Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2022 May 27.

Published in final edited form as:

Cell Rep. 2022 May 03; 39(5): 110769. doi:10.1016/j.celrep.2022.110769.

Topologically associating domains are disrupted by evolutionary genome rearrangements forming species-specific enhancer connections in mice and humans

Sarah E. Gilbertson[1], **Hannah C. Walter**[1], **Katherine Gardner**[1], **Spencer N. Wren**[1], **Golnaz Vahedi**[2], **Amy S. Weinmann**[1,3,*]

[1]Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[2]Department of Genetics, Institute of Immunology, Epigenetics Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA

[3]Lead contact

SUMMARY

Distinguishing between conserved and divergent regulatory mechanisms is essential for translating preclinical research from mice to humans, yet there is a lack of information about how evolutionary genome rearrangements affect the regulation of the immune response, a rapidly evolving system. The current model is topologically associating domains (TADs) are conserved between species, buffering evolutionary rearrangements and conserving long-range interactions within a TAD. However, we find that TADs frequently span evolutionary translocation and inversion breakpoints near genes with species-specific expression in immune cells, creating unique enhancer-promoter interactions exclusive to the mouse or human genomes. This includes TADs encompassing immune-related transcription factors, cytokines, and receptors. For example, we uncover an evolutionary rearrangement that created a shared LPS-inducible regulatory module between *OASL* and *P2RX7* in human macrophages that is absent in mice. Therefore, evolutionary genome rearrangements disrupt TAD boundaries, enabling sequence-conserved enhancer elements from divergent genomic locations between species to create unique regulatory modules.

In brief

It is currently unclear how evolutionary genome rearrangements affecting the mouse and human genomes influence the expression of genes important in immunity. Gilbertson et al. report that evolutionary genome rearrangements disrupt topologically associating domain boundaries,

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*Correspondence: weinmann@uab.edu.

AUTHOR CONTRIBUTIONS

A.S.W. conceived of the project. A.S.W., S.E.G., and H.C.W. designed the study and performed analyses. K.G. and S.N.W. performed analyses. G.V. performed analysis to define TADs for DP T cells. A.S.W. and S.E.G. wrote the manuscript. All authors contributed to editing the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

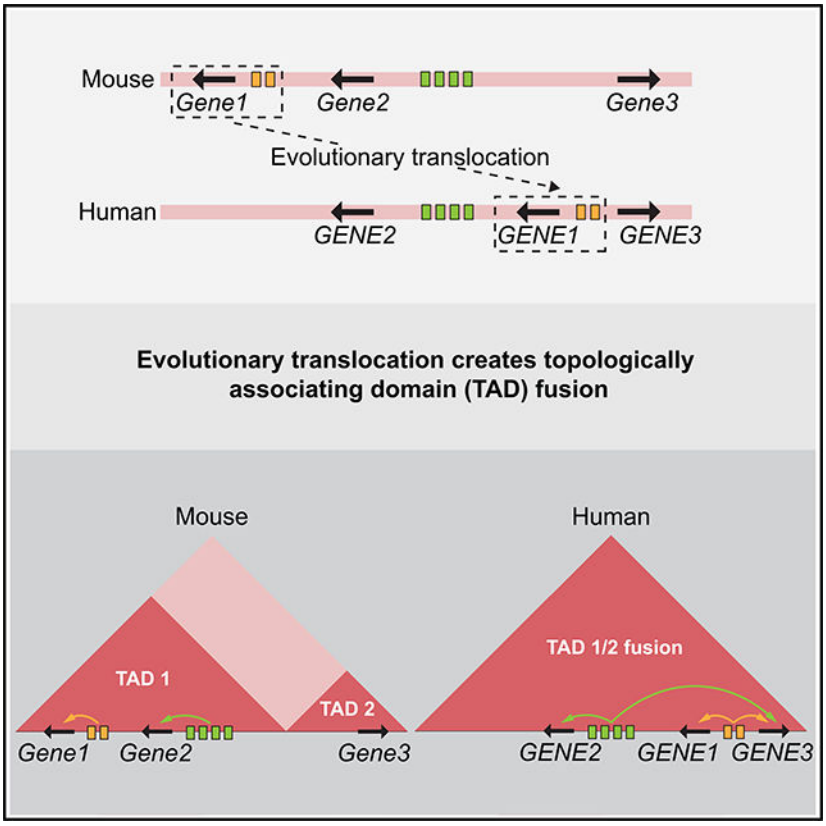
SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.110769.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

enabling sequence-conserved enhancer elements from divergent locations between species to create unique regulatory modules.

Graphical Abstract



INTRODUCTION

Mouse models have been invaluable tools to define mechanisms for immune cell differentiation and function. With the expansion of genomics technologies, many studies now incorporate comparisons between mouse and human immune cells into the experimental design (Choi et al., 2019; Moreira-Teixeira et al., 2020; Philip et al., 2017; Zilionis et al., 2019). Typically, these comparisons predominantly focus on the conserved aspects of gene programming potential between the two species. However, close inspection of these datasets has revealed a substantial fraction of non-conserved events as well, which is consistent with species-specific activities contributing to differences in immune responses between mice and humans (Schroder et al., 2012; Shay et al., 2013). Identifying the mechanistic principles that contribute to divergent activities between species has received less attention, even though this information has the potential to improve the utility of mice as a preclinical model of human immunity.

Genomic structural variation contributes to differences in gene expression potential between individuals and dysregulated expression in disease states (Audano et al., 2019; Spielmann

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

et al., 2018). Structural variation in the genome refers to chromosomal rearrangements such as translocations, inversions, amplifications, and insertions/deletions. Translocations and inversions are balanced genome rearrangements that alter the location and/or orientation of chromosomal content. In contrast, amplifications and insertions/deletions are unbalanced due to the addition or removal of genomic content. In disease states, structural variation has been shown to expose genes to a different composition of regulatory elements, causing aberrant gene expression patterns (Kloetgen et al., 2020; Lupianez et al., 2015; Spielmann et al., 2018; Wang et al., 2021).

Genomic structural variation between different species is referred to as evolutionary structural variation and is thought to contribute to acquiring (or losing) gene expression potential between species (Merot et al., 2020). Genomics studies performed in mice have defined gene-regulatory networks involved in the differentiation and activation of immune cells, with the goal of applying this information to humans. However, a significant barrier to accomplishing this goal is that we do not currently understand the functional impact of rearranged genetic content resulting from evolutionary structural variation. Shifting the regulatory landscape available to a given gene has the potential to cause cell-type and inducible gene expression differences between mice and humans.

Long-range enhancer-promoter interactions contribute to cell-type and inducible transcription, and these interactions can span multiple intervening genes (Zheng and Xie, 2019). Topologically associating domains (TADs) confine genomic interactions, with long-range interactions occurring more frequently inside of a TAD rather than between adjacent TADs (Rowley and Corces, 2018). This means that disrupting TAD boundaries has the potential to alter the gene-regulatory landscape. The current view in the field is that TADs are conserved between species and that the regulatory landscape available to genes in different species is largely maintained (Dixon et al., 2012; Krefting et al., 2018; Rao et al., 2014; Vietri Rudan et al., 2015). This leads to the assumption that evolutionary structural variation is somehow prevented from disrupting TAD boundaries, suggesting that the breakpoints for rearrangements occur at or near TAD boundaries. However, whether TAD boundaries are affected by evolutionary structural variation has not been thoroughly interrogated. This is especially true regarding rearrangements occurring in proximity to immune response genes, a class of genes that tends to evolve more rapidly between species (Sironi et al., 2015). Furthermore, close examination of the data underlying the view that TADs are conserved between species reveals that oftentimes the interpretation is more complicated than a simple conservation paradigm, as was expertly discussed in Eres and Gilad (2021). It is therefore currently unclear whether evolutionary structural variation can affect TAD composition similarly to how genomic rearrangements affect TADs in human diseases (Franke et al., 2016; Lupianez et al., 2015).

In this study, our goal was to determine whether evolutionary structural variation alters the genomic composition of TADs and disrupts underlying enhancer-promoter interactions for immune-related genes. We found that genes with differences in cell-type- and stimuli-dependent expression between mouse and human immune cells were often located in proximity to evolutionary structural variation, with translocations representing the most common type. We also revealed that evolutionary translocations and inversions frequently

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

altered TAD genomic content. Specifically, the locations of otherwise sequence-conserved enhancer elements diverged between the mouse and human genomes to the extent that this created new, long-range enhancer-promoter interactions. One example of this was an evolutionary inversion enabling the *P2RX7* locus to acquire inducible genomic connections with the *OASL* locus in the human genome. This coincided with the inducible co-regulation of *OASL* and *P2RX7* in human monocyte-derived macrophages, while *Oasl1*, but not *P2rx7*, was inducible in mouse macrophages. We also identified differences in genomic configuration and regulation of the *NOS2* locus (encodes inducible nitric oxide synthase [iNOS]) between the mouse and human genomes. Together, these findings provide insight into how evolutionary genome rearrangements affect regulatory mechanisms and contribute to species-specific differences in gene programming potential for immune cells in healthy and disease states.

RESULTS

Differentially expressed genes between mouse and human immune cells are found in proximity to evolutionary structural variation

We first wanted to determine whether genes with cell-type- or stimuli-dependent expression differences between mouse and human immune cells were found in proximity to evolutionary structural variation. To accomplish this, we first examined published datasets comparing mouse and human immune cells to identify curated lists of differentially expressed genes between species (Schroder et al., 2012; Shay et al., 2013). In total, there were over 400 genes from three sources: 169 genes with cell-type-specific differences in expression in resting immune cells (Shay et al., 2013), 134 genes with differences in LPS induction potential in macrophages (Schroder et al., 2012), and 175 genes with differences in activation potential in $CD4^+$ T cells (Shay et al., 2013). We confirmed the species-specific expression patterns for a subset of genes by using independent, publicly available single-cell or bulk RNA-seq datasets (Figures 1, and S1) (Heng et al., 2008; Manago et al., 2019; Tong et al., 2016). For example, there was enhanced expression of *JUN* in human B and T cells compared with mouse cells (Figures 1A and 1B). In contrast, *GIMAP6*, which encodes GTPase of immunity-associated protein 6, was expressed in T cells in humans, whereas in mice it was expressed in both B and T cells (Figures 1A and 1B). In addition, *IDO1* and *CXCL13* were induced by LPS in human macrophages, but these genes lacked LPS induction potential in mouse macrophages (Figures 1C–1E), whereas *SLC7A2* (encodes cationic amino acid transporter 2) displayed the opposite species-specific expression profile (Figure 1D). Consistent with the LPS-inducible expression profiles for these genes, there were species-specific LPS-inducible H3K27Ac peaks in proximity to the loci (Figures 1C–1E). It was interesting to note the clear difference in genomic configuration surrounding the human *CXCL13* locus compared with mouse *Cxcl13*, as well as the human *IDO1* locus compared with mouse *Ido1* (Figures 1C–1E). This indicates that these loci are in proximity to evolutionary structural variation.

Extending these findings, we interrogated the genomic configuration surrounding the full list of genes with species-specific expression differences to define how frequently these genes were located in proximity to evolutionary structural variation between the mouse and human

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

genomes. We discovered that approximately 41% of genes with cell-type-specific expression profiles were found in proximity to evolutionary structural variation between the mouse and human genomes (Figure 2A). Furthermore, we found that 55% and 60% of genes with LPS- or T cell receptor (TCR)-inducible specific expression were also in proximity, respectively (Figure 2A). Notably, translocations represented the most abundant form of evolutionary structural variation in proximity to genes with cell-type- or stimuli-dependent differential expression between mouse and human immune cells (Table S1).

Translocations are balanced genome rearrangements that cause changes to the location and/or orientation of genomic content while not actually altering the content itself. This means that the genes in proximity to evolutionary translocation breakpoints are the ones most likely to be exposed to a new regulatory landscape. Therefore, we wanted to determine how many genes separated the differentially expressed genes from the translocation breakpoint. Approximately 72%–87% of the genes with cell-type- or stimuli-dependent differences in expression between mouse and human immune cells were within four genes of the rearrangement breakpoint (Figure S2F). In contrast, only 7% of randomly selected genes were within this distance. This suggests that the differentially expressed genes may be exposed to a new composition of regulatory elements because they are located in proximity to the boundaries of evolutionary genome rearrangements.

Genes important for immune responses to infectious diseases are near breakpoints for evolutionary genome rearrangements

We next wanted to determine whether other genes involved in immune responses to infectious diseases were located near evolutionary translocation and inversion breakpoints. To address this, we examined the genomic configuration surrounding genes with roles in immunity in the mouse and human genomes by using the UCSC genome browser. In this analysis, we identified over 30 genes located in proximity to breakpoints for evolutionary rearrangements between the mouse and human genomes (e.g., Figures 2C, S2, and Table S2). This included transcription factors such as *EOMES*, *IRF4*, *FOXO1*, and *XBP1*, receptors such as *PDCD1* (encodes PD-1), *HAVCR2* (encodes Tim-3), *IL3RA*, *CSF2RA*, *CRLF2*, *IL9R*, *IL23R*, and *IL12RB2*, and cytokines such as *IL6*, *TSLP*, and the *IL1* cluster. Figures 2C and 2D show the evolutionary genome rearrangements surrounding the *IL6* and *GIMAP6* loci, with the coloring for the chromosomal content using the mouse genome as the reference. These data illustrate that the genetic content on both the 5′ and 3′ sides of human *IL6* diverged from the mouse genome because of evolutionary rearrangements, whereas there was an evolutionary translocation to the 3′ side of the *GIMAP6* locus. As another example, *IL3RA*, *CSF2RA* (encodes GM-CSF receptor), and *CRLF2* (encodes TSLP receptor) were clustered together on the pseudoautosomal region of the X and Y chromosomes in the human genome, whereas these genes were located on three different autosomal chromosomes in mice (Figure S2B). Together, the data indicate that many genes required for immune responses to infectious diseases are located near evolutionary rearrangement breakpoints, causing changes to the genetic content in their proximity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The human genome contains gene clusters with roles in the immune response

Primates and rodents diverged from a common ancestral tree over 80 million years ago. This means the genome rearrangements identified in our analysis might have occurred in either branch after the split. Therefore, we next wanted to define the species relationship for the evolutionary rearrangements to determine whether the mouse or human genome is more closely related to other species in the ancestral tree. To address this, we first examined the available genome builds for a variety of species to define the configuration of gene loci in proximity to the genome rearrangements. We compared these configurations with the patterns found in the mouse and human genomes. From an evolutionary perspective, we found that the genomic configuration for the *IL6* locus in human was similar to that of several species including hedgehog, dog, chicken, and primates. In contrast, the mouse configuration was partially similar only to closely related rodents. This suggests that the mouse genomic configuration diverged from the ancestral tree. Clusters of immune regulator genes in the human genome often represented the common configuration, including the *IL1* cluster as well as the cytokine receptor cluster on the pseudoautosomal region of the X and Y chromosomes. In contrast, the genes in these clusters were separated in the mouse genome. This indicates that evolutionary genome rearrangements in rodents disrupted these cytokine and cytokine receptor clusters, scattering the genes into distinct locations in the mouse genome.

