HHS Public Access

Author manuscript

Adv Immunol. Author manuscript; available in PMC 2015 November 26.

Published in final edited form as:

Adv Immunol. 2015 ; 128: 123–182. doi:10.1016/bs.ai.2015.07.003.

LONG RANGE REGULATION OF V(D)J RECOMBINATION

Charlotte Proudhon[1,*], **Bingtao Hao**[1,*], **Ramya Raviram**[1], **Julie Chaumeil**[3], and **Jane A. Skok**[1,2]

[1]Department of Pathology, New York University School of Medicine, 550 First Avenue, MSB 599, New York, NY10016, USA

[3]Institut Curie, CNRS UMR3215, INSERM U934, 26 rue d'Ulm, Paris 75005, France

Abstract

Given their essential role in adaptive immunity, antigen receptor loci have been the focus of analysis for many years and are among a handful of the most well studied genes in the genome. Their investigation led initially to a detailed knowledge of linear structure and characterization of regulatory elements that confer control of their rearrangement and expression. However, advances in DNA FISH and imaging combined with new molecular approaches that interrogate chromosome conformation have led to a growing appreciation that linear structure is only one aspect of gene regulation and in more recent years the focus has switched to analyzing the impact of locus conformation and nuclear organization on control of recombination. Despite decades of work and intense effort from numerous labs we are still left with an incomplete picture of how antigen receptor loci are regulated. This chapter summarizes our advances to date and points to areas that need further investigation.

Keywords

RAG; V(D)J recombination; allelic exclusion; ATM; homologous pairing; nuclear organization; pericentromeric heterochromatin; CTCF

1. OVERVIEW OF V(D)J RECOMBINATION

In total there are seven antigen receptor loci, four T cell receptor (*Tcr*) loci (*Tcrg*, *Tcrd*, *Tcrb* and *Tcra*) and three B cell specific immunoglobulin genes (*Igh*, *Igk* and *Igl*). B and T cells make use of this modest investment in DNA to generate an almost infinite assortment of different specificity receptors that can be used to combat a wide variety of invading pathogens. Somatic rearrangement of variable (V), diversity (D) and joining (J) gene segments arrayed along each locus generates this receptor diversity enabling specific recognition of foreign antigen, which is a fundamental feature of the adaptive immune response (Helmink and Sleckman, 2012; Tonegawa, 1983).

[2]To whom correspondence should be addressed. Tel: 212 263 0504; jane.skok@nyumc.org.
[*]These authors contributed equally

1.1 RAG binding

Recombination is mediated by the lymphoid-specific recombinase, consisting of RAG1 and RAG2 (the protein products of the *recombination activating genes 1* and *2*). The RAG1 protein, which harbors the endolytic activity, functions in conjunction with RAG2, a co-factor that is essential for recombinase activity (Mombaerts et al., 1992; Shinkai et al., 1992; Spanopoulou et al., 1994). The RAG1 protein cleaves specifically at highly conserved recombination sequences (RSSs) made up of heptamers and nonamer motifs separated by non-conserved spacers of either 12 or 23bps (Kim et al., 1999; Landree et al., 1999). The RAG1/2 complex preferentially binds two RSS sites of different spacer lengths, brings them together and cuts at the borders of these elements generating DSBs. RSSs, which flank the individual V, D and J gene segments, are distributed throughout each antigen receptor locus and synapse formation and cleavage can occur between regions that are many kilobases apart. The four broken ends (two coding ends and two signal ends) are held together in a RAG post cleavage complex that directs repair through the non homologous end joining (NHEJ) pathway, which is important for the maintenance of genome stability (Deriano et al., 2011; Helmink and Sleckman, 2012; Lee et al., 2004; Schatz and Swanson, 2011). Recent ground breaking analyses of the crystal structures of these two proteins indicates that the RAG1-RAG2 heterotetramer is Y-shaped, with a RAG1-RAG2 heterodimer constituting each arm (Kim et al., 2015). The structure explains numerous mutations known to be associated with immunodeficiencies.

According to ChIP-seq analysis, the binding profile of RAG1 and RAG2 overlaps with that of H3K4me3 (Ji et al., 2010). Promiscuous genome wide binding to this active chromatin mark is mediated via a plant homeodomain (PHD) in RAG2 (Liu et al., 2007b; Matthews et al., 2007). However, each RAG protein can bind in the absence of the other, and when RAG1 is bound without RAG2 it binds in an RSS specific manner and is not found at H3K4me3 enriched promoters (Ji et al., 2010). This finding suggests that binding of the proteins can occur individually at differential locations or together as a preformed RAG1/2 complex that directs both proteins to RSSs as well as H3K4me3 enriched regions.

The question of how and what controls RAG targeting at the locus and allele specific level on the individual antigen receptor loci continues to be an area under investigation. Moreover, there is the puzzle about how, in normal circumstances, other genes in the genome with the appropriate or cryptic recognition sequences are protected from being cleaved. Since cryptic RSSs are found every 1–2Kb in the genome, promiscuous RAG1 binding could contribute to off-target cleavage occurring within non-antigen receptor loci. Indeed RAG targeting has been linked to genetic defects in *IKZF1, Notch1, SIL-SCL, Bcl11b, PTEN, ETV6, BTG1, TBL1XR1*, and *CDKN2A-CDKN2B* that are associated with numerous B and T acute lymphoblastic leukemias (ALLs) (Mendes et al., 2014; Mullighan et al., 2008; Onozawa and Aplan, 2012; Papaemmanuil et al., 2014).

1.2 Lineage and stage specific rearrangement

Given the risks entailed by repeated cutting and pasting, V(D)J recombination is tightly regulated with respect to target gene accessibility, RAG expression and the activities of the DNA damage signaling and repair pathways. As the recombinase machinery (the RAG

proteins) and the DNA targets (RSSs) are the same for each antigen receptor locus in both lineages, lymphocytes restrict recombination by controlling the accessibility of the individual loci (Figure 1). First, rearrangement is restricted by lineage: *Ig* gene segments complete rearrangement only in B cells, and *Tcr* gene segments rearrange only in T cells. Second, rearrangement is ordered by stage within a given lineage: the *Ig heavy chain (Igh)* is rearranged at the pro-B cell stage of development prior to *Ig light chain (kappa* or *lambda*) rearrangement in pre-B cells. Furthermore, D_H-to-J_H recombination at the *Igh* locus must take place in pre-pro-B cells before V_H-to-DJ_H rearrangement can begin in pro-B cells.

In T cells the situation is more complex as productive rearrangement of the different *Tcr* loci gives rise to two distinct lineages: *Tcrg/Tcrd* and *Tcrb/Tcra* recombination leads to γδ and αβ T cells, respectively (Ciofani and Zuniga-Pflucker, 2010; Krangel, 2009). Nonetheless, recombination of the different loci overlaps such that *Tcrg*, *Tcrd* and *Tcrb* are all rearranged at the early $CD4^-CD8^-$ double negative DN2/3 stage of development, while *Tcra* recombination occurs later in double positive (DP) cells after successful *Tcrb* rearrangement (Livak et al., 1999). In addition, promiscuous D_H-to J_H rearrangement of the *Igh* locus occurs at low level in T lineage cells (predominantly the DN cell stage) (Chaumeil et al., 2013b; Kurosawa et al., 1981). Multi-locus rearrangement in the same developmental compartment increases the risks associated with recombination and the probability of aberrant repair (Chaumeil et al., 2013b). Regulation of recombination is further complicated because *Tcra* and *Tcrd*, which are rearranged in DN and DP cell stages, respectively, share the same chromosomal location with *Tcrd* embedded between the Vα and Jα gene segments of *Tcra*.

2. LINEAR STRUCTURE OF THE ANTIGEN RECEPTOR LOCI

Antigen receptor loci consist of large arrays of V gene segments (ranging from 34 in *Tcrb* to 183 segments in *Igh* that are dispersed over 0.67Mb and 2.4Mb, respectively). A much smaller proximal domain containing D, J and C gene segments that encompass potent enhancers, occupies genomic regions in the kb range (4kb in *Igk*, 25kb in *Tcrb*, 70kb in *Tcra* and 26kb in *Igh*). Although all the loci are comprised of the same basic units (V, D and J gene segments) that are flanked by RSSs and a constant region, each antigen receptor locus has a unique structure that impacts their regulation (Figure 2–5).

2.1. Igh

The murine *Igh* locus spans 2.75Mb (nearly a quarter of the yeast genome) and is located at the distal end of chromosome 12 in mouse. It contains a total of 113 functional V_H segments that are dispersed over 2.4Mb. *Igh* holds 10–15 functional D_H segments (depending on the mouse strain), 4 J_H gene segments and 8 different constant regions that are all preceded by switch regions with the exception of Cδ. These are used as substrates for class switch recombination (CSR) which generates different *Ig* isotypes that streamline antibody effector function after encounter with an antigen (IgE, IgG, IgA etc) (Figure 2A).

2.2 Igk

The *Igk* light chain locus is located on the mouse chromosome 6. It spans 3.17Mb and contains 92 functional Vκ segments, 4 functional Jκs and a single Cκ region. In contrast to *Igh*, *Igk* does not contain any D gene segments (Figure 3A). Another feature of the *Igk* locus is that half of the Vκs are in reverse orientation and are rearranged by non-destructive inversion, which leads to retention of the segments located between the joining Vκ and Jκ segments. This conserves Vκ segments for (i) secondary rearrangements that can occur with remaining downstream Jκs in the event of nonproductive rearrangement, and (ii) receptor editing which functions to eliminate self reactive receptors or enable IGK to associate with IGH (Feddersen et al., 1990; Halverson et al., 2004; Pelanda et al., 1997; Prak and Weigert, 1995; Tiegs et al., 1993). Ongoing rearrangement and receptor editing is possible because of the lack of D gene segments and recombination on each allele can continue until all the Jκ gene segments are used up. Based on the delayed activation of *Igl*, it is estimated that three rounds of rearrangement are possible for each *Igk* allele (Arakawa et al., 1996), which corresponds to the number of functional Jκ gene segments. While no specific order is determined for Vκ rearrangement (Nadel et al., 1998), primary rearrangement generally involve the most 5'Jκ segment, Jκ1 (Yamagami et al., 1999).

2.3 Tcrb

The T cell receptor beta locus, *Tcrb*, is encoded by 700kb of DNA on mouse chromosome 6. The vast majority of the locus (~624kb) is comprised of 22 functional Vβ gene segments. Except for Vβ31, which is localized downstream of the proximal domain in an inverted orientation, all the Vβ genes are located upstream of a duplicated cluster of '1 Dβ, 7 Jβ and 1 Cβ' of which 11 of the 14 Jβs are functional. In addition to this atypical proximal duplication, 2 clusters of trypsinogen genes, that are inactive in lymphocytes, separate the bulk of the Vβ array from the first Dβ segment on the 3' side (separation of 250kb) as well as from the first Vβ segment, Vβ1 located at the 5' end of the locus (Figure 4A).

2.4 Tcra

As mentioned above the most striking feature of the *Tcra* locus is that *Tcrd* is embedded within it and the two loci share a subset of V genes (Figure 5A). The whole locus spans 1.6Mb in the 129 mouse strain and 2.0Mb in C57BL/6. These differences stem from repeat regions within the V gene cluster of which there are two in strain 129 and three in C57BL/6. The *Tcrd* locus (which is located between the Vα and Jα gene segments of *Tcra*) harbors two Dδ, two Jδ genes and one Cδ gene segment. There are 5 Vδ specific genes located in the 3' unique Vα cluster and a single Vδ gene in reverse orientation that is located downstream of Cδ (Vδ5) adjacent to the Jα array that is comprised of 60 gene segments.

Tcrd rearrangement occurs in DN cells prior to *Tcra* rearrangement in DP cells. This order is important because the first round of *Tcra* rearrangement deletes the *Tcrd* gene. Unlike the other loci that contain D gene segments (*Igh* and *Tcrb*) *Tcrd* is not subjected to ordered rearrangement. Thus Vδ-to-Dδ and Dδ-to-Jδ rearrangement occur at the same time, which enables Dδ gene segments to recombine together to form DDδ gene rearrangements (Monroe et al., 1999). *Tcrd* makes use of only a subset of V gene segments including several *Tcrd* specific V genes (TRDV1, TRDV2-1, TRDV2-2, TRDV4, TRDV5) and some V gene

segments that are shared with *Tcra* (TRAV21/DV12, TRAV13-4/DV7, TRAV6-7/DV9, TRAV4-4/DV10, TRAV14D-3/DV8, TRAV16D/DV11, and four members of the TRAV15/DV6 family) (Hawwari and Krangel, 2005). TRDV4 rearranges specifically in fetal thymocytes and is repressed in adult thymocytes by constitutively high levels of a suppressive modification, H3K9me2 (Hao and Krangel, 2011). The local accessibility of Vδ gene segments determines which genes undergo rearrangement in DN cells and a recent study showed that replacing the promoter of a *Tcra* specific Vα gene, TRAV12 with the TRAV15/DV6 promoter increases the usage of the TRAV12 in *Tcrd* recombination (Hao and Krangel, 2011; Naik et al., 2015). *Tcra* recombination occurs after *Tcrd* recombination in DP cells and it has recently been shown that IL-7 signaling contributes to the control of stage specificity by preventing premature *Tcra* rearrangement in DN4 cells (Boudil et al., 2015).

3. NUCLEAR ORGANIZATION AND ITS IMPACT ON ACCESSIBILITY

With the exception of the stages in development where they recombine, antigen receptor loci are by default inaccessible to the RAG proteins. Opening up of the loci for rearrangement occurs at multiple levels including DNA demethylation, activation of chromatin, initiation of sense and antisense germline transcription, nucleosome repositioning, relocation of the loci from inaccessible repressive nuclear locations (peripheral lamin associated domains or pericentromeric heterochromatin) to accessible euchromatin, and locus contraction (which brings distal V gene segments into close physical contact with the proximal DJC domain enabling recombination between widely separated gene segments) as reviewed in (Chaumeil and Skok, 2013; Hewitt et al., 2010; Johnson et al., 2010). All of these changes are regulated by lineage and stage specific activation of *cis* regulatory elements (promoters, enhancers, chromatin insulators and others) that recruit transcription factors and structural proteins such as CTCF and cohesin that orchestrate changes which promote ordered recombination within each locus as outlined in **§4**. In this section we focus on changes in nuclear organization that occur during rearrangement.

3.1 Subnuclear localization

DNA FISH analyses by our lab and others revealed that there are links between the location of antigen receptor loci and their activation status. The first study to demonstrate that the proximity of a gene relative to the nuclear periphery is reflective of an inactive state focused on the *immunoglobulin loci* (Kosak et al., 2002). *Igh* alleles are located at the nuclear periphery in T lineage cells and they move inwards just prior to the onset of recombination at the pro-B cell stage. Further detailed analysis revealed that in pre-pro-B cells and T lineage cells the locus is anchored at the periphery through its 5' V_H gene segments while the 3' end, containing the $D_HJ_HC_H$ cluster of gene segments is more centrally located. This orientation is compatible with D_HJ_H segments having access to euchromatically located recombinase enzymes and rearrangement being restricted to D_H and J_H gene segments in both these cell types (Figure 6) (Fuxa et al., 2004). Peripherally located V_H gene segments are refractory to RAG and do not get recombined in these cells.

3.2 Locus contraction

The antigen receptor loci all occupy large expanses of DNA ranging from 1Mb (*Tcrb*) to over 3MB (*Igk*) that mostly encompass V gene segments. This presents a logistical problem since rearrangement requires the formation of a synapse between recombining regions that can be widely separated on the linear chromosome. The first 3D DNA FISH analyses of antigen receptor loci in developing lymphocytes revealed that the *Igh* locus is in an extended position in lymphocyte progenitors and non-B cells and that in recombining cells it is in a contracted conformation (Fuxa et al., 2004; Kosak et al., 2002). This is also the case for the *Igk*, *Tcrb* and *Tcra* loci, however unlike the other antigen receptor loci, *Igk* is contracted in pro-B cells, the stage prior to recombination, which may be a reflection of the fact that *Igk* can undergo low-level recombination in these cells (Roldan et al., 2005; Skok et al., 2007). In all cases, locus contraction brings widely dispersed V gene segments into contact with the proximal DJC domain through chromatin looping to provide every V gene an equal opportunity to rearrange. However, it should be noted that other factors, such as RSS sequence, chromatin status and transcriptional activity also influence which V gene is targeted for recombination.

Detailed FISH analysis by the Murre lab investigating loop formation of the *Igh* locus used multiple small probes combined with mathematical modeling to reveal that the V_H genes of the *Igh* locus are folded into two one megabase rosette-like structures that are connected by linkers. The rosettes are separated from each other in pre-pro-B and T cells consistent with a decontracted state and they interact with each other in pro-B cells when the locus is contracted (Jhunjhunwala et al., 2008). The rosettes are compatible with TAD structures that were defined four years later as a result of Hi-C chromosome conformation analyses that analyze interactions between everything and everything (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). TAD structures are highly self-interacting regions that are separated by distinct boundaries. 5C analyses performed in the Skok lab delineates two distinct TADs for the proximal and distal V_H gene domains that are consistent with the Murre labs findings and with a separation of their regulation (Figure 2B and 7) Contraction – or inter TAD association – is a reversible process and once recombination has taken place the loci are once again found in a decontracted conformation. This is discussed in more detail in the context of allelic exclusion (see **§5**).

Most of the early work focusing on understanding the mechanisms underlying locus contraction centered on the *Igh* locus, in part because the B cell specific transcription factor, Pax5 was the first factor to be identified as essential for altering locus conformation. Pax5 is upregulated at the pro-B cell stage after the initiation of D_H-to-J_H rearrangement, which starts at the earlier pre-pro-B cells stage (Fuxa et al., 2004). In the absence of Pax5, the two *Igh* alleles are found in an extended form in the center of the nucleus (Fuxa et al., 2004) and in this conformation only the 4 most proximal V_H genes out of a total of nearly 200, can undergo recombination, underlining the importance of this process in generating antibody diversity (Roldan et al., 2005). Pax5 is essential for mediating contraction in B cells. However, ectopic expression of this factor in T cells cannot induce a change in locus conformation although it does have an impact on relocalizing the two alleles from the periphery to the center of the nucleus (Figure 6). These findings indicate that another, yet

unidentified, factor that is present in B cells but not T cells, is required for contraction (Fuxa et al., 2004). Relocation to the center of the nucleus in T cells ectopically expressing Pax5 likely occurs as an indirect effect of Pax5 on upregulating another B cell specific transcription factor, EBF.

Based on these observations we put forward a two-step model for *Igh* activation (Figure 6) (Roldan et al., 2005). In lymphoid progenitor cells the *Igh* locus is found in an extended conformation and is anchored at the nuclear periphery via its 5' end. The central more accessible location of the D_HJ_H segments is compatible with D_H-to-J_H rearrangement occurring at low levels in T cells as previously reported (Chaumeil et al., 2013b; Kurosawa et al., 1981). Upregulation of EBF early in B cell development induces relocation of *Igh* to the center of the nucleus, which increases D_H-to-J_H rearrangement and allows V_H-to DJ_H recombination involving proximal V_H genes. Distal V_H gene rearrangement is not possible when the locus is in an extended conformation and thus Pax5 expression at the pro-B cell stages is required for inclusion of these segments in the antibody repertoire.

No lineage and stage specific factors have been identified as essential for locus contraction of *Igk*, *Tcrb* or *Tcra*. However, ubiquitously expressed YY1, which has been shown to be important for *Igh* locus contraction (Liu et al., 2007a; Medvedovic et al., 2013), has also been identified as important for mediating *Igk* contraction (Liu et al., 2007a; Pan et al., 2013). Furthermore, CCCTC-binding factor, CTCF and its binding partner cohesin also play important roles, as reviewed in (Chaumeil and Skok, 2012) and discussed in more detail below.

3.3 Allelic pairing and pericentromeric localization

Accessibility of antigen receptor loci has been linked to proximity to a second repressive compartment of the nucleus, pericentromeric heterochromatin (PCH). The Fisher lab was the first to show that transcriptionally inactive genes localize to pericentromeric heterochromatin in developing T cells (Brown et al., 1999) and subsequent studies led to the discovery that productively rearranged and non-productively rearranged *Igh* alleles are found in distinct nuclear compartments in mature activated B cells (Skok et al., 2001). Differences in nuclear localization of highly expressed productively rearranged versus low expressing non-productively rearranged alleles led to the idea that repositioning to pericentric regions could play a role in allelic exclusion (Roldan et al., 2005; Skok et al., 2007). Indeed it does, but not surprisingly this turns out to be just one aspect of control. In the case of the antigen receptor loci, pericentromeric localization is linked to homologous and heterologous antigen receptor allele pairing, which in turn is linked to control of allele and locus specific accessibility ensuring that breaks are introduced asynchronously on one allele or locus at a time.

Briefly, we discovered that RAG proteins enhance association of recombining homologous and heterologous loci in euchromatic regions of the nucleus. Pairing is mediated by RAG induced higher-order looping of one allele away from its respective chromosome territory. RAG-mediated cleavage is targeted to the looped out allele and once the break is introduced the DNA damage sensing factor, ATM (ataxia telengiectasia mutated) is recruited in *cis* to the site to initiate repair. Both the C-terminal portion of RAG2 and ATM perform the same

function and act in *trans* on other recombining alleles (homologues) or loci (heterologues e.g. *Igh* and *Tcra*) repositioning them to PCH and inhibiting further higher-order loop formation (Figure 8). This may be part of the mechanism that serves to prevent further cleavage until repair of the first break is completed (Chaumeil et al., 2013a; Chaumeil et al., 2013b; Chaumeil and Skok, 2013; Hewitt et al., 2009). Together these data support a model in which breaks occur on paired alleles, however since allelic association is likely transient and cleavage a rapid event with repair occurring over several hours, it is not surprising we find repair foci on both paired and unpaired alleles. Given these observations our model is difficult to prove without a live imaging system that can track these events over time.

Is pairing necessary for this level of control? We hypothesize that pairing is required for cleavage as well as feedback control of cleavage as close proximity could be important for coordinating *trans* control by ATM and the C-terminal portion of RAG2. Although RAG is found in abundance in euchromatic regions of the nucleus and is enriched at H3K4me3 marked chromatin (Ji et al., 2010), RAG-mediated cleavage is inherently inefficient (it would be dangerous if it were any other way) as RAG-mediated breaks are only detected in around 20% or less of recombining cells as determined by immuno-FISH analyses of breaks (Chaumeil et al., 2013a; Chaumeil et al., 2013b; Hewitt et al., 2009). It is thus not inconceivable that the chance of a break occurring improves as the local concentration of RAG increases. However, it could be argued that increasing the local concentration of RAG in recombination centers via association of RAG bound genes (Chaumeil et al., 2013b; Chaumeil and Skok, 2013) would increase the risk of cleavage on closely associated loci. Not so, if feedback control of RAG cleavage via an ATM-mediated mechanism occurs in a localized fashion (Figure 8).

Support for this idea comes from feedback control of Spo-11p mediated cleavage during meiosis: ATM appears to play a similar role in regulating the introduction of breaks in this process, however it is clear that cleavage control occurs in a localized fashion in meiosis (Garcia et al., 2015; Lange et al., 2011). Since feedback control of recombination appears to be conserved in meiosis and V(D)J recombination it likely shares common mechanistic features, the details of which have yet to be worked out. In both cases feedback control of cleavage is important for maintenance of genome stability and in the case of V(D)J recombination, asynchronous cleavage provides a means of (i) maintaining genome instability and preventing the generation of translocations and (ii) initiating allelic exclusion, ensuring that both alleles are not rearranged at the same time which could lead to the generation of two productively recombined alleles (discussed in **§5** below).

3.4 *Igk-Igh* allelic pairing and its impact on *Igh* locus contraction

A further example of pairing linked to PCH localization comes from a transient interaction between *Igh* and *Igk* that occurs at the pre-B cell stage of development (Hewitt et al., 2008). We discovered that association of one *Igk* allele with one *Igh* allele at PCH triggers (i) the repositioning of *Igh* to PCH and (ii) *Igh* locus decontraction. This serves to (i) reduce accessibility of partially recombined (DJ_H rearranged) *Igh* alleles that could otherwise go on being rearranged in pre-B cells, and to (ii) prevent ongoing mid and distal V_H rearrangement occurring during light chain rearrangement. Inter-locus *Igk-Igh* pairing and *Igh*

decontraction rely on the 3'Eκ *Igk* enhancer: in its absence there is reduced *Igk-Igh* pairing, reduced *Igh* localization at PCH and *Igh* remains in a contracted conformation increasing the level of mid and distal V_H rearrangement detected in pre-B cells (Figure 9) (Hewitt et al., 2008). Intriguingly, the intronic enhancer of *Igk*, MiEκ, has an antagonistic effect on *Igk-Igh* pairing, *Igh* localization at PCH and decontraction which are all increased in its absence in association with reduced levels of mid and distal V_H rearrangement in pre-B cells (Hewitt et al., 2008). Stopping ongoing rearrangement of *Igh* at the pre-B cell stage is important because a second productively rearranged *Igh* allele could potentially violate allelic exclusion.

