UNIVERSITY OF CALIFORNIA SAN DIEGO

Molecular Mechanisms that Underpin Antigen Receptor Recombination: Exploring the Role of E2A and BRD4 on V(D)J Recombination in B Cells

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Biology

by

Tatiana Fourfouris

Committee in charge:

Professor Cornelis Murre, Chair
Professor Matthew Daugherty, Co-Chair
Professor Li-Fan Lu

2019

The Thesis of Tatiana Fourfouris is approved, and it is acceptable in quality and form for publication on microfilm electronically:

Co-Chair

Chair

University of California San Diego
2019

DEDICATION

To all those who found me worthy of their love and encouragement, I will forever try and make you proud.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Cornelis Murre for all his help and support as my Committee Chair and Research Supervisor/Mentor. His guidance and encouragement throughout the process has been invaluable.

I would also like to acknowledge Dr. Edgar Mauricio Barajas-Mora for enduring many questions over the course of my year at the Murre Lab. Without him, I would not have learned how to ask the right questions in order to conclude and write my thesis. His love for science and research has been an inspiration.

This thesis is coauthored with Fourfouris, Tatiana; Murre, Cornelis; & Barajas-Mora, Edgar-Mauricio. The thesis author was the primary author of this paper.

ABSTRACT OF THE THESIS

Molecular Mechanisms that Underpin Antigen Receptor Recombination: Exploring the Role of E2A and BRD4 on V(D)J Recombination
in
B Cells

by

Tatiana Fourfouris

Master of Science in Biology

University of California San Diego, 2019

Professor Cornelis Murre, Chair
Professor Matthew Daugherty, Co-Chair

A diverse antibody repertoire is the backbone of a successful adaptive immune response. This diversity is created in developing B cells by the rearrangement of the V (variable), D (diversity), and J (joining) gene segments, termed V(D)J recombination, to produce unique antigen specific B cell receptors (BCRs). V(D)J recombination is regulated by an ensemble of transcription factors, co-activators as well as architectural proteins. We know that certain V gene regions are more highly rearranged in wild type cells, yet the mechanism for this bias is not well known. We propose that the enhancer-associated E2A, p300 and BRD4 proteins influence the frequency and pattern of gene rearrangement. Using CRISPR/Cas9 technology, we have successfully deleted E2A as well as BRD4 to determine the effects of each mutation on the frequency and pattern of rearrangement in the immunoglobulin kappa locus (Igκ) locus of an inducible cell line. We have examined the consequences of these deletions on Vκ rearrangements. Notably, our results indicate that E2A and BRD4 are responsible for influencing the pattern and frequency of Vκ gene rearrangement. These

data indicate that E2A and BRD4 are key factors in modulating nuclear architecture and the Igκ repertoire.

INTRODUCTION

The adaptable immune system is made possible by the generation of antigen receptors, including the BCR and TCR. The BCR, secreted as antibodies, is a Y-shaped transmembrane protein that recognizes and binds antigens. BCRs are composed of two immunoglobulin heavy chains (IgH) and two immunoglobulin light chains (IgL). Each pair is identical and includes a constant region at its C-terminal and a variable region at its N-terminal, oriented with the V regions pointing away from the body of the cell. The V region is responsible for binding antigens, and thus a crucial region for repertoire diversity. Composed of a rearranged sequence of DNA, antigen receptor (AgR) loci undergo what is known as V(D)J recombination; individual genes from the V, D, J, and constant (C) gene segments are rearranged to produce a single variant in a pool of billions.

With the help of recombination activation gene-1 (RAG1) and RAG2 working as a complex, the proteins introduce DNA breaks at recombination signal sequences to promote the joining of V, D and J gene segments. The cell's DNA repair machinery mends the breaks, piecing the DNA together to form a recombined segment. Antigen receptor locus rearrangement is dictated by developmental stages of B cell precursors, the Igh locus rearranges first followed by the IgL (Igκ and Igλ). Rearrangement of the IgH chain includes D_H-J_H rearrangement (pre-pro-B cell) followed by second V_H-DJ_H rearrangement. IgL undergoes one V-J rearrangement due to absence of a D region. Each chain is tested for auto-reactivity and must pass through multiple checkpoints prior to expression on the cell surface. A rearranged gene product may include sequences that are separated from each other by vast genomic distances, with the Vκ region alone spanning approximately 2.5-3 Mb. In order to facilitate such long-range interactions, the Igκ locus has been shown to undergo a form of DNA locus contraction, bringing the V genes in close spatial proximity with the J genes (Fuxa et al. 2004). Studies have shown that the rearrangement of the light chain's variable and joining segments (V_L-J_L) occurs at varying levels along the loci and can show a bias for certain genes, creating a pattern of rearrangement frequency and gene usage (Aoki-Ota et al. 2012; Kleiman, Xu, and Feeney 2018). However, the mechanism responsible for the pattern is not clear.

The necessity for accessibility between the Ig loci and recombination machinery was first proposed by researchers who noticed the lineage specific and stage-dependent non-coding

transcription driven by promoters in the V region and regulatory region of the H and L chains (Abarrategui and Krangel 2009; de Almeida, Hendriks, and Stadhouders 2015; Yancopoulos and Alt 1985). Prior to rearrangement, non-coding transcription, or germ-line transcription (GLT), occurs across Vκ and Jκ-Cκ gene segments, with much higher levels over Jκ-Cκ than Vκ segments, and this is driven by Igκ promoters and two germ-line transcription promoters (K^o) (de Almeida et al. 2015). GLT acts in part to "open" chromatin regions by recruitment of transcription factors and demethylation, allowing RAG proteins to access the recombination signal sequences (Abarrategui and Krangel 2006). Speculation of the function of non-coding RNA transcripts themselves in B cell development and V(D)J recombination has increased now that non-coding transcripts have proven important across many biological sub-fields (Murre 2019; Verma-Gaur et al. 2012). Overall, there is strong evidence in the field that GLT over the Jκ-Cκ region mediates Ig gene rearrangement by controlling chromatin accessibility (Sleckman and Oltz 2012).

Chromatin accessibility will be an important part of our hypothesis, however studies point to other factors in addition to chromatin structure that influence Ig rearrangement. Prominent amongst these are different RSS signal strength and transcription factor (TF) binding (Feeney 2009; Murre 2019; Stadhouders et al. 2014). Accessibility of the locus is associated with a large number of factors including locus compaction, acetylation and methylation of histones, DNA demethylation, and cis-regulatory elements like promoters, enhancers, and silencers (de Almeida et al. 2015).

Previous studies revealed that a transcriptional regulator, named Pax5, is particularly important in promoting Igh locus chromatin accessibility. Genes that are distally far away when the locus is linear are thought to undergo 3D locus restructuring to achieve a condensed interactive state (Fuxa et al. 2004). Fuxa et al. were able to illustrate the 3D DNA contraction and looping event that precedes rearrangement in the Igh and Igκ loci. Contraction refers to a specific fold that brings the distal region in proximity with the regulatory region, causing the two genomic domains to interact (Fuxa et al. 2004). It is now generally agreed upon that locus contraction and local looping is required to generate a diverse set of Vκ-Jκ sequences.

With consideration of enhancer-promoter regulation, insulator activity must also be considered. The effects of CCCTC-binding factor (CTCF), a transcriptional repressor protein, have

been shown to effect 3D folding of the genome by regulating the range of promoter and enhancer influence. Typically flanking topologically associated domains (TADs) and sub-TADs in a convergent orientation (Bastiaan Holwerda and de Laat 2013; Dixon et al. 2012; Rao et al. 2014; Tang et al. 2015; de Wit et al. 2015), CTCF sites monitor the effects of strong enhancer elements and control boundaries between active and inactive chromatin regions (de Almeida et al. 2015; Bastiaan Holwerda and de Laat 2013). CTCF sites between the Vκ and Jκ gene regions have been studied in the context of Igκ rearrangement, and found to inhibit the cell's preference for rearranging Vκ genes proximal to the Jκ regulatory region (Kleiman et al. 2018; Xiang, Park, and Garrard 2014), as well as to control the mechanism of locus contraction, both promoting the distribution of rearrangement activity throughout the locus (de Almeida et al. 2015; Choi and Feeney 2014; Fuxa et al. 2004; Jhunjhunwala et al. 2009).

A key element regulating Igκ locus rearrangement involves the intronic κ enhancer (iEκ). Located in the regulatory region of the Igκ locus, iEκ provided some insight into the role of enhancers on rearrangement frequency at Vκ genes (de Almeida et al. 2015). Although it was shown that iEκ itself is not necessary for rearrangement, when deleted along with another major regulatory enhancer element (3'Eκ), Igκ rearrangement was essentially eliminated (Inlay et al. 2002). In vivo studies have also supported this idea by illustrating that the deletion of iEκ leaves the cell with an overall decrease in locus demethylation, GLT from K^0, and Igκ rearrangement (Gorman et al. 1996; Inlay et al. 2002; Kleiman et al. 2018). Although several publications identified the role of iEκ in demethylation, promoter activity, and GLT (Inlay et al. 2002; Lichtenstein et al. 1994), the mechanism correlating it with rearrangement has not been fully elucidated.

