

GENERAL ARTICLE

Skewed allelic expression on X chromosome associated with aberrant expression of XIST on systemic lupus erythematosus lymphocytes

Yanfeng Zhang^{1,†}, Xinrui Li², Andrew Gibson², Jeffrey Edberg², Robert P. Kimberly² and Devin M. Absher^{1,*}

¹HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA and ²Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA

*To whom correspondence should be addressed at: 601 Genome Way, Huntsville, AL 35806, USA. Tel: +1 2563270649; Fax: +1 2563279833; Email: dabsher@hudsonalpha.org

Abstract

A common feature of autoimmune diseases, including systemic lupus erythematosus (SLE), is an increased prevalence in women. However, the molecular basis for sex disparity in SLE remains poorly understood. To examine the role of X-linked transcription in SLE adaptive immune cells, we performed RNA-seq in T cell and B cell subsets from either healthy donors or patients with SLE. Analyses of allelic expression (AE) profiles identified a pattern of increased allelic imbalance across the entire X chromosome in SLE lymphocytes. X-linked genes exhibiting AE in SLE had an extensive overlap with genes known to escape X chromosome inactivation (XCI). XIST RNA was overexpressed in SLE patients. Differential XIST expression correlated with AE profiles more positively at X-linked genes than the genome-wide background. Analysis of three independent RNA-seq data verified the XIST-associated skewed AE on X chromosome in SLE. Integrative analyses of DNA methylation profiles showed an increased variability of DNA methylation levels at these AE-related X-linked genes. In cultured lymphoblastic cells, knockdown of XIST specifically altered allelic imbalance patterns between X chromosomes. Our study provides genetic evidence that upregulation of XIST accompanied with more skewed allelic expression on X chromosome is associated with the pathogenesis of SLE and may provide mechanistic insights into the increased incidence of SLE in females.

Introduction

X chromosome inactivation (XCI), occurring specifically in female mammals, represents a sex-specific phenomenon that provides equalized gene expression between the sexes and dosage compensation. One of the X chromosomes is randomly silenced in normal female cells through a process involving long noncoding RNA (lncRNA), nucleic acid binding factors and

chromatin regulators (1,2). Once silenced, most genes on the inactive X chromosome (Xi) remain silent through subsequent somatic cell divisions, with few exceptions at genes that can escape XCI (eXCI) to some degree in somatic cells.

Due to the random choice of Xi in each cell, females have varying XCI ratios, defined as the proportions of cells expressing alleles from the maternal or the paternal X chromosome. In

[†]Yanfeng Zhang, <http://orcid.org/0000-0002-3859-3839>

Received: February 3, 2020. Revised: April 27, 2020. Accepted: June 11, 2020

the general female population, XCI ratios have a normal distribution, with an average of ~50:50. Skewed XCI is a bias of X-inactivation ratios towards cells overwhelmingly expressing one allele from two X chromosomes, e.g. a skewed ratio of >80:20 (3). Skewed XCI can occur in both normal females and those with X-linked diseases and can result from either a primary non-random inactivation occurring early in development or a secondary, acquired nonrandom inactivation in somatic tissues (4). More recently, skewed XCI has been reported in association with several autoimmune disorders (5,6) and cancers (7,8). Thanks to high-throughput sequencing techniques, determining skewed XCI is more readily accessible in various ways. For example, the presence of a heterozygous polymorphism on the two X chromosomes enables allelic expression (AE) mapping from deep RNA-sequencing (RNA-seq) data and has been a direct approach to examine allelic imbalance (AI) and XCI (9,10). Studies focusing on the epigenome have also uncovered the differences of DNA methylation and chromatin modification profiles on active and inactive X chromosomes (11,12).

X-inactive-specific transcript (XIST), a lncRNA that enables most, if not all, X-linked gene silencing by recruiting multiple proteins to the Xi (13), plays an essential role in the initiation and maintenance of XCI in female cells. Due to its role in XCI, XIST RNA is also one of the genes with a strong sex-specific expression pattern (14). This makes XIST lncRNA itself a strong candidate gene for sex-biased diseases.

Systemic lupus erythematosus (SLE) is an autoimmune disease characterized by the dysfunction of immune cells and the production of pathogenic autoantibodies. Like many other autoimmune disorders, there is a strong female bias in the incidence of SLE (15). However, the molecular basis for the sex disparity in SLE remains poorly understood, although both hormonal and genetic factors have been proposed (16–20). Recently, the role of XIST lncRNA itself in pathogenesis and sex bias of autoimmune disorders was explored. For example, using cellular imaging, Wang *et al.* reported different XIST RNA localization patterns but no alteration of XIST RNA levels, in B cells from pediatric patients with SLE (21). Syrett *et al.* revealed that dynamic maintenance of XCI is dependent on XIST RNA localization during murine B cell development (22). To pursue the role of XIST and XCI in the sex disparity observed in autoimmune diseases, we have conducted deep RNA-seq in lymphocytes and performed AE analyses in both CD4⁺ T and CD19⁺ B cells isolated from female adult SLE patients and healthy controls. Seventy-two RNA-seq datasets from three independent studies in SLE patients were included for validation. Our study reveals an overwhelming allele-specific expression pattern on X chromosome on SLE lymphocytes, which is positively correlated with the upregulation of XIST RNA in SLE lymphocytes. Functional studies further demonstrate that aberrant expression of XIST is responsible for the skewed AE between two X chromosomes in cultured lymphoblastic cells.

Results

Isolation and RNA-sequencing of T cell subsets

In the present study, 17 female participants, including five SLE patients and 12 controls, were recruited for the isolation of T cell subsets. All participants are adult with the average age of ~30 years (Supplementary Material, Table S1). In total, 66 cell populations across five types of CD4⁺ T cells were collected from the 17 individuals (Fig. 1A).

We performed RNA-sequencing on these cells. On average, nearly 130 million raw reads with ~90% mapping rate to reference human genome were obtained for each sample (Supplementary Material, Table S2). Such ultra-deep coverage data is favorable for AE analysis.

Allelic expression signals on the X chromosome across T cells

In a search for heterozygous loci that would be useful for measuring allelic expression levels, we identified 48 760 heterozygous single nucleotide polymorphisms (SNPs) occurring at least once in both case and control T cells (see Materials and Methods). Using two alignment strategies, the sequence reads were mapped to either the standard reference genome or the modified reference genome (see Materials and Methods); our results showed that the distribution of allelic ratios (ARs) for those heterozygotes was close to 0.5 for both mapping strategies (Supplementary Material, Fig. S1), indicating that there was no significant bias of variant versus reference alleles for either alignment strategy. It is important to note that the number of reads carrying the reference allele in the actual dataset may not necessarily be 50% (23). The presence of inherent biases, such as allele-specific loci due to uniparental imprinting, can tilt the numbers (23,24).

We then compared the global distribution of ARs at 48 760 heterozygous SNPs between SLE and controls. We found that there was an SLE-specific distribution of allelic ratios on the X chromosome but no significant difference on autosomes, in a comparison between the two groups (Fig. 1B). The result also showed a greater skewing pattern towards the two allelic ratio extremes (close to 0 or 1) for X-linked loci in female lupus T cells, relative to the pattern in controls. These results indicate that lupus T cells may exist as skewed XCI.

Next, based on those heterozygous SNPs ($n = 48\,760$) shared between cases and controls, we used a beta-binomial Bayesian model to identify loci with allelic bias in gene expression (see Materials and Methods). We defined heterozygous SNPs showing statistically significant ($P < 0.05$) allelic RNA expression imbalance in at least one type of T cell within SLE patients but not significant across any T cell subtypes among controls, as SLE-associated AE sites (SLE-AEs). On the basis of this definition, we identified 1718 SLE-AE sites across the genome. By calculating the proportion distributed on chromosomes for these SLE-AEs shown in Figure 1C, we found that there was a higher proportion of SLE-AEs ($n = 144$) on the X chromosome, relative to the proportions on other chromosomes. Additionally, we ruled out the possibility that overrepresentation of SLE-AE signals on the X chromosome was due to the biased distribution of 48 760 SNPs analyzed in this study. Together, these results suggest a nonrandom excess of SLE-AE signals on the X chromosome in female lupus patients.

