Ancient viral genomes reveal introduction of HBV and B19V to Mexico during the transatlantic slave trade.

Axel A. Guzmán-Solís[1]*, Daniel Blanco-Melo[2*‡], Viridiana Villa-Islas[1], Miriam J. Bravo-López[1], Marcela Sandoval-Velasco[3], Julie K. Wesp[4], Jorge A. Gómez-Valdés[5], María de la Luz Moreno-Cabrera[6], Alejandro Meraz-Moreno[6], Gabriela Solís-Pichardo[7], Peter Schaaf[8], Benjamin R. tenOever[2,9,10] & María C. Ávila-Arcos[1‡].

1.International Laboratory for Human Genome Research, Universidad Nacional Autónoma de México (México); 2.Department of Microbiology, Icahn School of Medicine at Mount Sinai (NY, USA); 3.Section for Evolutionary Genomics, The Globe Institute, Faculty of Health, University of Copenhagen (Denmark); 4.Department of Sociology and Anthropology, North Carolina State University (USA); 5.Escuela Nacional de Antropología e Historia (México); 6.Instituto Nacional de Antropología e Historia (México); 7.Laboratorio Universitario de Geoquímica Isotópica (LUGIS), Instituto de Geología, Universidad Nacional Autónoma de México (México); 8.LUGIS, Instituto de Geofísica, Universidad Nacional Autónoma de México (México); 9.Virus Engineering Center for Therapeutics and Research (VECToR), Icahn School of Medicine at Mount Sinai (NY, USA); 10.Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai (NY, USA). * These authors contributed equally to this work. ‡ Correspondence: daniel.blancomelo@mssm.edu (D.B-M.), mavila@liigh.unam.mx (M.C.A-A.).

ABSTRACT

After the European colonization of the Americas there was a dramatic population collapse of the Indigenous inhabitants caused in part by the introduction of new pathogens. Although there is much speculation on the etiology of the Colonial epidemics, direct evidence for the presence of specific viruses during the Colonial era is lacking. To uncover the diversity of viral pathogens during this period, we designed an enrichment assay targeting ancient DNA (aDNA) from viruses of clinical importance and applied it on DNA extracts from individuals found in a Colonial (16^{th} c. – 18^{th} c.) hospital and a Colonial chapel where records suggest victims of epidemics were buried during important outbreaks in Mexico City. This allowed us to reconstruct three ancient human parvovirus B19 genomes, and one ancient human hepatitis B virus genome from distinct individuals. The viral genomes are similar to African strains, consistent with the inferred morphological and genetic African ancestry of the hosts as well as with the isotopic analysis of the human remains, suggesting an origin on the African continent. This study provides direct molecular evidence of ancient viruses being transported to the Americas during the transatlantic slave trade and their subsequent introduction to New Spain. Altogether, our observations enrich the discussion about the etiology of infectious diseases during the Colonial period in Mexico.

INTRODUCTION

European colonization in the Americas resulted in a notable genetic exchange mainly between Native American populations, Europeans, and Africans[1–3]. Along

with human migrations, numerous new species were introduced to the Americas including bacterial and viral pathogens, which played a major role in the dramatic population collapse that afflicted the immunologically-naïve Indigenous inhabitants[4,5]. Among these pathogens, viral diseases, such as smallpox, measles and mumps have been proposed to be responsible for many of the devastating epidemics during the Colonial period[4]. Remarkably, the pathogen(s) responsible for the deadliest epidemics reported in New Spain (the Spanish viceroyalty that corresponds to Mexico, Central America, and the current US southwest states) remains unknown and is thought to have caused millions of deaths during the 16th century[4]. Indigenous populations were drastically affected by these mysterious epidemics, generically referred to as *Cocoliztli* ("pest" in Nahuatl)[6], followed by Africans and to a lesser extent European people[4,6,7]. Symptoms of the 1576 *Cocoliztli* epidemic were described in autopsy reports of victims treated at the "Hospital Real de San José de los Naturales" (HSJN)[6,8], the first hospital in Mexico dedicated specifically to treat the Indigenous population[6,8] (Figure 1a-b). The symptoms described included high fever, severe headache, neurological disorders, internal and external bleeding, hepatitis and intense jaundice[4,6,7]. This has led some scholars to postulate that the etiological agent of the *Cocoliztli* epidemic was a hemorrhagic fever virus[4,9], although others have suggested that the symptoms could be explained by bacterial infections[6,10].

The study of ancient viral genomes has revealed important insights into the evolution of specific viral families[11–19], as well as their interaction with human populations[20]. To explore the presence of viral pathogens in circulation during

epidemic periods in New Spain, we leveraged the vast historical and archeological information available for the Colonial HSJN. These include the skeletal remains of over 600 individuals recovered from mass burials associated with the hospital's architectural remnants (Figure 1b). Many of these remains were retrieved from burial contexts suggestive of an urgent and simultaneous disposal of the bodies, as in the case of an epidemic[8,21]. Prior bioarcheological research has shown that the remains of a number of individuals in the HSJN collection displayed dental modifications and/or morphological indicators typical of African ancestry[21], consistent with historical and archeological research that documents the presence of a large number of both free and enslaved Africans and their descendants in Colonial Mexico[3]. Indeed a recent paleogenomics study reported a Sub-Saharan African origin of three individuals from this collection[22].

Here we describe the screening for viral pathogens that circulated in New Spain during Colonial times, using ancient DNA (aDNA) techniques (Supplementary Figure 1). For this work, we sampled skeletal human remains recovered from the HSJN, and from a second site in Mexico City ("La Concepcion" chapel) belonging to one of the first catholic conversion centers in New Spain[23], where records suggest victims of epidemics were also buried (Figure 1a). We report the reconstruction of ancient hepatitis B virus (HBV) and human parvovirus B19 (B19V) genomes recovered from these remains. Our findings provide a direct molecular evidence of human viral pathogens of African origin being introduced to New Spain during the transatlantic slave trade.

RESULTS

We sampled the skeletal remains from two archeological sites, a Colonial Hospital and a Colonial chapel in Mexico City (Figure 1a). For the HSJN (as a part of a broader ongoing study of African ancestry), 21 dental samples (premolar and molar teeth) were selected based on previous morphometric analyses and dental modifications that suggested an African ancestry[21,24–26]. The African presence in the Indigenous Hospital might reflect an urgent response to an epidemic outbreak, since hospitals treated patients regardless of the origin of the affected individuals during serious public health crises[21]. Dental samples of five additional individuals were selected (based on their conservation state) from “La Concepción” chapel (COY), which is located 10 km south of the HSJN in Coyoacán, a Pre-Hispanic Indigenous neighborhood that later became the first Spanish settlement in Mexico City after the fall of Tenochtitlan[23]. Following strict aDNA protocols, we processed these dental samples to isolate aDNA for next-generation sequencing (NGS) (Methods). Teeth roots (which are vascularized) can be assumed to be a good source of pathogen DNA[27], especially in the case of viruses that are widespread in the bloodstream during a systemic infection. Accordingly, a number of previous studies have successfully recovered ancient viral DNA from teeth roots[14–16,22].

Metagenomic analysis with MALT[10] on the NGS data using the Viral NCBI RefSeq database as a reference[28], revealed that sixteen samples contained traces of viral DNA, particularly of sequences related to *Hepadnaviridae*, *Herpesviridae*, *Parvoviridae* and *Poxviridae* (Figure 1c, Supplementary Figure 2, Methods). These viral hits revealed the potential to recover ancient viral genomes from these

samples. We selected thirteen samples for further screening based on the availability of the NGS library and the quality of the detected viral hit (best hit when using megaBLAST[29] to NCBI nr database) to a clinically important virus (HBV, B19V, Papillomavirus, Smallpox). To isolate and enrich the viral DNA fraction in the sequencing libraries, biotinylated single-stranded (ss) RNA probes designed to capture sequences from diverse human viral pathogens were synthesized (Supplementary Table 1). The selection of the viruses included in the capture design considered the following criteria: 1) DNA viruses previously retrieved from archeological human remains (i.e. Hepatitis B virus, Human Parvovirus B19, Variola Virus), 2) representative viruses from families capable of integrating into the human genome (i.e. *Herpesviridae*, *Papillomaviridae*, *Polyomaviridae*, *Circoviridae*) or 3) RNA viruses with a DNA intermediate (i.e. *Retroviridae*). Additionally, a virus-negative aDNA library, which showed no hits to any viral family included in the capture assay (except for a frequent *Poxviridae*-like region identified as an Alu repeat[30]), was captured and sequenced as a negative control (HSJN177) to estimate the efficiency of our capture assay. Four post-capture libraries had a ~100-300-fold increase of HBV-like hits or a ~50-100-fold increase of B19V-like hits (Figure 1c, Supplementary Table 2) compared to their corresponding pre-capture libraries (Methods). In contrast, the captured negative control (HSJN177) presented a negligible enrichment of these viral hits (Figure 1c, Supplementary Table 2).

