely often results in compensatory phenotypes which obfuscate interpretations. These only become apparent when compared to acute knockdown strategies such as small interfering RNAs (siRNAs) [4]. The exact mechanisms are often multiplicitous, but generally can be summarized as the modified expression of network-adjacent genes or the accumulation of secondary mutations in rapidly dividing cells such as yeast [9, 14]. Secondly, gene knockouts are, by definition, unfeasible when studying essential genes.

Fortunately, newer advances in acute depletion strategies circumvent these issues. Auxin-induced degradation (AID), a system discovered in Arabidopsis which has been transgenetically adapted to yeast, fly and mammalian cells, has quickly one of the most popular methods. AID has two main advantages over its RNA-based counterparts: AID has fewer off target effects and offers near-complete removal of the target protein [7, 12]. This efficiency, however, requires the added complexity of two CRISPR knock-ins: AID utilizes the ubiquitin E3 ligase TIR1 from the rice plant Oryza sativa and its associated degron tag to be integrate into the genome and the target protein respectively [7]. Upon the presence of the small signaling molecule auxin (IAA), OsTIR1 becomes active and ubiquinates the degron-carrying protein, allowing for the use of endogenous degradation pathways. Very recently, the system had been improved upon using a bump-and-hole strategy for increasing the affinity between a modified auxin (5-Ph-IAA) and a mutated OsTIR1 (F74G), resulting in a quicker degradation, lower basal activity and requiring less of the ligand with a proven activity in a variety of mouse cell types [17]. Given these advantages, using AID to target the Mediator complex in mouse embryonic fibroblasts (MEFs) became an obvious choice.

In order to degrade the complex in its entirety, two core sub-units were initially chosen: MED14; the main scaffolding protein which link the tail, middle and head modules contiguously, and MED17; the scaffolding protein of the head module [16]. Despite MED14's central role in the complex, MED17 was chosen as it did not have multiple transcript isoforms with differing C-termini that MED14 showed in mouse [2, 1]. MED17 depletion has also been shown to be a more efficient abrogator of Mediator activity as it disrupts the possibility of the assembly of the pre-initiation complex being derived from an isolated head module [15].

Given the available technologies as discussed, in addition to the methods developed by Pryzhkova et al., the initial strategy for degrading Mediator in mouse embryonic fibroblasts can be summarized as follows [8]. Two sequential CRISPR knockins using the homologous repair pathway, one targeting the H11 safe harbour locus for the OsTIR1 and the other the C-terminus of MED17 for integrating the AID degron tag. Once a OsTIR1 harbouring clonal population had been established, a GFP-AID vector can be randomly integrated and whose degradation can be measured using both cell flow cytometry and western blot in order to validate its activity. Unsurprisingly, many intervening challenges presented themselves when attempting to genetically modify this cell type, thus requiring new strategies and adaptations which will be discussed in the following sections.

2.2 Methods

2.2.1 CRISPR Repair DNA Template Construction of AID Components

The OsTIR1 vector contains the CMV constitutive promoter for strong expression, V5 antibody tag and a SV40 poly(A) signal. The OsTIR1(F74G) itself was taken from the AID2 system provided by Yesbolatova et al. (addgene no. 140536) [17]. For selection, hygromycin resistance powered by the strong promoter PGK was used (addgene no. 17446). Homology arms both roughly 800 and 500 base pairs long were used for the 5' and 3' end respectively. Homology arms were amplified from genomic DNA with one end ten basepairs or less from the cut site.

The AID degron vector used the mini-AID tag from Yesbolatova et al. with an in-frame 6HA antibody tag. Both neomycin and blasticidin resistance was used in separate vectors to allow for the selection on both alleles. Both constructs used the SV40 promoter for high expression and SV40 poly(A) signals for RNA stability (from addgene no. 103833). The blasticidin used was fused with GFP was cloned from pTracer ™ -CMV/Bsd (Invitrogen product no.V883-01). Homology arms both approximately 800 and 1000 basepairs long were again derived mouse genomic DNA. Both AID degron and OsTIR1 containing vectors were assembled onto a pUC18 backbone.

All vectors were constructed using Nebuilder HIFI as per the manufacturer's protocol.

2.2.2 CRSIPR Knock-In of AID Components

Cell Culture

Large T-antigen immortalized C57/CAST MEFs were donated from Dr.Sauvageau. They were cultured in DMEM supplemented with MEM non-essential amino acids and 10% FBS and incubated at 37°C at 5% CO2.

OsTIR1 CRISPR Integration

Transfections were first attempted using Cas9 conjugated to streptavidin reverse transcribed from PCS2-Cas9-mSA (addgene no. 227473) [5]. Reverse transcription was performed using the mMESSAGE mMACHINE SP6 Transcription Kit (ThermoFisher) following the manufacturer's protocol. The resulting RNA was purified using RNeasy MinElute kit (Qiagen). Synthetic guide RNAs targeting the H11 locus were ordered from Synthego. Guide sequences were taken from Pryzhkova et al. [8].

Transfections were initially performed using the Amaxa MEF 2 Nucleofector Kit (Lonza) and the T20 program on the D Nucleofector Unit following the manufacturer's protocol. Increasing molar concentrations of transfection mixes as used were summarized in *table 1.* Nucleofected cells were allowed to recover 35mm plates for 24-48 hours depending on confluency before switching to selection media. Lipofection was also performed to compare with nucleofection using TransIT-LT1 (Mirus). Cells were plated with increasing cell densities as shown in Table 2.1 on a 35mm plate and incubated for two days. Cells were then incubated with TransIT liposomes for 24 hours before changing the media for selection media. Selection media against OsTIR1 construct included 1 mg/ml of hygromycin B (Wisent). Cells selected for a week until isolated colonies appeared at which point they were expanded and

Name	sgRNA:DNA:Cas9 molar Ratio	Total Mass (ug)	Transfection Technique	Additional Condition
N1x	4:1:5	2	Nucleofection	NA
N2x	4:1:5	4	Nucleofection	NA
N3x	4:1:5	6	Nucleofection	NA
N4x	4:1:5	8	Nucleofection	NA
N2xT	4:2:5	4.3	Nucleofection	NA
N2xC	4:1:5	4	Nucleofection	Double cell concentration
L0.8	4:2:3	2.25	Lipofection	0.8E5 Seed density
L1.6	4:2:3	2.25	Lipofection	1.6E5 Seed density
L3.2	4:2:3	2.25	Lipofection	3.2E5 Seed density

Table 2.1: Reagent Mixtures for OsTIR1 Knock-in

whole cell extracts were taken for western blot. Cells were sorted into four 96 well plates via FACS supplemented with 20% conditioned media. Clonal cell lines were screened again for OsTIR1 using western blot from confluent 12 well plates.

Western Blot

Whole cell extracts were lysed with RIPA buffer and supplemented with protease inhibitors. Prior to loading cell lysates were sheared using a 0.2mm needle or treated with DNAse (roughly 40 000 units). Extracts were quantified using the Pierce BCA Protein Assay Kit (ThermoFisher) and equal amounts were fractionated on a 8% SDS-PAGE gels.

Antibodies

All of the following antibodies were used for western blots unless stated otherwise. Mouse anti-V5 (Sigma V8012) was at a 1:2000 dilution ($5x10^{-5}$mg/ml in 5% milk. Mouse anti-HA (Santa Cruz Biotechnology, sc-7392) was diluted at 1:1000 in 5% BSA. Anti-GFP (Roche/Sigma, 11814460001) was diluted to 1:1000 in 5% milk. A rabbit hMED17 antibody (Bio-Rad, VPA00478) was diluted to 1:1000 in 5% BSA. Anti-tubulin (Sigma, T5168) was used at a 1:50000 dilution in 5% milk. Donkey anti-mouse IRDye 680 (Li-cor Biosc) at a 1:5000 dilution was used with all mouse antibodies. Donkey anti-rabbit IRDye 800CW (Li-Cor Biosc.) was used with the anti-hMED17 antibody in a 1:5000 dilution.

Immonuflorescence Imaging

Lipofected and nucleofected cells were seeded at a density of $2x10^5$ cells in a six well plate and incubated overnight. Cells were fixed with 3.7% formaldehyde solution for 15 minutes. Fixed cells were blocked with 5% w/v BSA in PBS for an hour at room temperature. Anti-V5 antibody (Sigma V8012) was diluted to 0.002 mg/ml in 1% BSA and 0.3% Triton X-100. Fixed cells were incubated with this solution overnight at 4C. Cells were washed with PBS and incubated for an hour with a 1:300 dilution of Cy5 AffiniPure Donkey Anti-Mouse IgG

Cassette/Locus	5' Sequence	3' Sequence	Size
MED17-mAID 5'	tgcccatgattcagtcccat	ttctccttctcgagccatcc	1.6kb
MED17-mAID 3'	ctcatgctggagttcttcgc	agacatctctcccttgtgct	230bp
MED17-exon 12	ggcctgctcttgaaatgctt	cagcagtgcacatgagtctg	435bp
MED17-exon 12	tcgtaacggacctgagagtg	cccgtgaaagattggagcag	338bp

Table 2.2: Oligonucleotide sequences for screening MED17-mAID CRISPR insertions.

(Cedarlane). Cells were washed and briefly incubated with DAPI before imaging.

AID Degron Integration at MED17

The CRISPR gRNA guide sequences used were AGGTCCGTTACGAACTATGG and AGGTCTTTGCACTGGTTACG. They were integrated into px458 by using Nebuilder HIFI from single stranded oligos as per the manufacturer's protocol. 1ug of each gRNA-Cas9 plasmid and template DNA was transfected using nucleofection as previously described. Cells were selected using 350ug/ml of G148 sulfate for seven days before screening by western blot. Populations were screened prior to single cell sorting using western blot and an anti-HA antibody. Positive colonies were identified by the presence of an HA band in line with a hMED17 band at roughly 95kDa. Cells were sorted into two 96 well plates supplemented with 50% conditioned media. A second vector harboring the mAID tag targeting MED17 but with blasticidin-GFP was transfected into heterozygous MED17-mAID$^{+/-}$ cells. The transfection method was identical to the first except selection was performed using 0.8ug/ml blasticidin for two weeks before bulk sorting for positive GFP expression using flow cytometry. Flow cytometry visualizations were generated by Cytoflow [3].

Genotyping CRISPR Integrations

All genomic DNA was extracted using the phenol-chloroform method. Oligos targeting the 5' and 3' end of either insertion are summarized in Table 2.2. Seven oligo pairs targeting the OsTIR1 H11 locus and several different combinations all failed to produce consistent PCR fragments.

OsTIR1 piggyBac Random Integration

OsTIR1 was randomly integrated using the piggyBac transposase system (Hera BioLabs). A vector from Yesbolatova et al. (addgene no. 296196) had the mAID-GFP component swapped for OsTIR1(F74G) (from addgene no.140536) using Nebuilder HIFI. A mCherry-T2A-BlastR fragment was also added downstream simultaneously from an in-house vector. Cells were transfected using 5ug of the OsTIR1-mCherry vector and 1ug of the transposase expression vector (Hera BioLabs, no. SPB-D10). Transfection was performed using the Amaxa MEF2 Nucleofection kit with the T20 program as per the manufacturer's recommendations. After three days post nucleofection, cells were treated to 0.8ug/ml blasticidin until colonies began to appear.

mAID-GFP piggyBac Random Integration

OsTIR1 activity was measured with a randomly integrated mAID-GFP cassette from Yesbolatova et al. (addgene no. 296196) [17]. The hygromycin resistance marker was swapped for neomycin using Nebuilder HIFI. The cassette was randomly integrated into OsTIR1 positive

cell lines using 1ug of a piggyBac transposase vector (Hera BioLabs, no. SPB-D10) and 5ug of the expression vector. The DNA was transfected using nucleofection as previously described. Cells were put under selection for ten days using G418 sulfate (350ug/ml).

Auxin Degradation Timecourses

Auxin timecourses were performed by incubating transfected cells in 10uM of 5-phenyl-indole-3-acetic (5-Ph-IAA) for up to 48hrs. GFP expression was measured at multiple time-points using both western blot and flow cytometry. One dimensional mixture modeling and visuals were generated using Cytoflow [3]. Auxin timecourses targeting MED17 was performed as described above. Degradation was measured using western blot following the disappearance of the HA tag and verified using a hMED17 antibody.

2.3 Results

2.3.1 Design and Construction of CRISPR Repair Templates

CRISPR knock-ins by way of the homology-directed repair (HDR) pathway has become a standard for genome engineering in mammalian cells and has been used by many groups to implement AID [8, 17]. This requires the use of long homology arms, typically between 500 to 1000 nucleotides, to flank either end of the repair cassette. These cassettes are summarized in Figure 2.1. Since the difficulty of the insertion scales with the length of the cassette, a simplified design for each was implemented with the minimal requirements for successful cloning.

To the author's knowledge, no C-terminal mAID tag vector was available at the time, nor was there a OsTIR1(F74G) vector with an appropriate antibiotic resistance for our chosen cell line. Given this, both vectors needed to be constructed from many constituent parts. Due to the complexity and the number of separate fragments in the design of these vectors (6 fragments each ranging from 160bp to 2.3kb), the cloning method of choice was NEBuilder® HiFi DNA Assembly Cloning Kit (New England BioLabs, E2621). HiFi offers many advantages over traditional cloning methods, most notably the ability to anneal multiple PCR fragments into a single reaction without the need of ligases or restriction sites. Although the method boasts high efficiency, experiences within the lab have found that assemblies can increase in difficulty with the number fragments being used and the diversity of fragment lengths. In order to alleviate these challenges, several rounds of successive cloning with smaller batches of fragments may be done. This was the case for the MED17-AID degron vector; after initially failing to assemble it in a single six fragment reaction, we opted for a sequential four (pUC18-5'-NeoR-3') and three fragment construction (Figure 2.1). What resulted was that all isolated colonies contained a vector that missed either the 5' or 3' homology arm. Rather than attempting to restart the process and isolate the intended sequence, we decided continue to add the missing fragments as originally planned but with a final two-fragment reaction to add the missing homology arm. In contrast to the three required of the MED17-mAID vector, the OsTIR1 vector was assembled in one reaction. All additional reactions based off these vectors were trivial, only requiring a maximum of three fragments to be assembled in a given reaction.

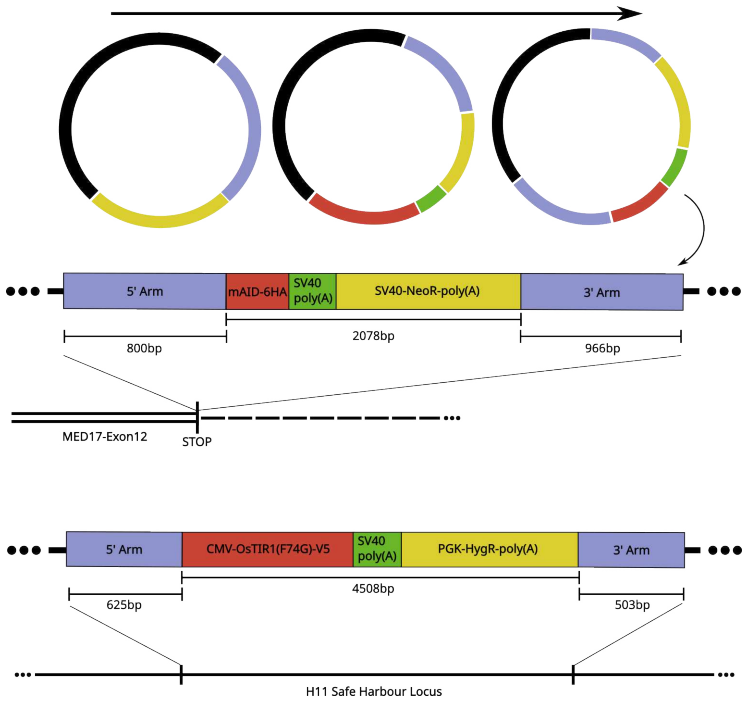

Figure 2.1: MED17-mAID three reaction HiFi vector cloning and OsTIR1 Cassette integration.

2.3.2 OsTIR1 Integration

First attempts at generating OsTIR1 expressing populations entirely relied on using synthetic guides and Cas9-mSA mRNA from Gu et al. [5]. After amplifying the repair cassettes with biotinylated primers, Cas9-mSA is expected to bind to the cassette and allow for higher integration efficiencies [5]. And while expression was observed in western blots for all the transfection mixes and conditions, it never became strong enough where we could be confident in using the population for a second round of transfection, assuming a low abundance of positive knock-ins (Figure 2.2, A). Following this, immunoflourescence imaging targeting the V5 antibody tag was performed to assess the probability of obtaining a positive clone after single-cell cloning. After subtracting background signal, it was found that roughly 1-2% of cells highly expressed the V5 tag (Figure 2.2, B). Following this, 384 cells were sorted and 140 clonal cell lines were screened, out of which five were kept for high OsTIR1 expression. Unfortunately a positive PCR confirming the genotype remained elusive. After testing seven different primer pairs, and sending one seemingly positive PCR fragment for full length sequencing, we concluded that the OsTIR1 gene must be randomly integrated and not in the intended locus.

2.3.3 OsTIR1 Degradation Activity

After isolating several clonal cell lines expressing OsTIR1, we decided test if the expression was strong enough to degrade a fluorescent marker. Using a randomly integrated mAID-GFP expression cassette, partial degradation was observed between 24 to 48 hours (Figure 2.2). With our experiences within the lab, a partial degradation is often indicative of a much

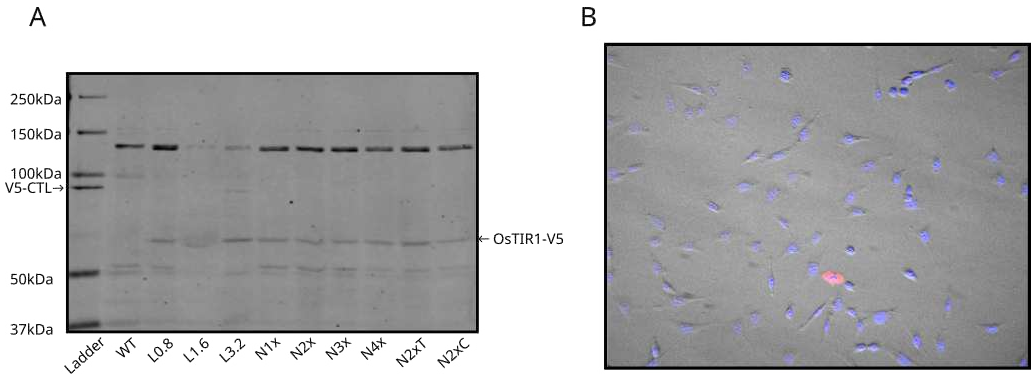

Figure 2.2: OsTIR1 Cassette Integration in MEFs. (A) Western Blot of transfected polyclonal populations for evaluating OsTIR1-V5 expression. Transfection mixes for the populations are listed in Table 2.1. (B) Immunofluorescence image showing OsTIR1 expression. Red channel is V5-Cy5 flourescence, blue is DAPI staining. The population is L3.2, listed in Table 2.1.

more efficient degradation of an endogenous protein. This is likely because the high dosage of the GFP which the OsTIR1 cannot keep up with. Given this result, we decided to pursue tagging MED17 with the degron.