LRCIs are thought to be mediated by enhancer regions throughout the locus (in addition to regulatory region enhancers) for their ability to attract specific protein activity in a tissue-specific manner, interacting with promoters throughout the region and compartmentalized by CTCF domains (Barajas-Mora et al. 2019; Murre 2019; Ong and Corces 2011). Studies have identified "hubs" of interaction across the Igκ locus in pro-B cells in correlation with increased occupancy of transcription factors like the E2A proteins (Lin et al. 2012).

The E2A proteins, also known as transcription factor 3 (tcf3), are members of the E-protein family of helix-loop-helix (H-L-H) proteins, and best known for their role in modulating enhancer activation. E2A proteins recognize E-box sites, a distinct sequence: CANNTG. The target E-box sites are found in many regions of the Ig locus, most notably in major enhancer elements like iEκ. E-proteins are known to remove methyl groups from CpG islands in enhancer regions and to recruit histone acetyltransferases like P300/CBP, or histone deacetylases (HDACs) across histone tails (Murre 2019; Murre, McCaw, and Baltimore 1989). E2A activity is closely coupled with other transcription factors like early B cell factor 1(EBF1) and paired box 5 (Pax5), both of which are associated with orchestrating early B cell development (de Almeida et al. 2015; Lin et al. 2010; O'Riordan and Grosschedl 1999).

The E2A proteins bind E-box sites located in the Igκ enhancer as well as many sites along the Vκ locus. They are frequently positioned within close proximity to V gene promoters and RSS sites (de Almeida et al. 2015; Bemark, Liberg, and Leanderson 1998; Brekke and Garrard 2004; Kleiman et al. 2018). Just as iEκ has been shown to regulate rearrangement, E2A knockouts demonstrate a similar decrease in rearrangement (Gorman et al. 1996; Inlay et al. 2002). Furthermore, recruitment and accumulation of E2A across the Igκ locus correlates well with DNA loop formation. Further data suggests that E2A mediates changes to chromatin topology before rearrangement and development can proceed (Lin et al. 2012). Just as its been shown that GLT provides chromatin marks over AgR genes for recombinase machinery to bind (Sleckman and Oltz 2012), we believe enhancers may act to tag chromatin in a similar and more specific manner. Enhancer activity, specifically mediated by E2A could act to tag DNA regions by causing the sequestration of transcription machinery to target gene sites, biasing the cells internal looping structure and thereby affecting rearrangement. Assessing the role of enhancer activity, specifically E2A binding, in regulating Vκ-Jκ rearrangement necessitates the consideration of potential downstream effects and their role in regulating chromatin topology and locus accessibility.

E2A occupancy is closely associated with P300 binding. Once recruited, P300 occupancy at E2A bound sites leads to acetylation of lysine residues on histone tails H3 and H4, a target for recognition by the chromatin reader bromodomain containing protein BRD4 (Murre 2019). BRD4 is

known to contain an intrinsically disordered region (IDR), comprised of repeat-sequences and low-amino acid complexity, often resulting in "blocks" of negative or positive charge (Banani et al. 2017). IDRs are unique because of the versatile nature of the protein they produce; intrinsically disordered proteins (IDPs) are inherently less rigid than other proteins, with a tendency for conformational disarray (Brangwynne, Tompa, and Pappu 2015). A study on the characteristics of proteins known to produce liquid coalescence and phase separation illuminated the importance of IDRs in the formation of weakly adhesive bonds, critical for phase separation, naming BRD4 as a prime example (Banani et al. 2017; Sabari et al. 2018).

To organize different biological processes, cells utilize a system of membrane-less compartmentalization, known as phase separation, to create a chemically favorable state for reactions to proceed without interruption (Brangwynne et al. 2015; Hyman, Weber, and Jülicher 2014; Strzyz 2018). This is made possible by a physiochemical boundary serving to encompass the reaction, and a physical state within the boundary where molecules can diffuse freely (Hyman et al. 2014). These fluid compartments are key contributors to the homeostatic conditions of the organism, and are thought to be closely coupled with chromatin remodeling and accessibility of the genome (Gibson et al. 2019). Although further studies are required it is conceivable that E2A and BRD4 promote phase separation across the Igκ locus as a result of E2A and BRD4 activity in the region.

In other words E2A, P300, and BRD4, might be involved in regulating V gene usage as a result of interactions with cis elements like enhancers, promoters and silencers throughout the locus by promoting a phase separated state. We believe these interactions regulate transcription events known create regions of "open" or active chromatin, providing rearrangement machinery with locus accessibility in a spatially controlled manner, and acting to flag target sites, influencing V gene rearrangement biases throughout the locus. Confinement of enhancer activity, dictated by sub-TAD regions, we believe may be a cause of long-range locus interactions and 3D locus looping driven by phase separation. In our experimental model, E2A influences target sites for BRD4 to bind and alter the 3D structure of the locus, creating looped domains by phase separation, and a natural pattern of gene usage as a result.

By using a differentiation-inducible cell line and CRISPR/Cas9 technology, we have examined the role of E2A and BRD4 in modulating Igκ locus rearrangement. We generated two deletions, the IDR region of BRD4, and the exons encoding the DNA binding motif of E2A. We next examined the impact of these deletions on V gene usage, changes to GLT, and chromatin structure. Using RNA-sequencing, we found that GLT in the Vκ locus was severely affected in E2A and BRD4 knockouts. In addition, we developed an assay to study Vκ-Jκ rearranged transcripts using deep sequencing technology to illustrate changes to gene usage patterns as a result of the deletions. We found that impaired E2A functioning caused an overall reduction in transcription of rearranged genes, and specifically causes the cells to skew the usage of V genes to the middle region of the Vκ locus and decreasing of Jκ-proximal gene usage. BRD4 knockouts demonstrated the opposite effect, decreasing the usage of the Vκ middle region and upregulating Jκ-proximal gene rearrangement. The changes to Vκ gene usage were confined in a domain-specific manner. Taken together, these data show that E2A and BRD4 are key factors in modulating Igκ locus rearrangement.

RESULTS

Determining CRISPR/Cas9 Target Deletion Sites

In order to understand the effects of E2A and BRD4 on GLT, rearrangement, and regulation of chromatin conformation, we used the CRISPR/Cas9 editing system to make targeted deletions to the genome of an Abl-transformed pro-B cell line (445.3-WT). Arrested between the pro-B cell and pre-B cell stage, this cell line can be stimulated to induce GLT and Igκ rearrangement with the use of the Abl kinase inhibitor (STI571). STI571 removes the block on differentiation and allows the cells to rearrange the Igκ locus (Muljo and Schlissel 2003). The 445.3 cell line was created on a Rag-/- background to eliminate any possibility of preemptive rearrangement affecting the integrity of the cells germline Igκ locus, a key feature for our assessment of changes to GLT and chromatin interaction and structure.

To effectively eliminate the activity of E12 and E47 transcription factors encoded by the E2A gene, we chose to target the exons responsible for DNA binding. E12 and E47 differ only in binding regions as a result of E2A gene splicing variants; two versions of exon 18/19 produces the

two variants (Figure 1A). We created a deletion spanning ~2.7kb in order to eliminate the entire region containing the two exon 18s (Figure 1). In order to address the phase separating abilities of the BRD4 protein, we designed guides to target a ~2kb intrinsic disordered region of the protein, identified through the use of the http://www.pondr.com website, a predictor of a protein or DNA sequences' natural disorder, and literature previously published depicting the intrinsic disordered region of the BRD4 protein (Sabari et al. 2018). For each deletion we generated two bi-allelic deletion clones referred to from hereinafter as 445.3-E2AΔ and 445.3-BRD4Δ (Figure 1).

