Origin and Health Status of First-Generation Africans from Early Colonial Mexico

Highlights

- Genomic and isotopes data suggest an African origin for the three individuals
- One ~14X *Treponema pallidum* sub. *pertenue* genome was recovered
- One ~1,500X hepatitis B virus genome was recovered
- Both pathogen genomes cluster together with present day pathogens from Africa

Authors

Rodrigo Barquera,
Thiseas C. Lamnidis,
Aditya Kumar Lankapalli, ..., Denise Kühnert,
Lourdes Márquez-Morfin,
Johannes Krause

Correspondence

kuehnert@shh.mpg.de (D.K.),
rlmorfin@gmail.com (L.M.-M.),
krause@shh.mpg.de (J.K.)

In Brief

Barquera et al. analyze individuals from a colonial period burial from Mexico. Using an interdisciplinary approach, they reconstruct the genetic ancestry, origins, and health status of three enslaved Africans. Genomes of pathogens recovered from them provide insight on infectious diseases brought to the Americas by the transatlantic slave trade.
Origin and Health Status of First-Generation Africans from Early Colonial Mexico

Rodrigo Barquera,1,2,13 Thiseas C. Lamnidis,1,13 Aditya Kumar Lankapalli,1,13 Arthur Kocher,3,13 Diana I. Hernández-Zaragoza,2,4,13 Elizabeth A. Nelson,1,13 Adriana C. Zamora-Herrera,6 Patxi Ramallo,7,8 Natalia Bernal-Felipe,7 Alexander Immel,1,12 Kirsten Bos,1 Victor Acuña-Alonzo,7 Chiara Barbieri,11,12 Patrick Roberts,7 Alexander Herbig,1 Denise Kühnert,3,* Lourdes Márquez-Morfin,6,* and Johannes Krause1,14,*

1Department of Archaeogenetics (DAG), Max-Planck Institute for the Science of Human History (MPI-SHH), Kahlaische Str. 10, 07745 Jena, Germany
2Molecular Genetics Laboratory, Escuela Nacional de Antropología e Historia (ENAH), Periférico Sur y Zapote s/n. Col. Isidro Fabela, Tlalpan, 14030 Mexico City, Mexico
3Transmission, Infection, Diversification & Evolution Group (TIDE), Max-Planck Institute for the Science of Human History (MPI-SHH), Kahlaische Str. 10, 07745 Jena, Germany
4Immunogenetics Unit, Técnicas Genéticas Aplicadas a la Clínica (TGAC), Calz. del Hueso 714, Coapa, Los Sauces, Coyoacán, 04940 Mexico City, CDMX, Mexico
5Institute for the Archaeological Sciences, University of Tübingen, Geschwister-Scholl-Platz, 72074 Tübingen, Germany
6Osteology Laboratory, Post Graduate Studies Division, Escuela Nacional de Antropología e Historia (ENAH), Periférico Sur y Zapote s/n. Col. Isidro Fabela, Tlalpan, 14030 Mexico City, Mexico
7Department of Archaeology (DA), Max-Planck Institute for the Science of Human History (MPI-SHH), Kahlaische Str. 10, 07745 Jena, Germany
8Faculty of Medicine and Nursing, University of the Basque Country (UPV/EHU), Arriola Pasealekua, 2, 20018 Donostia, Gipuzkoa, Spain
9Escuela Nacional de Antropología e Historia (ENAH), Periférico Sur y Zapote s/n. Col. Isidro Fabela, Tlalpan, 14030 Mexico City, Mexico
10Institute of Clinical Molecular Biology, Kiel University, Rosalind-Franklin-Straße 12, 24105 Kiel, Germany
11Department of Linguistic and Cultural Evolution (DLCE), Max-Planck Institute for the Science of Human History (MPI-SHH), Kahlaische Str. 10, 07745 Jena, Germany
12Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
13These authors contributed equally
14Lead Contact
*Correspondence: kuehnert@shh.mpg.de (D.K.), rlmorfin@gmail.com (L.M.-M.), krause@shh.mpg.de (J.K.)

SUMMARY

The forced relocation of several thousand Africans during Mexico’s historic period has so far been documented mostly through archival sources, which provide only sparse detail on their origins and lived experience. Here, we employ a bioarchaeological approach to explore the life history of three 16th century Africans from a mass burial at the San José de los Naturales Royal Hospital in Mexico City. Our approach draws together ancient genomic data, osteological analysis, strontium isotope data from tooth enamel, δ13C and δ15N isotope data from dentine, and ethnohistorical information to reveal unprecedented detail on their origins and health. Analyses of skeletal features, radiogenic isotopes, and genetic data from uniparental, genome-wide, and human leukocyte antigen (HLA) markers are consistent with a Sub-Saharan African origin for all three individuals. Complete genomes of Treponema pallidum sub. pertenue (causative agent of yaws) and hepatitis B virus (HBV) recovered from these individuals provide insight into their health as related to infectious disease. Phylogenetic analysis of both pathogens reveals their close relationship to strains circulating in current West African populations, lending support to their origins in this region. The further relationship between the treponemal genome retrieved and a treponemal genome previously typed in an individual from Colonial Mexico highlights the role of the transatlantic slave trade in the introduction and dissemination of pathogens into the New World. Putting together all lines of evidence, we were able to create a biological portrait of three individuals whose life stories have long been silenced by disreputable historical events.

INTRODUCTION

Almost 500 years ago, in 1518, Charles I of Spain issued an authorization to transport the first African slaves into the Viceroyalty of New Spain (which at its zenith comprised the entirety of present day Mexico, the Caribbean, parts of the United States and Canada, and all Central America except for Panama) [1]. 5 centuries later, the ancestry of those hundreds of thousands of forcefully abducted people has formed an integral part of the genetic heritage carried by a large number of people in Mexico and a highly visible part of its national cultural heritage. Slavery was the primary mechanism of immigration of Africans to Mexico
during the first half of the colonial period, which constitutes one of the most significant forced oppressive migrations in human history [2]. The increased demand for enslaved manual laborers, together with the establishment of the first European settlements in what is euro-centrically referred to as the “New World,” caused the growth and consolidation of the transatlantic slave trade. Despite uncertainties around absolute numbers, an estimated 10.6 to 19.4 million Africans had been forcibly deported from their homelands until slavery was finally abolished in most parts of the Americas in the 1860s. To add to these staggering figures, high mortality rates during their voyages meant that only 9.6 to 15.5 million actually arrived in the Americas [3–6]. These data include only the entries through the official ports, which ignores individuals that were smuggled, for whom there is no reliable record. In the case of the Viceroyalty of New Spain, between 130,000 and 150,000 Africans had arrived by 1779, the year in which the importation of slaves into New Spain was banned [7, 8]. Of these, approximately 70,000 entered between the years 1600 and 1640 [3]. The sudden increase in demand was in part due to a reduction in the indigenous labor force that resulted from both casualties in the many conflicts during the European conquest and from diseases (among them, smallpox, measles, and typhoid fever) [9–11] that devastated nearly 90% of the native population [10]. Creoles, Africans, mulattoes, and other African-descended groups were thought to have higher resistance to these diseases compared to Indigenous Americans and Europeans [7, 12–14] making them desirable assets. Further to this, Las Leyes Nuevas (The New Laws) of 1542 prohibited the use of Native American labor as slaves in New Spain [15].

The African genetic diversity has been studied in other regions of the Americas by means of genome-wide analysis of both ancient [16–18] and present-day [19, 20] populations. To better understand the lives and health status of first-generation African slaves in Colonial Mexico, historical, archaeological, isotopic, and genetic evidence needs to be considered. Here, we use a bioarchaeological approach that marries ancient human genetic data with osteological analyses, stable carbon and nitrogen isotope analysis, strontium isotope analysis, ethnohistorical information, as well as a molecular screening for potential pathogens to comment on the origins, health status, and life history of three putatively enslaved Africans from Mexico City from a mass burial at the Hospital Real de San José de los Naturales (San José de los Naturales Royal Hospital [SJN]) [10, 21]. Documentary evidence indicates this hospital serviced exclusively los naturales (i.e., the indigenous population) [22]; however, different sources and studies point to non-indigenous peoples as being treated and buried there [23–27]. Particular traits, such as culturally associated dental modification patterns (Figure 1; see also [6, 28]), suggested a possible African origin of three individuals. Indeed, our results confirm and characterize their African ancestry. We got important insights into these individuals’ health status and retrieved genomic material of two pathogens phylogenetically linked to strains from Sub-Saharan Africa. Based on isotopic evidence, we can suggest a non-Mexican origin for these individuals, as they are presumably first-generation enslaved Africans who died very early during the colonial period in Mexico.

RESULTS

Samples were processed according to protocols designed for ancient DNA work in dedicated facilities. We were able to extract DNA from tooth dentine powder for all three individuals, which was used for building genomic libraries and producing ~5 million reads for each of them. Pre-processed sequencing data were used for pathogen screening, and PCR reamplified, reconditioned libraries were enriched for human-specific genomic regions or genomes of pathogens. After sequencing the enriched genomic libraries, quality control (QC), population genetics, and pathogen genome reconstruction analyses were carried out.

Ancient DNA Authentication

We observed short average fragment length (~55 bp) and the characteristic increase in cytosine (C) to thymine (T) transitions toward the termini of the DNA fragments that support the presence of authentic ancient DNA (aDNA) in the dataset. Pre-processed sequences were mapped to the human genome assembly GRCh37 (hg19) [29]. For our uracil DNA glycosylase (UDG)-treated libraries, the rate of deamination damage in the first base on the 5’ and 3’ termini were ~7% and 6%, respectively. After enriching for 1,237,207 single-nucleotide polymorphisms (SNPs), we decreased the number of SNPs to place our coverage in the range of that we generated for the ~600,000 SNP positions present in the Human Origins dataset [30], where ~400,000 SNPs for the three enriched libraries
were obtained. Coverage comparisons of X and Y chromosomes SNPs [31] from the 1,240K capture set assigned all three individuals as genetically male. To estimate levels of nuclear contamination in our samples, we used ANGS2 to measure the rate of heterozygosity of polymorphic sites on the X chromosome [32]. Further to this, a reconstructed mtDNA consensus sequence was used to estimate mtDNA contamination levels using schmutzi [33]. Both methods yielded comparable contamination estimates below 2% for all three individuals (Table 1). The pairwise mismatch rates do not support genetic kinship among the three individuals (Table 2).