2.3.4 Dual Allele Tagging MED17 with mAID

Tagging MED17 in MEFs was a persistent challenge. Initially, tagging MED17 relied on the same methods as described previously when introducing OsTIR1. Three synthetic guides nearing the C-terminus of MED17 transfected with equimolar ratios in combination with Cas9 mRNA and linearized DNA template. This method, however, yielded cell populations who were resistant to the antibiotic (neomycin) but appeared to have none of the HA tag in western blots. After multiple failed attempts, it became necessary to develop new methods in tandem.

The first of which was to use a dual guide, Cas9-nickase approach [11]. This was done in order to increase the specificity of Cas9 as it requires two active enzymes to cleave at a site for integration. A D10A point mutation was added to the PCS2-Cas9-mSA vector using Nebuilder HIFI. Synthetic guides from Synthego were used in a PAM out orientation roughly 40bp surrounding the MED17 c-terminus. However, after nucleofections were performed the polyclonal populations failed to produce a positive signal in the subsequent western blots.

Another approach was to use a standard Cas9-gRNA plasmid (px458, addgene no. 48138) co-transfected with either one or two linear repair cassettes with differing antibiotic markers. New RNA guides were used which were much farther from the integration site but yielded much higher cutting and specificity scores according to the CRISPR prediction software CCTop [13]. Using this method, two heterozygous MED17-mAID$^{+/-}$ clones were isolated after screening 60 cell lines via western blot (Figure 2.4, A-B). Although the mAID was strongly expressed, the OsTIR1 expression decreased significantly (Figure 2.4, B).

Using this MED17-mAID $^{+/-}$ cell line, a preliminary auxin timecourse was performed to assess if the new quantity of OsTIR1 might be enough to degrade Med17 in a fully tagged cell line. Using a 10uM of 5-Ph-IAA, a 40-60% depletion of the HA tag was observed after

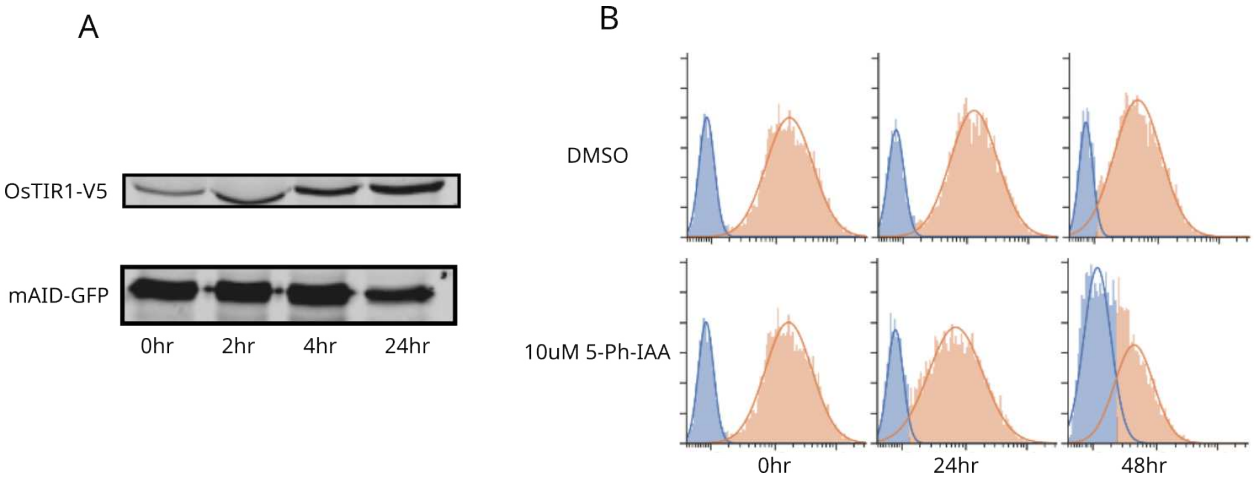

Figure 2.3: AID degradation assays using transposase integrated mAID-GFP cassette. (A) Western blot of an 10uM auxin timecourse following GFP expression. (B) Flow cytometry of GFP amplitude during a 48 hour auxin timecourse. A one dimensional mixture model of the gaussian distributions was fitted to the polyclonal population.

48 hours (Figure 2.5, B) . Endogenous MED17 levels also decreased by roughly 20% to 30% during the same timespan (Figure 2.5, C).

In order to improve the degradation efficiency, a twofold approach was attempted: randomly integrate an OsTIR1 cassette using a piggyBac transposase to increase expression in the tagged cell line, and secondly, tag the second allele with new cassette containing the mAID degron and a GFP-blasticidin conjugated resistance gene. The transposase mediated approach, adapted from Yesbolatova et al. showed very high efficiency when testing an OsTIR1$^{+/-}$ cell line (Figure 2.3). Cell line 3.5 (MED17-mAID$^{+/-}$), which showed some residual OsTIR1 expression (Figure 2.5, B), was first re-transfected in the same manner when tagging the first allele, but was selected using blasticidin and then bulk sorted for GFP expression using flow cytometry. The remaining population was then transfected with the transposase adapted cassette (OsTIR1-V5-mCherry-T2A-BlastR) (Figure 2.5, A). A preliminary western blot indicated that a roughly ten-fold increase in OsTIR1 was found, but relatively little change MED17-mAID-6HA (Figure 2.5, D). Regardless, cells were single-cell sorted once again using FACS for both fluorescent tags into six 96 well plates. Despite the strong expression of the both fluorescent tags, all 170 clones recovered from sorting still displayed a wild-type allele when screening with PCR (Figure 2.5, C,E).

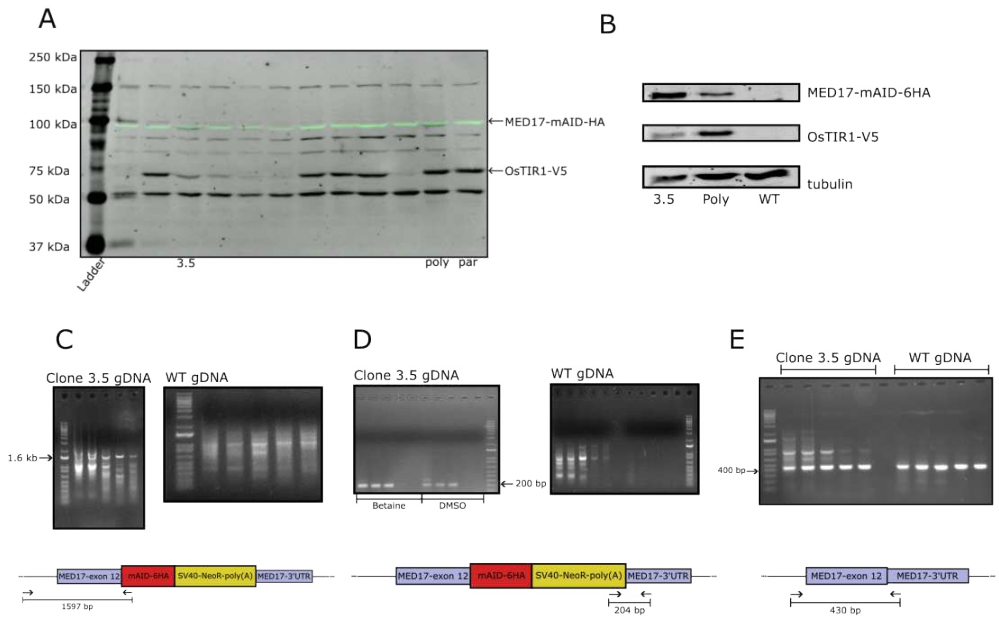

Figure 2.4: Evidence of a tagged MED17 allele with mAID-6HA degron. (A) Western blot of single cell clones derived from a tagged polyclonal population. HA and V5 antibodies are represented in gray. The hMED17 antibody is represented in green. *Poly* and *par* refer to extracts from the polyclonal population and the parental cell line prior to the transfection.(B) Western blot of clone 3.5 with *Poly* as the polyclonal population and *WT* as wild-type MEFs. (C-D) Gradient PCR genotyping of clone 3.5 and wild type gDNA. Wells are loaded with increasing annealing temperature from left to right, starting from 42 to 62°C. (E) PCR illustrating the remaining wild-type allele in the clone.

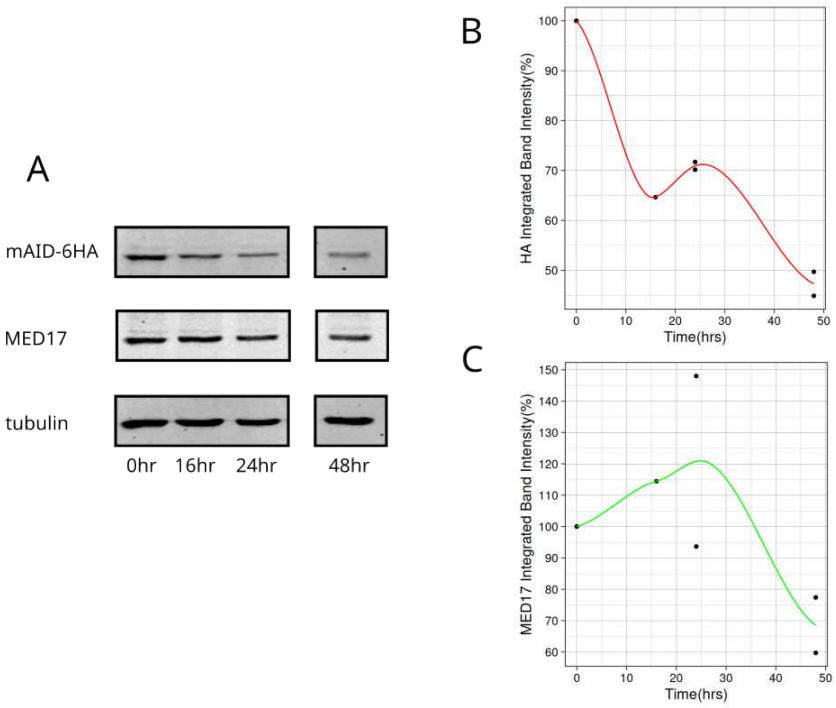

Figure 2.5: Auxin Timecourse of clone 3.5 (OsTIR1$^{+/-}$, MED17-mAID-6HA$^{+/-}$). (A) Western blot using 10 uM 5-Ph-IAA for the tagged Med17-mAID-6HA and the endogenous Med17. 48hr sample was loaded into a non-adjacent well. (B-C) Integrated intensity of bands normalized to tubulin relative to initial timepoint for two replicate auxin timecourses.

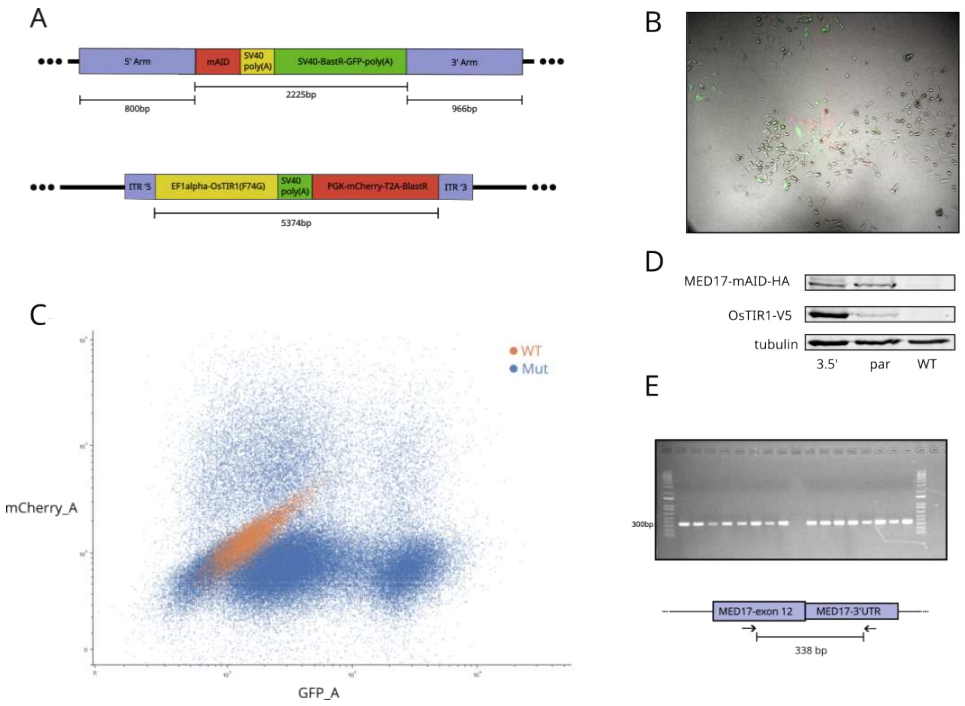

Figure 2.6: Dual tagging MED17 with mAID-6HA for Homozygous Expression and Enhancing OsTIR1 Expression. (A) Genetic constructs for tagging a second MED17 allele and randomly integrating an OsTIR1 expression cassette. (B) Microscopy image representing mCherry and GFP expression in transfected MEFs. (C) Flow Cytometry of transfected cells prior to single cell sorting. Data points in orange are wild-type, untransfected cells. Data in blue are double transfected with the constructs from (A). (D) Western blot of cell line 3.5' which has been transfected with constructs in (A). *Par* is the parental cell line 3.5 (Figure 2.4) and *WT* is a wild-type cell line.(E) Genotyping PCR representative of all isolated clonal lines from FACS sorting. A 330bp band is indicative of the presence of a wild-type allele.

2.4 Discussion

Initial attempts at generating a OsTIR1$^{+/-}$ MEF cell line, although partially unsuccessful, was informative in determining the format in which the genetic material was to be delivered. The use of mRNA CRISPR-Cas9 and synthetic gRNAs produced underwhelming results using either lipid-based transfection reagents or nucleofection (Figure 2.2). The mRNA approach was based on a protocol in the context of embryo micro-injections; it may be that for cultured cells, Cas9-mSA mRNA may simply be not stable enough to endure the transfection. Later DNA-plasmid based transfections were able to give positive results where mRNA could not.

Both tagging MED17 with the AID degron and integrating OsTIR1 in the H11 locus in mouse embryonic fibroblasts presented itself as a significant challenge. MEFs, while very easy to maintain when immortalized with the large-T antigen, appeared very reluctant to incorporating targeted CRISPR mediated knock-ins. MEFs within the lab's experience seemed particularly amenable to random integration; cell populations post-transfection would often form colonies after a week selection on antibiotics but show no apparent integration via western-blot or genotyping PCR. Even when MED17 was successfully tagged, only two out of sixty isolated colonies were positive (an integration efficiency of 3.33%). Large amounts of random integration was also apparent when attempting to tag the second Med17 allele. High GFP expressing cells were first bulk sorted, re-transfected with OsTIR1-mCherry and single cell sorted for both fluorescent tags, effectively performing positive selection twice all while maintaining a lower dosage of the selection antibiotic (Figure 2.5).

There are several immediate approaches that could potentially remove the possibility of random integration. Firstly, in all genetic constructs the antibiotic resistance was driven by a strong promoter so as to avoid the potential weak output from the endogenous one. In the case where the endogenous promoter is enough to drive selection, a in-frame protein cleavage site (such as T2A) followed by the resistance gene could potentially eliminate most random integration events. Secondly, in the event that a second allele needs to tagged sequentially, re-designing the cassette so that it contained a different antibody tag (such as FLAG) would be beneficial for screening at the population level prior to sorting.

It also remains a possibility that MED17 is simply not amenable to tagging with this particular tag in this cell type. It may be possible that a mAID-6HA tag in the C-terminus is a large enough growth disadvantage to facilitate selection, especially when both alleles are modified. To the author's knowledge, no C-terminal Med17 tag has been performed in mammalian cells. Given this, it could be prudent to target MED14 for degradation using this system despite the uncertainty of an N-terminal tag. Since starting the project, a MED14-mAID tagged cell line has been established in HCT116 cells [6]. The authors make use of a micro-homology mediated repair pathway which has been shown to have higher integration efficiency than homology-directed repair in some cell types [10]. Such alternatives to the HDR pathway could be worth testing given the known low integration efficiency in MEFs.

2.5 Conclusion

After multiple attempts and iterations, a MED17-mAID$^{+/-}$ OsTIR1$^{+/-}$ mouse embryonic fibroblast cell line had been isolated for the use of inducing the degradation of the Mediator

complex. A partial degradation was achieved when incubating with 10uM 5-Ph-IAA after 24hrs.

Bibliography

[1] Med14 mediator complex subunit 14 [Mus musculus (house mouse)] - Gene - NCBI, . URL https://www.ncbi.nlm.nih.gov/gene/26896.

[2] MED17 Polyclonal Antibody (11505-1-AP), . URL https://www.thermofisher.com/antibody/product/MED17-Antibody-Polyclonal/11505-1-AP.

[3] Brian Teague. Cytoflow: A Python Toolbox for Flow Cytometry. *bioRxiv*, page 2022.07.22.501078, January 2022. doi: 10.1101/2022.07.22.501078. URL http://biorxiv.org/content/early/2022/07/23/2022.07.22.501078.abstract.

[4] Mohamed A. El-Brolosy and Didier Y. R. Stainier. Genetic compensation: A phenomenon in search of mechanisms. *PLOS Genetics*, 13(7):e1006780, July 2017. ISSN 1553-7404. doi: 10.1371/journal.pgen.1006780. URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006780. Publisher: Public Library of Science.

[5] Bin Gu, Eszter Posfai, and Janet Rossant. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol*, 36 (7):632–637, August 2018. ISSN 1546-1696. doi: 10.1038/nbt.4166. URL http://www.nature.com/articles/nbt.4166. Number: 7 Publisher: Nature Publishing Group.

