Published in final edited form as:

Sci Immunol. 2020 February 07; 5(44): . doi:10.1126/sciimmunol.aay5864.

Noncoding RNA transcription alters chromosomal topology to promote isotype-specific class switch recombination

Gerson Rothschild^{1,*}, Wanwei Zhang^{1,*}, Junghyun Lim^{1,*}, Pankaj Kumar Giri^{1,†}, Brice Laffleur¹, Yiyun Chen², Mingyan Fang^{3,4}, Yuling Chen⁵, Lekha Nair¹, Zhi-Ping Liu⁶, Haiteng Deng⁵, Lennart Hammarström^{3,4}, Jiguang Wang², Uttiya Basu^{1,‡}

¹Department of Microbiology and Immunology, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA.

²Division of Life Science, Department of Chemical and Biological Engineering, Center for Systems Biology and Human Health, and State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Hong Kong SAR, China.

³BGI-Shenzhen, Shenzhen 518083, China.

⁴Division of Clinical Immunology and Transfusion Medicine, Department of Laboratory Medicine, Karolinska Institutet, Karolinska University Hospital Huddinge, SE-141 86 Stockholm, Sweden.

⁵School of Life Sciences, Tsinghua University, Beijing 100084, China.

⁶Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China.

Abstract

B cells undergo two types of genomic alterations to increase antibody diversity: introduction of point mutations into immunoglobulin heavy- and light-chain (*IgH* and *IgL*) variable regions by somatic hypermutation (SHM) and alteration of antibody effector functions by changing the expressed IgH constant region exons through IgH class switch recombination (CSR). SHM and

The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

[‡]Corresponding author. ub2121@cumc.columbia.edu.

[†]Present address: Faculty of Life Sciences and Biotechnology, South Asian University, New Delhi 110021, India.

*These authors contributed equally to this work.

Author contributions: G.R. designed, executed, analyzed, and interpreted the experiments and helped write the manuscript, assisted in all steps by J.L. who also undertook the deep sequencing and interaction studies. W.Z. undertook the bioinformatics analyses of most genomics experiments. P.K.G. performed experiments to determine lncRNA-CSR^{IgA} interacting proteins and analyzed mass spectrometry experiments. Yuling Chen and H.D. performed mass spectrometric analyses of lncRNA-CSR^{IgA} interacting proteins. B.L. analyzed the JH4 mutational analysis and executed the developmental studies. M.F. and L.H. provided data for the human homolog studies, and structural modeling thereof was contributed by Z.-P.L. Yiyun Chen and J.W. studied the interactions among multiple genomic regions affecting coregulatory relationships. U.B. participated in designing the research, interpreting the experiments, and authoring the manuscript. All authors contributed to writing the final manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All the detailed information of (i) each reagent used, (ii) all the pipelines and computational parameters used for analyses of high-throughput genomic datasets, (iii) the accession numbers and location of the raw sequence datasets themselves, (iv) the CRISPR-Cas9/Cas13 protocols used for genome or RNA editing, and (v) the results of the extraction of fecal RNA/DNA can be found in the Supplementary Materials.

SUPPLEMENTARY MATERIALS

immunology.sciencemag.org/cgi/content/full/5/44/eaay5864/DC1

CSR require the B cell–specific activation-induced cytidine deaminase (AID) protein, the transcription of germline noncoding RNAs, and the activity of the 3′ regulatory region (3′RR) super-enhancer. Although many transcription regulatory elements (e.g., promoters and enhancers) reside inside the *IgH* and *IgL* sequences, the question remains whether clusters of regulatory elements outside *IgH* control CSR. Using RNA exosome-deficient mouse B cells where long noncoding RNAs (lncRNAs) are easily detected, we identified a cluster of three RNA-expressing elements that includes *lncCSR^{IgA}* (that expresses lncRNA-CSR^{IgA}). B cells isolated from a mouse model lacking lncRNA-CSR^{IgA} transcription fail to undergo normal levels of CSR to IgA both in B cells of the Peyer’s patches and grown in ex vivo culture conditions. lncRNA-CSR^{IgA} is expressed from an enhancer site (*lncCSR^{IgA}*) to facilitate the recruitment of regulatory proteins to a nearby CTCF site (CTCF^{lncCSR}) that alters the chromosomal interactions inside the TAD^{lncCSR^{IgA}} and long-range interactions with the 3′RR super-enhancer. Humans with IgA deficiency show polymorphisms in the *lncCSR^{IgA}* locus compared with the normal population. Thus, we provide evidence for an evolutionarily conserved topologically associated domain (TAD^{lncCSR^{IgA}}) that coordinates IgA CSR in Peyer’s patch B cells through an lncRNA (lncRNA-CSR^{IgA}) transcription-dependent mechanism.

INTRODUCTION

A remarkable set of developmentally controlled DNA rearrangement and alteration steps in the immunoglobulin (Ig) locus of mammals leads to the generation of antibodies able to interact with the vast number of host-encountered antigens. First, developing progenitor B cells undergo VDJ recombination to accomplish antigen-independent diversity of the B cell receptor and secreted antibodies (1–3). Thereafter, diversity of the Ig loci is further increased through the somatic hypermutation (SHM) and class switch recombination (CSR) mechanisms, which are both antigen-driven processes (4–6).

Rapid progress has been made in understanding the spatial organization of the mammalian genome, including the identification of chromosomal subcompartments known as topologically associating domains (TADs) (7–9). A TAD is a genomic region enriched with regulatory DNA sequences that physically interact with each other more frequently than with sequences outside the TAD. DNA loops that encompass a TAD are often defined by CTCF (CCCTC-binding factor) sites at loop anchors that promote DNA element interactions during TAD formation by a process known as loop extrusion (7). The *IgH* locus requires several key interactions to accomplish productive DNA alterations that facilitate VDJ rearrangement, SHM, and CSR including interactions between and among (i) the two enhancer elements iE μ and 3′ regulatory region (3′RR); (ii) the promoters placed upstream to isotype switch sequences that drive germline transcription, a prerequisite for CSR; and (iii) CTCF/cohesin binding sites, some of which are located surrounding the 3′RR. The Ig heavy-chain locus (*IgH*) is restricted to a single TAD that spans from the *VDJ* exons to the 3′RR, with the 3′RR forming a sub-TAD within the *IgH* locus TAD (TAD^{IgH}). These various elements are involved in controlling DNA loops during *VDJ* recombination and promoting accessibility of recombination substrates via mechanisms involving chromatin scanning and formation of DNA double-strand breaks at CTCF/cohesin binding sites during TAD formation (10–13). The complex regulation of DNA rearrangements and transcription

control occurring inside the *IgH* locus provides a noteworthy opportunity to study various mechanisms that drive TAD regulation (9, 14). Whether regions outside the *IgH* locus play roles in controlling genomic rearrangements and whether these sequences control tissue context-specific antibody diversification processes remain unclear. Flanking the *IgH* locus, a series of TADs exists whose function in regulating *IgH* function is not understood and that have not been investigated previously.

In the literature, various protein factors and transregulatory elements have been shown to control CSR and SHM (15, 16), but there is more to learn about isotype-specific and tissue/microenvironment-specific regulation of antibody diversification, and although some evidence exists regarding inter-TAD and intra-TAD interactions, the physiological meaning of such interactions is only starting to emerge (17–20). In this regard, in the Peyer's patches (PPs) present in the small intestine of mice, there is robust CSR to the IgA-specific isotype, and thus, PPs provide a defined environment for identifying and characterizing tissue-specific antibody gene diversification elements. IgA is the most abundant antibody secreted in the intestinal tract. IgA interacts with a small plasma cell-derived polypeptide termed joining (J) chain to form IgA dimers that recognize polymeric Ig receptor (pIgR) on the basolateral surface of intestinal epithelial cells (IECs). pIgR facilitates the release of secretory IgA onto the surface of the gut that increases the stability of secretory IgA in the intestinal lumen and anchors secreted IgA (sIgA) to mucus. Secretory IgA favors both maintenance of noninvasive commensal bacteria and neutralization of invasive pathogens through multiple mechanisms. By using the V region of IgA, sIgA blocks certain bacterial epitopes from interacting with the apical surface of IECs. In addition, sIgA limits microbial motility by nonspecifically binding bacteria through glycans associated with the secretory component and constant region α (C α) of IgA. Besides neutralizing pathogens in the intestinal lumen, IgA can intercept microbes and toxins inside IECs. sIgA delivers these protective functions without activating the complement cascade, and it also impedes inflammatory damage to the epithelial barrier (21). Accordingly, it would be both interesting and beneficial to identify new mechanisms that control IgA class switching in PPs.

Here, we observe that an evolutionarily conserved TAD that is present about 2.6 Mb away from the *IgH* locus is important for controlling specific IgA CSR in mice. This newly identified conserved TAD (named TAD^{lncCSR IgA}) regulates IgA CSR by promoting specific intra-TAD interactions as well as interactions with the hypersensitivity sites of the 3'RR as previously shown (22). Our study is a rare example of molecular dissection of intra-TAD interactions during B cell development and function and identifies a role for TAD regulation in tissue-specific regulation of antibody diversification.