Characterization of African Ancestry
The carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) measurements of dentine (Table S2) provided values between $-19.9^{\circ}$ and $-18.8^{\circ}$ ($\delta^{13}C$) and from $13.2^{\circ}$ to $10.8^{\circ}$ ($\delta^{15}N$). Broadly speaking, these values show that the dietary protein of the three individuals was composed mainly of terrestrial proteins but with a significant portion of C4 plants and/or marine proteins, consistent with documented diets either in the arid interior or the western coast of West Africa [34]. Unfortunately, it is difficult to distinguish the two kinds of diet due to an absence of associated faunal remains with the individuals that would provide a local ecological “base-line.” A comparison between individuals shows higher nitrogen values ($\delta^{15}N = 13.2^{\circ}$) for individual SJN001 relative to the other two individuals ($\delta^{15}N = 10.8^{\circ}$ and $10.5^{\circ}$, for SJN002 and SJN003, respectively). Without any additional contextual information or isotopic data from fauna, it remains challenging to determine the exact nature of their individual diets. Distinguishing these sources has been a consistent problem in African archaeology, where significant C4 ecosystems coexist in proximity to the coast [35, 36].

Osteological analyses of the three individuals reveal evidence suggesting a life experience of conflict and hardship. Individual ML8 SL 150 (SJN001) was found with five buck shots and two healing needles (used in traditional medicine) in the thoracic cavity, as well as gunshot wounds. Both SJN001 and SJN003 (ML8 SLU9B 296) presented porotic hyperostosis and cribra orbitalis, two pathological changes associated with a skeletal response to nutritionally inadequate diets, anemia, parasitic infectious diseases, and blood loss [37–40]. Individual ML8 San José 214 (SJN002) displayed several skeletal changes associated with intense labor and heavy manual activity, including enthesopathies on the clavicle and scapulae as well as osteophytic lipping on the joint surfaces with some additional joint contour deformation at the sternoclavicular joint of the clavicle. Additionally, he suffered from a poorly aligned complete fracture in the right fibula and tibia, resulting in associated joint changes of the knee, including osteochondritis dissecans of the distal femoral surface with joint contour deformation and associated osteophytic lipping of the articular surface margin. Furthermore, this individual displayed osteoarthrosis of the lumbar vertebral in addition to signs of deficient oral health and cut marks on the frontal bone.

Skeletal analysis of cranial nonmetric traits (please refer to the STAR Methods section for further information) showed a strong African affinity for all three individuals (macromorphosophsic characteristics: 92.23% for SJN001; 95.24% for SJN002; 91.95% for SJN003). The dental modification patterns (i.e., decoration of the tooth by mechanical modification) are consistent with cultural practices previously reported for African enslaved individuals from different contexts [41, 42] but also observed in groups living in Western Africa today, most notably in the coastal area and nearby regions of equatorial Guinea, Gabon, Cameroon, the Republic of Congo, and the Democratic Republic of the Congo [43]. Uniparental haplogroups and human leucocyte antigen (HLA) haplotypes are representative of Sub-Saharan populations (Table 1). The mitochondrial lineages of the SJN individuals were compared to a set of published mitochondrial genomes from the same three haplogroups (Figure S1). The three mtDNA haplogroups found show their highest frequencies in Western and Central Africa [44], with haplogroup L1b (present in SJN001) concentrated in Western-Central Africa, particularly along the coastal areas [44, 45]; L3d (present in SJN002) being prevalent in western Sub-Saharan Africa [44, 46]; and L3e1 (present in SJN003) being prevalent in the central areas of the continent [44, 47, 48]. Regarding the paternal lineages, all three individuals carried the Y chromosome E3b1a (E-M2) lineage (but not the same haplotypes; Table 1), which is highly prevalent in modern Sub-Saharan populations (particularly in Western Africa) and equatorial Africans and is also the most common lineage among African Americans [49–53]. In order to get a glimpse at the immunogenetic diversity present in these individuals, we carried out immune capture to obtain information on the HLA allelic variation. HLA diversity is helpful in determining the susceptibility and resistance to specific diseases, both infectious and autoimmune. HLA typing poses technical difficulties, and it can be challenging to call HLA alleles from ancient DNA data. For that reason, we enriched the HLA region with an in-solution capture approach (please refer to the STAR Methods section for further detail), yielding 10x–100x coverage for the HLA region. Allele assignation was carried out using Optitype [34], a software specifically developed to deal with short reads coming from ancient DNA data. The HLA haplotypes described here have so far been reported only in African or African-descent populations, like Mozambique, Kenya, South Africa, and African Americans [55–59]. In the case of both HLA haplotypes of SJN01 and haplotype HLA-A*74:01~B*49:01~C*07:01~DRB1*04:05~DRB4*01:01~DQA1*03:02~DQB1*03:02~DPB1*02:01~DPB1*13:01, present in SJN002 (Table 1), none of these haplotypes were yet reported in any African populations, but only in mixed ancestry individuals of at least partial African descent [59, 60]. However, the observed HLA-A~B~C blocks (corresponding to the HLA class I genomic region, functionally devoted to present endogenous peptides to the immune system) have been reported in Sub-Saharan African populations, such as Luo from Kenya [57] and in Rwandan women [61].

Genomic data from the three individuals were merged to 593,124 autosomal SNPs of the Human Origins (HO) dataset [30] for downstream population genetics analyses. To further assess the genetic relationship of these African individuals to present-day populations from Africa, we carried out principal-component analysis (PCA). We first projected the data from the three individuals on PCs calculated on variation from 371 worldwide populations (Figure S2A) and then on variation from 534 individuals from 51 populations from Northern and Sub-Saharan Africa. We found that the SJN individuals cluster together with Sub-Saharan African populations like Mandenka, Mende,
Table 1. Basic QC and Genetic Markers of the Analyzed African Individuals from the Royal Hospital of San José de los Naturales

<table>
<thead>
<tr>
<th>Library</th>
<th>Sex</th>
<th>Modeled $^{14}$C Dating</th>
<th>Damage 1st Base 3'</th>
<th>Damage 1st Base 5'</th>
<th>Called SNPs on HO (Post-merging)</th>
<th>Contamination Estimates (Post-merging)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SJN001</td>
<td>M</td>
<td>AD 1453–1626</td>
<td>0.0678</td>
<td>0.0751</td>
<td>427,155</td>
<td>0.02 ± 0.01 0.54% ± 0.002%</td>
</tr>
<tr>
<td>SJN001</td>
<td></td>
<td>0.0309</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SJN002</td>
<td>M</td>
<td>AD 1450–1620</td>
<td>0.0710</td>
<td>0.0803</td>
<td>397,342</td>
<td>0.02 ± 0.01 0.44% ± 0.002%</td>
</tr>
<tr>
<td>SJN002</td>
<td></td>
<td>0.0776</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SJN003</td>
<td>M</td>
<td>AD 1436–1472*</td>
<td>0.0474</td>
<td>0.0730</td>
<td>446,940</td>
<td>0.02 ± 0.01 0.28% ± 0.002%</td>
</tr>
<tr>
<td>SJN003</td>
<td></td>
<td>0.0443</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual</th>
<th>mtDNA Haplogroup</th>
<th>Y-Chr Haplogroup</th>
<th>HLA A</th>
<th>HLA B</th>
<th>HLA C</th>
<th>HLA-DRB1</th>
<th>HLA-DRB3/4/5</th>
<th>HLA-DQA1</th>
<th>HLA-DQB1</th>
<th>HLA-DPA1</th>
<th>HLA-DPB1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SJN001</td>
<td>L1b2a</td>
<td>E1b1a1a1c1b</td>
<td>A*03:01</td>
<td>B*44:10</td>
<td>C*04:01</td>
<td>DRB1*07:01</td>
<td>DRB4*01:01</td>
<td>DQA1*02:01</td>
<td>DQB1*02:01</td>
<td>DPA1*02:01</td>
<td>DPB1*01:01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-M263.2</td>
<td>A*66:02</td>
<td>B*44:03</td>
<td>C*03:03</td>
<td>DRB1*16:02</td>
<td>DRB5*02:02</td>
<td>DQA1*01:02</td>
<td>DQB1*05:02</td>
<td>DPA1*02:02</td>
<td>DPB1*01:01</td>
</tr>
<tr>
<td>SJN002</td>
<td>L3d1a1a</td>
<td>E1b1a1a1d1</td>
<td>A*30:01</td>
<td>B*42:01</td>
<td>C*17:01</td>
<td>DRB1*03:02</td>
<td>DRB3*01:01</td>
<td>DQA1*04:01</td>
<td>DQB1*04:02</td>
<td>DPA1*02:01</td>
<td>DPB1*15:01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E-P278.1</td>
<td>A*74:01</td>
<td>B*49:01</td>
<td>C*07:01</td>
<td>DRB1*04:05</td>
<td>DRB4*01:01</td>
<td>DQA1*03:02</td>
<td>DQB1*03:02</td>
<td>DPA1*02:01</td>
<td>DPB1*13:01</td>
</tr>
<tr>
<td>SJN003</td>
<td>L3e1a1a</td>
<td>E1b1a1c1a1c1</td>
<td>A*30:01</td>
<td>B*42:01</td>
<td>C*17:01</td>
<td>DRB1*07:01</td>
<td>DRB4*01:01</td>
<td>DQA1*03:01</td>
<td>DQB1*02:02</td>
<td>DPA1*01:03</td>
<td>DPB1*04:01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(E-CTS8030)</td>
<td>A*66:02</td>
<td>B*58:01</td>
<td>C*07:01</td>
<td>DRB1*15:03</td>
<td>DRB5*01:01</td>
<td>DQA1*01:02</td>
<td>DQB1*06:02</td>
<td>DPA1*01:03</td>
<td>DPB1*02:01</td>
</tr>
</tbody>
</table>