We verified the authenticity of the viral sequences by querying the mapped reads against the non-redundant (nr) NCBI database using megaBLAST[29]. We only

retained reads for which the top hit was to either B19V or HBV, respectively (Supplementary Table 3). For each sample, the retained reads were mapped simultaneously to different available genotype sequences for their respective virus, and the genotype with the best sequence coverage was selected as the reference for final mapping (Supplementary Methods). To confirm the ancient origin of these viral reads, we evaluated the misincorporation damage patterns using the program mapDamage 2.0[31], which revealed an accumulation of C to T mutations towards their 5' terminal site with an almost symmetrical G to A pattern on the 3' end (Figure 2a, Supplementary Figure 3), as expected for aDNA[32]. Three ancient B19V genomes were reconstructed (Figure 2b, Supplementary Table 3) with sequence coverages between 92.37% and 99.1%, and average depths of 2.98-15.36X along their single stranded DNA (ssDNA) coding region, which excludes the double stranded DNA (dsDNA) hairpin regions at each end of the genome[33]. These inverse terminal repeats (ITRs) displayed considerably higher depth values compared to the coding region (<218X) consistent with the better *post-mortem* preservation of dsDNA compared to ssDNA[34] (Figure 2b). In addition, we reconstructed one ancient HBV genome (Figure 2c, Supplementary Table 3) at 30.8X average depth and with a sequence coverage of 89.9%, including its ssDNA region at a reduced depth (<10X). This genome presents a 6 nucleotide (nt) insertion in the core gene, that is characteristic of the genotype A[35]. Further phylogenetic analyses (Methods) revealed that the Colonial HBV genome clustered with modern sequences corresponding to sub-genotype A4 (previously named A6)[36] (Figure 3a, Supplementary Figure 4). The Genotype A (HBV/GtA) has a broad diversity in Africa reflecting its long history in this continent[35,37], while the

sub-genotype A4 has been recovered uniquely from African individuals in Belgium[38] and has never been found in the Americas. Regarding the three Colonial B19V genomes from individuals HSJN240, COYC4 and HSJNC81, these were phylogenetically closer to modern B19V sequences belonging to genotype 3 (Figure 3b, Supplementary Figure 5). This B19V genotype is divided into two sub-genotypes: 3a that is mostly found in Africa, and 3b, which is proposed to have spread outside Africa in the last decades[39]. The viral sequences from the individuals HSJN240 and COYC4 are similar to sub-genotype 3b genomes sampled from immigrants (Morocco, Egypt and Turkey) in Germany[40] (Figure 3b, Supplementary Figure 5); while the sequence of the individual HSJNC81 is more similar to a divergent sub-genotype 3a strain (Figure 3b, Supplementary Figure 5) retrieved from a child with severe anemia from France[41].

We recalibrated the molecular clock of HBV (Supplementary Figure 4b) and B19V (Supplementary Figure 5b) with a dated coalescent phylogeny analysis that included our ancient Colonial viral genomes, as well as other published ancient and modern HBV and B19V genomes from different genotypes in diverse geographical regions. The median evolutionary rate for HBV was estimated to be between 1.1×10^{-5} and 9.5×10^{-6} (mean of 7.1×10^{-6}) nucleotide substitutions per site per year (s/s/y) with a divergence time from last common ancestor between 8.3 and 13.61 (mean of 10.68) thousand years ago (ka) under a relaxed molecular clock model and a coalescent exponential population prior (Supplementary Figure 4b). For B19V, the estimated substitution rate is 1.3×10^{-5} and 9.9×10^{-6} (mean of 1.17×10^{-6}) s/s/y with a divergence time from last common ancestor between 9.45

and 11.83 (mean 10.6) ka under a strict molecular clock model and a coalescent Bayesian skyline population prior (Supplementary Figure 5b). The recalibrated estimations are similar to previous reports[15,16] that included only Eurasian viral genomes considerably older (~0.2–24.0 ka) than our Colonial viral genomes.

The similarity of our ancient genomes to African strains is consistent with the previously reported morphological indicators of African ancestry[21,24–26] of the corresponding HSJN individuals, as well as to an ancient DNA analysis based on three individuals from the HSJN collection[22]. We thus used the *de novo* generated sequence data to determine the mitochondrial haplogroup of the hosts, as well as their autosomal genetic ancestry using the 1000 Genomes Project[42] as a reference panel (Figure 3c, Supplementary Table 4). The nuclear genetic ancestry analysis showed that all three HSJN individuals, from which the reconstructed viral genomes were isolated, fall within African genetic variation in a Principal Component Analysis plot (Figure 3c), while their mitochondrial DNA belong to the L haplogroup, which has high frequency in African populations (Supplementary Table 4). Additionally, we performed $^{87}Sr/^{86}Sr$ isotopic analysis on two of the HSJN individuals using teeth enamel as well as phalange (HSJN240) or parietal bone (HSJNC81) to provide insights on the places of birth (adult enamel) and where the last years of life were spent (phalange/parietal). The $^{87}Sr/^{86}Sr$ ratios measured on the enamel of the individual HSJNC81 (0.71098) and HSJN240 (0.71109) are similar to average $^{87}Sr/^{86}Sr$ ratios found in soils and rocks from West Africa (average of 0.71044, Supplementary Figure 6, Supplementary Tables 5 and 6). In

contrast, the $^{87}Sr/^{86}Sr$ ratios on the parietal and phalange bones from the HSJNC81 (0.70672) and HSJN240 (0.70755), show lower values similar to those observed in the Trans Mexican Volcanic Belt where the Mexico City Valley is located (0.70420 - 0.70550, Supplementary Figure 6, Supplementary Tables 5 and 6). Moreover, radiocarbon dating of HSJN240 (1442-1608 CE, years calibrated for 1σ) and HSJN194 (1472-1625 CE, years calibrated for 1σ) (Supplementary Table 4, Supplementary Figure 7) indicates that these individuals arrived during the first decades of the Colonial period, when the number of enslaved individuals arriving from Africa was particularly high[3]. Strikingly, Colonial individual COYC4, who was also infected with an African B19V strain, clusters with present-day Latin-American populations from the 1000 Genomes Project (Figure 3c), suggesting that following introduction from Africa, the virus (B19V) spread and infected people of different ancestries during Colonial times.

DISCUSSION

In this study we reconstructed one HBV and three B19V ancient genomes from four different individuals using NGS metagenomics and in-solution targeted enrichment methods (Figure 2b, c, Supplementary Figure 1). Several lines of evidence support that these ancient viral genomes are authentic and not an environmental contamination nor a capture artifact. First, our negative control did not show a considerable enrichment for B19V or HBV hits (Figure 1c). For those samples that showed an enrichment in viral sequences after capture, the reads covered the reference genomes almost in their entirety and displayed deamination patterns at the terminal ends of the reads as expected for aDNA (Figure 2a).

Moreover, it is important to notice that B19V and HBV are blood-borne human pathogens that are not present in soil or the environment, and that DNA from these viruses had never been extracted before in the aDNA facilities used in this study.