To address whether SLE-AE signals are distributed regionally or chromosome widely, we plotted the X chromosome wide view of 144 SLE-AE signals (Fig. 1D). The result showed that these X-linked SLE-AE signals were broadly distributed on the entire X chromosome. We then used two metrics (see Materials and Methods) to quantify the AE difference between case and control T cells. As shown in Figure 1E, a 2- to 3-fold increasing on average of AE signals in all 5 T cell subtypes was detected for 144 SLE-AEs in patients, compared with control counterparts. In addition, we did not find a remarkable difference (more than 1.5-fold change) of AE profiles among T cell subtypes within SLE patients. Based on these results, we infer that changes in the skewed AE

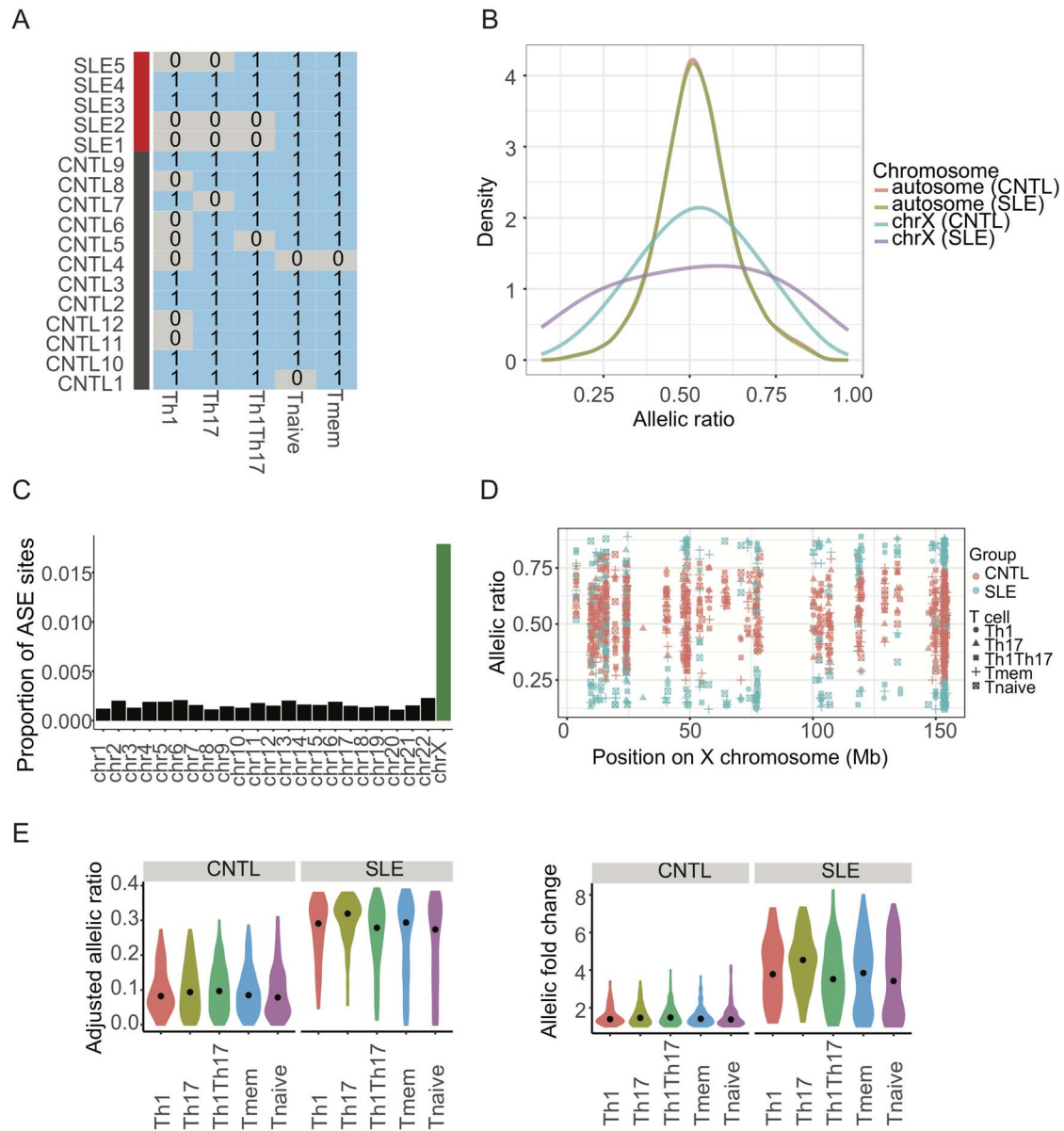


Figure 1. Allelic expression profiles on X chromosome in T cell subtypes from SLE patients. (A) Summary of 66 T cell samples collected from five SLE patients and 12 healthy controls (CNTL). 1 and 0 denote the collected or not collected T cell subtypes from participants, respectively. (B) Distribution of ARs for all heterozygotes identified on SLE and control T cells. Data on autosomes and X chromosome were plotted separately. (C) Plot of the proportion of SLE-AE sites on all chromosomes. The proportion (y-axis) was calculated by counting the number of SLE-AE SNPs divided by all heterozygous SNPs counted on each chromosome. (D) Dot plot showing the ARs of SLE-AE sites on the X chromosome for all informative T cells. (E) Plot of the adjusted ARs (left panel) and the allelic fold changes (right panel) across all T cell subsets in a comparison between control and SLE groups.

pattern on the X chromosome may exist prior to T cell lineage commitment in females with lupus.

Positive correlation between gene-based allelic RNA expression and XIST expression levels in T cells

By comparing allelic fold change between two sets of X-linked genes in normal T cells (see Materials and Methods), we evaluated the effectiveness of the gene-based allelic analysis. As expected, we observed that the set of genes with skewed XCI had significantly higher allelic expression levels than the other set of genes with eXCI ($P = 1.2 \times 10^{-7}$, two-sided Wilcoxon rank sum test, [Supplementary Material, Fig. S2A](#)). We also presented the

XCI status for five genes with known skewed XCI and three genes with known eXCI ([Supplementary Material, Fig. S2B](#)). Interestingly, these five genes showed high variability of XCI skewing in normal T cells, which is concordant with reports in other studies (25,26). Collectively, the method of gene-based allelic analysis is effective to determine the allelic difference of X-linked genes.

Next, we performed the gene-based allelic analysis for 144 SLE-AE sites that map onto 88 genes. Concordant with SNP-based allelic analysis, these 88 X-linked genes had a dramatic increase in allelic RNA expression ([Supplementary Material, Fig. S3A and B](#)) but didn't remarkably change the overall expression levels ([Supplementary Material, Fig. S4](#)), in T cells from SLE patients, relative to controls.

Interestingly, of these 88 genes, 64 (72.7%) were enriched (shown in blue, Fig. 2C) and had a significant overlap ($P=2.4 \times 10^{-5}$, two-sided binomial test) with genes showing escape (or variable escape) from XCI reported by previous studies in normal human tissues and cells, including lymphocytes (25). For example, two genes (*MSL3* and *ZFX*, Fig. 2C) known to eXCI previously identified in lymphoblastoid cells (27) also showed escapee in T cell subsets from healthy individuals.

Gene expression analysis identified a sparse but significant upregulation of XIST RNA in SLE T cell populations ($P=0.01$, Student's *t*-test, Fig. 2A). When stratified by T cell subsets, the XIST RNA was mainly upregulated in activated T cell subsets from SLE patients, relative to the corresponding T cells from controls. By calculating the Pearson's correlation coefficients (PCCs) between XIST expression levels and gene-wise allelic fold changes for these 88 genes across all T cell subsets, we detected that allelic fold changes on these 88 genes show much higher, although apparently weak (Fig. 2C), positive correlation with the XIST expression levels ($P=0.02$, two-sided Wilcoxon rank sum test), relative to the background sampling in same bin size (Fig. 2B). Taken together, our results suggest that elevated expression of XIST is associated with X-linked genes showing SLE-AE profiles among T cell subsets.

RNA-sequencing of naïve B cells

To further determine whether the XCI defect is present in other lymphocytes from patients with SLE, seven female donors (five SLE patients and two healthy controls) were recruited and their naïve B cells were sorted. By RNA-sequencing (RNA-seq) on these cells, we also obtained sequencing data with deep coverage. On average, approximately 73 million raw reads with ~90% mapping rate were acquired for these naïve B cells (Supplementary Material, Table S2).

Allelic expression signals in B lymphocytes

Using the same approaches shown above, we conducted single SNP-based AE analysis in B cells. Analysis of the chromosomal distribution of heterozygous SNPs with allelic ratios showed that the X chromosome, but not autosomes, harbors much denser AE in naïve B cells from SLE patients, compared with controls (Fig. 3A), a pattern that was also seen in lupus T cells.

Of 17 248 heterozygous SNPs analyzed in this study, we identified 933 SLE-AE sites in naïve B cells from lupus patients. We found a substantial excess of these SLE-AE signals on the X chromosome (Fig. 3B). These X-linked SLE-AE signals ($n=51$) were scattered across the X chromosome (Fig. 3C). Based on the same quantitative analyses, our results confirmed a significant difference with 2- to 3-fold increase in SLE B cells for these X-linked SLE-AE signals, relative to controls (Fig. 3D and E). These results are consistent with findings in the T cell data.

Allelic expression patterns are positively correlated with the XIST expression levels in B cells

By comparing the RNA levels of XIST between SLE and control B cells, we observed higher expression of XIST in four of five SLE B cell samples, relative to controls (Fig. 4A). RT-qPCR analysis in four additional naïve B cells further confirmed the elevated expression of XIST RNA in SLE B cells ($P=0.016$, Student's *t*-test, Fig. 4B).

