Role of the CTCF Binding Site in Human T-Cell Leukemia Virus-1 Pathogenesis

Ancy Joseph[1], Xiaogang Cheng[1], John Harding[1], Jacob Al-Saleem[4,5], Patrick Green[4,5], Deborah Veis[1,2], Daniel Rauch[1], Lee Ratner[1,3]

Departments of Medicine[1], Pathology & Immunology[2], and Molecular Microbiology[3], Washington University School of Medicine, St Louis, MO, USA

Center for Retrovirus Research[4] and Department of Veterinary Biosciences[5], The Ohio State University, Columbus, OH 43210

Corresponding Author: Lee Ratner MD PhD

660 S Euclid Ave, Box 8069

St Louis, MO 63110

Fax: 314-747-2120

Telephone: 314-362-8836

Email: lratner@wustl.edu

Running Title: CTCF Regulation of HTLV-1

Key Words: CTCF; HTLV; DNA Methylation; Epigenome; ATL

Word Count

Abstract: 291 words

Authors Summary: 115 words

Manuscript: 5044 words

Figures: 7

Supplemental Figures: 12

Supplemental Tables: 1

Abstract

During HTLV-1 infection, the virus integrates into the host cell genome as a provirus with a single CCCTC binding protein (CTCF) binding site (vCTCF-BS), which acts as an insulator between transcriptionally active and inactive regions. Previous studies have shown that the vCTCF-BS is important for maintenance of chromatin structure, regulation of viral expression, and DNA and histone methylation. Here, we show that the vCTCF-BS also regulates viral infection and pathogenesis *in vivo* in a humanized (Hu) mouse model of adult T-cell leukemia/lymphoma. Three cell lines were used to initiate infection of the Hu-mice, i) HTLV-1-WT which carries an intact HTLV-1 provirus genome, ii) HTLV-1-CTCF, which contains a provirus with a mutated vCTCF-BS which abolishes CTCF binding, and a stop codon immediate upstream of the mutated vCTCF-BS which deletes the last 23 amino acids of p12, and iii) HTLV-1-p12stop that contains the intact vCTCF-BS, but retains the same stop codon in p12 as in the HTLV-1-CTCF cell line. Hu-mice were infected with mitomycin treated or irradiated HTLV-1 producing cell lines. There was a delay in pathogenicity when Hu-mice were infected with the CTCF virus compared to mice infected with either p12 stop or WT virus. Proviral load (PVL), spleen weights, and CD4 T cell counts were significantly lower in HTLV-1-CTCF infected mice compared to HTLV-1-p12stop infected mice. Furthermore, we found a direct correlation between the PVL in peripheral blood and death of HTLV-1-CTCF infected mice. In cell lines, we found that the vCTCF-BS regulates Tax expression in a time-dependent manner. The scRNAseq analysis of splenocytes from infected mice suggests that the vCTCF-BS plays an important role in activation and expansion of T lymphocytes *in vivo*. Overall, these findings indicate that the vCTCF-BS regulates Tax expression, proviral load, and HTLV pathogenicity *in vivo*.

Author Summary

Human T-cell leukemia virus type 1 (HTLV-1) is a cause of leukemia and lymphoma, and several inflammatory medical disorders. The virus integrates in the host cell DNA, and it includes a single binding site for a cellular protein designated CTCF. This protein is important in regulation of many viruses, as well as properties of normal and malignant cells. In order to define the role of CTCF in HTLV-1 pathogenesis *in vivo*, we analyzed a mutant virus lacking the binding site in humanized mice. We found that this mutation slowed virus spread and attenuated the development of disease. Gene expression studies demonstrated a dynamic role of CTCF in regulating viral gene expression and T lymphocyte activation.

Introduction

Human T-cell leukemia virus type-1 (HTLV-1) is the cause of adult T-cell leukemia/lymphoma (ATLL) [1]. HTLV-1 is a delta-retrovirus which encodes plus (+) strand classical retrovirus genes, *gag, pol, pr, env*, as well as regulatory genes, *tax* and *rex*, auxillary genes, *p12, p30,* and *p13*, and minus (-) strand gene, *hbz.* The *tax* and *hbz* gene products both have oncogenic activity in tissue culture and mouse models [2, 3]. The Tax protein enhances viral and cellular gene transcription, and it has post-transcriptional roles inhibiting apoptosis and DNA repair, and promoting cellular proliferation [3]. Tax is expressed intermittently in a small proportion of ATLL cells at any given time [4, 5]. The Hbz protein represses multiple transcriptional pathways, whereas the *hbz* RNA promotes T-cell proliferation [2]. Hbz is expressed continuously by most ATLL cells, and the Hbz protein is critical for viral persistence and disease development [6].

Most ATLL cells have a single copy of the provirus integrated at a wide variety of different chromosomal sites [7]. The 5'portion of the integrated provirus is heavily DNA methylated with histone post-translational modifications consistent with epigenetic silencing [8]. In contrast, the 3'portion of the provirus exhibits little DNA methylation and has characteristic histone modifications of open chromatin. At the border is a binding site for the chromatin barrier element known as 11-zinc finger protein or CCCTC-binding factor (CTCF). There is a single viral CTCF-binding site (vCTCF-BS) in HTLV-1, which is conserved in other delta-retroviruses. In contrast, there are about 55,000 CTCF-binding sites in the cellular genome [9]. CTCF has been shown to have transcriptional suppression and DNA looping activity [10]. The latter is mediated through binding to the cohesin complex [11]. CTCF is important for regulation of latency, replication, and pathogenicity of many DNA viruses, including Kaposi sarcoma herpes virus (KSHV),Epstein-Barr virus (EBV), cytomegalovirus (CMV), herpes simplex virus (HSV), and adenovirus [12].

Studies of cell lines and primary cells infected with HTLV-1 *ex vivo* showed that the vCTCF-BS modulates transcription of the viral genome and cellular genes within several hundred bases of the provirus [8, 13, 14]. In order to assess the role of the vCTCF-BS in HTLV-1 replication and

pathogenesis, we examined the effect of vCTCF-BS mutation *in vivo*, using a humanized mouse model (Hu-mice). For this purpose, we used non-obese diabetic *scid* IL2 receptor gamma c null kit (NBSGW) mice injected intrahepatic with human cord blood CD34+ hematopoietic stem cells within the first three days of life, and allowed to engraft without irradiation [15]. Infection of these mice at 13-16 weeks of age perturbs human thymic alpha-beta T-cell development, resulting in expansion in the thymus of mature single-positive CD4+ and CD8+ lymphocytes at the expense of immature and double-positive (DP) thymocytes (Fig 1A) [16]. Human lymphocytes from the thymus, spleen, and lymph nodes are activated in this model, with increased expression of nuclear factor kappa-B (NF-□B)-dependent genes. These mice manifest hepatosplenomegaly, lymphadenopathy, and lymphoma.

Results

Role of vCTCF-BS in HTLV-1 Replication in Humanized Mice

Hu-mice, 13-16 weeks of age, were assessed for levels of human leukocytes in the peripheral blood by FACS analysis with an antibody to human CD45. Mice from each litter with at least 5% human CD45+ cells, were separated into equal groups based on sex and levels of human CD45+ cells. Mice were then inoculated intraperitoneally with lethally mitomycin-treated human 729B lymphoid cells infected with either wild type HTLV-1 (HTLV-1-WT) or an HTLV-1 mutant with a premature stop codon in the p12-coding gene that does not affect known function of p12, 24 codons from the 3'end of the 297 nucleotide long gene (HTLV-1-p12stop) [14]. An equal number of mice were also infected with HTLV-1 with the same mutation found in the HTLV-1-p12stop virus, as well as an additional mutation that abrogates the vCTCF-BS (HTLV-1-CTCF) [14]. Use of the p12stop mutant virus was a necessary control since the vCTCF-BS overlaps the p12 coding sequence, and mutation of the vCTCF-BS would also produce a mutation in the p12 protein if it had not been truncated. Mice were monitored clinically for up to 12.5 weeks after infection, with analysis of blood samples obtained every 2.5 weeks. Eight litters of mice were used to obtain a

sufficient number of mice for statistically valid results, with similar numbers, sexes, and levels of CD34+ cells in mice within each litter allocated for infection with HTLV-1-WT, HTLV-1-p12stop, or HTLV-1-CTCF.

Hu-mice infected with WT or p12 stop HTLV-1 showed disease development with a median survival of 5.0 and 4.3 weeks, respectively, whereas only 35% of the HTLV-1-CTCF mice developed disease within 12.5 weeks of infection (Figure 1B). Mice infected with HTLV-1-CTCF had lower mean spleen weights (162 vs 302 mg, P=0.0074; Figure 1C) and lower absolute lymphocyte counts at the time of necropsy (1038 vs 7600 cell/ul, P=0.03; Figure 1D) than those infected with HTLV-1-p12stop. In addition, the percent of CD4+ per CD45+ cells in the spleen, liver, and bone marrow were significantly lower in HTLV-1-CTCF infected than HTLV-1-p12stop infected mice (Figure 1E). No significant differences were seen in the proportion of CD4+ per CD45+ cells in the blood and tissues in comparison of HTLV-1-WT and HTLV-1-p12stop infected mice (Figure 1E).

Mice were also humanized via intratibial injection of CD133 hematopoietic progenitor cells, and infected 13-16 weeks later (Figure 1A). Infection with HTLV-1-CTCF (n=6) resulted in delayed onset of lymphoproliferative disease compared to mice infected with HTLV-1-WT (n=7) or HTLV-1-p12stop (n=3) (Figure S1A). At the time of necropsy there were no significant differences in spleen weight (Figure S1B), absolute lymphocyte counts (Figure S1C), or percentages of CD4+ lymphocytes in the liver in HTLV-1-CTCF compared to HTLV-1-WT infected mice (Figure S1D). However, significant differences in CD4+ lymphocytes were found in the blood, spleen, and bone marrow when compared to combined HTLV-1-WT and -p12stop infected CD133 humanized mice as compared to HTLV-1-CTCF infected mice. Differences in lymphocyte and neutrophil percentages were seen at necropsy in HTLV-1-CTCF compared to HTLV-1-WT infected mice (Figures S2A, B). There were insufficient mice infected with HTLV-p12stop for statistical analysis in these experiments. An example of a HTLV-1-WT infected Hu-mouse with ATLL-like flower cells and dramatic leukocytosis is shown in Figure S2C.