RESULTS

High-throughput chromosome capture-mediated identification of TAD^{lncCSR IgA}

We and others have reported previously that an 11-subunit complex, known as RNA exosome, regulates the levels of noncoding RNAs in mammalian cells (22–27). In the absence of RNA exosome activity, accumulation of noncoding RNAs eventually leads to formation of DNA/RNA hybrids at various transcription regulatory elements in the genome, leading to the onset of genomic instability (22, 28, 29). Exosome activity deletion in cells

substantially reduces the decay of long noncoding RNAs (lncRNAs) and enhancer RNAs (22, 23, 30). From the transcriptome of B cells that lack RNA exosome activity, we identified three bidirectionally transcribed regions that express RNA exosome-sensitive ncRNAs [schematized in Fig. 1A; one of the three loci, noncoding lncRNA-CSR^{IgA} expressing *lncCSR^{IgA}*, was previously identified in (22)]. Of all the ncRNAs that are expressed from mouse chromosome 12 at levels twofold or higher in RNA exosome-deficient B cells (in comparison with the wild-type B cell transcriptome), these bidirectionally transcribed regions are located closest to the *IgH* locus. We named these three regions as *lncCSR^{IgA}*, *locus A*, and *locus B*. We performed high-throughput chromosome capture (Hi-C) in B cells and found that these three regions exist in a TAD. For the purpose of continuity here, we named this topologically associating cluster as TAD^{lncCSR^{IgA}}. The TAD^{lncCSR^{IgA}} is located about 2.6Mb away from the *IgH* 3'RR (Fig. 1C), and other TADs are found to be present between TAD^{lncCSR^{IgA}} and TAD^{IgH} (fig. S1A). We looked closely at the TAD^{lncCSR^{IgA}} (Fig. 1D) and observed that the bidirectional transcription from *locus A* to *locus B* is placed next to regions that interact with the *lncCSR^{IgA}* locus (Fig. 1, B and E). Thus, we considered that locus A and locus B interact with *lncCSR^{IgA}* inside the TAD^{lncCSR^{IgA}}. These observations are reproducible both in the gut-derived B cell line CH12F3 and in splenic primary B cells (Fig. 1F), indicating the existence of a functional TAD across different forms of B cells.

***lncCSR^{IgA}* TAD is conserved in human cells**

The human and mouse genomes harbor a similar but not quite overlapping *IgH* locus structure (15, 31). We compared regions neighboring the 3'RR of mouse and human genomes for TAD formation in B cell lines and observed partly conserved TADs [Fig. 1G (left) for a mouse B cell line and Fig. 1G (right) for a human B cell line] using insulation scores as a parameter [details in legends of fig. S1 (B and C) and Supplementary Materials and Methods]. Thus, on the basis of the overall conservation of TAD distribution neighboring *IgH*, we wanted to investigate whether the TAD^{lncCSR^{IgA}} is conserved in humans. As shown in Fig. 1G and fig. S1 (B and C), the interactions of *lncCSR^{IgA}/A*, *lncCSR^{IgA}/B*, and A/B occur and are conserved in the mouse and human TADs. Similarly, the distribution of CTCF sites can be inferred from the comparison of fig. S1D and fig. S1E because the presence and interactions of conserved CTCF sites would be expected to exist in such conserved loops seen between *lncCSR^{IgA}*, *locus A*, and *locus B*. Conserved CTCF sites can be mapped and predicted to be convergently or divergently oriented (fig. S1, D and E). Overall, the TADs formed from TAD^{IgH} to TAD^{lncCSR^{IgA}} are quite conserved in structure/organization. The genes expressed in the TAD^{lncCSR^{IgA}} include *Ppp2r5c*, *Dync1h1*, *Hsp90aa1*, *Mok*, *orphan gene17000001k19Rik*, *Wdr20*, *Zfp839*, and an lncRNA neighboring the *Mok* gene lncRNA^{Mok1}. In some human patients with selective IgA deficiency (sIgAD), there is an accumulation of rare variant single-nucleotide polymorphisms and insertions/deletions (InDels) in the *lncCSR^{IgA}* locus (Fig. 1H and fig. S1F). These variants are different from those found in a well-known previous genome-wide association study that identified common variants (32).

TAD^{IncCSR}IgA is important for Ig gene diversification and IgA expression in gut PPs of mice

To dissect the role of TAD^{IncCSR}IgA, we generated a knockout mouse model in which *IncCSR*^{IgA} expression was genetically ablated. For a detailed schematic presentation of the knockout allele map and the genotyping, see fig. S2A. Figure S2B shows a sequencing test of the knockout of the *IncCSR*^{IgA} region in the mouse genome; highlighted areas represent the 3' arm and 5' arm of the targeting vector and the replacement of the *IncCSR*^{IgA} locus with an FRT site. Bone marrow-derived pro-B cells, pre-B cells, and IgM⁺B220⁺ immature B cells demonstrate comparable levels between littermate *IncCSR*^{IgA}^{-/-} and *IncCSR*^{IgA}^{+/+} mice (fig. S2D). Among splenic B cells, there was no difference in the distribution of follicular B cells and marginal zone B cells (fig. S2E). Thus, the *IncCSR*^{IgA}^{-/-} mice did not demonstrate any defects in B cell development. Analysis of circulating serum for antibody isotypes through an enzyme-linked immunosorbent assay approach showed reduced levels of IgA (fig. S2C) without reduction in the expression of IgG1 and IgG2a. However, differences of serum Ig levels are often diminished because of serum accumulation in the animal over long periods of time, which complicates the analyses of class switching isotypes. Thus, we wanted to investigate the specific effect on IgA CSR in the PP B cells in the intestine. PP B cells isolated from *IncCSR*^{IgA}^{-/-} mice show reduced IgA CSR (Fig. 2B). To ascertain whether the decreased level of IgA expression in the PP of *IncCSR*^{IgA}^{-/-} mice was indeed detectable in the B cells resident in the small intestine tissues, we performed immunostaining. It is quite clear that in both the PP (Fig. 2C, top) and the small intestine villi (Fig. 2C, bottom), there are reduced numbers of IgA⁺ B cells in the *IncCSR*^{IgA}^{-/-} mice, in comparison with the *IncCSR*^{IgA}^{WT/WT} mice. IgA expression is critical for microbiome homeostasis in the intestinal tract (33–36). It has been previously shown that the level of some bacteria, including segmented filamentous bacteria (SFB), is enriched in the gut of mice in the absence of IgA expression (37). Thus, we performed overall microbiome profiling and species-specific analyses. As shown in Fig. 2D, principal components analyses of overall gene expression of microbiota isolated from the *IncCSR*^{IgA}^{WT/WT} versus *IncCSR*^{IgA}^{-/-} mice clearly showed a difference because the fecal material 16S ribosomal RNA (rRNA) transcriptomes of *IncCSR*^{IgA}^{WT/WT} mice clustered separately from the *IncCSR*^{IgA}^{-/-} mice. These differences were observed in *IncCSR*^{IgA}^{WT/WT} and *IncCSR*^{IgA}^{-/-} littermates that were cohoused. Moreover, microbiota-specific primers were used to analyze the content of various well-established microbiome components colonized in these mice; we observed a clear increase in SFB levels and Enterobacteriaceae levels in the *IncCSR*^{IgA}^{-/-} mice (fig. S3A). Increased SFB and Enterobacteriaceae levels are correlated with inflammation in the gut (35, 38, 39). Thus, we looked for signs of increased inflammation in the gut of the *IncCSR*^{IgA}^{-/-} mice. We find that in *IncCSR*^{IgA}^{-/-} mice, the structure of the intestinal villi is disrupted (fig. S3B, left) and locations of clear inflammation were observed more frequently in intestinal Swiss rolls, compared with what is seen in *IncCSR*^{IgA}^{WT/WT} littermate control cohoused mice (fig. S3B, right).

To evaluate whether, in *IncCSR*^{IgA}^{-/-} mice, the defect in IgA CSR in PP B cells is cell intrinsic and whether CSR to other isotypes is perturbed, we performed cytokine-induced ex vivo CSR assays. We observed that *IncCSR*^{IgA}^{-/-} B cells have a clear CSR defect to IgA and IgG2b isotypes but not to IgG1 and IgG3 (Fig. 3, A and B). This was not due to changed AID expression (Fig. 3D) or altered B cell proliferation (Fig. 3C). Consistent with our

observation with primary B cells, transformed B cell lines that undergo CSR only to IgA (CH12F3) also show a decreased efficiency in CSR from IgM to IgA (fig. S4, A and B) (22). We wanted to investigate whether there is altered SHM frequency in the *IncCSR^{IgA}*^{-/-} mice. For this purpose, we isolated germinal center B220⁺GL7⁺ B cells from the PPs of *IncCSR^{IgA}*^{-/-} and cohoused littermate control *IncCSR^{IgA}*^{AWT/WT} mice [the fluorescence-activated cell sorting (FACS)-sorted population is shown in fig. S4C]. We isolated total PP B cell DNA and performed SHM sequencing (SHM-seq). As a control for the SHM-seq experiment, we also performed SHM-seq from *AID*^{-/-} PP B cells. As shown in fig. S4D, the level of mutation in the PP B cells obtained from the *IncCSR^{IgA}*^{AWT/WT} *JH4* region is higher than that seen in the PP B cells from *IncCSR^{IgA}*^{-/-} mice. The overall mutation frequency at AID hotspots for G to A conversion is lower in the *IncCSR^{IgA}*^{-/-} B cells, in comparison with *IncCSR^{IgA}*^{AWT/WT} cohoused littermate mice (fig. S4E). It is possible that the effect on SHM is due to overall decreased levels of IgA⁺ B cells, which, in littermate control mice, are selectively chosen for SHM.