*These years refer to pre-contact dates; however, a possible reservoir effect due to marine diet cannot be ruled out as discussed in the Results section.
Table 2. Pairwise Mismatch Ratio Analysis

<table>
<thead>
<tr>
<th>Individual 1</th>
<th>Individual 2</th>
<th>SNPs Tested</th>
<th>Mismatched SNPs</th>
<th>Mismatch Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>SJN001</td>
<td>SJN002</td>
<td>594,706</td>
<td>160,514</td>
<td>0.2699</td>
</tr>
<tr>
<td>SJN001</td>
<td>SJN003</td>
<td>661,252</td>
<td>178,194</td>
<td>0.2695</td>
</tr>
<tr>
<td>SJN002</td>
<td>SJN003</td>
<td>613,943</td>
<td>164,424</td>
<td>0.2678</td>
</tr>
</tbody>
</table>

Wambo, Ovambo, and Damara. When PC1 and PC2 are considered (Figure S2B), the three individuals are projected in the Bantu-speaking population cluster. When plotting PC1 versus PC3 (Figure 2A), the three individuals from SJN appear slightly separated from each other. SJN002 and SJN003 practically overlap with each other and are projected in close proximity to the Wambo population of Southern Africa. SJN001 is projected close to the other two individuals from SJN but falls even closer to the Mende population of Western Africa. PC3 separates hunter-gatherer populations like the Mbuti, Biaka, and Hadza from all other African populations. Admixture estimates (lowest cross validation error at K = 10) show that the SJN individuals look very similar, though not identical, in terms of genetic ancestry composition (Figure 2C). Predominantly, ancestry components that are shared with Sub-Saharan populations, such as Yoruba (non-Bantu speakers), Mandenka (non-Atlantic Niger-Congo linguistic family), and Bantu speakers from Kenya and South Africa, were found in all three individuals. An additional component shared with Mbuti and/or Biaka was also evident for SJN002 and SJN003. This extra component might affect the projection of these individuals in the PCA, especially when PC3 is considered. No ancestry related to other continental sources (especially Native Americans) is visible in the ADMIXTURE results (Figure 2C). Altogether, these data support a heterogeneous genetic makeup for the SJN individuals within Sub-Saharan Africa, with clear links to Central-Western Africa and possible links with populations who live in Southern Africa today.

**Evidence of the Place of Birth from Strontium Isotope Analysis**

Radiocarbon dating suggests that these individuals died shortly after the beginning of the colonial period in Mexico City (Table 1). To further assess the origin of these ancient individuals, we carried out strontium analysis on the molars of each individual, which forms early during childhood and thus is a good proxy for place of birth. For all three individuals, the strontium ratio values are above 0.7089 (SJN001: 0.71078 ± 0.00001; SJN002: 0.72045 ± 0.00001; SJN003: 0.72026 ± 0.00001), which is the highest value recorded for Mexico in the northern part of the Yucatán Peninsula (Tizimin and Dzibilchaltun) [63–65] and even higher than those found for central Mexico (Teotihuacan [0.7049], Xico [0.7045], Chapantongo [0.7052], and Mexico City [0.7062–0.7064]) [63, 66]. In fact, these values are consistent with a West African origin (0.70603–0.74143; Figures 2B and S4; see also [63, 67]).

**Paleomicrobiological Findings**

The three individuals were buried in a mass grave in the grounds located just outside of the hospital. Their mass grave contained skeletal remains of several individuals disposed in layers, which is consistent with burials made during epidemics, in which dead bodies rapidly outnumber the availability of single graves. We therefore decided to screen these individuals for potential pathogenic agents using a bioinformatic approach [68] to screen and filter reads from the genomic libraries, and this way, we obtained reads mapping to bacteria that could either be part of the oral microbiota (like Tannerella forsythia, Streptococcus mutans, and S. gordonii) or implicated in taphonomic processes (Cloestridium tetani). Apart from environmental bacteria, 51 reads were mapping specifically to Treponema pallidum genomes in individual SJN003, which were marked as potentially positive for treponemal infection. We also found evidence of hepatitis B virus (HBV) DNA in individual SJN001 (library SJN001.B0102), with eight sequencing reads mapping specifically to HBV.

The positive libraries were enriched for HBV (SJN001) and Treponema pallidum (SJN003), respectively, using a modified in-solution capture strategy followed by paired-end sequencing (Supplemental Information for further details). After capture, a total of 50,493 (deduplicated) reads mapping to the HBV reference genome were recovered from SJN001. Those exhibited typical patterns of ancient DNA damage, and the complete HBV circular
genome could be recovered, with a mean coverage of 1,567-fold. These results confirmed that SJN001 was infected with HBV. The ancient HBV genome exhibited a 6-nt insert at the carboxyl end of the core gene, which is characteristic of the HBV A genotype [69]. The phylogenetic analysis confirmed that SJN001 HBV belonged to genotype A (Figure 4) and further showed that it was included in quasi-subgenotype A3 (QS-A3) [70]. Today, QS-A3 is typically found in West Africa. More specifically, SJN001 HBV grouped with strains isolated from Gambia and Guinea.

From SJN003, we reconstructed a *T. pallidum* genome at a mean coverage of 14-fold with 96.66% of the reference genome (Nichols strain) covered at 5-fold average coverage. After aligning the SJN003 genome to previously reported ancient and modern *Treponema pallidum* genomes [71], we produced a maximum likelihood tree (Figure 5) that showed the SJN003 genome to cluster within *Treponema pallidum* subsp. pertenue (TPE), the causative agent of yaws. Surprisingly, we found that the SJN003 strain is most closely related to the ancient TPE genome 133 previously isolated from human remains in Colonial Mexico [71]. Both genomes are phylogenetically associated with strains isolated from patients in Ghana.

**DISCUSSION**

Our combination of genetic analyses, isotope data, ethnohistorical information, and osteobiographies led to the construction of unique life portraits for three individuals from a colonial hospital cemetery in Mexico City. We opted to treat these three individuals independently as opposed to predetermining our analysis on the assumption that they represent a singular cultural tradition and ethnic history; this approach was used based on the evidence of forced social intermixing as a means of social control and disintegration among the enslaved population [2, 3, 5, 72]. Osteological assessment, dental modification patterns, and their genetic ancestry are all consistent with a Western or Southern African origin for the three individuals. This finding is further supported by strontium isotope ratios that suggest a non-local origin outside Mexico. The lives of physical hardships and traumas revealed by their osteobiographies, coupled with the fact that most
Sub-Saharan Africans in the Americas from the early colonial period (consistent with SJN individuals’ radiocarbon dates) were taken there by force, make the natural conclusion that these individuals were actually enslaved Africans, even in absence of concrete archaeological evidence (such as the individual from the ex-Temple of Corpus Christi who was found in shackles) [73].

The genomic analyses suggest that all three individuals show unadmixed African ancestry that is similar to that of Western and Southern African populations. Combining those results with the strontium isotope ratios evidence allows us to conclude that all three individuals were born outside Mexico and likely originate from Western Africa. To the best of our knowledge, they are the earliest genetically identified first-generation Africans in the Americas. mtDNA lineages found in the individuals from SJN (L1 and L3) have been found, although not highly represented, among modern day Mexicans with mixed ancestry [74, 75]. The African HLA haplotypes found in the SJN individuals, or parts of their haplotypes, can be found throughout modern Mexico in frequencies ranging from 0.01% to 2.27% [76]. Alleles HLA-DRB1*07 and HLA-DQB1*02 have been previously found associated with decreased antibody response against HBV and non-responsiveness to HBV vaccination [77, 78]. Interestingly, those two alleles are present in individual SJN001, which was found to be infected with HBV. Those alleles are also present in SJN003 but joined by protective alleles [77, 79], namely HLA-DQB1*06:02 and HLA-DPB1*02:01, which could have aided this individual not to become infected, if he would have been exposed to the virus in the first place. There is currently no reported association between HLA alleles and treponemal infections (apart from HLA-DRB1*14 in Asians) [80]. The fact that three of the six haplotypes were not previously reported in any African population highlights the necessity for exhaustive sampling of HLA haplotypes in African populations. D-statistics identify the closest genetic matches for the three individuals in central West Africa, Western Africa, and Southern African populations (Bantu-speaking groups). Our analyses point to a high genetic diversity in the populations who were the initial sources of the slave trade as suggested by historical records [3, 7, 81]. Furthermore, it is possible that the genetic diversity of living African populations was shaped by substantial migrations through the past four centuries, which displaced human groups and genetic ancestries. The analysis of ancient DNA from early African slave migrants in the Americas therefore provides an alternative opportunity to look at the past genetic makeup of the African continent, where the climatic conditions are not ideal for DNA preservation.

Based on radiocarbon dating, the three individuals lived during the early years of the colonial period in Mexico City (14C range: AD 1436–1626). The osteobiography revealed non-specific markers of physiological stress, evidence of occupational stress and healed severe wounds. They were found in the context of a mass grave (stacked in several layers suggestive of catastrophe deposits made during periods of epidemics) in proximity to a hospital that served only indigenous people [21], from a time period notorious for major epidemics in the region.

Based on our molecular screening, two of the three individuals showed molecular evidence of infectious diseases. HBV is widespread today and poses a significant global health burden [82]. It is currently classified into ten different genotypes that are heterogeneous distributed around the globe [69], and its long history infecting humans dates back at least to the Late Neolithic and Bronze Age [83, 84]. The biogeographic distribution of this virus is tightly linked to human history and migration, given its ease of transmission through contact with infected body fluids and the fact that it can cause a chronic infection [82]. Dispersion of HBV in the past through forced African migration was suggested by the presence of extant African HBV subgenotypes in regions...
where slavery was practiced [69, 85]. Given that individual SJN001 carried the HBV QS-A3 genotype typically found in West Africa today, coupled with the data that support his origins in this region, provides the first direct evidence of HBV disease movement with a potential dissemination through the transatlantic slave trade. Today, Haiti is the only location outside Africa where the HBV QS-A3 subgenotype is found [69, 85]. Thus, our results reveal that this HBV lineage seems not to have established itself in Mexico.

