Article

Ancient genomes reveal insights into ritual life at Chichén Itzá

🔴 Check for updates

Rodrigo Barquera[1,2,20✉], Oana Del Castillo-Chávez[3,20✉], Kathrin Nägele[1,20], Patxi Pérez-Ramallo[4,5,6,18,20], Diana Iraíz Hernández-Zaragoza[1,2,20], András Szolek[7,8], Adam Benjamin Rohrlach[1,9], Pablo Librado[10], Ainash Childebayeva[1,19], Raffaela Angelina Bianco[1], Bridget S. Penman[11], Victor Acuña-Alonzo[2], Mary Lucas[4,6], Julio César Lara-Riegos[12], María Ermila Moo-Mezeta[13], Julio César Torres-Romero[12], Patrick Roberts[4,6], Oliver Kohlbacher[7,14,15,16], Christina Warinner[1,17] & Johannes Krause[1✉]

The ancient city of Chichén Itzá in Yucatán, Mexico, was one of the largest and most influential Maya settlements during the Late and Terminal Classic periods (AD 600–1000) and it remains one of the most intensively studied archaeological sites in Mesoamerica[1-4]. However, many questions about the social and cultural use of its ceremonial spaces, as well as its population's genetic ties to other Mesoamerican groups, remain unanswered[2]. Here we present genome-wide data obtained from 64 subadult individuals dating to around AD 500–900 that were found in a subterranean mass burial near the Sacred Cenote (sinkhole) in the ceremonial centre of Chichén Itzá. Genetic analyses showed that all analysed individuals were male and several individuals were closely related, including two pairs of monozygotic twins. Twins feature prominently in Mayan and broader Mesoamerican mythology, where they embody qualities of duality among deities and heroes[5], but until now they had not been identified in ancient Mayan mortuary contexts. Genetic comparison to present-day people in the region shows genetic continuity with the ancient inhabitants of Chichén Itzá, except at certain genetic loci related to human immunity, including the human leukocyte antigen complex, suggesting signals of adaptation due to infectious diseases introduced to the region during the colonial period.

The ancient Maya city of Chichén Itzá, centrally located in the northern part of the Yucatán Peninsula (Fig. 1a,b), ranks among the largest and most iconic archaeological sites in Mesoamerica but much about its origins and history remains poorly understood[1,2]. First rising to prominence during the Late Classic period (AD 600–800), Chichén Itzá became the dominant political centre of the northern Maya lowlands during the Terminal Classic (AD 800–1000), a period when most other Classic Maya sites in the southern and northern lowlands underwent a political collapse. Most of the inscribed calendar dates on carved monuments at Chichén Itzá fall between AD 850 and 875 and the northern ceremonial centre of the site, known as New Chichén, was largely constructed after AD 900 and includes the site's largest structure, *El Castillo*, also known as the Temple of Kukulkán. A *sacbe* (limestone causeway) was constructed to connect New Chichén to the Sacred Cenote[6], an enormous sinkhole containing abundant ritual offerings,

including the remains of more than 200 ritually sacrificed individuals, mostly children[1,3,7]. Evidence of ritual killing is extensive throughout the site of Chichén Itzá and includes both the physical remains of sacrificed individuals as well as representations in monumental art[8]. Elite activity at Chichén Itzá declined during the eleventh century AD, with a last inscribed calendar date of AD 998 (refs. 9,10), but the site continued to be a prominent ritual and pilgrimage centre during the colonial period and beyond[11-13].

In 1967, a repurposed *chultún* containing the remains of more than 100 subadults was discovered near the Sacred Cenote (Fig. 1c,d)[4,14,15]. As for cenotes, *chultúns* (underground cisterns) are associated with water storage and also ritual activity[16,17] and they share symbolism with caves[18]. Such subterranean features have long been associated with water, rain and child sacrifice[14,19,20] and they are widely viewed as access points to the Maya underworld[21,22]. Given the location and context of

[1]Department of Archaeogenetics, Max-Planck Institute for Evolutionary Anthropology (MPI-EVA), Leipzig, Germany. [2]Molecular Genetics Laboratory, Escuela Nacional de Antropología e Historia (ENAH), Mexico City, Mexico. [3]Centro INAH Yucatán, Instituto Nacional de Antropología e Historia (INAH), Mérida, Yucatán, Mexico. [4]isoTROPIC Research Group, Max Planck Institute of Geoanthropology, Jena, Germany. [5]University of the Basque Country (EHU), San Sebastián-Donostia, Spain. [6]Department of Archaeology, Max Planck Institute of Geoanthropology, Jena, Germany. [7]Applied Bioinformatics, Dept. for Computer Science, University of Tübingen, Tübingen, Germany. [8]Department of Immunology, Interfaculty Institute for Cell Biology, University of Tübingen, Tübingen, Germany. [9]School of Computer and Mathematical Sciences, University of Adelaide, Adelaide, South Australia, Australia. [10]Institut de Biologia Evolutiva (CSIC—Universitat Pompeu Fabra), Barcelona, Spain. [11]The Zeeman Institute and the School of Life Sciences, University of Warwick, Coventry, UK. [12]Chemistry Faculty, Universidad Autónoma de Yucatán (UADY), Mérida, Mexico. [13]Nursing Faculty, Universidad Autónoma de Yucatán (UADY), Mérida, Mexico. [14]Institute for Bioinformatics and Medical Informatics, University of Tübingen, Tübingen, Germany. [15]Quantitative Biology Center, University of Tübingen, Tübingen, Germany. [16]Translational Bioinformatics, University Hospital Tübingen, Tübingen, Germany. [17]Department of Anthropology, Harvard University, Cambridge, MA, USA. [18]Present address: Department of Archaeology and Cultural History, University Museum, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [19]Present address: Department of Anthropology, University of Texas at Austin, Austin, TX, USA. [20]These authors contributed equally: Rodrigo Barquera, Oana Del Castillo-Chávez, Kathrin Nägele, Patxi Pérez-Ramallo, Diana Iraíz Hernández-Zaragoza. ✉e-mail: rodrigo_barquera@eva.mpg.de; oana_delcastillo@inah.gob.mx; krause@eva.mpg.de
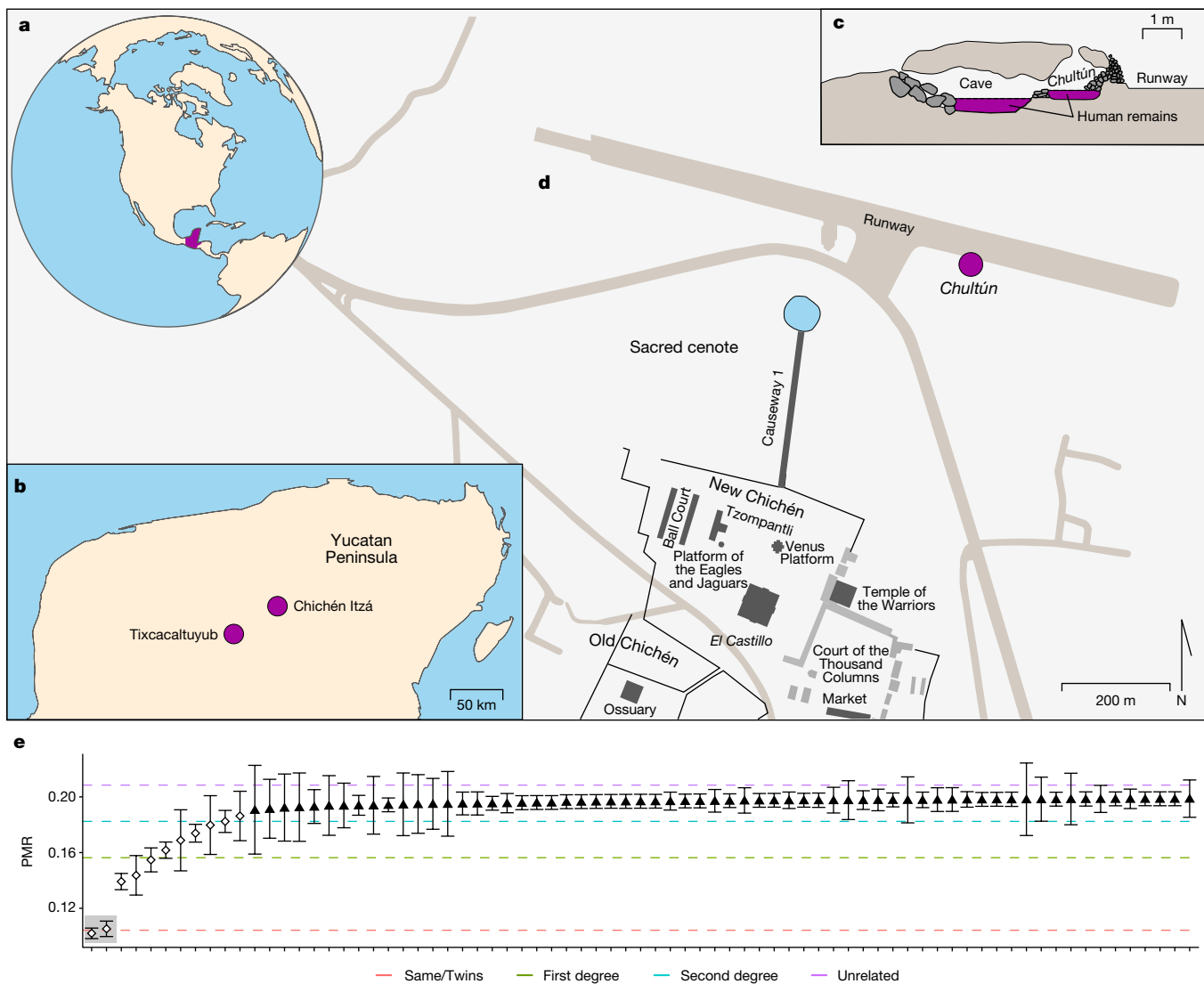
Fig. 1 | Geographical context for the groups analysed and biological relatedness in the *chultún*. a, Location of the Maya region in the Americas. **b**, Geographical locations of Chichén Itzá and Tixcacaltuyub in the Yucatan Peninsula. **c**, Stratigraphy for the *chultún* and the adjacent cave in which the burial was found (adapted from ref. 4). **d**, Location of the *chultún* within the archaeological site of Chichén Itzá and its relation to *El Castillo* (adapted from ref. 10). Modern roads are marked in light grey; the *chultún* abuts an airport runway. **e**, Genetic pairwise mismatch rate (PMR) for child pairs in the *chultún* identifies 11 close relative pairs (hollow diamonds), including two pairs of monozygotic twins (highlighted in grey). A low overall PMR for unrelated individuals (black triangles) confirms low genetic diversity in the population; only pairs with PMR < 0.20 are visualized in the plot. See Supplementary Fig. 2 for individual annotations.

the Chichén Itzá *chultún*, which was also connected to a small underground cave, it has been speculated to contain children sacrificed to support maize agricultural cycles[19] or given as offerings to the Maya rain deity Chaac[23]. Sixteenth century Spanish colonial accounts and early twentieth century investigations following the dredging of the Sacred Cenote popularized the notion that young women and girls were primarily sacrificed at the site[24,25] but more recent osteological analyses indicate that the bodies of both males and females were deposited in the Sacred Cenote[7,26]. Systematic investigations of sacrificial assemblages across the Maya region have confirmed that both males and females were subject to ritual killing but, because most sacrificed individuals at Classic Maya sites are juveniles, precise sex distributions cannot be determined using conventional osteological approaches alone[19,26]. Sixteenth century Spanish sources record that such children were obtained locally by kidnapping, purchase and gift exchange[19,27,28], although recent isotopic studies suggest that at least some individuals within the Sacred Cenote were non-local and may have originated as

far away as Honduras or Central Mexico[1]. Nevertheless, despite more than a century of research, much about child sacrifice and the ceremonial use of subterranean features as ritual mass graves at Chichén Itzá remains unknown.

To better understand the origin and biological relationships of the sacrificed children to each other and to present-day inhabitants of the region, here we used a combined bioarchaeological and genomics approach to investigate 64 subadults interred within a *chultún* near the Sacred Cenote (Fig. 1c) and compare them to 68 present-day Maya inhabitants of the nearby town of Tixcacaltuyub, as well as to other available ancient and contemporary genetic data from the region. The community of Tixcacaltuyub has been collaborating with our research team for many years and their perspectives informed the development of this manuscript (Supplementary Methods: 'Community engagement activities'). Our analyses, comprising of ancient human genetic data analysis, bone collagen stable isotope analysis of carbon and nitrogen and radiocarbon dating (Supplementary Table 1) show that all *chultún*

subadults were male and that close relatives were present in the mass burial, including two sets of monozygotic twins. Stable isotope analysis indicates that related children consumed more similar diets and that overall the diet of Chichén Itzá children was comparable to that of other Classic period populations throughout the Maya Lowlands (Supplementary Table 2 and Supplementary Fig. 4). Genetic comparison to other ancient and present-day people shows long-term genetic continuity in the Maya region but indicates allele frequency shifts in immunity genes at the human leukocyte antigen (HLA) class II locus, specifically an increase in HLA-DR4 alleles which provide greater resistance to *Salmonella enterica* infection, the causative agent of an enteric fever previously identified in a colonial-era mass grave in Oaxaca, southern Mexico, which was associated with the 1545 *cocoliztli* pandemic[29].

Genome and immune-genes data generation

Bone samples from ancient individuals found in the *chultún* burial at Chichén Itzá (from here onwards referred to as YCH) were collected, processed and analysed according to protocols designed for ancient DNA (aDNA) work in dedicated facilities[30]. Because not all skeletal elements could be unequivocally assigned to a single individual, only left petrous bones were collected to avoid sampling individuals more than once. Radiocarbon dating ($n = 26$) showed that the *chultún* was in use for at least 500 years from the initial florescence of the site in the early seventh century AD to its height during the tenth century AD until the mid-twelfth century (Supplementary Fig. 1 and Supplementary Table 1).