Proviral load (PVL) in CD34+ cell humanized mouse PBMCs were assessed by digital droplet PCR every 2.5 wks [17]. The assay measures the number of copies of provirus using primers within the *tax* gene, normalized to number of copies of human ribosomal P subunit p30 gene (Figure 2). At 2.5 wks post-infection, PVL levels in HTLV-1-WT and HTLV-1-p12stop infected mice varied between 0.1 and 1.8 copies/cell, with average levels of 0.43 and 0.78 copies/cell that were not significantly different (Figure 2A, B). However, in HTLV-1-CTCF mice, levels were between 0 and 0.02 copies/cell, with average level 0.02 copies/cell that was significantly lower than that in the HTLV-1-WT and HTLV-1-p12stop infected mice (P=0.0002). PVL was also measured at the time of necropsy in blood, spleen, liver, and bone marrow samples (Figure 2C). In each case, PVL was lower in HTLV-1-CTCF than HTLV-1-p12stop and HTLV-1-WT infected mice. The one exception was that the PVL was lower in HTLV-1-WT than HTLV-1-p12stop infected mouse liver, but the number of animals available for this analysis were small, and this may have been due to a sampling error as a result of heterogeneous levels of virus infected cells within the liver. Similarly, in CD133+ cell humanized mice, lower proviral loads were seen in HTLV-1-CTCF infected mice compared to HTLV-1-WT infected mice (Figure S3).

Viral gene expression was observed in infected splenocytes from 7 animals with lymphoproliferative disease submitted for single cell (sc) RNAseq (Figure S4). Although 10X scRNAseq reads are not strand specific, splice donor and splice acceptor sites for single-spliced and double-spliced, sense-strand transcripts and for spliced anti-sense transcripts enabled delineation and quantitation of *tax* and *hbz* transcripts in a subset of 4 representative samples (Figure S4A) compared to spliced transcripts of human actin in each sample (Figure S4B). Interestingly, unlike in infected cells in tissue culture [14], *hbz* transcripts were the most abundant viral mRNAs in infected splenocytes *in vivo* (Figure S4C) for both HTLV-1-CTCF (n=3) and HTLV-1-p12stop infected Hu-mice (n=4). There were no significant differences in the percent of TCR+ cells that are *hbz*+ in HTLV-1-CTCF compared HTLV-1-p12stop infected splenocytes (6.4±3.3 vs 4.6±1.6%, respectively, P=0.36). The sequence reads also provided confirmation that the

nucleotide substitutions used to create the HTLV-1-CTCF and HTLV-1-p12stop viruses were present in 100% of viral transcripts, and that reversion back to the WT sequence did not occur (Figure S4D).

Role of vCTCF-BS in HTLV-1 Pathogenesis

Pathological analysis of infected mice that succumbed from infection demonstrated a lymphoproliferative disorder, with diffuse infiltration in the spleen, liver, and lungs (Figures 3A, S5). The infiltrating cells were found to be predominantly CD4+ lymphocytes, as demonstrated by immunohistochemistry (Figure 3B). In comparison, control human tissues are shown highlighting CD4+ lymphocytes in normal tonsil, and in biopsies from lymphomatous tissues from two different HTLV-positive patients with ATLL.

The results for HTLV-1-CTCF mice were stratified into two groups, depending upon whether the blood absolute lymphocyte count at time of necropsy was greater than (HTLV-1-CTCF-1; n=8) or less than 400 cells/μl (HTLV-1-CTCF-2; n=7; Figure 4A). Levels of total lymphocytes at necropsy were higher in the blood of HTLV-1-CTCF-1 than HTLV-1-CTCF-2 infected mice, and similar to those seen in HTLV-1-WT and HTLV-1-p12 infected mice at the time of necropsy (Figure S6). In contrast, the number of neutrophils in the blood were higher at necropsy in HTLV-1-CTCF-2 than HTLV-1-CTCF-1 infected mice. Thirty eight percent of the HTLV-1-CTCF-1 infected mice died by 7.5 weeks post-infection, whereas no disease was seen in HTLV-1-CTCF-2 infected mice (Figure 4B). Furthermore, no pathological abnormalities were seen in HTLV-1-CTCF-2 infected mouse tissues. Spleen weight at necropsy (Figure 4C) and PVL was higher at 2.5 and 5 weeks post-infection in HTLV-1-CTCF-1 compared to HTLV-1-CTCF-2 infected mice (Figure 4D). The ratio CD4+ to CD45+ cells at necropsy in the blood, spleen, liver, and bone marrow were higher in HTLV-1-CTCF-1 than HTLV-1-CTCF-2 infected mice (Figure 4E). There was an inverse correlation between PVL at 5 wks and survival in HTLV-1-CTCF infected mice (p=0.0005; Figure 4F).

Effects of vCTCF-BS on Transcriptomic Profiles

In order to assess the effect of the vCTCF-BS on transcription *in vivo*, we performed single cell RNA-seq on splenocytes from 4 HTLV-1-p12stop and 3 HTLV-1-CTCF-2 humanized mice (Figure 5A). We focused our analysis on human transcripts, which were clearly separated from murine transcripts [18]. Although the 729B cells, used as a donor in humanized mice for HTLV-1 infection, contain the EBV genome [19], no EBV transcripts were detectable in the HTLV-1 infected humanized mice at the time of necropsy (data not shown). Human cells in the spleen of the hematopoietic stem cell transplanted NBSGW mice were exclusively lymphoid cells. They included clusters of CD4+, CD8+, CD25+, TCR+, and more rarely, NKT, and B lymphocytes (Figure 5B). Interestingly, the TCR+ cells of each T cell subset were enriched in T cell activation factors, protein tyrosine phosphatase receptor type C-associated protein (PTPRCAP) and interferon-induced transmembrane protein 1 (IFITM1) (Figure S7).

At the time of necropsy, reads corresponding to HTLV-1 transcripts were detected in a small subset of splenocytes (Figure 5B). Characteristics of predominant T cell clones, including CD4+, CD8+, Treg, and double CD4+CD8+ clones, varied in each sample, and detection of viral transcripts was not restricted to a single T cell clone. T cell clonality was high for all samples, with the Gini coefficient ranging from 0.54 to 0.67, and the Shannon Diversity Index ranging from 3.7 to 5.2 (Figure S8). There were no significant differences in clonality indices in HTLV-1-p12stop and HTLV-1-CTCF-2 infected humanized mouse splenocytes.

String network analysis of genes significantly upregulated in CD4+ T cells demonstrated expanded lymphocyte populations in both HTLV-1-CTCF and HTLV-1-p12stop infected mice (Figure 5C). The cells expressed genes previously associated with ATLL, including CADM1, IL2RA, FOXP3, BATF3, CD28, and CTLA4.

When comparing HTLV-1-CTCF-2 infected mice to control HTLV-1-p12stop infected mice, several patterns emerged. First, activated CD8+ T cells were much more abundant in the spleen

of HTLV-1-p12stop infected mice (Figure 5D). Granzyme B (GZMB) expressing CD8+ cells were enriched in transcripts for natural killer cell granule protein 7 (NKG7) and granulysin (GNLY), indicating these cells were cytotoxic effector T cells or cytotoxic vesicle releasing cells (Figure S9) [20]. Second, the expression of 5-lipoxygenase activating protein (ALOX5AP), a regulator of tumor immunity associated with "hot" tumors [21], was a distinguishing characteristic between mice infected HTLV-1-CTCF-2 and HTLV-1-p12stop (Figures 5E, S10). Third, double negative CD4-CD8-TCR+ T cells were more abundant in the spleens of HTLV-1-CTCF-2 mice (Figure S11). Although these cells lacked transcripts for CD4 or CD8, they were enriched in transcripts for calcium-binding helix-loop-helix S100A protein family members (including S100A4, A6, A10, and A11) and interferon-induced transmembrane protein (IFITM) family members (IFITM1 and M2) [22]. Finally, the number of cells in T cell clones carrying a glycine-leucine-alanine-glycine (GLAG) motif in CDR3 of T-cell receptor (TCR) β, previously identified as a Tax-specific epitope [23] was more abundant in HTLV-1-CTCF-2 than HTLV-1-p12stop infected mice (Figure 5F, Table S1). These data suggest that loss of the vCTCF-BS results in significant effects on gene expression and expansion of human T cell populations *in vivo*.

Effects of vCTCF-BS on Temporal Viral Gene Expression

In order to examine the effect of vCTCF-BS on the temporal expression of Tax, we transfected 293T cells with the molecular clones expressing HTLV-1-WT, HTLV-1-p12stop, or HTLV-1-CTCF (Figure 6A). After 48 hrs, these cells were producing equivalent quantities of HTLV-1 p19 antigen (5.6±1.2 and 5.4±1.0 ng/ml, in HTLV-1-p12stop and -CTCF expressing 293T cells, respectively). The transfected 293T cells were co-cultivated with Jurkat cells carrying a Tax-responsive red fluorescent protein (RFP) indicator (JET cells). IncuCyte analysis was performed to assess temporal changes in Tax expression, as measured by RFP fluorescence (Figure 6B, C). The number of RFP-positive cells was significantly higher in HTLV-1-p12stop than HTLV-1-CTCF infected cultures from 1-3.5 days of culture, but similar thereafter (Figure 6B). The total RFP

intensity was greater in HTLV-1-p12 stop than HTLV-1-CTCF infected cultures from 0-2.5 days of infection, but lower during 2.5-5 days of infection (Figure 6C). Cell viability was assessed using Cytolight rapid dye, and no differences were detected. Similar results were obtained after cocultivation of HTLV-1 infected 729B cells with JET cells (Figure S12). These results suggest that the vCTCF-BS has dynamic regulation of HTLV-1 gene expression.

Discussion

Our previous studies of the role of the vCTCF-BS examined in Jurkat cells and PBMCs, the role of mutation of the vCTCF-BS on virus gene expression [14]. We found that mutation of the vCTCF-BS did not disrupt the kinetics and levels of virus gene expression. Furthermore, there was no effect on the establishment of or reactivation from latency. Nevertheless, the mutation disrupted the epigenetic barrier function, resulting in enhanced DNA CpG methylation downstream of the vCTCF-BS on both strands of the integrated provirus. We also found enhanced methylation of histones H3 K4, K27, and K36 bound to the provirus.