We also described an unusual coverage pattern on the B19V genome, where the dsDNA hairpins at its terminal sites are highly covered reflecting a better stability of these regions over time (Figure 2b). Similarly, the partially circular dsDNA genome from HBV was poorly covered at the ssDNA region (Figure 2c), as in at least three previous ancient HBV genomes[14] (Supplementary Discussion 1). The variable coverage in both viruses argues against an integration event of these viruses, that would result as an uniform dsDNA coverage; further analysis are needed to elucidate if the aDNA retrieved in this and other studies is coming from circulating virions or from cell-free DNA intermediates[43] produced after viral replication in bone marrow or liver, for B19V and HBV, respectively[44,45].

The ancient B19V genomes were assigned to genotype 3. This genotype is considered endemic and the most prevalent in West Africa (Ghana: 100%, 11/11; Burkina Faso: 100%, 5/5)[39,46,47] and a potential African origin has been suggested[46]. This genotype has been also sporadically found outside Africa[46,47] in countries historically tied to this continent, like Brazil (50%, 6/12)[48,49], India (15.4%, 2/13)[50], France (11.4%, 9/79)[41,51], and USA (0.85%, 1/117)[47] as well as in immigrants from Morocco, Egypt, and Turkey in Germany (6.7%, 4/59)[40]. Two other genotypes, 1 and 2 exist for this virus. Genotype 1 is the most common and is found worldwide, while the almost extinct genotype 2 is mainly found in elderly people from Northern Europe[52]. Ancient genomes from genotypes 1 and 2 have

264 been recovered from Eurasian samples, including a genotype 2 B19V genome
265 from a 10th century Viking burial in Greenland[16]. $^{87}Sr/^{86}Sr$ isotopes on individuals
266 from the same burial revealed they were immigrants from Iceland[16], suggesting an
267 introduction of the genotype 2 to contiguous North America during Viking
268 explorations of Greenland. While serological evidence indicates that B19V
269 currently circulates in Mexico, only the presence of genotype 1 has been formally
270 described using molecular analyses[53]. Taken together, our results are consistent
271 with an introduction of the genotype 3 to New Spain as a consequence of the
272 transatlantic slave trade imposed by the European colonization. This hypothesis is
273 supported by the $^{87}Sr/^{86}Sr$ isotopic analysis, which indicates that the individuals
274 from the HSJN with B19V (HSJN240, HSJNC81) were born in West Africa and
275 spent their last years of life in New Spain (Supplementary Figure 6). Furthermore,
276 the radiocarbon ages of individuals HSJN240, HSJN194 (Supplementary Figure 7)
277 support this notion as they correspond to the Early Colonial period, during which
278 the number of enslaved Africans arriving was higher compared to later periods[3].
279 Remarkably, a similar B19V belonging to the genotype 3 was recovered from an
280 Indigenous individual (COYC4) in an independent archeological site 10 Km south
281 of the HSJN (Figure 1a), supporting that viral transmissions between African
282 individuals and people of different ancestries occurred during the Colonial period in
283 Mexico City.
284
285 The genotype A from HBV is highly diverse in Africa, reflecting its long evolutionary
286 history, originated somewhere between Africa, Middle East and Central Asia[37]. The

287 introduction of the genotype A from Africa to America has been proposed based on
288 phylogenetic analysis of modern strains for Brazil[37,49] and Mexico[54], and more
289 precisely to the sub-genotype A1 for sequences from Martinique[55], Venezuela[56],
290 Haiti[57] and Colombia[58]. Recently, a similar introduction pattern was proposed for
291 the quasi genotype A3 based on an ancient HBV genome recovered from African
292 individuals sampled in Mexico[22]. The Colonial ancient HBV genome reconstructed
293 in our work is assigned to the genotype A4 (Figure 3a, Supplementary Figure 4),
294 which has never been reported in the Americas. A similar introduction pattern has
295 been proposed for other human-infecting viruses such as smallpox[7,59], based on
296 historical records; or Yellow fever virus[60], HTLMV-1[61], Hepatitis C virus (genotype
297 2)[62] and human herpes simplex virus[63] based on phylogenetic analysis of modern
298 strains from Afro-descendant or admixed human populations.
299
300 Although we cannot assert where exactly the African-born individuals in this study
301 contracted B19V or HBV (Africa, America, or the Middle Passage) nor if the cause
302 of their deaths can be attributed to such infections, the identification of ancient
303 B19V and HBV in contexts associated with Colonial epidemics in Mexico City is still
304 relevant in light of their paleopathological marks and the clinical information
305 available for the closest sequences in the phylogenetic analyses. The
306 reconstructed ancient B19V genome from individual HSJNC81 is closest to the V9
307 strain, which was isolated from an individual with severe anemia[41] (AJ249437)
308 (Figure 3b, Supplementary Figure 5). In fact, individual HSJNC81 displayed cribra

orbitalia in the eye sockets and porotic hyperostosis on the cranial vault (Supplementary Figure 8); these morphological changes to the skeleton are typically associated with anemias of varying different causes[64]. It is acknowledged that B19V can result in a severe or even fatal anemia due to the low level of hemoglobin when present simultaneously with other blood disorders, as thalassemia, sickle-cell anemia, malaria and iron deficiency[45,65]. Therefore, since B19V infects precursors of the erythroid lineage[45], it is possible that the morphological changes found in HSJNC81 might be the result of a severe anemia caused or enhanced by a B19V infection (Supplementary Discussion 2). Moreover, the identification of ancient B19V in a Colonial context is noteworthy considering there are historical records that describe the treatment of an outbreak of measles at the HSJN in 1531[21], that numerous cases of measles were reported during this period[4,8,59], and that several recent reports reveal that measles-like cases were actually attributable to B19V[66,67] (Supplementary Discussion 2). Nevertheless, this hypothesis requires future and more comprehensive studies aimed to characterize the presence of measles and rubella viruses from ancient remains. A task that imposes difficult technical challenges given that RNA is known to degrade rapidly, in fact the most ancient viral RNA genomes (1912 CE) have been recovered only from formalin-fixed tissue[17,18]. Additionally, historical records of the autopsies of victims of the 1576 *Cocoliztli* epidemic treated at the HSJN, describe the observation of enlarged hard liver and jaundice[4,6,7,9,68], which could be explained by severe liver damage or epidemic hepatitis[4,6]. This is noteworthy given both viruses HBV and B19V proliferate in the liver, have been isolated from liver tissue and are associated with hepatitis and jaundice[44,45]. However, it is important to acknowledge

that both viruses have also been previously identified in aDNA datasets not necessarily associated with disease or epidemic contexts[11,12,14,15], thus establishing a direct link would require additional samples and a more comprehensive pathogen screening to rule out the involvement of other pathogens. Finally, although our data is not a strong evidence that the reported manifestations of liver damage in *Cocolitzi* autopsies were directly caused by HBV or B19V, the identification of these viruses in likely victims of epidemic outbreaks in the Colonial period opens up new opportunities for investigating the presence of these viruses in similar contexts. This type of research is particularly relevant when considering previous hypotheses favoring the synergistic action of different pathogens in these devastating Colonial epidemics[7] (Supplementary Discussion 3).

It is important to emphasize that our findings should be interpreted with careful consideration of the historical and social context of the transatlantic slave trade. This cruel episode in history involved the forced transportation of millions of individuals to the Americas (ca. 250,000 to New Spain[3]) under inhumane, unsanitary and overcrowded conditions that, with no doubt, favored the spread of infectious diseases[59]. Therefore, the introduction of these and other pathogens from Africa to the Americas should be attributed to the brutal and harsh conditions of the Middle Passage that enslaved Africans were subjected to by traders and colonizers, and not to the African peoples themselves. Moreover, the adverse life conditions for enslaved Africans and Native Americans, especially during the first decades after colonization, surely favored the spread of diseases and emergence

of epidemics[59]. Integrative and multidisciplinary approaches are thus needed to understand this phenomenon at its full spectrum.