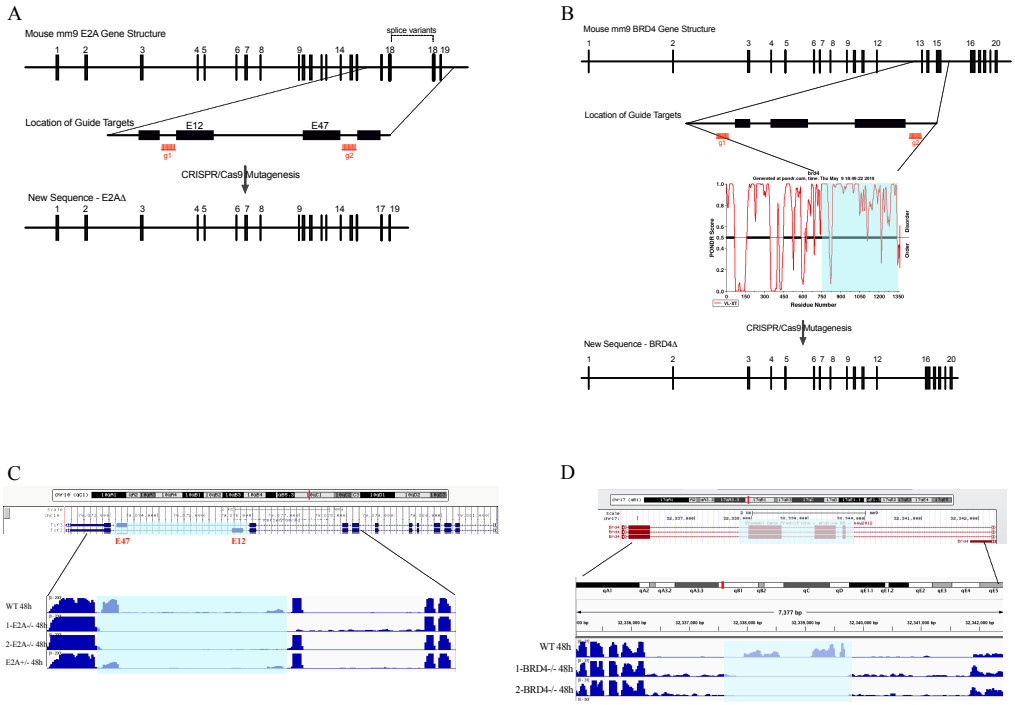

Figure 1 Generating E2A and BRD4 KO clones. (**A**) Diagrams the E2A locus (line 1), the location of the guide RNA targets labeled as g1 and g2 (line 2), and the structure of the gene following CRISPR/Cas9 editing (line 3). Exons are depicted as vertical lines, numbered above, and shown as 5' to 3' from right to left. (**B**) Diagrams the BRD4 locus (line 1), the location of the guide RNA target sites labeled with g1 and g2 (line 2), the corresponding predictor of natural disorder (PONDR) score of the region on line 3 (image taken from pondr.com), and the subsequent structure of the locus following CRISPR/Cas9 editing. (**C**) Data from RNA-seq track files illustrating the loss of transcription to the deletion region in the two E2A knockout clones compared to the transcription in the wild-type (WT) cells and one E2A heterozygous deletion clone at 48 hours following addition of STI571. (**D**) This image depicts the loss of transcription over the BRD4 deletion region in two BRD4 knockout clones compared to the WT at 48 hours post stimulation. Both C and D were created using images from UCSC's online genome browser, and RNA-seq results mapped in IGV 2.6.2.

E2A and BRD4 Cause Global Changes to the Transcriptome

There is long-standing evidence for correlations involving GLT, locus accessibility, and rearrangement. We began with RNA-Sequencing (RNA-seq) on the 445.3-WT, 445.3-BRD4Δ, and

445.3-E2AΔ cells at 0 hours and 48 hours following stimulation with STI571. Data from two independent knockout clones as well as two independent 445.3-WT clones analyzed. We hypothesized that there would be global changes to transcription levels due to our E2A and BRD4 deletions. We further hypothesized that there would be a specific decrease in Vκ gene transcription as a result of perturbed E2A binding at the regulatory region enhancers (iEκ and 3'Eκ), similar to the effects seen in other studies (de Almeida et al. 2015; Kleiman et al. 2018). When compared to data from 0 hours, we found that 445.3-E2AΔ clones showed marked differences compared to WT cells in transcription levels after 48 hours of stimulation with STI571, indicating their impact on developmentally regulated transcription in association with the activation of rearrangement (Figure 2A). Figures 2B and 2C are showing the distribution of fold change at 0 hours and at 48 hours individually for 445.3-E2AΔ compared to 445.3-WT.

We found that there seems to be more genes showing a reduction in expression upon E2A depletion 48 hours post stimulation. Using differential expression and gene ontology analyses, we identified distinct gene sets showing significantly dysregulated transcription in 445.3-E2AΔ. Using http://www.pantherdb.org, genes showing higher than 2-fold reduction in 445.3-E2AΔ when compared with 445.3-WT cells were sorted based on protein class, biological process, and molecular functioning. Calculated as percent of gene hit against total number of class hits, down-regulated genes in 445.3-E2AΔ cells were classified: 25.7% characterized as nucleic acid binding proteins (sorted by protein class), 41% associated with binding overall (sorted by biological function), and 36.5% affecting cellular processes (sorted by molecular function) (Figure 2D-2E).

445.3-BRD4Δ cells showed similar results, with highly affected transcription levels 48 hours post stimulation (Figure 3A). We found that at both time points, there were a large number of genes that were significantly affected, with a higher number of dysregulated genes after 48 hours (Figure 3A-C). Using the same tools described above, we found that 33.2% of transcripts that were more than 2-fold reduced in our mutant BRD4 cells were associated with nucleic acid binding proteins, 44.9% relating to binding overall, and 38.2% affecting cellular processes (Figure 3B-D). Notably, the 445.3-BRD4Δ cells had a higher percentage of down-regulated transcripts associated with the defense/immunity protein classes and nucleic acid binding than did the 445.3-E2AΔ cells. On the

other hand, 445.3-E2AΔ cells showed a greater percentage of decreased transcripts involved in transcription regulator activity when compared with 445.3-BRD4Δ cells. Notably, we observed a fairly substantial and uniform decrease in all but two histone encoding gene transcripts in 445.3-BRD4Δ cells. Overall, the RNA-seq data confirmed the expected: the deletions would affect overall cell functioning with special regard to transcription and chromatin.

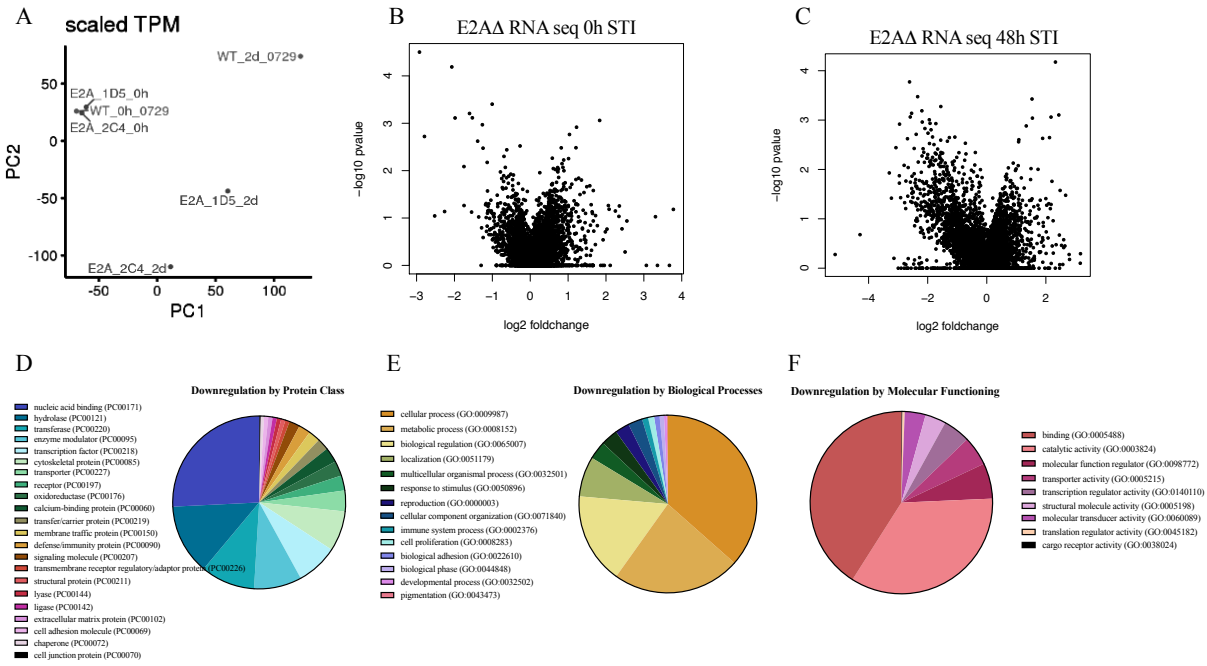

Figure 2 Dysregulation of global gene transcription levels in 445.3-E2AΔ cells. (**A**) Principle component analysis illustrating RNA seq data similarities between wild type cells and mutant cells at 0h and 48h after stimulation by STI571. Volcano plots illustrating distribution of fold change between 445.3-E2AΔ cells and 445.3-WT at time 0 hours (**B**) and time 48 hours (**C**). For all pie charts, gene ontology website pantherdb.org was used to generate gene groupings based on RNA-seq reads from transcripts showing greater than 2-fold reduction (mutant/wild type) at 48 hours following STI treatment. Data from this site was calculated as percent of gene hit against total number of function hits; grouping genes with lower transcription rates by protein class (**D**), associated biological processes (**E**), and molecular function (**F**).