In contrast, the variety of yaws identified in individual SJN003 shares a common origin with the strain identified in a 17th-century individual of Colonial Mexico City carrying a European mitochondrial haplogroup (H1c+152) [71]. The two strains also cluster with yaws subtypes identified in extant West African groups (although the nodes for the relevant sections of the tree have low statistical support; Figure 5). The osteopathological findings on SJN003 (i.e., cribra orbitalia and hyperostosis and a mild lesion in the frontal bone; please refer to the STAR Methods section for further details) are consistent with chronic infectious diseases and have been reported in suspected cases of individuals affected by treponemal diseases. The polymorphic nature of the lesions left on the bone by treponemal bacteria invites one to pay attention to any potential treponematosis case [71]. These data suggest the establishment of at least one disease of African origin in the local population. As yaws is a highly contagious skin infection that is associated with poor hygiene, it is unsurprising that skeletal signs of treponemal diseases have been reported in other African enslaved individuals from similar historic periods in other contexts in Europe [42]. Our findings add to the discussion of the dynamics of treponemal diseases in the colonial period of Mexico, where they were a prominent yet not completely understood health issue [86, 87]. Our current data, however, shed little light on the reason for their deposition in a presumed epidemic burial ground.

Here, we present the application of a comprehensive set of methods and techniques to display the first in-depth characterization of first-generation African slaves in Colonial Mexico, a historically oppressed group. This was accomplished through reconstruction of the origins and life histories of three putatively enslaved Africans from the beginning of the colonial period in central Mexico. We were able to describe their biological identity and explore elements of their health status by recovering high-quality ancient DNA from both the individuals and their identifiable pathogens. Our work adds important information on the origin, health status, and life histories of the first generation of African slaves that were forcefully relocated and transported to the New World during one of the most horrific cases of callous disregard for human life and violation of human rights during the colonial period. By investigating the origin and disease experience of these individuals through molecular methods and evaluating the skeleton for signs of life experience and cultural affinity, we illuminate, in some measure, the identity, culture, and life of these people whose history has largely been lost. Furthermore, by exploring the African diversity that first came into Mexico and

---

**Figure 4. Phylogenetic Tree of Hepatitis B Virus Based on Full Genomes, Including SJN001, Estimated by Maximum Likelihood**

Branch lengths represent numbers of substitutions. Bootstrap support values are written at the nodes. In the left panel, inter-genotype relationships are shown and the main continental distribution of each genotype is indicated [69]. The relationships within genotype A and the position of SJN001 are shown in the right panel.
Latin America, we can draw a sharper picture of the biological legacy and cultural roots of modern Latin Americans, explore the extent of this diversity through different disciplines, and also acknowledge how African populations contributed to the biological diversity of mixed ancestry populations of the Americas.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Archaeological sites and sample description
- **METHOD DETAILS**
  - Osteological assessment
  - Stable isotope analysis, diets, and the reservoir effect
  - Ancient DNA sample processing and quality control
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - QC and data processing for human PopGen
  - Pathogen genomes assemblies

**SUPPLEMENTAL INFORMATION**

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2020.04.002.

**ACKNOWLEDGMENTS**

The authors wish to thank Franziska Aron, Cacilia Freund, Rita Radzевичуте, Raphaela Stahl, Marta Burri, Antje Wissgott, Guido Brandt, András Szolek, Stephen Clayton, Perla del Carmen Ruı́z Albarrán, Laura Curiel Giles, and Ber-sal del Carmen Villegas Camposeco for their valuable technical and analytical support. We are grateful to Jorge A. Gómez Valdés, Stephan Schiffels, Choongwon Jeong, Arturo Pacho Varela, Kathrin Nägele, Vanessa Villalba, Cosimo Posth, Maria Spyrou, Felix Key, and Malte Rivotrat for insightful comments and helpful discussions, which are part of the finished version of this manuscript. We acknowledge the Council of Archaeology, National Institute of Anthropology and History (INAH) (Mexico City, Mexico) for the permit granted to analyze the individuals from the San José de los Naturales collection (project: Identificación de los sitios de origen de los individuos con limado dental característico de grupos africanos el Hospital Real de San José de los Naturales a través de análisis genético molecular; official notice number: 401.B(4)19.2015/36/1026; final report approval: 401.1S.3-2020/0112). C.B

**Figure 5. Maximum Likelihood Tree for Ancient Treponemal Genomes from SJN003 (T. pallidum sub. pertenue), 113 (T. pallidum sub. pertenue), and 94B SJN003 (T. pallidum sub. pallidum) and 30 Modern Treponema pallidum Genomes**

Orange, violet, and green tip labels represent T. pallidum subsp. pallidum (causative agent of syphilis), T. pallidum sub. pertenue (causative agent of yaws), and T. pallidum sub. endemicum (causative agent of bejel) strains, respectively. The scale represents the mean number of substitutions per site for the GTR+GAMMA substitution model. The strain found in individual SJN003 branches with T. pallidum sub. pertenue strains. Bootstrap support for each branch estimated from 1,000 replicates is provided. The branches with statistical support lower than 80 are presented in gray.
was supported by the University Research Priority Program of Evolution in Action of the University of Zurich. This work was financially supported by the Max Planck Society.

AUTHOR CONTRIBUTIONS


DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 17, 2020
Revised: March 3, 2020
Accepted: April 1, 2020
Published: April 30, 2020

REFERENCES


## STAR★METHODS

### KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biological Samples</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ancient individual</td>
<td>This study/ San José de los Naturales archaeological site</td>
<td>Sjin001/ ML8 SL 150</td>
</tr>
<tr>
<td>Ancient individual</td>
<td>This study/ San José de los Naturales archaeological site</td>
<td>Sjin002/ ML8 San José 214</td>
</tr>
<tr>
<td>Ancient individual</td>
<td>This study/ San José de los Naturales archaeological site</td>
<td>Sjin003/ ML8 SLU9B 296</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2x HI-RPM hybridization buffer</td>
<td>Agilent Technologies</td>
<td>Cat# 5190-0403</td>
</tr>
<tr>
<td>Herculase II Fusion DNA Polymerase</td>
<td>Agilent Technologies</td>
<td>Cat# 600679</td>
</tr>
<tr>
<td>Pfu Turbo Cx Hotstart DNA Polymerase</td>
<td>Agilent Technologies</td>
<td>Cat# 600412</td>
</tr>
<tr>
<td>D1000 ScreenTapes</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-5582</td>
</tr>
<tr>
<td>D1000 Reagents</td>
<td>Agilent Technologies</td>
<td>Cat# 5067-5583</td>
</tr>
<tr>
<td>0.5 M EDTA pH 8.0</td>
<td>BioExpress</td>
<td>Cat# E177</td>
</tr>
<tr>
<td>Sera-Mag Magnetic Speed-beads Carboxylate-Modified (1 mm, 3EDAC/PA5)</td>
<td>GE LifeScience</td>
<td>Cat# 651521050260</td>
</tr>
<tr>
<td>Bst DNA Polymerase2.0, large frag.</td>
<td>New England Biolabs</td>
<td>Cat# M0537</td>
</tr>
<tr>
<td>UGI</td>
<td>New England Biolabs</td>
<td>Cat# M0281</td>
</tr>
<tr>
<td>USER enzyme</td>
<td>New England Biolabs</td>
<td>Cat# M5505</td>
</tr>
<tr>
<td>PE buffer concentrate</td>
<td>Qiagen</td>
<td>Cat# 19065</td>
</tr>
<tr>
<td>1 M Tris-HCl pH 8.0</td>
<td>Sigma Aldrich</td>
<td>Cat# AM9856</td>
</tr>
<tr>
<td>1M NaOH</td>
<td>Sigma Aldrich</td>
<td>Cat# 71463</td>
</tr>
<tr>
<td>20% SDS</td>
<td>Sigma Aldrich</td>
<td>Cat# 5030</td>
</tr>
<tr>
<td>3M Sodium Acetate (pH 5.2)</td>
<td>Sigma Aldrich</td>
<td>Cat# S7899</td>
</tr>
<tr>
<td>5M NaCl</td>
<td>Sigma Aldrich</td>
<td>Cat# S5150</td>
</tr>
<tr>
<td>Ethanol</td>
<td>Sigma Aldrich</td>
<td>Cat# E7023</td>
</tr>
<tr>
<td>Guanidine hydrochloride</td>
<td>Sigma Aldrich</td>
<td>Cat# G3272</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>Sigma Aldrich</td>
<td>Cat# 650447</td>
</tr>
<tr>
<td>PEG-8000</td>
<td>Sigma Aldrich</td>
<td>Cat# 89510</td>
</tr>
<tr>
<td>Proteinase K</td>
<td>Sigma Aldrich</td>
<td>Cat# P6556</td>
</tr>
<tr>
<td>Tween-20</td>
<td>Sigma Aldrich</td>
<td>Cat# P9416</td>
</tr>
<tr>
<td>Water</td>
<td>Sigma Aldrich</td>
<td>Cat# W4502</td>
</tr>
<tr>
<td>10x Buffer Tango</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# BY5</td>
</tr>
<tr>
<td>50x Denhardt’s solution</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 750018</td>
</tr>
<tr>
<td>ATP</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# R0441</td>
</tr>
<tr>
<td>dNTP Mix</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# R1121</td>
</tr>
<tr>
<td>Dynabeads MyOne Streptavidin T1</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 65602</td>
</tr>
<tr>
<td>DyNaMo HS SYBR Green qPCR Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# F410L</td>
</tr>
<tr>
<td>GeneAmp 10x PCR Gold Buffer</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 4379874</td>
</tr>
<tr>
<td>Human Cot-I DNA</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 15279011</td>
</tr>
<tr>
<td>Salmon sperm DNA</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# 15632-011</td>
</tr>
<tr>
<td>SSC Buffer (20x)</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# AM9770</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# EL0011</td>
</tr>
<tr>
<td>T4 DNA polymerase</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# EP0062</td>
</tr>
<tr>
<td>T4 Polynucleotide Kinase</td>
<td>Thermo Fisher Scientific</td>
<td>Cat# EK0032</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCES AVAILABILITY

Lead Contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Johannes Krause (krause@shh.mpg.de).