[6] Christoph Neumayr, Vanja Haberle, Leonid Serebreni, Katharina Karner, Oliver Hendy, Ann Boija, Jonathan E. Henninger, Charles H. Li, Karel Stejskal, Gen Lin, Katharina Bergauer, Michaela Pagani, Martina Rath, Karl Mechtler, Cosmas D. Arnold, and Alexander Stark. Differential cofactor dependencies define distinct types of human enhancers. *Nature*, 606(7913):406–413, June 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04779-x. URL http://www.nature.com/articles/s41586-022-04779-x. Number: 7913 Publisher: Nature Publishing Group.

[7] Kohei Nishimura, Tatsuo Fukagawa, Haruhiko Takisawa, Tatsuo Kakimoto, and Masato Kanemaki. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods*, 6(12):917–922, December 2009. ISSN 1548-7105. doi: 10.1038/nmeth.1401. URL https://www.nature.com/articles/nmeth.1401. Number: 12 Publisher: Nature Publishing Group.

[8] Marina V. Pryzhkova, Michelle J. Xu, and Philip W. Jordan. Adaptation of the AID system for stem cell and transgenic mouse research. *Stem Cell Research*, 49:102078, December 2020. ISSN 1873-5061. doi: 10.1016/j.scr.2020.102078. URL https://www.sciencedirect.com/science/article/pii/S1873506120303792.

[9] Andrea Rossi, Zacharias Kontarakis, Claudia Gerri, Hendrik Nolte, Soraya Hölper, Marcus Krüger, and Didier Y. R. Stainier. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, 524(7564):230–233, August 2015. ISSN 1476-4687. doi: 10.1038/nature14580. URL https://www.nature.com/articles/nature14580. Number: 7564 Publisher: Nature Publishing Group.

[10] Tetsushi Sakuma, Shota Nakade, Yuto Sakane, Ken-Ichi T. Suzuki, and Takashi Yamamoto. MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc*, 11(1):118–133, January 2016. ISSN 1750-2799. doi: 10.1038/nprot.2015.140. URL https://www.nature.com/articles/nprot.2015.140. Publisher: Nature Publishing Group.

[11] Bin Shen, Wensheng Zhang, Jun Zhang, Jiankui Zhou, Jianying Wang, Li Chen, Lu Wang, Alex Hodgkins, Vivek Iyer, Xingxu Huang, and William C. Skarnes. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods*, 11(4):399–402, April 2014. ISSN 1548-7105. doi: 10.1038/nmeth.2857. URL https://www.nature.com/articles/nmeth.2857. Number: 4 Publisher: Nature Publishing Group.

[12] Ameet Shetty, Natalia I. Reim, and Fred Winston. Auxin-inducible degron system for depletion of proteins in S. cerevisiae. *Curr Protoc Mol Biol*, 128(1):e104, September 2019. ISSN 1934-3639. doi: 10.1002/cpmb.104. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6741457/.

[13] Manuel Stemmer, Thomas Thumberger, Maria del Sol Keyer, Joachim Wittbrodt, and Juan L. Mateo. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLOS ONE*, 10(4):e0124633, April 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0124633. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0124633. Publisher: Public Library of Science.

[14] Xinchen Teng, Margaret Dayhoff-Brannigan, Wen-Chih Cheng, Catherine E. Gilbert, Cierra N. Sing, Nicola L. Diny, Sarah J. Wheelan, Maitreya J. Dunham, Jef D. Boeke, Fernando J. Pineda, and J. Marie Hardwick. Genome-wide Consequences of Deleting Any Single Gene. *Molecular Cell*, 52(4):485–494, November 2013. ISSN 1097-2765. doi: 10.1016/j.molcel.2013.09.026. URL https://www.cell.com/molecular-cell/abstract/S1097-2765(13)00748-X. Publisher: Elsevier.

[15] Jason P Tourigny, Kenny Schumacher, Moustafa M Saleh, Didier Devys, and Gabriel E Zentner. Architectural Mediator subunits are differentially essential for global transcription in Saccharomyces cerevisiae. *Genetics*, 217(3):iyaa042, March 2021. ISSN 1943-2631. doi: 10.1093/genetics/iyaa042. URL https://doi.org/10.1093/genetics/iyaa042.

[16] Kuang-Lei Tsai, Chieri Tomomori-Sato, Shigeo Sato, Ronald C. Conaway, Joan W. Conaway, and Francisco J. Asturias. Subunit architecture and functional modular rearrangements of the transcriptional Mediator complex. *Cell*, 157(6):1430–1444, June 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2014.05.015. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4104964/.

[17] Aisha Yesbolatova, Yuichiro Saito, Naomi Kitamoto, Hatsune Makino-Itou, Rieko Ajima, Risako Nakano, Hirofumi Nakaoka, Kosuke Fukui, Kanae Gamo, Yusuke Tominari, Haruki Takeuchi, Yumiko Saga, Ken-ichiro Hayashi, and Masato T. Kanemaki. The auxin-inducible degron 2 technology provides sharp degradation control in yeast,

mammalian cells, and mice. *Nat Commun*, 11(1):5701, November 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-19532-z. URL https://www.nature.com/articles/s41467-020-19532-z. Number: 1 Publisher: Nature Publishing Group.

Chapter 3 Preamble

In the previous chapter, we focused on Mediator's role in transcritpional kinetics and attempted to establish an acute degradation system in mouse embryonic fibroblasts. In the next chapter we will turn enhancers, chromatin conformation and their relationship with bursting kinetics. One of Mediator's core functions is to bridge the relationship between enhancers and promoters. After examining the roles of enhancers in isolation, we will speculate on the impact of Mediator depletion and its effects on enhancers in the general discussion.

3

Enhancer-Promoter Contacts and their Relationship with Transcriptional Kinetics

List of Figures

List of Tables

3.1 Introduction

Enhancers are one of the most well defined metazoan regulatory sequences, capable of positively regulating transcription at promoters several hundreds of kilobases away [24]. While many enhancer-promoter relationships have been characterized, the exact mechanism by which these elements confer transcriptional activity remains elusive.

Transcription has been shown to occur in stochastic bursts in metazoans [4, 22, 30, 23, 1, 35]. Using a two-state model of transcription, combined with real-time imaging or single cell RNA sequencing technologies, many research groups have independently found how genetic components modulate the timing and the size of these bursts. Burst frequency has been well understood to be influenced by enhancers [2, 15, 8, 14]. Direct observation of this effect however has been limited to a few isolated genes. When a global approach has been attempted, enhancer activity is inferred via correlating burst frequency with single nucleotide polymorphism (SNP) density or H3K27ac enrichment in enhancers [8, 15]. A global approach has yet to be attempted that incorporates the 3D conformation of chromatin, specifically one which captures enhancers looping into proximity with their cognate promoters. In this study, we use public chromatin capture and kinetics data to concretize enhancers' role in bursting kinetics. Here, we re-confirm the positive effect enhancers have on burst frequency at all enhancer-promoter (EP) contacts. Additionally, there is a modest increase in burst size in the presence of an EP. Multi-way connections between promoters and multiple enhancers show that enhancers affect both bursting parameters differently, with burst frequency saturating at 2-3 enhancers and burst size increasing linearly. Additionally, using public topologically associated domain (TAD) annotations we show that EP linkages that cross TAD boundaries do not impact kinetics. Lastly, we examine how super-enhancers (SEs) effect the kinetics of proximal, overlapping and looped genes. Chromatin looping from the SE to the promoter appears sufficient but not necessary for regulating burst size at proximal genes. For burst frequency however, SEs appear to require chromatin contact to induce faster kinetics.

3.2 Results

3.2.1 Enhancer Promoter Contacts Positively Regulate Burst Kinetics

Chromatin confirmation capture via HiChIP using H3K27ac offers an efficient method for identifying enhancer-promoter contacts. Here we use processed contact maps from Yang et al., which identifies single point linkages between two H3K27ac peaks [34]. By combining EP pairs with scRNAseq data from Larsson et al. , two-state transcriptional kinetics data was re-calculated for roughly nine thousand murine genes and annotated for EP contacts. Promoters that are connected to at least one enhancer with this method (n = 609), have a significantly greater burst frequency than unconnected ones (Figure 3.1). When normalized with an average decay mRNA rate of 4.8hrs, EPs on average had waiting times between bursts that were 2.7 hours shorter than non-associated promoters. Burst size also increased significantly, but to a lesser degree with a Wilcoxon effect-size estimate of 0.0452 versus 0.0979 for burst frequency (Figure 3.1). EPs on average produced bursts with 0.66 more transcripts than unconnected ones. Intrinsic noise, which is the cell-to-cell variation in transcripts derived from inter-allele bursting, significantly decreases upon the the presence

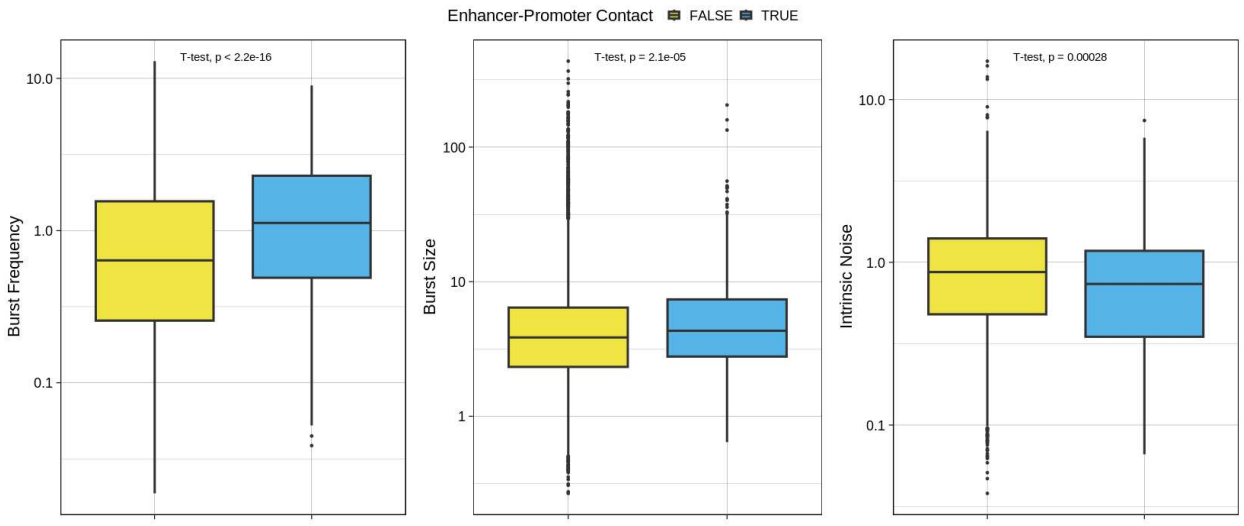

Figure 3.1: H3K27ac Hi-ChIP defined enhancer-promoter contacts induce significant changes in the kinetics of expression. Kinetics of enhancer-connected promoters (n = 609) versus unconnected. Burst frequency and burst size both increase significantly when a promoter is in contact with at least one enhancer ($p < 0.001$). Intrinsic noise decreases significantly upon enhancer contact.

of an connected enhancer (t.test, $p = 0.00028$, Figure 3.1). This follows previous observations of the strong inverse relationship between intrinsic noise and burst frequency [21].

3.2.2 Multivalent EP Connections Reveals a Dual-Role in Transcription

In metazoans, it is far from uncommon to have many enhancers interacting with one promoter [18, 5, 26]. Using H3K27ac Hi-ChIP data, we were able to identify these multivalent interactions and classify promoters on the number of interacting enhancers. From the 609 connected promoters, we observed a wide range of unique enhancer connections, reaching a maximum of 19 at the gene *Fosl2*. When we performed a gene ontology enrichment on the set of genes with eight or more connections (n=29), we found most terms were associated with cell communication, adherence and signaling (Supplemental Table 3.3). This falls in line with prior observations that enhancer hubs are implicated in tissue specificity and pleiotropy [19, 27]. We then examined the shape of the distributions that the kinetic parameters take over EP counts. Interestingly, burst frequency appears asymptotic at 2 to 3 enhancer contacts (Figure 3.2). When attempting to fit the data to linear model, a Ramsey RESET test determined that burst frequency does not vary linearly with contact frequency ($p < 0.0001$). Instead, a second degree polynomial was added to a bootstrapped OLS model which accommodated the apparent saturation (table 3.1).

Meanwhile, burst size increases linearly with EP contacts (Figure 3.2). A simple boostrapped linear model was performed on burst size as a function of contact frequency (β =1.043, 95% CI [1.021,1.066], Table 3.1). The linear combination of the two loess curves is apparent when examining the growth normalized read counts over increasing EPs (Figure

3.2). The increasing step-wise pattern of the medians alone suggest that enhancers act in a combinatorial manner and that HiChIP interactions are sensitive enough to distinguish transcriptional output at a single enhancer a time. If we attempt to model gene expression as function of enhancer contacts, a nearly 17% increase in transcription is expected per enhancer ($\beta = 1.169$, CI [1.13, 1.20], table 3.1).

The discrepancy in magnitude and shape of the fitted curves suggest enhancers facilitate transcription via multiple mechanisms. Namely, enhancers appear to act redundantly when initiating an active transcriptional state but act additively when maintaining one.

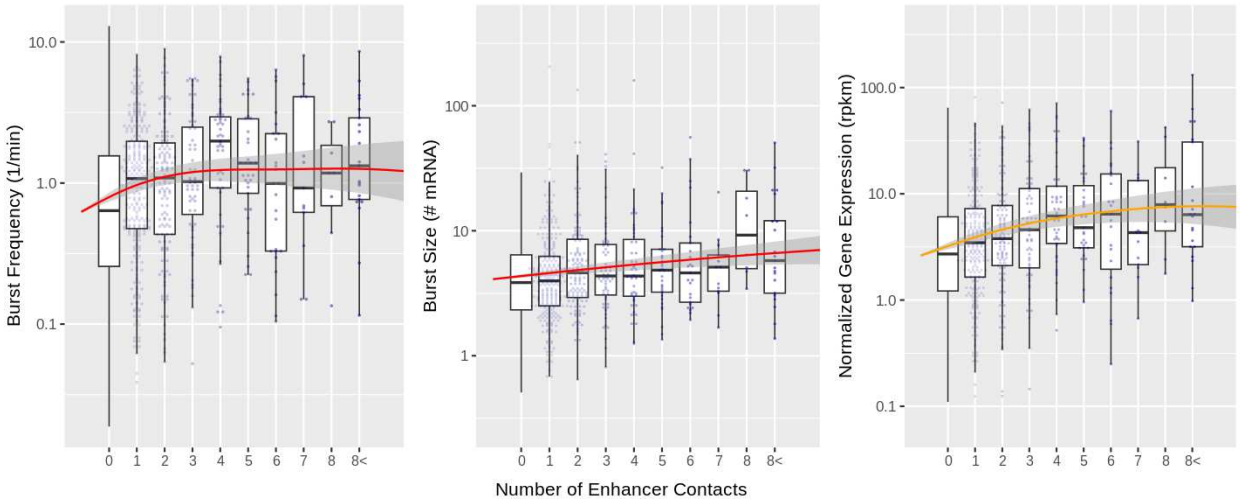

Figure 3.2: H3K27ac HiChIP defined EPs with increasing multivalency at the promoter. Red and orange traces are loess curves.

	Predictors	*log10(Estimate)*	*95% log10(CI)*	*P*
Burst Frequency (min^{-1})	(Intercept)	-0.13	-0.142 - -0.120	**<0.001**
	Freq	3.50	2.56 - 4.42	**<0.001**
	$Freq^2$	-1.54	-2.44 - -0.594	**0.002**
	R^2	0.008		
Burst Size (# mRNAs)	(Intercept)	0.647	0.639 - 0.656	**<0.001**
	Freq	0.018	0.010 - 0.028	**<0.001**
	R^2	0.003		
Expression (mRNAs)	(Intercept)	0.420	0.407 -0.433	**<0.001**
	Freq	0.068	0.054 - 0.082	**<0.001**
	R^2	0.011		
	Observations	7606		

Table 3.1: Mean coefficients of a boostrapped linear models of gene transcription as a function of Enhancer Contact Frequency. *Freq* refers the number of interacting enhancers with the promoter. Frequency values above eight are binned. All linear models were significant with $p < 0.001$

3.2.3 Super-Enhancer Activity Depends On Distance, Connectivity and Position in Gene Body

Super-Enhancers (SE) are large genomic regions composed of multiple enhancers responsible for cell identity genes [32]. Hallmarks of a super enhancer is the occupancy of Med1, a subunit of the transcriptional co-factor Mediator. ROSE (Rank Ordering of Super Enhancers) has long been considered the gold standard of SE identifying algorithms [32]. Using the original ChIP-seq dataset from Whyte et al., SEs were defined in mESCs and the kinetics were compared in proximal and non-proximal genes using data from Larsson et al. (Figure 3.3) [32, 15]. Proximal SE genes (n = 293) showed a greater burst frequency (t.test, p = 0.0013). Meanwhile, burst size at SEs remain unchanged (t.test, p=0.43). Here, SEs appear help initiate a state of permissive transcription but do little to help maintain it, at least at these loci which contain an over-abundance of Med1. Having performed SE calling in mESCs, we then decided to return to the MEF data and perform SE calling with H3K27ac ChIPseq reads. After excluding promoter regions, 422 super-enhancers were identified. Using the HiChIP data, 71 promoters were found to be in contact at least once with these SEs (Figure 3.4, A). Interestingly, the general relationship with EPs was re-established when examining connected super-enhancers: both burst size and frequency were upregulated ($p < 0.001$, $p = 0.0024$, respectively; Figure 3.4, B). Overlapping, non-linked genes however, appear to be regulated in the same manner to linked ones, suggesting that chromatin looping is dispensable for SE activity at these loci ($p = 0.85$, $p = 0.75$ for burst frequency and burst size respectively). Genes that are unlinked and proximal to the SE (TSS less than 50kb from the middle of the SE) appear benefit from an increase in burst size ($p = 0.0015$) while burst frequency does not improve ($p = 0.19$). If burst frequency is related to enhancer contact frequency as some groups suggest, then the HiChIP data combined with SEs reflects this [36]. However, it appears that SE activity on burst size is less dependant on contact frequency and is able to induce larger bursts at distance greater than what is captured by HiChIP.