In summary, our study provides direct aDNA evidence of HBV and B19V introduced to the Americas from Africa during the transatlantic slave trade. The isolation and characterization of these ancient HBV and B19V genomes represent an important contribution to the recently reported ancient viral genome reported in the Americas[22] (only one before the present study). Our results expand our knowledge on the viral agents that were in circulation during Colonial epidemics like *Cocoliztli*, some of which resulted in the catastrophic collapse of the immunologically-naïve Indigenous population. Although we cannot assign a direct causality link between HBV and B19V and *Cocoliztli*, our findings confirm that these potentially harmful viruses were indeed circulating in individuals found in archeological contexts associated with this epidemic outbreak. Further analyses from different sites and samples will help understand the possible role of these and other pathogens in Colonial epidemics, as well as the full spectrum of pathogens that were introduced to the Americas during European colonization.

METHODS

Sample selection and DNA extraction

Dental samples (premolars and molars) were obtained from twenty-one individuals from the skeletal collection of the HSJN, selected based on their African-related skeletal indicators[21,24–26]. Five additional samples were taken from “La Concepción” chapel, based on their conservation state. Permits to carry out this sampling and

aDNA analyses were obtained by the Archeology Council of the National Institute of Anthropology and History (INAH).

DNA extraction and NGS library construction

Bone samples were transported to a dedicated ancient DNA clean-room laboratory at the International Laboratory for Human Genome Research (LIIGH-UNAM, Querétaro, Mexico), were DNA extraction and NGS-libraries construction was performed under the guidelines on contamination control for aDNA studies[69]. Previously reported aDNA extraction protocols were used for the HSJN[70] and COY[71] samples. Double-stranded DNA (dsDNA) indexed sequencing libraries were constructed from the DNA extract, as previously reported[72]. In order to detect contaminants in reagents or by human manipulation, extraction and library constructions protocols included negative controls (NGS blanks) that were analyzed in parallel with the same methodology. The resulting NGS dsDNA indexed libraries were quantified with a Bioanalyzer 2100 (Agilent) and pooled into equimolar concentrations.

NGS sequencing

Pooled libraries were paired-end sequenced on an Illumina NextSeq550 at the "Laboratorio Nacional de Genómica para la Biodiversidad" (LANGEBIO, Irapuato, Mexico), with a Mid-output 2x75 format. The reads obtained (R1 and R2) were merged (>11bp overlap) and trimmed with AdapterRemoval 1.5.4[73]. Overlapping reads (>30 bp in length) were kept and mapped to the human genome (hg19) using BWA 0.7.13[74], mapped reads were used for further human analysis (genetic

ancestry, and mitochondrial haplogroup determination), whereas unmapped reads were used for metagenomic analysis and viral genome reconstruction.

Metagenomic analyses

The NCBI Viral RefSeq database was downloaded on February 2018; this included 7530 viral genomes. MALT 0.4.0[10] software was used to taxonomically classify the reads using the viral genomes database. The viral database was formatted automatically with malt-build once, and not human (unmapped) reads were aligned with malt-run (85 minimal percent identity). The produced RMA files with the viral abundances were normalized (default parameters) and compared to all the samples from the same archeological site with MEGAN 6.8.0[75].

Capture-enrichment Assay

Twenty-nine viruses were included in the in-solution enrichment design, the complete list of NCBI IDs is provided in Supplementary methods 5 and Supplementary Table 1. It contained viral genomes previously recovered from archeological remains like B19V, B19V-V9, and HBV (consensus genomes), selected VARV genes, as well as clinically important viral families that are able to integrate into the human genome, have dsDNA genomes, or dsDNA intermediates. The resulting design comprised 19,147 ssRNA 80 nt probes targeting with a 20 nt interspaced distance the whole or partial informative regions of eight viral families of clinical relevance (*Poxviridae, Hepadnaviridae, Parvoviridae, Herpesviridae, Retroviridae, Papillomaviridae, Polyomaviridae, Circoviridae*). To avoid a simultaneous false-positive DNA enrichment, low complexity regions and human-

like (hg38) sequences were removed (*in sillico*). The customized kit was produced by Arbor Biosciences (Ann Arbor, MI, USA). Capture-enrichment was performed on the indexed libraries based on the manufacturer's protocol (version 4) to pull-down aDNA with minimal modifications. Libraries were amplified with 18-20 cycles (Phusion U Hot Start DNA Polymerase by Thermo Fischer Scientific), purified with SPRISelect Magnetic Beads (Beckman Coulter) and quantified with a Bioanalyzer 2100 (Agilent). Amplified libraries were then pooled in different concentrations and deep sequenced yielding >1x10^6 non-human reads (Supplementary Table 4) in order to saturate the target viral genome. Reads generated from each enriched library were analyzed exactly as the shotgun (not-enriched) libraries. Normalized abundances between shotgun and captured libraries were compared in MEGAN 6.8.0[75] to evaluate the efficiency and specificity of the enrichment assay.

Viral datasets

The full list of accession numbers of the following datasets is given in Supplementary Methods 8.

HBV-Dataset-1 (HBV/DS1): comprises 35 HBV genomes from A-J human genotypes, 2 well-covered ancient HBV genomes (LT992443, LT992459) and a wholly monkey genome.

HBV-Dataset-2 (HBV/DS2): comprises 110 genomes based on a previous phylogenetic analysis[15], that included genomes from A-J genotypes as well as not-human primates HBV genomes (gibbon, gorilla, and chimpanzee), 16 ancient HBVgenomes[11,12,14,15] and one ancient HBV genome from this study (HSJN194).

B19V-Dataset-1 (B19V/DS1): comprises 13 B19V genomes from human genotypes 1-3 as well as a bovine parvovirus.

B19V-Dataset-2 (B19V/DS2): comprises 112 genomes from 1 to 3 B19V genotypes based on previous phylogenetic analysis[16], that included the 10 best-covered ancient genomes from genotype 1 and 2[16] as well as 3 ancient B19V from this study. Since many of the reported genomes are not complete, only the whole coding region (CDS) was used.

Genome Reconstruction and authenticity

HBV: Non-human reads were simultaneously mapped to HBV/DS1 with BWA (aln algorithm) with seedling disabled[76]. The reference sequence with the most hits was used to map uniquely to this reference and generate a BAM alignment without duplicates, from which damage patterns were determined and damaged sites rescaled using mapDamage 2.0[31], the rescaled alignment was used to produce a consensus genome. All the HBV mapped reads were analyzed through megaBLAST[29] using the whole NCBI nr database, in order to verify they were assigned uniquely to HBV (carried out with Krona 2.7[77]).

B19V: The reconstruction of the B19V ancient genome was done as previously reported from archeological skeletal remains[16], but increasing the stringency of some parameters as described next. Non-human reads were mapped against B19V/DS1 with BWA (aln algorithm) with seedling disabled[76], if more than 50% of the genome was covered, the sample was considered positive to B19V. Reads from the B19V-positive libraries were aligned with blastn (-evalue 0.001) to B19V/DS1 in order to recover all the parvovirus-like reads. To avoid local

alignments, only hits covering >85% of the read were kept and joined to the B19V mapped reads (from BWA), duplicates were removed. The resulting reads were analyzed with megaBLAST[29] using the whole NCBI nr database to verify the top hit was to B19V (carried out with Krona 2.7[77]). This pipeline was applied for two independent enrichments assays per sample and the filtered reads from the two capture rounds were joined. The merged datasets per sample were mapped using as a reference file the three known B19V genotypes with GeneiousPrime 2019.0.4[78] using median/fast sensibility and iterate up to 5 times. The genotype with the longest covered sequence was selected as the reference for further analysis. Deamination patterns for HBV and B19V were determined with mapDamage 2.0[31] and damaged sites were rescaled in the same program to produce a consensus whole genome using SAMtools 1.9[79].

Phylogenetic analyses

HBV/DS2 and B19V/DS2 were aligned independently in Aliview[80] (Muscle algorithm), and evolutionary models were tested under an AICc and BIC in jModelTest[81]. A neighbor joining tree with 1000 bootstraps was generated in GeneiousPrime 2019.0.4[78] using a Hasegawa-Kishino-Yano (HKY) model for both alignments. A maximum likelihood tree with 1000 bootstraps was produced in RAxML 8.2.10[82] using as a model a Generalized Time-Reversible (GTR)+G and GTR+G+I for B19V and HBV, respectively.