Ancient DNA was successfully retrieved from all 64 YCH individuals. In addition, DNA was collected from blood samples of 68 modern-day inhabitants of the town Tixcacaltuyub, Yucatán, Mexico (hereafter TIX) to compare modern and ancient inhabitants of the region. The extracted genomic material was built into either uracil DNA glycosylase (UDG)-treated (for YCH) or non-UDG-treated genomic libraries (for TIX) and sequenced to a depth of about 5–11 million reads to assess DNA preservation and authenticity. We then built 11 single-stranded, UDG-treated libraries to further increase the analysable data for a subset of YCH individuals (Supplementary Methods: 'Preparation of single-stranded libraries'). We performed quality control assessment to ensure acceptable (less than 5%) contamination amounts with two methods implemented as part of the nf-core/Eager pipeline[31]. All TIX and 56 YCH individuals yielded sufficient human DNA for analysis (more than 0.1%) and after a reconditioning procedure we further enriched these DNA libraries for a panel of 1.2 million ancestry-informative single nucleotide polymorphisms (SNPs, '1,240 K'), the mitochondrial genome (mtDNA) and a panel of immune genes[32–34]. After sequencing the enriched genomic libraries, we obtained about 40 million reads per library; on these data we performed further quality control and conducted population genetics analyses and HLA typing.

Genetic kinship and twins in the *chultún*

Coverage comparisons of X and Y chromosome SNPs assigned all YCH subadult individuals as genetically male and confirmed the recorded sex for all TIX participants[35]. The pairwise mismatch rate analyses[36] (Fig. 1e and Supplementary Fig. 2) for YCH support the presence of two pairs of identical twins (YCH018–YCH019 and YCH033–YCH054) and nine other close relative pairs (YCH016–YCH017, YCH017–YCH018, YCH017–YCH019, YCH034–YCH041, YCH036–YCH038, YCH042–YCH049, YCH047–YCH057, YCH049–YCH057 and YCH059–YCH060). Overall, 25% ($n = 16$) of the analysed children in the ritual interment are closely related to another child within the *chultún*.

Isotopic patterns in the children's diets

The $\delta^{13}C$ and $\delta^{15}N$ measurements of bone collagen (Supplementary Table 1 and Supplementary Fig. 3; also Supplementary Methods: 'Stable isotopes analyses') provided $\delta^{13}C$ values between −13.9‰ and −7.6‰ (mean and s.d. = −9.9‰ ± 1.5‰) and $\delta^{15}N$ values ranging from 5.9‰ to 14.0‰ (mean and s.d. = 9.7‰ ± 1.5‰). Overall, these values are similar to those reported at other Classic period Maya Lowland sites (Supplementary Table 2 and Supplementary Fig. 4) and are consistent with dietary evidence in the Yucatán Peninsula[37,38] and at other Classic Maya sites[39–42]. It is possible that the diets of some individuals with higher $\delta^{15}N$ values (for example, YCH004, YCH008, YCH023, YCH039, YCH047 and YCH061) may have included aquatic resources[37] but could also indicate other dietary variations related to social status[43] or result from breast-feeding influence[44]. Without any more contextual information or local baseline isotopic data from associated fauna, it remains challenging to more precisely determine individual diets. Nevertheless, the comparison of our data with published results (more than 450 individuals studied from the Late Classic and Terminal Classic periods)[45–47] suggest that the 54 *chultún* individuals consumed significant amounts of C_4 terrestrial resources combined with varying amounts of terrestrial C_3 protein and freshwater and/or marine resources. This is consistent with previous archaeological investigations focusing on documented Classic Maya diets[37–39,41,42,48]. The isotopic values of related individuals fall close to each other, suggesting dietary similarity (Supplementary Fig. 5).

Genetic continuity in the Maya region

We performed a principal component analysis (PCA) based on worldwide populations and present-day individuals from the Americas[49] (Fig. 2). As expected for Mesoamerican populations, YCH cluster closely together on a worldwide PCA with unadmixed Indigenous American populations. Some individuals from TIX are shifted towards Europeans, suggesting genetic admixture (Fig. 2). When both YCH and TIX are projected onto present-day Indigenous American populations from North, Central and South America[50] they cluster with present-day Maya (Fig. 2a). Unsupervised admixture analyses using a subset of populations from Africa, Europe, Oceania and the Americas (Fig. 2b,c) showed no signs of admixture in the YCH individuals and low contributions of European and African ancestry for the TIX individuals, with some ($n = 18$) showing no indications of non-Indigenous American genetic contributions. It is of interest to note that a genetic component maximized in ancient populations from the Caribbean region is present in both ancient Mayans from Belize[49,51] and YCH but it is nearly absent in the genetic make-up of present-day Mayans[52] and TIX. Admixture with other populations from Mesoamerica (where the component has not yet been detected), or genetic drift, could explain this component fading away in present-day Mayans.

To test the genetic affinities to present-day and ancient American populations, we calculated an outgroup F_3-statistic of the form f_3(outgroup, target, test), using Mbuti from Sub-Saharan Africa[52,53] as an outgroup, a panel of Indigenous American populations as the target and either TIX or YCH as the test. The highest genetic similarities detected for both YCH and TIX (Fig. 2) included Central and South American groups. TIX share the highest drift with the ancient individuals from Chichén Itzá (Fig. 2). Among the ancient populations tested, a 9,300-year-old individual from the Mayahak Cab Pek site in the Maya Mountains in southern Belize[49] and other analysed individuals, also from Belize but from a more recent context[51], are genetically most similar to the ancient Chichén Itzá individuals, suggesting long-term genetic continuity in the Maya region. We tested if YCH and TIX are closer to each other than to other Indigenous American populations using an F_4-statistic of the form f_4 (Mbuti, TIX, target, YCH), using a subset of the highest F_3-scoring Indigenous American populations. Our results indicate that a few members of the target Indigenous American populations we tested are more closely related to TIX (Extended Data Fig. 1). This may suggest that, even though TIX are genetically related to the ancient inhabitants of Chichén Itzá, the genetically closest population to Chichén Itzá may no longer exist or has not been sampled yet.
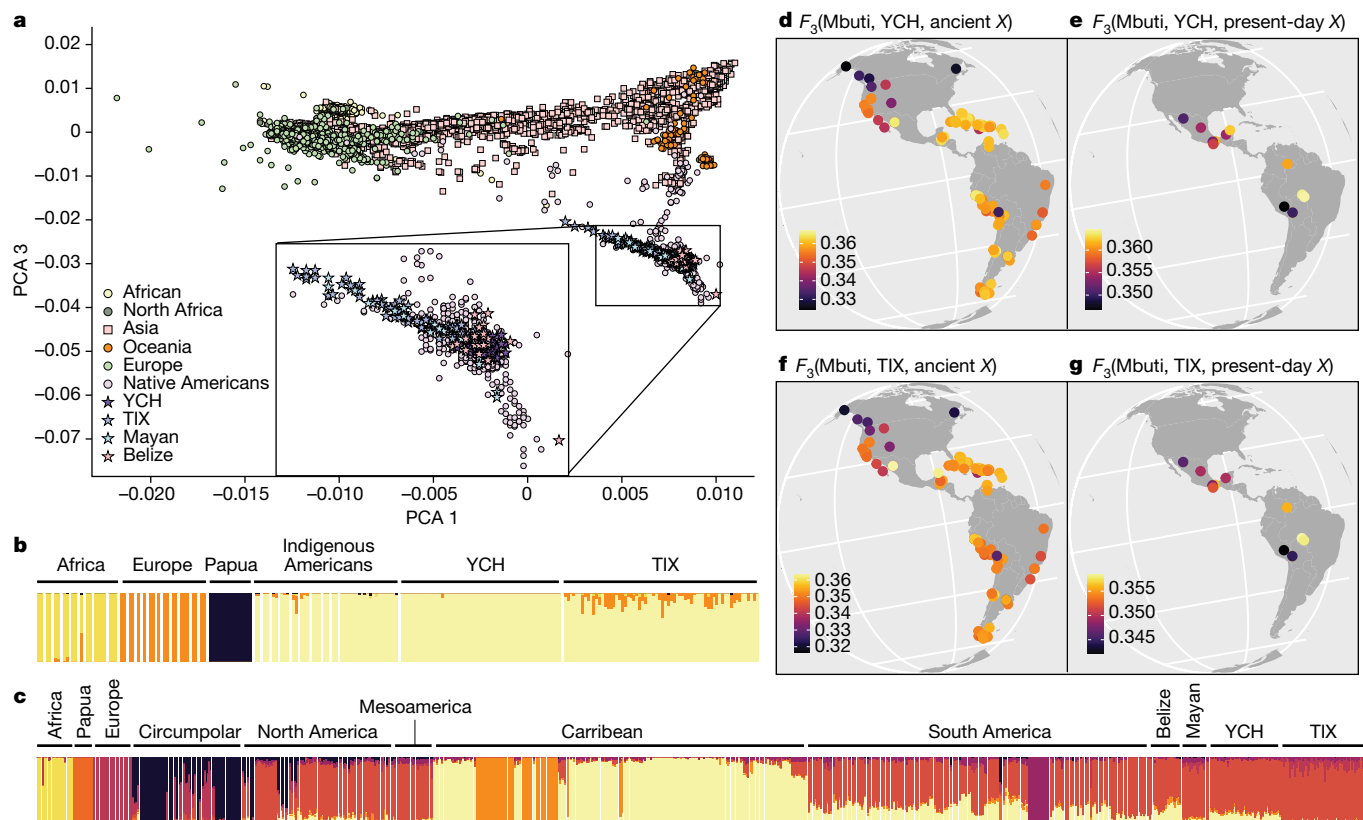
Fig. 2 | Genetic similarities between Chichén Itzá, Tixcacaltuyub and present-day and ancient American Indigenous groups. a, PCA showing ancient Chichén Itzá (YCH) individuals and present-day Tixcacaltuyub (TIX) in a worldwide PCA plot. **b,c**, Admixture analyses showing the clustering from $k = 4$ for YCH, TIX and different continental genetic sources (**b**) and the clustering for the newly produced samples at $k = 8$ (lowest cross-validation errors), with a more comprehensive list of Indigenous American populations (**c**). **d,e**, Outgroup F_3 statistics in the form f_3(Mbuti; YCH, X) where X are present-day groups from the Americas (**d**) and published ancient groups and individuals (**e**). High similarities are indicated by warmer colours; low similarities by darker colours. **f,g**, Outgroup F_3 statistics in the form f_3(Mbuti; TIX, X) where X are present-day groups from the Americas (**f**) and published ancient groups and individuals (**g**). High similarities are indicated by warmer colours, low similarities by darker colours. A list for the populations/individuals used in these analyses can be found in the Supplementary Information: source populations for the population genetics analyses.

We then applied qpWave[50], an ancestry composition modelling approach that evaluates the likelihood that two individuals/groups derive from the same ancestry source relative to a given set of population references. The resulting P values (Kaqchikel: 0.08, 'Maya' (from ref. 54): 0.755) suggest that YCH and present-day Mayan groups share the same ancestry. TIX can be modelled as a mixture of YCH, Spanish and Yoruba; the working model ($P = 0.11$) suggests a composition of 92% Indigenous American component, 7% of European genetic contribution and 0.03% African ancestry. Using a maximum-likelihood test for genetic continuity[55], we could formally test that TIX is a direct genetic descendant population from YCH (average genetic drift across individuals: about 0.5, P approximately 1.0 for each TIX individual). For 53 YCH and all 68 TIX individuals, the mtDNA haplogroup could be determined and the frequency of haplogroups (A, B, C and D) is almost identical between both groups. However, from the haplotypes, which represent a higher level of genetic resolution, it is clear that the diversity of mtDNA is higher in YCH than in TIX (Supplementary Fig. 6). The mtDNA haplogroups and haplotype lineages correspond to those previously reported for both ancient[56,57] and present-day[58,59] Maya. All Y chromosome (Y-Chr) haplotypes (Supplementary Fig. 6) recovered from the Chichén Itzá individuals ($n = 51$) are part of the Q family (prevalent among Indigenous Americans), whereas more than half of the TIX Y-Chr ($n = 19$) are European (47.37%) and Middle Eastern (5.26%), reflecting a strong sex bias in the admixture process during and after the colonial period, as has been previously described in other Latin American populations[60–63].

