THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Pierre CLAVEL

Le 27 septembre 2023

Santé et environnement à l'heure de la colonisation et des grandes pandémies du passé : apports des génomes et microbiomes anciens.

Ecole doctorale : **BSB - Biologie, Santé, Biotechnologies**

Spécialité : **ANTHROPOBIOLOGIE**

Unité de recherche :
CAGT - Centre d'Anthropobiologie et de Génomique de Toulouse

Thèse dirigée par
Ludovic ORLANDO et Andaine SEGUIN-ORLANDO

Jury

Mme Marie-Claude MARSOLIER-KERGOAT, Rapporteure
M. Etienne PATIN, Rapporteur
Mme Christine KEYSER, Rapporteure
M. Eric CRUBÉZY, Examinateur
Mme Evelyne HEYER, Examinatrice
M. Ludovic ORLANDO, Co-directeur de thèse
Mme Andaine SEGUIN-ORLANDO, Co-directrice de thèse

Remerciements

Je tiens avant tout à exprimer ma gratitude envers les membres du jury (Marie-Claude Marsolier-Kergoat, Etienne Patin, Christine Keyser, Éric Crubézy et Evelyne Heyer) qui ont accepté d'évaluer mon travail, en particulier les rapporteurs, pour le temps précieux qu'ils ont consacré à l'examen de mon manuscrit. Mes sincères remerciements vont également aux membres de mon comité de thèse (Macha Nikolski, Lluis Quintana-Murci, Virginie Rougeron et Sébastien Calvignac-Spencer) pour leur accompagnement et leurs conseils durant les deux premières années de ma thèse.

Je souhaite exprimer ma profonde reconnaissance envers Ludovic Orlando et Andaine Seguin-Orlando qui ont placé leur confiance en moi et m'ont permis de réaliser ce projet passionnant. Andaine, merci infiniment pour le temps que tu as passé à mes côtés à me former aux techniques et manipulations dans les laboratoires d'ADN ancien, pour m'avoir encadré dès ma première année de master et pour m'avoir transmis le goût de la recherche. Ludovic, je te remercie de ton écoute, ta patience et de ton aide indispensable à la réalisation de ce projet. Je te suis reconnaissant de m'avoir permis de participer à de nombreux travaux passionnants menés par le laboratoire. Tous les deux, vous m'avez poussé à continuer malgré les difficultés, vous m'avez guidé avec détermination tout au long de ces trois années de thèse. Votre soutien inébranlable m'a été d'une aide plus que précieuse lors des moments difficiles.

Je tiens bien évidemment à remercier toute l'équipe AGES (Archaeology, Genomics, Evolution and Societies) du CAGT (Centre for Anthoropobiology and Genomics of Toulouse) qui m'a accueilli comme un membre de cette grande famille. Je souhaite adresser mes remerciements particuliers à Laure Calvière-Tonasso, Stéphanie Schiavinato et Lorelei Chauvey qui m'ont encadré au cours de mes premières manipulations dans les laboratoires d'ADN ancien. Sans vous, je n'aurais jamais pu mener à bien ce projet. Je tiens également à remercier mes co-encadrants de stage de master (Jean Dumoncel, Clio Der Sarkissian, Andaine Seguin-Orlando et José Braga) pour leur bienveillance et leur patience, sans lesquels je n'aurais jamais entrepris un tel projet. Enfin, je tiens à remercier Lexane Louis, qui m'a initié à la bioinformatique durant son stage, et Perle Guarino-Vignon pour son aide précieuse dans l'analyse des séquences anciennes.

Mes remerciements vont également à tous les archéologues et collaborateurs pour nous avoir fourni un corpus exceptionnel d'échantillons. Je remercie tout particulièrement Benoit Clavel pour avoir consacré un temps précieux à l'échantillonnage de dents et de crânes, et pour m'avoir présenté Richard Jonvel, marquant le début d'une belle collaboration. Je remercie José Braga et Alexandre Ribéron pour leur enseignement passionnant durant le master ABI (AnthropoBiologie Intégrative) et pour m'avoir donné le goût de l'anthropobiologie.

Je tiens à remercier Julie, qui partage ma vie depuis un certain temps déjà, pour m'avoir toujours encouragé à me dépasser, à ne jamais abandonner, pour m'avoir soutenu dans les bons et les mauvais moments, et m'avoir supporté lors de mes sautes d'humeur et mes déprimes. Mes remerciements vont également à ma famille, qui a toujours été présente pour moi, ainsi qu'à la famille de Mollien-Lucas pour ces délicieux repas dominicaux qui réchauffent le cœur. Merci à tous mes amis du Gers, d'Aveyron, de Toulouse et plus encore pour les soirées jeux de société qui font tant de bien, pour les séances de sport qui libèrent l'esprit et pour l'oreille attentive qu'ils m'ont prêté lors de mes pertes de moral.

Résumé

Ce travail de thèse entre dans le cadre général de l'étude des épidémies du passé et leur impact biologique sur les populations humaines. Notre travail a plus spécifiquement porté sur l'analyse paléogénomique de squelettes humains historiques, retrouvés en France, dans des contextes archéologiques liés aux épidémies de peste des XVIIème et XVIIIème siècles. Les sites étudiés sont associés à la temporalité de la guerre de Trente Ans (XVIIème siècle) et à la grande épidémie de peste de Marseille de 1720-1722. Deux axes principaux ont structuré notre recherche : la caractérisation du génome de *Yersinia pestis*, agent infectieux ayant sévi pendant ces deux crises, mais aussi celle des individus ayant succombé à la maladie, en particulier les loci impliqués dans l'immunité innée. Notre travail a abouti au développement d'un protocole d'extraction d'ADN ancien optimisé pour le séquençage de pathogènes anciens tels que celui de la peste ; il a également permis de caractériser les souches pathogéniques qui sévissaient en France aux XVIIème et XVIIIème siècles, de séquencer les génomes des individus morts au cours de ces épidémies afin de caractériser leur ascendance génétique et d'évaluer si certains variants génétiques avaient pu augmenter leur susceptibilité à la maladie.

Pour caractériser l'agent responsable de ces épidémies, l'ADN a été extrait de la pulpe dentaire d'individus anciens, avant d'être immortalisé sous forme de banques d'ADN de nouvelle génération. Les banques d'ADN ont ensuite été enrichies pour leur contenu pathogénique en ciblant par capture en solution le génome de la bactérie *Y. pestis*. Cette approche se base sur l'utilisation de sondes ARN de synthèse de courte taille, choisies pour être complémentaires au génome d'intérêt. Si cette approche a prouvé son efficacité par le passé, elle n'est cependant pas dénuée de limitations, notamment au regard des qualités et quantités limitées de matériel disponible. Afin d'en optimiser l'efficacité, nous avons étudié la dynamique de libération de l'ADN de *Y. pestis* au cours de l'étape d'extraction d'ADN et avons établi qu'une prédigestion de la pulpe dentaire pendant une heure suffisait à libérer plus de 99 % de l'ADN du pathogène.

Nous avons ensuite cherché à caractériser génétiquement les souches de peste impliquées dans les épidémies de peste sévissant en France aux XVIIème et XVIIIème siècles. Cette approche nous a permis de déterminer les relations phylogénétiques avec les souches de peste responsables de la seconde pandémie, parmi lesquelles figurent la souche dite de la Peste

Noire, qui a décimé l'Europe au XIVème siècle. Nous avons également testé si la différence de mortalité observée entre les épisodes de peste de 1720 et de 1722 à Marseille pouvait s'expliquer par des variations génétiques modulant la virulence. Globalement, notre analyse a permis de caractériser 12 génomes de *Y. pestis* et a démontré que les souches du XVIIème siècle étaient étroitement liées à d'autres souches associées à la guerre de Trente Ans. D'autre part, les bactéries responsables de la grande peste de Marseille de 1720 apparaissent comme génétiquement très proches de celles de 1722. Enfin, cette thèse s'est aussi intéressée à celles et ceux qui sont morts de la peste à ces époques. Le séquençage de l'ADN contenu dans l'os pétreux dont les taux exceptionnels de préservation nous ont permis de caractériser la variation génétique à l'échelle du génome dans son entier. Au total, les 73 génomes humains ont été séquencés avec une profondeur de séquençage moyenne d'au moins 1X. Ces données ont ainsi permis d'explorer le profil génétique des victimes de la peste pour mieux comprendre qui mourrait au juste, au cours de ces épidémies.

En conclusion, ce travail de thèse apporte de nouvelles données sur l'impact biologique des épidémies passées sur les populations humaines. Il démontre l'utilité de l'approche paléogénomique pour étudier les pathogènes anciens et leurs interactions avec les populations humaines.

Abstract

This PhD work falls within the general framework of studying past epidemics and their biological impact on human populations. Specifically, this work focuses on the palaeogenomic analysis of historical human skeletons from France that were found in archaeological contexts related to the 17th and 18th centuries plague epidemics. The sites under study are associated with the Thirty Years War (17th century) and the great plague epidemic of Marseille from 1720-1722. Two main axes have structured our research: the characterization of the genome of *Yersinia pestis*, the infectious agent that accompanied both crises, and the characterization of individuals who succumbed to the disease, particularly their innate immunogenome. The present work resulted in (*i*) the development of an optimized ancient DNA extraction protocol for the sequencing of ancient pathogens such as *Y. pestis*, characterizing the pathogenic strains that prevailed in France during the 17th and 18th centuries, and (*ii*) the sequencing of the genomes of individuals who died during these epidemics, in order to characterize their genetic ancestry and assess whether certain genetic variants may have increased their susceptibility to the disease.

To characterize the agent responsible for these epidemics, DNA was extracted from the dental pulp of ancient individuals and then immortalized through next-generation DNA libraries. The DNA libraries were then enriched for their pathogenic content by targeting the genome of the *Y. pestis* bacterium through in-solution capture using synthetic short-length RNA probes chosen to be complementary to the genome of interest. While this approach has proven effective in the past, it is not without limitations, especially considering the limited quality and quantity of available material. To optimize its efficiency, the dynamics of *Y. pestis* DNA release during the DNA extraction step was investigated and it was established that pre-digestion of the dental pulp for one hour was sufficient to release over 99 % of the pathogen's DNA. Next, the genetic characterization of plague strains involved in the plague epidemics in France during the 17th and 18th centuries was undertaken. This approach allowed us, in particular, to determine the phylogenetic relationships with the plague strains responsible for the second pandemic, including the so-called Black Death strain that decimated Europe in the mid-14th century. We also tested whether the difference in mortality observed between the 1720 and 1722 Marseille plague episodes could be explained by genetic variations affecting virulence. Overall, the analysis characterized 12 *Y. pestis* genomes and demonstrated that 17th-century strains were closely related to other strains associated with the Thirty Years War. Furthermore, the bacteria

responsible for the plague epidemic of Marseille in 1720 appeared to be genetically very close to those of 1722. Finally, this PhD thesis also paid attention to individuals killed by plague during these outbreaks. The genomes of these individuals were sequenced from the DNA contained in their petrous bones, which showed exceptional preservation rates, allowing us to characterize nearly complete genomes. A total of 73 human genomes were sequenced with a minimum average depth-of-coverage of 1-fold. These data thus revealed the genetic profile of those individuals most often falling as victims during these epidemics.

In conclusion, this thesis provides new data on the biological impact of past epidemics on human populations. It demonstrates the power of palaeogenomic approaches for studying ancient pathogens, their interactions with and their impact on human populations.

Abréviations & Glossaire

ACP : Analyse en Composantes Principales ; PCA en anglais

ADN : Acide DésoxyriboNucléique

AP : Avant le présent ; *Before present* (BP) en anglais

ARN : Acide RiboNucléique

CAGT : Centre d'Anthropobiologie et de Génomique de Toulouse

cM : centimorgan

Fst : Indice de fixation

kb : Kilobase

NGS : *Next Generation Sequencing* en anglais ; Nouvelles techniques de séquençage en français

bp : Paires de bases ; base pairs en anglais

ROH : *Runs Of Homozygosity* ; « blocs d'homozygosité » en français

Rpm : Rotations Par Minute

SNP : *Single Nucleotide Polymorphism* en anglaise ; pouvant être traduit par « polymorphisme nucléotidique » en français

Table des matières

I) INTRODUCTION 12

1) UN BREF APERÇU DES RECHERCHES EN ADN ANCIEN 12

a) Les prémices d'une discipline 12
b) L'année 2010 : l'avènement de la paléogénomique 12

2) LES APPLICATIONS DE L'ADN ANCIEN 15

a) L'étude de la domestication 16
(i) La domestication animale 16
(ii) La domestication végétale 17
b) Les pathogènes humains 18
c) L'ADN environnemental 20

3) LES METHODES POUR L'ETUDE DE L'ADN ANCIEN 22

a) L'extraction de l'ADN ancien 24
b) Le traitement USER 27
c) La construction de banques d'ADN ancien 28
(i) Les banques dites « double brin » 29
(ii) Les banques dites « simple brin » 30
d) La capture par hybridation 31
e) L'authentification de l'ADN ancien 35
f) Les limites de l'ADN ancien 35

4) LES APPORTS DE L'ADN ANCIEN SUR LES EPIDEMIES PASSEES 36

a) Identifier les pathogènes du passé 36
b) L'origine des pathogènes 38
c) L'évolution des pathogènes 39
d) L'immunité face aux pathogènes 40

5) OBJECTIFS 44

a) Le contexte de la thèse 44
b) Optimisation de l'extraction de l'ADN de pathogènes anciens 46
c) Étude de la diversité des souches de peste 48
d) L'Immunité innée face à la peste 48

II) ARTICLE 1 49

1) CONTEXTE 49

2) IMPROVING THE EXTRACTION OF ANCIENT YERSINIA PESTIS GENOMES FROM THE DENTAL PULP 51

3) DISCUSSION 77

III) ÉTUDE DES GENOMES DES VICTIMES DE LA PESTE 79

1) INTRODUCTION 79

2) ANALYSE DES GENOMES HUMAINS ANCIENS 80

a) Procédure expérimentale 80
(i) Extraction de l'ADN 80
(ii) Construction de banques ADN et amplification par PCR 81
b) Analyse bioinformatique 82
(i) Traitement et alignement des séquences 82
(ii) Dégradations post-mortem de l'ADN 83
(iii) Estimation des taux d'erreur et de la contamination 85
(iv) Détermination du sexe, des haplogroupes des génomes uniparentaux, et calcul de la consanguinité individuelle 86
(v) Analyse en Composantes Principales 88

(vi) Analyse des variants en lien avec l'immunité innée 90
3) DISCUSSION 90

IV) PERSPECTIVES 91

V) ANNEXES 92

VI) REFERENCES 93

I) Introduction

1) Un bref aperçu des recherches en ADN ancien

a) Les prémices d'une discipline

Le milieu des années 1980 marque la naissance d'une discipline, la paléogénétique, qui consiste à étudier le matériel génétique provenant de restes archéologiques et paléontologiques. L'ambition de ce domaine est de caractériser des échantillons biologiques anciens pour en expliquer l'histoire évolutive comme celle d'espèces aujourd'hui disparues ou encore de l'espèce humaine et des organismes qui y sont étroitement liés. Aux prémices de cette discipline, le défi était de réussir à isoler des fragments d'ADN ancien pour en obtenir la séquence. Cependant, de nombreuses contraintes liées au matériel rendent la tâche particulièrement complexe. Les molécules d'ADN ancien sont présentes en très faible quantité et sont fortement endommagées suite à la mort de l'individu et à l'arrêt des processus cellulaires de réparation de l'ADN. Cela est amplifié par l'action du temps et des micro-organismes qui colonisent les restes archéologiques (Dabney et al., 2013b). Ainsi, la majorité du matériel génétique que l'on retrouve correspond souvent à des contaminants, tels que l'ADN des micro-organismes environnementaux ou celui des expérimentateurs, tandis qu'une faible fraction correspond à l'ADN de l'organisme étudié (ADN endogène, Green et al., 2010). Malgré ces contraintes, Higuchi et ses collaborateurs ont réussi à extraire en 1984 quelques courtes séquences d'ADN de 229 paires de bases (pb) qui provenaient d'un spécimen de musée d'une espèce d'équidé aujourd'hui éteinte, le quagga. Cette étude démontrait ainsi la possibilité de retrouver et d'étudier le matériel génétique d'échantillons anciens âgés d'au-moins 100 ans (Higuchi et al., 1984). Dans les décennies qui suivront, la paléogénomique s'est développée et les travaux sur l'ADN ancien se sont multipliés. L'essor de ce domaine a poussé des laboratoires à se spécialiser et adapter leurs infrastructures et leurs méthodes pour le traitement d'échantillons anciens (Knapp et al., 2012; Orlando et al., 2021a). Ces installations sont conçues pour permettre de manipuler les molécules d'ADN ancien fortement dégradées et de limiter au maximum toute contamination provenant de l'expérimentateur ou du matériel, en respectant une réglementation et des protocoles stricts.

b) L'année 2010 : l'avènement de la paléogénomique

Malgré ces avancées, il faudra attendre le développement des techniques de séquençage nouvelle génération (NGS) et l'année 2010 pour voir le premier génome ancien séquencé. C'est en effet au cours de cette année que l'ADN d'un paléo-Eskimo associé à la culture Saqqaq a été extrait à partir de cheveux préservés dans le permafrost pendant près de 4 000 ans (Rasmussen et al., 2010). C'est aussi cette même année, en 2010, qu'une première ébauche du génome de l'homme de Néandertal (*Homo neanderthalensis*) a été séquencée à partir de trois vestiges osseux âgés de 38 310 (± 2 130 ans) à 44 450 (± 550 ans) ans AP (avant le présent), provenant de la grotte de Vindija en Croatie (Green et al., 2010). Cette avancée majeure a ouvert la voie à l'exploration de nos origines et de nos relations évolutives avec nos plus proches parents. Cette étude démontre que les populations Néandertaliennes partagent un excès de mutations dérivées avec les populations humaines actuelles d'Eurasie que d'Afrique. Des échanges de matériel génétique et donc de métissage entre *Homo sapiens* et *Homo neanderthalensis* ont ainsi pu être mis en évidence. D'après les auteurs de ces travaux, confirmés depuis, les populations non-africaines auraient reçu une contribution génétique des Néandertaliens qui, selon les individus, représente une proportion variable allant de 1,8 % à 2,6 % de leur génome (Prüfer et al., 2014, 2017).

Toujours en 2010, un autre proche parent des humains anatomiquement modernes (*Homo sapiens*) et des Néandertaliens a été découvert grâce à la génétique (Reich et al., 2010). Cette découverte est le fruit de l'analyse de l'ADN extrait d'une phalange retrouvée dans la grotte de Denisova qui est localisée dans l'Altaï, en Russie. Cet unique os nous livrait un génome appartenant à un groupe qui partageait une origine commune avec les populations Néandertaliennes : les Dénisoviens (Reich et al., 2010). Selon les modélisations génomiques, cette humanité aujourd'hui disparue dériverait d'un ancêtre commun avec les populations néandertaliennes datant de 381 000 à 473 000 ans AP (Meyer et al., 2016). Par cette étude, la possibilité d'explorer une diversité qui nous était jusqu'alors inconnue a été rendue possible. Cette humanité a elle aussi laissé des traces dans le génome de certaines populations actuelles et plus particulièrement chez les Mélanésiens (Reich et al., 2011; Qin and Stoneking, 2015; Choin et al., 2021). En effet, selon les populations, jusqu'à 3 % du génome des individus proviennent des Dénisoviens. Des événements d'introgression, c'est-à-dire de transfert de matériel génétique entre deux espèces ou populations, de fragments génomiques d'homininés archaïques dans le génome humain font partie de l'histoire évolutive de notre espèce et constituent un héritage qui a façonné les populations humaines actuelles (Racimo et al., 2017). Des allèles Néandertaliens peuvent en effet être partagés avec certaines populations non-

africaines et s'y retrouver avec des fréquences très importantes. Cet héritage peut être lourd de conséquences d'un point de vue de la santé publique. Par exemple, un ensemble de 76 polymorphismes a pu être repéré pour se retrouver à des fréquences alléliques importantes parmi les populations péruviennes et mexicaines (Racimo et al., 2017). Ces allèles sont localisés dans une région de 120 kb (kilobases) situé sur le chromosome 5 à proximité du gène *LARS* codant pour une enzyme cytoplasmique (la aminoacyl-tRNA synthétase, Giles et al., 1980). Ces variants seraient impliqués dans le syndrome d'insuffisance hépatique (Casey et al., 2012), une maladie assez rare (fréquence dans les populations africaines < 1 %) qui présente un fort taux de mortalité infantile (McClean and Davison, 2003). Ces marques d'introgressions archaïques se retrouvent en fréquence anormalement haute dans ces populations (populations péruviennes > 50 %, mexicaines > 20 %) et pourraient avoir de graves conséquences sur le développement de maladies hépatiques qui sont la cinquième cause de mortalité au Mexique (Méndez-Sánchez et al., 2005). Ce phénomène a aussi eu lieu au niveau d'une autre région du chromosome 10 mesurant 40 kb où 11 allèles présents chez les Dénisoviens se retrouvent très fréquemment dans les populations du Sud et de l'Est de l'Asie (> 50 %) à l'inverse des populations africaines (< 1 %). Cette région contient deux gènes, *LIPA* qui code pour une lipase (Warner et al., 1980), et *CH25H*, qui code pour une hydroxylase. Le premier est impliqué dans le métabolisme du cholestérol (Klima et al., 1993) et la maladie de Wolman (Aslanidis et al., 1996) ; le second dans le métabolisme du cholestérol (Lund et al., 1998), la maladie d'Alzheimer (Shibata et al., 2006) ; il est connu aussi pour son activité antivirale (Liu et al., 2013). Les conséquences de ces phénomènes à l'échelle des populations sont peu connues, mais les risques et enjeux de santé publique associés sont bien réels. Sans ces investigations, qui ont été rendues possibles par le développement des NGS et de la paléogénomique, l'origine de tels mécanismes nous aurait échappé. Ces études nous ont ainsi permis d'ouvrir une fenêtre sur un passé qui jusqu'alors nous était inconnu, et semble néanmoins affecter la santé des populations humaines actuelles.

L'année 2010 a marqué un tournant majeur pour la paléogénomique en ouvrant de multiples perspectives de recherche dans ce domaine. Depuis, le nombre d'études et de génomes anciens séquencés s'est drastiquement accru. La décennie suivante a vu émerger des projets à grandes échelles, avec le séquençage non plus d'un unique individu mais de dizaines voire de centaines (Allentoft et al., 2015; Haak et al., 2015; Damgaard et al., 2018; de Barros Damgaard et al., 2018; Narasimhan et al., 2019; Margaryan et al., 2020). Depuis les années 2000, l'essor et la démocratisation des NGS ont entraîné une forte diminution du coût et du temps de séquençage (van Dijk et al., 2018), offrant ainsi la possibilité aux chercheurs de séquencer un plus grand

nombre d'individus. Ces avancées ont offert l'opportunité de plonger dans les origines des populations humaines et de leur évolution au cours du temps afin d'explorer les nombreuses problématiques que l'anthropologie classique et l'archéologie peinent à résoudre. Ces études ont fondé la base d'une discipline qui ne cesse de repousser ses frontières temporelles et géographiques, malgré les contraintes de préservation de l'ADN. Les génomes d'homininés archaïques pré-néandertaliens (Meyer et al., 2016) et des humains modernes du paléolithique supérieur (Fu et al., 2013a, 2014; Seguin-Orlando et al., 2014; Sikora et al., 2017; Yang et al., 2017) ont pu être en partie séquencés. Malgré une concentration des études sur les populations Européennes et plus généralement les populations de l'ouest de l'Eurasie, les recherches paléogénomiques se sont ouvertes à d'autres zones géographiques du monde entier qui ne présentent pas toujours des conditions optimales de préservation de l'ADN. C'est le cas par exemple de l'Afrique (Llorente et al., 2015; Schlebusch et al., 2017; Fregel et al., 2018; Wang et al., 2020), de l'Océanie (Rasmussen et al., 2011; Lipson et al., 2018) ou encore de l'Amérique du Sud (Posth et al., 2018; Nakatsuka et al., 2020; Roca-Rada et al., 2021). Les techniques de biologie moléculaire employées pour le séquençage de génomes anciens ont été testées, adaptées et appliquées à tous types d'échantillons datant d'époques variées. Ces échantillons peuvent être de différentes natures (Orlando et al., 2021a) comme par exemple des tissus calcifiés (os, dents), des tissus mous, des tissus végétaux (bois, graines, herbiers, etc.) ou encore des sédiments (coprolithes ou carottes sédimentaires). Ils peuvent provenir de sites couvrant une large gamme temporelle dont l'âge peut aller de 100 ans à plusieurs centaines de milliers d'années. Aujourd'hui l'organisme le plus ancien jamais séquencé à partir de macrofossiles est un mammouth datant de 1.65 millions d'années (van der Valk et al., 2021) et un florilège de plantes et animaux datant de 2 millions d'années a pu être identifié à partir d'ADN contenu dans des sédiments du Groenland (Kjær et al., 2022).

2) Les applications de l'ADN ancien

Les études en ADN ancien ont permis de révolutionner notre compréhension de l'histoire évolutive des humains et d'une grande diversité d'organismes. Ces dernières années, les applications de l'ADN ancien se sont multipliées, allant de l'analyse des génomes des espèces éteintes à la reconstruction des migrations humaines préhistoriques. Dans le cadre de ce travail, j'ai choisi d'illustrer certaines des applications les plus couramment étudiées de l'ADN ancien,

telles que l'étude de la domestication animale et végétale, des pathogènes humains, ou encore de l'ADN environnemental.

a) L'étude de la domestication

(i) La domestication animale

La recherche en paléogénomique ne se limite pas à l'étude des humains ou de leurs proches parents. De nombreux autres organismes sont étroitement liés à notre histoire évolutive et ont eu un rôle prépondérant dans le développement des sociétés humaines. La domestication, qu'elle soit animale ou végétale, est un bon exemple de relation forte entre les humains et d'autres espèces et représente un tournant majeur dans l'histoire de nos sociétés. Ce processus complexe est le fruit d'une relation à long terme, façonnée par des facteurs écologiques, biologiques et culturels, entre des groupes d'humains et des plantes ou des animaux (Price and Bar-Yosef, 2011; Harlan et al., 2012; Frantz et al., 2020). Le mécanisme de la domestication a longtemps été sujet à débat en particulier concernant son déroulement, sa localisation ainsi que les ancêtres des espèces domestiques actuelles (Larson et al., 2014). La complexité de ce phénomène rend son étude difficile par le biais de l'archéologie et peut aboutir à des interprétations erronées. Un exemple marquant concerne l'origine de la domestication du cheval qui a longtemps été débattue. En effet, les archéologues avaient émis l'hypothèse que les chevaux domestiques actuels trouvaient leur origine chez les chevaux associés à la culture Botaï qui datent de 3 500 ans AP et proviennent du Kazakhstan (Outram et al., 2009). Cependant, le séquençage des chevaux de Botaï a démontré qu'ils sont en réalité les ancêtres directs des chevaux de Przewalski. Ils représentaient jusqu'alors les derniers chevaux sauvages vivant sur Terre, alors qu'ils descendaient en fait de chevaux domestiques (Gaunitz et al., 2018; Fages et al., 2019). Une étude récente a démontré que les ancêtres des chevaux domestiques actuels ont émergé des steppes de l'ouest de la Russie il y a 4 200 ans (Librado et al., 2021).

Le processus de domestication implique de nombreux mécanismes évolutifs tels que des remplacements importants de populations, des épisodes sélections ou encore, parfois, des métissages entre animaux domestiques et sauvages (Frantz et al., 2020). Cependant, grâce aux avancées récentes en paléogénomique, il est maintenant possible de mieux caractériser ces processus. En somme, la paléogénomique apporte de nouveaux éclairages sur la complexité de

la domestication animale, ouvrant ainsi de nouvelles perspectives pour la compréhension des relations entre les humains et les animaux domestiques.

(ii) La domestication végétale

L'origine des plantes domestiques est un sujet largement abordé en paléogénomique de par l'influence de l'agriculture sur l'histoire humaine, les enjeux actuels majeurs pour l'économie et les implications futures, notamment face aux changements climatiques. Le processus de domestication des plantes est lui aussi très complexe. En effet, malgré leur immobilité, on peut observer des flux génétiques importants et des croisements avec des espèces sauvages ou d'autres variétés que l'humain ne peut totalement contrôler, notamment durant premières étapes de la domestication (Oliveira et al., 2021). De plus, une certaine plasticité phénotypique peut être observée chez ces organismes avec de fortes influences environnementales (Gratani, 2014) ; cependant, cela ne reflète pas forcément la présence d'une plus grande diversité de cultivars, et peut amener les recherches archéologiques à des interprétations fausses à propos de l'histoire naturelle de ces lignées domestiques.

De nombreuses plantes domestiques ont fait l'objet de recherches paléogénomiques (raisin, blé, etc.), avec le maïs (*Zea mays* subsp. *mays*) comme l'une des espèces végétales les plus étudiées. Aujourd'hui, cette plante est répandue mondialement et le rendement des cultures peut avoir un fort impact économique dans les différentes régions où elle est cultivée (Rouf Shah et al., 2016). Le maïs a été domestiqué à partir de la téosinte (*Zea mays* subsp. *parviglumus*) il y a environ 9 000 ans dans le sud du Mexique par les populations locales (Piperno et al., 2009). Par la suite, une dispersion s'est opérée en Amérique Centrale il y a 7 500 ans dans un premier temps (Dickau et al., 2007), puis en Amérique du Sud 500 ans plus tard (Grobman et al., 2012; Lombardo et al., 2020) et enfin en Amérique du Nord il y a 4 000 ans AP (Merrill et al., 2009). Il faudra attendre les premiers contacts entre les Européens et les natifs Américains au XVIème siècle pour voir une dispersion mondiale de cette plante (Rebourg et al., 2003). Aux prémices de la domestication, certains phénotypes avantageux pour la culture et la consommation humaine auraient été favorisés par les conditions environnementales comme de faibles teneurs en CO_2 ou encore des températures favorables (Piperno, 2017; Piperno et al., 2019). Ces caractères auraient ensuite été fixés dans les populations par l'action humaine via la sélection artificielle des plants avantageux pour la culture. Certains gènes impliqués dans l'architecture

des inflorescences et des graines, dans le cycle circadien ou encore dans la biosynthèse du glycogène ont pu être préférentiellement sélectionnés par l'homme (Ramos-Madrigal et al., 2016; Vallebueno-Estrada et al., 2016). L'histoire de la domestication du maïs est complexe et les dispersions par vagues ainsi que les croisements avec les variétés sauvages rendent son étude difficile par le seul biais de l'archéologie. En effet, des événements d'introgressions et de croisement avec des sous-espèces sauvages (*Zea mays* subsp. *parviglumus*, *Zea mays* subsp. *mexicana*) ont aussi eu lieu lors des phases précoces de la domestication du maïs (van Heerwaarden et al., 2011; Wang et al., 2017) et de manière continue par la suite (Staller et al., 2006). L'analyse de l'ADN ancien provenant d'échantillons retrouvés sur des sites qui couvrent des périodes et des zones géographiques variées permet une meilleure compréhension de ces processus. L'étude de la domestication animale comme végétale ne peut se faire que par le biais d'approches pluridisciplinaires afin d'en appréhender toute la complexité et les implications sur les sociétés humaines.

b) Les pathogènes humains

L'ADN ancien, en plus d'offrir des opportunités d'étudier le génome des humains du passé, nous permet également d'explorer l'évolution des micro-organismes qu'ils portaient. Qu'ils soient commensaux ou pathogènes ces organismes partagent avec nous une partie de leur histoire évolutive. Parmi ces micro-organismes, les pathogènes représentent un intérêt particulier car ils sont la cause d'épidémies qui frappent encore les populations humaines actuelles, et leur étude est donc naturellement d'un grand intérêt pour la santé publique. Ces agents pathogènes évoluent depuis des millénaires en parallèle des populations humaines et ont été la cause de nombreuses épidémies historiques dévastatrices. Ils peuvent être d'origine virale, bactérienne ou eucaryote et la compréhension des mécanismes de leur évolution représente un enjeu majeur pour la microbiologie, l'épidémiologie et la médecine mais aussi pour l'anthropologie, l'archéologie et l'histoire. Aujourd'hui, si nous connaissons les formes modernes de ces pathogènes, leur origine et leur évolution restent souvent débattues. La nature même de ces micro-organismes anciens est difficile à déterminer en se basant uniquement sur les récits historiques et l'absence de stigmates sur les squelettes présents sur les sites archéologiques ne facilite pas cette tâche. En effet, la grande majorité des pathogènes provoquent des infections aigues et ne produisent pas de séquelles osseuses, il est donc difficile de caractériser la cause de ces épidémies anciennes. Il arrive néanmoins que des infections

chroniques puissent léser les tissus osseux et ainsi donner des indications sur l'origine infectieuse. La tuberculose par exemple, peut provoquer des lésions osseuses dans 1 à 5 % des cas observés (Resnick and Niwayama, 1987). Ces maladies chroniques n'entraînent pas forcément la mort de l'individu ce qui rend difficile l'étude de l'origine certaines épidémies.

Le premier génome de pathogène historique a été séquencé en 2005 et correspond au virus de la grippe espagnole de 1918 (Taubenberger et al., 2005), responsable de la mort de plus de 50 millions de personnes à travers le monde (Tumpey et al., 2005). Cette étude a permis d'appuyer l'hypothèse de l'origine aviaire de ce virus qui serait donc une zoonose comme c'est le cas pour d'autres virus grippaux (*Othomyxoviridae*). Le premier génome de pathogène ancien séquencé à l'aide des NGS a été quant à lui publié en 2011 (Bos et al., 2011) et correspond à l'agent causatif de la peste, la bactérie *Yersinia pestis*, qui décima la population européenne à partir du XIVème siècle. Ces travaux ont offert l'opportunité d'étudier en profondeur les pathogènes du passé. Depuis, l'étude des pathogènes anciens s'est étendue à une grande diversité d'agents infectieux tels que des bactéries comme *Treponema pallidum* (syphilis ; Schuenemann et al., 2018b), *Mycobacterium leprae* (lèpre, Mendum et al., 2014; Krause-Kyora et al., 2018a; Schuenemann et al., 2013, 2018a), *Mycobacterium tuberculosis* (Chan et al., 2013; Bos et al., 2014; Kay et al., 2015), *Vibrio cholerae* (Devault et al., 2014) ; des virus comme le virus de la variole (VARV, Duggan et al., 2016), le virus de l'hépatite B (VHB ; et des pathogènes eucaryotes tels que *Plasmodium falciparum* et *Plasmodium vivax* (Gelabert et al., 2016; Marciniak et al., 2016).

La révolution des techniques de séquençage et le développement de méthodes robustes de biologie moléculaire ont permis d'accéder au matériel génétique de ces micro-organismes qui est le plus souvent présent en fraction encore inférieure à celle de son hôte, pourtant elle-même déjà très faible dans les squelettes anciens. Ainsi, il a pu être démontré qu'au cours de transitions majeures dans leur mode de vie, les sociétés humaines ont été exposées à de nouveaux agents infectieux. C'est le cas de la révolution Néolithique qui, avec la domestication de nombreuses espèces animales et la croissance démographique qui lui est associée, a entraîné l'apparition de zoonoses qui se sont transmises rapidement et massivement au travers de ces groupes avec pour résultat des épidémies dévastatrices (Armelagos et al., 1991; Barrett et al., 1998). Cependant, tous les pathogènes ne sont pas d'origine zoonotique et c'est parfois même le contraire. *Mycobacterium tuberculosis*, l'agent infectieux responsable de la tuberculose, nous en donne un bon exemple. En effet, l'hypothèse traditionnelle expliquant l'apparition de cette maladie

supposait un transfert zoonotique au cours de la révolution Néolithique (Clark et al., 1987; Manchester, 1984). Cette hypothèse était soutenue par le fait que de nombreuses espèces de *Mycobacterium* peuvent infecter des animaux domestiques pour ensuite être transmises aux humains. C'est le cas notamment de *Mycobacterium tuberculosis* subsp. *bovis*, qui est la plus ancienne preuve paléopathologique d'infection par *M. tuberculosis* et qui date du Néolithique (Smith et al., 2009). L'étude de restes d'animaux plus anciens a suggéré que ces bactéries étaient capables de les infecter bien avant la période Néolithique, c'est-à-dire avant leur domestication (Rothschild et al., 2001). Cependant, des études plus récentes ont démontré qu'au contraire les souches infectant les animaux descendent de la souche humaine qui aurait émergé il y a 70 000 ans et aurait accompagné les migrations des populations humaines depuis leur sortie d'Afrique (Comas et al., 2013).

Avec l'arrivée des colons Européens aux Amériques, de nouveaux pathogènes ont été introduits dans des populations présentant une immunité naïve face à ces agents infectieux. Des épidémies dramatiques ont décimé ces populations, ce qui a précipité le déclin des civilisations indigènes (Dobyns, 1993; Crosby, 1994; Acuna-Soto et al., 2004). Le « cocoliztli » est un autre exemple d'épidémie n'ayant pas une origine directement animale. Cette maladie infectieuse a ravagé les populations de l'actuel Mexique au cours du XVIème siècle (Brown, 2019). L'origine de cette épidémie restait ambiguë (Warinner et al., 2012) jusqu'en 2018 lorsque la paléogénomique a permis de proposer un candidat sérieux comme source de cet événement dramatique. Le séquençage de 10 individus du XVIème siècle provenant d'un cimetière situé à Oaxaca au Mexique a en effet permis de mettre en évidence la présence de la bactérie *Salmonella enterica* subsp. *enterica* et plus particulièrement le sérovar Paratyphi C, responsable de la fièvre entérique (Vågene et al., 2018). Même s'il est possible que ce pathogène ne soit pas responsable de la totalité des décès observés, il reste un candidat robuste pour expliquer la source de l'épidémie de « cocoliztli » au cours des années 1545-1550 au Mexique. Ce type d'étude permet donc d'explorer la question plus large de l'émergence de nouveaux pathogènes dans le passé. Les sources de ces épidémies ont été sujet à débat pendant des siècles et la paléogénomique offre l'opportunité d'explorer ces questions irrésolues par l'histoire et l'archéologie.

c) L'ADN environnemental

Il est possible, depuis 20 ans, d'analyser l'ADN environnemental contenu dans les sédiments et les résultats qui ont été générés au cours des dernières années sont très prometteurs. Ce matériel particulier a pour avantage d'offrir une fenêtre sur le passé d'une autre échelle. En effet, il est devenu possible de recenser la composition taxonomique des communautés entières d'organismes, mais aussi de révéler la présence d'espèces en l'absence de matériel archéologique et/ou paléontologique. En 2003, des sédiments provenant du permafrost sibérien et d'une grotte de Nouvelle-Zélande ont été analysés pour leur contenu en ADN (Willerslev et al., 2003). Ces travaux ont permis de déceler la présence d'espèces aujourd'hui éteintes ; il ne pouvait donc s'agir de contaminations. L'opportunité de caractériser la composition des écosystèmes passés en l'absence de fossile repose désormais sur des techniques qui permettent de focaliser l'effort de séquençage sur des biomarqueurs spécifiques aux taxons recherchés. Cependant, il existe certaines contraintes qui sont liées au matériel et une attention particulière doit être portée à l'élimination de la contamination présente sur le site d'échantillonnage (Der Sarkissian et al., 2017). D'autres phénomènes peuvent entraîner des difficultés pour l'étude de l'ADN sédimentaire et notamment le lessivage stratigraphique qui peut introduire des erreurs conséquentes lors des analyses (Haile et al., 2007). En effet, des mouvements de l'ADN présent dans les différentes strates sédimentaires peuvent avoir lieu à cause de mouvement du sédiment en lui-même ou encore de fluides, comme l'eau. Le résultat de ce phénomène est la présence d'ADN de certaines espèces dans des couches géologiques où elles ne devraient normalement pas se trouver. Ce phénomène ne pose pas de problème sur des sédiments congelés (permafrost), mais pour des environnements tempérés par exemple, cela peut fortement compliquer les analyses voire invalider certains résultats (Poinar et al., 1998; Hofreiter et al., 2000).

La recherche d'ADN ancien dans les sédiments constitue un domaine à part entière qui nécessite des approches et/ou méthodes spécifiques à ce matériau. Les premières analyses employaient le clonage moléculaire de métabarcodes amplifiés par PCR. Cette méthode permet ainsi de révéler la présence d'espèces mais reste limitée et peut introduire certains biais (Ruppert et al., 2019), notamment en raison de la fragmentation importante des molécules anciennes (Ziesemer et al., 2015). L'utilisation des NGS a permis d'accéder à toute une diversité moléculaire de plantes et d'animaux en séquençant en parallèle des quantités importantes d'ADN (Binladen et al., 2007). Aujourd'hui de nombreuses techniques se sont développées pour accroître l'efficacité de séquençage et en diminuer les coûts, notamment via l'utilisation de méthodes de capture ciblant le génome (entier ou en partie) des espèces d'intérêt (Slon et al., 2017; Murchie et al., 2021). L'un des avantages majeurs de l'ADN environnemental

est la détection d'organismes appartenant à tous les domaines du vivant en l'absence même de macrofossiles, pourvu que suffisamment d'ADN soit préservé (Haile et al., 2007; Murchie et al., 2021). Cependant, la forte fragmentation de l'ADN environnemental ainsi que l'utilisation de bases de données incomplètes peuvent limiter la résolution de l'identification taxonomique (Winther Pedersen et al., 2015).

L'origine de l'ADN contenu dans les sédiments reste floue et il est supposé qu'il provienne de la dégradation de tissus, de fèces, d'urine, de microfossiles et de fragments microscopiques d'os et de dents confondus dans le sédiment (Willerslev et al., 2003; Slon et al., 2017). L'analyse minutieuse des sédiments présents dans la grotte de Denisova a cependant révélé que la répartition de l'ADN provenant de mammifères varie considérablement, à l'échelle du millimètre (Massilani et al., 2022). De plus, l'ADN est particulièrement concentré dans des particules de petite taille, comme des fragments d'os ou des coprolithes, qui pourraient être des sources importantes d'ADN sédimentaire. Dans tous les cas, il reste que l'étude de l'ADN environnemental est un outil puissant pour la paléontologie et la paléoécologie car elle permet d'explorer la diversité des écosystèmes anciens mais aussi de mettre en évidence la succession des espèces au cours du temps dans un endroit donné. Certaines études ont ainsi permis de retrouver la présence de mammouths ou de chevaux en Alaska (Haile et al., 2009) ainsi que dans le Yukon (Murchie et al., 2021), à des époques où ils étaient supposés avoir disparu (Wang et al., 2021). L'ADN environnemental a aussi permis de caractériser la répartition spatio-temporelle de certaines espèces, y compris des homininés archaïques tels que Néandertal et Denisova (Slon et al., 2017). Il est également possible d'étudier les effets des transitions environnementales majeures sur les écosystèmes telles que des changements climatiques (Coolen and Overmann, 1998; Willerslev et al., 2007, 2014; Pedersen et al., 2016; Seersholm et al., 2020). Ce domaine de la recherche en paléogénétique nous offre désormais l'opportunité d'étudier les successions d'espèces sur de grandes échelles spatio-temporelles. Malgré les limites en termes de contamination ou encore d'inhibiteurs enzymatiques, les résultats obtenus au cours des dernières années en font un champ de la paléogénomique très prometteur avec de nombreuses applications pour l'écologie et notamment pour l'étude de l'impact des activités humaines et du changement climatique sur les écosystèmes passés et actuels.