In our previous study, we also examined the role of CTCF in clonal latently infected Jurkat cell lines carrying the HTLV-1 provirus at different integration sites [14]. For this purpose, we induced viral gene expression with phorbol ester and ionomycin in the presence of a shRNA to repress CTCF expression or a control shRNA. In the majority of these cell lines, knockdown of CTCF resulted in enhanced plus strand gene expression. However in a minority of cell lines, knockdown of CTCF had no effect on plus strand gene expression. We did not identify cell lines in which knockdown of CTCF decreased virus gene expression. Knockdown of CTCF had no effect on virus gene expression from cell lines with mutation of the vCTCF-BS. Moreover, no effects on minus strand gene expression were seen in any of these cell lines. We found that cell lines manifesting enhanced plus gene expression with CTCF knockdown also exhibited decreased DNA CpG methylation downstream of the CTCF binding site. However, no significant changes

were seen in DNA CpG methylation in cell lines not exhibiting alterations of gene expression with CTCF knockdown.

In the current work, we examined the effects of vCTCF-BS mutation *in vivo* in a humanized mouse model [16]. In this model, human CD34+ cells were injected into the liver of newborn mice. After 13-16 weeks, sufficient lymphoid reconstitution occurred to allow HTLV-1 infection, replication, and lymphoproliferative disease. However, we have not detected HTLV-1 antibodies in this model system, suggesting at least partially compromised immune responses to viral infection (not shown).

The vCTCF-BS overlaps with the p12 and Hbz coding genes. Mutation of the vCTCF-BS to abrogate binding of CTCF required conservative mutations in these overlapping genes. The p12 mutation truncates the predicted protein product from 99 to 76 amino acids. This truncated protein is similar to that expressed from simian T cell leukemia virus type 1 [24]. Previously, we demonstrated that deletion of the C-terminus of p12 did not affect its ability to functionally enhance nuclear factor of activated T cells (NFAT) [14]. We also showed that the conservative mutation in Hbz had no effect on its ability to repress Tax-mediated viral *trans*-activation or canonical NFκB activity. The mutations in p12 and Hbz used in this study did not have a significant effect on HTLV-1 replication and pathogenicity in Hu-mice, based on similar results with HTLV-1-WT and HTLV-1-p12stop infected animals.

Mutation of the vCTCF-BS delayed virus spread and delayed or abrogated lymphoproliferative disease in infected Hu-mice (Figure 1). However, the lymphoproliferative disease occurring at late time points in the minority of HTLV-1-CTCF infected Hu-mice were derived from CD4+ lymphocytes as in the case of HTLV-1-WT and HTLV-1-p12stop mice (Figure 3). There was no qualitative change in the characteristics of the lymphoproliferative disease occurring in HTLV-1-CTCF-1 infected Hu-mice compared to that present in HTLV-1-WT and HTLV-1-p12stop infected mice.

Single cell RNAseq is a powerful tool for evaluating human lymphocytes within the spleen of infected Hu-mice. The presence of human T cell subsets confirmed that CD34+ hematopoietic stem cells were capable of differentiating into CD4+ and CD8+ T cells *in vivo*. Mature (TCR+) cells were consistently enriched in T cell activation factors PTPRCAP and IFITM1. PTPRCAP is a transmembrane phosphoprotein specifically associated with CD45, a key regulator of T cell activation and differentiation. Along with CD45, CD71, and lymphocyte-specific protein tyrosine kinase (LCK), PTPRCAP (also known as lymphocyte phosphatase-associated phosphoprotein, LPAP) is known to be a major component of the CD4 receptor complex [25]. IFITM1 is a member of a family of interferon-inducible transmembrane proteins that can confer resistance to viral infections, regulate adaptive immunity, and regulate T cell differentiation [22]. Multi-omic evaluation of TCR sequences offered clear evidence of extensive clonal T cell expansion in this model, established that viral gene expression could be detected in expanded clones, and confirmed that the expanded CD4+ T cells were enriched in genes frequently expressed in ATLL cells, including CD25 and cell adhesion molecular 1 (CADM1).

Surprisingly, the most significant difference in the spleens of Hu-mice infected with virus carrying the vCTCF-BS mutation was discovered in the CD8+ T cell population in which the abundance and activity of CD8+ T cells was suppressed relative to control. There were fewer CD8+ T cells and the CD8+ T cells expressed less Granzyme B and less ALOX5AP. Granzyme B is a serine protease and abundant component of cytotoxic granules which when released results in caspase-independent pyroptosis or caspase-dependent apoptosis [26]. Granzyme B is an essential component of immunity and wound healing, and it is also capable of causing injury to healthy tissue or even elevated risk of death [27]. ALOX5AP (aka FLAP) is required for leukotriene synthesis; it has been implicated in inflammatory responses, stroke, and myocardial infarction; and it is an indicator for predicting high CD8+ tumor infiltration and a “hot” tumor microenvironment [21]. These data establish that scRNAseq can effectively detect human and viral gene expression in mouse spleen in infected Hu-mice, confirm that the expanded lymphocyte

populations in this model retain characteristics similar to those described in ATLL, and supports the hypothesis that the vCTCF-BS is involved in viral regulation of immunity and pathogenesis in vivo.

The results of the current study suggest that CTCF binding to the HTLV-1 provirus regulates Tax expression in a time-dependent manner (Figure 6, 7). Initially, CTCF promotes higher levels of Tax, which results in enhanced plus strand gene transcription, antigen expression, virus production, and enhanced clonal expansion of lymphocytes. However, high levels of Tax and other plus strand genes are associated with enhanced senescence, apoptosis, and immune-mediated responses to virus-infected cells. We conjecture that this results in rapid onset of acute disease in Hu-mice infected with HTLV-1 possessing the vCTCF-BS. In the absence of the vCTCF-BS, there are lower levels of Tax, plus strand gene expression, and virus production, diminished senescence and cytotoxicity, and more gradual lymphocyte expansion resulting in slower development of disease, if disease develops at all.

The mechanism for the effects of CTCF on HTLV-1 transcription could be related to its known silencer effects on initiation or elongation of RNA. This may be a result of monomeric CTCF binding to the provirus or dimeric CTCF-cohesin complexes promoting chromatin looping. High levels of Tax and plus strand transcription at early time points after infection, promote clonal expansion of infected lymphocytes and enhanced viral particle and viral antigen production. Under these conditions, Tax has been shown to induce cellular senescence [28], whereas the viral envelope may induce fusion [29], and multiple viral proteins can induce cell death through direct effects or through immune-mediated cytotoxicity [30]. In contrast, in the absence of CTCF binding to the provirus, there is a more gradual level of Tax expression, resulting in delayed onset or smoldering disease, or asymptomatic infection. It is possible that CTCF activity may contribute to differences in disease subtypes seen in infected patients.

Previous studies by Bangham and colleagues, reported that removal of the vCTCF-BS had no discernible impact on virus transcription or epigenetic modifications in 2 different cell clones of

HTLV-1 infected lymphocytes [31]. However, mutation of the vCTCF-BS resulted in altered clone-specific transcription in *cis* at non-contiguous loci up to more than 300 kb from the integration site, suggested to be due to disruption of chromatin loops [13]. Our previous studies contrast with those of Miura *et al*, in that we examined the effects on viral transcription with loss of vCTCF-BS in a large number of cells *ex vivo* and *in vivo.* In our previous studies with infected rabbits, mutation of the vCTCF-BS did not affect virus replication or spread [32]. However, it is notable that there was a decreased HTLV-1-specific antibody response in this model. Perhaps HTLV-1-specific antibody responses reflect the peak levels of plus strand transcription after infection. It is notable that lymphoproliferative disease does not occur in the rabbit model.

These data suggest a working model in which CTCF regulates Tax, that assimilates the *in vivo* survival data, and that HTLV-1-CTCF is much less pathogenic than HTLV-1-WT or HTLV-1-p12stop. The scRNAseq gene expression data of 12.5 week old diseased HTLV-1-CTCF-2 infected mice look very similar to 2.5 week old diseased HTLV-1-p12stop mice. The IncuCyte data showed that Tax activity per infected cell is elevated on day 1 and depressed on day 3 in pHTLV-1-p12stop but not HTLV-1-CTCF infected cells, and the number of Tax-positive cells shows rapid expansion in HTLV-1-p12stop but not HTLV-1-CTCF infected mice. Previously published data on latent cell lines show that CTCF is a suppressor of Tax in latent or already suppressed cell clones. This model also explains the observation that although HTLV-1-CTCF and HTLV-1p12stop virus particles are equally infectious, HTLV-1-p12stop infected Hu-mice develop much higher viral loads in the peripheral blood at an accelerated rate compared to HTLV-1-CTCF infected Hu-mice, because the vCTCF-BS present in HTLV-1p12stop infected Hu-mice is an enhancer of Tax and virus production.

In summary, our results demonstrate an important role of CTCF binding to the HTLV-1 provirus in dynamic regulation of virus replication and pathogenicity, and support a potentially new discovery that CTCF regulates TAX *in vivo.*

Materials and Methods

Plasmids and Sources of Cells

The infectious HTLV-1 ACH wild type and mutant clones pHTLV-1(WT), pHTLV-1(p12stop) and pHTLV-1(CTCF) were used in this study [14]. JET cells (JET WT35) and stable 729B HTLV-1 producer cell lines, 729B-HTLV-1(WT), 729B-HTLV-1(p12stop) and 729B-HTLV-1(CTCF) were described in our previous study [14].

Mononuclear cells were isolated by density gradient centrifugation using Ficoll-Paque premium (Sigma Aldrich) and 50 ml SepMate tubes (STEM CELL Technologies) according to the manufacturer's protocol, and CD34+ hematopoietic stem cells (HSCs) were isolated from these mononuclear cells using CD34 microbead kit (Miltenyi Biotec CD34 MicroBead Kit, Human). The purity of isolated CD34+ cells were accessed by flow cytometry using mouse anti human CD34 (BD Bioscience).

Ethics statement

All the experiments in mice were performed in accordance with ethical and regulatory standards set by NIH for animal experimentation. The animal use protocol (20180321) was approved by Washington University Department of Comparative Medicine. Cord blood samples obtained in this study were obtained from Cleveland Cord Blood Center (CCBC). Informed consent was obtained from all the donors.