Gene-wise AE analysis aggregating heterozygous SNPs at 36 genes showing SLE-AE signals confirmed a tremendous

difference of AE profiles between SLE and controls (Fig. 4C). It should be noted that the AE patterns were variable within patient B cells, particularly for the subject with SLE, which was the SLE patient with the lowest XIST expression levels (arrow shown in Fig. 4A). Similar to T cells, we did not observe large total expression changes (greater than 2-fold change) at these 36 genes in lupus B cells (Supplementary Material, Fig. S5), except for XIST RNA itself. Thus, the allelic ratio changes we observe are not due to elevated overall expression.

Consistent with findings in the study of T cells, we observed a considerably higher positive correlation ($P=1.4 \times 10^{-7}$, two-sided Kolmogorov-Smirnov test) between XIST RNA levels and allelic fold changes for 36 genes, relative to genome-wide background (Fig. 4D).

To further determine whether these genes showing SLE-AE signals observed in naïve B cells could be related to B cell maturation, we compared the AE profiles of 36 genes between naïve and transitional B cells within patients of SLE. We observed a slight difference of AE signals between naïve and transitional B cells (Supplementary Material, Fig. S6A and B). Meanwhile, the AE pattern also showed a greater positive correlation with the XIST RNA levels, relative to genome-wide background ($P=0.004$, two-sided Kolmogorov-Smirnov test, Supplementary Material, Fig. S6C-E). Similarly, among these 36 genes, no individual showed differential overall expression between B cell subsets (Supplementary Material, Fig. S6F). Taken together, our data indicate that XIST expression is directly related to AE on the X chromosome and is a common feature across lymphocyte populations.

Validation and refinement of results from independent RNA-seq data

To further validate the XIST-associated increase of AE on the X chromosome in lupus lymphocytes, we collected and analyzed a series of 83 RNA-seq data from lupus B or T cell types in three independent studies (Supplementary Material, Table S3), where 11 RNA-seq data were excluded due to the presence of genome-wide AE pattern with unknown reason (see an example in Supplementary Material, Fig. S7), resulting in 72 RNA-seq data for AE analysis.

Analyzing 56 published RNA-seq data on two independent cohorts of SLE B cells ($n=38$ for one study with ID, GSE118254; $n=18$ for the other study with ID, GSE110999, Supplementary Material, Table S3), we observed a significant enhanced expression of XIST in lupus B cells, compared with healthy controls (Supplementary Material, Figs S8A and S9A). In line with our RNA-seq study, we found that there was a trend showing the gradual increasing of XIST RNA levels in B cell differentiation from transitional to naïve stage within lupus patients (Supplementary Material, Fig. S8B). Based on AE profiling analyses from six SLE patients across three B cell subsets and three SLE patients across two B cell subtypes, we discovered that there were present of two different AE patterns on X chromosomes during lupus B cell maturation. One pattern was increased AE between two X chromosomes in association with XIST RNA levels (Supplementary Material, Figs S8C and S9B), which was consistent with the results from our RNA-seq data. The other pattern appeared to be XIST expression-independent AE on the X chromosome (Supplementary Material, Figs S8D and S9C).

We further analyzed 16 RNA-seq data in pediatric SLE from 2 T cell types (accession ID, GSE109843, Supplementary Material, Table S3) and observed that XIST expression levels were significantly increased in follicular helper T (T_{fh}) cells, relative

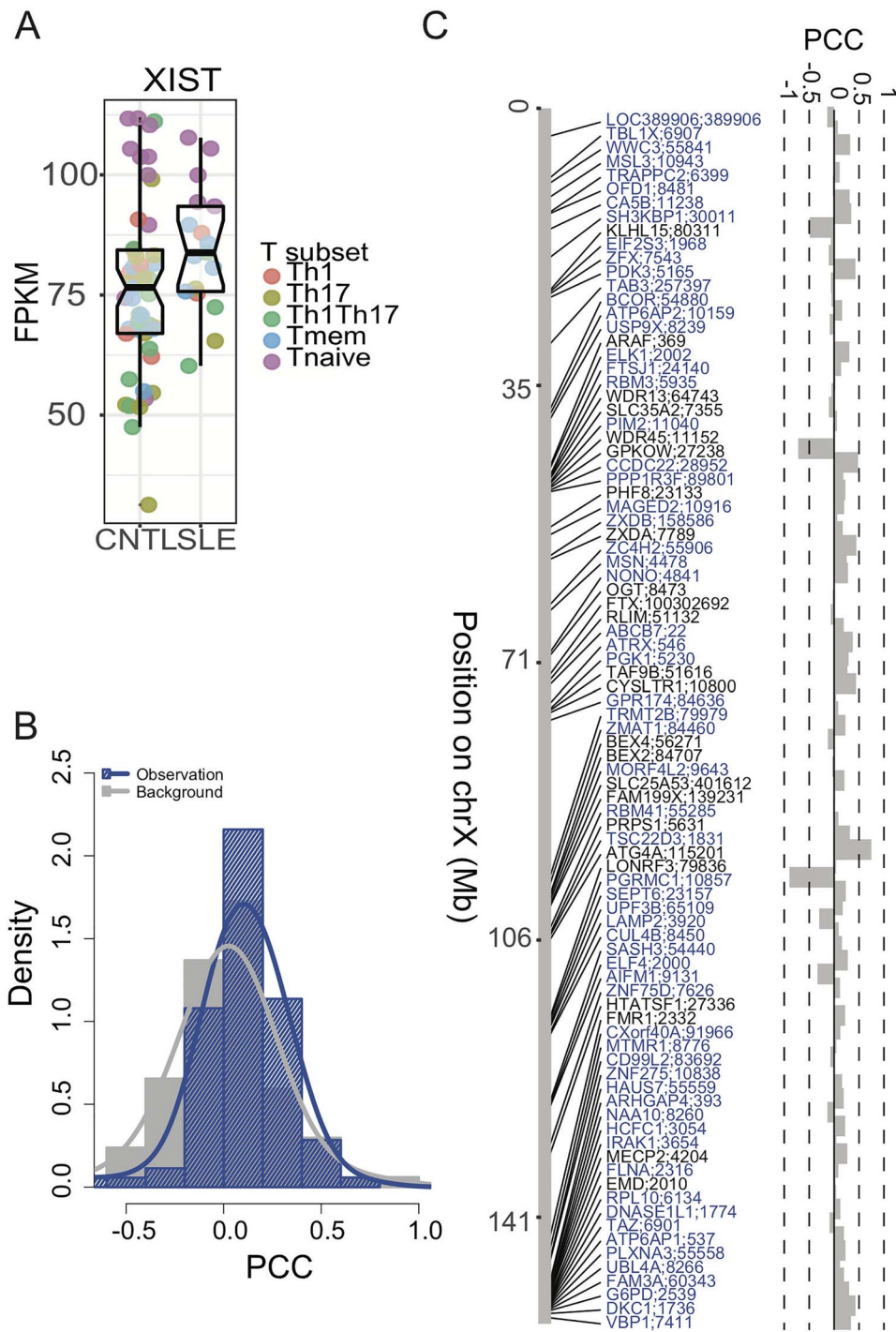


Figure 2. Association of allelic expression profiles on X chromosome with XIST expression levels across T cell subtypes. (A) Comparison of XIST expression levels in T cell subtypes from SLE patients and controls. (B) Distribution of PCCs between gene-based allelic fold changes and XIST expression levels for the observation and the background. (C) Barplot of the Pearson correlations between gene-based allelic fold changes and XIST expression levels for those X-linked genes harboring SLE-AE signals. The resulting PCCs are also seen as the observation that is used below. Genes blue highlighted are the previously reported genes showing eXCI.

to memory T cells (T_{mem}) in 8 SLE patients (Supplementary Material, Fig. S10A). Comparative analysis of AE patterns on X chromosome indicated that there was a significant increase of X-linked AE signals in T_{th} cells, relative to the pattern in T_{mem} cells (Supplementary Material, Fig. S10B). In contrast, there was no difference of AE profiles on autosomes between those two T cell populations. We also observed a

significantly higher positive correlation ($P=8.9 \times 10^{-8}$, two-sided Kolmogorov-Smirnov test) between XIST RNA levels and allelic fold changes for X-linked genes, relative to genome-wide background (Supplementary Material, Fig. S10C). Taken together, these results confirm the observation of increased AE on X chromosome in lymphocyte differentiation for lupus patients.