Since a temporal signal has been described for the coding region of B19V[16] and the whole genome of HBV[14,15], a Bayesian tree was generated to estimate the impact of the Colonial viral genomes on the divergence time from the most recent

499 common ancestor (MRCA). We used BEAST 1.8.4[83], a Bayesian skyline
500 population prior and a relaxed lognormal or strict molecular clock for B19V and a
501 Coalescent Exponential Population prior with relaxed lognormal molecular clock for
502 HBV as previously tested[15,16]. All parameters were mixed and converged into an
503 estimated sample size (ESS) >150 analyzed in Tracer 1.7[84]. The first 25% of trees
504 where discarded (burn in) and a Maximum Clade Credibility Tree was created with
505 TreeAnnotator[83]. The generated trees were visualized and edited in FigTree 1.4.3
506 with a midpoint root.
507
508 *Human population genetic analyses*
509 Human-mapped reads (BWA aln) obtained from the pre-capture sequence data of
510 viral-positive samples were used to infer the genetic ancestry of the hosts. A
511 Principal Components Analysis (PCA) was carried out using 16 populations from
512 the 1000 Genomes Project[42] reference panel including genotype data of 1,562,771
513 single nucleotide variants (SNVs) from 2,504 individuals (phase 3). Genomic
514 alignments of each ancient individual (HSJNC81, HSJN240, HSJN194 and
515 COYC4) were intersected with the positions of the SNVs present in the reference
516 panel genotype data. Pseudo haploid genotypes were called by randomly selecting
517 one allele at each intersected site and filtering by a base quality >30. Pseudo
518 haploid genotypes were also called for the complete reference panel. PCA was
519 performed on the merged ancient and modern dataset with smartpca (EIGENSOFT
520 package)[85,86] using the option *lsqproject* to project the ancient individuals into the
521 PC space defined by the modern individuals.
522

Mitochondrial haplogroup and sex determination

NGS reads were mapped to the human mitochondrial genome reference (rCRS) with BWA (aln algorithm, -l default), the alignment file was then used to generate a consensus mitochondrial genome with program Schmutzi[87] The assignment of the mitochondrial haplogroup was carried out with Haplogrep[88,89] using the consensus sequence as the input. Assignment of biological sex was inferred based on the fraction of reads mapped to the Y-chromosome (Ry) compared to those mapping to the Y and X-chromosome[90]. Ry<0.016 and Ry>0.075 were considered XX or XY genotype, respectively. The resulting sex was coherent with the inferred morphologically (Supplementary Methods).

ACKNOWLEDGMENTS

This work was funded by the Wellcome Trust Sanger grant number 208934/Z/17/Z, and by project IA201219 PAPIIT-DGAPA-UNAM. D.B-M is an Open Philanthropy Fellow of the Life Sciences Research Foundation (LSRF). We thank the INAH Archeology Council for the sample permissions for this study. We are grateful with Teodoro Hernández Treviño, Gerardo Arrieta García from the “Laboratorio Universitario de Geoquímica Isotópica” (LUGIS-UNAM) for their technical support in performing the $^{87}Sr/^{86}Sr$ analyses and to Luis Alberto Aguilar Bautista, Alejandro de León Cuevas, Carlos Sair Flores Bautista and Jair Garcia Sotelo from the “Laboratorio Nacional de Visualización Científica Avanzada” (LAVIS/UNAM) who stored our data and provided the computational resources to perform this study.

We thank Alejandra Castillo Carbajal and Carina Uribe Díaz for technical support throughout the project.

DATA AVAILABILITY

Reconstructed genomes from this study are available in Genbank under accession number MT108214, MT108215, MT108216, MT108217. Accession number of sequences used in phylogenetic analysis are indicated in supplementary information. NGS data is available upon reasonable request.

555 **REFERENCES:**

556

557 1. Rotimi, C. N., Tekola-Ayele, F., Baker, J. L. & Shriner, D. The African
558 Diaspora: History, Adaptation and Health. *Curr. Opin. Genet. Dev.* **41**, 77–84
559 (2016).
560 2. Salas, A. *et al.* The African Diaspora: Mitochondrial DNA and the Atlantic
561 Slave Trade. *Am. J. Hum. Genet.* **74**, 454–465 (2004).
562 3. Aguirre-Beltrán, G. La presencia del negro en México. *Rev. CESLA* **7**, 351–
563 367 (2005).
564 4. Acuna-Soto, R., Stahle, D. W., Therrell, M. D., Griffin, R. D. & Cleaveland,
565 M. K. When half of the population died: the epidemic of hemorrhagic fevers of 1576
566 in Mexico. *FEMS Microbiol. Lett.* **240**, 1–5 (2004).
567 5. Lindo, J. *et al.* A time transect of exomes from a Native American population
568 before and after European contact. *Nat. Commun.* **7**, 13175 (2016).
569 6. Malvido, E. & Viesca, C. La epidemia de cocoliztli de 1576. in *Ensayos*
570 *sobre la historia de las epidemias en México.* (eds. Florescano, E. & Malvido, E.)
571 27–32 (Instituto Mexicano del Seguro Social, 1982).
572 7. Somolinos d'Árdois, G. Las epidemias en México durante el siglo XVI. in
573 *Ensayos sobre la historia de las epidemias en México.* (eds. Florescano, E. &
574 Malvido, E.) 138–143 (Instituto Mexicano del Seguro Social, 1982).
575 8. Wesp, J. K. Caring for Bodies or Simply Saving Souls: The Emergence of
576 Institutional Care in Spanish Colonial America. in *New Developments in the*
577 *Bioarchaeology of Care: Further Case Studies and Expanded Theory* (eds. Tilley,
578 L. & Schrenk, A. A.) 253–276 (Springer International Publishing, 2017).
579 9. Marr, J. S. & Kiracofe, J. B. Was the huey cocoliztli a haemorrhagic fever?
580 *Med. Hist.* **44**, 341–362 (2000).
581 10. Vågene, Å. J. *et al.* Salmonella enterica genomes from victims of a major
582 sixteenth-century epidemic in Mexico. *Nat. Ecol. Evol.* **2**, 520 (2018).
583 11. Kahila Bar-Gal, G. *et al.* Tracing hepatitis B virus to the 16th century in a
584 Korean mummy. *Hepatol. Baltim. Md* **56**, 1671–1680 (2012).
585 12. Patterson Ross, Z. *et al.* The paradox of HBV evolution as revealed from a

16th century mummy. *PLoS Pathog.* **14**, e1006750 (2018).

13. Duggan, A. T. *et al.* 17th Century Variola Virus Reveals the Recent History of Smallpox. *Curr. Biol. CB* **26**, 3407–3412 (2016).

14. Krause-Kyora, B. *et al.* Neolithic and medieval virus genomes reveal complex evolution of hepatitis B. *eLife* **7**, (2018).

15. Mühlemann, B. *et al.* Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature* **557**, 418–423 (2018).

16. Mühlemann, B. *et al.* Ancient human parvovirus B19 in Eurasia reveals its long-term association with humans. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 7557–7562 (2018).

17. Xiao, Y.-L. *et al.* High-throughput RNA sequencing of a formalin-fixed, paraffin-embedded autopsy lung tissue sample from the 1918 influenza pandemic. *J. Pathol.* **229**, 535–545 (2013).

18. Düx, A. *et al. The history of measles: from a 1912 genome to an antique origin*. http://biorxiv.org/lookup/doi/10.1101/2019.12.29.889667 (2019) doi:10.1101/2019.12.29.889667.

19. Pajer, P. *et al.* Characterization of Two Historic Smallpox Specimens from a Czech Museum. *Viruses* **9**, (2017).

20. Spyrou, M. A., Bos, K. I., Herbig, A. & Krause, J. Ancient pathogen genomics as an emerging tool for infectious disease research. *Nat. Rev. Genet.* **20**, 323–340 (2019).