Generation of CD34+ humanized mice (HuMice)

NBSGW (NOD.Cg-KitW-41J Tyr + Prkdcscid Il2rgtm1Wjl/ThomJ), hereafter referred as NSGBW mice were purchased from Jackson laboratories. All mice were kept in animal housing in a pathogen-free environment with ambient temperature, humidity and controlled light cycles. The NSGBW mice breeding colonies were produced in house. After birth, 0-3 day old pups were injected, using a 27 gauge insulin syringe, intra-hepatically with $5x10^4$ CD 34+ hematopoietic stem

cells (HSCs); which were isolated from cord samples collected from full term deliveries (Miltenyi Biotec CD34 MicroBead Kit, Human). Human $CD45^+$ levels were assessed at 13-16 weeks post transplantation (wpt) by flow cytometry analysis.

Generation of CD133+ humanized mice (HuMice)

After birth, 4- 5 weeks old pups were anesthetized and each mouse was injected with $5x10^4$ CD 133+ hematopoietic stem cells (HSCs), by intra tibial injection, which were isolated from cord samples collected from full term deliveries (Miltenyi Biotec CD133 MicroBead Kit, Human). Human $CD45^+$ levels were assessed at 13-16 weeks post transplantation (wpt) by flow cytometry analysis.

Cell culture and infection with HTLV-1

Stable 729B HTLV-1 producer cell lines: WT HTLV-1(WT), HTLV-1(CTCF), or HTLV-1(p12stop), which were generated previously [14], were used in this study. Cell lines were maintained in RPMI media (Sigma) supplemented with 10% Fetal Bovine Serum ,100 u/ml Penicillin and Streptomycin (Gibco). One million cells/ml were plated in 12 well plates and HTLV-1 p19 antigen in the supernatant was assessed (ZeptoMetrix HTLV p19 Antigen ELISA kit) after 24 hrs. of culture according to the manufacture's protocol. Based on p19 values, cell numbers corresponding to 70 ng p19 /mice were used for infection. Before infection HTLV -1 producing cell lines were treated for 90 min with 20 ug/ml of mitomycin C (Sigma Aldrich) to inhibit replication/proliferation of producer cells. Mice were monitored for a period of 12.5 weeks post infection (wpi) for signs of disease. Mice were anesthetized and necropsied when the body weight dropped by 20% or more of their initial body mass prior to infection. Blood, bone marrow, spleen, liver, tumors, and enlarged lymph nodes were collected at the time of necropsy. Complete blood counts (CBC) and Giemsa staining was performed on the peripheral blood smears at the time of necropsy.

Flow Cytometry

Peripheral blood was collected by mandibular cheek bleed every 2.5 weeks post infection and at time of necropsy by cardiac puncture after anesthesia with 100mg/kg ketamine and 20mg/kg xylazine). Single cell suspensions were made from spleen and liver by, crushing the organs using a wide 1ml tip and then passing the cell suspension through a sterile 100µm mesh. PBS supplemented with 2% FBS was used as media. Bone marrow was collected from both femurs by dissection, and then flushing the bones with PBS. All the collected cells were treated with RBC lysis buffer (Sigma-Aldrich) and stained using PE mouse anti-human 45 (BD bioscience) and APC mouse anti-human CD4 antibodies (BD Bioscience). Flow cytometry was performed using BD FACScan (BD Biosciences) and data was analyzed using FlowJo software.

DNA isolation and proviral load analysis

DNA was extracted from peripheral blood and bone marrow by conventional phenol-chloroform method and Blood and Tissue kit (Qiagen) was used to extract DNA from spleen and liver. A minimum of 50ng of DNA was used to quantify proviral load. Proviral loads were measured by digital droplet PCR as previously described [14, 17].

Histopathologic analysis

Tissue samples were fixed using neutral buffered formalin (Fisher Scientific) for 24 hours, parafilm embedded and stained with Hematoxylin and Eosin (H&E) solution To detect the presence of human CD4+ cells, immunohistochemistry was performed using anti-CD4 (SP35) rabbit monoclonal primary antibody (Ventana Medical systems) according to manufacturer's instruction with slight modification in cell conditioning for 64 min followed by antibody incubation for 40 min at 36 °C. CD4 staining was performed using BenchMark Ultra staining module. Stained

sections were observed under a light microscope, and images of whole sections were captured (Nanozoomer) and viewed using the NDP 2.00 viewer (Hamamatsu, Japan).

JET cell infection, imaging and analysis using IncuCyte system

HEK293T cells were transfected with 2µg of either pHTLV-1(p12stop) or pHTLV-1(CTCF) using 1mg/ml polyetheliamine (PEI40K, Polyscience) by a ratio of 3:1 plasmid concentration. After 48 hrs, cells were irradiated (30 Gy) and co-cultured with JET cells [33] and placed in an IncuCyte live cell S3 analysis system (Sartorius). The cells were then continuously imaged for RFP every 3 hrs for 5 days. The IncuCyte software was used to calculate the read mean intensity and total red object count.

Single cell RNAseq and analysis

Samples of viably cryopreserved mouse splenocytes stored in liquid nitrogen were retrieved immediately before sample processing and submission. Cells were thawed partially in a 37ºC water bath and then placed on ice immediately. Single cell suspensions were revived in ice cold medium by gently adding the cell suspension to 10 ml of RPMI medium supplemented with 10% FBS. Cells were centrifuged and gently washed with PBS with 2% FBS and passed through a 70 µM cell strainer to avoid clumps while processing samples. Cells were stained with 0.4% trypan blue to quantify viability and submitted to the McDonnell Genome Institute for processing for scRNAseq using the 10X Genomics 5’ GEX plus TCR V(D)J enrichment. Reads were mapped to the human genome, the HTLV-1 genome, and the EBV genome and data files obtained included 10x Cell Ranger scRNAseq-FASTQ files, Cell Ranger output, and matrix files. Custom analysis, differential expression, and creation of feature plots was performed using Loupe Browser.

Statistics

P values were determined by unpaired t tests. using GraphPad Prism version 10.0.0 for Windows, GraphPad Software (Boston, Massachusetts). The correlation method (www.graphpad.com. Spearman) was used to determine statistically significant correlation between proviral load and survival curve. Results were considered to be significant if the p value was ≤0.05 (* indicates p value ≤ 0.05 , ** for ≤0.01, *** for ≤ 0.001, **** for ≤ 0.0001).

Acknowledgements

This work was supported by Public Health Service grants to L.R. (R01 CA258359, R21 CA252869) and P.G. (P01 CA100730). We thank the Alvin J. Siteman Cancer Center at Washington University School of Medicine and Barnes-Jewish Hospital in St. Louis, MO. and the Institute of Clinical and Translational Sciences (ICTS) at Washington University in St. Louis, for the use of the Genome Technology Access Center, which provided scRNAseq services and support. The Siteman Cancer Center is supported in part by an NCI Cancer Center Support Grant CA091842 and the ICTS is funded by the National Institutes of Health's NCATS Clinical and Translational Science Award (CTSA) program grant #UL1 TR002345.

Financial Statement

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Bangham CRM, Matsuoka M. Human T-cell leuekemia virus types 1 and 2. In: Howley PM, Knipe DM, Whelan S, Freed EO, Cohen JL, editors. Fields Virology. RMA Vorises. Philadelphia, PA: Lippincott Williams & Wilkins; 2023. p. 527-57.

2. Matsuoka M, Mesnard J-M. HTLV-1 bZIP factor; the key viral gene for pathogenesis. Retrovirology. 2020;17:2.

3. Mohanty S, Harhaj EW. Mechanisms of oncogeneis by HTLV-1 Tax. Pathogens. 2020;9:543.

4. Billman MR, Rueda D, Bangham CRM. Single-cell heterogeneity and cell-cycle-related viral gene bursts in the human leukaemia virus HTLV-1. Welcome Open Research. 2017;2:87.

5. Mahgoub M, Yasunaga JI, Iwami S, Nakaoka S, Koizumi Y, Shimura K, et al. Sporadic on/off switching of HTLV-1 Tax expression is crucial to maintain the whole population of virus-induced leukemic cells. Proceedings of the National Academy of Sciences, USA. 2018;115:E1269-E78.

6. Matsuoka M, Green PL. The HBZ gene, a key player in HTLV-1 pathogenesis. Retrovirology. 2009;6:71.

7. Cook LB, Melamed A, Niederer H, Valganon M, Laydon D, Foroni L, et al. The role of HTLV-1 clonality, proviral structure, and genomic integration site in adult T-cell leukemia/lymphoma. Blood. 2014;123:3925-31.

8. Satou Y, Miyazato P, Ishihara K, Yaguchi H, Melamed A, Miura M, et al. The retrovirus HTLV-1 inserts an ectopic CTCF-binding site into the human genome. Proceedings of the National Academy of Sciences, USA. 2016;113:3054-9.

9. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al. Analysis of the vertebrate insulator protein CTCF binding sites in the human genome. Cell 2008;128:1231-45.

10. Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell. 2009;137:1194-211.

11. Holwerda SJB, deLaat W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. Philosophical Transactions of the Royal Society of London. 2013;368:20120369.

12. Pentland I, Parish JL. Targeting CTCF to control virus gene expression: a common theme amongst diverse DNA viruses. Viruses. 2015;6:3574-85.

13. Melamed A, Yaguchi H, Miura M, Witkover A, Fitzgeerald TW, Birney E, et al. The human leukemia virus HTLV-1 alters the structure and transcription of host chromatin in cis. Elife. 2018;7:e36245.

14. Cheng X, Joseph A, Castro V, Chien-Liaw A, Skidmore Z, Ueno T, et al. Epigenomic regulation of human T-cell leukemia virus by chromatin-insulator CTCF. PLoS Pathogens. 2021;17:e1009577. PMID: 34019588.

15. Cosgun KN, Rahmig S, Mende N, Reinke S, Hauber I, Schafer C, et al. Kit regulates HSC engraftment across the human-mouse species barrier. Cell Stem Cell. 2014;15:227-38.

16. Villaudy J, Wencker M, Gadot N, Scoazec J-Y, Gazzolo L, Manz MG, et al. HTLV-1 propels thymic human T cell development in "human immune system" Rag2-/-IL-2Rgammac-/- mice. PLoS Pathog 2011;7:e1002231.

17. Brunetto GS, Massoud R, Leibovitch EC, Caruso B, Johnson K, Ohayon J, et al. Digital droplet PCR (ddPCR) for the precise quantification of human T-lymphotropic virus 1 proviral loads in peripheral blood and cerebrospinal fluid of HAM/TSP patients and identification of viral mutations. Journal of Neurovirology. 2014;20:341-51.

18. Callari M, Batra AS, Batra RN, Sammut S-J, Greenwood W, Clifford H, et al. Computational approach to discriminate human and mouse sequences in patient-derived tumour xenografts. BMC Genomics. 2018;19:19.