Materials Availability
This study did not generate new unique reagents.

Data and Code Availability
Data are available at the European Nucleotide Archive (ENA) under study accession number: PRJEB37490 (including the HBV genome recovered from individual SJN001 and the Treponema pallidum sub. pertenue genome recovered from individual SJN003).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The San José de los Naturales Royal Hospital (originally Hospital Real de Sanct Joseph de los Naturales) was founded between 1529 and 1531 [97–99] and dedicated to care exclusively for the indigenous population of the Viceroyalty of the New Spain. The establishment of this hospital was largely motivated by the need for medical facilities to aid the victims of the smallpox outbreaks that occurred early during the Colonial Period [100]. The vast majority of the patients were Nahuatl and Otomi speakers, which could tell us that most patients were from the central region of Mexico, including the valley of Morelos to the south and the Basin of the Balsas River to the west, as well as the boroughs of La Candelaria, Santo Tomás, San Pablo, San Antonio Abad and the towns of Jamaica and...
Ixtacalco in Mexico City [22, 24]. The mortuary patterns of the archaeological context (multiple human skeletons in 13 out of 16 excavation units) suggests that most of the individuals probably died due to the epidemics that devastated Mexico City and the central plateau of New Spain which resulted in mass deaths [21, 101, 102].

Archaeological sites and sample description

Archaeological context
The samples used in this study were obtained from three individuals whose skeletons were recovered from the archaeological context of the San José de los Naturales Royal Hospital. These almost complete skeletons were recovered from the cemetery associated with the hospital during the excavations for a new subway line in downtown Mexico City between 1988 and 1994, under the supervision of the archaeologists Salvador Pulido Méndez and María de Jesús Sánchez Vázquez [21, 24]; during the season of 1992 the skeletons were found in the third stage of the excavation, suggesting they belonged to the oldest structures within the archaeological context.

Individuals and samples
Due to the possibility that these three individuals are of distinct ethnic origin (given that ethnic heterogeneity was promoted to generate social disintegration as a way to prevent the creation of communities and social organization among African slaves [2, 3, 5]). We collected one tooth from each of the following individuals: ML8 SL 150 (age at death: 25-30 years old, male [25]; first right molar of the mandible, genomic library SJN001), ML8 San José 214 (age at death: 30-35 years old, male [25]; second right molar of the mandible), and ML8 SLUBB 296 (age at death: 25-30 years old, male [25]; second right molar of the mandible, genomic library SJN003). Samples were obtained under controlled conditions in the Osteology Laboratory of the Post Graduate Studies Division at the National School of Anthropology and History (ENAH) with a protocol devoted to minimize any possibility of contamination. Photographic records of the samples were obtained throughout the whole procedure. These individuals were previously described of potential African origin [6, 25, 23, 25] because of dental modifications (Figure 1) consistent with those practiced by sub-Saharan ethnic groups [25, 43, 103], as well as apparent dental and skeletal features with suggested affinity to sub-Saharan human groups [26, 101]. The radiocarbon dates and strontium isotope analyses of the three individuals’ samples were processed and analyzed at The Curt-Engelhorn-Centre for Archaeometry (Mannheim, Germany). Radiocarbon dates are reported showing the lab codes MAMS, and were processed using ultra filtrated collagen (fraction > 30kD) and dated using the MICADAS-AMS of the Klaus-Tschira-Archäometrie Zentrum. Strontium isotopes determination was performed in a high-resolution ICP mass spectrometer in a clean room laboratory at the same laboratory. Since this is quite a precise determination, as reflected by the reproducibility of NBS-987 measurements, we can confidently report and trust values to the fifth decimal place [63]. Osteological analyses were conducted on the skeletons in search for skeletal evidence of pathologies or trauma using the health indicators proposed by Goodman and Martin as well as references on osteological analyses of pathology [86, 104–106]. This methodology includes the identification of bone modifications in the skeleton present in the sites of muscle or ligament insertion caused by the hyperactivity of the main muscles responsible for the movement [107]. Dental modification patterns were assessed by comparing the dental decoration present in the SJN Africans with previously observed patterns [following the classic works by Reinaldo de Almeida (1953, 1957), as cited by references [41, 106]]. The individual biological affinity was estimated by cranial non-metric traits ancestry estimation, ranked by their degree of expression, with which percentages of biological affinity were obtained according to four biogeographical groups: Native American, African, Asian and European [109–111].

METHOD DETAILS

Osteological assessment

Osteobiography of individual ML8 SL 150 (SJN001)
Estimation of sex was performed by morphological analysis of sexually dimorphic features of the os coxa (pre-auricular sulcus, greater sciatic notch, subpubic morphology). Results indicate that this individual was likely a biologically male individual [112]. The age at death was estimated around 24 years and was obtained by analyzing the scalene tubercle and the first rib face (range 24.28 to 51.86 years) [113]; the pubic symphysis (range: 22-25 years) [114] and the auricular surface of the ilium (range: 24-28 years) [115], which is consistent with previous estimates [25]. Health: The skull shows moderate cribra orbitalia and severe porotic hyperostosis. These abnormal skeletal changes are often in response to conditions associated with diet such as anemia and malnutrition, as well as parasitic infections and blood loss [38–40]. Observations investigating dental health reveal some signs of periodontal disease including receding of the alveoli with abnormal margins with associated alveolar inflammation. The tibiae and fibulae display dense sclerotic periosteal deposition suggestive of well-healed periosteal reaction. It is possible that these pathological changes are associated with conditions of poor hygiene, infectious processes, and malnutrition [40, 105, 107]. Occupational activity: The left clavicle presents enthesophyte development with associated raised margins at the insertion point of the coracoclavicular ligament, this may be attributed to remodelling of bone that occurs due to a microtraumas associated with repetitive use and has been reported to be found among individuals that they carry weighted loads on their shoulders [107]. The right clavicle presents exostosis in the anterolateral aspect located in the insertion of the coracoclavicular ligament (Figure S5A). The seventh thoracic vertebra presents Schmör’s node (Figure S5B) on the inferior aspect of the vertebral body resulting from the compression and subsequent hemia of inter-vertebral discs. Likewise, the ninth to the twelfth vertebrae present Schmör’s nodes on the superior and inferior aspects of the vertebral bodies. These nodes, or hemias, are the result of intense gradual compression and excessive
biomechanical stress [116]. Cultural practices: This individual shows dental modification in the four maxillary incisors (Figure 1 in the main text) made with two techniques: percussion fracture and filing, but also it can be seen some decoration work done in the mandibular canines [25, 107]. The patterns seen in the maxillary incisors are consistent with a “V” shaped modification present in Fang and other ethnic groups from the coastal region of central-Western Africa (present day Equatorial Guinea, northern Gabon, and southern Cameroon) [41, 43, 108]. The mandibular canines, however, present a dental modification pattern seen during 19th-century expeditions in the Kingdom of Loango (present day Republic of the Congo), a multi-ethnic kingdom in the central-western part of Africa; and in Bakongo people from the Congo region [43]. Taphonomy: We observed green coloration on multiple skeletal elements acquired by contact with copper in the burial associated with the body. This green coloration was observed on the fifth and seventh cervical vertebra (Figure S5C) and the fourth and fifth left ribs (Figure S5D), there is a pattern of continuous inclined lines on the coloration. These indications are considered by other authors [23, 28] as a result of gun impacts. It is likely this individual had been shot and buried with the fragments in his body thus resulting in the green coloration of the skeleton. Biological affinity: The biological affinity of this individual was estimated from macromorphoscopic characteristics [109, 111], these characters are established by their degree of expression. For this work, the values obtained from the observation were analyzed using the software Osteomics [110], which bases its classification by a simple Bayesian method and proposes a probabilistic result taking into account the percentages of biological affinity according to four biogeographic groups: Native American, African, Asian and European. Individual 150 presented a large interorbital width (IOW); large nasal aperture width (NAW); as well as deep post-bregmatic depression (PBD); intermediate anterior nasal spine (ANS); inferior nasal aperture (INA) with border; a pronounced malar tubercle (MT); nasal bone contour (NBC) with intermediate plateau; supranasal suture (SPS) closed but visible; and zygomaticomaxillary suture (ZS) with an angle in the middle part; nasal overgrowth (NO) and transverse palatine suture (TPS) were not analyzed due to them being fractured. According to the classification analysis of Osteomics, a 92.23% probability of belonging to the African biogeographical group was found.