21. Meza, A. Presencia africana en el cementerio del Hospital Real de San Jóse de los Naturales. *Arqueol. Mex.* **119**, 40–44 (2013).

22. Barquera, R. *et al.* Origin and Health Status of First-Generation Africans from Early Colonial Mexico. *Curr. Biol.* S0960982220304826 (2020) doi:10.1016/j.cub.2020.04.002.

23. Moreno-Cabrera, M. de la L., Meraz-Moreno, A. & Cervantes-Rosado, J. Relleno aligerado con vasijas cerámicas en el templo de la Inmaculada Concepción, en Coyoacán. in *Boletín de monumentos históricos 35* vol. 35 121–134 (Instituto Nacional de Antropología e Historia, 2015).

24. Ruíz-Albarrán, P. *Estudio de variabilidad biológica en la colección*

esquelética Hospital Real de San José de los Naturales. Un acercamiento a tráves de la técnica de morfometría geométrica. (Escuela Nacional de Antropología e Historia., 2012).

25. Hernández-Lopez, P. E. & Negrete, S. *¿Realmente Eran Indios? Afinidad biológica entre las personas atendidas en el Hospital Real San Jose de los Naturales, siglos XVI - XVIII.* (Escuela Nacional de Antropología e Historia., 2012).

26. Karam-Tapia, C. E. *Estimación del Mestizaje Mediante la Morfología Dental en la Ciudad de México (Siglo XVI al XIX).* (Escuela Nacional de Antropología e Historia., 2012).

27. Key, F. M., Posth, C., Krause, J., Herbig, A. & Bos, K. I. Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends Genet. TIG* **33**, 508–520 (2017).

28. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35**, D61-65 (2007).

29. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

30. Tithi, S. S., Aylward, F. O., Jensen, R. V. & Zhang, L. FastViromeExplorer: a pipeline for virus and phage identification and abundance profiling in metagenomics data. *PeerJ* **6**, e4227 (2018).

31. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinforma. Oxf. Engl.* **29**, 1682–1684 (2013).

32. Briggs, A. W. *et al.* Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14616–14621 (2007).

33. Luo, Y. & Qiu, J. Human parvovirus B19: a mechanistic overview of infection and DNA replication. *Future Virol.* **10**, 155–167 (2015).

34. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).

35. Kramvis, A. Genotypes and Genetic Variability of Hepatitis B Virus. *Intervirology* **57**, 141–150 (2014).

36. Pourkarim, M. R., Amini-Bavil-Olyaee, S., Kurbanov, F., Van Ranst, M. & Tacke, F. Molecular identification of hepatitis B virus genotypes/subgenotypes: revised classification hurdles and updated resolutions. *World J. Gastroenterol.* **20**, 7152–7168 (2014).

37. Kostaki, E.-G. *et al.* Unravelling the history of hepatitis B virus genotypes A and D infection using a full-genome phylogenetic and phylogeographic approach. *eLife* **7**, (2018).

38. Pourkarim, M. R., Lemey, P., Amini-Bavil-Olyaee, S., Maes, P. & Van Ranst, M. Novel hepatitis B virus subgenotype A6 in African-Belgian patients. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* **47**, 93–96 (2010).

39. Hübschen, J. M. *et al.* Phylogenetic analysis of human parvovirus b19 sequences from eleven different countries confirms the predominance of genotype 1 and suggests the spread of genotype 3b. *J. Clin. Microbiol.* **47**, 3735–3738 (2009).

40. Schneider, B. *et al.* Simultaneous persistence of multiple genome variants of human parvovirus B19. *J. Gen. Virol.* **89**, 164–176 (2008).

41. Nguyen, Q. T. *et al.* Novel Human Erythrovirus Associated with Transient Aplastic Anemia. *J. Clin. Microbiol.* **37**, 2483–2487 (1999).

42. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

43. Cheng, A. P. *et al.* A cell-free DNA metagenomic sequencing assay that integrates the host injury response to infection. *Proc. Natl. Acad. Sci.* **116**, 18738–18744 (2019).

44. Yuen, M.-F. *et al.* Hepatitis B virus infection. *Nat. Rev. Dis. Primer* **4**, 18035 (2018).

45. Broliden, K., Tolfvenstam, T. & Norbeck, O. Clinical aspects of parvovirus B19 infection. *J. Intern. Med.* **260**, 285–304 (2006).

46. Candotti, D., Etiz, N., Parsyan, A. & Allain, J.-P. Identification and Characterization of Persistent Human Erythrovirus Infection in Blood Donor Samples. *J. Virol.* **78**, 12169–12178 (2004).

47. Rinckel, L. A., Buno, B. R., Gierman, T. M. & Lee, D. C. Discovery and

679 analysis of a novel parvovirus B19 Genotype 3 isolate in the United States.
680 *Transfusion (Paris)* **49**, 1488–1492 (2009).
681 48. Sanabani, S., Neto, W. K., Pereira, J. & Sabino, E. C. Sequence Variability
682 of Human Erythroviruses Present in Bone Marrow of Brazilian Patients with
683 Various Parvovirus B19-Related Hematological Symptoms. *J. Clin. Microbiol.* **44**,
684 604–606 (2006).
685 49. Freitas, R. B. *et al.* Molecular characterization of human erythrovirus B19
686 strains obtained from patients with several clinical presentations in the Amazon
687 region of Brazil. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* **43**, 60–65 (2008).
688 50. Jain, P. *et al.* Prevalence and genotypic characterization of human
689 parvovirus B19 in children with hemato-oncological disorders in North India. *J.*
690 *Med. Virol.* **87**, 303–309 (2015).
691 51. Servant, A. *et al.* Genetic Diversity within Human Erythroviruses:
692 Identification of Three Genotypes. *J. Virol.* **76**, 9124–9134 (2002).
693 52. Pyöriä, L. *et al.* Extinct type of human parvovirus B19 persists in tonsillar B
694 cells. *Nat. Commun.* **8**, 14930 (2017).
695 53. Valencia Pacheco, G. *et al.* Serological and molecular analysis of parvovirus
696 B19 infection in Mayan women with systemic lupus erythematosus in Mexico.
697 *Colomb. Médica* **v48**, 105–112 (2017).
698 54. Roman, S. *et al.* Occult hepatitis B in the genotype H-infected Nahuas and
699 Huichol native Mexican population. *J. Med. Virol.* **82**, 1527–1536 (2010).
700 55. Brichler, S. *et al.* African, Amerindian and European hepatitis B virus strains
701 circulate on the Caribbean Island of Martinique. *J. Gen. Virol.* **94**, 2318–2329
702 (2013).
703 56. Quintero, A. *et al.* Molecular epidemiology of hepatitis B virus in Afro-
704 Venezuelan populations. *Arch. Virol.* **147**, 1829–1836 (2002).
705 57. Andernach, I. E., Nolte, C., Pape, J. W. & Muller, C. P. Slave trade and
706 hepatitis B virus genotypes and subgenotypes in Haiti and Africa. *Emerg. Infect.*
707 *Dis.* **15**, 1222–1228 (2009).
708 58. Alvarado-Mora, M. V. *et al.* Phylogenetic analysis of complete genome
709 sequences of hepatitis B virus from an Afro-Colombian community: presence of