Genomics of metabolic pathways in Mayans

Using the SNP data generated for both populations, we calculated locus-specific branch lengths (LSBL)[64,65] to test for genome-wide selection for both YCH and TIX. We performed two LSBL comparisons: first, YCH versus Iberians from Spain (from the 1000 Genomes project[66]) and Han from China (from the 1000 Genomes project); and second, TIX versus Iberians and YCH to test for selection separate from YCH and Iberians. Among the top 0.5% annotated SNPs, we found 29 genes involved in lipid metabolism for YCH, including the previously reported fatty acid desaturases (*FAD*s) genes[67] and we found 20 genes for TIX, including FTO alpha-ketoglutarate dependent dioxygenase (*FTO*) and transcription factor 7 like 2 (*TCF7L2*), both of which have been associated with metabolic traits in Latin American and particularly Mayan, populations[68–70] (Supplementary Tables 3–5). The SNPs of certain genes, like those belonging to the adenylate cyclase family (*ADCY*), were also among the top 0.5% for TIX, which is consistent with previous reports[71], but not among YCH, which may point to differential selection before and after the colonial period. We then searched for enriched gene ontology pathways using GoWinda[72]. Although both groups exhibit enriched (adjusted $P < 0.05$) gene ontology terms associated with metabolic pathways (Supplementary Tables 6 and 7), YCH shows an increase in fertility-associated biological processes (such as oogenesis, steroid hormone mediated signalling pathway, ovulation cycle and oestrous cycle, among others), whereas cholesterol- and lipid-metabolic pathway terms (such as negative regulation of lipid
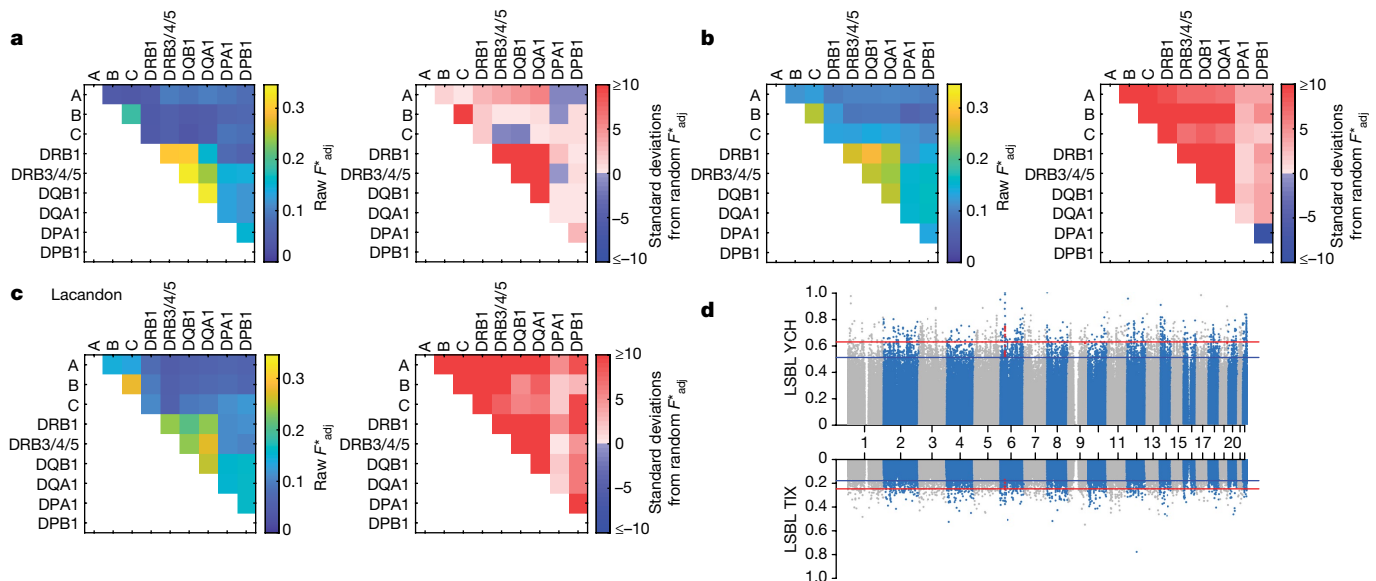
Fig. 3 | Evidence for selection in the HLA region. a–c, Pattern of non-overlap associations for the genes in the HLA region as measured by F^*_{adj} (heatmaps on the left) and the difference between those observed amounts of non-overlap and randomized allelic associations, in units of standard deviation (heatmaps on the right), for ancient Mayans from Chichén Itzá (YCH) (**a**), present-day Mayans from Tixcacaltuyub (TIX) (**b**) and present-day Lacandon Mayans from Chiapas (southeast Mexico)[76] (**c**). **d**, Manhattan plots of genome-wide LSBL values for YCH (top) and TIX (bottom). HLA loci are shown in red. Top 0.1% (red line) and 0.5% (blue line) of all statistics are indicated.

biosynthetic process, cholesterol homeostasis and sterol homeostasis) appear more prominently in TIX.

HLA genes point to shifts in immunity

For genes involved in immunity, we could detect 15 and 7 HLA region SNPs among the top 0.5% annotated SNPs for YCH and TIX, respectively, showing signs of positive selection (Fig. 3). None of the SNPs were shared by both YCH and TIX and none was found to be under selection in a previous study in ancient and modern Indigenous Americans from the Northwest Coast of North America[73]. The SNPs found on the YCH individuals are located in the *HLA-B, -DRB1, -DQA1, -DQA2, -S, -X, HLA-DOA* and *-DQB1* genes or nearby regions, whereas TIX SNPs are found in or around the *HLA-C, -DQA1, -DQA2* and *-DQB1* genes (Supplementary Table 8). Using a multilocus model of host–pathogen co-evolution with allele-specific adaptive immunity, it has been shown that, if selection from a pathogen maintains associations between host recognition loci (such as in the HLA system), alleles at those loci would not only be in linkage disequilibrium but might also exhibit non-overlapping associations[74]. For that reason, we analysed the pattern of non-overlapping associations[74,75] (Fig. 3 and Supplementary Figs. 7 and 8) to test for pathogen-driven selection on HLA associations in YCH, TIX and the previously analysed Lacandon Mayan from the highlands of Chiapas, southeast Mexico[76]. We measured the f^*_{adj} metric (a parameter used to rank the strength of non-overlapping associations; Supplementary Methods: 'Non-overlapping HLA associations')[74] between different pairs of HLA loci. We also measured the difference between our observed amounts of non-overlap and that which would be observed for randomized allelic associations, in units of standard deviations.

When compared with the ancient data, the present-day Maya data seem to have a higher degree of non-overlap (Fig. 3a–c), which would suggest that the analysed modern populations have experienced selection on the HLA locus, potentially due to pathogen exposure. When we compare the HLA allelic frequencies (Supplementary Tables 9–19) of YCH and TIX, we detect a statistically significant shift ($P_{corr} < 0.05$) in eight alleles: three HLA class I alleles and five HLA class II alleles, after

correcting for several comparisons. There is a significant decrease in the frequency of alleles HLA-B*40:02 (0.2447 versus 0.0821; $P_{corr} = 0.0296$), HLA-DQA1*03:03 (0.1277 versus 0.0224; $P_{corr} < 0.0218$) and HLA-DQB1*04:02 (0.1809 versus 0.0299; $P_{corr} = 0.0249$), whereas HLA-A*68:03 (0.0532 versus 0.2687; $P_{corr} = 0.0003$), HLA-B*39:05 (0.0532 versus 0.2687; $P_{corr} = 0.0437$), HLA-C*07:02 (0.2021 versus 0.3955; $P_{corr} = 0.0369$), HLA-DQB1*03:02 (0.4894 versus 0.7015; $P_{corr} = 0.0177$) and HLA-DRB1*04:07 (0.2340 versus 0.4627; $P_{corr} = 0.0114$) increased in frequency in TIX when compared to YCH. For TIX, we found that 88% of HLA haplotypes have been reported previously in Indigenous American populations, 10% of which are probably of European origin and 2% of which represent African haplotypes. All YCH HLA haplotypes are consistent with those found in Indigenous American populations, particularly among Maya[76–78].

The HLA class II region has been previously implicated in selection events before and after the contact with Europeans in the sixteenth century in the Americas[73] and in the resistance to *S. enterica* infection[79–81]. Hence, we were interested to test how the alleles for which there is a significant shift would react to peptides derived from *Salmonella*. To do so, we selected 18 *Salmonella* proteins for which there is evidence of immunogenicity in humans and presented in silico peptides derived from them to the HLA class II molecules found both in YCH and TIX (Supplementary Table 20 and Supplementary Methods: 'In silico binding prediction assays') using the NetMHCIIPan binding prediction method[82] implemented in the Immune Epitope Database (IEDB) Analysis Resource (http://tools.iedb.org/mhcii). Strong binding means the bound peptide is more likely to initiate an immune response against that peptide, whereas weaker binders are not as successful in eliciting an immune response. The best HLA-DR strong binders were HLA-DRB1*14:02, HLA-DRB1*04:07, HLA-DRB1*16:02 and HLA-DRB1*04:17; whereas the best strong binders for HLA-DQ were HLA-DQA1*05:01/DQB1*03:01, HLA-DQA1*05:05/DQB1*03:03 and HLA-DQA1*03:01/DQB1*03:03. The HLA molecules with the fewest strongly bound peptides are represented by HLA-DRB1*14:06, HLA-DRB1*04:04, HLA-DRB1*04:05 for HLA-DR and HLA-DQA1*03:03/DQB1*04:02, HLA-DQA1*04:01/DQB1*04:02 and HLA-DQA1*03:01/DQB1*03:02 for HLA-DQ. Overall, the lowest number of peptides, either strongly or weakly presented, by HLA class II

Article

molecules correspond to HLA-DR alleles DRB1*08:02, DRB1*04:04 and DRB1*14:06 and the same HLA-DQ molecules listed before. Broadly, the frequencies of HLA-DRB1*04:07 (strong binder) and HLA-DQB1*03:02 (weak binder, in linkage disequilibrium with HLA-DRB1*04:07) go up in TIX, whereas HLA-DQA1*03:03 and HLA-DQB1*04:02 go down in TIX and are weak binders.

Discussion

Archaeogenetics offers an opportunity to investigate aspects of past Maya ritual practice, biological kinship, sex and genetic history that can be difficult to infer using more conventional archaeological methods. Here we investigated the subadult remains of 64 individuals who were ritually interred over a period of 500 years in a Classic period *chultún* located near the Sacred Cenote at the ancient Maya city of Chichén Itzá. In contrast to the human remains in the Sacred Cenote, we found that all analysed subadults in the *chultún* were male, demonstrating a strong preference for the ritual sacrifice of male children in this context. Genetic analysis also showed the presence of related individuals within the *chultún*, including two sets of monozygotic twins and nine other close relative pairs. As such twins occur spontaneously in only 0.4% of the general population[83], the presence of two sets of identical twins in the *chultún* is much higher than would be expected by chance. Overall, 25% of the children had a close relative within the assemblage, suggesting that the sacrificed children may have been specifically selected for their close biological kinship. Moreover, this may underestimate the true number of relatives present in *chultún* as only 64 of the estimated 106 individuals in the *chultún* had a preserved petrous portion of the left temporal bone available for analysis. The further finding that the closely related children in each set seem to have consumed a similar diet and died at a similar age suggests that they have been sacrificed during the same ritual event as a pair or twin sacrifice.

Twins are especially auspicious in Mayan mythology and twin sacrifice is a central theme in the sacred K'iche' Mayan Book of Council, the *Popol Vuh*[84], a book whose antecedents can be traced to the Maya Preclassic period[85]. In the *Popol Vuh*, the twins Hun Hunapu and Vucub Hunahpu descend into the underworld and are sacrificed by the gods following defeat in a ballgame. The head of Hun Hunapu is then hung in a calabash tree, where it impregnates a maiden who gives birth to a second set of twins, Hunapu and Xbalanque. These twins, known as the Hero Twins, then go on to avenge their father and uncle by undergoing repeated cycles of sacrifice and resurrection to outwit the gods of the underworld. The Hero Twins and their adventures are amply represented in Classic Maya art[5] and given that subterranean structures were viewed as entrances to the underworld, the twin and relative sacrifices within the *chultún* at Chichén Itzá may recall rituals involving the Hero Twins.

In comparing the subadults in the *chultún* to other ancient and present-day populations in the Maya region, we find evidence of long-term genetic continuity, which also suggests that the sacrificed children and sibling pairs at Chichén Itzá were obtained from nearby ancient Maya communities. Among present-day individuals at TIX, we detect evidence of European and African admixture since the Contact period. Although ancestry contributions from non-Indigenous American sources are low at the genome-wide level, they are strongly asymmetrical with respect to uniparental markers. Whereas all mtDNA haplogroups are Indigenous American, more than half of the TIX Y-Chr are non-Indigenous American (mostly European and Middle Eastern/Mediterranean in origin, consistent with previous reports[60,61]), indicating that a strong male bias towards non-Indigenous American ancestries occurred during the Contact period.

The genetic similarity among Maya groups determined in the population genetic analyses (admixture profiles, F_3 and F_4 and genetic continuity test) allowed us to explore changes in genomic regions encoding functional variants to test for selection in ancient (YCH) and present-day (TIX) Maya. Our findings support previous hypotheses that both lipid metabolism[86,87] and fertility[86] are traits that have been recently selected for in Indigenous Americans, probably due to the strong bottlenecks and caloric restrictions experienced by these populations during the early colonial and settlement periods[86,88]. The observed standard deviation of the $\delta^{15}N$ values (s.d. = 1.5) obtained from the individuals analysed at Chichén Itzá is the highest of all the Late Classic and Terminal Classic Maya sites analysed so far (Supplementary Figs. 3 and 4 and Supplementary Table 2). The overall picture from the reconstruction of palaeodiet in the Maya region shows the consumption of significant amounts of C_4 foods, probably maize but with geographic variations reflected in microenvironmental differences in available foods and variability in trade networks[45,47]. To explain this variability, it is possible that a considerable portion of the individuals sacrificed at Chichén Itzá may be non-local Maya consuming slightly different diets. Alternatively, previous research indicates that the dietary patterns of the Classic Maya elite tend to exhibit greater variability than those of the general population over time[43,47,89]. This variability is also reflected in the standard deviations of $\delta^{13}C$ and $\delta^{15}N$ observed in other locations, such as Altun Ha or Baking Pot (Supplementary Fig. 4 and Supplementary Table 2). Therefore, the differences in protein intake observed in the analysed Chichén Itzá individuals studied could also indicate variations in social status. It also cannot necessarily be excluded that some of the observed $\delta^{15}N$ variability is a result of breastfeeding influence[90], as the sampled remains come from individuals estimated to be between 3 and 6 years of age. Thus, caution in interpreting specific dietary variations is warranted in the absence of other contextual information. Our data indicate that those individuals whose DNA showed a close genetic relationship had more similar $\delta^{13}C$ and $\delta^{15}N$ values, suggesting that they may have been raised in an extended family network that provided similar care and feeding (Supplementary Figs. 3 and 5 and Supplementary Table 1).