Osteobiography of individual ML8 San José 214 (SJN002)
The biological sex estimation of the individual was performed through observations of sexually dimorphic traits. Morphological analysis of the os coxa (preauricular sulcus, sciatic notch, composite arch, anterior edge form of the ischiopubic ramus and the ischiopubic ratio) [112] with particular attention to the morphology of the features associated with the pubic region (ventral arch, subpubic concavity, and the medial aspect of the ischiopubic ramus) resulted in the estimation of a male individual [117]. The age at death of the individual was estimated to be approximately 25 years and was calculated through age associated morphological observations of the scalene tubercle and the first rib face (range of 24.28 to 51.86 years) [113]; the pubic symphysis (range: 22-28 years) [114] and the auricular surface of the ilium (range: 25-29 years) [115]. Health: This individual presents signs of dental pathology including an abscess in the second lower right molar (Figure S6I), as well as periodontitis. The diaphyses of the lower limbs display periostea deposition at the inferior most aspects. The right femora display sclerotic bone deposition creating an irregular surface which is suggestive of a healed peristoeal reaction in life. Likewise, the tibiae and fibulae display peristoeal reaction active at time of death; this is most apparent on the right tibia and fibula where the periostea bone is deposited near and around the point of fracture with associated dense sclerotic and osteomyelitic bone development. Degenerative diseases: Degenerative disease of the joints is present and can be observed on the distal articular surface of the femur (Figure S6A) as osteochondritis dissecans, a defect of the articular surface of the bone due to compromise of the synovial capsule, and associated signs of osteoarthritis i.e., eburnation, lipping of the joint margins, and irregular joint surface [105]. Osteophytic development is present on the eighth, eleventh and twelfth thoracic vertebrae, and third and fourth lumbar vertebrae (Figure S6H); although these are age associated skeletal changes, development in younger individuals is associated with intense physical activity and excessive mechanical stress [107]. Here we see extreme osteophytic development and compression of the vertebrae. Trauma: Trauma refers to any sudden physical injury resulting from an external force. This individual, presents a healed fracture of the right tibia and fibula in the distal third of the diaphysis (Figures S6B–S6D) that extends from the lateral aspect to the medial aspect covered by and dense bony callus which is consistent with a healed oblique fracture and cross-section caused by the combination of axial compression force and angulation [107]. Healing of the fracture occurred without correct alignment of the fractured bones, which resulted in poor bone consolidation. The skull also presents a cut mark in the frontal bone with signs of bone regeneration (Figure S6J), which is indicative of the lesion happening ante mortem [118]. Occupational activity: The left and right clavicles display skeletal changes at the costoclavicular ligament, however, the right clavicle presents a higher degree of change with enthesal changes including depressed grained surface displaying raised irregular margins at the costoclavicular ligament (Figures S6E–S6G). The ulna presents enthesophyte development on the olecranon at the insertion of the triceps brachii, it is associated with continued biomechanical stress of the arm [107]. The lumbar vertebrae present Schmorl’s nodes on the superior and inferior aspects of the vertebral bodies, resulting from the compression exerted on the inter-vertebral discs. In this individual, multiple Schmorl’s nodes can be observed from the first to the fourth lumbar vertebrae. These vertebrae display skeletal changes (Figure S6H), as a result of an intense and gradual compression from biomechanical requirements including osteophytes on the superior and inferior surface and severe compression of the vertebra bodies in response to compression over time. The linea aspera of both femora show rugose enthesophyte development, a skeletal change associated with activity patterns; the significant enthesophytic development on the right femur is likely due to alterations in biomechanical activity resulting from the unaligned fracture of the proximal tibia and fibula. Cultural practices: This individual also presented dental modification in the four maxillary incisors (Figure 1 in the main text) from what appeared to be the fracture by percussion technique [107]. The patterns seen in the maxillary incisors are consistent with the “pattern No. 2” (resembling the “T” shape pattern proposed by Reichart et al. [105]) proposed by Wasterlain et al. [41], consisting in the removal of both mesial and distal incisal angles. This pattern was reported in D’zem (Njem) people, an ethnic group far to the east of the Dscha Valley (present day Cameroon) and their neighbors [43], but also
remarkably similar to that found in the individual no. 81 from the deposit of urban waste of the Valle da Gafaria (Lagos, Portugal) [41]. Biological affinity: The biological affinity of the individual was estimated from macromorphoscopic characteristics [109–111]. The individual shows a large interorbital width (IOW), a medium nasal aperture width (NAW), as well as deep post-bregmatic depression (PBD); small anterior nasal spine (ANS); inferior nasal aperture (INA) with a little border; a not so pronounced malar tubercle (MT); a rounded nasal bone contour (NBC); a closed but visible supranasal suture (SPS); and zygomaticomaxillary suture (ZS) with an angle in the middle part; nasal overgrowth (NO) and transverse palatine suture (TPS) were not analyzed due to them being fractured [109, 111]. According to the classification analysis of Osteomics, a 95.24% probability of belonging to the African biogeographical group was found.

Osteobiography of individual ML8 SLU9B 296 (SJN003)

The sex estimation of individual SJN003 was performed through morphological observations of the os coxa (preauricular sulcus, sciatic notch, composite arch, anterior edge form of the ischiopubic ramus and the ischiopubic ratio) [112] including the morphology of the pubic region (ventral arch, subpubic concavity, and the medial aspect of the ischiopubic ramus) [117]. All observations resulted in estimations of a biologically male individual. The age at death was estimated to be approximately 25.26 years and was calculated by analyzing the scalene tubercle and the first rib face [113], which is consistent with previous estimates [25]. It was not possible to obtain age using another method due to preservation issues of the pubic symphysis and the auricular surface of the ilium. Health: The skull presents moderate cribra orbitalia. The number of permanent teeth observed is 13 and an alveolus; abscesses or injured teeth were not present and periodontitis is observed. The presence of cribra orbitalia and hyperostosis in the skeleton can be due to three factors: diet such as malnutrition and anemia, parasitic infections, and blood loss [86]. The diaphyses of the long bones present slight infectious processes, with the exception of the femurs in which osteomyelitis can be observed on the anterior and posterior sides of the shaft (Figure S7A). Periosteal changes could be associated with a variety of physiological and environmental stressors including diet or infectious processes [86]. Occupational activity: The left clavicle diaphysis (Figure S7B) presents diffuse cortical irregularity and a roughened, raised appearance at the muscular insertion of the deltoid [86]. Cultural practices: This individual presents dental modification of maxillary and mandibular canines and central and lateral incisors (Figure 1 in the main text). The patterns seem resemble the “pattern No. 2” (resembling the “T” shape pattern proposed by Reichart et al. [108]) proposed by Wasterlain et al. [41], consisting in the removal of both mesial and distal incisal angles. This pattern was reported in D’zem (Njem) people, an ethnic group far to the east of the Dsch a Valley (present day Cameroon) and their neighbors [43], and very similar to that found in the individual no. 81 from the deposit of urban waste of the Valle da Gafaria (Lagos, Portugal) [41]. Biological affinity: The biological affinity of the individual was estimated from macromorphoscopic characteristics [109–111]. The individual shows a large interorbital width (IOW), a medium nasal aperture width (NAW), as well as deep post-bregmatic depression (PBD); medium anterior nasal spine (ANS); inferior nasal aperture (INA) with a pronounced border; a not so pronounced malar tubercle (MT); a slightly rounded nasal bone contour (NBC) with high walls; a closed but visible supranasal suture (SPS); transverse palatine suture (TPS) goes all way through and reaches the medium line, where it projects to both the anterior and the posterior ends drawing a M shaped line; and zygomaticomaxillary suture (ZS) with an angle in the middle part; nasal overgrowth (NO) was not analyzed due to it being fractured [109, 111]. According to the classification analysis of Osteomics, a 91.95% probability of belonging to the African biogeographical group was found.

Stable isotope analysis, diets, and the reservoir effect

Stable isotope analysis of human bone can provide insights into the significance of different resources to the protein component of the diet, this includes the potential reliance on marine resources and their impact on 14C radiocarbon measurements through the ‘reservoir effect’ [119, 120]. Here, we analyzed dentine collagen from the same teeth used to conduct the radiocarbon dating and aDNA analysis. These teeth were first (SJN001) and second molars (SJN002 and SJN003), providing dietary information for the years of tooth formation (crown-enamel formation at 2-2.5 years old for the first molar, and between 7-8 years old for the second molar) [121, 122]. Sample SJN001, as a first molar, should be treated with caution as high 15N values could be the product of breastfeeding which results in a higher trophic level appearance [34]. We obtained 250 mg of dentine powder from the crown of the analyzed teeth using a diamond-tipped drill. Collagen was then extracted from this powder using standard procedures [123]. The powder was demineralised in 10 mL aliquots of 0.5M HCl at 4°C. The acid was changed until CO2 stopped evolving. The residue was rinsed three times in deionised water before being gelatinized in pH = 3 HCl at 75°C for 48 hours. The resulting solution was filtered, with the supernatant then being lyophilized over a period of 24 hours. After calculating the collagen yield, all purified collagen samples (1 mg) were located in tin capsules to be analyzed in duplicate at the Department of Archaeology, Max Planck Institute for the Science of Human History using a Thermo Fisher™ Elemental Analyzer coupled to a Thermo Fisher™ Delta V™ Advantage Mass Spectrometer via a ConFloIV system. Accuracy was determined by measurements of international standard reference materials within each analytical run. These were USGS40 10Craw = −26.4 ± 0.1, 13Ctrue = −26.4 ± 0.0, 15Ntrue = −4.4 ± 0.1, 15Ntrue = −4.5 ± 0.2; IAEA N2 15Ntrue = 20.2 ± 0.1, 15Ntrue = 20.3 ± 0.2; IAEA C6 13Ctrue = −10.9 ± 0.1, 13Ctrue = −10.8 ± 0.0. The atomic C:N ratio along with the collagen yields were used in order to determine the quality of collagen preservation. Collagen yields over 1 wt% are considered acceptable for carbon and nitrogen values [124], while the C:N ratio should range from 2.9 to 3.6 [125]. Stable carbon and nitrogen isotope ratios of human tissues are expressed in δ notation relative to established international standards and shown in parts per thousand (‘‰). The delta values are obtained by the following equation:
where the R represents the ratio between the heavier (13C, 15N) and lighter (12C, 14N) isotopes [126]. δ13C and δ15N analysis of human bone collagen primarily reflects the isotopic values of the protein input to the diet, with minimal contribution of lipids and carbohydrates, meaning that it will be heavily influenced by protein-rich foods [127].

Stable carbon isotope variability in terrestrial ecosystems is primarily driven by two dominant photosynthetic pathways, C3 and C4, which differ in their net discrimination against 13C during CO2 fixation [128]. In C3 plants, strong discrimination against 13C results in lower 13C values in virtually all trees, shrubs, and temperate grasses, including wheat, than in C4 plants such as maize and millet [129]. C3 13C values vary from c. −24 to −36‰ (global mean −26.5‰), while C4 values range from c. −9 to −17‰ (global mean −12‰) [129]. C3 and C4 plants thus have distinct and non-overlapping 13C values [130]. These distinctions are reflected in the tissues of consumers, with small trophic level effects of 1-2‰ [127].

Stable nitrogen isotope ratios vary with trophic level, and 15N trophic shifts of +2-6‰ from plants to herbivores, and from herbivores to carnivores, is well documented in marine and terrestrial systems [35, 131]. This trophic effect is most likely linked to the loss of 15N-depleted excretion products [132], although diet-tissue distinctions are highly variable between animals [133]. The long length of marine food chains, leads to distinctively high 15N in marine foods and consumers compared to their terrestrial counterparts [134]. Freshwater foods also tend to have high 15N though 13C does not follow the same trend toward higher measurements as in marine food chains due to different sources of carbon dioxide for primary producers [135].