710 HBV F3/A1 recombinant strain. *Virol. J.* **9**, 244 (2012).
711 59. Mandujano-Sánchez, A., Solache, L. C. & Mandujano, M. A. Historia de las
712 Epidemias en el México Antiguo: Algunos Aspectos Biológicos y Sociales. in
713 *Ensayos sobre la historia de las epidemias en México.* (eds. Florescano, E. &
714 Malvido, E.) 9–21 (Instituto Mexicano del Seguro Social, 1982).
715 60. Bryant, J. E., Holmes, E. C. & Barrett, A. D. T. Out of Africa: A Molecular
716 Perspective on the Introduction of Yellow Fever Virus into the Americas. *PLoS*
717 *Pathog.* **3**, (2007).
718 61. Gadelha, S. R. *et al.* The origin of HTLV-1 in the South Bahia by
719 phylogenetic, mitochondrial DNA and β-globin analysis. *Retrovirology* **11**, P49
720 (2014).
721 62. Markov, P. V. *et al.* Phylogeography and molecular epidemiology of hepatitis
722 C virus genotype 2 in Africa. *J. Gen. Virol.* **90**, 2086–2096 (2009).
723 63. Forni, D. *et al.* Recent out-of-Africa migration of human herpes simplex
724 viruses. *Mol. Biol. Evol.* msaa001 (2020) doi:10.1093/molbev/msaa001.
725 64. Angel, J. L. Porotic hyperostosis, anemias, malarias, and marshes in the
726 prehistoric Eastern Mediterranean. *Science* **153**, 760–763 (1966).
727 65. Heegaard, E. D. & Brown, K. E. Human Parvovirus B19. *Clin. Microbiol.*
728 *Rev.* **15**, 485–505 (2002).
729 66. Rezaei, F. *et al.* Prevalence and genotypic characterization of Human
730 Parvovirus B19 in children with measles- and rubella-like illness in Iran. *J. Med.*
731 *Virol.* **88**, 947–953 (2016).
732 67. De Los Ángeles Ribas, M. *et al.* Identification of human parvovirus B19
733 among measles and rubella suspected patients from Cuba. *J. Med. Virol.* **91**,
734 1351–1354 (2019).
735 68. Acuna-Soto, R., Stahle, D. W., Cleaveland, M. K. & Therrell, M. D.
736 Megadrought and Megadeath in 16th Century Mexico. *Emerg. Infect. Dis.* **8**, 360–
737 362 (2002).
738 69. Warinner, C. *et al.* A Robust Framework for Microbial Archaeology. *Annu.*
739 *Rev. Genomics Hum. Genet.* **18**, 321–356 (2017).
740 70. Dabney, J. *et al.* Complete mitochondrial genome sequence of a Middle

741 Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl.*
742 *Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).
743 71. Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth.
744 *Nat. Protoc.* **2**, 1756–1762 (2007).
745 72. Meyer, M. & Kircher, M. Illumina Sequencing Library Preparation for Highly
746 Multiplexed Target Capture and Sequencing. *Cold Spring Harb. Protoc.* **2010**, 1–10
747 (2010).
748 73. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid
749 adapter trimming, identification, and read merging. *BMC Res. Notes* **9**, (2016).
750 74. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–
751 Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
752 75. Huson, D. H. *et al.* MEGAN Community Edition - Interactive Exploration and
753 Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput. Biol.* **12**,
754 e1004957 (2016).
755 76. Schubert, M. *et al.* Improving ancient DNA read mapping against modern
756 reference genomes. *BMC Genomics* **13**, 178 (2012).
757 77. Ondov, B. D., Bergman, N. H. & Phillippy, A. M. Interactive metagenomic
758 visualization in a Web browser. *BMC Bioinformatics* **12**, 385 (2011).
759 78. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop
760 software platform for the organization and analysis of sequence data.
761 *Bioinformatics* **28**, 1647–1649 (2012).
762 79. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools.
763 *Bioinformatics* **25**, 2078–2079 (2009).
764 80. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for
765 large datasets. *Bioinformatics* **30**, 3276–3278 (2014).
766 81. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more
767 models, new heuristics and high-performance computing. *Nat. Methods* **9**, 772
768 (2012).
769 82. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-
770 analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
771 83. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian

772 Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973
773 (2012).
774 84. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A.
775 Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.*
776 **67**, 901–904 (2018).
777 85. Price, A. L. *et al.* Principal components analysis corrects for stratification in
778 genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
779 86. Patterson, N., Price, A. L. & Reich, D. Population Structure and
780 Eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
781 87. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of
782 contamination and endogenous mitochondrial consensus calling for ancient DNA.
783 *Genome Biol.* **16**, 224 (2015).
784 88. Kloss-Brandstätter, A. *et al.* HaploGrep: a fast and reliable algorithm for
785 automatic classification of mitochondrial DNA haplogroups. *Hum. Mutat.* **32**, 25–32
786 (2011).
787 89. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup
788 classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**,
789 W58–W63 (2016).
790 90. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex
791 identification of ancient human remains using DNA shotgun sequencing. *J.*
792 *Archaeol. Sci.* **40**, 4477–4482 (2013).
793

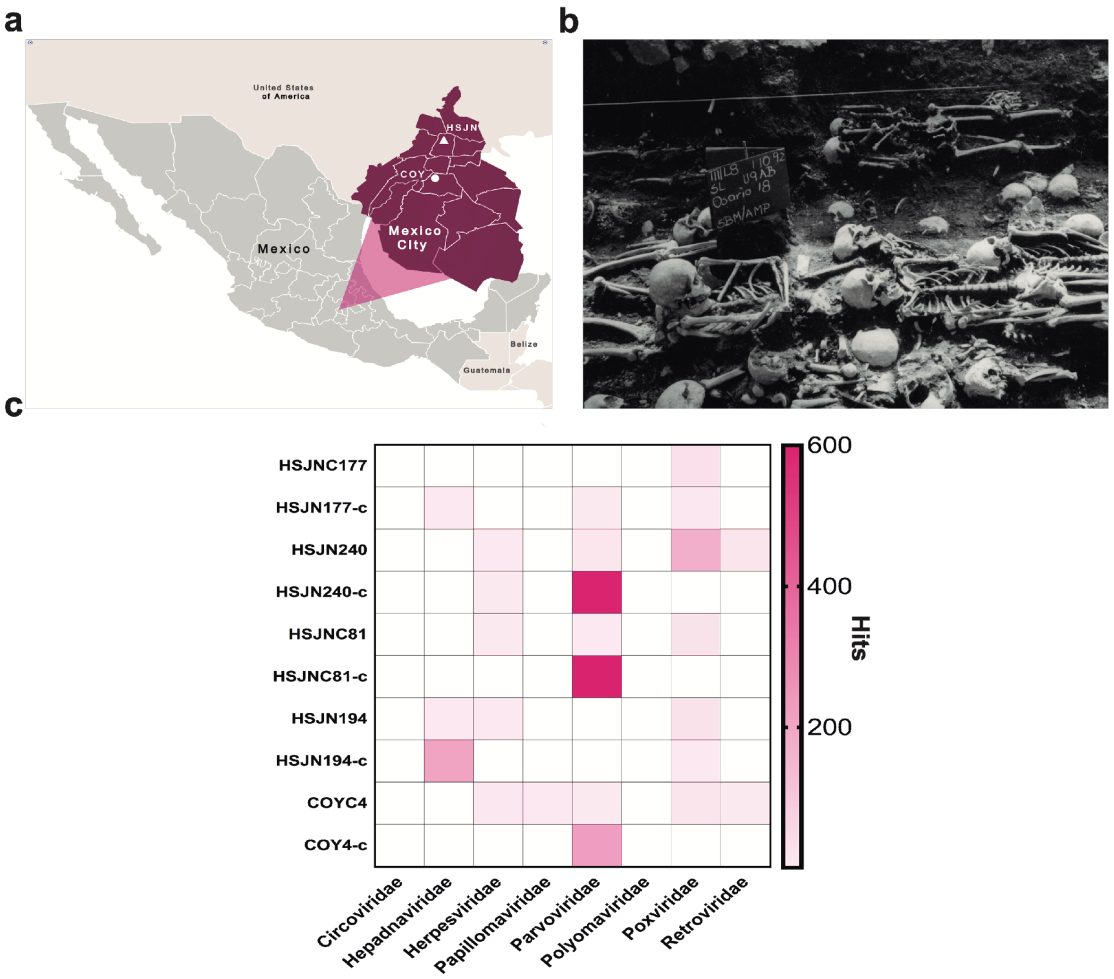

Figure 1. Metagenomic analysis of Colonial individuals reveal HBV-like and B19V-like hits. a, Location of the archeological sites used in this study, HSJN (19.431704, -99.141740) is shown as a triangle and COY (19.347079, -99.159017) as a circle, lines in pink map show current division of Mexico City. **b**, Several individuals discovered in massive burials archaeologically associated with the HSJN and colonial epidemics (photo courtesy of "Secretaria de Cultura INAH, SINAFO, Fototeca DSA"). **c**, Metagenomic analysis performed with MALT 0.4.0 based on the Viral NCBI RefSeq. Viral abundancies were compared and normalized automatically in MEGAN between shotgun (samplename) and capture (samplename-c) NGS data. Only targeted viral families are shown for HBV or B19V positive samples (all samples analysis shown in Supplementary Figure 3), a capture negative control (HSJN177) is shown as well that presented one B19V-like hit and three HBV-like hits after capture.