The values for the genetic drift obtained from the genetic continuity test mean that the TIX ancestors went through a serious population decline sometime in the last 1,000 years or so. Over the course of the sixteenth century, it has been demonstrated that wars, famines and epidemics resulted in a population decline, potentially as high as 90%, from 10–20 million Indigenous people living in Mexico at the time of European contact to only 2 million people by the end of the sixteenth century[91–96]. Infectious diseases, such as smallpox, measles, mumps, influenza, *tabardillo* or *matlalzahuatl* (typhus), typhoid and enteric fever, rubella, pertussis, *garrotillo* (severe diphtheria), endemic dysentery, tertiary fevers (malaria) and syphilis[29,97–101] are argued to have caused large-scale outbreaks in colonial Mesoamerica, significantly contributing to population decline and possibly causing selection at immune-related loci. The HLA class II region has been reported previously to have undergone selection events during the colonial period in the Americas[73]. Notably, we find that three of the alleles whose frequencies change when we compare YCH to TIX are part of the HLA class II region, a finding that is further supported by the SNPs found in the LSBL analyses. One of those alleles (HLA-DRB1*04:07) belongs to an allele group (HLA-DR4) that has been previously reported to be associated with resistance to enteric fever caused by *S. enterica* subspecies in South America[81] and East Asia[80]. Recently, an archaeogenetic study identified the presence of *S. enterica* sp. Paratyphi C in a mass burial associated with the AD 1545 *cocoliztli* pandemic, suggesting that it was at least one of the causative agents of this pandemic[29], which had the highest mortality of all recorded colonial-era epidemics. Genomic analysis of ancient *S. enterica* strains strongly supports an introduction to the Americas during the sixteenth century[102].

The observed increase of the HLA-DR4 allele group among present-day Maya and Mexicans in general[103], is consistent with selection caused by an epidemic event and subsequent sustained exposure to the pathogen. Further examination of the non-overlapping associations between HLA alleles likewise agrees with the present-day Maya

populations having undergone pathogen selection, which has driven their HLA associations to be less overlapping than for the YCH. The in silico binding prediction assays also point to HLA-DRB1*04:07 as a strong binder for *Salmonella*-derived peptides, whereas HLA-DQB1*04:02 and HLA-DQB1*03:03, significantly reduced in present-day Mayans, are weaker binders for the same peptides. Considered together, each line of evidence points to selection event(s) occurring in the HLA region in response to an epidemic event during the colonial period. Such selection would be expected in the face of the intensity of the 1545 *cocoliztli* pandemic and the high number of epidemic events documented in the Maya region since the beginning of the sixteenth century[104–106]. Although it cannot be ruled out that haplotypes with non-overlapping combinations of alleles survived a non-disease related bottleneck by chance, such a scenario would probably result in an increased frequency of alleles already present in the ancient population, which is not what we observe.

Our study shows an intimate portrait of Late and Terminal Classic Maya children at Chichén Itzá, suggesting that the genomic legacy of the ancient inhabitants of this site is still present in communities inhabiting the region surrounding this ancient city. The discovery of two sets of identical twins, as well as other close relatives, in a ritual mass burial of male children suggests that young boys may have been selected for sacrifice because of their biological kinship and the importance of twins in Maya mythology. We show that, at a genome-wide level, the present-day Maya of Tixcacaltuyub exhibit genetic continuity with the ancient Maya who once inhabited Chichén Itzá and we demonstrate through several lines of evidence the involvement of the HLA region in a pathogen-driven selection event(s) probably caused by infectious diseases brought into the Americas by Europeans during the colonial period.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07509-7.

1. Price, T. D., Tiesler, V. & Freiwald, C. Place of origin of the sacrificial victims in the Sacred Cenote, Chichén Itzá, Mexico. *Am. J. Phys. Anthropol.* **170**, 98–115 (2019).
2. Kristan-Graham, C. et al. *Twin Tollans: Chichén Itzá, Tula and the Epiclassic to Early Postclassic Mesoamerican World* (Dumbarton Oaks, 2007).
3. Beck, L. A. & Sievert, A. K. in *Interacting with the Dead: Perspectives on Mortuary Archaeology for the New Millennium* (eds Rakita, G. et al.) 290–304 (Univ. Press Florida, 2005).
4. Márquez Morfín, L. & Schmidt, P. in *Investigaciones Recientes en el área Maya.* Vol. II. *Memorias de la XVII Mesa Redonda (1981)* 89–104 (Sociedad Mexicana de Antropología, 1984).
5. Coe, M. D. in *The Maya Vase Book: A Corpus of Rollout Photographs of Maya Vases* (eds Kerr, B. & Kerr, J.) 161–184 (Kerr Associates, 1989).
6. Cobos, R. & Winemiller, T. L. The Late and Terminal Classic-period causeway systems of Chichen Itza, Yucatan, Mexico. *Anc. Mesoam.* **12**, 283–291 (2001).
7. Tiesler, V. in *Stone Houses and Earth Lords: Maya Religion in the Cave Context* (eds Prufer, K. M. & Brady, J. E.) 341–364 (Univ. Press Colorado, 2005).
8. Miller, V. E. in *New Perspectives on Human Sacrifice and Ritual Body Treatments in Ancient Maya Society* (eds Tiesler, V. & Cucina, A.) 165–189 (Springer New York, 2007).
9. Graña-Behrens, D., Prager, C. & Wagner, E. The hieroglyphic inscription of the 'High Priest's Grave' at Chichén Itzá, Yucatán, Mexico. *Mexicon* **21**, 61–66 (1999).
10. Dahlin, B. H. Climate change and the end of the Classic period in Yucatán. Resolving a paradox. *Anc. Mesoam.* **13**, 327–340 (2002).
11. Roys, R. L. *The Book of Chilam Balam of Chumayel* (Univ. Oklahoma Press, 1967).
12. Tozzer, A. M. *Chichen Itza and its Cenote of Sacrifice: A Comparative Study of Contemporaneous Maya and Toltec* (Harvard Univ., 1957).
13. Anda, G. D., Tiesler, V. & Zabala, P. in *Los Investigadores de la Cultura Maya* Vol. 12, 376–386 (Universidad Autónoma de Campeche, 2004).
14. Márquez Morfín, L. in *Los Niños Como Actores Sociales Ignorados. Levantando el Velo, una Mirada al Pasado* (ed. Márquez Morfín, L.) 253–282 (ENAN-INAH, 2006).
15. Baudez, C.-F. & Latsanopoulos, N. Political structure, military training and ideology at Chichen Itza. *Anc. Mesoam.* **21**, 1–20 (2010).
16. Cagnato, C. Underground pits (chultunes) in the southern Maya lowlands: excavation results from Classic period Maya sites in northwestern Petén. *Anci. Mesoam.* **28**, 75–94 (2017).
17. Puleston, D. E. An experimental approach to the function of Classic Maya chultuns. *Am. Antiq.* **36**, 322–335 (1971).
18. Brady, J. & Layco, W. Maya cultural landscapes and the subterranean: assessing a century of chultun research. *Int. J. Archaeol.* **6**, 46–55 (2018).
19. Prout, M. G. & Brady, J. E. Paleodemographics of child sacrifice at Midnight Terror Cave: reformulating the emphasis of Maya sacrificial practices. *Archaeol. Discov.* **6**, 1–20 (2018).
20. Ardren, T. Empowered children in Classic Maya sacrificial rites. *Child. Past* **4**, 133–145 (2011).
21. MacLeod, B. & Puleston, D. E. in *Tercera Mesa Redonda de Palenque* (eds Robinson, M. G. & Jeffers, D. C.) 71–77 (Herald Printers, 1978).
22. Moyes, H. & Brady, J. E. in *Sacred Darkness: A Global Perspective on the Ritual Use of Caves* (ed. Moyes, H.) 151–170 (Univ. Press Colorado, 2014).
23. del Castillo Chávez, O. & Williams-Beck, L. *Rituales k'ex al Dios Del Maíz En Chichén Itzá y Mayapán: Una Tradición Ritual Del Clásico al Postclásico* (Centro Regional Yucatán del INAH, 2020).
24. Willard, T. A. *The City of the Sacred Well* (Century, 1926).
25. Arnold, C. & Frost, F. J. T. *The American Egypt: A Record of Travel in Yucatan* (Hutchinson, 1909).
26. Cucina, A. & Tiesler, V. in *The Bioarchaeology of Space and Place* (ed. Wrobel, G. D.) 225–254 (Springer, 2014).
27. de Landa, D., Tozzer, A. M. & Bowditch, C. P. *Landa's Relación de Las Cosas de Yucatan: A Translation* (Harvard Univ., 1941).
28. Clendinnen, I. *Ambivalent Conquests. Maya and Spaniard in Yucatan, 1517–1570* (Cambridge Univ. Press, 2003).
29. Vågene, Å. J. et al. *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat. Ecol. Evol.* **2**, 520–528 (2018).
30. Llamas, B. et al. From the field to the laboratory: controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *Sci. Technol. Archaeol. Res.* **3**, 1–14 (2017).
31. Fellows Yates, J. A. et al. Reproducible, portable and efficient ancient genome reconstruction with nf-core/eager. *PeerJ* **9**, e10947 (2021).
32. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
33. Immel, A. et al. Analysis of genomic DNA from medieval plague victims suggests long-term effect of *Yersinia pestis* on human immunity genes. *Mol. Biol. Evol.* **38**, 4059–4076 (2021).
34. Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
35. Lamnidis, T. C. et al. Ancient Fennoscandian genomes reveal origin and spread of Siberian ancestry in Europe. *Nat. Commun.* **9**, 5018 (2018).
36. Lamnidis, T. C. pMMRCalculator. GitHub https://github.com/TCLamnidis/pMMRCalculator (2020).
37. Götz, C. M. La alimentación de los mayas prehispánicos vista desde la zooarqueología. *An. Antropol.* **48**, 167–199 (2014).
38. Herrera Flores, D. A. & Markus Götz, C. La alimentación de los antiguos mayas de la península de Yucatán: consideraciones sobre la identidad y la cuisine en la época prehispánica. *Estud. Cult. Maya* **XLIII**, 69–98 (2014).
39. Kennett, D. J. et al. Early isotopic evidence for maize as a staple grain in the Americas. *Sci. Adv.* **6**, eaba3245 (2020).
40. Staller, J. E., Tykot, R. H. & Benz, B. F. in *Histories of Maize in Mesoamerica* (eds Staller, J. et al.) Ch. 11 (Routledge, 2010).
41. Fernández Souza, L., Toscano, L. & Zimmermann, M. *De Maíz y de Cacao: Aproximaciones a la Cocina de las Elites Mayas en Tiempos Prehispánicos. In: Hernández Álvarez, H., et al. Estética y Poder en la Ciencia y la Tecnología, Mérida* pp 107–130 (UADY, 2017).
42. Zimmermann, M. Starch grain extraction in lime-plastered archaeological floors. *Sci. Technol. Archaeol. Res.* **7**, 31–42 (2021).
43. Somerville, A. D., Fauvelle, M. & Froehle, A. W. Applying new approaches to modeling diet and status: isotopic evidence for commoner resiliency and elite variability in the Classic Maya lowlands. *J. Archaeol. Sci.* **40**, 1539–1553 (2013).
44. White, C. D. & Schwarcz, H. P. Ancient Maya diet: as inferred from isotopic and elemental analysis of human bone. *J. Archaeol. Sci.* **16**, 451–474 (1989).
45. Wright, L. E. & White, C. D. Human biology in the Classic Maya collapse: evidence from paleopathology and paleodiet. *J. World Prehist.* **10**, 147–198 (1996).
46. Tykot, R. H. in *Archaeological Chemistry* Vol. 831 (ed. Jakes, K.) 214–230 (American Chemical Society, 2002).
47. Scherer, A. K. Bioarchaeology and the skeletons of the Pre-Columbian Maya. *J. Archaeol. Res.* **25**, 133–184 (2017).
48. Mansell, E. B., Tykot, R. H., Freidel, D. A., Dahlin, B. H. & Ardren, T. in *Histories of Maize: Multidisciplinary Approaches to the Prehistory, Linguistics, Biogeography, Domestication and Evolution of Maize* (eds Staller, J. et al.) 173–185 (Elsevier, 2010).
49. Posth, C. et al. Reconstructing the deep population history of Central and South America. *Cell* **175**, 1185–1197 (2018).
50. Reich, D. et al. Reconstructing Native American population history. *Nature* **488**, 370–374 (2012).
51. Kennett, D. J. et al. South-to-north migration preceded the advent of intensive farming in the Maya region. *Nat. Commun.* **13**, 1530 (2022).
52. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
53. Bergström, A. et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* https://doi.org/10.1126/science.aay5012 (2020).
54. Mallick, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).
55. Schraiber, J. G. Assessing the relationship of ancient and modern populations. *Genetics* **208**, 383–398 (2018).
56. González-Oliver, A., Márquez-Morfín, L., Jiménez, J. C. & Torre-Blanco, A. Founding Amerindian mitochondrial DNA lineages in ancient Maya from Xcaret, Quintana Roo. *Am. J. Phys. Anthropol.* **116**, 230–235 (2001).
57. Ochoa-Lugo, M. I. et al. Genetic affiliation of pre-Hispanic and contemporary Mayas through maternal lineage. *Hum. Biol.* **88**, 136–167 (2016).
58. Mizuno, F. et al. Characterization of complete mitochondrial genomes of indigenous Mayans in Mexico. *Ann. Hum. Biol.* **44**, 652–658 (2017).

59. Bodner, M. et al. The mitochondrial DNA landscape of modern Mexico. *Genes* **12**, 1453 (2021).

60. Ruiz-Linares, A. et al. Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet.* **10**, e1004572 (2014).

61. Chacón-Duque, J. C. et al. Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. *Nat. Commun.* **9**, 5388 (2018).

62. Ongaro, L. et al. Evaluating the impact of sex-biased genetic admixture in the Americas through the analysis of haplotype data. *Genes* **12**, 1580 (2021).