Previous studies have defined evolutionary breakpoints between blocks of synteny in genomes, along with the species relationships for the genome rearrangements associated with the breakpoints (Bourque et al., 2004; Larkin et al., 2009). This infers the evolutionary history of rearrangements and the branch in the evolutionary tree they occurred. In the context of the rodent-primate split, rodents have accumulated at least two times more rearrangement breakpoints than primates (Bourque et al., 2004; Larkin et al., 2009). Therefore, we next performed the reciprocal analysis to Figure 2A in order to reveal whether the rodent- or primate-lineage breakpoints described in Larkin et al., (2009) were found in proximity to genes with differences in expression between species. Approximately 50% of the rodent-specific breakpoints and 39% of mouse-specific breakpoints were in proximity to one of the genes with known expression differences between mouse and human immune cells (Figure 2B). In contrast, about 25% and 26% of the primate- or hominoidae-specific breakpoints, respectively, had a known differentially expressed gene in proximity. This suggests that rodent- and mouse-specific genome rearrangements represent an important source of genetic diversity that accumulated after the evolutionary split with primates.

The proximity of differentially expressed genes to rearrangement breakpoints raised the question of whether genes with similar functions clustered together in either the mouse or human genomes, as we had observed with the cytokine clusters in humans, which might suggest opportunities for shared regulation in one species. To address this, we identified the four closest genes on the 5′ and 3′ sides of the subset of genes with roles in the immune response to infectious diseases that we found in proximity to evolutionary translocations and inversions (Table S2). We then performed separate GSEA pathway analyses on the gene sets from the mouse and human genomes. The analysis from the human genome preferentially enriched for pathways associated with immune responses compared with the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

pathway analysis in mice (Figure S2A). This included enrichment of both the FDR q value and absolute number of genes in pathways such as the defense response and immune system processes. Collectively, the data indicate that genes involved in the immune response cluster together more frequently in the human genome compared with the mouse genome for this subset of immune regulator genes.

Evolutionary translocations between the mouse and human genomes affect TAD genomic composition

The current view in the field is that TADs are conserved between species, even in the context of evolutionary structural variation (Dixon et al., 2012; Vietri Rudan et al., 2015). TADs could be preserved in this context if genomic rearrangements occurred only at TAD boundaries, which would constrain the effects of evolutionary structural variation. Long-range genomic interactions preferentially occur within a TAD, meaning that maintenance of TAD boundaries would limit the acquisition (or loss) of enhancer-promoter interactions between species. However, the underlying data used to support this model suggest that this interpretation is oversimplified, making it necessary to revisit this question (Eres and Gilad, 2021). Therefore, we next wanted to determine whether the TADs containing genes with differential expression between mouse and human immune cells included genomic content spanning evolutionary translocation breakpoints. We found numerous cases, including surrounding the *IL6*, *GIMAP6*, and *TLR9* loci, in which TAD genomic content spanned the breakpoints for evolutionary genome rearrangements, with many TADs incorporating a new composition of genes between mouse and human cells (Figures 2E, 2F, and S2). In fact, for the evolutionary rearrangements in proximity to differentially expressed immune genes, approximately 71% and 68% of TADs spanned the translocation breakpoints in the mouse or human genomes, respectively (Figure 2G). This means that these TADs contained rearranged genomic content between mouse and human cells. This number increased to approximately 84% when considering whether the TADs encompassing these regions differed in genomic content in at least one species (Figure 2H). Notably, the TADs spanning evolutionary translocation breakpoints contained the differentially expressed immune gene in the vast majority of cases (Figure 2H).

A limitation for these types of analyses is the computational programs that designate TAD boundaries are constantly being refined and can affect the assignment of boundaries (Dali and Blanchette, 2017; Vahedi, 2021). Another consideration is TADs might differ between cell types, in contrast to the current model hypothesizing that TADs are consistent in diverse cell populations (Zheng and Xie, 2019). To explore whether these variables affected our interpretations, we assessed whether TADs called with the same program in a variety of human immune cell types showed evidence of differences in genomic content between mouse and human cells (Figures S2G and S2H, and data not shown). We also examined whether TADs called by independent laboratories with different programs had similar differences in TAD content between the mouse and human genomes (Figures 2G and 2H). Consistent with previous results, these technical and biological variables did cause modest differences in the location of TAD boundaries. However, in both scenarios, the majority of TADs spanned the evolutionary breakpoints, creating differences in genomic content between mouse and human cells (Figures 2G and 2H compared with S2G, S2H). This means

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that TAD genomic content frequently varies between mice and humans in proximity to the evolutionary translocation breakpoints adjacent to differentially expressed immune genes.

An evolutionary inversion in primates created a TAD fusion containing *OASL* and *P2RX7*

Differences in TAD genomic content between species might create new enhancer-promoter interactions that alter the cell-type or inducible expression pattern of genes. In support of this, we uncovered the acquisition of a shared inducible regulatory module between *OASL* and *P2RX7* in human macrophages that did not occur in mice. *OASL* encodes the 2′–5′-oligoadenlyate synthetase-like protein that binds double-stranded RNA and has antiviral activity (Zhu et al., 2014), and *P2RX7* encodes a purinergic receptor for ATP that plays a role in inflammatory responses (Di Virgilio et al., 2017). Both *OASL* and *P2RX7* expression were highly induced by LPS in human macrophages (Figure 3A). In contrast, *Oasl1* and *P2rx7* were not co-regulated in mouse macrophages, with LPS inducing only *Oasl1* expression (Figure 3B).

We first explored the genomic configuration surrounding *OASL* and *P2RX7* in the mouse and human genomes to define mechanisms that might account for the differences in LPS-induction potential between species. In the mouse genome, approximately 8 Mb separated the *Oasl1* and *P2rx7* loci on mouse chromosome 5, and consistent with this separation, *Oasl1* and *P2rx7* were contained within two distinct TADs (Figure 3C, top). In humans, *OASL* and *P2RX7* were found adjacent to each other on chromosome 12, with both genes contained in the same TAD (Figure 3C, bottom). These data indicate that a genome rearrangement affecting the human and mouse *P2RX7* and *OASL* orthologs occurred after the rodent-primate split. In species prior to primates, there were two separate *Oasl* genes (*Oasl1* and *Oasl2*), and these genes were not located in proximity to *P2rx7*. As the genome configuration evolved, the *Oasl1*/*Oasl2* gene neighbors frequently changed, whereas *P2rx7* typically had the same neighboring genes prior to primates. This suggests that the *Oasl1*/*Oasl2* loci were in an unstable region, rearranging at several different times in evolution. After the early primate bushbaby, an evolutionary inversion brought *OASL* and *P2RX7* in proximity (Figure 3C, middle), with the breakpoint disrupting *Oasl2* to leave a single intact *OASL* gene. Consistent with our analysis, *P2RX7* is located at a primate-specific breakpoint (Larkin et al., 2009). Together, the data suggest that an evolutionary inversion in primates brought *OASL* and *P2RX7* into proximity and merged TAD content between species, potentially contributing to *P2RX7* acquiring a shared regulatory module with *OASL*.

Both sequence and location dictate enhancer conservation across species

By definition, the location of an enhancer does not strictly limit its activity (i.e., enhancers are able to function at different distances and orientations relative to the gene(s) they regulate) (Long et al., 2016). Rather, TAD boundaries confine the potential enhancer search space to a specific subset of genes. This implies that sequence-conserved enhancer elements will maintain the potential to regulate a specific suite of genes in different species as long as TAD boundaries are conserved. The pervasive assumption that TADs are conserved between species has meant that researchers typically do not consider enhancer location as an important variable when translating mechanisms between species. However, our data challenging this assumption suggest that information about both enhancer sequence and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

location conservation are needed to predict functional potentials for enhancers between species.

To explore this concept, we examined the enhancer landscape available to *P2RX7* and *OASL* in the mouse and human genomes, using elements originating from mouse immune cells as the reference. We first identified the open chromatin regions (OCRs) contained within the two distinct TADs encompassing *Oasl1* and *P2rx7* from a variety of mouse immune cell types (Figure 3C). This defined the potential enhancer landscape for *Oasl1* (Figure 3C top, blue elements) compared with *P2rx7* (Figure 3C top, green elements) in mice. We next performed a liftOver analysis to determine whether the sequences for the OCRs were conserved in the human genome, with this analysis also providing the location of the sequence-conserved elements. Approximately 68% and 72% of the OCRs from the *Oasl1* and *P2rx7* TADs, respectively, were sequence conserved at a 0.1 threshold in the human genome. Notably, 47% of the sequence-conserved OCRs from the *Oasl1* TAD in mice relocated into the novel human TAD containing *OASL* and *P2RX7* (Figure 3C bottom, blue elements), while 49% of the sequence-conserved OCRs from the mouse *P2rx7* TAD maintained their location in this TAD (Figure 3C bottom, green elements). These data demonstrate that disrupting TAD boundaries at the breakpoints of evolutionary genome rearrangements creates hybrid TADs composed of sequence-conserved elements originating from two separate regions of the genome in the other species.

These findings led us to ask how generally sequence and location conservation contribute to species differences in TAD composition at evolutionary genome rearrangement breakpoints. To address this, we performed a similar analysis to the one described for *OASL*/*P2RX7* to assess the mouse TADs in proximity to the breakpoints from Table S1. For this group of TADs, approximately 67% of OCRs had sequence conservation at a 0.1 threshold between the mouse and human genomes (Figure S3A). We next determined whether the sequence-conserved elements were found at a single location in the human genome or rather were distributed between multiple distinct locations, as observed for the *OASL*/*P2RX7* breakpoint. For reference, we designated the location in the human genome with the most sequence-conserved elements as the “majority location,” and then defined the elements that diverged from this location. Approximately 66% of these TADs from the mouse genome had sequence-conserved elements that diverged from the majority location in the human genome (Figure S3B), which is consistent with the percentage of TADs spanning evolutionary breakpoints (Figure 2G). The percentage of sequence-conserved elements diverging from the majority location ranged from 1% to 62% (Figure S3B), with the percentage dependent on the proportional total distance from the translocation breakpoint to each TAD boundary. There are complexities for interpreting the location conservation analysis (see Discussion), but the data demonstrate the importance of accounting for both sequence and location conservation when translating regulatory events between species in rearranged regions of the genome.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The LPS-inducible co-regulation of *P2RX7* and *OASL* in humans was acquired through an evolutionary inversion

We next wanted to define whether the acquisition of the LPS-inducible shared regulatory module for *P2RX7* and *OASL* in humans might have been caused by the relocation of sequence-conserved elements from mice into the hybrid TAD in humans. To address this, we examined the long-range genomic interactions for *P2RX7* in human monocyte-derived macrophages either mock infected or infected with H5N1 (Heinz et al., 2018). Notably, the *P2RX7* promoter formed long-range interactions with the *OASL* locus and an upstream element in H5N1-infected human monocyte-derived macrophages, but not in mock-infected cells (Figure 3D). We performed a liftOver analysis to determine if the inducible upstream H3K27Ac enhancer element in LPS-stimulated human macrophages was present in the mouse genome, and, as predicted, the sequence for this element diverged in location between mice and humans because it was derived from the 5′ region of the *Oasl2* gene found in mice. Collectively, the data suggest that the inducible co-regulation of *OASL* and *P2RX7* in humans was acquired through an evolutionary inversion that brought the *OASL* and *P2RX7* loci into proximity, causing the TAD to merge content from the two sides of the inversion breakpoint. Rearranged genomic content within this TAD created a human-specific *OASL/P2RX7*-regulatory module that does not exist in mice.

The P2RX7 receptor is induced in tissue-resident memory T (Trm) cells in mice, but the role for it in human Trm cells is less clear (Borges da Silva et al., 2020; Kumar et al., 2017). Therefore, we wanted to address whether *P2RX7* is expressed similarly in human as it is in mouse Trm cells. To accomplish this, we examined human *P2RX7* and mouse *P2rx7* expression in published RNA-seq datasets from Trm cells (Kumar et al., 2017; Mackay et al., 2016). *P2rx7* was highly induced in mouse Trm cells compared with central memory T (Tcm) cells and effector memory T (Tem) cells (Figure S3C and data not shown). In contrast, *P2RX7* was not induced in human Trm cells compared with circulating memory T cells (Figure S3C). Consistent with the species differences in expression patterns, a high proportion of the long-range genomic interactions in human $CD4^+$ and $CD8^+$ T cells spanned the evolutionary rearrangement breakpoint, indicating that these interactions were specific to the human genome (Figure S3D). Collectively, these data indicate that the evolutionary inversion in proximity to the *P2RX7* locus affects the regulatory events controlling its expression in mouse and human immune cells.