19. Glassy MC, Handley HH, Hagiwara H, Royston I. UCD 729-6, a human lymphoblastoid B-cell line useful for generating antibody-secreting human-human hybridomas. Proceedings of the National Academic of Sciences USA. 1983;80:6327-31.

20. Tewary P, Yang D, delaRosa G, Li Y, Finn MW, Krensky AM, et al. Granulysin activates antigen-presenting cells through TLR4 and acts as immune alarmin. Blood. 2010;116:3465-74.

21. Chen Y, Zeng C, Zhang X, Hua Q. ALOIX5AP is an indicator for high CD8 lymphocyte infiltration and "hot" tumor microenvironment in osteosarcomaA: a bioinformatic study. Biochemical Genetics. 2023;61:2363-81.

22. Yanez DC, Ross S, Crompton T. The ITITM protein family in adaptive immunity. Immunology. 2020;159:365-72.

23. Ishihara Y, Tanaka Y, Kobayashi S, Kawamura K, Nakasone H, Comyo A, et al. A unique T-cell receptor amino acid sequence selected by human T-cell lymphotropic virus type 1 Tax 301-309-speicific cytotoxic T cells in HLA-A24:02-positive asymptomatic carriers and adult T-cell leukemia/lymphoma patients. Journal of Virology. 2017;91:e00974-17.

24. Iniguez AM, Gastaldello R, Gallego S, Otsuki K, Vicente ACP. HTLV-1 p12-I protein sequences from South America: truncated proteins and common genetic signatures. AIDS Research and Human Retroviruses. 2006;22:466-9.

25. Krotov GI, Krutikova MP, Zgoda VG, Filatov AV. Profiling of the CD4 receptor complex proteins. Biochemistry. 2007;72:1216-24.

26. Cullen SP, Martin SJ. Mechanisms of granule-dependent killing. Cell Death & Differentiation. 2008;15:251-62.

27. Santos-Zas I, Lemaire J, Zlatanova I, Cachanado M, Seghezzi J-C, Benamer H, et al. Cytotoxic CD8+ T cells promote granzyme B-dependent adverse post-ischemic cardiac remodeling. Nature Communications. 2021;12:1483.

28. Zhi H, Yang L, Kuo YL, Ho YK, Shih HM, Giam CZ. NF-kB hyperactivation by HTLV-1 tax induces cellular senescence, but can be alleviated by the viral anti-sense protein HBZ. PLOS Pathogens. 2011;7:e1002025.

29. Delamarre L, Pique C, Pham D, Tursz T, Dokhelar M-C. Identification of functional regions in the human T-cell leukemia virus type 1 SU glycoprotein. Journal of Virology. 1994;68:3544-9.

30. Mohanty S, Harhaj EW. Mechanisms of innate immune sensing of HTLV-1 and viral immune evasion. Pathogens. 2023;12:735.

31. Miura M, Miyazato P, Satou Y, Tanaka Y, Bangham CRM. Epigenetic changes around the pX region and spontaneous HTLV-1 transcription are CTCF-independent. Wellcome Open Research. 2018;3:105.

32. Martinez MP, Al-Saleem J, Cheng X, Panfil A, Dirksen WP, Joseph A, et al. HTLV-1 CTCF-binding site is dispensable for in vitro immortalization and persistent infection in vivo. Retrovirolgy. 2019;16:44. PMID:31864373.

33. Fujisawa T, Toita M, Yoshida M. A unique enhancer element for the trans activator (p40tax) of human T-cell leukemia virus type I that is distinct from cyclic AMP- and 12-O-tetradecanoylphorbol-13-acetate-responsive elements. . Journal of Virology. 1989;63:3234-9.

Figure Legends

Figure 1. Infection of CD 34+ humanized mice resulted in decrease in pathogenicity in CTCF Infected mice. A. Schematic representation of CD34+ and CD133+ hematopoietic stem cell humanization and experimental flow. B. Survival curve of HTLV-1-WT, HTLV-1-p12stop, and HTLV-1-CTCF infected CD34+ Hu-mice. C. Spleen weights at time of necropsy/death were significantly lower in HTLV-1-CTCF compared to HTLV-1-p12stop infected mice. D. Absolute lymphocyte counts at time of necropsy in peripheral blood. E. Percentage of CD4+ T cells among total CD45+ cells in blood, spleen, and bone marrow cells (* indicates p value lower than 0.05; ** lower than 0.01: *** lower than 0.001).

Figure 2. Repressed proviral loads in Hu-mice infected with HTLV-1-CTCF. A. Individual Hu-mice infected with HTLV-1-WT or HTLV-1-p12stop had a high PVL at 2.5 wpi, whereas most mice infected with HTLV-1-CTCF had low PVL at all time points up to 10 wpi. B. Comparison of average PVL in HTLV-1-WT, HTLV-1-p12stop and HTLV-1-CTCF infected mice at 2.5 wpi. C. PVLs in blood, spleen, liver, and bone marrow cells of infected mice at time of necropsy. (* indicates p value lower than 0.05 ; ** lower than 0.01; *** lower than 0.001).

Figure 3. Histopathological changes in spleen and liver of infected Hu-mice. A. Hematoxylin and Eosin staining (original magnification 10 X) of spleen showing infiltrating lymphocytes in spleen (top panel), and liver showing lymphoid infiltration into the periportal, midzonal and centrilobular region in the liver (lower panel) Infiltration of lymphoid cells are marked with arrow. B. IHC stain for human CD4 inpatient samples as well as humanized mice infected with HTLV-1(original magnification 20 X). Controls for IHC include human tonsil, and samples from two patients who had HTLV-1-associated lymphoma, including a paratracheal lymph node (HTLV-

patient 1) and spheoid mass (HTLV-patient 2). The lower panel shows CD4 staining in HTLV-1-WT, HTLV-1-p12stop, and HTLV-1-CTCF infected Hu-mice.

Figure 4. Comparison of pathogenicity in HTLV-1-**CTCF Hu-mice based on absolute lymphocyte count**. A. HTLV-1-CTCF infected Hu-mice were separated into two groups based on absolute lymphocyte count at time of necropsy (Group 1 > 400 cells/µl; Group 2 < 400 cells/ul). B. The mice in Group 2 survived until the end point of the experiment. C. Spleen weights were significantly lower in HTLV-1-CTCF-2 than HTLV-1-CTCF-1 infected mice. D. Proviral load in blood was significantly lower in HTLV-1-CTCF-2 than HTLV-1-CTCF-1 infected mice at 2.5 and 5 wpi. E. Total CD4+ T cell counts in blood and spleen were significantly lower in HTLV-1-CTCF-2 than HTLV-1-CTCF-1 infected Hu-mice. F. Correlation of peripheral blood PVL and survival in HTLV-1-CTCF infected Hu-mice. (* indicates p value lower than 0.05 ; ** lower than 0.01; *** lower than 0.001).

Figure 5. vCTCF-BS determines the effect of HTLV-1 on survival and expansion of lymphocytes *in vivo.*

A. Single cell RNAseq was performed on splenocytes harvested from Hu-mice infected with HTLV-1-CTCF or HTLV-1-p12stop. B. TSNE plots of a representative data set confirm that TCR+ clusters overlap with T cell markers and viral gene expression, and that human CD4+ and CD8+ T cells clusters can be readily identified. C. String network (string-db.org) of genes significantly upregulated in CD4+ T cells confirm that expanded lymphocyte populations in both HTLV-1-CTCF and HTLV-1-p12stop infected Hu-mice express genes associated with ATLL, including IL2RA, FOXP3, BATF3, CD28, and CTLA4. Symbols within circles represent schematic protein structures. D. Heatmap comparing the relative abundance of CD4+ and CD8+ T cells in each sample, normalized against total TCR+ cells in each sample. E. Expression of ALOX5AP shown

as the ratio of ALOX5AP+/ALOX5AP- in CD8 vs. CD4 in each sample. F. Number of cells with GLAG peptide in TRB CDR3 of T cell clones.

Figure 6. Temporal expression of Tax in JET cells infected with HTLV-1-CTCF and HTLV-1-p12stop. A. Schematic flow of the experiment, shows transfection of 293T cells with pHTLV-1(WT), pHTLV-1(p12stop), or pHTLV-1(CTCF) plasmid. Transfected cells were co-cultured with JET cells, carrying a Tax-dependent td tomato (RFP) indicator. After 72 hrs of infection the total number of td tomato+ cells in HTLV-1-CTCF infected cells was lower when compared to HTLV-1-p12stop infected cells. B. In order to examine the time course of infection, co-cultured cells were placed in the IncuCyte and were observed for 5 days. The total red object count is shown for duplicate samples of cells infected with HTLV-1-p12stop or HTLV-1-CTCF. C. The total red intensity in infected cells is shown. (** indicates p value lower than 0.01).

Figure 7. Working Hypothesis. A. Schematic depiction of the HTLV-1 genome indicating the location of the vCTCF-BS and major transcripts. B. Working model of the role of the vCTCF-BS in viral gene expression and pathogenesis. Early infection, before methylation of the integrated provirus and suppression of (+) strand transcription from the 5'LTR promoter, the vCTCF-BS acts as an enhancer, Tax expression is elevated, and the effects of TAX protein (virus production, lymphocyte expansion, Tax antigen presentation, cytotoxicity) are elevated. Cytotoxicity resulting from virus production and Tax expression applies selective pressure for (+) sense-strand suppression (DNA hypermethylation) that converts the CTCF binding site from an enhancer to a barrier element and suppresses Tax expression and activity. This rhythm maximizes virus production early followed by entrance into latency to preserve cellular viability, and leads to a burst of initial viremia *in vivo* followed by lymphocyte expansion and death in the infected Hu-mouse with no adaptive immunity (rapid onset acute disease) or a low-level / undetectable steady state plateau in an infected immunocompetent human (latency). The loss of the vCTCF-BS

656 dysregulates this rhythm, and leads to low level viremia, smoldering infection, and asymptomatic

657 or delayed pathogenesis in the humanized mouse.