63. Ongaro, L. et al. The genomic impact of European colonization of the Americas. *Curr. Biol.* **29**, 3974–3986 (2019).

64. Shriver, M. D. et al. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. *Hum. Genom.* **1**, 274 (2004).

65. Duforet-Frebourg, N., Luu, K., Laval, G., Bazin, E. & Blum, M. G. B. Detecting genomic signatures of natural selection with principal component analysis: application to the 1000 Genomes data. *Mol. Biol. Evol.* **33**, 1082–1093 (2016).

66. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

67. Amorim, E. et al. Genetic signature of natural selection in first Americans. *Proc. Natl Acad. Sci. USA* **114**, 2195–2199 (2017).

68. Villalobos-Comparán, M. et al. The FTO gene is associated with adulthood obesity in the Mexican population. *Obesity* **16**, 2296–2301 (2008).

69. Ortega, P. E. N. et al. Association of rs9939609-FTO with metabolic syndrome components among women from Mayan communities of Chiapas, Mexico. *J. Physiol. Anthropol.* **40**, 11 (2021).

70. Lara-Riegos, J. C. et al. Diabetes susceptibility in Mayas: evidence for the involvement of polymorphisms in HHEX, HNF4α, KCNJ11, PPARγ, CDKN2A/2B, SLC30A8, CDC123/CAMK1D, TCF7L2, ABCA1 and SLC16A11 genes. *Gene* **565**, 68–75 (2015).

71. Ojeda-Granados, C. et al. Dietary, cultural and pathogens-related selective pressures shaped differential adaptive evolution among Native Mexican populations. *Mol. Biol. Evol.* **39**, msab290 (2022).

72. Kofler, R. & Schlötterer, C. Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. *Bioinformatics* **28**, 2084–2085 (2012).

73. Lindo, J. et al. A time transect of exomes from a Native American population before and after European contact. *Nat. Commun.* **7**, 13175 (2016).

74. Penman, B. S., Ashby, B., Buckee, C. O. & Gupta, S. Pathogen selection drives nonoverlapping associations between HLA loci. *Proc. Natl Acad. Sci. USA* **110**, 19645–19650 (2013).

75. Buckee, C. O., Gupta, S., Kriz, P., Maiden, M. C. J. & Jolley, K. A. Long-term evolution of antigen repertoires among carried meningococci. *Proc. R. Soc. B* **277**, 1635–1641 (2010).

76. Barquera, R. et al. Diversity of HLA Class I and Class II blocks and conserved extended haplotypes in Lacandon Mayans. *Sci. Rep.* **10**, 3248 (2020).

77. Gonzalez-Galarza, F. F. et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res.* **48**, D783–D788 (2020).

78. Gómez-Casado, E. et al. Origin of Mayans according to HLA genes and the uniqueness of Amerindians. *Tissue Antigens* **61**, 425–436 (2003).

79. Dunstan, S. J. et al. Genes of the class II and class III major histocompatibility complex are associated with typhoid fever in Vietnam. *J. Infect. Dis.* **183**, 261–268 (2001).

80. Dunstan, S. J. et al. Variation at HLA-DRB1 is associated with resistance to enteric fever. *Nat. Genet.* **46**, 1333–1336 (2014).

81. de Vries, R. R. P., Meera Khan, P., Bernini, L. F., van Loghem, E. & van Rood, J. J. Genetic control of survival in epidemics. *J. Immunogenet.* **6**, 271–287 (1979).

82. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454 (2020).

83. Fuchs, K. M. & D'Alton, M. E. in *Obstetric Imaging: Fetal Diagnosis and Care* (eds Copel, J. A. et al.) 639–641 (Elsevier, 2018).

84. Christenson, A. J. *Popol Vuh. Sacred Book of the Quiché Maya People* (Transl.) (Univ. Oklahoma Press, 2007).

85. Saturno, W. A., Stuart, D. & Taube, K. in *XVIII Simposio de Investigaciones Arqueológicas en Guatemala* (eds LaPorte, J. P. et al.) Ch. 60 (Museo Nacional de Guatemala, 2004).

86. Acuña-Alonzo, V. et al. A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. *Hum. Mol. Genet.* **19**, 2877–2885 (2010).

87. León-Mimila, P. et al. Genome-wide association study identifies a functional SIDT2 variant associated with HDL-C (high-density lipoprotein cholesterol) levels and premature coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.* **41**, 2494–2508 (2021).

88. Fritz, J., Lopez-Ridaura, R., Choudhry, S., Razo, C. & Lamadrid-Figueroa, H. The association of Native American genetic ancestry and high-density lipoprotein cholesterol: a representative study of a highly admixed population. *Am. J. Hum. Biol.* **32**, e23426 (2020).

89. Scherer, A. K., Wright, L. E. & Yoder, C. J. Bioarchaeological evidence for social and temporal differences in diet at Piedras Negras, Guatemala. *Latin Am. Antiq.* **18**, 85–104 (2007).

90. Wright, L. E. & Schwarcz, H. P. Stable carbon and oxygen isotopes in human tooth enamel: identifying breastfeeding and weaning in prehistory. *Am. J. Phys. Anthropol.* **106**, 1–18 (1998).

91. Acuña-Soto, R., Stahle, D. W., Therrell, M. D., Griffin, R. D. & Cleaveland, M. K. When half of the population died: the epidemic of hemorrhagic fevers of 1576 in Mexico. *FEMS Microbiol. Lett.* **240**, 1–5 (2004).

92. Cook, S. F. & Simpson, L. B. *The Population of Central Mexico in the Sixteenth Century* Vol. 31 (Univ. California Press, 1948).

93. Marr, J. S. & Kiracofe, J. B. Was the huey cocoliztli a haemorrhagic fever? *Med. Hist.* **44**, 341–362 (2000).

94. Cook, N. *Born to Die: Disease and New World Conquest, 1492–1650* (Cambridge Univ. Press, 1998).

95. Cook, S. F. & Borah, W. *The Indian Population of Central Mexico, 1531–1610* (Univ. California Press, 1960).

96. Zambardino, R. A. Mexico's population in the sixteenth century: demographic anomaly or mathematical illusion? *J. Interdiscipl. Hist.* **11**, 1–27 (1980).

97. Crosby, A. W. Virgin soil epidemics as a factor in the aboriginal depopulation in America. *William Mary Q.* **33**, 289–299 (1976).

98. Crosby Jr, A. W. *The Columbian Exchange: Biological and Cultural Consequences of 1492* (Greenwood, 1972).

99. Real Academia de la Historia & del Paso & Troncoso, F. *Relaciones Geograficas de La Diocesis de Mexico: Manuscritos de La Real Academia de La Historia de Madrid y Del Archivo de Indias En Sevilla. Anos 1579–1582* (Est. tipografico 'Sucesores de Rivadeneyra,' Sevilla, 1905).

100. Viesca, C. in *Ensayos Sobre la Historia de las Epidemias en México* (eds Florescano, E. & Elsa, M.) 157–165 (Instituto Mexicano del Seguro Social, 1982).

101. Márquez Morfín, L. La sífilis y su carácter endémico en la ciudad de México. *Hist. Mex.* **64**, 1099–1161 (2015).

102. Key, F. M. et al. Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization process. *Nat. Ecol. Evol.* **4**, 324–333 (2020).

103. Barquera, R. et al. The immunogenetic diversity of the HLA system in Mexico correlates with underlying population genetic structure. *Hum. Immunol.* **81**, 461–474 (2020).

104. de Buenaventura, J., Guerrero, G., Solís Robleda, G. & Bracamonte y Sosa, P. *Historias de la Conquista del Mayab, 1511–1697* (Universidad Autónoma de Yucatán, Facultad de Ciencias Antropológicas, 1994).

105. Góngora-Biachi, R. A. La fiebre amarilla en Yucatán durante las épocas precolombina y colonial. *Rev. Biomed.* **11**, 301–307 (2000).

106. Patch, R. W. Sacraments and disease in Mérida, Yucatán, Mexico, 1648–1727. *Historian* **58**, 731–743 (1996).

Methods

Archaeological and geographical context

Chichén Itzá was one of the largest and most influential Maya cities of the Terminal Classic Period and today stands as an iconic archaeological site among the most representative of pre-Hispanic monumental architecture in the Americas (Fig. 1). The architecture of the site consists of several different styles, with some of the structures resembling those found in Central Mexico (Teotihuacan, Tula) and in the Puuc and Chenes regions (northeastern and central Yucatan, respectively) of the Maya lowlands[107]. The Maya Puuc style represents the second phase of the Florescent period in the Yucatan Peninsula (AD 750–800)[107,108]. Ceramic types and epigraphic and radiometric data suggest that the city emerged as a chief regional political centre in the first half of the nineth century and stayed as such until probably the first half of the eleventh century[108,109]. The ancient samples used in this study belonged to individuals whose skeletons were recovered from the archaeological excavations of a *chultún* (cistern) connected to a natural cave within the archaeological site of Chichén Itzá, Yucatán, Mexico. *Chultúns* were mostly used from the Middle Preclassic to the Late Classic-Postclassic periods (AD 1000–1600) in Mesoamerica from the northern Maya lowlands of the Yucatan peninsula to modern-day Belize and Guatemala in the southern Maya lowlands[16,110]. The sampled skeletons, which have been studied using anthropometry and osteology guidelines at the Centro INAH Yucatán, were recovered between April and June 1967, during the construction of a new runway close to *El Castillo* ('The Castle', the main basament of the archaeological site and a building with special calendrical significance[4,111]) and 300 m northeast to the Sacred Cenote of Chichén Itzá. The *chultún* had stones that had been worked in the style that characterized the Florescent period, which would indicate that it had been built around the Late or Terminal Classic periods of Chichén Itzá[4,110]. The chamber and the adjacent cave contained many skeletal remains, some of them in anatomical position. The many skeletons and the position in which they were found, as well as them being covered by mostly undisturbed bark and limestone powder[4], are mostly consistent with a mixed secondary/primary burial with periodic clearance, which points to the burial having a cultural motivation, probably as an offering[4] related to an important and recurring ceremony[4,14]. There was a minimum of 106 skeletons of infants and young children (according to the number of temporal bones) without any apparent ordering or anatomical context, except for a group of skulls and skullcaps aligned from north to south from the wall separating the chamber and the cave to the southern limit of the cave[14]. Many ceramic and clay objects, as well as animal bone remains, were also found in the cave[4]. The ceramic styles and the archaeological context indicate that the *chultún* was used during the Florescent period, contemporary to the use of another ritual cave near Chichén Itzá known as Balankanché (AD 860 ± 130)[4,112]. A radiocarbon test on a sample of the bark covering the bones from the *chultún* burial has also suggested a use date of around AD 920 ± 60 (ref. 107). However, our ^{14}C dates (Supplementary Fig. 2) suggest that the ritual burial was performed both before (around AD 600) and after (around AD 1100) the Florescent period. According to the age determination by tooth eruption method[113], most (about 50%) of the subadult individuals were found to belong to the age group of 3–6 years old[4]. The fact that the age of the children differs from the age group normally related to death in children associated with infectious diseases according to palaeodemography[19,113–115] further supports the interpretation of this burial as a sacrificial offering[4].

The biological samples from present-day Maya were obtained from individuals of self-reported Native American Maya ancestry from Tixcacaltuyub. The village of Tixcacaltuyub, part of the municipality of Yaxcabá, is located 90 km southeast of Merida, 16 km off Sotuta and is located 55 km southeast of Chichén Itzá (Fig. 1). It has around 2,100 inhabitants. Ethical approval for the collection of blood samples from individuals from Tixcacaltuyub, Yucatán, Mexico, was granted by the Committee of Ethics and Research, Autonomous University of Yucatán (UADY), Mexico. All subjects were informed about the objectives and methods used and signed an informed consent form. To avoid potential bias, participants with any HLA-associated clinical condition or cancer[116,117] were excluded.

Inclusion and ethics

The community of Tixcacaltuyub self-identifies as a Mayan community and has been in a years-long cooperative relationship with the Chemistry and Nursing Faculties of UADY, Mérida, Yucatán, following projects investigating the relationship between health and lifestyle in the community. J.C.L.R., M.E.M.M. and J.C.T.R., together with C. Tzec-Puch, have been working in collaboration with the community to communicate results of clinical tests and to develop health interventions in Tixcacaltuyub. As a result of these interventions, healthcare barriers and opportunities have been identified and work is now underway with the active participation of the community to implement a co-responsible model of healthcare. In April 2023, J.C.L.R. and R.B. visited Tixcacaltuyub to hold meetings with participants and elementary, high school and UADY students to return the results of the genetic findings and to collect community feedback, including how they reconcile these results with their own views. Feedback from these engagements has been incorporated into the final manuscript (Supplementary Methods: 'Community engagement activities'). As part of our strategies for outreach and making our results available to a broader audience, we included a non-peer reviewed, Spanish translation of the main manuscript (Supplementary Note 1).

