Multiple stages of evolutionary change in anthrax toxin receptor expression in humans

Lauren A. Choate[1,2], Gilad Barshad[1], Pierce W. McMahon[1], Iskander Said[2], Edward J. Rice[1], Paul R. Munn[1], James J. Lewis[1,*], and Charles G. Danko[1,3,*]

[1] Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853.
[2] Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853.
[3] Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853.

Address correspondence to:
Charles G. Danko, Ph.D.
Baker Institute for Animal Health
Cornell University
Hungerford Hill Rd.
Ithaca, NY 14853
Phone: (607) 256-5620
E-mail: dankoc@gmail.com

James J. Lewis
E-mail: jjl336@cornell.edu

Abstract

The advent of animal husbandry and hunting increased human exposure to zoonotic pathogens. To understand how a zoonotic disease influenced human evolution, we studied changes in human expression of anthrax toxin receptor 2 (*ANTXR2*), which encodes a cell surface protein necessary for *Bacillus anthracis* virulence toxins to cause anthrax disease. In immune cells, *ANTXR2* was 8-fold down-regulated in all available human samples compared to non-human primates, indicating regulatory changes early in the evolution of modern humans. We also observed multiple genetic signatures consistent with recent positive selection driving a European-specific decrease in *ANTXR2* expression in several non-immune tissues affected by anthrax toxins. Our observations fit a model in which humans adapted to anthrax disease following early ecological changes associated with hunting and scavenging, as well as a second period of adaptation after the rise of modern agriculture.

Despite abundant evidence that infectious diseases have driven human adaptation (*1–8*), we understand relatively little about how exposure to new pathogens can shape human host genetics. Here we study the evolution of human host gene expression to *Bacillus anthracis* infection, a bacterium that causes anthrax disease (*9, 10*), as a model for how repeated intermittent exposure to an infectious disease can drive recurrent adaptive alteration. *B. anthracis* primarily afflicts ruminants during the course of its natural life-cycle, and has driven selective pressures in cattle and sheep (*11, 12*). While anthrax disease is believed to rarely affect most primate species, with exceptions (*13, 14*), anthrax disease has been a notable source of mortality in humans (*9, 10*). Genetic diversity of extant *Bacillus* strains suggests that anthrax-causing *Bacillus* pathogens evolved in sub-Saharan Africa (*15, 16*), likely well before human radiation around the globe (*17*). *B. anthracis* then radiated from Africa into the middle east and Europe approximately 3-6ka, possibly following the spread of Neolithic agricultural practices (*16*). *B. anthracis* strains endemic to Europe later spread around the globe through trade and colonization (*16*). Thus, although anthrax disease has been associated with the rise of animal husbandry (*18–20*), the putative African origin of anthrax disease-associated *Bacillus* species may have provided opportunities for early humans to encounter *B. anthracis*, or the ancestral *B. cereus,* while hunting or scavenging ruminate game species well before the advent of modern farming. This led us to hypothesize that humans may have adapted to anthrax disease during several periods of human evolution—first to early ecological changes associated with hunting and scavenging, followed by a second period of adaptation after the rise of agriculture.

To identify evolutionary differences associated with anthrax susceptibility in humans, we studied changes in CD4+ T-cells between humans and non-human primates. We confirmed that in T-cells anthrax toxins inhibit cellular activation using an ELISA for the T-cell activation marker IL2 (**Fig. S1**), consistent with studies indicating that anthrax toxins inhibit immune responses that could clear the pathogen (*21–23*). We then studied transcriptional changes between human and non-human primate CD4+ T-cells using PRO-seq (Pol II loading) and RNA-seq (mature mRNA). We focused our analysis on the *ANTXR2* gene, which encodes a ubiquitously expressed transmembrane receptor that aids in the cellular entry of toxins secreted by the *B. anthracis* bacterium (*24–26*) (**Fig. 1A**). Consistent with our hypothesis, *ANTXR2* was 8-fold downregulated in human CD4+ T-cells compared to non-human primates at the level of both Pol II loading and mRNA (**Fig. 1B-C; Fig. S2A-B**). Expanding our analysis of RNA-seq data from CD4+ T-cells to 85 humans confirmed the loss of *ANTXR2* expression in all of the available human data (*27*) (**Fig. 1D**). Moreover, analysis of RNA-seq from peripheral blood mononuclear cells supported decreased *ANTXR2* expression in Batwa (a hunter gatherer population historically in the Great Lakes region of Africa) and Bakiga (agriculturalists from neighboring Rwanda and Uganda), compared to rhesus macaque (*28, 29*) (**Fig. S3**). Collectively, analysis of *ANTXR2* expression data was consistent with our hypothesis of ancient evolutionary changes affecting *ANTXR2* expression during the early divergence of humans from other primates.

To confirm that *ANTXR2* expression changes of the magnitude observed between human and non-human primates can affect sensitivity to anthrax toxins, we overexpressed *ANTXR2* using CRISPR activation (CRISPRa) (*30*). CRISPRa increased *ANTXR2* mRNA levels by 10-fold relative to an empty vector control 24 hours after transfection, a change that was similar in magnitude to the difference observed between humans and non-human primates (**Fig. 1E**). *ANTXR2* overexpressing cells had a significantly lower viability following treatment with recombinant anthrax toxins: protective antigen (PA) and FP59, a potent analog of lethal factor (LF) (*31*) (**Fig. 1F**). Thus a 10-fold increase in human *ANTXR2* expression, similar to changes found in non-human primates, increased the effect of anthrax toxins on cellular phenotypes.

We asked whether cis-regulatory changes could account for transcriptional divergence in *ANTXR2* between species. We used enhancer RNA transcription detected using PRO-seq data to identify candidate cis-regulatory elements (CREs) near *ANTXR2* in CD4+ T-cells of each primate species (*32*). dREG annotated 14 candidate CREs near *ANTXR2*, including four in introns
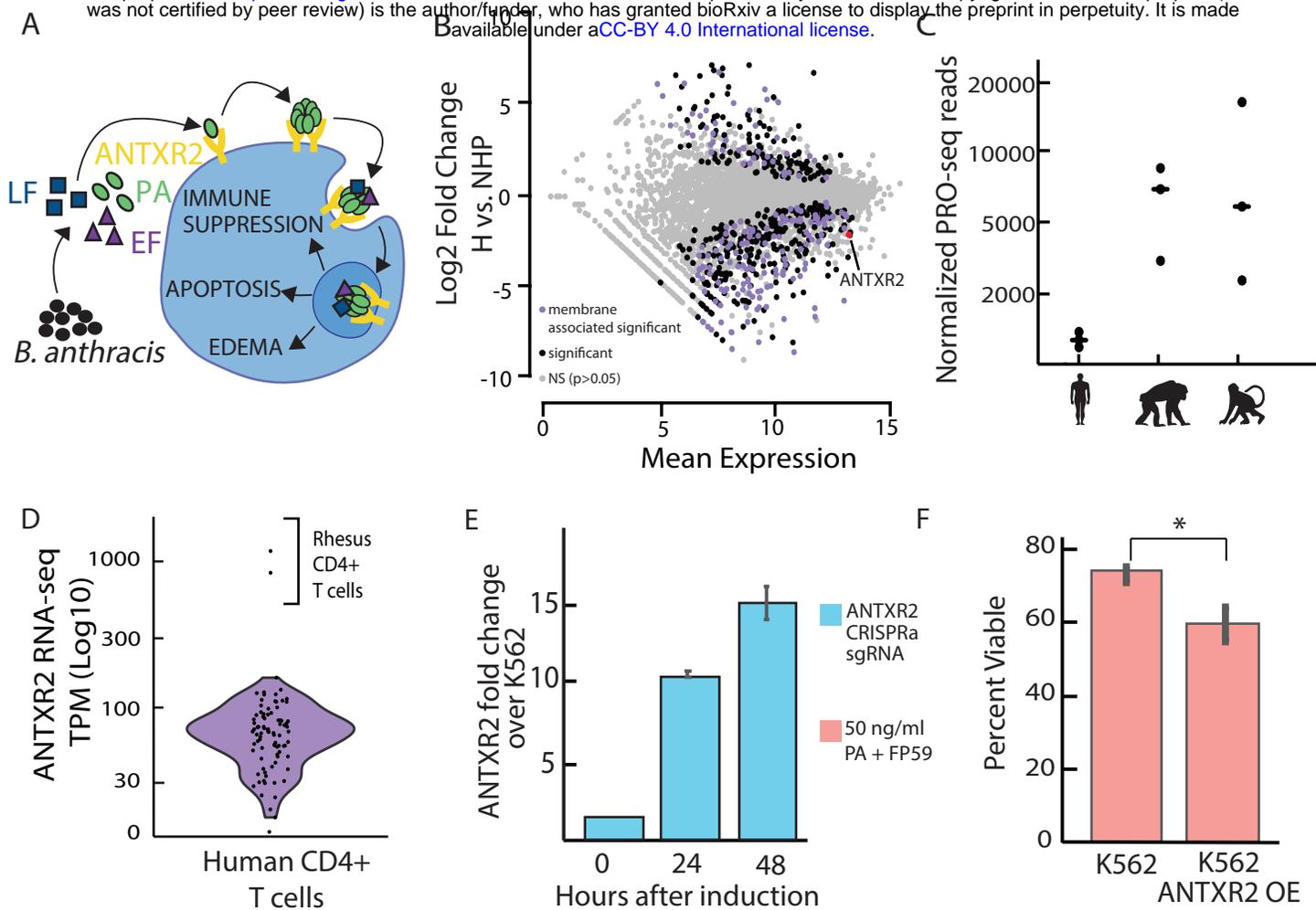
Figure 1. Decreased *ANTXR2* in humans can affect anthrax toxin sensitivity in blood cells. A) B. anthracis toxins lethal factor (LF) and edema factor (EF) cause apoptosis, edema, and immune suppression in target host cells. **B)** Differentially transcribed protein-coding genes between humans and non-human primates in CD4+ T cells based on PRO-seq data. The GO term 'integral component of the membrane' is enriched in differentially transcribed genes. **C)** CD4+ T cell PRO-seq data shows that *ANTXR2* is transcribed 8-fold lower in humans than chimpanzee and rhesus macaque. **D)** Comparison of *ANTXR2* RNA expression levels among a large set of humans (n=85) to rhesus macaque. Human variation in expression does not overlap with rhesus macaque expression. **E)** CRISPRa induction in K562 cells results in increased ANTXR2 expression as measured using qRT-PCR. F) CRISPRa K562 cells that overexpress *ANTXR2* are less viable after an anthrax toxin challenge as measured by the alamar blue viability stain.