The human *NOS2* locus lacks TLR-inducible regulatory mechanisms found in mice

Nos2 encodes iNOS, which has been the focus of many studies of infectious diseases and cancer in mouse models (Bogdan, 2015; Eyler et al., 2011; Formaglio et al., 2021; Goldberg et al., 2018). *Nos2* expression is highly induced by Toll-like receptor (TLR) ligands and IFNγ signaling in mouse macrophages and DCs, and *Nos2*-deficient mice have been used to define roles for iNOS in the clearance of intracellular bacteria (Bogdan, 2015; Mattner et al., 2004; Miller et al., 2004). In contrast, studies from independent groups have shown that *NOS2* is not induced by TLR ligands in human macrophages (Gross et al., 2014; Young et al., 2018). To confirm this finding, we reanalyzed several independent published RNA-seq datasets from unstimulated or TLR ligand-simulated mouse and human macrophages (Link et al., 2018; Manago et al., 2019; Moreira-Teixeira et al., 2020; Park et al., 2017; Tong et al.,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2016). These data confirmed that *Nos2* was highly induced by TLR ligands in mouse, but not human, macrophages (Figures 4A–4D).

To explore the divergence in *Nos2* induction, we examined whether there were any differences in regulatory events at the locus between the mouse and human genomes. First, we identified evolutionary structural variation in proximity to *NOS2*. We uncovered a 75-kb insertion directly adjacent to the 3′ end of the *NOS2* gene in the human genome that was not found in mice (Figures 4E and 4F). In addition, there was an inversion of the region containing the human *NOS2* locus compared with mouse *Nos2*. In human diseases, inversions and insertions/deletions can affect TAD boundaries, and exposure to new TAD content impacts the functional consequence of the structural variation (Spielmann et al., 2018). Therefore, we next analyzed the TADs surrounding the human *NOS2* and mouse *Nos2* locus in immune cells. We found that a single TAD encompassed *Nos2* and the surrounding genes in mice, whereas there were two separate TADs containing this region in the human genome (Figures 4E and 4F). The divergent TAD boundaries in the human genome were adjacent to the insertion and several CTCF sites exclusive to the human genome (Figure 4F). These data suggest that evolutionary structural variation in this region disrupted TAD boundaries.

We next characterized the LPS-inducible H3K27Ac enhancer landscape in mouse and human macrophages contained within the species-specific TADs (Cuartero et al., 2018; Novakovic et al., 2016). We identified two LPS-inducible H3K27Ac enhancer elements located upstream of the *Nos2* gene in ChIP-seq datasets from mouse macrophages (Figure 4E, red box). However, the sequences for these H3K27Ac enhancer elements were not conserved in the human genome (Figure 4E, conservation track). Consistent with these findings, there were no LPS-inducible H3K27Ac enhancer elements in proximity to *NOS2* in human macrophages (Figure 4F). Taken together, these data indicate that TAD boundaries, CTCF sites, and H3K27Ac enhancer elements diverge between the human *NOS2* and mouse *Nos2* locus, correlating with the mouse-specific inducible expression of *Nos2*.

NOS2 and genes near evolutionary breakpoints diverge in expression between species in disease states

Nos2, and its protein product iNOS, have been widely studied in mouse models in the context of the microbiome, bacterial infections, and tumor responses (Bogdan, 2015; Fritz et al., 2011; Goldberg et al., 2018). Therefore, we next analyzed whether there was any evidence that *NOS2* could be expressed in these settings in human immune cells. In contrast to mouse models, there was no evidence for *NOS2* expression in IgA^+ plasma cells in the human ileum (Figure S4A), circulating blood cells in TB patients (Figure 4C) (Moreira-Teixeira et al., 2020), or tumor-infiltrating immune cells in non-small-cell lung cancer patients (Figure S4B) (Zilionis et al., 2019). Importantly, *NOS2* can be expressed in epithelial cells of the human small intestine, demonstrating there is not an overall defect in the gene (data not shown) (Kayisoglu et al., 2021). Collectively, these data indicate that species-specific differences in *NOS2* expression in immune cells extend to pathogenic and disease states.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The fact that *NOS2* expression was altered in immune cells in both healthy and disease states led us to explore whether there were other genes with this pattern. To address this, we examined published single-cell RNA-seq datasets for tumor-infiltrating immune cells from human non-small-cell lung cancer patients and a mouse model of lung adenocarcinoma (Zilionis et al., 2019). Many of the genes with robust cell-type or stimuli-dependent differences in expression between mice and humans also displayed differences in their expression patterns between tumor-infiltrating immune cells from human and mouse lung cancer (Figures 4G, 4H, and S4D). Examples of this included the genes shown in Figure 1, such as *JUN*, *GIMAP6*, *CXCL13*, *IDO1*, and *SLC7A2,* as well as genes including *IRF4* and *CCL20*. Furthermore, we compared gene enrichment patterns for the differentially expressed immune genes and the corresponding genes on either side of the nearby evolutionary translocation breakpoints in the tumor-infiltrating immune cells from mice and humans. There were 175 genes in this category enriched in at least one tumor-infiltrating immune cell type, and notably, 84% had differences in their enrichment patterns between mouse and human tumor-infiltrating immune cells, with T and B cells showing the most variability between species (Figure S4C). As examples, *CCR5*, *CD55*, *CD163*, *CXCR4*, *NAMPT*, and *XBP1* were enriched in different types of tumor-infiltrating immune cells in humans compared with mice (Figures 4G, 4H, S4D, and data not shown). We also examined the expression patterns for genes in proximity to rodent- and primate-specific breakpoints from Figure 2B and identified several additional genes, including *DUSP4*, *CREM*, *FOSL2, KLRB1, TLE1* and *SPON2*, with differences in expression between mouse and human tumor-infiltrating immune cells (Figure S4D and data not shown). Together, these data indicate that genes near breakpoints for evolutionary genome rearrangements often have species-specific expression patterns in both healthy and disease states.

DISCUSSION

In this study, we explore how evolutionary structural variation alters TAD boundaries and composition, with the goal of identifying regulatory mechanisms that contribute to species-specific differences in cell-type and inducible gene expression in immune cells. Typically, research focuses solely on the sequence conservation of regulatory elements to predict whether genes will have similar expression patterns between species. This is a reasonable approach for syntenic regions of the genome without evolutionary rearrangements. However, this approach is often improperly applied to regions of the genome with evolutionary structural variation, likely because of the pervasive assumption that TADs are conserved between species. Under this assumption, tolerated breakpoints for evolutionary genome rearrangements would preferentially occur and insert at TAD boundaries. This model also implies that TADs constrain the impact of evolutionary structural variation by limiting the exchange of regulatory elements between TADs. However, our analysis of a wide range of published genome-wide interaction datasets challenges this model by revealing that TAD composition is frequently altered by evolutionary structural variation in proximity to genes involved in immunity. This means that, in rearranged regions of the genome, sequence-conserved enhancer elements can relocate into hybrid TADs between species. Divergence in TAD content and enhancer location likely contribute to the acquisition (or loss) of cell-type and inducible gene expression potential between mouse and human

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

immune cells. Therefore, conservation of the regulatory landscape available to genes in each species encompasses both sequence and location conservation. Evolutionary changes like the ones identified in this study have been and will continue to be missed when only using linear sequence conservation analyses.

The genome topology field has expanded over the last decade, with technological advances and sophisticated studies enhancing our understanding of genome regulation (Kempfer and Pombo, 2020; Schoenfelder and Fraser, 2019; Zheng and Xie, 2019). In agreement with our findings, analyses from some prior genome topology studies have raised the issue that TADs diverge between species, but this point tends to be lost in citations (discussed in Eres and Gilad, 2021). It is unclear why the assumption that all TAD content is conserved between species became so pervasive in the gene regulation literature and why it has dominated the design of genomics experiments translating mechanisms between mice and humans. Indeed, this oversimplification has hindered genomics researchers from discovering mechanistic principles contributing to differences in gene expression potential between species, a topic that is essential when animal models of human diseases are being used (Gilbertson and Weinmann, 2021). A thorough discussion of possible reasons for the oversimplification of the TAD conservation model is found in Eres and Gilad (2021), but a key point for this discussion is that there are many interpretations of the word “conservation.” Variation in its interpretation has likely influenced the discussion of this and other complex topics. For example, it is true the data indicate a majority (54%) of TADs are conserved between mouse and human cells (Dixon et al., 2012), and the conservation of TAD structures through evolutionary time contributes to similarities in gene regulation mechanisms between species (Szabo et al., 2019). However, as shown here, it is also true that TAD boundaries diverge between species at a large number of locations. Therefore, the word “conserved” in this setting is most likely referring to the presence of similar TAD structures and the conservation of a subset of TAD boundaries across species, but this should not be interpreted as absolute in terms of all TAD boundaries between species. Importantly, this means that each region of the genome needs to be examined individually to determine whether it contains conserved or divergent TAD boundaries between species. More generally, genomics studies must appreciate the complexities of the diverse mechanisms regulating different aspects of the genome rather than minimizing them (Tong et al., 2016). In this light, our study suggests that evolutionary rearranged regions of the genome should be analyzed differently from the blocks of synteny in genomics research.

Oversimplifying the TAD conservation model has conceptually limited the interrogation of regulatory events contributing to species-specific gene expression patterns. This has likely been the reason for a surprising lack of consideration for enhancer location as a relevant variable when translating mechanisms between mice and humans. In this study, we show that the majority of TADs diverged between mouse and human immune cells at the breakpoints for evolutionary rearrangements near genes with species-specific expression patterns. This caused sequence-conserved enhancers to relocate to divergent regions of the genome in the other species, creating unique TADs with species-specific enhancer-promoter interactions. This demonstrates the value of redefining current models to incorporate the concept that TADs can be dynamic between species, especially in rearranged regions of genomes. It also highlights the importance of assessing how balanced evolutionary genome

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

rearrangements, which change the location or orientation of genomic content, contribute to differences in gene regulation potential between species.

When one is assessing location conservation between species, it is important to account for the "perspective" of each gene within the TAD. Consider an example where a single gene in the mouse genome is located on one side of an evolutionary breakpoint and is included in a TAD with nine genes from the other side of the breakpoint. The location conservation analysis might show that 5% of OCRs (i.e., potential enhancer elements) within the TAD maintain both sequence and location conservation with respect to the single gene in the human genome, while 95% of sequence-conserved elements maintain location conservation with respect to the other nine genes. In this case, the "dominant side" interpretation is that 95% of the sequence-conserved elements maintain location conservation. However, from the perspective of the single gene, only 5% of sequence-conserved enhancer elements from the original mouse TAD maintain location conservation in human, and importantly, this means the single gene lost the potential to interact with 95% of the elements from the mouse TAD because of the genome rearrangement. Therefore, each gene will have its own perspective for enhancer element location conservation.

Both balanced and unbalanced forms of structural variation can impact TAD genomic content by altering TAD boundaries, and in our study we also uncovered the effects of unbalanced evolutionary structural variation at the *NOS2* locus. An insertion of sequence near *NOS2* coincided with the split of a single TAD in mice into two neo-TADs in humans. In addition to divergent TADs, the underlying sequences for two TLR-inducible H3K27Ac elements were not conserved in the human genome. Although it is unclear whether one or both of these evolutionary events contribute to the lack of TLR-inducible expression of *NOS2* in human macrophages, these data illustrate the importance of assessing conservation at the level of sequence, location, TAD formation, and genomic interactions for studies in mouse and human immune cells. Research performed in mice has suggested a role for iNOS in immune responses to bacterial infections and tumor responses (Bogdan, 2015; Goldberg et al., 2018). Yet, the lack of conservation shown here between mice and humans suggests it is imperative to look more closely at whether iNOS plays a similar role in human immune responses.

Our study also demonstrates the powerful approach of using published genomics datasets to address mechanistic questions with high importance. Individual laboratories do not have the technical or financial resources to generate the breadth of genomics datasets across cell types and species that are needed to simultaneously interrogate the range of one- and three-dimensional events contributing to genome regulation. This has limited interpretations, especially in regard to translating discoveries made in preclinical models to human diseases. To address this, research funding agencies have invested substantial monetary resources into consortiums to perform specialized genomics studies in model organisms and the diverse human population (Consortium, 2011; Genomes Project et al., 2015; Heng et al., 2008). Many individual laboratories have also applied their extensive expertise to complex biological problems, generating a plethora of diverse genomic resources (Kim et al., 2021; Moreira-Teixeira et al., 2020; Wang et al., 2018; Zilionis et al., 2019). Although each individual study explores a sophisticated question, those analyses represent only a small

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

fraction of the information contained in genomics datasets. Therefore, our study represents an example of addressing questions of high research value with published datasets and resources. The approach presented here allowed us to minimize bias in interpretations by cross-comparing datasets from a variety of sources to identify reproducible regulatory events between independently generated datasets. Related to this point, one reason that the model for TAD conservation has been hard to address is that technical, biological, and informatics differences can affect interpretations of TAD boundaries (Dali and Blanchette, 2017). Here, we were able to mitigate these variables, an endeavor not possible without the collective efforts of the scientific research community.

Mice are the most widely used mechanistic model of human immunity, and many fundamental principles in immunology have been defined with mouse models. However, many promising preclinical studies are not recapitulated in humans (Eynde et al., 2020), and not all aspects of the human immune response are observed in mice. Therefore, more knowledge is needed to predict which regulatory events translate between species to refine preclinical models. Incorporating the concepts of evolutionary structural variation and TAD divergence into species comparisons will empower researchers to better utilize and interpret model systems.