3) Les méthodes pour l'étude de l'ADN ancien

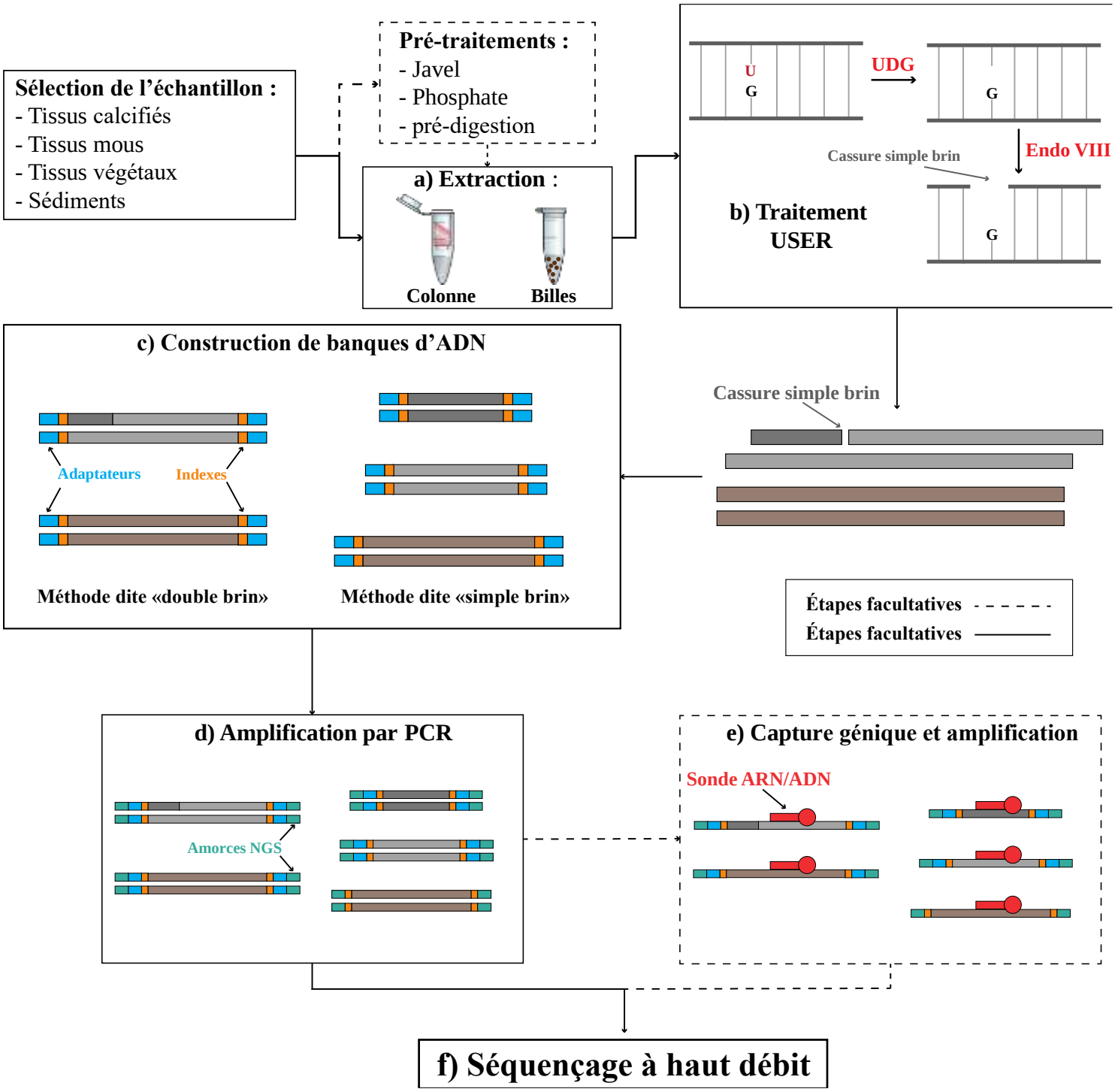


Figure 1 : Les différentes étapes expérimentales au traitement de l'ADN ancien. (a) L'extraction de l'ADN après échantillonnage des restes archéologiques. (b) Le traitement USER partiel ou complet pour exciser les uraciles, dérivés de la désamination post-mortem des cytosines et incorporés *in situ* dans les séquences anciennes. (c) La construction de banques d'ADN selon les méthodes dites « double brin » et « simple brin », est suivie d'une (d) étape d'amplification par PCR. (e) La capture du génome ciblé à l'aide de sondes (ADN ou ARN)

suivie d'une amplification par PCR. (f) Les banques d'ADN amplifiées et/ou capturées sont enfin séquencées.

a) L'extraction de l'ADN ancien

L'ADN ancien présente des caractéristiques très spécifiques qui rendent son extraction compliquée à partir de restes anciens souvent très dégradés. La fragmentation est une des spécificités majeures des séquences anciennes qui mesurent en général 35 à 100 pb seulement (Briggs et al., 2007; Dabney et al., 2013b). Dans la majorité des cas, les fragments d'intérêt sont représentés en faible abondance en comparaison à la fraction d'ADN contaminant. En effet, dans la plupart des cas, le taux d'ADN endogène dans les extraits est inférieur à 1 % (Carpenter et al., 2013; Damgaard et al., 2015), même si d'exceptionnelles conditions de conservation peuvent offrir des taux bien plus importants pouvant atteindre 70-80 %, voire plus (Rasmussen et al., 2010; Reich et al., 2010). Des techniques spécifiques ont été développées pour permettre non seulement de récupérer l'ADN ancien mais aussi de réduire la fraction de molécules contaminantes qui pourraient fausser les analyses ou inhiber les réactions enzymatiques impliquées dans le processus de traitement de l'ADN (Höss and Pääbo, 1993; Green et al., 2009; Korlević et al., 2015).

L'extraction de l'ADN peut être réalisée sur différents types d'échantillons dont la conservation et la concentration en ADN varient selon les conditions environnementales, l'individu ou encore le tissu étudié. Les travaux en paléogénomique se focalisent majoritairement sur les dents et les os qui sont retrouvés fréquemment sur les sites archéologiques et peuvent être de bonnes sources d'ADN. En effet, ces tissus calcifiés peuvent résister à la dégradation durant des siècles voire des millénaires selon les conditions de préservations offertes par le milieu. À l'inverse, les tissus mous vont être les premières victimes de la dégradation et sont, par conséquent, retrouvés peu fréquemment ; ils sont aussi souvent, à l'exception des poils et des cheveux, sujets à la contamination (Seguin-Orlando et al., 2015). Des études ont démontré que certaines parties du squelette renferment des taux d'ADN endogène plus importants que d'autres. C'est le cas par exemple du cément dentaire (le tissu recouvrant la dentine) par rapport à la dentine qui était jusqu'alors préférentiellement employée (Adler et al., 2011; Higgins et al., 2013; Damgaard et al., 2015). Un os a été particulièrement étudié car il présente souvent des taux exceptionnels d'ADN endogène (Gamba et al., 2014;

Pinhasi et al., 2015), supérieurs à ceux du cément (Hansen et al., 2017) : c'est la partie pétreuse de l'os temporal qui porte l'oreille interne. Appelé os pétreux, c'est la partie la plus dense du squelette des mammifères (Lam et al., 1999), ce qui lui conférerait une résistance importante à la contamination (Gamba et al., 2014). Ces deux éléments, le cément et l'os pétreux, sont privilégiés dans les études paléogénomiques. Il existe cependant une forte variation interindividuelle dans la conservation de l'ADN et la rareté de certains échantillons restreint, bien entendu, les analyses au matériel disponible. En plus des os et des dents, les techniques d'extraction ont été testées sur de nombreux types de tissus comme les cheveux (Barnes et al., 2002; Gilbert et al., 2004), fèces (Poinar et al., 1998), ou encore les tissus momifiés (Pääbo, 1985; Ermini et al., 2008; Gamba et al., 2016).

L'extraction de l'ADN ancien consiste premièrement à générer de la poudre à partir des os (ou des dents selon le matériel étudié). La production mécanique de la poudre peut s'effectuer de différentes manières (forage, mortier/pilon, broyeur cryogénique), mais la vitesse de forage et la chaleur provoquée par la friction peuvent provoquer des dégradations supplémentaires de l'ADN (Adler et al., 2011). La seconde étape implique l'utilisation d'un tampon de lyse afin de digérer la poudre générée et ses constituants. Il est généralement composé d'EDTA (acide éthylène diamine tétra-acétique), afin de chélater les ions bivalents pour libérer la matrice osseuse, et de protéinase K pour désintégrer le collagène et l'hydroxyapatite présents dans les dents et les os. Pour les cheveux, du SDS (Sodium Dodecyl Sulfate) et du DTT (dithiothréitol) en plus de la protéinase K sont employés pour dénaturer les protéines, y compris la kératine. Un prétraitement peut avoir lieu avant l'étape de lyse afin de réduire la part d'ADN contaminant un utilisant un tampon de phosphate ou de sodium hypochlorite, cela peut néanmoins réduire le taux d'ADN endogène (Warinner et al., 2014; Korlević et al., 2015; Basler et al., 2017; Nieves-Colón et al., 2018). Une étape de prédigestion de 15 à 30 minutes (Damgaard et al., 2015), suivie d'une digestion complète (Gamba et al., 2016) sont aussi souvent utilisées dans le but d'accroître les taux d'ADN endogène. En effet, on pense qu'une importante concentration de micro-organismes contaminants est présente en surface des échantillons anciens alors que l'ADN endogène est quant à lui confiné à des « micro niches » osseuses (Ginolhac et al., 2012). L'ajout d'une courte étape de prédigestion en amont de l'extraction permet d'éliminer une grande partie de ces contaminations et d'accroître la part de la fraction d'intérêt. La dernière étape est la plus critique car elle consiste à purifier l'ADN ancien. Elle détermine ainsi le taux de récupération de l'ADN, c'est-à-dire la quantité d'ADN purifiée à l'issue de l'extraction, et l'efficacité de la réduction de la fraction contaminante.

Les premières méthodes de purification développées utilisaient le phénol chloroforme par des séparations répétées en phases hydrophobes et aqueuses (Hagelberg et al., 1989; Hagelberg and Clegg, 1991). Cette méthode est efficace car elle permet l'élimination à moindre coût des inhibiteurs enzymatiques, ce qui l'a rendue particulièrement populaire (Miller et al., 2008; Keller et al., 2012; Hofmanová et al., 2016). Cependant l'utilisation de phénol chloroforme comporte des risques pour l'expérimentateur liés à sa toxicité importante et d'autres méthodes efficaces existent pour réduire l'effet de ces inhibiteurs (Moreira, 1998; Kemp et al., 2006). Par la suite, des approches plus simples ont été développées et se basent sur l'utilisation de colonnes de silice. La silice, de par sa forte affinité pour l'ADN en milieu acide, permet de concentrer et de purifier l'ADN contenu dans les extraits (Yang et al., 1998). Une méthode alternative basée sur l'utilisation de particules de silice en solution a été développée une décennie plus tard (Rohland and Hofreiter, 2007a). Ces particules sont en solution dans un tampon dit de « liaison » (Binding buffer) composé de sels chaotropiques qui favorisent la liaison des molécules d'ADN à la silice. Des lavages successifs à l'éthanol permettent d'éliminer les produits chimiques et les sels. L'ADN fixé aux particules sera ensuite séparé de ces dernières lors de l'élution. Cette technique est très efficace pour la récupération de l'ADN et permet des taux de succès d'amplification par PCR plus importants. Trois ans plus tard, cette méthode sera encore améliorée avec une simplification des étapes de lavage et d'élution avec l'utilisation de colonnes de silice disponibles dans le commerce. Grâce au développement et la simplification de ces approches, le nombre d'échantillons traités simultanément a pu être fortement augmenté (Rohland et al., 2010). Ce protocole a ensuite été adapté pour de plus grands volumes afin d'accroître le rendement de l'extraction d'ADN. L'ajout d'un tampon très concentré en chlorure de guanidium a également permis d'augmenter l'efficacité de l'extraction des fragments mesurant 35 à 150 pb, ce qui correspond à la gamme de taille des séquences d'ADN ancien (Dabney et al., 2013a; Gamba et al., 2016). De nos jours, ce protocole est certainement le plus employé par les laboratoires d'ADN ancien. Il existe donc deux types de protocoles d'extraction qui diffèrent par leurs méthodes de purification qui sont basées sur une purification sur colonne ou en solution. Leur efficacité varie selon le matériel mais la purification sur colonne semble plus efficace et offre une diversité de séquences plus importante. Le protocole de purification en solution est limité en raison de la perte de courtes séquences qui sont souvent majoritaires dans les échantillons anciens. Cependant, il peut permettre d'ajuster la quantité de silice sans être limitée par la taille de la colonne lors de l'extraction de grandes quantités d'ADN.

b) Le traitement USER

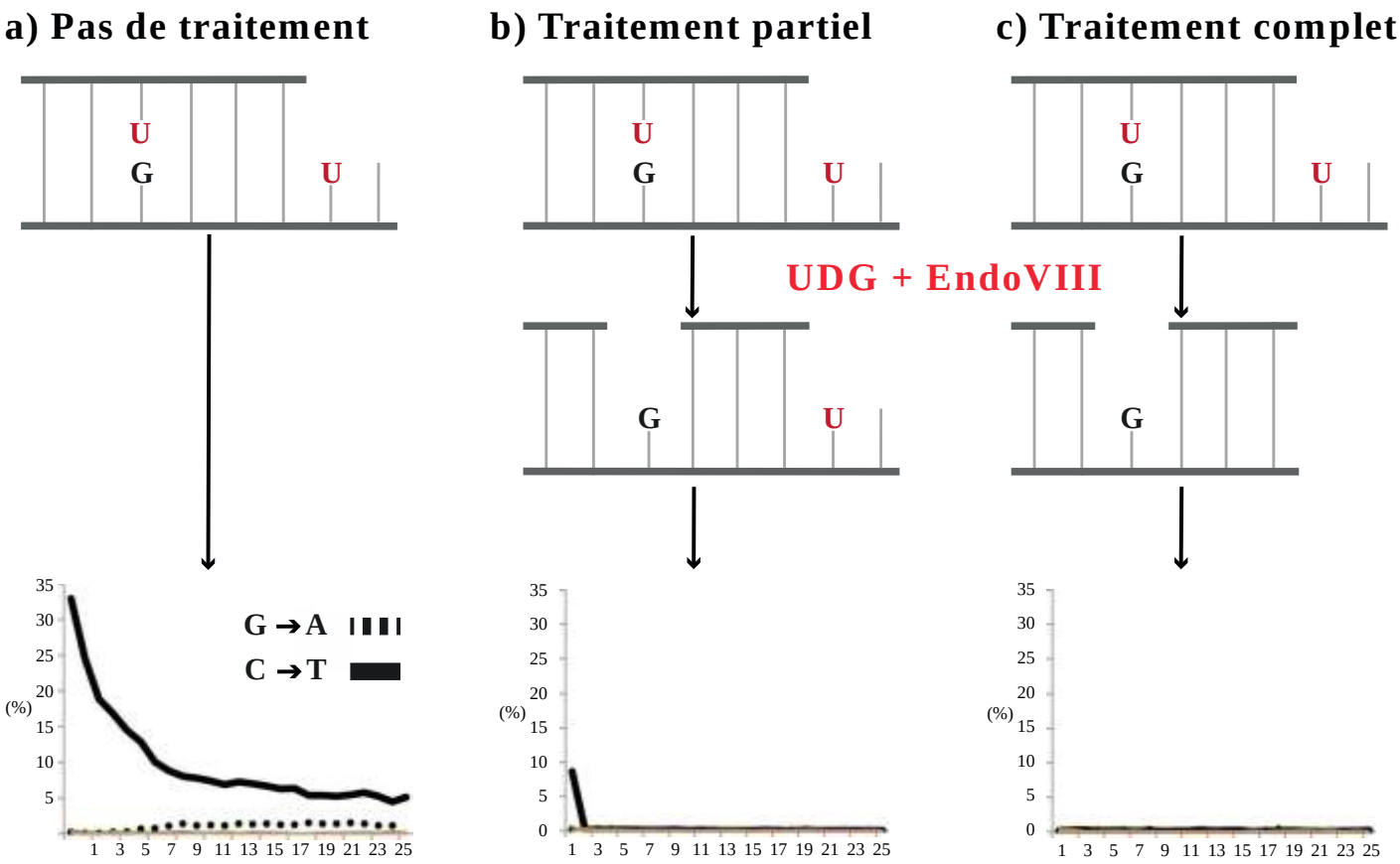


Figure 2 : Principe du traitement USER et profils de substitutions (C → T et G → A) attendus selon le traitement (les graphiques sont issus de Rohland et al. 2015). (a) Pas de traitement USER. (b) Traitement USER partiel pour retirer les uraciles situés à l'intérieur des fragments d'ADN (mais pas à leurs extrémités). (c) Traitement USER complet pour retirer la majorité des cytosines désaminées en uraciles.

L'ADN ancien présente des dégradations post-mortem qui peuvent introduire des biais lors des analyses. Ces dommages sont dus à l'arrêt des processus physiologiques suite à la mort de l'individu. L'une des modifications principales est la désamination des cytosines en uracile (Briggs et al., 2007), ou en thymine lorsque ces cytosines sont méthylées (Pedersen et al., 2014). Ces mutations artificielles produisent des taux d'erreur importants et apparaissent plus particulièrement aux extrémités simple brin des molécules d'ADN ancien. Ces dommages s'accumulent au cours du temps (Sawyer et al., 2012), la cinétique de dégradation est cependant principalement déterminée par les conditions de conservation offertes par l'environnement

(Orlando et al., 2013; Wagner et al., 2018). Les biais introduits par ces mutations artificielles peuvent mener à des interprétations incorrectes (Ho et al., 2007; Axelsson et al., 2008). Pour en limiter l'impact, des traitements enzymatiques ont été développés et notamment le traitement à l'uracile-DNA glycosylase (UDG) couplé à l'endonucléase VIII (Figure 2). Ces deux enzymes permettent respectivement de retirer les cytosines désaminées et d'exciser les sites abasiques. Ce traitement, appelé USER (New England Biolabs), intervient suite à l'étape d'extraction et a obtenu de bons résultats dans la réduction du nombre de mutations de ce type (Briggs et al., 2010). Il est aussi possible de réaliser un traitement partiel des extraits qui affecte uniquement les bases situées à l'intérieur des séquences et non pas celles localisées à leurs extrémités (Figure 2b). Cette approche permet un gain du temps tout en authentifiant l'ADN ancien qui présentera un excès de mutations artificielles aux extrémités (Krause et al., 2010; Jónsson et al., 2013; Rohland et al., 2015) qui pourront être éliminées facilement lors des analyses bioinformatiques. Elle offre néanmoins l'inconvénient de réduire la taille des molécules d'ADN ancien, et ainsi de conduire à la perte d'une fraction de l'ADN extrait, qui se retrouvera, post-séquençage, trop court pour pouvoir être aligné avec certitude le long de génomes de référence.

c) La construction de banques d'ADN ancien

Les méthodes de construction de banques d'ADN ancien reposaient initialement sur le clonage moléculaire des produits de PCR (Polymerase Chain Reaction) à l'aide de plasmides bactériens (Noonan et al., 2005). Cette approche présentait cependant certaines limites car elle ne permettait pas de séquencer un grand nombre de séquences en parallèle et son coût restait important. Aujourd'hui, la construction de banques ADN compatibles avec les plateformes de séquençage est largement utilisée ; elle se base sur l'ajout de courtes séquences connues aux extrémités des fragments appelés adaptateurs ou indexes. Ces adaptateurs permettent l'amplification par PCR et peuvent contenir des indexes uniques facilitant leur reconnaissance par séquençage, ce qui garantit l'identification des échantillons lorsque différentes banques sont regroupées ensemble (on parle de (dé-)multiplexage). Cela assure aussi un contrôle de la contamination au cours des différentes étapes expérimentales suivant la construction des banques car seules les séquences correspondant au produit de l'extraction portent ces marqueurs. Deux méthodes principales de construction de banques d'ADN sont communément utilisées dans l'étude de l'ADN ancien. La première méthode, qui est la plus ancienne, s'appuie

sur des molécules d'ADN double brin (Figure 3a), tandis que la seconde s'appuie sur des molécules réduites à l'état simple brin (Figure 3b), par exemple par chauffage ou dénaturation chimique.

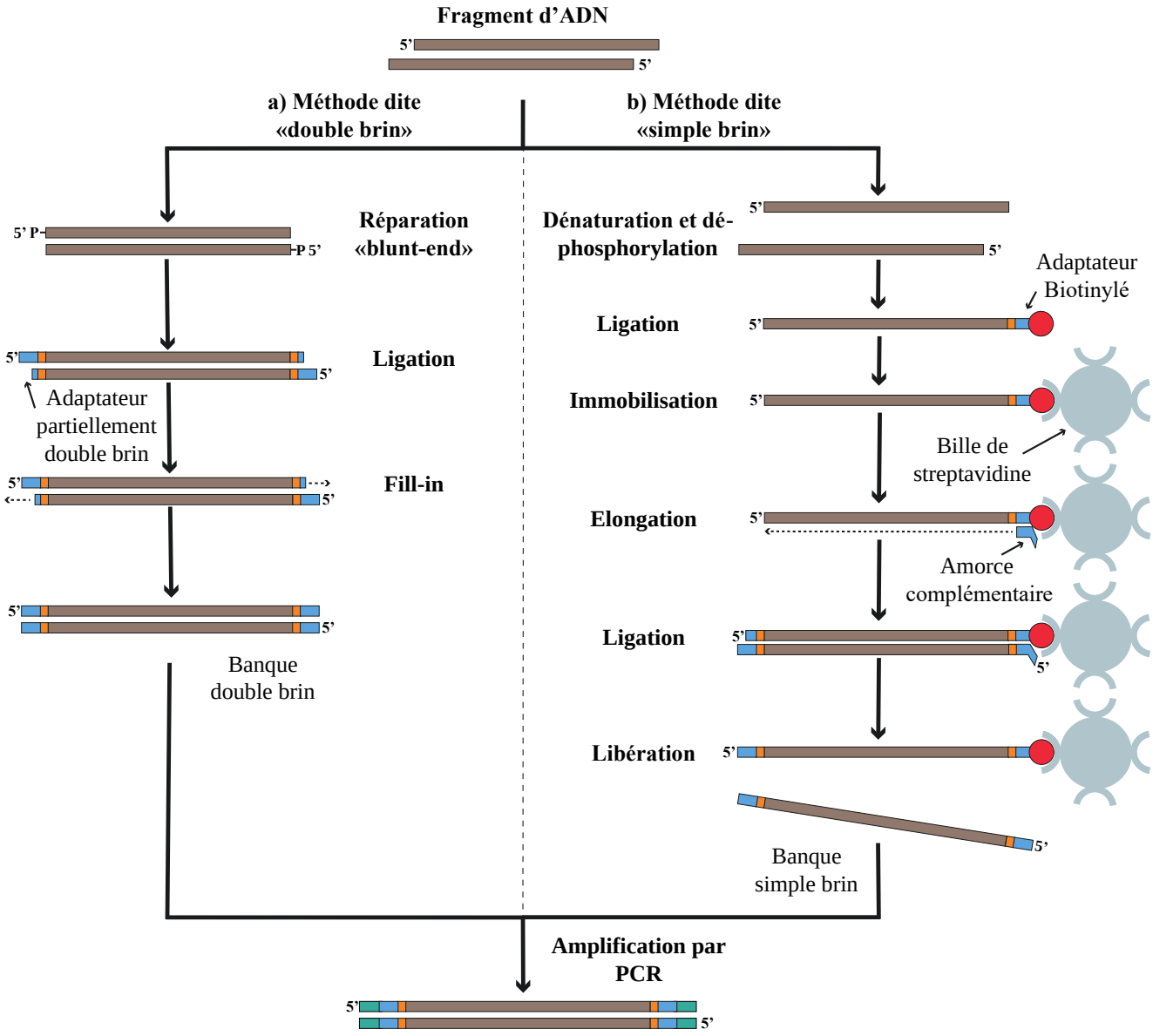


Figure 3 : Les différentes méthodes de construction de banques d'ADN ancien. (a) La méthode dite « double brin ». (b) La méthode dite « simple brin ».

(i) Les banques dites « double brin »

Les premières étapes de la construction des banques d'ADN consistent à réparer les extrémités des fragments (« blunt end repair ») pour ensuite y lier un adaptateur double brin selon deux méthodes différentes (« ligation »). La première dérive d'une technique prévue à l'origine pour le séquençage 454 (Margulies et al., 2005), et consiste en une ligation de deux adaptateurs partiellement double brin aux séquences dont les extrémités ont été « réparées » au préalable. Par la suite, cette méthode a été adaptée pour le séquençage sur les plateformes Illumina en utilisant des adaptateurs compatibles avec ces dernières (Meyer and Kircher, 2010). Elle a ensuite été modifiée en diverses versions, notamment par l'emploi d'adaptateurs en forme de Y avec une extrémité protrudente 5' composées d'une thymine qui pourra se lier aux brins d'ADN complémentaires protrudents démarrant par une adénine. La méthode peut donc introduire des biais dans la composition en base des fragments introduits par ligation, en particulier, elle limite la capacité de ligation des adaptateurs aux séquences d'ADN ancien commençant par des thymines, et donc des cytosines désaminées (Seguin-Orlando et al., 2013). Son utilisation est aujourd'hui dépréciée.

(ii) Les banques dites « simple brin »

La méthode simple brin a été développée plus récemment et présente une sensibilité plus importante pour les échantillons présentant une faible quantité d'ADN. Cette approche permet de lier les adaptateurs aux fragments présentant des extrémités 3'-OH, comme celles résultant de coupures simple brin. La méthode s'appuie par ailleurs sur l'utilisation d'adaptateurs biotinylés qui permettent d'éviter la perte trop importante d'ADN lors des purifications post-ligation car les séquences vont être liées à des billes enrobées de streptavidine à chaque étape de réaction. Enfin, les coûts de cette méthode ont pu être considérablement réduits avec le remplacement de la CircLigase par la T4DNA ligase qui est aussi compatible avec l'automatisation (Gansauge et al., 2017). L'utilisation de la T4DNA ligase permet de lier les adaptateurs en plus grande concentration et à moindre coût que la CircLigase, et donc d'optimiser le rendement de la ligation. Les banques d'ADN simple brin permettent ainsi d'accroître le nombre de séquences anciennes converties en banques ADN et de sélectionner les molécules présentant des dommages post-mortem typiques de l'ADN ancien, cela aboutit à une réduction des coûts de séquençage. L'efficacité de cette approche a été optimisée grâce à l'utilisation d'enzymes bon marché et la réduction du temps de manipulation. L'efficacité de

conversion, à savoir le taux de séquences converties en banques ADN par rapport aux séquences non-converties, a fortement augmenté.

d) La capture par hybridation

Les échantillons anciens présentent une limite majeure, c'est la faible fraction en ADN endogène qu'ils contiennent. Cela restreint l'efficacité du séquençage et peut rendre les manipulations très coûteuses lorsque l'on cherche à atteindre de fortes profondeurs de séquençage (le nombre moyen de séquences couvrant chaque position du génome) ainsi qu'une bonne couverture des régions ciblées (la proportion de positions dans le génome qui ont été séquencées avec une certaine profondeur). Un moyen d'accroître la fraction d'ADN endogène est l'utilisation de méthodes d'enrichissement ou de capture ciblant les génomes ou des régions d'intérêt. Ces approches de capture par hybridation de sondes ADN ou ARN permettent de cibler des régions spécifiques comme des génomes d'organelles (mitochondries, chloroplastes), des exomes, des chromosomes entiers, voire même le génome entier (Briggs et al., 2009; Burbano et al., 2010; Maricic et al., 2010; Carpenter et al., 2013; Fu et al., 2013b; Castellano et al., 2014; Enk et al., 2014). La fraction d'ADN endogène peut être considérablement accrue sur les portions ciblées du génome. Cela permet de produire à moindre coût des quantités de données suffisantes à partir d'échantillons de mauvaise qualité. Par exemple, Enk et ses collaborateurs ont réussi à obtenir plus de 70 % d'ADN endogène après la capture d'une banque d'ADN qui ne contenait initialement que 5 % d'ADN endogène (Enk et al., 2014).

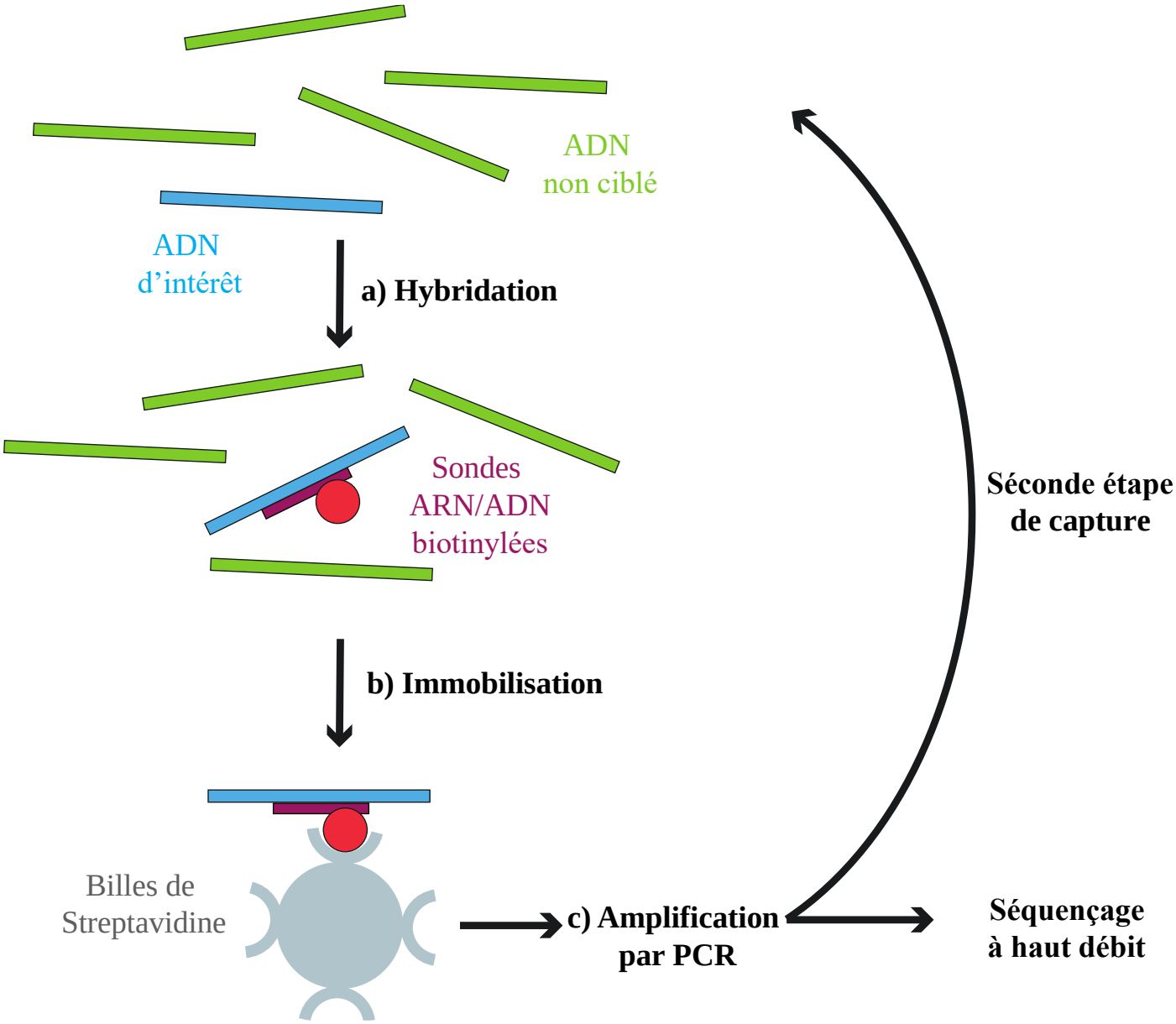


Figure 4 : Le principe de la capture génique. Les différentes étapes de la capture sont schématisées avec (a) l'étape d'hybridation aux molécules d'ADN complémentaires (ADN d'intérêt) des sondes ARN ou ADN biotinylées, (b) l'étape d'immobilisation à l'aide de billes recouvertes de streptavidine et de purification des complexes sondes-banque, (c) l'étape d'amplification des séquences d'intérêts suivie d'une seconde capture ou du séquençage des produits de cette dernière.

Les premières approches de capture qui ont été développées utilisent des plaques appelées « microarrays » qui contiennent des sondes immobilisées dessus (Burbano et al., 2010). Actuellement, il existe de nombreuses techniques de capture en solution qui reposent sur l'emploi de sondes ADN ou ARN. Ces méthodes peuvent être utilisées selon le type

d’échantillon à analyser, sa préservation ou encore les loci d’intérêts. La méthode hyRAD (hybridation by Restriction site Associated DNA markers) a prouvé son efficacité pour la capture de séquences fortement dégradées qui proviennent par exemple de collections de muséums. Cette méthode repose sur l’utilisation d’enzymes de restriction pour produire des sondes à partir d’ADN frais de l’espèce d’intérêt. Les fragments ainsi obtenus sont convertis en banques d’ADN amplifiables (ou exprimables par transcritpion) à volonté pour produire les sondes ADN (ARN) à utiliser pour la capture. Après en avoir retiré les indexes, ces sondes sont biotynilées pour pouvoir récupérer les produits de capture à l’aide de billes magnétiques enrobées de streptavidine. La capture hyRAD est une méthode efficace à coût limité pour étudier l'ADN ancien car elle offre une forte sensibilité, une réduction de la contamination et une couverture génomique élevée (Suchan et al., 2016, 2022a, 2022b). Cette approche présente néanmoins des limites qui peuvent être notamment dues à la qualité de l'ADN de départ et au degré de divergence entre les spécimens étudiés, qui se traduit par un polymorphisme de la séquence d'ADN au niveau des sites de restriction ciblés pouvant conduire à une perte d’une partie d’entre eux entre des clades divergents. Cette limitation réduit considérablement le nombre de loci orthologues qui peuvent être étudiés sur l'ensemble des spécimens analysés, ce qui peut conduire à des estimations biaisées de la diversité génétique (Arnold et al., 2013; Davey et al., 2013; Gautier et al., 2013). D’autres techniques existent et ciblent des régions spécifiques du génome comme par exemple la capture MBD (Methyl Binding Domains). Cette méthode n’emploie pas de sondes et se base sur l’affinité des protéines MBD à se lier aux dinucléotides CpG méthylés. Elle permet ainsi de cibler des régions génomiques selon leur degré de méthylation. Cette approche a été testée avec succès et a permis d’enrichir des régions méthylées du génome sur des types d’échantillons variés datant de différentes époques (Seguin-Orlando et al., 2015). Cependant, des biais importants sont à souligner car ce procédé cible principalement les longues séquences et les sites riches en dinucléotides CpG. De plus, la désamination des cytosines au niveau des allèles méthylés introduit des thymines en lieu et place des cytosines ; elle représente donc une contrainte supplémentaire, en atténuant l’efficacité d’appariement sonde-cible. Cette approche est donc plus appropriée pour l’étude d’échantillons bien conservés qui présentent une fragmentation et une désamination de l’ADN moins importante.

Les méthodes de capture sollicitant l’hybridation de sondes biotinylées produites par synthèse chimique avec les banques d’ADN sont en fait celles qui sont les plus utilisées. La liaison de ces sondes avec les séquences complémentaires d’intérêt produit des complexes

sondes-banque. Ces structures pourront ensuite être immobilisées pour être purifiées, ce qui élimine les fragments non-complémentaires aux sondes. Les fragments d'intérêt sont ensuite amplifiés par PCR pour éliminer les sondes et produire suffisamment de matériel en vue du séquençage. La capture permet ainsi de réduire les coûts de séquençage tout en augmentant son efficacité (Mamanova et al., 2010). En général, deux séries d'hybridation sont réalisées, chacune suivie d'une étape d'amplification et de purification. La capture peut néanmoins entraîner une hausse de la clonalité (Ávila-Arcos et al., 2011, 2015), c'est-à-dire une augmentation du nombre de doublons de PCR. Deux étapes de capture sont suffisantes pour limiter la proportion de séquences non ciblées (Hernandez-Rodriguez et al., 2018), mais elles entraînent souvent une forte augmentation de la clonalité.

Ces techniques de capture ont prouvé leur efficacité et ont permis de reconstruire, au moins en partie, des génomes à partir d'échantillons récalcitrants, ou encore d'obtenir de larges volumes de données pour des polymorphismes orthologues (« 1,240k » SNP capture ; Mathieson et al., 2015) mais ces techniques peuvent présenter certaines limites. En effet, il est difficile de concevoir des sondes pour le génome entier ou des régions du génome d'espèces éteintes sans information au préalable, ni de données modernes provenant d'espèce étroitement liées (Lauterbur, 2019). Certains biais peuvent également apparaître et notamment le biais détermination (« *ascertainment bias* »). En effet, lorsque des variants (SNPs) ou certaines régions du génome sont préférentiellement ciblés, il peut arriver de s'écarter de la représentation moyenne du génome, et la perte d'information concernant des régions non ciblées peut entraîner des erreurs lors des analyses (McTavish and Hillis, 2015). Enfin, malgré la forte sensibilité de la capture par hybridation (capacité à récupérer les séquences ciblées dans un échantillon contenant de multiples sources d'ADN) ainsi que son efficacité (capacité à récupérer de manière sélective les séquences ciblées) varient selon différents facteurs et paramètres de capture. Une grande variabilité peut ainsi être observée dans les taux d'enrichissement (mesure quantitative permettant de déterminer l'enrichissement ou l'amplification d'une séquence spécifique par rapport aux séquences non ciblées) entre les individus au sein d'une même expérience (Carpenter et al., 2013; Enk et al., 2013, 2014). Les facteurs responsables de cette variabilité peuvent être de différentes natures et dépendent de la divergence sonde/cible, de la température d'hybridation ou de lavage, de la longueur et de la concentration des molécules ciblées, de la complexité des séquences, du type de sonde ou encore de la stratégie de conception des sondes (Ávila-Arcos et al., 2011; Bodi et al., 2013; Ávila-Arcos et al., 2015; Paijmans et al., 2016; Cruz-Dávalos et al., 2017).

e) L'authentification de l'ADN ancien

L'une des préoccupations majeures des études en paléogénomique est l'authentification de l'ancienneté de l'ADN qui implique un argumentaire pour différencier les séquences anciennes des séquences modernes. L'ADN ancien présente de nombreuses modifications biochimiques qui compliquent sa manipulation en laboratoire. Cependant, ces modifications fournissent la base à des arguments puissants d'authentification. Un premier argument consiste à étudier la distribution en taille des fragments anciens qui sont très fragmentés (<100 pb) contrairement à leurs homologues modernes (Dabney et al., 2013b). Il est aussi possible d'étudier la présence de dommages spécifiques par l'analyse des modifications de séquences qu'ils entraînent, comme la fréquence des substitutions de C→T (et de G→A selon la méthode choisie pour construire les banques d'ADN), qui s'accumulent aux extrémités des fragments anciens. Ces altérations post-mortem peuvent être mesurées à l'aide d'outils bioinformatiques tels que mapDamage2 (Jónsson et al., 2013), d'autant plus facilement que les extraits ADN n'ont pas été traités au USER ou n'ont été que partiellement traités. Il est aussi possible d'utiliser PMDtools (Skoglund et al., 2014) pour évaluer la désamination des cytosines sur des extraits qui ont subi un traitement USER complet en se focalisant sur les dinucléotides CpG (Briggs et al., 2010; Orlando et al., 2021b). En effet, ces positions sont fortement méthylées le long du génome des vertébrés et sont protégées du traitement USER car l'enzyme UDG est spécifique des uraciles ; elle n'élimine donc pas les thymines résultant de la désamination des cytosines méthylées. Enfin, il est indispensable d'estimer directement le degré de contamination des échantillons anciens. Les outils disponibles se basent, par exemple, sur une estimation des niveaux d'hétérozygosité sur les chromosomes haploïdes, attendus donc comme monomorphe, comme le chromosome X chez les mâles (ANGSD, Korneliussen et al., 2013), ou le génome mitochondrial (schmutzi, Renaud et al., 2015). De nouvelles méthodes émergent, telles que AuthentiCT (Peyrégne and Peter, 2020), afin de permettre d'estimer la contamination sans restriction due à l'organisme, le sexe ou encore le chromosome étudié.

f) Les limites de l'ADN ancien

La limite majeure en ADN ancien est la quantité d'ADN endogène présente dans les échantillons. En effet, les ossements datant de siècles voir de millénaires sont rapidement

colonisés par les micro-organismes du sol jusqu'à constituer le plus souvent la majorité de l'ADN contenu dans les ossements (Fortes and Paijmans, 2015). Par exemple, le génome Néandertalien partiellement séquencé en 2010 présentait un taux d'ADN endogène allant de 1 à 5 % seulement (Green et al., 2010). D'autres sources de contamination existent et peuvent provenir de la manipulation des ossements par les expérimentateurs au cours des différentes étapes expérimentales. Les laboratoires qui se sont spécialisés dans l'étude de l'ADN ancien obéissent à des règles strictes afin de limiter ces risques au maximum. Cela inclut l'utilisation d'équipement de protection personnel, de salles blanches en surpression, d'instruments de décontamination, de javel et UV ainsi que des blancs de contrôle à chaque étape expérimentale (Handt et al., 1996; Cooper, 1997; Cooper and Poinar, 2000; Pääbo et al., 2004). Cependant, il est impossible de garantir l'absence de toute contamination, il est donc primordial de la quantifier et d'en identifier les potentielles sources à l'issue du séquençage.

4) Les apports de l'ADN ancien sur les épidémies passées

La compréhension de l'histoire des pathogènes est un enjeu majeur à l'heure où de nouvelles maladies émergent au sein de nos sociétés. Les données issues du séquençage de ces organismes anciens peuvent fournir de précieuses informations sur la dynamique de leur évolution et apporter des réponses aux débats sur l'origine de tels événements que l'archéologie et l'histoire ne peinent à résoudre.

a) Identifier les pathogènes du passé

Les archéologues découvrent depuis des décennies des charniers ou fosses multiples qui témoignent d'épidémies passées sans en connaître néanmoins l'origine exacte. Les sources historiques ne permettent pas non plus d'identifier avec certitude l'agent responsable du nombre important de morts au cours de ces crises. En effet, les symptômes décrits pour ces maladies restent souvent vagues et recoupent souvent différents pathogènes. De nombreux historiens ont par exemple essayé d'identifier le pathogène qui est responsable de la peste Antonine du IIème siècle de notre ère. En s'appuyant sur les récits de Gallien, un médecin de l'époque, les études historiques ont dénombré une multitude de pathogènes pouvant être responsables de cette épidémie. Les candidats vont de la peste bubonique, la variole, la fièvre jaune, le typhus, la fièvre typhoïde, la méningite ou encore la rougeole (Littman and Littman, 1973; Cunha and

Cunha, 2008). Mais, faute d'identification biologique, la confusion règne concernant cet épisode épidémique et l'organisme qui en est responsable est encore aujourd'hui inconnu.

Traditionnellement, l'étude anthropologique des maladies infectieuses anciennes est réalisée par le biais d'analyses paléopathologiques d'assemblages osseux (D. J. Ortner, 2003; Buikstra and Roberts, 2011). Cette approche reste cependant limitée car la plupart des maladies infectieuses aigües ne causent aucune lésion visible sur les os, ces dernières n'apparaissant le plus souvent qu'au cours d'infections chroniques. De plus, l'observation de tels stigmates ne révèle pas systématiquement la cause de la mort de l'individu. Effectivement, même si une personne était en bonne santé au moment de son décès, elle peut tout de même présenter des lésions résultant d'une infection antérieure. Cela est connu sous le nom de paradoxe ostéologique. (DeWitte and Stojanowski, 2015).

Depuis les années 1990 et plus particulièrement au cours des 10 dernières années, le développement de l'étude de l'ADN ancien a offert des perspectives intéressantes pour l'étude des maladies infectieuses. La détection de pathogènes était initialement basée sur l'amplification de l'ADN microbien cible par PCR (Spigelman and Lemma, 1993; Arriaza et al., 1995; Drancourt et al., 1998; Zink et al., 2001). Cette approche restait limitée car elle ne permettait pas toujours de distinguer les microbes anciens des modernes (Wilbur et al., 2009; Warinner et al., 2017). Le développement des NGS a permis d'accroître considérablement l'efficacité du séquençage et la masse de données provenant d'échantillons anciens. Le premier génome de pathogène ancien caractérisé à l'aide des NGS est celui de l'agent causatif de la peste, la bactérie *Yersinia pestis* (Bos et al., 2011). Les individus victimes de cette bactérie proviennent d'un site archéologique datant de 1348-1350 CE en Angleterre (East-Smithfield, Londres) associé à la peste noire, cette épidémie redoutable qui a décimé près de la moitié de la population européenne entre 1347 et 1351 CE (Benedictow, 2004; Green, 2020). La détection de ce même pathogène sur des individus datant du début du Moyen Âge a permis de confirmer que *Y. pestis* était responsable de la première pandémie de peste, à savoir la peste Justinienne (Wagner et al., 2014; Keller et al., 2019). Depuis ces premiers succès, la paléogénomique est devenue un outil puissant et efficace pour caractériser les pathogènes anciens, ou du moins en attester la présence. Un bon exemple du niveau de résolution que l'on peut atteindre à l'aide de tels outils est une étude de 2018 qui s'est intéressée à la tréponématose (Schuenemann et al., 2018b). Des individus retrouvés sur un site au Mexique datant du XVIIème et du XVIIIème siècle présentaient des lésions osseuses (périostite) pouvant être associées à la tréponématose. Le

séquençage de ces échantillons a démontré la présence de deux sous-espèces de *Treponema*, *T. pallidum* subsp. *pallidum* responsable de la syphilis et *T. pallidum* subsp *pertenue* responsable du pian. Les indices ostéologiques n'auraient jamais pu permettre aux anthropologues de différencier ces deux sous-espèces. Seule l'étude paléogénomique permet d'atteindre une résolution suffisante pour distinguer à cette échelle l'origine de maladies présentant des symptômes similaires.

b) L'origine des pathogènes

L'étude de l'ADN ancien a fourni des résultats pour le moins surprenants concernant l'origine des épidémies historiques en caractérisant les vecteurs et leurs origines dans le temps. Un des exemples de maladie infectieuse les plus connues est la peste, dont l'agent causatif a été caractérisé par Alexandre Yersin en 1894 (Treille and Yersin, 1894). Ce pathogène sévit encore aujourd'hui et reste très proche d'une autre bactérie, *Yersinia pseudotuberculosis*. Dans un premier temps, la date de divergence entre ces deux bactéries a été estimée sur la base des données génomiques entre 2 600 et 28 000 ans AP (Achtman et al., 1999, 2004; Cui et al., 2013; Wagner et al., 2014). Cependant, le manque de résolution temporelle, la variabilité du taux de mutation selon les souches de *Y. pestis* (Eaton et al., 2023) ainsi que l'utilisation de modèles inappropriés rendent la datation de leur ancêtre commun difficile. Des estimations plus récentes ont été réalisées et repoussent la date de divergence de 13 000 à 79 000 ans avant notre ère. De plus, des souches de *Y. pestis* aujourd'hui éteintes ont été identifiées chez des individus anciens datant du néolithique et de l'âge du bronze dans toute l'Eurasie, entre 2 800 et 5 000 ans AP (Rasmussen et al., 2015a; Andrades Valtueña et al., 2017; Spyrou et al., 2018; Rascovan et al., 2019; Andrades Valtueña et al., 2022). Contre toute attente pour les sources historiques qui faisaient remonter l'origine de la peste à la première pandémie associée à la peste Justinienne, la peste humaine remonte en réalité à au-moins 5 000 ans.