4

of *ANTXR2,* a broad promoter, six CREs situated within the gene desert that separated *ANTXR2* and *PRDM8*, and three CREs within the region surrounding the *PRDM8* transcription unit (**Fig. 2A**). Notably, a comparison with DNase-I hypersensitivity and H3K27ac ChIP-seq data in humans did not identify additional candidate CREs in this locus (**Fig. S4**). Testing each putative CRE for differential transcription, a hallmark of regulatory activity at enhancers (*33*), revealed a strong bias for decreased transcription in the human lineage across the *ANTXR2* locus (median 2.3-fold lower transcription in human; p = 1e-3; Wilcoxon rank sum test; **Fig. S5**). At least three CREs were differentially transcribed on the human lineage using a conservative test for differential transcription (median CRE 16-fold lower in human; FDR corrected p-value < 0.05). Notably, the majority of CREs in the gene desert or near either the *ANTXR2* or *PRDM8* promoter were found in orthologous positions in both non-human primate species. Since the majority of genome-wide CRE activity changed between chimpanzee and rhesus macaque (*34*), the deep conservation of gene desert CREs may indicate an important biological function. The significant degree of cis-regulatory divergence between human and non-human primates at multiple annotated CREs thus led us to speculate that cis-acting loci, rather than trans-acting factors, underlie much of the divergence in *ANTXR2* expression.

While species-specific variation in *ANTXR2*-associated CREs is suggestive of CRE-mediated transcriptional evolution, we aimed to explicitly link distal CREs to the *ANTXR2* gene. To verify that annotated CREs interact with the *ANTXR2* promoter, we performed *in situ* Hi-C (*35*) in CD4+ T-cells from human and rhesus macaque and tested for enrichment of contacts in the *ANTXR2* locus. Our Hi-C data revealed that *ANTXR2*, *PRDM8*, and all candidate CREs were found within the same topological associated domain (TAD), a structure reported to insulate distal enhancers from affecting expression (*36, 37*) (**Fig. 2B**). Moreover, Hi-C data revealed focal contacts between the *ANTXR2* promoter and candidate CREs in the gene desert upstream (**Fig. 2B**). The main focal contacts observed in human Hi-C data had a higher contact frequency with the *ANTXR2* promoter in rhesus macaque, potentially reflecting species-specific differences in *ANTXR2* transcription (**Fig. 2C; Fig. S6**). Virtual 4C-seq analysis confirmed that several of the distal candidate CREs had increased contact frequency with the *ANTXR2* promoter relative to flanking regions (**Fig. 2C; Fig. S7**). Additionally, we identified an *ANTXR2* expression quantitative trait locus (eQTL) that overlapped CREs in the gene desert and near *PRDM8* using RNA-seq data from human CD4+ T-cells (*27*) (**Fig. S8**). Taken together, these findings indicate that many of the evolutionarily conserved CREs throughout the upstream gene desert interact with the *ANTXR2* promoter and affect *ANTXR2* expression in CD4+ T-cells.