Limitations of the study

This study was not designed to address whether there were differences in TAD genomic content between cell types, as we made interpretations based on the aggregate findings from multiple cell types. Therefore, it will be interesting in future studies to define whether there are any cell type constraints on TAD boundaries in rearranged regions of the genome. Related to this point, our study focused on the breakpoints for evolutionary genome rearrangements near genes with differential expression in immune cells between mice and humans. We did not explore species-specific differences in gene expression in other biological systems or how those systems might be affected by rearranged regions of the genome. As mentioned, the immune system is thought to evolve more rapidly between species than other developmental systems, meaning mechanisms that regulate the immune response might be more prone to diverge between species. However, we did note that several recognizable metabolism and cancer-related genes were near evolutionary breakpoints. It will be important in future studies to define how the principles described here might affect translating mechanisms between mice and humans in cancer models as well as other biological systems.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Amy Weinmann (weinmann@uab.edu).

Materials availability—This study did not generate new unique reagents.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Data and code availability

- This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the Key resources table.
- This study does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the Lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The human and mouse datasets used in this study were from published studies and publicly available resources. Please see original publications, NCBI gene expression omnibus (GEO) entries, and references for details on each of the human and mouse experimental models as cited in the Methods Details section and Key Resources Table.

METHOD DETAILS

Defining differential gene expression between mouse and human immune cells—Genes with differences in expression between mouse and human immune cells were extracted from published studies designed to compare 1) expression profiles in different cell types (Shay et al., 2013), 2) $CD4^+$ T cells activated with αCD3 and αCD28 (Shay et al., 2013), or 3) LPS stimulated mouse bone-marrow derived macrophages and human monocyte-derived macrophages (Schroder et al., 2012). For macrophages, only genes that were induced at least 5-fold in at least one species were included. In total, 169 genes with cell-type-specific differences in expression in resting immune cells (Shay et al., 2013), 175 genes with differences in TCR and co-stimulation potential in $CD4^+$ T cells (Shay et al., 2013), and 134 genes with differences in LPS induction potential in macrophages (Schroder et al., 2012) were identified from these datasets.

To confirm that genes were differentially expressed between mouse and human immune cells, we analyzed the expression of a subset of genes in independently generated and publicly available datasets. Gene expression profiles were downloaded using the Broad Single Cell portal for single cell RNA-seq datasets from the Immunological Genome Project (ImmGen) for blood mononuclear cells from two human donors (13,316 cells) for Figure 1A: (https://singlecell.broadinstitute.org/single_cell/study/SCP345/ica-blood-mononuclear-cells-2-donors-2-sites?scpbr=immune-cell-atlas#study-visualize), or mouse whole CD45 + splenocytes from C57BL/6 mice (HMS) (10,651 cells) for Figure 1B: (https://singlecell.broadinstitute.org/single_cell/study/SCP306/whole-cd45-splenocytes-from-b6-mice-10x-hms?scpbr=immunological-genome-project#study-visualize), or human ileum lamina propria (39,563 cells) for Figure S4: (https://singlecell.broadinstitute.org/single_cell/study/SCP359/ica-ileum-lamina-propria-immunocytes-sinai?scpbr=immune-cell-atlas#study-visualize). For Figures 1C–1E, 3A, 3B, 4A, and 4B, RNA-seq fastq files were obtained from GSE67355 (Tong et al., 2016) and GSE135753 (Manago et al., 2019). A data analysis pipeline consisting of Trim Galore!, FastQC, RNA STAR (Dobin et al., 2013), featureCounts (Liao et al., 2014), and DESeq2 (Love et al., 2014) was performed with the publicly available Galaxy server (Afgan et al., 2018). For Figures 4C and 4D, raw count files were downloaded from the online

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

platform https://ogarra.shinyapps.io/tbtranscriptome for human (GSE107995) and mouse (GSE140945) blood from healthy controls or tuberculosis infection (Moreira-Teixeira et al., 2020; Singhania et al., 2018), and were processed using DESeq2. DESeq2 uses the following statistical tests: Wald test for calculating p values, and a Benjamini-Hochberg test for calculating adjusted p values.

Determination of evolutionary structural variation between the mouse and human genomes—The hg38 and mm10 genome releases for human and mouse, respectively, were used for the evolutionary structural variation analysis. Gene annotations and species conservation tracks for each genome were displayed with the University of California Santa Cruz (UCSC) genome browser (Kent et al., 2002). We first identified the chromosomal location for each differentially expressed gene in the human and mouse genomes. We next identified the genes found on the 5′ and 3′ sides within an approximately 1–2 Mb interval in each genome. We also defined gaps in genome sequence conservation between mouse and human in this interval using the sequence conservation tracks provided by the UCSC genome browser. Evolutionary translocations were defined if there was a gap in sequence conservation between the mouse and human genomes, with the gene on one side of the gap common between the mouse and human genomes, while the gene on the other side of the gap was different between the two species. These conservation gaps were then defined as the evolutionary translocation breakpoints, with the common gene and divergent gene most proximal to the breakpoint considered the breakpoint genes for the TAD analysis. To determine the nature of the evolutionary translocation, the divergent breakpoint gene for each species was located in the genome of the other species, along with the next four genes, to confirm the origin of the genome configuration in mice and humans (See Table S2). Of note, in some instances, amplifications of species-specific genes were found proximal to the breakpoint genes in one species, possibly indicating an unstable region in the genome. Inversions were defined as an evolutionary translocation in which the divergent breakpoint gene was located on the same chromosome and the intervening genes could be traced back to the common breakpoint gene without interruption by another translocation (see *P2RX7*/*OASL* as an example). Evolutionary amplifications were defined as a region with differences in the number of related genes between the mouse and human genomes. Most often, these caused a disruption in the sequence conservation track predominantly in either the mouse or human genome. Evolutionary insertions/deletions were defined as regions with large gaps in sequence conservation (see *NOS2* example) or the addition of a gene specific to the mouse or human genome.

For Figures 2C, 2D, and S2, evolutionary translocations between the mouse and human genome are depicted from the perspective of the mouse as the reference genome. This is not intended to imply direct relationships or directionality of the evolutionary genome rearrangements, but rather to represent the relative location of homologous content for human chromosomes in the mouse genome. For the human chromosomes, the breakpoint genes, the proximal 3′ and 5′ neighboring genes, and immune genes of interest are displayed below the chromosome schematic. For mouse, the immune gene and the genes proximal to the breakpoint are shown.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Ancestral and gene set enrichment analysis (GSEA) for evolutionary translocations and inversions—The genomic configuration for the genes most proximal to the common and divergent breakpoint genes for genome rearrangements were defined in several species with the genome builds available in the UCSC genome browser. This included the genome builds for several primates such as rhesus, baboon, marmoset, and bushbaby, as well as rodents such as rat, Chinese hamster and guinea pig. We also examined other species such as cow, horse, dog, and chicken. The genomic configuration surrounding the evolutionary rearrangement breakpoint genes from each species was then compared to the human and mouse configuration to determine whether the species more closely resembled the human or mouse genome in these regions. For the GSEA (Subramanian et al., 2005) between the mouse and human genomes, we identified the 4 most proximal genes on the 5′ and 3′ sides of the evolutionary breakpoint for the subset of 30+ genes with roles in the immune response to infectious diseases in proximity to the breakpoints of evolutionary translocations and inversions (Table S2). We then performed GSEA on the gene sets from the mouse genome, and an independent GSEA for the gene sets from the human genome.

Analysis of rodent- and primate-specific breakpoints—A liftOver analysis between the human hg17 and hg19 genomes was performed on the coordinates for rodent-, mouse-, primate- and hominoidae-lineage specific evolutionary rearrangement breakpoints from Larkin et al., (2009), followed by a GREAT analysis (McLean et al., 2010) to identify the nearest 5′ and 3′ genes to the breakpoint. We then determined the overlap between the list of species-specific breakpoint genes and the evolutionary rearrangement breakpoint genes from Table S1.

Defining relationships between TADs and evolutionary rearrangement breakpoints—We started this analysis by compiling a series of bed files containing the coordinates for TADs from published studies. For mice, TAD datasets included fetal liver hematopoietic stem cells (GSE119347) (Chen et al., 2019), bone marrow hematopoietic stem cells (GSE119347) (Chen et al., 2019), embryonic stem cells (GSE113339) (Di Giammartino et al., 2019), and mouse embryonic fibroblasts (GSE113339) (Di Giammartino et al., 2019). For humans, TAD datasets included 3 independent peripheral T cell samples (GSE134761) (Kloetgen et al., 2020), embryonic stem cells (GSE106687) (Bertero et al., 2019), mesoderm, (GSE106687) (Bertero et al., 2019), cardiomyocytes (GSE106687) (Bertero et al., 2019) and cardiac progenitors (GSE106687) (Bertero et al., 2019). The original publications and Gene Expression Omnibus (GEO) entries for each study contain detailed information about the data processing parameters used to define TADs, and the processed TAD files used in our analysis are available in the GEO accession entries for the studies as indicated above. Briefly, for TAD determinations, Chen et al., (2019) used the GMAP algorithm (Yu et al., 2017), Di Giammartino et al., (2019) used the HiCExplorer software (Ramirez et al., 2018), Kloetgen et al., (2020) used the hic-bench algorithm (Lazaris et al., 2017), and Bertero et al., (2019) used the directionality index (Dixon et al., 2012). This provided a list of TADs called from diverse informatics methods by independent laboratories for our analysis. In addition, we also performed an analysis to define TADs for $CD4^+CD8^+$ double-positive (DP) T cells (GSE141847) (Fasolino et al., 2020). For this, TADs and boundaries were identified with insulation score (Crane et al., 2015), which

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was implemented in cooltools “diamond-insulation” function (Abdennur and Mirny, 2020). We calculated the genome-wide insulation score at 20 Kb resolution with 500 Kb window size utilizing cooltools diamond-insulation. The insulation score result was converted to a bedgraph file, which was further converted to a bigwig file for visualization purposes. Boundaries were identified with the bins whose boundary strength >0.1 and were further used to define TADs with custom python script.

For our analysis, we examined the evolutionary rearrangement breakpoints in proximity to the genes with differences in expression between mouse and human immune cells, as well as the evolutionary translocations and inversions in proximity to the subset of over 30 genes involved in the immune response to infectious diseases. This represented approximately 120 evolutionary rearrangement breakpoints (Table S1). We first determined whether a TAD crossed the evolutionary rearrangement breakpoint by defining whether the genes proximal to the 5′ and 3′ sides of the breakpoint were contained within the same TAD. In this scenario, both the common and divergent sides of the rearrangement breakpoint were merged into the same TAD. Effectively this meant the genomic content in the TAD must be different between the mouse and human genomes because the content was derived from rearranged (non-syntenic) regions of the genomes. Therefore, for TADs at evolutionary rearrangement breakpoints, the interpretation that a TAD diverged between species did not strictly depend on directly comparing TADs between mouse and human cells. For TADs that contained only one of the genes most proximal to the breakpoint, we inspected each one to determine whether the TAD contained sequence that spanned the evolutionary rearrangement breakpoint, even though it did not reach the nearest gene on the other side of the breakpoint. In the cases where a TAD spanned the evolutionary translocation or inversion breakpoint, we then determined whether the gene with differences in expression between mouse and human immune cells was contained within the TAD that spanned the breakpoint. To ensure interpretations were reproducible and not dependent on informatics programs, cell types, or differences in individual datasets, the final determinations for whether a TAD spanned an evolutionary rearrangement breakpoint were based on the findings from the majority of TAD datasets. If a determination was ambiguous, or did not reach a clear majority, they were excluded from the final tally (17 for human and 7 for mouse fit into this criteria). Datasets from the mouse and human genomes were analyzed separately as indicated in Figures 2G and 2H.

As a secondary approach, we analyzed independent human Hi-C and Hi-ChIP datasets processed and displayed with the publicly available 3D-genome interactive viewer (3DIV) (Kim et al., 2021; Yang et al., 2018). We used the directionality index (DI) parameters with a window size of 2 Mb. We analyzed the TADs from 10 different human immune cell types and stimulation conditions. This included monocyte-derived macrophages (MDM) either mock infected or infected with influenza H5N1 (GSM3112387, GSM3112388, GSM3112391, GSM3112392) (Heinz et al., 2018), two donors for naive $CD4^+$ T cells (GSM2827786, GSM2827787) (Johanson et al., 2018) or naive $CD8^+$ T cells (GSM2827788, GSM2827789) (Johanson et al., 2018), purified human naive B cells (GSM2225739, GSM2225740) (Bunting et al., 2016), purified human germinal center B cells (GSM2225741, GSM2225742) (Bunting et al., 2016), spleen (GSM2322556, GSM2322557) (Schmitt et al., 2016) and thymus (GSM1419083) (Leung et al., 2015).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Determinations were performed as described above and data from this analysis are summarized in Figures S2G and S2H. The interaction contact matrix heatmap from the analysis of $CD8^+$ T cells for *IL6* (Figure 2E) and *GIMAP6* (Figure 2F) or monocyte-derived macrophages for *TLR9* (Figures S2E) are shown, with TADs indicated by blue triangles.

Long-range genomic interactions with *P2RX7* promoter—We analyzed long-range genomic connections with the *P2RX7* promoter in human MDM samples that were either mock infected (GSM3112391, GSM3112392) or infected with influenza H5N1 (GSM3112387, GSM3112388) (Heinz et al., 2018) with the 3D-genome interactive viewer using the Hi-C analysis interaction visualization tool. The default parameters for a distance normalized interaction frequency threshold of two were used for the analysis. We overlayed the interaction plot with H3K27Ac ChIP-seq tracks from unstimulated and LPS stimulated human macrophages (GSE85245; GSM2262977, GSM2263011) (Novakovic et al., 2016). The liftOver program available on the UCSC genome browser was used to determine the relative location of the H3K27Ac ChIP-seq peak located in between the *OASL* and *P2RX7* loci. We also analyzed long-range genomic connections with the *P2RX7* promoter in human $CD4^+$ T cells (GSM2827787) (Johanson et al., 2018) and $CD8^+$ T cells (GSM2827788) (Johanson et al., 2018) by the same approach. The TADs in Figure 3 are from (C) mouse DP T cells (Fasolino et al., 2020) or (C, D) human spleen (Schmitt et al., 2016).