Des preuves génomiques attestent de la présence d'autres pathogènes à des époques très anciennes. C'est le cas du virus de l'hépatite B pour lequel la preuve d'infection la plus ancienne date du Néolithique, il y a 7 000 ans (Krause-Kyora et al., 2018b). De plus, on retrouve des traces du pathogène sur des individus datant de l'âge du bronze, de l'âge du fer jusqu'au XVIème siècle, soulignant une persistance lointaine de ce pathogène dans les populations humaines (Bar-Gal et al., 2012; Mühlemann et al., 2018; Ross et al., 2018). La paléogénomique permet

de placer un cadre temporel à l'apparition des pathogènes anciens et donne de précieuses informations sur leur émergence au sein des populations humaines. Ceci est d'autant plus précieux qu'en l'absence de fossiles facilement observables, il est difficile, voire impossible, de calibrer les taux de substitution sous-tendant l'évolution des microbes.

c) L'évolution des pathogènes

La compréhension des mécanismes qui régissent l'évolution des agents infectieux représente un des défis majeurs des études paléopathologiques. La bactérie *Y. pestis* est le pathogène le plus étudié en paléogénomique (Spyrou et al., 2019a) et elle est responsable des trois grandes épidémies de peste qu'ont connues les sociétés humaines. Cette bactérie trouve ses origines au Néolithique mais elle ne présentait pas les caractéristiques que les souches modernes arborent. À l'origine, *Y. pestis* n'est pas adaptée à l'homme et ses hôtes primaires sont les rongeurs sauvages, tels que les marmottes, où la bactérie peut se maintenir de manière continue ou intermittente (Anisimov et al., 2004; Gage and Kosoy, 2005; Riehm et al., 2011). L'identification de la peste à partir de restes humains datant du Néolithique et de l'âge du Bronze suggère que l'infection humaine était possible il y a plus de 5 000 ans. Cependant, la bactérie ne possédait dans son génome que certains facteurs de virulence contrairement aux souches actuelles et historiques. Le gène *ymt* (yersinia murine toxine) est un facteur de virulence déterminant pour la bactérie car il lui offre la capacité à coloniser le tube digestif des puces, dont le rôle dans la transmission de la peste bubonique est primordial (Sun et al., 2014). Ce gène est cependant absent dans le génome des souches du Néolithique (Figure 5) ce qui a pu suggérer qu'elles étaient incapables de produire la forme bubonique de la peste, qui dérive des morsures de puce. La souche la plus ancienne présentant ce gène date de 3 800 ans et cohabitait certainement avec d'autres souches qui en étaient dépourvues (Spyrou et al., 2018). Ces souches anciennes possédaient également les formes actives des gènes *rcsA*, *PDE3*, *PDE2* et *ureD* qui inhibent la production de biofilm (Figure 5; Chouikha and Hinnebusch, 2014; Sun et al., 2014). Les bactéries ne peuvent ainsi s'agréger dans le proventricule de la puce et bloquer le flux sanguin de l'hôte vers l'estomac de la puce. Au contraire, lorsque ces gènes sont inactifs, comme chez les souches modernes et historiques plus récentes, la formation d'un biofilm entraîne un reflux de sang contaminé par les bactéries vers l'hôte au niveau du site de piqûre, accroissant considérablement la transmission de la bactérie (Bacot, 1915; Hinnebusch and Erickson, 2008). Ce mécanisme est un tournant évolutif pour *Y. pestis* et a certainement joué

un rôle très important dans l'émergence d'épidémies massives de peste. Enfin, le gène *flhD* est présent sous la forme de pseudogène, c'est-à-dire sous forme non-fonctionnelle, chez les souches plus récentes. Les souches très anciennes possèdent une version active de ce gène qui est un puissant inducteur de réponse innée chez l'hôte (Minnich and Rohde, 2007). Ainsi, durant les premiers stades de son évolution, *Y. pestis* ne semblait pas être capable de se transmettre par le biais des puces. Ce type de transmission est pourtant indispensable au développement de la forme la plus fréquente de la peste, la peste bubonique. Il a donc été suggéré que cette forme de la maladie n'était pas présente au Néolithique (Rasmussen et al., 2015a; Hinnebusch et al., 2016). La première preuve connue à ce jour de transmission par les puces a été mise en évidence sur deux squelettes datant de 3 800 ans provenant de Russie (dans la région de Samara ; Spyrou et al., 2018).

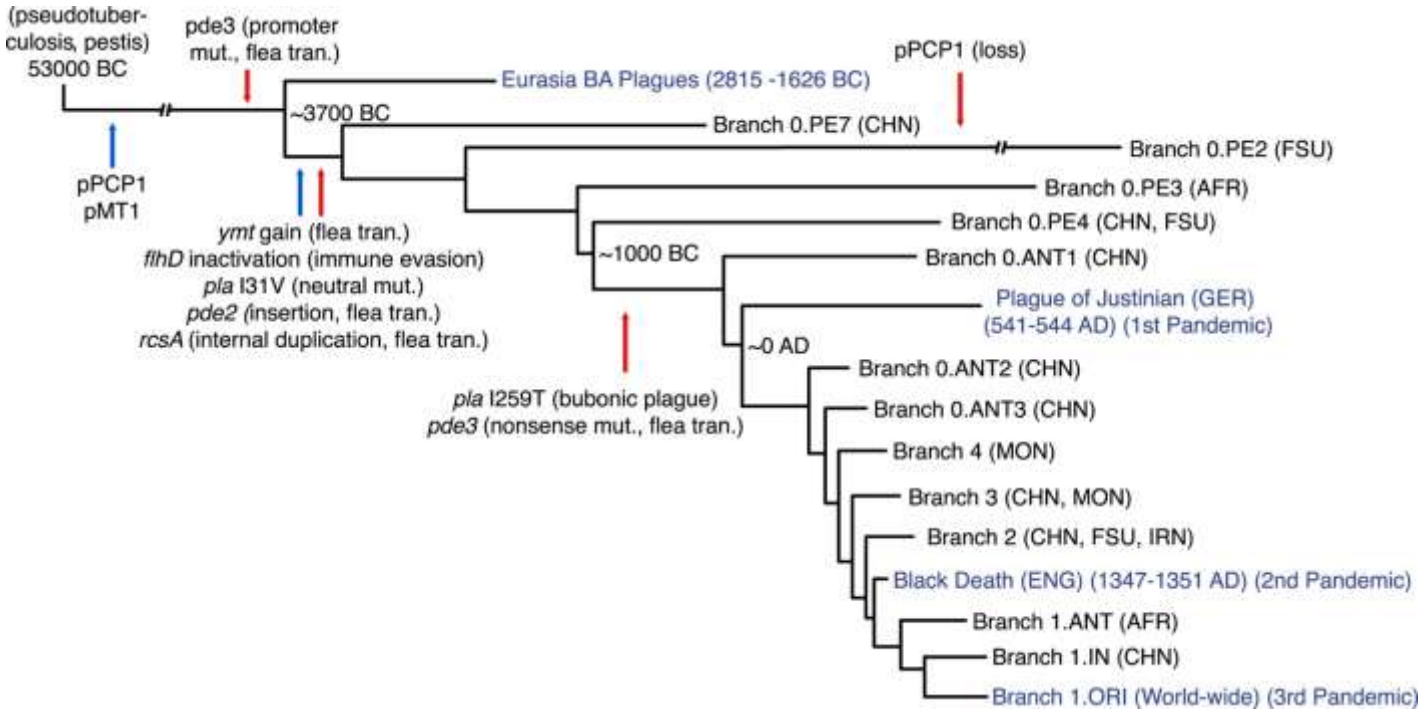


Figure 5 : Arbre phylogénétique des pestes anciennes et modernes représentant l'évolution du génome de la bactérie *Y. pestis* (figure issue de l'article de Rasmussen et al., 2015b). Cette phylogénie représente les événements évolutifs importants depuis la divergence entre *Y. pestis* et *Y. pseudotuberculosis.* Les gains (bleu) et les pertes ou mutations perte de fonction (rouge) sont indiqués par des flèches.

d) L'immunité face aux pathogènes

Les pathogènes actuels sont le fruit d'une longue évolution au cours de laquelle plusieurs formes ont pu exister, cohabiter et s'éteindre. L'étude paléogénomique des agents infectieux du passé nous permet de reconstruire leur histoire évolutive, mais il est important de s'intéresser également à leurs hôtes. La relation entretenue entre le pathogène et son hôte est le fruit d'une longue et étroite coévolution. L'étude des épidémies peut ainsi être abordée par le biais des individus qui en sont les victimes afin de comprendre les facteurs génétiques qui rendent les populations susceptibles ou résistantes à ces crises. Les maladies infectieuses représentent l'une des pressions de sélection les plus fortes exercées sur les populations humaines et les autres espèces animales (Fumagalli et al., 2011; Inhorn and Brown, 1990). Les symptômes et la susceptibilité des individus face à une maladie varient fortement (Casanova, 2015). La réponse immunitaire à une infection peut être non spécifique voire absente selon les individus, ce qui peut entretenir une certaine confusion dans l'analyse classique des pathologies anciennes (Steinbock, 1978; Wood et al., 1992; D. Ortner, 2003; Roberts and Manchester, 2005).

La paléogénomique est un outil puissant et complémentaire à l'épidémiologie et à la génétique clinique pour identifier des variants génétiques associés aux risques de développer des formes graves de maladies infectieuses (Casanova and Abel, 2022). À l'inverse, il est aussi possible d'identifier des variants qui confèrent une protection accrue face à certaines pathologies. L'exposition aux pathogènes varie en fonction des régions du monde et la sélection de mutations avantageuses pour l'immunité innée sera spécifique des populations (Quintana-Murci and Clark, 2013; Karlsson et al., 2014). L'étude de 1 000 génomes européens datant des 10 000 dernières années a révélé que le variant TYK2 P1104A (« tuberculosis risk variant ») est aujourd'hui porté par 3 % des personnes d'origine européenne. Cette étude démontre que ce variant a évolué sous une forte sélection négative, certainement induite par *Mycobacterium tuberculosis* durant les deux derniers millénaires, avec une fréquence en chute libre depuis l'âge du Bronze (Kerner et al., 2021). L'étude de l'ADN ancien peut aussi permettre d'identifier des mutations avantageuses qui ont été sélectionnées suite à une exposition de longue durée à un pathogène. En effet, au cours des 10 000 dernières années et plus particulièrement depuis l'âge du Bronze, une sélection positive et récurrente de variants a été exercée sur des gènes de l'immunité tels que les gènes *TLR10, TRL1, TLR6* ainsi que certains gènes codants pour des antigènes associés aux leucocytes humains (Mathieson et al., 2015; Kerner et al., 2021, 2023). Des marques de sélection ont aussi été détectées pour le gène *CCR5* (Bamshad et al., 2002) qui est impliqué, entre autres, dans l'infection du VIH-1. Néanmoins, cette sélection ne résulte pas

de la pression de sélection exercée par le VIH car il n'est apparu que récemment dans les populations humaines (depuis les années 1950-60 ; Worobey et al., 2016), ce serait donc le fruit d'une pression exercée par d'autres pathogènes. L'identification d'événements passés de sélection pour des variants a été facilitée grâce au développement des NGS et ces avancées ont permis de retracer l'évolution temporelle de l'immunité des populations humaines. La paléogénomique fournit donc un cadre précieux pour identifier les formes de sélection qui ont pu contribuer à l'adaptation des populations humaines face aux épidémies et à l'évolution de leur système immunitaire.

La peste Noire a décimé près de la moitié de la population européenne au XIVème siècle (Benedictow, 2004; Green, 2020) et s'est installée durablement en Europe jusqu'au XVIIIème siècle (Spyrou et al., 2019a). N'ayant pas été exposées à cette peste auparavant, les populations européennes étaient vraisemblablement immunologiquement naïves, peu ou pas adaptées à *Y. pestis*. Le taux de mortalité élevé suggère que des variants génétiques conférant une protection contre l'infection par *Y. pestis* auraient pu faire l'objet d'une très forte sélection au cours de cette période. De plus, les épidémies récurrentes au cours des quatre siècles qui ont suivi la peste noire en Europe n'ont pas toujours été associées à des taux de mortalité importants (Benedictow, 1987). Les causes potentielles expliquant ces variations dans les taux de mortalité peuvent être de différentes natures, allant de l'émergence de pratiques culturelles plus hygiéniques à la mise en place de protocoles pour limiter la propagation de la maladie. L'évolution du pathogène pourrait aussi en être à l'origine. Il semble cependant que cette bactérie n'ait évolué que très lentement au cours de la seconde pandémie, à en croire la variabilité génomique très faible observée chez les souches anciennes séquencées à ce jour (Duchêne et al., 2016; Eaton et al., 2023). La diminution de la mortalité observée au cours de certaines épidémies pourrait donc être due au moins en partie à l'adaptation de la population européenne à ce pathogène.

Au cours de ces deux dernières années, plusieurs travaux ont étudié si des variants dans les gènes de l'immunité chez l'homme avaient été sélectionnés au cours de la seconde pandémie de peste (Immel et al., 2021; Di et al., 2022; Klunk et al., 2022). L'une de ces études a présenté l'analyse de données génomiques d'individus contemporains à la peste noire provenant de Londres et du Danemark (Klunk et al., 2022). Ce corpus est particulièrement intéressant car les individus qui le constituent sont morts avant, pendant ou après cet épisode de peste. Ainsi, un variant spécifique (rs2549794) a été détecté à proximité du gène *ERAP2*, un gène impliqué dans

la présentation d'antigènes, et pourrait induire une meilleure protection vis-à-vis de la peste. D'après ces analyses, ce variant présente une forte signature de sélection positive, avec un coefficient de sélection de 0.4, un coefficient 5 fois plus important que celui observé à ce jour pour l'allèle de persistance de la lactase, le plus important coefficient de sélection connu à ce jour (Liebert et al., 2017). Cela signifierait que les personnes étant homozygotes pour ce variant étaient 40 % plus aptes à survivre et à se reproduire durant l'épidémie de peste noire que les individus homozygotes pour les allèles non protecteurs. Cependant, ces résultats font aujourd'hui débat après qu'une équipe ait ré-analysé les données étudiées par Klunk et ses collaborateurs (Barton et al., 2023). Dans un premier temps, les auteurs ont souligné que les changements de fréquence observés au niveau des gènes de l'immunité pourraient être dus à des différences de couvertures entre les loci immuns et neutres investigués. Cela a pour conséquence de gonfler artificiellement les valeurs de Fst (indice permettant de mesurer la différenciation des populations à partir du polymorphisme génétique) observées. Le signal indicatif de la sélection ne signerait donc pas à coup sûr la sélection. Ensuite, les fréquences alléliques auraient été estimées de manière incorrecte. Après correction, plus aucun des loci étudiés ne semblait plus passer les seuils indicatifs de sélection. De plus, comme des stratégies pour éviter les faux positifs dans la détection de variants (limitation à des sites polymorphes connus, à des transversions ou utilisation d'un traitement à l'Uracil-DNA glycosilase) n'ont pas été utilisées, la majorité des variants identifiés pourraient donc être des faux positifs. Les seuils employés pour filtrer les données n'ont pas non plus été corrigés correctement pour les tests multiples. Barton et al. n'ont mis en évidence aucun changement de fréquence significatif pour le variant rs2549794 du gène *ERAP2*, que ce soit dans les données rapportées par Klunk et al. ou dans les données publiées sur une période de 2 000 ans. Enfin, la taille de l'échantillon semble également trop faible pour évaluer la sélection car l'incertitude est trop grande. Pour conclure, la question de la sélection de variants conférant une immunité accrue vis-à-vis de la peste fait débat. Vilgalys et ses collaborateurs ont par la suite réalisé à nouveau les analyses en tenant compte des critiques énoncées précédemment (Vilgalys et al., 2023). D'après leurs derniers résultats, il semblerait que les préoccupations énoncées par Barton et al., 2023 n'ont qu'un impact limité sur les conclusions initiales et en utilisant la méthode préconisée, il semblerait que le variant rs2549794 près d'*ERAP2* continue d'émerger comme le candidat le plus fort pour la sélection.

Il est difficile d'affirmer avec certitude qu'une sélection ait eu lieu suite à ces épisodes épidémiques. Des jeux de données de meilleure qualité sont nécessaires pour mener à bien ces

investigations tout en tenant compte des biais statistiques introduits dans les données anciennes. Malgré les résultats contradictoires dans l'investigation de preuve de sélection de variants protecteurs face à la peste, il est important de poursuivre les investigations pour déterminer de potentielles ayant eu un impact sur la survie des individus au cours de ces épidémies.

5) Objectifs

a) Le contexte de la thèse

La brutalité et la persistance des épidémies de peste ont fortement marqué les populations humaines, en particulier les populations européennes qui ont massivement été touchées entre le XIVème et le XVIIIème siècle. Aujourd'hui disparue d'Europe, la bactérie *Y. pestis* continue de sévir dans 21 pays d'Afrique, d'Asie et d'Amérique. Ce ne sont pas moins de 26 237 cas qui ont été répertoriés par l'OMS (Organisation Mondiale de la Santé) dans ces régions entre 2000 et 2018 (Bertherat, 2016, 2019), l'écrasante majorité (97 %) en Afrique (Barbieri et al., 2020) et plus particulièrement en République démocratique du Congo et à Madagascar. Les ressources économiques et sanitaires dont disposent ces pays ne permettent pas de lutter efficacement contre ces épidémies récurrentes de peste. Il est donc crucial d'étudier cette bactérie sous tous ses aspects afin de comprendre son évolution, son épidémiologie et éventuellement trouver des parades à son éventuelle réémergence et propagation. Mon projet de thèse se focalise sur les épidémies historiques de peste et plus particulièrement les vagues tardives ayant frappé la France aux XVIIème et au XVIIIème siècles. Les sites étudiés sont associés à la temporalité de la guerre de Trente Ans (1618-1648) et à la grande épidémie de peste de Marseille de 1720-1722.

Les épisodes de guerre s'accompagnent souvent d'épidémies qui se propagent par le biais des mouvements de troupes. De nombreux facteurs peuvent influencer la circulation des pathogènes en contexte militaire, cela peut passer évidemment par le contexte sanitaire qui se retrouve bouleversé (famine, blessures, etc.), la promiscuité d'un grand nombre d'individus ou encore les mouvements de population provoqués par ces crises. Ces épidémies aggravent considérablement le bilan déjà lourd de certaines guerres et peuvent avoir un rôle important dans l'issue d'un conflit. Les sites étudiés ici datant du XVIIème siècle proviennent des Alpes, d'Amiens et de Beauvais. Le cimetière de Lariey Puy-Saint-Pierre (Alpes) est un petit village qui a été touché par une épidémie de peste en 1629-1630 (Tzortzis and Signoli, 2016). Cette

épidémie s'est propagée lors de mouvements de troupes en lien avec la guerre de Trente Ans. La bactérie ayant provoqué la mort des 34 individus présents dans le cimetière est génétiquement très étroitement apparentée aux souches qui ont été caractérisées à San Procolo a Naturno en Italie et a Brandenburg en Allemagne (Spyrou et al., 2019b; Guellil et al., 2020; Seguin-Orlando et al., 2021a). Les sites d'Amiens et de Beauvais n'ont pas de datation précise mais selon les données céramologiques, les échantillons étudiés proviendraient du milieu du XVIIème siècle, entre 1630-1670. La fouille de l'espace funéraire de l'Hôtel-Dieu Saint-Jean-Baptiste d'Amiens a permis de mettre en évidence les squelettes de pas moins de 259 individus ayant succombé au cours d'un ou plusieurs épisodes épidémiques. Le site de la maladrerie Saint-Lazare de Beauvais a été fondé entre la fin du XIème siècle et le début du XIIème siècle. Probablement en raison du recul important de la lèpre, une partie de cet espace a, par la suite, été utilisée pour l'hospitalisation des malades et l'enterrement des victimes de la peste. Les archives historiques indiquent qu'à partir de 1623, la partie sud-est de la maladrerie a été dédiée aux pestiférés. Les documents historiques ne fournissent pas d'information concernant la cause de la mort des individus provenant des sites d'Amiens et Beauvais. Cependant, à cette période, la présence militaire était importante dans le nord de la France notamment à cause de la guerre de Trente Ans, les mouvements de troupes auraient ainsi pu accroître la propagation des maladies infectieuses hautement transmissibles. Le bilan de la guerre de Trente Ans est lourd pour de nombreux pays européens, tant sur le plan démographique que sur le plan économique, politique et culturel et les épidémies, notamment celles de peste, ont certainement joué un rôle prépondérant dans ce bilan désastreux (Wilson, 2009).

Les différents sites archéologiques datant du XVIIIème siècle correspondent à des fosses multiples de la grande peste de Marseille (1720-1722). Les sites de la Major (1720) et de l'Observance (1722) sont localisés à Marseille tandis que les sites de Delos (1720-1721) et les Rayettes (1720-1721) sont situés à Martigues. La grande peste de Marseille est le dernier épisode de peste en France et est connue pour avoir débuté avec l'arrivée du grand Saint Antoine dans la ville (Signoli, 2022). Ce navire marchand en provenance du Levant transportait le bacille de la peste en plus de sa précieuse cargaison d'étoffes. Des cas de peste se sont rapidement déclarés au sein de l'équipage et de ses passagers, et le bilan à l'issue du voyage s'élèvera à 7 morts (Beauvieux et al., 2022). À l'arrivée du navire à Marseille, l'équipage est placé en quarantaine dans le Lazaret d'Arenc et, en dépit des cas de peste, la cargaison du navire sera tout de même délivrée sous la pression exercée par les propriétaires marseillais. Malgré certaines précautions sanitaires imposées par les autorités, des défaillances imputables à la

mauvaise application du règlement en vigueur ont permis la diffusion de la peste dans toute la ville au cours des jours et semaines qui suivirent. L'épidémie se répandit ensuite dans toute la Provence et dans les régions adjacentes, provoquant la mort de dizaines de milliers d'individus. Le bilan de cette épidémie fut très lourd, avec près de la moitié de la population marseillaise et un quart de la population de la région de Provence qui succomba à la maladie en l'espace de quelques semaines. L'année 1720 fut la plus meurtrière avec des taux de mortalité pouvant atteindre 1 000 morts par jour entre août et septembre (Carrière et al., 2008). L'épidémie recula à partir du mois d'octobre 1720 et la ville reprit la plupart de ses activités au début de l'année 1721. Une rechute épidémique eut lieu au printemps 1722, cependant le nombre de morts resta limité, avec seulement 213 victimes dénombrées au sein des archives historiques (Signoli et al., 1998). La peste disparaît de Marseille et de la Provence à l'été 1722, ce qui marqua la fin d'un épisode traumatique pour les populations locales.

De nombreuses questions concernant ces événements font toujours débat et l'archéologie et l'anthropologie classique peinent à les résoudre. L'approche paléogénomique peut permettre de préciser certains aspects de ces problématiques pour mieux comprendre l'histoire de ces épidémies passées. Ainsi, à l'aide des outils de biologie moléculaire, il est devenu possible de caractériser les souches de peste qui sévissaient dans le passé et d'en préciser l'origine phylogéographique. Cette approche peut aussi permettre de mesurer la virulence de certaines souches ou alors d'étudier le génome des individus morts de cette maladie pour expliquer la part des susceptibilités génétiques dans les chances de survie individuelles. La caractérisation de variants génétiques, que ce soit au niveau du génome de la bactérie ou de celui de l'hôte, peut permettre de définir des facteurs génétiques qui vont influencer la virulence de la bactérie ou la résistance immunitaire de son hôte humain. Notre travail de recherche s'est ainsi structuré autour de deux axes principaux : la caractérisation du génome de *Y. pestis*, agent infectieux ayant sévi au cours des XVIIème et XVIIIème siècles ; et l'étude du génome des individus ayant succombé à la maladie, en particulier leur immunogénome inné.

b) Optimisation de l'extraction de l'ADN de pathogènes anciens

Les méthodes d'extraction de l'ADN provenant d'échantillons anciens ont été étudiées et comparées à de nombreuses reprises afin d'en améliorer les performances (Prado et al., 2002; Hummel, 2003; Rohland and Hofreiter, 2007b; Rohland et al., 2010; Gamba et al., 2016;

Nieves-Colón et al., 2018). Aussi, il existe aujourd'hui de multiples approches permettant d'optimiser la récupération de l'ADN de l'hôte selon le matériel étudié et le génome à extraire (Orlando et al., 2021a). Cependant, il n'existait pas, au démarrage de ce travail, d'étude visant à optimiser l'extraction de l'ADN des pathogènes pour ensuite capturer et séquencer leurs génomes. La plupart des recherches sur les pathogènes anciens emploient des protocoles d'extraction standards, et la peste n'échappe pas à cette règle. La méthode la plus utilisée se base sur une prédigestion puis une digestion complète la poudre provenant des os ou des dents (Gamba et al., 2016).

Il faut savoir que le bacille de la peste se développe principalement dans le système lymphatique et, plus particulièrement, les ganglions lymphatiques. C'est la forme bubonique de la peste qui est majoritaire et qui résulte d'une transmission par le biais de piqûres de puces infectées. Cette forme de la maladie provoque la mort dans environ 70 % des cas sans traitement. La peste pulmonaire est une autre forme de la maladie qui découle d'une transmission interhumaine par les aérosols. Elle est moins fréquente que la peste bubonique mais elle demeure la forme la plus dangereuse de la peste avec des taux de mortalité avoisinant les 100 % en absence de traitement. Ces deux formes représentent la majorité des cas rencontrés au cours des épidémies de peste, cependant, il est difficile de les étudier dans le cadre de la recherche paléogénomique. En effet, l'absence de tissus mous (poumons, ganglions lymphatiques, lymphe, etc.) sur les sites archéologiques restreint les études aux tissus calcifiés. Il a été démontré dans les années 2000 qu'il était possible de mettre en évidence la présence de ce pathogène à partir de la pulpe dentaire (Drancourt et al., 1998). Ce matériel est très intéressant pour étudier les pathogènes circulant dans le sang et notamment la dernière forme de peste, à savoir la peste septicémique qui découle souvent de la forme bubonique, et qui correspond au stade final de la maladie où le pathogène rejoint la circulation sanguine générale. Elle représente environ 10 à 20 % des cas et est mortelle sans traitement mais non contagieuse (Barbieri et al., 2020). Or, la pulpe dentaire est richement vascularisée, à l'inverse des os à l'âge adulte (Drancourt et al., 1998; Drancourt and Raoult, 2004, 2005) : elle maximise donc les chances de retrouver le matériel génétique du pathogène. Cependant, alors que la plupart des études sur des pathogènes anciens, et plus particulièrement la peste, emploient des protocoles standards d'extraction de l'ADN, la dynamique de libération de l'ADN des pathogènes anciens au cours du temps de digestion de la poudre issue de la pulpe dentaire est, à ce jour, peu connue. Au cours de ce projet de recherche, nous avons réalisé des tests d'extraction en vue de modéliser la libération de l'ADN du pathogène séquencé suite à une capture de son génome. L'objectif de

cette expérience a été de développer une méthode fiable et robuste, permettant de réduire les coûts expérimentaux liés à l'extraction ainsi que le temps des manipulations et donc le risque de contamination des extraits anciens.

c) Étude de la diversité des souches de peste

Nous avons ensuite cherché à caractériser d'un point de vue génétique les souches de peste impliquées dans les épidémies qui ont sévi en France aux XVIIème et XVIIIème siècles. Dans un premier temps, l'analyse phylogénétique des bactéries identifiées a eu pour objectif de déterminer les relations entretenues entre les souches étudiées et les souches précédemment publiées, notamment avec les souches anciennes de la seconde pandémie ainsi que les souches modernes de la branche 1, qui est actuellement la lignée la plus répandue dans le monde et est associée à la seconde et la troisième pandémie de peste (Yang et al., 2023). Nous avons également examiné si les souches de *Y. pestis* identifiées dans cette étude possédaient des caractéristiques génétiques distinctes des autres souches connues, notamment les plus proches. Pour évaluer ces spécificités, nous nous sommes intéressés à la charge plasmidique en étudiant les proportions moyennes des séquences de bonne qualité alignées contre le chromosome circulaire (CO92) et les 3 plasmides (pCD1, pMT1 et pPCP1) en fonction des proportions de sondes utilisées pour cibler ces différents éléments du génome de la bactérie. La charge plasmidique peut effectivement jouer un rôle dans la virulence (Wang et al., 2016), notamment le plasmide pCD1 qui porte les gènes codant pour le système de sécrétion de type 3 (T3SS) permettant à la bactérie d'injecter des toxines dans la cellule hôte (Demeure et al., 2019). Enfin, nous avons cherché à caractériser la présence de variants génétiques (substitutions et insertions/délétions) dans les génomes étudiés. Ces derniers peuvent en effet avoir des conséquences potentiellement importantes pour la bactérie s'ils concernent des gènes impliqués dans la virulence (Yu et al., 2019), en particulier s'il s'agit de variants non-synonymes (c'est-à-dire qui changent la chaîne d'acides aminés finalement traduite).

d) L'Immunité innée face à la peste

L'étude de l'immunité innée humaine face aux souches historiques de peste, et plus particulièrement lors de la peste noire, fait débat dans la communauté scientifique. Des équipes ont essayé de répondre à cette problématique mais la question de la sélection de variants

conférant une immunité accrue vis-à-vis de la peste reste à approfondir. Le manque de données de bonne qualité ainsi que la faible taille des échantillons étudiés rendent la tâche particulièrement difficile. Il faut cependant persister dans l'exploration de ces hypothèses en continuant de séquencer des génomes humains dans ces contextes d'épidémies de peste pour accroître la masse de données et la résolution temporelle des échantillons étudiés. Dans le but d'approfondir cette question, le séquençage des génomes des individus provenant des différents sites étudiés a été réalisé à partir de l'ADN extrait de l'os pétreux. Dans un premier temps, l'étude de ces génomes humains a eu pour objectif de caractériser l'état des populations touchées lors de ces épidémies de peste des XVIIème et XVIIIème siècles. Dans un second temps, cette approche a visé à chercher de nouveaux variants candidats qui pourraient être impliqués dans l'immunité face à cette maladie.

II) <u>Article 1</u>

1) Contexte

La détection et la caractérisation de pathogènes anciens sur des restes archéologiques sont des défis que la recherche en ADN ancien a su relever de nombreuses fois au cours de la dernière décennie (Spyrou et al., 2019a; Malyarchuk et al., 2022). La paléogénomique a ainsi pu révéler la présence de pathogènes figurant parmi les plus meurtriers de l'histoire humaine, tels que la variole, la tuberculose ou encore la peste. Malgré le développement et la multiplication des recherches sur les épidémies du passé, retrouver le génome entier d'un pathogène datant de plusieurs siècles, voire plusieurs millénaires, peut s'avérer difficile notamment à cause de la faible quantité de matériel génétique présent dans les ossements archéologiques et de sa qualité médiocre. Comme souvent en ADN ancien, le génome d'intérêt n'est présent qu'en faible quantité en comparaison aux multiples sources de contaminations dans les tissus étudiés. Cela est d'autant plus vrai pour les pathogènes, dont la fraction est encore inférieure à celle du génome de son hôte. Le séquençage de l'entièreté de l'ADN présent dans des extraits (« *shotgun sequencing* ») est la meilleure manière de garantir une quantité de données suffisante pour étudier les génomes issus des restes archéologiques en général. Cependant, lorsqu'il s'agit d'étudier des pathogènes, ce sont les techniques de « capture génique » ou « enrichissement » du génome d'intérêt qui sont privilégiées. Cette approche se base sur l'utilisation de sondes ADN ou ARN complémentaires au génome d'intérêt. Elles ont

gagné en popularité au cours de ces dernières années car elles garantissent une efficacité plus importante pour le séquençage du génome d'intérêt tout en diminuant les coûts expérimentaux, et plus particulièrement pour les échantillons de mauvaise qualité.

Le développement de techniques permettant de se focaliser sur une partie du matériel présent dans les extraits d'ADN ancien a permis de réduire drastiquement les coûts d'analyse, tout en augmentant l'efficacité du séquençage de l'ADN d'intérêt et notamment celui des pathogènes. De nombreuses études ont été réalisées au cours des dernières décennies afin de caractériser les paramètres maximisant l'efficacité des techniques de capture du génome de certains organismes mais de telles études faisaient cruellement défaut pour l'analyse de pathogènes. C'est dans ce contexte que nous avons réalisé un protocole expérimental afin de tester la dynamique de libération de l'ADN pour la capture du génome de *Y. pestis*. Le but de ces travaux était de développer une méthode d'extraction favorisant la libération des molécules d'ADN du pathogène pour les capturer par affinité de séquences avec des sondes synthétiques, et ainsi focaliser l'effort de séquençage sur le génome *Y. pestis*. L'article publié par iScience rapporte ce travail, en particulier le temps de digestion optimal pour l'extraction de l'ADN des pathogènes. Dans un premier temps, nous avons appliqué un protocole classique d'extraction de la pulpe dentaire qui consiste en une prédigestion de 60 minutes, suivie d'une digestion complète de la poudre restante pendant une nuit. L'analyse de la fraction correspondant à la prédigestion a révélé que le nombre de séquences uniques correspondant au génome de la bactérie après capture était plus important que dans celle issue de la digestion complète. Afin d'étudier la dynamique de libération de l'ADN du pathogène au cours de la prédigestion, nous avons effectué des prélèvements du surnageant toutes les 10 minutes pendant une heure, puis à intervalles réguliers au cours des 10 premières minutes de digestion. Les résultats issus de la capture de l'ADN de *Y. pestis* nous ont permis de modéliser la dynamique de libération de l'ADN au cours du temps et d'estimer un temps optimal de digestion de la pulpe dentaire. Ainsi après une prédigestion d'une heure, près de 95 % de l'ADN du pathogène contenu dans la pulpe dentaire a été extrait. Cette approche nous a permis de caractériser 12 génomes de peste provenant des XVIIème et XVIIIème siècles. Ces données génomiques nous ont ensuite permis de reconstruire les relations phylogénétiques que ces souches entretiennent avec les souches précédemment publiées. Enfin, nous avons pu caractériser les particularités génétiques de ces bactéries en explorant certains paramètres tels que la présence de variants au niveau de gènes de virulence ou de pathogénicité, ou encore mesurer la charge plasmidique relative à la proportion de sondes utilisées lors de la capture. Au final, cette étude a apporté des informations

nouvelles au sujet des épidémies de peste en France et a considérablement amélioré les outils expérimentaux permettant d'extraire de l'ADN de *Y. pestis* à partir de la pulpe dentaire, et potentiellement d'autres pathogènes circulant dans le sang. Il est crucial de poursuivre ces recherches afin de perfectionner les approches adaptées à ces problèmes spécifiques pour permettre de gagner en efficacité lors des études d'épidémies anciennes.

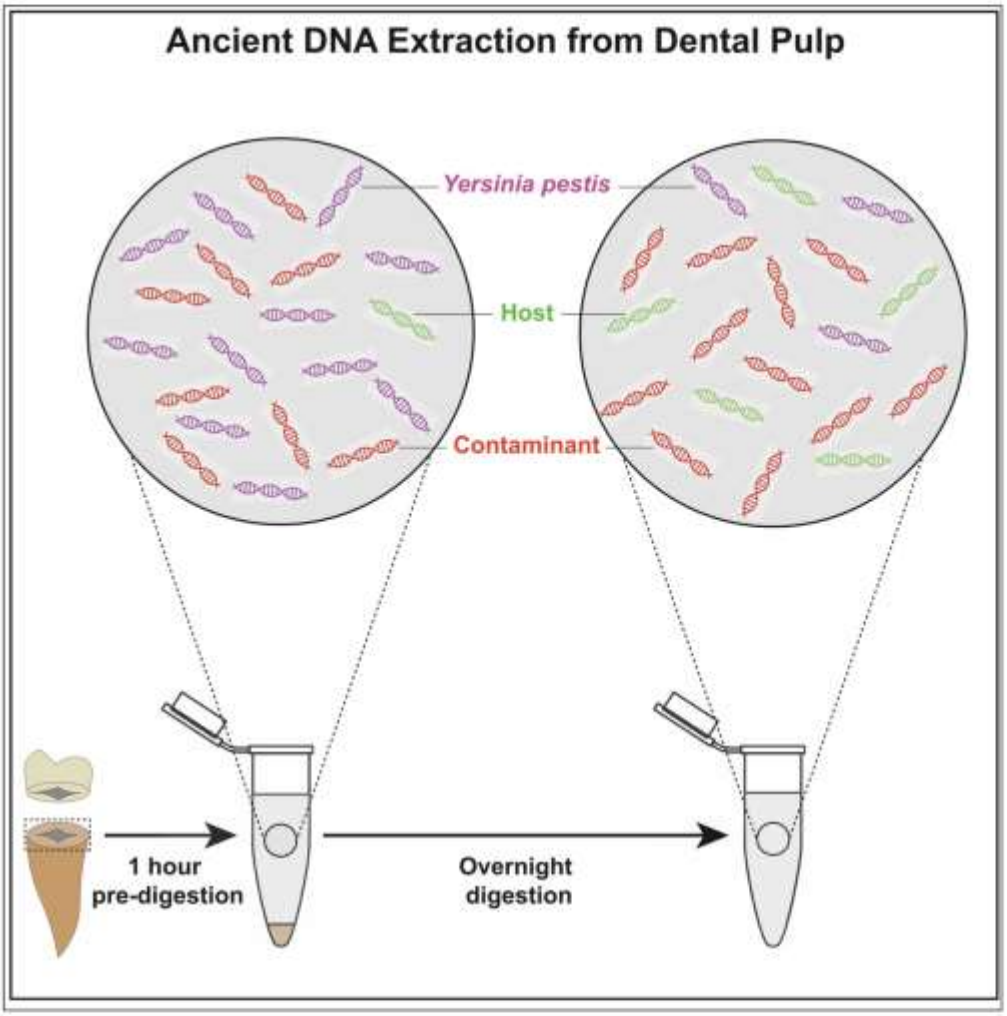
2) Improving the extraction of ancient Yersinia pestis genomes from the dental pulp

iScience

CellPress
OPEN ACCESS

Article

Improving the extraction of ancient *Yersinia pestis* genomes from the dental pulp



Pierre Clavel, Lexane Louis, Clio Der Sarkissian, ..., Maria A. Spyrou, Andaine Seguin-Orlando, Ludovic Orlando

ludovic.orlando@univ-tlse3.fr

Highlights

Short tooth powder digestion times improve recovery of *Yersinia pestis* ancient DNA

Variable *Y. pestis* detection success across teeth from the same individual

Sequencing of 12 second plague pandemic genomes from France

Different strains during the Thirty Years' War and the Great Plague of Marseille

Clavel et al., iScience 26, 106787
May 19, 2023 © 2023 The Author(s).
https://doi.org/10.1016/j.isci.2023.106787

iScience

CellPress
OPEN ACCESS

Article

Improving the extraction of ancient *Yersinia pestis* genomes from the dental pulp

Pierre Clavel,[1] Lexane Louis,[1] Clio Der Sarkissian,[1] Catherine Thèves,[1] Claudia Gillet,[1] Lorelei Chauvey,[1] Gaétan Tressières,[1] Stéphanie Schiavinato,[1] Laure Calvière-Tonasso,[1] Norbert Telmon,[1] Benoît Clavel,[2] Richard Jonvel,[3] Stéfan Tzortzis,[4] Laetitia Bouniol,[5] Jean-Marc Fémolant,[5] Jennifer Klunk,[6] Hendrik Poinar,[7,8,9] Michel Signoli,[10] Caroline Costedoat,[10] Maria A. Spyrou,[11] Andaine Seguin-Orlando,[1] and Ludovic Orlando[1,12,*]

SUMMARY

Ancient DNA preserved in the dental pulp offers the opportunity to characterize the genome of some of the deadliest pathogens in human history. However, while DNA capture technologies help, focus sequencing efforts, and therefore, reduce experimental costs, the recovery of ancient pathogen DNA remains challenging. Here, we tracked the kinetics of ancient *Yersinia pestis* DNA release in solution during a pre-digestion of the dental pulp. We found that most of the ancient *Y. pestis* DNA is released within 60 min at 37°C in our experimental conditions. We recommend a simple pre-digestion as an economical procedure to obtain extracts enriched in ancient pathogen DNA, as longer digestion times release other types of templates, including host DNA. Combining this procedure with DNA capture, we characterized the genome sequences of 12 ancient *Y. pestis* bacteria from France dating to the second pandemic outbreaks of the 17th and 18th centuries Common Era.

INTRODUCTION

Ancient DNA preserved within archaeological and paleontological remains provides genetic information about our evolutionary past.[1] The increasing throughput of next-generation sequencing instruments has considerably improved our capacity to characterize ancient genomes, despite environmental microbes often producing the dominant fraction of the DNA extracts.[2] Specific types of remains, including petrosal bones,[3] ossicles,[4] and the tooth cementum,[5] are known to be generally less prone to environmental DNA contamination. Together with improved decontamination and DNA extraction procedures, involving various pre-digestion treatments of sample powder through bleaching[6,7] and/or washing buffers,[8–10] these biological remains have facilitated the characterization of ancient genomes. The development of DNA capture methodologies targeting million-scale SNP (Single Nucleotide Polymorphism) panels,[11,12] whole chromosomes[13] or even the entire genome[14,15] have provided a further opportunity to focus sequencing efforts on the DNA fraction of interest, therefore, increasing sensitivity and reducing sequencing costs. As a result, there has been an acceleration since the first ancient human genomes have been sequenced in 2010[16–18] to the thousands of ancient human genomes characterized, with recent human paleogenomic studies typically including hundreds of individuals.[19,20]

Besides host and environmental microbial material, ancient DNA extracts can also preserve molecules originating from pathogens that circulated in the blood system when the host died.[21] In cases where the pathogenic load is high and environmental contamination is low, shotgun sequencing has allowed for the successful characterization of the entire genome sequence from ancient pathogens, including viruses (e.g., smallpox,[22,23] hepatitis B[24,25]), bacteria (e.g., *Mycobacterium leprae*[26]), and protozoan parasites (e.g., *Plasmodium vivax*[27]). Reconstructing ancient pathogen genomes, however, most commonly requires DNA capture technologies, both due to the molecular complexity of ancient DNA extracts and the generally limited genome size of pathogens, relative to their hosts.[21,28] Additionally, the type of material considered for ancient pathogen genome characterization is critical, as tissues favoring host DNA preservation, such as petrosal bones, show virtually no bacterial pathogen DNA.[29,30] Despite these limitations, more than five hundred ancient pathogen genomes have been sequenced,[31] revealing important insights into the epidemiological outbreaks with substantive morbidity and mortality.

[1]Centre d'Anthropobiologie et de Génomique de Toulouse (CAGT), CNRS UMR5288, Université Paul Sabatier, 37 allées Jules Guesde, 31000 Toulouse, France

[2]Archéozoologie, Archéobotanique: Sociétés, Pratiques et Environnements (AASPE), CNRS-UMR7209, Muséum national d'histoire naturelle, 55 Rue Buffon, 75005 Paris, France

[3]Amiens Métropole Service Archéologie Préventive, 2 rue Colbert, 80000 Amiens, France

[4]Service Régional de l'Archéologie, 21 allée Claude Forbin, 13100 Aix-en-Provence, France

[5]Service archéologique de la ville de Beauvais, 1 rue Desgroux, 60021 Beauvais, France

[6]Daicel Arbor Biosciences, Ann Arbor, MI 48103 USA

[7]McMaster Ancient DNA Centre, Departments of Anthropology, Biology and Biochemistry, McMaster University, Hamilton, ON L8S 4L9, Canada

[8]Michael G. DeGroote Institute of Infectious Disease Research, McMaster University, Hamilton, ON L8S, 4L9, Canada

[9]Humans and the Microbiome Program, Canadian Institute for Advanced Research, Toronto, ON, Canada

[10]Aix-Marseille Université, CNRS, EFS, ADES, 13005 Marseille, France

[11]Institute for Archaeological Sciences, Eberhard Karls University of Tübingen, Tübingen, Germany

[12]Lead contact

Continued

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Plague represents the best-studied infectious disease from an ancient DNA perspective.[37] DNA analyses of human individuals buried during the Justinian plague and the Black Death have confirmed *Y. pestis* as the etiological agent responsible for massive human loss in the sixth-eighth century Common Era (c. CE)[33,34] and the 14th c. CE,[35] respectively. Such analyses have, however, revealed that plague already circulated during the late Neolithic and early Bronze Age,[36–41] and have started reconstructing the evolutionary trajectory of key genetic changes underlying virulence and transmission.[38,42] Genetic evidence has also tracked the expansion of the second plague pandemic from a Central Asian homeland[43] and the following diversification of several lineages through space and time.[44]

Despite such advances, the characterization of ancient pathogen genomes remains difficult. Relatively low detection rates and minimal target ancient DNA contents are commonly observed in mass graves, even when the association with known pandemic, such as the second plague pandemic, is documented historically.[45] While the performance of various DNA capture systems is well established,[46–48] no single study has attempted to optimize the molecular toolkit for pathogen DNA extraction and capture. In this study, we investigated the temporal dynamics of plague and host DNA release from dental pulp, using a total of 120 human teeth dated to the 17th and 18th c. CE. These remains originated from five French archaeological sites associated with the second plague pandemic. We found that the vast majority of the pathogen DNA can be recovered following a short pre-digestion wash of the dental pulp, while the recovery of the host DNA benefits from further digestion. Therefore, focusing on the pre-digestion wash improves the ratio of *Y. pestis*-to-host DNA. Whether such an approach could also enhance the recovery of other ancient pathogens in different preservation contexts and/or tissue types remains to be tested. Combined with targeted capture, pre-digestion washes helped us characterize 12 ancient *Y. pestis* genomes, including from material showing limited DNA preservation. Our work contributes to ongoing efforts aimed at mapping the genetic diversity of second pandemic plague strains.