Fig 1

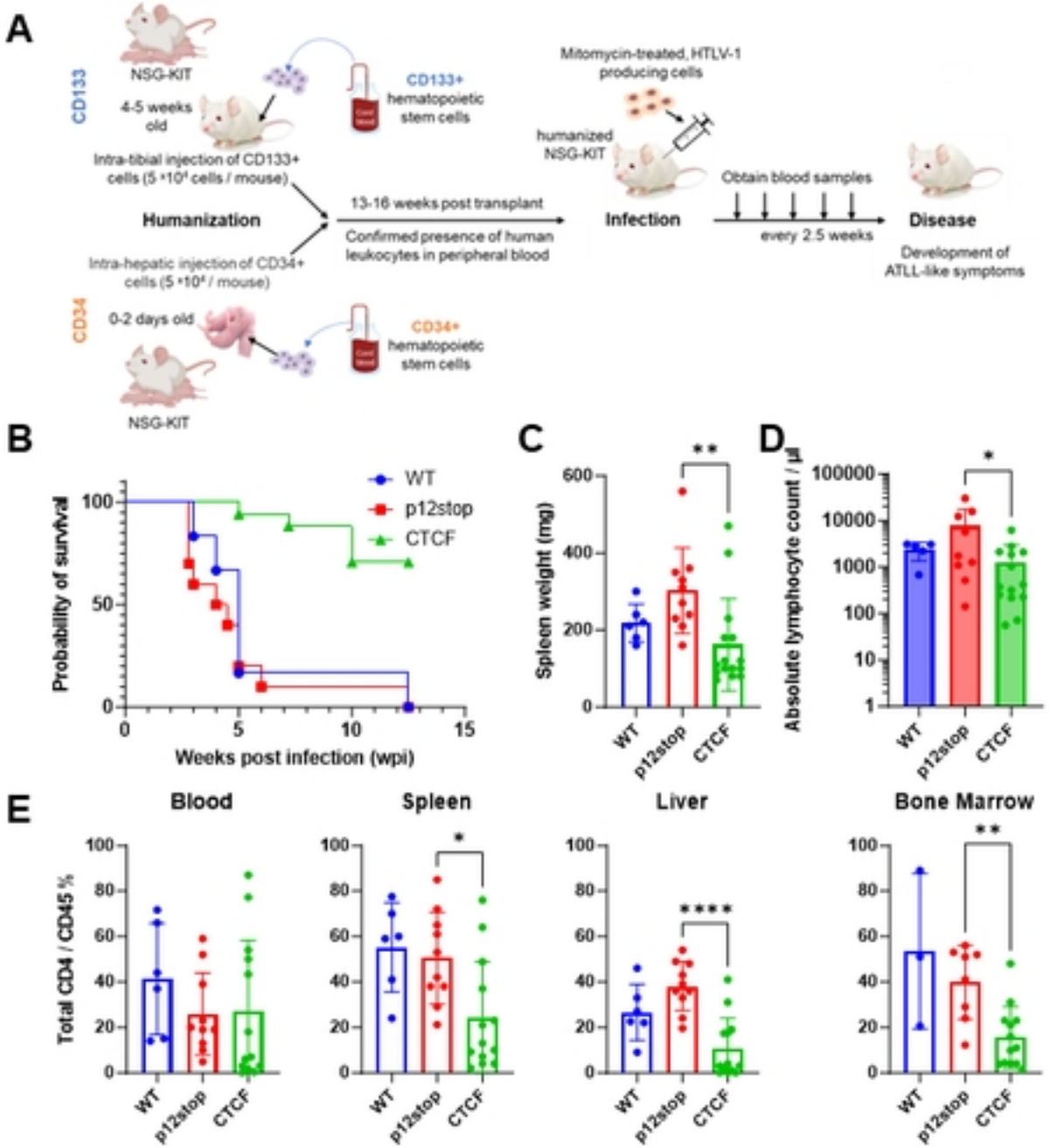

Figure 1

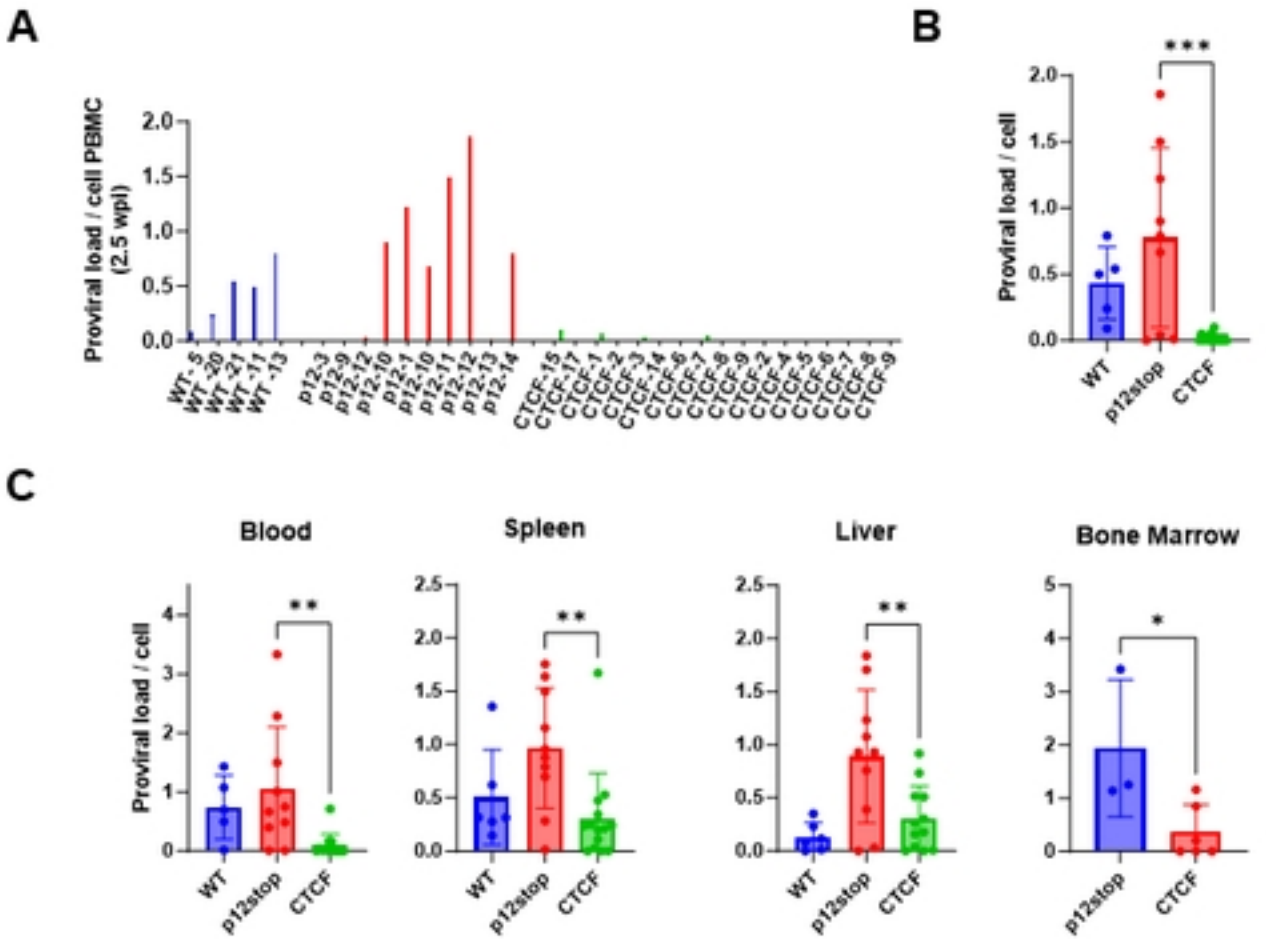
A
Proviral load / cell PBMC (2.5 wpi)
2.0
1.5
1.0
0.5
0.0
WT-5
WT-20
WT-21
WT-11
WT-13
p12-3
p12-9
p12-12
p12-10
p12-1
p12-10
p12-11
p12-12
p12-13
p12-14
CTCF-15
CTCF-17
CTCF-1
CTCF-2
CTCF-3
CTCF-14
CTCF-6
CTCF-7
CTCF-8
CTCF-9
CTCF-2
CTCF-4
CTCF-5
CTCF-6
CTCF-7
CTCF-8
CTCF-9
B
Proviral load / cell
***
WT
p12stop
CTCF
C
Blood
Spleen
Liver
Bone Marrow
Proviral load / cell
**
*
WT
p12stop
CTCF