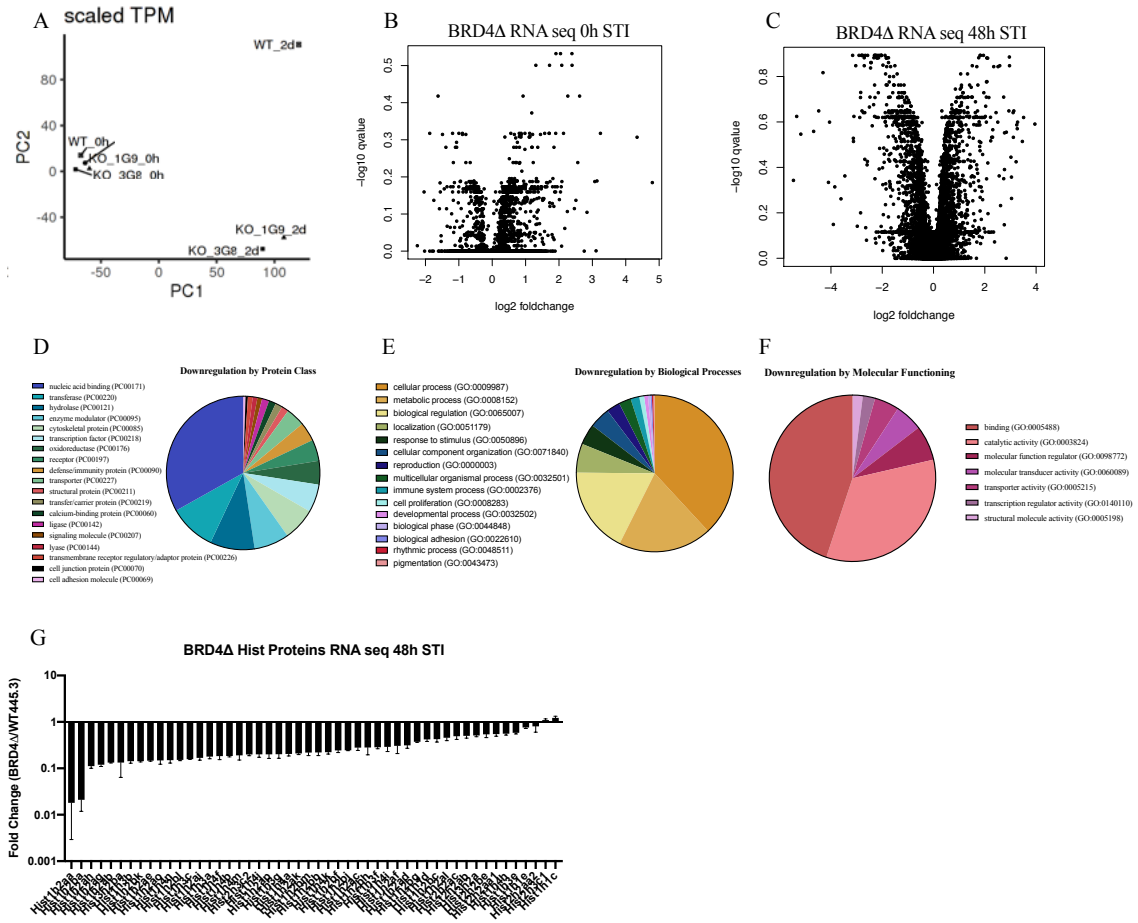

Figure 3 Dysregulation of global transcription levels in 445.3-BRD4Δ cells. (**A**) Principle component analysis illustrating RNA-seq data similarities between wild type cells and mutant cells at 0h and 48h after stimulation by STI571. Volcano plots illustrating distribution of fold change between 445.3-BRD4Δ cells and 445.3-WT at time 0 hours (**B**) and time 48 hours (**C**). For all pie charts, gene ontology website pantherdb.org was used to generate gene groupings based on RNA-seq reads from transcripts showing greater than 2-fold reduction (mutant/wild type) following STI treatment. Data from this site was calculated as percent of gene hit against total number of function hits; grouping genes with lower transcription rates by protein class (**D**), associated biological processes (**E**), and molecular function (**F**). (**G**) Bar graph showing impressive fold reduction in transcripts that encode for histone proteins. Error bars reported as standard error of the mean (SEM).

E2A and BRD4 affect Igκ Locus GLT

Next, we decided to take a look at GLT specifically in the Igκ locus and its correlation with E2A binding, taken from ChIP data already published (Barajas-Mora et al. 2019). We observed locus-wide changes to transcription levels in the Vκ region in both 445.3-BRD4Δ and 445.3-E2AΔ cells as compared to 445.3-WT cells. 445.3-E2AΔ cells showed decreased consistency in pattern of GLT variance in Jκ-distal regions and a notable decrease in GLT when approaching the regulatory region. On the other hand, 445.3-BRD4Δ cells showed the opposite pattern with genes transcribing more

frequently in areas proximal to the regulatory region, and less so in the distal region. In accordance with the literature, a decrease in E2A binding in the iEκ region would cause decrease functioning of the enhancer, which is known to affect transcription over proximal genes to the regulatory region (de Almeida et al. 2015).

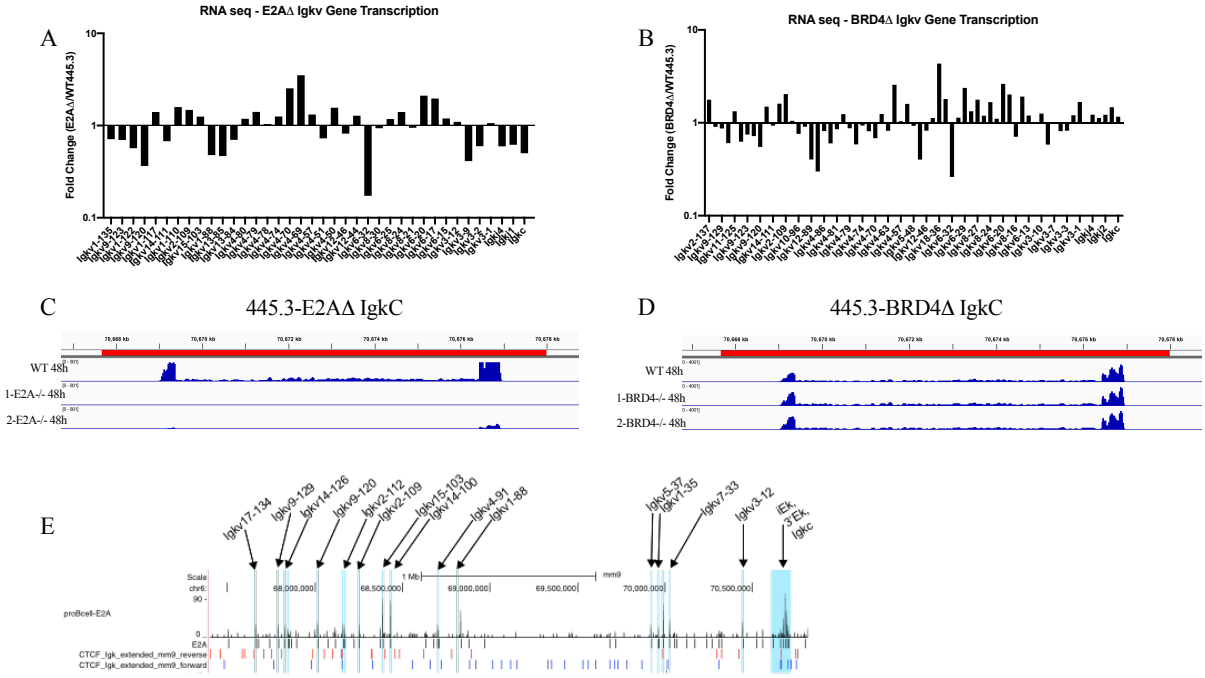

Figure 4 Changes to GLT in the Vκ locus in 445.3-E2AΔ and 445.3-BRD4Δ cells. Bar graphs illustrate mean as a ratio between the average of two mutant clones and two wildtypes, calculated from the percentage of reads per gene over total reads. Error bars are reported as SEM and the locus is depicted geographically following the gene order of the locus, distal (left) to proximal (right) in relation to the regulatory region. (**A**) Bars are depicting transcription changes over annotated genes in the Igκ locus in 445.3-E2AΔ cells. (**B**) Bars are depicting transcription changes over annotated genes in the Igκ locus in 445.3-BRD4Δ cells. (**C**) This is an image of the RNA-seq track files for 445.3-WT and two 445.3-E2AΔ -/- clones illustrating the depletion of transcription over the IgκC region. (**D**) This is an image of the RNA-seq track files for 445.3-WT and two 445.3-BRD4Δ -/- clones illustrating that there is no change to transcription over the IgκC region as a result of the deletion.