RESULTS AND DISCUSSION

Sample characteristics

We collected 120 teeth from 89 human individuals that were excavated from French mass graves and cemeteries related to the second plague pandemic (Figure 1, Data S1). More specifically, a total of 56 teeth originated from the sites of La Major (1720 CE, Marseille) and Les Rayettes (1721 CE, Martigues), both associated with the last major plague outbreak in France. This epidemic started after the ship *Grand Saint Antoine* departed from the Levant and landed on the Marseille harbor in 1720 CE. It rapidly decimated half of the population of the city, before it spread to the Provence region, hitting Marseille a second time (causing considerably lower mortality), and finally vanished in 1722 CE.[49] Overall, it is estimated that this plague outbreak killed more than a quarter of the population of the Provence region. The remaining 64 teeth were collected from 58 individuals buried during the 17th c. CE, including from the Lariey Puy-Saint-Pierre cemetery (1629–1630 CE), as well as the Maladrerie Saint-Lazare (1630–1670 CE, Beauvais) and the Hôtel-Dieu Saint-Jean-Baptiste (1630–1670 CE, Amiens). While the Lariey cemetery was previously analyzed for *Y. pestis* and human ancient DNA,[45] the latter two sites were genetically investigated here for the first time. These sites correspond to plague mass graves formed at the time of the Thirty Years' War.

A total of 33 teeth were first subjected to the standard procedure for ancient DNA extraction routinely applied at the Centre for Anthropobiology and Genomics of Toulouse (CAGT) ancient DNA facilities. Briefly, the dental pulp powder is collected by cutting the tooth at the cementum-enamel junction in order to drill the pulp chamber from inside with a diamond ball bit. The powder is first pre-digested in a DNA extraction buffer for 1 h at 37°C. The undigested tooth pellets are then collected and digested until completion in a fresh DNA extraction buffer, following overnight incubation with agitation at 42°C. This procedure was previously optimized for recovering ancient DNA from the host,[50] but nonetheless led to the successful characterization of two ancient human and *Y. pestis* genomes from the Lariey cemetery, following shotgun sequencing on Illumina platforms.[45] Stringent alignment of shallow sequencing data against the *Y. pestis* reference genome[34,51] revealed limited, if any, *Y. pestis* DNA preservation across all the samples investigated (< 0.93%), regardless of the archaeological site considered (Table S1). DNA capture procedures were, thus, implemented to attempt to enrich for *Y. pestis* and human mtDNA templates, using the synthetic RNA 80-mers probes designed by Wagner and colleagues.[33] In this experiment, a total of 5 and 55 amplified DNA libraries were selected from individuals excavated at Lariey and La Major, respectively, as representatives of the entire range of DNA preservation conditions. Bioinformatic demultiplexing and collapsing of capture DNA reads provided between 14,386 and 2,250,565 sequences per

*Correspondence:
ludovic.orlando@univ-tlse3.fr

https://doi.org/10.1016/j.isci.2023.106787



Figure 1. Sample corpus

(A) Sample location and dates. The archaeological sites analyzed in this study are all from France and are shown in purple, with respect to those previously published, following the coloring scheme from Spyrou and colleagues,[43] for consistency. Pictures of the mass graves studied with the exception of the archaeological site of Lariey (12 teeth), which was investigated by Seguin-Orlando and colleagues[45] (the number of teeth investigated, N, is indicated on each picture).

(B) La Major (Marseille, 1720 CE).

(C) Les Rayettes (Martigues, 1721 CE).

(D) Hôtel-Dieu Saint-Jean-Baptiste (Amiens, 1630–1670 CE).

(E) Maladrerie Saint-Lazare (Beauvais, 1630–1670 CE). See also Tables S1 and S2.

individual, of which only those from two individuals from Lariey (LAR8 and LAR11) provided sufficient coverage of the *Y. pestis* genome (Table S1). Further metagenomics analyses in MetaPhlAn 4[52] confirmed a greater *Y. pestis* content in capture versus shotgun collapsed reads for 12 libraries prepared from six individual teeth (Figures 2A and 2B). This supported numbers of classified microbial species dropping with capture, as expected for a procedure aimed at focusing sequencing efforts on *Y. pestis* only. Accordingly, the fraction of environmental and oral bacteria identified following capture was reduced relative to shotgun experiments (Figure 2B).

While targeted DNA capture improved the *Y. pestis* DNA recovery rate, the vast majority of the sequence data for other individuals consisted of PCR duplicates (37.72–98.03%), which limited the cost-effectiveness as an approach for characterizing the plague genome. The high proportion of PCR duplicates indicated that the total amount of plague DNA templates present in the original extracts was extremely limited. We, thus, decided to identify DNA extraction procedures that could enhance the recovery of ancient plague DNA molecules.

Pre-digestion DNA release

Previous work reported increased proportions of host DNA in libraries constructed following complete digestion, relative to those obtained pre-digestion.[8,9] This suggested that pre-digestion could wash away a fraction of the environmental microbial DNA in contrast to the host DNA, which would be preserved deeper into the calcified matrix, and thus would require longer digestion time to be released.[53] We reasoned that the *Y. pestis* DNA would be mostly restricted to the blood circulation and would not

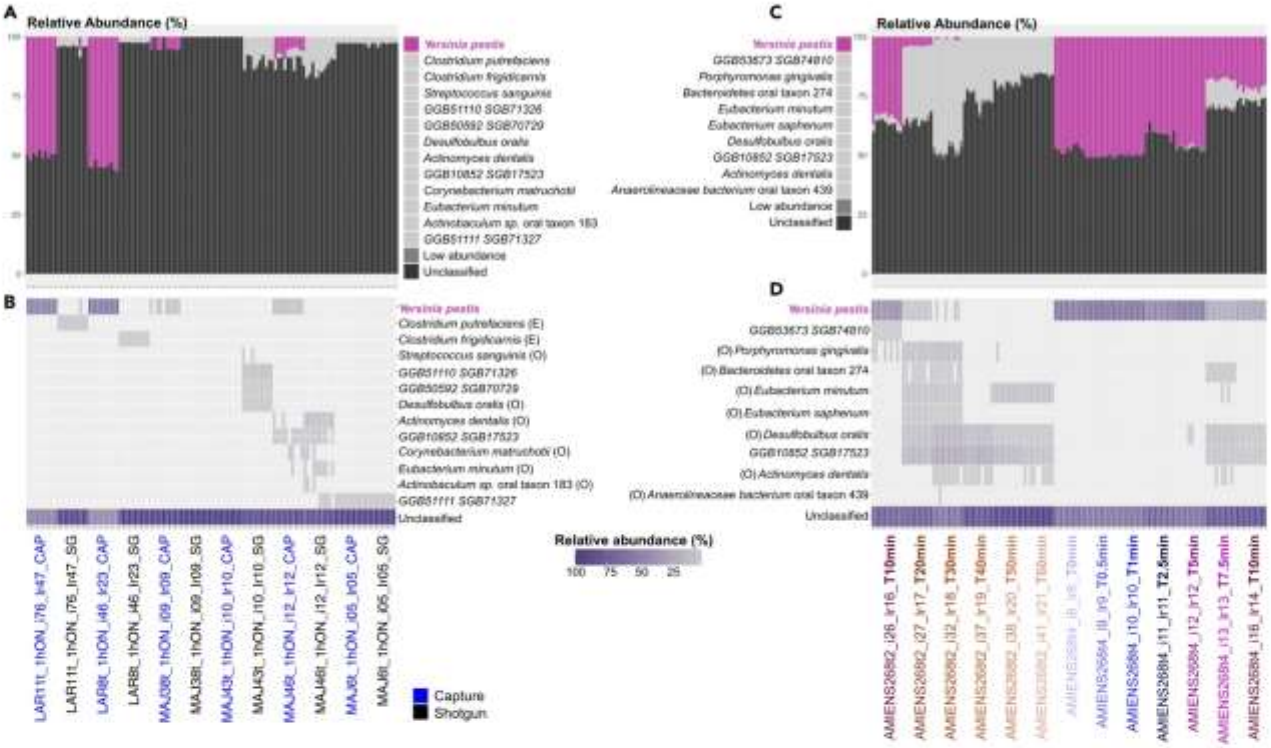


Figure 2. MetaPhlAn 4 microbial taxonomic profiles of the tooth material analyzed

(A and B) Comparison of microbial diversity in shotgun versus capture sequence datasets. These experiments were carried out on the extracts resulting from a 1-h pre-digestion, followed by an overnight digestion of the remaining pellets (1h+ON). (A, B) Relative abundances of microbial species.

(C and D) Same as A, B except that different pre-digestion times were contrasted from 10 to 60 min and from 0 to 10 min. The analyses were restricted to two teeth from the same sample (AMIENS268t4 and AMIENS268t2).

penetrate the deeper extracellular matrix. Under this assumption the pathogen DNA would follow release kinetics similar to the environmental microbial DNA. To test this theory, we purified the DNA from 17 1-h pre-digested fractions (1h) to prepare libraries for the capture of *Y. pestis* and human mtDNA genomes. The DNA contents of the pre-digested samples were compared to those obtained following capture of the DNA libraries prepared after complete pellet digestion (1h+O/N; Figures 3A and 3B). Normalizing for sequencing efforts after *Y. pestis* whole-genome capture, we found that most of the libraries constructed following 1 h pre-digestion (1h, without overnight digestion) contained larger proportions of *Y. pestis* sequences. We also observed inverse trends with clonality, which is reduced in 1-h pre-digested samples (Figure S1). Combined, these results support the presence of more *Y. pestis* DNA molecules in the pre-digested fraction.

Human mtDNA content followed no particular trend, representing lower, similar or greater proportions of the sequence data depending on the specimen considered. This indicated that only some of the remains analyzed still contained a substantial fraction of human DNA in the undigested pellets and that DNA preservation was considerably more limited in others.

Encouraged by the increased proportions of *Y. pestis* DNA present in the pre-digested fraction, we further investigated the kinetics of DNA release during the 60 min (min) pre-digestion wash. To achieve this, we generated fresh dental pulp powder for four additional teeth from individuals that showed positive plague identification in our previous experiments. The powder was then subjected to six consecutive partial digestion steps at 37°C of 10 min each (10 min, 20 min, 30 min, 40 min, 50 min, and 60 min; Figures 3C and 3D). The remaining undigested pellets were collected after each step, and incubated further in the same volume of a fresh extraction buffer. The resulting six pre-digested fractions were purified and used for library preparation, amplification, DNA capture, and sequencing (Figures 3C and 3D). Normalizing for sequencing efforts after *Y. pestis* whole-genome capture, we found that most of the *Y. pestis* DNA was released in the

extraction buffer during the first 10 min pre-digestion. The proportion of *Y. pestis* DNA released during further pre-digestion steps was consistently lower than that released during the previous incubations. This indicated rapid *Y. pestis* DNA release in the extraction buffer, and diminishing returns with increasing pre-digestion times. Human mtDNA release followed a similar trend for one tooth (MAJ46t9), but exhibited no consistent trend across the four teeth investigated. Interestingly, MetaPhlAn 4 taxonomic profiling of microbial diversity for the single of the four teeth for which sufficient sequencing data were generated (≥ 100,000 collapsed reads; AMIENS268t2), revealed increasingly abundant off-target microbial hits with pre-digestion times (Figures 2C and 2D). This translated into an increasing number of classified microbial species, consistent with capture probes not being saturated by *Y. pestis* DNA templates only allowing aspecific annealing, including to oral bacterial DNA (e.g., *Desulfobulbus oralis*; Figure 2D).

We repeated the same experiment for four additional teeth belonging to the same four individuals, except that shorter pre-digestion times were considered (0.5 min, 1 min, 2.5 min, 5 min, 7.5 min, and 10 min; Figures 3E and 3F). The proportions of *Y. pestis* and human mtDNA present in each resulting pre-digested fraction were compared to those found in the absence of pre-digestion (0 min), which were estimated after suspending the dental pulp powder in the pre-heated extraction buffer, and immediately centrifugating the pellets to collect the supernatant.

Two of the teeth examined were found to contain almost no *Y. pestis* DNA considering pre-digestion times up to 10 min (MAJ46t11 and SQ8723t4; Figures 3E and 3F). Other teeth from these individuals, however, were tested positive for *Y. pestis* in previous experimental conditions. This highlights that the success of *Y. pestis* detection strongly depends on the tooth material considered, possibly due to differences in DNA preservation across tooth microenvironments and/or tooth physical integrity conditioning colonization success from the oral and environmental communities.

The other two of the four teeth examined showed inconsistent kinetics of DNA release, with one sample (AMIENS268t4) releasing roughly twice as much *Y. pestis* DNA between 2.5 and 5 min than during the following 2.5 min (between 5 and 7.5 min), but the same proportion again between 7.5 and 10 min. For the last tooth (LAR27t4), the proportion of *Y. pestis* DNA collected in the last experimental condition was 6.43-fold–10.10-fold larger than during the two previous steps, despite lasting for 2.5 min each. It was also equivalent to the proportion of *Y. pestis* DNA obtained following a 1 min pre-digestion (Figure 3E). The AMIENS268t4 tooth provided sufficient numbers of collapsed reads for metagenomic profiling with MetaPhlAn 4, which revealed the presence of other microbes than *Y. pestis* in some but not all pre-digestion conditions (Figures 2C and 2D). Combined, these results indicated stochastic effects dominating DNA release of both *Y. pestis* and other microbes during the first 10 min of pre-digestion. No specific trend could again be observed for the release of human mtDNA.

As the pre-digestion conditions between 10 and 60 min provided consistent results across the four teeth investigated, we decided to further characterize the underlying kinetics of DNA release in the pre-digestion buffer. We found that the proportion of unique stringent alignments characterized decreased exponentially through time (Figures 4A and 4B). This was true for both *Y. pestis* and human mtDNA, which followed trends that were not significantly different (ANCOVA test, p values ≥ 0.19). As a result, the cumulative proportion of *Y. pestis* and human mtDNA templates sequenced rapidly reached saturation, leaving only 5% (1%) of unreleased templates after 28.0–38.5 min (43.1–59.2 min) of pre-digestion (Figures 4C, 4D, and S2). This indicated that pre-digesting the dental pulp powder for 30 to 60 min in the experimental conditions tested would be sufficient to release most of the *Y. pestis* DNA present. We, thus, performed the DNA analysis of the remaining teeth following a conservative pre-digestion for 60 min at 37°C to extract the DNA fraction of the pathogen almost entirely. The supernatant was purified and concentrated for preparing a total of 176 DNA libraries for *Y. pestis* and human mtDNA enrichment.

Phylogenetic analyses

The improved DNA extraction procedures developed in this study allowed us to characterize the genome sequence of *Y. pestis* strains present in 12 ancient individuals. Patterns of nucleotide mis-incorporations and base compositional profiles at the genomic positions preceding and following read termini were characteristic of authentic ancient DNA data generated following USER-treatment of extracts (Figure S3). Specifically, USER enzymes cleave those cytosine residues that are deaminated postmortem, leaving only slightly inflated C→T (G→A) mis-incorporation rates at read starts (ends), as well as an excess of ancient

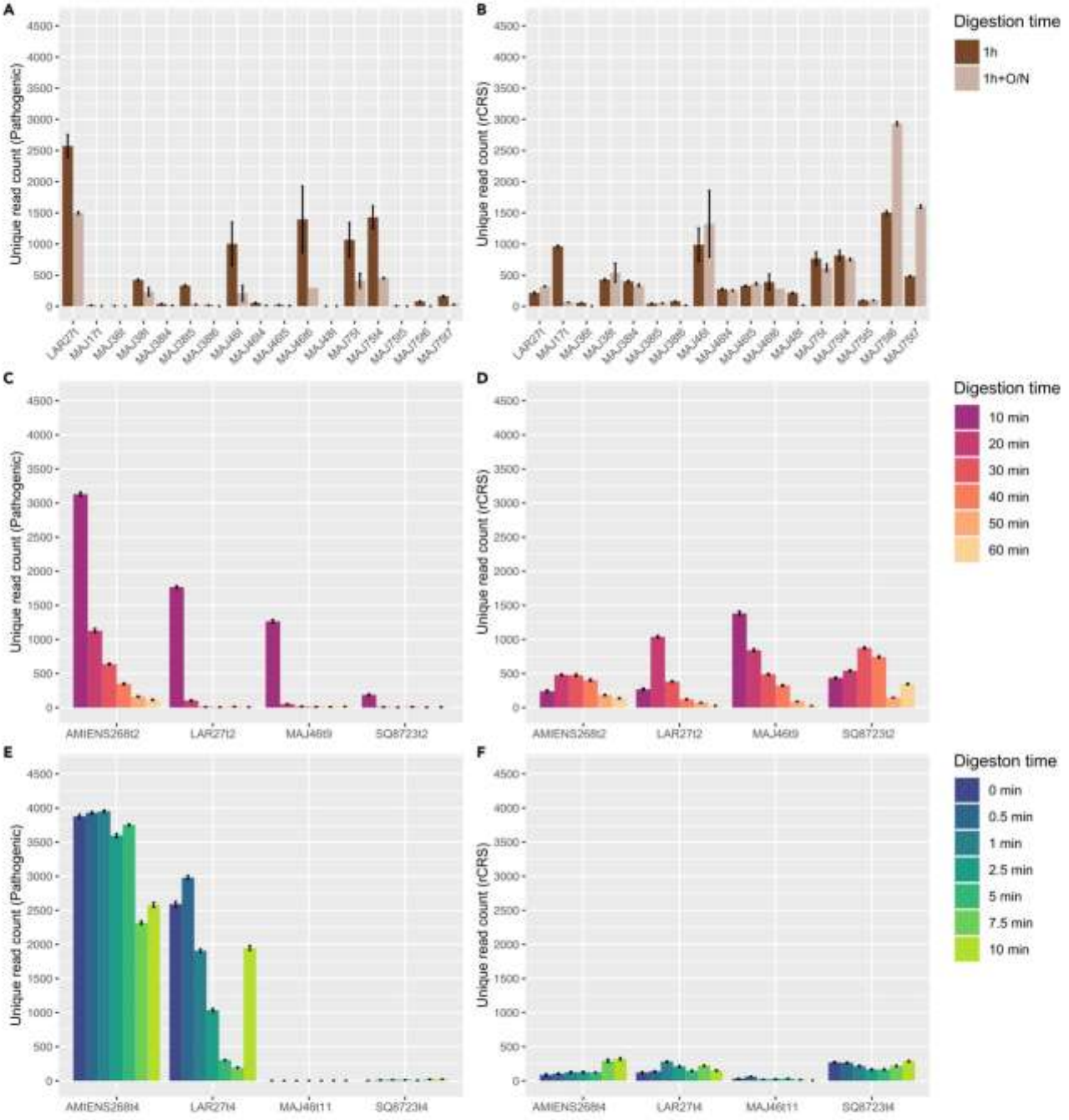


Figure 3. Number of unique high-quality alignments identified following different extraction conditions

(A and B) One hour of pre-digestion wash (1h, dark brown) versus complete pellet digestion overnight (1h+O/N, light brown).

(C and D) 10, 20, 30, 40, 50, and 60 min of pre-digestion wash (from purple to yellow).

(E and F) 0, 0.5, 1, 2.5, 5, 7.5, and 10 min of pre-digestion wash (from blue to green). Analyses are repeated 10 times, following random sampling of 4,557 collapsed reads in each experimental condition for normalizing sequencing efforts. Data are represented as mean ± SEM. Alignments are shown against the *Y. pestis* CO92 genome (left; A, C, E, pathogenic) and the human mtDNA genome (right; B, D, F, rCRS). See also Figure S1.

DNA templates starting immediately after cytosine residues.[2] Read-to-reference edit distance distributions confirmed a closer genetic proximity to *Y. pestis* relative to the closest outgroup, *Yersinia pseudotuberculosis* (Figure S4).

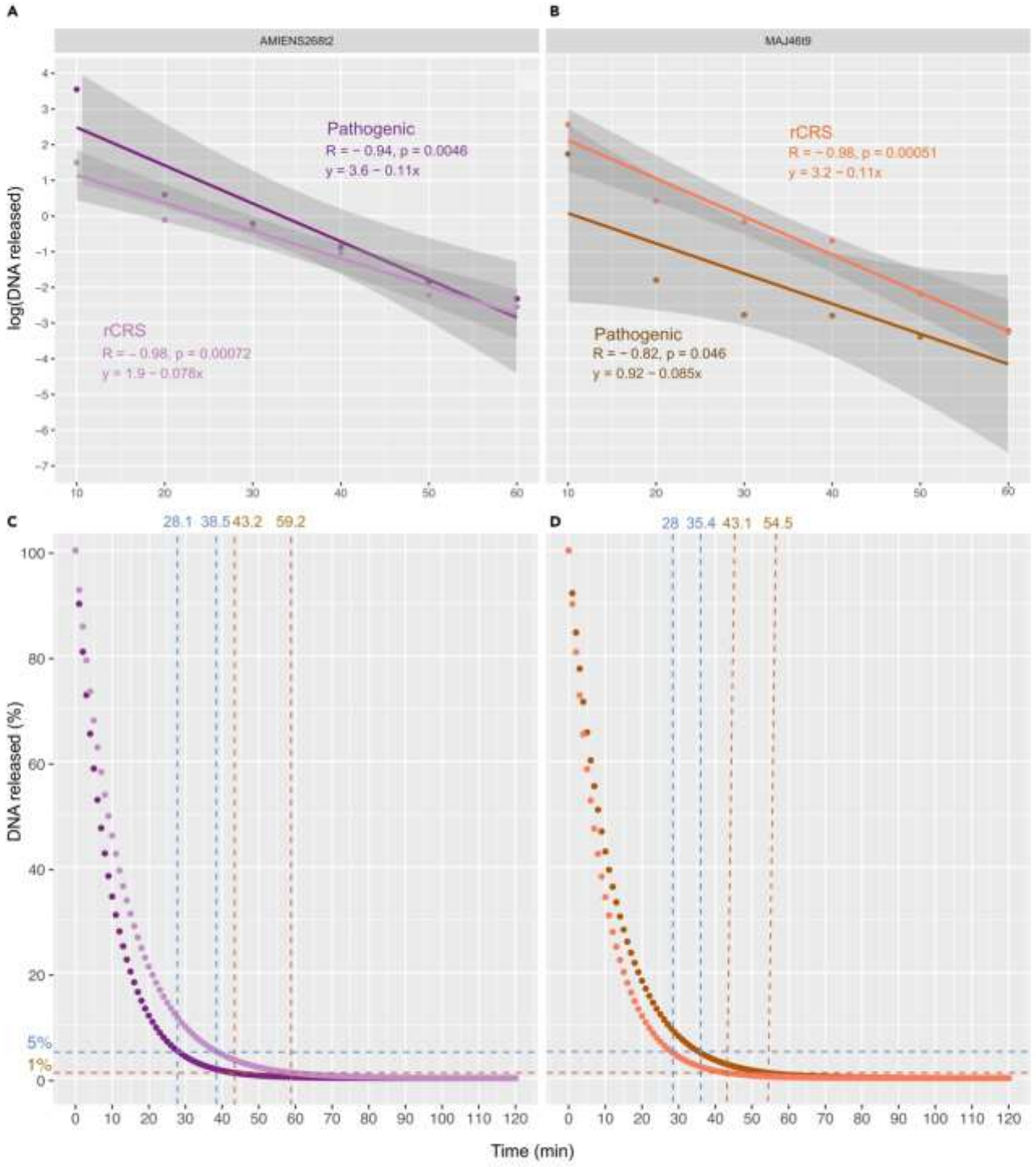


Figure 4. Kinetics of DNA release during pre-digestion wash

(A and B) Percentage of DNA released following pre-digestion wash of incremental time (10 min–1 h). The results are shown for the AMIENS268t2 (A) and MAJ46t9 (B) teeth (see Figure S2 for LAR27t2 and SQ8723t2 results). High-quality sequence alignments are aligned against the pathogen genome (CO92 and plasmids, dark purple, and dark orange) or the human mtDNA genome (Revised Cambridge Reference Sequence rCRS, light purple, and light orange). (C and D) Prediction for the DNA fraction remaining to be released following incremental pre-digestion wash. Predictions are based on the linear models shown in panels A and B for *Y. pestis* (dark purple and dark orange) and human mtDNA (light purple and light orange). The blue and dark orange dashed lines indicate the times when respectively 95% and 99% of the DNA has been released. See also Figure S2.

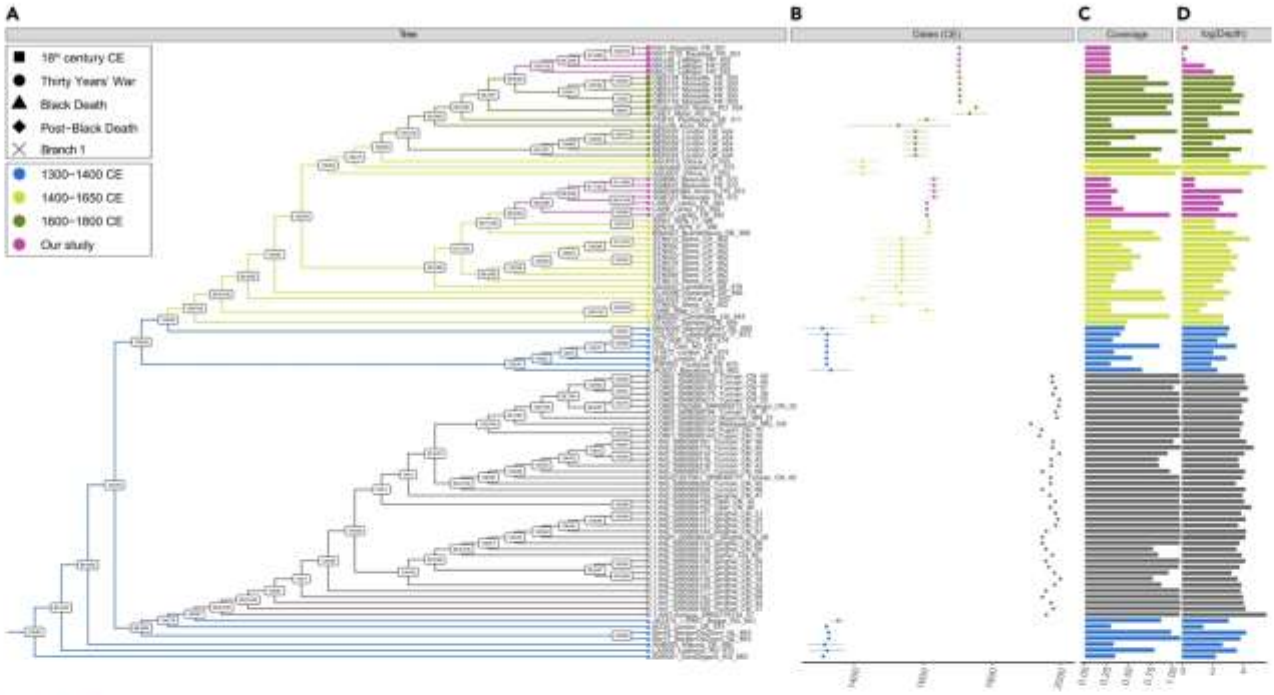


Figure 5. Phylogenetic reconstruction

(A) Maximum Likelihood tree (IQ-TREE), considering 29,609 genome-wide variant positions known to provide reliable phylogenetic signal. This analysis included 185 previously published genomes, but the tree shown is restricted to second and third pandemic genomes, disregarding branch lengths, for clarity (the tree with branch length is shown in Figure S5). Node supports are estimated from SH-aLRT calculations (left) and ultrafast bootstrap approximation (right).

(B) Sample dates CE, with error margins representing the lower and upper boundaries provided by archaeological contexts.

(C and D) represent the coverage and the logarithm of the average sequencing depth-of-coverage achieved for the different CO92 genomes considered. The colors illustrated follow the coloring scheme from Spyrou et al., 2022,[43] and the tip shapes indicate the historical context of archaeological samples investigated with respect to Figure 1. See also Figures S5 and S6.

Phylogenetic reconstructions revealed that most of the second pandemic *Y. pestis* genomes clustered together within a monophyletic group, including strains from the Black Death to the 18th c. CE (Figures 5 and S5, Table S2). Within this cluster, the 12 ancient *Y. pestis* genomes characterized in this study grouped together with all the other post-Black Death ancient *Y. pestis* genomes hitherto sequenced. This was true genome-wide (Figure S6), or when considering only a subset of 29,609 parsimony informative sites located outside non-core, repetitive or highly conserved regions, which are reported to be less prone to environmental contaminant mis-mapping (Figures 5 and S5).[54] The latter analysis confirmed the 1338–1339 CE genome from KaraDjigach, Kyrgystan, as basal to all subsequent strains sequenced, belonging to the second and third pandemic. Phylogenetic reconstructions also supported all modern branch 1 genomes responsible for the third pandemic descending from a subset of 14th c. CE strains identified in the Netherlands (BergenOpZoom), the United Kingdom (London), and Russia (Bolgar), as previously reported.[35,55,56] Additionally, the LAR27 genome from Lariey characterized in this study (1629–1630 CE) appeared closely related to the two genomes previously reported from the same site (LAR8 and LAR11), confirming their genetic proximity with contemporary *Y. pestis* genomes from San Procolo a Naturno (Italy, 1636 CE).[45] The *Y. pestis* genomes from Beauvais and Amiens (1630–1670 CE) also clustered with this group, suggesting common epidemiological origins. This lineage further spread during or in the aftermath of the Thirty Years' War (Figures 5 and S5), which is recognized as one of the most destructive conflicts in European history, it affected Central Europe between 1618 and 1648 CE, a time contemporary with the ancient *Y. pestis* genome from Brandenburg, which is located basally within this group.

The *Y. pestis* genome sequences retrieved from the samples of La Major (Marseille, 1720 CE) and Les Rayettes (Martigues, 1721 CE) formed a monophyletic group together with the previously reported genomes from Observance (Marseille, 1722 CE).[57] This supports historical sources that describe an initial

introduction of the plague in Marseille by the ship *Grand Saint Antoine* on May 1720 CE,[49] and a subsequent rapid spread to the entire Provence region. Interestingly, our phylogenetic reconstructions indicated that the ancient *Y. pestis* genomes from Observance do not directly descend from those entering Marseille two years earlier and present at La Major. Taken at face value, this may have provided support for the introduction of two closely related plague strains. However, we caution that due to extreme DNA degradation, the genomes from La Major and Les Rayettes could only be characterized with limited coverage (between 1.19-fold and 5.25-fold), despite the improved methodology described in this study. Furthermore, the majority of ancient strains presented in our phylogenetic analysis appear to have excessively long terminal branches, even when compared to modern genetic diversity (Figures S5 and S6). The combination of low-genome coverage, remnant sequencing errors due to DNA degradation, and residual contamination from environmental microbes, can result in excessive terminal branch lengths in ancient genomes. The presence of such long branches can introduce significant biases in phylogenetic placement,[58,59] possibly impacting fine-scale evolutionary predictions. Additionally, we note that the Observance genomes were characterized using a different DNA capture system,[57] introducing subtle technical batch effects, also potentially impacting phylogenetic reconstruction. This can be illustrated by the single Lariey genome characterized in this study, following DNA capture, which appears phylogenetic close, but does not form a monophyletic clade with those previously characterized by shotgun sequencing in Seguin-Orlando and colleagues.[45] In light of these caveats, we consider our phylogenetic reconstructions consistent with the abundance of historical sources supporting a unique introduction of the plague in Marseille in 1720 CE.[49] Determining whether the underlying strain first emerged in the Levant, or derived from a European source re-entering Europe through the *Grand Saint-Antoine*, requires further work. For example, mapping the genetic structure of *Y. pestis* strains at a finer-grained resolution during the 17th and the 18th c. CE may clarify the geographic sources and spread of related plague outbreaks. Integrating historical evidence will however be essential, given that the mutation rate of the *Y. pestis* genome provides genetic resolution at the 10-year timescale,[60] which is considerably slower than the pace of human circulation and trade.

Genome analyses

We further tested whether the ancient *Y. pestis* strains sequenced in this study showed specific genetic characteristics compared to their closest phylogenetic relatives. The average depth-of-coverage of the 12 individuals sequenced here ranged between 0.16 and 56.16-fold for the circular chromosome CO92 (median = 3.97-fold, average = 11.67-fold), while it ranged between 0.79 and to221.45-fold, 0.20–to73.46-fold and 5.71– to1092.67-fold for the pCD1, pMT1, and pPCP1 plasmids, respectively. Interestingly, the average proportion of high-quality alignments against the CO92 chromosome was lower-than-expected on the basis of the fraction of CO92 probes used for capture. The average proportions of high-quality alignment against the three different plasmids were, conversely, larger-than-expected on the basis of the probe distribution (Figures 6A–6D). This is in line with bacterial cells containing multiple copies of plasmids. The excess of plasmid DNA relative to probe proportions was, however, significantly more limited in the specimens dated from the 17th c. CE than in those from the 18th c. CE (Figures 6B–6D). This suggests that the latter contained a greater number of plasmid copies per cell than the former. Interestingly, the plasmid copy number can have an impact on virulence,[61] in particular the pCD1 plasmid, which carries the T3SS genes responsible for delivering *Yersinia* outer membrane proteins into the host cells.[62] The number of plasmids carried by the 17th and 18th c. CE strains may, thus, have driven differences in their virulence. Whether this genetic feature played a role in the excessive mortality observed during the Great Plague of Marseille, which decimated half the city's population in only a few months,[49] remains to be tested. Whether it also had equally dramatic consequences on the populations of rodent natural hosts, thus, precipitating disease disappearance is also unknown.

Furthermore, a 49-kb region from the CO92 chromosome was found to be largely absent in the strains from La Major and Les Rayettes (Figures 7A–7E). The *cspE* gene (cspE_1883773–1883985), which is involved in the survival and the multiplication at low temperature,[63] returned spurious non-null coverage, due to sporadic alignments resulting from environmental *Yersinia* bacteria. The 49-kb region spans the *mgtA* (mgtA_1888533–1891232, previously named *mgtB*) and *mgtC* (YPORS09280_1887434–1888132) genes involved in the survival of the bacteria under condition with low concentrations of Mg^{2+}.[64] It was previously reported to also be missing from the genomes of their closest phylogenetic relatives, including 18th c. CE strains from France (Observance, Marseille), Russia (Maist), and Sweden (Pestbacken), as well as mid-16th/mid-17th c. CE strains from the United Kingdom (London).[51] Moreover, a 47-kb region overlapping 11

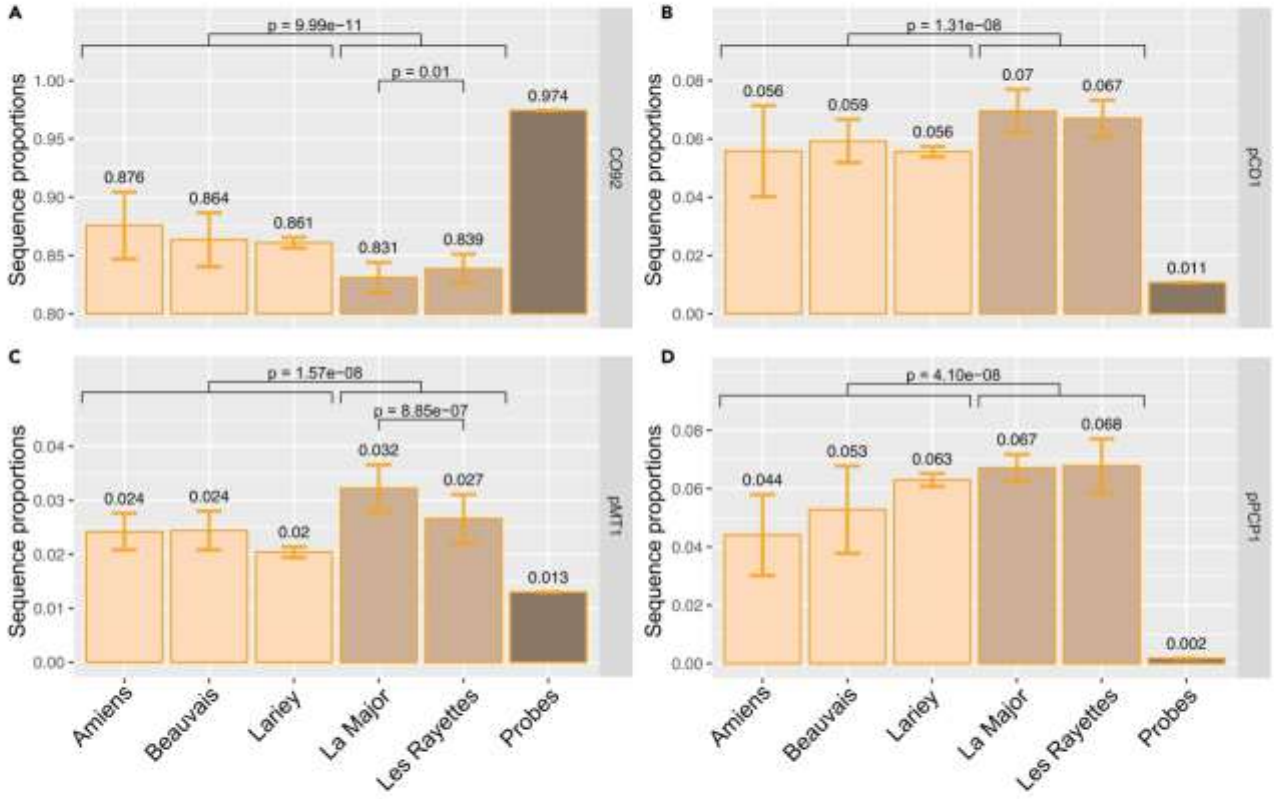


Figure 6. Plasmid copy numbers per bacterial cell

(A–D)Average proportions of sequences mapping against the CO92 circular chromosome (A) and the plasmids (B, C, and D) for the different archaeological sites from the 17th c. CE (orange, Amiens, Beauvais, and Lariey) and the 18th c. CE (light brown, La Major, and Les Rayettes). Probe proportions targeting the CO92 genome and the three plasmids are indicated by the brown bars. Data are represented as mean ± SEM. Statistically significant values (Wilcoxon test) for inter- and intra-century comparisons are indicated. Only those libraries prepared on the DNA extracted following 1-h pre-digestion were considered to avoid technical batch effects, disregarding those for which an insufficient number of high-quality read alignments could be identified (*i.e.* < 500).

virulence and pathogenicity factors (inv_2040296–2042350, fyuA_2140840–2142861, ybtE_2142992–2144569, ybtT_2144573–2145376, ybtU_2145373–2146473, irp1_2146470–2155961, irp2_2156049–2162156, ybtA_2162347–2163306, ybtP_2163563–2165275, ybtQ_2165262–2167064, ybtX_2167057–2168337) was largely under-covered in all the genomes from La Major, Beauvais, Lariey and Amiens reported here (Figure 7A). They were, however, covered in two genomes from Lariey that were previously characterized using shotgun sequencing. We, thus, consider that the limited coverage observed most likely reflects under-performing probe-template annealing in those regions than genuine genomic rearrangements.

We next used snpToolkit[56] to characterize the single nucleotide mutations found among the new 17th and 18th c. CE strains sequenced here, and their closest phylogenetic relatives. This analysis revealed a total of 375 SNPs, comprising 148 non-synonymous, 97 synonymous, and 130 intergenic mutations (Figure S7, Table S3). Interestingly, the iron ABC transporter encoded by the *YbtP* gene, which is essential for the bacterial physiology[65] and can deplete *Y. pestis* avirulence in mice,[66] was found to be polymorphic in both strains from the 17th and 18th c. CE. No mutations were otherwise found in the 235 virulence and pathogenicity loci considered in the analyses above. The genomes from Amiens and Beauvais *Y. pestis* strains (1630–1670 CE) showed remarkable genetic similarity with those from Lariey (1629–1630 CE), as they shared only eight non-synonymous variants not found in the latter (Table S3). One such variant affected the autotransporter *YapC* gene, which mediates adhesion to human cells and, thus, may affect host innate immune response.[67] In contrast, the number of non-synonymous differences found between the 1720–1722 CE genomes from La Major, Les Rayettes, and Observance was larger. Whether this is indicative of currently unidentified minor mis-mapping from other bacteria than *Y. pestis*, or faster mutation rates in this specific lineage requires further clarification through additional sequencing of complete genomes from *Y. pestis* strains responsible for the plague epidemic that spreads to the Provence region in the early 18th c. CE.

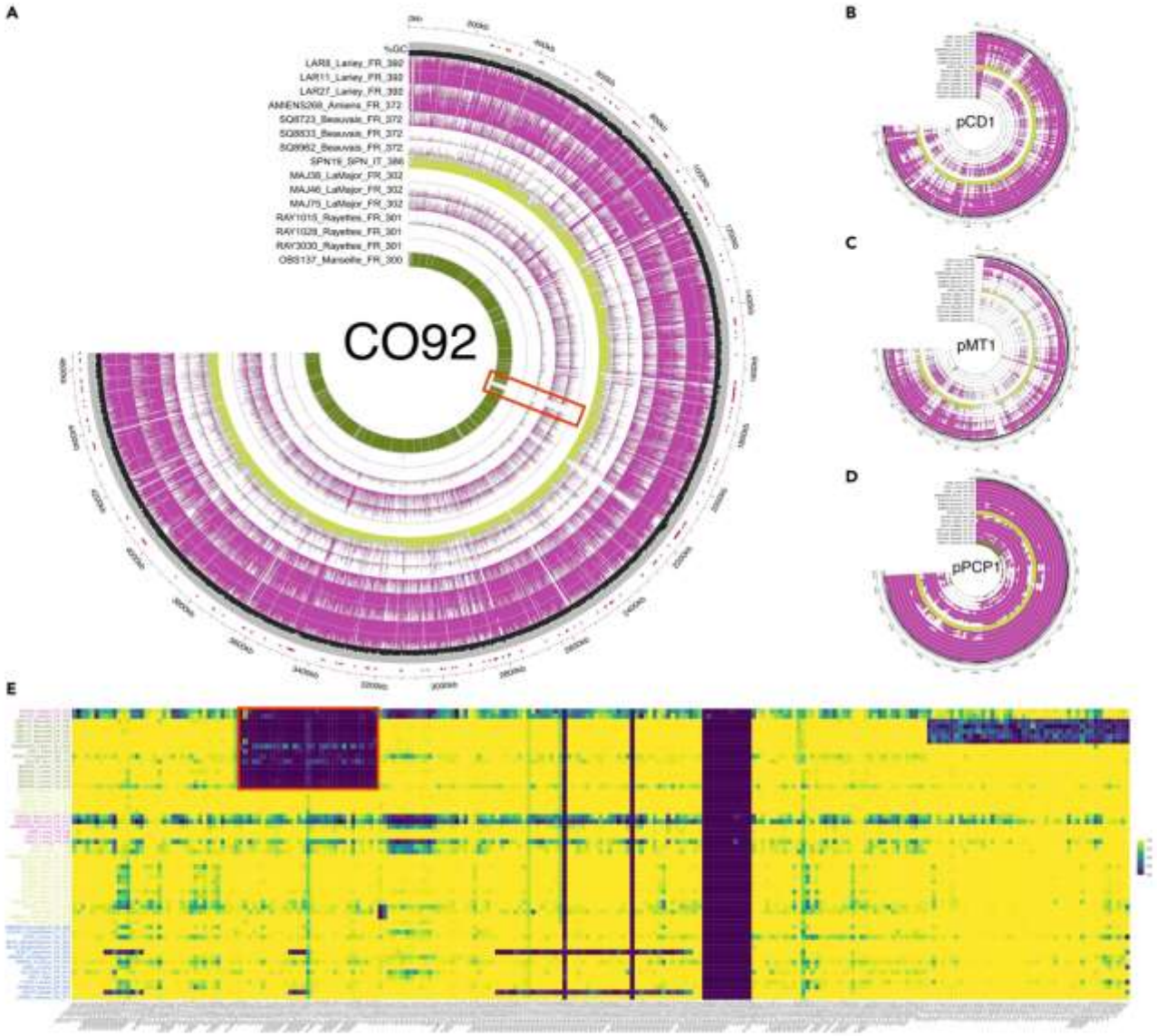


Figure 7. Genome coverage and %GC variation

(A) CO92 *Y. pestis* reference genome.

(B) pCD1 plasmid.

(C) pMT1 plasmid.

(D) pPCP1 plasmid. A total of 13 samples sequenced in this study are shown (purple) together with two close relatives from the 17th c. CE (yellow) and the 18th c. CE (green) for comparison.

(E) Fraction of the gene covered at least once for 235 virulence and pathogenicity loci. Second pandemic *Y. pestis* genomes were considered as long as they were sequenced to a minimum 2-fold average depth-of-coverage. The red box highlights the 49-kb deletion identified, with "X" indicating those virulence and pathogenicity loci showing best Blast hits against environmental *Yersinia* species, and, thus, likely reflecting false positive alignments against *Y. pestis*. See also Figures S3, S4, and S7 and Table S3.