Organization of TAD^{IncCSR-IgA}

Our efforts turned toward understanding the molecular mechanism(s) underpinning the role of the *IncCSR^{IgA}* locus and the TAD^{IncCSR-IgA} in altering antibody gene diversification in the gut. First, to evaluate whether the interaction of the *IncCSR^{IgA}* locus regulates CSR by interacting with *locus A* and/or *locus B* (see Fig. 1 for details of *IncCSR^{IgA}* interaction with *A* and *B* loci), we generated multiple *locus A* and *locus B* knockout/mutant CH12F3 cell lines. Figure 4A demonstrates the locations of *locus A* and *locus B*, with *locus A* located 128 kb and *locus B* located 250 kb away from *IncCSR^{IgA}*. Both *locus A* and *locus B* have neighboring CTCF binding sites that may have a role in promoting *IncCSR^{IgA}* interaction with *locus A* and *locus B*. Using chromatin immunoprecipitation sequencing (ChIP-seq) data for CTCF and RAD21 (cohesin subunit) DNA binding in B cells, we found that there are CTCF- and RAD21-bound CTCF sites neighboring *IncCSR^{IgA}*, *locus A*, and *locus B*. (Fig. 4A). CTCF sites flank *locus A* (*locus A*^{CTCF}) and *locus B* (*locus B*^{CTCF}), whereas 7 kb further from *IncCSR^{IgA}*, a CTCF (*IncCSR^{IgA}*^{CTCF})– and a RAD21-occupied site exist (Fig. 4A). To characterize the *IncCSR^{IgA}* sequence for regulatory properties, we checked deoxyribonuclease I (DNase I) hypersensitivity, enhancer marks (H3K27Ac and H3K4me1), and MED1 binding. We found that the *IncCSR^{IgA}* locus shows all these marks (as does the 3'RR with which it interacts), establishing its role as a potential transcription enhancer sequence (Fig. 4B).

The interaction of CTCF sites flanking *IncCSR^{IgA}*, *locus A*, and *locus B* sequences regulates TAD dynamics that drive IgA CSR

We next wanted to determine whether *IncCSR^{IgA}* and associated regions physically interact with the 3'RR of the TAD^{IgH}, located about 2.6 Mb away. Using 3C experiments, we were able to show that the CTCF^{IncCSR} (neighboring the *IncCSR^{IgA}* locus) is able to interact with the HS4 region of 3'RR, with deletion of *IncCSR^{IgA}* leading to a decrease in this interaction frequency (fig. S5A). These inter-TAD interaction frequencies are weaker in comparison with intra-TAD interactions but are clearly detectable in our 3C assays (fig. S5A). Overlap of the *IncCSR^{IgA}* regions with published DNA break sites, as observed in END-seq assays, clearly shows that both the *IncCSR^{IgA}* locus and the CTCF^{IncCSR-IgA} accumulate DNA

double-strand breaks in an etoposide treatment-dependent fashion (which prevents repair of Top2B-mediated DNA breaks) (fig. S5B). All these observations point toward a long-range interaction between the TAD^{lncCSRIgA} and the TAD^{IgH} that potentially collaborates or competes with short-range intra-TAD^{lncCSRIgA} interactions.

We wished to understand in greater detail the molecular interactions between *lncCSRIgA*, *locus A*, *locus B*, and 3'RR in the regulation of IgA CSR in the TAD^{lncCSRIgA}. Therefore, we wanted to investigate whether the loss of *lncCSRIgA* (in *lncCSRIgA*^{-/-} B cells) effectively leads to altered interaction of the neighboring CTCF^{lncCSRIgA} site to the locus B region. We generated Hi-C maps of B cells obtained from *lncCSRIgA*^{AWT/WT} and *lncCSRIgA*^{-/-} mice and observed very comparable Hi-C map patterns in the overall B cell genome. Thus, the *lncCSRIgA*^{-/-} mutation does not have a genome-wide defect in B cell genome organization (fig. S5D). Next, we compared Hi-C maps of *lncCSRIgA*^{AWT/WT} and *lncCSRIgA*^{-/-} mice at genome coordinates covering the TAD^{lncCSRIgA} and TAD^{IgH}. We found an increased interaction of the *lncCSRIgA*/CTCF^{lncCSRIgA} region with the B locus in *lncCSRIgA*^{-/-} B cells compared with *lncCSRIgA*^{AWT/WT} B cells (Fig. 4, C to E), the observation being quantitated over multiple independent experiments in Fig. 4E using B cells of littermate *lncCSRIgA*^{AWT/WT} and *lncCSRIgA*^{-/-} mice. These observations indicate that the CTCF neighboring the *lncCSRIgA* locus (CTCF^{lncCSRIgA}) interacts with both the *locus A* and the *locus B* sequences, with the interaction with the *locus B* sequence becoming stronger when the *lncCSRIgA* locus is deleted.

Next, we developed a 4C sequencing (4C-seq) assay with baits immediately adjacent to the *lncCSRIgA* locus (H3 bait) and one next to the CTCF site located 3 to 4 kb downstream (H4 bait). In our 4C-seq assays, we used the H3 bait and learned that the *lncCSRIgA* region physically interacts with the *locus B* region. Consistent with the Hi-C data of Fig. 4 (C to E), this interaction of *lncCSRIgA* and *locus B* is increased in the *lncCSRIgA*^{-/-} B cells (Fig. 4, F and G, top). The interactions between *lncCSRIgA* and *locus B* should be occurring via neighboring CTCF sites [portrayed in Fig. 4 (A and B)]. To interrogate CTCF interactions, the H4 bait (neighboring the CTCF site) was found to bind with the *locus B* sequence more efficiently in *lncCSRIgA*^{-/-} B cells (Fig. 4, F and G, bottom). Both in the Hi-C assays and the 4C-seq assays, the effect of CTCF^{lncCSRIgA} to *locus A* was not detectably altered (fig. S5C, top). In addition, from RNA sequencing (RNA-seq) of PP B cells, we did not observe a marked change in local gene expression overlapping locus A or locus B in *lncCSRIgA*^{-/-} B cells (fig. S5C, bottom). Overall, these observations point toward an interaction of CTCF^{lncCSRIgA} with CTCF^{locusB}, with transcription of the *lncCSRIgA* locus being important for weakening the strength of this interaction in vivo (in *lncCSRIgA*^{-/-} cells).

Effect on CSR in TAD^{lncCSRIgA} mutants

We wanted to understand what the effects of the loss of *locus A*, *locus B*, and CTCF^{lncCSRIgA} on IgA CSR are. *Locus A* and *locus B* are placed 120 and 250 kb from the *lncCSRIgA* locus, respectively (Fig. 5A), allowing us to make deletion mutants using CRISPR-Cas9 gene editing technology. We generated CH12F3 cell lines that lack each of these elements. We performed CSR assays from each of these lines to evaluate the effect on CSR; loss of *locus B* sequence leads to increased IgA CSR (Fig. 5B), whereas loss of *locus*

A sequence leads to decreased IgA CSR (Fig. 5C). These observations were obtained from three independently isolated CH12F3 clones. Individual flow cytometry plots demonstrating the experiments in Fig. 5 (B and C) are shown in fig. S6 (A and B). Lack of *IncCSR^{IgA}* or *locus A* sequence does not ablate CSR but, rather, reduces efficiency. Last, in the absence of CTCF^{IncCSR}, there is a decreased IgA CSR (Fig. 5, D and E), indicating that CTCF^{IncCSR^{IgA}} could be a pivotal site for directing interactions inside the TAD^{IncCSR} so that efficient IgA CSR can occur.

IncRNA-CSR^{IgA} recruits CTCF cofactors to the CTCF^{IncCSR^{IgA}} site and partly regulates IgA CSR

At this point, we wanted to ask whether the lncRNA-CSR^{IgA} (exosome-sensitive transcript from the *IncCSR^{IgA}* locus) is important for orchestrating IgA CSR. Using lentiviral transduction, we introduced lncRNA-CSR^{IgA} in *IncCSR^{IgA}*^{-/-} CH12F3 cells [as control, we separately inserted an inverted sequence to lncRNA-CSR^{IgA} (INV) in CH12F3 cells]. The transduced lncRNA-CSR^{IgA} vector expressed green fluorescent protein (GFP), and these cells were therefore easily identified by flow cytometry. The overexpression of lncRNA-CSR^{IgA} in *IncCSR^{IgA}*^{-/-} cells rescued the CSR defect (Fig. 6, A and B), but neither the vector control alone nor the inverted version of the lncRNA-CSR^{IgA} was able to do so (Fig. 6, A and B). The expression level of the lncRNA-CSR^{IgA} is 10-fold or higher than the endogenous levels of lncRNA-CSR^{IgA} expressed in these cells (Fig. 6C). We then wanted to determine whether the knockdown of the lncRNA-CSR^{IgA} affects CSR. We introduced CRISPR-Cas13a along with two different sets of lncRNA-CSR^{IgA}-directed guide RNAs (gRNAs) into the normal CH12F3 cells. As can be seen in Fig. 6D, both sets of gRNAs successfully knocked down lncRNA-CSR^{IgA} in the CH12F3 cells. The expression of IgSμ germline transcripts and AID expression was not perturbed by the introduction of the Cas13a-based knockdown system, but the expression of lncRNA-CSR^{IgA} was significantly reduced (Fig. 6D). In these conditions, IgA CSR was decreased by 30 to 45% (Fig. 6E). The 30 to 45% decrease of IgA CSR after lncRNA-CSR^{IgA} knockdown was consistent with the loss of the complete DNA locus *IncCSR^{IgA}*, where the effect is slightly stronger (in the range of 50 to 70%). Thus, both through overexpression (Fig. 6, A to C) and knockdown (Fig. 6, D and E) of lncRNA-CSR^{IgA}, we were able to show a role of the RNA moiety expressed in the *IncCSR^{IgA}* DNA locus in influencing CSR.