4. *CIS* ACTING ELEMENTS THAT CONTROL ACCESSIBILITY AND RECOMBINATION

In this section we aim to highlight functions of *cis* acting elements and their role in regulating accessibility, locus conformation and ordered recombination. In particular we will focus on the most recent work identifying regulatory elements that play an important role in all of these aspects of control. All of these elements are depicted in Figures 2–5. In particular, in these figures we have focused on CTCF binding elements, (CBEs) as these are an important component of long range interactions. For all loci we have analyzed the orientation of the CBEs as a recent study demonstrates that loop bases involve a pair of CTCF sites in a head to head orientation (Rao et al., 2014), and it is possible that the directionality of the CTCF sites determines who interacts with whom. In addition we analyzed the location of the closest CBE relative to the TSS of each V gene and have marked whether these are upstream, downstream or overlapping with it and whether there is any pattern to this organization on the individual loci. Segment annotations with coordinates, strand orientation and functional status as well as coordinates for regulatory elements are provided in Table 1. Annotations were collected from NCBI (*Igh* gene ID: 111507; *Igk* gene ID: 243469; *Tcrb* gene ID: 21577; *Tcra* Gene ID: 21473) and IMGT/LIGM-DB databases (Giudicelli et al., 2006) using the mm10 genome build that uses the C57BL/6 strain as the reference genome.

4.1 Important *cis* acting elements and their function in *Igh* rearrangement

4.1.1 The intronic enhancer Eμ—The intronic enhancer Eμ, located in the 700kb region that separates the J_H and the C_H clusters is a combination of a 220bp core enhancer element (cEμ) and two 310–350bp flanking matrix attachment regions (MARs). Deletion of Eμ has been shown to impair both D_H to J_H and V_H to DJ_H recombination. In Eμ knockout mouse models, sense μ0 (initiated at the D_HQ52 region), lμ transcripts (which originate 3' of the Eμ core (Lennon and Perry, 1985; Perlot et al., 2005) and antisense transcripts in the J_H and D_H regions (Afshar et al., 2006) are severely impaired. However, despite the defect in V(D)J recombination and a partial block in B cell development at the pro-B cell stage, Eμ deletion (core or full length) does not severely affect germline sense or antisense transcription in the V_H region or V_H gene usage (Afshar et al., 2006; Perlot et al., 2005). Moreover, this enhancer does appear to be important for efficient Igμ-chain expression and strong signaling through the pre-BCR and BCR (Marquet et al., 2014)

4.1.2 The 3’ regulatory region—The 3’ regulatory region (3’RR), located 200kb downstream of the C_H cluster, spans 30kb and contains multiple enhancer elements with strict B-lineage specificity (HS3a, HS1–2, HS3b and HS4) and a proposed insulator region containing CTCF binding sites (HS5, 6, 7 and 8) which also bind cohesin and likely act as a *Igh* 3′ chromatin boundary (Degner et al., 2011; Garrett et al., 2005). These hypersensitive sites mostly show occupancy by transcription factors in mature B cells as this enhancer is not implicated in V(D)J regulation but controls CSR and somatic hypermutation, which take place in mature germinal center B cells after encounter with antigen (Khamlichi et al., 2000; Pinaud et al., 2011; Rouaud et al., 2013). In line with its role in these late events, the binding profile of Pax5 to the 3’RR is altered during CSR leading to enrichment on HS1–2, HS4 and HS7 (Chatterjee et al., 2011).

4.1.3 PAIR elements—The Busslinger lab identified 14 PAIR elements (Pax5-Activated Intergenic Repeats), within the distal V_H region that contain functional CTCF, E2A and Pax5 binding sites (Ebert et al., 2011). 11 out of the 14 PAIR elements are found immediately upstream of V_H3609 genes interspersed within the distal V_HJ558 gene family. Detailed investigation of PAIRS 4, 6 and 7 demonstrate binding of Pax5, E2A and CTCF in pro-B cells. In contrast, at the later pre-B cell stage there is depletion of Pax5 at these sites (Ebert et al., 2011). Pax5 binding at the pro-B cell stage correlates with the presence of antisense transcripts, that are distinct from those identified by the Corcoran lab (Bolland et al., 2004). These PAIR elements are implicated in locus contraction since Pax5 binding at the pro-B cell stage coincides with contraction, and Pax5 depletion in pre-B cells corresponds to a decontracted state (Roldan et al., 2005). It is of interest that neither of the *Igh* enhancers (Eμ and the 3’RR region) or the IGCR1 insulator site have any impact on locus contraction, suggesting that only elements within the V_H cluster are required (Medvedovic et al., 2013). The involvement of PAIRs in *Igh* contraction will need to be confirmed with genetic approaches that target these elements.

4.1.4 The intergenic control region 1 (IGCR1)—The intergenic control region 1 (IGCR1) is located within a 100kb-long intergenic region, which separates the V_H and D_H gene segments *of Igh*. It spans 4.3kb and lies between V_H81× (Ighv5-1) and DFL16.1 (2.1kb upstream of DFL16.1 also named Ighd1-1). IGCR1 consists of six hypersensitive (HS) sites (HS 1 to 6). Two conserved CTCF binding sites, HS4/5 that exhibit enhancer blocking activity, mark a sharp boundary of antisense transcription that stops at least 40kb from the V_H genes (Featherstone et al., 2010). In T cells and early pre-pro-B cells undergoing D_H-J_H rearrangement, the two CTCF sites separate regions of active and inactive chromatin in the D_H and V_H regions, respectively. Antisense transcription, which occurs at high level within this region in these cells is reduced in pro-B cells where V_H-DJ_H recombination takes place. Thus, it was proposed that the two CTCF sites act as an insulator preventing the spreading of chromatin activation and transcription into the V_H region during D_H-J_H rearrangement (Featherstone et al., 2010).

Deletion of the 4.1kb fragment (named the IGCR1) encompassing both CTCF binding elements (CBE1/2) alongside potential binding sites for other regulators (YY1 and PU.1) demonstrated that mutant alleles in a RAG deficient background were indeed associated

with upregulation of proximal V_H7183 and V_HQ52 transcripts and an enrichment of active histone marks in pro-B cells. Increased accessibility/transcription of proximal V_H genes is linked to preferential rearrangement of V_H7183 and V_HQ52 at the expense of distal V_HJ558 gene rearrangement in recombination competent IGCR1 targeted cells. Furthermore, mutant CBE alleles can undergo V_H-D_H rearrangement prior to D_H-J_H rearrangement, indicating a role for these elements in regulating ordered rearrangement (Guo et al., 2011). Additionally, mutant *Igh* alleles can undergo V_H-DJ_H recombination in developing thymocytes, in contrast to wild-type counterparts which normally only undergo D_H-J_H rearrangement. Thus the two CBE sites have a role in regulating lineage specific, ordered rearrangement as reviewed in (Chaumeil and Skok, 2012).

In more recent studies the Alt lab extended their analysis of the CBE sites by scrambling each element to separately assess their individual contributions to these processes (Lin et al., 2015). They demonstrate that scrambling of CBE1 but not CBE2 impacts allele expression such that in F1 mice harboring a 129 IgMa CBE1 mutated allele and a C57BL/6 IgMb WT allele, resulting B cells were found to express half as many IgMa compared to IgMb alleles. However, this defect was much more severe in F1 mice if both CBEs were mutated simultaneously. In line with their previous findings (Guo et al., 2011), mutation of either CBE1 or CBE2 led to a decrease in distal V_H gene rearrangement but this defect was more pronounced in CBE1$^{-/-}$ mice. Double CBE1/CBE2$^{-/-}$ mice however had the most severe defect.

These mutant mice also display defects in ordered rearrangement such that direct V_H-to-DJ_H joins were detected in CBE1$^{-/-}$ mice and variably in CBE2$^{-/-}$ mice, but again this was most pronounced in the double CBE1/CBE2$^{-/-}$ mutants. Finally, mutant CBE1 and 2 mice displayed low levels of proximal V_H-to-DJ_H rearrangements in T cells but lineage inappropriate rearrangement was much more severe in the double CBE1/CBE2$^{-/-}$ mice. No defects in allelic inclusion were observed in any of the three mutant mice. However, in line with what was previously observed in the double CBE1/CBE2$^{-/-}$ mutant mice (Guo et al., 2011), the presence of a productively rearranged V1–8 knockin allele did not suppress proximal V_H-to-DJ_H rearrangement. These rearrangements were detected in spleen on CBE2$^{-/-}$ alleles but more so on CBE1$^{-/-}$ alleles, however they were predominantly non-productive. No distal V_H gene rearrangements were observed likely because locus contraction was impaired as we previously showed in mice harboring a rearranged transgene that skip past the pro-B cell stage of development (Roldan et al., 2005).

Together these studies indicate that CBE1 has a more pronounced effect on ordered rearrangement and feedback control than CBE2 and the Alt lab suggest that this could be explained in two ways. First, CBE1 contains binding sites for PU.1 and YY1 and the presence of these binding sites could impact these functions. Second, and more interesting is the observation that the orientation of the CTCF site within CBE1 is in the opposite direction to the 60 V_H CTCF sites, while the CTCF site within CBE2 is in the opposite orientation to the 10 3’ CTCF sites (Figure 2), with the implication that CBE2 interacts with the 3’RR region, promoting D_H-to J_H rearrangement and inhibiting direct V_H-to-DJ_H joins. CBE1 on the other hand could interact with CBEs in V_H genes. It is of note that deletion of both CBE1/2 sites does not alter locus contraction as determined by 4C-seq from the Eμ

viewpoint, that measures interactions specifically with this region alone (Medvedovic et al., 2013). However, there has been no 4C-seq analyses from the viewpoint of CBE1/2 so we do not know what impact combined or individual deletion of these CTCF insulator sites has on surrounding interactions, and whether the change in chromatin boundaries that accompanies their deletion is matched by a change in interaction boundaries, as shown by us in the *Hoxa* locus (Narendra et al., 2015). A detailed analysis of looping in wild-type versus mutant CBE1 and 2 B cells could help resolve these issues. Furthermore, it would be interesting to find out what effect reversing the orientation of the two CBE sites has on regulation of the *Igh* locus.

4.2 Important *cis* acting elements and their function in *Igk* rearrangement

4.2.1 Enhancers—*Igk* possesses three powerful B cell-specific transcriptional enhancers: the matrix attachment region (MAR) and the intronic enhancer, iEκ, (together known as the MiEκ) are located between the Jκ and Cκ gene segments while two additional enhancers, 3'Eκ and Edκ, are found 8.5kb and 15.5kb downstream of the constant region (Liu et al., 2002; Meyer et al., 1990; Zhou et al., 2010). MiEκ and 3'Eκ are both important for rearrangement and deletion of either one leads to a reduction in the ratio of κ/λ expressing B cells, **w**hile the double mutant is sufficient to abrogate *Igk* recombination altogether (Inlay et al., 2002; Inlay et al., 2006). In contrast, an absence of both the 3'Eκ and Edκ leads to a dramatic reduction in germline and rearranged transcription, a reduction in active chromatin marks, increased DNA methylation and reduced levels of rearrangement. Furthermore, in mature cells IGλ is exclusively expressed on the cell surface despite functional rearrangement of *Igk*. This indicates that in the absence of both the 3'Eκ and Edκ the intronic enhancer is incapable of triggering *Igk* transcription (Zhou et al., 2010). Conditional knockout of the 3'Eκ in mature cells with an Edκ deletion leads to complete silencing of the *Igk* locus (Zhou et al., 2013). In these mice the mature B cells partially dedifferentiate, inducing RAG1/2 expression along with other pro-B cell makers and re-differentiate after triggering *Igl* gene rearrangement. These findings demonstrate that 3'Eκ and Edκ are essential for both the establishment and maintenance of transcriptional activity of *Igk*.

4.2.2 Promoters that influence Jκ usage—*Igk* germline transcription is initiated from two promoters located 150bp (proximal) and 3.5kb (distal) upstream of Jκ1 that give rise to the κ^0 transcripts (Schlissel, 2004) (Figure 3). The κ0 0.8 and κ^0 1.1 germ line transcripts are initiated from the proximal and distal promoters respectively and spliced to the Cκ region (Engel et al., 1999; Martin and van Ness, 1990). Germline transcription from these promoters is linked to rearrangement, and deletion of a 4kb region encompassing the two promoters has a marked impact on rearrangement of the allele bearing the deletion (Cocea et al., 1999). Recent studies from the Schlissel lab demonstrate a role for the proximal promoter in directing primary rearrangements to Jκ1, thereby ensuring the retention of other Jκ segments that can be used in subsequent rounds of recombination for receptor editing. They show that the distal but not the proximal promoter is active in both recombining and editing cells. Deletion of the proximal promoter leads to increased breaks on Jκ2 and decreased usage of Jκ1 and this correlates with an increase in H3K4me3 levels and transcription in the Jκ1 region. Thus the proximal promoter acts as a suppressor of accessibility and secondary recombination. Since it is inactive in recombining B cells the

Schlissel lab propose that it could be a result of promoter interference that is also found in *Tcra* where the active TEA suppresses activity of downstream Jα promoters (Abarrategui and Krangel, 2006, 2007).

4.2.3 Sis and Cer—Two additional regulator elements within *Igk*, Sis - hypersensitivity sites 3–6 (HS3–6) and Cer -hypersensitivity sites 1–2 (HS1–2), reside in the 18kb intervening Vκ-Jκ sequence (Liu et al., 2002). Sis (Silencer in the Intervening sequence) is a recombination silencer and heterochromatin targeting element. It binds both Ikaros and CTCF and directs the repositioning of *Igk* to PCH in pre-B cells (Liu et al., 2006). Deletion of Sis leads to reduced distal Vκ and enhanced proximal Vκ usage (Ribeiro de Almeida et al., 2011; Xiang et al., 2011). The neighboring CTCF binding site, Cer (Contracting element for recombination) plays a role in *Igk* locus contraction (Xiang et al., 2013). Like Sis, deletion of Cer increases proximal and diminishes distal Vκ usage although it has no impact on germline transcription or chromatin. Additionally, an absence of Cer leads to rearrangement of *Igk* in T cells. This is somewhat surprising since, unlike *Igh*, there is no evidence for *Igk* activation in T cells. Double deletion of both Cer and Sis gives rise to increased transcription of proximal Vκ in both pre-B and splenic B cells (Xiang et al., 2014). In this respect Cer and Sis behave in a similar manner to the CTCF binding elements, CBE1/2 in the *Igh* locus, although mutation of the IGCR1 does not appear to impact *Igh* locus contraction as determined by 4C-seq (Medvedovic et al., 2013) (see **§4.1.3** above), while in contrast DNA FISH analyses of *Igk* alleles with deleted Cer (or double deleted Cer and Sis) demonstrate a dramatic effect on *Igk* contraction (Xiang et al., 2013, 2014). It is difficult to compare the effects of the 4C-seq and DNA-FISH analyses in these two studies as neither give a complete picture of how interactions are altered across each locus in entirety. The 4C-seq analysis was performed from the Eμ viewpoint in IGCR1 mutated cells and this serves to highlight interactions from Eμ alone (Medvedovic et al., 2013) while the FISH analyses provide information on the distances separating three points on the *Igk* locus and offers no details of intra-locus interactions (Xiang et al., 2013, 2014).

4.2.4 Pre-BCR signaling and its impact on long-range interactions—Functional rearrangement of one *Igh* allele in pro-B cells leads to cell surface expression of the pre-BCR, which is comprised of IGH paired with surrogate light chain. Pre-BCR signaling in large pre-B cells drives proliferation and subsequent differentiation to the small pre-B cell stage where cells exit cell cycle and *Igk* rearrangement is initiated. To examine changes in *Igk* locus conformation by 4C–seq during the transition from the pro- to the pre-B cell stage, the Hendriks lab used pre-BCR signaling mutants of increasing severity (mice lacking *Btk*, *Slp65* or both together) on a RAG deficient background (Stadhouders et al., 2014). These analyses revealed that pre-BCR signaling reduces interactions of the three enhancers with *Igk* flanking sequences and increases interactions of the 3'Eκ with the Vκ regions, without altering Vκ interactions with the MiEκ (these are already in close contact at the pro-B cell stage). It is of note that in all cases the enhancers interact more frequently with functional versus non-functional Vκs in pre-B cells. The Sis element also displays an altered interaction pattern within the *Igk* locus in pro-B and pre-B cells, interacting much more with the proximal domain (JκCκ) in pro-B cells compared to pre-B cells. In the latter, the interaction profile spreads to the Vκ gene region, which may be a reflection of a change in

transcriptional activity although transcriptional profiles in pro-B cells are not shown in this study (Stadhouders et al., 2014). Vκ interactions correlate strongly with binding of E2A and Ikaros that are frequently found close to promoters and bind to the 3'Eκ, MiEκ and Sis regulatory elements (Bossen et al., 2012; Kil et al., 2012; Ribeiro de Almeida et al., 2011; Ribeiro de Almeida et al., 2012). Furthermore, interactions occur preferentially if both E2A and Ikaros are present together, versus Ikaros alone and the presence of both factors is linked to frequency of Vκ gene usage.

4.3 Important *cis* acting elements and their function in *Tcrb* rearrangement

4.3.1 The Eβ enhancer and promoters—Eβ is the sole known enhancer of *Tcrb*. It spans 550bp and is located 6kb downstream of the Cβ2 region and about 3kb upstream Vβ31 (Figure 4). Eβ facilitates activation of promoters flanking each of the two Dβ segments (McMillan and Sikes, 2008; Sikes et al., 1998). PDβ1, positioned immediately 5' of the Dβ1 12-RSS, was the first germline *Tcrb* promoter discovered. It uses a TATA element situated in the RSS spacer to initiate transcription at Dβ1. PDβ1 is bound by T cell-restricted transcription factors including SP1, GATA-3, and members of the ETS, RUNX and bHLH families. Most of these factors also bind Eβ (Doty et al., 1999; Sikes et al., 1998; Tripathi et al., 2000). Deletion of Eβ or PDβ1 dramatically affects T cell development in the following way. An Eβ deficiency gives rise to a similar phenotype as RAG deficiency in terms of *Tcrb* rearrangement. A total absence of germline transcription in the proximal DJCβ1-DJCβ2 domain leads to a failure of *Tcrb* rearrangement and a block in T cell development at the DN3 stage (Bories et al., 1996; Bouvier et al., 1996). In contrast, targeted deletion of PDβ1 specifically attenuates DJβ1 rearrangement without affecting DJβ2 and V-DJβ2 rearrangements (Whitehurst et al., 1999). These phenotypes demonstrate the importance of Eβ in modulating chromatin accessibility across both DJCβ regions, while the contribution of each promoter is specifically directed towards their associated DJβ clusters.

The DJβ2 comprise two promoters, one upstream (5'PDβ2) and one downstream (3'PDβ2) of the Dβ2 segment. Analogous to PDβ1, 5' PDβ2 is located immediately 5' to Dβ2 and binds GATA-3, RUNX1 and E47 (McMillan and Sikes, 2009). However, PDβ2 is inactive prior to Dβ2-Jβ2 recombination. Germline transcription at the DJβ2 cluster is driven by the NFκB dependant promoter 3'PDβ2, located several hundred bp downstream of Dβ2 (McMillan and Sikes, 2008). Repression of 5'PDβ2 is ensured by USF-1, a constitutively expressed bHLH protein that binds in the spacer of the Dβ2 12-RSS. It has recently been shown that the introduction of DNA DSBs relieves the USF-mediated repression of Dβ2 (Stone et al., 2012). Following DJβ recombination, 5'PDβ2 is activated and both DJβ clusters are transcribed and can rearrange with distant Vβ genes.

Little is known about the developmental regulation of Vβ promoters. They are responsible for germline as well as rearranged transcription of Vβ elements. Similarly they are involved in regulating recombinase accessibility as deletion of the Vβ31 promoter leads to a 10 fold decrease in Vβ31 rearrangement (Ryu et al., 2004). Unlike proximal promoters, Vβ promoters do not appear to require Eβ for their transcriptional activation in DN3 (Mathieu et al., 2000). However this enhancer can increase expression of the most highly transcribed subset of *Tcrb* Vβ segments in DN thymocytes.

4.3.2 Long range interactions—A recent paper from the Oltz lab investigated the role of enhancers and an insulator in shaping the interaction landscape of *Tcrb* that is so important for ensuring diversification of the *Tcrb* repertoire (Majumder et al., 2015). Using 3C (one to one interaction analysis) with an anchor on Eβ they show that an absence of this enhancer leads to reduced interactions with the rest of the locus, however interactions from this view-point to the mid Vβ gene region (Vβ12–13, Vβ14, Vβ16 and Vβ20) are maintained, although the 3C signal is lower than controls. In line with contraction analyses of other enhancer deficient antigen receptor loci, the Oltz lab find that deletion or inactivation of Eβ (through introduction of mutations in RUNX binding sites) does not disrupt the interaction between Dβ clusters and the Vβ gene segments despite ablation of germline transcription and reduced H3K4me3 levels in the region. This indicates that even transcription is dispensable for long-range interactions between Vβ, Dβ and Jβ gene segments. Nonetheless Eβ could still have an impact on Vβ gene repertoire as it alters germline transcription of a subset of Vβ genes, but this is not testable because Eβ is essential for activation and recombination of the DβJβ region.

Their studies demonstrate that the promoter PDβ1 is important for interactions between the Dβ2 region and distal Vβ genes because a 3.5kb deletion impacts the 3C signal on these Vβ genes when Dβ2 is used as an anchor. This deletion also reduces CTCF levels at distal Vβ genes without impacting transcription or cohesin levels. Reduced CTCF binding may explain alterations in Vβ interaction frequency with the proximal DβJβ domain. It is of note that the PDβ1 deletion does not alter interactions from the distal Vβ5 viewpoint. In addition, proximal Vβ gene interactions with the DβJβ domain do not require the PDβ1, however interactions between proximal and distal Vβ genes do. In contrast, a minimal deletion of the PDβ1 promoter does not impact long-range Vβ to DβJβ interactions, despite its impact on Dβ1 transcription. Interactions between the distal Vβs and the DβJβ thus rely on a 3kb region just upstream of the PDβ1 promoter. However distal V gene interactions occur most robustly with the 5'Prss2-CTCF (5'PC) site, which is intact in the 3.5kb deleted promoter allele and furthermore CTCF binding is not altered at the 5'PC if this region is deleted. The 5'PC can be distinguished by a 5' repetitive tract (which contains a viral LTR that is expressed at low levels in DN cells that harbors insulator properties) and a pair of CTCF/RAD21 binding sites. The Prss2 gene is normally inactive in WT, minimal PDβ1 deletion or Eβ mutated alleles. However it is activated if the entire 3.5kb promoter is deleted and the chromatin around the promoter region is enriched for H3K4me3. This mark spreads from the PDβ1 and PDβ2 region all the way up to the 5'PC in the PDβ1 mutant suggesting that a chromatin boundary has been disrupted. This chromatin barrier appears to be required for mediating interactions between the distal Vβ gene segments (where, in contrast to the proximal domain there is robust CTCF binding) and the PDβ region. Future studies will be required to identify the transcription factors that are involved in this interaction. Certainly, it is clear that these insulator sites are a common feature of antigen receptor loci as they are found in *Tcrb*, *Igk* and *Igh*. In the case of *Igk* and *Igh*, deletion of these elements increases transcription and recombination in the proximal domain, perhaps disrupting interactions with distal Vβ gene segments that have not yet been resolved with current methods of analysis (see **§6.2**).

4.4 Important *cis* acting elements and their function in *Tcra/d* rearrangement

4.4.1 Enhancers and promoters—There are two enhancers in the *Tcra/d* locus: Eδ and Eα regulate *Tcrd* and *Tcra* rearrangement, respectively (Figure 5) (Krangel, 2009). The Eδ, which is located between the Jδ and Cδ gene segments, regulates germline transcription of the promoter pDδ in the DJδ region. The Eδ functions locally in adult DN cells and its deletion Eδ reduces *Tcrd* rearrangement by ten fold although it does not alter accessibility of the Vδ genes (Hao and Krangel, 2011). In addition, Eδ appears to be dispensable for *Tcrd* expression after rearrangement (Monroe et al., 1999).

The Eα is essential for *Tcra* rearrangement and it also regulates expression from rearranged *Tcrd* alleles (Krangel, 2009; Monroe et al., 1999). Deletion of Eα blocks *Tcra* rearrangement and T cell development (Sleckman et al., 1997). Eα regulates a “T early α” promoter (TEAp) located just upstream of the Jα array, and this in turn activates the Jα array. Specifically the TEA promoter targets primary rearrangement at the extreme 5’ of the Jα array by opening up the RSSs of these genes (Hawwari et al., 2005). This maximizes the use of Jα segments during secondary rearrangement as use of the 3’ Jαs in the first round of rearrangement could result in deletion of intervening segments leaving few substrates for subsequent recombination events. The Eα also regulates germline transcription and accessibility of the proximal Vα genes via long-range interactions (over 500kb) (Hawwari et al., 2005). In addition, the Eα mediates interactions between the proximal Vα and Jα gene segments, which are essential for synapsis and rearrangement (Shih et al., 2012).