Long-range genomic interactions with the *TLR9* promoter—We analyzed long-range genomic interactions with the *TLR9* promoter in MDMs with the 3D genome interactive viewer similar to the methods described for the *P2RX7* promoter (Figure S2E). As a comparison, we also analyzed long-range genomic interactions with the *TLR9* promoter in M0 macrophages (Javierre et al., 2016) using the 3D genome browser Capture Hi-C tool (Figure S2D) (Wang et al., 2018). Both analyses identified long-range genomic interactions with the *TLR9* promoter that spanned the evolutionary translocation breakpoint indicating these genomic interactions are specific to the human genome as compared to mouse.

Sequence and location conservation analysis—We first determined the mouse mm10 genomic coordinates for the TADs containing the evolutionary rearrangement breakpoints (Table S1) from mouse DP T cells (Fasolino et al., 2020). Using these genomic coordinates, we then identified the open chromatin regions (OCRs) contained with each individual TAD from a variety of immune cells types utilizing the ImmGen chromatin databrowser (http://rstats.immgen.org/Chromatin/chromatin.html) (Yoshida et al., 2019). The immune cell types used in our analysis included splenic naive $CD4^+$ T cells, splenic activated $CD4^+$ T cells, splenic naive $CD8^+$ T cells, splenic terminal effector and memory precursor $CD8^+$ T cells from LCMV infection day 7, splenic effector and central memory $CD8^+$ T cells from LCMV infection day 180, gut $CD8^+$ intraepithelial lymphocytes (IEL) from LCMV infection day 7, splenic $CD4^+$ regulatory T cells, splenic B cells, splenic B cell germinal center centroblasts and centrocytes, splenic memory B cells, splenic marginal zone B cells, splenic B cell plasmablasts and plasma cells, splenic $CD4^+$ dendritic cells (DCs), splenic $CD8^+$ DCs, splenic plasmacytoid DCs, bone marrow and splenic neutrophils, unstimulated and polyIC stimulated alveolar macrophages, peritoneal and lamina propria

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

macrophages, Ly6C high and low blood monocytes, and $CD34^-$ and $CD34^+$ multipotent long term hematopoietic stem cells. For each individual TAD, we downloaded the table containing the mm10 genomic coordinates for the OCRs in the TAD and then sorted this list to only include the genomic coordinates for the OCRs that ImmGen utilized in their systemic analyses (i.e. minus log10 best p value greater than 1.5). We then used the Galaxy informatics online platform to perform a liftOver analysis between the mm10 and hg38 genomes using a conservation threshold of 0.1. This analysis was performed separately for the OCRs contained within each individual TAD. The liftOver analysis provided information about whether the OCRs from immune cells in mice were mapped in the human genome (i.e. sequence conserved), as well as the hg38 genomic coordinates for the sequence conserved OCRs. We next assessed whether the sequence conserved elements were located in a single continuous region in the human genome, or rather they were located in two or more non-continuous regions in the human genome. We designated the region containing the most sequence conserved elements as the 'majority' location, with each additional cluster of sequence conserved elements found in a discrete, non-continuous location as a 'minority' location. A single element mapped to a 'minority' location was excluded from the analysis as a likely false positive for sequence conservation. For each individual TAD, we then calculated the percentage of sequence conserved OCRs found in the 'majority' location in comparison to the total conserved OCRs in the TAD. We also calculated the sequence conserved OCRs found in 'minority' locations in comparison to the total conserved OCRs in the TAD.

Determination of LPS-inducible H3K27Ac elements in mouse and human macrophages—H3K27Ac ChIP-seq bigwig tracks from unstimulated or LPS stimulated mouse macrophages (GSM2913416, GSM2913418) (Cuartero et al., 2018) or human macrophages (GSM2262977, GSM2263011) (Novakovic et al., 2016) were visualized with the UCSC genome browser. For the *NOS2* locus, CTCF ChIP-seq and TAD tracks from mouse (GSE60482, GSE141847) (Fasolino et al., 2020; Vahedi et al., 2015) or human (GSE115893; GSM3967075, GSE134761; GSM3967114) (Kloetgen et al., 2020) immune cells were also shown along with the species sequence conservation tracks provided by the UCSC genome browser.

Differential gene expression in human and mouse tumor infiltrating immune cells—Gene expression profiles were extracted from single cell RNA-seq datasets from tumor infiltrating immune cells from human non-small cell lung cancer patients (GSE127465) (Zilionis et al., 2019) or the KP1.9 mouse model of lung adenocarcinoma (GSE127465) (Zilionis et al., 2019) using the SPRING Viewer single cell visualization tool for the human samples: (https://kleintools.hms.harvard.edu/tools/springViewer_1_6_dev.html?datasets/Zilionis2019/human/NSCLC_immune) or mouse samples: (https://kleintools.hms.harvard.edu/tools/springViewer_1_6_dev.html?datasets/Zilionis2019/mouse/all_CD45_cells). The original cell type annotations were used. The human samples represent aggregate data from the tumors of 7 patients (34,450 cells), and the mouse samples represent aggregate data from 2 tumor-bearing mice (9201 cells). We also compared gene enrichment patterns between the tumor infiltrating immune cells from the mouse and human lung cancer samples (Zilionis et al., 2019). For this, we started

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with the list of the differentially expressed immune genes in proximity to evolutionary translocations. We also included in the list the genes on either side of the evolutionary translocation breakpoints in their proximity from the mouse or human genomes. This created a list of approximately 471 total genes in these categories. We next selected individual immune cell types, including neutrophils, monocytes/macrophages/DCs, NK cells, T cells and B cells, and downloaded the top 1,000 enriched genes in each cell type from the mouse and human samples. We then compared the genes from the translocation list to the enrichment lists for each cell type in mouse and human. From these comparisons, we defined 175 genes from the translocation list that were enriched in at least one immune cell type in mice or humans. For these genes, we determined whether each individual gene was enriched in the cell type(s) from humans, mice, or both humans and mice.

QUANTIFICATION AND STATISTICAL ANALYSIS

Reprocessing of RNA-seq datasets was performed as described in the Methods Details section with differential analyses utilizing DESeq2. DESeq2 uses the Wald test for calculating p values, and a Benjamini-Hochberg test for calculating adjusted p values. For graphs, error bars represent standard error of the mean (SEM). Number of biological replicates are found in Methods Details, GEO entries, and figure legends. The statistical tests for published and publicly available datasets can be found in the original publications and GEO entries associated with each dataset as well as the 3D interactive genome viewer as referenced in the Methods Details section.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We would like to thank the NIH for funding (AI061061 to A.S.W. and HL145754 to G.V.). We are also grateful for the high-quality published genomics datasets deposited into the GEO database by many research groups and the freely available user-friendly visualization and analysis tools such as the UCSC genome browser, 3DIV, and 3D genome browsers for Hi-C, and the ImmGen and Galaxy platforms.

REFERENCES

Abdennur N, and Mirny LA (2020). Cooler: scalable storage for Hi-C data and other genomically labeled arrays. Bioinformatics 36, 311–316. 10.1093/bioinformatics/btz540. [PubMed: 31290943]

Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Gruning BA, et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 46, W537–W544. 10.1093/nar/gky379. [PubMed: 29790989]

Audano PA, Sulovari A, Graves-Lindsay TA, Cantsilieris S, Sorensen M, Welch AE, Dougherty ML, Nelson BJ, Shah A, Dutcher SK, et al. (2019). Characterizing the major structural variant alleles of the human genome. Cell 176, 663–675.e19. 10.1016/j.cell.2018.12.019. [PubMed: 30661756]

Bertero A, Fields PA, Ramani V, Bonora G, Yardimci GG, Reinecke H, Pabon L, Noble WS, Shendure J, and Murry CE (2019). Dynamics of genome reorganization during human cardiogenesis reveal an RBM20-dependent splicing factory. Nat. Commun 10, 1538. 10.1038/s41467-019-09483-5. [PubMed: 30948719]

Bogdan C (2015). Nitric oxide synthase in innate and adaptive immunity: an update. Trends Immunol. 36, 161–178. 10.1016/j.it.2015.01.003. [PubMed: 25687683]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Borges da Silva H, Peng C, Wang H, Wanhainen KM, Ma C, Lopez S, Khoruts A, Zhang N, and Jameson SC (2020). Sensing of ATP via the purinergic receptor P2RX7 promotes CD8+ Trm cell generation by enhancing their sensitivity to the cytokine TGF-β. Immunity 53, 158–171.e6. 10.1016/j.immuni.2020.06.010. [PubMed: 32640257]

Bourque G, Pevzner PA, and Tesler G (2004). Reconstructing the genomic architecture of ancestral mammals: lessons from human, mouse, and rat genomes. Genome Res. 14, 507–516. 10.1101/gr.1975204. [PubMed: 15059991]

Bunting KL, Soong TD, Singh R, Jiang Y, Beguelin W, Poloway DW, Swed BL, Hatzi K, Reisacher W, Teater M, et al. (2016). Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. Immunity 45, 497–512. 10.1016/j.immuni.2016.08.012. [PubMed: 27637145]

Chen C, Yu W, Tober J, Gao P, He B, Lee K, Trieu T, Blobel GA, Speck NA, and Tan K (2019). Spatial genome Re-organization between fetal and adult hematopoietic stem cells. Cell Rep. 29, 4200–4211.e7. 10.1016/j.celrep.2019.11.065. [PubMed: 31851943]

Choi J, Baldwin TM, Wong M, Bolden JE, Fairfax KA, Lucas EC, Cole R, Biben C, Morgan C, Ramsay KA, et al. (2019). Haemopedia RNA-seq: a database of gene expression during haematopoiesis in mice and humans. Nucleic Acids Res. 47, D780–D785. 10.1093/nar/gky1020. [PubMed: 30395284]

Consortium EP (2011). A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 9, e1001046. 10.1371/journal.pbio.1001046. [PubMed: 21526222]

Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J, and Meyer BJ (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature 523, 240–244. 10.1038/nature14450. [PubMed: 26030525]

Cuartero S, Weiss FD, Dharmalingam G, Guo Y, Ing-Simmons E, Masella S, Robles-Rebollo I, Xiao X, Wang YF, Barozzi I, et al. (2018). Control of inducible gene expression links cohesin to hematopoietic progenitor self-renewal and differentiation. Nat. Immunol 19, 932–941. 10.1038/s41590-018-0184-1. [PubMed: 30127433]

Dali R, and Blanchette M (2017). A critical assessment of topologically associating domain prediction tools. Nucleic Acids Res. 45, 2994–3005. 10.1093/nar/gkx145. [PubMed: 28334773]

Di Giammartino DC, Kloetgen A, Polyzos A, Liu Y, Kim D, Murphy D, Abuhashem A, Cavaliere P, Aronson B, Shah V, et al. (2019). KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. Nat. Cell. Biol 21, 1179–1190. 10.1038/s41556-019-0390-6. [PubMed: 31548608]

Di Virgilio F, Dal Ben D, Sarti AC, Giuliani AL, and Falzoni S (2017). The P2X7 receptor in infection and inflammation. Immunity 47, 15–31. 10.1016/j.immuni.2017.06.020. [PubMed: 28723547]

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, and Ren B (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380. 10.1038/nature11082. [PubMed: 22495300]

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. 10.1093/bioinformatics/bts635. [PubMed: 23104886]

Eres IE, and Gilad Y (2021). A TAD skeptic: is 3D genome topology conserved? Trends Genetics 37, 216–223. 10.1016/j.tig.2020.10.009.

Eyler CE, Wu Q, Yan K, MacSwords JM, Chandler-Militello D, Misuraca KL, Lathia JD, Forrester MT, Lee J, Stamler JS, et al. (2011). Glioma stem cell proliferation and tumor growth are promoted by nitric oxide synthase-2. Cell 146, 53–66. 10.1016/j.cell.2011.06.006. [PubMed: 21729780]

Van den Eynde BJ, van Baren N, and Baurain J-F (2020). Is there a clinical future for Ido1 inhibitors after the failure of epacadostat in melanoma? Annu. Rev. Cancer Biol 4, 241–256. 10.1146/annurev-cancerbio-030419-033635.