Multiple CREs changed Pol II loading in human, which raised the possibility that multiple CREs contributed to changes in *ANTXR2* expression. To determine whether changes in the transcriptional activity at multiple CREs reflect independent DNA sequence changes, we used a luciferase assay to measure the activity of CREs using DNA sequences from human, chimpanzee, and rhesus macaque. We focused on 9 CREs in the upstream gene desert, the *ANTXR2* promoter, and in the *PRDM8* transcription unit, most of which were active in orthologous positions in multiple species, providing evidence of evolutionary constraint indicative of a biological function. We introduced nine CREs into Jurkat human leukemic CD4+ T-cells, a model trans-environment which recapitulates the pattern of transcription in the *ANTXR2* locus observed in primary T-cells (**Fig. S9**). Five of the 9 candidate CREs showed higher luciferase activity in chimpanzee or rhesus macaque than in human (p < 0.05, t-test; **Fig. 2D; Fig. S10**). Previous studies have noted multiple, compensatory changes in the activity of sequences within a CREs (*38–41*). Consistent with this idea, breaking CREs into separate regions in some cases uncovered compensatory changes within distinct regions of the CREs, for instance within the complex *ANTXR2* promoter (**Fig. S11**). Nevertheless, decreased overall activity in five of nine CREs, with no CREs increased in human, is unlikely to happen by chance (p = 0.03, binomial test; p = 0.01, multinomial test assuming 36% of CREs change (*41*)). Thus, our results may best fit an adaptive
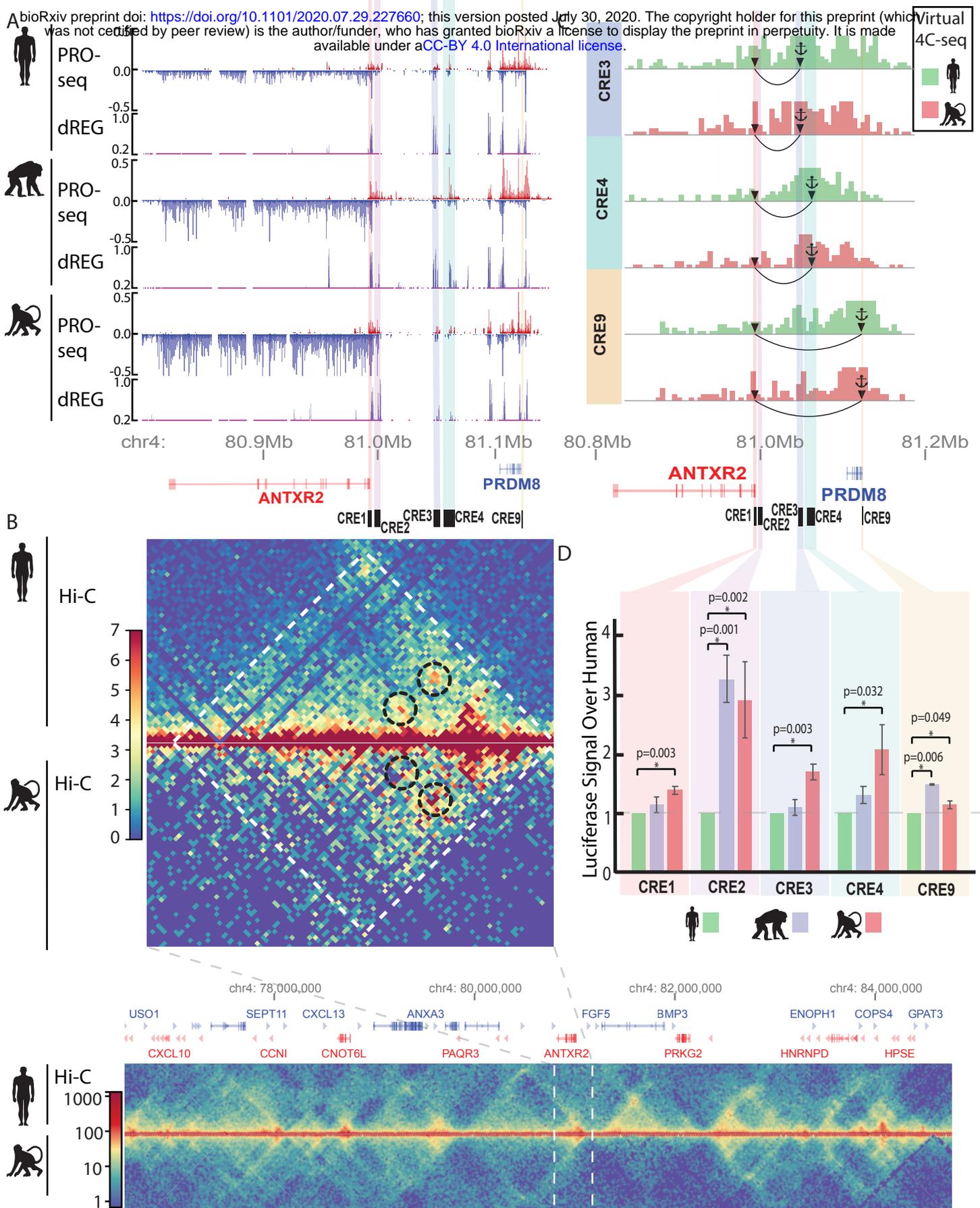
Figure 2. Changes in *ANTXR2* cis regulatory element activity and chromatin structure. A) Genome browser shot of CD4+ T cell PRO-seq (normalized by reads per million) in human, chimpanzee, and rhesus macaque and regulatory elements and regulatory elements predicted by dREG. **B)** Hi-C in CD4+ T cells from human and rhesus macaque. *ANTXR2* is located in the same topological domain (TAD) as the upstream gene *PRDM8*. TADs are marked with white dotted lines. Focal contacts with the *ANTXR2* promoter are circled with black dotted lines. **C)** irtual 4C-seq plots for distal regulatory elements tested in the luciferase assay show contacts to the *ANTXR2* promoter. **D)** Luciferase assay performed in Jurkat cells to test the activity of regulatory elements in human, chimpanzee, and rhesus macaque shows an increase in activity for either chimpanzee or rhesus macaque compared to humans for CRE1, CRE2, CRE , CRE4, and CRE9.

model in which human ancestors had an ancient adaptive change for reduced *ANTXR2* expression that is now fixed in humans.

In addition to changes compared with other primates, *ANTXR2* expression varied by one order of magnitude within humans, suggesting that a number of genetic variants may still exist within human populations associated with anthrax susceptibility. Most extant *B. anthracis* strains outside of Africa show genetic similarity to those in Europe, leading to a model in which *B. anthracis* that was endemic to Europe during the Neolithic era spread throughout the globe by trade and colonization during the past several thousand years (*16, 42–44*). Indeed, previous reports have noted lower anthrax toxin sensitivity in Europeans (*45*). To determine how selection has influenced modern human populations at the *ANTXR2* locus, we computed the composite-likelihood-ratio (CLR) of a selective sweep across chromosome 4 in four human populations (Europeans [CEU], East Asians [CHB and JPT], and Africans [YRI]) using SweepFinder2 (*46*). Consistent with the hypothesis that anthrax disease in Europeans led to a relatively recent population specific adaptive response, we found a candidate selective sweep in the gene desert upstream of the *ANTXR2* locus (**Fig. 3A**). The selective sweep had a higher CLR in European (CEU) than 98% of other loci on chromosome 4 (**Fig. S12**). Moreover, the CLR was substantially higher in Europeans compared to 1000 Genomes populations representative of East Asian (CHB, JPT) or African (YRI) ancestry (**Fig. 3A**). Surprisingly, this sweep had minimal overlap with CREs significantly associated with decreased *ANTXR2* expression between humans and non-human primates, and suggests a complex evolutionary response driven by recurrent selection around *ANTXR2*.

The discovery of a selective sweep at the *ANTXR2* locus that excluded loci associated with decreased *ANTXR2* expression in CD4+ T-cells led us to speculate that a recent hard sweep may have occurred in response to anthrax-related selection pressures on additional organs. Anthrax disease is a systemic disorder affecting multiple organs in different manners. The main effects of anthrax disease are apoptosis (by anthrax lethal toxin), edema (by anthrax edema toxin), and a suppressed immune response (*26, 47*). The main targets of anthrax lethal toxin are cardiomyocytes and smooth muscle cells, whereas the main targets of anthrax edema toxin are the dermis, hepatocytes, and lungs (*48, 49*). Anthrax disease progresses with attacks on multiple organ systems, but the cause of host death is often by the induction of apoptosis in the heart (*50, 51*). We therefore examined human DNase-I hypersensitivity data in 42 tissues from the Epigenome Roadmap project (*52*). We found substantial numbers of DNase-I hypersensitive sites overlapping the interval of high CLR values in heart, lung, liver, and dermis (**Fig. 3B, Fig. S13**). Of these, multiple putative CREs were not found in blood DNase-I samples, supporting the hypothesis that the recent selective sweep affected tissues susceptible to apoptosis or edema.

Evidence of decreased expression between populations in non-blood tissues would be consistent with our prediction of recurrent selection on multiple disease-affected organs. To test whether *ANTXR2* expression varies among populations at anthrax disease targeted tissues, we divided GTEx RNA-seq data into European (n = 209, 306, 135, 60 for heart, skin, lung, and liver, respectively) and African (n = 28, 43, 17, 10 for heart, skin, lung, and liver) ancestry based on the mitochondrial haplotype (*53*). Although we were underpowered in most tissues due to the small number of individuals of African ancestry, *ANTXR2* expression was reduced in all four tissue types in Europeans relative to African samples and significantly reduced in lung (Mann-Whitney U test p = 0.07, 0.15, 0.04, 0.08 in heart, skin, lung, and liver; **Fig. 3C**). Combining p-values across tissues using Fisher's method also supported decreased *ANTXR2* expression in Europeans (p = 0.0085, Fisher's method). While additional study will be needed, we speculate that Europeans have, on average, decreased expression of *ANTXR2* in several tissues that are central to anthrax disease pathogenesis via recent selection on tissue specific CREs in the heart, skin, lung, and liver.