E2A and BRD4 Impacts Igκ Rearrangement

Following assessment of global transcription and GLT in the Igκ locus, we used a retrovirus expressing Rag1 to transduce our Rag-/- pro-B cell cell populations: 445.3-WT, 445.3-BRD4Δ, and 445.3-E2AΔ cells. After allowing enough time for the virus to incorporate, we add our differentiation stimulant, STI571, to induce rearrangement. Using a Vκall primer designed to bind and amplify a large array of rearranged gene products, we performed quantitative PCR (qPCR) on 445.3- BRD4Δ and 445.3-E2AΔ cells. We quantified the levels of rearrangement as fold change between mutant and

wild type cells, thereby assessing whether or not we see a substantial change to total rearrangement before further investigation of individual gene frequencies to determine usage patterns. 445.3-E2AΔ was found to have an overall reduction in Vκ gene rearrangement when compared to wild type cells, and 445.3- BRD4Δ an overall increase in Vκ gene rearrangement. Fold changes were almost identical in 445.3-BRD4Δ cells for cDNA and gDNA showing an ~2 fold induction in overall rearrangement, while the 445.3-E2A cells showed a greater fold reduction in amplified rearranged gene products, cDNA, than in gremline rearrangement, measured as gDNA for the same cells (Figure 5E).

We next evaluated individual Vκ gene rearrangement frequencies in our wild type and mutant cell lines. We note that the Vκ-Jκ rearrangement assay was developed using the same Vκall primer mentioned above. This primer amplifies a large number of rearranged gene products, which we utilized to create a PCR library of the repertoire. Illumina tru-seq adapter sequences were added to the individual libraries and pooled for next generation sequencing. Reads for both 445.3-BRD4Δ and 445.3-E2AΔ cell populations were analyzed to determine the percent gene usage, calculated as a function of reads per gene over total number of reads, and then compared with that of 445.3-WT cells to determine fold change.

We found that gene usage was affected in a domain-like manner in both 445.3-E2AΔ, and 445.3-BRD4Δ cells. Bound by sub-TADs in the region (Barajas-Mora et al. 2019), the bar graphs below indicate fold change plotted on a $\log_{10}$ scale when compared with wild type. 445.3-E2AΔ cells seemed to favor rearrangement of genes located in the middle region of the Vκ locus, while the genes most proximal to the regulatory region showed a severe decrease in gene usage when compared to 445.3-WT (Figure 5A). Different domains showed varying consistency in gene usage pattern, with some regions showing smaller sub domains, like that of the blue region in Figure 5A where the pattern seems to switch at Igκ6-15. Mann-Whitney plots allowed us to determine the significance of gene usage between domains, showing significant changes between the Jκ-proximal domains but not between the distal groups (Figure 5B). On the other hand, 445.3-BRD4Δ cells showed less variance in gene usage patterns within domains, like that of the most Jκ-proximal region denoted in “color” where almost all the genes seem to show increased rearrangement when compared to wild type cells (Figure 5D). Exhibiting the opposite pattern in gene rearrangement frequency, BRD4 showed a striking loss

in rearrangement of a large portion of middle genes of the locus, with increased variance approaching the distal end. The Mann-Whitney plot shows a similar trend as the 445.3-E2AΔ cells, with the proximal groups showing a significant difference between domains that are absent in the distal regions. In sum, E2A and BRD4 play an important role in regulating Vκ gene rearrangement, established in their ability to manipulate the locus in a domain-specific manner.

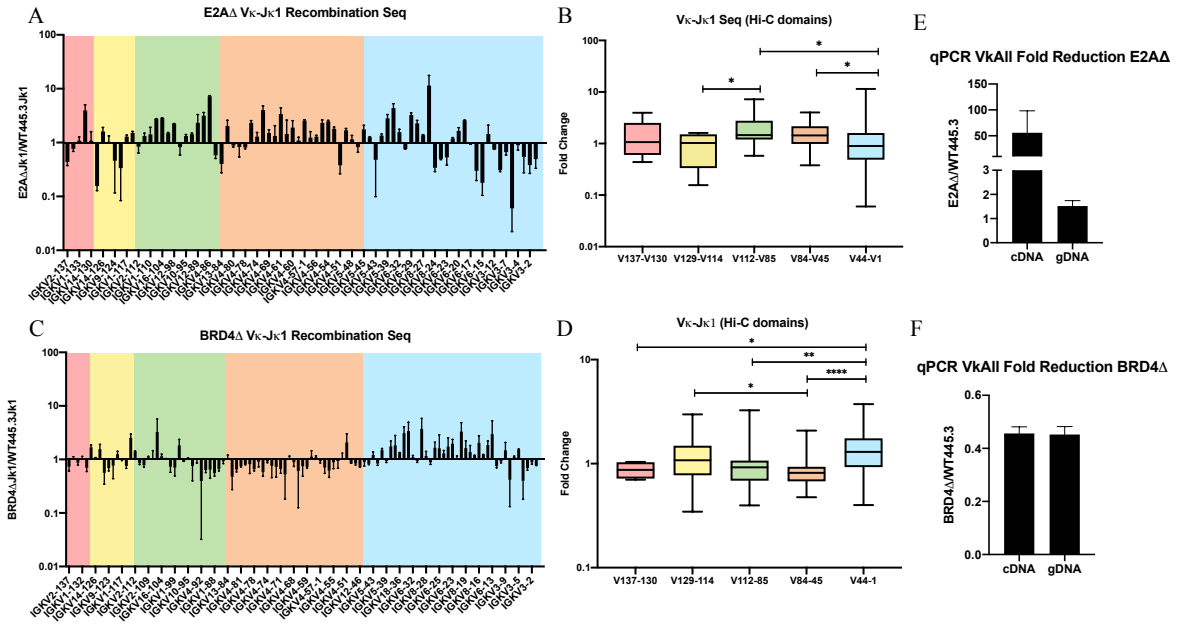

Figure 5 Changes to Vκ gene rearrangement along the locus of 445.3-BRD4Δ and 445.3-E2AΔ cells. (**A**) This bar graph shows the ratio of sequencing reads from 445.3-E2AΔ cells over reads from 445.3-WT cells. Fold change is plotted on a $\log_{10}$ scale and includes data from two biological replicates for both KO and WT collected 48 hours after stimulation with STI571. Includes data from 2 biological replicates for both KO and WT. (**B**) Mann-Whitney test plot of 5 domains in 5A bounded by CTCF sites and characterized by Hi-C in previous studies*. (**C**) qPCR data illustrating fold reduction from cDNA and gDNA extracts after 48h STI stimulation with a VκAll primer was used to determine overall rearrangement in the locus. (**D**) This bar graph shows the ratio of sequencing reads from 445.3-BRD4Δ cells over reads from 445.3-WT cells. The fold change is plotted on a $\log_{10}$ scale and includes data from two biological replicates for both KO and WT. (**E**) Mann-Whitney test plot of domains in 5D bounded by CTCF sites and characterized by Hi-C in previous studies*. (**F**) qPCR data illustrating fold reduction from 445.3- BRD4Δ cDNA and gDNA extracts after 48h STI stimulation with a VκAll primer was used to determine overall rearrangement in the locus. All error bars are reported as SEM in this figure. *(Barajas-Mora et al. 2019)

Perturbed Phase Separation Affects the Pattern of Vκ Gene Usage

We hypothesized that 445.3-BRD4Δ would stunt phase separation in the Igκ locus, making it more difficult for the cell to make long-range chromatin interactions and thereby decreasing Vκ rearrangement of distal Vκ genes as a result. To take a more direct look at the effects of phase separation on the frequency of Vκ gene rearrangement, we decided to run the same Vκ-Jκ assay on

cells treated with compounds known to disrupt cellular phase separation, sodium chloride (NaCl) and ammonium acetate (AmAc) (Jain and Vale 2017; Sabari et al. 2018). Cells were treated with 100mM of NaCl or AmAc at 3 hours after the addition of STI571. At 6 hours of stimulation (3 hours after addition of treatment), the cells were washed and plated with fresh media. After a total of 24 hours stimulation, the cells were harvested for RNA and gDNA. cDNA was created from RNA yields and Vκ-Jκ rearrangement sequencing was performed as above described. Our results indicate that AmAc and NaCl treatment disrupt Vκ rearrangement in different ways, reflected in their contrasting frequency patterns. AmAc shows similarities to the pattern of gene usage seen in 445.3-E2AΔ cells in its upregulation of middle region genes, while NaCl seem to only resemble 445.3-BRD4Δ cells in the minor downregulation of middle region rearrangement. However, we note that the data from these treatment assays does not depict the same striking domain groupings as seen in the E2A or BRD4 deletions. Taken together, these observations indicate that reagents known to affect phase separation affect Igκ locus rearrangement.