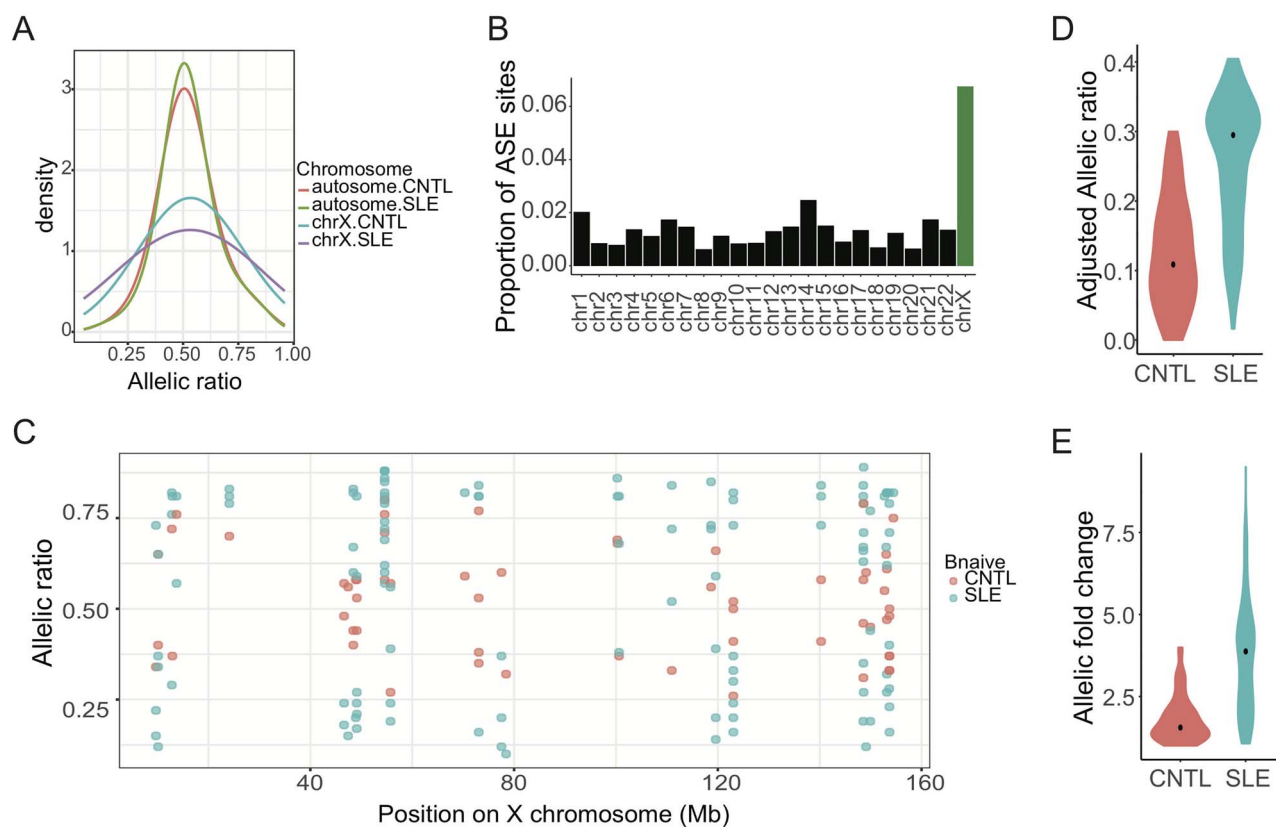


Figure 3. Allelic expression profiles on SLE naïve B cells. (A) Distribution of ARs for heterozygotes identified on autosomes and X chromosome in Naïve B cells from patients with SLE and controls. (B) Barplot showing the proportional distribution of SLE-AE sites on chromosomes. (C) Dotplot showing ARs for SLE-AE signals on the whole X chromosome. (D and E) Violin plot showing the adjusted ARs and the allelic fold change in a comparison between controls and SLE B naïve cells.

Reduced allelic imbalance on two X chromosomes when blocking XIST function

To further demonstrate the causal relationship between XIST expression and allelic imbalance of X-linked genes, we performed experiments to modulate levels of XIST RNA. We inhibited XIST activity in REH lymphocytic cells via knockdown of XIST expression by short hairpin (sh) RNA. We confirmed that the RNA level of XIST was greatly reduced in shRNA-XIST cells by both RNA-seq (Fig. 5A) and qPCR analyses (Fig. 5B), relative to cells transfected with a scrambled shRNA.

A comparative analysis showed that there was a remarkable decline of allelic fold changes ($P=0.004$, two-sided Wilcoxon rank sum test) for these X-linked polymorphisms when XIST lncRNA expression was inhibited (Fig. 5C). In contrast, we did not observe a significant change of AE on autosomes when XIST expression was blocked (Fig. 5C). Collectively, our data provide evidence that aberrant expression of XIST could specifically alter the AE pattern between X chromosomes.

Greater variability of DNA methylation levels on X-linked genes exhibiting AE signals in lupus patients

Finally, by integrating the independent dataset from four case-control association studies ($n=287$, including 148 cases and 139 controls, Supplementary Material, Table S4) of the DNA methylation profiling in patients with lupus, we explored whether DNA methylation levels were associated with XIST-induced AE patterns on Xs. For those AE-harboring X-linked genes detected in patients with lupus, we did not find a remarkable difference

of DNA methylation levels between cases and healthy controls in all four datasets (data not shown). The result is in agreement with the findings that the overall expression levels for those genes did not differ between lupus and control lymphocytes in this study.

However, under disease states like SLE, in which the XCI fidelity has been compromised, X-linked genes may have greater variation in DNA methylation levels among patients. We tested this hypothesis by comparing the variability of DNA methylation levels between SLE and controls by comparing the coefficient of variation (CV) of DNA methylation levels on those X-linked genes exhibiting AE signals detected in lupus T and B cells. We observed increased CVs in SLE patients compared with controls in independent datasets, where DNA methylation profiles were measured in CD4⁺ T cells and CD19⁺ B cells (Fig. 6A and B). However, we did not observe increased variability of DNA methylation levels in another two datasets (Supplementary Material, Fig. S11) in which myeloid cells (monocytes and neutrophils) were analyzed. These results together suggest a skewing towards greater variability of DNA methylation levels on those SLE-associated AE-harboring X-linked genes in SLE lymphoid but not myeloid cells, compared to the controls.

Discussion

Allelic expression analysis is a powerful approach to identify cis-regulatory variations and skewed XCI patterns on the X chromosome. By analyzing AE patterns from deep RNA-sequencing data, we have observed allele-specific expression patterns in lupus T

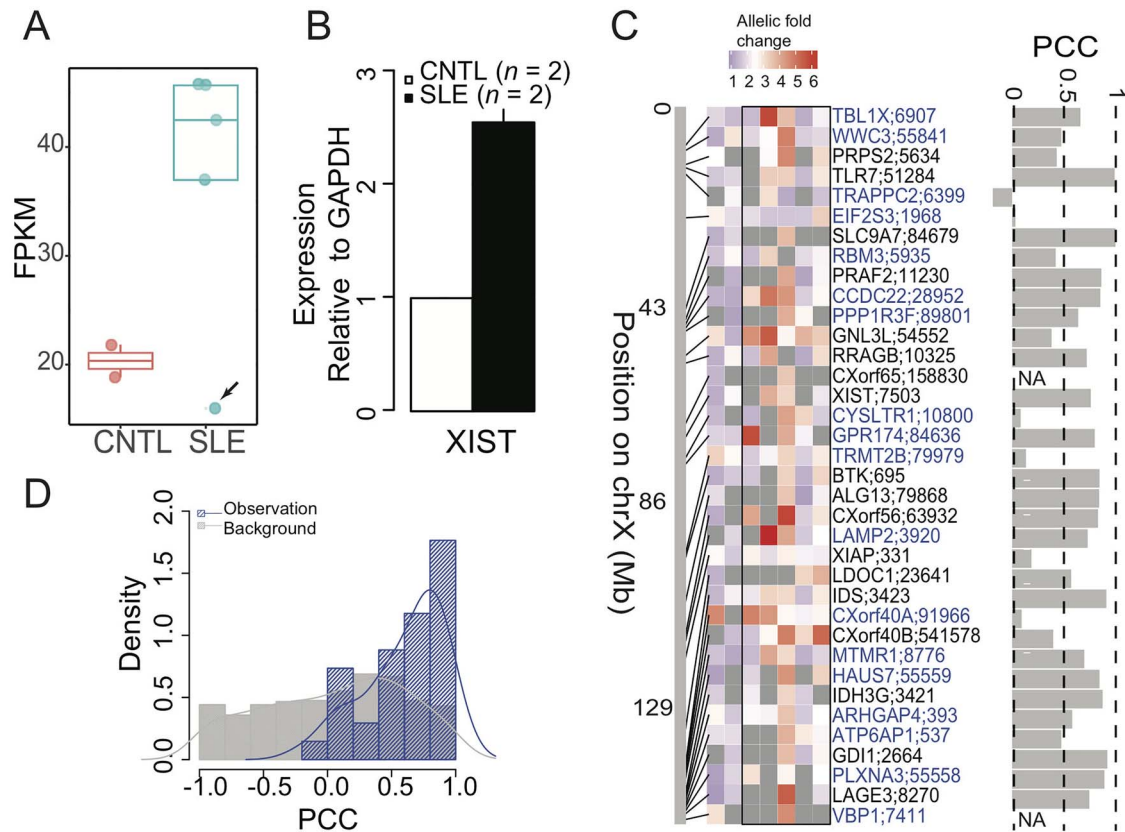


Figure 4. Association of allelic expression profiles on X chromosome with XIST expression levels across B cell subtypes. (A) Comparison of XIST expression levels (represented with FPKM values) in naïve B cells from SLE and controls. Arrow points to a SLE sample with the lowest XIST expression level. (B) Quantitative RT-PCR assay shows an upregulation of XIST on SLE naïve B cells. Error bars indicate the standard deviations from three replicates. (C) Heatmap showing gene-based AE profiles for seven B cell samples, where five patients with SLE were grouped and shown in rectangle. Genes overlapped with the genes detected in T cell subtypes (Fig. 2C) are highlighted in blue. The PCCs between gene-based allelic fold changes and XIST expression levels were shown on the top in a range of -1 (negative correlation) to 1 (positive correlation). (D) Distribution of PCCs between gene-based allelic fold changes and XIST expression levels for the observation and the background.