At this point, we wished to determine the role of the *IncCSR^{IgA}* locus in modulating interaction of CTCF^{IncCSR^{IgA}} with locus A and locus B sequences. For this purpose, we identified *IncCSR^{IgA}*-transcribed lncRNA (lncRNA-CSR^{IgA}) interacting proteins by expressing a stem-loop tagged lncCSR^{IgA} in human embryonic kidney (HEK) 293 T cells and performing biochemical purification of associated proteins in 293 T cells that express MS2 protein (MS2 protein interacts with stem-loops; see schematic in fig. S7A). To distinguish proteins that bind to the MS2 bead matrix or to the stem-loop tag from proteins that bind to the lncRNA-CSR^{IgA} ncRNA bait, we performed controls of purification with expression of the untagged lncCSR^{IgA} (lncRNA-CSR^{IgA} alone) or the MS2 tag fused with a GFP RNA (MS2-GFP-alone), along with the experimental tagged lncRNA-CSR^{IgA} (MS2-GFP-lncCSR^{IgA}). We purified the RNA-protein complex and evaluated interacting proteins using mass spectrometry. This approach provides a tool for evaluating proteins that bind to

ncRNAs independent of transcription and locus positioning. We eliminated nonspecific binding proteins, considering the controls used, and mapped the proteins that specifically interacted with lncRNA-CSR^{IgA} ncRNA. We found that Poly [ADP-Ribose] polymerase I (PARP1), SUPT16H, and structural maintenance of chromosomes (SMC3) are specific interactors with lncRNA-CSR^{IgA} (fig. S7, B and C, and table S4). We evaluated whether these interactions could be confirmed through an immunoprecipitation-coupled Western blot. As can be seen in Fig. 6F, MS2-fused lncCSR^{IgA} interacts with SUPT16H, SMC3, and PARP1 in coimmunoprecipitation assays. SUPT16H is a component of the FACT complex that is important for recruiting epigenetic regulators at promoters and enhancers (40). PARP1 is important for promoting DNA double-strand break formation at DNA regulatory elements (41), and SMC3 is a component of the cohesin complex (42). To ascertain whether, in B cells, these proteins regulate the activity of the CTCF^{lncCSR} site, we performed ChIP with antibodies against CTCF and all three interacting proteins in *lncCSR*^{IgA^{WT/WT} and *lncCSR*^{IgA^{-/-}} B cells (Fig. 6, G and H). We confirmed that in *lncCSR*^{IgA^{-/-}} B cells, there is a modest decrease in CTCF occupancy at the CTCF^{lncCSR} site but no detectable change at CTCF^A or CTCF^B or control V gene-associated CTCF sites (Fig. 6G). We found a significant loss of SMC3, SUPT16H, and PARP1 at CTCF^{lncCSR} in the *lncCSR*^{IgA^{-/-}} B cells (Fig. 6H). Moreover, in cells that have overexpression of lncRNA-CSR^{IgA} (overexpressing cells sorted through the coexpression of GFP), there is a substantial rescue of SUPT16H, SMC3, and PARP1 recruitment at the CTCF^{lncCSR} region (Fig. 6H). Together, these observations indicate that recruitment of CTCF cofactors (FACT subunit SUPT16H, PARP1, and cohesin subunit SMC3) may be provoked upon expression of lncRNA-CSR^{IgA} inside the TAD^{lncCSR} and may be important for modulating CSR. Perhaps, the overexpression of lncRNA-CSR^{IgA} through artificial means overrides its requirement of localized expression inside the TAD^{lncCSR}. When expressed at physiological levels, the lncRNA-CSR^{IgA} should be expressed at close proximity to the CTCF^{lncCSR} site to influence SUPT16H, PARP1, and cohesin recruitment.}

DISCUSSION

Although transcription at enhancer sites and the ensuing enhancer RNA transcripts at various regions of the mammalian genome have been extensively identified and reported on, the requirement of such for gene regulation and organism development is incompletely understood (22, 24, 43–46). In this study, we provide evidence of both a bidirectionally transcribed enhancer locus and an associated ncRNA having roles in controlling TAD interaction dynamics, recruitment of genome organization regulatory factors at critical sites, and antibody gene diversification mechanisms. We show that the *lncCSR*^{IgA} locus that transcribes the ncRNA (lncRNA-CSR^{IgA}) has particular characteristics of an enhancer element that interacts with two different loci: *locus A* and *locus B* inside the TAD^{lncCSR}. Molecular dissection and genome topological analyses of the *lncCSR*^{IgA} locus using genetically manipulated mouse models and mutated cell lines identifies a potential mechanism where interaction of *lncCSR*^{IgA} with *locus B* is reduced after B cell stimulation through an ncRNA transcription-dependent fashion, eventually leading to physiological levels of robust IgA CSR and SHM. The availability of the *lncCSR*^{IgA} locus (free of interaction with *locus B*) may promote its interaction with the IgG 3'RR locus HS4 site (24),

eventually leading to robust IgA and IgG2b CSR [modeled in fig. S7 (D and E)]. All these interactions are facilitated by CTCF sites flanking the *locus A*, *locus B*, and two sites next to *IncCSR^{IgA}*. Ablation of the pivotal CTCF sites, CTCF^{IncCSR^{IgA}}, used in IncCSR^{IgA}/*locus A* and IncCSR^{IgA}/*locus B* interactions, leads to defective IgA CSR. Molecular examination of the lncRNA-CSR^{IgA} ncRNA transcript aligns it with a mechanism of recruitment of cohesion subunit SMC3 protein, FACT subunit SUPT16H, and PARP1 at the pivotal CTCF^{IncCSR^{IgA}} of TAD^{IncCSR^{IgA}} [Fig. 6 and modeled in fig. S7 (D and E)]. All three proteins could be associated with the regulation of CTCF site biology. Because of topological stress at enhancers and CTCF sites, there are etoposide-sensitive DNA double-strand breaks (fig. S5B) (47), and PARP1 may play a role in resolving these DNA breaks. SMC3 is a component of the cohesion complex that interacts with CTCF to regulate loop extrusion (48–50), and the FACT complex (40) often functions to regulate local DNA topology. Future studies will provide additional information regarding the mechanistic aspects of these proteins binding RNA at a CTCF site. At this moment, we do not have tools to completely quantitatively segregate the effect of the IncCSR^{IgA} DNA element and the lncRNA-CSR^{IgA} RNA element on CSR. It is possible that in *IncCSR^{IgA}* *IgA*^{−/−} B cells, the loss of regulation of the CTCF^{IncCSR} site and the decreased efficiency of *IncCSR^{IgA}* enhancer to interact with 3'RR together cause the defective IgA CSR. However, detailed mutational analyses of *IncCSR^{IgA}*, CTCF^{IncCSR}, *locus A*, *locus B*, and the HS4 site of 3'RR will be required to fully understand the underlying mechanisms of the phenotypes that we report in this study. Together, this study provides a rare example of the role of a conserved TAD, an enhancer sequence, and an enhancer-associated ncRNA in regulating intra-TAD interactions that ultimately affect an important biological process (adaptive immunity).

Last, we have identified a regulatory element that expresses a functional ncRNA to control *IgH* recombination in B cells that undergo CSR in the gut. The regulatory element is conserved in humans and contains polymorphisms at a frequency higher than normal controls in IgA-deficient patients (Fig. 1H and fig. S1F). We note that only a subset of patients with IgAD carry the *IncCSR^{IgA}* locus mutant alleles; the genetics and mechanism of induction may well differ in different patients depending on the distribution of many other variant alleles genome-wide. For example, there are three major major histocompatibility complex susceptibility haplotypes seen in patients with IgAD, and these may potentially biologically interact with distinct sets of coding and noncoding sequences, including *IncCSR^{IgA}*. In IncCSR^{IgA} *IgA*^{−/−} mice, IgA expression and SHM efficiency are reduced but not completely abolished in PP B cells. Thus, we think that the inter-TAD and intra-TAD interactions controlled by *IncCSR^{IgA}* locus transcription and lncRNA-CSR^{IgA} expression have a regulatory effect on antibody gene diversification, but other compensatory mechanisms that control IgA CSR exist. The reduced IgA expression and SHM in B cells in the gut were sufficient to alter the homeostatic levels of microbiome components. It will be of interest to characterize other TADs surrounding the *IgH* or *IgL* loci to determine whether different regulatory DNA elements in tissue-specific antibody gene diversification processes can be identified.