Finding a putative selective sweep specific to Europeans led us to ask whether the *ANTXR2* locus differentiates CEU from other human populations. To address this, we calculated
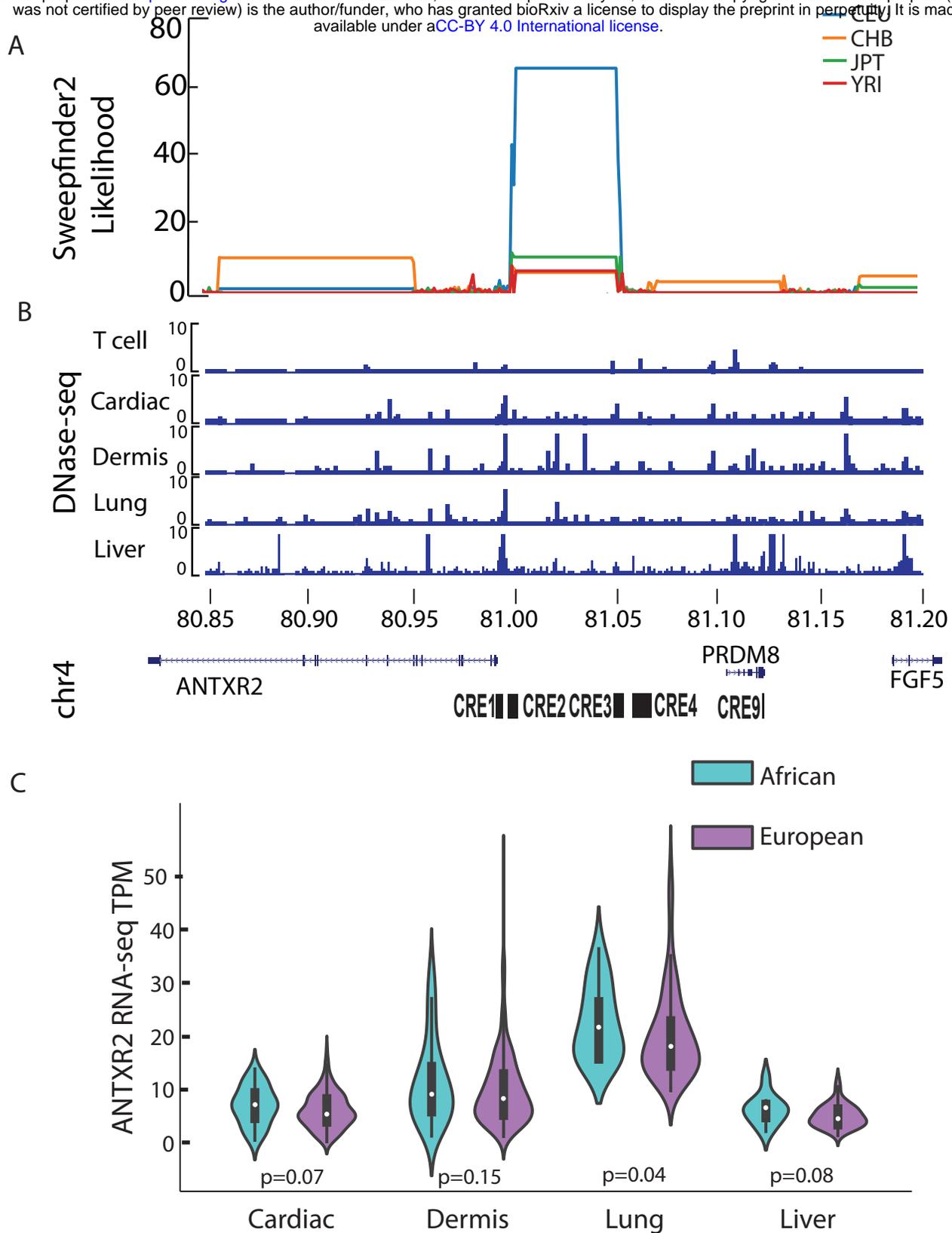
Figure . ard selective sweep affects ANTXR2 expression in multiple tissues. A) Sweepfinder2 shows a predicted selective sweep in the CEU population upstream of *ANTXR2*. **B)** ENCODE DNase-I-seq data for T cells, cardiac cells, dermis, lung, and liver show differential regulatory landscapes within the predicted selective sweep in CEU. **C)** GTEX data for cardiac, dermis, lung, and liver show increased *ANTXR2* expression in African American individuals compared to European individuals.

relative population differentiation (Fst) between Europeans and either East Asians or Africans. The entire *ANTXR2* locus showed elevated differentiation between European and non-European populations, with a median value of 0.14-0.33 (64th-96th percentile) (**Fig. 4A**). The greatest signal of differentiation occurred outside of the putative selective sweep, but overlapped upstream CREs (including CRE2) that diverged between human and non-human primates (**Fig. 4B**). One particularly interesting haplotype, directly adjacent to the region of high CLR, had a very high Fst in Europeans relative to all other ethnic groups. This region included several SNPs overlapping CREs upstream of *ANTXR2*, including rs41407844—the allele frequency of which was correlated with reported variation in anthrax toxin sensitivity in lymphoblastoid cell lines (*45*): high frequency of the derived allele in Europeans (~0.85), intermediate in East Asians (~0.38), and low frequency in African populations (~0.17; **Fig. 4C-D**). This result, and the observation of elevated Fst at CRE3, CRE4, and CRE9 in the absence of high CLR profile, suggests that Europeans have maintained genetic separation at loci derived from both incomplete soft sweeps on genetic variation from early human divergence and a hard sweep associated with recent European anthrax exposure.

Our work is consistent with the hypothesis that humans adapted to anthrax exposure during several periods of human evolution (**Fig. 4E**). First, ecological changes associated with increased hunting and scavenging early in the evolution of modern humans may have led to increased interactions with *Bacillus* species in Africa. Indeed, *Bacillus* species are believed to have originated in Africa (*55*), and remain endemic to Africa, where they are a source of mortality for wildlife, including wild chimpanzees (*13, 14*). It stands to reason that human ancestors who adopted hunting and scavenging behaviors, especially of ruminant species in endemic areas, would have faced a higher burden of anthrax disease than neighboring chimpanzee relatives. Second, we found evidence of more recent selective pressures acting in Europeans that are consistent with historical records of recent outbreaks within Europe (*18, 43, 56*). The nature of human interactions with anthrax disease might make the *ANTXR2* locus particularly susceptible to soft selective sweeps over time. Both prehistoric and modern anthrax exposure was likely highly sporadic and caste specific—dependent upon a variety of factors including regional endemic areas of the globe, the availability of ruminant species, and the roles of individuals within the population. Indeed, anthrax disease disproportionately affected agricultural and textile workers during the 19th century (*10*). Moreover, *ANTXR2* expression was slightly higher in blood cells isolated from hunter gatherers than in nearby agricultural populations within Africa (*28*), potentially consistent with agricultural populations being at higher risk of exposure to anthrax disease. In summary, our findings imply a complex evolutionary history in the *ANTXR2* locus, supporting a model in which continued hard and soft selection on multiple alleles has continued to drive changes in *ANTXR2* expression within human populations.
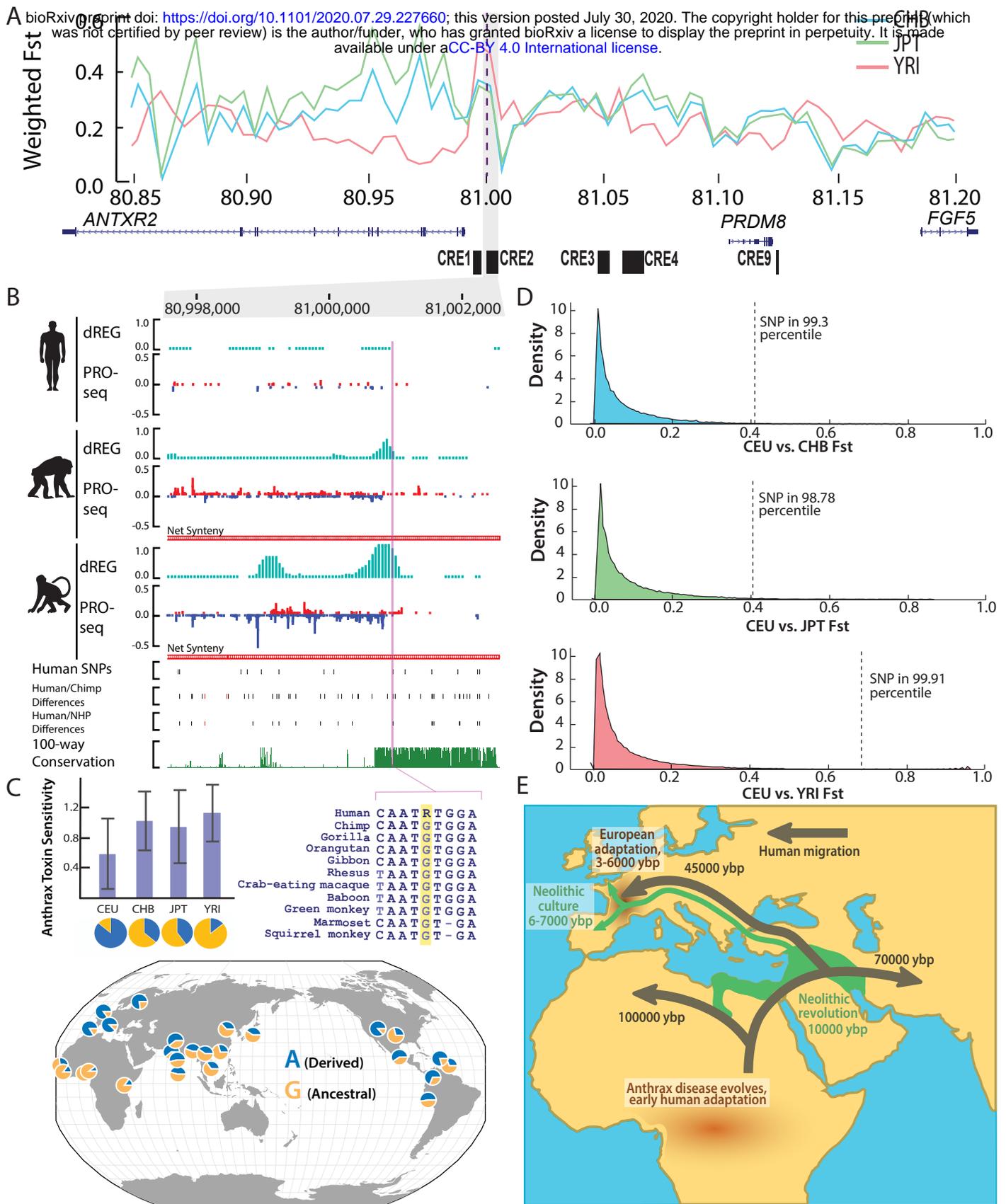
Figure 4. Genetic differentiation in European populations consistent with multiple phases of selection driving changes in *ANTXR2*. A) Weighted Fst comparisons between CEU and CHB, JPT, and YRI HapMap populations at the *ANTXR2* locus show a peak around *ANTXR2*. **B)** PRO-seq (normalized by reads per million) and dREG signal at CRE2. rs41407844 falls within a dREG peak shared by chimpanzee and rhesus macaque and the ancestral allele G is conserved within the primate lineage. Conservation of CRE2 between humans and non-human primate tracks show human-specific SNPs, human/chimpanzee differences (SNPs in black, INDELS in red), and human/non-human primate differences (SNPs in black, INDELS in red), and 100-way conservation. **C)** Martchenko et al. (45) showed that CEU lymphoblastoid cell lines have the lowest sensitivity to anthrax toxins. rs41407844 allele frequencies across populations. **D)** rs41407844 is above the 98% percentile for Fst comparisons between CEU and CHB, JPT, and YRI HapMap populations. **E)** Model demonstrating the pattern of human migration (dark green) compared to the spread of Neolithic culture (light green) and anthrax disease distribution (brown).