Fasolino M, Goldman N, Wang W, Cattau B, Zhou Y, Petrovic J, Link VM, Cote A, Chandra A, Silverman M, et al. (2020). Genetic variation in type 1 diabetes reconfigures the 3D chromatin organization of T cells and alters gene expression. Immunity 52, 257–274.e11. 10.1016/j.immuni.2020.01.003. [PubMed: 32049053]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Formaglio P, Alabdullah M, Siokis A, Handschuh J, Sauerland I, Fu Y, Krone A, Gintschel P, Stettin J, Heyde S, et al. (2021). Nitric oxide controls proliferation of Leishmania major by inhibiting the recruitment of permissive host cells. Immunity 54, 2724–2739.e10. 10.1016/j.immuni.2021.09.021. [PubMed: 34687607]

Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schopflin R, Kraft K, Kempfer R, Jerkovic I, Chan WL, et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature 538, 265–269. 10.1038/nature19800. [PubMed: 27706140]

Fritz JH, Rojas OL, Simard N, McCarthy DD, Hapfelmeier S, Rubino S, Robertson SJ, Larijani M, Gosselin J, Ivanov II, et al. (2011). Acquisition of a multifunctional IgA+ plasma cell phenotype in the gut. Nature 481, 199–203. [PubMed: 22158124]

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, and Abecasis GR (2015). A global reference for human genetic variation. Nature 526, 68–74. 10.1038/nature15393. [PubMed: 26432245]

Gilbertson SE, and Weinmann AS (2021). Conservation and divergence in gene regulation between mouse and human immune cells deserves equal emphasis. Trends Immunol. 42, 1077–1087. 10.1016/j.it.2021.10.007. [PubMed: 34740529]

Goldberg MF, Roeske EK, Ward LN, Pengo T, Dileepan T, Kotov DI, and Jenkins MK (2018). Salmonella persist in activated macrophages in T cell-sparse granulomas but are contained by surrounding CXCR3 ligand-positioned Th1 cells. Immunity 49, 1090–1102.e7. 10.1016/j.immuni.2018.10.009. [PubMed: 30552021]

Gross TJ, Kremens K, Powers LS, Brink B, Knutson T, Domann FE, Philibert RA, Milhem MM, and Monick MM (2014). Epigenetic silencing of the human NOS2 gene: rethinking the role of nitric oxide in human macrophage inflammatory responses. J. Immunol 192, 2326–2338. 10.4049/jimmunol.1301758. [PubMed: 24477906]

Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L, et al. (2018). Transcription elongation can affect genome 3D structure. Cell 174, 1522–1536.e22. 10.1016/j.cell.2018.07.047. [PubMed: 30146161]

Heng TS, Painter MW, Immunological Genome Project C, Lukacs-Kornek V, Mauermann N, Turley SJ, Koller D, Kim FS, Wagers AJ, Asinovski N, et al. (2008). The Immunological Genome Project: networks of gene expression in immune cells. Nat. Immunol 9, 1091–1094. 10.1038/ni1008-1091. [PubMed: 18800157]

Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, Cairns J, Wingett SW, Varnai C, Thiecke MJ, et al. (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell 167, 1369–1384.e19. 10.1016/j.cell.2016.09.037. [PubMed: 27863249]

Johanson TM, Coughlan HD, Lun ATL, Bediaga NG, Naselli G, Garnham AL, Harrison LC, Smyth GK, and Allan RS (2018). Genome-wide analysis reveals no evidence of trans chromosomal regulation of mammalian immune development. Plos Genet. 14, e1007431. 10.1371/journal.pgen.1007431. [PubMed: 29883495]

Kayisoglu O, Weiss F, Niklas C, Pierotti I, Pompaiah M, Wallaschek N, Germer CT, Wiegering A, and Bartfeld S (2021). Location-specific cell identity rather than exposure to GI microbiota defines many innate immune signalling cascades in the gut epithelium. Gut 70, 687–697. 10.1136/gutjnl-2019-319919. [PubMed: 32571970]

Kempfer R, and Pombo A (2020). Methods for mapping 3D chromosome architecture. Nat. Rev. Genet 21, 207–226. 10.1038/s41576-019-0195-2. [PubMed: 31848476]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, and Haussler AD (2002). The human genome browser at UCSC. Genome Res. 12, 996–1006. 10.1101/gr.229102. [PubMed: 12045153]

Kim K, Jang I, Kim M, Choi J, Kim MS, Lee B, and Jung I (2021). 3DIV update for 2021: a comprehensive resource of 3D genome and 3D cancer genome. Nucleic Acids Res. 49, D38–D46. 10.1093/nar/gkaa1078. [PubMed: 33245777]

Kloetgen A, Thandapani P, Ntziachristos P, Ghebrechristos Y, Nomikou S, Lazaris C, Chen X, Hu H, Bakogianni S, Wang J, et al. (2020). Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. Nat. Genet 52, 388–400. 10.1038/s41588-020-0602-9. [PubMed: 32203470]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Krefting J, Andrade-Navarro MA, and Ibn-Salem J (2018). Evolutionary stability of topologically associating domains is associated with conserved gene regulation. BMC Biol. 16, 87. 10.1186/s12915-018-0556-x. [PubMed: 30086749]

Kumar BV, Ma W, Miron M, Granot T, Guyer RS, Carpenter DJ, Senda T, Sun X, Ho SH, Lerner H, et al. (2017). Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. Cell Rep. 20, 2921–2934. 10.1016/j.celrep.2017.08.078. [PubMed: 28930685]

Larkin DM, Pape G, Donthu R, Auvil L, Welge M, and Lewin HA (2009). Breakpoint regions and homologous synteny blocks in chromosomes have different evolutionary histories. Genome Res. 19, 770–777. 10.1101/gr.086546.108. [PubMed: 19342477]

Lazaris C, Kelly S, Ntziachristos P, Aifantis I, and Tsirigos A (2017). HiC-bench: comprehensive and reproducible Hi-C data analysis designed for parameter exploration and benchmarking. BMC genomics 18, 22. 10.1186/s12864-016-3387-6. [PubMed: 28056762]

Leung D, Jung I, Rajagopal N, Schmitt A, Selvaraj S, Lee AY, Yen CA, Lin S, Lin Y, Qiu Y, et al. (2015). Integrative analysis of haplotype-resolved epigenomes across human tissues. Nature 518, 350–354. 10.1038/nature14217. [PubMed: 25693566]

Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. 10.1093/bioinformatics/btt656. [PubMed: 24227677]

Link VM, Duttke SH, Chun HB, Holtman IR, Westin E, Hoeksema MA, Abe Y, Skola D, Romanoski CE, Tao J, et al. (2018). Analysis of genetically diverse macrophages reveals local and domain-wide mechanisms that control transcription factor binding and function. Cell 173, 1796–1809.e17. 10.1016/j.cell.2018.04.018. [PubMed: 29779944]

Long HK, Prescott SL, and Wysocka J (2016). Ever-changing landscapes: transcriptional enhancers in development and evolution. Cell 167, 1170–1187. 10.1016/j.cell.2016.09.018. [PubMed: 27863239]

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

Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell 161, 1012–1025. 10.1016/j.cell.2015.04.004. [PubMed: 25959774]

Mackay LK, Minnich M, Kragten NAM, Liao Y, Nota B, Seillet C, Zaid A, Man K, Preston S, Freestone D, et al. (2016). Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. Science 352, 459–463. 10.1126/science.aad2035. [PubMed: 27102484]

Manago A, Audrito V, Mazzola F, Sorci L, Gaudino F, Gizzi K, Vitale N, Incarnato D, Minazzato G, Ianniello A, et al. (2019). Extracellular nicotinate phosphoribosyltransferase binds Toll like receptor 4 and mediates inflammation. Nat. Commun 10, 4116. 10.1038/s41467-019-12055-2. [PubMed: 31511522]

Mattner J, Wandersee-Steinhauser A, Pahl A, Rollinghoff M, Majeau GR, Hochman PS, and Bogdan C (2004). Protection against progressive leishmaniasis by IFN-β. J. Immunol 172, 7574–7582. 10.4049/jimmunol.172.12.7574. [PubMed: 15187137]

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, and Bejerano G (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol 28, 495–501. 10.1038/nbt.1630. [PubMed: 20436461]

Merot C, Oomen RA, Tigano A, and Wellenreuther M (2020). A roadmap for understanding the evolutionary significance of structural genomic variation. Trends Ecol. Evol 35, 561–572. 10.1016/j.tree.2020.03.002. [PubMed: 32521241]

Miller BH, Fratti RA, Poschet JF, Timmins GS, Master SS, Burgos M, Marletta MA, and Deretic V (2004). Mycobacteria inhibit nitric oxide synthase recruitment to phagosomes during macrophage infection. Infect. Immun 72, 2872–2878. 10.1128/iai.72.5.2872-2878.2004. [PubMed: 15102799]

Moreira-Teixeira L, Tabone O, Graham CM, Singhania A, Stavropoulos E, Redford PS, Chakravarty P, Priestnall SL, Suarez-Bonnet A, Herbert E, et al. (2020). Mouse transcriptome reveals potential

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

signatures of protection and pathogenesis in human tuberculosis. Nat. Immunol 21, 464–476. 10.1038/s41590-020-0610-z. [PubMed: 32205882]

Novakovic B, Habibi E, Wang SY, Arts RJW, Davar R, Megchelen-brink W, Kim B, Kuznetsova T, Kox M, Zwaag J, et al. (2016). β-Glucan reverses the epigenetic state of LPS-induced immunological tolerance. Cell 167, 1354–1368.e14. 10.1016/j.cell.2016.09.034. [PubMed: 27863248]

Park SH, Kang K, Giannopoulou E, Qiao Y, Kang K, Kim G, Park-Min KH, and Ivashkiv LB (2017). Type I interferons and the cytokine TNF cooperatively reprogram the macrophage epigenome to promote inflammatory activation. Nat. Immunol 18, 1104–1116. 10.1038/ni.3818. [PubMed: 28825701]

Philip M, Fairchild L, Sun L, Horste EL, Camara S, Shakiba M, Scott AC, Viale A, Lauer P, Merghoub T, et al. (2017). Chromatin states define tumour-specific T cell dysfunction and reprogramming. Nature 545, 452–456. 10.1038/nature22367. [PubMed: 28514453]

Ramirez F, Bhardwaj V, Arrigoni L, Lam KC, Gruning BA, Villaveces J, Habermann B, Akhtar A, and Manke T (2018). High-resolution TADs reveal DNA sequences underlying genome organization in flies. Nat. Commun 9, 189. 10.1038/s41467-017-02525-w. [PubMed: 29335486]

Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, and Aiden EL (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680. 10.1016/j.cell.2014.11.021. [PubMed: 25497547]

Rowley MJ, and Corces VG (2018). Organizational principles of 3D genome architecture. Nat. Rev. Genet 19, 789–800. 10.1038/s41576-018-0060-8. [PubMed: 30367165]

Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL, and Ren B (2016). A compendium of chromatin contact maps reveals spatially active regions in the human genome. Cell Rep. 17, 2042–2059. 10.1016/j.celrep.2016.10.061. [PubMed: 27851967]

Schoenfelder S, and Fraser P (2019). Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Genet 20, 437–455. 10.1038/s41576-019-0128-0. [PubMed: 31086298]

Schroder K, Irvine KM, Taylor MS, Bokil NJ, Le Cao KA, Masterman KA, Labzin LI, Semple CA, Kapetanovic R, Fairbairn L, et al. (2012). Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. Proc. Natl. Acad. Sci. U S A 109, E944–E953. 10.1073/pnas.1110156109. [PubMed: 22451944]

Shay T, Jojic V, Zuk O, Rothamel K, Puyraimond-Zemmour D, Feng T, Wakamatsu E, Benoist C, Koller D, Regev A, and ImmGen C (2013). Conservation and divergence in the transcriptional programs of the human and mouse immune systems. Proc. Natl. Acad. Sci. U S A 110, 2946–2951. 10.1073/pnas.1222738110. [PubMed: 23382184]

Singhania A, Verma R, Graham CM, Lee J, Tran T, Richardson M, Lecine P, Leissner P, Berry MPR, Wilkinson RJ, et al. (2018). A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. Nat. Commun 9, 2308. 10.1038/s41467-018-04579-w. [PubMed: 29921861]

Sironi M, Cagliani R, Forni D, and Clerici M (2015). Evolutionary insights into host-pathogen interactions from mammalian sequence data. Nat. Rev. Genet 16, 224–236. 10.1038/nrg3905. [PubMed: 25783448]

Spielmann M, Lupianez DG, and Mundlos S (2018). Structural variation in the 3D genome. Nat. Rev. Genet 19, 453–467. 10.1038/s41576-018-0007-0. [PubMed: 29692413]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]

Szabo Q, Bantignies F, and Cavalli G (2019). Principles of genome folding into topologically associating domains. Sci. Adv 5, eaaw1668. 10.1126/sciadv.aaw1668. [PubMed: 30989119]

Tong AJ, Liu X, Thomas BJ, Lissner MM, Baker MR, Senagolage MD, Allred AL, Barish GD, and Smale ST (2016). A stringent systems approach uncovers gene-specific mechanisms regulating inflammation. Cell 165, 165–179. 10.1016/j.cell.2016.01.020. [PubMed: 26924576]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Vahedi G (2021). Remodeling the chromatin landscape in T lymphocytes by a division of labor among transcription factors. Immunol. Rev 300, 167–180. 10.1111/imr.12942. [PubMed: 33452686]

Vahedi G, Kanno Y, Furumoto Y, Jiang K, Parker SCJ, Erdos MR, Davis SR, Roychoudhuri R, Restifo NP, Gadina M, et al. (2015). Super-enhancers delineate disease-associated regulatory nodes in T cells. Nature 520, 558–562. 10.1038/nature14154. [PubMed: 25686607]

Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, and Hadjur S (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell Rep. 10, 1297–1309. 10.1016/j.celrep.2015.02.004. [PubMed: 25732821]

Wang X, Xu J, Zhang B, Hou Y, Song F, Lyu H, and Yue F (2021). Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. Nat. Methods 18, 661–668. 10.1038/s41592-021-01164-w. [PubMed: 34092790]

Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, Li D, Choudhary MNK, Li Y, Hu M, et al. (2018). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol. 19, 151. 10.1186/s13059-018-1519-9. [PubMed: 30286773]

Yang D, Jang I, Choi J, Kim MS, Lee AJ, Kim H, Eom J, Kim D, Jung I, and Lee B (2018). 3DIV: a 3D-genome interaction viewer and database. Nucleic Acids Res. 46, D52–D57. 10.1093/nar/gkx1017. [PubMed: 29106613]

Yoshida H, Lareau CA, Ramirez RN, Rose SA, Maier B, Wroblewska A, Desland F, Chudnovskiy A, Mortha A, Dominguez C, et al. (2019). The cis-regulatory atlas of the mouse immune system. Cell 176, 897–912.e20. 10.1016/j.cell.2018.12.036. [PubMed: 30686579]