This study entices us to question just how general is the possibility that enhancer locus expressed ncRNAs or other ncRNAs cause changes in genome topology by altering CTCF site action. Given our example here of *IncCSR^{IgA}* affecting control of a neighboring CTCF

site, we looked genome-wide to see how often CTCF sites are found next to noncoding RNA-expressing unidirectional or bidirectional elements (fig. S8A). More often than not, in various TAD datasets available (fig. S8B), three-way interacting CTCF sites are neighbored by an ncRNA-expressing element (fig. S8, C and D, and table S5). We postulate that at least in the case of *lncCSR^{IgA}* transcription and regulation of CTCF^{*lncCSR*} residing inside the TAD^{*lncCSR*}, the local concentration of ncRNA is increased to influence biological events and such effects could also be artificially reconstituted by overexpressing the ncRNA from a different region of the genome [as in Fig. 6 (A and B)]. The possibility exists that these ncRNAs provide some secondary structural properties that promote macromolecular complex formation. It is too premature to conclude but it is worth speculating whether ncRNAs control CTCF biology and cofactor interactions genome-wide by localized regulation, but this study does raise this question and opens up avenues for interrogating the role of ncRNA biology in genome topology control.

MATERIALS AND METHODS

Hi-C sequencing, 4c-seq, and RNA-seq genomic datasets

Genomic datasets have been uploaded in the National Center for Biotechnology Information for public access after publication (accession no. PRJNA486392).

Experimental models and protocols

Mouse models—*Exosc3^{COIN/COIN}*, *Rosa26 CreERT2* and *lncCSR^{IgA}*^{−/−} mouse models were generated and bred according to the Institutional Animal Care and Use Committee guidelines at Columbia University, New York.

Cloning and site-directed mutagenesis for generation of stem-loop structure fused lncRNA-CSR^{IgA} construct—To identify and characterize lncRNA-CSR^{IgA} interacting proteins in ex vivo conditions, we prepared *Plasmid-A* lncCSR^{IgA}-MS2 (a chimeric RNA containing lncCSR^{IgA} followed at the 3′ end by 24 copies of the MS2 hairpin). Restriction site Bgl II of the pSP72 cloning vector containing lncCSR^{IgA} was converted into an Spe I site using a protocol reported earlier (51) to accommodate 24× MS2 loops at the 3′ end of lncCSR^{IgA}, with the 24× MS2 loop fragment digested from another vector with Bam HI and Spe I restriction enzymes. *Plasmid-B* hemagglutinin (HA)-pMS2 (pEGFP-1 Clontech vector backbone) expressed the MS2 coat protein containing the N terminus HA-tag and C terminus GFP fusion proteins.

lncRNA-CSR^{IgA} immunoprecipitation—We cotransfected plasmids A and B into HEK-293 T cells using PEI reagents (2). We also used two cotransfection controls, namely, (i) plasmid-A alone and (ii) plasmid-B alone, respectively, in HEK-293 T cells under identical conditions. HEK-293 T cells were harvested 72 hours after transfection, washed twice with cold phosphate-buffered saline (PBS), pelleted and lysed with 500 μl per 25 million to 30 million cells of lysis buffer [20 mM tris-HCl (pH 7.5), 100 mM KCl, 5 mM MgCl₂, 5% glycerol, 0.5% NP-40, 1× protease inhibitors (Roche), RiboLock RNase (ribonuclease) inhibitor (Fermentas; 40 U/ml), and 1 mM dithiothreitol (DTT)]. After a 30-min incubation on ice, cells were slowly homogenized using a glass douncer (300 to 500

times) to achieve at least 85 to 95% cell lysis and subsequently pelleted at 20,000g for 20 min at 4°C. The protein lysate was collected and incubated with equilibrated anti-HA agarose and Protein G Sepharose beads. Cell lysate along with beads was incubated overnight at 4°C, with rotation.

Subsequently, beads were pelleted and washed thrice using wash buffer [20 mM tris-HCl (pH 7.5), 100 mM KCl, 5 mM MgCl₂, 0.01% NP-40, 1× protease inhibitors (Roche), RiboLock RNase inhibitor (Fermentas; 40 U/ml), and 1 mM DTT] for 15 min at 4°C and pelleted at 500g for 5 min. After pulldown and washes, beads were divided into two parts: One-half was used for protein analysis; the other half was used for the isolation of lncRNA-CSR^{IgA}. For RNA analysis, the pulldown material was digested with DNase and proteinase, and the RNA was precipitated with sodium acetate/ethanol for further analysis. Pulldown enriched with lncRNA-CSR^{IgA} was detected by quantitative reverse transcription polymerase chain reaction (qRT-PCR) amplification. For protein experiments, lncRNA-CSR^{IgA} was eluted using 250 µl of HA-peptide (Thermo Fisher Scientific, catalog no. 26184; 1 mg/ml) per elution based on principles originally described (52). Supernatant containing the lncRNA-CSR^{IgA} and its interacting proteins was collected. The protein complexes recovered from purification were fractionated on a 4 to 20% SDS–polyacrylamide gel electrophoresis (SDS-PAGE) gradient gel. Unknown proteins were directly excised from the gel, and mass spectrometry was performed.

Western blot analysis for lncRNA-CSR^{IgA} interacting proteins—The protein complexes recovered from lncRNA-CSR^{IgA} immunoprecipitation were fractionated on a 4 to 20% SDS-PAGE gradient gel and transferred to polyvinylidene difluoride membrane. Immunoprecipitation products were analyzed using anti-SUPT16H (Abcam, 204303), PARP1 (Active Motif, 39559), and SMC3 (Abcam, 9263).

lncCSR^{IgA} mouse generation—With CRISPR-Cas9 assistance, small homology arms were designed for gene targeting. A colony PCR approach was used to amplify both the 5′ and 3′ homology arms (with a length of about 1.0 and 0.95 kb) from a positively identified C57BL/6 BAC clone (RP23: 209D11). The FRT-flanked neomycin cassette replaced 1121 base pairs (bp) of the genomic sequence encoding the lncCSR-IgA. The targeting vector was confirmed by restriction analysis after each modification step and by sequencing using primers inside the selection cassette into the 3′ end of the 5′ homology arm and likewise into the 5′ end of the 3′ arm. Primers in the vector sequence were also used to read into the 5′ and 3′ homology arms. Ten micrograms of the targeting vector was linearized and then transfected by electroporation into iTL C57Bl6 (IC1) embryonic stem (ES) cells. After selection with G418 antibiotic, surviving clones were expanded for PCR analysis to identify recombinant ES clones. Expanded clones were reconfirmed by PCR, and then, secondary confirmation of positive clones identified by PCR was performed by Southern blot analysis. Briefly, DNA was digested with Afl II, separated on an agarose gel, and transferred to nylon membrane. The digested DNA was hybridized with a probe targeted against the 3′ external region. Positive clones were further confirmed by Southern blot analysis using an internal probe that involved digestion of the DNA with Bam HI, electrophoretic separation on an agarose gel, transfer to nylon membrane, and hybridization with a probe targeted against the

5' internal region. A primary and a backup clone were identified and injected. Recombinant ES cells were injected into BALB/c blastocysts, and chimerism was determined visually by the distribution of the black coat color on the pups. Four chimeras were identified and mated with FLPe deleter mice to excise the neomycin cassette, and the resulting black pups from this mating were genotyped. These somatic neo deleted mice were then mated with wild-type C57BL/6 mice. Positive progeny were termed germline neo deleted (GND) mice and were also confirmed for absence of the FLP transgene and bred further for experiments. All generations of these GND mice were confirmed by PCR genotyping using the following primers: 28647 Seq For and 286473crv with 36 cycles of PCR performed at 94°30", 60°30", and 72°1'30" using Ex Taq. Animals with lncCSR-IgA intact generate a PCR product of 1677 bp, whereas animals with lncCSR-IgA deleted and only an Frt site remaining will generate a PCR fragment of 635 bp. All mouse experiments were performed in adherence with protocols approved by the Columbia University Institutional Animal Care and Use Committee.

Chromatin immunoprecipitation—A total of 3×10^6 CH12 lncCSR^{IgAWT/WT} or CH12 lncCSR^{IgA^{-/-}} cells or primary mouse lncCSR^{IgAWT/WT} or lncCSR^{IgA^{-/-}} B cells were isolated and cultured as required and subjected to addition of formaldehyde [27 µl/ml of cells of a 37% solution (Sigma-Aldrich, F8775; 500 ml)] for 10 min with gentle inversion/rotation before the addition of 100 µl/ml of 1.375 M glycine to quench the formaldehyde with further gentle inversion/rotation for 10 min. Cells were washed several times with PBS before being suspended in SDS lysis buffer with protease inhibitors [1% SDS, 10 mM EDTA, and 50 mM tris plus 1 "complete mini" EDTA-free protease inhibitor solution tablet (Roche) per 10 ml of lysis buffer]. Sonication was performed on ice using a Branson Sonifier 250 apparatus for 25 cycles, each cycle comprising 20 s of sonication at duty cycle 30%, followed by a 2-min rest period. After gel electrophoresis to gauge approximate amounts of material, the lysate was centrifuged at maximum speed for 10 min at 4°C, and 2% of the material was snap-frozen as input. Lysates were diluted appropriately and precleared with protein A/G plus agarose solution and ChIP-grade rabbit IgG for 1 hour. Immunoprecipitation of lysates was performed overnight at 4°C with gentle rotation using indicated antibodies (anti-SMC3, anti-PARP1, and anti-SUPT16H). Sonicated herring sperm DNA (Promega, D1815) and protein A/G plus agarose solution were added for 90 min with continued rotation. Subsequently, the beads were washed by the standard series of washes [low salt, high salt, LiCl, and Tris-EDTA (TE)], and ChIP products were eluted, followed by RNase A treatment overnight at 60°C and proteinase K treatment for 12 hours at 37°C and followed by 2 hours at 55°C. ChIP DNA was purified by phenol-chloroform-isoamyl alcohol addition (and recovered using phase-lock gel heavy tubes) and then precipitated using ethanol. ChIP products were evaluated with appropriate primers that were specific for the regions interrogated (e.g., CTCF^{lncCSRIgA}).