Conclusion

In this study, we report that ancient *Y. pestis* DNA is almost entirely released during a 60 min pre-digestion wash of the dental pulp. This finding might have important methodological implications for the workflow of ancient pathogen studies, as DNA extracts obtained following pre-digestion appear to maximize the ratio of ancient pathogen DNA to human DNA. This contrasts to the extracts obtained following complete pellet digestion, which are generally used for the study of ancient plagues, despite containing a much greater

ratio of human to pathogen DNA. In the experimental conditions described herein, a pre-digestion of the dental pulp powder for 60 min at 37°C appears sufficient to release at least 99% of the ancient *Y. pestis* DNA generally preserved. Applying targeted capture to these DNA extracts allowed us to characterize 12 ancient *Y. pestis* genomes from five French archaeological sites dated to the 17th and 18th c. CE. We find that the number of plasmid copies within *Y. pestis* bacteria from the 17th c. CE was lower than in those from the 18th c. CE. The Thirty Years' War likely facilitated the spread of a plague lineage throughout Germany, Italy, and France, all of which were eventually dead-end emergences. The last plague outbreak in France appears to follow a single emergence in Marseille, before spreading to the Provence region.

Limitations of the study

In this study, genome variation among ancient *Y. pestis* strains was only investigated from sequence alignments against one single reference genome. The full extent of structural variation and their possible consequences on transmission, pathogenicity and virulence remains to be characterized. Moreover, the quality of the data generated for most of the ancient genomes presented was limited, which precluded the examination of fine-scale strain evolution and divergence dates. Only a small number of samples were subjected to pre-digestion tests, they were also confined to a single geographic area and, only encompassed preservation conditions from a relatively recent period. Even if our results highlight a strong pattern for the DNA release of *Y. pestis*, experiments should be extended to various areas, periods and epidemics. Finally, whether the recommended methodology shows similar performance on other types of non-dental tissues, and will equally apply to the entire range of pathogenic bacteria that can circulate in the blood, remains untested.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Samples and archaeological sites
- METHOD DETAILS
  - Ancient DNA extraction
  - Library construction and PCR amplification
  - Capture and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106787.

ACKNOWLEDGMENTS

We thank all the members of the AGES group at CAGT for fruitful discussion and Alice Rabasse for proofreading the manuscript. This research was funded by the CNRS MITI (*Mission pour les Initiatives Transverses et Interdisciplinaires*) IndigenousHealth program; the ANR LifeChange; the ANR GenIn; the Simone and Cino Del Duca Foundation (Subventions scientifiques 2020, HealthTime Travel) and the European Research Council (ERC), under the European Union's Horizon 2020 research and innovation program (grant agreement 681605).

AUTHOR CONTRIBUTIONS

L.O. conceived the project and designed the research. N.T., B.C., R.J., S.T., L.B., J.M.F., M.S., and C.C. provided samples and information about historical and archaeological context. P.C., C.D.S., C.T., C.G., and A.S.O. carried out ancient DNA laboratory work with input from L.C., G.T., S.S., and L.C.T., who managed the sequencing runs. P.C., L.L., C.D.S., and L.O. carried out computational analyses, with input from A.S.O. and L.O. provided reagents and material. J.K. and H.P. provided the probe design. P.C., C.D.S., A.S.O., and L.O. wrote the paper with input from J.K., H.P., and M.A.S., and all the coauthors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 22, 2022
Revised: February 11, 2023
Accepted: April 26, 2023
Published: May 2, 2023

REFERENCES

1. Llamas, B., Willerslev, E., and Orlando, L. (2017). Human evolution: a tale from ancient genomes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20150484. https://doi.org/10.1098/rstb.2015.0484.

2. Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., and Warinner, C. (2021). Ancient DNA analysis. Nat. Rev. Methods Primers *1*, 14–26. https://doi.org/10.1038/s43586-020-00011-0.

3. Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., et al. (2015). Optimal ancient DNA yields from the inner ear part of the human petrous bone. PLoS One *10*, e0129102. https://doi.org/10.1371/journal.pone.0129102.

4. Sirak, K., Fernandes, D., Cheronet, O., Harney, E., Mah, M., Mallick, S., Rohland, N., Adamski, N., Broomandkhoshbacht, N., Callan, K., et al. (2020). Human auditory ossicles as an alternative optimal source of ancient DNA. Genome Res. *30*, 427–436. https://doi.org/10.1101/gr.260141.119.

5. Harney, É., Cheronet, O., Fernandes, D.M., Sirak, K., Mah, M., Bernardos, R., Adamski, N., Broomandkhoshbacht, N., Callan, K., Lawson, A.M., et al. (2021). A minimally destructive protocol for DNA extraction from ancient teeth. Genome Res. *31*, 472–483. https://doi.org/10.1101/gr.267534.120.

6. Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., and Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques *59*, 87–93. https://doi.org/10.2144/000114320.

7. Boessenkool, S., Hanghøj, K., Nistelberger, H.M., Der Sarkissian, C., Gondek, A.T., Orlando, L., Barrett, J.H., and Star, B. (2017). Combining bleach and mild predigestion improves ancient DNA recovery from bones. Mol. Ecol. Resour. *17*, 742–751. https://doi.org/10.1111/1755-0998.12623.

8. Ginolhac, A., Vilstrup, J., Stenderup, J., Rasmussen, M., Stiller, M., Shapiro, B., Zazula, G., Froese, D., Steinmann, K.E., Thompson, J.F., et al. (2012). Improving the performance of true single molecule sequencing for ancient DNA. BMC Genom. *13*, 177. https://doi.org/10.1186/1471-2164-13-177.

9. Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., and Allentoft, M.E. (2015). Improving access to endogenous DNA in ancient bones and teeth. Sci. Rep. *5*, 11184. https://doi.org/10.1038/srep11184.

10. Essel, E., Korlević, P., and Meyer, M. (2021). A method for the temperature-controlled extraction of DNA from ancient bones. Biotechniques *71*, 382–386. https://doi.org/10.2144/btn-2021-0025.

11. Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503. https://doi.org/10.1038/nature16152.

12. Rohland, N., Mallick, S., Mah, M., Maier, R., Patterson, N., and Reich, D. (2022). Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. Genome Res. *32*, 2068–2078. https://doi.org/10.1101/gr.276728.122.

13. Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H.A., Kelso, J., and Pääbo, S. (2013). DNA analysis of an early modern human from Tianyuan Cave, China. Proc. Natl. Acad. Sci. USA *110*, 2223–2227. https://doi.org/10.1073/pnas.1221359110.

14. Carpenter, M.L., Buenrostro, J.D., Valdiosera, C., Schroeder, H., Allentoft, M.E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S., Nekhrizov, G., et al. (2013). Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. Am. J. Hum. Genet. *93*, 852–864. https://doi.org/10.1016/j.ajhg.2013.10.002.

15. Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.-M., and Poinar, H.N. (2014). Ancient whole genome enrichment using baits built from modern DNA. Mol. Biol. Evol. *31*, 1292–1294. https://doi.org/10.1093/molbev/msu074.

16. Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., et al. (2010). A draft sequence of the neandertal genome. Science *328*, 710–722. https://doi.org/10.1126/science.1188021.

17. Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., et al. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature *463*, 757–762. https://doi.org/10.1038/nature08835.

18. Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., et al. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature *468*, 1053–1060. https://doi.org/10.1038/nature09710.

19. Allentoft, M.E., Sikora, M., Refoyo-Martínez, A., Irving-Pease, E.K., Fischer, A., Barrie, W., Ingason, A., Stenderup, J., Sjögren, K.-G., Pearson, A., et al. (2022). Population genomics of stone Age eurasia. Preprint at bioRxiv 2022. https://doi.org/10.1101/2022.05.04.490594.

20. Lazaridis, I., Alpaslan-Roodenberg, S., Acar, A., Açıkkol, A., Agelarakis, A., Aghikyan, L., Akyüz, U., Andreeva, D., Andrijašević, G., Antonović, D., et al. (2022). The genetic history of the southern arc: a bridge between west asia and Europe. Science *377*, eabm4247. https://doi.org/10.1126/science.abm4247.

21. Spyrou, M.A., Bos, K.I., Herbig, A., and Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. Nat. Rev. Genet. *20*, 323–340. https://doi.org/10.1038/s41576-019-0119-1.

22. Pajer, P., Dresler, J., Kabickova, H., Písa, L., Aganov, P., Fucik, K., Elleder, D., Hron, T., Kuzelka, V., Velemínsky, P., et al. (2017). Characterization of two historic smallpox specimens from a Czech museum. Viruses *9*, E200. https://doi.org/10.3390/v9080200.

23. Ferrari, G., Neukamm, J., Baalsrud, H.T., Breidenstein, A.M., Ravinet, M., Phillips, C., Rühli, F., Bouwman, A., and Schuenemann, V.J. (2020). Variola virus genome sequenced from an eighteenth-century museum specimen supports the recent origin of smallpox. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190572. https://doi.org/10.1098/rstb.2019.0572.

24. Kahila Bar-Gal, G., Kim, M.J., Klein, A., Shin, D.H., Oh, C.S., Kim, J.W., Kim, T.-H., Kim, S.B., Grant, P.R., Pappo, O., et al. (2012). Tracing hepatitis B virus to the 16th century in a Korean mummy. Hepatology *56*, 1671–1680. https://doi.org/10.1002/hep.25852.

25. Krause-Kyora, B., Susat, J., Key, F.M., Kühnert, D., Bosse, E., Immel, A., Rinne, C., Kornell, S.-C., Yepes, D., Franzenburg, S., et al. (2018). Neolithic and medieval virus

genomes reveal complex evolution of hepatitis B. Elife *7*, e36666. https://doi.org/10.7554/eLife.36666.

26. Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jäger, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., et al. (2013). Genome-wide comparison of medieval and modern Mycobacterium leprae. Science *341*, 179–183. https://doi.org/10.1126/science.1238286.

27. van Dorp, L., Gelabert, P., Rieux, A., de Manuel, M., de-Dios, T., Gopalakrishnan, S., Carøe, C., Sandoval-Velasco, M., Fregel, R., Olalde, I., et al. (2020). Plasmodium vivax malaria viewed through the lens of an eradicated European strain. Mol. Biol. Evol. *37*, 773–785. https://doi.org/10.1093/molbev/msz264.

28. Marciniak, S., and Poinar, H.N. (2019). Ancient pathogens through human history: a paleogenomic perspective. In Paleogenomics: Genome-Scale Analysis of Ancient DNA Population Genomics, C. Lindqvist and O.P. Rajora, eds. (Springer International Publishing), pp. 115–138. https://doi.org/10.1007/13836_2018_52.

29. Margaryan, A., Hansen, H.B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., et al. (2018). Ancient pathogen DNA in human teeth and petrous bones. Ecol. Evol. *8*, 3534–3542. https://doi.org/10.1002/ece3.3924.

30. Duchêne, S., Ho, S.Y.W., Carmichael, A.G., Holmes, E.C., and Poinar, H. (2020). The recovery, interpretation and use of ancient pathogen genomes. Curr. Biol. *30*, R1215–R1231. https://doi.org/10.1016/j.cub.2020.08.081.

31. Fellows Yates, J.A., Andrades Valtueña, A., Vågene, Å.J., Cribdon, B., Velsko, I.M., Borry, M., Bravo-Lopez, M.J., Fernandez-Guerra, A., Green, E.J., Ramachandran, S.L., et al. (2021). Community-curated and standardised metadata of published ancient metagenomic samples with AncientMetagenomeDir. Sci. Data *8*, 31. https://doi.org/10.1038/s41597-021-00816-y.

32. Malyarchuk, A.B., Andreeva, T.V., Kuznetsova, I.L., Kunizheva, S.S., Protasova, M.S., Uralsky, L.I., Tyazhelova, T.V., Gusev, F.E., Manakhov, A.D., and Rogaev, E.I. (2022). Genomics of ancient pathogens: first advances and prospects. Biochemistry *87*, 242–258. https://doi.org/10.1134/S0006297922030051.

33. Wagner, D.M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J.W., Enk, J., Birdsell, D.N., Kuch, M., Lumibao, C., et al. (2014). Yersinia pestis and the plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect. Dis. *14*, 319–326. https://doi.org/10.1016/S1473-3099(13)70323-2.

34. Keller, M., Spyrou, M.A., Scheib, C.L., Neumann, G.U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., Ribera i Lacomba, A., Raynaud, C., et al. (2019). Ancient Yersinia pestis genomes from across western Europe reveal early diversification during the first pandemic (541–750). Proc. Natl. Acad. Sci. USA *116*, 12363–12372. https://doi.org/10.1073/pnas.1820447116.

35. Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., DeWitte, S.N., Meyer, M., Schmedes, S., et al. (2011). A draft genome of Yersinia pestis from victims of the Black Death. Nature *478*, 506–510. https://doi.org/10.1038/nature10549.

36. Rasmussen, S., Allentoft, M.E., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K.G., Pedersen, A.G., Schubert, M., Van Dam, A., Kapel, C.M.O., et al. (2015). Early divergent strains of Yersinia pestis in eurasia 5,000 Years ago. Cell *163*, 571–582. https://doi.org/10.1016/j.cell.2015.10.009.

37. Andrades Valtueña, A., Mittnik, A., Key, F.M., Haak, W., Allmäe, R., Belinskij, A., Daubaras, M., Feldman, M., Jankauskas, R., Janković, I., et al. (2017). The stone Age plague and its persistence in eurasia. Curr. Biol. *27*, 3683–3691.e8. https://doi.org/10.1016/j.cub.2017.10.025.

38. Spyrou, M.A., Tukhbatova, R.I., Wang, C.-C., Valtueña, A.A., Lankapalli, A.K., Kondrashin, V.V., Tsybin, V.A., Khokhlov, A., Kühnert, D., Herbig, A., et al. (2018). Analysis of 3800-year-old Yersinia pestis genomes suggests Bronze Age origin for bubonic plague. Nat. Commun. *9*, 2234. https://doi.org/10.1038/s41467-018-04550-9.

39. Rascovan, N., Sjögren, K.G., Kristiansen, K., Nielsen, R., Willerslev, E., Desnues, C., and Rasmussen, S. (2019). Emergence and spread of basal lineages of Yersinia pestis during the neolithic decline. Cell *176*, 295–305.e10. https://doi.org/10.1016/j.cell.2018.11.005.

40. Yu, H., Spyrou, M.A., Karapetian, M., Shnaider, S., Radzeviciūtė, R., Nägele, K., Neumann, G.U., Penske, S., Zech, J., Lucas, M., et al. (2020). Paleolithic to Bronze Age siberians reveal connections with first Americans and across eurasia. Cell *181*, 1232–1245.e20. https://doi.org/10.1016/j.cell.2020.04.037.

41. Susat, J., Lübke, H., Immel, A., Brinker, U., Macāne, A., Meadows, J., Steer, B., Tholey, A., Zagorska, I., Gerhards, G., et al. (2021). A 5,000-year-old hunter-gatherer already plagued by Yersinia pestis. Cell Rep. *35*, 109278. https://doi.org/10.1016/j.celrep.2021.109278.

42. Barbieri, R., Signoli, M., Chevé, D., Costedoat, C., Tzortzis, S., Aboudharam, G., Raoult, D., and Drancourt, M. (2020). Yersinia pestis: the natural history of plague. Clin. Microbiol. Rev. *34*. e000444-19. https://doi.org/10.1128/CMR.00044-19.

43. Spyrou, M.A., Musralina, L., Gnecchi Ruscone, G.A., Kocher, A., Borbone, P.-G., Khartanovich, V.I., Buzhilova, A., Djansugurova, L., Bos, K.I., Kühnert, D., et al. (2022). The source of the Black Death in fourteenth-century central Eurasia. Nature *606*, 718–724. https://doi.org/10.1038/s41586-022-04800-3.

44. Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L.A., Wang, Z., Guo, Z., Xu, L., et al. (2013). Historical variations in mutation rate in an epidemic pathogen, Yersinia pestis. Proc. Natl. Acad. Sci. USA *110*, 577–582. https://doi.org/10.1073/pnas.1205750110.

45. Seguin-Orlando, A., Costedoat, C., Der Sarkissian, C., Tzortzis, S., Kamel, C., Telmon, N., Dalén, L., Thèves, C., Signoli, M., and Orlando, L. (2021). No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. iScience *24*, 102383. https://doi.org/10.1016/j.isci.2021.102383.

46. Cruz-Dávalos, D.I., Llamas, B., Gaunitz, C., Fages, A., Gamba, C., Soubrier, J., Librado, P., Seguin-Orlando, A., Pruvost, M., Alfarhan, A.H., et al. (2017). Experimental conditions improving in-solution target enrichment for ancient DNA. Mol Ecol Res *17*, 508–522. https://doi.org/10.1111/1755-0998.12595.

47. Ziesemer, K.A., Ramos-Madrigal, J., Mann, A.E., Brandt, B.W., Sankaranarayanan, K., Ozga, A.T., Hoogland, M., Hofman, C.A., Salazar-Garcia, D.C., Frohlich, B., et al. (2019). The efficacy of whole human genome capture on ancient dental calculus and dentin. Am. J. Phys. Anthropol. *168*, 496–509. https://doi.org/10.1002/ajpa.23763.

48. Suchan, T., Kusliy, M.A., Khan, N., Chauvey, L., Tonasso-Calvière, L., Schiavinato, S., Southon, J., Keller, M., Kitagawa, K., Krause, J., et al. (2022). Performance and automation of ancient DNA capture with RNA hyRAD probes. Mol Ecol Res *22*, 891–907. https://doi.org/10.1111/1755-0998.13518.

49. Signoli, M. (2022). History of the plague of 1720-1722, in Marseille. Presse Med. *51*, 104138. https://doi.org/10.1016/j.lpm.2022.104138.

50. Gamba, C., Hanghøj, K., Gaunitz, C., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Bradley, D.G., and Orlando, L. (2016). Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. Mol Ecol Res *16*, 459–469. https://doi.org/10.1111/1755-0998.12470.

51. Spyrou, M.A., Keller, M., Tukhbatova, R.I., Scheib, C.L., Nelson, E.A., Andrades Valtueña, A., Neumann, G.U., Walker, D., Alterauge, A., Carty, N., et al. (2019). Phylogeography of the second plague pandemic revealed through analysis of historical Yersinia pestis genomes. Nat. Commun. *10*, 4470. https://doi.org/10.1038/s41467-019-12154-0.

52. Blanco-Miguez, A., Beghini, F., Cumbo, F., McIver, L.J., Thompson, K.N., Zolfo, M., Manghi, P., Dubois, L., Huang, K.D., Maltez Thomas, A., et al. (2022). Extending and improving metagenomic taxonomic profiling with uncharacterized species with MetaPhlAn 4. Preprint at bioRxiv. https://doi.org/10.1101/2022.08.22.504593.

53. Der Sarkissian, C., Ermini, L., Jónsson, H., Alekseev, A.N., Crubezy, E., Shapiro, B., and Orlando, L. (2014). Shotgun microbial

profiling of fossil remains. Mol. Ecol. *23*, 1780–1798. https://doi.org/10.1111/mec.12690.

54. Andrades Valtueña, A., Neumann, G.U., Spyrou, M.A., Musralina, L., Aron, F., Beisenov, A., Belinskiy, A.B., Bos, K.I., Buzhilova, A., Conrad, M., et al. (2022). Stone Age Yersinia pestis genomes shed light on the early evolution, diversity, and ecology of plague. Proc. Natl. Acad. Sci. USA *119*, e2116722119. https://doi.org/10.1073/pnas.2116722119.

55. Spyrou, M.A., Tukhbatova, R.I., Feldman, M., Drath, J., Kacki, S., Beltrán de Heredia, J., Arnold, S., Sitdikov, A.G., Castex, D., Wahl, J., et al. (2016). Historical Y. pestis genomes reveal the European Black Death as the source of ancient and modern plague pandemics. Cell Host Microbe *19*, 874–881. https://doi.org/10.1016/j.chom.2016.05.012.

56. Namouchi, A., Guellil, M., Kersten, O., Hänsch, S., Ottoni, C., Schmid, B.V., Pacciani, E., Quaglia, L., Vermunt, M., Bauer, E.L., et al. (2018). Integrative approach using Yersinia pestis genomes to revisit the historical landscape of plague during the Medieval Period. Proc. Natl. Acad. Sci. USA *115*, E11790–E11797. https://doi.org/10.1073/pnas.1812865115.

57. Bos, K.I., Herbig, A., Sahl, J., Waglechner, N., Fourment, M., Forrest, S.A., Klunk, J., Schuenemann, V.J., Poinar, D., Kuch, M., et al. (2016). Eighteenth century Yersinia pestis genomes reveal the long-term persistence of an historical plague focus. Elife *5*, e12994. https://doi.org/10.7554/eLife.12994.

58. Bos, K.I., Kühnert, D., Herbig, A., Esquivel-Gomez, L.R., Andrades Valtueña, A., Barquera, R., Giffin, K., Kumar Lankapalli, A., Nelson, E.A., Sabin, S., et al. (2019). Paleomicrobiology: diagnosis and evolution of ancient pathogens. Annu. Rev. Microbiol. *73*, 639–666. https://doi.org/10.1146/annurev-micro-090817-062436.

59. Susko, E., and Roger, A.J. (2021). Long branch attraction biases in phylogenetics. Syst. Biol. *70*, 838–843. https://doi.org/10.1093/sysbio/syab001.

60. Eaton, K., Featherstone, L., Duchene, S., Carmichael, A.G., Varlık, N., Golding, G.B., Holmes, E.C., and Poinar, H.N. (2023). Plagued by a cryptic clock: insight and issues from the global phylogeny of Yersinia pestis. Commun. Biol. *6*, 23. https://doi.org/10.1038/s42003-022-04394-6.

61. Wang, H., Avican, K., Fahlgren, A., Erttmann, S.F., Nuss, A.M., Dersch, P., Fallman, M., Edgren, T., and Wolf-Watz, H. (2016). Increased plasmid copy number is essential for Yersinia T3SS function and virulence. Science *353*, 492–495. https://doi.org/10.1126/science.aaf7501.

62. Demeure, C.E., Dussurget, O., Mas Fiol, G., Le Guern, A.-S., Savin, C., and Pizarro-Cerdá, J. (2019). Yersinia pestis and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. Gene Immun. *20*, 357–370. https://doi.org/10.1038/s41435-019-0065-0.

63. Yu, T., Keto-Timonen, R., Jiang, X., Virtanen, J.-P., and Korkeala, H. (2019). Insights into the phylogeny and evolution of cold shock proteins: from enteropathogenic yersinia and escherichia coli to eubacteria. Int. J. Mol. Sci. *20*, 4059. https://doi.org/10.3390/ijms20164059.

64. Groisman, E.A., Hollands, K., Kriner, M.A., Lee, E.-J., Park, S.-Y., and Pontes, M.H. (2013). Bacterial Mg2+ homeostasis, transport, and virulence. Annu. Rev. Genet. *47*, 625–646. https://doi.org/10.1146/annurev-genet-051313-051025.

65. Delepelaire, P. (2019). Bacterial ABC transporters of iron containing compounds. Res. Microbiol. *170*, 345–357. https://doi.org/10.1016/j.resmic.2019.10.008.

66. Fetherston, J.D., Bertolino, V.J., and Perry, R.D. (1999). YbtP and YbtQ: two ABC transporters required for iron uptake in Yersinia pestis. Mol. Microbiol. *32*, 289–299. https://doi.org/10.1046/j.1365-2958.1999.01348.x.

67. Felek, S., Lawrenz, M.B., and Krukonis, E.S. (2008). The Yersinia pestis autotransporter YapC mediates host cell binding, autoaggregation and biofilm formation. Microbiology *154*, 1802–1812. https://doi.org/10.1099/mic.0.2007/010918-0.

68. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193.

69. Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, R., Kircher, M., McCue, M., et al. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat. Protoc. *9*, 1056–1082. https://doi.org/10.1038/nprot.2014.063.

70. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88. https://doi.org/10.1186/s13104-016-1900-2.

71. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

72. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274. https://doi.org/10.1093/molbev/msu300.

73. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

74. Dolle, D., Fages, A., Mata, X., Schiavinato, S., Tonasso-Calvière, L., Chauvey, L., Wagner, S., Der Sarkissian, C., Fromentier, A., Seguin-Orlando, A., and Orlando, L. (2020). CASCADE: a custom-made archiving system for the conservation of ancient DNA experimental data. Front. Ecol. Evol. *8*. https://doi.org/10.3389/fevo.2020.00185.

75. Guellil, M., Kersten, O., Namouchi, A., Luciani, S., Marota, I., Arcini, C.A., Iregren, E., Lindemann, R.A., Warfvinge, G., Bakanidze, L., et al. (2020). A genomic and historical synthesis of plague in 18th century Eurasia. Proc. Natl. Acad. Sci. USA *117*, 28328–28335. https://doi.org/10.1073/pnas.2009677117.

76. Tzortzis, S., and Signoli, M. (2016). Characterization of the funeral groups associated with plague epidemics. Microbiol. Spectr. *4*. https://doi.org/10.1128/microbiolspec.PoH-0011-2015.

77. Neumann, G.U., Andrades Valtueña, A., Fellows Yates, J.A., Stahl, R., and Brandt, G. (2020). Tooth Sampling from the Inner Pulp Chamber for Ancient DNA Extraction V1 (protocols.io.Bakqicvw). https://doi.org/10.17504/protocols.io.bakqicvw.

78. Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M., Seguin-Orlando, A., Owens, I.J., Felkel, S., Bignon-Lau, O., et al. (2018). Ancient genomes revisit the ancestry of domestic and Przewalski's horses. Science *360*, 111–114. https://doi.org/10.1126/science.aao3297.

79. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20130624. https://doi.org/10.1098/rstb.2013.0624.

80. Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., McCrory Constantz, C., Gamba, C., Al-Rasheid, K.A.S., Albizuri, S., et al. (2019). Tracking five millennia of horse management with extensive ancient genome time series. Cell *177*, 1419–1435.e31. https://doi.org/10.1016/j.cell.2019.03.049.

81. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*. pdb.prot5448. https://doi.org/10.1101/pdb.prot5448.

82. Poullet, M., and Orlando, L. (2020). Assessing DNA sequence alignment methods for characterizing ancient genomes and methylomes. Front. Ecol. Evol. *8*.

83. Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T., Prentice, M.B., Sebaihia, M., James, K.D., Churcher, C., Mungall, K.L., et al. (2001). Genome sequence of Yersinia pestis, the causative agent of plague. Nature *413*, 523–527. https://doi.org/10.1038/35097083.

84. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R.

Bioinformatics *30*, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

85. Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. J. Comput. Biol. *7*, 203–214. https://doi.org/10.1089/10665270050081478.

86. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). Genome project data processing subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

87. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

88. Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. *35*, 518–522. https://doi.org/10.1093/molbev/msx281.

89. Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321. https://doi.org/10.1093/sysbio/syq010.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Osteological remain	This study	MAJ6t
Osteological remain	This study	MAJ17t
Osteological remain	This study	MAJ19t
Osteological remain	This study	MAJ33t
Osteological remain	This study	MAJ36t
Osteological remain	This study	MAJ38t
Osteological remain	This study	MAJ38t2
Osteological remain	This study	MAJ38t3
Osteological remain	This study	MAJ38t4
Osteological remain	This study	MAJ38t5
Osteological remain	This study	MAJ38t6
Osteological remain	This study	MAJ43t
Osteological remain	This study	MAJ44t
Osteological remain	This study	MAJ46t
Osteological remain	This study	MAJ46t2
Osteological remain	This study	MAJ46t3
Osteological remain	This study	MAJ46t4
Osteological remain	This study	MAJ46t5
Osteological remain	This study	MAJ46t6
Osteological remain	This study	MAJ46t7
Osteological remain	This study	MAJ46t8
Osteological remain	This study	MAJ46t9
Osteological remain	This study	MAJ46t11
Osteological remain	This study	MAJ48t
Osteological remain	This study	MAJ55t2
Osteological remain	This study	MAJ70t
Osteological remain	This study	MAJ72t
Osteological remain	This study	MAJ75t
Osteological remain	This study	MAJ75t3
Osteological remain	This study	MAJ75t4
Osteological remain	This study	MAJ75t5
Osteological remain	This study	MAJ75t6
Osteological remain	This study	MAJ75t7
Osteological remain	This study	MAJ75t8
Osteological remain	This study	MAJ75t9
Osteological remain	This study	MAJ86t
Osteological remain	This study	MAJ87t
Osteological remain	This study	MAJ89t
Osteological remain	This study	MAJ104t
Osteological remain	This study	MAJ1001t

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Osteological remain	This study	RAY1002t
Osteological remain	This study	RAY1005t
Osteological remain	This study	RAY1009t
Osteological remain	This study	RAY1015t
Osteological remain	This study	RAY1015t2
Osteological remain	This study	RAY1015t3
Osteological remain	This study	RAY1028t
Osteological remain	This study	RAY1028t2
Osteological remain	This study	RAY3009t
Osteological remain	This study	RAY3010t
Osteological remain	This study	RAY3015t
Osteological remain	This study	RAY3023t
Osteological remain	This study	RAY3024t
Osteological remain	This study	RAY3027t
Osteological remain	This study	RAY3030t
Osteological remain	This study	RAY3030t2
Osteological remain	This study	LAR8t
Osteological remain	This study	LAR9t
Osteological remain	This study	LAR11t
Osteological remain	This study	LAR14t
Osteological remain	This study	LAR23t
Osteological remain	This study	LAR24t
Osteological remain	This study	LAR26t
Osteological remain	This study	LAR27t
Osteological remain	This study	LAR27t2
Osteological remain	This study	LAR27t4
Osteological remain	This study	LAR30t
Osteological remain	This study	LAR31t
Osteological remain	This study	SQ7044t
Osteological remain	This study	SQ8504t
Osteological remain	This study	SQ8579t
Osteological remain	This study	SQ8833t
Osteological remain	This study	SQ8844t
Osteological remain	This study	SQ8950t
Osteological remain	This study	SQ8519t
Osteological remain	This study	SQ8637t
Osteological remain	This study	SQ8670t
Osteological remain	This study	SQ8772t
Osteological remain	This study	SQ8893t
Osteological remain	This study	SQ8918t
Osteological remain	This study	SQ8962t
Osteological remain	This study	SQ8723t
Osteological remain	This study	SQ8723t2
Osteological remain	This study	SQ8723t4
Osteological remain	This study	AMIENS109t

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Osteological remain	This study	AMIENS121t
Osteological remain	This study	AMIENS122t
Osteological remain	This study	AMIENS133t
Osteological remain	This study	AMIENS134t
Osteological remain	This study	AMIENS146t
Osteological remain	This study	AMIENS148t
Osteological remain	This study	AMIENS177t
Osteological remain	This study	AMIENS183t
Osteological remain	This study	AMIENS209t
Osteological remain	This study	AMIENS222t
Osteological remain	This study	AMIENS255t
Osteological remain	This study	AMIENS260t
Osteological remain	This study	AMIENS266t
Osteological remain	This study	AMIENS268t
Osteological remain	This study	AMIENS268t2
Osteological remain	This study	AMIENS268t4
Osteological remain	This study	AMIENS32t
Osteological remain	This study	AMIENS41t
Osteological remain	This study	AMIENS46t
Osteological remain	This study	AMIENS54t
Osteological remain	This study	AMIENS55t
Osteological remain	This study	AMIENS63t
Osteological remain	This study	AMIENS69t
Osteological remain	This study	AMIENS74t
Osteological remain	This study	AMIENS75t
Osteological remain	This study	AMIENS78t
Osteological remain	This study	AMIENS81t
Osteological remain	This study	AMIENS82t
Osteological remain	This study	AMIENS85t
Osteological remain	This study	AMIENS86t
Osteological remain	This study	AMIENS91t
Osteological remain	This study	AMIENS93t
Osteological remain	This study	AMIENS94t
Osteological remain	This study	AMIENS97t
Osteological remain	This study	AMIENS99t
Deposited data		
Sequencing dataset	This study	ENA: PRJEB59233
Chemicals, peptides, and recombinant proteins		
N-Lauroylsarcosine solution 30% 500ml	Dutscher	Cat#348533
Proteinase K 10MG	Thermo Fisher Scientific	Cat#AM2542
H2O, Molecular Biology Grade, Fisher BioReagents	Thermo Fisher Scientific	Cat#10977015
Tween 20 100ML	Thermo Fisher Scientific	Cat#13464259
Ethanol, Absolute, Mol Biology Grade	Thermo Fisher Scientific	Cat#16606002
USER Enzyme	New England Biolabs	Cat#M5505L
NEBNext End Repair Module	New England Biolabs	Cat#E6050L

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bst DNA Polymerase	New England Biolabs	Cat#M0275L
NEBNext Quick Ligation Module	New England Biolabs	Cat#E6056L
BSA Molecular Biology Grade	New England Biolabs	Cat#B9000S
ACCUPRIME PFX DNA POLYMERASE 100mL	Thermo Fisher Scientific	Cat#10472482
Agencourt AMPure XP - 60ml	Beckman Coulter	Cat#A63881
Buffer PE	QIAGEN	Cat#19065
Buffer PB	QIAGEN	Cat#19066
Buffer EB	QIAGEN	Cat#19086
dNTP Set 100mM 100mL	Thermo Fisher Scientific	Cat#10297018
EDTA 0.5M pH 8.0 Fisher Bioreagents 500ML	Thermo Fisher Scientific	Cat#AM9261
Critical commercial assays		
MinElute PCR Purification kit	QIAGEN	Cat#28006
Tapestation screenTape D1000 HS	Agilent	Cat#5067-5584
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat#Q32854
MiniSeq High Output Reagent Kit (150-cycles)	Illumina	Cat#FC-420-1002
myBaits Custom DNA-Seq	Arbor Biosciences	Cat#300596.v5
Deposited data		
Raw and analyzed data	This study	ENA: PRJEB59233
Software and algorithms		
mapDamage2	Jónsson et al.,[68]	https://ginolhac.github.io/mapDamage
PALEOMIX	Schubert et al.[69]	https://github.com/MikkelSchubert/paleomix
AdapterRemoval	Schubert et al.[70]	https://github.com/MikkelSchubert/adapterremoval
Samtools	Li and Durbin,[71]	https://github.com/samtools/samtools
snpToolkit v2.3.0	Namouchi et al.[54]	https://github.com/Amine-Namouchi/snpToolkit
IQ-Tree v1.6.12	Nguyen et al.[72]	https://github.com/iqtree/iqtree1
Bowtie2	Langmead and Salzberg,[73]	https://github.com/BenLangmead/bowtie2
BWA	Li and Durbin,[71]	https://github.com/lh3/bwa
seqtk	-	https://github.com/lh3/seqtk
MetaPhlAn 4	Blanco-Miguez et al.[52]	https://github.com/biobakery/MetaPhlAn

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and materials may be directed to lead contact, Prof. Ludovic Orlando (ludovic.orlando@univ-tlse3.fr).

Materials availability

This study did not generate new unique reagents or materials.

Data and code availability

- Sequencing data can be found on the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) and are publicly available as of the date of publication under project accession number ENA: PRJEB59233.
- This paper does not report original code.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Samples and archaeological sites

The 120 teeth analyzed in this study originate from 89 individual skeletons buried in five archaeological sites (Figure 1, Table S1; see Data S1 for further description of the archaeological contexts). Prior to destructive sub-sampling, precise virtual 3D models were constructed for the majority of teeth investigated (94/120 teeth), using the ArtecMicro© surface scanner (Artec 3D, 1 mm voxel size). Samples and laboratory procedures were registered in the Laboratory Information Management System CASCADE.[74]

The plague cemetery of Lariey Puy-Saint-Pierre dating to 1629-1630 CE is located in the French Alps, and was excavated in 2002 CE (Minimal number of individuals, MNI = 34). Two out of 10 individuals (LAR8 and LAR11) investigated in this study were previously sequenced by Seguin-Orlando and colleagues in 2021,[45] who demonstrated that they were infected with a *Y. pestis* strain genetically close to those infecting 1636 CE individuals from San Procolo a Naturno, Italy (1636 CE),[75] in the context of the Thirty Year's War.

The excavations at the Hôtel-Dieu Saint-Jean-Baptiste of Amiens, France, were carried out in 2017 CE by the *Service archéologie préventive d'Amiens Métropole*, under the supervision of Richard Jonvel, who provided access to the city's ancient plague cemetery (MNI = 259). The 34 individuals studied here come from multiple graves dating to 1630-1670 CE, when Amiens experienced notable military and epidemic turmoil.

The site of the Maladrerie Saint-Lazare (Beauvais, France) was excavated between 2002 and 2014 CE by the *Service archéologique municipal de Beauvais* (MNI = 490). The discovery of a plague cemetery established since 1623 CE allowed the analysis of 14 individuals coming from the mass grave labelled 8559.

The parish cemetery of La Major cathedral (La Major), excavated in 2008 CE[76] was created at the beginning of September 1720 CE during the last great wave of plague in Marseille (1720-1722 CE). A total of 19 individuals were DNA tested, but showed poor macroscopic preservation, consistent with that observed across the many other individuals buried at the same site (MNI = 106).

The site of Les Rayettes (Capucins de Ferrières trenches, Martigues, France, MNI = 208), was excavated in 2002 CE. It is located less than 40 kilometers away from La Major. The 12 individuals investigated here were found in multiple mass graves dated to 1720-1721 CE and characteristic of the epidemic ongoing locally at the time.

METHOD DETAILS

Ancient DNA extraction

All molecular work, from dental pulp sampling to library construction and PCR mix preparation, was carried out in the ancient DNA facilities of the CAGT (CNRS UMR 5288/Université Paul Sabatier), following strict experimental standards to avoid and monitor contamination, including use of disposable personal protection equipment, positive air pressure, bleach/UV surface and instrument decontamination, and negative control blanks. PCR amplification, purification, quantitation as well as DNA capture and sequencing were carried out in post-PCR laboratories that are located in a building physically isolated from the ancient DNA facilities.

Between 15 and 80 mg of dental pulp were collected for each of the 120 teeth analyzed (incisors, canines, premolars and molars), following the method developed by Neumann and colleagues,[77] which is routinely used at CAGT.[45] The extraction was performed following the protocol described by Seguin-Orlando and colleagues,[45] which corresponds to a modified version of the Y2 protocol described by Gamba and colleagues.[50] The standard protocol makes use of 963 µL of lysis buffer (0.45M EDTA, 0.25 mg/mL proteinase K and 0.5% sodium lauroyl sarcosinate) to carry out a first pre-digestion step of 1 hour at 37°C. The supernatant is then collected following centrifugation at 12,000 rotation per minute (rpm) for 30 seconds and a 963 µL volume of a fresh, pre-heated, lysis buffer is added to proceed to full digestion of the remaining pellets at 42°C overnight. After centrifugation at 12,000 rpm for 30 seconds, 200 µL of the lysis buffer is purified through QIAgen© Minelute columns and concentrated in a final elution buffer consisting of 23 µL EB

supplemented with 0.05 % Tween 20 (later referred to as EB-Tween buffer). A total of 22.8 μL purified extract was treated with 7 μL of USER® enzyme (NEB®) for 3 hours at 37°C, following,[78] to reduce spurious nucleotide mis-incorporation at those deaminated cytosine residues formed by post-mortem degradation. The pre-digestion supernatant was also purified and processed in parallel to the overnight digestion fraction, to contrast the DNA content following shotgun DNA sequencing and/or DNA capture.

The standard extraction protocol was modified to assess the kinetics of DNA release in the lysis buffer. This was achieved through two main experiments, each focusing on two different teeth from four individuals that were previously tested positive for *Y. pestis* DNA (AMIENS268, tooth 2 (t2) and t4; LAR27 t2 and t4; SQ8723 t2 and t4, and; MAJ46 t9 and t11). The first experiment followed the standard extraction protocol, with the exception that the whole supernatant was collected every 10 min and replaced by 963 μL of a fresh lysis buffer. Supernatant collection was carried every 10 min until 1 hour of pre-digestion was completed. The second experiment followed the same procedure as the first, except that the supernatants were collected directly after suspending the dental pulp powder in the lysis buffer (0 min), and 0.5, 1, 2.5, 5, 7.5 and 10 min after. Each collected supernatant was further processed following the methodology described above to obtain 29.8 μL of USER-treated DNA extract.

Library construction and PCR amplification

Triple-indexed double-stranded DNA libraries compatible with Illumina sequencing were prepared following the protocol originally described by Rohland et al., 2015,[79] as modified by Fages and colleagues.[80] This protocol includes a unique combination of two barcodes, each located 3′ of the P5 and P7 library adapters. Due to their location immediately downstream of the sequencing primers, these "internal" barcodes form the first seven nucleotide positions within each sequencing read. Each library was amplified in eight parallel PCRs for eight cycles in the same conditions as described by Seguin-Orlando and colleagues[45]: in a total reaction volume of 25 μL, using 0.4 μL of AccuPrime™ Pfx DNA polymerase (one unit), 2.5 μL 10X Accuprime™ Pfx reaction mix, 2 μL of DNA library, 1 μL of BSA (20 mg/mL), 17.9 μL of H2O, 0.2 μL of inPE1 primer (25 μM) and 1 μL indexing primer (5 μM) containing a unique 6-bp "external" barcode. The latter corresponds to a third barcode located within the PCR primer.[81] The same external index was used in each of the eight parallel amplifications carried out on the same original library. Libraries were then amplified through a second additional PCR round to obtain sufficient DNA mass as input for capture enrichment with a custom myBaits® hybridization capture kit (Daicel Arbor Biosciences). To achieve this, two amplification products for a given library were merged and purified together using Agencourt Ampure XP beads (1:1.2 or 1:1.4 as DNA:beads ratio), and eluted in 11 μL of EB-Tween buffer. This provided four purified amplification products per library, which were further split and processed through two parallel PCR amplifications, using 1 μL of the eluate for each PCR reaction, for 10 cycles to obtain sufficient DNA material for DNA capture. These amplifications formed the second PCR round and were carried out in 25 μL reaction volumes, following the same conditions as for the first amplification round, except that IS5 and IS6 were used as PCR primers.[82] The resulting eight PCR products were co-purified on a single MinElute column (QIAGEN®), and eluted in 12 μL of EB-Tween buffer. Library concentration and/or size distribution were measured on a Tapestation 4200 instrument using a High sensitivity D1000 ScreenTapeAssay (Agilent technologies) as well as on a Qubit HS dsDNA assay (Invitrogen).

Capture and sequencing

A total of four to eight libraries from different extracts with unique internal and external barcodes were pooled and concentrated together on a MinElute column (QIAGEN®). This resulted in pools of 7 μL, containing between 2-12 μg of DNA, as recommended in the myBaits® High Sensitivity Protocol Version 5.00. DNA capture was performed according to the aforementioned protocol, consisting of two rounds of 16 to 24 hour hybridization with the probes described by Wagner and colleagues.[33] Captured libraries were sequenced using the Paired-End mode on an Illumina MiniSeq instrument (2x81 cycles) at CAGT. A similar procedure was followed for shotgun sequencing, except that only one amplification round of 8-12 PCR cycles was performed on each library, before purification, quantification, pooling and DNA sequencing.

QUANTIFICATION AND STATISTICAL ANALYSIS

Read demultiplexing and collapsing, as well as adapter and poor-quality end trimming were performed using AdapterRemoval2 version 2.3.0,[70] tolerating at most one mismatch in each internal barcode (–barcode-mm-r[12] 1 –minadapteroverlap 3 –mm 5). Read mapping against the human reference genome

(hg19, GRCh37) and the revised Cambridge Reference Sequence (rCRS, Accession Number = NC_012920.1) were carried out using Bowtie2 version 2.3.4.1,[73] and the recommended parameters from Poullet and Orlando, 2020.[82] Read mapping against the *Y. pseutotuberculosis* reference genome (Accession Number NC_006155.1), the *Y. pestis* reference genome (strain CO92, Accession Number = NC_003143.1)[83] and the three plasmids pCD1 (Accession Number NC_003131.1), pMT1 (Accession Number NC_003134.1) and pPCP1 (Accession Number AL109969.1) were performed using BWA version 0.7.17 "backtrack" mode,[71] and the "stringent" parameters defined by Spyrou and colleagues.[51] Read (re-)alignment, PCR duplicate removal, quality filtering and mapDamage profiling[68] were carried out using the automated PALEOMIX pipeline v1.2.13.2.[69] The same procedures were followed using previously published read sequence data (Table S2) to obtain the comparative panel of *Y. pestis* genomes and plasmids used for phylogenetic reconstructions and/or for assessing the presence/absence of specific virulence genes and lineage-specific mutations, which was summarized using snpToolkit v2.3[56] (see below).

Basic alignment statistics for each sample, DNA library and experimental conditions are provided in Table S1. When comparing the respective performance of each experimental condition, *Y. pestis* and human DNA content of each library were estimated for normalized sequencing efforts. This was achieved by considering the fraction of collapsed reads only, and down-sampling the sequencing data 10 times for each library to the number collected for the least sequenced library.

The investigation of pathogen and mitochondrial unique reads released through time was performed on the first set of extraction tests on four teeth, which were digested in lysis buffer during one hour, sampling and replacing the supernatant with fresh lysis buffer every 10 min. Results were normalized by down-sampling the number of reads for each digestion time by the lowest number of reads sequenced across the different digestion times for each sample. As the percentage of unique reads mapping against the pathogen genomes (CO92, pCD1, pMT1 and pPCP1) over digestion time approximatively fit an inverse exponential function, data have been log-linearized. Linear models of the logarithm of the percentage of unique reads mapping against pathogen and mitochondrial DNA (mtDNA) were performed using the R *lm* command. The remaining *Y. pestis* and human mtDNA to be released at a given time could be predicted using *predict* command in R, according to these models.

The proportions of unique reads mapping against the circular chromosome and the different plasmids of the pathogen were calculated for each site to compare the plasmid load between 17th and 18th c. CE strains. These were contrasted to the proportion of probes targeting these different genetic materials. Calculations were based on libraries subjected to a one-hour digestion, as long as at least 500 high-quality alignments against the CO92 genome were found. Inter and intra-century differences in these proportions were compared for 17th and 18th c. CE samples using non-parametric Wilcoxon and Dunn tests in R.