10

Materials & Methods

EXPERIMENTAL METHODS
Isolation of CD4+ T cells from humans and non-human primates
All human and animal experiments were done in compliance with Cornell University IRB and IACUC guidelines. We obtained peripheral blood samples (60–80 mL) from healthy adult male humans, chimpanzees, and rhesus macaques. Informed consent was obtained from all human subjects. To account for within-species variation in gene transcription we used three individuals to represent each primate species. Blood was collected into purple top EDTA tubes. Human samples were maintained overnight at 4C to mimic shipping non-human primate blood samples. Blood was mixed 50:50 with phosphate buffered saline (PBS). Peripheral blood mononuclear cells (PBMCs) were isolated by centrifugation (750× g) of 35 mL of blood:PBS over 15 mL Ficoll-Paque for 30 minutes at 20C. Cells were washed three times in ice cold PBS. CD4+ T-cells were isolated using CD4 microbeads (Miltenyi Biotech, 130-045-101 [human and chimp], 130-091-102 [rhesus macaque]). Up to 10^8 PBMCs were resuspended in binding buffer (PBS with 0.5% BSA and 2mM EDTA). Cells were bound to CD4 microbeads (20uL of microbeads/107 cells) for 15 minutes at 4C in the dark. Cells were washed with 1–2 mL of PBS/BSA solution, resuspended in 500uL of binding buffer, and passed over a MACS LS column (Miltenyi Biotech, 130-042-401) on a neodymium magnet. The MACS LS column was washed three times with 2mL PBS/BSA solution, before being eluted off the neodymium magnet. Cells were counted in a hemocytometer.

Luciferase assays
Genomic DNA was isolated from human, chimp, and rhesus macaque PBMCs depleted for CD4+ cells using a Quick-DNA Miniprep Plus Kit (#D4068S; Zymo research) following the manufacturer's instructions. Putative enhancer regions were amplified from the genomic DNA, restriction digested with KpnI and MluI, and cloned into the pGL3-promoter vector (Promega). The same orthologous regions were amplified from all three species with identical primers where possible or species-specific primers covering orthologous DNA in diverged regions. The media (RPMI-1640) was changed in Jurkat cells one day before transfection. RPMI-1640 with 20% FBS was equilibrated in plates in a 37C incubator prior to transfection. On the day of transfection, Jurkat cells were centrifuged at 100xg for 10 minutes, washed with PBS, and centfrigured again. After centrifugation, Jurkat cells (2 million per reaction) were resuspended in 100 ul room temperature Mirus electroporation solution. Vectors were co-transfected with pRL-SV40 Renilla (Promega) in a 20:1 ratio (2ug pGL3 to 100ng pRL-SV40). Electroporation was done in a Lonza Nucleofector 2b device using program X-001. Immediately after electroporation, 1 ml equilibrated media was added to the cuvette and then the cell mixture was added to 6 well plates and incubated at 37C. 18 hours post-transfection, luminescence was measured in triplicate using the Dual-Luciferase® Reporter Assay System (Promega).

CRISPRa in K562
Generation of dCas9-KRAB K562 line—
Lentivirus was made using lipofectamine 3000 from Invitrogen. Phoenix Hek cells (grown in DMEM with 10% FBS and antibiotics) were seeded in a 6-well plate at 400,000 cells/plate. Cells were grown until ~90% confluent. 1ug of pHAGE_EF1a_dCas9-KRAB plasmid from addgene (#50919) plasmid was transfected. 24 hours later 3ml/well of virus was mixed with 10ug/ml polybrene and incubated for 5 minutes at room temperature. This mix was added to 300,000 K562 cells and centrifuged for 40 minutes at 800g at 32C. 12–24 hours later the virus was removed and fresh media was added. 24–48 hours later the cells were selected with 150ug/ml Hygromycin B for 2 weeks. The K562 dCas9-KRAB stable cell lines was grown and maintained in Hygromycin B.

sgRNA cloning—

Primers for sgRNAs were designed using ChopChop (http://chopchop.cbu.uib.no/) for the *ANTXR2* gene and scrambled controls. Primers were located -400 to -50 bp away from the TSS for CRISPRa. A G was added at the 5' end of primers for use with a U6 promoter, along with restriction sites for cloning. Forward and reverse sgRNAs were synthesized separately by IDT and annealed. T4 Polynucleotide Kinase (NEB) was used to phosphorylate the forward and reverse sgRNA during the annealing. 10× T4 DNA Ligase Buffer, which contains 1mM ATP, was incubated for 30 minutes at 37°C and then at 95C for 5 minutes, decreasing by 5°C every 1 minute until 25°C. Oligos were diluted 1:200 using Molecular grade water. sgRNAs were inserted into the pLenti SpBsmBI sgRNA Hygro plasmid from addgene (#62205) by following the authors protocol (26501517). The plasmid was linearized using BsmBI digestion (NEB) and purified using gel extraction (QIAquick Gel Extraction Kit). The purified linear plasmid was then dephosphorylated using Alkaline Phosphatase Calf Intestinal (CIP) (NEB) to ensure the linear plasmid did not ligate with itself. A second gel extraction was used as before to purify the linearized plasmid. The purified dephosphorylated linear plasmid and phosphorylated annealed oligos were ligated together using the Quick Ligation Kit (NEB). The ligated product was transformed into One Shot Stbl3 Chemically Competent E. coli (ThermoFisher Scientific). 100ul of the transformed bacteria were plated on Ampicillin (200ug/ml) plates. Single colonies were picked, sequenced, and the plasmid was isolated using endo free midi-preps from Omega.