CSR protocol—CH12F3 cells were plated at a concentration of 5×10^4 cells in 1 ml of appropriate medium [RPMI supplemented with 10% fetal bovine serum (FBS) and nonessential amino acids, sodium pyruvate, L-glutamine, Hepes, penicillin-streptomycin, and β-mercaptoethanol] and placed in one well of a cluster 12-well tissue culture plate (Corning Costar, 3512). Cells were plated in triplicate to reduce variation. Cells receiving

stimulation were administered lipopolysaccharide (LPS), interleukin 4 (IL4), and transforming growth factor- β (TGF- β) (20 μ g/ml, 20 ng/ml, and 1 ng/ml, respectively). All cells were placed in a 37°C tissue culture incubator with 5% CO₂. At the time of examination, 200 μ l of cells was removed from the well, washed with PBS/2% FBS, spun down, and resuspended in 100 μ l of 2% FBS in PBS containing 0.1 μ l of anti-IgA-phycoerythrin (PE) antibody. Cells were incubated with the antibody at 4°C for 30 min before being centrifuged and resuspended in 2% FBS/PBS and interrogated on an LSRFortessa (BD Biosciences) FACS machine.

CSR protocol of splenic B cells for Ig isotypes CSR—CD43-negative B cells from spleen were stimulated with LPS (20 μ g/ml) (Sigma-Aldrich) for IgG3; LPS (20 μ g/ml) and IL4 (20 ng/ml) (PeproTech) for IgG1; and LPS (10 μ g/ml) (Sigma-Aldrich), TGF- β (2 ng/ml) (R&D Systems), and anti-IgD dextran (0.33 μ g/ml) (Fina Biosolutions) for IgG2b and IgA CSR. CSR efficiency of each isotype was measured after 3 to 4 days. B220-PE (BioLegend)/IgG3-FITC (fluorescein isothiocyanate) (BD Biosciences) and B220-FITC (BD Biosciences)/IgG2b-PE (BioLegend), IgA-PE (eBioscience), or IgG1-PE (BD Biosciences) were used for staining cells for FACS analyses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank T. Honjo (Kyoto University) for providing CH12-F3 cells and AID^{-/-} mice, B. R. San Martin (IGCMB, Strasbourg) for ChIP-seq datasets, E. L. Aiden and A. Presser (Baylor College of Medicine) for advice regarding Hi-C data, and members of the Basu laboratory for valuable discussions; the Columbia Genome Center for high-throughput genome sequencing and the Department of Microbiology and Immunology core FACS facility as well as the Flow Cytometry Core, Columbia Stem Cell facility; and the Ingenious Targeting Laboratory for mouse model generation.

Funding: This work was supported by grants to U.B. [NIAID (1R01AI099195 and Ro1AI134988), the Leukemia and Lymphoma Society, and the Pershing Square Sohn Cancer Research Alliance], B.L. (EMBO fellowship ALTF 906–2015), and L.N. (NIH grant T32 AI106711).

REFERENCES AND NOTES

1. Schatz DG, Swanson PC, V(D)J recombination: Mechanisms of initiation. *Annu. Rev. Genet* 45, 167–202 (2011). [PubMed: 21854230]
2. Bednarski JJ, Sleckman BP, At the intersection of DNA damage and immune responses. *Nat. Rev. Immunol* 19, 231–242 (2019). [PubMed: 30778174]
3. Bossen C, Mansson R, Murre C, Chromatin topology and the regulation of antigen receptor assembly. *Annu. Rev. Immunol* 30, 337–356 (2012). [PubMed: 22224771]
4. Methot SP, Di Noia JM, Molecular mechanisms of somatic hypermutation and class switch recombination. *Adv. Immunol* 133, 37–87 (2017). [PubMed: 28215280]
5. Keim C, Kazadi D, Rothschild G, Basu U, Regulation of AID, the B-cell genome mutator. *Genes Dev.* 27, 1–17 (2013). [PubMed: 23307864]
6. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B, Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* 152, 417–429 (2013). [PubMed: 23374339]
7. Dekker J, Mirny L, The 3D genome as moderator of chromosomal communication. *Cell* 164, 1110–1121 (2016). [PubMed: 26967279]

8. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680 (2014). [PubMed: 25497547]
9. Vian L, P kowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, Huang SC, el Khattabi L, Dose M, Pruett N, Sanborn AL, Canela A, Maman Y, Oksanen A, Resch W, Li X, Lee B, Kovalchuk AL, Tang Z, Nelson S, di Pierro M, Cheng RR, Machol I, St Hilaire BG, Durand NC, Shamim MS, Stamenova EK, Onuchic JN, Ruan Y, Nussenzweig A, Levens D, Aiden EL, Casellas R, The energetics and physiological impact of cohesin extrusion. *Cell* 173, 1165–1178.e20 (2018). [PubMed: 29706548]
10. Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng H-L, Hansen E, Despo O, Bossen C, Vettermann C, Bates JG, Richards N, Myers D, Patel H, Gallagher M, Schlissel MS, Murre C, Busslinger M, Giallourakis CC, Alt FW, CTCF-binding elements mediate control of V(D)J recombination. *Nature* 477, 424–430 (2011). [PubMed: 21909113]
11. Birshtein BK, Epigenetic regulation of individual modules of the immunoglobulin heavy chain locus 3' regulatory region. *Front. Immunol* 5, 163 (2014). [PubMed: 24795714]
12. Birshtein BK, The role of CTCF binding sites in the 3' immunoglobulin heavy chain regulatory region. *Front. Genet* 3, 251 (2012). [PubMed: 23162572]
13. Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, Loukinov D, Eckhardt LA, Lobanenkov VV, Birshtein BK, Chromatin architecture near a potential 3' end of the igh locus involves modular regulation of histone modifications during B-cell development and in vivo occupancy at CTCF sites. *Mol. Cell. Biol* 25, 1511–1525 (2005). [PubMed: 15684400]
14. Volpi SA, Verma-Gaur J, Hassan R, Ju Z, Roa S, Chatterjee S, Werling U, Hou H Jr., Will B, Steidl U, Scharff M, Edelman W, Feeney AJ, Birshtein BK, Germline deletion of Igh 3' regulatory region elements hs 5, 6, 7 (hs5–7) affects B cell-specific regulation, rearrangement, and insulation of the Igh locus. *J. Immunol* 188, 2556–2566 (2012). [PubMed: 22345664]
15. Pinaud E, Marquet M, Fiancette R, Péron S, Vincent-Fabert C, Denizot Y, Cogné M, The IgH locus 3' regulatory region: Pulling the strings from behind. *Adv. Immunol* 110, 27–70 (2011). [PubMed: 21762815]
16. Rouaud P, Vincent-Fabert C, Saintamand A, Fiancette R, Marquet M, Robert I, Reina-San-Martin B, Pinaud E, Cogné M, Denizot Y, The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J. Exp. Med* 210, 1501–1507 (2013). [PubMed: 23825188]
17. Beagrie RA, Scialdone A, Schueler M, D. C. A., Chotalia M, Xie SQ, Barbieri M, de Santiago I, Lavitas L-M, Branco MR, Fraser J, Dostie J, Game L, Dillon N, Edwards PAW, Nicodemi M, Pombo A, Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543, 519–524 (2017). [PubMed: 28273065]
18. Will AJ, Cova G, Osterwalder M, Chan W-L, Wittler L, Brieske N, Heinrich V, de Villartay J-P, Vingron M, Klopocki E, Visel A, Lupiáñez DG, Mundlos S, Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nat. Genet* 49, 1539–1545 (2017). [PubMed: 28846100]
19. Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, Jung I, Wu H, Zhai Y, Tang Y, Lu Y, Wu Y, Jia Z, Li W, Zhang MQ, Ren B, Krainer AR, Maniatis T, Wu Q, CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162, 900–910 (2015). [PubMed: 26276636]
20. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B, Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128, 1231–1245 (2007). [PubMed: 17382889]
21. Gutzeit C, Magri G, Cerutti A, Intestinal IgA production and its role in host-microbe interaction. *Immunol. Rev* 260, 76–85 (2014). [PubMed: 24942683]
22. Pefanis E, Wang J, Rothschild G, Lim J, Kazadi D, Sun J, Federation A, Chao J, Elliott O, Liu ZP, Economides AN, Bradner JE, Rabadan R, Basu U, RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* 161, 774–789 (2015). [PubMed: 25957685]
23. Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, Economides AN, Basu U, Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* 514, 389–393 (2014). [PubMed: 25119026]