As with the other loci, long distance interactions between *cis*-elements are essential for *Tcra* and *Tcrd* rearrangement. Chromatin organizers like Cohesin and CTCF play an important role in mediating long range interactions (Seitan et al., 2011; Shih et al., 2012) and reviewed in (Chaumeil and Skok, 2012). Cohesin binds to the *Tcra* locus control region (LCR), the Eα enhancer, the Jα49 promoter, the TEAp, sites located between *Tcrd* and the first Vα segments and to Vα gene promoters in DP cells. Deletion of Rad21, one of the cohesin complex components, impairs the interaction between the Eα and TEAp, which in turn impacts activation of distal 3’ Jα genes and impairs secondary *Tcra* rearrangements (Seitan et al., 2011). CTCF also mediates interaction between the Eα and TEAp and binds to the proximal Vα gene promoters, which may assemble a rosette with Vα, Eα and Jα (Shih et al., 2012).

5. ALLELIC EXCLUSION

Antigen receptors are expressed from only one allele in individual lymphocytes to ensure unique receptor specificity. This is fundamental to the proper functioning of the adaptive immune response, which relies on clonal expansion of lymphocytes expressing receptors that specifically recognize an invading pathogen. Elucidating the mechanisms underlying monospecific receptor expression - allelic exclusion - has proved to be a challenging puzzle, likely because the process involves multiple levels of control. However, allelic exclusion is not infallible as dual *Tcr* or *Ig* receptor expressing T and B cells are found at low frequency in the periphery (Fournier et al., 2012; Pelanda, 2014). Although tolerance mechanisms exist to restrain dual receptor cells with a self-reactive receptor (Fournier et al., 2012) these cells can become activated and cause autoimmune disease if the non-self-reactive receptor

recognizes a pathogenic antigen (Ercolini and Miller, 2009; Flodstrom-Tullberg, 2003; Ji et al., 2010; Pelanda, 2014). Nonetheless, the effects of these dual receptor cells are not altogether negative as their presence can be beneficial in counteracting infection because allelically included cells expand the receptor repertoire and may in some instances be important for combatting an invading pathogen (He et al., 2002); Thus evolution may tolerate a certain frequency of allelically included dual receptor cells, balancing an autoimmune outcome with that of counteracting infection (Figure 10).

All the antigen receptor loci are regulated in a unique manner and in particular they all are subjected to different controls when it comes to allelic inclusion. The *Igh* locus is subject to stringent allelic exclusion and only 2–4% of mouse spleen B cells contain two in-frame rearrangements with 0.01% expressing dual receptors. *Igk* and *Tcrb* are also subject to fairly stringent controls and allelic inclusion occurs at a frequency of 1–7% and 1–3%, respectively. *Tcra* on the other hand, can rearrange both alleles prior to differentiation, but the frequency of allelic inclusion on the cell surface is only due to 10%, which may be due to post-translation control (Brady et al., 2010).

As mentioned above, allelic exclusion ensures the expression of only one productively rearranged allele (Jung et al., 2006; Vettermann and Schlissel, 2010). The other allele can be non functional for one of three reasons: (i) it remains in germline configuration (*Igk* or *Tcra*) or is partially recombined having undergone D-to-J but not V-to-DJ rearrangement (*Tcrb* or *Igh*), (ii) the allele has an out of frame rearrangement and the mRNA is degraded by the nonsense mediated decay (NMD) pathway, (iii) the allele encodes a protein that cannot pair with its partner (ie *Igh* with *Igk* or *Tcrb* with *Tcra*) and thus a receptor cannot be assembled on the surface. In this way allelic exclusion is very different from other well known mono-allelically expressed genes such as olfactory receptors or those resulting from X inactivation, or parental imprinting.

As a general rule, allelic exclusion is enforced during the process of V(D)J rearrangement (Figure 11). However, in some cells with dual rearrangements, the product of only one allele is expressed at the cell surface as a result of post-translational silencing, and in this case allelic exclusion is enforced by a later event (Alam and Gascoigne, 1998).

5.1 Asynchronous rearrangement

Early models proposed that asynchronous recombination occurred as a result of low efficiency recombination (RAG breaks are introduced in around 20% or less cells at any one time) which reduces the chances of rearrangement occurring on the two alleles at the same time. Added to this, the imprecise nature of junctions results in a high failure rate of rearrangements (two out of three will be non-productive) and this in itself will contribute to the initiation of allelic exclusion. Whilst these facts are indisputable it is also now well established that breaks are introduced in a regulated asynchronous manner on all antigen receptor alleles analyzed (*Igh*, *Igk* and *Tcra*) (figure 8) (Chaumeil et al., 2013a; Chaumeil and Skok, 2013; Hewitt et al., 2009). Rearrangement on one allele at a time involves regulation *in trans* and allelic communication, which may or may not be reliant on pairing (discussed in **§3.3**). As described in **§3.3**, the introduction of RAG-mediated cleavage on one allele recruits ATM to the site of the break and this acts in *trans* on the other allele

preventing the introduction of further breaks by a mechanism that involves repositioning of the other allele to repressive pericentromeric heterochromatin and curtailment of higher-order looping (Figure 8).

DNA FISH analyses has revealed that in most cells the two *Igh* and *Tcra* alleles are both located in euchromatic regions of the nucleus prior to the onset of recombination, and thus by this means of assessment homologues appear to be equivalently accessible to RAG. Despite this observation, our data suggest that mono-allelic targeting of *Igh* and *Tcra* occurs preferentially on highly transcribed alleles that are looped outside of their respective chromosome territories and that for both these loci loop formation occurs on only one allele at a time (Chaumeil et al., 2013a; Chaumeil et al., 2013b). We have not yet examined RAG targeting of *Igk* and *Tcrb* in the context of higher-order loop formation, however we do know that in contrast, to *Igh* and *Tcra*, prior to recombination, one *Igk* and *Tcrb* allele are associated with repressive pericentomeric heterochromatin and / or the nuclear lamina in pre-B and DN T cells respectively, while the other allele is found in a euchromatic location where RAG targeting occurs (Roldan et al., 2005; Schlimgen et al., 2008; Skok et al., 2007). Thus differential accessibility of the two alleles may play a role in determining which allele is targeted (Figure 12). However these studies provide no information about whether differential positioning of the two alleles is heritably transmitted or whether the two alleles are equally likely to find themselves in opposite locations in the same, or a subsequent cell cycle.

5.1.1 Replication timing—Studies from the Bergman lab support a deterministic model of accessibility that relies on the observation that homologous antigen receptor alleles are asynchronously replicated (Mostoslavsky et al., 2001). As a general principal, early replicating loci are more active than late replicating loci. Thus, based on this premise, differences in replication timing likely reflect differences in activation status of homologues and differences in accessibility that may predispose one allele to recombine before the other. According to their data, allelic choice is a random process that mirrors the process of X inactivation. Through lineage tracing experiments they determined that allelic choice (which correlates with differences in replication timing), could be imposed early on in lymphoid development at the common lymphoid progenitor (CLP) stage; a subgroup of single CLP cells gives rise to mature B cells that all express *Igk* from the same allele (Farago et al., 2012). Thus commitment occurs prior to rearrangement, but once allelic differences are imposed they are heritably transmitted through development.

5.1.2 The impact of non-productive rearrangements—A recent study by the Barreto lab presents data that disagree with the Bergman lab's findings (Alves-Pereira et al., 2014). In this study clonal analysis of reconstituted single IgMa/IgMb heterozygous hematopoietic stem cells (HSCs) in irradiated RAG deficient recipients consistently generated equal numbers of IgMa and IgMb expressing B cells in each animal. Moreover, PCR analysis showed the expected differences in the retention of a V_H-D_H intergenic fragment (60% in the VDJ_H/DJ_H and 40% in VDJ_H/VDJ_H configuration). In contrast CLP-derived clones were completely skewed to either IgMa or IgMb expressing cells and highly skewed clones were found more frequently in Ly6d+ B cell progenitors compared to the more uncommitted

Ly6d- cells. Furthermore, Ly6d+ cells and pro-B cells have a similar capability to skew clones. Analyses of *Igh* rearrangement status on both the productive and silent alleles in skewed clones indicate that the bias to rearrange one allele can be explained by the impact of non-productive rearrangement. Thus, their findings support the idea that the two *Igh* alleles are synchronously competent to undergo rearrangement. Furthermore, in contrast to the Bergman lab's findings, they demonstrate that the bias observed for *Igh* is not reproduced for *Igk* suggesting that this locus is not pre-committed in the CLP stage of B cell development. Taken together, these data suggest that allelic exclusion of *Ig* loci differs from X-chromosome inactivation as no stable epigenetic mark is propagated until pro-B cells start rearranging. The key difference in the Barreto and Bergman lab's studies is that the Bergman lab did not analyze the rearrangement status of the silent allele.

5.2 Feedback Inhibition

5.2.1 Feedback inhibition through the introduction of a DSB break—As discussed in **§3.3** above, feedback inhibition occurs at the level of breaks (Figure 8 and 11). A break in one allele or locus inhibits further breaks during repair as a result of ATM and RAG2 mediated control. In the absence of the C terminus of RAG2 and ATM bi-allelic and bi-locus breaks are introduced and this can lead to the generation of intra-locus translocations (Chaumeil et al., 2013a; Chaumeil et al., 2013b; Chaumeil and Skok, 2013; Deriano et al., 2011; Hewitt et al., 2009 6058) which are a hallmark of ATM deficiency. Clearly controlling the number of breaks that are introduced per cell at any one time is important for maintenance of genome stability and thwarting the occurrence of translocations. However feedback control also contributes to the initiation of allelic exclusion by preventing the simultaneous rearrangement of homologues that could lead to allelic inclusion. Support for this comes from the observation that ATM deficient mice have increased allelic inclusion of *Igh, Igk* and *Tcrb* (Steinel et al., 2014; Steinel et al., 2013). The C terminus of RAG2 and ATM also inhibit the introduction of bi-allelic breaks on *Tcra* (Chaumeil et al., 2013a) even though this locus is not subjected to stringent enforcement of allelic exclusion. Thus ATM mediated control of cleavage appears to be a common mechanism that is shared by different loci in recombining lymphocytes as well as in cells undergoing meiosis (see **§3.3**) (Lange et al., 2011).

5.2.2 Control of recombination via regulation of RAG expression – implications for allelic exclusion—It is clearly critical to have mechanisms in place to control RAG activity to ensure that cleavage does not occur across cell cycle as this could lead to genome instability. Productive rearrangement of an *Igh* or *Tcrb* allele in pro-B or DN cells, respectively leads to cell surface expression of the pre-BCR or pre-TCR. Signaling through these two receptors results in a proliferative burst and subsequent differentiation to the pre-B or DP cell stage where *Igk, Igl* or *Tcra* are recombined. There are two known mechanisms that have evolved to prevent the introduction of breaks during cell cycle. The first involves degradation of RAG2 protein (Lee and Desiderio, 1999) and the second involves control of *Rag1* expression (Johnson et al., 2012). Both mechanisms are also important for preventing the introduction of further breaks on the second allele, which could violate allelic exclusion.

5.2.3 Feedback Inhibition by productive mRNA—Two thirds of transcripts generated by rearrangement are out-of-frame. In contrast to mRNAs from productively rearranged alleles, these are degraded by the nonsense-mediated mRNA decay (NMD) pathway, which selectively degrades transcripts harboring premature termination codons (Figure 11) (Weischenfeldt et al., 2008). Recent studies suggest that mRNA from productively rearranged alleles can have a role in suppressing rearrangement and initiating allelic exclusion. The evidence for this comes from two complementary mouse models. The first harbors a dominant-negative mutation of Rent1/hUpf1, an essential *trans*-effector of the NMD pathway. This mutation induces premature shut-off of *Tcrb* rearrangement causing a block in T cell development at the DN stage. This defect can be rescued with a productively rearranged *Tcrb* transgene, suggesting that mRNA has a function in V(D)J recombination independent of its protein product (Frischmeyer-Guerrerio et al., 2011).

Further support for a negative regulatory role for mRNA in recombination comes from a mouse model that has the endogenous $DQ52J_H$ cluster replaced by a $V_HB1–8$ VDJ exon rendered nonproductive by the introduction of a termination codon at position 5 on one allele (Ter5 allele) (Lutz et al., 2011). Transcription of the targeted *Igh* allele is driven by its physiological *Igh* chain promoter in one mouse line (Ter5hi), while in the other transcription is driven by a weak, truncated DQ52 promoter (Ter5lo). This results in the production of Ter5 high and low amounts of stable *Igh* transcripts, respectively that do not encode protein. Thus, stable Igμ mRNA is separated from translation into IGH protein. The presence of stable Ter5hi transcripts leads to a severe block in B cell differentiation at the pro-B cell stage and a corresponding decrease in the pre-B cell compartment. Importantly, the block in development is linked to a decreased frequency of recombined *Igh* alleles, while the RAG recombinase remains unaffected (Lutz et al., 2011) (Figure 11). Recombination of the wild type allele is inhibited in the Ter5hi heterozygous cells, preventing the generation of a productively rearranged *Igh* allele that could drive development to the pre-B cell stage. In contrast, there is a significantly higher frequency of *Igh* rearrangement on the non-targeted allele in the heterozygous Ter5lo mice. Thus it appears that the difference in mRNA stability allows pro-B cells to distinguish between productive and non productive *Ig* gene rearrangements and that the presence of stable Igμ transcripts contributes to *Igh* chain allelic exclusion.

5.2.4 Feedback Inhibition at the level of protein—It has been known for some time that cell surface expression of a productively rearranged IGH prevents ongoing rearrangement on remaining DJ_H rearranged *Igh* alleles (Figure 12). This is in large part because productive rearrangement drives development forward and any antigen receptor locus moving to a new developmental compartment will be subject to changes in signaling pathways and transcriptional profiles that do not support its continued accessibility and ongoing rearrangement. This is exemplified in mice that express a pre-rearranged *Igh* (knockin or transgenic) allele, which drives B cell development forward skipping out the pro-B cell stage where *Igh* recombination normally takes place. As a result B cells reach the pre-B cell stage without an opportunity to fully open up V_H genes via IL-7/STAT5 signaling (Bertolino et al., 2005; Chowdhury and Sen, 2003) and without undergoing locus contraction (Roldan et al., 2005). At the pre-B cell stage there is a reduction in IL-7/STAT5

signaling and Pax5 binding at PAIRs is reduced to an extent that may impact the ability of the *Igh* locus to contract (Ebert et al., 2011). Indeed, we found that rearrangement on decontracted endogenous *Igh* loci in these mice is limited to the 4 most proximal V_H gene segments (Roldan et al., 2005).

Further evidence for this idea comes from mice with mutations in components of the pre-BCR and pre-TCR signaling pathways that result in a partial or total block in B and T cell development that forces cells to remain in the pro-B or DN cell stages, respectively. In this situation the presence of mRNA or protein originating from productive *Igh* or *Tcrb* rearrangement is not sufficient to prevent ongoing rearrangement on accessible alleles that can be targeted by RAG leading to a violation of exclusion. To some extent staying in the same compartment indefinitely may help to overcome constraints that are not 100% efficient at preventing recombination on the second allele.

5.3 Maintenance of allelic exclusion

As summarized above, there are multiple levels at which allelic exclusion is enforced starting with differential accessibility of the two alleles (which may or may not be heritably predetermined), and moving on to feedback control at the level of (i) ATM and RAG2 mediated regulation of asynchronous cleavage, (ii) stable mRNA production and (iii) protein production. Protein expression of a functionally rearranged *Igh* or *Tcrb* drives differentiation forward and at the subsequent stages of development these loci are subjected to different signaling pathways and transcriptional factor profiles that alter their accessibility to maintain allelic exclusion. It is particularly important for the *Igh* or *Tcrb* loci that accessibility is reduced in pre-B and DP cells as RAG expression is once again up-regulated and this could potentially target unrearranged alleles during *Igk/Igl* and *Tcra* recombination, respectively that could then lead to allelic inclusion. There are many transcription factors involved in regulating rearrangement but only a few (for example E2A) which are reported to function in allelic exclusion (Hauser et al., 2014). In DN thymocytes E2A binds to the *Tcrb* DJβ region, Eβ and some Vβ gene promoters and activates germ-line transcription (Belle and Zhuang, 2014). The E2A inhibitor, ID3 is upregulated downstream of pre-TCR signaling and E2A binding to *Tcrb* is reduced, which in turn reduces accessibility. In contrast, enforced expression of E2A in thymocytes overrides allelic exclusion in mice expressing a rearrarranged *Tcrb* transgene (Agata et al., 2007). As mentioned in **§5.4.2** above, IL-7/STAT5 signaling has been implicated in regulating allelic exclusion of the *Igh* locus. Thus, in mice expressing constitutively active STAT5a, accessibility of the *Igh* locus is maintained in pre-B cells supporting ongoing rearrangement (Hewitt et al., 2009).

In addition to changes in accessibility, *Igh* and *Tcrb* both undergo decontraction in pre-B and DP cells, respectively but this is not sufficient to block rearrangement on proximal V genes. The lineage specific factors that induce locus contraction and decontraction of *Tcrb* have not been identified. In contrast, this is much better understood for *Igh* (see **§3.4.4**) although there are details that need further clarification. We know for example that a 3'Eκ mediated interaction between *Igh* and *Igk* at the pre-B cell stage is important for relocating unrearranged *Igh* alleles to repressive pericentromeric heterochromatin and for inducing locus decontraction (Figure 9). Going forward it will be important to determine if the 3'Eκ

enhancer is important for reducing Pax5 binding to PAIR elements at the pre-B cell stage. Currently, it is not known why Pax5 binding to PAIR elements is altered in cells where Pax5 expression remains at high levels.

In conclusion, V(D)J recombination is tightly regulated at multiple levels in order to limit the possible hazards associated with the introduction of DSBs. As discussed in this section many of the mechanisms controlling accessibility and cleavage also contribute to allelic exclusion. Finally, although great progress has been made in recent years we still have some way to go before we fully understand the process.

6. FUTURE DIRECTIONS

6.1 Long range interactions in V(D)J recombination

Long range interactions appear to be an important component of V(D)J recombination (Skok, 2010) and studies that have focused on this aspect of control have contributed to our understanding of gene regulation as a whole. Indeed, recombination involves (i) intra-locus interactions that are important for the generation of repertoire diversity, (ii) homologue and heterologue pairing which is linked to feedback control of cleavage via a mechanism that involves ATM and RAG2, and (iii) enhancer-mediated inter-loci interactions that is important for reducing accessibility and inducing locus decontraction after productive rearrangement (between *Igh* and *Igk* for example, see §3.4). For now these are the ones we know about, but it is likely that future studies will reveal other long range interactions of relevance.

6.2 A holistic approach for analyzing interactions

Of these long range interactions, intra-locus interactions are the most well-studied as demonstrated by the multitude of papers that have been published on this topic. Nonetheless, we still do not have a holistic understanding of the structure of each locus and the impact of *cis* and *trans* acting factors on looping. This is because most studies reported here have analyzed locus conformation at the molecular level using 3C or 4C-seq. These approaches provide limited information about the structure of the locus as a whole because interactions are only examined from a particular viewpoint and this gives no information on TAD structure. In addition, the available Hi-C data is too low resolution and does not provide sufficient information on looping. For a complete picture, a Capture C or High resolution Hi-C (Rao et al., 2014) approach will be required to generate a high resolution matrix of inter and intra-TAD interactions. Next generation approaches that are currently emerging (Hughes et al., 2013; Kolovos et al., 2014) and constantly being improved upon will be useful for providing the information we are lacking. Furthermore, use of these approaches will allow us to compare one locus with another and to determine the impact of regulatory elements on structure and regulation using gene targeted mouse models, information that we currently do not have.

6.3 The impact of RAG on organization of antigen receptor loci

Finally, RAG, which is such a fundamental component of recombination, is largely eliminated in 3C and 4C analyses because of the impact of recombination on locus structure.

However, we cannot rule out that RAG itself has no influence on locus conformation as we already know from our studies that RAG has a significant impact on the organization of each locus. For example we know that RAG is required for pairing of homologous and heterologous antigen receptor alleles and for bringing RAG bound loci together in the nucleus (Brandt et al., 2010; Chaumeil et al., 2013a; Chaumeil et al., 2013b; Chaumeil and Skok, 2013; Hewitt et al., 2009). In addition we have shown that the presence of RAG is important for inducing the formation of higher-order loops that separate the 3' end of the locus from the chromosome territory which is linked to enrichment of RAG and active histone modifications as well as directed cleavage in this region. Thus future studies should not ignore the impact of this important player.

Acknowledgements

JAS is a Leukemia & Lymphoma Society (LLS) scholar and is supported by NIH grants R01 GM086852, NIAID R56 A1099111 and R01GM112192.

REFERENCES

Abarrategui I, Krangel MS. Regulation of T cell receptor-alpha gene recombination by transcription. Nature Immunology. 2006; 7:1109–1115. [PubMed: 16936730]

Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor alpha recombination. EMBO J. 2007; 26:4380–4390. [PubMed: 17882258]

Afshar R, Pierce S, Bolland DJ, Corcoran A, Oltz EM. Regulation of IgH gene assembly: role of the intronic enhancer and 5'DQ52 region in targeting DHJH recombination. J Immunol. 2006; 176:2439–2447. [PubMed: 16456003]

Agata Y, Tamaki N, Sakamoto S, Ikawa T, Masuda K, Kawamoto H, Murre C. Regulation of T cell receptor beta gene rearrangements and allelic exclusion by the helix-loop-helix protein, E47. Immunity. 2007; 27:871–884. [PubMed: 18093539]

Alam SM, Gascoigne NR. Posttranslational regulation of TCR Valpha allelic exclusion during T cell differentiation. J Immunol. 1998; 160:3883–3890. [PubMed: 9558094]

Alves-Pereira CF, de Freitas R, Lopes T, Gardner R, Marta F, Vieira P, Barreto VM. Independent recruitment of Igh alleles in V(D)J recombination. Nat Commun. 2014; 5:5623. [PubMed: 25517887]

Arakawa H, Shimizu T, Takeda S. Re-evaluation of the probabilities for productive arrangements on the kappa and lambda loci. International Immunology. 1996; 8:91–99. [PubMed: 8671593]

Belle I, Zhuang Y. E proteins in lymphocyte development and lymphoid diseases. Curr Top Dev Biol. 2014; 110:153–187. [PubMed: 25248476]

Bertolino E, Reddy K, Medina KL, Parganas E, Ihle J, Singh H. Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. Nat Immunol. 2005; 6:836–843. [PubMed: 16025120]

Bolland DJ, Wood AL, Johnston CM, Bunting SF, Morgan G, Chakalova L, Fraser PJ, Corcoran AE. Antisense intergenic transcription in V(D)J recombination. Nat Immunol. 2004; 5:630–637. [PubMed: 15107847]

Bories JC, Demengeot J, Davidson L, Alt FW. Gene-targeted deletion and replacement mutations of the T-cell receptor beta-chain enhancer: the role of enhancer elements in controlling V(D)J recombination accessibility. Proc Natl Acad Sci U S A. 1996; 93:7871–7876. [PubMed: 8755569]

Bossen C, Mansson R, Murre C. Chromatin topology and the regulation of antigen receptor assembly. Annu Rev Immunol. 2012; 30:337–356. [PubMed: 22224771]

Boudil A, Matei IR, Shih HY, Bogdanoski G, Yuan JS, Chang SG, Montpellier B, Kowalski PE, Voisin V, Bashir S, et al. IL-7 coordinates proliferation, differentiation and Tcra recombination during thymocyte beta-selection. Nat Immunol. 2015

Bouvier G, Watrin F, Naspetti M, Verthuy C, Naquet P, Ferrier P. Deletion of the mouse T-cell receptor beta gene enhancer blocks alphabeta T-cell development. Proc Natl Acad Sci U S A. 1996; 93:7877–7881. [PubMed: 8755570]

Brady BL, Steinel NC, Bassing CH. Antigen receptor allelic exclusion: an update and reappraisal. J Immunol. 2010; 185:3801–3808. [PubMed: 20858891]

Brandt VL, Hewitt SL, Skok JA. It takes two: Communication between homologous alleles preserves genomic stability during V(D)J recombination. Nucleus. 2010; 1:23–29. [PubMed: 21327101]

Brown KE, Baxter J, Graf D, Merkenschlager M, Fisher AG. Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. Mol Cell. 1999; 3:207–217. [PubMed: 10078203]

Chatterjee S, Ju Z, Hassan R, Volpi SA, Emelyanov AV, Birshtein BK. Dynamic changes in binding of immunoglobulin heavy chain 3' regulatory region to protein factors during class switching. J Biol Chem. 2011; 286:29303–29312. [PubMed: 21685395]

Chaumeil J, Micsinai M, Ntziachristos P, Deriano L, Wang JM, Ji Y, Nora EP, Rodesch MJ, Jeddeloh JA, Aifantis I, et al. Higher-Order Looping and Nuclear Organization of Tcra Facilitate Targeted RAG Cleavage and Regulated Rearrangement in Recombination Centers. Cell reports. 2013a

Chaumeil J, Micsinai M, Ntziachristos P, Roth DB, Aifantis I, Kluger Y, Deriano L, Skok JA. The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. Nat Commun. 2013b; 4:2231. [PubMed: 23900513]

Chaumeil J, Skok JA. The role of CTCF in regulating V(D)J recombination. Current Opinion in Immunology. 2012; 24:153–159. [PubMed: 22424610]

Chaumeil J, Skok JA. A new take on v(d)j recombination: transcription driven nuclear and chromatin reorganization in rag-mediated cleavage. Frontiers in immunology. 2013; 4:423. [PubMed: 24367365]

Chowdhury D, Sen R. Transient IL-7/IL-7R signaling provides a mechanism for feedback inhibition of immunoglobulin heavy chain gene rearrangements. Immunity. 2003; 18:229–241. [PubMed: 12594950]

Ciofani M, Zuniga-Pflucker JC. Determining gammadelta versus alphass T cell development. Nature reviews Immunology. 2010; 10:657–663.