Transfection of sgRNA plasmid—

The day prior to transfection, the media (RPMI-1640 with 10% media) was changed in K562 cells and cells were diluted to a concentration of 1 million cells/mL. On the day of transfection, K562 cells were centrifuged at 100xg for 10 minutes, washed with PBS, and centrifuged again. RPMI-1640 with 10% FBS was equilibrated in plates in a 37C incubator prior to transfection. After centrifugation, K562 cells (1 million per reaction) were resuspended in 100 ul room temperature Mirus electroporation solution. 2ug of the sgRNA plasmid was added to each reaction. Electroporation was done in a Lonza Nucleofector 2b device using the program for K562 cells. Immediately after electroporation, 1 ml equilibrated media was added to the cuvette and then the cell mixture was added to 6 well plates and incubated at 37C. 12 hours after transfection, 2ug/ml doxycycline was added to cells to activate dCas9-KRAB expression. To confirm overexpression of *ANTXR2*, 4 hours after the addition of doxycycline a portion of the cells were collected for RNA extraction using Trizol. cDNA was generated from RNA samples using the Thermo Fisher High Capacity RNA-to-cDNA kit and qPCR was performed using SsoAdvanced Universal SYBR Green master mix with primers to assay *ANTXR2* expression.

Toxin viability assays

K562 cells transfected with CRISPRa plasmids were confirmed to have overexpression of *ANTXR2*. After confirmation and within the 12 hour half-life window of the activating doxycycline, 50,000 K562 cells were added to wells of 96-well plates in 100 ul RPMI-1640, supplemented with 10% RPMI. Anthrax toxin PA (List Biological Laboratories 171D) was added to wells at a concentration of 1ug/ml (or vehicle control). FP59, a recombinant anthrax lethal factor fused to the Pseudomonas Exotoxin A Catalytic Domain (Kerafast ENH013), which is capable of killing blood cells, was added at a concentration of 50ng/ml (or vehicle control). 20 hours after the addition of PA and FP59, 10ul of Alamar Blue (Thermo Fisher #DAL1025) was added to each well. Four hours after the addition of Alamar Blue, fluorescence was measured using a plate reader with an excitation of 570 and emission of 610.

Activation of CD4+ T cells

CD4+ T cells were isolated using the above procedure for two human subjects. After equilibration in RPMI-1640 with 10% FBS, cells were stimulated with 25ng/mL PMA and 1mM Ionomycin (P/I or π) or vehicle control (2.5uL EtOH and 1.66uL DMSO in 10mL of culture media). Thirty minutes after activation or addition of the vehicle control, cells were treated with 2.5ug/ml Recombinant PA (List Biological Laboratories 171D) and 500ng/ml Recombinant LF (List Biological Laboratories 172A) or vehicle control. 24 hours later, media from non-activated, non-activated with toxin treatment, activated, and activated with toxin treatment wells was collected for ELISA.

IL-2 ELISA on CD4+ T cells

ELISA was done using the R&D Human IL-2 DuoSet Kit (Catalog #DY202).

Plate preparation—16 hours prior to the ELISA experiment, the capture antibody was added to 96-well plates at a concentration of 0.5ug/ul in 100ul of PBS. Plates were sealed and left at room temperature. The following day, the diluted capture antibody was aspirated and washed with 400ul Wash Buffer three times. 300ul of Block Buffer was added to each well and incubated for at least one hour. After incubation, plates were washed with 400ul Wash Buffer.

ELISA assay—100ul of RPMI from the samples or standards diluted in the Reagent Diluent were added to the prepared 96-well plate. The plate was covered and left to incubate for 2 hours at room temperature. 100ul of the Detection Antibody diluted at 1:60 with the Reagent Diluent was added to each well. The plate was washed with Wash Buffer. 100ul of Streptavidin HRP diluted at 1:40 in the Reagent Diluent was added to each well. The plate was covered, protected from light, and incubated for 20 minutes at room temperature. The plate was washed and 100ul of the Substrate Solution was added to each well, the plate was covered, protected from light, and incubated for 20 minutes at room temperature. 50ul of Stop Solution was added to each well and mixed. Fluorescence was measured on a plate reader at 450nm and wavelength corrected at 540nm.

Hi-C library preparation

Cell preparation—

CD4+ T cells were isolated according to the above procedure. After >1 hour of equilibration in RPMI-1640 supplemented with 10% FBS, cells were centrifuged at 300xg, washed with PBS, and centrifuged again. Cells were resuspended in a mixture of 1% paraformaldehyde in 1x PBS. Cells were incubated at room temperature for 10 min on a rocker. Paraformaldehyde was quenched by the addition of 2.5M Glycine to a Cf=0.2M. Cells were incubated for room temperature for 5 minutes on a rocker. Cells were centrifuged at 4C, washed in cold PBS, centrifuged, and PBS was aspirated. Pellets were flash frozen using dry ice and stored at -80C prior to library preparation.

Hi-C—

The protocol detailed in (_35_) was followed with the following adjustments. After the addition of lysis buffer, cells were incubated on ice for 30 min. MboI (NEB #R0147) was used for restriction digestion. Following DNA purification, unligated biotin was removed using a mixture of 0.5uL of 10mM dATP, 0.5uL dGTP, 20uL 3000U/ml T4 DNA polymerase (NEB #M02030) for each sample. Samples were incubated for 4 hours at room temperature and then T4 DNA polymerase was inactivated at 72C for 20min. Shearing was done using a Bioruptor sonicator using the LOW setting 30S ON/ 90S OFF for 2 cycles of 10 minutes. Libraries were prepared with the NEBNext Ultra II Library Preparation Kit (NEB #E7103). Samples were sequenced on a combination of Illumina's NovaSeq 6000 and HiSeq 4000 at Novogene.

DATA ANALYSIS
Hi-C analysis and visualization
Individual Hi-C samples were mapped to hg19 (for human samples) or rheMac8 (for rhesus macaque samples) using Juicer (*57*). Replicates for each species were combined using Juicer's mega function. The combined rhesus macaque aligned dataset was lifted over to hg19 using Crossmap (*58*).

Mapping orthologs between species
Cross-species comparison of genomic coordinates and genes was based on the methods using in (*34*). Briefly, all datasets for chimpanzee and rhesus macaque were converted to the human assembly (hg19) using CrossMap (*58*). Reciprocal-best (rbest) nets were used to convert genomic coordinates between genome assemblies using (*59*).

PRO-seq and RNA-seq differential expression
PRO-seq data from human, chimpanzee, and rhesus macaque were published in ref (*34*). We mapped PRO-seq reads using standard informatics tools. Our PRO-seq mapping pipeline begins by removing reads that fail Illumina quality filters and trimming adapters using cutadapt with a 10% error rate. Reads were mapped with BWA (*60*) to the appropriate reference genome (either hg19, panTro4, or rheMac3) and a single copy of the Pol I ribosomal RNA transcription unit (GenBank ID# U13369.1). Mapped reads were converted to bigWig format for analysis using BedTools (*61*) and the bedGraphToBigWig program in the Kent Source software package (*62*). The location of the RNA polymerase active site was represented by the single base, the 3′ end of the nascent RNA, which is the position on the 5′ end of each sequenced read.

ANTXR2 RNA-seq analysis in PBMCs
Bakiga and Batwa PBMC RNA-seq data (*28, 29*)) was downloaded for the control condition (GEO series GSE120502). Rhesus macaque RNA-seq data was downloaded for the control condition (NCBI BioProject PRJNA246101). RNA-seq data was mapped using Salmon (*63*) and NCBI RefSeq genes to hg19. Rhesus macaque RNA-seq data was mapped to hg19 using the same parameters. *ANTXR2* transcripts per million were compared between samples for transcript NM_001145794.1.

DICE RNA-seq and eQTL analysis
RNA-seq from 85 human CD4+ naive T cell samples from DICE (*27*) was downloaded under dbGap protocol 23187. RNA-seq data was mapped using Salmon (*63*) and NCBI RefSeq genes to hg19. Rhesus macaque RNA-seq data was mapped to hg19 using the same parameters. *ANTXR2* transcripts per million were compared between samples for transcript NM_001145794.1. We retrieved eQTLs for *ANTXR2* from the DICE online database.

Sweepfinder2 CLR scan
Human genome data was taken from phase3 of the 1,000 Genomes Project(*64*). VCFs were subsetted based on their population group. The b-value maps that are used to compute the effect of background selection on the human genome were taken from(*65*). Recombination maps from the deCODE database(*66*) were used. Sweepfinder2(*46*) was used to calculate the composite-likelihood-ratio for CEU, JPT, CHB, and YRI on chromosome 4.