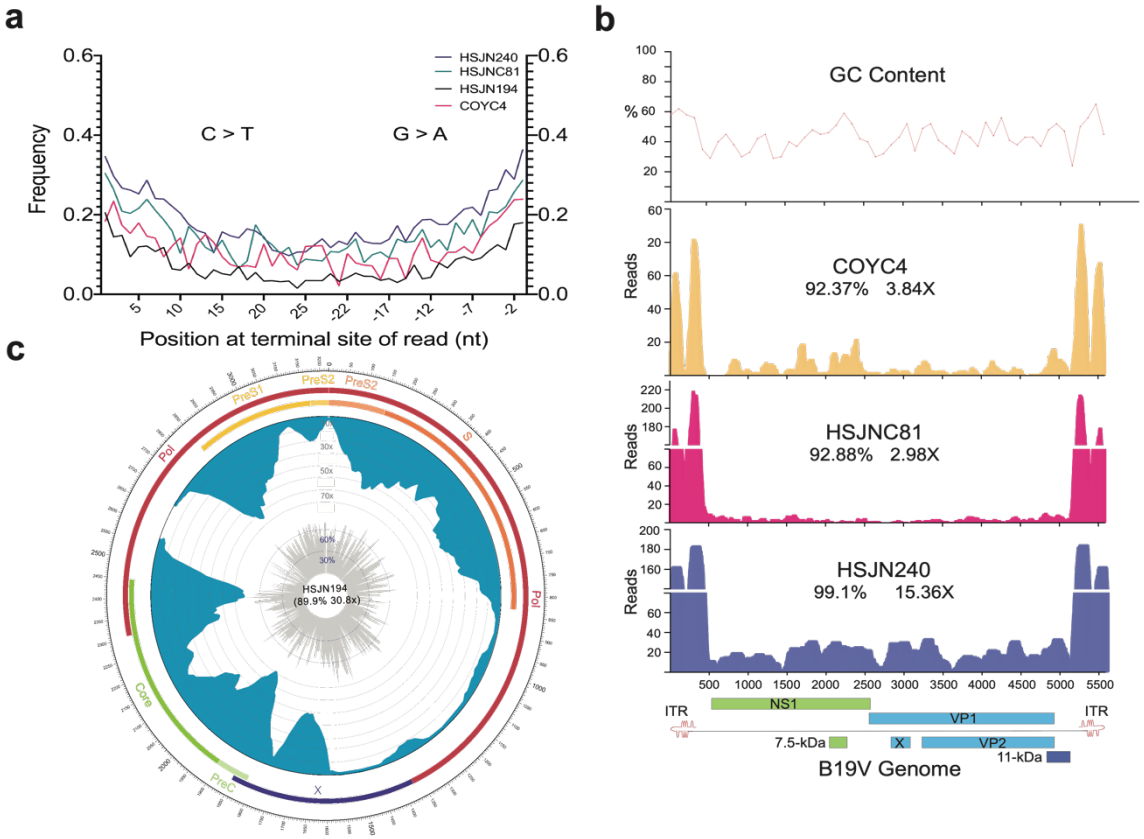

Figure 2. Ancient B19V and HBV ancient genomes. a, Superimposed damage patterns of ancient HBV (HSJN194) and B19V (HSJNC81, HSJN240, COYC4), X axis shows the position (nt) on the 5' (left) and 3' (right) end of the read, Y axis shows the damage frequency (raw damage patterns are shown on Supplementary Figure 4). **b**, B19V ssDNA linear genome, X axis shows position (nt) based on the reference genome (AB550331), and Y axis shows depth (as number of reads), GC content is shown as a percentage of each 100 bp windows, CDS coverage and average depth are shown under each individual ID. Schematic of the B19V genome is shown at the bottom. Highly covered regions correspond to dsDNA ITRs shown as a red hairpin. **c**, HBV circular genome, outer numbers show position (nt) based on reference genome (GQ331046), outer bars show genes with names, blue bars represent coverage and gray bars shows GC content each 10 bp windows. Coverage and average depth are shown in the center. Low covered region between S and X overlaps with ssDNA region.

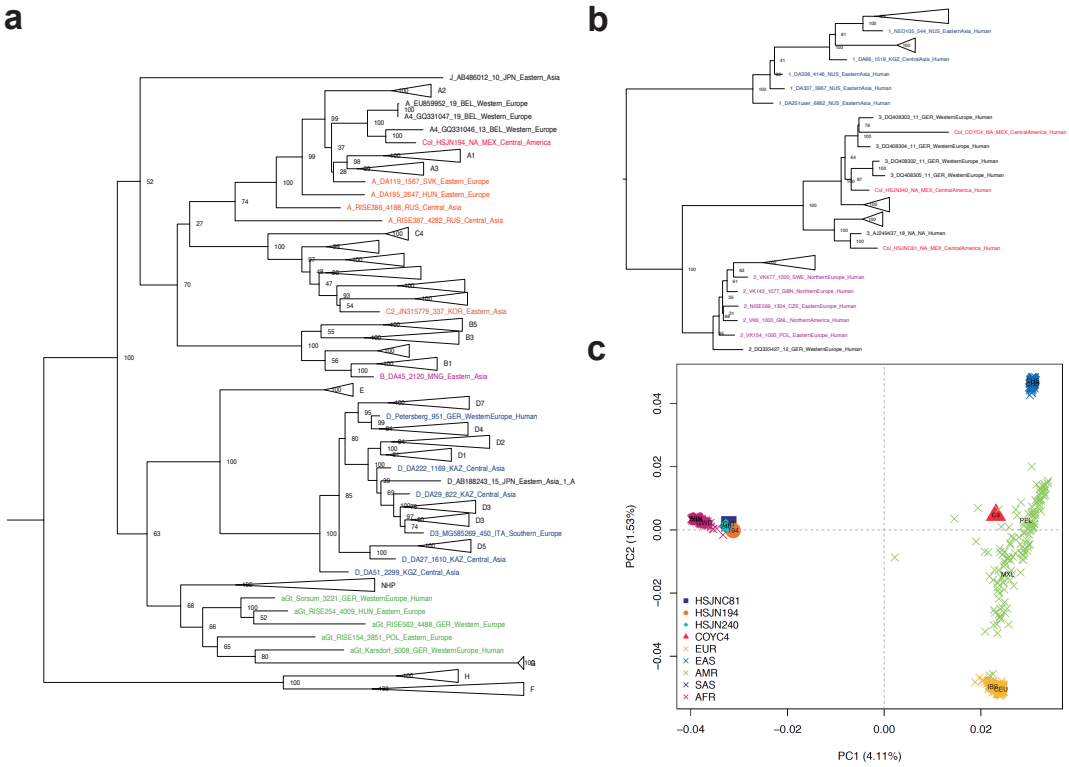

Figure 3. Viral Colonial genomes and their human hosts cluster with modern African genetic diversity. Maximum likelihood tree performed on RAxML 8.2.10 (1000 bootstraps) with a midpoint root based on the HBV whole genome (**a**) and B19V CDS (**b**), previously reported ancient genomes are colored, Colonial samples are shown in red. Sequences are named as follows: genotype_ID_sampling.year_country.of.origin_area.of.origin_host. HBV genotype nomenclature is based on letters, while for B19V is on numbers. **c**, PCA showing genetic affinities of ancient human hosts compared to 1000 Genomes Project reference panel, “x” show individuals from the reference panel while other shapes show human hosts from which ancient HBV (HSJN194) and B19V (HSJNC81, HSJN240, COYC4) were recovered.