Clonally stable V\(\kappa\) allelic choice instructs Ig\(\kappa\) repertoire

Rena Levin-Klein\(^1\), Shira Fraenkel\(^1\), Michal Lichtenstein\(^1\), Louise S. Matheson\(^2\), Osnat Bartok\(^3\), Yuval Nevo\(^4\), Sebastian Kadener\(^3\), Anne E. Corcoran\(^2\), Howard Cedar\(^1\) & Yehudit Bergman\(^1\)

Although much has been done to understand how rearrangement of the Ig\(\kappa\) locus is regulated during B-cell development, little is known about the way the variable (V) segments themselves are selected. Here we show, using B6/Cast hybrid pre-B-cell clones, that a limited number of V segments on each allele is stochastically activated as characterized by the appearance of non-coding RNA and histone modifications. The activation states are clonally distinct, stable across cell division and developmentally important in directing the Ig repertoire upon differentiation. Using a new approach of allelic ATAC-seq, we demonstrate that the Ig\(\kappa\) V alleles have differential chromatin accessibility, which may serve as the underlying basis of clonal maintenance at this locus, as well as other instances of monallelic expression throughout the genome. These findings highlight a new level of immune system regulation that optimizes gene diversity.

\(^1\)Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel–Canada, Hebrew University Medical School, Jerusalem 91120, Israel. \(^2\)Nuclear Dynamics Programme, Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, UK. \(^3\)Biological Chemistry Department, Silberman Institute of Life Sciences, Edmund J. Safra Campus, The Hebrew University, Jerusalem 91904, Israel. \(^4\)Computation Center, Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel. Correspondence and requests for materials should be addressed to Y.B. (email: yehuditb@ekmd.huji.ac.il).
Rearrangement of immune receptor loci in B and T lymphocytes takes place in an ordered developmental manner using transcription factors and regulatory elements to open up and turn on the rearrangement process at each individual cluster during its specific stage of differentiation. In the B-cell lineage, the IgH locus is activated first in pro-B cells, whereas the Igk region gets turned on and rearranged only at a later stage of development in the small pre-B-cell compartment. This activation occurs initially on only one allele, which undergoes J–C region demethylation and proceeds with rearrangement seemingly choosing from the full range of V segments.

Originally, it was thought that at the time of rearrangement the two k alleles in each cell are equal substrates for activation, with the choice being made in a stochastic manner. Previous work in our laboratory, however, has indicated that this is probably not the case and the decision is actually of an instructive nature, with the two alleles first becoming marked by asynchronous replication at the early lymphoid progenitor stage followed later by opening of the k J–C region specifically on the early allele. Through the use of pre-B-cell clones, it was then demonstrated that it is this same allele that undergoes the first rearrangement in each cell.

The k locus is distributed over a large 3 Mb region carrying ~140 different V segments and this domain already has an accessible chromatin conformation at the pre-B-cell stage even prior to the initiation of rearrangement. However, the actual chromatin structure and transcription pattern of individual V segments on the two alleles has not yet been identified.

In this study, we use hybrid C57BL/6/Castaneous (B6/Cast) pre-B-cell clones to examine the chromatin and transcriptional state of the k locus V segments in an allele-specific manner. The results indicate that each parental chromosome independently activates a select number of V segments. Once chosen, these activity states are then maintained in clonal populations probably through their highly stable accessible chromatin structure. In the case of the Igk locus, this ‘choice’ of V segments seems to generate alternate recombination patterns on each allele, thus providing a mechanism for enhancing the chances of each B cell to produce functional antibodies. Furthermore, this same chromatin-based model may also serve as the basis for the maintenance of differential expression at a large number of monoallelic loci present in the genome.

Results

V region allele-specific histone modification states. To determine the pattern of V region activation states, we analysed histone acetylation over select V segments in pre-B-cell clones derived from chimeric B6/Cast mice. Since, in general, the sequences of the two alleles differ by about 1% genome wide, we were able to identify many polymorphic sites that could be used to determine the histone acetylation pattern of each allele separately. We first chose a single clone (E9-3) and carried out anti-histone H3Ac ChIP, which was then assayed by PCR analysis of various V segments within the k locus, using polymorphisms at restriction-enzyme binding sites to distinguish between the alleles (Fig. 1a). In a striking manner, it appears that individual Vs are acetylated in a monoallelic manner. Nonetheless, because they represent a mixture of individual clones that choose to assay ncRNA transcription for a small repertoire of 20 different Vs (Fig. 2a). In any given clone, we found that each V was either silenced on both alleles, expressed preferentially on one allele (B6 or Cast) or displayed a biallelic pattern (Fig. 2b). This suggests that the original choice of Vs in each clonal population may be based on a stochastic process. Strikingly, no two clones had the exact same transcription pattern of the V segments.