Cocea L, De Smet A, Saghatchian M, Fillatreau S, Ferradini L, Schurmans S, Weill JC, Reynaud CA. A targeted deletion of a region upstream from the Jkappa cluster impairs kappa chain rearrangement in cis in mice and in the 103/bcl2 cell line. J Exp Med. 1999; 189:1443–1450. [PubMed: 10224284]

Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, Vettermann C, Lin YC, Ju Z, Schulz D, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. ProcNatlAcadSciUSA. 2011; 108:9566–9571.

Deriano L, Chaumeil J, Coussens M, Multani A, Chou Y, Alekseyenko AV, Chang S, Skok JA, Roth DB. The RAG2 C terminus suppresses genomic instability and lymphomagenesis. Nature. 2011; 471:119–123. [PubMed: 21368836]

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485:376–380. [PubMed: 22495300]

Doty RT, Xia D, Nguyen SP, Hathaway TR, Willerford DM. Promoter element for transcription of unrearranged T-cell receptor beta-chain gene in pro-T cells. Blood. 1999; 93:3017–3025. [PubMed: 10216098]

Ebert A, McManus S, Tagoh H, Medvedovic J, Salvagiotto G, Novatchkova M, Tamir I, Sommer A, Jaritz M, Busslinger M. The distal V(H) gene cluster of the Igh locus contains distinct regulatory elements with Pax5 transcription factor-dependent activity in pro-B cells. Immunity. 2011; 34:175–187. [PubMed: 21349430]

Engel H, Rolink A, Weiss S. B cells are programmed to activate kappa and lambda for rearrangement at consecutive developmental stages. European Journal of Immunology. 1999; 29:2167–2176. [PubMed: 10427979]

Ercolini AM, Miller SD. The role of infections in autoimmune disease. Clinical and experimental immunology. 2009; 155:1–15. [PubMed: 19076824]

Farago M, Rosenbluh C, Tevlin M, Fraenkel S, Schlesinger S, Masika H, Gouzman M, Teng G, Schatz D, Rais Y, et al. Clonal allelic predetermination of immunoglobulin-kappa rearrangement. Nature. 2012; 490:561–565. [PubMed: 23023124]

Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. The Journal of biological chemistry. 2010; 285:9327–9338. [PubMed: 20100833]

Feddersen RM, Martin DJ, Van Ness BG. Novel recombinations of the IG kappa-locus that result in allelic exclusion. J Immunol. 1990; 145:745–750. [PubMed: 2114448]

Flodstrom-Tullberg M. Viral infections: their elusive role in regulating susceptibility to autoimmune disease. Microbes and infection / Institut Pasteur. 2003; 5:911–921. [PubMed: 12919859]

Fournier EM, Velez MG, Leahy K, Swanson CL, Rubtsov AV, Torres RM, Pelanda R. Dual-reactive B cells are autoreactive and highly enriched in the plasmablast and memory B cell subsets of autoimmune mice. J Exp Med. 2012; 209:1797–1812. [PubMed: 22927551]

Frischmeyer-Guerrerio PA, Montgomery RA, Warren DS, Cooke SK, Lutz J, Sonnenday CJ, Guerrerio AL, Dietz HC. Perturbation of thymocyte development in nonsense-mediated decay (NMD)-deficient mice. Proc Natl Acad Sci U S A. 2011; 108:10638–10643. [PubMed: 21670277]

Fuxa M, Skok J, Souabni A, Salvagiotto G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev. 2004; 18:411–422. [PubMed: 15004008]

Garcia V, Gray S, Allison RM, Cooper TJ, Neale MJ. Tel1-mediated interference suppresses clustered meiotic double-strand-break formation. Nature. 2015

Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, Loukinov D, Eckhardt LA, Lobanenkov VV, Birshtein BK. Chromatin architecture near a potential 3' end of the igh locus involves modular regulation of histone modifications during B-Cell development and in vivo occupancy at CTCF sites. Molecular and Cellular Biology. 2005; 25:1511–1525. [PubMed: 15684400]

Giudicelli V, Duroux P, Ginestoux C, Folch G, Jabado-Michaloud J, Chaume D, Lefranc MP. IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucleic Acids Res. 2006; 34:D781–D784. [PubMed: 16381979]

Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng HL, Hansen E, Despo O, Bossen C, Vettermann C, et al. CTCF-binding elements mediate control of V(D)J recombination. Nature. 2011; 477:424–430. [PubMed: 21909113]

Halverson R, Torres RM, Pelanda R. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. Nat Immunol. 2004; 5:645–650. [PubMed: 15156139]

Hao B, Krangel MS. Long-distance regulation of fetal V(delta) gene segment TRDV4 by the Tcrd enhancer. J Immunol. 2011; 187:2484–2491. [PubMed: 21784972]

Hauser J, Grundstrom C, Grundstrom T. Allelic exclusion of IgH through inhibition of E2A in a VDJ recombination complex. J Immunol. 2014; 192:2460–2470. [PubMed: 24470503]

Hawwari A, Bock C, Krangel MS. Regulation of T cell receptor alpha gene assembly by a complex hierarchy of germline Jalpha promoters. Nature Immunology. 2005; 6:481–489. [PubMed: 15806105]

Hawwari A, Krangel MS. Regulation of TCR delta and alpha repertoires by local and long-distance control of variable gene segment chromatin structure. J Exp Med. 2005; 202:467–472. [PubMed: 16087716]

He X, Janeway CA Jr, Levine M, Robinson E, Preston-Hurlburt P, Viret C, Bottomly K. Dual receptor T cells extend the immune repertoire for foreign antigens. Nat Immunol. 2002; 3:127–134. [PubMed: 11812989]

Helmink BA, Sleckman BP. The response to and repair of RAG-mediated DNA double-strand breaks. Annu Rev Immunol. 2012; 30:175–202. [PubMed: 22224778]

Hewitt SL, Chaumeil J, Skok JA. Chromosome dynamics and the regulation of V(D)J recombination. Immunological Reviews. 2010; 237:43–54. [PubMed: 20727028]

Hewitt SL, Farmer D, Marszalek K, Cadera E, Liang HE, Xu Y, Schlissel MS, Skok JA. Association between the Igk and Igh immunoglobulin loci mediated by the 3' Igk enhancer induces 'decontraction' of the Igh locus in pre-B cells. Nat Immunol. 2008; 9:396–404. [PubMed: 18297074]

Hewitt SL, Yin B, Ji Y, Chaumeil J, Marszalek K, Tenthorey J, Salvagiotto G, Steinel N, Ramsey LB, Ghysdael J, et al. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. Nat Immunol. 2009; 10:655–664. [PubMed: 19448632]

Hughes JR, Lower KM, Dunham I, Taylor S, De Gobbi M, Sloane-Stanley JA, McGowan S, Ragoussis J, Vernimmen D, Gibbons RJ, et al. High-resolution analysis of cis-acting regulatory networks at the alpha-globin locus. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2013; 368:20120361. [PubMed: 23650635]

Inlay M, Alt FW, Baltimore D, Xu Y. Essential roles of the kappa light chain intronic enhancer and 3' enhancer in kappa rearrangement and demethylation. Nat Immunol. 2002; 3:463–468. [PubMed: 11967540]

Inlay MA, Lin T, Gao HH, Xu Y. Critical roles of the immunoglobulin intronic enhancers in maintaining the sequential rearrangement of IgH and Igk loci. J Exp Med. 2006; 203:1721–1732. [PubMed: 16785310]

Jhunjhunwala S, van Zelm MC, Peak MM, Cutchin S, Riblet R, van Dongen JJ, Grosveld FG, Knoch TA, Murre C. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. Cell. 2008; 133:265–279. [PubMed: 18423198]

Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. Cell. 2010; 141:419–431. [PubMed: 20398922]

Johnson K, Chaumeil J, Micsinai M, Wang JM, Ramsey LB, Baracho GV, Rickert RC, Strino F, Kluger Y, Farrar MA, et al. IL-7 functionally segregates the pro-B cell stage by regulating transcription of recombination mediators across cell cycle. J Immunol. 2012; 188:6084–6092. [PubMed: 22581861]

Johnson K, Chaumeil J, Skok JA. Epigenetic regulation of V(D)J recombination. Essays Biochem. 2010; 48:221–243. [PubMed: 20822496]

Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annu Rev Immunol. 2006; 24:541–570. [PubMed: 16551259]

Khamlichi AA, Pinaud E, Decourt C, Chauveau C, Cogne M. The 3' IgH regulatory region: a complex structure in a search for a function. Adv Immunol. 2000; 75:317–345. [PubMed: 10879288]

Kil LP, de Bruijn MJ, van Nimwegen M, Corneth OB, van Hamburg JP, Dingjan GM, Thaiss F, Rimmelzwaan GF, Elewaut D, Delsing D, et al. Btk levels set the threshold for B-cell activation and negative selection of autoreactive B cells in mice. Blood. 2012; 119:3744–3756. [PubMed: 22383797]

Kim DR, Dai Y, Mundy CL, Yang W, Oettinger MA. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. Genes Dev. 1999; 13:3070–3080. [PubMed: 10601033]

Kim MS, Lapkouski M, Yang W, Gellert M. Crystal structure of the V(D)J recombinase RAG1-RAG2. Nature. 2015; 518:507–511. [PubMed: 25707801]

Kolovos P, van de Werken HJ, Kepper N, Zuin J, Brouwer RW, Kockx CE, Wendt KS, van IWF, Grosveld F, Knoch TA. Targeted Chromatin Capture (T2C): a novel high resolution high throughput method to detect genomic interactions and regulatory elements. Epigenetics & chromatin. 2014; 7:10. [PubMed: 25031611]

Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, Singh H. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science. 2002; 296:158–162. [PubMed: 11935030]

Krangel MS. Mechanics of T cell receptor gene rearrangement. Current Opinion in Immunology. 2009; 21:133–139. [PubMed: 19362456]

Kurosawa Y, von Boehmer H, Haas W, Sakano H, Trauneker A, Tonegawa S. Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes. Nature. 1981; 290:565–570. [PubMed: 6783962]

Landree MA, Wibbenmeyer JA, Roth DB. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. Genes Dev. 1999; 13:3059–3069. [PubMed: 10601032]

Lange J, Pan J, Cole F, Thelen MP, Jasin M, Keeney S. ATM controls meiotic double-strand-break formation. Nature. 2011

Lee GS, Neiditch MB, Salus SS, Roth DB. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. Cell. 2004; 117:171–184. [PubMed: 15084256]

Lee J, Desiderio S. Cyclin A/CDK2 Regulates V(D)J Recombination by Coordinating RAG-2 Accumulation and DNA Repair. Immunity. 1999; 11:771–781. [PubMed: 10626899]

Lennon GG, Perry RP. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5'-nontranslatable exon. Nature. 1985; 318:475–478. [PubMed: 3934561]

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

Liu H, Schmidt-Supprian M, Shi Y, Hobeika E, Barteneva N, Jumaa H, Pelanda R, Reth M, Skok J, Rajewsky K, et al. Yin Yang 1 is a critical regulator of B-cell development. Genes Dev. 2007a; 21:1179–1189. [PubMed: 17504937]

Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity. 2007b; 27:561–571. [PubMed: 17936034]

Liu Z, Widlak P, Zou Y, Xiao F, Oh M, Li S, Chang MY, Shay JW, Garrard WT. A recombination silencer that specifies heterochromatin positioning and ikaros association in the immunoglobulin kappa locus. Immunity. 2006; 24:405–415. [PubMed: 16618599]

Liu ZM, George-Raizen JB, Li S, Meyers KC, Chang MY, Garrard WT. Chromatin structural analyses of the mouse Igkappa gene locus reveal new hypersensitive sites specifying a transcriptional silencer and enhancer. J Biol Chem. 2002; 277:32640–32649. [PubMed: 12080064]

Livak F, Tourigny M, Schatz DG, Petrie HT. Characterization of TCR gene rearrangements during adult murine T cell development. JImmunol. 1999; 162:2575–2580. [PubMed: 10072498]

Lutz J, Heideman MR, Roth E, van den Berk P, Muller W, Raman C, Wabl M, Jacobs H, Jack HM. Pro-B cells sense productive immunoglobulin heavy chain rearrangement irrespective of polypeptide production. ProcNatlAcadSciUSA. 2011; 108:10644–10649.

Majumder K, Koues OI, Chan EA, Kyle KE, Horowitz JE, Yang-Iott K, Bassing CH, Taniuchi I, Krangel MS, Oltz EM. Lineage-specific compaction of Tcrb requires a chromatin barrier to protect the function of a long-range tethering element. J Exp Med. 2015; 212:107–120. [PubMed: 25512470]

Marquet M, Garot A, Bender S, Carrion C, Rouaud P, Lecardeur S, Denizot Y, Cogne M, Pinaud E. The Emu enhancer region influences H chain expression and B cell fate without impacting IgVH repertoire and immune response in vivo. J Immunol. 2014; 193:1171–1183. [PubMed: 24965776]

Martin DJ, van Ness BG. Initiation and processing of two kappa immunoglobulin germ line transcripts in mouse B cells. Mol Cell Biol. 1990; 10:1950–1958. [PubMed: 2109186]

Mathieu N, Hempel WM, Spicuglia S, Verthuy C, Ferrier P. Chromatin remodeling by the T cell receptor (TCR)-beta gene enhancer during early T cell development: Implications for the control of TCR-beta locus recombination. J Exp Med. 2000; 192:625–636. [PubMed: 10974029]

Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D, Gallardo M, Carney D, Cheung P, Ciccone DN, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature. 2007; 450:1106–1110. [PubMed: 18033247]

McMillan RE, Sikes ML. Differential activation of dual promoters alters Dbeta2 germline transcription during thymocyte development. J Immunol. 2008; 180:3218–3228. [PubMed: 18292546]

McMillan RE, Sikes ML. Promoter activity 5' of Dbeta2 is coordinated by E47, Runx1, and GATA-3. Mol Immunol. 2009; 46:3009–3017. [PubMed: 19592096]

Medvedovic J, Ebert A, Tagoh H, Tamir IM, Schwickert TA, Novatchkova M, Sun Q, Huis In 't Veld PJ, Guo C, Yoon HS, et al. Flexible long-range loops in the VH gene region of the Igh locus

facilitate the generation of a diverse antibody repertoire. Immunity. 2013; 39:229–244. [PubMed: 23973221]

Mendes RD, Sarmento LM, Cante-Barrett K, Zuurbier L, Buijs-Gladdines JG, Povoa V, Smits WK, Abecasis M, Yunes JA, Sonneveld E, et al. PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events. Blood. 2014; 124:567–578. [PubMed: 24904117]

Meyer KB, Sharpe MJ, Surani MA, Neuberger MS. The importance of the 3'-enhancer region in immunoglobulin kappa gene expression. Nucleic Acids Res. 1990; 18:5609–5615. [PubMed: 2120679]

Mombaerts P, Iacomini J, Johnson RS, Herrup K, Tonegawa S, Papaioannou VE. RAG-1-deficient mice have no mature B and T lymphocytes. Cell. 1992; 68:869–877. [PubMed: 1547488]

Monroe RJ, Sleckman BP, Monroe BC, Khor B, Claypool S, Ferrini R, Davidson L, Alt FW. Developmental regulation of TCR delta locus accessibility and expression by the TCR delta enhancer. Immunity. 1999; 10:503–513. [PubMed: 10367896]

Mostoslavsky R, Singh N, Tenzen T, Goldmit M, Gabay C, Elizur S, Qi P, Reubinoff BE, Chess A, Cedar H, et al. Asynchronous replication and allelic exclusion in the immune system. Nature. 2001; 414:221–225. [PubMed: 11700561]

Mullighan CG, Miller CB, Radtke I, Phillips LA, Dalton J, Ma J, White D, Hughes TP, Le Beau MM, Pui CH, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature. 2008; 453:110–114. [PubMed: 18408710]

Nadel B, Tang A, Escuro G, Lugo G, Feeney AJ. Sequence of the spacer in the recombination signal sequence affects V(D)J rearrangement frequency and correlates with nonrandom Vkappa usage in vivo. J Exp Med. 1998; 187:1495–1503. [PubMed: 9565641]

Naik AK, Hawwari A, Krangel MS. Specification of Vdelta and Valpha usage by Tcra/Tcrd locus V gene segment promoters. J Immunol. 2015; 194:790–794. [PubMed: 25472997]

Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D. Transcription. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science. 2015; 347:1017–1021. [PubMed: 25722416]

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012; 485:381–385. [PubMed: 22495304]

Onozawa M, Aplan PD. Illegitimate V(D)J recombination involving nonantigen receptor loci in lymphoid malignancy. Genes, chromosomes & cancer. 2012; 51:525–535. [PubMed: 22334400]

Pan X, Papasani M, Hao Y, Calamito M, Wei F, Quinn Iii WJ, Basu A, Wang J, Hodawadekar S, Zaprazna K, et al. YY1 controls Igkappa repertoire and B-cell development, and localizes with condensin on the Igkappa locus. EMBO J. 2013; 32:1168–1182. [PubMed: 23531880]

Papaemmanuil E, Rapado I, Li Y, Potter NE, Wedge DC, Tubio J, Alexandrov LB, Van Loo P, Cooke SL, Marshall J, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. Nat Genet. 2014; 46:116–125. [PubMed: 24413735]

Pelanda R. Dual immunoglobulin light chain B cells: Trojan horses of autoimmunity? Curr Opin Immunol. 2014; 27:53–59. [PubMed: 24549093]

Pelanda R, Schwers S, Sonoda E, Torres RM, Nemazee D, Rajewsky K. Receptor editing in a transgenic mouse model: site, efficiency, and role in B cell tolerance and antibody diversification. Immunity. 1997; 7:765–775. [PubMed: 9430222]

Perlot T, Alt FW, Bassing CH, Suh H, Pinaud E. Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci U S A. 2005; 102:14362–14367. [PubMed: 16186486]

Pinaud E, Marquet M, Fiancette R, Peron S, Vincent-Fabert C, Denizot Y, Cogne M. The IgH locus 3' regulatory region: pulling the strings from behind. Adv Immunol. 2011; 110:27–70. [PubMed: 21762815]

Prak EL, Weigert M. Light chain replacement: a new model for antibody gene rearrangement. J Exp Med. 1995; 182:541–548. [PubMed: 7629511]

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell. 2014; 159:1665–1680. [PubMed: 25497547]

Ribeiro de Almeida C, Stadhouders R, de Bruijn MJ, Bergen IM, Thongjuea S, Lenhard B, van Ijcken W, Grosveld F, Galjart N, Soler E, et al. The DNA-binding protein CTCF limits proximal Vkappa recombination and restricts kappa enhancer interactions to the immunoglobulin kappa light chain locus. Immunity. 2011; 35:501–513. [PubMed: 22035845]

Ribeiro de Almeida C, Stadhouders R, Thongjuea S, Soler E, Hendriks RW. DNA-binding factor CTCF and long-range gene interactions in V(D)J recombination and oncogene activation. Blood. 2012; 119:6209–6218. [PubMed: 22538856]

Roldan E, Fuxa M, Chong W, Martinez D, Novatchkova M, Busslinger M, Skok JA. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol. 2005; 6:31–41. [PubMed: 15580273]

Rouaud P, Vincent-Fabert C, Saintamand A, Fiancette R, Marquet M, Robert I, Reina-San-Martin B, Pinaud E, Cogne M, Denizot Y. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. J Exp Med. 2013; 210:1501–1507. [PubMed: 23825188]

Ryu CJ, Haines BB, Lee HR, Kang YH, Draganov DD, Lee M, Whitehurst CE, Hong HJ, Chen J. The T-cell receptor beta variable gene promoter is required for efficient V beta rearrangement but not allelic exclusion. Mol Cell Biol. 2004; 24:7015–7023. [PubMed: 15282302]

Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. AnnuRevGenet. 2011; 45:167–202.

Schlimgen RJ, Reddy KL, Singh H, Krangel MS. Initiation of allelic exclusion by stochastic interaction of Tcrb alleles with repressive nuclear compartments. Nat Immunol. 2008; 9:802–809. [PubMed: 18536719]

Schlissel MS. Regulation of activation and recombination of the murine Igkappa locus. Immunol Rev. 2004; 200:215–223. [PubMed: 15242407]

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. Nature. 2011; 476:467–471. [PubMed: 21832993]

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell. 2012; 148:458–472. [PubMed: 22265598]

Shih HY, Verma-Gaur J, Torkamani A, Feeney AJ, Galjart N, Krangel MS. Tcra gene recombination is supported by a Tcra enhancer- and CTCF-dependent chromatin hub. Proc Natl Acad Sci U S A. 2012; 109:E3493–3502. [PubMed: 23169622]

Shinkai Y, Rathbun G, Lam K-P, Oltz EM, Stewart V, Mendelsohn M, Charron J, Datta M, Young F, Stall AM, et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell. 1992; 68:855–867. [PubMed: 1547487]

Sikes ML, Gomez RJ, Song J, Oltz EM. A developmental stage-specific promoter directs germline transcription of D beta J beta gene segments in precursor T lymphocytes. J Immunol. 1998; 161:1399–1405. [PubMed: 9686603]

Skok J. V(D)J recombination: a paradigm for studying chromosome interactions in mammalian cells. Epigenomics. 2010; 2:176–177.

Skok JA, Brown KE, Azuara V, Caparros ML, Baxter J, Takacs K, Dillon N, Gray D, Perry RP, Merkenschlager M, et al. Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. Nat Immunol. 2001; 2:848–854. [PubMed: 11526401]

Skok JA, Gisler R, Novatchkova M, Farmer D, de Laat W, Busslinger M. Reversible contraction by looping of the Tcra and Tcrb loci in rearranging thymocytes. Nature Immunology. 2007; 8:378–387. [PubMed: 17334367]

Sleckman BP, Bardon CG, Ferrini R, Davidson L, Alt FW. Function of the TCR alpha enhancer in alphabeta and gammadelta T cells. Immunity. 1997; 7:505–515. [PubMed: 9354471]

Spanopoulou E, Roman CAJ, Corcoran L, Schlissel M, Silver DP, Nemazee D, Nussenzweig M, Shinton SA, Hardy RR, Baltimore D. Functional immunoglobulin transgenes guide ordered B cell differentiation in RAG-1-deficient mice. 1994; 8:1030–1042.

Stadhouders R, de Bruijn MJ, Rother MB, Yuvaraj S, Ribeiro de Almeida C, Kolovos P, Van Zelm MC, van Ijcken W, Grosveld F, Soler E, et al. Pre-B cell receptor signaling induces immunoglobulin kappa locus accessibility by functional redistribution of enhancer-mediated chromatin interactions. PLoS Biol. 2014; 12:e1001791. [PubMed: 24558349]

Steinel NC, Fisher MR, Yang-Iott KS, Bassing CH. The Ataxia Telangiectasia Mutated and Cyclin D3 Proteins Cooperate To Help Enforce TCRbeta and IgH Allelic Exclusion. J Immunol. 2014; 193:2881–2890. [PubMed: 25127855]

Steinel NC, Lee BS, Tubbs AT, Bednarski JJ, Schulte E, Yang-Iott KS, Schatz DG, Sleckman BP, Bassing CH. The ataxia telangiectasia mutated kinase controls Igkappa allelic exclusion by inhibiting secondary Vkappa-to-Jkappa rearrangements. J Exp Med. 2013; 210:233–239. [PubMed: 23382544]

Stone JL, McMillan RE, Skaar DA, Bradshaw JM, Jirtle RL, Sikes ML. DNA double-strand breaks relieve USF-mediated repression of Dbeta2 germline transcription in developing thymocytes. J Immunol. 2012; 188:2266–2275. [PubMed: 22287717]

Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. 1993; 177:1009–1020.