Despite the long history of stable isotope research in African archaeology, there is a significant inequality in the application of these studies across the regions of Africa [36]. While South and East Africa have been well covered, other regions, especially West Africa, remain relatively unexplored [38]. In the context of 16th-19th century slavery, historical records are often lacking in terms of insights into the origins of enslaved individuals and their subsistence strategies, focusing on coastal embarkation points into the Atlantic Trade rather than places of residence which may, in many cases, have been hundreds of kilometres away [136]. Stable carbon and nitrogen isotope analysis of 16th-19th century slave populations in southern Africa [137] and the Caribbean [67] have already shown the efficacy of this methodology for uncovering insights into individual origins.

In the context of 16th-19th West Africa, stable isotopic insights into broad subsistence bases may also provide geographic information. For example, along the west coast there was a well-documented focus on rice (C3) [138], in the south of Ghana there was a greater focus on yams and other root crops (C3), and in the semi-arid interior there was often a contribution of C4 crops such as sorghum and millet [139]. Thus, the different regional agriculture techniques of subsistence may be useful in establishing the origin of those individuals enslaved in the West of Africa [67]. That said, existing isotopic information from slaves of African origin in the Americas indicate an adult diet of C4 crops and marine resources [64, 67, 140, 141]. This is particularly significant in our context given its potential influence on establishing a secure chronology for human remains.

Marine and freshwater environments are reservoirs of 14C. This means that in a marine or freshwater environment, 14C concentrations in CO2 are lower than contemporary atmospheric values, leading to ‘older’ radiocarbon dates in those organisms relying on freshwater or marine systems [142]. Aquatic plants and phytoplankton fix this CO2, transferring relatively low concentrations of 14C up the foodchain. The direct or indirect consumption of marine proteins by human beings will thus impact 14C measurements of their tissues, making them seem ‘older’ through the marine reservoir effect [143–145]. Thus, before the conversion of the radiocarbon date results to a calendar age, it is important to determine the degree of marine consumption in an individual.

**Ancient DNA sample processing and quality control**

**Sampling**

All samples were processed in dedicated laboratories at the Max Planck Institute for the Science of Human History in Jena, Germany. Tooth powder (~80 mg) was obtained by cutting each dental piece at the junction between the root and crown and sampling the dental pulp from the crown (A libraries) and the root (B libraries). These procedures were carried out after a bleach/rinse and UV decontamination protocol. After sampling, the remaining root was used for radiocarbon dating, while the crown was sampled for Sr determinations as previously described.

**DNA extraction**

DNA was obtained from bone powder after overnight extraction at 37°C with a decalcifying buffer (900 μL EDTA 0.5M, 75μL H2O and 25 μL Proteinase K) [146]. The whole mixture was then centrifuged to pellet the remaining tooth powder and the liquid transferred into a binding buffer as previously described [147]. DNA was then purified by a silica column-based method using silica columns for high purity and quantified using qPCR with the IS7 and IS8 primers in a quantification assay using a DyNAmo SYBR Green qPCR Kit (Thermo Fisher Scientific, L and frozen at −20°C until library preparation. Negative blanks and cave bear positive controls were included for each step of the procedure.

**Library preparation**

We built one non-uracil DNA glycosylase (non-UDG) treated library using 15 μL of each DNA extract to assess the authenticity of the extracted DNA after obtaining the characteristic damage plots associated with ancient DNA [148]. We then used 20 μL of each DNA extract to build UDG-half libraries with Illumina-specific adapters following a modified double-stranded library preparation protocol as previously described [149, 150]. Each library was treated independently despite being from the same individual. Libraries were quantified using qPCR with the IS7 and IS8 primers in a quantification assay using a DyNAmo SYBR Green qPCR Kit (Thermo Fisher Scientific, Pleasanton, CA).

Current Biology 30, 2078–2091.e1–e11, June 8, 2020 e6
To capture the remaining 5%, a set of 162 Nucleotide database [157]. Introns were also used in the case of MIC genes. Given the polymorphic nature of these genes, we added sequences for these genes were extracted from the human genome build [29] using the hg19 [29] and RefSeqGene records from the NCB/ Nucleotide database [157]. Introns were also used in the case of MIC genes. Given the polymorphic nature of these genes, we added alternative alleles for HLA, MIC, TAP and KIR genomic regions, which were obtained from the IMGT/HLA database [158]. For HLA class I and KIR genes the intronic regions were also included. For the HLA and MIC genes a set of 83 representative alleles with full-length gene sequences was chosen that encompasses the major allelic groups [159] and covers 95% of the known polymorphism. To capture the remaining 5%, a set of 162 × 160 base pairs (bp) consensus sequences was designed. A 60-bp probe set was designed at every 5bp interval along the target sequences. The final 52-bp probe sequences were mapped to hg19 using RazerS3 [160] with a minimum threshold of 95% identity. The probe set was triplicated to complete capacity of the one-million feature SureSelect DNA Capture Array, which was turned into an in-solution DNA capture library as described elsewhere [153]. We enriched the reads mapping to the HLA region with the described in-solution capture approach, yielding 10X-100X coverage for this genomic region. For dealing with the allele assignment coming from short reads, we used Optitype [54], a software specifically developed to deal with short reads coming from ancient DNA data. Our capture approach was validated using previously typed samples coming from an external quality control program (UCLA’s International HLA DNA Exchange), for which the DNA was shredded using the Covaris system to provide short DNA fragments for a proper validation of the results. The fact that the alleles typed in our individuals are commonly found in Sub-Saharan African populations further supports our findings.

Treponema pallidum capture design

For targeted enrichment of Treponema pallidum DNA probes were designed on the basis of Treponema pallidum subsp. pallidum, T. pallidum subsp. endemicum and T. pallidum subsp. pertenue strains. The probes were designed with a 1bp tiling and a length of 52 bp with an additional 8bp linker sequence (CAGTCGCG) as described previously [153]. Duplicated probes and probes with low sequence complexity were removed. This resulted in 1,125,985 unique probe sequences. The probe set was spread on two Agilent one-million feature SureSelect DNA Capture Arrays. The capacity of the two arrays was filled by randomly duplicating probes from the probe set. The arrays were turned into an in-solution DNA capture library as described elsewhere [153].

List of Treponema pallidum genomes IDs used for probe design

Treponema pallidum subsp. pallidum strains Nichols (NC_000919.1), SS14 (NC_021508.1), Sea 81-4 (NZ_CP003679.1), Mexico A (NC_018722.1), T. pallidum subsp. endemicum strain Bosnia A (NC_CP007548.1), and T. pallidum subsp. pertenue strain Fribourg-Blanc (NC_021179.1).

Hepatitis B Virus (HBV) capture design

For targeted HBV DNA enrichment, probes were designed on the basis of a worldwide set of HBV full genomes (see list of GenBank accessions below). The probes were designed with a 1bp tiling and a length of 52 bp with an additional 8bp linker sequence (CAGTCGCG) as described previously [153]. Duplicated probes and probes with low sequence complexity were removed. This resulted in 221,190 unique probe sequences. This probe set was replicated four times on an Agilent one-million feature SureSelect DNA Capture Array, which was turned into an in-solution DNA capture library as described elsewhere [153].

List of HBV genome IDs used for probe design

Petersberg, Sorsum, Karsdorf, LT992459, LT992455, LT992454, LT992448, LT992447, LT992444, LT992443, LT992442, LT992441, LT992440, LT992439, LT992438, JN315779, MG585269, AB076679, AB116084, AB453988, AY738142, GQ477499, AY934764,
were reconditioned by further amplifying with IS5/IS6 primers and was performed on 5.25 of 200–400 ng/
were enriched via in-solution capture from the total DNA in the sequencing libraries [152–154]. Briefly, UDG-half- treated libraries
respective libraries. Using in-solution capture based on modified immortalized probe sequences [153], target pathogen genomes
pools were paired-end sequenced on the Illumina Hiseq 4000 (Illumina, Inc., San Diego, CA) with 75 cycles providing on average
10 million reads per sample for the first round of capture and 20 million (Treponema) and 13 million (HBV) reads per sample for the
second round of capture.

Pathogen genome capture
After finding reads mapping to HBV (SJN001) and Treponema pallidum (SJN003), we selectively enriched for these pathogens in their respective libraries. Using in-solution capture based on modified immortalized probe sequences [153], target pathogen genomes were enriched via in-solution capture from the total DNA in the sequencing libraries [152–154]. Briefly, UDG-half- treated libraries were reconditioned by further amplifying with IS5/IS6 primers and Herculase II Fusion DNA Polymerase to reach a concentration of 200–400 ng/µL as measured on a NanoDrop 8000 spectrophotometer (Thermo Fisher Scientific Inc.; Waltham, MA). Capture was performed on 5.25 µL of each re-conditioned library using that volume for each capture procedure [153, 154]. For targeted HBV and Treponema genome DNA enrichment, we used the previously described sets of probes. After enrichment, captured library pools were paired-end sequenced on the Illumina Hiseq 4000 (Illumina, Inc., San Diego, CA) with 75 cycles providing on average
10 million reads per sample for the first round of capture and 20 million (Treponema) and 13 million (HBV) reads per sample for the second round of capture.

QUANTIFICATION AND STATISTICAL ANALYSIS

QC and data processing for human PopGen
We performed an analysis of the human captured sequence data using EAGER ver. 1.92.55 [89], Clip&Merge [90] and AdapterRemoval v.2 [88] were used to trim adaptor sequences and to remove adaptor dimers and low-quality sequence reads (min length = 30; min base quality = 20) from the reads resulting from the sequencing of the aDNA libraries. Pre-processed sequences were mapped to the human genome assembly GRCh37 (hg19) from the Genome Reference Consortium [29] using BWA ver. 0.7.12 [89] and a default seed length of 32.

aDNA Authentication
The C to T misincorporation frequencies typical of aDNA were obtained using mapDamage 2.0 [92] to assess the authenticity of the ancient DNA fragments from the non-UDG treated libraries. Characteristic short average fragment length (49–55 base pairs) and the increased proportion of miscoding lesions due to deamination at the molecule extremes were found on the six UDG-half libraries analyzed (Table S2).