As a further proof of this principle, we next adapted a method for genome-wide analysis of the total nuclear RNA fraction and quantified RNA levels over the entire V region of both alleles (Fig. 3). The overall pattern of transcription across the locus was distinct in each of the clones analysed (Fig. 3a, Supplementary Fig. 2a) and in every case, we also observed striking differences between the two parental alleles (Fig. 3b, Supplementary Fig. 2b,c), generating a kaleidoscope of gene expression (Supplementary Fig. 3). Furthermore, these profiles remain stable across multiple divisions and cell passages as confirmed by principal component analysis (PCA) (Fig. 3c,d). Interestingly, clone 3 apparently underwent deletion of the Vk locus Cast allele in late passages. Nonetheless, the pattern of V segment expression on the B6 allele remained constant (Supplementary Fig. 4). This suggests that each allele profile is maintained independently, with no crosstalk between the alleles being required.

V region allele-specific chromatin accessibility. Since expression is always correlated with local histone acetylation on an allelic basis, it appears that, in general, some V segments are relatively open and accessible within the nucleus while others are in a more closed configuration (Fig. 1). To test this idea, we carried out

---

**V region allele-specific transcription.** To obtain an expanded picture of V region choice, we developed a multiplex PCR method to assay ncRNA transcription for a small repertoire of 20 different Vs (Fig. 2a). In any given clone, we found that each V was either silenced on both alleles, expressed preferentially on one allele (B6 or Cast) or displayed a biallelic pattern (Fig. 2b). This suggests that the original choice of Vs in each clonal population may be based on a stochastic process. Strikingly, no two clones had the exact same transcription pattern of the V segments.

---

**V region allele-specific histone modification states.** To determine the pattern of V region activation states, we analysed histone acetylation over select V segments in pre-B-cell clones derived from chimeric B6/Cast mice. Since, in general, the sequences of the two alleles differ by about 1% genome wide, we were able to identify many polymorphic sites that could be used to determine the histone acetylation pattern of each allele separately. We first chose a single clone (E9-3) and carried out anti-histone H3Ac ChIP, which was then assayed by PCR analysis of various V segments within the k locus, using polymorphisms at restriction-enzyme binding sites to distinguish between the alleles (Fig. 1a). In a striking manner, it appears that individual Vs are acetylated in a monoallelic manner. This was confirmed by principal component analysis (PCA) (Fig. 3c,d). Interestingly, clone 3 apparently underwent deletion of the Vk locus Cast allele in late passages. Nonetheless, the pattern of V segment expression on the B6 allele remained constant (Supplementary Fig. 4). This suggests that each allele profile is maintained independently, with no crosstalk between the alleles being required.

---

**V region allele-specific chromatin accessibility.** Since expression is always correlated with local histone acetylation on an allelic basis, it appears that, in general, some V segments are relatively open and accessible within the nucleus while others are in a more closed configuration (Fig. 1). To test this idea, we carried out

---

**V region allele-specific transcription.** To obtain an expanded picture of V region choice, we developed a multiplex PCR method to assay ncRNA transcription for a small repertoire of 20 different Vs (Fig. 2a). In any given clone, we found that each V was either silenced on both alleles, expressed preferentially on one allele (B6 or Cast) or displayed a biallelic pattern (Fig. 2b). This suggests that the original choice of Vs in each clonal population may be based on a stochastic process. Strikingly, no two clones had the exact same transcription pattern of the V segments.

---

**V region allele-specific histone modification states.** To determine the pattern of V region activation states, we analysed histone acetylation over select V segments in pre-B-cell clones derived from chimeric B6/Cast mice. Since, in general, the sequences of the two alleles differ by about 1% genome wide, we were able to identify many polymorphic sites that could be used to determine the histone acetylation pattern of each allele separately. We first chose a single clone (E9-3) and carried out anti-histone H3Ac ChIP, which was then assayed by PCR analysis of various V segments within the k locus, using polymorphisms at restriction-enzyme binding sites to distinguish between the alleles (Fig. 1a). In a striking manner, it appears that individual Vs are acetylated in a monoallelic manner. This was confirmed by principal component analysis (PCA) (Fig. 3c,d). Interestingly, clone 3 apparently underwent deletion of the Vk locus Cast allele in late passages. Nonetheless, the pattern of V segment expression on the B6 allele remained constant (Supplementary Fig. 4). This suggests that each allele profile is maintained independently, with no crosstalk between the alleles being required.