Young R, Bush SJ, Lefevre L, McCulloch MEB, Lisowski ZM, Muriuki C, Waddell LA, Sauter KA, Pridans C, Clark EL, and Hume DA (2018). Species-specific transcriptional regulation of genes involved in nitric oxide production and arginine metabolism in macrophages. ImmunoHorizons 2, 27–37. 10.4049/immunohorizons.1700073. [PubMed: 30467554]

Yu W, He B, and Tan K (2017). Identifying topologically associating domains and subdomains by Gaussian Mixture model and Proportion test. Nat. Commun 8, 535. 10.1038/s41467-017-00478-8. [PubMed: 28912419]

Zheng H, and Xie W (2019). The role of 3D genome organization in development and cell differentiation. Nat. Rev. Mol. Cell Biol 20, 535–550. 10.1038/s41580-019-0132-4. [PubMed: 31197269]

Zhu J, Zhang Y, Ghosh A, Cuevas RA, Forero A, Dhar J, Ibsen MS, Schmid-Burgk JL, Schmidt T, Ganapathiraju MK, et al. (2014). Antiviral activity of human OASL protein is mediated by enhancing signaling of the RIG-I RNA sensor. Immunity 40, 936–948. 10.1016/j.immuni.2014.05.007. [PubMed: 24931123]

Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saatcioglu HD, Krishnan I, Maroni G, Meyerovitz CV, Kerwin CM, et al. (2019). Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. Immunity 50, 1317–1334.e10. 10.1016/j.immuni.2019.03.009. [PubMed: 30979687]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- Comparative analysis of mouse and human genome topology and gene expression datasets
- Divergence in immune cell gene expression near evolutionary rearrangement breakpoints
- TAD divergence between mice and humans in rearranged regions of the genome
- Both sequence and location dictate enhancer conservation across species

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

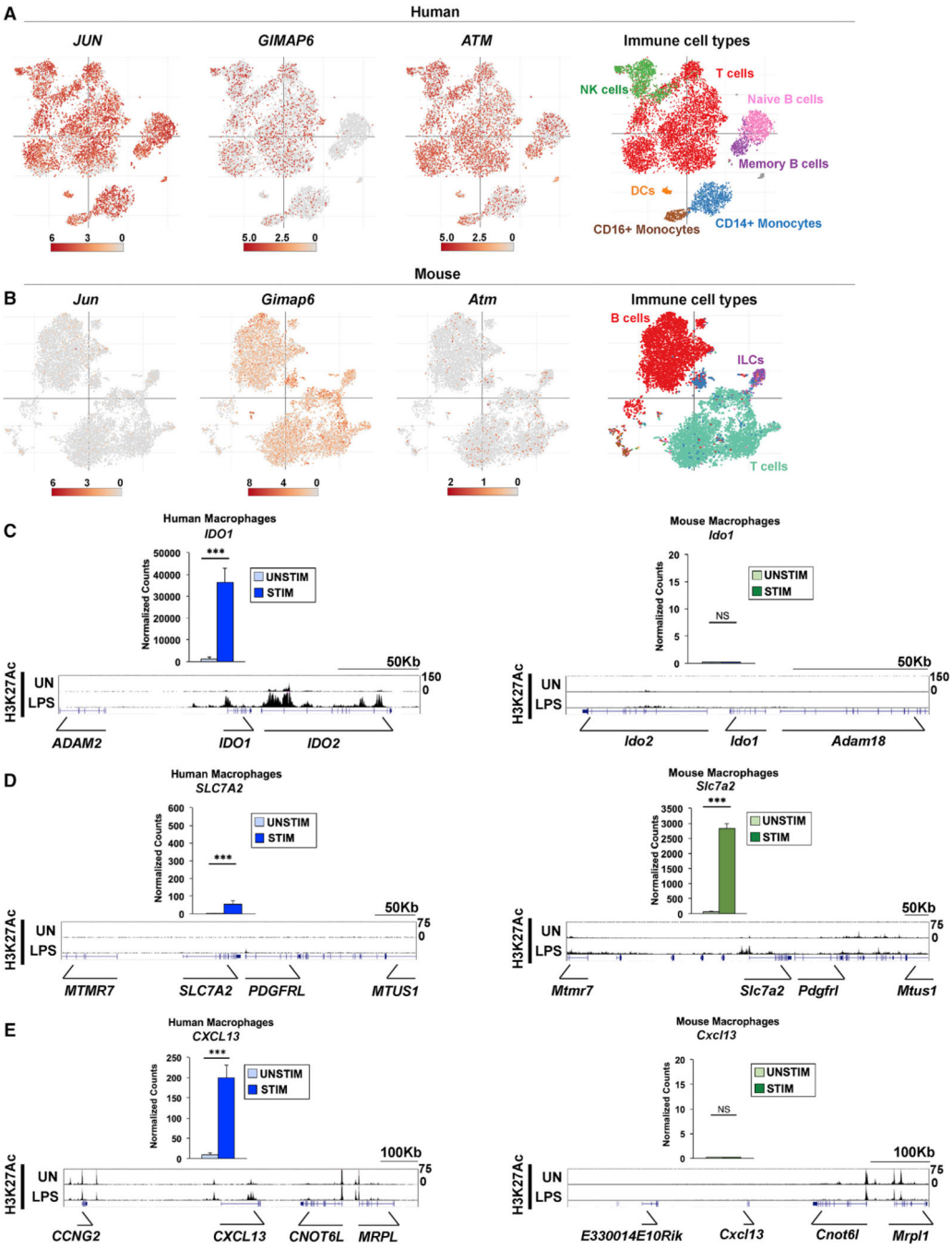


Figure 1. Differences in cell-type and stimulus-dependent gene expression potential between human and mouse immune cells

(A and B) Displayed is single-cell RNA-seq (sc-RNA-seq) datasets depicting gene expression profiles in (A) human or (B) mouse immune cells (from ImmGen; see STAR Methods).

(C–E) Graphs depicting mean normalized counts for RNA-seq datasets from unstimulated compared with LPS-stimulated macrophages from either humans (blue) or mice (green) (GSE135753, GSE67355). Error bars represent SEM, and DESeq2 calculated the adjusted p value with a Benjamini-Hochberg test (***padj < 0.001, NS, not significant); n = 3

(humans) or 2 (mice) biological replicates. Below the graphs are UCSC genome browser tracks displaying H3K27Ac ChIP-seq datasets from unstimulated and LPS-stimulated macrophages from humans or mice (GSE108805, GSE85245). Data are representative of two biological replicates. See also Figure S1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

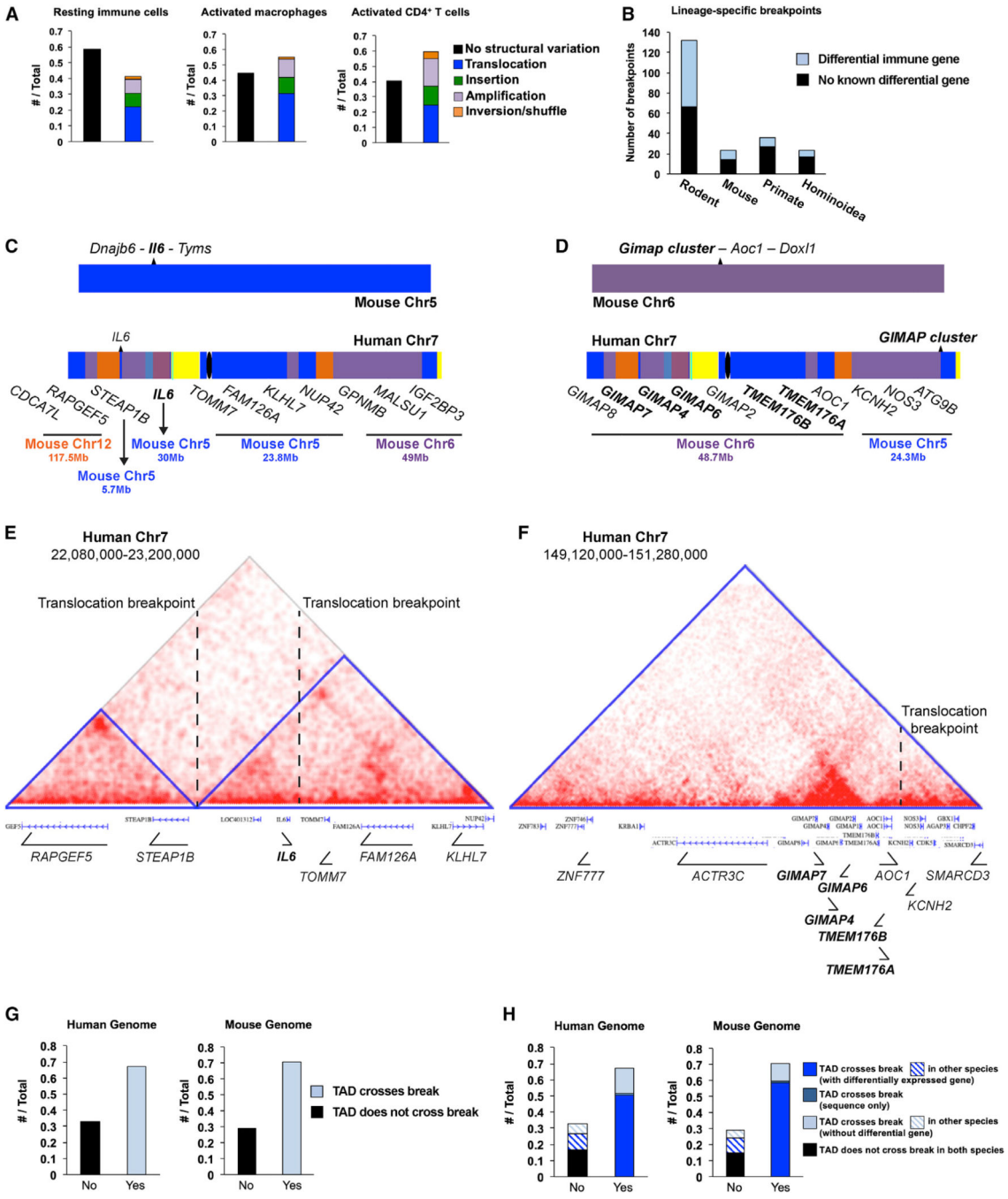


Figure 2. Evolutionary translocations affect genome configuration and TAD genomic content

(A) Graphs representing percentages of genes with differences in expression between human and mouse immune cells found in proximity to evolutionary structural variation between the mouse and human genomes; n = 169 (resting immune cells), 134 (macrophages), or 175 (activated $CD4^+$ T cells) total genes in each category.

(B) Graphical representation of the total number of evolutionary genome rearrangement breakpoints specific to the rodent, mouse, primate, or hominoidae lineages as identified in Larkin et al. (2009). Breakpoints in proximity to genes with differential expression between

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

mouse and human immune cells (blue) or without known differential genes (black) are indicated.

(C) Shown is the relative position of the *Il6* locus on mouse chromosome 5 and *IL6* locus on human chromosome 7, with 3′ and 5′ neighboring genes for human *IL6* listed below the schematic. The colors depict homology to different mouse chromosomal locations.

(D) Shown is a cluster of *GIMAP* genes on chromosome 6 in mice and chromosome 7 in humans, with display arranged similarly to (C).

(C and D) Differentially expressed genes between species are indicated by bold text.

(E and F) Interaction contact matrix heatmap for (E) *IL6* or (F) *GIMAP* cluster in human $CD8^+$ T cells (GSE105776) using 3DIV, with TADs indicated by blue triangles. Gene positions are shown with arrows below the browser, and vertical dashed lines indicate the location of evolutionary translocation breakpoints between the mouse and human genomes. Data are representative of two biological replicates.

(G) Graph depicting the percentage of TADs spanning evolutionary translocation breakpoints in proximity to differentially expressed genes between species from (A); n = 101 (human) or 116 (mouse) total breakpoints.