Data processing
Duplicates were removed with DeDup [90], which removes identical reads. UDG-half libraries were trimmed for the first and last three positions to reduce the impact of deamination-induced misincorporations during genotyping. We used samtools mpileup (parameters –q 30 –Q 30 –B) to generate a pileup file from the merged sequence data of each individual, and used a custom script [pileupCaller ver. 8.2.2 [161]] to genotype the individuals, using a pseudo-haploid random draw approach. For each position on our capture panel, a random read was drawn for each individual and the allele of that read was assumed to be the homozygous genotype of the individual at that position. In order to compare with available data from African populations, were merged our SNPs to the approx. 600,000 SNPs of the Human Origins dataset [30]. Genetic sex of the three samples was assigned using SNP capture data by calculating the ratio of average X chromosomal and Y chromosomal coverage to average autosomal coverage at the targeted SNPs [31]. Samples with an X rate between 0.35 and 0.55 and a Y rate between 0.4 and 0.7 were confirmed male. Since the three analyzed individuals were male, ANGSW was run to measure the rate of heterozygosity of polymorphic sites on the X chromosome after accounting for sequencing errors in the flanking regions to estimate the nuclear contamination, since males are expected to have only one allele at each position [32]. Reads mapping to the human mitochondrial DNA (mtDNA) were used to reconstruct a mtDNA consensus sequence and estimate contamination levels with schmutz [33]. Present-day human contamination estimates were performed using a comparative database of 197 modern-day worldwide mtDNA sequences provided with the software package. After assessing the quality of the independent libraries, libraries A and B of each individual were merged using SAMtools ver. 1.3 [94].

Uniparental markers
Mitochondrial DNA (mtDNA) genomes were determined mapping reads to the revised Cambridge reference sequence [162]. For the resulting sequences, we filtered positions with likelihoods above 10, 20 or 30 and used HaploGrep2 [163] and HAPLOFIND [164] to assign and confirm the corresponding mtDNA haplogroups. The aligned mtDNA genomes were compared to other available genomes from literature which belong to one of the three haplogroups retrieved (153 sequences for L1b, 208 for L3d and 415 for L3e). The mtDNA sequences were aligned with MAFFT (http://mafft.cbrc.jp/alignment/software/) [165] and manually checked with Bioedit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The two poly-C regions (positions: 303–315, 16,183–16,194) were excluded from the analysis. Haplotype based networks (Figure S1) were built with R with the command haploNet of package PEGAS [166]. Sequenced reads mapped to Y-Chromosome (Y-Chr) SNPs within our capture panel that are also present in the Y-DNA...
Haplogroup Tree 2018 ver. 13.174 [167] were used to assign Y-Chr haplogroups for each individual. Haplogroup assignment was manually confirmed by looking at the most downstream SNP retrieved after assessing the presence of upstream mutations along the Y-Chr haplogroup phylogeny.

**Principal components analysis**

smartpca (version 16000) from the *Eigensoft* package [168] was used to calculate Principal Components (PCs) of variation in the present-day populations from the HO dataset, using the options “lsqproject: YES” and “shrinkmode: YES.” Initially, we projected the three ancient individuals on PCs calculated on the genetic variation in 371 worldwide populations, to access the continental-level ancestries in the ancient individuals (Figure S2). We then projected the ancient individuals on PCs calculated on variation from 534 Africans from 51 populations, both from Northern and sub-Saharan Africa (Figures 2 and S2).

**Worldwide populations**


**African and related populations**

Algerian, Bantu (Kenya), Bantu (South Africa), Bantu (Herero), Bantu (Ovambo), Bedouin A, Bedouin B, Biaka, Canary Islander, Damara, Datog, Dinka, Egyptian, Esan, French, Gambian, Gana, Gui, Hadza, Haiom, Han, Himba, Hoan, Igbo, Ethiopian Jew, Libyan Jew, Moroccan Jew, Tunisian Jew, Ju hoan North, Ju hoan South, Kgalagadi, Khamani, Khigh, Kikuyu, Kongo,
Lemande, Libyan, Luhya, Luo, Mandenka, Masai, Mbuti, Mende, Mixe, Moroccan, Mozabite, Nama, Naro, New Guinea, Oromo, Sahrawi, Sandawe, Shua, SJN, Somali, Spanish, Taa East, Taa North, Taa West, Tshwa, Tswana, Tunisian, Wambo, Xuu, Yoruba, Zapotec.

**F₃ and D tests**

To assess the genetic relationships and admixtures suggested in the PCA and ADMIXTURE analysis, we carried out F-statistics using the programs qp3Pop ver. 412 and qpDstat ver. 711 in the ADMIXTOOLS suite [172] for F₃- and D-statistics, respectively. We tested each individual independently, as they most probably were of different ancestries as explained before. We used a F₃-statistics of the form f₃ (Outgroup; X; Y) to measure the amount of shared genetic drift of populations X and Y after their divergence from a non-African outgroup (Figure 3); were X is each of the SJN individuals, Y is each of the populations in our database, and Outgroup is the genome of the Ust'-Ishim individual [62]. D-statistics of the form D (Chimp, SJN00X; Y; Target) were used to demonstrate if our African individuals are related to a target population or shared an excess of alleles with any population in position Y. A negative value implies that either Chimp and Target, or SJN and Y share more alleles than expected under the null hypothesis of a symmetrical relationship between Y and Target. Target was each of the populations with the highest top F₃ values for each SJN individual. For SJN001 (Figure S3A) Mende were used as Target; for SJN002 (Figure S3B), Bantu-speaking Ovambo were used as Target; for SJN003 (Figure S3C) Lemande were used as Target. To test for cladality between each of the individuals from SJN and their respective Target populations, we computed D statistics of the form D (Chimp; X; Target, SJN00X). We find that none of the ancient individuals are truly cladal with the Target populations (Table S1).

**Pathogen genomes assemblies**

**Hepatitis B Virus genome**

After HBV DNA capture and sequencing, resulting reads were merged and aligned against an HBV genome (GenBank accession: KC875253) using Geneious v. 9 with medium sensitivity settings. A 90% consensus sequence was generated with a coverage threshold of 3x. Reads were remapped against the consensus sequence within the Eager pipeline [90] to perform final coverage and damage assessment. The sequence was aligned with a worldwide set of HBV full genomes (see accession numbers below), using MAFFT [165]. The alignment was cleaned using Gblocks [173], removing positions with more than 50% gaps. A phylogenetic tree was constructed using RAxML v. 8 [174] with GTRCAT substitution model and the rapid bootstrap algorithm (Figure 4).

**Treponema pallidum sub. pertenue genome and phylogenetic analyses**

Shotgun reads sequenced from both libraries of SJN003 were screened for pathogens in MALT [175]. Reads mapping to Treponema pallidum were identified and their ancient origin was validated from deamination of cytosine to thymine bases at the reads’ ends. A whole genome-wide capture with respect to the reference genome was created (Table S3).

To compare the ancient and modern treponemal genomes, a genome wide alignment and concatenated SNP alignment of Treponema pallidum sub. pallidum (n = 5), Treponema pallidum sub. endemicum (n = 2) and Treponema pallidum sub. pertenue (n = 26) genomes was generated. A maximum likelihood tree was reconstructed using RAxML [174], for 1000 replicates with a GTR model of substitution and a GAMMA distribution of rate heterogeneity model for eight categories. SNPs in SJN003 and 133 were analyzed in stringent setting (-n) and mapping quality of 37 followed by exclusion of duplicated reads using Mark Duplicates [176]. GATK identified genotypes for all sites and variants were estimated. Variants from all vcf. files were compared and combined for reference Nichols genome using MultiVCFAnalyser v. 0.87 (https://github.com/alexherbig/MultiVCFAnalyzer [177]). For each position, an allele is called if covered by a minimum of three reads, having a mapping quality of 30 and a frequency of 90% or greater among all reads covering the position. Otherwise, the position is considered as “N.” All variant sites are concatenated to a SNP alignment. Additionally, a genome wide alignment with respect to the reference genome was created (Table S3).

**List of HBV genomes GenBank accession numbers used in the phylogenetic analysis**

<table>
<thead>
<tr>
<th>Accession Numbers</th>
<th>Reference Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB076679</td>
<td>AY116084, AB453988, AY738142, GQ477499, AY934764, FJ692556, FJ692598, FJ692611, GQ161813, GQ331046, AB073858, AB033555, AB219429, AB219430, AP011106, AP011108, AB111946, AB121066, AB112472, DQ089767, X75656, X75665, AB048704, AB048705, AF241141, AP011100, AP011102, AP011109, AP011108, AP011106, FJ899792, JN642140, GQ477453, GQ477455, JN642160, JN642163, JN688710, JN688711, GQ922005, HE974378, KJ470893, KJ470896, KJ470898, FJ904430, FJ904436, AB033559, AB048701, AB048702, AB188243, AB210818, AM494716, AY796031, AY92768, DQ315779, X80925, X75657, X75664, AY090458, AY116654, FJ657525, AY090455, AY311369, DQ899144, DQ899146, AB116549, X75663, AF229362, AB166850</td>
</tr>
</tbody>
</table>
AB056513, AB064312, AF405706, AB059660, AB375163, AY090454, AY090457, AB486012, AY330911, AJ131571, AY781180, U46935, AJ131567, AF193863, EU155824, FJ899779, AB049609, FJ904399, AY721612, AY741797, AB270543, EU594409, AB109476, AB555496, GQ205377, HM363593, JN792922, AB562463, FJ023669, EU835241, AY934764.

**List of Treponema spp. genomes used in the phylogenetic analysis**

133 (from Santa Isabel, colonial Mexico [71]), A10, A12, IGU, Bosnia_A, CDC 2575, CDC-2, CDC-1, Fribourg-Blanc, Gauthier, Ghana-051, HATO, Iraq B, Kampung Dalan K363, LMNP-1, M2, M3, Nichols, Sea81-4, ERS945418, ERS945420, ERS945424, ERS945426, ERS945430, ERS945436, ERS945437, ERS945442, Samoa D, Sei Geringging K403, SS14, Mexico A, 94B (from Santa Isabel, colonial Mexico [71]).