---

**V region allele-specific chromatin accessibility.** Since expression is always correlated with local histone acetylation on an allelic basis, it appears that, in general, some V segments are relatively open and accessible within the nucleus while others are in a more closed configuration (Fig. 1). To test this idea, we carried out

---

**V region allele-specific transcription.** To obtain an expanded picture of V region choice, we developed a multiplex PCR method to assay ncRNA transcription for a small repertoire of 20 different Vs (Fig. 2a). In any given clone, we found that each V was either silenced on both alleles, expressed preferentially on one allele (B6 or Cast) or displayed a biallelic pattern (Fig. 2b). This suggests that the original choice of Vs in each clonal population may be based on a stochastic process. Strikingly, no two clones had the exact same transcription pattern of the V segments.

---

**V region allele-specific histone modification states.** To determine the pattern of V region activation states, we analysed histone acetylation over select V segments in pre-B-cell clones derived from chimeric B6/Cast mice. Since, in general, the sequences of the two alleles differ by about 1% genome wide, we were able to identify many polymorphic sites that could be used to determine the histone acetylation pattern of each allele separately. We first chose a single clone (E9-3) and carried out anti-histone H3Ac ChIP, which was then assayed by PCR analysis of various V segments within the k locus, using polymorphisms at restriction-enzyme binding sites to distinguish between the alleles (Fig. 1a). In a striking manner, it appears that individual Vs are acetylated in a monoallelic manner. This was confirmed by principal component analysis (PCA) (Fig. 3c,d). Interestingly, clone 3 apparently underwent deletion of the Vk locus Cast allele in late passages. Nonetheless, the pattern of V segment expression on the B6 allele remained constant (Supplementary Fig. 4). This suggests that each allele profile is maintained independently, with no crosstalk between the alleles being required.

---

**V region allele-specific chromatin accessibility.** Since expression is always correlated with local histone acetylation on an allelic basis, it appears that, in general, some V segments are relatively open and accessible within the nucleus while others are in a more closed configuration (Fig. 1). To test this idea, we carried out
genome-wide ATAC-seq analysis on each of the different pre-B-cell clones. In this assay, a transposase is used to insert a marker sequence into multiple sites in the genome which are then detected through sequencing analysis, with the degree of integration being proportional to the level of accessibility at each individual locus. By restricting our analysis to sequences located exclusively in the Vκ region, we were able to construct a local allelic ‘openness’ map of specific V segments (Fig. 4). Vκ1–135 for...
example, was shown to be accessible on both the Cast and B6 alleles in a pool of pre-B cells. In clone 4, however, only the Cast allele was open, while the B6 allele showed a relatively small number of sequence reads indicating that the Cast allele is in a more accessible configuration, consistent with the finding that this V is preferentially transcribed on the same allele. In contrast, V2–112 is more open on the B6 allele in this same clone, while V9–124 is accessible on both alleles in equal measure. In clone 8, however, the pattern of chromatin accessibility is distinctly different, with both V9–124 and V2–112 being accessible on the B6 allele, while V1–135 is accessible on both alleles, with a skew towards Cast. Analysis of other pre-B-cell lines indicated that this
allelic pattern is different in each individual clone (Supplementary Fig. 5). It should be noted that while this assay is quite robust, allowing one to identify a large number of accessible Vs, we were not always able to determine their allelic distribution, mainly because the relevant marker sequences were often too small to include nearby polymorphisms (see example V\textsubscript{2–137} (Fig. 4)).