Individuals and samples

We acknowledge the Council of Archaeology, National Institute of Anthropology and History (INAH, Mexico City, Mexico) for the permit granted to analyse the individuals from the Chichén Itzá Chultun collection (project: *Caracterización genética de una muestra de población de Chichén Itzá para el periodo Clásico Tardío a partir del análisis del ADN*; official notice no. 401.1 S.3-2017/482; 10.03.2017). We analysed the material recovered from the left petrous bones of 64 individuals (YCH) from Chichén Itzá. The bone specimens were collected under controlled conditions in the Osteology Laboratory of the Centro INAH Yucatán with a protocol devoted to minimize contamination. Photographic records of the samples were maintained throughout the whole procedure. The present-day DNA was extracted from blood samples obtained from 68 unrelated adult individuals (TIX) from Tixcacaltuyub, Yucatán, Mexico under a protocol approved by the Committee of Ethics and Research, UADY, granting permit to collect blood samples and carry out analyses on the genetic material obtained from such samples (project: *Bienestar Comunitario: Proyecto de capacitación para la autogestión de la salud de personas con DT2 y sus familias, en la comunidad de Tixcacaltuyub y Yaxcabá*; official notice no. F-FENC-SAC-14/REV: 04; registry no. 09/17) and performed according to the requisites of the Helsinki Declaration (2008) and the General Health Law of Mexico. For YCH, bone powder (about 90 mg) was obtained from the densest part of each petrous bone in facilities dedicated to ancient DNA protocols and after a bleach/rinse and ultraviolet decontamination protocol[30]. For TIX, 5 ml of peripheral blood were obtained by venepuncture after each participant was informed about the procedure, the aim of the protocol and the potential risks, and after signing a consent letter. For TIX, the DNA was extracted in a dedicated laboratory at UADY. Their samples were anonymized from this point onwards.

[14]C dating

According to the report issued by The Curt-Engelhorn-Centre for Archaeometry (Mannheim, Germany), the portions of the petrous bone of 26 YCH samples were pretreated and analysed using a standardized procedure. Collagen was extracted from the bone samples (approximately 1 g, using a modified version of the Longin method[118]), purified

Article

by ultrafiltration (fraction greater than 30 kD) and freeze-dried. The collagen was then combusted to CO_2 in an elemental analyser. The CO_2 was then converted catalytically to graphite and analysed using a MICADAS-type AMS system (Ionplus AG). The isotopic ratios $^{14}C/^{12}C$ and $^{13}C/^{12}C$ of samples, calibration standard (oxalic acid II), blanks and control standards were measured simultaneously in the AMS system. ^{14}C-ages are normalized to $\delta^{13}C = -25‰$ (ref. 119) with a typical uncertainty of 2‰ and calibrated using the dataset IntCal20 and the software Oxcal (v.4.3.2)[120,121].

Stable isotope analyses

Direct insight into dietary trends among past populations enables the investigation of connections between diet and social status, cultural customs linked to food, environmental impacts on subsistence and perhaps even individual mobility[122,123]. Measurements of bone collagen $\delta^{13}C$ and $\delta^{15}N$ disproportionately reflect the protein component of an individual's diet during the period of tissue formation and, to a lesser extent, the lipid and carbohydrate sources[124]. To reconstruct the diet of the subadult individuals from Chichén Itzá studied, we analysed bone collagen from temporal and petrous bones (Supplementary Fig. 4). Collagen was extracted using a standard procedure[125] (Supplementary Methods). The atomic C:N ratio along with the collagen yields were used to determine the quality of collagen preservation. Collagen yields of more than 1% in weight are considered acceptable for carbon and nitrogen values[126], whereas the C:N ratio should range from 2.9 to 3.6 for archaeological samples[127] (Supplementary Table 1).

DNA extraction

For YCH, the bone powder was decalcified and proteins were digested by an overnight incubation (more than 16 h) at 37 °C in a buffer containing EDTA and Proteinase K[30]. The DNA was purified from the supernatant by a silica column-based method using a silica column for high volumes assay (High Pure Viral Nucleic Acid Large Volume Kit, Roche Molecular Systems). DNA was eluted in 100 μl of TET (10 mM Tris, 1 mM EDTA and 0.05% Tween) and frozen at −20 °C until library preparation[128]. For the contemporary participants, 5 ml of peripheral blood were collected in BD Vacutainer blood collection tubes containing K_2-EDTA (Becton, Dickinson and Company) and the DNA was extracted using the Quick-DNA Miniprep Plus kit (Zymo Research Corporation) following the developers' instructions. Because our protocols are optimized for short-length aDNA and to avoid potential bias through laboratory methods, we sheared the DNA extracted from modern individuals using ultrasonic DNA shearing to an average length comparable to that of aDNA. Therefore, 50 μl of a 50 ng μl^{-1} dilution of each of the modern DNA samples were sheared to an average fragment length of 150 base pairs using a Covaris M220 Focused ultrasonicator (Covaris).

Library preparation

We built non-UDG-treated libraries using 15 μl of each YCH extract to assess the authenticity of the extracted DNA after obtaining the characteristic damage plots associated with aDNA[129]. We then used 20 μl of each YCH extract to build UDG-half libraries with Illumina-specific adaptors following a modified double-stranded library preparation protocol as previously described[130]. We built non-UDG-treated libraries for TIX extracts using 20 μl of each modern DNA extract. Both YCH and TIX libraries were quantified using quantitative PCR (qPCR) with the IS7 and IS8 primers in a quantification assay using a DyNAmo SYBR Green qPCR Kit (Thermo Fisher Scientific) on the LightCycler 96 platform (Roche Diagnostics).

For DNA extracts from the individuals for which the data obtained by sequencing the UDG-treated libraries was not enough, we used a new library protocol that uses directional splinted ligation of Illumina P5 and P7 adaptors to convert natively single-stranded DNA and heat-denatured double-stranded DNA into sequencing libraries in a single enzymatic reaction using the fully automated version of the protocol[131–133]. We followed the protocol as described by the authors using 25 μl of DNA extract for each individual (Supplementary Methods: 'Preparation of single-stranded libraries').

Each library was identified with the respective pair of indexes in double-100 μl reactions using PfuTurbo DNA Polymerase (Agilent Technologies) for the double-stranded libraries and AccuPrime Pfx DNA Polymerase (Thermo Fisher Scientific). The indexed products for each library were pooled, purified over silica columns using the MinElute PCR Purification Kit (QIAGEN N.V.), eluted in 44 μl of TET and again qPCR quantified, now using the IS5 and IS6 primers. Conditioning for sequencing included the amplification of the purified product in 4 × 100 μl reactions using Herculase II Fusion DNA Polymerase (Agilent Technologies) following the manufacturer's specifications with 0.3 μM of each IS5/IS6 primer, following a purification over silica columns also using the MinElute PCR Purification Kit and elution in a final volume of 22 μl of TET. A total of 2 ml of the conditioned product were diluted 1:10 and quantified using the Agilent 2100 Bioanalyzer DNA 1000 protocol (Agilent Technologies). Independent, equimolar (10 mM final concentration) pools of YCH libraries, TIX libraries and extraction and libraries blanks were then prepared for shotgun sequencing on the Illumina HiSeq 4000 Systems platform (Illumina)[134,135]. YCH libraries were sequenced to 5 million reads depth, whereas TIX libraries were sequenced to 11 million reads depth and then all of them were analysed to obtain basic quality control parameters using nf-core/eager v.2.3.4 (ref. 31).

Whole-genome and immune-genes captures

Using an in-solution capture approach based on modified immortalized probe sequences[136], target immunity genes sequences, mtDNA, Y chromosome or a panel of 1,237,207 SNPs were enriched from the total DNA in the sequencing libraries[32,33,137–140]. After enrichment, captured library pools were paired-end sequenced on the Illumina Hiseq 4000 (Illumina) platform with 75 cycles providing about 20 million reads per assay.

Quality control and data processing

For both YCH and TIX libraries, we performed analyses of the captured sequence data using nf-core/eager v.2.3.4 (ref. 31). AdapterRemoval v.2 (ref. 141) was used to trim adaptor sequences and to remove adaptor dimers and low-quality sequence reads (minimum length, 30; minimum base quality, 20). Preprocessed sequences were mapped to the human genome assembly GRCh37 (hg19) from the Genome Reference Consortium[142] using BWA v.0.7.12 (ref. 143) and a seed length of 32. In the case of the ancient samples, the C to T misincorporation frequencies typical of aDNA were obtained using mapDamage 2.0 (ref. 144) to assess the authenticity of the aDNA fragments from the half-UDG-treated libraries. Genetic sex of the analysed individuals was assigned using SNP capture data by calculating the ratio of average X chromosomal and Y chromosomal coverage to the average autosomal coverage normalized by the chromosome length at the targeted SNPs[35]. Samples with an X rate between 0.35 and 0.55 and a Y rate between 0.4 and 0.7 were confirmed male. Analysis of next generation sequencing data(ANGSD) was used to estimate nuclear contamination, as males are expected to be homozygous at each X chromosome position[145]. 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 v.8.2.2; ref. 146) 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. To compare with available data from Mesoamerican and Central American populations, we merged our SNPs to the 593,124 SNPs of the Human Origins dataset[52]. Only individuals retaining more than 20,000 SNPs on this assay were kept for downstream analyses.

Uniparental markers

The mtDNA haplogroups were determined mapping reads to the revised Cambridge reference sequence[147]. For the resulting sequences, we used HaploGrep2 (ref. 148) and HAPLOFIND[149] to assign and confirm the corresponding mtDNA haplogroups (Supplementary Fig. 6). For the Y-haplogroup assignments, we created pileups of reads for each individual which mapped to Y chromosome (Y-Chr) SNPs as listed on the ISOGG Y-DNA Haplogroup Tree (v.15.73; https://isogg.org/tree/). We then manually assigned Y-Chr haplogroups for each individual based on the most downstream SNP retrieved after evaluating the presence of upstream mutations along the Y-Chr haplogroup phylogeny (Supplementary Fig. 6). Full mtDNA haplogroups and Y-Chr genotypes can be found in Supplementary Tables 9 (YCH) and 10 (TIX).

Genome-wide selection scans

Using the ancient (YCH; $n = 64$) and present-day (TIX; $n = 68$) genomic data, we performed two LSBL comparisons: first YCH versus Iberians from Spain ($n = 102$) and Han from China ($n = 103$) from the 1000 Genomes project[66] and then TIX versus Iberians and YCH to test for selection separate from YCH and Iberians. After quality control filtering, we removed individual YCH046 based on its position on the PCA. Sex chromosomes were not analysed, SNPs with minor allele frequencies less than 0.01 were excluded, as well as individuals with more than 97% missing data and finally we pruned SNPs for linkage disequilibrium using the following command: --indep-pairwise 200 25 0.4. The top 0.5% SNPs were annotated with biomaRt (release 3.15)[150]. LSBLs x, y and z are calculated using pairwise F_{ST} distances, d_{AB}, d_{BC} and d_{AC}, where $x = (d_{AB} + d_{AC} - d_{BC})/2$, $y = (d_{AB} + d_{BC} - d_{AC})/2$, $z = (d_{AC} + d_{BC} - d_{AB})/2$ and A, B and C are the three populations under consideration (Extended Data Fig. 2).

HLA allele assignment, haplotype reconstruction, statistical analyses and in silico binding prediction assays

We applied a development version of OptiType[151] on sequence data from the immune-captured libraries, mapped against a custom HLA reference panel containing 1,025 alleles with 'common' or 'intermediate' CIWD 3.0 designation (https://github.com/FRED-2/OptiType, tag GRG). The number of possible best matches was set to infinity. In this way we obtained HLA class I and class II alleles for the genes *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1* and *HLA-DPB1*. Of the 882 raw allele calls, 72 (8.1%) with anomalous coverage patterns induced by reads cross-mapping to several loci were overruled in favour of secondary allele predictions. Haplotypes were assigned on the basis of previously reported frequencies and properties such as linkage disequilibrium[77,152,153] and kinship when possible. To compare the HLA frequencies between precontact Chichén Itzá and postcontact Tixcacaltuyub, we performed Fisher exact tests[154] and adjusted the *P* values for several comparisons according to the Benjamini–Hochberg correction[155] which control the false discovery rate (FDR). We use the FDR adjustment as, compared to methods such as the Bonferroni correction, the FDR approach controls for a low proportion of false positives, instead of ensuring the existence of absolutely no false positives, resulting in increased statistical power[155]. It has been argued[74] that certain types of non-overlapping association between HLA loci may be a signature of pathogen selection. The f^*_{adj} metric[74,75,156] was calculated for each pairwise combination of HLA loci in the dataset, using YCH and TIX haplotypes for which we had data for every HLA locus. To provide a point of comparison for the f^*_{adj} score of each locus pair for both YCH and TIX, we generated 5,000 random permutations of the order of the total alleles at one of the loci in each pair and recalculated f^*_{adj} for each set of randomized data. We generated distributions of possible f^*_{adj} scores for each pair of loci in the dataset that would be obtained if the alleles at those loci were associated entirely at random. We then calculated the difference between the f^*_{adj} value calculated

from each dataset for each pair of loci and the mean of the f^*_{adj} scores calculated from the randomized data for that pair of loci, then divided that difference by the standard deviation of f^*_{adj} calculated from the randomized data. The resulting scores allowed us to rank the HLA pairs in order of how extreme an amount of non-overlap they showed relative to entirely random associations between the same alleles (Supplementary Methods: 'Non-overlapping associations'). To further assess the role of HLA class II alleles in resistance to *S. enterica* infection, we ran in silico binding predictions for the *HLA-DRB1/3/4/5* alleles and -*DQA1/DQB1* allele pairs found in YCH individuals and peptides derived from 18 *Salmonella* spp. proteins that have been previously reported to be highly immunogenic in humans (Supplementary Methods: 'In silico binding prediction assays') using NetMHCIIpan-4.0 (ref. 82) as implemented in the IEDB Analysis Resource virtual machine image[157,158]. Binders were classified as strong (rank less than 2%) or weak (2% ≤ rank ≤ 10%) on the basis of the adjusted rank values recommended by the authors and previously published research[82,159].

Principal components analysis

Smartpca (v.16000) from the Eigensoft package[160] was used to calculate principal components (PCs) of variation in the dataset, using the options 'lsqproject: YES' and 'shrinkmode: YES'. Initially, we projected the ancient individuals on PCs calculated on the genetic variation in 371 worldwide populations, to access the continental-level ancestries in the ancient individuals. We then projected the ancient individuals on PCs calculated on variation from 172 Native American populations, both modern and ancient (Supplementary Methods).