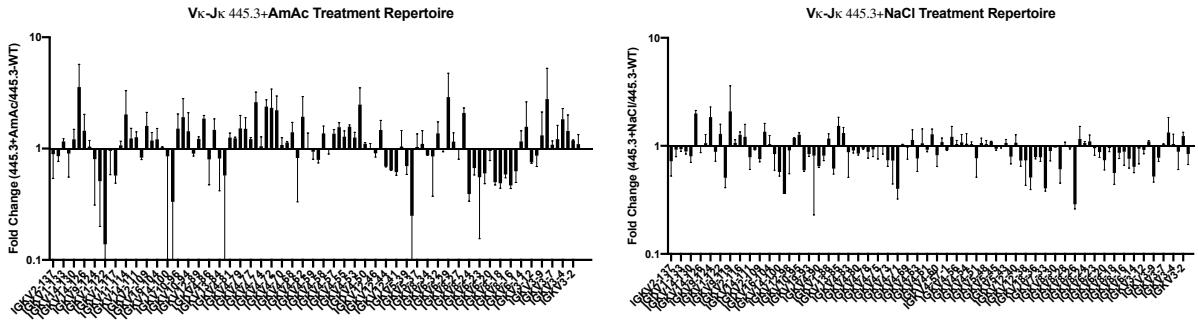

Fig. 6 Disruption of phase separation alters the natural Vκ-Jκ repertoire. All error bars are reported as SEM. Wild-type cells treated with AmAc (left) and NaCl (right) were assayed for Vκ rearrangement using the same protocol described as Vκ-Jκ seq in figure 5. Bars are plotted as mean fold change on a $\log_{10}$ scale using two biological replicates for both treated and wild type cells; illustrating the ratio of gene usage, calculated as percent reads per gene over total reads for the sample and compared between populations.

DISCUSSION

Using CRISPR/Cas9 selective knockouts in E2A and BRD4 gene regions, we disrupted E2A binding and BRD4 phase separation activity in the Igκ locus. We confirmed that not only are these genes crucial for normal cellular functioning, but hindering their functionality specifically translates to altered germline transcription and rearrangement of the Igκ locus. Our data supports findings that

enhancer activity throughout the locus, not only in the regulatory region (e.g. iEκ) act to control V gene usage during rearrangement activation at the pro-B and pre-B cell stage (Barajas-Mora et al. 2019). In addition, our data suggests that locus-wide enhancer binding causes digression of the natural pattern of rearrangement with effects contained to domains currently thought to be bounded by CTCF boundary sites. We believe that enhancer activation and recruitment of transcription machinery like E2A and BRD4 to key sites throughout the locus acts to bias recombinase activity in a domain specific manner.

The similarity between 445.3-E2AΔ and 445.3-BRD4Δ cells in affected gene groups illustrated by RNA-seq and gene ontology analysis indicates that they are involved in the same or similar regulatory processes and pathways (Figure 2). It is now well established that the E2A proteins are associated with enhancer marks and recruits P300 to recruit BRD4 (Murre 2019). Our deletion of the DNA-binding region of E2A, and the IDR region of BRD4 was meant to disrupt this process and reveal at what stage, if any, rearrangement would be affected. From this we would be able to ascertain to what extent certain mechanisms might be involved in regulating frequency of gene usage in the process of V(D)J recombination. By RNA-seq, we found that GLT is affected in the Vκ locus as well. Overall, it is clear that there are similarities in affected regions between E2A and BRD4 knockouts, like in the areas surrounding genes IGKV6-32, IGKV4-86, and IGKV9-120 that show approximately 3-5 fold reduction in both knockouts when compared to wild type. However, when it comes to domains, the knockouts seem to show trends following that of rearrangement in each corresponding gene region. Hence we conclude that both E2A and BRD4 work to influence rearrangement of certain genomic regions.

However, it was clear that there is a distinction in the mechanism behind their individual roles as seen in the opposing fold change patterns; E2A showing increases in the middle region and decreases in the Jκ-proximal region, and BRD4 showing decreases in the middle region and increases in the Jκ-proximal region (Figure 5). With the evidence that Vκ GLT and rearrangement is affected in similar regions in both E2A and BRD4 knockouts, we note that enhancer driven transcription may be involved in influencing rearrangement by domain, and not simply in the regulation of indivdual genes.

As described above, the deletions in E2A and BRD4 caused opposite phenotypic patterns, illustrating domains that are dependent on successful E2A and BRD4 binding for normal rearrangement levels in the region. Taking a closer look at rearrangement by domain, in 445.3-E2A cells we conclude that E2A binding is not required for the rearrangement of middle region genes seeing that there is no depletion of rearrangement in the region. CTCF sites and their corresponding transcriptional silencing effects have been shown to limit enhancer range (Kleiman et al. 2018; Ribeiro de Almeida et al. 2011), providing a mechanistic explanation for the containment of the impact of E2A onto certain regions of the genome. There also seem to be more fine-tuned effects that create sub-domains within gene regions, seen in their opposing trends when compared with the rest of the genes in their corresponding domain. An example of this can be seen in the proximal blue region in Figure 5A, where genes switch around IGKV8-24 from heightened to depleted rearrangement when approaching the regulatory region. In our 445.3-E2AΔ cells, we see an overall reduction in rearrangement of the Igκ locus quantified with qPCR. When using cDNA, the effect is much more pronounced than when using gDNA. This is most likely a result of the loss of function of the iEκ enhancer caused by the absence of E2A activity. The cells transcribe the rearranged products at a higher level in the WT cells than in the E2A knockouts as a direct result of the deletion, causing a bias in the assay and showing a higher reduction overall in rearrangement between populations. In reality, the relative difference in rearrangement is more accurately depicted in our gDNA qPCR data. The cell may act to compensate for the loss of rearrangement in E2A-dependent domains by creating the effect seen in Figure 5A where the middle most region of genes is up-regulated in rearrangement frequency when compared with wild type.

Our data also appears to conflict with previous literature indicating that E2A depletion causes a complete loss of rearrangement in the Igκ region. A study using a similar Abelson transformed cell line and an E2A deletion of exons 17 and 18 (one additional exon than our deletion), showed impaired GLT in the Igκ region as well as complete knock out of rearrangement. The details of their assay were as described here except for the use of signal break ends (SBE) to determine levels of rearrangement in mutant and wild-type cells. The complete absence of detectable SBE led to the conclusion that E2A is absolutely essential for the proper rearrangement of the Igκ locus. They were able to rescue the

effect of their deletion on transcription and rearrangement with the ectopic expression of E47. In addition, they found that E2A is involved in the hyper acetylation of intronic enhancer elements at the start of locus activation and rearrangement. They suggest that E2A works to modify chromatin landscapes by regulating acetylation, and thereby affecting rearrangement and locus accessibility. This is where our model may overlap, with the downstream effects of acetylation by E2A activity leading to BRD4 binding(Lazorchak, Schlissel, and Zhuang 2006). We believe a possible explanation for the discrepancy in results could be the difference in number of exon deletions, as Lazorchak et al., 2006 deleted one additional exon, exon 17, compared to our model.

To try and address the differences between the two phenotypes seen in our deletion clones, we further consider the possibility that phase separation may be acting to bias the cells rearrangement in a different manner than E2A mediated recruitment and sequestration of transcription machinery. Therefore, we treated wild type cells with compounds known to effectively disrupt phase separation, sodium chloride (NaCl) and ammonium acetate (AmAc) (Jain and Vale 2017). Considering the fact that our specific BRD4 deletion was designed to target the IDR region and therefore disrupt its phase separating abilities, we expected to see similar effects in both treatments as in the rearrangement pattern of 445.3-BRD4 cells. In Figure 6 we show the results of this assay; NaCl vaguely mimicking the effects of our BRD4 deletion, and AmAc slightly resembling the effects of our E2A deletion. Neither bearing the same striking domain-like effects as our deletions, we conclude that there must be a more specified mechanism by which E2A and BRD4 mediate V(D)J recombination involving specific enhancer activity and marker regions.

In summary, using CRISPR/Cas9 engineering we have generated B cell progenitors that are depleted for E2A and BRD4 expression and have examined the consequences of these deletions on Vκ rearrangements. The data reveal that E2A and BRD4 play crucial roles in modulating Igκ locus rearrangement and V-region gene usage.

METHODS

Cells

A Rag-/- pro-B cell line transformed with Abelson virus (445.3) was used for all in vitro experiments, referred to in this paper as 445.3-WT. RPMI media (Gibco™ 11875-093) was used with addition of 10% Corning Fetal Bovine Serum (FBS)(Corning 35-015-CV), 1% Pennicillin/Streptomycin (Gibco™ 10378016), 2mM L-glutamine (Gibco™ 25030081), and 0.05mM b-mercaptoethanol (Aldrich M6250) to maintain and culture the 445.3 cells, with exception of the 24 hours following Neon transfection in which the cells were given media containing all the above listed components except penicillin-streptomycin.

CRISPR-Cas9 Mutagenesis

The E2A deletion is a ~2.7kb deletion of exon 18/19 in E12/E47, the DNA sequence encoding the DNA binding region. The BRD4 deletion is ~2kb and covers the IDR region (as depicted by (Sabari et al., 2018), eclipsing exons 13-15/20. CRISPR/Cas9 technology facilitated site-specific mutagenesis. An eSpCas9(1.1) vector (Addgene Plasmid #71814) from Addgene was used to create the knockout clones, and guides were designed with http://benchling.com. Two guides were used for each transfection, with each Cas9 vector containing a single guide, which were then combined for the Neon procedure. By Neon electroporation, the guides were introduced into the cells along with a plasmid expressing GFP EGFP max (pCAG-EGxxFP). For transfection, 10ug of guide RNA plus an additional 1ug GFP plasmid were added to 2.5 million cells, and electroporated with Neon program 4. After 24hrs, the cells were single cell sorted for GFP expression to ensure successful transfection. Genotyping by PCR identified the clones with the desired bi-allelic deletion, and sequencing confirmed the cut sites.