24. Lim J, Giri PK, Kazadi D, Laffleur B, Zhang W, Grinstein V, Pefanis E, Brown LM, Ladewig E, Martin O, Chen Y, Rabadan R, Boyer F, Rothschild G, Cogné M, Pinaud E, Deng H, Basu U, Nuclear proximity of Mtr4 to RNA exosome restricts DNA mutational asymmetry. *Cell* 169, 523–537.e15 (2017). [PubMed: 28431250]
25. Belair C, Sim S, Kim KY, Tanaka Y, Park IH, Wolin SL, The RNA exosome nuclease complex regulates human embryonic stem cell differentiation. *J. Cell Biol* 218, 2564–2582 (2019). [PubMed: 31308215]
26. Lloret-Llinares M, Karadoulama E, Chen Y, Wojenski LA, Villafano GJ, Bornholdt J, Andersson R, Core L, Sandelin A, Jensen TH, The RNA exosome contributes to gene expression regulation during stem cell differentiation. *Nucleic Acids Res.* 46, 11502–11513 (2018). [PubMed: 30212902]
27. Chiu AC, Suzuki HI, Wu X, Mahat DB, Kriz AJ, Sharp PA, Transcriptional pause sites delineate stable nucleosome-associated premature polyadenylation suppressed by U1 snRNP. *Mol. Cell* 69, 648–663.e7 (2018). [PubMed: 29398447]
28. Wahba L, Amon JD, Koshland D, Vuica-Ross M, RNase H and multiple RNA biogenesis factors cooperate to prevent RNA: DNA hybrids from generating genome instability. *Mol. Cell* 44, 978–988 (2011). [PubMed: 22195970]
29. Domingo-Prim J, Endara-Coll M, Bonath F, Jimeno S, Prados-Carvajal R, Friedländer MR, Huertas P, Visa N, EXOSC10 is required for RPA assembly and controlled DNA end resection at DNA double-strand breaks. *Nat. Commun* 10, 2135 (2019). [PubMed: 31086179]
30. Kilchert C, Wittmann S, Vasiljeva L, The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol* 17, 227–239 (2016). [PubMed: 26726035]
31. Sepulveda MA, Garrett FE, Price-Whelan A, Birshtein BK, Comparative analysis of human and mouse 3' Igh regulatory regions identifies distinctive structural features. *Mol. Immunol* 42, 605–615 (2005). [PubMed: 15607820]
32. Bronson PG, Chang D, Bhangale T, Seldin MF, Ortmann W, Ferreira RC, Urcelay E, Pereira LF, Martin J, Plebani A, Lougaris V, Friman V, Freiburger T, Litzman J, Thon V, Pan-Hammarström Q, Hammarström L, Graham RR, Behrens TW, Common variants at PVT1, ATG13–AMBRA1, AHI1 and CLEC16A are associated with selective IgA deficiency. *Nat. Genet* 48, 1425–1429 (2016). [PubMed: 27723758]
33. Longman RS, Yang Y, Diehl GE, Kim SV, Littman DR, Microbiota: Host interactions in mucosal homeostasis and systemic autoimmunity. *Cold Spring Harb. Symp. Quant. Biol* 78, 193–201 (2013). [PubMed: 24913313]
34. Cho H, Jaime H, de Oliveira RP, Kang B, Spolski R, Vaziri T, Myers TG, Thovarai V, Shen Z, Fox JG, Leonard WJ, Kelsall BL, Defective IgA response to atypical intestinal commensals in IL-21 receptor deficiency reshapes immune cell homeostasis and mucosal immunity. *Mucosal Immunol.* 12, 85–96 (2019). [PubMed: 30087442]
35. Kurashima Y, Kiyono H, Mucosal ecological network of epithelium and immune cells for gut homeostasis and tissue healing. *Annu. Rev. Immunol* 35, 119–147 (2017). [PubMed: 28125357]
36. Fagarasan S, Muramatsu M, Suzuki K, Nagaoka H, Hiai H, Honjo T, Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora. *Science* 298, 1424–1427 (2002). [PubMed: 12434060]
37. Suzuki K, Meek B, Doi Y, Muramatsu M, Chiba T, Honjo T, Fagarasan S, Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. *Proc. Natl. Acad. Sci. U.S.A* 101, 1981–1986 (2004). [PubMed: 14766966]
38. Fadlallah J, El Kafsi H, Sterlin D, Juste C, Parizot C, Dorgham K, Autaa G, Gouas D, Almeida M, Lepage P, Pons N, Le Chatelier E, Levenez F, Kennedy S, Galleron N, de Barros J-PP, Malphettes M, Galicier L, Boutboul D, Mathian A, Miyara M, Oksenhendler E, Amoura Z, Doré J, Fieschi C, Ehrlich SD, Larsen M, Gorochov G, Microbial ecology perturbation in human IgA deficiency. *Sci. Transl. Med* 10, eaan1217 (2018).
39. Macpherson AJ, Yilmaz B, Limenitakis JP, Ganai-Vonarburg SC, IgA function in relation to the intestinal microbiota. *Annu. Rev. Immunol* 36, 359–381 (2018). [PubMed: 29400985]
40. Formosa T, The role of FACT in making and breaking nucleosomes. *Biochim. Biophys. Acta* 1819, 247–255 (2013). [PubMed: 24459727]

41. Ju B-G, Rosenfeld MG, A breaking strategy for topoisomerase II β /PARP-1-dependent regulated transcription. *Cell Cycle* 5, 2557–2560 (2006). [PubMed: 17106262]
42. Uhlmann F, SMC complexes: From DNA to chromosomes. *Nat. Rev. Mol. Cell Biol* 17, 399–412 (2016). [PubMed: 27075410]
43. Wu X, Sharp PA, Divergent transcription: A driving force for new gene origination? *Cell* 155, 990–996 (2013). [PubMed: 24267885]
44. Rothschild G, Basu U, Lingering questions about enhancer RNA and enhancer transcription-coupled genomic instability. *Trends Genet.* 33, 143–154 (2017). [PubMed: 28087167]
45. Li W, Notani D, Rosenfeld MG, Enhancers as non-coding RNA transcription units: Recent insights and future perspectives. *Nat. Rev. Genet* 17, 207–223 (2016). [PubMed: 26948815]
46. Kim T-K, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME, Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187 (2010). [PubMed: 20393465]
47. Canela A, Maman Y, Jung S, Wong N, Callen E, Day A, Kieffer-Kwon KR, Pekowska A, Zhang H, Rao SSP, Huang SC, Mckinnon PJ, Aplan PD, Pommier Y, Aiden EL, Casellas R, Nussenzweig A, Genome organization drives chromosome fragility. *Cell* 170, 507–521.e18 (2017). [PubMed: 28735753]
48. Srinivasan M, Scheinost JC, Petela NJ, Gligoris TG, Wissler M, Ogushi S, Collier JE, Voulgaris M, Kurze A, Chan KL, Hu B, Costanzo V, Nasmyth KA, The cohesin ring uses its hinge to organize DNA using non-topological as well as topological mechanisms. *Cell* 173, 1508–1519.e18 (2018). [PubMed: 29754816]
49. Nasmyth K, Cohesin: A catenase with separate entry and exit gates? *Nat. Cell Biol* 13, 1170–1177 (2011). [PubMed: 21968990]
50. Rowley MJ, Corces VG, Organizational principles of 3D genome architecture. *Nat. Rev. Genet* 19, 789–800 (2018). [PubMed: 30367165]
51. Giri PK, Tang X, Thangamani S, Shenoy RT, Ding JL, Swaminathan K, Sivaraman J, Modifying the substrate specificity of Carcinoscorpis rotundicauda serine protease inhibitor domain 1 to target thrombin. *PLOS ONE* 5, e15258 (2010).
52. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Séraphin B, A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol* 17, 1030–1032 (1999). [PubMed: 10504710]
53. van de Werken HJ, de Vree PJ, Splinter E, Holwerda SJ, Klous P, de Wit E, de Laat W, 4C technology: Protocols and data analysis. *Methods Enzymol.* 513, 89–112 (2012). [PubMed: 22929766]
54. Didion JP, Martin M, Collins FS, Atropos: Specific, sensitive, and speedy trimming of sequencing reads. *PeerJ* 5, e3720 (2017). [PubMed: 28875074]
55. Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* 3, 95–98 (2016). [PubMed: 27467249]
56. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293 (2009). [PubMed: 19815776]
57. Langmead B, Salzberg SL, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359 (2012). [PubMed: 22388286]
58. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP, Integrative genomics viewer. *Nat. Biotechnol* 29, 24–26 (2011). [PubMed: 21221095]
59. Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J, Meyer BJ, Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* 523, 240–244 (2015). [PubMed: 26030525]
60. Ye BH, Cattoretti G, Shen Q, Zhang J, Hawe N, de Waard R, Leung C, Nouri-Shirazi M, Orazi A, Chaganti RSK, Rothman P, Stall AM, Pandolfi P-P, Dalla-Favera R, The BCL-6 proto-oncogene

controls germinal-centre formation and Th2-type inflammation. *Nat. Genet* 16, 161–170 (1997). [PubMed: 9171827]

61. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T, Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553–563 (2000). [PubMed: 11007474]

Fig. 1. *IncCSR^{IgA}* exists in a TAD containing two additional interactors (loci A and B). (A) A linear map of the location of the IgH, specifically the IgM and IgG1 regions. The *IncCSR^{IgA}* region, locus A, and locus B are transcribed regions that express RNA exosome substrate ncRNAs. (B) RNA-seq used for identification of bidirectionally transcribed *IncCSR^{IgA}*, locus A, and locus B in (A). The expressed ncRNAs are identified in RNA exosome-proficient wild-type (WT) B cells (*Exosc3^{WT/WT}*) and RNA exosome-deficient B cells (*Exosc3^{COIN/COIN}*). (C) Map of the TAD *IncCSR^{IgA}* separated from the IgH constant region genes; the 3' RR forms a TAD with intronic enhancer E μ . (D and E) Visualizations of