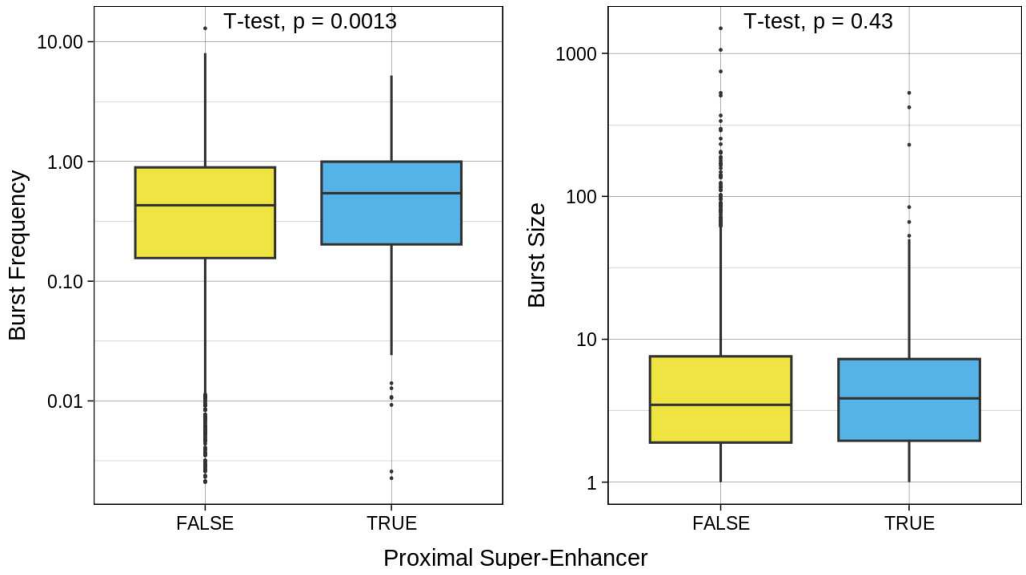

Figure 3.3: Transcriptional kinetics of genes proximal to super-enhancers in mESCs. Consitituent enhancers are first defined by the co-occupancy of Oct4, Sox2 and Nanog. Clustered enhancers with high Med1 occupancy were stitched together and those passing a rank threshold defined by ROSE were labeled as a super-enhancer. Proximal genes are those with transcription start sites at least 50kb from the stitched SE.

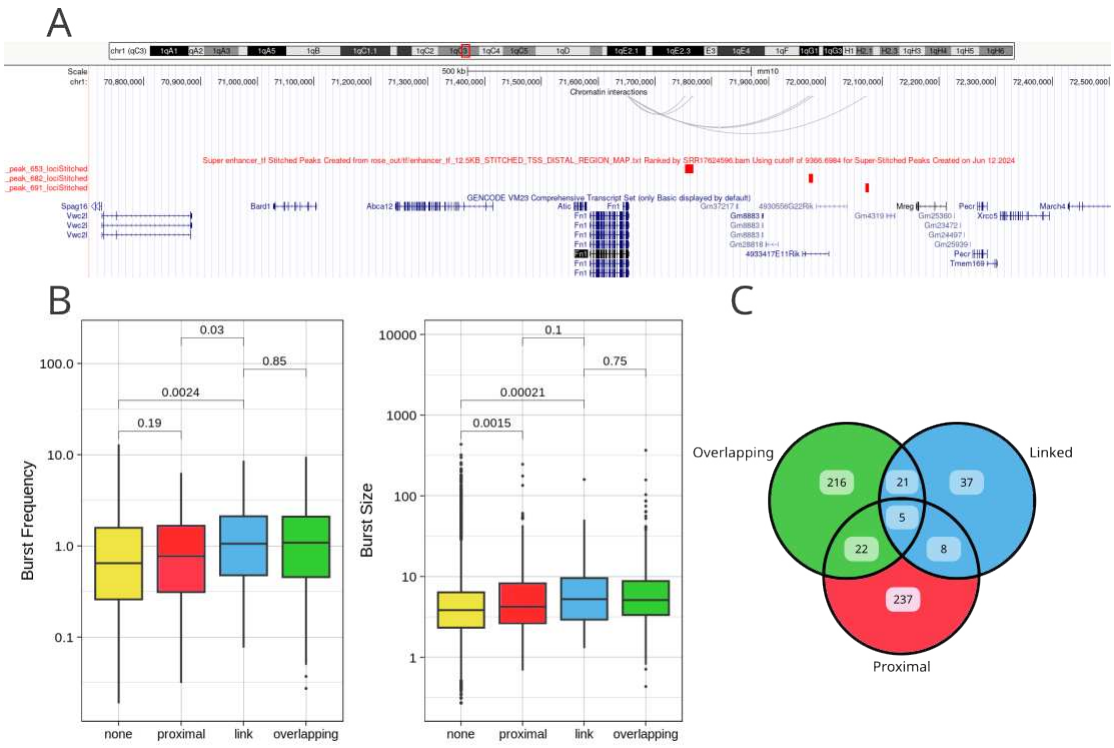

Figure 3.4: MEF super-enhancer kinetics and their relationship with gene promoters. (A) UCSC gene track of HiChIP interactions of gene Fn1 with annotated super-enhancers. (B) Kinetic activity of SE associated genes. Linked genes are those identified to have at least one HiChIP linkage with an SE (as in A). Proximal genes are those with a TSS at least 50kb from a SE and without a HiChIP. Overlapping genes are those without a HIChIP connection but intersect with an SE. All p-values are derived from a Wilcoxon rank sum test.(C) Relationship categories and gene counts of the groups in B. Colours denote the set operations of each group, e.g "Proximal" is the difference of the union of "Overlapping" and "Linked" with proximal genes.

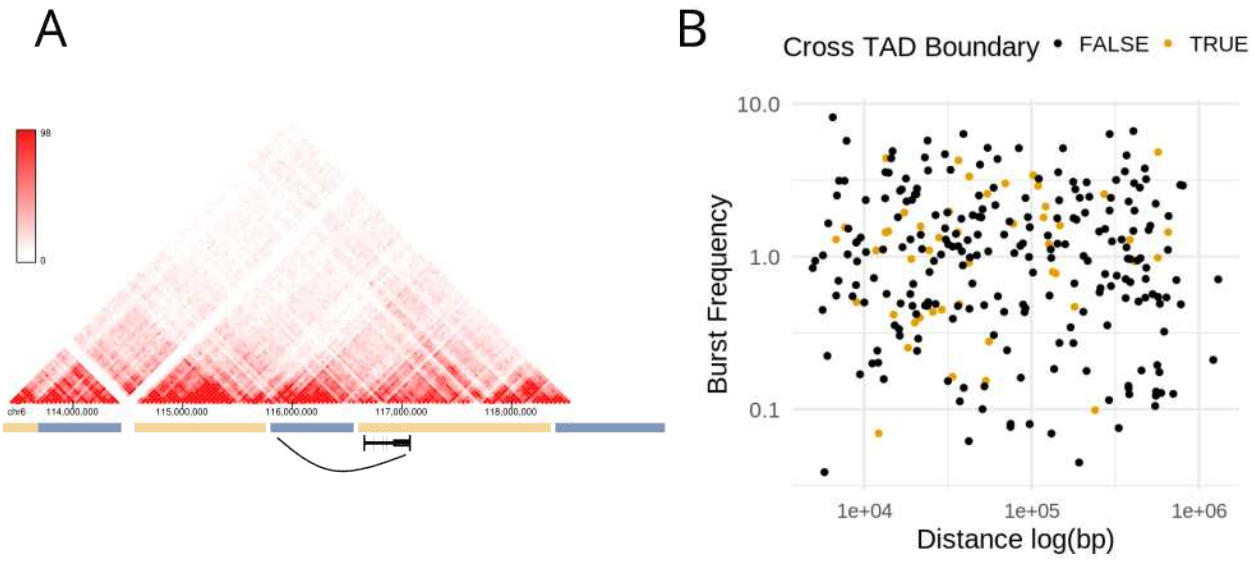

Figure 3.5: The presence of topologically associated boundaries does not alter kinetics for single EP linkages. (A). Example of an EP crossing an TAD boundary. MEF TAD annotations are derived from HI-C from Li et al. (B) Burst Frequency of single linked EPs versus genomic distance between enhancer and promoter.

3.2.4 Cross Topologically Associating Domain Linkages Are Kinetically Invariant

Topologically associating domains (or TADs) are typically defined as large, megabase sized regions, which have high self-interacting frequencies which taper off beyond its boundaries [11]. A recent study has promoted model where TAD boundaries act as insulators facilitating enhancer specificity, particularly by inhibiting burst frequency [36]. Others have found TAD disruptions often leads to very little difference in gene expression [10]. Here we investigate how TAD boundaries might influence transcriptional kinetics by taking promoters that are linked to only one enhancer and intersecting the regions between linked enhancers and promoters with TAD annotations in MEFs from Li et al. [17]. Single EP linkages that cross TAD boundaries (n=79) show no difference in kinetic output (Figure 3.5 B). Furthermore, genomic distance between enhancers and promoters, which ranges from kilobases to megabases in size, do not alter burst frequency (Figure 3.5, B).

3.3 Methods

3.3.1 SmartSeq3 scRNA-seq

Raw Smart-Seq3 reads for C57_Bl6/CAST_EiJ hybrid MEFs were downloaded from ArrayExpress (Accession no. E-MTAB-10148). A mixed allele C57/CAST genome was generated using SNPsplit (https://github.com/FelixKrueger/SNPsplit) by N-masking a mm10 (GRCm38) assembly at verified CAST SNP positions. The reads were then mapped to this genome by STAR with the parameters "–limitSjdbInsertNsj 2000000 –clip3pAdapterMMp 0.1 0.1 –clip3pAdapterSeq CTGTCTCTTATACACATCT CTGTCTCTTATACACATCT".

UMIs were collapsed to a hamming distance of 1 and expression profiles were generated for both 5' ends and combined full length data was performed using zUMIs (v2.9.7b/c). Alignment files, barcodes and gene expression data were merged from two plates using in-house scripts. Reads were assigned to their allele of origin by an R script provided by the Sandberg Lab (https://github.com/sandberg-lab/Smart-seq3/tree/master/allele_level_expression). Genes which had expression (reads) but no assigned allelic reads were marked as missing data. Finally, transcriptional bursting parameters were estimated using python scripts found at https://github.com/sandberg-lab/txburst provided by the Sandberg lab. Full pipeline is illustrated in supplemental Figure 3.6.

3.3.2 Estimating Intrinsic Transcriptional Noise

Low abundance cells were first removed if the library size was less than 5000 UMIs. UMI counts per gene were then removed if they differed significantly between alleles to avoid SNP induced or monoallelic expression (poisson test $p < 0.001$). Library size normalization was then performed such that:

$$U_i' = (\frac{U_i}{C_j}) * n$$

Where U is the UMI count per gene i, C is the library size per cell j and n is an arbitrary scaling factor. Normalized allele specific UMI counts were then used to calculate intrinsic noise [21]:

$$\eta_{int}^2 = \frac{\langle (U_{c57}' - U_{cast}')^2 \rangle}{2\langle U_{c57}' \rangle \langle U_{cast}' \rangle}$$

Where U_{c57} and U_{cast} represent the normalized UMI counts per gene each at allele. Angle brackets denote the mean across cells. Total noise was calculated as follows [21]:

$$\eta_{tot}^2 = \frac{\langle U_{c57}^2 + U_{cast}^2 \rangle - 2\langle U_{c57} \rangle \langle U_{cast} \rangle}{2\langle U_{c57} \rangle \langle U_{cast} \rangle}$$

At genes where intrinsic noise was greater than the total noise the gene was discarded from analysis.

3.3.3 H3K27ac HiChIP

Processed HiChIP data was downloaded from Yang et al. (GSE193079) [34]. Left and right defined anchors were separately intersected with promoter annotations provided by the Eukaryotic Promoter Database. The 60 bp annotations were expanded with a 300bp window. Self-loops and promoter-promoter contacts were removed. Genes with multiple promoter annotations had the maximally connected annotation retained. Only unique promoter-enhancer annotations were kept and joined with the C57 allele kinetics data. MEF TAD annotations were downloaded from GSE167579 and intersecting HiChIP linkages were assigned using bedtools and the parameter "-f 0.95 -v" to force 95% of the linkage to be crossing a TAD boundary.

3.3.4 ChIPseq and Identifying Super Enhancers

Med1, Nanog, Sox2 and Oct4 ChIPseq reads were downloaded from GSE44288 [32]. Files were aligned to mm10 genome using bowtie2 with the parameter "-p 10" followed by quality filtering using samtools view and the parameters "-b -h -F 3844 -q 10". Peaks were called using MACS2 with a p-value cutoff of 1e-5. Super-Enhancers were defined using ROSE with Med1 aligned reads and constituent enhancers defined by Nanog, Sox2 and Oct4 co-peaks. SEs were associated with genes by being at least 12kb proximal to the TSS. mESC kinetics were downloaded from https://github.com/sandberg-lab/txburst. Bl6xCAST H3K27ac raw files were downloaded from GSE193727 and aligned to an N-masked genome using bowtie2 with the same parameters as above. Peaks were called again with MACS2 with a p-value cutoff of 1e-9. ROSE was performed using H3K27ac peaks and aligned reads from SRR17624596. Identified SEs were then intersected with HiChIP anchors and mapped to those with interactions at a promoter.

3.3.5 Statistical Analysis

All methods and visualizations were performed in R. Linear modeling was performed using ordinary least squares using the base "lm" function. Bootstrapping was performed using the "boot" package and function of the same name with 5000 iterations. Gene Ontology enrichment was performed using the "gProfiler" package in R. Scripts, data and figures can be found at: https://github.com/Benjamin-R-Clark/EPs-and-Kinetics.

3.4 Discussion

Proximity-ligation techniques such as HiChIP provides an interesting glimpse into the dynamics of chromatin folding. Although only 609 EPs were identified here, other more permissive techniques such as 5C have identified that active TSS comes in contact with an average of four distal regions [26]. H3K27ac loops here represent a much more restricted set of criteria, namely the presence of a region of open chromatin flanked by the histone mark [34]. These linkages identify a subset of genes that are much more active than the average (Figure 3.1). As shown in Figure 3.2, expression appears to increase step-wise with the number of enhancer contacts, suggesting that most enhancers operate in the classical additive manner [31, 19]. What expression data does not show is the disjointed response in the underlying kinetic parameters. The rapid saturation of burst frequency and the slow progression of burst size suggest enhancers operate with multiple functions. Reflecting on burst frequency, enhancers appear to modulate the capacity of the promoter to enter a permissive transcriptional state, a fact well known in the literature [15, 8, 2, 14]. What is novel is how redundant enhancers appear after three connections, only requiring one or two to induce permissivity (Figure 3.2). Elsewhere enhancers are able to maintain transcription by inducing larger bursts sizes via a linear accumulation of enhancer contacts (Figure 3.2). Transcription factor and co-factor concentrations around the promoter have been shown to induce greater burst sizes, therefore it would follow that enhancer multiplicity could recruit and maintain larger concentrations of these factors [20]. Studies have shown that longer TF dwell times can induce longer bursts and that these dwell times themselves can mark the temporal boundaries of bursts in yeast [28, 6]. Protein condensates have also been implicated wherein enhancer clusters primarily act in facilitating capturing these local concentrations

of factors [13, 33, 12]. How the biophysics of enhancers might do so remains undeveloped.

An interesting paralogue with enhancer hubs are canonically defined super-enhancers. SEs currently are the focal point for imaging studies in mammalian cells for studying endogenous bursting dynamics [7]. SEs are often made up of many constituent enhancers, suggesting they might share some kinetic properties with what was shown in Figure 3.1. Here we show Med1 enriched SEs induce greater bursting frequencies in isolation of burst size in neighboring genes in mESCs (Figure 3.3). When we use chromatin conformation to identify SE-interacting genes in MEFs again, we see different relationship emerging. Namely, SEs induce a similar kinetic effect to regular enhancers in that they impact both parameters (Figure 3.4, B). This reflects the importance of chromatin conformation on enhancers and their regulatory capacity. When looking at the individual parameters, HiChIP linkages appear sufficient but not necessary for inducing larger burst sizes (Figure 3.4 B). Again, if enhancers help maintain transcription by sequestering local concentrations of TFs, SEs might be able to do so at a distance that is larger than the capture radius of HiChIP, but less than what 50kb of genomic distance provides. Further analysis will be needed to determine at what genomic distance do promoters lose this benefit in burst size. For burst frequency, close proximity in 3D space appears necessary for non-overlapping genes to induce more frequent transitions into an ON state (Figure 3.4, B). This somewhat reflects our understanding of what constitutes burst frequency, namely, that it is a function of contact frequency between enhancer and promoter [36]. The curious dynamics of overlapping genes, however, still remains to be fully understood. The majority of genes that were captured here do not have a TSS proximal to an SE but rather the SE resides more than 50kb downstream (Figure 3.4, C). Despite this, these SEs can induce greater burst frequencies without contacting the promoter in a manner that can be captured by HiChIP. This result somewhat contradicts our previous conclusion and raises an interesting question, how do SEs regulate burst frequency without looping to the promoter of their overlapping gene? It is possible that these SEs are recruiting factors which interact directly with the moving body of polymerases and mRNA, circumventing the need to loop frequently with the promoter in their resident gene. Otherwise, it may be that these regions are not acting like enhancers at all and ROSE is merely identifying segments of highly active genes with enrichment in H3K27ac. Future work using TF ChIPseq data could help narrow the definition of super-enhancers or help determine what factors are at play in the gene body.

Topologically associating domains have largely been considered hallmarks of eukaryotic chromatin landscape, facilitating enhancer selection [29]. TADs have been shown to be highly conserved between species and in many isolated examples these boundaries are essential for development [19]. In other contexts TAD boundaries appear to be much more transient and permeable [9, 3]. Live-cell imaging suggests that TADs may persist in a semi-extruded state most of the time, with a full TAD loop lasting from anywhere between ten to 30 minutes [9]. Here we find linkages that cross these barriers in MEFs, reinforcing the latter claim. Furthermore these linkages do not alter kinetics, despite evidence by Zuin et al. that TADs inhibit contact frequency of enhancers and burst frequency [36]. Zuin et al. also found using their enhancer shuffling strategy that EP distance has a real effect on kinetics and expression within the context of a single TAD [36]. Rather, here we expect that for HiChIP defined loops a much more stable looping mechanism must be at play which enables shorter distances across 3D space rather than the rapid and transient ones captured by micro-C. Furthermore,

if TADs mostly exist in a semi-extruded state, EPs may be able to cross these perceived boundaries. Some models suggest that the mobile dynamics of cohesin complexes may even facilitate EP formation [25, 33]. Regardless, evidence here and elsewhere suggests we should think more broadly about TADs as dynamic structures rather than discrete containers.