Allele-specific Ig\textsubscript{k} rearrangement repertoires. Having shown at the global level that each clone maintains a different V region accessibility and expression pattern, we then asked whether this may serve as the basis for determining the rearrangement patterns that take place following differentiation to B cells. To this end, we developed a technique to pick up RNA from rearranged molecules by using a C\textsubscript{k} primer together with 3’ ligation-mediated PCR and then quantitatively assayed all detectable rearrangements by large-scale sequencing (Supplementary Fig. 6). This data yielded a clonal recombination map for each allele independently (Fig. 5a,b), and with each individual clone producing a different pattern of rearrangement. Furthermore, by comparing the ratios between the B6 and Cast alleles at each V segment, we found that the degree of rearrangement following differentiation is directly correlated with the amount of ncRNA transcription for each V segment in pre-B-cell clones (Fig. 5c). It thus appears that the choice of V segment ‘opening’ ultimately affects the pattern of rearrangement on each allele.

The data we have presented suggest that while the choice of Vs on each allele is mainly stochastic, some sites appear to have a fixed bias for either B6 or Cast, suggesting that genetic factors may also play a role in this choosing process. To address this question in a general manner, we carried out PCA on the ncRNA spectrum measured in several different pre-B-cell populations (Fig. 3). Strikingly, we found that all the individual clones have ncRNA patterns far removed from the profile observed in a pool of pre-B cells derived from bone marrow. This provides a good indication that genetic background only plays a relatively minor role in V region activation choice. The same idea appears to be true for the rearrangement process, as well, as careful analysis of the results in Fig. 5 demonstrates that less than 35% of the Vs are skewed (>80%) for one allele or the other in the pool.

The data in Fig. 5 demonstrate that following 48 h of induction, each clone appears to have undergone rearrangement on both alleles. To understand how this comes about, we followed the kinetics of this process for one individual clone (clone 4). Previous studies in our laboratory have already shown that the two alleles in pre-B cells are differentially marked by replication
showing that a large number of genomic loci are expressed monoallelic expression profiles. On the basis of previous studies accessibility in a clonal manner and in this way support stable segments are able to maintain an allelic pattern of chromatin transcription on an allelic basis (showed that a large percentage of them are actually located in 

was found to be stable and distinct in each clone, as verified by 

between the two alleles. Strikingly, analysis of all of the accessible taking advantage of multiple polymorphisms to distinguish 

stability, it seemed likely that the basis for allelic differences may 

Global allele-specific transcription. These experiments present an intriguing picture of the Vx region whereby individual gene segments are able to maintain an allelic pattern of chromatin accessibility in a clonal manner and in this way support stable monoallelic expression profiles. On the basis of previous studies showing that a large number of genomic loci are expressed monoallelically in differentiated ES cells22,23, we next asked whether this may also be true for the pre-B cells used in this study. Indeed, global analysis of RNA-seq data from different passages of the pre-B-cell clones used in this study indicates that the allelic transcription pattern is clonally stable and distinct from the other clones (Supplementary Fig. 8a,b). Each clone has between 2,000 and 4,000 genes, which are expressed in a monoallelic manner, consistent with the reports from other cell types22–24. Strikingly, while some genes are monoallelic in all clones analysed, others are monoallelic in only some clones, and biallelically transcribed in others, giving each clone a unique set of monoallelically expressed genes (Supplementary Fig. 8c,d). These genes had extremely diverse functions. Gene ontology analysis on the subset of genes monoallelically transcribed in all clones demonstrated an enrichment of glycoproteins (P = 2.4 × 10^{-7}), as well as proteins involved in signal transduction (P = 9.3 × 10^{-9}). These results suggest that the V-segment allelic choice may actually be part of a much wider phenomenon involving many other regions of the genome.

Global allele-specific chromatin accessibility. Given clonal stability, it seemed likely that the basis for allelic differences may lie in the ability to maintain fixed alternate chromatin structures over many cell generations. To test this idea, we analysed our genome-wide ATAC-seq data from these same B6/Cast clones, taking advantage of multiple polymorphisms to distinguish between the two alleles. Strikingly, analysis of all of the accessible sites in the genome showed that there are indeed thousands of loci that show differential availability in individual clones even though they appear to be accessible on both alleles in a pool of pre-B cells (Fig. 6a). Furthermore, the allelic accessibility profile was found to be stable and distinct in each clone, as verified by hierarchical clustering and PCA (Fig. 6b,c).

Mapping the locations of these monoallelically accessible sites showed that a large percentage of them are actually located in gene promoters (Fig. 6d). Furthermore, RNA analysis indicated that there is a strong correlation between open chromatin and transcription on an allelic basis (χ^2 test P value < 10^{-15} for each clone individually) (Fig. 7f). For example, the promoter of Gng12 is inaccessible and not expressed in clone 3, is accessible and expressed on both alleles in clone 4 and is accessible and transcribed only on the Cast allele in clone 8 (Fig. 7a).