Tonegawa S. Somatic generation of antibody diversity. Nature. 1983; 302:575–581. [PubMed: 6300689]

Tripathi RK, Mathieu N, Spicuglia S, Payet D, Verthuy C, Bouvier G, Depetris D, Mattei MG, Hempe LW, Ferrier P. Definition of a T-cell receptor beta gene core enhancer of V(D)J recombination by transgenic mapping. Mol Cell Biol. 2000; 20:42–53. [PubMed: 10594007]

Vettermann C, Schlissel MS. Allelic exclusion of immunoglobulin genes: models and mechanisms. Immunol Rev. 2010; 237:22–42. [PubMed: 20727027]

Weischenfeldt J, Damgaard I, Bryder D, Theilgaard-Monch K, Thoren LA, Nielsen FC, Jacobsen SE, Nerlov C, Porse BT. NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. 2008; 22:1381–1396. [PubMed: 18483223]

Whitehurst CE, Chattopadhyay S, Chen J. Control of V(D)J recombinational accessibility of the D beta 1 gene segment at the TCR beta locus by a germline promoter. Immunity. 1999; 10:313–322. [PubMed: 10204487]

Xiang Y, Park SK, Garrard WT. Vkappa gene repertoire and locus contraction are specified by critical DNase I hypersensitive sites within the Vkappa-Jkappa intervening region. J Immunol. 2013; 190:1819–1826. [PubMed: 23296705]

Xiang Y, Park SK, Garrard WT. A major deletion in the Vkappa-Jkappa intervening region results in hyperelevated transcription of proximal Vkappa genes and a severely restricted repertoire. J Immunol. 2014; 193:3746–3754. [PubMed: 25187654]

Xiang Y, Zhou X, Hewitt SL, Skok JA, Garrard WT. A multifunctional element in the mouse Igkappa locus that specifies repertoire and Ig loci subnuclear location. JImmunol. 2011; 186:5356–5366. [PubMed: 21441452]

Yamagami T, ten Boekel E, Andersson J, Rolink A, Melchers F. Frequencies of multiple IgL chain gene rearrangements in single normal or kappaL chain-deficient B lineage cells. Immunity. 1999; 11:317–327. [PubMed: 10514010]

Zhou X, Xiang Y, Ding X, Garrard WT. Loss of an Igkappa gene enhancer in mature B cells results in rapid gene silencing and partial reversible dedifferentiation. Mol Cell Biol. 2013; 33:2091–2101. [PubMed: 23508106]

Zhou X, Xiang Y, Garrard WT. The Igkappa gene enhancers, E3' and Ed, are essential for triggering transcription. JImmunol. 2010; 185:7544–7552. [PubMed: 21076060]

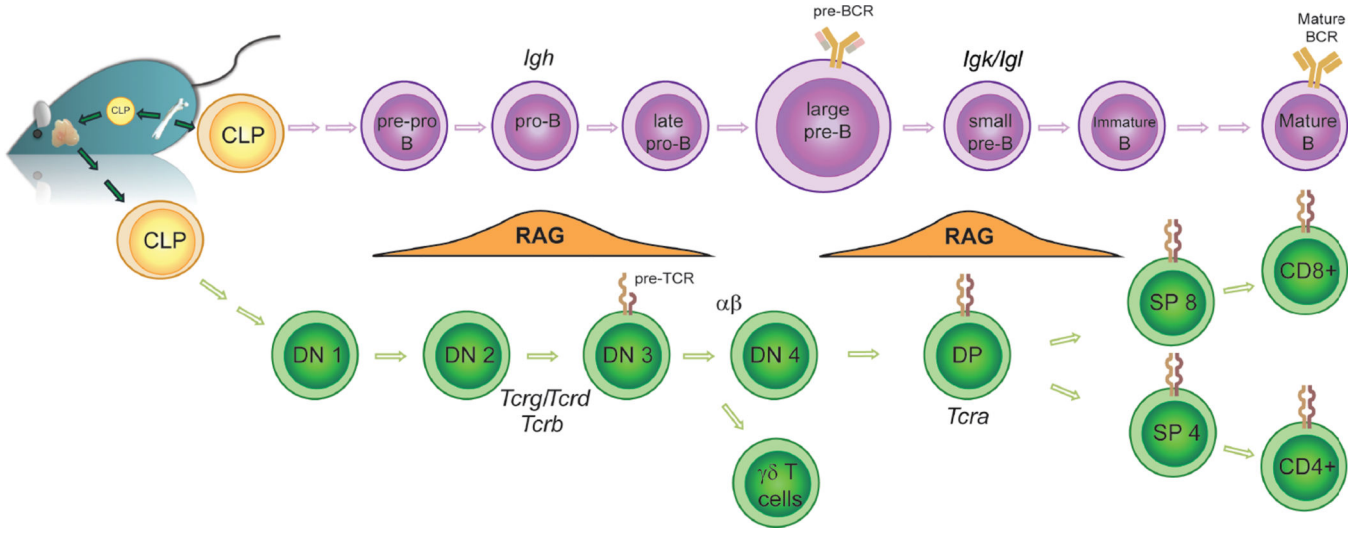


Figure 1.

Scheme showing the different stages of B and T cell development where rearrangement of the *Ig* or *Tcr* loci take place.

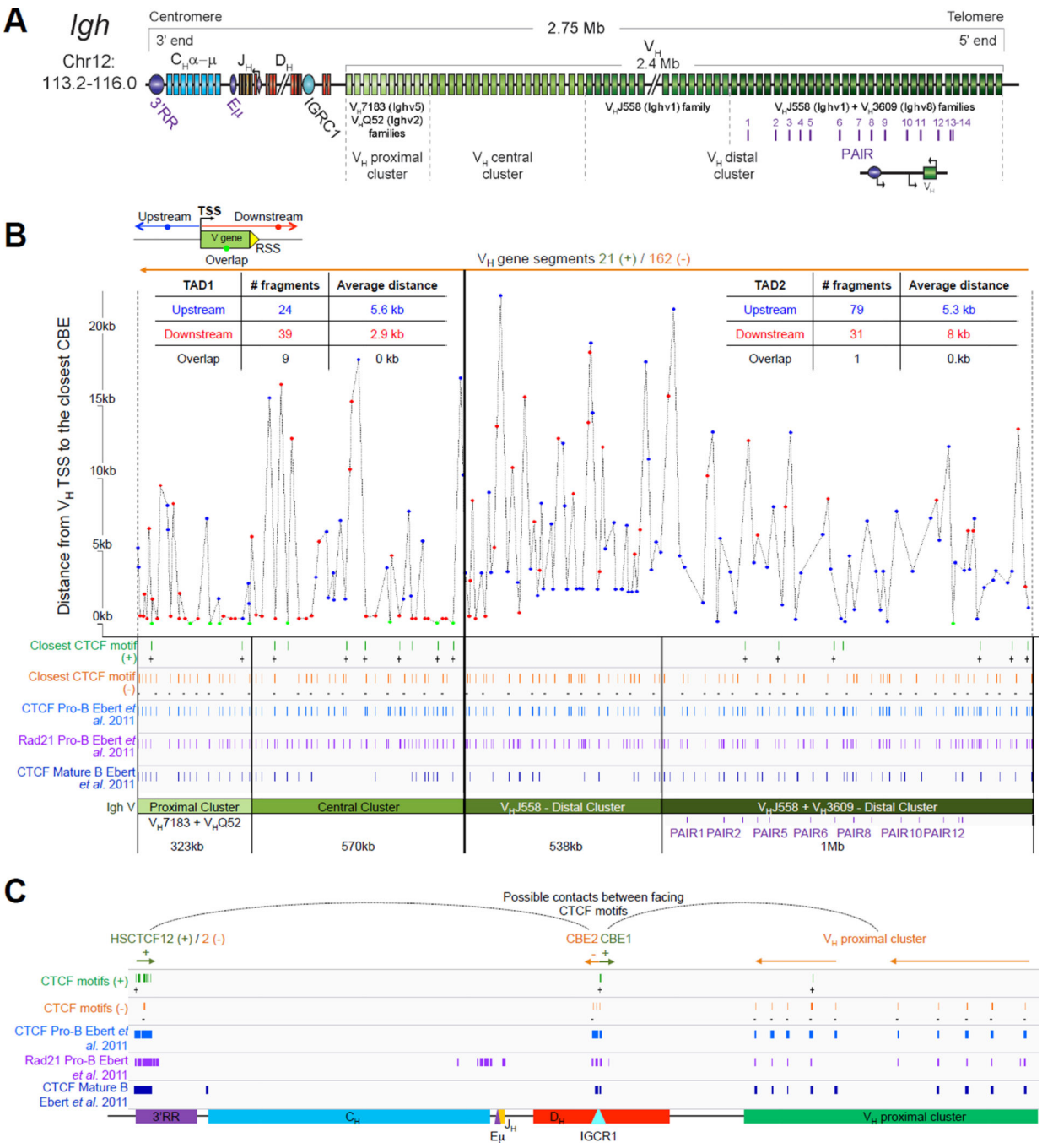


Figure 2.

A. *Igh* linear structure and its *cis* acting elements. *Igh* spans 2.75Mb. V_H, D_H, J_H and C_H segments are organized in separate clusters with all segments in the same 5'-to-3' orientation on the minus strand of chromosome 12. *Igh* contains 183 V_H segments (113 functional), 20 D_H segments (10 functional), 4 J_H and 8 C_H (all functional). It is of note that 2 D_H segments (one of which is functional) are located 5' to the intergenic insulator IGCR1. The V_H array contains V_H sub-clusters determined by the type of V_H families represented: the proximal cluster enriched for V_H7183 (IghV5) and V_HQ52 (IghV2) segments; the

central cluster which does not include specific family types and the distal cluster enriched for V_HJ558 (IghV1) and V_H3609 (IghV8) segments. **B. Distribution of CTCF binding elements (CBEs) within the *Igh* V_H gene region**. Distance from the TSS of each V_H gene to the closest CBE is reported and marked as upstream, downstream or overlapping depending on the CBE location relative to the TSS. Closest CBEs have been selected among motifs falling within CTCF peaks called from pro-B cells ChIP-seq data (Ebert et al., 2011). CTCF motifs were called using FIMO (part of the MEME suite) with p-value < 10e-4. This analysis identifies 125 CBEs (54 in TAD1 and 71 in TAD2). The vast majority of V_H segments (162 of the 183) are associated to a CBE on the minus strand (+), pointing towards the 3' end of the *Igh* locus. The average distance between a V_H segment and its closest CBE is around 5kb and overall there is no relationship to upstream or downstream localization of the motif. However, the Murre lab described two sub-domains constituting the V_H array (annotated TAD1 and TAD2 here) (Jhunjhunwala et al., 2008) which show specificity in localization of the closest CBE associated with different distances. TAD1 contains more downstream sites which are closer than upstream ones. In contrast, TAD2 contains more upstream sites which are closer than the downstream ones. **C. Zoomed in region of the 3'end of *Igh to* highlighting the orientation of CBEs**. CTCF binding motifs have been selected by intersection with CTCF peaks called from published pro-B cell ChIP-seq data (Ebert et al., 2011). CBE1 and CBE2 are pointing away from one another i(Lin et al., 2015) potentially enabling loop formation between convergent motifs on CBE1 and the 5' V_H cluster and CBE2 and the 3'RR. Segment annotations, with coordinates, strand orientation and functional status as well as coordinates for regulatory elements are provided in Table 1. Annotations correspond to the mm10/GRCm38 genome assembly which uses the C57BL/6 strain as genome reference and were collected from NCBI (*Igh* gene ID: 111507; *Igk* gene ID: 243469; *Tcrb* gene ID: 21577; *Tcra* Gene ID: 21473) and IMGT/LIGM-DB databases (Giudicelli et al., 2006). V_H segments (green), D_H segments (red), J_H segments (orange) and C_H constant region (blue), enhancers (purple), insulators (aqua).

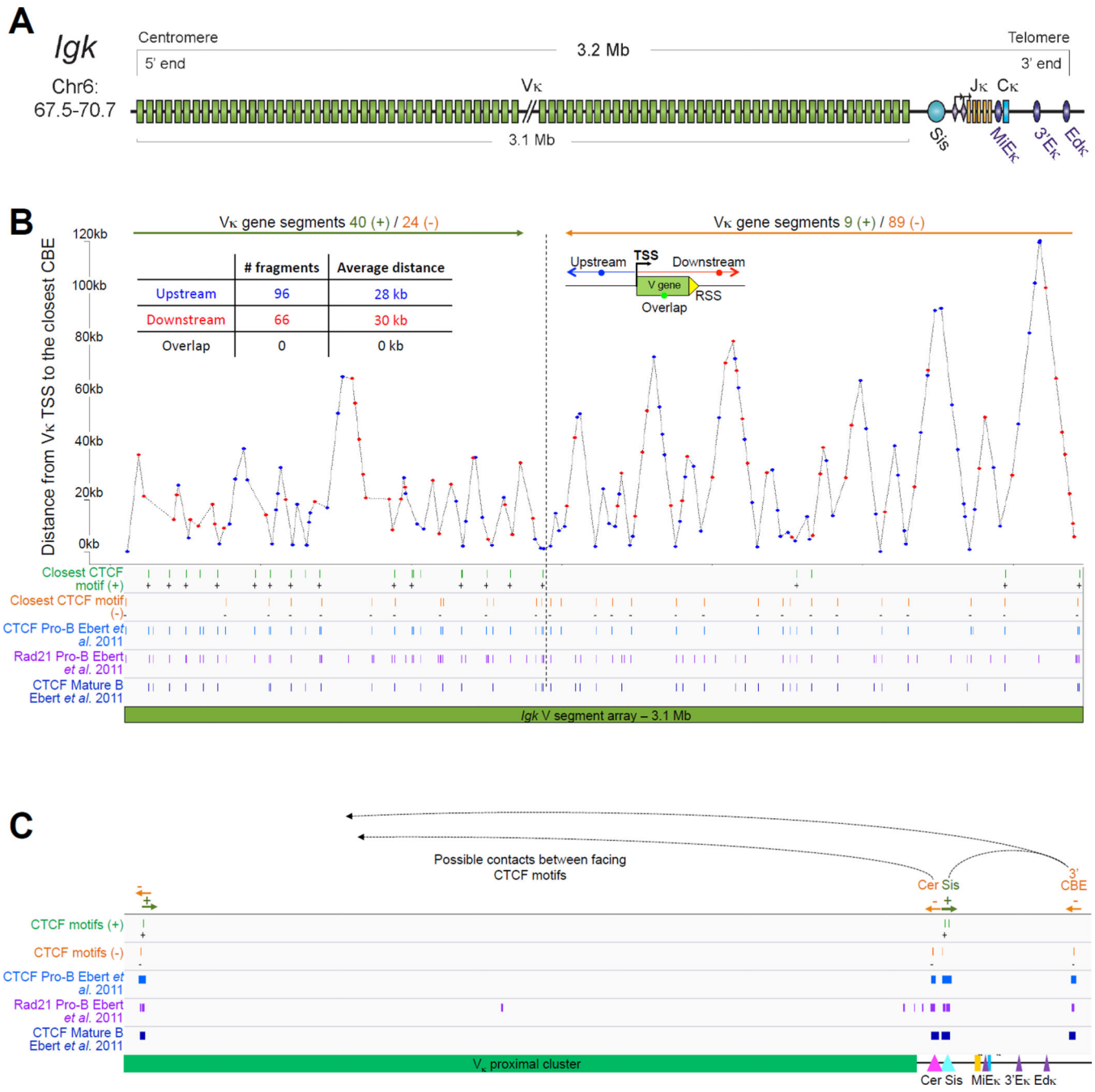


Figure 3.

A. *Igk* linear structure and its *cis* acting regulatory elements. *Igk* spans 3.17Mb on the murine chromosome 6. It contains 162 Vκ segments (92 functional), 5 Jκs (4 functional, Jκ3 is a pseudogene with a mutated RSS not recognized by RAG) and a single Cκ region. Half of the Vκs are positioned in reverse orientation and are rearranged by non-destructive inversion. All the other segments of the locus follow a (+) strand orientation. **B. CTCF binding element (CBE) distribution within the Vκ region**. Closest CTCF binding motifs from the TSS of each Vκ gene segment were called using CTCF peaks from pro-B cells ChIP-seq data (Ebert et al., 2011). The Vκ cluster harbors 59 closest CBEs which display

alternative orientation but there is an enrichment of Vκ segments associated to (+) motifs at the distal part and (−) motifs at the proximal part of the Vκ cluster (40 (+) / 24 (−) versus 9 (+) / 89 (−) respectively). It should be noted that several segments can be associated with the same CTCF motif. The fact that Vκ rearrangement can occur via inversion could explain this non specific orientation that is in contrast to what is seen for the *Igh* locus. There is no apparent correlation between the CTCF motif orientation and the Vκ segment orientation. However there are more segments associated to a CTCF motif with the same orientation (62%) as opposed to an inverse orientation (38%). The average distance between a Vκ segment and its closest CTCF motif is much larger than for *Igh* - 30kb versus 5kb – and there is no relationship to upstream or downstream localization of the motif. **C. Zoomed in region of the proximal domain of *Igk* to highlight the orientation of CBEs**. The CTCF motif at the 3' boundary is directed towards the *Igk* locus, which could facilitate intra-locus contacts. *Cer* and *Sis* display a similar organization to IGCR1, which encompasses CBE1 and 2 in *Igh* that have opposite orientations pointing away from each other towards the 5' end and the 3' end respectively. *Sis* and the 3' boundary anchor display constitutive CTCF binding throughout B cell development. They contain motifs with a head to head orientation that could promote proximal domain segregation, which is important for limiting proximal Vκ recombination and restricting *Igk* enhancer interactions to the *Igk* locus outside of its rearrangement stages (Ribeiro de Almeida et al., 2011; Ribeiro de Almeida et al., 2012; Xiang et al., 2011).

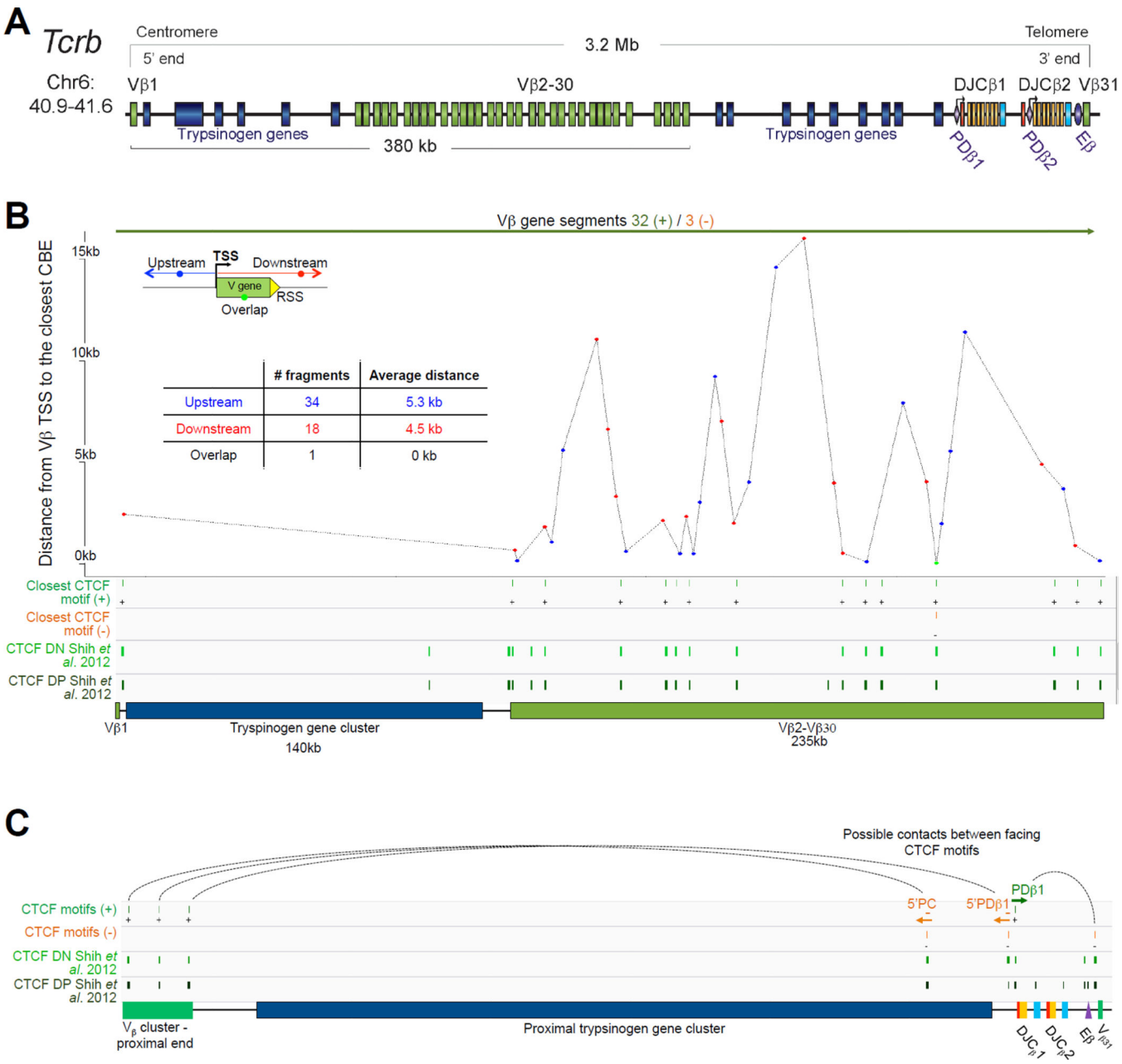


Figure 4.

A. *Tcrb* linear structure and its *cis* acting regulatory elements. *Tcrb* encompasses 700kb on the murine chromosome 6. It contains 35 Vβ gene segments (22 functional) spread out over 624kb, with the exception of Vβ31 that is localized at the 3' end of the locus in an inverted orientation. All the other segment of the locus follow a (+) strand orientation. The proximal domain is duplicated with a total of 2 Dβ, 14 Jβ (11 functional) and 2 Cβ gene segments. Two clusters of trypsinogen genes separate the bulk of Vβ genes from the first Dβ segment on the 3' side (separation of 250kb) as well as from the Vβ1 segment located at the 5' end. **B. CTCF binding element (CBE) distribution along *Tcrb* Vβ gene segments (excluding Vβ31)**. Closest CTCF binding motifs from the TSS of each Vβ gene segment

were called using CTCF peaks from DN cells ChIP-seq data (Shih et al., 2012). The Vβ cluster harbors 21 closest CBEs. 33 of the 35 Vβ segments are associated to CTCF motifs in the same orientation ((+) orientation for Vβ1 to Vβ30; (−) orientation for Vβ31). Motifs associated with Vβ1 to Vβ30 point towards the 3' end of the *Tcrb* locus and could establish contact with the facing motif in the 5'PDβ1 region. The average distance between a Vβ segment and its closest CTCF motif is around 5kb with no relationship to upstream or downstream localization of the motif. **C. Zoomed in region of the proximal domain of *Tcrb* to highlight the orientation of CBEs**. The CTCF motif associated to PDβ1 (+) is facing the 3' end motif (−) located between Eβ and Vβ31. In contrast, the 5'PDβ1 and 5'PC (5'Prss2-CTCF) motifs (−) face motifs located at Vβ segments (+). This is one more example where the region between the V array and the proximal domain harbors a composite element containing CTCF motifs pointing away from each other towards the 5' end and the 3' end of the locus.

Figure 5.

A. *Tcra* linear structure and its *cis* acting regulatory elements. *Tcra* spans 1,65Mb on the murine chromosome 14. It contains 130 Vα segments (108 functional) spread out over 1.55Mb and located upstream of 60 Jα genes (38 functional) and a single Cα gene. In the C57BL/6 background, the Vα array is composed of triplicated clusters located in the center with 8 and 10 unique segments on each side respectively. *Tcra* shares V segments with the *Tcrd* locus that is embedded within its locus. These 10 V segments, annotated Trav*-dv*, rearrange either to Jα or to Dδ. **B. CTCF binding element (CBE) distribution along *Tcra* V segments array**. Closest CTCF binding motifs from the TSS of each Vα gene segment

were called using CTCF peaks from DP cells ChIP-seq data (Shih et al., 2012). The Vα cluster harbors 124 closest CBEs. 121 of the 130 Vα segments are associated to CTCF motifs in the same orientation ((+) orientation) pointing towards the 3' end of the *Tcra* locus. The average distance between a Vα segment and its closest CTCF motif is around 3kb for motifs located upstream and 9kb for motifs located downstream. **C. Zoomed in region of the proximal domain of *Tcra* to highlight the orientation of CBEs**. The CTCF motif associated to TEAp (−) is facing the (+) motifs of the V segments towards the 5' end of the locus.