The read-to-reference edit distance distributions (Figure S2) were generated for each of the new genomes characterized in this study from the NM:I field obtained while running the *samtools view* command, considering either the *Y. pestis* (CO92, NC_003143.1), or the *Y. pseudotuberculosis* (IP 32953, NC_006155.1) as a reference. Depth-of-coverage variation was calculated using the PALEOMIX[69] *coverage* command and non-overlapping sliding windows of 1,000 bp for the CO92 chromosome, or 100 bp for the pCD1, pMT1 and pPCP1 plasmids. GC content was estimated for the same windows using seqtk (https://github.com/lh3/seqtk), and visualized together with coverage estimates through circular plots produced using the R *circlize* library.[84] The fraction of the positions that are covered at least once within a given gene was also calculated using the PALEOMIX *depth* command to assess the presence or absence of 235 plague virulence and pathogenicity loci, 190 of which located on the CO92 chromosome, and 45 on the three bacterial plasmids.[43,45,54] Gene coverage plots for virulence genes were generated using the heatmap function from the R ggplot2 library (https://cran.r-project.org/web/packages/ggplot2/index.html). A 49-kb region of CO92 was found absent in most of the phylogenetically-close *Y. pestis* strains from New Churchyard (BED038, BED034, BED028, BED024, BED030), Azov (Azov38), CHE1, Rostov2033, Observance (OBS107, OBS116, OBS137, OBS110, OBS124), and La Major (MAJ75 and MAJ46). However, the coverage of some genes within this region was sporadically greater than zero in a subset of strains, suggesting the possible presence of spurious alignments, despite stringent alignment parameters. This was further confirmed using BlastN[85] against the nucleotide collection (nt), showing best hits against environmental *Yersinia* strains than *Y. pestis*. Those cases are highlighted on the heatmap presented on Figure 7E by crossing the cell presenting the corresponding coverage estimate. We used snpToolkit v2.3[56] and the individual vcf files

generated as described above to annotate the new *Y. pestis* genome and plasmid sequences characterized in this study for synonymous and non-synonymous sequence polymorphisms present in at least 90% of the reads. The minimal depth threshold was set to 3 (-d 3) for bases showing Phred scores above or equal to 30. The samples considered for this analysis were merged by archaeological site using samtools to provide sufficient coverage for the detection of strain-specific SNPs.

Microbial taxonomic profiles were obtained using MetaPhlAn 4[52] on collapsed sequencing unique reads after filtering for those mapping against the human nuclear and mitochondrial genomes following the procedure presented above. To accommodate for variable sequencing efforts across DNA libraries, we random-sampled a number of 100,000 reads. This procedure was repeated 10 times to assess the robustness of the taxonomic assignments, given the limited of reads considered. Reads were mapped against the MetaPhlAn 4 biomarker database using Bowtie2 v.2.3.4.1[73] with default parameters, and further filtered for minimal mapping quality of 25, and PCR duplicates using the *view* and *rmdup* modules from samtools v1.11,[85] respectively. Those species showing relative abundances <1% in a given dataset were defined as "low abundance".

Maximum Likelihood phylogenetic trees were constructed using IQ-TREE v1.6.12[72] and a panel of *Y. pestis* genomes, including 124 modern accessions and 61 second pandemic samples, 12 of which were sequenced here for the first time. Individual base calling was carried out using BCFtools[87] (–ploidy 1), requiring minimum base, mapping and genotype Phred quality scores of 30, and filtering those variants located within 10 bp of indels. The best mutational model was selected based on AICc and BIC values with the Model Finder Plus module (MFP), as part of IQ-TREE. Node support was estimated from a total of 1,000 pseudo-replicates using the ultrafast bootstrap (UFBoot) approximation,[88] optimized by nearest neighbor interchange (NNI), as well as the SH-like approximate likelihood ratio (aLRT) test.[89] The final sequence alignments corresponded, on the first hand, to the entire CO92 chromosome, considering polymorphic sites only, or, on the other hand, were restricted to 29,609 genome-wide variant positions previously characterized to provide robust phylogenetic signal.[43,54] Two sequence alignments were considered, requiring a conservative minimum 3-fold average depth-of-coverage per sample. Accessions for the sequence data underlying the comparative panel of 185 *Yersinia pestis* and 1 *Yersinia pseudotuberculosis* genomes are provided in Table S2.

3) Discussion

Cette étude présente une méthode simple et économique en termes de temps et de ressources pour l'extraction de l'ADN de pathogènes anciens à partir de la pulpe dentaire. Les travaux qu'elle présente ont permis de caractériser la variabilité génétique parmi 12 génomes historiques de la bactérie *Y. pestis* et, ainsi, d'accroître considérablement le nombre de génomes anciens de souches originaires de France. Malgré ces avancées, certaines limites et points importants à développer dans de futures études sont à souligner. La variation du génome des souches anciennes de *Y. pestis* n'a été étudiée qu'à partir d'alignements de séquences par rapport à un seul génome de référence. L'étendue complète des variations structurelles et leurs conséquences possibles sur la transmission, la pathogénicité et la virulence restent donc à caractériser. En outre, la qualité des données générées pour la plupart des génomes anciens présentés est limitée, en raison de la présence d'ADN anciens particulièrement dégradés. Cela restreint l'analyse de l'évolution de ces souches et la précision des dates de divergences. Il est par ailleurs important de rappeler que seul un nombre restreint d'échantillons a été soumis aux tests de prédigestion, or la préservation de l'ADN peut fortement varier d'un individu à l'autre (Parker et al., 2020) et il n'est pas exclu que la dynamique de libération de l'ADN du pathogène ne soit pas toujours la même d'un échantillon à l'autre. Il est donc important de réaliser ces tests sur un panel d'échantillons plus important, provenant de diverses périodes et zones géographiques afin de définir l'impact des conditions de conservation sur notre méthode. Même si nos résultats mettent en évidence un schéma solide pour la libération de l'ADN de *Y. pestis*, les expériences devraient être également étendues à d'autres pathogènes, car les propriétés chimiques des parois cellulaires sont connues pour affecter les niveaux de préservation de l'ADN (comme décrit par exemple chez *Mycobacterium leprae* ; Schuenemann et al., 2013).

Notre méthode a été testée uniquement sur la pulpe dentaire contenue dans les dents des individus infectés et il serait intéressant de tester si cette méthode produit des résultats similaires sur d'autres types de tissus. La préservation de l'ADN des agents infectieux anciens est variable selon les éléments anatomiques étudiés. En effet, elle dépend du tropisme tissulaire, de sa voie de transmission et d'entrée dans le corps ainsi que des symptômes liés à l'infection (Spyrou et al., 2019a). Les registres archéologiques peuvent comprendre une multitude d'échantillons de natures différentes et l'ADN microbien a pu être extrait de la plupart de ces tissus y compris les échantillons provenant de collections de muséums (Raxworthy and Smith, 2021; Straube et al.,

2021; Hawkins et al., 2022). Le succès de la récupération de l'ADN ancien des pathogènes dépend de la prise en compte de certains facteurs liés à l'infection. Parmi ces facteurs, la localisation et la dissémination des microbes au cours de l'infection ou encore la réponse immunitaire de son hôte peuvent jouer un rôle important dans notre aptitude à retrouver les séquences de l'agent pathogène. Il apparaît donc important d'intégrer ces paramètres pour définir la stratégie d'échantillonnage la plus appropriée (Duchêne et al., 2020). Les méthodes d'extraction doivent donc être adaptées aux types de tissus et aux pathogènes de manière spécifique. Par exemple, les agents pathogènes qui provoquent des infections limitées aux tissus mous ne peuvent généralement pas être récupérés à partir du matériel squelettique (Kolling et al., 2012). Pour ces agents pathogènes, la récupération de l'ADN ancien nécessite des échantillons de tissus mous, des restes fécaux, des échantillons de sédiments ou encore des échantillons provenant d'archives médicales ou histologiques (Devault et al., 2014). Dans des cas exceptionnels, les agents pathogènes entériques peuvent se transmettre par le sang, ce qui pourrait expliquer la découverte de *Salmonella enterica* serovar Paratyphi C (responsable de la fièvre entérique bactérienne) dans les dents des victimes d'une épidémie mésoaméricaine du XVIème siècle (Eng et al., 2015; Vågene et al., 2018) ainsi que dans la Norvège du XIIIème siècle (Key et al., 2020; Zhou et al., 2018).

Pour le cas particulier de la peste, il semble difficile de retrouver des traces de son génome dans d'autres parties du squelette que la pulpe dentaire. En effet, les os à l'âge adulte ne sont que très peu vascularisés, il est donc difficile d'y retrouver des traces de l'ADN de pathogènes circulant dans le sang. Des comparaisons ont été réalisées entre l'os pétreux et les dents pour le séquençage de pathogènes anciens, et notamment de *Y. pestis* (Margaryan et al., 2018). Des dents et des os pétreux de 10 individus positifs à la peste ont été séquencés pour comparer le matériel génétique présent dans ces deux éléments anatomiques. Les résultats de cette étude soulignent que l'os pétreux ne permet pas de retrouver de séquences de la peste. Ces résultats indiquent également que les profils métagénomiques provenant des dents et de l'os pétreux sont radicalement différents pour un même individu. Une diminution de la diversité microbienne dans les os pétreux est à souligner, cela signifie que cet os est moins prompt à la colonisation par des microorganismes environnementaux probablement grâce à sa très forte densité. L'os pétreux est l'élément anatomique le plus recherché pour les études paléogénomiques mais il ne semble donc pas approprié pour la recherche de pathogènes circulant dans le sang tels que la bactérie *Y. pestis* (il reste cependant compatible avec la caractérisation d'autres pathogènes, tels que les virus, dont le virus de l'hépatite B ; Kocher et

al., 2021). La recherche de tels organismes doit donc se faire préférentiellement sur les dents et peut être couplée au séquençage de l'os pétreux pour reconstruire le génome de l'hôte et en étudier les origines et caractéristiques, notamment d'un point de vue de l'immunité. Pour finir, rappelons que certaines pathologies peuvent induire des lésions osseuses ; c'est le cas par exemple de la tuberculose (Resnick and Niwayama, 1987). Ces maladies pourraient être de bons candidats pour tester ce nouveau protocole sur des os pour en déterminer la dynamique de libération de l'ADN du pathogène à partir de ce type de tissus.

III) Étude des génomes des victimes de la peste

1) Introduction

Tout au long de l'évolution humaine, la sélection de ces mutations génétiques dans de nombreux gènes est souvent spécifique à certaines populations (Quintana-Murci and Clark, 2013; Karlsson et al., 2014). Au cours de ces dernières années, nombreuses sont les études portées sur l'identification de variants génétiques sous sélection causée par la pression exercée par les pathogènes (Quintana-Murci, 2019). Ces études ont pu révéler la présence de mutations dans les génomes des populations humaines et passées qui ont une influence sur la susceptibilité des individus face à ces maladies. Les implications de telles recherches sont majeures pour la biologie et la santé publique car de nombreuses populations restent exposées à des pathogènes mortels. Il est important de caractériser les spécificités des populations et d'en comprendre l'origine pour pouvoir lutter de manière efficace contre les épidémies.

La communauté scientifique a des débats sur l'étude de l'immunité innée humaine contre les souches historiques de peste, et plus particulièrement celle de la peste noire. Bien que certaines équipes aient tenté de répondre à cette question, l'étude de la sélection de variants génétiques qui confèrent une immunité ou une susceptibilité accrue face à la peste reste à approfondir en raison du manque de données de qualité et de la petite taille des échantillons étudiés. Pour résoudre cette problématique, il est important de continuer à séquencer les génomes humains dans les contextes d'épidémies de peste afin d'augmenter la quantité de données et la résolution temporelle des échantillons étudiés. Dans ce but, les génomes des individus provenant de différents sites ont été séquencés à partir de l'ADN extrait de l'os pétreux. Cette étude vise à caractériser l'état des populations touchées lors des épidémies de peste des XVIIème et XVIIIème

siècles et à rechercher de nouveaux candidats qui pourraient être impliqués dans l'immunité innée face à cette maladie. Les résultats préliminaires de l'analyse de ces génomes sont présentés dans les parties suivantes. Au cours des derniers mois de ce projet de thèse, l'objectif sera d'approfondir les analyses concernant l'immunité innée des individus victimes de la peste au cours de ces crises, et de les présenter sous la forme d'un article scientifique. Ces résultats seront présentés au cours de la soutenance de thèse et intégrés au manuscrit à l'issue de cette dernière.

2) Analyse des génomes humains anciens

Les os pétreux analysés au cours de cette étude proviennent de 132 individus retrouvés dans les 6 sites archéologiques décrits précédemment (Table S4). Les données provenant de la capture du génome de *Y. pestis* à partir des dents provenant de 93 individus ont également été analysées (Table S1 ; Clavel et al., 2023). Les os pétreux et les dents étudiés au cours de ce travail représentent au total 149 individus. Entre 40 et 210 mg de poudre ont été générés à partir des os pétreux. Les expérimentations de biologie moléculaire, allant de l'échantillonnage de l'os à la construction des banques ADN et à la préparation du mix PCR, ont été effectuées dans les laboratoires d'ADN ancien du CAGT (CNRS UMR 5288/Université Paul Sabatier). Les manipulations ont été réalisées en suivant des normes expérimentales strictes pour éviter, réduire et contrôler la contamination, y compris l'utilisation d'équipements de protection individuelle jetables, la pression d'air positive, la décontamination des surfaces et des instruments à l'eau de Javel/UV, ainsi que des blancs de contrôle négatifs. L'amplification, la purification et la quantification des banques d'ADN ainsi que leur séquençage à faible profondeur ont été effectués dans des laboratoires post-PCR situés dans un bâtiment physiquement isolé des laboratoires d'ADN ancien. Une première série de séquençage a été réalisée au sein du CAGT pour évaluer les taux d'ADN endogènes présents dans nos banques ADN. Dans un second temps, les échantillons qui présentaient les meilleurs taux d'ADN endogène ont été séquencés en profondeur sur un instrument Illumina NovaSeq 6000, dans les laboratoires SciLifeLab (Science for Life Laboratory, Stockholm, Suède).

a) Procédure expérimentale

(i) Extraction de l'ADN

L'extraction a été réalisée selon le protocole décrit par Seguin-Orlando et al. (Seguin-Orlando et al., 2021c), qui correspond à une version modifiée du protocole Y2 décrit par Gamba et ses collègues (Gamba et al., 2016). Une première étape de prédigestion de 1 heure à 37 °C est réalisée sur la poudre provenant des os pétreux dans 963 µL de tampon de lyse (EDTA 0,45M, protéinase K 0,25 mg/mL et lauroyl sarcosinate de sodium 0,5 %). Le surnageant est ensuite recueilli après centrifugation à 12 000 rotations par minute (rpm) pendant 30 secondes. Une digestion complète des culots est ensuite réalisée durant une nuit à 42 °C dans un volume de 963 µL d'un tampon de lyse frais et préchauffé. Après centrifugation à 12 000 rpm pendant 30 secondes, 200 µL du tampon de lyse ont été purifiés à travers les colonnes QIAgen© MinElute et concentrés dans un tampon d'élution final composé de 23 µL d'EB complété par 0,05 % de Tween 20 (appelé par la suite tampon EB-Tween). Un total de 22,8 µL d'extrait purifié a été traité avec 7 µL d'enzyme USER® (NEB®) pendant 3 heures à 37 °C (Gaunitz et al., 2018), pour réduire les erreurs d'incorporation de nucléotides au niveau des cytosines désaminés formés par les dégradations post-mortem.

(ii) Construction de banques ADN et amplification par PCR

Des banques d'ADN « double brin » contenant trois indexes ont été préparées selon un protocole (Rohland et al., 2015) modifié par Fages et ses collaborateurs (Fages et al., 2019). Ce protocole fait intervenir une combinaison unique de deux indexes internes P5 et P7 compatibles avec le séquençage Illumina, chacun situé en 3' des paires de séquences anciennes. En raison de leur emplacement immédiatement en aval des amorces de séquençage, ces codes-barres "internes" forment les sept premières positions nucléotidiques de chaque séquence. Chaque banque a été amplifiée par PCR pendant 12 cycles dans les mêmes conditions que celles décrites par Seguin-Orlando et ses collègues (Seguin-Orlando et al., 2021b). Un mix de PCR de 25 µL a été réalisé et contient 0.4 µL of AccuPrime™ Pfx DNA polymerase (une unité), 2.5 µL de mix de réaction 10X Accuprime™ Pfx, 3 µL de banque d'ADN, 1 µL de BSA (20 mg/mL), 16.9 µL de H2O, 0.2 µL d'amorces inPE1 (25 µM) et 1 µL d'amorces (5 µM) contenant un code-barres "externe" unique de 6 pb. Ces dernières correspondent à un troisième code-barres situé à l'intérieur de l'amorce PCR. Les banques amplifiées ont ensuite été purifiées à l'aide de billes Agencourt Ampure XP (1:1.2 ou 1:1.4 comme rapport ADN/billes), et éluées dans 20 µL de tampon EB-Tween. La concentration et la distribution de la taille des banques ont été

mesurées sur un instrument Tapestation 4200 à l'aide d'une High sensitivity D1000 ScreenTapeAssay (Agilent technologies) (Agilent technologies) ainsi que sur un Qubit HS dsDNA (Invitrogen). Les banques ont ensuite été séquencées sur un séquenceur Illumina MiniSeq en mode Paired-End (2x81 cycles) au CAGT pour en évaluer le taux d'ADN endogène. Les banques qui présentaient un taux d'ADN endogène supérieur à 11,15 % ont par la suite été amplifiées à nouveaux à hauteur de 2 amplifications indépendantes pour chacune de ces banques. Ces dernières ont par la suite été séquencées sur un NovaSeq S4 (2x150 bp reads) dans les laboratoires SciLifeLab (Stockholm, Suède).

b) Analyse bioinformatique

(i) Traitement et alignement des séquences

Les séquences ADN obtenues suite au séquençage ont été démultiplexées sur la base des séquences d'indexes internes de 7 pb présentes au début des paires de séquences, en utilisant une distance d'édition maximale de 1 mésappariement pour chaque index. Les séquences démultiplexées ont ensuite été éliminées pour les séquences d'adaptateurs (-mm 5) et les extrémités de mauvaise qualité (scores de qualité Phred ≤ 2), et regroupées en séquences uniques à l'aide de AdapterRemoval2. Les paires de séquences regroupées, tronquées regroupées et tronquées non regroupées ont été alignées avec le génome humain de référence (GRCh37, hg19) et le génome mitochondrial de référence (rCRS) à l'aide de Bowtie2 (version 2.4.1 ; Langmead and Salzberg, 2012), en suivant les paramètres recommandés dans des études antérieures (Poullet and Orlando, 2020). Les séquences alignées de taille inférieure à 25 pb et les alignements présentant une qualité strictement inférieure à 30 ont été écartées. Les doublons PCR ont été supprimés à l'aide de MarkDuplicates de Picard Tools (version 2.18.0, http://broadinstitute.github.io/picard/). Le réalignement local autour des indels a été effectué à l'aide de GATK (version 3.8.1 ; McKenna et al., 2010). Le regroupement, l'élagage et l'alignement des séquences, ainsi que l'élimination des doublons de PCR et le réalignement local ont été effectués à l'aide du pipeline de calcul automatisé Paleomix (version 1.2.13.2 ; Schubert et al., 2014). Les statistiques de séquençage, y compris le nombre de séquences, les taux en ADN endogène, la clonalité, ainsi que la longueur et la couverture moyennes des séquences, sont fournies dans la Table S4.

(ii) Dégradations post-mortem de l’ADN

Le pipeline Paleomix a produit automatiquement des profils mapDamage2 (Jónsson et al., 2013), qui permettent d'évaluer les dommages post mortem causés à l'ADN à l’aide de l'analyse des erreurs d'incorporation de nucléotides et de la composition des bases dans les séquences et dans leurs régions génomiques flanquantes. Les profils d'erreurs d'incorporation de nucléotides pour chaque banque d'ADN individuelle étaient similaires et conformes aux attentes après un traitement enzymatique USER partiel. Des exemples illustrant un excès limité de transitions C→T au début de l'alignement, reflété par un excès équivalent de transitions G→A à la fin de l'alignement, sont présentés dans la Figure 6. La composition des bases dans les séquences est uniforme, tandis que la position génomique précédant les alignements est significativement enrichie en C.

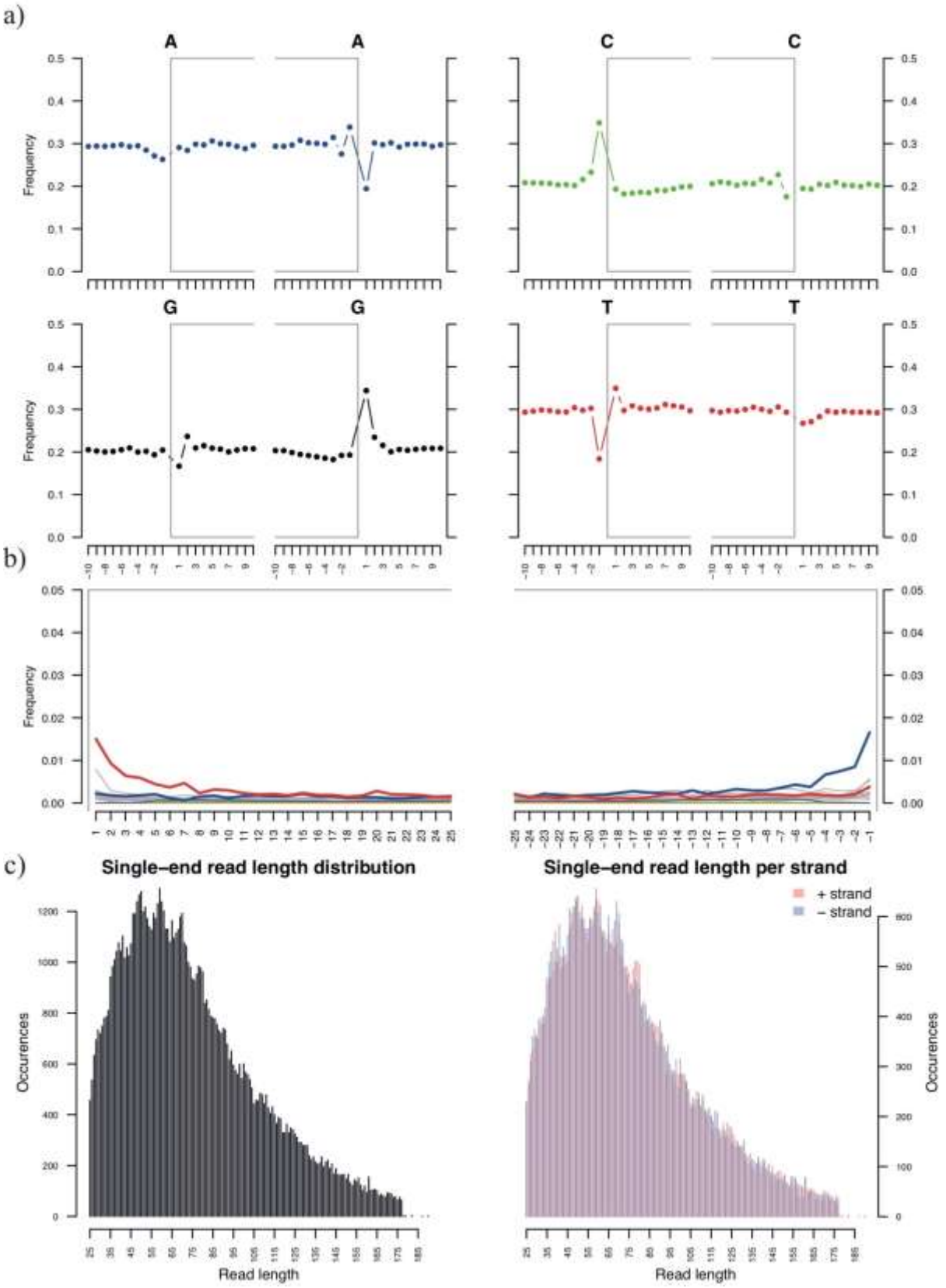


Figure 6 : (a) Profils de fréquence des bases nucléotidiques selon la position au sein des séquences d'ADN ancien, (b) de fréquence des substitutions de nucléotides aux extrémités 5'

et 3’ et (c) des tailles des séquences analysées pour l’échantillon LAR8p produits par MapDamage2 avec les paramètres par défaut à l’exception du paramètre *-y 0.05* pour définir une nouvelle échelle pour le panel (b).

(iii) Estimation des taux d'erreur et de la contamination

Les taux d'erreur ont été calculés à l'aide d'ANGSD (version 0.933-86-g3fefdc4 ; Korneliussen et al., 2013) en suivant la méthodologie entièrement décrite dans des travaux antérieurs (Orlando et al., 2013). Cette approche emploie des alignements comprenant un génome « outgroup » (le génome humain ancestral), un génome de haute qualité (le génome d'un individu Sarde moderne; numéro d'accession : SS6004474 ; Meyer et al., 2012) et chaque génome ancien comme génome test. La différence relative entre la longueur de la branche menant au génome ancien et le génome moderne fournit une mesure directe des taux d'erreur relatifs dans le génome ancien, à la fois globalement et pour chaque classe de substitution individuelle (Table S4). Les taux d’erreurs globaux de chaque génome individuel avec une couverture moyenne de supérieure à 1X (N=73) et une couverture partielle (0.16-0.95X ; N=6) se situent en moyenne entre 0,04 % et 0,08 % substitutions par base pour les données issues du séquençage des os pétreux. Ces taux sont inférieurs aux taux d'erreur présentés d’autres études comparatives (Lazaridis et al., 2014a; Broushaki et al., 2016; Hofmanová et al., 2016; Scheib et al., 2019; Brunel et al., 2020). Cependant, certains génomes issus du séquençage à faible profondeur des os pétreux (N=53 sur un total 132) présentent des taux d’erreurs plus importants mais correspondent systématiquement aux échantillons avec les couvertures moyennes du génome les plus faibles (<0.01X). Les échantillons qui proviennent du séquençage des dents ont également des taux d’erreurs plus importants dus aux faibles couvertures de génomes qu’ils présentent (N=93). Ces échantillons qui présentent des couvertures moyennes de génomes très faibles n’ont pas été inclus dans les analyses suivantes.

Les taux de contamination basés sur les données mitochondriales ont été estimés à l'aide de Schmutzi (Renaud et al., 2015). Les estimations de la contamination ont été évaluées dans un intervalle de confiance de 1 % à 2 % pour la plupart des individus présentant une couverture moyenne supérieure à 0.16X (N=79, Table S4) à l’exception de 8 échantillons avec des taux de contamination plus importants allant de 3 % à 24 % (AMIENS46p, AMIENS55p, AMIENS134p, LAR9p, LAR33p, SQ8579p, SQ8772 et SQ8950p). Afin de contrôler la

véracité des taux de contamination, nous avons également estimé la contamination à l'aide de deux méthodes supplémentaires ANGSD et hapCon. Les taux de contamination ont également été estimés pour les hommes à l'aide de mesures de l'hétérozygotie sur des sites polymorphes présents sur le chromosome X, en suivant la méthodologie de Rasmussen et ses collaborateurs (Rasmussen et al., 2011) et en l'implémentant dans ANGSD (Korneliussen et al., 2013) ainsi qu'en utilisant hapCON (Huang and Ringbauer, 2022). Les substitutions de transition et les sites couverts une fois ou plus de 200 fois n'ont pas été pris en compte. Les estimations mesurées à l'aide de ces méthodes sont relativement limitées et ne dépassent pas 0,71 % pour la majorité des échantillons. Cependant, 5 échantillons présentaient des taux plus importants (>10 % ; AMIENS85p, LAR16p, LAR25p, LAR33p, et SQ8579p) et n'ont pas été retenus pour les analyses qui ont suivi. Enfin, pour 3 échantillons (AMIENS55p, AMIENS134p et LAR9p) qui présentaient une forte contamination mitochondriale, les estimations mesurées par ANGSD et hapCon se sont avérées très faibles, cela est certainement dû à une surestimation des taux de contamination par schmutzi et/ou un différentiel important entre le nombre de molécules anciennes d'origine nucléaire et d'origine mitochondriale (Green et al., 2009).

(iv) Détermination du sexe, des haplogroupes des génomes uniparentaux, et calcul de la consanguinité individuelle

Le sexe des individus a été déterminé à la fois sur la base de la couverture des séquences du chromosome X par rapport aux séquences autosomales $(\frac{Nombre\ de\ séquences\ autosome}{Taille\ autosome})/(\frac{Nombre\ de\ séquences\ chromosome\ X}{Taille\ chromosome\ X})$, ainsi qu'en utilisant la méthodologie de Skoglund et ses collaborateurs (Skoglund et al., 2013). Ces approches nous ont permis d'identifier 44 femmes et 65 hommes (Table S4). Le sexe n'a pas pu être déterminé pour 12 échantillons ou reste incertain pour 28 échantillons en raison de la trop faible quantité de données. Les haplogroupes mitochondriaux ont été déterminés en utilisant haplogrep (version v2.4.0 ; Kloss-Brandstätter et al., 2011) avec des seuils minimaux de scores d'alignement de 30 et de qualité de base de 30. Brièvement, les alignements des séquences contre le génome mitochondrial de référence rCRS (numéro d'accession GenBank NC_120920.1) ont été convertis en un fichier VCF à l'aide des modules *mpileup*, *call* (-ploidy 1) et *filter* de bcftools (version 1.8-31-g9ba4024,), en ne tenant pas compte des sites couverts par moins de 5 séquences indépendantes. Le fichier VCF résultant a ensuite été directement traité par haplogrep, en calculant les 100 meilleurs résultats. Les haplotypes du chromosome Y ont été identifiés à l'aide du progiciel statistique Yleaf (Ralf et

al., 2018). La parenté entre les individus a été évaluée à l'aide de lcMLkin sur les données pseudo-haploïdes (Lipatov et al., 2015). Seuls les SNP (Single Nucleotide Position) autosomiques correspondant aux données 1240K (Mathieson et al., 2015 ; dataset v54.1.p1_1240K disponible à ce lien : https://reichdata.hms.harvard.edu/pub/datasets/amh_repo/curated_releases/V54/V54.1.p1/SHARE/public.dir/v54.1.p1_1240K_public.tar) ont été prises en compte. Avant d'exécuter lcMLkin, nous avons filtré la matrice VCF résultant du script SNPbam2vcf.py en utilisant VCFtools (version 0.1.17) et les paramètres *-thin* 100 000 et *-maf* 0.05 afin de conditionner les analyses aux sites situés au moins à 100 kb l'un de l'autre et présents à des fréquences alléliques minimales de 5 %. La parenté entre les individus a également été mesurée à l'aide de READ (Kuhn et al., 2018) en utilisant les données pseudo-haploïdes restreintes aux SNP de la base de données 1240K. Comme les seuils de distance par paire fournis par défaut dans READ se sont avérés efficaces lorsque l'on considère au moins quatre individus, les analyses ont été effectuées en regroupant les individus par site archéologique. Seuls deux sites ont révélé des relations de parenté parmi les sites archéologiques étudiés, le site de Lariey et celui de Delos (Table S4). Les résultats générés par READ ont suggéré que les individus DEL26 et DEL34 seraient apparentés au deuxième degré, cependant les valeurs des coefficients de parenté issus de lcMLkin restent faibles (k1=0.009). Un nombre plus important d'individus apparentés au sein du site de Lariey a pu être observé avec 2 apparentements au premier degré (LAR17 et LAR26, LAR11 et LAR28) et 5 au second degré (LAR34 et LAR9, LAR23 et LAR9, LAR 20 et LAR31, LAR20 et LAR24, LAR17 et LAR18). Enfin, afin de déterminer des marques de consanguinité dans les génomes étudiés, nous avons détecté les ROH (Runs of homozygosity) à l'aide de hapROH (Ringbauer et al., 2021). Comme recommandé, nous avons analysé des individus avec plus de 400 000 SNP détectés sur le panel 1240k, et nous avons identifié des ROH (runs of homozygosity) d'une longueur supérieure à 4 cM (Figure 7). Les mesures des ROH pour les différents sites archéologiques étudiés se sont révélées limitées, à l'exception du site de Lariey. Ces résultats sont en accord avec le contexte de ce site, à savoir un petit village Alpin, pour lequel on peut s'attendre à des relations consanguines dues à l'isolement géographique, à la faible taille de la population et à la proximité génétique des individus. Ainsi, d'après les résultats issus de hapROH, un individu serait le fruit d'une relation entre des cousins au premier degré (LAR22p), et 7 individus pourraient provenir d'une relation entre cousins du second degré (LAR23p, LAR30p, LAR34p, LAR15p, LAR8p, LAR12p).

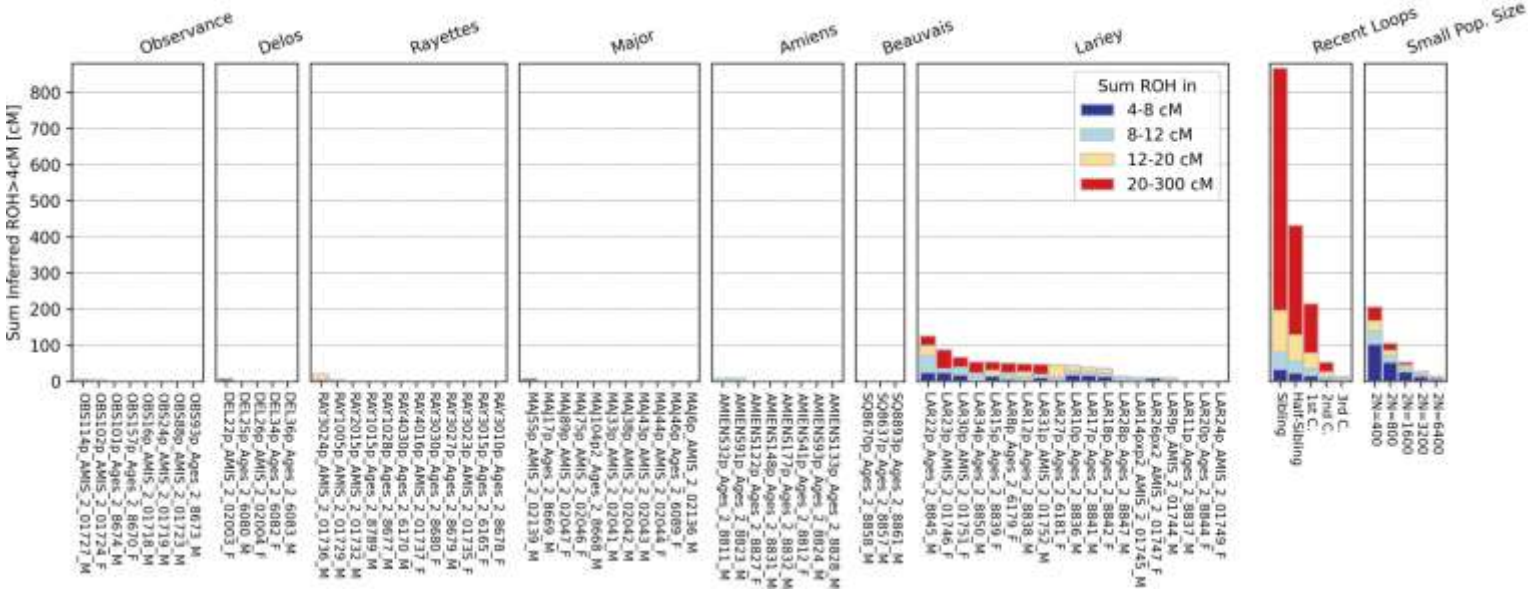


Figure 7 : Mesure des ROH avec une longueur supérieure à 4 cM pour les sites de la Major (Marseille, 1720), l'Observance (Marseille, 1722), Delos (Martigues, 1720-1721), les Rayettes (Martigues, 1720-1721), Amiens (1630-1670), Beauvais (1630-1670) et Lariey (1629-1630).

(v) Analyse en Composantes Principales

L'Analyse en composantes Principales (ACP, Figure 8) a été réalisée à l'aide du panel de référence « *Human Origin* » contenant 592 998 genotypes autosomaux provenant de 9 990 individus anciens et 10 513 individus modernes du monde entier. Les génotypes ont été téléchargés à partir du site internet de David Reich (Mathieson et al., 2015 ; dataset v54.1.p1_1240K disponible à ce lien : https://reichdata.hms.harvard.edu/pub/datasets/amh_repo/curated_releases/V54/V54.1.p1/SHARE/public.dir/v54.1.p1_HO_public.tar). Les génotypes pseudohaploïdes de 63 échantillons de pestiférés ont été générés à l'aide de ANGSD et des génotypes de la base de données de référence « Human Origin » en utilisant le module *-doHaploCall 1* et *-doCounts 1* afin d'échantillonner aléatoirement les bases et de les compter après le passage des filtres de qualités d'alignement et de base (≥ 30). L'Analyse en Composantes Principales a été réalisée sur les échantillons anciens ainsi que 7 359 individus modernes du monde entier (Figure 8a) ou 796 individus modernes européens (Figure 8b ; Lazaridis et al., 2014b) provenant du panel de référence en utilisant *smartPCA* du logiciel EIGENSOFT version 7.2.1 avec les options *lsqproject* et *shrink* pour projeter les individus anciens sur les deux premières composantes principales (PC) obtenues à partir des individus modernes de référence (Patterson et al., 2006).

L'Analyse en Composantes Principales a révélé que les individus anciens étudiés se situent dans l'espace de projection au niveau des populations européennes modernes (Figure 8a). Plus précisément, les individus provenant des sites de Marseille (La Major, Observance), de Martigues (Les Rayettes et Delos) ainsi que de Lariey sont localisés au niveau des populations du Sud-Ouest de l'Europe (Italie, Espagne et France) tandis que les échantillons d'Amiens et Beauvais sont situés entre les individus modernes originaires de France et de Grande-Bretagne (Figure 8b). Ces résultats sont peu surprenants étant donné que ces individus proviennent d'une période relativement récente et suggèrent une certaine continuité génétique dans ces régions.

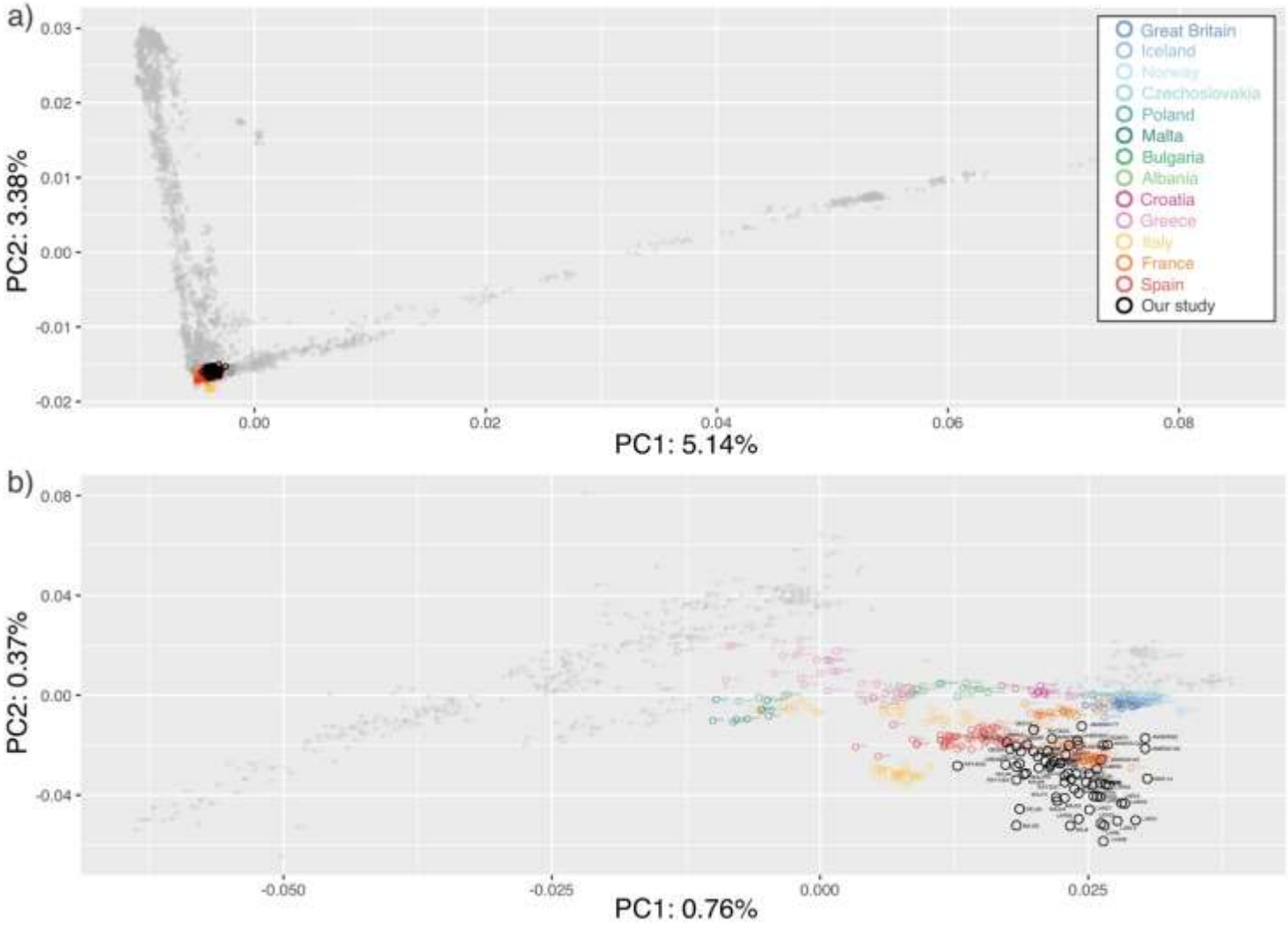


Figure 8 : Analyse en Composantes Principales avec la projection de 63 individus anciens victimes de la peste ainsi que (a) 7 359 individus modernes du monde entier et (b) 796 individus modernes européens du panel de référence « *Human Origin* ».

(vi) Analyse des variants en lien avec l'immunité innée

L'analyse préliminaire des variants en lien avec l'immunité innée des victimes de la peste a été réalisée sur les génotypes diploïdes imputés pour 63 individus. Afin de déceler d'éventuels candidats pouvant avoir eu un effet délétère sur l'immunité vis-à-vis de la peste, nous avons filtré les SNP issus du panel de référence. Les variants situés dans les régions de 356 gènes de l'immunité innée définies par Klunk et ses collaborateurs (Table S5, Klunk et al., 2022) ont été isolés. 7 750 variants alternatifs ont été détectés au niveau de ces gènes pour les génomes des victimes de la pestes. Les fréquences alléliques des variants ont été calculées pour ces individus anciens ainsi que pour le panel de référence issu de la phase 3 du 1000 Genome Project, qui contient plus de 88 millions de variants génomiques chez 2 504 individus modernes provenant de 26 populations différentes (Auton et al., 2015). Une comparaison des fréquences observées dans les populations anciennes et les populations modernes issues du panel de référence a été réalisée par un test du chi2. Aucun de ces variants n'a été retrouvé avec des fréquences statistiquement différentes dans les génomes des pestiférés. Cependant, les génomes étudiés présentent des couvertures partielles de certaines régions du génome et il est fort probable que des variants n'aient pas été détectées au cours de nos analyses. Il est donc nécessaire d'approfondir la profondeur et la couverture de séquençage moyenne pour les génomes des victimes de la peste afin d'accroître la probabilité de retrouver des variants intéressants à étudier par la suite.

3) Discussion

Le séquençage des individus provenant des 6 sites archéologiques datant des XVIIème et XVIIIème siècles a permis de produire 73 génomes de bonne qualité (couvertures moyennes >1 X). Cependant, l'estimation de la contamination par le biais de différentes méthodes a réduit ce nombre à 63 génomes exploitables pour nos analyses. L'étude préliminaire de ces génomes a permis d'identifier l'ascendance génétique de ces individus ainsi que leur consanguinité. De plus, ces analyses ont permis d'identifier 7 750 variants alternatifs présents dans des régions liées à l'immunité. Cependant, elles n'ont pas permis de déceler d'éventuels candidats qui pourraient avoir un rôle délétère dans la résistance face à la peste et leur détection reste limitée par le faible nombre d'individus présents dans notre échantillon ainsi que la couverture moyenne parcellaire de leurs génomes. L'objectif de la fin de mon projet de thèse est

d'approfondir les analyses des génomes humains anciens, notamment en étendant la liste des régions génomiques à étudier en ciblant par exemple les promoteurs et les gènes définis au cours d'études antérieures (Deschamps et al., 2016). La quantité de données produite par le séquençage des os pétreux ne permet pas non plus d'étudier l'effet potentiel des mécanismes épigénétiques sur l'expression des gènes de l'immunité innée. En effet, pour explorer ce type de processus, un couverture moyenne minimale des génomes et/ou des positions connues pour leurs signatures épigénétiques de 20X est nécessaire (Hanghøj et al., 2019). La préservation et donc la qualité de l'ADN ancien des génomes étudiés restent limitées et ne permettent pas l'étude de ces processus. Une approche de capture en solution pour des loci immuns, des promoteurs de gènes de l'immunité innée ou encore des sites de méthylation pourrait permettre d'atteindre des profondeurs de séquençage suffisantes pour approfondir notre problématique à condition que la qualité des séquences anciennes puisse permettre un enrichissement efficace pour ces régions génomiques. Enfin, il serait également pertinent d'avoir un ensemble de génomes contemporains provenant de divers contextes autres que des épidémies. À l'instar de l'étude réalisée par Klunk et ses collaborateurs (Klunk et al., 2022), la comparaison de génomes provenant d'individus décédés dans d'autres contextes pourrait permettre d'étudier des traces de sélection spécifiquement liées aux crises épidémiques, et de comprendre l'effet de la pression exercée par ces pathogènes sur l'immunité des populations humaines.