3.5 Conclusions

Enhancer-promoter contacts greatly influence the kinetics of transcription. These contacts primarily induce more frequent switching into an active promoter state. This effect does not scale with increasing EPs but saturates after the second enhancer contact. Burst size however increases linearly with EP contacts, suggesting enhancers have multiple mechanisms which emerge through cooperativity. Super-enhancers appear to share similar properties to regular enhancers in that they modulate both parameters. They are also able to extend activity to unlinked genes, suggesting chromatin loops are dispensable at proximal and overlapping genes. Transcriptional kinetics data, in combination with enhancer silencing screens could help reveal precisely how different enhancer cooperativity models impact the particular mechanisms which compose transcription itself.

3.6 Supplemental Figures

ID	Description	GeneRatio	BgRatio	p.adjust	qvalue
GO:0007044	cell-substrate junction assembly	5/29	98/23012	0.000119	6.46e-05
GO:0150115	cell-substrate junction organization	5/29	105/23012	0.000119	6.46e-05
GO:0031589	cell-substrate adhesion	7/29	368/23012	0.000119	6.46e-05
GO:0007160	cell-matrix adhesion	6/29	235/23012	0.000119	6.46e-05
GO:0002062	chondrocyte differentiation	5/29	123/23012	0.000119	6.46e-05
GO:0071560	cellular response to transforming growth factor beta stimulus	6/29	248/23012	0.000124	6.69e-05
GO:0071559	response to transforming growth factor beta	6/29	252/23012	0.000124	6.69e-05
GO:0030099	myeloid cell differentiation	7/29	476/23012	0.000221	0.00012
GO:0034329	cell junction assembly	7/29	479/23012	0.000221	0.00012
GO:0042593	glucose homeostasis	6/29	300/23012	0.000221	0.00012
GO:0033500	carbohydrate homeostasis	6/29	301/23012	0.000221	0.00012
GO:0048732	gland development	7/29	489/23012	0.000225	0.000122
GO:0048146	positive regulation of fibroblast proliferation	4/29	79/23012	0.000305	0.000165
GO:0007369	gastrulation	5/29	186/23012	0.000328	0.000178
GO:0030325	adrenal gland development	3/29	25/23012	0.000375	0.000203
GO:0042060	wound healing	6/29	370/23012	0.000499	0.00027
GO:0007229	integrin-mediated signaling pathway	4/29	101/23012	0.000618	0.000334
GO:0051216	cartilage development	5/29	223/23012	0.000618	0.000334
GO:0048638	regulation of developmental growth	6/29	404/23012	0.000663	0.000359
GO:0035987	endodermal cell differentiation	3/29	33/23012	0.000663	0.000359

Table 3.2: Top 20 enriched Gene Ontology terms for genes with eight or more enhancer contacts.

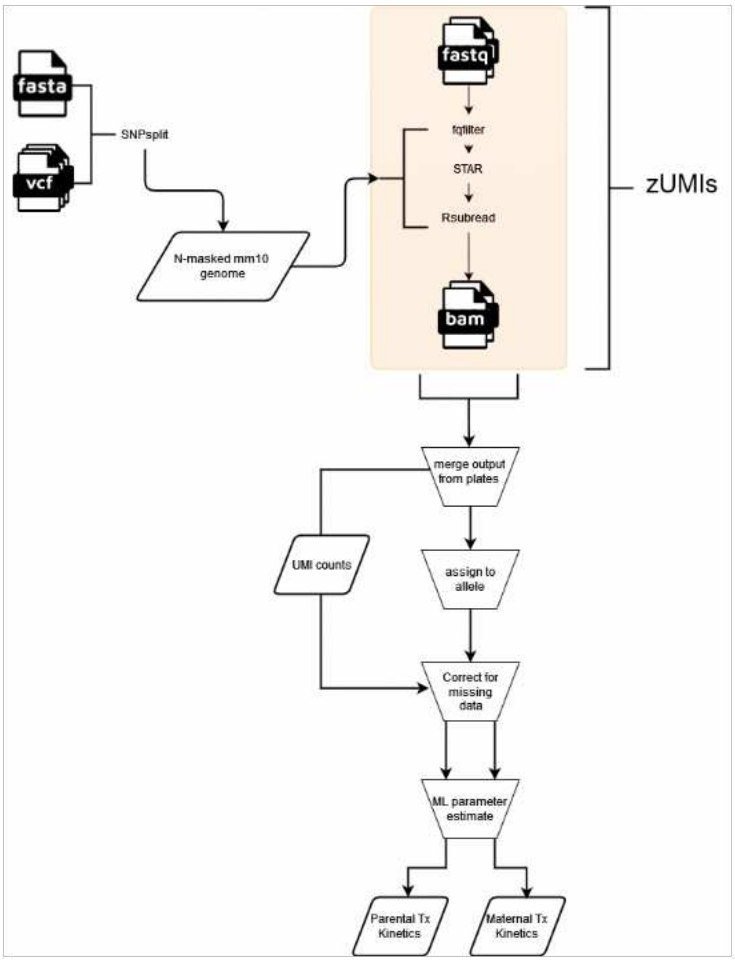

Figure 3.6: Transcriptional Kinetics Pipeline from SmartSeq3 scRNAseq Data

Bibliography

[1] Keren Bahar Halpern, Sivan Tanami, Shanie Landen, Michal Chapal, Liran Szlak, Anat Hutzler, Anna Nizhberg, and Shalev Itzkovitz. Bursty Gene Expression in the Intact Mammalian Liver. *Molecular Cell*, 58(1):147–156, April 2015. ISSN 1097-2765. doi: 10.1016/j.molcel.2015.01.027. URL https://www.sciencedirect.com/science/article/pii/S1097276515000507.

[2] Caroline R. Bartman, Sarah C. Hsu, Chris C. S. Hsiung, Arjun Raj, and Gerd A. Blobel. Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell*, 62(2):237–247, April 2016. ISSN 1097-2765. doi: 10.1016/j.molcel.2016.03.007. URL https://www.sciencedirect.com/science/article/pii/S1097276516001854.

[3] Shreeta Chakraborty, Nina Kopitchinski, Ariel Eraso, Parirokh Awasthi, Raj Chari, and Pedro P Rocha. High affinity enhancer-promoter interactions can bypass CTCF/cohesin-mediated insulation and contribute to phenotypic robustness. preprint, Developmental Biology, January 2022. URL http://biorxiv.org/lookup/doi/10.1101/2021.12.30.474562.

[4] Jonathan R. Chubb, Tatjana Trcek, Shailesh M. Shenoy, and Robert H. Singer. Transcriptional Pulsing of a Developmental Gene. *Current Biology*, 16(10):1018–1025, May 2006. ISSN 0960-9822. doi: 10.1016/j.cub.2006.03.092. URL https://www.sciencedirect.com/science/article/pii/S0960982206014266.

[5] Wouter de Laat and Denis Duboule. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature*, 502(7472):499–506, October 2013. ISSN 1476-4687. doi: 10.1038/nature12753. URL https://www.nature.com/articles/nature12753. Publisher: Nature Publishing Group.

[6] Benjamin T Donovan, Anh Huynh, David A Ball, Heta P Patel, Michael G Poirier, Daniel R Larson, Matthew L Ferguson, and Tineke L Lenstra. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *The EMBO Journal*, 38(12):e100809, June 2019. ISSN 0261-4189. doi: 10.15252/embj.2018100809. URL https://www.embopress.org/doi/full/10.15252/embj.2018100809. Publisher: John Wiley & Sons, Ltd.

[7] Manyu Du, Simon Hendrik Stitzinger, Jan-Hendrik Spille, Won-Ki Cho, Choongman Lee, Mohammed Hijaz, Andrea Quintana, and Ibrahim I. Cissé. Direct observation of a condensate effect on super-enhancer controlled gene bursting. *Cell*, 0(0), January 2024. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2023.12.005. URL https://www.cell.com/cell/abstract/S0092-8674(23)01337-5. Publisher: Elsevier.

[8] Takashi Fukaya, Bomyi Lim, and Michael Levine. Enhancer Control of Transcriptional Bursting. *Cell*, 166(2):358–368, July 2016. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2016.05.025. URL https://www.cell.com/cell/abstract/S0092-8674(16)30573-6. Publisher: Elsevier.

[9] Michele Gabriele, Hugo B. Brandão, Simon Grosse-Holz, Asmita Jha, Gina M. Dailey, Claudia Cattoglio, Tsung-Han S. Hsieh, Leonid Mirny, Christoph Zechner, and Anders S. Hansen. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, 376(6592):496–501, April 2022. doi: 10.1126/science.abn6583. URL https://www-science-org.proxy3.library.mcgill.ca/doi/10.1126/science.abn6583. Publisher: American Association for the Advancement of Science.

[10] Yad Ghavi-Helm, Aleksander Jankowski, Sascha Meiers, Rebecca R. Viales, Jan O. Korbel, and Eileen E.M. Furlong. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat Genet*, 51(8):1272–1282, August 2019. ISSN 1061-4036. doi: 10.1038/s41588-019-0462-3. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7116017/.

[11] Antonina Hafner and Alistair Boettiger. The spatial organization of transcriptional control. *Nat Rev Genet*, 24(1):53–68, January 2023. ISSN 1471-0064. doi: 10.1038/s41576-022-00526-0. URL https://www.nature.com/articles/s41576-022-00526-0. Publisher: Nature Publishing Group.

[12] Kota Hamamoto and Takashi Fukaya. Molecular architecture of enhancer–promoter interaction. *Current Opinion in Cell Biology*, 74:62–70, February 2022. ISSN 0955-0674. doi: 10.1016/j.ceb.2022.01.003. URL https://www.sciencedirect.com/science/article/pii/S0955067422000047.

[13] Denes Hnisz, Krishna Shrinivas, Richard A. Young, Arup K. Chakraborty, and Phillip A. Sharp. A Phase Separation Model for Transcriptional Control. *Cell*, 169(1):13–23, March 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.02.007. URL https://www.sciencedirect.com/science/article/pii/S009286741730185X.

[14] Koji Kawasaki and Takashi Fukaya. Functional coordination between transcription factor clustering and gene activity. *Molecular Cell*, 83(10):1605–1622.e9, May 2023. ISSN 1097-2765. doi: 10.1016/j.molcel.2023.04.018. URL https://www.sciencedirect.com/science/article/pii/S1097276523002897.

[15] Anton J. M. Larsson, Per Johnsson, Michael Hagemann-Jensen, Leonard Hartmanis, Omid R. Faridani, Björn Reinius, Åsa Segerstolpe, Chloe M. Rivera, Bing Ren, and Rickard Sandberg. Genomic encoding of transcriptional burst kinetics. *Nature*, 565(7738):251–254, January 2019. ISSN 1476-4687. doi: 10.1038/s41586-018-0836-1. URL https://www.nature.com/articles/s41586-018-0836-1. Number: 7738 Publisher: Nature Publishing Group.

[16] Anton J. M. Larsson, Christoph Ziegenhain, Michael Hagemann-Jensen, Björn Reinius, Tina Jacob, Tim Dalessandri, Gert-Jan Hendriks, Maria Kasper, and Rickard Sandberg. Transcriptional bursts explain autosomal random monoallelic expression and affect allelic imbalance. *PLOS Computational Biology*, 17(3):e1008772, March 2021. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1008772. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008772. Publisher: Public Library of Science.

[17] Diyan Li, Mengnan He, Qianzi Tang, Shilin Tian, Jiaman Zhang, Yan Li, Danyang Wang, Long Jin, Chunyou Ning, Wei Zhu, Silu Hu, Keren Long, Jideng Ma, Jing Liu, Zhihua Zhang, and Mingzhou Li. Comparative 3D genome architecture in vertebrates. *BMC Biol*, 20(1):99, May 2022. ISSN 1741-7007. doi: 10.1186/s12915-022-01301-7.

[18] Guoliang Li, Xiaoan Ruan, Raymond K. Auerbach, Kuljeet Singh Sandhu, Meizhen Zheng, Ping Wang, Huay Mei Poh, Yufen Goh, Joanne Lim, Jingyao Zhang, Hui Shan Sim, Su Qin Peh, Fabianus Hendriyan Mulawadi, Chin Thing Ong, Yuriy L. Orlov, Shuzhen Hong, Zhizhuo Zhang, Steve Landt, Debasish Raha, Ghia Euskirchen, Chia-Lin Wei, Weihong Ge, Huaien Wang, Carrie Davis, Katherine I. Fisher-Aylor, Ali Mortazavi, Mark Gerstein, Thomas Gingeras, Barbara Wold, Yi Sun, Melissa J. Fullwood, Edwin Cheung, Edison Liu, Wing-Kin Sung, Michael Snyder, and Yijun Ruan. Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. *Cell*, 148(1):84–98, January 2012. ISSN 0092-8674. doi: 10.1016/j.cell.2011.12.014. URL https://www.sciencedirect.com/science/article/pii/S0092867411015170.

[19] Hannah K. Long, Sara L. Prescott, and Joanna Wysocka. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell*, 167(5): 1170–1187, November 2016. ISSN 0092-8674. doi: 10.1016/j.cell.2016.09.018. URL https://www.sciencedirect.com/science/article/pii/S009286741631251X.

[20] Joseph V. W. Meeussen and Tineke L. Lenstra. Time will tell: comparing timescales to gain insight into transcriptional bursting. *Trends in Genetics*, 40(2):160–174, February 2024. ISSN 0168-9525. doi: 10.1016/j.tig.2023.11.003. URL https://www.cell.com/trends/genetics/abstract/S0168-9525(23)00258-5. Publisher: Elsevier.

[21] Hiroshi Ochiai, Tetsutaro Hayashi, Mana Umeda, Mika Yoshimura, Akihito Harada, Yukiko Shimizu, Kenta Nakano, Noriko Saitoh, Zhe Liu, Takashi Yamamoto, Tadashi Okamura, Yasuyuki Ohkawa, Hiroshi Kimura, and Itoshi Nikaido. Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Science Advances*, 6(25):eaaz6699, June 2020. doi: 10.1126/sciadv.aaz6699. URL https://www.science.org/doi/full/10.1126/sciadv.aaz6699. Publisher: American Association for the Advancement of Science.

[22] Arjun Raj and Alexander van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, 135(2):216–226, October 2008. ISSN 0092-8674. doi: 10.1016/j.cell.2008.09.050. URL https://www.sciencedirect.com/science/article/pii/S0092867408012439.

[23] Arjun Raj, Charles S. Peskin, Daniel Tranchina, Diana Y. Vargas, and Sanjay Tyagi. Stochastic mRNA Synthesis in Mammalian Cells. *PLOS Biology*, 4(10):e309, September 2006. ISSN 1545-7885. doi: 10.1371/journal.pbio.0040309. URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0040309. Publisher: Public Library of Science.

[24] Helen Ray-Jones and Mikhail Spivakov. Transcriptional enhancers and their communication with gene promoters. *Cell. Mol. Life Sci.*, 78(19):6453–6485, October 2021. ISSN 1420-9071. doi: 10.1007/s00018-021-03903-w. URL https://doi.org/10.1007/s00018-021-03903-w.

[25] William F. Richter, Shraddha Nayak, Janet Iwasa, and Dylan J. Taatjes. The Mediator complex as a master regulator of transcription by RNA polymerase II. *Nat Rev Mol Cell Biol*, 23(11):732–749, November 2022. ISSN 1471-0080. doi: 10.1038/s41580-022-00498-3. URL https://www.nature.com/articles/s41580-022-00498-3. Publisher: Nature Publishing Group.

[26] Amartya Sanyal, Bryan R. Lajoie, Gaurav Jain, and Job Dekker. The long-range interaction landscape of gene promoters. *Nature*, 489(7414):109–113, September 2012. ISSN 1476-4687. doi: 10.1038/nature11279. URL https://www.nature.com/articles/nature11279. Publisher: Nature Publishing Group.

[27] Stefan Schoenfelder, Mayra Furlan-Magaril, Borbala Mifsud, Filipe Tavares-Cadete, Robert Sugar, Biola-Maria Javierre, Takashi Nagano, Yulia Katsman, Moorthy Sakthidevi, Steven W. Wingett, Emilia Dimitrova, Andrew Dimond, Lucas B. Edelman, Sarah Elderkin, Kristina Tabbada, Elodie Darbo, Simon Andrews, Bram Herman, Andy Higgs, Emily LeProust, Cameron S. Osborne, Jennifer A. Mitchell, Nicholas M. Luscombe, and Peter Fraser. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.*, 25(4):582–597, April 2015. ISSN 1088-9051, 1549-5469. doi: 10.1101/gr.185272.114. URL https://genome-cshlp-org.proxy3.library.mcgill.ca/content/25/4/582. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[28] Adrien Senecal, Brian Munsky, Florence Proux, Nathalie Ly, Floriane E. Braye, Christophe Zimmer, Florian Mueller, and Xavier Darzacq. Transcription Factors Modulate c-Fos Transcriptional Bursts. *Cell Reports*, 8(1):75–83, July 2014. ISSN 2211-1247. doi: 10.1016/j.celrep.2014.05.053. URL https://www.cell.com/cell-reports/abstract/S2211-1247(14)00447-1. Publisher: Elsevier.

[29] Fei Sun, Constantinos Chronis, Michael Kronenberg, Xiao-Fen Chen, Trent Su, Fides D. Lay, Kathrin Plath, Siavash K. Kurdistani, and Michael F. Carey. Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. *Molecular Cell*, 73 (2):250–263.e5, January 2019. ISSN 1097-2765. doi: 10.1016/j.molcel.2018.10.039. URL https://www.sciencedirect.com/science/article/pii/S109727651830933X.

[30] David M. Suter, Nacho Molina, David Gatfield, Kim Schneider, Ueli Schibler, and Felix Naef. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science*, 332(6028):472–474, April 2011. doi: 10.1126/science.1198817. URL http://www.science.org/doi/10.1126/science.1198817. Publisher: American Association for the Advancement of Science.

[31] Christopher M. Uyehara and Effie Apostolou. 3D enhancer-promoter interactions and multi-connected hubs: Organizational principles and functional roles. *Cell Reports*, 42(4):112068, April 2023. ISSN 22111247. doi: 10.1016/j.celrep.2023.112068. URL https://linkinghub.elsevier.com/retrieve/pii/S2211124723000797.

[32] Warren A. Whyte, David A. Orlando, Denes Hnisz, Brian J. Abraham, Charles Y. Lin, Michael H. Kagey, Peter B. Rahl, Tong Ihn Lee, and Richard A. Young. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell*, 153(2):307–319, April 2013. ISSN 0092-8674. doi: 10.1016/j.cell.2013.03.035. URL https://www.sciencedirect.com/science/article/pii/S0092867413003929.