Figure 2

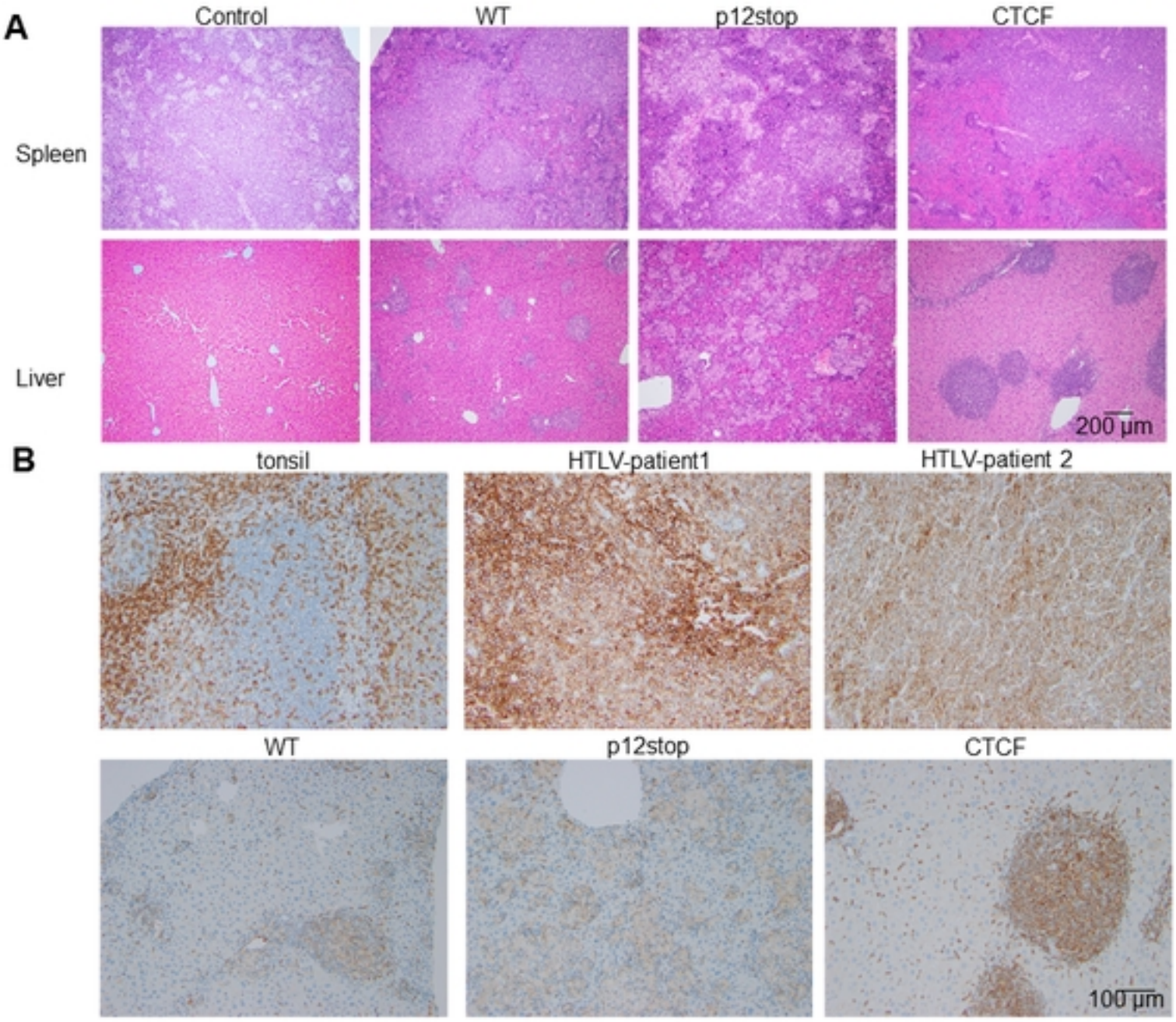

Fig 3

Figure 3

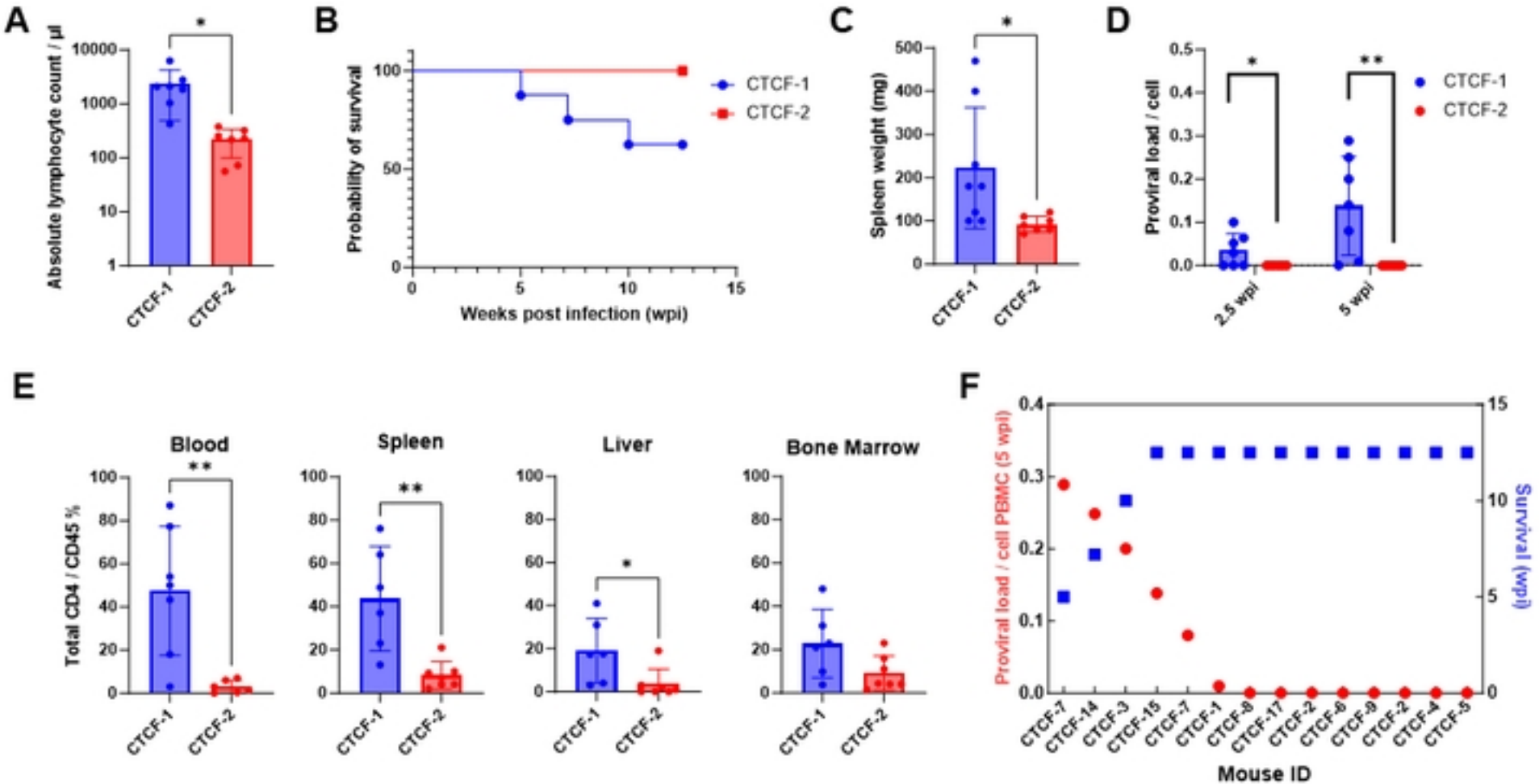
A
Absolute lymphocyte count / µl
10000
1000
100
10
1
*
CTCF-1
CTCF-2
B
Probability of survival
100
50
0
0
5
10
15
Weeks post infection (wpi)
CTCF-1
CTCF-2
C
Spleen weight (mg)
500
400
300
200
100
0
*
CTCF-1
CTCF-2
D
Proviral load / cell
0.5
0.4
0.3
0.2
0.1
0.0
*
**
2.5 wpi
5 wpi
CTCF-1
CTCF-2
E
Total CD4 / CD45 %
Blood
Spleen
Liver
Bone Marrow
100
80
60
40
20
0
**
**
*
CTCF-1
CTCF-2
F
Proviral load / cell PBMC (5 wpi)
Survival (wpi)
0.4
0.3
0.2
0.1
0.0
15
10
5
0
CTCF-7
CTCF-14
CTCF-3
CTCF-15
CTCF-7
CTCF-1
CTCF-8
CTCF-17
CTCF-2
CTCF-6
CTCF-9
CTCF-2
CTCF-4
CTCF-5
Mouse ID

Figure 4

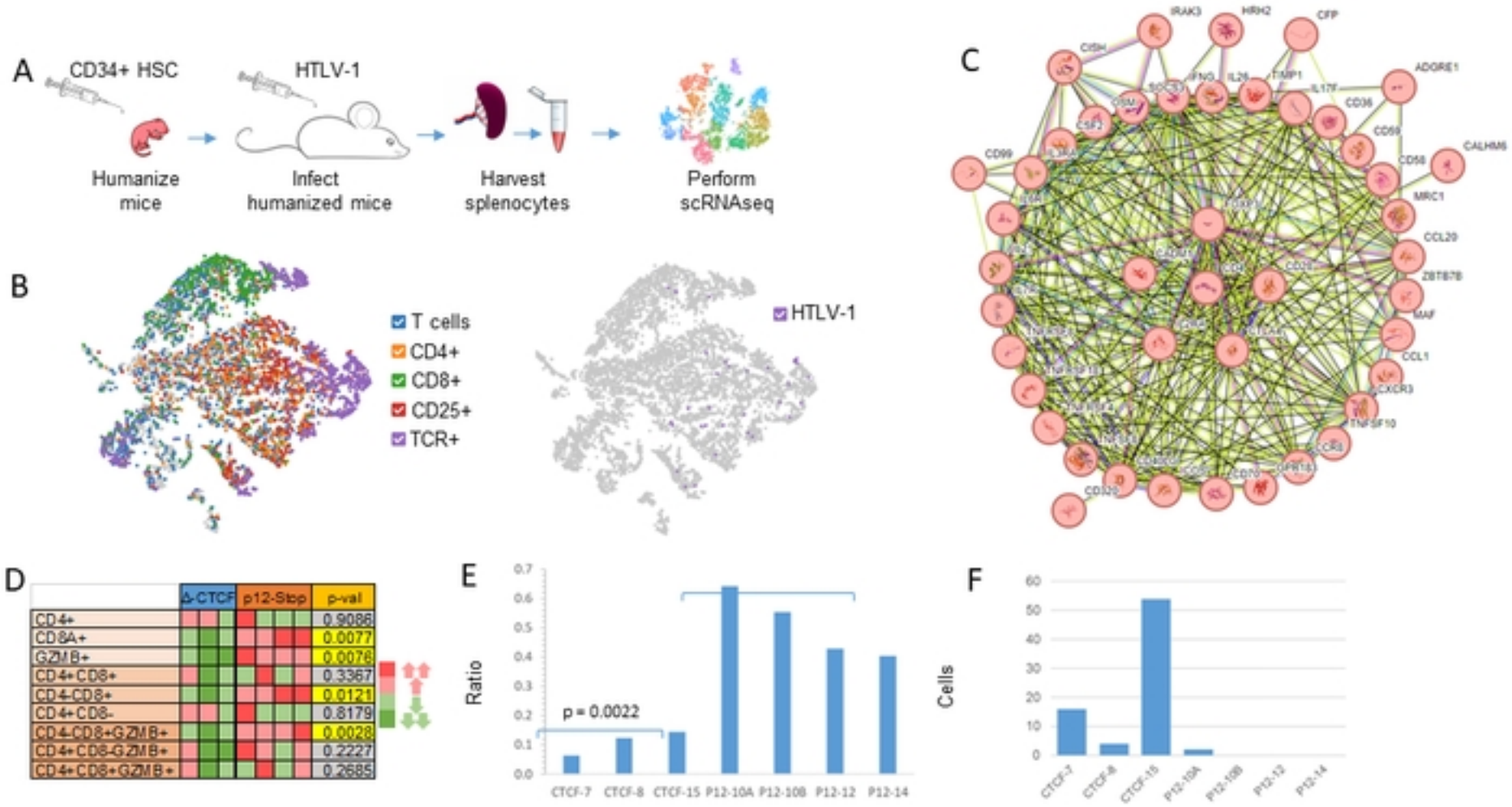
A
CD34+ HSC
HTLV-1
Humanize mice
Infect humanized mice
Harvest splenocytes
Perform scRNAseq
B
T cells
CD4+
CD8+
CD25+
TCR+
HTLV-1
C
D
Δ-CTCF
p12-Stop
p-val
CD4+ 0.9086
CD8A+ 0.0077
GZMB+ 0.0076
CD4+CD8+ 0.3367
CD4-CD8+ 0.0121
CD4+CD8- 0.8179
CD4-CD8+GZMB+ 0.0028
CD4+CD8-GZMB+ 0.2227
CD4+CD8+GZMB+ 0.2685
E
Ratio
p = 0.0022
CTCF-7 CTCF-8 CTCF-15 P12-10A P12-10B P12-12 P12-14
F
Cells
CTCF-7 CTCF-8 CTCF-15 P12-10A P12-10B P12-12 P12-14