TAD^{lncCSR^{IgA}} at different resolutions. The green arrowhead indicates *locus A*; the orange arrowhead indicates *lncCSR^{IgA}*, and the blue arrowhead indicates *locus B*. **(F)** The TAD^{lncCSR^{IgA}} is identified in both CH12F3 cells and primary B cells, with primary B cell TAD on the left and CH12F3 cell TAD on the right. The black square indicates interacting area of *lncCSR^{IgA}* and *locus A*, and the blue square indicates interacting area of *lncCSR^{IgA}* and *locus B*. **(G)** Interactions between *lncCSR^{IgA}*, *locus A*, and *locus B* are conserved in mouse (left-hand side) and human (right-hand side) B cell-derived cell lines. The regions mm9 chr12:111,615,000–112,170,000 of the CH12-LX cell line and hg19 chr14:102,100,000–102,900,000 of the GM12878 cell line Hi-C contact matrix were visualized in Juicebox. **(H)** Rare variants [single-nucleotide variant (SNV) and InDels] in selective IgA-deficient human patients align in remarkable congruence with locus *lncCSR^{IgA}*. Ten patients with sIgAD having rare variants observed on both alleles of the *lncCSR^{IgA}* locus are shown; the frequency of these rare variants is two- to fourfold higher than the local population (Sweden) and is shown to be significant ($^{+}P < 0.05$).

Fig. 2. *IncCSR*^{IgA} knockout B cells have specific attenuated IgA CSR in PPs and physiological consequences.

(A) A mouse model was generated in which the *IncCSR*^{IgA} locus was ablated [targeting strategy shown in (A) and fig. S2A]. (B) B cells isolated from the PPs of these mice show a significant decrease in CSR efficiency ($P = 0.01$, $n = 5$ littermate pairs). Left: IgA CSR in B cells isolated from PPs of *IncCSR*^{IgA^{WT/WT} and *IncCSR*^{IgA^{-/-} mice, along with CSR analyzed in PP B cells from AID^{-/-} mice. Right: Data quantitation, where $**P = 0.01$, by Student's *t* test. (C) Immunohistochemical staining for IgA in the small intestine of *IncCSR*^{IgA^{WT/WT} and *IncCSR*^{IgA^{-/-} mice; staining was performed from sections showing PP}}}}

(top) and intestinal villi (bottom). **(D)** A principal components analysis of the metagenomics of the intestinal bacterial flora of *lncCSR*^{Ig^{AWT}/WT, *lncCSR*^{Ig^A-/-}, and *AID*^{-/-} mice by high-throughput sequencing of 16S rRNA. Note that *lncCSR*^{Ig^{AWT}/WT}, *lncCSR*^{Ig^A-/-}, and *AID*^{-/-} metagenomics data cosegregate.}

Fig. 3. *IncCSR^{IgA}* knockout B cells have decreased IgA and IgG2b CSR defects in ex vivo cytokine-stimulated cultures.

(A) Purified splenocytes from *IncCSR^{IgA}AWT/WT*, *IncCSR^{IgA}-/-*, and *AID^{-/-}* mice were cultured for 72 hours in media with components to promote expression of the different isotypes indicated. IgA expression in the *IncCSR^{IgA}-/-* mice is specifically attenuated (along with partnering IgG2b). (B) Quantitation of data presented in (A). ns, not significant; ***P* 0.01 and ****P* 0.001, by Student's *t* test (C) Cell division of *IncCSR^{IgA}AWT/WT*, *IncCSR^{IgA}-/-*, and *AID^{-/-}* B cells grown with cytokine stimulation in ex vivo culture

conditions measured by VPD450 dye dilution method. **(D)** Failure to undergo CSR in the *IncCSR^{IgA-/-}* mice compared with the *IncCSR^{IgAWT/WT}* mice is not owed to a diminution of AID protein expression as shown by Western blot against AID.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4. CTCF binding sites neighboring locus A and *IncCSR*^{lgA} interact to promote CSR, whereas sites neighboring locus B and *IncCSR*^{lgA} interact to suppress CSR.

(A) The expression of ncRNAs from *Exosc3*^{WT/WT} and *Exosc3*^{COIN/COIN} cells and the presence of CTCF, RAD21 (cohesin), and MED1 binding. (B) Detection of enhancer marks H3K4me1 and H3K27Ac overlapping the *IncCSR*^{lgA} locus and CTCF, MED1, and RAD21 binding. (C and D) (D) is a zoom-in of (C) showing the enhanced CTCF^{locusB}-CTCF^{IncCSR} interaction (black arrow) in B cells obtained from *IncCSR*^{lgA} mice in comparison with those obtained from *IncCSR*^{lgAWT/WT} mice. (E) A quantitative estimate of CTCF^{locusB}-

CTCF^{lncCSR} interaction of (C) from three biological repeats. (F) Quantitation of the results of CTCF^{locusB}-CTCF^{lncCSR} interactions with 4C analyses with baits placed immediately outside of lncCSR^{IgA} (H3) or next to CTCF^{lncCSR}IgA (H4 bait). (G) Quantitation of 4C-seq results from (F).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 5. Characterization of various TAD^{IncCSR} mutants for CSR.

(A) A cartoon demonstrating the locations of *locus A*, *IncCSR^{IgA}*, and *locus B* relative to the 3'RR of the *IgH* locus. (B) CSR efficiency increases in locus B mutants in comparison with WT (CH12F3) controls, but this increase is lost in *IncCSR^{IgA}-/-*, *locus B^{-/-}* double mutants. Three individual clones are shown (***P* = 0.01, by Student's *t* test). (C) CSR to IgA is decreased in CH12F3 cells that are deleted of locus A (***P* = 0.01, and ****P* = 0.001, by Student's *t* test); three independent clones are portrayed. (D and E) CRISPR-Cas9-mediated editing of CTCF^{IncCSR^{IgA}} site leads to decreased CSR in CH12F3 cells. (D) FACS plots of

one experiment. (E) Representation of four separate biological repeats of (D). The experiment was done with four independently derived clones ($n = 4$; each clone was assayed three times).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 6. The noncoding enhancer RNA, lncRNA CSR^{IgA}, regulates the pivotal function of CTCF^{lncCSR}.

(A) IgA CSR deficiency in *lncCSR*^{IgA-/-} CH12F3 cells can be rescued by overexpression of lentivirally transduced lncRNA-CSR^{IgA}. FACS plots of cytokine-stimulated CH12F3 B cells that are parental clones (CH12WT), CH12/*lncCSR*^{IgA-/-}, and rescued either by an inverted sequence of lncCSR (lncRNA-CSRInv) or by the physiological lncRNA-CSR^{IgA} (representative of three independent clones) are shown. (B) Plot summarizing multiple experiments that evaluate CSR rescue of *lncCSR*^{IgA-/-} CH12 cells transduced with lncCSR-

RNA^{IgA}. The number of independently performed experiments is shown in the figure by black circles. ****** $P = 0.01$ and ******* $P = 0.001$, by Student's t test. **(C)** Quantitation by qRT-PCR of *lncCSR^{IgA}* expression in *lncCSR^{IgA}*^{-/-} cells after lentiviral infection of the *lncCSR^{IgA}*^{-/-} cells. Note the logarithmic scale on the vertical axis. **(D)** Degradation of lncRNA-CSR^{IgA} with a Cas13a/gRNA approach reduces IgA CSR efficiency. gRNAs specific to the lncRNA-CSR^{IgA} were constructed to specifically knock down the lncRNA-CSR^{IgA} (right) but did not alter the AID expression levels (left) or IgS μ transcription (center). ****** $P = 0.01$ and ******* $P = 0.001$, by Student's t test **(E)** Decreased IgA CSR in CH12F3 cells that express two different gRNA pairs that knocked down lncRNA-CSR^{IgA} expression: FACS plots (left), percent CSR decrease (middle), and ratio of CSR decrease in comparison with Cas13 with empty gRNA vector control (right). **(F)** Coimmunoprecipitation of a stem-loop-associated lncRNA-CSR^{IgA} after expression in 293 T cells validates interaction with SMC3 (top blot), PARP1 (middle blot), and SUPT16H (bottom blot). **(G)** CTCF protein association to CTCF binding sites flanking locus A (CTCF^{LocusA}), lncCSR^{IgA} (CTCF^{lncCSR^{IgA}}), and locus B (CTCF^{LocusB}) was analyzed by ChIP-seq experiments from B cells isolated from *lncCSR^{IgA}*^{WT/WT} and *lncCSR^{IgA}*^{-/-} littermate mice. **(H)** ChIP experiment to confirm that the recruitment of lncRNA-CSR^{IgA} interacting factors (identified by mass spectrometry) SMC3, SUPT16H, and PARP1 to CTCF^{lncCSR^{IgA}} is strongly decreased. Reintroduction of lncRNA-CSR^{IgA} in the *lncCSR^{IgA}*^{-/-} CH12F3 cells partly rescues the recruitment of the three factors to CTCF^{lncCSR^{IgA}}.