and B cells. We demonstrate that aberrant expression of XIST RNA is associated with and likely drives the observed AE pattern. Given the crucial role of XIST in XCI, and the concentration of lupus-specific AE on the X chromosome, our results likely indicate a dysfunctional XCI process in lupus lymphocytes. Our results are consistent with previous findings in mouse models that demonstrate that altered expression levels of XIST RNA can trigger X-linked gene silencing (28). There have also been studies indicating skewed XCI in several autoimmune disorders, including autoimmune thyroid diseases (6) and scleroderma (5). In this regard, these data contribute to a growing body of literature supporting the importance of altered XCI feature in the pathogenesis of many autoimmune diseases. Given that XCI is limited to female cells, these finding may explain the female predisposition to autoimmune diseases, including SLE, if a defective XCI process increases the propensity to generate autoreactive lymphocytes or reduces the ability of patients to suppress or eliminate autoreactive lymphocytes.

Since not all of X-linked genes showed an AE pattern in SLE lymphocytes, it indicates that skewed XCI is unlikely to have occurred early in embryogenesis, when X-inactivation patterns are first established. Furthermore, the observation that AE was detected across all T cell subsets within SLE, including in naïve T cells, suggests that XCI defects arise at an early hematopoietic stage, potentially through the aberrant expression of XIST.

Two X-linked regions, spanning the *PRPS2*-*TLR7* and *IRAK1*-*MECP2* genes, have been reported to be associated with SLE (29,30). Our result indicated that these two risk loci escaped the XCI in healthy controls but are subjected to XCI in SLE patients. Another interesting X-linked gene, *CCDC22*, showing a strong skewed XCI in SLE lymphocytes in the present study, is associated with SLE susceptibility (31). The *CCDC22* protein functions in the regulation of $\text{NF-}\kappa\text{B}$ signaling (32), a well-known regulator being aberrantly activated in lupus disease (33,34). Therefore, X-linked SLE-related genes may serve as a selective target in patients as a failure to maintain XCI fidelity that could create a population of dysfunctional lymphocytes carrying aberrant expression of specific risk alleles at these genes.

In the study of Wang *et al.*, they reported that altered distribution of XIST RNA patterns, rather than expression levels, could contribute to silenced X chromosome reactivation in pediatric SLE B cells (21). On the one hand, we observed a comparable result in a SLE patient where reduced AE pattern on X chromosome but no change of XIST expression was detected in memory B cells, relative to naïve B cells (Supplementary Material, Fig. S9C). Thus, our analyses partially support their findings. On the other hand, distinct from their findings, we discovered skewed AE associated with aberrant expression of XIST in lupus patients. Some studies have suggested a pronounced difference of clinical manifestations between pediatric SLE and adult SLE (35–37). For

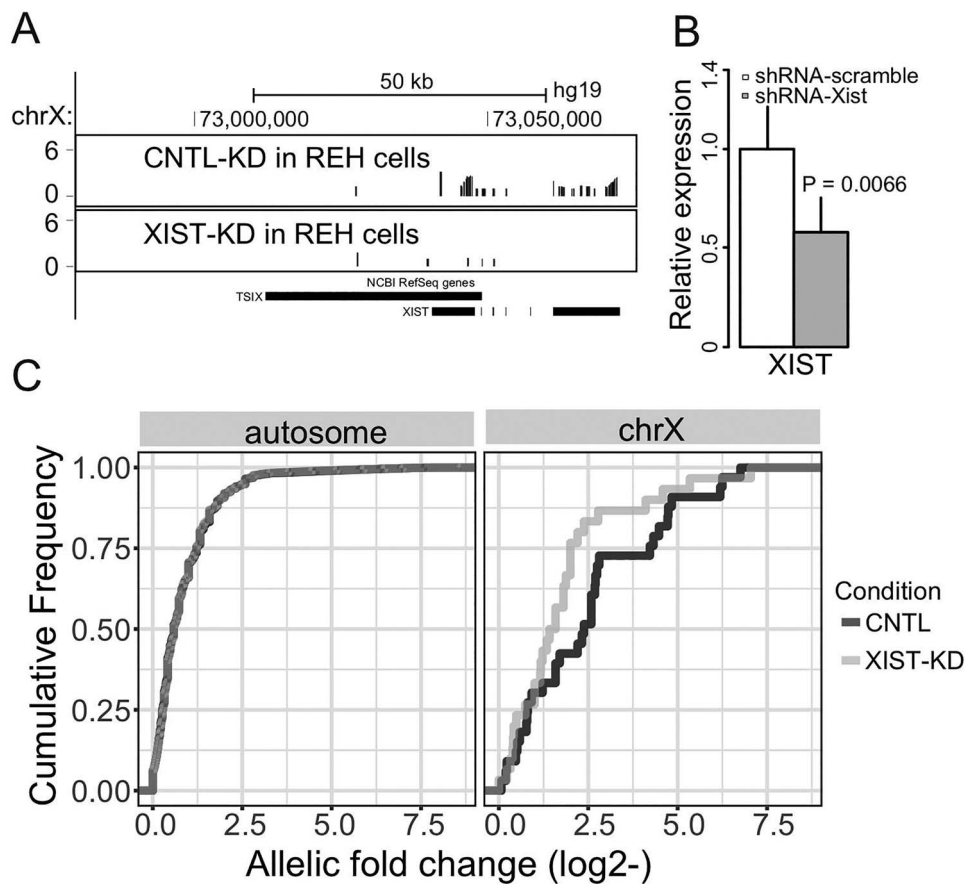


Figure 5. Reduced allelic expression pattern when blocking the XIST function in REH cells. (A) Comparison of expression levels of XIST between scrambled shRNA and shRNA-XIST REH lymphoblastic cells. Two tracks are the read coverage in 20-bp bins across the XIST and its adjacent TSIX locus. (B) Quantitative RT-PCR assay quantifies the expression levels of XIST in shRNA-XIST cells, relative to control cells. (C) Empirical cumulative distribution of allelic fold change between scrambled shRNA and shRNA-XIST for heterozygous SNPs mapped onto autosomes and X chromosome.

example, childhood-onset SLE has a more severe clinical course than that seen in adults, with a higher prevalence of lupus nephritis, hematologic and immunologic anomalies. Meanwhile, the gender bias towards female is commonly observed in adults with SLE, while the gender ratio shows even in childhood-onset SLE (36). This difference may provide a possible explanation of the difference of our analysis of adult-onset SLE with Wang *et al.* findings in pediatric SLE.

Yildirim *et al.* reported that XIST RNA is a potential tumor suppressor of hematologic cancer in mouse model (38). In contrast, Yao *et al.* found that XIST expression was upregulated in human glioblastoma stem cells and its downregulation suppressed tumor growth (39). These results suggest that XIST not only plays a role in the XCI but also has roles in cell growth and proliferation. Engreitz *et al.* proposed a model that XIST RNA coats the X chromosome via chromatin conformation interaction (40). This means that aberrant expression of XIST might also alter the DNA topological architecture on the X chromosome in SLE lymphocytes. Thus, we speculated that the up-expression of XIST might have some additional influences to SLE patient cells and may alter the proliferative potential of clonal populations that carry aberrant XIST levels.

One caveat to our study is that expression levels of XIST RNA within female lupus lymphocytes and skewed AE profiles were highly variable. This result implies that some additional factors linked with SLE pathogenesis and gender bias are present.