ADMIXTURE analysis

We used ADMIXTURE v.1.3.0 (ref. 161), a maximum-likelihood based clustering algorithm, to estimate the genetic structure present in our samples, after excluding variants with minor allele frequency of 0.01 and following linkage disequilibrium pruning using Plink (v.1.90) with a step size of 5, a window size of 200 and an R^2 threshold of 0.5 (ref. 161). For $K = 2$ to $K = 13$, we estimated the cross-validation error with 100 bootstrap replicates in an unsupervised model using either a panel of populations that are comprised of four ancestry clusters of African, European, Papuan and American origin based on the Simons Genome Diversity Project dataset[54] plus other relevant populations or a set of the aforementioned populations plus a set of 66 modern and ancient Native American populations to account for different sources of genetic contribution (Supplementary Methods). The lowest cross-validation errors corresponded to $K = 4$ and $K = 8$, respectively.

F-statistics and ancestry modelling tests

To assess the genetic relationships and patterns of admixtures suggested by the PCA and ADMIXTURE analysis, we carried out *F*-statistics analyses using the Xerxes CLI software (v.0.3.0.0) from the Poseidon (v.2.5.0) framework[162]. We calculated F_3-statistics of the form f_3 (Outgroup; target, X) to measure the amount of shared genetic drift of populations target and X after their divergence from an African outgroup; where X is either YCH or TIX and Outgroup corresponds to Mbuti from Congo[52]. We then used a subset of Native American populations (those with the top F_3 scores from our previous test) to perform F_4 tests of the form f_4(Mbuti, YCH; test, TIX) to investigate if the YCH and TIX groups are more closely related to each other or shared an excess of alleles with any population in position test. A negative value implies that either YCH and test or TIX and Mbuti share more alleles than expected under the null hypothesis of a symmetrical relationship between YCH and TIX. On the other hand, a positive value suggests that YCH and TIX share an excess of alleles between themselves. To test whether the newly sequenced individuals were genetically similar enough to be grouped into one group, we used qpWave[50] v.420, a modelling approach which evaluates the likelihood of two individuals to derive from the same ancestry, relative to a given set of references

Article

including worldwide populations. A value below 0.05 indicates a low likelihood, from which we conclude that one of the individuals has a different ancestry profile. For most tests, the respective individuals are consistent with deriving from the same ancestry tested relative to the populations Mbuti.DG, Onge.DG, Papuan.DG, Han.DG, Russia_MA1_HG.SG, USA_Ancient_Beringian.SG, USA_Anzick.SG, Mixe.DG, Mexico_Zapotec.DG, Belize_MayahakCabPek_9300BP, Karitiana.DG, Piapoco.DG, Peru_LaGalgada_4100BP, Pima, Cabecar. There are few exceptions, in which an individual shows low likelihood of sharing ancestry with several other individuals. Particularly high rates of pairs with a low likelihood were produced by the individuals YCH006, YCH011, YCH037, YCH046, YCH59 and YCH064. To investigate which ancestries differentiate the respective individual from most individuals, we calculated the differential affinities using an F_4 statistics of the form f_4(Mbuti.DG, X; YCH Group, Y), in which X would be the respective individual and Y present-day groups[50] and published ancient groups and individuals[49,163–172]. In this test, we would assume a positive test score, if the group or individual in Y has a higher affinity to the tested individual (X) compared to the group. Unexpectedly, there is no ancient or modern group or individual that could explain the behaviour in the qpWave analysis. To test whether present-day Mayan speakers ('Mayan' per The Simons Genome Diversity Project[54], Kaqchikel) were consistent with deriving from the same ancestry as the ancient individuals from Chichén Itzá, we tested them relative to the reference set of tested relative to the populations Mbuti.DG, Onge.DG, Papuan.DG, Han.DG, Russia_MA1_HG.SG, USA_Ancient_Beringian.SG, USA_Anzick.SG, Mixe.DG, Mexico_Zapotec.DG, Belize_MayahakCabPek_9300BP and Karitiana.DG. In both cases, the resulting P values (Kaqchikel: 0.08, 'Mayan': 0.755) suggest that the ancient and present-day groups share the same ancestry. The same test was not successful for other present-day populations such as Gen-Pano-Carib or Chibchan-Paezan speakers as well as for the seemingly unadmixed individuals from Tixcacaltuyub, possibly attributable to European admixture not detected in the model-based clustering analyses (Fig. 2 and Extended Data Fig. 1). When modelling their ancestry composition relative to the populations Mbuti.DG, Onge.DG, Papuan.DG, Han.DG, Russia_MA1_HG.SG, USA_Ancient_Beringian.SG, USA_Anzick.SG, Mixe.DG, Mexico_Zapotec.DG, Belize_MayahakCabPek_9300BP, Karitiana.DG as a mixture of Yoruba.DG, Spanish.DG and 'Mayan', the working model ($P = 0.11$) suggests a composition of 92% Mayan, 7% Spanish and 0.03% Yoruban ancestry.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the genomic data (including nuclear DNA, mtDNA and HLA alignment sequences) for the ancient Chichén Itzá individuals (YCH) are archived in the European Nucleotide Archive database (accession no. PRJEB73567). Present-day genomic data of Tixcacaltuyub individuals (TIX) are archived in the European Genome-Phenome Archive database (dataset no. EGAD50000000426) and will be made available on request from R.B., K.N., J.C.L.R. and J.K. and subject to a signed agreement to restrict usage to anonymized studies of population history. All HLA data from our sample sets, both frequencies and individual genotypes, can be found at The Allele Frequency Net Database website (www.allelefrequencies.net) under accession nos. 3791 (YCH) and 3790 (TIX).

107. Potter, D. F. Prehispanic architecture and sculpture in Central Yucatan. *Am. Antiq.* **41**, 430–448 (1976).
108. Andrews IV, E. W. in *Handbook of Middle American Indians: Guide to Ethnohistorical Sources* Vol. 13 (ed. Wauchope, R.) 288–330 (Univ. Texas Press, 1964).
109. Andrews, A. P., Andrews, E. W. & Castellanos, F. R. The northern Maya collapse and its aftermath. *Anc. Mesoam.* **14**, 151–156 (2003).
110. Hermes, B. & Calderón, Z. in *El Sitio Maya de Topoxté. Investigaciones en una Isla del Lago Yaxhá, Petén, Guatemala* (ed. Wurster, W. W.) 66–74 (Verlag Philipp von Zabern, 2000).

111. Milbrath, S. in *New Directions in American Archaeoastronomy (Proc. 46th International Congress of Americanists)* (ed. Aveni, A. F.) 57–79 (BAR International, 1988).
112. Andrews, E. W. *Balankanche, Throne of the Tiger Priest* (Middle American Research Institute, 1970).
113. Ubelaker, D. H. *Human Skeletal Remains: Excavation, Analysis, Interpretation* (Aldine, 1978).
114. Hernández Espinosa, P. O. *La Regulación Del Crecimiento de La Población En El México Prehispánico* (Instituto Nacional de Antropología e Historia, 2006).
115. Verdugo, C. et al. Implications of age and sex determinations of ancient Maya sacrificial victims at Midnight Terror Cave. *Int. J. Osteoarchaeol.* **30**, 458–468 (2020).
116. Karnes, J. H. et al. Phenome-wide scanning identifies multiple diseases and disease severity phenotypes associated with HLA variants. *Sci. Transl. Med.* **9**, eaai8708 (2017).
117. Fernández-Torres, J., Flores-Jiménez, D., Arroyo-Pérez, A., Granados, J. & López-Reyes, A. HLA-B*40 allele plays a role in the development of Acute Leukemia in Mexican population: a case-control study. *BioMed. Res. Int.* **2013**, 705862 (2013).
118. Brown, T. A., Nelson, D. E., Vogel, J. S. & Southon, J. R. Improved collagen extraction by modified Longin method. *Radiocarbon* **30**, 171–177 (1988).
119. Stuiver, M. & Polach, H. A. Discussion reporting of ¹⁴C data. *Radiocarbon* **19**, 355–363 (1977).
120. Reimer, P. J. et al. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).
121. Ramsey, C. B. & Lee, S. Recent and planned developments of the program OxCal. *Radiocarbon* **55**, 720–730 (2013).
122. Błaszczyk, D. et al. Social status and diet. Reconstruction of diet of individuals buried in some early medieval chamber graves from Poland by carbon and nitrogen stable isotopes analysis. *J. Archaeol. Sci. Rep.* **38**, 103103 (2021).
123. Pérez-Ramallo, P. et al. Stable isotope analysis and differences in diet and social status in northern Medieval Christian Spain (9th–13th centuries CE). *J. Archaeol. Sci. Rep.* **41**, 103325 (2022).
124. Ambrose, S. H. & Norr, L. in *Prehistoric Human Bone: Archaeology at the Molecular Level* (eds Lambert, J. B. & Grupe, G.) 1–37 https://doi.org/10.1007/978-3-662-02894-0_1 (Springer, 1993).
125. Richards, M. P. & Hedges, R. E. M. Stable isotope evidence for similarities in the types of marine foods used by late Mesolithic humans at sites along the Atlantic coast of Europe. *J Archaeol. Sci.* **26**, 717–722 (1999).
126. van Klinken, G. J. Bone collagen quality indicators for palaeodietary and radiocarbon measurements. *J. Archaeol. Sci.* **26**, 687–695 (1999).
127. DeNiro, M. J. Postmortem preservation and alteration of in vivo bone collagen isotope ratios in relation to palaeodietary reconstruction. *Nature* **317**, 806–809 (1985).
128. Rohland, N. & Hofreiter, M. Comparison and optimization of ancient DNA extraction. *Biotechniques* **42**, 343–352 (2007).
129. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
130. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl Acad. Sci. USA* **104**, 14616–14621 (2007).
131. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737 (2013).
132. Gansauge, M.-T. & Meyer, M. in *Ancient DNA: Methods and Protocols* (eds Shapiro, B. et al.) 75–83 (Springer New York, 2019).
133. Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020).
134. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **5**, pdb.prot5448 (2010).
135. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* **370**, 20130624–20130624 (2014).
136. Gnirke, A. et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* **27**, 182–189 (2009).
137. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl Acad Sci USA* **110**, 2223–2227 (2013).
138. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
139. Barquera, R. et al. Origin and health status of first-generation Africans from early Colonial Mexico. *Curr. Biol.* **30**, 2078–2091 (2020).
140. Rohrlach, A. B. et al. Using Y-chromosome capture enrichment to resolve haplogroup H2 shows new evidence for a two-path Neolithic expansion to Western Europe. *Sci. Rep.* **11**, 15005 (2021).
141. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification and read merging. *BMC Res. Notes* **9**, 88 (2016).
142. Lander, E. S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
143. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2009).
144. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. MapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
145. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinf.* **15**, 356 (2014).
146. Schiffels, S. sequenceTools. GitHub https://github.com/stschiff/sequenceTools.git (2018).
147. Andrews, R. M. et al. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147 (1999).
148. Weissensteiner, H. et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).
149. Vianello, D. et al. HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. *Hum. Mutat.* **34**, 1189–1194 (2013).

150. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

151. Szolek, A. et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014).

152. Cao, K. et al. Analysis of the frequencies of HLA-A, B and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum. Immunol.* **62**, 1009–1030 (2001).

153. Yunis, E. J. et al. In *Encyclopedia of Molecular Cell Biology and Molecular Medicine, Vol. 13* (ed. Meyers, R. A.) pp 192–215 (John Wiley & Sons, 2006).

154. R Core Team. *R: A Language and Environment for Statistical Computing* https://www.r-project.org/ (Foundation for Statistical Computing, 2021).

155. Benjamini, Y., Kenigsberg, E., Reiner, A. & Yekutieli, D. FDR adjustments of microarray experiments (FDR-AME). *MathTauAcIl* **1**, 1–3 (2005).

156. Penman, B. S. & Gupta, S. Detecting signatures of past pathogen selection on human HLA loci: are there needles in the haystack? *Parasitology* **145**, 731–739 (2018).

157. Wang, P. et al. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinf.* **11**, 568 (2010).

158. Moutaftsi, M. et al. A consensus epitope prediction approach identifies the breadth of murine T(CD8 +)-cell responses to vaccinia virus. *Nat. Biotechnol.* **24**, 817–819 (2006).

159. Di, D., Simon Thomas, J., Currat, M., Nunes, J. M. & Sanchez-Mazas, A. Challenging ancient DNA results about putative HLA protection or susceptibility to *Yersinia pestis*. *Mol. Biol. Evol.* **39**, msac073 (2022).

160. Bryc, K. et al. Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc. Natl Acad. Sci. USA* **107**, 8954–8961 (2010).

161. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

162. Schiffels, S. Xerxes CLI software. GitHub https://github.com/poseidon-framework/poseidon-framework.github.io/blob/master/xerxes.md (2022).

163. Nägele, K. et al. Genomic insights into the early peopling of the Caribbean. *Science* **369**, 456–460 (2020).

164. Nakatsuka, N. et al. A paleogenomic reconstruction of the deep population history of the Andes. *Cell* **181**, 1131–1145 (2020).

165. Scheib, C. L. et al. Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **360**, 1024–1027 (2018).

166. Fernandes, D. M. et al. A genetic history of the pre-contact Caribbean. *Nature* **590**, 103–110 (2021).

167. John, L. et al. The genetic prehistory of the Andean highlands 7000 years BP though European contact. *Sci. Adv.* **4**, eaau4921 (2018).

168. Capodiferro, M. R. et al. Archaeogenomic distinctiveness of the Isthmo-Colombian area. *Cell* **184**, 1706–1723 (2021).

169. Moreno-Mayar, J. V. et al. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**, 203–207 (2018).

170. Moreno-Mayar, J. V. et al. Early human dispersals within the Americas. *Science* **362**, eaav2621 (2018).

171. Rasmussen, M. et al. The ancestry and affiliations of Kennewick Man. *Nature* **523**, 455–458 (2015).

172. Schroeder, H. et al. Origins and genetic legacies of the Caribbean Taino. *Proc. Natl Acad. Sci. USA* **115**, 201716839 (2018).