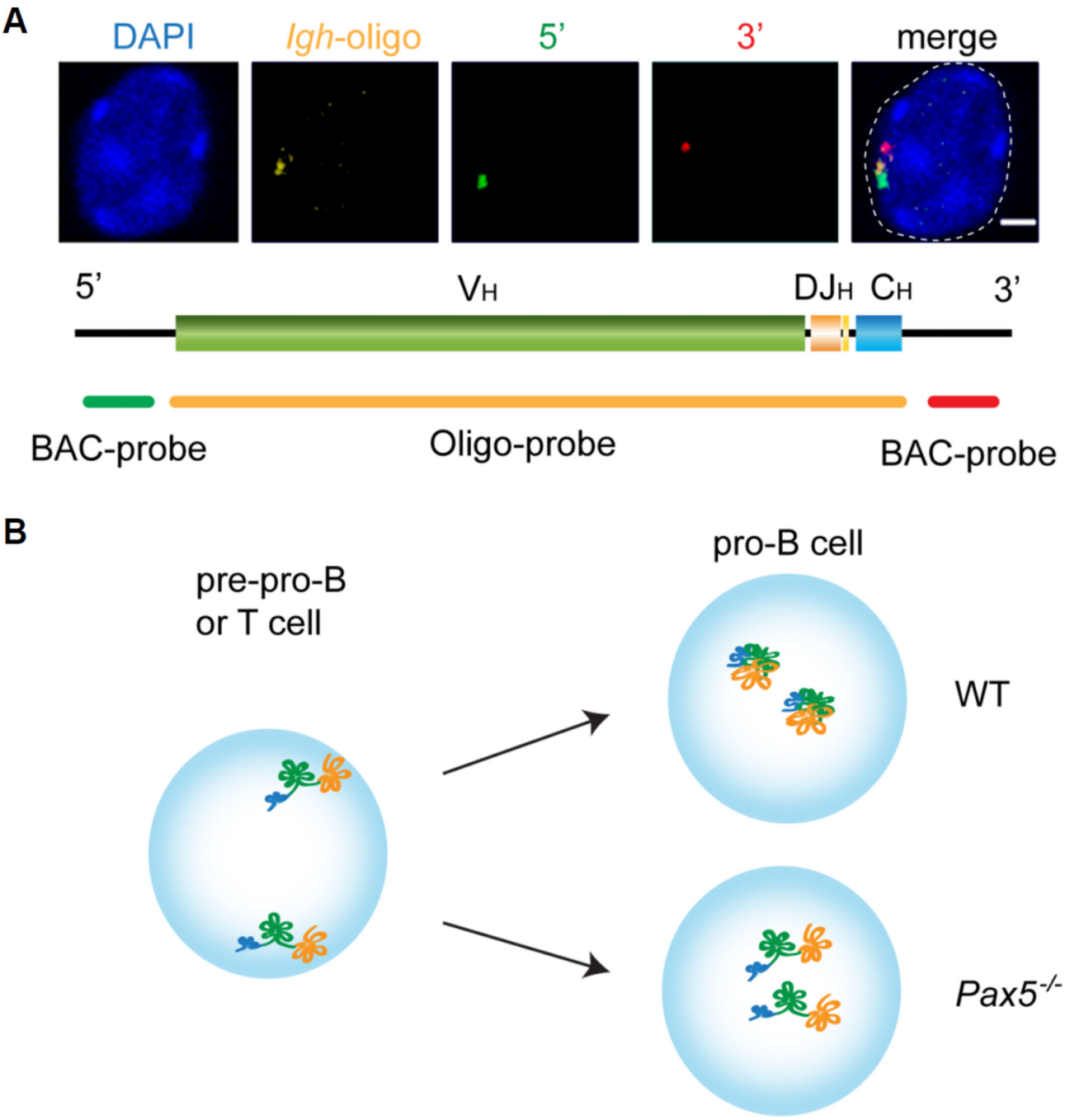


Figure 6.

The *Igh* locus is activated by a two step mechanism that involves relocation to the center of the nucleus and Pax5-mediated locus contraction at the time of recombination in Pro-B cells. A. 3D DNA FISH showing the location of the *Igh* locus and its orientation at the nuclear periphery when it is in a decontracted state. A probe scheme is shown below the FISH images identifying the location of 5' and 3' BAC probes relative to an oligonucleotide probe that covers the entire *Igh* locus. **B.** Scheme showing the two-step activation of *Igh*.

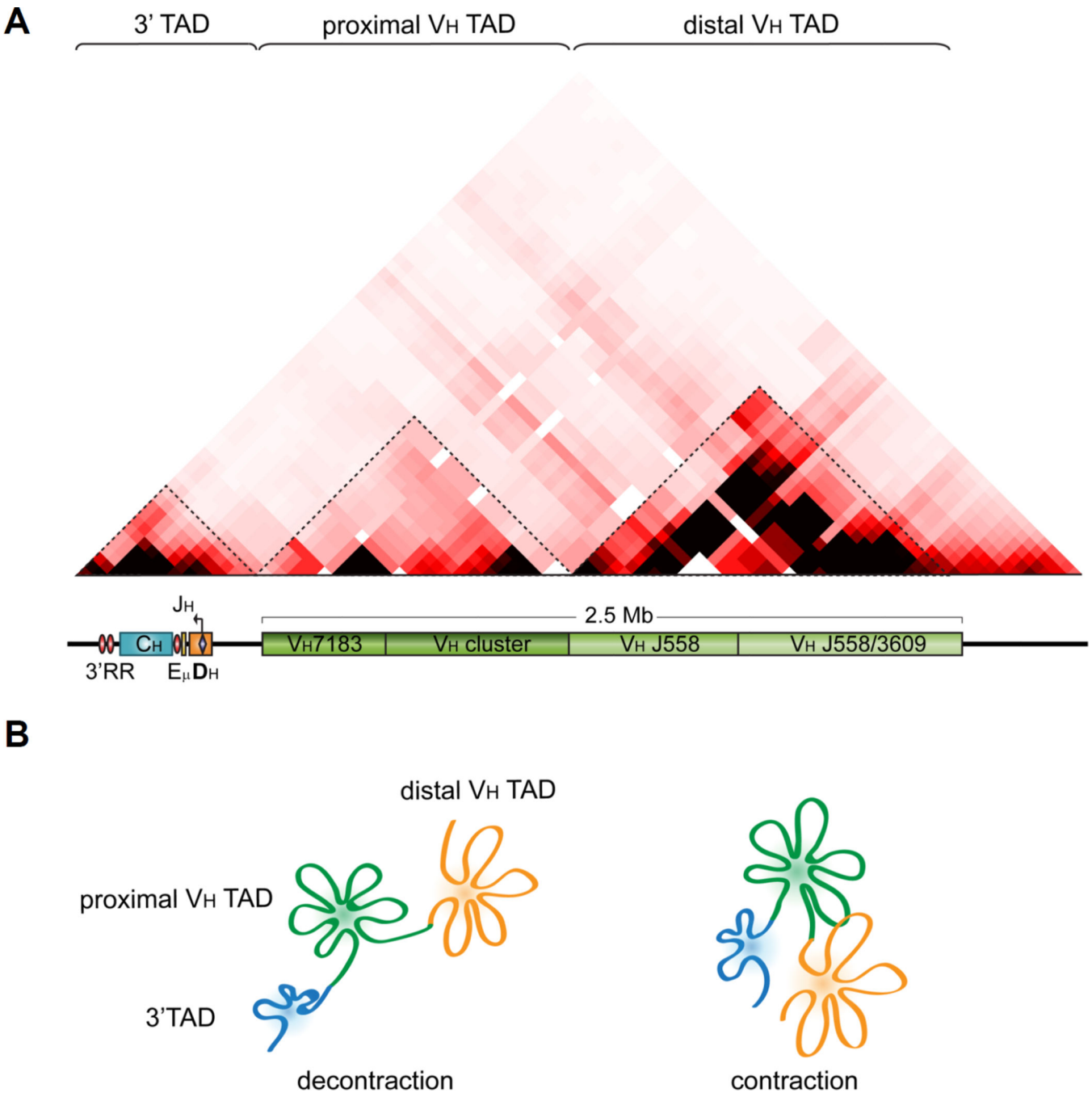


Figure 7.

Locus contraction of the *Igh* locus involves interaction of two TADs or rosette structures that encompass the V_H gene region. A. Scheme showing the various gene segments of the *Igh* locus relative to a 5C matrix of *Igh* interactions in DP cells (unpublished BH and JS). **B.** Scheme showing rosette like structure of the *Igh* locus as identified by DNA FISH analyses (Jhunjhunwala et al., 2008).

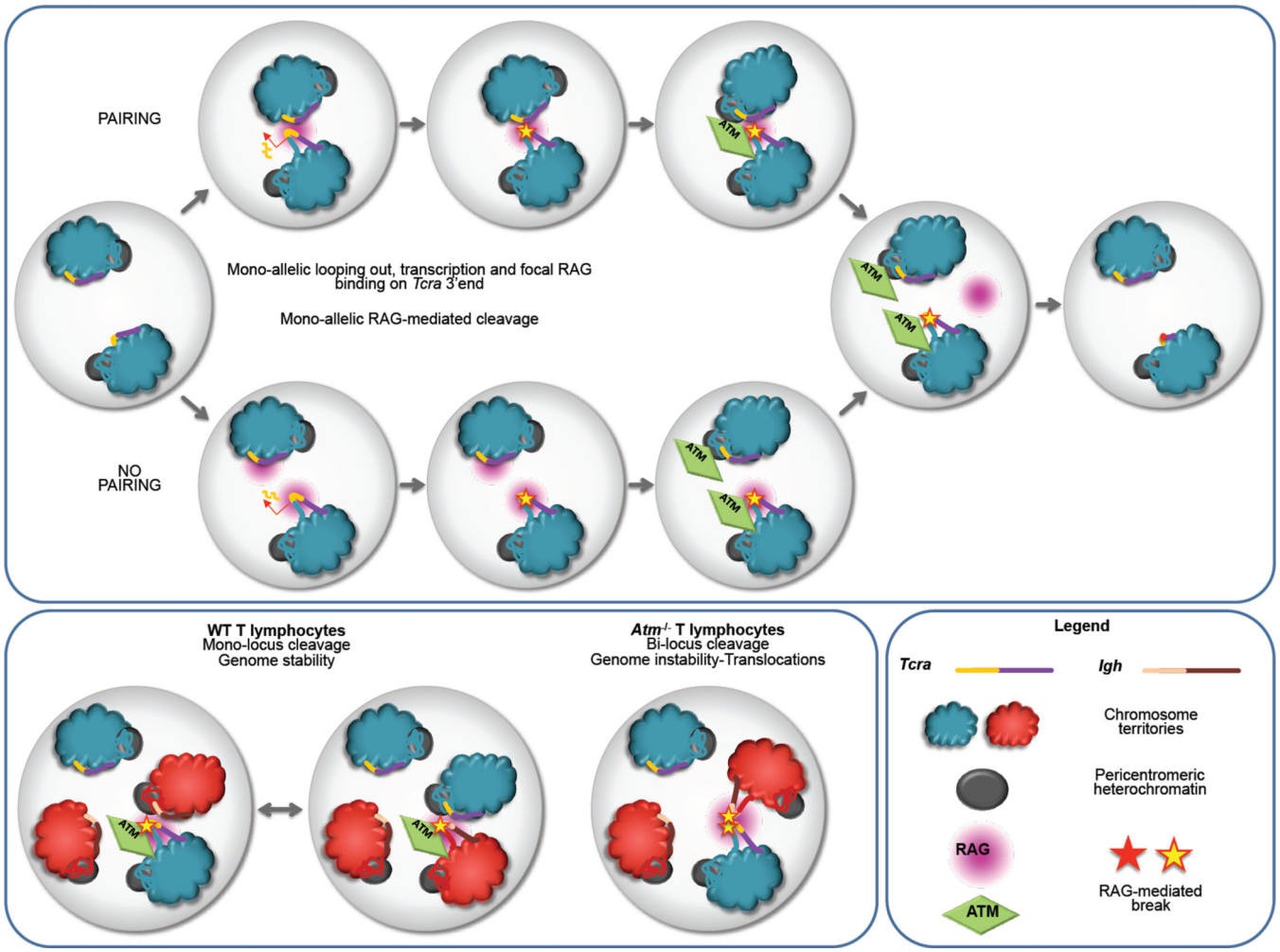


Figure 8.

Model of **ATM-mediated control of cleavage**. Changes in nuclear accessibility of the antigen-receptor loci are linked with mono-allelic and mono-locus rearrangement and maintenance of genome integrity.

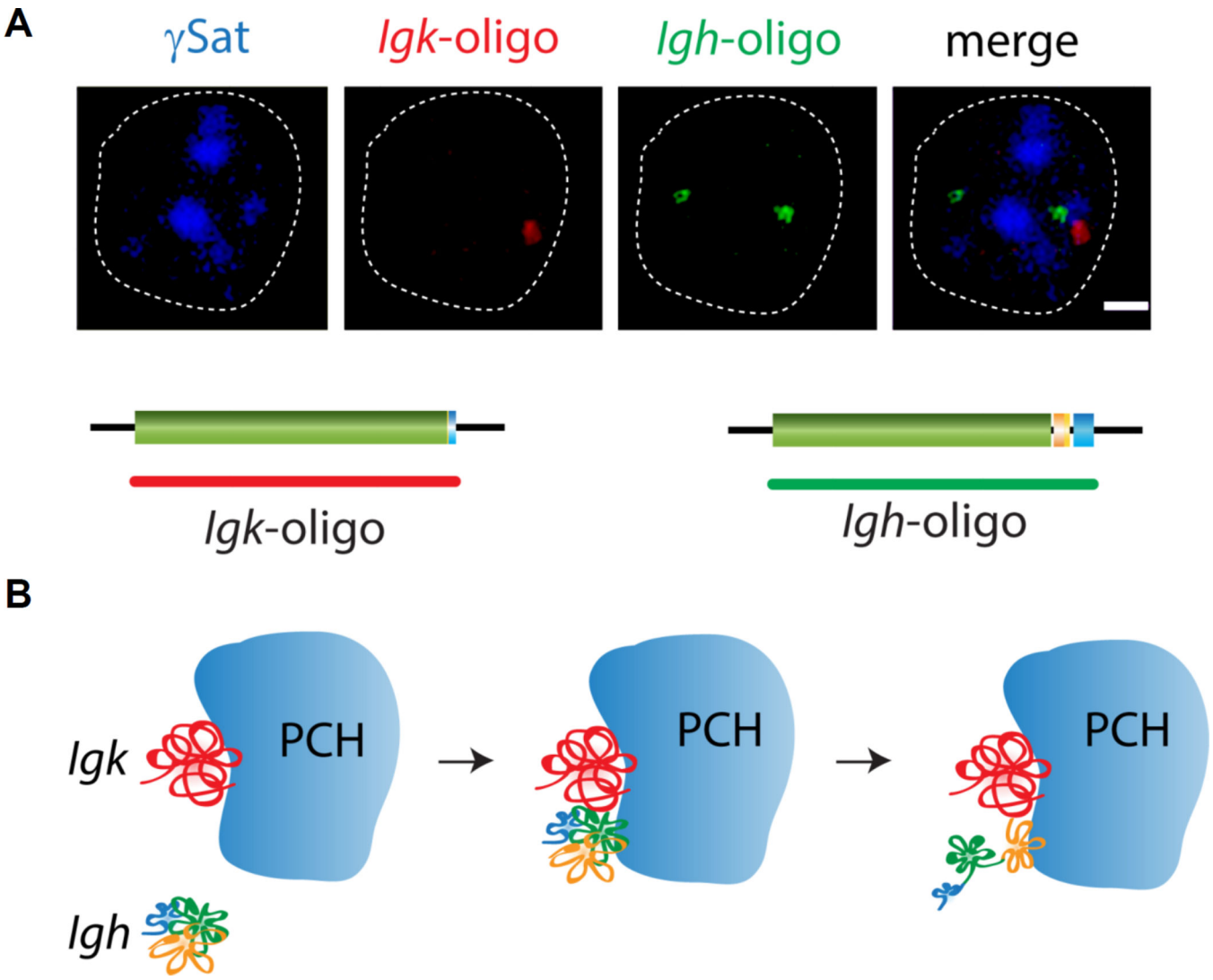


Figure 9.

An interaction between *Igk* and *Igh* is responsible for *Igh* PCH association and *Igh* locus decontraction post recombination, at the pre-B cell stage. A. 3D DNA FISH showing the *Igk-Igh* interaction at PCH in pre-B cells. Oligonucleotide probes encompassing the entire *Igk* and *Igh* loci were used for this analysis. **B.** Scheme showing *Igk-Igh* interaction at PCH leading to *Igh* locus decontraction.

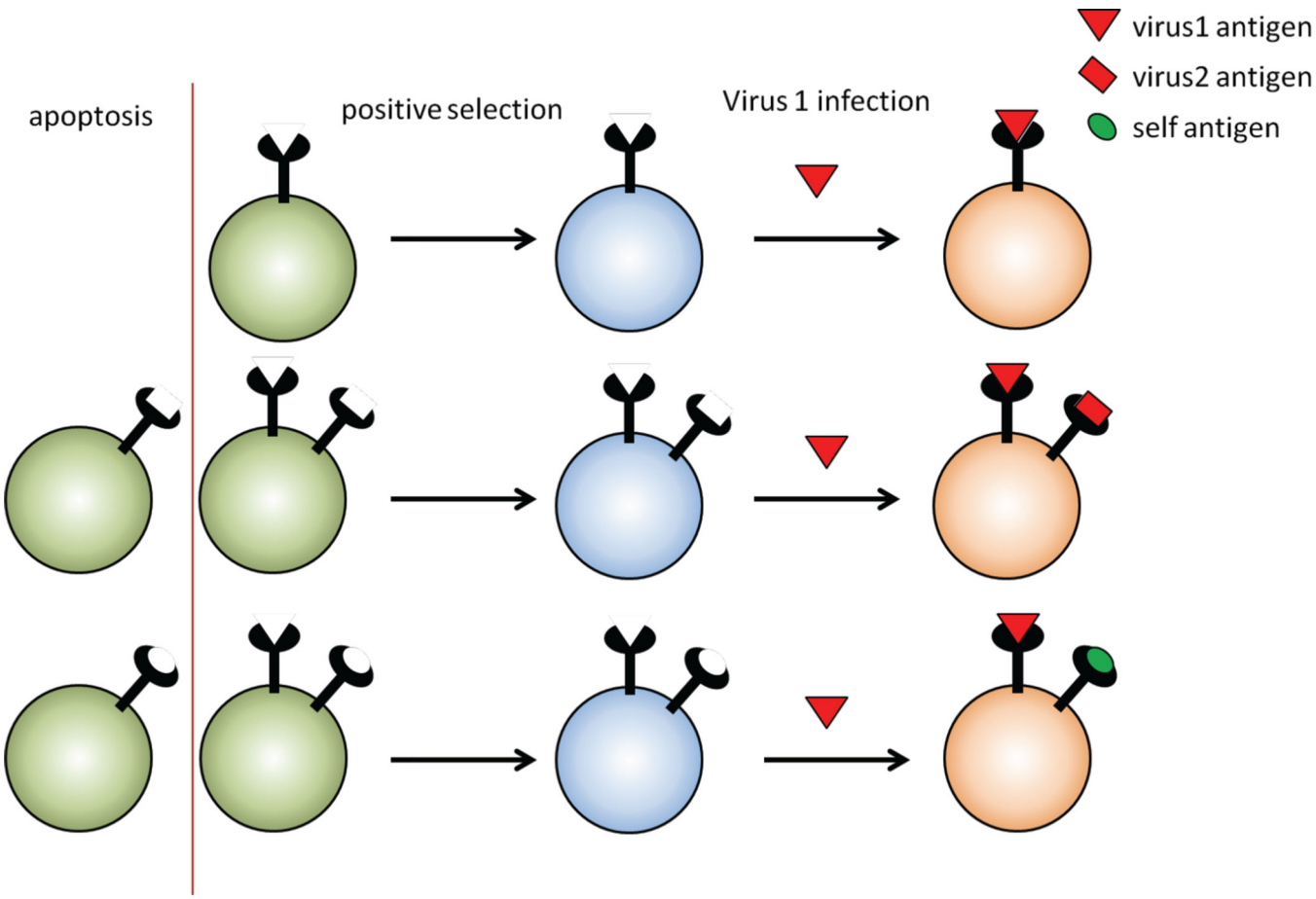


Figure 10.
Evolution may tolerate a certain frequency of allelically included dual receptor cells, balancing an autoimmune outcome with that of counteracting infection

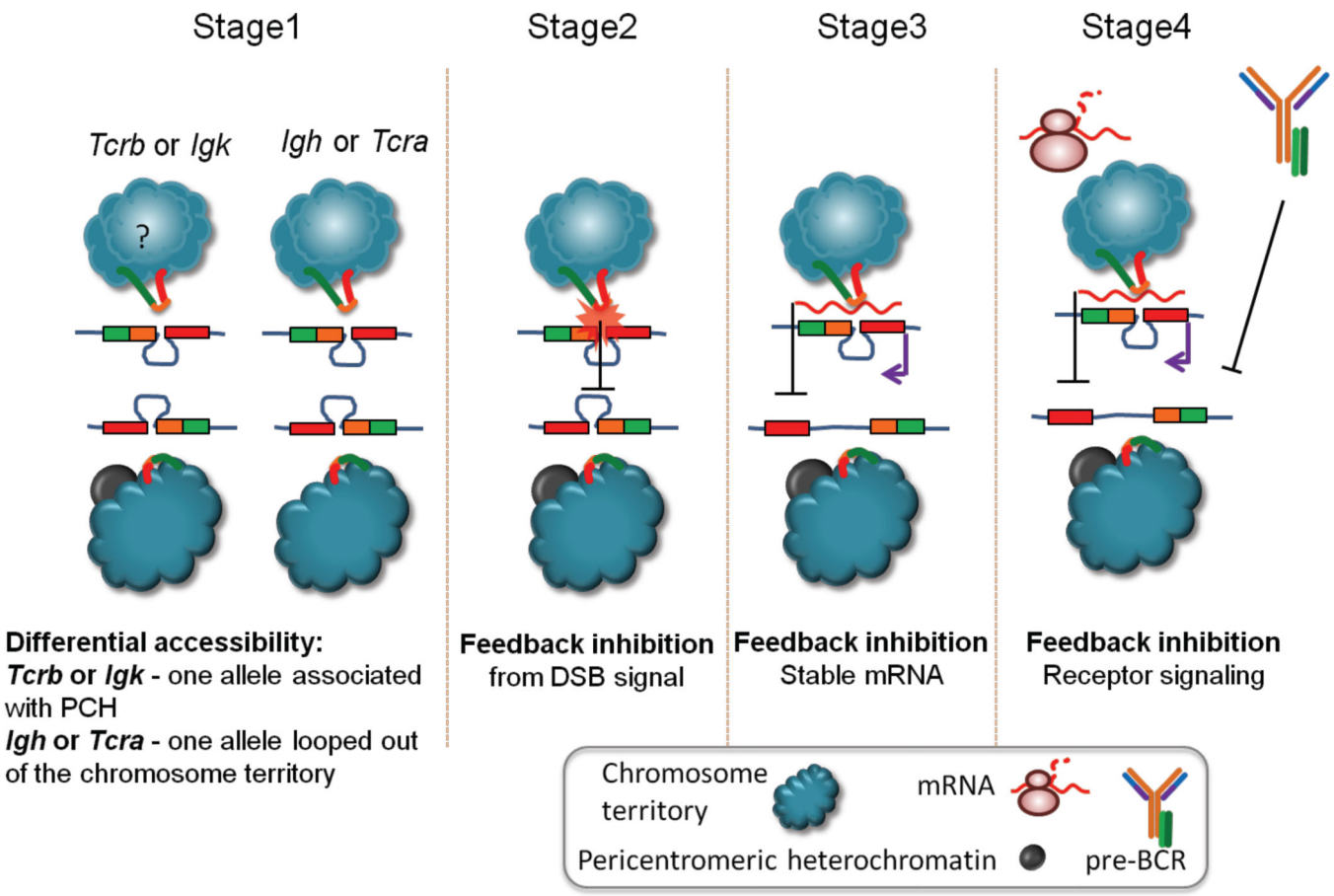


Figure 11.
Allelic exclusion is enforced at multiple stages of development

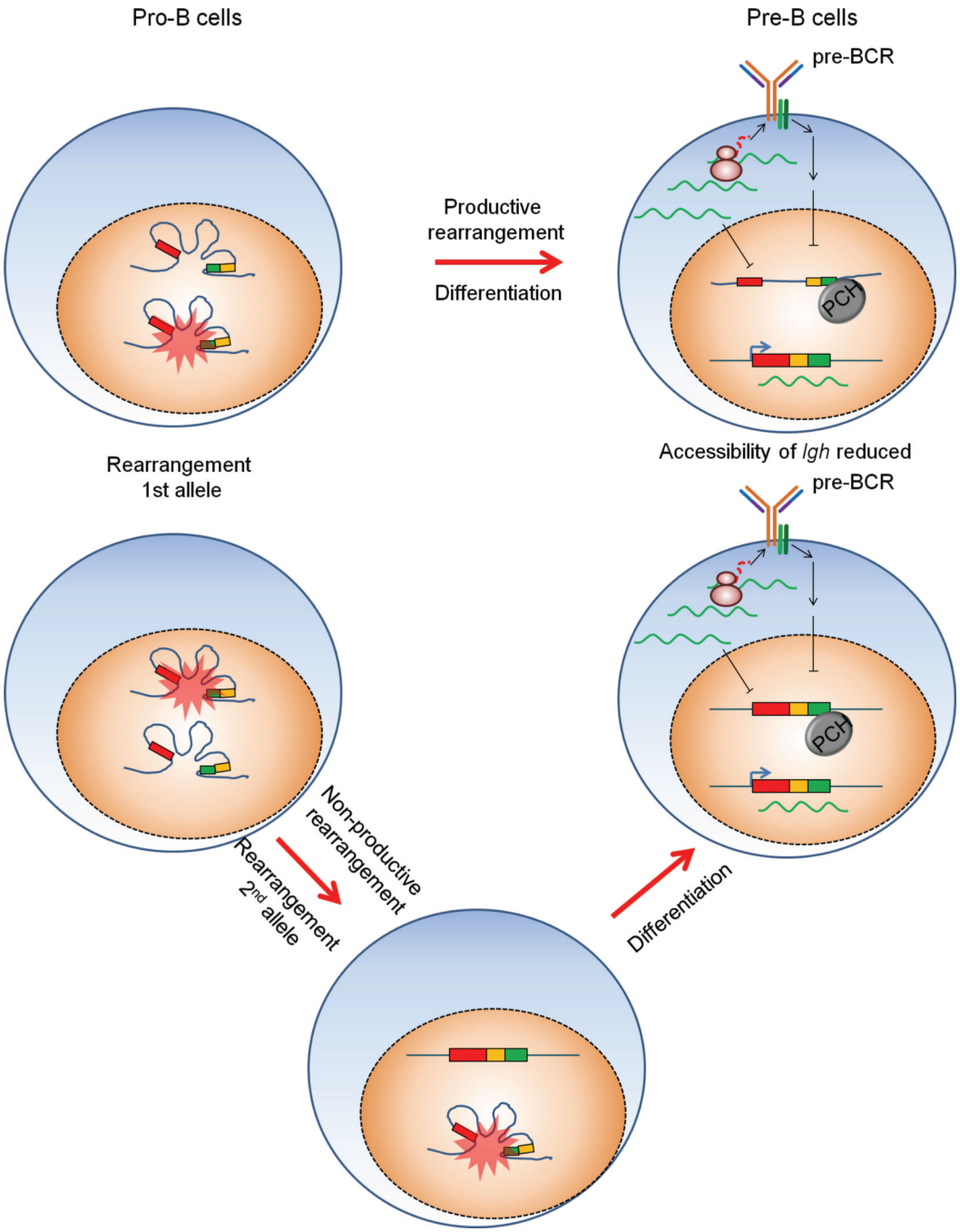


Figure 12.