However, the statistical power in the present study is insufficient to link the AE with SLE clinical features. Indeed, besides genetic components, several environmental factors, including infections, endogenous and exogenous hormones (19,20,41), have been proposed to explain the sex difference in SLE. Thus, we anticipate that AE mapping across more hematopoietic cell states and larger samples, as well as systematic analyses by integrating genetic, epigenetic and disease-linked phenotypic data, will yield new insight into the variable XCI and the potential factors (or cofactors) associated with sex bias in autoimmune diseases.

In conclusion, our study provides evidence that XIST-altered AE on X chromosome may be implicated in the pathogenesis of female patients with SLE and may also help explain the significant female predisposition to the autoimmunity.

Materials and Methods

Sorting of lymphocyte subtypes

Peripheral blood mononuclear cells (PBMC) and neutrophils were isolated from blood following fractionation by Ficoll density centrifugation. CD14⁺ monocytes were isolated from the PBMC using anti-CD14 coated beads (Invitrogen/Thermo Fisher). CD4⁺ cells were isolated using anti-CD4 coated beads, detached from the beads and stained with a panel of fluorochrome-conjugated

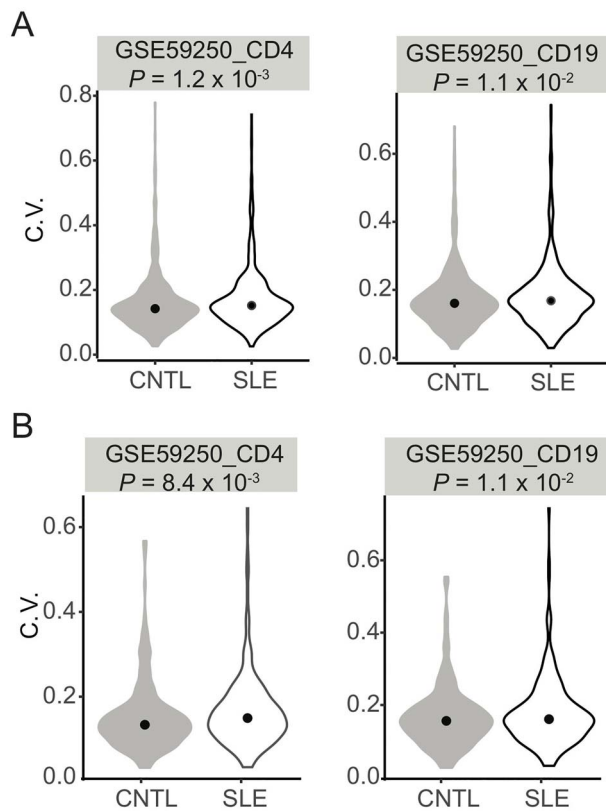


Figure 6. Variability of DNA methylation levels on promoters of X-linked genes exhibiting AE. (A and B) Comparison of the coefficient of variation (CV) of DNA methylation levels on promoters of SLE-AE X-linked genes discovered in T cell subsets (A) and naive B cells (B) between patients with SLE and controls from two independent studies (head shown).

antibodies (CD4, CD56, CD45RA, CD45RO, CD25, CD127, CXCR3, CCR4 and CCR6 from BD Pharmingen and eBioscience). Cells were then sorted by FACS analysis to obtain CD56⁻ populations representing naïve (CD45RO⁻, CD45RA⁺), memory (CD45RO⁺, CD45RA⁻) and Treg (CD25⁺, CD127⁻) T cells. Memory T cells were further sorted to obtain Th1 (CXCR3⁺, CCR4⁻, CCR6⁻), Th1Th17 (CXCR3⁺, CCR4⁻, CCR6⁺) and Th17 (CXCR3⁻, CCR4⁺, CCR6⁺) subsets.

CD19⁺ B cells were isolated from the PBMC using anti-CD19-coated beads. CD19⁺ cells were detached from the beads and then incubated with anti-CD27-coated beads to separate memory and plasma cells (CD19⁺, CD27⁺) from immature, transitional and naïve B cells (CD19⁺, CD27⁻). Following detachment from the beads, memory and plasma cells were then stained with fluorochrome-conjugated anti-CD19, anti-CD27 and anti-CD24 and sorted to obtain memory cells (CD19⁺, CD27⁺, CD24⁺) and plasma cells (CD19⁺, CD27⁺, CD24⁻). The mixed population of immature, transitional and naïve cells was further stained with anti-IgD and sorted to isolate immature cells (CD19⁺, CD27⁻, IgD⁻) from the transitional and naïve cells (CD19⁺, CD27⁻, IgD⁺). This latter mixed population was then stained with anti-CD38 and anti-CD24 and sorted to separate transitional cells (CD19⁺, CD27⁻, IgD⁺, CD38^{hi}, CD24^{hi}) from naïve cells (CD19⁺, CD27⁻, IgD⁺, CD38^{lo}, CD24^{lo}). The purity of the cell subsets was ascertained by FACS analysis of a portion of the sorted cells to determine the percentage that fell within the gate initially used to target that population, and typically purity was 91% or greater.

Plasmid constructs

pSIREN-RetroQ scrambled construct (5'-GCAAGCTGCCCGTGCCCTG-3', as control (42)) was a gift from Dr Gilmore (Boston University, USA). Short hairpin RNAs (shRNAs) for XIST sequences were designed using the shRNA Sequence Designer (Clontech) and were subcloned into the pSIREN-RetroQ retroviral vector (Clontech) according to the manufacturer's protocol.

Cell culture, RNA isolation and real-time RT-PCR

REH (purchased from ATCC, cat. no. CRL-8286) pre-B lymphoblastic cells were cultured in RPMI-1640 medium, supplemented with 10% FBS (Thermo Fisher Scientific, Waltham, MA, USA), 2 mM L-glutamine and 1% penicillin–streptomycin at 37°C with 5% CO₂. Transfections were carried out using Lipofectamine 2000 Transfection Reagent (Thermo Fisher) according to the manufacturer's instruction. 72 h after transfection, the cells were washed once with PBS. Then, RNA was isolated from cells using TRIzol Reagent (Invitrogen) according to the manufacturer's protocol. 1 µg total RNA was reverse transcribed using SuperScript III (Thermo Fisher Scientific) and Oligo-dT primer. One-tenth of the RT reaction was used as a template for real-time PCR using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) on a QuantStudio 6 system. Relative expression was calculated with 2^{-ΔΔCt} using the average value of housekeeping gene GAPDH. Data are presented as mean ± SD of three replicates. All qPCR primers are listed in [Supplementary Material, Table S5](#).

RNA-seq library construction

For RNA-sequencing, indexed libraries were constructed from ~500 ng total RNA using the NEBNext Ultra RNA Library Prep Kit for Illumina with the Poly(A) mRNA Magnetic Isolation Module (NEB). We confirmed the final cDNA library size range by gel electrophoresis on a 1.7% agarose gel. Samples were barcoded during library preparation and sequenced on a HiSeq 2500 or NextSeq (Illumina) instrument as 2 × 100 or 2 × 75 bp paired-end libraries at the HudsonAlpha Institute Sequencing Facility.

RNA-seq data processing

RNA-seq data were analyzed as described elsewhere (43). Briefly, reads were demuxed based on their index sequence at the Genomic Services Laboratory at HudsonAlpha. After removal of adapter sequences, low-quality reads and trimmed reads that were shorter than 20 bp by using cutadapt (v1.3.1), filtered reads were aligned with Tophat2 with default parameters. Mapped reads in BAM format were assembled with Cufflinks (v2.2.1). We used FPKM values for representing the expression levels of genes. We retained genes with mean FPKM ≥ 1 for fold change analysis.

Identification of allelic expression polymorphic sites

We used a similar approach described in previous study to call genetic variants (44). Briefly, mapped reads in BAM format were first subject to the removal of PCR duplicates and then were realigned and recalibrated, and genetic variants were called in a multiple-sample joint manner implemented in the GATK toolkit (version 3.3). We next filtered out variants as follows: (1) mapping quality score < 20, (2) ≥ 3 SNPs detected within 10 bp distance, (3) variant confidence/quality by depth < 2, (4) strand bias score > 50, (5) genotype score < 15 and (6) read depth < 10. Then, we extracted SNPs annotated from dbSNP (Build 142) that

were called as heterozygotes for each sample. For a reasonable comparison, those heterozygous SNPs identified at least once in both case and control cells were retained for allelic analysis.