Fst analysis

All Fst values were computed using VCFtools (*67*) using the --weir-fst command with a window size of 5kb and a step size of 5kb. In addition to the Fst calculations for bins of 5kb, Fst was calculated for each SNP of chromosome 4. Fst was calculated for all pairwise comparisons of CEU, YRI, JPT and CHB.

Acknowledgements

References

1. J. B. S. Haldane, The rate of mutation of human genes. *Hereditas*. **35**, 267–273 (1949).

2. J. B. S. Haldane, in *Malaria: Genetic and Evolutionary Aspects*, K. R. Dronamraju, P. Arese, Eds. (Springer US, Boston, MA, 2006), pp. 175–187.

3. M. Fumagalli, M. Sironi, U. Pozzoli, A. Ferrer-Admetlla, L. Pattini, R. Nielsen, Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet.* **7**, e1002355 (2011).

4. C. Kosiol, T. Vinar, R. R. da Fonseca, M. J. Hubisz, C. D. Bustamante, R. Nielsen, A. Siepel, Patterns of positive selection in six Mammalian genomes. *PLoS Genet.* **4**, e1000144 (2008).

5. D. Enard, D. A. Petrov, Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. *Cell*. **175**, 360–371.e13 (2018).

6. D. Enard, D. A. Petrov, Ancient RNA virus epidemics through the lens of recent adaptation in human genomes. *bioRxiv* (2020), p. 2020.03.18.997346.

7. A. C. Allison, Protection afforded by sickle-cell trait against subtertian malareal infection. *Br. Med. J.* **1**, 290–294 (1954).

8. P. C. Sabeti, The International HapMap Consortium, P. Varilly, B. Fry, J. Lohmueller, E. Hostetter, C. Cotsapas, X. Xie, E. H. Byrne, S. A. McCarroll, R. Gaudet, S. F. Schaffner, E. S. Lander, Genome-wide detection and characterization of positive selection in human populations. *Nature*. **449** (2007), pp. 913–918.

9. S. M. Kamal, A. K. M. M. Rashid, M. A. Bakar, M. A. Ahad, Anthrax: an update. *Asian Pac. J. Trop. Biomed.* **1**, 496–501 (2011).

10. R. C. Spencer, Bacillus anthracis. *J. Clin. Pathol.* **56**, 182–187 (2003).

11. F. Zhao, S. McParland, F. Kearney, L. Du, D. P. Berry, Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet. Sel. Evol.* **47**, 49 (2015).

12. F.-H. Lv, S. Agha, J. Kantanen, L. Colli, S. Stucki, J. W. Kijas, S. Joost, M.-H. Li, P. Ajmone Marsan, Adaptations to climate-mediated selective pressures in sheep. *Mol. Biol. Evol.* **31**, 3324–3343 (2014).

13. F. H. Leendertz, H. Ellerbrok, C. Boesch, E. Couacy-Hymann, K. Mätz-Rensing, R. Hakenbeck, C. Bergmann, P. Abaza, S. Junglen, Y. Moebius, L. Vigilant, P. Formenty, G. Pauli, Anthrax kills wild chimpanzees in a tropical rainforest. *Nature*. **430**, 451–452 (2004).

14. C. Hoffmann, F. Zimmermann, R. Biek, H. Kuehl, K. Nowak, R. Mundry, A. Agbor, S. Angedakin, M. Arandjelovic, A. Blankenburg, G. Brazolla, K. Corogenes, E. Couacy-Hymann, T. Deschner, P. Dieguez, K. Dierks, A. Düx, S. Dupke, H. Eshuis, P. Formenty, Y. G. Yuh, A. Goedmakers, J. F. Gogarten, A.-C. Granjon, S. McGraw, R. Grunow, J. Hart, S. Jones, J. Junker, J. Kiang, K. Langergraber, J. Lapuente, K. Lee, S. A. Leendertz, F. Léguillon, V. Leinert, T. Löhrich, S. Marrocoli, K. Mätz-Rensing, A. Meier, K. Merkel, S. Metzger, M. Murai, S. Niedorf, H. De Nys, A. Sachse, J. van Schijndel, U. Thiesen, E. Ton, D. Wu, L. H. Wieler, C. Boesch, S.

R. Klee, R. M. Wittig, S. Calvignac-Spencer, F. H. Leendertz, Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest. *Nature*. **548**, 82–86 (2017).

15. T. Pearson, J. D. Busch, J. Ravel, T. D. Read, S. D. Rhoton, J. M. U'Ren, T. S. Simonson, S. M. Kachur, R. R. Leadem, M. L. Cardon, M. N. Van Ert, L. Y. Huynh, C. M. Fraser, P. Keim, Phylogenetic discovery bias in Bacillus anthracis using single-nucleotide polymorphisms from whole-genome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13536–13541 (2004).

16. M. N. Van Ert, W. R. Easterday, L. Y. Huynh, R. T. Okinaka, M. E. Hugh-Jones, J. Ravel, S. R. Zanecki, T. Pearson, T. S. Simonson, J. M. U'Ren, S. M. Kachur, R. R. Leadem-Dougherty, S. D. Rhoton, G. Zinser, J. Farlow, P. R. Coker, K. L. Smith, B. Wang, L. J. Kenefic, C. M. Fraser-Liggett, D. M. Wagner, P. Keim, Global genetic population structure of Bacillus anthracis. *PLoS One*. **2**, e461 (2007).

17. F. G. Priest, M. Barker, L. W. J. Baillie, E. C. Holmes, M. C. J. Maiden, Population structure and evolution of the Bacillus cereus group. *J. Bacteriol.* **186**, 7959–7970 (2004).

18. F. M. Laforce, Woolsorters' disease in England. *Bull. N. Y. Acad. Med.* **54**, 956–963 (1978).

19. J. A. Witkowski, L. C. Parish, The story of anthrax from antiquity to the present: a biological weapon of nature and humans. *Clin. Dermatol.* **20**, 336–342 (2002).

20. M. Schwartz, Dr. Jekyll and Mr. Hyde: a short history of anthrax. *Mol. Aspects Med.* **30**, 347–355 (2009).

21. J. E. Comer, A. K. Chopra, J. W. Peterson, R. König, Direct inhibition of T-lymphocyte activation by anthrax toxins in vivo. *Infect. Immun.* **73**, 8275–8281 (2005).

22. S. R. Paccani, F. Tonello, R. Ghittoni, M. Natale, L. Muraro, M. M. D'Elios, W.-J. Tang, C. Montecucco, C. T. Baldari, Anthrax toxins suppress T lymphocyte activation by disrupting antigen receptor signaling. *J. Exp. Med.* **201**, 325–331 (2005).

23. C. T. Baldari, F. Tonello, S. R. Paccani, C. Montecucco, Anthrax toxins: A paradigm of bacterial immune suppression. *Trends Immunol.* **27**, 434–440 (2006).

24. H. M. Scobie, G. J. A. Rainey, K. A. Bradley, J. A. T. Young, Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5170–5174 (2003).

25. J. Sun, P. Jacquez, Roles of Anthrax Toxin Receptor 2 in Anthrax Toxin Membrane Insertion and Pore Formation. *Toxins* . **8**, 34 (2016).

26. M. Moayeri, S. H. Leppla, The roles of anthrax toxin in pathogenesis. *Curr. Opin. Microbiol.* **7**, 19–24 (2004).

27. B. J. Schmiedel, D. Singh, A. Madrigal, A. G. Valdovino-Gonzalez, B. M. White, J. Zapardiel-Gonzalo, B. Ha, G. Altay, J. A. Greenbaum, G. McVicker, G. Seumois, A. Rao, M. Kronenberg, B. Peters, P. Vijayanand, Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell*. **175**, 1701–1715.e16 (2018).

28. G. F. Harrison, J. Sanz, J. Boulais, M. J. Mina, J.-C. Grenier, Y. Leng, A. Dumaine, V. Yotova, C. M. Bergey, S. L. Nsobya, S. J. Elledge, E. Schurr, L. Quintana-Murci, G. H. Perry, L. B.

Barreiro, Natural selection contributed to immunological differences between hunter-gatherers and agriculturalists. *Nature Ecology & Evolution*. **3**, 1253–1264 (2019).