Reduced accessibility of *Igh* downstream of pre-BCR signaling occurs as a result of differentiation and an alteration in signaling pathways that do not support continued accessibility of this locus

Table 1

Segment	Chr	Start	End	Strand	Size	Functional status
HSCTCF	chr12	113215906	113224658	.	8752	
HS3	chr12	113233598	113235248	.	1650	
HS12	chr12	113243680	113245543	.	1863	
HS3a	chr12	113249772	113251171	.	1399	
Igha	chr12	113258768	113260236	-	1468	Functional
Ighe	chr12	113269263	113273248	-	3985	Functional
Ighg2c	chr12	113287389	113288932	-	1543	Functional
Ighg2b	chr12	113304314	113307933	-	3619	Functional
Ighg1	chr12	113326544	113330523	-	3979	Functional
Ighg3	chr12	113357442	113361232	-	3790	Functional
Ighd	chr12	113407535	113416324	-	8789	Functional
Ighm	chr12	113418826	113422730	-	3904	Functional
Emu	chr12	113427403	113427623	.	220	
Ighj4	chr12	113428514	113428567	-	53	Functional
Ighj3	chr12	113429085	113429132	-	47	Functional
Ighj2	chr12	113429468	113429515	-	47	Functional
Ighj1	chr12	113429781	113429833	-	52	Functional
Ighd4-1	chr12	113430528	113430538	-	10	Functional
Ighd3-2	chr12	113448214	113448229	-	15	Functional
Ighd5-6	chr12	113449588	113449597	-	9	D ORF
Ighd2-8	chr12	113450851	113450867	-	16	Functional
Ighd5-5	chr12	113454942	113454951	-	9	D ORF
Ighd2-7	chr12	113456720	113456736	-	16	Functional
Ighd5-8	chr12	113459864	113459892	-	28	D ORF
Ighd5-4	chr12	113460101	113460110	-	9	D ORF
Ighd2-6	chr12	113461369	113461385	-	16	Functional
Ighd5-7	chr12	113464524	113464552	-	28	D ORF
Ighd5-3	chr12	113464761	113464770	-	9	D ORF
Ighd2-5	chr12	113466027	113466043	-	16	Functional
Ighd5-2	chr12	113469426	113469435	-	9	D ORF
Ighd2-4	chr12	113470694	113470710	-	16	Functional
Ighd6-2	chr12	113474875	113474903	-	28	pseudo
Ighd2-3	chr12	113475400	113475416	-	16	Functional
Ighd6-1	chr12	113480143	113480171	-	28	pseudo
Ighd1-1	chr12	113482170	113482192	-	22	Functional
IGCR1	chr12	113485095	113485503	.	408	
Ighd3-1	chr12	113525313	113525329	-	16	Functional
Ighd5-1	chr12	113526800	113526809	-	9	pseudo
Ighv5-1	chr12	113572929	113573222	-	293	pseudo
Ighv2-1	chr12	113574247	113574540	-	293	pseudo
Ighv5-2	chr12	113578504	113578990	-	486	Functional
Ighv2-2	chr12	113588267	113588700	-	433	Functional
Ighv5-3	chr12	113589967	113590243	-	276	pseudo
Ighv5-4	chr12	113597448	113597741	-	293	Functional
Ighv6-1	chr12	113603460	113603928	-	468	pseudo
Ighv2-3	chr12	113611184	113611476	-	292	Functional
Ighv5-5	chr12	113612965	113613442	-	477	pseudo
Ighv5-6	chr12	113625508	113625801	-	293	Functional

Ighv5-7	chr12	113634686	113634981	-	295	pseudo
Ighv2-4	chr12	113653291	113653583	-	292	Functional
Ighv5-8	chr12	113654967	113655245	-	278	pseudo
Ighv5-9	chr12	113661771	113662228	-	457	Functional
Ighv5-10	chr12	113669673	113669968	-	295	pseudo
Ighv2-5	chr12	113685482	113685774	-	292	Functional
Ighv5-11	chr12	113687229	113687505	-	276	pseudo
Ighv5-12	chr12	113702126	113702419	-	293	Functional
Ighv2-6	chr12	113716672	113716962	-	290	Functional
Ighv5-9-1	chr12	113736113	113736406	-	293	Functional
Ighv5-12-4	chr12	113762251	113762544	-	293	Functional
Ighv2-9-1	chr12	113769857	113770142	-	285	Functional
Ighv5-13	chr12	113794063	113794358	-	295	pseudo
Ighv2-6-8	chr12	113796138	113796430	-	292	Functional
Ighv2-7	chr12	113807314	113807747	-	433	Functional*
Ighv5-15	chr12	113826648	113826941	-	293	Functional
Ighv5-16	chr12	113838528	113838821	-	293	Functional
Ighv5-17	chr12	113859149	113859442	-	293	Functional
Ighv5-18	chr12	113875772	113876239	-	467	pseudo
Ighv2-8	chr12	113877361	113877659	-	298	pseudo
Ighv2-9	chr12	113879097	113879388	-	291	Functional
Ighv5-19	chr12	113884761	113885342	-	581	pseudo
Ighv7-1	chr12	113896408	113896946	-	538	Functional*
Ighv7-2	chr12	113912025	113912483	-	458	Functional*
Ighv14-1	chr12	113931953	113932246	-	293	Functional
Ighv4-1	chr12	113948284	113948577	-	293	Functional
Ighv3-1	chr12	113964390	113964683	-	293	Functional
Ighv11-1	chr12	113981879	113982174	-	295	Functional
Ighv14-2	chr12	113994469	113994898	-	429	Functional*
Ighv4-2	chr12	114013144	114013439	-	295	Functional*
Ighv3-2	chr12	114033803	114034097	-	294	Functional*
Ighv11-2	chr12	114048241	114048704	-	463	Functional
Ighv14-3	chr12	114059845	114060273	-	428	Functional
Ighv16-1	chr12	114068828	114069126	-	298	Functional
Ighv6-2	chr12	114089387	114089681	-	294	pseudo
Ighv9-1	chr12	114093928	114094221	-	293	Functional
Ighv12-1	chr12	114107259	114107563	-	304	pseudo
Ighv9-2	chr12	114109001	114109430	-	429	Functional
Ighv12-2	chr12	114127533	114127984	-	451	pseudo
Ighv9-3	chr12	114140692	114141122	-	430	Functional
Ighv7-3	chr12	114153180	114153644	-	464	Functional
Ighv15-1	chr12	114158025	114158315	-	290	pseudo
Ighv14-4	chr12	114176438	114176867	-	429	Functional
Ighv3-3	chr12	114196439	114196875	-	436	Functional*
Ighv7-4	chr12	114222788	114223254	-	466	Functional
Ighv3-4	chr12	114253617	114253915	-	298	Functional
Ighv3-5	chr12	114262652	114262950	-	298	Functional
Ighv13-1	chr12	114267607	114267906	-	299	Functional*
Ighv3-6	chr12	114288152	114288447	-	295	Functional

Ighv9-4	chr12	114299961	114300254	-	293	Functional
Ighv3-7	chr12	114314186	114314479	-	293	pseudo
Ighv5-21	chr12	114320030	114320324	-	294	pseudo
Ighv3-8	chr12	114322376	114322666	-	290	Functional
Ighv8-1	chr12	114332796	114333086	-	290	pseudo
Ighv1-1	chr12	114351705	114351995	-	290	pseudo
Ighv13-2	chr12	114357761	114358060	-	299	Functional
Ighv12-3	chr12	114366526	114366819	-	293	Functional
Ighv6-3	chr12	114391712	114392010	-	298	Functional
Ighv6-4	chr12	114406474	114406773	-	299	Functional
Ighv6-5	chr12	114416596	114416895	-	299	Functional
Ighv6-6	chr12	114434788	114435087	-	299	Functional
Ighv6-7	chr12	114455626	114455925	-	299	Functional
Ighv8-2	chr12	114462290	114462589	-	299	V ORF
Ighv1-2	chr12	114469075	114469371	-	296	pseudo
Ighv10-1	chr12	114479007	114479451	-	444	Functional
Ighv1-3	chr12	114481615	114481916	-	301	pseudo
Ighv1-4	chr12	114487136	114487429	-	293	Functional
Ighv10-2	chr12	114496003	114496300	-	297	pseudo
Ighv1-5	chr12	114513330	114513623	-	293	Functional
Ighv10-3	chr12	114523441	114523887	-	446	Functional
Ighv1-6	chr12	114532972	114533264	-	292	pseudo
Ighv1-7	chr12	114538495	114538788	-	293	Functional
Ighv10-4	chr12	114547258	114547556	-	298	pseudo
Ighv1-8	chr12	114555610	114555902	-	292	pseudo
Ighv15-2	chr12	114564578	114564874	-	296	Functional
Ighv1-9	chr12	114583569	114583862	-	293	Functional
Ighv1-10	chr12	114597878	114598175	-	297	pseudo
Ighv1-11	chr12	114612244	114612536	-	292	Functional
Ighv1-12	chr12	114615850	114616143	-	293	Functional
Ighv1-13	chr12	114630680	114630973	-	293	pseudo
Ighv1-14	chr12	114646572	114646865	-	293	Functional
Ighv1-15	chr12	114657353	114657646	-	293	Functional
Ighv1-16	chr12	114665872	114666165	-	293	V ORF
Ighv1-17-1	chr12	114671968	114672262	-	294	Functional
Ighv1-17	chr12	114676749	114677042	-	293	pseudo
Ighv1-18	chr12	114682632	114682925	-	293	Functional
Ighv1-19-1	chr12	114704147	114704440	-	293	pseudo
Ighv1-19	chr12	114708648	114708941	-	293	Functional
Ighv1-20	chr12	114723772	114724065	-	293	Functional
Ighv1-21-1	chr12	114736245	114736538	-	293	pseudo
Ighv1-21	chr12	114740383	114740840	-	457	pseudo
Ighv1-22	chr12	114746273	114746566	-	293	Functional
Ighv1-23	chr12	114764450	114764743	-	293	V ORF
Ighv1-24	chr12	114772928	114773221	-	293	V ORF
Ighv1-25	chr12	114781409	114781702	-	293	pseudo
Ighv1-26	chr12	114788372	114788665	-	293	Functional
Ighv1-27	chr12	114804616	114804909	-	293	pseudo
Ighv1-28	chr12	114809421	114809714	-	293	pseudo

Ighv1-29	chr12	114812321	114812605	-	284	pseudo	
Ighv1-30	chr12	114817099	114817390	-	291	pseudo	
Ighv1-31	chr12	114829264	114829557	-	293	Functional	
Ighv1-32	chr12	114835223	114835516	-	293	pseudo	
Ighv1-33	chr12	114843693	114844083	-	390	pseudo	
Ighv1-34	chr12	114851190	114851483	-	293	Functional	
Ighv1-35	chr12	114875302	114875597	-	295	pseudo	
Ighv1-36	chr12	114879888	114880181	-	293	Functional	
Ighv1-37	chr12	114896238	114896531	-	293	Functional	
Ighv1-38	chr12	114910011	114910303	-	292	pseudo	
Ighv1-39	chr12	114914599	114914892	-	293	Functional	
Ighv1-40	chr12	114924941	114925233	-	292	pseudo	
Ighv1-41	chr12	114932371	114932664	-	293	pseudo	
Ighv1-42	chr12	114937154	114937544	-	390	Functional*	
Ighv1-43	chr12	114945950	114946243	-	293	Functional	
Ighv1-44	chr12	114963380	114963674	-	294	pseudo	
Ighv1-45	chr12	114970055	114970348	-	293	pseudo	
Ighv1-46	chr12	114977693	114977986	-	293	pseudo	
Ighv1-47	chr12	114991108	114991401	-	293	Functional	
Ighv8-3	chr12	115003418	115003712	-	294	pseudo	
Ighv8-4	chr12	115024026	115024318	-	292	Functional	
Ighv1-48	chr12	115038227	115038520	-	293	pseudo	
Ighv1-49	chr12	115055223	115055516	-	293	Functional	
Ighv8-5	chr12	115067560	115067860	-	300	Functional	
PAIR1	chr12	115070994	115072036	.	1042		PAIR1
Ighv1-50	chr12	115119748	115120041	-	293	Functional	
Ighv1-51	chr12	115131375	115131668	-	293	pseudo	
Ighv1-52	chr12	115145487	115145780	-	293	Functional	
Ighv1-53	chr12	115158445	115158835	-	390	Functional	
Ighv8-6	chr12	115165777	115166077	-	300	Functional	
PAIR2	chr12	115171765	115172222	.	457		PAIR2
Ighv1-54	chr12	115193675	115193968	-	293	Functional	
Ighv1-55	chr12	115208137	115208527	-	390	Functional	
PAIR3	chr12	115222694	115223160	.	466		PAIR3
Ighv1-56	chr12	115242802	115243095	-	293	Functional	
Ighv8-7	chr12	115258139	115258432	-	293	pseudo	
PAIR4	chr12	115262443	115262909	.	466		PAIR4
Ighv1-57	chr12	115268424	115268715	-	291	pseudo	
Ighv8-8	chr12	115294066	115294362	-	296	Functional	
PAIR5	chr12	115297471	115298510	.	1039		PAIR5
Ighv1-58	chr12	115312166	115312597	-	431	Functional*	
Ighv1-59	chr12	115335057	115335513	-	456	Functional	
Ighv1-60	chr12	115344537	115344830	-	293	pseudo	
Ighv1-61	chr12	115359140	115359433	-	293	Functional	
Ighv1-62	chr12	115371986	115372278	-	292	pseudo	
Ighv1-62-1	chr12	115386707	115386988	-	281	Functional	
PAIR6	chr12	115407442	115407908	.	466		PAIR6
Ighv1-62-2	chr12	115446417	115446710	-	293	Functional	
Ighv1-62-3	chr12	115460999	115461431	-	432	Functional*	

Ighv8-9	chr12	115468331	115468632	-	301	Functional*	
PAIR7	chr12	115474319	115474776	.	457		PAIR7
Ighv1-63	chr12	115495625	115496058	-	433	Functional	
Ighv1-64	chr12	115507587	115507977	-	390	Functional	
Ighv8-10	chr12	115519340	115519640	-	300	pseudo	
PAIR8	chr12	115524113	115524608	.	495		PAIR8
Ighv1-65	chr12	115532055	115532347	-	292	pseudo	
Ighv8-11	chr12	115567149	115567449	-	300	Functional	
PAIR9	chr12	115574340	115574807	.	467		PAIR9
Ighv1-66	chr12	115593110	115593403	-	293	Functional	
Ighv1-67	chr12	115603940	115604233	-	293	Functional	
Ighv1-68	chr12	115611340	115611633	-	293	pseudo	
Ighv1-69	chr12	115623203	115623595	-	392	Functional	
Ighv8-12	chr12	115647945	115648245	-	300	Functional	
PAIR10	chr12	115655807	115656267	.	460		PAIR10
Ighv1-70	chr12	115691889	115692182	-	293	pseudo	
PAIR11	chr12	115703584	115704051	.	467		PAIR11
Ighv1-71	chr12	115742205	115742506	-	301	Functional	
Ighv1-72	chr12	115757984	115758277	-	293	Functional	
Ighv8-13	chr12	115765335	115765635	-	300	V ORF	
PAIR12	chr12	115771193	115771646	.	453		PAIR12
Ighv1-73	chr12	115790593	115790887	-	294	pseudo	
Ighv1-74	chr12	115802690	115803080	-	390	Functional	
Ighv8-14	chr12	115808427	115808727	-	300	pseudo	
PAIR13	chr12	115812742	115813210	.	468		PAIR13
PAIR14	chr12	115820652	115821118	.	466		PAIR14
Ighv1-75	chr12	115833950	115834243	-	293	Functional	
Ighv8-15	chr12	115844016	115844309	-	293	pseudo	
Ighv1-76	chr12	115847881	115848174	-	293	Functional	
Ighv8-16	chr12	115858008	115858301	-	293	pseudo	
Ighv1-77	chr12	115861867	115862160	-	293	Functional	
Ighv1-78	chr12	115868773	115869066	-	293	Functional	
Ighv1-79	chr12	115888096	115888389	-	293	pseudo	
Ighv1-80	chr12	115912344	115912637	-	293	Functional	
Ighv1-81	chr12	115920279	115920572	-	293	Functional	
Ighv1-82	chr12	115952538	115952831	-	293	Functional	
Ighv1-83	chr12	115963817	115964109	-	292	pseudo	
Ighv1-84	chr12	115980702	115981134	-	432	Functional	
Ighv1-85	chr12	116000028	116000321	-	293	Functional	
Ighv1-86	chr12	116009659	116009954	-	295	pseudo	

Segment	Chr	Start	End	Strand	Size	Functional status
Igkv2-137	chr6	67555636	67556216	+	580	Functional*
Igkv1-136	chr6	67593876	67594308	+	432	Pseudo
Igkv1-135	chr6	67609745	67610508	+	763	Functional
Igkv14-134-1	chr6	67711174	67711628	-	454	Pseudo
Igkv17-134	chr6	67720729	67721225	-	496	Pseudo
Igkv1-133	chr6	67724914	67725661	+	747	Functional
Igkv1-132	chr6	67759700	67760413	+	713	Functional
Igkv1-131	chr6	67766036	67766772	-	736	V ORF
Igkv14-130	chr6	67791046	67791511	+	465	Functional
Igkv9-129	chr6	67839793	67840266	+	473	Functional
Igkv9-128	chr6	67847409	67847874	+	465	Pseudo
Igkv17-127	chr6	67861159	67861659	+	500	Functional
Igkv14-126-1	chr6	67876011	67876477	+	466	Pseudo
Igkv14-126	chr6	67896177	67896642	+	465	Pseudo
Igkv11-125	chr6	67913573	67914052	+	479	Functional
Igkv9-124	chr6	67942076	67942540	-	464	Functional
Igkv9-123	chr6	67954230	67954690	-	460	Functional
Igkv1-122	chr6	68016742	68017488	+	746	Functional
Igkv17-121	chr6	68036818	68037318	+	500	Functional
Igkv9-120	chr6	68049983	68050454	+	471	Functional
Igkv9-119	chr6	68056345	68056810	+	465	Pseudo
Igkv14-118-2	chr6	68066369	68066828	+	459	Pseudo
Igkv14-118-1	chr6	68082547	68083145	+	598	Pseudo
Igkv11-118	chr6	68105414	68105880	+	466	Pseudo
Igkv1-117	chr6	68121091	68121825	+	734	Functional
Igkv2-116	chr6	68151885	68152617	+	732	Pseudo
Igkv1-115	chr6	68160797	68161716	+	919	Pseudo
Igkv11-114	chr6	68164459	68164923	+	464	Pseudo
Igkv2-113	chr6	68178801	68179564	+	763	Pseudo
Igkv2-112	chr6	68219981	68220705	+	724	Functional
Igkv14-111	chr6	68256404	68256869	+	465	Functional
Igkv1-110	chr6	68270527	68271265	+	738	Functional
Igkv2-109	chr6	68302439	68303158	+	719	Functional
Igkv1-108	chr6	68312054	68312359	+	305	Pseudo
Igkv2-107	chr6	68326098	68326752	+	654	Pseudo
Igkv11-106	chr6	68339527	68339995	+	468	Pseudo
Igkv2-105	chr6	68348661	68349441	+	780	Pseudo
Igkv16-104	chr6	68425605	68426072	+	467	Functional
Igkv15-103	chr6	68437468	68437925	+	457	V ORF
Igkv15-102	chr6	68466126	68466590	-	464	Pseudo
Igkv20-101-2	chr6	68474888	68475098	+	210	?
Igkv15-101-1	chr6	68479082	68479573	+	491	Pseudo
Igkv15-101	chr6	68481421	68481718	-	297	Pseudo
Igkv14-100	chr6	68519012	68519477	+	465	Functional
Igkv1-99	chr6	68541658	68542422	+	764	Functional
Igkv12-98	chr6	68570763	68571235	+	472	Functional
Igkv15-97	chr6	68591380	68591832	+	452	Pseudo

Igkv10-96	chr6	68631965	68632430	-	465	Functional*
Igkv2-95-2	chr6	68647999	68648531	-	532	Pseudo
Igkv2-95-1	chr6	68670752	68671335	+	583	Pseudo
Igkv10-95	chr6	68680379	68680848	+	469	Functional
Igkv10-94	chr6	68704508	68704978	-	470	Functional
Igkv2-93-1	chr6	68712925	68713508	-	583	Pseudo
Igkv19-93	chr6	68736297	68736764	-	467	Functional
Igkv4-92	chr6	68755038	68755593	-	555	Functional
Igkv4-91	chr6	68768555	68769103	-	548	Functional
Igkv4-90	chr6	68807180	68807708	-	528	Functional
Igkv13-89-1	chr6	68810495	68810543	+	48	Pseudo
Igkv12-89	chr6	68834846	68835307	-	461	Functional
Igkv1-88	chr6	68862265	68863031	-	766	Functional
Igkv13-87	chr6	68902801	68903266	+	465	Pseudo
Igkv4-86	chr6	68910411	68910938	-	527	Functional
Igkv13-85	chr6	68930269	68930734	-	465	Functional
Igkv13-84	chr6	68939602	68940067	+	465	Functional
Igkv4-83	chr6	68961378	68962019	-	641	Pseudo
Igkv13-82	chr6	68979198	68979663	+	465	Pseudo
Igkv4-81	chr6	68990758	68991294	-	536	Functional
Igkv13-80-1	chr6	69009143	69009513	+	370	Pseudo
Igkv4-80	chr6	69016558	69017080	-	522	Functional
Igkv4-79	chr6	69042972	69043505	-	533	Functional
Igkv13-78-1	chr6	69051444	69051794	+	350	Pseudo
Igkv4-78	chr6	69059690	69060224	-	534	Functional
Igkv4-77	chr6	69110904	69111432	-	528	Pseudo
Igkv13-76	chr6	69137866	69138332	+	466	Pseudo
Igkv4-75	chr6	69156118	69156659	-	541	Pseudo
Igkv13-74-1	chr6	69177416	69177787	+	371	Pseudo
Igkv4-74	chr6	69184826	69185361	-	535	Functional
Igkv13-73-1	chr6	69190073	69190282	+	209	Pseudo
Igkv4-73	chr6	69197583	69198116	-	533	Functional
Igkv4-72	chr6	69226854	69227383	-	529	Functional
Igkv13-71-1	chr6	69235927	69236158	+	231	Pseudo
Igkv4-71	chr6	69243160	69243688	-	528	Functional
Igkv4-70	chr6	69267888	69268412	-	524	Functional
Igkv4-69	chr6	69283794	69284319	-	525	Functional
Igkv4-68	chr6	69304834	69305363	-	529	Functional
Igkv12-67	chr6	69324177	69324646	-	469	Pseudo
Igkv12-66	chr6	69334604	69335076	-	472	Pseudo
Igkv4-65	chr6	69342472	69343004	-	532	Pseudo
Igkv13-64	chr6	69362722	69363187	+	465	Pseudo
Igkv4-63	chr6	69377944	69378472	-	528	Functional
Igkv13-62-1	chr6	69392439	69392815	+	376	Pseudo
Igkv4-62	chr6	69399812	69400344	-	532	V ORF
Igkv13-61-1	chr6	69409499	69409875	+	376	Pseudo
Igkv4-61	chr6	69416893	69417425	-	532	Functional
Igkv4-59	chr6	69438218	69438741	-	523	Functional