Similar correlations can be seen for other specific genes, such as Htatip2 and Smx20 (Fig. 7a). At the global level, as well (Fig. 7b), allele-specific promoter accessibility accurately predicts the transcription state for each allele in every clone. It thus seems likely that promoter chromatin accessibility may represent an underlying mechanism for inheritable allelic gene transcription, and this has been confirmed by other studies, as well25.

Discussion

Following a programmed series of recombination events at the heavy chain locus, B cells then carry out rearrangement of the light chain region. This is done in an ordered manner with one allele being chosen in pre-B cells to undergo histone acetylation over the J–C region8. It is then this same allele that undergoes rearrangement by recombining with upstream V segments on the same allele15, but little was known about how they are chosen. Indeed, it has always been assumed that this reaction occurs in a stochastic manner with all of the V segments equally available for recombination. In this paper, we have used Cast/B6 hybrid mice and taken advantage of genetic differences to study each allele separately. We demonstrate that in pre-B cells, the V region on both alleles undergoes a process of opening which is reflected in ncRNA transcription, histone acetylation and increased chromatin accessibility, but only a portion of the segments become activated, with the other Vs remaining relatively closed and less capable of recombination. This represents a new level of immune receptor regulation.

Our analysis suggests that in each clone, only about 30–40% of the Vs are actually activated and by examining the distribution of these sites, it appears that this is established through a stochastic process with each allele representing an independent substrate. As a result, in any one clone some Vs are activated on both alleles, some on only one allele, either B6 or Cast and the rest remain closed on both alleles. The distribution we observe is, in general, consistent with a mathematical model in which the Vs on each allele have a relatively fixed probability of activation, but these choices are also influenced by sequence biases built into the B6 and Cast genomes and may be affected by additional epigenetic markers such as replication timing, which differ between the two alleles in each clone.

Many other monoallelic genes throughout the genome also appear to be activated in a similar manner. This stochastic, as opposed to instructive system of choice, is presumably based on the idea that the factors needed to turn on these genes are present in limited quantities which are insufficient for interaction with all the targets present within the nucleus11,26. This type of mechanism is usually associated with single-gene sequences, but several examples of choice within clusters, such as those containing NK receptors27,28 or cytokine genes29 may also be subject to kinetics similar to that of the Vc locus.

The most striking aspect of allelic choice seen in the Vc region is that these patterns are preserved in a clonal manner. This implies that the initial factor-directed activation of V segments probably takes place at an earlier stage just prior to the formation of pre-B clones, which are then no longer capable of choosing but continue to maintain the initial decision of allele specificity through each cell division. It is likely that this is carried out through the formation of some sort of chromatin structure, perhaps by means of autonomously maintainable histone modifications30–33. Consistent with this idea, we have demonstrated that individual V segments are either ‘open’ or ‘closed’ as determined by their histone modification pattern and degree of accessibility (ATAC).

By examining pre-B cells from interspecific B6/Cast hybrid mice we have discovered a new level of regulation within the κ light chain locus that serves to set up this region in a manner that directs the subsequent pattern of recombination. This mechanism may actually play a role in ensuring that B cells maximize their ability to produce useful antibodies. Prior to rearrangement, a subset of V segments are opened on both alleles. Initially, only the early replicating allele activates the process of
recombination, which can then be carried out with any of the Vs that have already been opened. If this results in a productive antibody, the pre-B cell has completed its task and no longer attempts further recombination. In the event that no productive antibody is generated, this same allele may undergo editing by switching to a more distal available V and to a different J segment. If this also fails, the same B cell attempts to carry out recombination on the second late-replicating allele (Supplementary Fig. 7)\textsuperscript{34}. Because of the stochastic nature built into the mechanism for opening V segments, this second allele only has a limited choice of V segment partners and most of these are different than those that were activated on the first-chosen early allele. For this reason, the second allele only has a small probability of recombining with the same Vs that have already proven to be incompatible, and will mostly partner with a different set of Vs, thereby increasing its chances of generating a productive antibody. Thus, the way in which V segments are initially opened contributes to the robustness of the B cell’s immune potential.