Retroviral Transduction

Retrovirus was obtained from pMSCV-IRES-Bsr-Rag1 transfected 293T cells, including the full-length Rag1 sequence and a blasticidin selectable marker along with a pEco plasmid for ecotropic envelope protein production. Supernatant was collected post transfection at 48 hours and 72 hours.

The harvested retrovirus was used to transduce 445.3 cells; pMSCV-IRES-Bsr-Rag1 virus was added to cells along with 4ug/ml of polybrene, followed by a spinoculation cycle of 2,000 rpm for 60 minutes at 32 ºC in a swinging bucket centrifuge. After 48 hours, blasticidin was added to the virus+cell culture at a concentration of 20ug/ml. Cells were replenished with media+blasticidin for 7 days to ensure selection for successfully transduced cells, and thus a population with consistent Rag1 expression. The transduced cells were frozen after 7 days in order to maintain the integrity of all rearrangement assays, eliminating as much background rearrangement as possible.

RNA and gDNA Extraction

All RNA extractions were performed using the Qiagen RNEasy Plus Mini Kit (Cat. 74104), and all gDNA extractions performed with the Qiagen DNEasy Blood and Tissue Kit. Prior to processing cells for RNA, up to 5 x 10^6 harvested cells were run through a QIAshredder (Cat. 79654) to aid in overall extraction yield. For the cDNA used in qPCR assays as well as in library preparation and sequencing, Qiagen's QuantiTect Reverse Transcription Kit was used.

Germline Transcription Assay

RNA was isolated as described above from 445.3-WT, 445.3-BRD4Δ ,and 445.3-E2AΔ cells Rag-/- cells after 0 hours and 48 hours stimulation with STI571. The UCSD Institute for Genomic Medicine (IGM) performed ribodepleted RNA stranded HiSeq4000 on the samples with up to 25 million reads per sample.

Vκ-Jκ Rearrangment Sequencing Assay

Cells used for the Vκ-Jκ rearrangment assay were transduced with the Rag1 retrovirus as previously described. 445.3 Rag+ cells were then stimulated with STI571 and harvested for RNA and gDNA after a period of 48 hours. An Igκ primer cocktail (Vκall) was used to amplify a large array of rearranged gene products by qPCR. Based on CT value, the minimum number of cycles was determined to produce a visible PCR product in a 1% agarose gel. Using the same primers, Q5 high-fidelity Hot start master mix, and the predetermined cycle value (CT per sample), the cDNA or gDNA

was amplified for Vκ-Jκ gene rearrangement. Combining eight 20ul PCR reactions per sample, cleanup was conducted using Qiagen's QIAquick PCR Purification Kit. Following cleanup, the DNA was library prepped using the KAPA Hyper Prep Kit (Cat. NC0855928 (KK8500 - (Kapa Biosystems)) Library barcoding was done with Illumina TruSeq DNA adapter sequences from the TruSeq DNA Sample Prep Kit (Cat. 20015960). Sequencing performed with HiSeq4000 platform.

AmAc and NaCl Treatement

Cells in culture were treated with STI571 stimulant to induce differentiation and activation of Igκ rearrangement. These cells were then split into separate plates for treatment at 3 hours post stimulation. Adding AmAc (Invitrogen Cat. AM9070G) and NaCl to a final concentration of 100mM in solution, the cells were left for another 3 hours before a wash step. At 6 hours post STI571 treatment, the cells were centrifuged and the media containing treatment was aspirated. The cells were plated in new media and harvested at 24 hours for RNA, and 48 hours for harvesting gDNA. RNA was converted to cDNA as above described and used for Vκ-Jκ rearrangement sequencing.

HiC Protocol

Hi-C samples were processed using the Arima-HiC Kit (A410030) for mammalian cell lines. The sonication step specified in the manufacturers protocol was performed by the UCSD Institute for Genomic Medicine (IGM) using the Covaris ultrasonicator. The HiC library preparation was made using the Arima-HiC Protocol using KAPA Hyper Prep Kit with Illumina TruSeq DNA adapter sequences from the TruSeq DNA Sample Prep Kit (Cat. 20015960) kit, and a KAPA Library Quant Kit (Illumina) Universal qPCR Mix (Cat. KK4824).

Statistical Analysis and Software

Vκ-Jκ seq data was analyzed using MiXCR and the default parameters for V, J, ad C alignment and paired end reads overlap (https://mixcr.readthedocs.io/en/master/align.html). Percent gene usage and all fold changes were calculated with Excel. All schematics in Figure 1 were made using Prism 8, as

well as all bar graphs, pie charts, and Mann-Whitney plots. Significant statistical values represented as * ($P \leq 0.05$), ** ($P \leq 0.01$), *** ($P \leq 0.001$), **** ($P \leq 0.0001$).

Primers

Primers used to create repertoire libraries.

VκAll: GGCTGCAGSTTCAGTGGCAGTGGRTCWGGRAC

Cκ: TGTTCAAGAAGCACACGACTGA

AF944: GGATTCTACTTACGTTTGATTTCCA

Primers used to create CRISPR deletions.

E47/E2A G1 S :CACC GCAGCAGGATGGAACAGACG

E47/E2A G1 AS :AAAC CGTCTGTTCCATCCTGCTGC

E12/E2A G2 S :CACC GAGACAGTGAGGTGATGTGG

E12/E2A G2 AS :AAAC CCACATCACCTCACTGTCTC

BRD4 G1 S :CACC GCCTGATAAATACGGAACAG

BRD4 G1 AS :AAAC CTGTTCCGTATTTATCAGGC

BRD4 G2 S :CACC GCGAAGAAGACATGGACATG

BRD4 G2 AS :AAAC CATGTCCATGTCTTCTTCGC

This thesis is coauthored with Fourfouris, Tatiana; Murre, Cornelis; & Barajas-Mora, Edgar-Mauricio. The thesis author was the primary author of this paper.

REFERENCES

Abarrategui, Iratxe and Michael S. Krangel. 2006. “Erratum: Regulation of T Cell Receptor-α Gene Recombination by Transcription (Nature Immunology (2006)).” *Nature Immunology* 7(10):1115.

Abarrategui, Lratxe and Michael S. Krangel. 2009. “Germline Transcription: A Key Regulator of Accessibility and Recombination.” *Advances in Experimental Medicine and Biology* 650(D):93–102.

de Almeida, Claudia Ribeiro, Rudi W. Hendriks, and Ralph Stadhouders. 2015. *Dynamic Control of Long-Range Genomic Interactions at the Immunoglobulin κ Light-Chain Locus*. Vol. 128. 1st ed. Elsevier Inc.

Aoki-Ota, Miyo, Ali Torkamani, Takayuki Ota, Nicholas Schork, and David Nemazee. 2012. “Skewed Primary Igκ Repertoire and V–J Joining in C57BL/6 Mice: Implications for Recombination Accessibility and Receptor Editing.” *The Journal of Immunology* 188(5):2305–15.

Banani, Salman F., Hyun O. Lee, Anthony A. Hyman, and Michael K. Rosen. 2017. “Biomolecular Condensates: Organizers of Cellular Biochemistry.” *Nature Reviews Molecular Cell Biology* 18(5):285–98.

Barajas-Mora, E. Mauricio, Eden Kleiman, Jeffrey Xu, Nancy C. Carrico, Hanbin Lu, Eugene M. Oltz, Cornelis Murre, and Ann J. Feeney. 2019. “A B-Cell-Specific Enhancer Orchestrates Nuclear Architecture to Generate a Diverse Antigen Receptor Repertoire.” *Molecular Cell* 73(1):48-60.e5.

Bastiaan Holwerda, Sjoerd Johannes and Wouter de Laat. 2013. “CTCF: The Protein, the Binding Partners, the Binding Sites and Their Chromatin Loops.” *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Bemark, Mats, David Liberg, and Tomas Leanderson. 1998. “Conserved Sequence Elements in K Promoters from Mice and Humans: Implications for Transcriptional Regulation and Repertoire Expression.” *Immunogenetics*.

Brangwynne, Clifford P., Peter Tompa, and Rohit V. Pappu. 2015. “Polymer Physics of Intracellular Phase Transitions.” *Nature Physics* 11(11):899–904.

Brekke, Katherine M. and William T. Garrard. 2004. “Assembly and Analysis of the Mouse Immunoglobulin Kappa Gene Sequence.” *Immunogenetics*.

Choi, Nancy M. and Ann J. Feeney. 2014. “CTCF and NcRNA Regulate the Three-Dimensional Structure of Antigen Receptor Loci to Facilitate V(D)J Recombination.” *Frontiers in Immunology* 5(FEB):1–8.

Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. “Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions.” *Nature*.