Igkv4-60	chr6	69463288	69463725	-	437	Pseudo
Igkv4-58	chr6	69500254	69500758	-	504	Functional
Igkv13-57-2	chr6	69523739	69524016	+	277	Pseudo
Igkv4-57-1	chr6	69544368	69544892	-	524	Functional
Igkv13-57-1	chr6	69569737	69570109	+	372	Pseudo
Igkv4-57	chr6	69575975	69576500	-	525	Functional
Igkv13-56-1	chr6	69581244	69581452	+	208	Pseudo
Igkv4-56	chr6	69587295	69587810	-	515	Pseudo
Igkv13-55-1	chr6	69599886	69600262	+	376	Pseudo
Igkv4-55	chr6	69607275	69607801	-	526	Functional
Igkv13-54-1	chr6	69617048	69617423	+	375	Pseudo
Igkv4-54	chr6	69631648	69632178	-	530	V ORF
Igkv4-53	chr6	69648824	69649354	-	530	Functional
Igkv4-51	chr6	69681406	69681939	-	533	Functional
Igkv4-50	chr6	69700767	69701287	-	520	Functional
Igkv12-49	chr6	69716498	69716956	-	458	Pseudo
Igkv5-48	chr6	69726573	69727125	-	552	Functional
Igkv12-47	chr6	69750854	69751310	-	456	Pseudo
Igkv12-46	chr6	69764523	69764992	-	469	Functional
Igkv5-45	chr6	69775750	69776306	-	556	Functional
Igkv12-44	chr6	69814631	69815100	-	469	Functional
Igkv5-43	chr6	69823355	69823911	-	556	Functional
Igkv12-42	chr6	69834792	69835245	-	453	Pseudo
Igkv12-41	chr6	69858420	69858884	-	464	Functional
Igkv5-40-1	chr6	69868595	69869193	-	598	Pseudo
Igkv12-40	chr6	69879350	69879819	-	469	Pseudo
Igkv5-39	chr6	69900424	69900977	-	553	Functional
Igkv12-38	chr6	69943186	69943648	-	462	Functional
Igkv5-37	chr6	69963312	69963861	-	549	Functional
Igkv18-36	chr6	69992465	69992977	-	512	Functional
Igkv1-35	chr6	70010949	70011673	-	724	V ORF
Igkv8-34	chr6	70044112	70044678	-	566	Functional
Igkv7-33	chr6	70058632	70059199	-	567	Functional
Igkv6-32	chr6	70074024	70074584	-	560	Functional
Igkv8-31	chr6	70105860	70106179	-	319	Pseudo
Igkv8-30	chr6	70117061	70117617	-	556	Functional
Igkv6-29	chr6	70138462	70139082	-	620	Functional
Igkv8-28	chr6	70143593	70144161	-	568	Functional
Igkv8-27	chr6	70171809	70172270	-	461	Functional
Igkv8-26	chr6	70193228	70193797	+	569	V ORF
Igkv6-25	chr6	70215433	70215957	+	524	Functional
Igkv8-24	chr6	70216858	70217421	-	563	Functional
Igkv8-23-1	chr6	70240427	70240710	+	283	V ORF
Igkv6-23	chr6	70260409	70260933	-	524	Functional
Igkv8-22	chr6	70297554	70298119	-	565	Pseudo
Igkv8-21	chr6	70314895	70315452	-	557	Functional
Igkv6-20	chr6	70335841	70336479	-	638	Functional
Igkv8-19	chr6	70340876	70341448	-	572	Functional

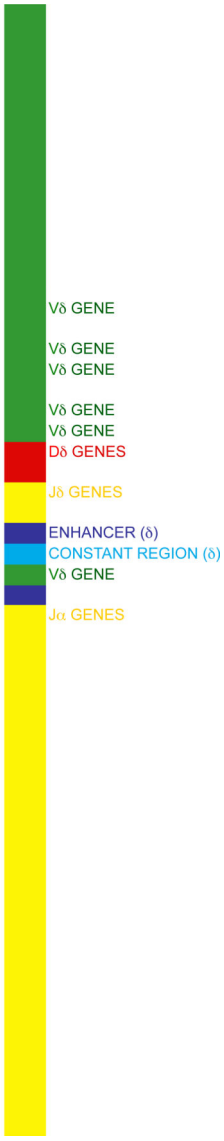
Igkv8-18	chr6	70355877	70356445	+	568	V ORF
Igkv6-17	chr6	70371469	70371993	+	524	Functional
Igkv8-16	chr6	70386672	70387238	-	566	Functional
Igkv6-15	chr6	70406469	70406992	-	523	Functional
Igkv6-14	chr6	70434952	70435476	-	524	Functional
Igkv6-13	chr6	70457503	70458036	-	533	Functional
Igkv3-12-1	chr6	70497662	70498342	-	680	?
Igkv3-12	chr6	70518250	70518849	+	599	Functional
Igkv3-11	chr6	70553274	70554440	+	1166	Pseudo
Igkv3-10	chr6	70572633	70573230	+	597	Functional
Igkv3-9	chr6	70588189	70588777	+	588	Functional
Igkv3-8	chr6	70588871	70589452	+	581	Pseudo
Igkv3-7	chr6	70607437	70608036	+	599	Functional
Igkv3-6	chr6	70642405	70643565	+	1160	Pseudo
Igkv3-5	chr6	70663298	70663895	+	597	Functional
Igkv3-4	chr6	70671789	70672377	+	588	Functional
Igkv3-3	chr6	70686946	70687534	+	588	Functional
Igkv3-2	chr6	70698468	70699067	+	599	Functional
Igkv3-1	chr6	70703578	70704177	+	599	Functional
Cer	chr6	70709374	70710025	.	651	
Sis	chr6	70712108	70715840	.	3732	
Igkj1	chr6	70722562	70722599	+	37	Functional
Igkj2	chr6	70722916	70722954	+	38	Functional
Igkj3	chr6	70723223	70723260	+	37	Pseudo
Igkj4	chr6	70723548	70723585	+	37	Functional
Igkj5	chr6	70723886	70723923	+	37	Functional
MiEk	chr6	70725210	70725956	.	746	
Igkc	chr6	70726434	70726754	+	320	Functional
3Ek	chr6	70735257	70736522	.	1265	
Edk	chr6	70743710	70744950	.	1240	

Segment	Chr	Start	End	Strand	Size	Functional status
Trbv1	chr6	40891296	40891885	+	589	Functional
Trbv2	chr6	41047556	41047995	+	439	Functional
Trbv3	chr6	41048394	41048824	+	430	Functional
Trbv4	chr6	41059443	41059886	+	443	Functional
Trbv5	chr6	41062359	41062803	+	444	Functional
Trbv6	chr6	41066871	41067154	+	283	Pseudo
Trbv7	chr6	41080253	41080654	+	401	Pseudo
Trbv8	chr6	41084669	41085107	+	438	Pseudo
Trbv9	chr6	41087991	41088288	+	297	Pseudo
Trbv10	chr6	41091935	41092228	+	293	Pseudo
Trbv11	chr6	41106805	41107303	+	498	Pseudo
Trbv12-1	chr6	41113567	41114067	+	500	Functional
Trbv13-1	chr6	41116036	41116468	+	432	Functional
Trbv12-2	chr6	41118864	41119364	+	500	Functional
Trbv13-2	chr6	41121396	41121832	+	436	Functional
Trbv12-3	chr6	41127592	41128076	+	484	Pseudo
Trbv13-3	chr6	41130147	41130585	+	438	Functional
Trbv14	chr6	41135167	41135616	+	449	Functional
Trbv15	chr6	41141211	41141658	+	447	Functional
Trbv16	chr6	41151791	41152230	+	439	Functional
Trbv17	chr6	41163092	41163556	+	464	Functional
Trbv18	chr6	41175137	41175608	+	471	Pseudo
Trbv19	chr6	41178589	41179048	+	459	Functional
Trbv20	chr6	41188273	41188977	+	704	Functional
Trbv21	chr6	41202587	41203044	+	457	V ORF
Trbv22	chr6	41212075	41212534	+	459	Pseudo
Trbv23	chr6	41216073	41216526	+	453	Functional
Trbv24	chr6	41218091	41218557	+	466	Functional
Trbv25	chr6	41221674	41222210	+	536	Pseudo
Trbv26	chr6	41227525	41227913	+	388	Functional
Trbv27	chr6	41258121	41258579	+	458	Pseudo
Trbv28	chr6	41266697	41267150	+	453	Pseudo
Trbv29	chr6	41271403	41271881	+	478	Functional
Trbv30	chr6	41281376	41281990	+	614	Functional
Trbd1	chr6	41533201	41533212	+	11	Functional
Trbj1-1	chr6	41533864	41533911	+	47	Functional
Trbj1-2	chr6	41534001	41534048	+	47	Functional
Trbj1-3	chr6	41534323	41534372	+	49	Functional
Trbj1-4	chr6	41534811	41534861	+	50	Functional
Trbj1-5	chr6	41535084	41535133	+	49	Functional
Trbj1-6	chr6	41535554	41535606	+	52	Pseudo
Trbj1-7	chr6	41535642	41535687	+	45	Pseudo
Trbc1	chr6	41538219	41539665	+	1446	Functional
Trbd2	chr6	41542163	41542176	+	13	Functional
Trbj2-1	chr6	41542754	41542803	+	49	Functional
Trbj2-2	chr6	41542957	41543007	+	50	Functional
Trbj2-3	chr6	41543223	41543271	+	48	Functional
Trbj2-4	chr6	41543362	41543410	+	48	Functional
Trbj2-5	chr6	41543453	41543501	+	48	Functional
Trbj2-6	chr6	41543598	41543645	+	47	Pseudo
Trbj2-7	chr6	41543810	41543856	+	46	Functional
Trbc2	chr6	41546730	41548181	+	1451	Functional
Eb	chr6	41554187	41554745	.	558	
Trbv31	chr6	41557693	41558371	.	678	Functional

Segment	Chr	Start	End	Strand	Size	Functional status
Trav1	chr14	52427967	52428876	+	909	Functional
Trav2	chr14	52567293	52568053	+	760	Functional
Trav3-1	chr14	52580762	52581213	+	451	Functional
Trav4-1	chr14	52607307	52607804	+	497	Pseudo
Trav5-1	chr14	52622566	52623082	+	516	Functional
Trav6-1	chr14	52638515	52638962	+	447	Functional
Trav7-1	chr14	52654817	52655329	+	512	Functional
Trav6-2	chr14	52667411	52667864	+	453	Functional
Trav7d-2	chr14	52683964	52684472	+	508	Functional
Trav4d-2	chr14	52710917	52711402	+	485	Pseudo
Trav6d-3	chr14	52725308	52726919	+	1611	Functional
Trav7d-3	chr14	52744314	52744838	+	524	Functional
Trav6d-4	chr14	52753398	52753846	+	448	Functional
Trav7d-4	chr14	52769863	52770392	+	529	Functional
Trav8d-1	chr14	52778510	52778995	+	485	Functional
Trav9d-1	chr14	52792316	52792776	+	460	Functional
Trav6d-5	chr14	52795165	52795633	+	468	Functional
Trav10d	chr14	52810997	52811496	+	499	Functional*
Trav6d-6	chr14	52821922	52822359	+	437	Functional
Trav11d	chr14	52824357	52824913	+	556	Functional
Trav7d-5	chr14	52835861	52836387	+	526	Functional
Trav12d-1	chr14	52844560	52845041	+	481	Functional
Trav13d-1	chr14	52851355	52851866	+	511	Functional
Trav14d-1	chr14	52855665	52856166	+	501	Functional*
Trav15d-1-dv6d-1	chr14	52863365	52863933	+	568	Functional
Trav3d-2	chr14	52887587	52888080	+	493	Pseudo
Trav9d-2	chr14	52892186	52892618	+	432	Functional*
Trav4d-3	chr14	52899972	52900470	+	498	Functional
Trav5d-2	chr14	52914417	52914702	+	285	Pseudo
Trav12d-2	chr14	52915915	52916426	+	511	Functional
Trav9d-3	chr14	52920678	52921110	+	432	Functional
Trav5d-3	chr14	52928320	52928532	+	212	Pseudo
Trav12d-3	chr14	52929789	52930282	+	493	Functional*
Trav13d-2	chr14	52942764	52943289	+	525	Functional
Trav14d-2	chr14	52948668	52949169	+	501	Functional
Trav15d-2-dv6d-2	chr14	52956828	52957400	+	572	Functional
Trav3d-3	chr14	52972532	52973038	+	506	Functional
Trav9d-4	chr14	52983456	52983770	+	314	Functional
Trav4d-4	chr14	52990889	52991389	+	500	Functional
Trav5d-4	chr14	53001713	53002231	+	518	Functional*
Trav6d-7	chr14	53007632	53008098	+	466	Functional
Trav7d-6	chr14	53020459	53020991	+	532	Functional
Trav13d-3	chr14	53032902	53033419	+	517	Functional
Trav8d-2	chr14	53042428	53042891	+	463	Functional
Trav16d-dv11	chr14	53047287	53047821	+	534	Functional
Trav13d-4	chr14	53072763	53073275	+	512	Functional
Trav14d-3-dv8	chr14	53078547	53079045	+	498	Functional
Trav15d-3	chr14	53082063	53082522	+	459	Pseudo
Trav7n-4	chr14	53091346	53091875	+	529	Functional
Trav9n-1	chr14	53101801	53102265	+	464	Pseudo
Trav6n-5	chr14	53104871	53105344	+	473	Functional
Trav10n	chr14	53122114	53122611	+	497	Functional
Trav6n-6	chr14	53132700	53133137	+	437	Functional
Trav11n	chr14	53135122	53135678	+	556	Pseudo
Trav7n-5	chr14	53146922	53147448	+	526	Functional
Trav12n-1	chr14	53155607	53156088	+	481	Functional

Trav13n-1	chr14	53162399	53162910	+	511	Functional	
Trav14n-1	chr14	53166677	53167178	+	501	Functional	
Trav15n-1	chr14	53174325	53174896	+	571	Functional	
Trav3n-2	chr14	53198525	53199018	+	493	Pseudo	
Trav9n-2	chr14	53203122	53203544	+	422	Functional	
Trav4n-3	chr14	53210906	53211407	+	501	Functional	
Trav5n-2	chr14	53225393	53225669	+	276	Pseudo	
Trav12n-2	chr14	53226884	53227400	+	516	Functional	
Trav9n-3	chr14	53231668	53232102	+	434	Functional	
Trav5n-3	chr14	53239284	53239563	+	279	Pseudo	
Trav12n-3	chr14	53240751	53241244	+	493	Pseudo	
Trav13n-2	chr14	53253792	53254317	+	525	Functional	
Trav14n-2	chr14	53259678	53260179	+	501	Functional	
Trav15n-2	chr14	53267963	53268535	+	572	Functional	
Trav3n-3	chr14	53283669	53284175	+	506	Functional	
Trav9n-4	chr14	53294596	53295030	+	434	Functional	
Trav4n-4	chr14	53302059	53302559	+	500	Functional	
Trav5n-4	chr14	53312904	53313422	+	518	Functional	
Trav6n-7	chr14	53318818	53319284	+	466	Functional	
Trav7n-6	chr14	53324676	53325208	+	532	Functional	
Trav13n-3	chr14	53337122	53337639	+	517	Functional	
Trav8n-2	chr14	53345961	53346424	+	463	Functional	
Trav16n	chr14	53351090	53351621	+	531	Functional	
Trav13n-4	chr14	53362368	53364106	+	1738	Functional	
Trav14n-3	chr14	53370077	53370575	+	498	Functional	
Trav15n-3	chr14	53374821	53375278	+	457	Pseudo	
Trav7-2	chr14	53390633	53391136	+	503	Functional	
Trav4-2	chr14	53418388	53418873	+	485	Functional	
Trav6-3	chr14	53428761	53430377	+	1616	Functional	
Trav7-3	chr14	53443314	53443839	+	525	Functional	
Trav6-4	chr14	53454327	53454784	+	457	Functional	
Trav7-4	chr14	53461209	53461738	+	529	Functional	
Trav8-1	chr14	53469756	53470231	+	475	Functional	
Trav9-1	chr14	53488106	53488567	+	461	Functional	
Trav6-5	chr14	53491152	53491622	+	470	Functional	
Trav10	chr14	53505790	53506286	+	496	Functional*	
Trav6-6	chr14	53516929	53517366	+	437	Functional	
Trav11	chr14	53519303	53519859	+	556	Functional*	
Trav7-5	chr14	53530786	53531313	+	527	Functional	
Trav12-1	chr14	53538266	53538738	+	472	Functional*	
Trav13-1	chr14	53545014	53545525	+	511	Functional	
Trav14-1	chr14	53554057	53554558	+	501	Functional	
Trav15-1-dv6-1	chr14	53559676	53560247	+	571	Functional	
Trav3-2	chr14	53586471	53586964	+	493	Pseudo	
Trav9-2	chr14	53591080	53591504	+	424	Functional	
Trav4-3	chr14	53598911	53599410	+	499	Functional	
Trav5-2	chr14	53614936	53615224	+	288	Pseudo	
Trav12-2	chr14	53616397	53616914	+	517	Functional	
Trav12-3	chr14	53621752	53622245	+	493	Functional	
Trav13-2	chr14	53634888	53635399	+	511	Functional	
Trav14-2	chr14	53640775	53641225	+	450	Functional	
Trav15-2-dv6-2	chr14	53649425	53649994	+	569	Functional	
Trav3-3	chr14	53666005	53666506	+	501	Functional	
Trav9-4	chr14	53676196	53676628	+	432	Functional	
Trav4-4-dv10	chr14	53683677	53684177	+	500	Functional	
Trav5-4	chr14	53704007	53704514	+	507	V ORF	
Trav6-7-dv9	chr14	53709943	53710419	+	476	Functional*	

Trav7-6	chr14	53716760	53717291	+	531	Functional	
Trav13-3	chr14	53729558	53730074	+	516	Functional*	
Trav8-2	chr14	53738375	53738838	+	463	Functional	
Trav16n	chr14	53743174	53743705	+	531	Functional	
Trav13-4-dv7	chr14	53757410	53757921	+	511	Functional	
Trav14-3	chr14	53763196	53763694	+	498	Functional	
Trav15-3	chr14	53766735	53767179	+	444	Pseudo	
Trav3-4	chr14	53777071	53777576	+	505	Functional	
Trav12-4	chr14	53778603	53778845	+	242	Pseudo	
Trav13-5	chr14	53795455	53795963	+	508	Functional	
Trav17	chr14	53806639	53807115	+	476	Functional	
Trav18	chr14	53831105	53831827	+	722	V ORF	
Trav19	chr14	53845322	53845827	+	505	Functional*	
Trav20	chr14	53863110	53863661	+	551	Pseudo	
Trav21/dv12	chr14	53876016	53876752	+	736	Functional	
Trdv1	chr14	53881612	53882217	+	605	Functional	Vδ GENE
Trav22	chr14	53926978	53927595	+	617	Pseudo	
Trdv2-1	chr14	53946073	53946660	+	587	Functional	Vδ GENE
Trdv2-2	chr14	53960993	53961602	+	609	Functional	Vδ GENE
Trav23	chr14	53977153	53977678	+	525	Pseudo	
Trdv3	chr14	54000948	54001253	+	305	Pseudo	Vδ GENE
Trdv4	chr14	54075004	54075515	+	511	Functional	Vδ GENE
Trdd1	chr14	54113468	54113476	+	8	Functional	Dδ GENES
Trdd2	chr14	54122226	54122241	+	15	Functional	
Trdj1	chr14	54123138	54123188	+	50	Functional	Jδ GENES
Trdj2	chr14	54136779	54136837	+	58	Functional	
Ed	chr14	54138465	54138512	+	47		ENHANCER (δ)
Trdc	chr14	54142851	54146108	+	3257	Functional	CONSTANT REGION (δ)
Trdv5	chr14	54148662	54149201	-	539	Functional	Vδ GENE
TEAp	chr14	54152995	54153117	+	122		
Traj61	chr14	54155005	54155076	+	71	Pseudo	Jα GENES
Traj60	chr14	54155933	54155985	+	52	Pseudo	
Traj59	chr14	54156133	54156194	+	61	J ORF	
Traj58	chr14	54157280	54157342	+	62	Functional	
Traj57	chr14	54158507	54158569	+	62	Functional	
Traj56	chr14	54159263	54159325	+	62	Functional	
Traj55	chr14	54161439	54161501	+	62	Pseudo	
Traj54	chr14	54161990	54162043	+	53	Pseudo	
Traj53	chr14	54162644	54162709	+	65	Functional	
Traj52	chr14	54165316	54165381	+	65	Functional	
Traj51	chr14	54166239	54166281	+	42	Pseudo	
Traj50	chr14	54167590	54167652	+	62	Functional	
Traj49	chr14	54168686	54168744	+	58	Functional	
Traj48	chr14	54169808	54169868	+	60	Functional	
Traj47	chr14	54171802	54171856	+	54	J ORF	
Traj46	chr14	54172338	54172398	+	60	J ORF	
Traj45	chr14	54172831	54172890	+	59	Functional	
Traj44	chr14	54173690	54173750	+	60	J ORF	
Traj43	chr14	54174742	54174798	+	56	Functional	
Traj42	chr14	54175773	54175836	+	63	Functional	
Traj41	chr14	54176271	54176325	+	54	J ORF	
Traj40	chr14	54177921	54177981	+	60	Functional	
Traj39	chr14	54179962	54180024	+	62	Functional	
Traj38	chr14	54180574	54180635	+	61	Functional	
Traj37	chr14	54181518	54181577	+	59	Functional	
Traj36	chr14	54182402	54182464	+	62	Pseudo	
Traj35	chr14	54183774	54183838	+	64	Functional	



Traj34	chr14	54184699	54184756	+	57	Functional
Traj33	chr14	54185358	54185414	+	56	Functional
Traj32	chr14	54186101	54186166	+	65	Functional
Traj31	chr14	54187895	54187951	+	56	Functional
Traj30	chr14	54189866	54189924	+	58	Functional
Traj29	chr14	54190946	54191005	+	59	J ORF
Traj28	chr14	54191661	54191725	+	64	Functional
Traj27	chr14	54192303	54192361	+	58	Functional
Traj26	chr14	54194490	54194549	+	59	Pseudo
Traj25	chr14	54194815	54194871	+	56	J ORF
Traj24	chr14	54195645	54195700	+	55	Functional
Traj23	chr14	54196081	54196140	+	59	Functional
Traj22	chr14	54197248	54197307	+	59	Functional
Traj21	chr14	54198790	54198846	+	56	Functional
Traj20	chr14	54199441	54199498	+	57	J ORF
Traj19	chr14	54200386	54200446	+	60	J ORF
Traj18	chr14	54200777	54200842	+	65	Functional
Traj17	chr14	54201775	54201837	+	62	Functional
Traj16	chr14	54203134	54203194	+	60	Functional
Traj15	chr14	54204422	54204481	+	59	Functional
Traj14	chr14	54205111	54205144	+	33	Pseudo
Traj13	chr14	54205741	54205797	+	56	Functional
Traj12	chr14	54206552	54206610	+	58	Functional
Traj11	chr14	54207139	54207197	+	58	Functional
Traj9	chr14	54209393	54209450	+	57	Functional
Traj8	chr14	54209918	54209954	+	36	Pseudo
Traj7	chr14	54211470	54211528	+	58	J ORF
Traj6	chr14	54212688	54212749	+	61	Functional
Traj5	chr14	54213777	54213838	+	61	Functional
Traj4	chr14	54216291	54216353	+	62	J ORF
Traj3	chr14	54217292	54217357	+	65	J ORF
Traj2	chr14	54217836	54217901	+	65	Functional
Traj1	chr14	54218814	54218842	+	28	Pseudo
Trac	chr14	54220521	54224198	+	3677	Functional
Ea	chr14	54227374	54227584	+	210