[33] Jin H. Yang and Anders S. Hansen. Enhancer selectivity in space and time: from enhancer–promoter interactions to promoter activation. *Nat Rev Mol Cell Biol*, pages 1–18, February 2024. ISSN 1471-0080. doi: 10.1038/s41580-024-00710-6. URL https://www.nature.com/articles/s41580-024-00710-6. Publisher: Nature Publishing Group.

[34] Marty G Yang, Emi Ling, Christopher J Cowley, Michael E Greenberg, and Thomas Vierbuchen. Characterization of sequence determinants of enhancer function using natural genetic variation. *eLife*, 11:e76500, August 2022. ISSN 2050-084X. doi: 10.7554/eLife.76500. URL https://doi.org/10.7554/eLife.76500. Publisher: eLife Sciences Publications, Ltd.

[35] Benjamin Zoller, Damien Nicolas, Nacho Molina, and Felix Naef. Structure of silent transcription intervals and noise characteristics of mammalian genes. *Molecular Systems Biology*, 11(7):823, July 2015. ISSN 1744-4292. doi: 10.15252/msb.20156257. URL https://www-embopress-org.proxy3.library.mcgill.ca/doi/full/10.15252/msb.20156257. Publisher: John Wiley & Sons, Ltd.

[36] Jessica Zuin, Gregory Roth, Yinxiu Zhan, Julie Cramard, Josef Redolfi, Ewa Piskadlo, Pia Mach, Mariya Kryzhanovska, Gergely Tihanyi, Hubertus Kohler, Mathias Eder, Christ Leemans, Bas van Steensel, Peter Meister, Sebastien Smallwood, and Luca Giorgetti. Nonlinear control of transcription through enhancer–promoter interactions. *Nature*, 604(7906):571–577, April 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04570-y. URL https://www.nature.com/articles/s41586-022-04570-y. Number: 7906 Publisher: Nature Publishing Group.

Discussion and Conclusions

4

Discussion

4.0.1 Mediator in Depleted Contexts: What do we know?

Exploring the kinetics of transcription in the context of Mediator activity presents a unique perspective for developing a more fundamental understanding of gene expression. Mediator is active in many aspects transcription which seemingly influence disparate kinetic outputs. Its capacity to deliver regulatory signals from distant enhancers is one of its most defining features. Equivalently, its ability to facilitate PolII PIC release is thought to contribute to the complex's essentiality. How the transcriptome might react to Mediator degradation within the context of kinetics is still outside of our understanding, but similar studies might offer some clues.

For Mediator's activity at enhancers, one group utilized a high-throughput enhancer activity screen (STARR-Seq) to assess co-factor specificity in mammalian cells. Using a degron tagged Med14 in HCT116 cells, researchers found a substantial subgroups of enhancers that showed no transcriptional response to Mediator depletion [49]. Intuitively, genes which showed the greatest negative response to Mediator depletion (such as MYC) likewise had increased levels of Med1 at enhancers. For Mediator independent enhancers, Neumayr et al. found that these loci were enriched in both p53 and BRD4, suggesting a possible rescue mechanism [49]. In this study and elsewhere, Mediator abrogation is accompanied with a near global down-regulation of transcripts [49, 35]. In Jaeger et al.'s work, super-enhancers, particularly among self-regulating transcription factors (such as MYC again), down-regulation was at its most pronounced [35]. Similarly, a rescue phenotype was found whereby the pause-release factor pTEFB became over-abundant when Mediator was depleted which partially mitigated transcriptional disruption [35]. The authors conclude that SE dependant genes are, in some unknown capacity, unable to facilitate this rescue mechanism [35]. Although here Jaeger et al. suggest an inhibitory role by Mediator on pTEFB, other groups suggest the opposite, particularly through the activity of select subunits such as Med23 and Med26 [78, 88].

Given this line of evidence, how might we imagine the kinetics of transcription in the absence of Mediator? Fundamentally these observations seem revolve around the sensitivity of super-enhancers. In our own analysis, proximal genes to SEs appear to be upregulated in burst frequency rather than burst size. These SEs are definitionally enriched in Med1 signal, which appears to be dependant on the presence of the canonical complex at these loci [35]. When depleted, we could expect Mediator to be unable to initiate a permissive transcriptional state, thus being realized in a decrease in burst frequency. Enhancers which appear independent of Mediator status could be invariant in burst frequency at their cognate genes but suffer from depleted burst sizes from the lack of PolII turnover. It would be interesting to see which genetic contexts offer a rescue phenotype; either from an abundance of pTEFb at promoters or p53 at enhancers. Unraveling such dynamics would provide the needed clues to uncover the sequence specificity for Mediator dependence and activity at large.

4.0.2 Enhancers and Mediator Condensates

Mediator activity at SEs is often associated with clusters of phase-separated complexes [35, 19]. In our work, SEs defined by an overabundance of Med1 showed an increase in burst frequency in proximal genes. In a recent imaging experiment, labeled Mediator condensates entering in proximity of the Sox9 enhancer/promoter locus was accompanied by a burst of transcription [19]. Although isolated, this line of evidence suggest Mediator facilitates transcription by inducing more frequent transitions to an active state. It may also be possible that these genes (primarily developmental) have evolved to be responsive to Mediator in its condensate form, and such a state is associated with burst frequency via the condensate's ability to draw enhancers into proximity with the promoter. There is some evidence of this dosage requirement for enhancer recruitment; DNA strands can be coerced into proximity by the force induced from the surface-tension of phase-separated protein condensates [74, 59]. On the other hand, chromatin confirmation capture methods (HiChIP, Hi-C, etc ...) repeatedly show a lack of structural changes when Mediator is depleted [35, 21]. The order of causation suggested from these experiments make already formed EPs the scaffold by which Mediator phase-separates onto, facilitated by the loop extrusion activity of cohesin stalled at CTCF sites or other DNA binding elements [21]. Alternatively, It may also be possible that 3C methods in general are underestimating the number of EP contacts. Proximity ligation necessitates the assumption that the capture radius is similar to the functional EP distance [93]. If EPs are in some proportion mediated by large macromolecular condensates, those of which can exceed as much as three times the size of the ligation capture proximity, that proportion will not be incorporated into the output [93]. If Hi-C is unable to capture this class of interaction (200-300nm in distance), then a Mediator perturbation will likely be underestimated in the contact map. Future studies with regards to Mediator, chromatin conformation and bursting will require careful designs in order to tease out the causative properties of these relationships.

In a similar light, enhancer hubs independent of their 'super' status have very interesting properties in of themselves. Most notably from a promoter-centric perspective, increasing number of enhancer contacts induces different responses for burst size and frequency. Promoters seemingly only need a single enhancer contact to induce faster switching into an ON state. The observed abundance of enhancers connected at some genes suggest redundancy,

which has been observed in some circumstances with gene expression in general [62]. Alternatively, enhancer hubs have been shown to assemble hierarchically whereby one 'hub' enhancer communicates with multiple neighboring enhancers and their cognate promoter [34]. Further enhancer knockdowns experiments in the context of kinetics could help to enlighten how individual enhancer hubs work how prevalent these models are.

Models of enhancer cooperativity have been shown elsewhere in other chromatin capture studies [62]. Cooperation has been seen to drive an additive model in their transcriptional output, particularly in lineage defining genes [68]. Here we not only can we reproduce these findings but propose a more nuanced model. The linear accumulation of burst size with enhancers is likely the main contributing factor to the general additive model. What these enhancers are doing with regards to promoting this active state is still unclear. Some groups suggest enhancer hubs and SEs are capable of capturing condensates, Mediator or otherwise [33, 32]. Our evidence suggests that the constituents which maintain an ON period are withheld at the site of expression by groups of enhancers. Imaging studies of labelled chromatin and condensates could help reveal to what extent this is true.

4.0.3 Topologically Associating Domains: A Permeable Barrier?

Lastly, topologically associating domains have quickly become a controversial category of chromatin organization. While largely accepted as one of the main drivers for enhancer selection, some select examples tend to break with this consensus. Chakraborty et al. in their focused analysis of the Sox2 enhancer in mESCs found that introducing new TADs boundaries in between the Sox2 promoter and the SCR produced very little difference in transcription [11]. Additionally, work by Ghavi-Helm et al. using a highly rearranged *Drosphila* 'balancer' genome in combination with allele-specific Hi-C and RNAseq show that only a minority of genes are affected by TAD disruptions despite the lower contact frequency [28]. Polymer simulations and *in-vitro* testing have shown that there is a sigmoidal relationship between contact frequency and burst frequency, suggesting that promoters do not require extensive interactions to initiate bursts [24, 99] . It follows therefore that even with a reduced contact frequency prompted by cohesin activity burst initiation would remain largely unattenuated. Our results show that some TAD boundaries are permissible to HiChIP H3K27ac EPs and largely do not impact bursting kinetics. This follows the notion that TADs are not necessary for EP formation nor do they explicitly insulate all EPs. Again we can only speculate on Mediator's role on facilitating these interactions. If Mediator or other co-factors play more of a structural role in maintaining these types of linkages, we should expect a loss of burst frequency at these sites when the acting co-factor is depleted in combination with a near absolute disruption of contact frequencies. As we have seen with other Mediator depletion experiments in the literature, this may have an uneven effect across the genome. Future studies will need to incorporate these many aspects in order to determine the insulating properties of TADs and their selective impact on neighboring genes.

5

Conclusions and Future Work

The discovery and characterization of transcription is decades old, it is ever more surprising that new discoveries are continually reshaping our understanding of this most fundamental process. In this project we explored kinetics of transcription and how enhancers play a fundamental role in shaping them. We also speculated on the role of Mediator, an essential transcriptional co-factor, on its impact on the two state model of gene expression. Additionally, acute partial Mediator degradation was achieved in a mouse embryonic fibroblast cell line which could be used to estimate kinetics in depleted conditions using scRNAseq.

Our understanding of Mediator and enhancer's regulatory potential is still nascent. For Mediator, beyond degrading the entire complex, individual components may also have different kinetic potentials which could be observed when degraded. As discussed previously, Med1 appears to play an essential role in regulating super-enhancers an cell-fate genes. Depletion of this subunit could potentially target these genes for kinetic perturbations, specifically burst frequency. Med11, as previously explored by Tantale et al. specifically impacts the number of PolII loaded into a convoy during a burst [79]. Again, a disjointed kinetic response could be observed by an attenuated burst size. In these examples and more, it would be interesting to see the extent by which these perturbations are gene specific and what genetic material constitutes subunit sensitivity at these loci.

There are a number enhancer hub models each with their own set of experimental evidence [84, 44]. From multiple enhancers acting additively, hierarchically or redundantly, our evidence leaves room for each model from a global perspective. Crucially, however, these enhancer models seem to be reflected differently in either kinetic parameter: burst size appears additive upon increasing enhancer contacts while burst frequency does not. Future work will need to focus where this general trend holds true. Primarily this could be done using enhancer knockdown strategies in combination with scRNAseq or nascent RNA imaging. Enhancer silencing strategies using CRISPRi have been used previously to target specific loci [34]. High-throughput strategies such as Mosaic-seq using similar methods but leverage barcoded gRNAs and lentiviral transfection for a global approach [92]. Using a combination

of these strategies, maps of enhancer relationships across organisms and cell-types could be constructed. In combination co-factor knockdowns, the regulatory landscape could further be characterized not only in their transcriptional output but in the molecular components that underpin it.

Lastly, intersection between chromatin conformation and transcriptional kinetics is underdeveloped. Although both the depletion of Cohesin sub-units and CTCF have both been performed separately, both have yet to be linked to a kinetic output. Chromatin insulation has been shown to a have profound effect on bursting on a single reporter gene in *Drosophila*, despite the mild difference in expression in mammalian cells when TADs are removed [25, 37] [52, 69]. A whole genome analysis could help reveal where insulation matters and how they regulate the particulars of transcription.

As with most aspects of fundamental research, future applications and paths of discovery radiate from the questions we ask in the present. In this work, we hope to contribute to these ever expanding potentials.

Acronyms

- **AID**: **A**uxin **I**nducible **D**egradation
- **CAK**: **C**DK **A**ctivating **K**inase module
- **CDK**:**C**yclin **D**ependant **K**inase
- **CKM**: **C**DK8 **K**inase **M**odule
- **CTD**: **C**-**T**erminal **D**omain
- **EP**: **E**nhancer **P**romoter
- **ER**: **E**strogen **R**eceptor
- **HDR**: **H**omology **D**irected **R**epair
- **IDR**: **I**ntrinsically **D**isordered **R**egion
- **IAA**: **I**ndole-3-**a**cetic **a**cid
- **MEF**: **M**ouse **E**mbryonic **F**ibroblast
- **mESC**: **m**ouse **E**mbryonic **S**tem **C**ell
- **MLE**: **M**aximum **L**ikelihood **E**stimator
- **OsTIR1**: **O**ryza **s**ativa **T**ransport **I**nhibitor **R**esponse **1**
- **PDF**: **P**robability **D**ensity **F**unction
- **PIC**: **P**re-**I**nitiation **C**omplex
- **PolII**: RNA **Pol**ymerase **II**
- **scRNA-seq**: **s**ingle **c**ell **R**NA **seq**uencing
- **SE**: **S**uper **E**nhancer
- **siRNA**: **s**mall **i**nterfering **R**NA
- **smFISH**: **s**ingle **m**olecule **F**luorescence **I**n **S**itu **H**ybridization
- **SEC**: **S**uper **E**longation **C**omplex

- **SMT**: **S**ingle **M**olecule **T**racking
- **SNP**: **S**ingle **N**ucleotide **P**olymorphisms
- **TAD**: **T**opologically **A**ssociating **D**omain
- **TAFs**: **T**BP **A**ssociating **F**actor**s**
- **TBP**: **T**ATA **B**inding **P**rotein
- **TF**: **T**ranscription **F**actor
- **TSS**: **T**ranscription **S**tart **S**ite

References

[1] Med14 mediator complex subunit 14 [Mus musculus (house mouse)] - Gene - NCBI, . URL https://www.ncbi.nlm.nih.gov/gene/26896.

[2] MED17 Polyclonal Antibody (11505-1-AP), . URL https://www.thermofisher.com/antibody/product/MED17-Antibody-Polyclonal/11505-1-AP.

[3] Benjamin L. Allen and Dylan J. Taatjes. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol*, 16(3):155–166, March 2015. ISSN 1471-0080. doi: 10.1038/nrm3951. URL https://www.nature.com/articles/nrm3951. Number: 3 Publisher: Nature Publishing Group.

[4] Keren Bahar Halpern, Sivan Tanami, Shanie Landen, Michal Chapal, Liran Szlak, Anat Hutzler, Anna Nizhberg, and Shalev Itzkovitz. Bursty Gene Expression in the Intact Mammalian Liver. *Molecular Cell*, 58(1):147–156, April 2015. ISSN 1097-2765. doi: 10.1016/j.molcel.2015.01.027. URL https://www.sciencedirect.com/science/article/pii/S1097276515000507.

[5] Caroline R. Bartman, Sarah C. Hsu, Chris C. S. Hsiung, Arjun Raj, and Gerd A. Blobel. Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell*, 62(2):237–247, April 2016. ISSN 1097-2765. doi: 10.1016/j.molcel.2016.03.007. URL https://www.sciencedirect.com/science/article/pii/S1097276516001854.

[6] William J. Blake, Mads KÆrn, Charles R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, April 2003. ISSN 1476-4687. doi: 10.1038/nature01546. URL http://www.nature.com/articles/nature01546. Number: 6932 Publisher: Nature Publishing Group.

[7] Marc Boehning, Claire Dugast-Darzacq, Marija Rankovic, Anders S. Hansen, Taekyung Yu, Herve Marie-Nelly, David T. McSwiggen, Goran Kokic, Gina M. Dailey, Patrick Cramer, Xavier Darzacq, and Markus Zweckstetter. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol*, 25(9):833–840, September 2018. ISSN 1545-9985. doi: 10.1038/s41594-018-0112-y. URL https://www.nature.com/articles/s41594-018-0112-y. Publisher: Nature Publishing Group.

[8] Ann Boija, Isaac A. Klein, Benjamin R. Sabari, Alessandra Dall'Agnese, Eliot L. Coffey, Alicia V. Zamudio, Charles H. Li, Krishna Shrinivas, John C. Manteiga, Nancy M. Hannett, Brian J. Abraham, Lena K. Afeyan, Yang E. Guo, Jenna K. Rimel,

Charli B. Fant, Jurian Schuijers, Tong Ihn Lee, Dylan J. Taatjes, and Richard A. Young. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*, 175(7):1842–1855.e16, December 2018. ISSN 0092-8674. doi: 10.1016/j.cell.2018.10.042. URL https://www.sciencedirect.com/science/article/pii/S0092867418313989.

[9] Brian Teague. Cytoflow: A Python Toolbox for Flow Cytometry. *bioRxiv*, page 2022.07.22.501078, January 2022. doi: 10.1101/2022.07.22.501078. URL http://biorxiv.org/content/early/2022/07/23/2022.07.22.501078.abstract.

[10] Greg T. Cantin, Jennitte L. Stevens, and Arnold J. Berk. Activation domain–mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proceedings of the National Academy of Sciences*, 100(21):12003–12008, October 2003. doi: 10.1073/pnas.2035253100. URL https://www-pnas-org.proxy3.library.mcgill.ca/doi/10.1073/pnas.2035253100. Publisher: Proceedings of the National Academy of Sciences.

[11] Shreeta Chakraborty, Nina Kopitchinski, Ariel Eraso, Parirokh Awasthi, Raj Chari, and Pedro P Rocha. High affinity enhancer-promoter interactions can bypass CTCF/cohesin-mediated insulation and contribute to phenotypic robustness. preprint, Developmental Biology, January 2022. URL http://biorxiv.org/lookup/doi/10.1101/2021.12.30.474562.

[12] Hongtao Chen, Michal Levo, Lev Barinov, Miki Fujioka, James B. Jaynes, and Thomas Gregor. Dynamic interplay between enhancer–promoter topology and gene activity. *Nat Genet*, 50(9):1296–1303, September 2018. ISSN 1546-1718. doi: 10.1038/s41588-018-0175-z. URL https://www.nature.com/articles/s41588-018-0175-z. Number: 9 Publisher: Nature Publishing Group.

[13] Won-Ki Cho, Jan-Hendrik Spille, Micca Hecht, Choongman Lee, Charles Li, Valentin Grube, and Ibrahim I. Cisse. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*, 361(6400):412–415, July 2018. doi: 10.1126/science.aar4199. URL https://www-science-org.proxy3.library.mcgill.ca/doi/10.1126/science.aar4199. Publisher: American Association for the Advancement of Science.