A number of different studies have documented the existence of many genes that are expressed monoallelically in a variety of different cell types, with a pattern consistent with the idea that they are activated in a stochastic manner by suboptimal concentrations of key factors\textsuperscript{22,23,24,35}. We have added a new dimension to this picture by showing that these same genes have a differential chromatin structure, with one allele being more accessible than the other. It is likely that monoallelic expression represents a readout of this stable chromatin structure, perhaps explaining how these genes retain their monoallelic pattern in a clonal manner. As in the immune system, this mechanism appears to be geared at optimizing diversity to allow the expression of different alleles in each cell. The identity and functionality of the monoallelically-regulated genes vary greatly in different clones, and this probably contributes to the distinct cellular responsiveness of each cell. This may allow the immune system to effectively and robustly contend with the myriad of challenges that the organism is bound to encounter over its lifespan.

**Methods**

**Animals and cells.** C57BL/6 (B6) mice (Harlan) were crossed with wild-type M. castaneous (Cast) mice (Jackson Laboratories) to generate B6/Cast\textsuperscript{+} hybrid mice. Mice were housed and cared for under specific pathogen-free conditions, and all animal procedures were approved by the Animal Care and Use Committee of the Hebrew University of Jerusalem.

Cells isolated from the bone marrow of F1 female B6/Cast mice (8–12 weeks) were grown in RPMI 1640 media (Gibco) supplemented with 10% fetal bovine serum (HyClone), penicillin–streptomycin (Gibco), l-glutamine (Gibco) and 50 μM of β-Mercaptoethanol (Gibco) on irradiated ST2 feeder cells. IL-7 was added to
conditioned medium collected from J558-IL7 secreting cells (as provided by A. Rolink) to select for pre-B cell populations. After 10–14 days of IL-7-mediated positive selection, cells were plated on 96-well plates in limiting dilutions to generate single-cell-derived pre-B-cell clones. Igκ locus rearrangement was induced by removal of IL-7 from the culture media for 48 h.

Chromatin immunoprecipitation. Chromatin immunoprecipitation was performed as described previously36. Briefly, cells were fixated in formaldehyde and resuspended in RIPA buffer and the resulting chromatin was sonicated with a water bath sonicator to sizes ranging from 300 to 800 bp, incubated overnight with the specified antibody at 4°C (H3Ac Millipore 06-599, H3K27me3 Millipore 07-449, H3K4me1abcam ab8895, H3K4me2 abcam ab3254) and then incubated for 3 h with protein A agarose beads (Millipore). The beads were then washed repeatedly with RIPA buffer supplemented with increasing levels of NaCl and 3 h with protein A agarose beads (Millipore). The beads were then washed immediately downstream of the second exon (RSS sequence), giving rise to an expected PCR product ~400 bp long. The primers had the following sequences (Supplementary Table 1): Fo 5'-GAGTTCTCAGTCGGACGCATC-3', R 5'-CCCTGGCAACCAGGAAATTCCA-3', corresponding to the Illumina small RNA adapters. Following PCR amplification, excess primer was removed using Ampure XT bead size selection (Beckman-Coulter). The resulting fragments were subjected to a second round of PCR, adding full length Illumina small RNA adapters. (F 5'-AATGTACAGGCGACACCGAGATCTACAGTCCGACGATC-3', R 5'-CAAGCAGAAGACGGCATACGAGAT-[NNNNNN]-GTGACTGGAGTTCCTTGGCACCCGAGAATTCA-3') with the N nucleotides signifying a six base index sequence used to differentially mark samples for demultiplexing. After the addition of the adapters, PCR fragments of the correct size (~500 bp) that do not contain the Vκ segment families were used to amplify 20 different Vκ segments (Supplementary Table 2). The forward primer was located within the first exon of the Vκ segment (leader sequence), while the reverse primer targeted the area immediately downstream of the second exon (RSS sequence), giving rise to an expected PCR product ~400 bp long. The primers had the following sequences:

**Allelic non-coding RNA analysis.** RNA was extracted from IL-7 dependent pre-B cells, treated with DNase (Ambion) for 1 h to remove traces of genomic DNA and cDNA then prepared with the qScript RT kit (Quanta), with Vκ segments being amplified using specific primers spanning the RSS sequence to ensure it had not undergone rearrangement. PCR products were cut with allele-specific restriction enzymes, and visualized on 8% polyacrylamide TBE gels. Allelic ratios were computed based on band strength, with genomic DNA being used as a biallelic control.