(H) Data from (G) depicted to represent whether TADs spanning evolutionary breakpoints include (dark blue) or do not include (light blue) the differentially expressed gene between species. Hatched bars represent TADs that cross the evolutionary breakpoint in the genome for the other species. See also Figure S2, Tables S1, and S2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

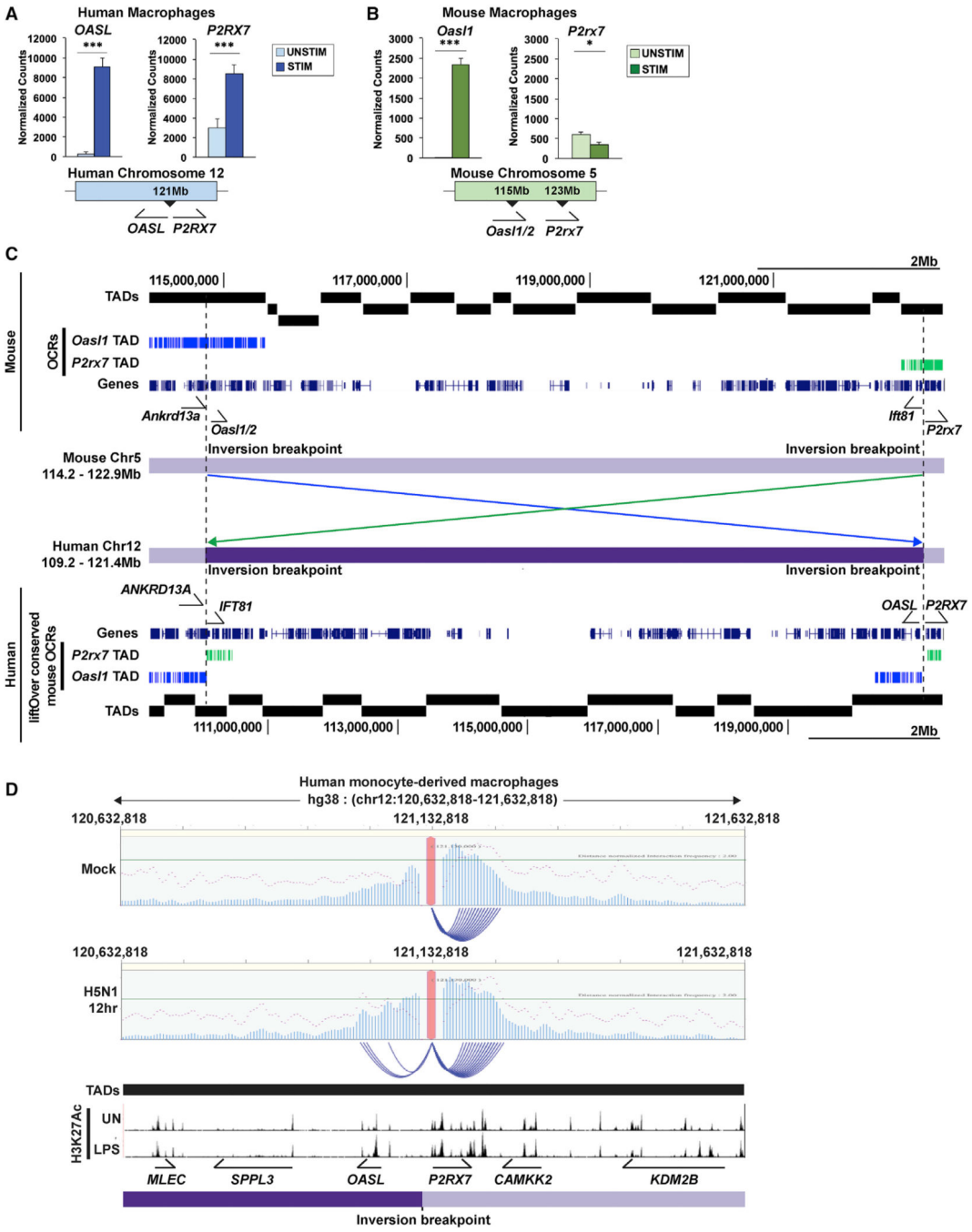


Figure 3. An evolutionary inversion influences the regulation of *P2RX7* between mouse and human immune cells

(A and B) Mean normalized counts for *P2RX7*/*P2rx7* and *OASL*/*Oasl1* in unstimulated and LPS-stimulated (A) human or (B) mouse macrophages (GSE135753, GSE67355). Error bars represent SEM and DESeq2 calculated the adjusted p value with a Benjamini-Hochberg test (***padj < 0.001, *padj < 0.05), n = 3 (humans) or 2 (mice) biological replicates.

(C) Top: Open chromatin regions (OCRs) residing within the TADs containing *Oasl1* (bright blue) or *P2rx7* (green) are shown in UCSC genome browser tracks. The genes (dark blue) and TADs (black) spanning ~8.7 Mb of mouse chromosome 5 (114,189,207–122,859,196)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

between the *Oasl1* and *P2rx7* TADs are also shown. Middle: Schematic representation of the primate-specific evolutionary inversion juxtaposing the *OASL* and *P2RX7* loci. Dark purple indicates the evolutionary inversion, and colored arrows indicate the directional relocation of OCRs from (C, top). Bottom: Locations for sequence-conserved OCRs from the mouse *Oasl1* (bright blue) and *P2rx7* (green) TADs in the human genome are shown. Genes (dark blue) and TADs (black) within human chromosome 12 (109,172,199–121,400,882) are also shown. Dotted lines represent the location of the evolutionary inversion breakpoints.
(D) Genomic interactions with the *P2RX7* promoter in mock-infected or H5N1-infected human monocyte-derived macrophages (GSE113703) using 3DIV. Interactions shown have a distance-normalized interaction frequency of ≥2. Also shown are UCSC genome browser tracks displaying H3K27Ac ChIP-seq from unstimulated and LPS-stimulated human macrophages as in Figure 1. See also Figure S3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

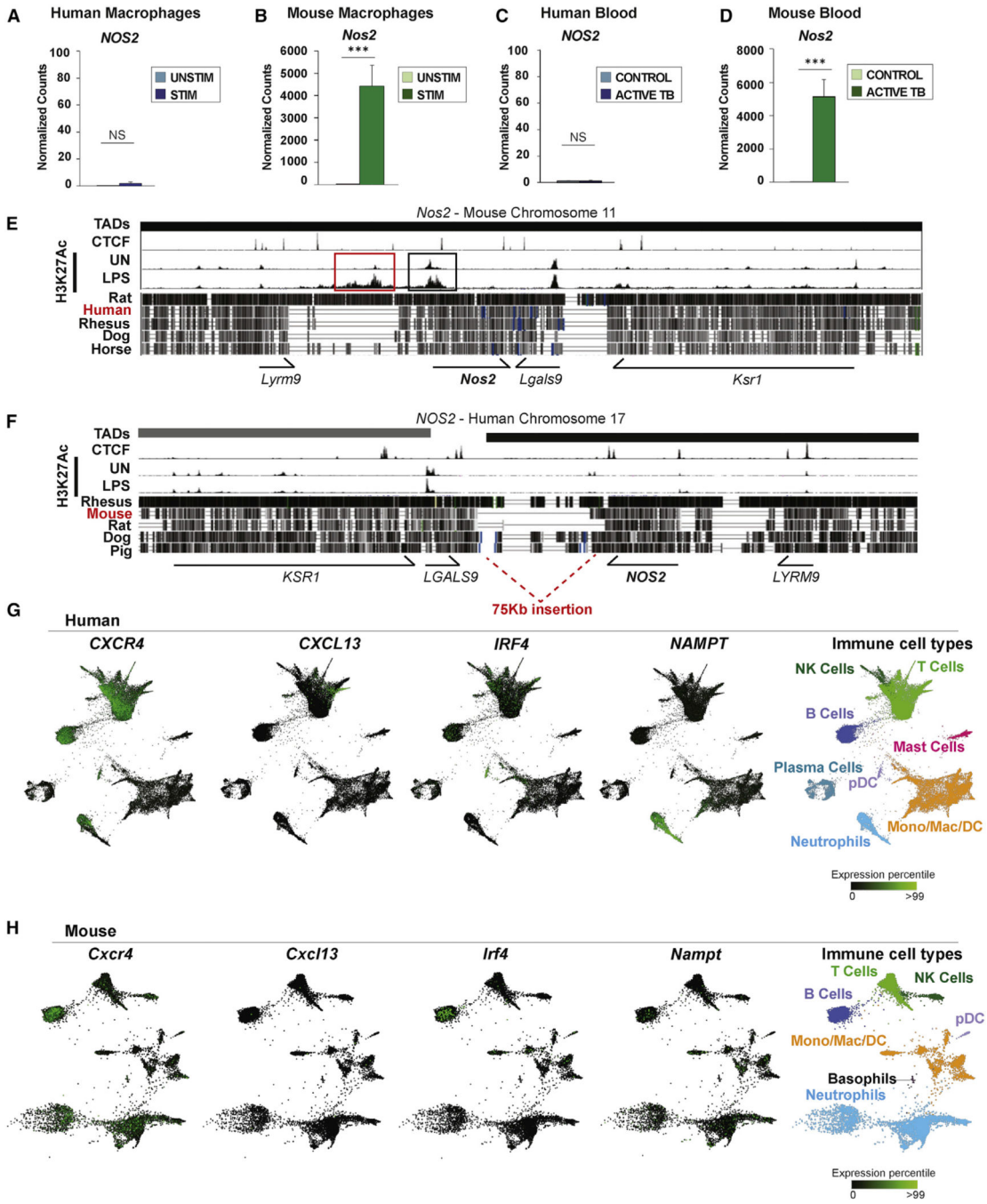


Figure 4. *NOS2*-regulatory events diverge between mouse and human macrophages

(A–D) Mean normalized counts for *NOS2*/*Nos2* in (A and B) unstimulated and LPS-stimulated (A) human or (B) mouse macrophages (GSE135753, GSE67355) or from (C and D) control and TB infection in (C) human or (D) mouse blood samples (GSE107995, GSE140945). Error bars represent SEM and DESeq2 calculated the adjusted p value with a Benjamini-Hochberg test (***padj < 0.001, NS, not significant); (A and B) n = 3 (humans) or 2 (mice) biological replicates or (C and D) n = 20 active TB patients and 12 controls (humans), or 5 infected and 4 controls (mice) biological replicates.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E and F) *NOS2* and surrounding loci in the (E) mouse or (F) human genome are displayed from the UCSC genome browser, with species sequence conservation tracks shown. TAD locations as well as CTCF and H3K27Ac ChIP-seq datasets are also shown (GSE134761, GSE141847, GSE115893, GSE60482, GSE85245, GSE108805). Gene locations and evolutionary structural variation are indicated below the browser image. (G and H) Displayed is sc-RNA-seq depicting the expression of the indicated genes in tumor-infiltrating immune cells from (G) human patients or (H) a mouse model of lung cancer (GSE127465; see STAR Methods). See also Figure S4.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
RNA-seq; *Mus musculus* bone marrow-derived macrophages	Tong et al. (2016); NCBI GEO	GSE67355
RNA-seq; *Homo sapiens* macrophages	Manago et al. (2019); NCBI GEO	GSE135753
RNA-seq; *Homo sapiens* whole blood	Moreira-Teixeira et al. (2020); Singhania et al. (2018); NCBI GEO	GSE107995
RNA-seq; *Mus musculus* blood	Moreira-Teixeira et al. (2020); NCBI GEO	GSE140945
Hi-C; *Mus musculus* fetal liver and bone marrow HSCs	Chen et al. (2019); NCBI GEO	GSE119347
Hi-C; *Mus musculus* embryonic stem cells and embryonic fibroblasts	Di Giammartino et al. (2019); NCBI GEO	GSE113339
Hi-C; *Homo sapiens* RUES2 cell line differentiation	Bertero et al. (2019); NCBI GEO	GSE106687
HiChIP-seq; *Mus musculus* $CD4^+CD8^+$ thymocytes	Fasolino et al. (2020); NCBI GEO	GSE141847
Hi-C; *Homo sapiens* monocyte-derived macrophages	Heinz et al. (2018); NCBI GEO	GSE113703
Hi-C; *Homo sapiens* $CD4^+$ T cells, $CD8^+$ T cells, B cells	Johanson et al. (2018); NCBI GEO	GSE105776
Hi-C; *Homo sapiens* naive B cells, GC B cells	Bunting et al. (2016); NCBI GEO	GSE84022
Hi-C; *Homo sapiens* spleen, other tissues	Schmitt et al. (2016); NCBI GEO	GSE87112
Hi-C; *Homo sapiens* thymus, other tissues	Leung et al. (2015); NCBI GEO	GSE58752
ChIP-seq; *Homo sapiens* monocyte-derived macrophages	Novakovic et al. (2016); NCBI GEO	GSE85245
ChIP-seq; *Mus musculus* bone marrow-derived macrophages	Cuartero et al. (2018); NCBI GEO	GSE108805
ChIP-seq; *Mus musculus* $CD4^+$ Th1 cells	Vahedi et al. (2015); NCBI GEO	GSE60482
ChIP-seq; *Homo sapiens* peripheral blood T cells	Kloetgen et al. (2020); NCBI GEO	GSE115893
Hi-C; *Homo sapiens* peripheral blood T cells	Kloetgen et al. (2020); NCBI GEO	GSE134761
scRNA-seq; *Homo sapiens* and *Mus musculus* tumor infiltrating immune cells	Zilionis et al. (2019); NCBI GEO	GSE127465
RNA-seq; *Homo sapiens* $CD4^+$ and $CD8^+$ T cells	Kumar et al. (2017); NCBI GEO	GSE94964
RNA-seq; *Mus musculus* $CD8^+$ T cells	Mackay et al. (2016); NCBI GEO	GSE70813
ATAC-seq; *Mus musculus* variety of immune cell types	Yoshida et al. (2019); NCBI GEO	GSE100738
Software and algorithms		
TrimGalore!	Galaxy; https://bio.tools/trim_galore	v. 0.4.3.1
RNA STAR	Galaxy; Dobin et al. (2013)	v. 2.6.0b-1
featureCounts	Galaxy; Liao et al. (2013)	v. 1.6.4 + galaxy
DESeq2	Galaxy; Love et al. (2014)	v 2.11.40.6 + galaxy1

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
annotateMyIDs	Galaxy; https://github.com/markdunning/galaxy-annotateMyIDs	v. 3.7.0 + galaxy2
GREAT analysis	McLean et al. (2010)	v. 3.0.0
liftOver	Galaxy; Afgan et al. (2018); UCSC; Kent et al. (2002)	ucsc-liftover v 357
SPRING viewer	Klein lab tools; Zilionis et al. (2019)	Zilionis et al. (2019)
3D Genome Interactive Viewer and Database	Yang et al. (2018); Kim et al. (2021)	http://3div.kr
ImmGen Databrowsers	Heng et al. (2008)	https://www.immgen.org
UCSC Genome Browser	Kent et al. (2002)	https://genome.ucsc.edu
3D Genome Browser Capture Hi-C tool	Wang et al. (2018)	http://3dgenome.fsm.northwestern.edu
ShinyApp	Moreira-Teixeira et al. (2020)	https://ogarra.shinyapps.io/tbtranscriptome
Gene Set Enrichment Analysis	Subramanian et al. (2005)	https://www.gsea-msigdb.org/gsea/index.jsp
Galaxy	Afgan et al. (2018)	https://usegalaxy.org
cooler; cooltools	Abdennur and Mirny (2020); https://open2c.github.io/	https://doi.org/10.5281/zenodo.5214125