29. A. V. Zimin, A. S. Cornish, M. D. Maudhoo, R. M. Gibbs, X. Zhang, S. Pandey, D. T. Meehan, K. Wipfler, S. E. Bosinger, Z. P. Johnson, G. K. Tharp, G. Marçais, M. Roberts, B. Ferguson, H. S. Fox, T. Treangen, S. L. Salzberg, J. A. Yorke, R. B. Norgren Jr, A new rhesus macaque assembly and annotation for next-generation sequencing analyses. *Biol. Direct*. **9**, 20 (2014).

30. P. Perez-Pinera, D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, K. W. Leong, F. Guilak, G. E. Crawford, T. E. Reddy, C. A. Gersbach, RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods*. **10**, 973–976 (2013).

31. N. Arora, K. R. Klimpel, Y. Singh, S. H. Leppla, Fusions of anthrax toxin lethal factor to the ADP-ribosylation domain of Pseudomonas exotoxin A are potent cytotoxins which are translocated to the cytosol of mammalian cells. *J. Biol. Chem.* **267**, 15542–15548 (1992).

32. Z. Wang, T. Chu, L. A. Choate, C. G. Danko, Identification of regulatory elements from nascent transcription using dREG. *Genome Res.* **29**, 293–303 (2018).

33. T.-K. Kim, M. Hemberg, J. M. Gray, A. M. Costa, D. M. Bear, J. Wu, D. A. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P. F. Worley, G. Kreiman, M. E. Greenberg, Widespread transcription at neuronal activity-regulated enhancers. *Nature*. **465**, 182–187 (2010).

34. C. G. Danko, L. A. Choate, B. A. Marks, E. J. Rice, Z. Wang, T. Chu, A. L. Martins, N. Dukler, S. A. Coonrod, E. D. Tait Wojno, J. T. Lis, W. L. Kraus, A. Siepel, Dynamic evolution of regulatory element ensembles in primate CD4+ T cells. *Nature Ecology & Evolution* (2018), doi:10.1038/s41559-017-0447-5.

35. S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. **159**, 1665–1680 (2014).

36. J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. **485**, 376–380 (2012).

37. O. Symmons, V. V. Uslu, T. Tsujimura, S. Ruf, S. Nassari, W. Schwarzer, L. Ettwiller, F. Spitz, Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).

38. M. Z. Ludwig, C. Bergman, N. H. Patel, M. Kreitman, Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature*. **403**, 564–567 (2000).

39. G. Kalay, P. J. Wittkopp, Nomadic enhancers: tissue-specific cis-regulatory elements of yellow have divergent genomic positions among Drosophila species. *PLoS Genet.* **6**, e1001222 (2010).

40. E. Cannavò, N. Koelling, D. Harnett, D. Garfield, F. P. Casale, L. Ciglar, H. E. Gustafson, R. R. Viales, R. Marco-Ferreres, J. F. Degner, B. Zhao, O. Stegle, E. Birney, E. E. M. Furlong, Genetic variants regulating expression levels and isoform diversity during embryogenesis. *Nature* (2016), doi:10.1038/nature20802.

41. S. Uebbing, J. Gockley, S. K. Reilly, A. A. Kocher, E. Geller, N. Gandotra, C. Scharfe, J. Cotney, J. P. Noonan, Massively parallel discovery of human-specific substitutions that alter neurodevelopmental enhancer activity. *bioRxiv* (2019), p. 865519.

42. G. B. Van Ness, Ecology of anthrax. *Science*. **172**, 1303–1307 (1971).

43. G. Schmid, A. Kaufmann, Anthrax in Europe: its epidemiology, clinical characteristics, and role in bioterrorism. *Clin. Microbiol. Infect.* **8**, 479–488 (2002).

44. D. Kollek, "The history of anthrax", by Sternbach G. *J. Emerg. Med.* **26** (2004), p. 354; author reply 354.

45. M. Martchenko, S. I. Candille, H. Tang, S. N. Cohen, Human genetic variation altering anthrax toxin sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2972–2977 (2012).

46. M. DeGiorgio, C. D. Huber, M. J. Hubisz, I. Hellmann, R. Nielsen, SweepFinder2: increased sensitivity, robustness and flexibility. *Bioinformatics*. **32**, 1895–1897 (2016).

47. S. Liu, M. Moayeri, S. H. Leppla, Anthrax lethal and edema toxins in anthrax pathogenesis. *Trends Microbiol.* **22**, 317–325 (2014).

48. S. Liu, Y. Zhang, M. Moayeri, J. Liu, D. Crown, R. J. Fattah, A. N. Wein, Z.-X. Yu, T. Finkel, S. H. Leppla, Key tissue targets responsible for anthrax-toxin-induced lethality. *Nature*. **501**, 63–68 (2013).

49. D. E. Lowe, I. J. Glomski, Cellular and physiological effects of anthrax exotoxin and its relevance to disease. *Front. Cell. Infect. Microbiol.* **2**, 76 (2012).

50. H. B. Golden, L. E. Watson, H. Lal, S. K. Verma, D. M. Foster, S.-R. Kuo, A. Sharma, A. Frankel, D. E. Dostal, Anthrax toxin: pathologic effects on the cardiovascular system. *Front. Biosci.* . **14**, 2335–2357 (2009).

51. J. Brojatsch, A. Casadevall, D. L. Goldman, Molecular determinants for a cardiovascular collapse in anthrax. *Front. Biosci.* . **6**, 139–147 (2014).

52. Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y.-C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K.-H. Farh, S. Feizi, R. Karlic, A.-R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. M. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L.-H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes. *Nature*. **518**, 317–330 (2015).

53. T. Cohen, C. Mordechai, A. Eran, D. Mishmar, mtDNA eQTLs and the m1A 16S rRNA modification explain mtDNA tissue-specific gene expression pattern in humans. *bioRxiv* (2018), p. 495838.

54. G. Sternbach, The history of anthrax. *J. Emerg. Med.* **24**, 463–467 (2003).

55. P. Keim, A. Kalif, J. Schupp, K. Hill, S. E. Travis, K. Richmond, D. M. Adair, M. Hugh-Jones, C. R. Kuske, P. Jackson, Molecular evolution and diversity in Bacillus anthracis as detected by amplified fragment length polymorphism markers. *J. Bacteriol.* **179**, 818–824 (1997).

56. F. W. Eurich, THE HISTORY OF ANTHRAX IN THE WOOL INDUSTRY OF BRADFORD, AND OF ITS CONTROL. *Lancet.* **207**, 107–110 (1926).

57. N. C. Durand, M. S. Shamim, I. Machol, S. S. P. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems.* **3**, 95–98 (2016).

58. H. Zhao, Z. Sun, J. Wang, H. Huang, J.-P. Kocher, L. Wang, CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics.* **30**, 1006–1007 (2014).

59. W. J. Kent, R. Baertsch, A. Hinrichs, W. Miller, D. Haussler, Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11484–11489 (2003).

60. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* **26**, 589–595 (2010).

61. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* **26**, 841–842 (2010).

62. R. M. Kuhn, D. Haussler, W. J. Kent, The UCSC genome browser and associated tools. *Brief. Bioinform.* **14**, 144–161 (2013).

63. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods.* **14**, 417–419 (2017).

64. 1000 Genomes Project Consortium, A. Auton, L. D. Brooks, R. M. Durbin, E. P. Garrison, H. M. Kang, J. O. Korbel, J. L. Marchini, S. McCarthy, G. A. McVean, G. R. Abecasis, A global reference for human genetic variation. *Nature.* **526**, 68–74 (2015).

65. G. McVicker, D. Gordon, C. Davis, P. Green, Widespread genomic signatures of natural selection in hominid evolution. *PLoS Genet.* **5**, e1000471 (2009).

66. A. Kong, G. Thorleifsson, D. F. Gudbjartsson, G. Masson, A. Sigurdsson, A. Jonasdottir, G. B. Walters, A. Jonasdottir, A. Gylfason, K. T. Kristinsson, S. A. Gudjonsson, M. L. Frigge, A. Helgason, U. Thorsteinsdottir, K. Stefansson, Fine-scale recombination rate differences between sexes, populations and individuals. *Nature.* **467**, 1099–1103 (2010).

67. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics.* **27**, 2156–2158 (2011).