We employed a beta-binomial model to identify the allele-specific expression SNPs that has been widely used elsewhere in read-count datasets (45). Briefly, the number of reads for one allele at heterozygous loci follows a binomial distribution, $X \sim \text{Bin}(n, p)$, and then $P(X = k | n, p) = L(p|k)$, where n represents the total number of reads at each heterozygous SNP, k represents the number of reads for one allele in that heterozygous SNP, and P is the probability of the observation of read coverage on an allele. The binomial test typically assumes a fixed probability ($P = 0.5$), which tends to overestimate the number of allele-specific signals. Thus, we assume that p is a random variable that follows a beta distribution, $p \sim \text{Beta}(\alpha, \beta)$. By calculating the allelic ratio (see the below formula) for all heterozygous SNPs, the two shape parameters, α and β , were estimated by using maximum likelihood approach implemented in *optim* built-in function in R software:

$$\text{Allelic ratio} = \frac{\text{Depth}_{\text{Ref}}}{\text{Depth}_{\text{Ref}} + \text{Depth}_{\text{Alt}}}$$

where Ref and Alt denote the reference and alternative alleles, respectively. By integrating the prior $\text{Beta}(\alpha, \beta)$ and binomial likelihood $P(X = k | n, p)$, we obtained the probability density function of the beta-binomial distribution:

$$f(X = k | n, \alpha, \beta) = \binom{n}{k} \frac{\text{Beta}(k + \alpha, n - k + \beta)}{\text{Beta}(\alpha, \beta)}$$

Finally, based on the read depth on bi-allele SNPs, P -value could be estimated using the following formula:

$$\Pr(x_i \leq k \leq n) = \sum_{x_i}^n f(X = k | n, \alpha, \beta)$$

where $x_i = \max(\text{Depth}_{\text{Ref}}, \text{Depth}_{\text{Alt}})$. All these analyses were calculated by implementing in the R package called *LearnBayes* (46). The allele-specific expression signals were identified with P -value < 0.05 . Two statistics—adjusted allelic ratio and allelic fold change, were also calculated using formulas shown as follows:

$$\text{Adjusted allelic ratio} = \frac{\max(\text{Depth}_{\text{Ref}}, \text{Depth}_{\text{Alt}})}{\text{Depth}_{\text{Ref}} + \text{Depth}_{\text{Alt}}} - 0.5$$

$$\text{Allelic fold change} = \frac{\max(\text{Depth}_{\text{Ref}}, \text{Depth}_{\text{Alt}})}{\min(\text{Depth}_{\text{Ref}}, \text{Depth}_{\text{Alt}})}$$

We also applied alternative alignment strategies described in the study of Smith *et al.* to identify heterozygous SNPs and their allelic ratios (9). Briefly, we built a modified hg19 reference genome by replacing the reference allele with the variant SNP allele annotated in dbSNP 142 using the GATK FastaAlternateReference tool. We remapped all reads to this modified reference genome using the same parameters as used for the standard reference genome. Then, ARs and Bayesian model-derived significance for allelic bias was calculated as above. Finally, we defined those heterozygous polymorphisms with the posterior probability for allelic ratio with $P < 0.05$ estimated from both the standard alignment and the modified genome alignment as AE sites.

By collapsing heterozygous SNPs mapped to Refseq genes by averaging the allelic fold changes, we calculated the gene-wise allelic fold change. We collected a dataset from Tukiainen *et al.* study (10), where they identified two sets of X-linked genes being subject to XCI or eXCI across various human normal tissues. By comparing gene-wise allelic fold change between two sets of genes in T cells from healthy controls, we evaluated the effectiveness of this approach.

Correlation analysis

We calculated the Pearson correlations between gene-based allelic fold changes and XIST RNA levels using PCC. The background Pearson correlations were calculated between gene-based allelic fold changes, and XIST RNA levels for all other genes and the same size were randomly selected for histogram.

Public data collection and analysis

We collected RNA-seq data (Accession ID, GSE110999, GSE118254 and GSE109843) from three independent studies in SLE patients of B and T cells (47–49). We downloaded the raw reads data from the Sequence Read Archive (SRA) database. We then used the same method described above to process the RNA-seq data and AE analysis.

The Infinium Human Methylation450 (HM450) microarray dataset used in this study was acquired from the Gene Expression Omnibus (GEO) database (50). We used the processed beta values to represent the DNA methylation levels. We first annotated CpG probes and collected those probes mapped to promoter regions (transcription start sites ± 1 kb) of these X-linked genes exhibiting ASE sites. We then calculated the coefficient of variation (CV), a measurement for the relative variability, for each CpG probe within gene promoters in cases and controls for each of five independent studies. For statistical analysis, we employed one-sided Wilcoxon rank sum test to compare the difference of CV distribution between lupus and controls.

Supplementary Material

Supplementary Material is available at HMG online.

Availability of data and materials

The RNA-seq data generated during the current study are available in the GEO database under accession number [GSE152767].

Acknowledgements

We thank the members of the Genomic Services Laboratory at HudsonAlpha for providing sequence data. We thank Dr Le Su, Dr Jun Song and Dr Kenneth Day for critical reading and suggestive discussions. This study was supported by the HudsonAlpha institutional funds. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Conflict of Interest statement The authors declare that they have no competing interests.

Authors' contributions

Y.Z. conceived and designed the allelic expression study, performed the experiments, collected and analyzed the data, prepared figures and wrote and revised the manuscript. X.L., A.G.,

J.E., R.K. and D.A. designed and collected the lymphocytes. D.A. conceived and designed the RNA-seq study involved in lymphocyte subsets and edited the manuscript. All authors read, interpreted and approved the final manuscript.

References

1. Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E. and Chang, H.Y. (2015) Systematic discovery of Xist RNA binding proteins. *Cell*, **161**, 404–416.
2. Wutz, A. (2011) Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat. Rev. Genet.*, **12**, 542–553.
3. Amos-Landgraf, J.M., Cottle, A., Plenge, R.M., Friez, M., Schwartz, C.E., Longshore, J. and Willard, H.F. (2006) X chromosome; inactivation patterns of 1,005 phenotypically unaffected females. *Am. J. Hum. Genet.*, **79**, 493–499.
4. Minks, J., Robinson, W.P. and Brown, C.J. (2008) A skewed view of X chromosome inactivation. *J. Clin. Invest.*, **118**, 20–23.
5. Özbalkan, Z., Bağışlar, S., Kiraz, S., Akyerli, C.B., Ozer, H.T., Şule, Y., Merih, B.A., Meral, Ç. and Tayfun, Ö. (2005) Skewed X chromosome inactivation in blood cells of women with scleroderma. *Arthritis Rheum.*, **52**, 1564–1570.
6. Ozcelik, T., Uz, E., Akyerli, C.B., Bagislar, S., Mustafa, C.A., Gursoy, A., Akarsu, N., Toruner, G., Kamel, N. and Gullu, S. (2006) Evidence from autoimmune thyroiditis of skewed X-chromosome inactivation in female predisposition to autoimmunity. *Eur. J. Hum. Genet.*, **14**, 791.
7. Kristiansen, M., Langerod, A., Knudsen, G.P., Weber, B.L., Borresen-Dale, A.L. and Orstavik, K.H. (2002) High frequency of skewed X inactivation in young breast cancer patients. *J. Med. Genet.*, **39**, 30–33.
8. Li, G., Jin, T., Liang, H., Tu, Y., Zhang, W., Gong, L., Su, Q. and Gao, G. (2013) Skewed X-chromosome inactivation in patients with esophageal carcinoma. *Diagn. Pathol.*, **8**, 55.
9. Smith, R.M., Webb, A., Papp, A.C., Newman, L.C., Handelman, S.K., Suhy, A., Mascarenhas, R., Oberdick, J. and Sadee, W. (2013) Whole transcriptome RNA-Seq allelic expression in human brain. *BMC Genomics*, **14**, 571.
10. Tukiainen, T., Villani, A.C., Yen, A., Rivas, M.A., Marshall, J.L., Satija, R., Aguirre, M., Gauthier, L., Fleharty, M., Kirby, A. et al. (2017) Landscape of X chromosome inactivation across human tissues. *Nature*, **550**, 244–248.
11. Cotton, A.M., Price, E.M., Jones, M.J., Balaton, B.P., Kobor, M.S. and Brown, C.J. (2015) Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Hum. Mol. Genet.*, **24**, 1528–1539.
12. Sharp, A.J., Stathaki, E., Migliavacca, E., Brahmachary, M., Montgomery, S.B., Dupre, Y. and Antonarakis, S.E. (2011) DNA methylation profiles of human active and inactive X chromosomes. *Genome Res.*, **21**, 1592–1600.
13. Lee, J.T. and Bartolomei, M.S. (2013) X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell*, **152**, 1308–1323.
14. Ellegren, H. and Parsch, J. (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.*, **8**, 689–698.
15. van Vollenhoven, R.F. (2009) Sex differences in rheumatoid arthritis: more than meets the eye. *BMC Med.*, **7**, 12.
16. McDonald, G., Cabal, N., Vannier, A., Umiker, B., Yin, R.H., Orjalo, A.V., Johansson, H.E., Han, J.-H. and Imanishi-Kari, T. (2015) Female bias in systemic lupus erythematosus is associated with the differential expression of X-linked toll-like receptor 8. *Front. Immunol.*, **6**, 457.
17. Brooks, W.H. and Renaudineau, Y. (2015) Epigenetics and autoimmune diseases: the X chromosome-nucleolus nexus. *Front. Genet.*, **6**, 22.
18. Jacob, C.O., Zhu, J., Armstrong, D.L., Yan, M., Han, J., Zhou, X.J., Thomas, J.A., Reiff, A., Myones, B.L., Ojwang, J.O. et al. (2009) Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 6256–6261.
19. Quintero, O.L., Amador-Patarroyo, M.J., Montoya-Ortiz, G., Rojas-Villarraga, A. and Anaya, J.-M. (2012) Autoimmune disease and gender: plausible mechanisms for the female predominance of autoimmunity. *J. Autoimmun.*, **38**, J109–J119.
20. Oliver, J.E. and Silman, A.J. (2009) Why are women predisposed to autoimmune rheumatic diseases? *Arthritis Res. Ther.*, **11**, 252.
21. Wang, J., Syrett, C.M., Kramer, M.C., Basu, A., Atchison, M.L. and Anguera, M.C. (2016) Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E2029–E2038.
22. Syrett, C.M., Sindhava, V., Hodawadekar, S., Myles, A., Liang, G., Zhang, Y., Nandi, S., Cancro, M., Atchison, M. and Anguera, M.C. (2017) Loss of Xist RNA from the inactive X during B cell development is restored in a dynamic YY1-dependent two-step process in activated B cells. *PLoS Genet.*, **13**, e1007050.
23. Satya, R.V., Zavaljevski, N. and Reifman, J. (2012) A new strategy to reduce allelic bias in RNA-Seq readmapping. *Nucleic Acids Res.*, **40**, e127.
24. Degner, J.F., Marioni, J.C., Pai, A.A., Pickrell, J.K., Nkadori, E., Gilad, Y. and Pritchard, J.K. (2009) Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics*, **25**, 3207–3212.
25. Zhang, Y., Castillo-Morales, A., Jiang, M., Zhu, Y., Hu, L., Urrutia, A.O., Kong, X. and Hurst, L.D. (2013) Genes that escape X-inactivation in humans have high intraspecific variability in expression, are associated with mental impairment but are not slow evolving. *Mol. Biol. Evol.*, **30**, 2588–2601.
26. Cotton, A.M., Ge, B., Light, N., Adoue, V., Pastinen, T. and Brown, C.J. (2013) Analysis of expressed SNPs identifies variable extents of expression from the human inactive X chromosome. *Genome Biol.*, **14**, R122.
27. Garieri, M., Stamoulis, G., Blanc, X., Falconnet, E., Ribaux, P., Borel, C., Santoni, F. and Antonarakis, S.E. (2018) Extensive cellular heterogeneity of X inactivation revealed by single-cell allele-specific expression in human fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 13015–13020.
28. Gayen, S., Maclary, E., Hinten, M. and Kalantry, S. (2016) Sex-specific silencing of X-linked genes by Xist RNA. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E309–E318.
29. Kaufman, K.M., Zhao, J., Kelly, J.A., Hughes, T., Adler, A., Sanchez, E., Ojwang, J.O., Langefeld, C.D., Ziegler, J.T., Williams, A.H. et al. (2013) Fine mapping of Xq28: both MECP2 and IRAK1 contribute to risk for systemic lupus erythematosus in multiple ancestral groups. *Ann. Rheum. Dis.*, **72**, 437–444.
30. Zhang, Y., Zhang, J., Yang, J., Wang, Y., Zhang, L., Zuo, X., Sun, L., Pan, H.-F., Hirankarn, N., Wang, T. et al. (2015) Meta-analysis of GWAS on two Chinese populations followed by