[14] Jonathan R. Chubb, Tatjana Trcek, Shailesh M. Shenoy, and Robert H. Singer. Transcriptional Pulsing of a Developmental Gene. *Current Biology*, 16(10):1018–1025, May 2006. ISSN 0960-9822. doi: 10.1016/j.cub.2006.03.092. URL https://www.sciencedirect.com/science/article/pii/S0960982206014266.

[15] Wouter de Laat and Denis Duboule. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature*, 502(7472):499–506, October 2013. ISSN 1476-4687. doi: 10.1038/nature12753. URL https://www.nature.com/articles/nature12753. Publisher: Nature Publishing Group.

[16] Qiaolin Deng, Daniel Ramsköld, Björn Reinius, and Rickard Sandberg. Single-Cell RNA-Seq Reveals Dynamic, Random Monoallelic Gene Expression in Mammalian Cells.

Science, 343(6167):193–196, January 2014. doi: 10.1126/science.1245316. URL https://www.science.org/doi/10.1126/science.1245316. Publisher: American Association for the Advancement of Science.

[17] Siddharth S. Dey, Jonathan E. Foley, Prajit Limsirichai, David V. Schaffer, and Adam P. Arkin. Orthogonal control of expression mean and variance by epigenetic features at different genomic loci. *Molecular Systems Biology*, 11(5):806, May 2015. ISSN 1744-4292. doi: 10.15252/msb.20145704. URL https://www.embopress.org/doi/10.15252/msb.20145704. Publisher: John Wiley & Sons, Ltd.

[18] Benjamin T Donovan, Anh Huynh, David A Ball, Heta P Patel, Michael G Poirier, Daniel R Larson, Matthew L Ferguson, and Tineke L Lenstra. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *The EMBO Journal*, 38(12):e100809, June 2019. ISSN 0261-4189. doi: 10.15252/embj.2018100809. URL https://www.embopress.org/doi/full/10.15252/embj.2018100809. Publisher: John Wiley & Sons, Ltd.

[19] Manyu Du, Simon Hendrik Stitzinger, Jan-Hendrik Spille, Won-Ki Cho, Choongman Lee, Mohammed Hijaz, Andrea Quintana, and Ibrahim I. Cissé. Direct observation of a condensate effect on super-enhancer controlled gene bursting. *Cell*, 0(0), January 2024. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2023.12.005. URL https://www.cell.com/cell/abstract/S0092-8674(23)01337-5. Publisher: Elsevier.

[20] Mohamed A. El-Brolosy and Didier Y. R. Stainier. Genetic compensation: A phenomenon in search of mechanisms. *PLOS Genetics*, 13(7):e1006780, July 2017. ISSN 1553-7404. doi: 10.1371/journal.pgen.1006780. URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006780. Publisher: Public Library of Science.

[21] Laila El Khattabi, Haiyan Zhao, Jens Kalchschmidt, Natalie Young, Seolkyoung Jung, Peter Van Blerkom, Philippe Kieffer-Kwon, Kyong-Rim Kieffer-Kwon, Solji Park, Xiang Wang, Jordan Krebs, Subhash Tripathi, Noboru Sakabe, Débora R. Sobreira, Su-Chen Huang, Suhas S.P. Rao, Nathanael Pruett, Daniel Chauss, Erica Sadler, Andrea Lopez, Marcelo A. Nóbrega, Erez Lieberman Aiden, Francisco J. Asturias, and Rafael Casellas. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell*, 178(5):1145–1158.e20, August 2019. ISSN 00928674. doi: 10.1016/j.cell.2019.07.011. URL https://linkinghub.elsevier.com/retrieve/pii/S0092867419307767.

[22] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. Stochastic Gene Expression in a Single Cell. *Science*, 297(5584):1183–1186, August 2002. doi: 10.1126/science.1070919. URL https://www.science.org/doi/full/10.1126/science.1070919. Publisher: American Association for the Advancement of Science.

[23] Cyril Esnault, Yad Ghavi-Helm, Sylvain Brun, Julie Soutourina, Nynke Van Berkum, Claire Boschiero, Frank Holstege, and Michel Werner. Mediator-Dependent Recruitment of TFIIH Modules in Preinitiation Complex. *Molecular Cell*, 31(3):337–346,

August 2008. ISSN 1097-2765. doi: 10.1016/j.molcel.2008.06.021. URL https://www.sciencedirect.com/science/article/pii/S109727650800470X.

[24] Yi Fu, Finnegan Clark, Sofia Nomikou, Aristotelis Tsirigos, and Timothee Lionnet. Connecting Chromatin Structures to Gene Regulation Using Dynamic Polymer Simulations. *bioRxiv*, page 2023.11.07.566032, November 2023. doi: 10.1101/2023.11.07.566032. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10659377/.

[25] Takashi Fukaya, Bomyi Lim, and Michael Levine. Enhancer Control of Transcriptional Bursting. *Cell*, 166(2):358–368, July 2016. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2016.05.025. URL https://www.cell.com/cell/abstract/S0092-8674(16)30573-6. Publisher: Elsevier.

[26] Michele Gabriele, Hugo B. Brandão, Simon Grosse-Holz, Asmita Jha, Gina M. Dailey, Claudia Cattoglio, Tsung-Han S. Hsieh, Leonid Mirny, Christoph Zechner, and Anders S. Hansen. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, 376(6592):496–501, April 2022. doi: 10.1126/science.abn6583. URL https://www-science-org.proxy3.library.mcgill.ca/doi/10.1126/science.abn6583. Publisher: American Association for the Advancement of Science.

[27] Matthew D. Galbraith, Mary A. Allen, Claire L. Bensard, Xiaoxing Wang, Marie K. Schwinn, Bo Qin, Henry W. Long, Danette L. Daniels, William C. Hahn, Robin D. Dowell, and Joaquín M. Espinosa. HIF1A Employs CDK8-Mediator to Stimulate RNAPII Elongation in Response to Hypoxia. *Cell*, 153(6):1327–1339, June 2013. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2013.04.048. URL https://www-cell-com.proxy3.library.mcgill.ca/cell/abstract/S0092-8674(13)00524-2. Publisher: Elsevier.

[28] Yad Ghavi-Helm, Aleksander Jankowski, Sascha Meiers, Rebecca R. Viales, Jan O. Korbel, and Eileen E.M. Furlong. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat Genet*, 51(8):1272–1282, August 2019. ISSN 1061-4036. doi: 10.1038/s41588-019-0462-3. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7116017/.

[29] Bin Gu, Eszter Posfai, and Janet Rossant. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol*, 36 (7):632–637, August 2018. ISSN 1546-1696. doi: 10.1038/nbt.4166. URL http://www.nature.com/articles/nbt.4166. Number: 7 Publisher: Nature Publishing Group.

[30] Yang Eric Guo, John C. Manteiga, Jonathan E. Henninger, Benjamin R. Sabari, Alessandra Dall'Agnese, Nancy M. Hannett, Jan-Hendrik Spille, Lena K. Afeyan, Alicia V. Zamudio, Krishna Shrinivas, Brian J. Abraham, Ann Boija, Tim-Michael Decker, Jenna K. Rimel, Charli B. Fant, Tong Ihn Lee, Ibrahim I. Cisse, Phillip A. Sharp, Dylan J. Taatjes, and Richard A. Young. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*, 572(7770): 543–548, August 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1464-0. URL

https://www.nature.com/articles/s41586-019-1464-0. Publisher: Nature Publishing Group.

[31] Antonina Hafner and Alistair Boettiger. The spatial organization of transcriptional control. *Nat Rev Genet*, 24(1):53–68, January 2023. ISSN 1471-0064. doi: 10.1038/s41576-022-00526-0. URL https://www.nature.com/articles/s41576-022-00526-0. Publisher: Nature Publishing Group.

[32] Kota Hamamoto and Takashi Fukaya. Molecular architecture of enhancer–promoter interaction. *Current Opinion in Cell Biology*, 74:62–70, February 2022. ISSN 0955-0674. doi: 10.1016/j.ceb.2022.01.003. URL https://www.sciencedirect.com/science/article/pii/S0955067422000047.

[33] Denes Hnisz, Krishna Shrinivas, Richard A. Young, Arup K. Chakraborty, and Phillip A. Sharp. A Phase Separation Model for Transcriptional Control. *Cell*, 169(1):13–23, March 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.02.007. URL https://www.sciencedirect.com/science/article/pii/S009286741730185X.

[34] Jialiang Huang, Kailong Li, Wenqing Cai, Xin Liu, Yuannyu Zhang, Stuart H. Orkin, Jian Xu, and Guo-Cheng Yuan. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat Commun*, 9(1):943, March 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-03279-9. URL https://www.nature.com/articles/s41467-018-03279-9. Publisher: Nature Publishing Group.

[35] Martin G. Jaeger, Björn Schwalb, Sebastian D. Mackowiak, Taras Velychko, Alexander Hanzl, Hana Imrichova, Matthias Brand, Benedikt Agerer, Someth Chorn, Behnam Nabet, Fleur M. Ferguson, André C. Müller, Andreas Bergthaler, Nathanael S. Gray, James E. Bradner, Christoph Bock, Denes Hnisz, Patrick Cramer, and Georg E. Winter. Selective Mediator dependence of cell-type-specifying transcription. *Nat Genet*, 52(7):719–727, July 2020. ISSN 1546-1718. doi: 10.1038/s41588-020-0635-0. URL https://www.nature.com/articles/s41588-020-0635-0. Publisher: Nature Publishing Group.

[36] Yuchao Jiang, Nancy R. Zhang, and Mingyao Li. SCALE: modeling allele-specific gene expression by single-cell RNA sequencing. *Genome Biol*, 18(1):1–15, December 2017. ISSN 1474-760X. doi: 10.1186/s13059-017-1200-8. URL https://genomebiology.biomedcentral.com/articles/10.1186/s13059-017-1200-8. Number: 1 Publisher: BioMed Central.

[37] Koji Kawasaki and Takashi Fukaya. Functional coordination between transcription factor clustering and gene activity. *Molecular Cell*, 83(10):1605–1622.e9, May 2023. ISSN 1097-2765. doi: 10.1016/j.molcel.2023.04.018. URL https://www.sciencedirect.com/science/article/pii/S1097276523002897.

[38] Jong Kyoung Kim and John C. Marioni. Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. *Genome Biology*, 14(1):R7, January 2013. ISSN 1474-760X. doi: 10.1186/gb-2013-14-1-r7. URL https://doi.org/10.1186/gb-2013-14-1-r7.

[39] Élie Lambert, Kavindu Puwakdandawa, Yi Fei Tao, and François Robert. From structure to molecular condensates: emerging mechanisms for Mediator function. *The FEBS Journal*, n/a(n/a). ISSN 1742-4658. doi: 10.1111/febs.16250. URL http://onlinelibrary.wiley.com/doi/abs/10.1111/febs.16250. _eprint: https://febs.onlinelibrary.wiley.com/doi/pdf/10.1111/febs.16250.

[40] Anton J. M. Larsson, Per Johnsson, Michael Hagemann-Jensen, Leonard Hartmanis, Omid R. Faridani, Björn Reinius, Åsa Segerstolpe, Chloe M. Rivera, Bing Ren, and Rickard Sandberg. Genomic encoding of transcriptional burst kinetics. *Nature*, 565 (7738):251–254, January 2019. ISSN 1476-4687. doi: 10.1038/s41586-018-0836-1. URL https://www.nature.com/articles/s41586-018-0836-1. Number: 7738 Publisher: Nature Publishing Group.

[41] Anton J. M. Larsson, Christoph Ziegenhain, Michael Hagemann-Jensen, Björn Reinius, Tina Jacob, Tim Dalessandri, Gert-Jan Hendriks, Maria Kasper, and Rickard Sandberg. Transcriptional bursts explain autosomal random monoallelic expression and affect allelic imbalance. *PLOS Computational Biology*, 17(3):e1008772, March 2021. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1008772. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008772. Publisher: Public Library of Science.

[42] Diyan Li, Mengnan He, Qianzi Tang, Shilin Tian, Jiaman Zhang, Yan Li, Danyang Wang, Long Jin, Chunyou Ning, Wei Zhu, Silu Hu, Keren Long, Jideng Ma, Jing Liu, Zhihua Zhang, and Mingzhou Li. Comparative 3D genome architecture in vertebrates. *BMC Biol*, 20(1):99, May 2022. ISSN 1741-7007. doi: 10.1186/s12915-022-01301-7.

[43] Guoliang Li, Xiaoan Ruan, Raymond K. Auerbach, Kuljeet Singh Sandhu, Meizhen Zheng, Ping Wang, Huay Mei Poh, Yufen Goh, Joanne Lim, Jingyao Zhang, Hui Shan Sim, Su Qin Peh, Fabianus Hendriyan Mulawadi, Chin Thing Ong, Yuriy L. Orlov, Shuzhen Hong, Zhizhuo Zhang, Steve Landt, Debasish Raha, Ghia Euskirchen, Chia-Lin Wei, Weihong Ge, Huaien Wang, Carrie Davis, Katherine I. Fisher-Aylor, Ali Mortazavi, Mark Gerstein, Thomas Gingeras, Barbara Wold, Yi Sun, Melissa J. Fullwood, Edwin Cheung, Edison Liu, Wing-Kin Sung, Michael Snyder, and Yijun Ruan. Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. *Cell*, 148(1):84–98, January 2012. ISSN 0092-8674. doi: 10.1016/j.cell.2011.12.014. URL https://www.sciencedirect.com/science/article/pii/S0092867411015170.

[44] Hannah K. Long, Sara L. Prescott, and Joanna Wysocka. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell*, 167(5): 1170–1187, November 2016. ISSN 0092-8674. doi: 10.1016/j.cell.2016.09.018. URL https://www.sciencedirect.com/science/article/pii/S009286741631251X.

[45] Andreas Mayer, Heather M. Landry, and L. Stirling Churchman. Pause & Go: from the discovery of RNA polymerase pausing to its functional implications. *Curr Opin Cell Biol*, 46:72–80, June 2017. ISSN 0955-0674. doi: 10.1016/j.ceb.2017.03.002. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5505790/.

[46] Joseph V. W. Meeussen and Tineke L. Lenstra. Time will tell: comparing timescales to gain insight into transcriptional bursting. *Trends in Genetics*, 40(2):160–174, February 2024. ISSN 0168-9525. doi: 10.1016/j.tig.2023.11.003. URL https://www.cell.com/trends/genetics/abstract/S0168-9525(23)00258-5. Publisher: Elsevier.

[47] Oscar L. Miller Jr. and Steven L. McKnight. Post-replicative nonribosomal transcription units in D. melanogaster embryos. *Cell*, 17(3):551–563, July 1979. ISSN 00928674. doi: 10.1016/0092-8674(79)90263-0. URL https://linkinghub.elsevier.com/retrieve/pii/0092867479902630.

[48] Mthabisi B. Moyo, J. Brandon Parker, and Debabrata Chakravarti. Altered chromatin landscape and enhancer engagement underlie transcriptional dysregulation in MED12 mutant uterine leiomyomas. *Nat Commun*, 11(1):1019, February 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-14701-6. URL https://www.nature.com/articles/s41467-020-14701-6. Publisher: Nature Publishing Group.

[49] Christoph Neumayr, Vanja Haberle, Leonid Serebreni, Katharina Karner, Oliver Hendy, Ann Boija, Jonathan E. Henninger, Charles H. Li, Karel Stejskal, Gen Lin, Katharina Bergauer, Michaela Pagani, Martina Rath, Karl Mechtler, Cosmas D. Arnold, and Alexander Stark. Differential cofactor dependencies define distinct types of human enhancers. *Nature*, 606(7913):406–413, June 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04779-x. URL http://www.nature.com/articles/s41586-022-04779-x. Number: 7913 Publisher: Nature Publishing Group.

[50] Damien Nicolas, Nick E. Phillips, and Felix Naef. What shapes eukaryotic transcriptional bursting? *Mol. BioSyst.*, 13(7):1280–1290, 2017. ISSN 1742-206X, 1742-2051. doi: 10.1039/C7MB00154A. URL http://xlink.rsc.org/?DOI=C7MB00154A.

[51] Kohei Nishimura, Tatsuo Fukagawa, Haruhiko Takisawa, Tatsuo Kakimoto, and Masato Kanemaki. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods*, 6(12):917–922, December 2009. ISSN 1548-7105. doi: 10.1038/nmeth.1401. URL https://www.nature.com/articles/nmeth.1401. Number: 12 Publisher: Nature Publishing Group.

[52] Elphège P. Nora, Anton Goloborodko, Anne-Laure Valton, Johan H. Gibcus, Alec Uebersohn, Nezar Abdennur, Job Dekker, Leonid A. Mirny, and Benoit G. Bruneau. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*, 169(5):930–944.e22, May 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.05.004. URL https://www.sciencedirect.com/science/article/pii/S0092867417305317.

[53] Hiroshi Ochiai, Tetsutaro Hayashi, Mana Umeda, Mika Yoshimura, Akihito Harada, Yukiko Shimizu, Kenta Nakano, Noriko Saitoh, Zhe Liu, Takashi Yamamoto, Tadashi Okamura, Yasuyuki Ohkawa, Hiroshi Kimura, and Itoshi Nikaido. Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Science Advances*, 6(25):eaaz6699, June 2020. doi: 10.1126/sciadv.aaz6699. URL https://www.science.org/doi/full/10.1126/sciadv.aaz6699. Publisher: American Association for the Advancement of Science.

[54] Simona Patange, David A. Ball, Yihan Wan, Tatiana S. Karpova, Michelle Girvan, David Levens, and Daniel R. Larson. MYC amplifies gene expression through global changes in transcription factor dynamics. *Cell Reports*, 38(4), January 2022. ISSN 2211-1247. doi: 10.1016/j.celrep.2021.110292. URL https://www.cell.com/cell-reports/abstract/S2211-1247(21)01807-6. Publisher: Elsevier.

[55] Jean Peccoud and Bernard Ycart. Markovian modeling of gene-product synthesis. *Theoretical Population Biology*, 48(2):222, April 1995. doi: 10.1006/tpbi.1995.1027.

[56] Virginia L. Pimmett, Matthieu Dejean, Carola Fernandez, Antonio Trullo, Edouard Bertrand, Ovidiu Radulescu, and Mounia Lagha. Quantitative imaging of transcription in living Drosophila embryos reveals the impact of core promoter motifs on promoter state dynamics. *Nat Commun*, 12(1):4504, July 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-24461-6. URL https://www.nature.com/articles/s41467-021-24461-6. Number: 1 Publisher: Nature Publishing Group.