Author contributions R.B., O.C.C., K.N., D.I.H.Z., C.W. and J.K. conceived the project. R.B., O.C.C. and D.I.H.Z. performed the registration and photographic recording for YCH. O.C.C. performed the osteological analyses for YCH. R.B., O.C.C. and D.I.H.Z. performed the field work sampling of YCH. J.C.L.R., M.E.M.M., J.C.T.R. and C.I.T.P. performed the sampling and DNA extraction for TIX. R.B. and R.A.B. performed the aDNA laboratory work and the library preparation for TIX. R.B., O.C.C., R.A.B., P.P.R., P.R. and C.W. analysed and discussed the archaeological data, including the radiocarbon dates. P.P.R., M.L., A.B.R., P.R. and C.W. performed and discussed the stable isotopes analyses. R.B., K.N., A.B.R. and P.L. performed the genetic continuity tests. R.B., D.I.H.Z., B.S.P., A.S., A.B.R., A.C. and O.K. performed the selection and HLA analyses. B.S.P. contributed to the HLA non-overlap analysis. R.B., O.C.C., K.N., D.I.H.Z., A.B.R., R.A.B., V.A.A., C.W. and J.K. analysed the population genetics data and uniparental markers results. All authors discussed the results. R.B., O.C.C., K.N., D.I.H.Z., P.P.R., C.W. and J.K. wrote the paper with contributions from all authors. All authors read and approved the final version. R.B., O.C.C., K.N., D.I.H.Z. and V.A.A. wrote and submitted the final report (in Spanish) for the Council of Archaeology (INAH) based on the present manuscript, as well as the Spanish translation for the main manuscript.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-07509-7.
Correspondence and requests for materials should be addressed to Rodrigo Barquera, Oana Del Castillo-Chávez or Johannes Krause.
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.
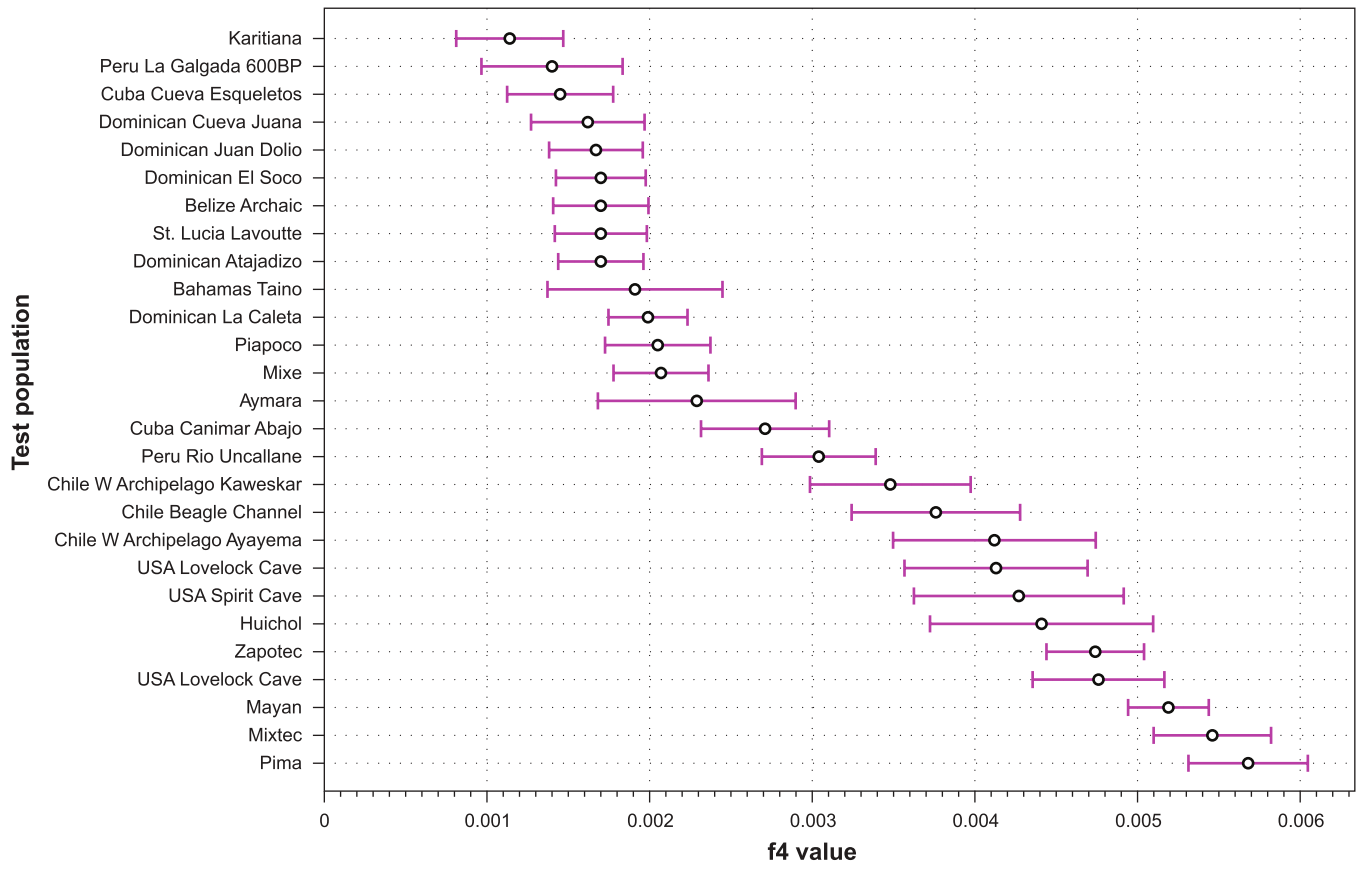
Extended Data Fig. 1 | F4 test of the form f_4**(Mbuti, YCH; *test*, TIX).** Values are plotted (circle) with their corresponding error bars (in purple). For all tested populations, the value of the test is positive, indicating that YCH And TIX shared more alleles between them than to any of the Native American populations tested.
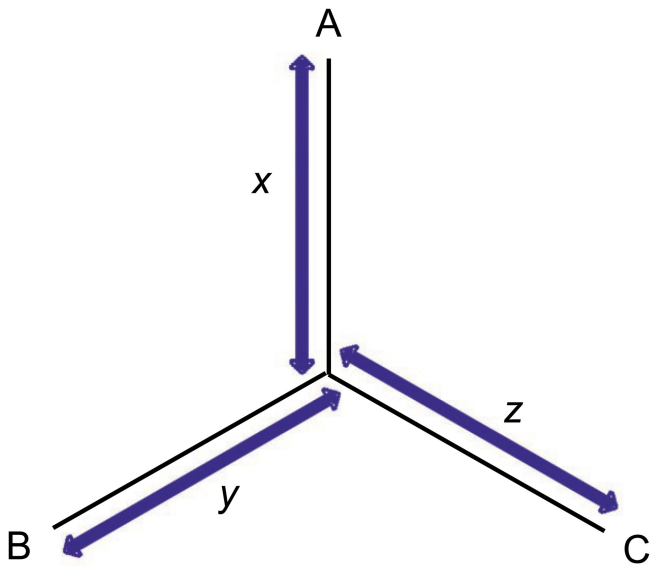
Extended Data Fig. 2 | A graphical representation for the locus-specific branch lengths comparisons. The variables x, y and z, (representing either YCH vs. Iberians from Spain and Han from China, or TIX vs. Iberians and YCH, to independently test for selection) are calculated using pairwise F_{ST} distances, d_{AB}, d_{BC} and d_{AC}, where $x = (d_{AB} + d_{AC} - d_{BC})/2$, $y = (d_{AB} + d_{BC} - d_{AC})/2$, $z = (d_{AC} + d_{BC} - d_{AB})/2$ and A, B and C are the three populations under consideration.

nature portfolio

Corresponding author(s): Rodrigo Barquera, Oana del Castillo-Chávez, Johannes Krause.

Last updated by author(s): 11/03/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	nf-core/eager ver. 2.3.4 Nextflow v20.10.0 FastQC v0.11.9 MultiQC v1.9 AdapterRemoval v2.3.1 fastP v0.20.1 BWA v0.7.12 Bowtie2 v2.4.1 mapDamage 2.0 Samtools v1.9 endorS.py v0.4 DeDup v0.12.7 Picard MarkDuplicates v2.22.9 Qualimap v2.2.2-dev Preseq v2.0.3 pileupCaller ver. 8.2.2
Data analysis	Oxcal ver. 4.3 HaploGrep2 HAPLOFIND biomaRt Release 3.15 OptiType, tag GRG NetMHCIIpan-4.0 Smartpca ver. 16000

```
ADMIXTURE ver. 1.3.0
Plink ver. 1.90
Xerxes CLI software v. 0.3.0.0
Poseidon v. 2.5.0
qpWave ver. 420
```

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All of the genomic data (including nuclear DNA, mtDNA, and HLA alignment sequences) for the ancient Chichén Itzá individuals (YCH) are archived in the European Nucleotide Archive (ENA) database (accession number: PRJEB73567). Present-day genomic data of Tixcacaltuyub individuals (TIX) are archived in the (EGA) database (accession number: 10489) and will be made available on request to R.B., J.C.L.R. and J.K., and subject to a signed agreement to restrict usage to anonymised studies of population history. All HLA data from our sample sets, both frequencies and individual genotypes, can be found at The Allele Frequency Net Database website (www.allelefrequencies.net) under accession numbers: 3791 (YCH) and 3790 (TIX).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to determine ancient DNA sample size a priori. The number of genomes analyzed in this study depends on available human remains associated with hunter-gatherer individuals with signature of preserved ancient DNA. Those specimens are very limited because of the scarce availability and poor molecular preservation of human remains from that period.
Data exclusions	For ancient DNA screening analysis, libraries with less than ~0.1% human DNA and/or with no ancient DNA damage pattern are not carried on for SNPs capture. For alignment, sequencing reads with fragment length <30bp and mapping quality/base quality <30 are excluded from genotyping. For ancient DNA authentication, libraries with indication of substantial contamination levels or individually-analyzed libraries with marginal contamination levels are filtered to analyze only sequencing reads that carry signs of post-mortem DNA damage. For whole genome analysis, individuals with less than 20,000 usable SNPs are excluded.
Replication	Multiple libraries (1-6) are prepared for a subset of samples as replication. The genotypes from different libraries (both single stranded and double stranded) are merged for downstream analysis after confirming of similar statistical behaviour. Each sample is analyzed for up to 1.24 million markers across the human genome that represent an internal replication of the findings.
Randomization	Randomization is not relevant to this study. Samples are grouped based on sampling locations, dates and genetic affinities.
Blinding	Blinding is not applicable for ancient specimens as the sampling locations and dates are known as prior. In downstream data analysis blinding is also not relevant since the newly generated ancient genomes are co-analyzed with previously published present-day and ancient human genomes.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Palaeontology and Archaeology

Specimen provenance	The ancient samples used in this study belonged to individuals whose skeletons were recovered from the archaeological excavations of a chultún (cistern) connected to a natural cave within the archaeological site of Chichén Itzá, Yucatán, Mexico.
Specimen deposition	The skeletal remains analyzed in this study derive from a skeletal collection curated by O.C.C. and her team, stored at the Centro INAH Yucatan, in Merida, Yucatan, Mexico. The genetic libraries generated from DNA extracted from the skeletal remains are stored at the Max Planck Institute for Evolutionary Anthropology (MPI-EVA) in Jena and Leipzig, Germany.
Dating methods	According to the report issued by The Curt-Engelhorn-Centre for Archaeometry (Mannheim, Germany), the portions of the petrous bone of 26 YCH samples were pre-treated and analysed using a standardised procedure. Collagen was extracted from the bone samples (approx. 1 g, using a modified version of the Longin method), purified by ultrafiltration (fraction >30kD) and freeze-dried. The collagen was then combusted to CO_2 in an Elemental Analyzer (EA). The CO_2 was then converted catalytically to graphite and analysed using a MICADAS-type AMS system. The isotopic ratios 14C/12C and 13C/12C of samples, calibration standard (Oxalic Acid-II), blanks and control standards were measured simultaneously in the AMS system. 14C-ages are normalised to delta13C=-25‰120 with a typical uncertainty of 2‰, and calibrated using the dataset IntCal20 and the software Oxcal (ver. 4.3.2).

☒ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	No ethical approval or guidance was required for the analyses carried on the archaeological material.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	The village of Tixcacaltuyub, part of the municipality of Yaxcabá, is located 90 km southeast of Merida, 16 km off Sotuta, and is located 55 Km southeast of Chichén Itzá. The community of Tixcacaltuyub self-identifies as a Mayan community and has been in a years-long cooperative relationship with the Chemistry and Nursing Faculties of the Universidad Autónoma de Yucatán (UADY), Mérida, Yucatán, following projects investigating the relationship between health and lifestyle in the community.
Recruitment	Ethical approval for the collection of blood samples from the individuals from Tixcacaltuyub, Yucatán, Mexico, was granted by the Committee of Ethics and Research, Autonomous University of Yucatán (UADY), Mexico (Project: Bienestar Comunitario: Proyecto de capacitación para la autogestión de la salud de personas con DT2 y sus familias, en la comunidad de Tixcacaltuyub y Yaxcabá; official notice number: F-FENC-SAC-14/REV: 04; Registry number: 09/17), and performed according to the requisites of the Helsinki Declaration (2008) and the General Health Law of Mexico. All subjects were informed about the objectives and methods used, and signed an informed consent form.
Ethics oversight	Committee of Ethics and Research, Autonomous University of Yucatán (UADY), Mexico.

Note that full information on the approval of the study protocol must also be provided in the manuscript.