Figure 5

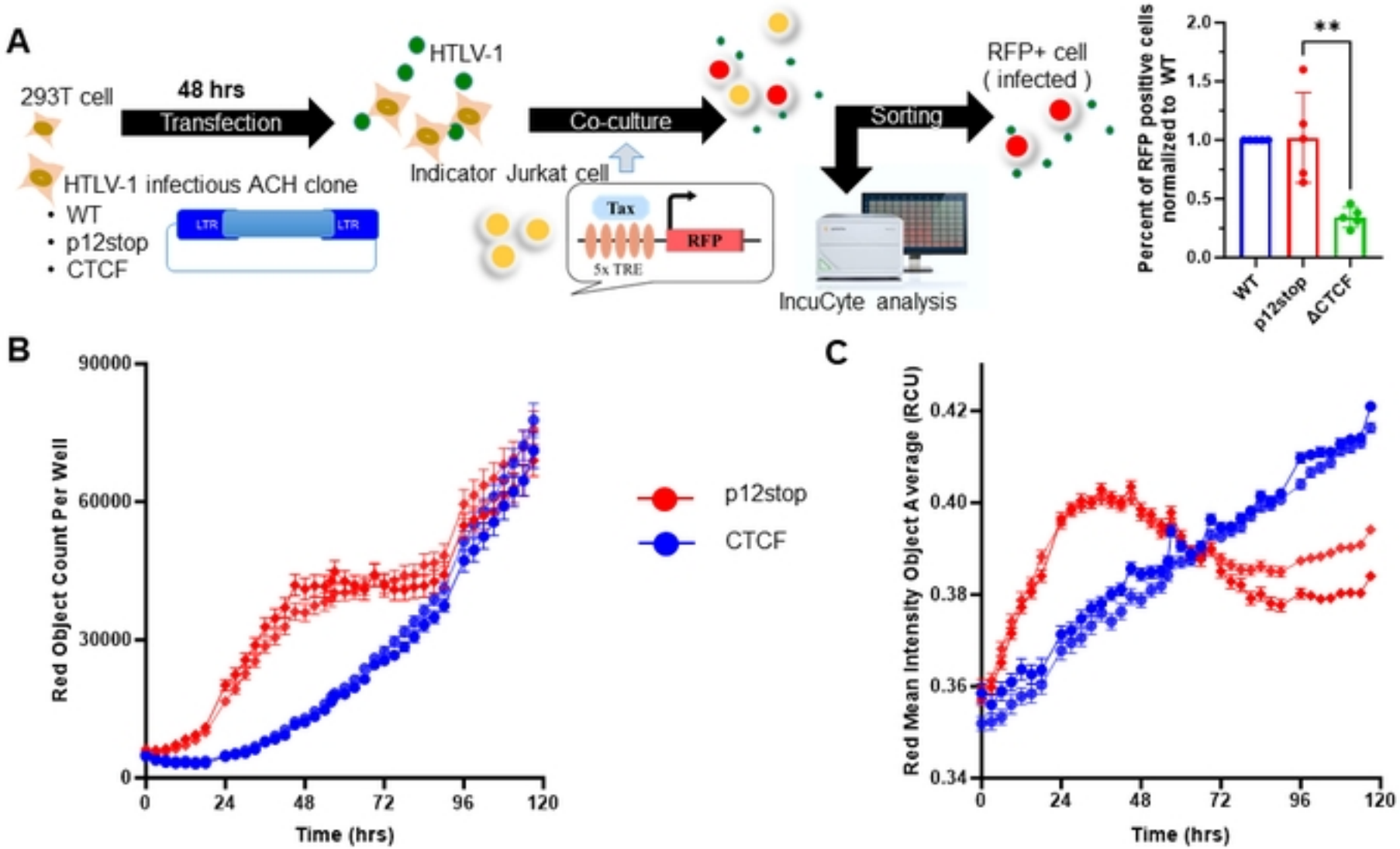
A
293T cell
48 hrs
Transfection
HTLV-1
HTLV-1 infectious ACH clone
• WT
• p12stop
• CTCF
LTR
LTR
Indicator Jurkat cell
Co-culture
Tax
RFP
5x TRE
Sorting
IncuCyte analysis
RFP+ cell
(infected)
Percent of RFP positive cells normalized to WT
2.0
1.5
1.0
0.5
0.0
**
WT
p12stop
ΔCTCF
B
Red Object Count Per Well
90000
60000
30000
0
0
24
48
72
96
120
Time (hrs)
p12stop
CTCF
C
Red Mean Intensity Object Average (RCU)
0.42
0.40
0.38
0.36
0.34
0
24
48
72
96
120
Time (hrs)

Figure 6

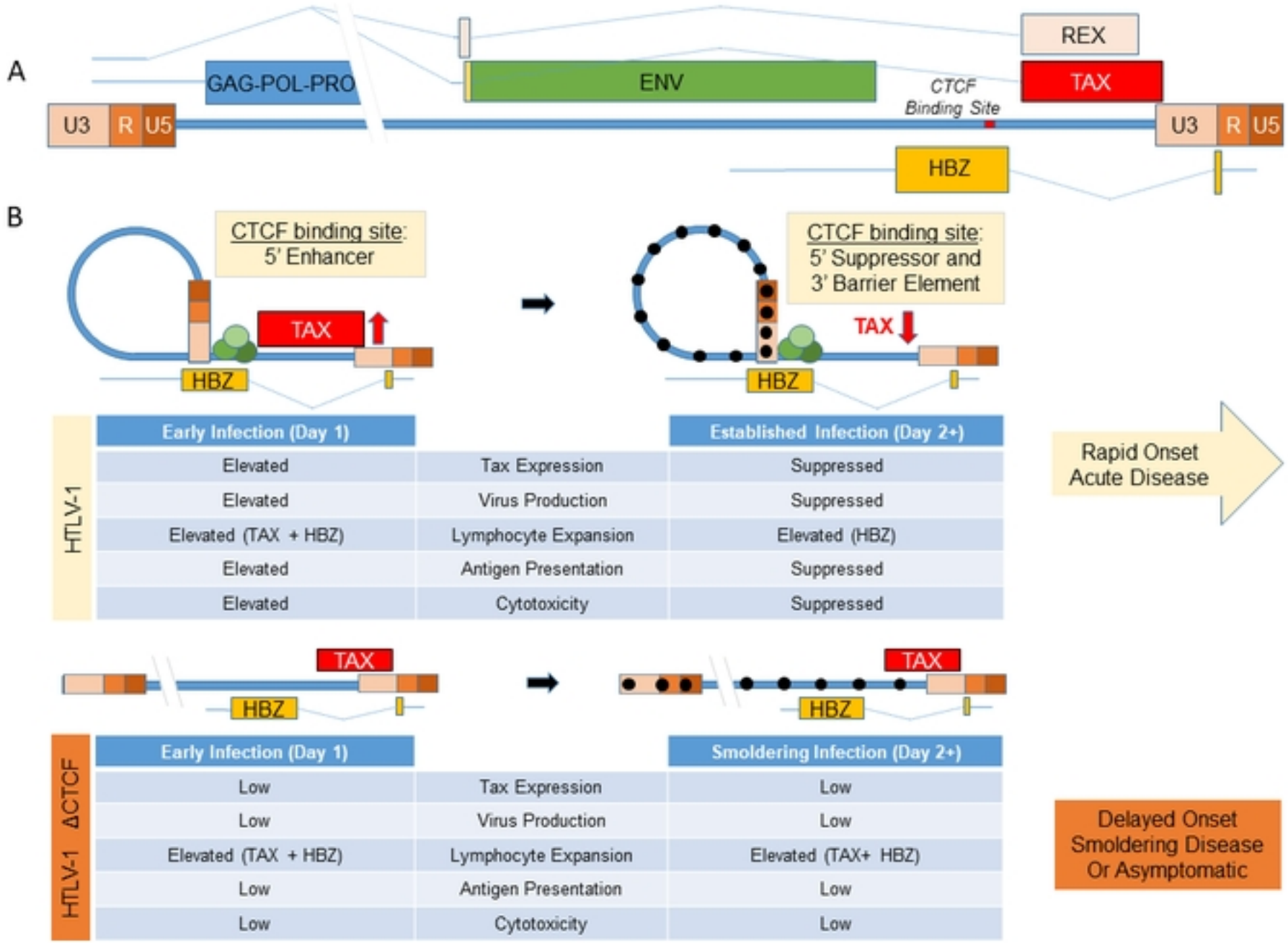
A
GAG-POL-PRO
ENV
REX
TAX
CTCF
Binding Site
U3
R
U5
U3
R
U5
HBZ
B
CTCF binding site:
5' Enhancer
TAX
HBZ
CTCF binding site:
5' Suppressor and
3' Barrier Element
TAX
HBZ
HTLV-1
Early Infection (Day 1)
Established Infection (Day 2+)
Elevated | Tax Expression | Suppressed
Elevated | Virus Production | Suppressed
Elevated (TAX + HBZ) | Lymphocyte Expansion | Elevated (HBZ)
Elevated | Antigen Presentation | Suppressed
Elevated | Cytotoxicity | Suppressed
Rapid Onset
Acute Disease
TAX
HBZ
TAX
HBZ
HTLV-1 ΔCTCF
Early Infection (Day 1)
Smoldering Infection (Day 2+)
Low | Tax Expression | Low
Low | Virus Production | Low
Elevated (TAX + HBZ) | Lymphocyte Expansion | Elevated (TAX+ HBZ)
Low | Antigen Presentation | Low
Low | Cytotoxicity | Low
Delayed Onset
Smoldering Disease
Or Asymptomatic

Figure 7