[57] Mark E. Pownall, Liyun Miao, Charles E. Vejnar, Ons M'Saad, Alice Sherrard, Megan A. Frederick, Maria D. J. Benitez, Curtis W. Boswell, Kenneth S. Zaret, Joerg Bewersdorf, and Antonio J. Giraldez. Chromatin expansion microscopy reveals nanoscale organization of transcription and chromatin. *Science*, 381(6653):92–100, July 2023. doi: 10.1126/science.ade5308. URL https://www-science-org.proxy3.library.mcgill.ca/doi/10.1126/science.ade5308. Publisher: American Association for the Advancement of Science.

[58] Marina V. Pryzhkova, Michelle J. Xu, and Philip W. Jordan. Adaptation of the AID system for stem cell and transgenic mouse research. *Stem Cell Research*, 49:102078, December 2020. ISSN 1873-5061. doi: 10.1016/j.scr.2020.102078. URL https://www.sciencedirect.com/science/article/pii/S1873506120303792.

[59] Thomas Quail, Stefan Golfier, Maria Elsner, Keisuke Ishihara, Vasanthanarayan Murugesan, Roman Renger, Frank Jülicher, and Jan Brugués. Force generation by protein–DNA co-condensation. *Nat. Phys.*, 17(9):1007–1012, September 2021. ISSN 1745-2481. doi: 10.1038/s41567-021-01285-1. URL https://www.nature.com/articles/s41567-021-01285-1. Publisher: Nature Publishing Group.

[60] Arjun Raj and Alexander van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, 135(2):216–226, October 2008. ISSN 0092-8674. doi: 10.1016/j.cell.2008.09.050. URL https://www.sciencedirect.com/science/article/pii/S0092867408012439.

[61] Arjun Raj, Charles S. Peskin, Daniel Tranchina, Diana Y. Vargas, and Sanjay Tyagi. Stochastic mRNA Synthesis in Mammalian Cells. *PLOS Biology*, 4(10):e309, September 2006. ISSN 1545-7885. doi: 10.1371/journal.pbio.0040309. URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0040309. Publisher: Public Library of Science.

[62] Helen Ray-Jones and Mikhail Spivakov. Transcriptional enhancers and their communication with gene promoters. *Cell. Mol. Life Sci.*, 78(19):6453–6485, October 2021.

ISSN 1420-9071. doi: 10.1007/s00018-021-03903-w. URL https://doi.org/10.1007/s00018-021-03903-w.

[63] William F. Richter, Shraddha Nayak, Janet Iwasa, and Dylan J. Taatjes. The Mediator complex as a master regulator of transcription by RNA polymerase II. *Nat Rev Mol Cell Biol*, 23(11):732–749, November 2022. ISSN 1471-0080. doi: 10.1038/s41580-022-00498-3. URL https://www.nature.com/articles/s41580-022-00498-3. Publisher: Nature Publishing Group.

[64] Andrea Rossi, Zacharias Kontarakis, Claudia Gerri, Hendrik Nolte, Soraya Hölper, Marcus Krüger, and Didier Y. R. Stainier. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, 524(7564):230–233, August 2015. ISSN 1476-4687. doi: 10.1038/nature14580. URL https://www.nature.com/articles/nature14580. Number: 7564 Publisher: Nature Publishing Group.

[65] Sarah Sainsbury, Carrie Bernecky, and Patrick Cramer. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol*, 16(3):129–143, March 2015. ISSN 1471-0080. doi: 10.1038/nrm3952. URL https://www.nature.com/articles/nrm3952. Publisher: Nature Publishing Group.

[66] Tetsushi Sakuma, Shota Nakade, Yuto Sakane, Ken-Ichi T. Suzuki, and Takashi Yamamoto. MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc*, 11(1):118–133, January 2016. ISSN 1750-2799. doi: 10.1038/nprot.2015.140. URL https://www.nature.com/articles/nprot.2015.140. Publisher: Nature Publishing Group.

[67] Amartya Sanyal, Bryan R. Lajoie, Gaurav Jain, and Job Dekker. The long-range interaction landscape of gene promoters. *Nature*, 489(7414):109–113, September 2012. ISSN 1476-4687. doi: 10.1038/nature11279. URL https://www.nature.com/articles/nature11279. Publisher: Nature Publishing Group.

[68] Stefan Schoenfelder, Mayra Furlan-Magaril, Borbala Mifsud, Filipe Tavares-Cadete, Robert Sugar, Biola-Maria Javierre, Takashi Nagano, Yulia Katsman, Moorthy Sakthidevi, Steven W. Wingett, Emilia Dimitrova, Andrew Dimond, Lucas B. Edelman, Sarah Elderkin, Kristina Tabbada, Elodie Darbo, Simon Andrews, Bram Herman, Andy Higgs, Emily LeProust, Cameron S. Osborne, Jennifer A. Mitchell, Nicholas M. Luscombe, and Peter Fraser. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.*, 25(4):582–597, April 2015. ISSN 1088-9051, 1549-5469. doi: 10.1101/gr.185272.114. URL https://genome-cshlp-org.proxy3.library.mcgill.ca/content/25/4/582. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[69] Wibke Schwarzer, Nezar Abdennur, Anton Goloborodko, Aleksandra Pekowska, Geoffrey Fudenberg, Yann Loe-Mie, Nuno A. Fonseca, Wolfgang Huber, Christian H. Haering, Leonid Mirny, and Francois Spitz. Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, 551(7678):51–56, November 2017. ISSN

1476-4687. doi: 10.1038/nature24281. URL https://www.nature.com/articles/nature24281. Publisher: Nature Publishing Group.

[70] Adrien Senecal, Brian Munsky, Florence Proux, Nathalie Ly, Floriane E. Braye, Christophe Zimmer, Florian Mueller, and Xavier Darzacq. Transcription Factors Modulate c-Fos Transcriptional Bursts. *Cell Reports*, 8(1):75–83, July 2014. ISSN 2211-1247. doi: 10.1016/j.celrep.2014.05.053. URL https://www.cell.com/cell-reports/abstract/S2211-1247(14)00447-1. Publisher: Elsevier.

[71] Edgar Serfling, Maria Jasin, and Walter Schaffner. Enhancers and eukaryotic gene transcription. *Trends in Genetics*, 1:224–230, January 1985. ISSN 0168-9525. doi: 10.1016/0168-9525(85)90088-5. URL https://www.cell.com/trends/genetics/abstract/0168-9525(85)90088-5. Publisher: Elsevier.

[72] Bin Shen, Wensheng Zhang, Jun Zhang, Jiankui Zhou, Jianying Wang, Li Chen, Lu Wang, Alex Hodgkins, Vivek Iyer, Xingxu Huang, and William C. Skarnes. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods*, 11(4):399–402, April 2014. ISSN 1548-7105. doi: 10.1038/nmeth.2857. URL https://www.nature.com/articles/nmeth.2857. Number: 4 Publisher: Nature Publishing Group.

[73] Ameet Shetty, Natalia I. Reim, and Fred Winston. Auxin-inducible degron system for depletion of proteins in S. cerevisiae. *Curr Protoc Mol Biol*, 128(1):e104, September 2019. ISSN 1934-3639. doi: 10.1002/cpmb.104. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6741457/.

[74] Yongdae Shin, Yi-Che Chang, Daniel S. W. Lee, Joel Berry, David W. Sanders, Pierre Ronceray, Ned S. Wingreen, Mikko Haataja, and Clifford P. Brangwynne. Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. *Cell*, 175(6):1481–1491.e13, November 2018. ISSN 0092-8674. doi: 10.1016/j.cell.2018.10.057. URL https://www.sciencedirect.com/science/article/pii/S0092867418314569.

[75] Manuel Stemmer, Thomas Thumberger, Maria del Sol Keyer, Joachim Wittbrodt, and Juan L. Mateo. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLOS ONE*, 10(4):e0124633, April 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0124633. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0124633. Publisher: Public Library of Science.

[76] Fei Sun, Constantinos Chronis, Michael Kronenberg, Xiao-Fen Chen, Trent Su, Fides D. Lay, Kathrin Plath, Siavash K. Kurdistani, and Michael F. Carey. Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. *Molecular Cell*, 73(2):250–263.e5, January 2019. ISSN 1097-2765. doi: 10.1016/j.molcel.2018.10.039. URL https://www.sciencedirect.com/science/article/pii/S109727651830933X.

[77] David M. Suter, Nacho Molina, David Gatfield, Kim Schneider, Ueli Schibler, and Felix Naef. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science*, 332(6028):472–474, April 2011. doi: 10.1126/science.1198817. URL http://

www.science.org/doi/10.1126/science.1198817. Publisher: American Association for the Advancement of Science.

[78] Hidehisa Takahashi, Tari J. Parmely, Shigeo Sato, Chieri Tomomori-Sato, Charles A. S. Banks, Stephanie E. Kong, Henrietta Szutorisz, Selene K. Swanson, Skylar Martin-Brown, Michael P. Washburn, Laurence Florens, Chris W. Seidel, Chengqi Lin, Edwin R. Smith, Ali Shilatifard, Ronald C. Conaway, and Joan W. Conaway. Human Mediator Subunit MED26 Functions as a Docking Site for Transcription Elongation Factors. *Cell*, 146(1):92–104, July 2011. ISSN 0092-8674. doi: 10.1016/j.cell.2011.06.005. URL https://www.sciencedirect.com/science/article/pii/S0092867411006477.

[79] Katjana Tantale, Florian Mueller, Alja Kozulic-Pirher, Annick Lesne, Jean-Marc Victor, Marie-Cécile Robert, Serena Capozi, Racha Chouaib, Volker Bäcker, Julio Mateos-Langerak, Xavier Darzacq, Christophe Zimmer, Eugenia Basyuk, and Edouard Bertrand. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun*, 7(1):12248, July 2016. ISSN 2041-1723. doi: 10.1038/ncomms12248. URL https://www.nature.com/articles/ncomms12248. Number: 1 Publisher: Nature Publishing Group.

[80] Xinchen Teng, Margaret Dayhoff-Brannigan, Wen-Chih Cheng, Catherine E. Gilbert, Cierra N. Sing, Nicola L. Diny, Sarah J. Wheelan, Maitreya J. Dunham, Jef D. Boeke, Fernando J. Pineda, and J. Marie Hardwick. Genome-wide Consequences of Deleting Any Single Gene. *Molecular Cell*, 52(4):485–494, November 2013. ISSN 1097-2765. doi: 10.1016/j.molcel.2013.09.026. URL https://www.cell.com/molecular-cell/abstract/S1097-2765(13)00748-X. Publisher: Elsevier.

[81] Xuhui Tong, Rong Tang, Jin Xu, Wei Wang, Yingjun Zhao, Xianjun Yu, and Si Shi. Liquid–liquid phase separation in tumor biology. *Sig Transduct Target Ther*, 7(1):1–22, July 2022. ISSN 2059-3635. doi: 10.1038/s41392-022-01076-x. URL https://www.nature.com/articles/s41392-022-01076-x. Publisher: Nature Publishing Group.

[82] Jason P Tourigny, Kenny Schumacher, Moustafa M Saleh, Didier Devys, and Gabriel E Zentner. Architectural Mediator subunits are differentially essential for global transcription in Saccharomyces cerevisiae. *Genetics*, 217(3):iyaa042, March 2021. ISSN 1943-2631. doi: 10.1093/genetics/iyaa042. URL https://doi.org/10.1093/genetics/iyaa042.

[83] Kuang-Lei Tsai, Chieri Tomomori-Sato, Shigeo Sato, Ronald C. Conaway, Joan W. Conaway, and Francisco J. Asturias. Subunit architecture and functional modular rearrangements of the transcriptional Mediator complex. *Cell*, 157(6):1430–1444, June 2014. ISSN 0092-8674. doi: 10.1016/j.cell.2014.05.015. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4104964/.

[84] Christopher M. Uyehara and Effie Apostolou. 3D enhancer-promoter interactions and multi-connected hubs: Organizational principles and functional roles. *Cell Reports*, 42(4):112068, April 2023. ISSN 22111247. doi: 10.1016/j.celrep.2023.112068. URL https://linkinghub.elsevier.com/retrieve/pii/S2211124723000797.

[85] Trung Nghia Vu, Quin F. Wills, Krishna R. Kalari, Nifang Niu, Liewei Wang, Matthias Rantalainen, and Yudi Pawitan. Beta-Poisson model for single-cell RNA-seq data analyses. *Bioinformatics*, 32(14):2128–2135, July 2016. ISSN 1367-4803. doi: 10.1093/bioinformatics/btw202. URL https://doi.org/10.1093/bioinformatics/btw202.

[86] Bin Wang, Lei Zhang, Tong Dai, Ziran Qin, Huasong Lu, Long Zhang, and Fangfang Zhou. Liquid–liquid phase separation in human health and diseases. *Sig Transduct Target Ther*, 6(1):1–16, August 2021. ISSN 2059-3635. doi: 10.1038/s41392-021-00678-1. URL https://www.nature.com/articles/s41392-021-00678-1. Publisher: Nature Publishing Group.

[87] Gang Wang, Michael A. Balamotis, Jennitte L. Stevens, Yuki Yamaguchi, Hiroshi Handa, and Arnold J. Berk. Mediator Requirement for Both Recruitment and Postrecruitment Steps in Transcription Initiation. *Molecular Cell*, 17(5):683–694, March 2005. ISSN 1097-2765. doi: 10.1016/j.molcel.2005.02.010. URL https://www.sciencedirect.com/science/article/pii/S1097276505010877.

[88] Wei Wang, Xiao Yao, Yan Huang, Xiangming Hu, Runzhong Liu, Dongming Hou, Ruichuan Chen, and Gang Wang. Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. *Transcription*, 4(1):39–51, January 2013. ISSN 2154-1264. doi: 10.4161/trns.22874. URL https://doi.org/10.4161/trns.22874. Publisher: Taylor & Francis _eprint: https://doi.org/10.4161/trns.22874.

[89] Warren A. Whyte, David A. Orlando, Denes Hnisz, Brian J. Abraham, Charles Y. Lin, Michael H. Kagey, Peter B. Rahl, Tong Ihn Lee, and Richard A. Young. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell*, 153(2):307–319, April 2013. ISSN 0092-8674. doi: 10.1016/j.cell.2013.03.035. URL https://www.sciencedirect.com/science/article/pii/S0092867413003929.

[90] Quin F Wills, Kenneth J Livak, Alex J Tipping, Tariq Enver, Andrew J Goldson, Darren W Sexton, and Chris Holmes. Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. *Nat Biotechnol*, 31(8):748–752, August 2013. ISSN 1087-0156, 1546-1696. doi: 10.1038/nbt.2642. URL https://www.nature.com/articles/nbt.2642.

[91] Shwu-Yuan Wu, Tianyuan Zhou, and Cheng-Ming Chiang. Human Mediator Enhances Activator-Facilitated Recruitment of RNA Polymerase II and Promoter Recognition by TATA-Binding Protein (TBP) Independently of TBP-Associated Factors. *Mol Cell Biol*, 23(17):6229–6242, September 2003. ISSN 0270-7306. doi: 10.1128/MCB.23.17.6229-6242.2003. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC180944/.

[92] Shiqi Xie, Jialei Duan, Boxun Li, Pei Zhou, and Gary C. Hon. Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Molecular Cell*, 66 (2):285–299.e5, April 2017. ISSN 1097-2765. doi: 10.1016/j.molcel.2017.03.007. URL https://www.cell.com/molecular-cell/abstract/S1097-2765(17)30174-0. Publisher: Elsevier.

[93] Jin H. Yang and Anders S. Hansen. Enhancer selectivity in space and time: from enhancer–promoter interactions to promoter activation. *Nat Rev Mol Cell Biol*, pages 1–18, February 2024. ISSN 1471-0080. doi: 10.1038/s41580-024-00710-6. URL https://www.nature.com/articles/s41580-024-00710-6. Publisher: Nature Publishing Group.

[94] Marty G Yang, Emi Ling, Christopher J Cowley, Michael E Greenberg, and Thomas Vierbuchen. Characterization of sequence determinants of enhancer function using natural genetic variation. *eLife*, 11:e76500, August 2022. ISSN 2050-084X. doi: 10.7554/eLife.76500. URL https://doi.org/10.7554/eLife.76500. Publisher: eLife Sciences Publications, Ltd.

[95] Aisha Yesbolatova, Yuichiro Saito, Naomi Kitamoto, Hatsune Makino-Itou, Rieko Ajima, Risako Nakano, Hirofumi Nakaoka, Kosuke Fukui, Kanae Gamo, Yusuke Tominari, Haruki Takeuchi, Yumiko Saga, Ken-ichiro Hayashi, and Masato T. Kanemaki. The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat Commun*, 11(1):5701, November 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-19532-z. URL https://www.nature.com/articles/s41467-020-19532-z. Number: 1 Publisher: Nature Publishing Group.

[96] Natalya Yudkovsky, Jeffrey A. Ranish, and Steven Hahn. A transcription reinitiation intermediate that is stabilized by activator. *Nature*, 408(6809):225–229, November 2000. ISSN 1476-4687. doi: 10.1038/35041603. URL https://www.nature.com/articles/35041603. Publisher: Nature Publishing Group.

[97] Sharon Yunger, Liat Rosenfeld, Yuval Garini, and Yaron Shav-Tal. Single-allele analysis of transcription kinetics in living mammalian cells. *Nat Methods*, 7(8):631–633, August 2010. ISSN 1548-7091, 1548-7105. doi: 10.1038/nmeth.1482. URL https://www.nature.com/articles/nmeth.1482.

[98] Benjamin Zoller, Damien Nicolas, Nacho Molina, and Felix Naef. Structure of silent transcription intervals and noise characteristics of mammalian genes. *Molecular Systems Biology*, 11(7):823, July 2015. ISSN 1744-4292. doi: 10.15252/msb.20156257. URL https://www-embopress-org.proxy3.library.mcgill.ca/doi/full/10.15252/msb.20156257. Publisher: John Wiley & Sons, Ltd.

[99] Jessica Zuin, Gregory Roth, Yinxiu Zhan, Julie Cramard, Josef Redolfi, Ewa Piskadlo, Pia Mach, Mariya Kryzhanovska, Gergely Tihanyi, Hubertus Kohler, Mathias Eder, Christ Leemans, Bas van Steensel, Peter Meister, Sebastien Smallwood, and Luca Giorgetti. Nonlinear control of transcription through enhancer–promoter interactions. *Nature*, 604(7906):571–577, April 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-04570-y. URL https://www.nature.com/articles/s41586-022-04570-y. Number: 7906 Publisher: Nature Publishing Group.