For amplicon sequencing, 10 semi-degenerate primer pairs specific for a number of Vκ segment families were used to amplify 20 different Vκ segments (Supplementary Table 2). The forward primer was located within the first exon of the Vκ segment (leader sequence), while the reverse primer targeted the area immediately downstream of the second exon (RSS sequence), giving rise to an expected PCR product ~400 bp long. The primers had the following sequences:

- **Gng12 (a pro-inflammatory G protein γ subunit), Htatip2 (a pro-apoptotic protein) and Snx20 (an endosome membrane trafficking protein)** in pre-B-cell clones. Location of the gene promoter is marked in grey. Total reads, regardless of allelic origin are marked in green, reads identified as originating from the Cast allele based on SNPs are marked in red, reads identified as originating from the B6 allele based on SNPs are marked in blue, and reads identified as originating from different pre-B-cell clones. The values are the averages of biological duplicates. Red to blue heatmap indicates linear Cast to B6 levels. Red to blue heatmap indicates linear Cast to B6 levels on a scale from 0 (100% Cast) to 1 (100% B6).
and allelic ratios were calculated for Vβ fragments with minimum depth of 20 reads. Read depth varied from 0 to 45,000 reads for any given Vβ segment.

Stranded nuclear RNA-seq was carried out as follows. Sixteen million cells (cultured pre-B-cell clones: clones 4 and 8 in biological replicates) were washed with PBS, resuspended in cold RLN-Igepal (30 mM Tris pH 7.5, 140 mM NaCl, 1.5 mM MgCl2, 1 mM DTT, 0.4% v/v Igepal CA-630) for 5 min at a ratio of 4 million cells per ml and washed with cold RLN (30 mM Tris pH 7.5, 140 mM NaCl, 1.5 mM MgCl2, 1 mM DTT) to obtain purified nuclei. RNA was then extracted (4 million per ml and washed with cold RLN (50 mM Tris pH 7.5, 140 mM NaCl, 1.5 mM MgCl2, 0.1 M acetic acid) and the single strand cDNA then purified using Silane beads (Life Technologies). A 3’TR3 adapter (5’-GGACTGAGCACAGCATG-CTG-TGGATG-3’) was ligated to the 3’ end by overnight incubation with T4 DNA ligase at 22 °C, and the cDNA then purified from excess adapter with Silane beads and PCR amplified for 12 cycles using the reverse complement of the 3’TR3 adapter as the forward primer and the upstream Cx as the reverse primer with the partial TruSeq Illumina adapter added to the beginning (5’-TACACGAGCGCTCTTCTCCGATCT-ACTGGATGGTGGGAAGATTAGATGTTG-3’). The PCR product was cleaned with 0.7 x ampure XT beads, amplified with indexed universal Illumina adapter primers for an additional seven cycles to obtain ~550 bp libraries, which were sequenced (Miseq, 2 x 200 x 150 bp paired end).

The resulting sequences were quality trimmed and aligned to a hybrid B6/Cast genome assembly using bowtie2 (ref. 37), with Cast polymorphic sites substituted based on the Sanger Mouse Genome Project database (release 1505). Reads over each Vκ segment were counted (HTseq-count) and normalized to the total mapped rearranged fragment number to allow comparison of the Vκ repertoire contribution between different libraries.

**Data availability**

Sequence data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) with the primary accession code GSE97148. Other data that support the findings of this study are available from the corresponding author upon request.

**References**


Acknowledgements

We thank Ofer Hakon and members of his group for assistance in setting up the ATAC-seq assay, Hagit Masika and Marganit Farago-Tirosh for helping in revising the manuscript and all the members of the Bergman group for helpful discussions. This research was supported by grants from the European Research Council (grant no. 268614 to H.C.), the Israel Cancer Research Fund (H.C. and Y.B.), the US-Israel Binational Foundation (Y.B., grant no. 2015506 to S.K.), the Rosetrees Foundation (H.C.), the Biological and Biotechnological Scientific Research Council (A.E.C. and L.M.), the I-CORE Program of ISF (Y.B. and H.C.), The Israel-China Binational ISF-NSFC (Y.B.) and the Emanuel Rubin Chair in Medical Science (Y.B.).

Author contributions


Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Levin-Klein, R. et al. Clonally stable Vk allelic choice instructs Igk repertoire. Nat. Commun. 8, 15575 doi: 10.1038/ncomms15575 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Additional information

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/