- replication identifies novel genetic variants on the X chromosome associated with systemic lupus erythematosus. *Hum. Mol. Genet.*, **24**, 274–284.
31. D'Amico, F., Skarmoutsou, E., Lo, L.J., Granata, M., Trovato, C., Rossi, G.A., Bellocchi, C., Marchini, M., Scorza, R., Mazzarino, M.C. et al. (2017) Association between rs2294020 in X-linked CCDC22 and susceptibility to autoimmune diseases with focus on systemic lupus erythematosus. *Immunol. Lett.*, **181**, 58–62.
 32. Starokadomskyy, P., Gluck, N., Li, H., Chen, B., Wallis, M., Maine, G.N., Mao, X., Zaidi, I.W., Hein, M.Y., McDonald, F.J. et al. (2013) CCDC22 deficiency in humans blunts activation of proinflammatory NF-kappaB signaling. *J. Clin. Invest.*, **123**, 2244–2256.
 33. Wong, H.K., Kammer, G.M., Dennis, G. and Tsokos, G.C. (1999) Abnormal NF-kappa B activity in T lymphocytes from patients with systemic lupus erythematosus is associated with decreased p65-RelA protein expression. *J. Immunol.*, **163**, 1682–1689.
 34. Zhang, W., Shi, Q., Xu, X., Chen, H., Lin, W., Zhang, F., Zeng, X., Zhang, X., Ba, D. and He, W. (2012) Aberrant CD40-induced NF-kappaB activation in human lupus B lymphocytes. *PLoS One*, **7**, e41644.
 35. Harry, O., Yasin, S. and Brunner, H. (2018) Childhood-onset systemic lupus erythematosus: a review and update. *J. Pediatr.*, **196**, 22–30.e22.
 36. Mina, R. and Brunner, H.I. (2010) Pediatric lupus—are there differences in presentation, genetics, response to therapy, and damage accrual compared with adult lupus? *Rheum. Dis. Clin. N. Am.*, **36**, 53–80 (vii–viii).
 37. Brunner, H.I., Gladman, D.D., Ibanez, D., Urowitz, M.D. and Silverman, E.D. (2008) Difference in disease features between childhood-onset and adult-onset systemic lupus erythematosus. *Arthritis Rheum.*, **58**, 556–562.
 38. Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T. and Lee, J.T. (2013) Xist RNA is a potent suppressor of hematologic cancer in mice. *Cell*, **152**, 727–742.
 39. Yao, Y., Ma, J., Xue, Y., Wang, P., Li, Z., Liu, J., Chen, L., Xi, Z., Teng, H., Wang, Z. et al. (2015) Knockdown of long non-coding RNA XIST exerts tumor-suppressive functions in human glioblastoma stem cells by up-regulating miR-152. *Cancer Lett.*, **359**, 75–86.
 40. Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., Lander, E.S. et al. (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science*, **341**, 1237973.
 41. Ercolini, A.M. and Miller, S.D. (2009) The role of infections in autoimmune disease. *Clin. Exp. Immunol.*, **155**, 1–15.
 42. Garbati, M.R., Thompson, R.C., Haery, L. and Gilmore, T.D. (2011) A rearranged EP300 gene in the human B-cell lymphoma cell line RC-K8 encodes a disabled transcriptional co-activator that contributes to cell growth and oncogenicity. *Cancer Lett.*, **302**, 76–83.
 43. Zhang, Y., Wagner, E.K., Guo, X., May, I., Cai, Q., Zheng, W., He, C. and Long, J. (2016) Long intergenic non-coding RNA expression signature in human breast cancer. *Sci. Rep.*, **6**, 37821.
 44. Zhang, Y., Song, J., Day, K. and Absher, D. (2017) dCATCH-Seq: improved sequencing of large continuous genomic targets with double-hybridization. *BMC Genomics*, **18**, 811.
 45. Chen, J., Rozowsky, J., Galeev, T.R., Harmanci, A., Kitchen, R., Bedford, J., Abyzov, A., Kong, Y., Regan, L. and Gerstein, M. (2016) A uniform survey of allele-specific binding and expression over 1000-genomes-project individuals. *Nat. Commun.*, **7**, 11101.
 46. Albert, J. (2009) *Bayesian Computation with R*. Springer Publishing Company, Incorporated.
 47. Scharer, C.D., Blalock, E.L., Mi, T., Barwick, B.G., Jenks, S.A., Deguchi, T., Cashman, K.S., Neary, B.E., Patterson, D.G., Hicks, S.L. et al. (2019) Epigenetic programming underpins B cell dysfunction in human SLE. *Nat. Immunol.*, **20**, 1071–1082.
 48. Wang, S., Wang, J., Kumar, V., Karnell, J.L., Naiman, B., Gross, P.S., Rahman, S., Zerrouki, K., Hanna, R., Morehouse, C. et al. (2018) IL-21 drives expansion and plasma cell differentiation of autoreactive CD11c(hi)T-bet(+) B cells in SLE. *Nat. Commun.*, **9**, 1758.
 49. Caielli, S., Veiga, D.T., Balasubramanian, P., Athale, S., Domic, B., Murat, E., Banchereau, R., Xu, Z., Chandra, M., Chung, C.H. et al. (2019) A CD4(+) T cell population expanded in lupus blood provides B cell help through interleukin-10 and succinate. *Nat. Med.*, **25**, 75–81.
 50. Edgar, R., Domrachev, M. and Lash, A.E. (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30**, 207–210.