Feeney, Ann J. 2009. “Genetic and Epigenetic Control of Vgene Rearrangement Frequency.” *Advances in Experimental Medicine and Biology* 650(D):73–81.

Fuxa, Martin, Jane Skok, Abdallah Souabni, Giorgia Salvagiotto, Esther Roldan, and Meinrad Busslinger. 2004. “Pax5 Induces.” *Genes & Development* (D):411–22.

Gibson, Bryan A., Lynda K. Doolittle, Maximillian W. G. Schneider, Liv E. Jensen, Nathan Gamarra, Lisa Henry, Daniel W. Gerlich, Sy Redding, and Michael K. Rosen. 2019. "Organization of Chromatin by Intrinsic and Regulated Phase Separation." *Cell* 179(2):470-484.e21.

Gorman, James R., Nienke Van Der Stoep, Robert Monroe, Michel Cogne, Laurie Davidson, and Frederick W. Alt. 1996. "The Igκ 3' Enhancer Influences the Ratio of Igκ versus Igλ B Lymphocytes." *Immunity* 5(3):241–52.

Hyman, Anthony A., Christoph A. Weber, and Frank Jülicher. 2014. "Liquid-Liquid Phase Separation in Biology." *Annual Review of Cell and Developmental Biology* 30(1):39–58.

Inlay, Matthew, Frederick W. Alt, David Baltimore, and Yang Xu. 2002. "Essential Roles of the κ Light Chain Intronic Enhancer and 3′ Enhancer in κ Rearrangement and Demethylation." *Nature Immunology* 3(5):463–68.

Jain, Ankur and Ronald D. Vale. 2017. "RNA Phase Transitions in Repeat Expansion Disorders." *Nature* 546(7657):243–47.

Jhunjhunwala, Suchit, Menno C. van Zelm, Mandy M. Peak, and Cornelis Murre. 2009. "Chromatin Architecture and the Generation of Antigen Receptor Diversity." *Cell*.

Kleiman, Eden, Jeffrey Xu, and Ann J. Feeney. 2018. "Cutting Edge: Proper Orientation of CTCF Sites in Cer Is Required for Normal Jκ-Distal and Jκ-Proximal Vκ Gene Usage." *The Journal of Immunology*.

Lazorchak, A. S., M. S. Schlissel, and Y. Zhuang. 2006. "E2A and IRF-4/Pip Promote Chromatin Modification and Transcription of the Immunoglobulin Locus in Pre-B Cells." *Molecular and Cellular Biology* 26(3):810–21.

Lichtenstein, Michal, Gil Keini, Howard Cedar, and Yehudit Bergman. 1994. "B Cell-Specific Demethylation: A Novel Role for the Intronic κ Chain Enhancer Sequence." *Cell* 76(5):913–23.

Lin, Yin C., Christopher Benner, Robert Mansson, Sven Heinz, Kazuko Miyazaki, Masaki Miyazaki, Vivek Chandra, Claudia Bossen, Christopher K. Glass, and Cornelis Murre. 2012. "Global Changes in the Nuclear Positioning of Genes and Intra-and Interdomain Genomic Interactions That Orchestrate B Cell Fate." *Nature Immunology* 13(12):1196–1204.

Lin, Yin C., Suchit Jhunjhunwala, Christopher Benner, Sven Heinz, Eva Welinder, Robert Mansson, Mikael Sigvardsson, James Hagman, Celso A. Espinoza, Janusz Dutkowski, Trey Ideker, Christopher K. Glass, and Cornelis Murre. 2010. "A Global Network of Transcription Factors, Involving E2A, EBF1 and Foxo1, That Orchestrates B Cell Fate." *Nature Immunology*.

Muljo, Stefan A. and Mark S. Schlissel. 2003. "A Small Molecule Abl Kinase Inhibitor Induces Differentiation of Abelson Virus-Transformed Pre-B Cell Lines." *Nature Immunology* 4(1):31–37.

Murre, Cornelis. 2019. "Helix–Loop–Helix Proteins and the Advent of Cellular Diversity: 30 Years of Discovery." *Genes and Development* 33(1–2):6–25.

Murre, Cornelis, Patrick Schonleber McCaw, and David Baltimore. 1989. "A New DNA Binding and Dimerization Motif in Immunoglobulin Enhancer Binding, Daughterless, MyoD, and Myc Proteins." *Cell* 56(5):777–83.

O'Riordan, Mary and Rudolf Grosschedl. 1999. "Coordinate Regulation of B Cell Differentiation by the Transcription Factors EBF and E2A." *Immunity*.

Ong, Chin Tong and Victor G. Corces. 2011. “Enhancer Function: New Insights into the Regulation of Tissue-Specific Gene Expression.” *Nature Reviews Genetics* 12(4):283–93.

Rao, Suhas S. P., Miriam H. Huntley, Neva C. Durand, Elena K. Stamenova, Ivan D. Bochkov, James T. Robinson, Adrian L. Sanborn, Ido Machol, Arina D. Omer, Eric S. Lander, and Erez Lieberman Aiden. 2014. “A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping.” *Cell*.

Ribeiro de Almeida, Claudia, Ralph Stadhouders, Marjolein J. W. De Bruijn, Ingrid M. Bergen, Supat Thongjuea, Boris Lenhard, Wilfred Van IJcken, Frank Grosveld, Niels Galjart, Eric Soler, and Rudi W. Hendriks. 2011. “The DNA-Binding Protein CTCF Limits Proximal Vκ Recombination and Restricts κ Enhancer Interactions to the Immunoglobulin κ Light Chain Locus.” *Immunity*.

Sabari, Benjamin R., Alessandra Dall’Agnese, Ann Boija, Isaac A. Klein, Eliot L. Coffey, Krishna Shrinivas, Brian J. Abraham, Nancy M. Hannett, Alicia V. Zamudio, John C. Manteiga, Charles H. Li, Yang E. Guo, Daniel S. Day, Jurian Schuijers, Eliza Vasile, Sohail Malik, Denes Hnisz, Tong Ihn Lee, Ibrahim I. Cisse, Robert G. Roeder, Phillip A. Sharp, Arup K. Chakraborty, and Richard A. Young. 2018. “Coactivator Condensation at Super-Enhancers Links Phase Separation and Gene Control.” *Science*.

Sleckman, Barry P. and Eugene M. Oltz. 2012. “Preparing Targets for V(D)J Recombinase: Transcription Paves the Way.” *The Journal of Immunology* 188(1):7–9.

Stadhouders, Ralph, Marjolein J. W. de Bruijn, Magdalena B. Rother, Saravanan Yuvaraj, Claudia Ribeiro de Almeida, Petros Kolovos, Menno C. Van Zelm, Wilfred van Ijcken, Frank Grosveld, Eric Soler, and Rudi W. Hendriks. 2014. “Pre-B Cell Receptor Signaling Induces Immunoglobulin κ Locus Accessibility by Functional Redistribution of Enhancer-Mediated Chromatin Interactions.” *PLoS Biology* 12(2).

Strzyz, Paulina. 2018. “Concentrating on Intrinsic Disorder.” *Nature Reviews Genetics* 19(9):534.

Tang, Zhonghui, Oscar Junhong Luo, Xingwang Li, Meizhen Zheng, Jacqueline Jufen Zhu, Przemyslaw Szalaj, Pawel Trzaskoma, Adriana Magalska, Jakub Wlodarczyk, Blazej Ruszczycki, Paul Michalski, Emaly Piecuch, Ping Wang, Danjuan Wang, Simon Zhongyuan Tian, May Penrad-Mobayed, Laurent M. Sachs, Xiaoan Ruan, Chia Lin Wei, Edison T. Liu, Grzegorz M. Wilczynski, Dariusz Plewczynski, Guoliang Li, and Yijun Ruan. 2015. “CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription.” *Cell*.

Verma-Gaur, Jiyoti, Ali Torkamani, Lana Schaffer, Steven R. Head, Nicholas J. Schork, and Ann J. Feeney. 2012. “Noncoding Transcription within the Igh Distal VH Region at PAIR Elements Affects the 3D Structure of the Igh Locus in Pro-B Cells.” *Proceedings of the National Academy of Sciences of the United States of America*.

de Wit, Elzo, Erica S. M. Vos, Sjoerd J. B. Holwerda, Christian Valdes-Quezada, Marjon J. A. M. Verstegen, Hans Teunissen, Erik Splinter, Patrick J. Wijchers, Peter H. L. Krijger, and Wouter de Laat. 2015. “CTCF Binding Polarity Determines Chromatin Looping.” *Molecular Cell*.

Xiang, Yougui, Sung-Kyun Park, and William T. Garrard. 2014. “A Major Deletion in the Vκ–Jκ Intervening Region Results in Hyperelevated Transcription of Proximal Vκ Genes and a Severely Restricted Repertoire.” *The Journal of Immunology*.

Yancopoulos, George D. and Frederick W. Alt. 1985. “Developmentally Controlled and Tissue-Specific Expression of Unrearranged VH Gene Segments.” *Cell* 40(2):271–81.