IV) Perspectives

Cette thèse s'inscrit dans le cadre de l'étude des épidémies du passé et leur impact biologique sur les populations humaines. Plus spécifiquement, elle se concentre sur l'analyse paléogénomique de squelettes humains historiques liés aux épidémies de peste des XVIIème et XVIIIème siècles en France. Les sites étudiés sont associés à la guerre de Trente Ans et à la grande épidémie de peste de Marseille de 1720-1722. Ma recherche s'est concentrée sur deux axes principaux : la caractérisation du génome de *Y. pestis*, l'agent infectieux responsable de ces deux crises, ainsi que celle des individus qui ont succombé à la maladie, en particulier les loci impliqués dans l'immunité innée. Au cours de ces travaux, nous avons pu reconstruire 12 génomes provenant de la bactérie *Y. pestis* dans les différents contextes présentés à l'aide d'un nouveau protocole d'extraction de l'ADN de pathogènes issu de la pulpe dentaire. Également, 63 génomes humains anciens de bonne qualité (couverture moyenne supérieure à 1X) ont été reconstitués à partir d'ossements provenant des victimes de ces crises.

En caractérisant les génomes de *Y. pestis*, cette thèse permet de mieux comprendre la diversité génétique de la bactérie et son évolution au fil du temps. Cela pourrait aider à mieux appréhender les mécanismes de transmission, de virulence et de résistance aux antibiotiques de la peste, et peut être à terme mieux prévoir et contrôler les futures épidémies. En se concentrant sur des sites associés à la guerre de Trente Ans et à la grande épidémie de peste de Marseille, cette thèse permet de relier les données génétiques aux événements historiques et sociaux de l'époque. Cela offre une perspective globale sur les conséquences sanitaires et sociétales des épidémies passées, permettant une meilleure compréhension de leur impact sur les populations humaines et les structures sociales. L'analyse paléogénomique des squelettes humains permet de reconstruire l'histoire biologique des individus qui ont succombé à la peste. En étudiant leur immunité innée, cette recherche offre la possibilité de comprendre comment les populations humaines ont réagi face à ces épidémies passées. Cela peut fournir des informations précieuses sur les mécanismes immunitaires en jeu, les facteurs de susceptibilité ou de résistance, ainsi que les effets à long terme sur la diversité génétique des populations touchées. Cette thèse offre également des perspectives sur l'histoire des maladies infectieuses de manière générale. En appliquant des techniques d'extraction d'ADN similaires à celles développées dans ce projet de thèse, il serait possible de rechercher des pathogènes dans des échantillons provenant d'autres périodes historiques ou de contextes différents. Cela permettrait de retracer l'histoire des maladies infectieuses, d'identifier les agents pathogènes anciens et de comprendre comment ils ont évolué au fil du temps.

En résumé, cette thèse ouvre des perspectives prometteuses pour approfondir notre compréhension des épidémies passées, de la peste en particulier, et de leur impact biologique, épidémiologique, historique et social. Ces connaissances peuvent avoir des implications importantes pour la prévention, le contrôle et la gestion des futures épidémies, ainsi que pour la compréhension de l'évolution humaine et de l'interaction entre les pathogènes et leurs hôtes.

V) Annexes

Document S1 : Figures S1–S5
Données S1 : Description des sites archéologiques.
Table S1 : Informations sur les échantillons et les résultats du séquençage de l'ADN après capture génique.
Table S2 : Informations sur les échantillons du panel comparatif et leurs numéros accessions.

Table S3 : Annotation des séquences avec snpToolkit des génomes caractérisés à Beauvais, Amiens, Les Rayettes et La Major, par rapport à leurs plus proches parents phylogénétiques.
Table S4 : Résultats du séquençage à faible profondeur et à forte profondeur des os pétreux, estimations des taux d'erreur, de la contamination et détermination du sexe, des haplogroupes, des génomes uniparentaux et des liens de parenté.
Table S5 : Panel comparatif des gènes de l'immunité innée défini par Klunk et al. 2022.

VI) Références

Achtman, M., Morelli, G., Zhu, P., Wirth, T., Diehl, I., Kusecek, B., Vogler, A.J., Wagner, D.M., Allender, C.J., Easterday, W.R., Chenal-Francisque, V., Worsham, P., Thomson, N.R., Parkhill, J., Lindler, L.E., Carniel, E., Keim, P., 2004. Microevolution and history of the plague bacillus, Yersinia pestis. Proc Natl Acad Sci U S A 101, 17837–17842. https://doi.org/10.1073/pnas.0408026101

Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A., Carniel, E., 1999. Yersinia pestis, the cause of plague, is a recently emerged clone of Yersinia pseudotuberculosis. Proceedings of the National Academy of Sciences 96, 14043–14048. https://doi.org/10.1073/pnas.96.24.14043

Acuna-Soto, R., Stahle, D.W., Therrell, M.D., Griffin, R.D., Cleaveland, M.K., 2004. When half of the population died: the epidemic of hemorrhagic fevers of 1576 in Mexico. FEMS Microbiol Lett 240, 1–5. https://doi.org/10.1016/j.femsle.2004.09.011

Adler, C.J., Haak, W., Donlon, D., Cooper, A., 2011. Survival and recovery of DNA from ancient teeth and bones. Journal of Archaeological Science 38, 956–964. https://doi.org/10.1016/j.jas.2010.11.010

Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., Malaspinas, A.-S., Margaryan, A., Higham, T., Chivall, D., Lynnerup, N., Harvig, L., Baron, J., Casa, P.D., Dąbrowski, P., Duffy, P.R., Ebel, A.V., Epimakhov, A., Frei, K., Furmanek, M., Gralak, T., Gromov, A., Gronkiewicz, S., Grupe, G., Hajdu, T., Jarysz, R., Khartanovich, V., Khokhlov, A., Kiss, V., Kolář, J., Kriiska, A., Lasak, I., Longhi, C., McGlynn, G., Merkevicius, A., Merkyte, I., Metspalu, M., Mkrtchyan, R., Moiseyev, V., Paja, L., Pálfi, G., Pokutta, D., Pospieszny, Ł., Price, T.D., Saag, L., Sablin, M., Shishlina, N., Smrčka, V., Soenov, V.I., Szeverényi, V., Tóth, G., Trifanova, S.V., Varul, L., Vicze, M., Yepiskoposyan, L., Zhitenev, V., Orlando, L., Sicheritz-Pontén, T., Brunak, S., Nielsen, R., Kristiansen, K., Willerslev, E., 2015. Population genomics of Bronze Age Eurasia. Nature 522, 167–172. https://doi.org/10.1038/nature14507

Andrades Valtueña, A., Mittnik, A., Key, F.M., Haak, W., Allmäe, R., Belinskij, A., Daubaras, M., Feldman, M., Jankauskas, R., Janković, I., Massy, K., Novak, M., Pfrengle, S., Reinhold, S., Šlaus, M., Spyrou, M.A., Szécsényi-Nagy, A., Tõrv, M., Hansen, S., Bos, K.I., Stockhammer, P.W., Herbig, A., Krause, J., 2017. The Stone Age Plague and Its Persistence in Eurasia. Current Biology 27, 3683-3691.e8. https://doi.org/10.1016/j.cub.2017.10.025

Andrades Valtueña, A., Neumann, G.U., Spyrou, M.A., Musralina, L., Aron, F., Beisenov, A., Belinskiy, A.B., Bos, K.I., Buzhilova, A., Conrad, M., Djansugurova, L.B., Dobeš, M., Ernée, M., Fernández-Eraso, J., Frohlich, B., Furmanek, M., Hałuszko, A., Hansen, S., Harney, É., Hiss, A.N., Hübner, A., Key, F.M., Khussainova, E., Kitov, E., Kitova, A.O., Knipper, C., Kühnert, D., Lalueza-Fox, C., Littleton, J., Massy, K., Mittnik, A., Mujika-Alustiza, J.A., Olalde, I., Papac, L., Penske, S., Peška, J., Pinhasi, R., Reich, D., Reinhold, S., Stahl, R., Stäuble, H., Tukhbatova, R.I., Vasilyev, S., Veselovskaya, E., Warinner, C., Stockhammer,

P.W., Haak, W., Krause, J., Herbig, A., 2022. Stone Age Yersinia pestis genomes shed light on the early evolution, diversity, and ecology of plague. Proceedings of the National Academy of Sciences 119, e2116722119. https://doi.org/10.1073/pnas.2116722119
Anisimov, A.P., Lindler, L.E., Pier, G.B., 2004. Intraspecific diversity of Yersinia pestis. Clin Microbiol Rev 17, 434–464. https://doi.org/10.1128/CMR.17.2.434-464.2004
Armelagos, G.J., Goodman, A.H., Jacobs, K.H., 1991. The origins of agriculture: Population growth during a period of declining health. Popul Environ 13, 9–22. https://doi.org/10.1007/BF01256568
Arnold, B., Corbett-Detig, R.B., Hartl, D., Bomblies, K., 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Mol Ecol 22, 3179–3190. https://doi.org/10.1111/mec.12276
Arriaza, B.T., Salo, W., Aufderheide, A.C., Holcomb, T.A., 1995. Pre-Columbian tuberculosis in Northern Chile: Molecular and skeletal evidence. American Journal of Physical Anthropology 98, 37–45. https://doi.org/10.1002/ajpa.1330980104
Aslanidis, C., Ries, S., Fehringer, P., Büchler, C., Klima, H., Schmitz, G., 1996. Genetic and biochemical evidence that CESD and Wolman disease are distinguished by residual lysosomal acid lipase activity. Genomics 33, 85–93. https://doi.org/10.1006/geno.1996.0162
Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B., Gibbs, R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S., Lee, C., Lehrach, H., Mardis, E.R., Marth, G.T., McVean, G.A., Nickerson, D.A., Schmidt, J.P., Sherry, S.T., Wang, J., Wilson, R.K., Gibbs, R.A., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J.G., Zhu, Y., Wang, J., Chang, Y., Feng, Q., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Lan, T., Li, G., Li, J., Li, Y., Liu, S., Liu, X., Lu, Y., Ma, X., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Xu, X., Yin, Y., Zhang, D., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Lander, E.S., Altshuler, D.M., Gabriel, S.B., Gupta, N., Gharani, N., Toji, L.H., Gerry, N.P., Resch, A.M., Flicek, P., Barker, J., Clarke, L., Gil, L., Hunt, S.E., Kelman, G., Kulesha, E., Leinonen, R., McLaren, W.M., Radhakrishnan, R., Roa, A., Smirnov, D., Smith, R.E., Streeter, I., Thormann, A., Toneva, I., Vaughan, B., Zheng-Bradley, X., Bentley, D.R., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Lehrach, H., Sudbrak, R., Albrecht, M.W., Amstislavskiy, V.S., Borodina, T.A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.-L., Mardis, E.R., Wilson, R.K., Fulton, L., Fulton, R., Sherry, S.T., Ananiev, V., Belaia, Z., Beloslyudtsev, D., Bouk, N., Chen, C., Church, D., Cohen, R., Cook, C., Garner, J., Hefferon, T., Kimelman, M., Liu, C., Lopez, J., Meric, P., O'Sullivan, C., Ostapchuk, Y., Phan, L., Ponomarov, S., Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang, H., McVean, G.A., Durbin, R.M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T.M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Schmidt, J.P., Davies, C.J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Campbell, C.L., Kong, Y., Marcketta, A., Gibbs, R.A., Yu, F., Antunes, L., Bainbridge, M., Muzny, D., Sabo, A., Huang, Z., Wang, J., Coin, L.J.M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Y., Li, Z., Lin, H., Liu, B., Luo, R., Shao, H., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Alkan, C., Dal, E., Kahveci, F., Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Fung Leong, W., Stromberg, M., Ward, A.N., Wu, J., Zhang, M., Daly, M.J., DePristo, M.A., Handsaker, R.E., Altshuler, D.M., Banks, E., Bhatia, G., del Angel, G., Gabriel, S.B., Genovese, G., Gupta, N., Li, H., Kashin, S., Lander, E.S., McCarroll, S.A., Nemesh, J.C., Poplin, R.E., Yoon, S.C., Lihm, J., Makarov, V., Clark, A.G., Gottipati, S., Keinan, A., Rodriguez-Flores, J.L., Korbel, J.O., Rausch, T., Fritz, M.H., Stütz, A.M., Flicek, P., Beal, K., Clarke, L., Datta, A., Herrero, J., McLaren, W.M., Ritchie, G.R.S., Smith, R.E., Zerbino, D., Zheng-Bradley, X., Sabeti, P.C., Shlyakhter, I., Schaffner, S.F., Vitti, J., Cooper, D.N., Ball, E.V., Stenson, P.D., Bentley, D.R., Barnes, B., Bauer, M., Keira

Cheetham, R., Cox, A., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J., Shaw, R., Kenny, E.E., Batzer, M.A., Konkel, M.K., Walker, J.A., MacArthur, D.G., Lek, M., Sudbrak, R., Amstislavskiy, V.S., Herwig, R., Mardis, E.R., Ding, L., Koboldt, D.C., Larson, D., Ye, K., Gravel, S., The 1000 Genomes Project Consortium, Corresponding authors, Steering committee, Production group, Baylor College of Medicine, BGI-Shenzhen, Broad Institute of MIT and Harvard, Coriell Institute for Medical Research, European Molecular Biology Laboratory, E.B.I., Illumina, Max Planck Institute for Molecular Genetics, McDonnell Genome Institute at Washington University, US National Institutes of Health, University of Oxford, Wellcome Trust Sanger Institute, Analysis group, Affymetrix, Albert Einstein College of Medicine, Bilkent University, Boston College, Cold Spring Harbor Laboratory, Cornell University, European Molecular Biology Laboratory, Harvard University, Human Gene Mutation Database, Icahn School of Medicine at Mount Sinai, Louisiana State University, Massachusetts General Hospital, McGill University, National Eye Institute, N., 2015. A global reference for human genetic variation. Nature 526, 68–74. https://doi.org/10.1038/nature15393

Ávila-Arcos, M.C., Cappellini, E., Romero-Navarro, J.A., Wales, N., Moreno-Mayar, J.V., Rasmussen, M., Fordyce, S.L., Montiel, R., Vielle-Calzada, J.-P., Willerslev, E., Gilbert, M.T.P., 2011. Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. Sci Rep 1, 74. https://doi.org/10.1038/srep00074

Ávila-Arcos, M.C., Sandoval-Velasco, M., Schroeder, H., Carpenter, M.L., Malaspinas, A.-S., Wales, N., Peñaloza, F., Bustamante, C.D., Gilbert, M.T.P., 2015. Comparative performance of two whole-genome capture methodologies on ancient DNA Illumina libraries. Methods in Ecology and Evolution 6, 725–734. https://doi.org/10.1111/2041-210X.12353

Axelsson, E., Willerslev, E., Gilbert, M.T.P., Nielsen, R., 2008. The Effect of Ancient DNA Damage on Inferences of Demographic Histories. Molecular Biology and Evolution 25, 2181–2187. https://doi.org/10.1093/molbev/msn163

Bacot, A.W., 1915. LXXXI. Further notes on the mechanism of the transmission of plague by fleas. J Hyg (Lond) 14, 774-776.3.

Bamshad, M.J., Mummidi, S., Gonzalez, E., Ahuja, S.S., Dunn, D.M., Watkins, W.S., Wooding, S., Stone, A.C., Jorde, L.B., Weiss, R.B., Ahuja, S.K., 2002. A strong signature of balancing selection in the 5′ cis-regulatory region of CCR5. Proceedings of the National Academy of Sciences 99, 10539–10544. https://doi.org/10.1073/pnas.162046399

Barbieri, R., Signoli, M., Chevé, D., Costedoat, C., Tzortzis, S., Aboudharam, G., Raoult, D., Drancourt, M., 2020. Yersinia pestis: the Natural History of Plague. Clinical Microbiology Reviews 34, e00044-19. https://doi.org/10.1128/CMR.00044-19

Bar-Gal, G.K., Kim, M.J., Klein, A., Shin, D.H., Oh, C.S., Kim, J.W., Kim, T.-H., Kim, S.B., Grant, P.R., Pappo, O., Spigelman, M., Shouval, D., 2012. Tracing hepatitis B virus to the 16th century in a Korean mummy. Hepatology 56, 1671–1680. https://doi.org/10.1002/hep.25852

Barnes, I., Matheus, P., Shapiro, B., Jensen, D., Cooper, A., 2002. Dynamics of Pleistocene Population Extinctions in Beringian Brown Bears. Science 295, 2267–2270. https://doi.org/10.1126/science.1067814

Barrett, R., Kuzawa, C.W., McDade, T., Armelagos, G.J., 1998. EMERGING AND RE-EMERGING INFECTIOUS DISEASES: The Third Epidemiologic Transition. Annual Review of Anthropology 27, 247–271. https://doi.org/10.1146/annurev.anthro.27.1.247

Barton, A.R., Santander, C.G., Skoglund, P., Moltke, I., Reich, D., Mathieson, I., 2023. Insufficient evidence for natural selection associated with the Black Death. https://doi.org/10.1101/2023.03.14.532615

Basler, N., Xenikoudakis, G., Westbury, M.V., Song, L., Sheng, G., Barlow, A., 2017. Reduction of the contaminant fraction of DNA obtained from an ancient giant panda bone.

BMC Res Notes 10, 754. https://doi.org/10.1186/s13104-017-3061-3
Beauvieux, F., Bertrand, R., Costedoat, C., Buti, G., Dutour, O., Écard, P., El Hadj, J., Goury, M., Mellinand, P., Michaudel, B., Paone, F., Riani, A., Signoli, M., Tzortzis, S., 2022. Marseille en temps de peste 1720-1722, Snoeck. ed. Snoeck.
Benedictow, O.J., 2004. The Black Death, 1346-1353: The Complete History. Boydell & Brewer.
Benedictow, O.J., 1987. Morbidity in Historical Plague Epidemics. Population Studies 41, 401–431. https://doi.org/10.1080/0032472031000142976
Bertherat, E., 2019. Plague around the world in 2019/La peste dans le monde en 2019. Weekly Epidemiological Record 94, 289–293.
Bertherat, E., 2016. Plague around the world, 2010–2015. Wkly Epidemiol Rec 91, 89–93.
Binladen, J., Gilbert, M.T.P., Bollback, J.P., Panitz, F., Bendixen, C., Nielsen, R., Willerslev, E., 2007. The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. PLOS ONE 2, e197. https://doi.org/10.1371/journal.pone.0000197
Bodi, K., Perera, A.G., Adams, P.S., Bintzler, D., Dewar, K., Grove, D.S., Kieleczawa, J., Lyons, R.H., Neubert, T.A., Noll, A.C., Singh, S., Steen, R., Zianni, M., 2013. Comparison of Commercially Available Target Enrichment Methods for Next-Generation Sequencing. J Biomol Tech 24, 73–86. https://doi.org/10.7171/jbt.13-2402-002
Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant, J.M., Harris, S.R., Schuenemann, V.J., Campbell, T.J., Majander, K., Wilbur, A.K., Guichon, R.A., Wolfe Steadman, D.L., Cook, D.C., Niemann, S., Behr, M.A., Zumarraga, M., Bastida, R., Huson, D., Nieselt, K., Young, D., Parkhill, J., Buikstra, J.E., Gagneux, S., Stone, A.C., Krause, J., 2014. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature 514, 494–497. https://doi.org/10.1038/nature13591
Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., DeWitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J.D., Herring, D.A., Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature 478, 506–510. https://doi.org/10.1038/nature10549
Briggs, A.W., Good, J.M., Green, R.E., Krause, J., Maricic, T., Stenzel, U., Lalueza-Fox, C., Rudan, P., Brajkovic, D., Kucan, Z., Gusic, I., Schmitz, R., Doronichev, V.B., Golovanova, L.V., de la Rasilla, M., Fortea, J., Rosas, A., Pääbo, S., 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science 325, 318–321. https://doi.org/10.1126/science.1174462
Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., Pääbo, S., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. PNAS 104, 14616–14621. https://doi.org/10.1073/pnas.0704665104
Briggs, A.W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., Pääbo, S., 2010. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res 38, e87. https://doi.org/10.1093/nar/gkp1163
Broushaki, F., Thomas, M.G., Link, V., López, S., van Dorp, L., Kirsanow, K., Hofmanová, Z., Diekmann, Y., Cassidy, L.M., Díez-del-Molino, D., Kousathanas, A., Sell, C., Robson, H.K., Martiniano, R., Blöcher, J., Scheu, A., Kreutzer, S., Bollongino, R., Bobo, D., Davoudi, H., Munoz, O., Currat, M., Abdi, K., Biglari, F., Craig, O.E., Bradley, D.G., Shennan, S., Veeramah, K.R., Mashkour, M., Wegmann, D., Hellenthal, G., Burger, J., 2016. Early Neolithic genomes from the eastern Fertile Crescent. Science 353, 499–503. https://doi.org/10.1126/science.aaf7943
Brown, N., 2019. Choc et échange épidémiologique : Indiens et Espagnols au Mexique (1520-1596), in: Priotti, J.-P. (Ed.), Identités et territoires dans les mondes hispaniques : (XVIe-XXe

siècles), Histoire. Presses universitaires de Rennes, Rennes, pp. 121–148. https://doi.org/10.4000/books.pur.90169
Brunel, S., Bennett, E.A., Cardin, L., Garraud, D., Barrand Emam, H., Beylier, A., Boulestin, B., Chenal, F., Ciesielski, E., Convertini, F., Dedet, B., Desbrosse-Degobertiere, S., Desenne, S., Dubouloz, J., Duday, H., Escalon, G., Fabre, V., Gailledrat, E., Gandelin, M., Gleize, Y., Goepfert, S., Guilaine, J., Hachem, L., Ilett, M., Lambach, F., Maziere, F., Perrin, B., Plouin, S., Pinard, E., Praud, I., Richard, I., Riquier, V., Roure, R., Sendra, B., Thevenet, C., Thiol, S., Vauquelin, E., Vergnaud, L., Grange, T., Geigl, E.-M., Pruvost, M., 2020. Ancient genomes from present-day France unveil 7,000 years of its demographic history. Proceedings of the National Academy of Sciences 117, 12791–12798. https://doi.org/10.1073/pnas.1918034117
Buikstra, J., Roberts, C., 2011. The Global History of Paleopathology: Pioneers and Prospects. Oxford University Press.
Burbano, H.A., Hodges, E., Green, R.E., Briggs, A.W., Krause, J., Meyer, M., Good, J.M., Maricic, T., Johnson, P.L.F., Xuan, Z., Rooks, M., Bhattacharjee, A., Brizuela, L., Albert, F.W., de la Rasilla, M., Fortea, J., Rosas, A., Lachmann, M., Hannon, G.J., Pääbo, S., 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. Science 328, 723–725. https://doi.org/10.1126/science.1188046
Carpenter, M.L., Buenrostro, J.D., Valdiosera, C., Schroeder, H., Allentoft, M.E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S., Nekhrizov, G., Leshtakov, K., Dimitrova, D., Theodossiev, N., Pettener, D., Luiselli, D., Sandoval, K., Moreno-Estrada, A., Li, Y., Wang, J., Gilbert, M.T.P., Willerslev, E., Greenleaf, W.J., Bustamante, C.D., 2013. Pulling out the 1%: Whole-Genome Capture for the Targeted Enrichment of Ancient DNA Sequencing Libraries. The American Journal of Human Genetics 93, 852–864. https://doi.org/10.1016/j.ajhg.2013.10.002
Carrière, C., Courdurié, M., Rebuffat, F., 2008. Marseille ville morte : La peste de 1720. Editions Autres Temps, Gémenos.
Casanova, J.-L., 2015. Human genetic basis of interindividual variability in the course of infection. Proceedings of the National Academy of Sciences 112, E7118–E7127. https://doi.org/10.1073/pnas.1521644112
Casanova, J.-L., Abel, L., 2022. From rare disorders of immunity to common determinants of infection: Following the mechanistic thread. Cell 185, 3086–3103. https://doi.org/10.1016/j.cell.2022.07.004
Casey, J.P., McGettigan, P., Lynam-Lennon, N., McDermott, M., Regan, R., Conroy, J., Bourke, B., O'Sullivan, J., Crushell, E., Lynch, S., Ennis, S., 2012. Identification of a mutation in LARS as a novel cause of infantile hepatopathy. Mol Genet Metab 106, 351–358. https://doi.org/10.1016/j.ymgme.2012.04.017
Castellano, S., Parra, G., Sánchez-Quinto, F.A., Racimo, F., Kuhlwilm, M., Kircher, M., Sawyer, S., Fu, Q., Heinze, A., Nickel, B., Dabney, J., Siebauer, M., White, L., Burbano, H.A., Renaud, G., Stenzel, U., Lalueza-Fox, C., de la Rasilla, M., Rosas, A., Rudan, P., Brajković, D., Kucan, Ž., Gušic, I., Shunkov, M.V., Derevianko, A.P., Viola, B., Meyer, M., Kelso, J., Andrés, A.M., Pääbo, S., 2014. Patterns of coding variation in the complete exomes of three Neandertals. Proceedings of the National Academy of Sciences 111, 6666–6671. https://doi.org/10.1073/pnas.1405138111
Chan, J.Z.-M., Sergeant, M.J., Lee, O.Y.-C., Minnikin, D.E., Besra, G.S., Pap, I., Spigelman, M., Donoghue, H.D., Pallen, M.J., 2013. Metagenomic analysis of tuberculosis in a mummy. N Engl J Med 369, 289–290. https://doi.org/10.1056/NEJMc1302295
Choin, J., Mendoza-Revilla, J., Arauna, L.R., Cuadros-Espinoza, S., Cassar, O., Larena, M., Ko, A.M.-S., Harmant, C., Laurent, R., Verdu, P., Laval, G., Boland, A., Olaso, R., Deleuze, J.-F., Valentin, F., Ko, Y.-C., Jakobsson, M., Gessain, A., Excoffier, L., Stoneking, M., Patin,

E., Quintana-Murci, L., 2021. Genomic insights into population history and biological adaptation in Oceania. Nature 592, 583–589. https://doi.org/10.1038/s41586-021-03236-5
Chouikha, I., Hinnebusch, B.J., 2014. Silencing urease: A key evolutionary step that facilitated the adaptation of Yersinia pestis to the flea-borne transmission route. Proceedings of the National Academy of Sciences 111, 18709–18714. https://doi.org/10.1073/pnas.1413209111
Clark, G.A., Kelley, M.A., Grange, J.M., Hill, M.C., Katzenberg, M.A., Klepinger, L.L., Lovell, N.C., McGrath, J., Micozzi, M.S., Steinbock, R.T., 1987. The Evolution of Mycobacterial Disease in Human Populations: A Reevaluation [and Comments and Reply]. Current Anthropology 28, 45–62. https://doi.org/10.1086/203490
Clavel, P., Louis, L., Sarkissian, C.D., Thèves, C., Gillet, C., Chauvey, L., Tressières, G., Schiavinato, S., Calvière-Tonasso, L., Telmon, N., Clavel, B., Jonvel, R., Tzortzis, S., Bouniol, L., Fémolant, J.-M., Klunk, J., Poinar, H., Signoli, M., Costedoat, C., Spyrou, M.A., Seguin-Orlando, A., Orlando, L., 2023. Improving the extraction of ancient Yersinia pestis genomes from the dental pulp. iScience 26, 106787. https://doi.org/10.1016/j.isci.2023.106787
Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat Genet 45, 1176–1182. https://doi.org/10.1038/ng.2744
Coolen, M.J.L., Overmann, J., 1998. Analysis of Subfossil Molecular Remains of Purple Sulfur Bacteria in a Lake Sediment. Applied and Environmental Microbiology 64, 4513–4521. https://doi.org/10.1128/AEM.64.11.4513-4521.1998
Cooper, A., 1997. Reply to Stoneking: ancient DNA--how do you really know when you have it? Am J Hum Genet 60, 1001–1003.
Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139. https://doi.org/10.1126/science.289.5482.1139b
Crosby, A.W., 1994. Virgin Soil Epidemics as a Factor in the Aboriginal Depopulation of America, in: Germs, Seeds and Animals: Routledge.
Cruz-Dávalos, D.I., Llamas, B., Gaunitz, C., Fages, A., Gamba, C., Soubrier, J., Librado, P., Seguin-Orlando, A., Pruvost, M., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Scheu, A., Beneke, N., Ludwig, A., Cooper, A., Willerslev, E., Orlando, L., 2017. Experimental conditions improving in-solution target enrichment for ancient DNA. Molecular Ecology Resources 17, 508–522. https://doi.org/10.1111/1755-0998.12595
Cui, Y., Yu, C., Yan, Y., Li, D., Li, Yanjun, Jombart, T., Weinert, L.A., Wang, Z., Guo, Z., Xu, L., Zhang, Y., Zheng, H., Qin, N., Xiao, X., Wu, M., Wang, X., Zhou, D., Qi, Z., Du, Z., Wu, H., Yang, X., Cao, H., Wang, H., Wang, Jing, Yao, S., Rakin, A., Li, Yingrui, Falush, D., Balloux, F., Achtman, M., Song, Y., Wang, Jun, Yang, R., 2013. Historical variations in mutation rate in an epidemic pathogen, Yersinia pestis. Proceedings of the National Academy of Sciences 110, 577–582. https://doi.org/10.1073/pnas.1205750110
Cunha, C.B., Cunha, B.A., 2008. Great Plagues of the Past and Remaining Questions. Paleomicrobiology 1–20. https://doi.org/10.1007/978-3-540-75855-6_1
Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., Meyer, M., 2013a. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences 110, 15758–15763. https://doi.org/10.1073/pnas.1314445110
Dabney, J., Meyer, M., Pääbo, S., 2013b. Ancient DNA damage. Cold Spring Harb Perspect Biol 5, a012567. https://doi.org/10.1101/cshperspect.a012567

Damgaard, P. de B., Marchi, N., Rasmussen, S., Peyrot, M., Renaud, G., Korneliussen, T., Moreno-Mayar, J.V., Pedersen, M.W., Goldberg, A., Usmanova, E., Baimukhanov, N., Loman, V., Hedeager, L., Pedersen, A.G., Nielsen, K., Afanasiev, G., Akmatov, K., Aldashev, A., Alpaslan, A., Baimbetov, G., Bazaliiskii, V.I., Beisenov, A., Boldbaatar, B., Boldgiv, B., Dorzhu, C., Ellingvag, S., Erdenebaatar, D., Dajani, R., Dmitriev, E., Evdokimov, V., Frei, K.M., Gromov, A., Goryachev, A., Hakonarson, H., Hegay, T., Khachatryan, Z., Khaskhanov, R., Kitov, E., Kolbina, A., Kubatbek, T., Kukushkin, A., Kukushkin, I., Lau, N., Margaryan, A., Merkyte, I., Mertz, I.V., Mertz, V.K., Mijiddorj, E., Moiyesev, V., Mukhtarova, G., Nurmukhanbetov, B., Orozbekova, Z., Panyushkina, I., Pieta, K., Smrčka, V., Shevnina, I., Logvin, A., Sjögren, K.-G., Štolcová, T., Taravella, A.M., Tashbaeva, K., Tkachev, A., Tulegenov, T., Voyakin, D., Yepiskoposyan, L., Undrakhbold, S., Varfolomeev, V., Weber, A., Wilson Sayres, M.A., Kradin, N., Allentoft, M.E., Orlando, L., Nielsen, R., Sikora, M., Heyer, E., Kristiansen, K., Willerslev, E., 2018. 137 ancient human genomes from across the Eurasian steppes. Nature 557, 369–374. https://doi.org/10.1038/s41586-018-0094-2

Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., Allentoft, M.E., 2015. Improving access to endogenous DNA in ancient bones and teeth. Sci Rep 5, 11184. https://doi.org/10.1038/srep11184

Davey, J.W., Cezard, T., Fuentes-Utrilla, P., Eland, C., Gharbi, K., Blaxter, M.L., 2013. Special features of RAD Sequencing data: implications for genotyping. Mol Ecol 22, 3151–3164. https://doi.org/10.1111/mec.12084

de Barros Damgaard, P., Martiniano, R., Kamm, J., Moreno-Mayar, J.V., Kroonen, G., Peyrot, M., Barjamovic, G., Rasmussen, S., Zacho, C., Baimukhanov, N., Zaibert, V., Merz, V., Biddanda, A., Merz, I., Loman, V., Evdokimov, V., Usmanova, E., Hemphill, B., Seguin-Orlando, A., Yediay, F.E., Ullah, I., Sjögren, K.-G., Iversen, K.H., Choin, J., de la Fuente, C., Ilardo, M., Schroeder, H., Moiseyev, V., Gromov, A., Polyakov, A., Omura, S., Senyurt, S.Y., Ahmad, H., McKenzie, C., Margaryan, A., Hameed, A., Samad, A., Gul, N., Khokhar, M.H., Goriunova, O.I., Bazaliiskii, V.I., Novembre, J., Weber, A.W., Orlando, L., Allentoft, M.E., Nielsen, R., Kristiansen, K., Sikora, M., Outram, A.K., Durbin, R., Willerslev, E., 2018. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. Science 360, eaar7711. https://doi.org/10.1126/science.aar7711

Demeure, C.E., Dussurget, O., Mas Fiol, G., Le Guern, A.-S., Savin, C., Pizarro-Cerdá, J., 2019. Yersinia pestis and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. Genes Immun 20, 357–370. https://doi.org/10.1038/s41435-019-0065-0

Der Sarkissian, C., Pichereau, V., Dupont, C., Ilsøe, P.C., Perrigault, M., Butler, P., Chauvaud, L., Eiríksson, J., Scourse, J., Paillard, C., Orlando, L., 2017. Ancient DNA analysis identifies marine mollusc shells as new metagenomic archives of the past. Mol Ecol Resour 17, 835–853. https://doi.org/10.1111/1755-0998.12679

Deschamps, M., Laval, G., Fagny, M., Itan, Y., Abel, L., Casanova, J.-L., Patin, E., Quintana-Murci, L., 2016. Genomic Signatures of Selective Pressures and Introgression from Archaic Hominins at Human Innate Immunity Genes. Am J Hum Genet 98, 5–21. https://doi.org/10.1016/j.ajhg.2015.11.014

Devault, A.M., Golding, G.B., Waglechner, N., Enk, J.M., Kuch, M., Tien, J.H., Shi, M., Fisman, D.N., Dhody, A.N., Forrest, S., Bos, K.I., Earn, D.J.D., Holmes, E.C., Poinar, H.N., 2014. Second-pandemic strain of Vibrio cholerae from the Philadelphia cholera outbreak of 1849. N Engl J Med 370, 334–340. https://doi.org/10.1056/nejmoa1308663

DeWitte, S.N., Stojanowski, C.M., 2015. The Osteological Paradox 20 Years Later: Past Perspectives, Future Directions. J Archaeol Res 23, 397–450. https://doi.org/10.1007/s10814-015-9084-1

Di, D., Simon Thomas, J., Currat, M., Nunes, J.M., Sanchez-Mazas, A., 2022. Challenging Ancient DNA Results About Putative HLA Protection or Susceptibility to Yersinia pestis. Molecular Biology and Evolution 39, msac073. https://doi.org/10.1093/molbev/msac073
Dickau, R., Ranere, A.J., Cooke, R.G., 2007. Starch grain evidence for the preceramic dispersals of maize and root crops into tropical dry and humid forests of Panama. Proc Natl Acad Sci U S A 104, 3651–3656. https://doi.org/10.1073/pnas.0611605104
Dobyns, H.F., 1993. Disease Transfer at Contact. Annual Review of Anthropology 22, 273–291.
Drancourt, M., Aboudharam, G., Signoli, M., Dutour, O., Raoult, D., 1998. Detection of 400-year-old Yersinia pestis DNA in human dental pulp: An approach to the diagnosis of ancient septicemia. Proceedings of the National Academy of Sciences 95, 12637–12640. https://doi.org/10.1073/pnas.95.21.12637
Drancourt, M., Raoult, D., 2005. Palaeomicrobiology: current issues and perspectives. Nat Rev Microbiol 3, 23–35. https://doi.org/10.1038/nrmicro1063
Drancourt, M., Raoult, D., 2004. Molecular detection of Yersinia pestis in dental pulp. Microbiology (Reading) 150, 263–264. https://doi.org/10.1099/mic.0.26885-0
Duchêne, S., Ho, S.Y.W., Carmichael, A.G., Holmes, E.C., Poinar, H., 2020. The Recovery, Interpretation and Use of Ancient Pathogen Genomes. Current Biology 30, R1215–R1231. https://doi.org/10.1016/j.cub.2020.08.081
Duchêne, S., Holt, K.E., Weill, F.-X., Le Hello, S., Hawkey, J., Edwards, D.J., Fourment, M., Holmes, E.C., 2016. Genome-scale rates of evolutionary change in bacteria. Microb Genom 2, e000094. https://doi.org/10.1099/mgen.0.000094
Duggan, A.T., Perdomo, M.F., Piombino-Mascali, D., Marciniak, S., Poinar, D., Emery, M.V., Buchmann, J.P., Duchêne, S., Jankauskas, R., Humphreys, M., Golding, G.B., Southon, J., Devault, A., Rouillard, J.-M., Sahl, J.W., Dutour, O., Hedman, K., Sajantila, A., Smith, G.L., Holmes, E.C., Poinar, H.N., 2016. 17th Century Variola Virus Reveals the Recent History of Smallpox. Curr Biol 26, 3407–3412. https://doi.org/10.1016/j.cub.2016.10.061
Eaton, K., Featherstone, L., Duchene, S., Carmichael, A.G., Varlık, N., Golding, G.B., Holmes, E.C., Poinar, H.N., 2023. Plagued by a cryptic clock: insight and issues from the global phylogeny of Yersinia pestis. Commun Biol 6, 1–13. https://doi.org/10.1038/s42003-022-04394-6
Eng, S.-K., Pusparajah, P., Ab Mutalib, N.-S., Ser, H.-L., Chan, K.-G., Lee, L.-H., 2015. Salmonella: A review on pathogenesis, epidemiology and antibiotic resistance. Frontiers in Life Science 8, 284–293. https://doi.org/10.1080/21553769.2015.1051243
Enk, J., Rouillard, J.-M., Poinar, H., 2013. Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. BioTechniques 55, 300–309. https://doi.org/10.2144/000114114
Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.-M., Poinar, H.N., 2014. Ancient Whole Genome Enrichment Using Baits Built from Modern DNA. Molecular Biology and Evolution 31, 1292–1294. https://doi.org/10.1093/molbev/msu074
Ermini, L., Olivieri, C., Rizzi, E., Corti, G., Bonnal, R., Soares, P., Luciani, S., Marota, I., De Bellis, G., Richards, M.B., Rollo, F., 2008. Complete mitochondrial genome sequence of the Tyrolean Iceman. Curr Biol 18, 1687–1693. https://doi.org/10.1016/j.cub.2008.09.028
Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., McCrory Constantz, C., Gamba, C., Al-Rasheid, K.A.S., Albizuri, S., Alfarhan, A.H., Allentoft, M., Alquraishi, S., Anthony, D., Baimukhanov, N., Barrett, J.H., Bayarsaikhan, J., Benecke, N., Bernáldez-Sánchez, E., Berrocal-Rangel, L., Biglari, F., Boessenkool, S., Boldgiv, B., Brem, G., Brown, D., Burger, J., Crubézy, E., Daugnora, L., Davoudi, H., de Barros Damgaard, P., de los Ángeles de Chorro y de Villa-Ceballos, M., Deschler-Erb, S., Detry, C., Dill, N., do Mar Oom, M., Dohr, A., Ellingvåg, S., Erdenebaatar, D., Fathi, H., Felkel, S., Fernández-

Rodríguez, C., García-Viñas, E., Germonpré, M., Granado, J.D., Hallsson, J.H., Hemmer, H., Hofreiter, M., Kasparov, A., Khasanov, M., Khazaeli, R., Kosintsev, P., Kristiansen, K., Kubatbek, T., Kuderna, L., Kuznetsov, P., Laleh, H., Leonard, J.A., Lhuillier, J., Liesau von Lettow-Vorbeck, C., Logvin, A., Lõugas, L., Ludwig, A., Luis, C., Arruda, A.M., Marques-Bonet, T., Matoso Silva, R., Merz, V., Mijiddorj, E., Miller, B.K., Monchalov, O., Mohaseb, F.A., Morales, A., Nieto-Espinet, A., Nistelberger, H., Onar, V., Pálsdóttir, A.H., Pitulko, V., Pitskhelauri, K., Pruvost, M., Rajic Sikanjic, P., Rapan Papeša, A., Roslyakova, N., Sardari, A., Sauer, E., Schafberg, R., Scheu, A., Schibler, J., Schlumbaum, A., Serrand, N., Serres-Armero, A., Shapiro, B., Sheikhi Seno, S., Shevnina, I., Shidrang, S., Southon, J., Star, B., Sykes, N., Taheri, K., Taylor, W., Teegen, W.-R., Trbojević Vukičević, T., Trixl, S., Tumen, D., Undrakhbold, S., Usmanova, E., Vahdati, A., Valenzuela-Lamas, S., Viegas, C., Wallner, B., Weinstock, J., Zaibert, V., Clavel, B., Lepetz, S., Mashkour, M., Helgason, A., Stefánsson, K., Barrey, E., Willerslev, E., Outram, A.K., Librado, P., Orlando, L., 2019. Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. Cell 177, 1419-1435.e31. https://doi.org/10.1016/j.cell.2019.03.049

Fortes, G.G., Paijmans, J.L.A., 2015. Analysis of Whole Mitogenomes from Ancient Samples, in: Kroneis, T. (Ed.), Whole Genome Amplification: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 179–195. https://doi.org/10.1007/978-1-4939-2990-0_13

Frantz, L.A.F., Bradley, D.G., Larson, G., Orlando, L., 2020. Animal domestication in the era of ancient genomics. Nat Rev Genet 21, 449–460. https://doi.org/10.1038/s41576-020-0225-0

Fregel, R., Méndez, F.L., Bokbot, Y., Martín-Socas, D., Camalich-Massieu, M.D., Santana, J., Morales, J., Ávila-Arcos, M.C., Underhill, P.A., Shapiro, B., Wojcik, G., Rasmussen, M., Soares, A.E.R., Kapp, J., Sockell, A., Rodríguez-Santos, F.J., Mikdad, A., Trujillo-Mederos, A., Bustamante, C.D., 2018. Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. Proceedings of the National Academy of Sciences 115, 6774–6779. https://doi.org/10.1073/pnas.1800851115

Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S.M., Bondarev, A.A., Johnson, P.L.F., Aximu-Petri, A., Prüfer, K., de Filippo, C., Meyer, M., Zwyns, N., Salazar-García, D.C., Kuzmin, Y.V., Keates, S.G., Kosintsev, P.A., Razhev, D.I., Richards, M.P., Peristov, N.V., Lachmann, M., Douka, K., Higham, T.F.G., Slatkin, M., Hublin, J.-J., Reich, D., Kelso, J., Viola, T.B., Pääbo, S., 2014. Genome sequence of a 45,000-year-old modern human from western Siberia. Nature 514, 445–449. https://doi.org/10.1038/nature13810

Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H.A., Kelso, J., Pääbo, S., 2013a. DNA analysis of an early modern human from Tianyuan Cave, China. Proceedings of the National Academy of Sciences 110, 2223–2227. https://doi.org/10.1073/pnas.1221359110

Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H.A., Kelso, J., Pääbo, S., 2013b. DNA analysis of an early modern human from Tianyuan Cave, China. Proceedings of the National Academy of Sciences 110, 2223–2227. https://doi.org/10.1073/pnas.1221359110

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admetlla, A., Pattini, L., Nielsen, R., 2011. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. PLoS Genet 7, e1002355. https://doi.org/10.1371/journal.pgen.1002355

Gage, K.L., Kosoy, M.Y., 2005. Natural history of plague: perspectives from more than a century of research. Annu Rev Entomol 50, 505–528. https://doi.org/10.1146/annurev.ento.50.071803.130337

Gamba, C., Hanghøj, K., Gaunitz, C., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Bradley, D.G., Orlando, L., 2016. Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. Molecular Ecology Resources 16, 459–469. https://doi.org/10.1111/1755-0998.12470

Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T.F.G., Hofreiter, M., Bradley, D.G., Pinhasi, R., 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nat Commun 5, 5257. https://doi.org/10.1038/ncomms6257
Gansauge, M.-T., Gerber, T., Glocke, I., Korlevic, P., Lippik, L., Nagel, S., Riehl, L.M., Schmidt, A., Meyer, M., 2017. Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. Nucleic Acids Res 45, e79. https://doi.org/10.1093/nar/gkx033
Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M., Seguin-Orlando, A., Owens, I.J., Felkel, S., Bignon-Lau, O., de Barros Damgaard, P., Mittnik, A., Mohaseb, A.F., Davoudi, H., Alquraishi, S., Alfarhan, A.H., Al-Rasheid, K.A.S., Crubézy, E., Benecke, N., Olsen, S., Brown, D., Anthony, D., Massy, K., Pitulko, V., Kasparov, A., Brem, G., Hofreiter, M., Mukhtarova, G., Baimukhanov, N., Lõugas, L., Onar, V., Stockhammer, P.W., Krause, J., Boldgiv, B., Undrakhbold, S., Erdenebaatar, D., Lepetz, S., Mashkour, M., Ludwig, A., Wallner, B., Merz, V., Merz, I., Zaibert, V., Willerslev, E., Librado, P., Outram, A.K., Orlando, L., 2018. Ancient genomes revisit the ancestry of domestic and Przewalski's horses. Science 360, 111–114. https://doi.org/10.1126/science.aao3297
Gautier, M., Gharbi, K., Cezard, T., Foucaud, J., Kerdelhué, C., Pudlo, P., Cornuet, J.-M., Estoup, A., 2013. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. Mol Ecol 22, 3165–3178. https://doi.org/10.1111/mec.12089
Gelabert, P., Sandoval-Velasco, M., Olalde, I., Fregel, R., Rieux, A., Escosa, R., Aranda, C., Paaijmans, K., Mueller, I., Gilbert, M.T.P., Lalueza-Fox, C., 2016. Mitochondrial DNA from the eradicated European Plasmodium vivax and P. falciparum from 70-year-old slides from the Ebro Delta in Spain. Proceedings of the National Academy of Sciences 113, 11495–11500. https://doi.org/10.1073/pnas.1611017113
Gilbert, M.T.P., Wilson, A.S., Bunce, M., Hansen, A.J., Willerslev, E., Shapiro, B., Higham, T.F.G., Richards, M.P., O'Connell, T.C., Tobin, D.J., Janaway, R.C., Cooper, A., 2004. Ancient mitochondrial DNA from hair. Curr Biol 14, R463-464. https://doi.org/10.1016/j.cub.2004.06.008
Giles, R.E., Shimizu, N., Ruddle, F.H., 1980. Assignment of a human genetic locus to chromosome 5 which corrects the heat sensitive lesion associated with reduced leucyl-tRNA synthetase activity in ts025Cl Chinese hamster cells. Somatic Cell Genet 6, 667–687. https://doi.org/10.1007/BF01538645
Ginolhac, A., Vilstrup, J., Stenderup, J., Rasmussen, M., Stiller, M., Shapiro, B., Zazula, G., Froese, D., Steinmann, K.E., Thompson, J.F., AL-Rasheid, K.A., Gilbert, T.M., Willerslev, E., Orlando, L., 2012. Improving the performance of true single molecule sequencing for ancient DNA. BMC Genomics 13, 177. https://doi.org/10.1186/1471-2164-13-177
Gratani, L., 2014. Plant Phenotypic Plasticity in Response to Environmental Factors. Advances in Botany 2014, 1–17. https://doi.org/10.1155/2014/208747
Green, M.H., 2020. The Four Black Deaths. The American Historical Review 125, 1601–1631. https://doi.org/10.1093/ahr/rhaa511
Green, R.E., Briggs, A.W., Krause, J., Prüfer, K., Burbano, H.A., Siebauer, M., Lachmann, M., Pääbo, S., 2009. The Neandertal genome and ancient DNA authenticity. EMBO J 28, 2494–2502. https://doi.org/10.1038/emboj.2009.222
Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., Hansen, N.F., Durand, E.Y., Malaspinas, A.-S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan,

P., Brajkovic, D., Kucan, Ž., Gušic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L.F., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S., 2010. A Draft Sequence of the Neandertal Genome. Science 328, 710–722. https://doi.org/10.1126/science.1188021
Grobman, A., Bonavia, D., Dillehay, T.D., Piperno, D.R., Iriarte, J., Holst, I., 2012. Preceramic maize from Paredones and Huaca Prieta, Peru. Proceedings of the National Academy of Sciences 109, 1755–1759. https://doi.org/10.1073/pnas.1120270109
Guellil, M., Kersten, O., Namouchi, A., Luciani, S., Marota, I., Arcini, C.A., Iregren, E., Lindemann, R.A., Warfvinge, G., Bakanidze, L., Bitadze, L., Rubini, M., Zaio, P., Zaio, M., Neri, D., Stenseth, N.C., Bramanti, B., 2020. A genomic and historical synthesis of plague in 18th century Eurasia. Proceedings of the National Academy of Sciences 117, 28328–28335. https://doi.org/10.1073/pnas.2009677117
Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., Fu, Q., Mittnik, A., Bánffy, E., Economou, C., Francken, M., Friederich, S., Pena, R.G., Hallgren, F., Khartanovich, V., Khokhlov, A., Kunst, M., Kuznetsov, P., Meller, H., Mochalov, O., Moiseyev, V., Nicklisch, N., Pichler, S.L., Risch, R., Rojo Guerra, M.A., Roth, C., Szécsényi-Nagy, A., Wahl, J., Meyer, M., Krause, J., Brown, D., Anthony, D., Cooper, A., Alt, K.W., Reich, D., 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. Nature 522, 207–211. https://doi.org/10.1038/nature14317
Hagelberg, E., Clegg, J.B., 1991. Isolation and characterization of DNA from archaeological bone. Proc Biol Sci 244, 45–50. https://doi.org/10.1098/rspb.1991.0049
Hagelberg, E., Sykes, B., Hedges, R., 1989. Ancient bone DNA amplified. Nature 342, 485–485. https://doi.org/10.1038/342485a0
Haile, J., Froese, D.G., MacPhee, R.D.E., Roberts, R.G., Arnold, L.J., Reyes, A.V., Rasmussen, M., Nielsen, R., Brook, B.W., Robinson, S., Demuro, M., Gilbert, M.T.P., Munch, K., Austin, J.J., Cooper, A., Barnes, I., Möller, P., Willerslev, E., 2009. Ancient DNA reveals late survival of mammoth and horse in interior Alaska. Proceedings of the National Academy of Sciences 106, 22352–22357. https://doi.org/10.1073/pnas.0912510106
Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M.T.P., Nielsen, R., Munch, K., Ho, S.Y.W., Shapiro, B., Willerslev, E., 2007. Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? Mol Biol Evol 24, 982–989. https://doi.org/10.1093/molbev/msm016
Handt, O., Krings, M., Ward, R.H., Pääbo, S., 1996. The retrieval of ancient human DNA sequences. Am J Hum Genet 59, 368–376.
Hanghøj, K., Renaud, G., Albrechtsen, A., Orlando, L., 2019. DamMet: ancient methylome mapping accounting for errors, true variants, and post-mortem DNA damage. Gigascience 8, giz025. https://doi.org/10.1093/gigascience/giz025
Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. PLoS One 12, e0170940. https://doi.org/10.1371/journal.pone.0170940
Harlan, J.R., Gepts, P., Famula, T.R., Bettinger, R.L., Brush, S.B., Damania, A.B., McGuire, P.E., Qualset, C.O., 2012. Biodiversity in Agriculture: Domestication, Evolution, and Sustainability. Cambridge University Press.
Hawkins, M.T.R., Flores, M.F.C., McGowen, M., Hinckley, A., 2022. A comparative analysis of extraction protocol performance on degraded mammalian museum specimens. Frontiers in Ecology and Evolution 10.
Hernandez-Rodriguez, J., Arandjelovic, M., Lester, J., de Filippo, C., Weihmann, A., Meyer, M., Angedakin, S., Casals, F., Navarro, A., Vigilant, L., Kühl, H.S., Langergraber, K.,

Boesch, C., Hughes, D., Marques-Bonet, T., 2018. The impact of endogenous content, replicates and pooling on genome capture from faecal samples. Molecular Ecology Resources 18, 319–333. https://doi.org/10.1111/1755-0998.12728
Higgins, D., Kaidonis, J., Townsend, G., Hughes, T., Austin, J.J., 2013. Targeted sampling of cementum for recovery of nuclear DNA from human teeth and the impact of common decontamination measures. Investig Genet 4, 18. https://doi.org/10.1186/2041-2223-4-18
Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C., 1984. DNA sequences from the quagga, an extinct member of the horse family. Nature 312, 282–284. https://doi.org/10.1038/312282a0
Hinnebusch, B.J., Chouikha, I., Sun, Y.-C., 2016. Ecological Opportunity, Evolution, and the Emergence of Flea-Borne Plague. Infect Immun 84, 1932–1940. https://doi.org/10.1128/IAI.00188-16
Hinnebusch, B.J., Erickson, D.L., 2008. Yersinia pestis biofilm in the flea vector and its role in the transmission of plague. Curr Top Microbiol Immunol 322, 229–248. https://doi.org/10.1007/978-3-540-75418-3_11
Ho, S.Y.W., Heupink, T.H., Rambaut, A., Shapiro, B., 2007. Bayesian Estimation of Sequence Damage in Ancient DNA. Molecular Biology and Evolution 24, 1416–1422. https://doi.org/10.1093/molbev/msm062
Hofmanová, Z., Kreutzer, S., Hellenthal, G., Sell, C., Diekmann, Y., Díez-del-Molino, D., van Dorp, L., López, S., Kousathanas, A., Link, V., Kirsanow, K., Cassidy, L.M., Martiniano, R., Strobel, M., Scheu, A., Kotsakis, K., Halstead, P., Triantaphyllou, S., Kyparissi-Apostolika, N., Urem-Kotsou, D., Ziota, C., Adaktylou, F., Gopalan, S., Bobo, D.M., Winkelbach, L., Blöcher, J., Unterländer, M., Leuenberger, C., Çilingiroğlu, Ç., Horejs, B., Gerritsen, F., Shennan, S.J., Bradley, D.G., Currat, M., Veeramah, K.R., Wegmann, D., Thomas, M.G., Papageorgopoulou, C., Burger, J., 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. Proceedings of the National Academy of Sciences 113, 6886–6891. https://doi.org/10.1073/pnas.1523951113
Hofreiter, M., Poinar, H.N., Spaulding, W.G., Bauer, K., Martin, P.S., Possnert, G., Pääbo, S., 2000. A molecular analysis of ground sloth diet through the last glaciation. Mol Ecol 9, 1975–1984. https://doi.org/10.1046/j.1365-294x.2000.01106.x
Höss, M., Pääbo, S., 1993. DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Res 21, 3913–3914.
Huang, Y., Ringbauer, H., 2022. hapCon: estimating contamination of ancient genomes by copying from reference haplotypes. Bioinformatics 38, 3768–3777. https://doi.org/10.1093/bioinformatics/btac390
Hummel, S., 2003. aDNA extraction, in: Hummel, S. (Ed.), Ancient DNA Typing: Methods, Strategies and Applications. Springer, Berlin, Heidelberg, pp. 57–80. https://doi.org/10.1007/978-3-662-05050-7_3
Immel, A., Key, F.M., Szolek, A., Barquera, R., Robinson, M.K., Harrison, G.F., Palmer, W.H., Spyrou, M.A., Susat, J., Krause-Kyora, B., Bos, K.I., Forrest, S., Hernández-Zaragoza, D.I., Sauter, J., Solloch, U., Schmidt, A.H., Schuenemann, V.J., Reiter, E., Kairies, M.S., Weiß, R., Arnold, S., Wahl, J., Hollenbach, J.A., Kohlbacher, O., Herbig, A., Norman, P.J., Krause, J., 2021. Analysis of Genomic DNA from Medieval Plague Victims Suggests Long-Term Effect of Yersinia pestis on Human Immunity Genes. Molecular Biology and Evolution 38, 4059–4076. https://doi.org/10.1093/molbev/msab147
Inhorn, M.C., Brown, P.J., 1990. The Anthropology of Infectious Disease. Annual Review of Anthropology 19, 89–117.
Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Karlsson, E.K., Kwiatkowski, D.P., Sabeti, P.C., 2014. Natural selection and infectious disease in human populations. Nat Rev Genet 15, 379–393. https://doi.org/10.1038/nrg3734
Kay, G.L., Sergeant, M.J., Zhou, Z., Chan, J.Z.-M., Millard, A., Quick, J., Szikossy, I., Pap, I., Spigelman, M., Loman, N.J., Achtman, M., Donoghue, H.D., Pallen, M.J., 2015. Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. Nat Commun 6, 6717. https://doi.org/10.1038/ncomms7717
Keller, A., Graefen, A., Ball, M., Matzas, M., Boisguerin, V., Maixner, F., Leidinger, P., Backes, C., Khairat, R., Forster, M., Stade, B., Franke, A., Mayer, J., Spangler, J., McLaughlin, S., Shah, M., Lee, C., Harkins, T.T., Sartori, A., Moreno-Estrada, A., Henn, B., Sikora, M., Semino, O., Chiaroni, J., Rootsi, S., Myres, N.M., Cabrera, V.M., Underhill, P.A., Bustamante, C.D., Vigl, E.E., Samadelli, M., Cipollini, G., Haas, J., Katus, H., O'Connor, B.D., Carlson, M.R.J., Meder, B., Blin, N., Meese, E., Pusch, C.M., Zink, A., 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat Commun 3, 698. https://doi.org/10.1038/ncomms1701
Keller, M., Spyrou, M.A., Scheib, C.L., Neumann, G.U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., Ribera i Lacomba, A., Raynaud, C., Cessford, C., Durand, R., Stadler, P., Nägele, K., Bates, J.S., Trautmann, B., Inskip, S.A., Peters, J., Robb, J.E., Kivisild, T., Castex, D., McCormick, M., Bos, K.I., Harbeck, M., Herbig, A., Krause, J., 2019. Ancient Yersinia pestis genomes from across Western Europe reveal early diversification during the First Pandemic (541–750). Proceedings of the National Academy of Sciences 116, 12363–12372. https://doi.org/10.1073/pnas.1820447116
Kemp, B.M., Monroe, C., Smith, D.G., 2006. Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. Journal of Archaeological Science 33, 1680–1689. https://doi.org/10.1016/j.jas.2006.02.015
Kerner, G., Laval, G., Patin, E., Boisson-Dupuis, S., Abel, L., Casanova, J.-L., Quintana-Murci, L., 2021. Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2,000 years. Am J Hum Genet 108, 517–524. https://doi.org/10.1016/j.ajhg.2021.02.009
Kerner, G., Neehus, A.-L., Philippot, Q., Bohlen, J., Rinchai, D., Kerrouche, N., Puel, A., Zhang, S.-Y., Boisson-Dupuis, S., Abel, L., Casanova, J.-L., Patin, E., Laval, G., Quintana-Murci, L., 2023. Genetic adaptation to pathogens and increased risk of inflammatory disorders in post-Neolithic Europe. Cell Genomics 3, 100248. https://doi.org/10.1016/j.xgen.2022.100248
Key, F.M., Posth, C., Esquivel-Gomez, L.R., Hübler, R., Spyrou, M.A., Neumann, G.U., Furtwängler, A., Sabin, S., Burri, M., Wissgott, A., Lankapalli, A.K., Vågene, Å.J., Meyer, M., Nagel, S., Tukhbatova, R., Khokhlov, A., Chizhevsky, A., Hansen, S., Belinsky, A.B., Kalmykov, A., Kantorovich, A.R., Maslov, V.E., Stockhammer, P.W., Vai, S., Zavattaro, M., Riga, A., Caramelli, D., Skeates, R., Beckett, J., Gradoli, M.G., Steuri, N., Hafner, A., Ramstein, M., Siebke, I., Lösch, S., Erdal, Y.S., Alikhan, N.-F., Zhou, Z., Achtman, M., Bos, K., Reinhold, S., Haak, W., Kühnert, D., Herbig, A., Krause, J., 2020. Emergence of human-adapted Salmonella enterica is linked to the Neolithization process. Nat Ecol Evol 4, 324–333. https://doi.org/10.1038/s41559-020-1106-9
Kjær, K.H., Winther Pedersen, M., De Sanctis, B., De Cahsan, B., Korneliussen, T.S., Michelsen, C.S., Sand, K.K., Jelavić, S., Ruter, A.H., Schmidt, A.M.A., Kjeldsen, K.K., Tesakov, A.S., Snowball, I., Gosse, J.C., Alsos, I.G., Wang, Y., Dockter, C., Rasmussen, M., Jørgensen, M.E., Skadhauge, B., Prohaska, A., Kristensen, J.Å., Bjerager, M., Allentoft, M.E., Coissac, E., Rouillard, A., Simakova, A., Fernandez-Guerra, A., Bowler, C., Macias-Fauria, M., Vinner, L., Welch, J.J., Hidy, A.J., Sikora, M., Collins, M.J., Durbin, R., Larsen, N.K., Willerslev, E., 2022. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. Nature 612, 283–291. https://doi.org/10.1038/s41586-022-05453-y

Klima, H., Ullrich, K., Aslanidis, C., Fehringer, P., Lackner, K.J., Schmitz, G., 1993. A splice junction mutation causes deletion of a 72-base exon from the mRNA for lysosomal acid lipase in a patient with cholesteryl ester storage disease. J Clin Invest 92, 2713–2718. https://doi.org/10.1172/JCI116888
Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G., Kronenberg, F., 2011. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Human Mutation 32, 25–32. https://doi.org/10.1002/humu.21382
Klunk, J., Vilgalys, T.P., Demeure, C.E., Cheng, X., Shiratori, M., Madej, J., Beau, R., Elli, D., Patino, M.I., Redfern, R., DeWitte, S.N., Gamble, J.A., Boldsen, J.L., Carmichael, A., Varlik, N., Eaton, K., Grenier, J.-C., Golding, G.B., Devault, A., Rouillard, J.-M., Yotova, V., Sindeaux, R., Ye, C.J., Bikaran, M., Dumaine, A., Brinkworth, J.F., Missiakas, D., Rouleau, G.A., Steinrücken, M., Pizarro-Cerdá, J., Poinar, H.N., Barreiro, L.B., 2022. Evolution of immune genes is associated with the Black Death. Nature 611, 312–319. https://doi.org/10.1038/s41586-022-05349-x
Knapp, M., Clarke, A.C., Horsburgh, K.A., Matisoo-Smith, E.A., 2012. Setting the stage – Building and working in an ancient DNA laboratory. Annals of Anatomy - Anatomischer Anzeiger, Special Issue: Ancient DNA 194, 3–6. https://doi.org/10.1016/j.aanat.2011.03.008
Kocher, A., Papac, L., Barquera, R., Key, F.M., Spyrou, M.A., Hübler, R., Rohrlach, A.B., Aron, F., Stahl, R., Wissgott, A., van Bömmel, F., Pfefferkorn, M., Mittnik, A., Villalba-Mouco, V., Neumann, G.U., Rivollat, M., van de Loosdrecht, M.S., Majander, K., Tukhbatova, R.I., Musralina, L., Ghalichi, A., Penske, S., Sabin, S., Michel, M., Gretzinger, J., Nelson, E.A., Ferraz, T., Nägele, K., Parker, C., Keller, M., Guevara, E.K., Feldman, M., Eisenmann, S., Skourtanioti, E., Giffin, K., Gnecchi-Ruscone, G.A., Friederich, S., Schimmenti, V., Khartanovich, V., Karapetian, M.K., Chaplygin, M.S., Kufterin, V.V., Khokhlov, A.A., Chizhevsky, A.A., Stashenkov, D.A., Kochkina, A.F., Tejedor-Rodríguez, C., de Lagrán, Í.G.-M., Arcusa-Magallón, H., Garrido-Pena, R., Royo-Guillén, J.I., Nováček, J., Rottier, S., Kacki, S., Saintot, S., Kaverzneva, E., Belinskiy, A.B., Velemínský, P., Limburský, P., Kostka, M., Loe, L., Popescu, E., Clarke, R., Lyons, A., Mortimer, R., Sajantila, A., de Armas, Y.C., Hernandez Godoy, S.T., Hernández-Zaragoza, D.I., Pearson, J., Binder, D., Lefranc, P., Kantorovich, A.R., Maslov, V.E., Lai, L., Zoledziewska, M., Beckett, J.F., Langová, M., Danielisová, A., Ingman, T., Atiénzar, G.G., de Miguel Ibáñez, M.P., Romero, A., Sperduti, A., Beckett, S., Salter, S.J., Zilivinskaya, E.D., Vasil'ev, D.V., von Heyking, K., Burger, R.L., Salazar, L.C., Amkreutz, L., Navruzbekov, M., Rosenstock, E., Alonso-Fernández, C., Slavchev, V., Kalmykov, A.A., Atabiev, B.Ch., Batieva, E., Calmet, M.A., Llamas, B., Schultz, M., Krauß, R., Jiménez-Echevarría, J., Francken, M., Shnaider, S., de Knijff, P., Altena, E., Van de Vijver, K., Fehren-Schmitz, L., Tung, T.A., Lösch, S., Dobrovolskaya, M., Makarov, N., Read, C., Van Twest, M., Sagona, C., Ramsl, P.C., Akar, M., Yener, K.A., Ballestero, E.C., Cucca, F., Mazzarello, V., Utrilla, P., Rademaker, K., Fernández-Domínguez, E., Baird, D., Semal, P., Márquez-Morfín, L., Roksandic, M., Steiner, H., Salazar-García, D.C., Shishlina, N., Erdal, Y.S., Hallgren, F., Boyadzhiev, Y., Boyadzhiev, K., Küßner, M., Sayer, D., Onkamo, P., Skeates, R., Rojo-Guerra, M., Buzhilova, A., Khussainova, E., Djansugurova, L.B., Beisenov, A.Z., Samashev, Z., Massy, K., Mannino, M., Moiseyev, V., Mannermaa, K., Balanovsky, O., Deguilloux, M.-F., Reinhold, S., Hansen, S., Kitov, E.P., Dobeš, M., Ernée, M., Meller, H., Alt, K.W., Prüfer, K., Warinner, C., Schiffels, S., Stockhammer, P.W., Bos, K., Posth, C., Herbig, A., Haak, W., Krause, J., Kühnert, D., 2021. Ten millennia of hepatitis B virus evolution. Science 374, 182–188. https://doi.org/10.1126/science.abi5658
Kolling, G., Wu, M., Guerrant, R., 2012. Enteric pathogens through life stages. Frontiers in Cellular and Infection Microbiology 2.

Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., Meyer, M., 2015. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. BioTechniques 59, 87–93. https://doi.org/10.2144/000114320
Korneliussen, T.S., Moltke, I., Albrechtsen, A., Nielsen, R., 2013. Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics 14, 289. https://doi.org/10.1186/1471-2105-14-289
Krause, J., Briggs, A.W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., Pääbo, S., 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol 20, 231–236. https://doi.org/10.1016/j.cub.2009.11.068
Krause-Kyora, B., Nutsua, M., Boehme, L., Pierini, F., Pedersen, D.D., Kornell, S.-C., Drichel, D., Bonazzi, M., Möbus, L., Tarp, P., Susat, J., Bosse, E., Willburger, B., Schmidt, A.H., Sauter, J., Franke, A., Wittig, M., Caliebe, A., Nothnagel, M., Schreiber, S., Boldsen, J.L., Lenz, T.L., Nebel, A., 2018a. Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. Nat Commun 9, 1569. https://doi.org/10.1038/s41467-018-03857-x
Krause-Kyora, B., Susat, J., Key, F.M., Kühnert, D., Bosse, E., Immel, A., Rinne, C., Kornell, S.-C., Yepes, D., Franzenburg, S., Heyne, H.O., Meier, T., Lösch, S., Meller, H., Friederich, S., Nicklisch, N., Alt, K.W., Schreiber, S., Tholey, A., Herbig, A., Nebel, A., Krause, J., 2018b. Neolithic and medieval virus genomes reveal complex evolution of hepatitis B. eLife 7, e36666. https://doi.org/10.7554/eLife.36666
Kuhn, J.M.M., Jakobsson, M., Günther, T., 2018. Estimating genetic kin relationships in prehistoric populations. PLOS ONE 13, e0195491. https://doi.org/10.1371/journal.pone.0195491
Lam, Y.M., Chen, X., Pearson, O.M., 1999. Intertaxonomic Variability in Patterns of Bone Density and the Differential Representation of Bovid, Cervid, and Equid Elements in the Archaeological Record. American Antiquity 64, 343–362. https://doi.org/10.2307/2694283
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923
Larson, G., Piperno, D.R., Allaby, R.G., Purugganan, M.D., Andersson, L., Arroyo-Kalin, M., Barton, L., Climer Vigueira, C., Denham, T., Dobney, K., Doust, A.N., Gepts, P., Gilbert, M.T.P., Gremillion, K.J., Lucas, L., Lukens, L., Marshall, F.B., Olsen, K.M., Pires, J.C., Richerson, P.J., Rubio de Casas, R., Sanjur, O.I., Thomas, M.G., Fuller, D.Q., 2014. Current perspectives and the future of domestication studies. Proceedings of the National Academy of Sciences 111, 6139–6146. https://doi.org/10.1073/pnas.1323964111
Lauterbur, M.E., 2019. orthoCapture: Facilitating Gene Capture Probe Design for Non-Model Species (preprint). Bioinformatics. https://doi.org/10.1101/703942
Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Gallego Romero, I., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D.,

Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C.A., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J., 2014a. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513, 409–413. https://doi.org/10.1038/nature13673
Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Gallego Romero, I., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C.A., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J., 2014b. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513, 409–413. https://doi.org/10.1038/nature13673
Librado, P., Khan, N., Fages, A., Kusliy, M.A., Suchan, T., Tonasso-Calvière, L., Schiavinato, S., Alioglu, D., Fromentier, A., Perdereau, A., Aury, J.-M., Gaunitz, C., Chauvey, L., Seguin-Orlando, A., Der Sarkissian, C., Southon, J., Shapiro, B., Tishkin, A.A., Kovalev, A.A., Alquraishi, S., Alfarhan, A.H., Al-Rasheid, K.A.S., Seregély, T., Klassen, L., Iversen, R., Bignon-Lau, O., Bodu, P., Olive, M., Castel, J.-C., Boudadi-Maligne, M., Alvarez, N., Germonpré, M., Moskal-del Hoyo, M., Wilczyński, J., Pospuła, S., Lasota-Kuś, A., Tunia, K., Nowak, M., Rannamäe, E., Saarma, U., Boeskorov, G., Lõugas, L., Kyselý, R., Peške, L., Bălăşescu, A., Dumitraşcu, V., Dobrescu, R., Gerber, D., Kiss, V., Szécsényi-Nagy, A., Mende, B.G., Gallina, Z., Somogyi, K., Kulcsár, G., Gál, E., Bendrey, R., Allentoft, M.E., Sirbu, G., Dergachev, V., Shephard, H., Tomadini, N., Grouard, S., Kasparov, A., Basilyan, A.E., Anisimov, M.A., Nikolskiy, P.A., Pavlova, E.Y., Pitulko, V., Brem, G., Wallner, B., Schwall, C., Keller, M., Kitagawa, K., Bessudnov, A.N., Bessudnov, A., Taylor, W., Magail, J., Gantulga, J.-O., Bayarsaikhan, J., Erdenebaatar, D., Tabaldiev, K., Mijiddorj, E., Boldgiv, B., Tsagaan, T., Pruvost, M., Olsen, S., Makarewicz, C.A., Valenzuela Lamas, S., Albizuri Canadell, S., Nieto Espinet, A., Iborra, M.P., Lira Garrido, J., Rodríguez González, E., Celestino, S., Olària, C., Arsuaga, J.L., Kotova, N., Pryor, A., Crabtree, P., Zhumatayev, R., Toleubaev, A., Morgunova, N.L., Kuznetsova, T., Lordkipanize, D., Marzullo, M., Prato, O., Bagnasco Gianni, G., Tecchiati, U., Clavel, B., Lepetz, S., Davoudi, H., Mashkour, M., Berezina, N.Y., Stockhammer, P.W., Krause, J., Haak, W., Morales-Muñiz, A., Benecke, N., Hofreiter, M., Ludwig, A., Graphodatsky, A.S., Peters, J., Kiryushin, K.Y., Iderkhangai, T.-O., Bokovenko, N.A., Vasiliev, S.K., Seregin, N.N., Chugunov, K.V., Plasteeva, N.A., Baryshnikov, G.F., Petrova, E., Sablin, M., Ananyevskaya, E., Logvin, A., Shevnina, I., Logvin, V., Kalieva, S., Loman, V., Kukushkin, I., Merz, I., Merz, V., Sakenov, S., Varfolomeyev, V., Usmanova, E., Zaibert, V., Arbuckle, B., Belinskiy, A.B., Kalmykov,

A., Reinhold, S., Hansen, S., Yudin, A.I., Vybornov, A.A., Epimakhov, A., Berezina, N.S., Roslyakova, N., Kosintsev, P.A., Kuznetsov, P.F., Anthony, D., Kroonen, G.J., Kristiansen, K., Wincker, P., Outram, A., Orlando, L., 2021. The origins and spread of domestic horses from the Western Eurasian steppes. Nature 598, 634–640. https://doi.org/10.1038/s41586-021-04018-9
Liebert, A., López, S., Jones, B.L., Montalva, N., Gerbault, P., Lau, W., Thomas, M.G., Bradman, N., Maniatis, N., Swallow, D.M., 2017. World-wide distributions of lactase persistence alleles and the complex effects of recombination and selection. Hum Genet 136, 1445–1453. https://doi.org/10.1007/s00439-017-1847-y
Lipatov, M., Sanjeev, K., Patro, R., Veeramah, K.R., 2015. Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. https://doi.org/10.1101/023374
Lipson, M., Skoglund, P., Spriggs, M., Valentin, F., Bedford, S., Shing, R., Buckley, H., Phillip, I., Ward, G.K., Mallick, S., Rohland, N., Broomandkhoshbacht, N., Cheronet, O., Ferry, M., Harper, T.K., Michel, M., Oppenheimer, J., Sirak, K., Stewardson, K., Auckland, K., Hill, A.V.S., Maitland, K., Oppenheimer, S.J., Parks, T., Robson, K., Williams, T.N., Kennett, D.J., Mentzer, A.J., Pinhasi, R., Reich, D., 2018. Population Turnover in Remote Oceania Shortly after Initial Settlement. Current Biology 28, 1157-1165.e7. https://doi.org/10.1016/j.cub.2018.02.051
Littman, R.J., Littman, M.L., 1973. Galen and the Antonine Plague. The American Journal of Philology 94, 243–255. https://doi.org/10.2307/293979
Liu, C.-T., Monda, K.L., Taylor, K.C., Lange, L., Demerath, E.W., Palmas, W., Wojczynski, M.K., Ellis, J.C., Vitolins, M.Z., Liu, S., Papanicolaou, G.J., Irvin, M.R., Xue, L., Griffin, P.J., Nalls, M.A., Adeyemo, A., Liu, J., Li, G., Ruiz-Narvaez, E.A., Chen, W.-M., Chen, F., Henderson, B.E., Millikan, R.C., Ambrosone, C.B., Strom, S.S., Guo, X., Andrews, J.S., Sun, Y.V., Mosley, T.H., Yanek, L.R., Shriner, D., Haritunians, T., Rotter, J.I., Speliotes, E.K., Smith, M., Rosenberg, L., Mychaleckyj, J., Nayak, U., Spruill, I., Garvey, W.T., Pettaway, C., Nyante, S., Bandera, E.V., Britton, A.F., Zonderman, A.B., Rasmussen-Torvik, L.J., Chen, Y.-D.I., Ding, J., Lohman, K., Kritchevsky, S.B., Zhao, W., Peyser, P.A., Kardia, S.L.R., Kabagambe, E., Broeckel, U., Chen, G., Zhou, J., Wassertheil-Smoller, S., Neuhouser, M.L., Rampersaud, E., Psaty, B., Kooperberg, C., Manson, J.E., Kuller, L.H., Ochs-Balcom, H.M., Johnson, K.C., Sucheston, L., Ordovas, J.M., Palmer, J.R., Haiman, C.A., McKnight, B., Howard, B.V., Becker, D.M., Bielak, L.F., Liu, Y., Allison, M.A., Grant, S.F.A., Burke, G.L., Patel, S.R., Schreiner, P.J., Borecki, I.B., Evans, M.K., Taylor, H., Sale, M.M., Howard, V., Carlson, C.S., Rotimi, C.N., Cushman, M., Harris, T.B., Reiner, A.P., Cupples, L.A., North, K.E., Fox, C.S., 2013. Genome-wide association of body fat distribution in African ancestry populations suggests new loci. PLoS Genet 9, e1003681. https://doi.org/10.1371/journal.pgen.1003681
Llorente, M.G., Jones, E.R., Eriksson, A., Siska, V., Arthur, K.W., Arthur, J.W., Curtis, M.C., Stock, J.T., Coltorti, M., Pieruccini, P., Stretton, S., Brock, F., Higham, T., Park, Y., Hofreiter, M., Bradley, D.G., Bhak, J., Pinhasi, R., Manica, A., 2015. Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa. Science 350, 820–822. https://doi.org/10.1126/science.aad2879
Lombardo, U., Iriarte, J., Hilbert, L., Ruiz-Pérez, J., Capriles, J.M., Veit, H., 2020. Early Holocene crop cultivation and landscape modification in Amazonia. Nature 581, 190–193. https://doi.org/10.1038/s41586-020-2162-7
Lund, E.G., Kerr, T.A., Sakai, J., Li, W.P., Russell, D.W., 1998. cDNA cloning of mouse and human cholesterol 25-hydroxylases, polytopic membrane proteins that synthesize a potent oxysterol regulator of lipid metabolism. J Biol Chem 273, 34316–34327. https://doi.org/10.1074/jbc.273.51.34316

Malyarchuk, A.B., Andreeva, T.V., Kuznetsova, I.L., Kunizheva, S.S., Protasova, M.S., Uralsky, L.I., Tyazhelova, T.V., Gusev, F.E., Manakhov, A.D., Rogaev, E.I., 2022. Genomics of Ancient Pathogens: First Advances and Prospects. Biochemistry Moscow 87, 242–258. https://doi.org/10.1134/S0006297922030051
Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., Turner, D.J., 2010. Target-enrichment strategies for next-generation sequencing. Nat Methods 7, 111–118. https://doi.org/10.1038/nmeth.1419
Manchester, K., 1984. Tuberculosis and leprosy in antiquity: an interpretation. Medical History 28, 162–173. https://doi.org/10.1017/S0025727300035705
Marciniak, S., Prowse, T.L., Herring, D.A., Klunk, J., Kuch, M., Duggan, A.T., Bondioli, L., Holmes, E.C., Poinar, H.N., 2016. Plasmodium falciparum malaria in 1st-2nd century CE southern Italy. Curr Biol 26, R1220–R1222. https://doi.org/10.1016/j.cub.2016.10.016
Margaryan, A., Hansen, H.B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., Saag, L., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2018. Ancient pathogen DNA in human teeth and petrous bones. Ecology and Evolution 8, 3534–3542. https://doi.org/10.1002/ece3.3924
Margaryan, A., Lawson, D.J., Sikora, Martin, Racimo, F., Rasmussen, S., Moltke, I., Cassidy, L.M., Jørsboe, E., Ingason, A., Pedersen, M.W., Korneliussen, T., Wilhelmson, H., Buś, M.M., de Barros Damgaard, P., Martiniano, R., Renaud, G., Bhérer, C., Moreno-Mayar, J.V., Fotakis, A.K., Allen, M., Allmäe, R., Molak, M., Cappellini, E., Scorrano, G., McColl, H., Buzhilova, A., Fox, A., Albrechtsen, A., Schütz, B., Skar, B., Arcini, C., Falys, C., Jonson, C.H., Błaszczyk, D., Pezhemsky, D., Turner-Walker, G., Gestsdóttir, H., Lundstrøm, I., Gustin, I., Mainland, I., Potekhina, I., Muntoni, I.M., Cheng, J., Stenderup, J., Ma, J., Gibson, J., Peets, J., Gustafsson, J., Iversen, K.H., Simpson, L., Strand, L., Loe, L., Sikora, Maeve, Florek, M., Vretemark, M., Redknap, M., Bajka, M., Pushkina, T., Søvsø, M., Grigoreva, N., Christensen, T., Kastholm, O., Uldum, O., Favia, P., Holck, P., Sten, S., Arge, S.V., Ellingvåg, S., Moiseyev, V., Bogdanowicz, W., Magnusson, Y., Orlando, L., Pentz, P., Jessen, M.D., Pedersen, A., Collard, M., Bradley, D.G., Jørkov, M.L., Arneborg, J., Lynnerup, N., Price, N., Gilbert, M.T.P., Allentoft, M.E., Bill, J., Sindbæk, S.M., Hedeager, L., Kristiansen, K., Nielsen, R., Werge, T., Willerslev, E., 2020. Population genomics of the Viking world. Nature 585, 390–396. https://doi.org/10.1038/s41586-020-2688-8
Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L.I., Jarvie, T.P., Jirage, K.B., Kim, J.-B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380. https://doi.org/10.1038/nature03959
Maricic, T., Whitten, M., Pääbo, S., 2010. Multiplexed DNA Sequence Capture of Mitochondrial Genomes Using PCR Products. PLOS ONE 5, e14004. https://doi.org/10.1371/journal.pone.0014004
Massilani, D., Morley, M.W., Mentzer, S.M., Aldeias, V., Vernot, B., Miller, C., Stahlschmidt, M., Kozlikin, M.B., Shunkov, M.V., Derevianko, A.P., Conard, N.J., Wurz, S., Henshilwood, C.S., Vasquez, J., Essel, E., Nagel, S., Richter, J., Nickel, B., Roberts, R.G., Pääbo, S., Slon, V., Goldberg, P., Meyer, M., 2022. Microstratigraphic preservation of ancient faunal and hominin DNA in Pleistocene cave sediments. Proc Natl Acad Sci U S A 119, e2113666118. https://doi.org/10.1073/pnas.2113666118

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., Sirak, K., Gamba, C., Jones, E.R., Llamas, B., Dryomov, S., Pickrell, J., Arsuaga, J.L., de Castro, J.M.B., Carbonell, E., Gerritsen, F., Khokhlov, A., Kuznetsov, P., Lozano, M., Meller, H., Mochalov, O., Moiseyev, V., Guerra, M.A.R., Roodenberg, J., Vergès, J.M., Krause, J., Cooper, A., Alt, K.W., Brown, D., Anthony, D., Lalueza-Fox, C., Haak, W., Pinhasi, R., Reich, D., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528, 499–503. https://doi.org/10.1038/nature16152
McClean, P., Davison, S.M., 2003. Neonatal liver failure. Semin Neonatol 8, 393–401. https://doi.org/10.1016/S1084-2756(03)00095-2
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110
McTavish, E.J., Hillis, D.M., 2015. How do SNP ascertainment schemes and population demographics affect inferences about population history? BMC Genomics 16, 266. https://doi.org/10.1186/s12864-015-1469-5
Méndez-Sánchez, N., Villa, A.R., Chávez-Tapia, N.C., Ponciano-Rodriguez, G., Almeda-Valdés, P., González, D., Uribe, M., 2005. Trends in liver disease prevalence in Mexico from 2005 to 2050 through mortality data. Annals of Hepatology 4, 52–55. https://doi.org/10.1016/S1665-2681(19)32086-1
Mendum, T.A., Schuenemann, V.J., Roffey, S., Taylor, G.M., Wu, H., Singh, P., Tucker, K., Hinds, J., Cole, S.T., Kierzek, A.M., Nieselt, K., Krause, J., Stewart, G.R., 2014. Mycobacterium leprae genomes from a British medieval leprosy hospital: towards understanding an ancient epidemic. BMC Genomics 15, 270. https://doi.org/10.1186/1471-2164-15-270
Merrill, W.L., Hard, R.J., Mabry, J.B., Fritz, G.J., Adams, K.R., Roney, J.R., MacWilliams, A.C., 2009. The diffusion of maize to the southwestern United States and its impact. Proceedings of the National Academy of Sciences 106, 21019–21026. https://doi.org/10.1073/pnas.0906075106
Meyer, M., Arsuaga, J.-L., de Filippo, C., Nagel, S., Aximu-Petri, A., Nickel, B., Martínez, I., Gracia, A., de Castro, J.M.B., Carbonell, E., Viola, B., Kelso, J., Prüfer, K., Pääbo, S., 2016. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. Nature 531, 504–507. https://doi.org/10.1038/nature17405
Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010, pdb.prot5448. https://doi.org/10.1101/pdb.prot5448
Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., Sudmant, P.H., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, M., Green, R.E., Bryc, K., Briggs, A.W., Stenzel, U., Dabney, J., Shendure, J., Kitzman, J., Hammer, M.F., Shunkov, M.V., Derevianko, A.P., Patterson, N., Andrés, A.M., Eichler, E.E., Slatkin, M., Reich, D., Kelso, J., Pääbo, S., 2012. A High-Coverage Genome Sequence from an Archaic Denisovan Individual. Science 338, 222–226. https://doi.org/10.1126/science.1224344
Miller, W., Drautz, D.I., Ratan, A., Pusey, B., Qi, J., Lesk, A.M., Tomsho, L.P., Packard, M.D., Zhao, F., Sher, A., Tikhonov, A., Raney, B., Patterson, N., Lindblad-Toh, K., Lander, E.S., Knight, J.R., Irzyk, G.P., Fredrikson, K.M., Harkins, T.T., Sheridan, S., Pringle, T., Schuster, S.C., 2008. Sequencing the nuclear genome of the extinct woolly mammoth. Nature 456, 387–390. https://doi.org/10.1038/nature07446
Minnich, S.A., Rohde, H.N., 2007. A rationale for repression and/or loss of motility by

pathogenic Yersinia in the mammalian host. Adv Exp Med Biol 603, 298–310. https://doi.org/10.1007/978-0-387-72124-8_27
Moreira, D., 1998. Efficient removal of PCR inhibitors using agarose-embedded DNA preparations. Nucleic Acids Res 26, 3309–3310.
Mühlemann, B., Jones, T.C., Damgaard, P. de B., Allentoft, M.E., Shevnina, I., Logvin, A., Usmanova, E., Panyushkina, I.P., Boldgiv, B., Bazartseren, T., Tashbaeva, K., Merz, V., Lau, N., Smrčka, V., Voyakin, D., Kitov, E., Epimakhov, A., Pokutta, D., Vicze, M., Price, T.D., Moiseyev, V., Hansen, A.J., Orlando, L., Rasmussen, S., Sikora, M., Vinner, L., Osterhaus, A.D.M.E., Smith, D.J., Glebe, D., Fouchier, R.A.M., Drosten, C., Sjögren, K.-G., Kristiansen, K., Willerslev, E., 2018. Ancient hepatitis B viruses from the Bronze Age to the Medieval period. Nature 557, 418–423. https://doi.org/10.1038/s41586-018-0097-z
Murchie, T.J., Kuch, M., Duggan, A.T., Ledger, M.L., Roche, K., Klunk, J., Karpinski, E., Hackenberger, D., Sadoway, T., MacPhee, R., Froese, D., Poinar, H., 2021. Optimizing extraction and targeted capture of ancient environmental DNA for reconstructing past environments using the PalaeoChip Arctic-1.0 bait-set. Quaternary Research 99, 305–328. https://doi.org/10.1017/qua.2020.59
Nakatsuka, N., Luisi, P., Motti, J.M.B., Salemme, M., Santiago, F., D'Angelo del Campo, M.D., Vecchi, R.J., Espinosa-Parrilla, Y., Prieto, A., Adamski, N., Lawson, A.M., Harper, T.K., Culleton, B.J., Kennett, D.J., Lalueza-Fox, C., Mallick, S., Rohland, N., Guichón, R.A., Cabana, G.S., Nores, R., Reich, D., 2020. Ancient genomes in South Patagonia reveal population movements associated with technological shifts and geography. Nat Commun 11, 3868. https://doi.org/10.1038/s41467-020-17656-w
Narasimhan, V.M., Patterson, N., Moorjani, P., Rohland, N., Bernardos, R., Mallick, S., Lazaridis, I., Nakatsuka, N., Olalde, I., Lipson, M., Kim, A.M., Olivieri, L.M., Coppa, A., Vidale, M., Mallory, J., Moiseyev, V., Kitov, E., Monge, J., Adamski, N., Alex, N., Broomandkhoshbacht, N., Candilio, F., Callan, K., Cheronet, O., Culleton, B.J., Ferry, M., Fernandes, D., Freilich, S., Gamarra, B., Gaudio, D., Hajdinjak, M., Harney, É., Harper, T.K., Keating, D., Lawson, A.M., Mah, M., Mandl, K., Michel, M., Novak, M., Oppenheimer, J., Rai, N., Sirak, K., Slon, V., Stewardson, K., Zalzala, F., Zhang, Z., Akhatov, G., Bagashev, A.N., Bagnera, A., Baitanayev, B., Bendezu-Sarmiento, J., Bissembaev, A.A., Bonora, G.L., Chargynov, T.T., Chikisheva, T., Dashkovskiy, P.K., Derevianko, A., Dobeš, M., Douka, K., Dubova, N., Duisengali, M.N., Enshin, D., Epimakhov, A., Fribus, A.V., Fuller, D., Goryachev, A., Gromov, A., Grushin, S.P., Hanks, B., Judd, M., Kazizov, E., Khokhlov, A., Krygin, A.P., Kupriyanova, E., Kuznetsov, P., Luiselli, D., Maksudov, F., Mamedov, A.M., Mamirov, T.B., Meiklejohn, C., Merrett, D.C., Micheli, R., Mochalov, O., Mustafokulov, S., Nayak, A., Pettener, D., Potts, R., Razhev, D., Rykun, M., Sarno, S., Savenkova, T.M., Sikhymbaeva, K., Slepchenko, S.M., Soltobaev, O.A., Stepanova, N., Svyatko, S., Tabaldiev, K., Teschler-Nicola, M., Tishkin, A.A., Tkachev, V.V., Vasilyev, S., Velemínský, P., Voyakin, D., Yermolayeva, A., Zahir, M., Zubkov, V.S., Zubova, A., Shinde, V.S., Lalueza-Fox, C., Meyer, M., Anthony, D., Boivin, N., Thangaraj, K., Kennett, D.J., Frachetti, M., Pinhasi, R., Reich, D., 2019. The formation of human populations in South and Central Asia. Science 365, eaat7487. https://doi.org/10.1126/science.aat7487
Nieves-Colón, M.A., Ozga, A.T., Pestle, W.J., Cucina, A., Tiesler, V., Stanton, T.W., Stone, A.C., 2018. Comparison of two ancient DNA extraction protocols for skeletal remains from tropical environments. Am J Phys Anthropol 166, 824–836. https://doi.org/10.1002/ajpa.23472
Noonan, J.P., Hofreiter, M., Smith, D., Priest, J.R., Rohland, N., Rabeder, G., Krause, J., Detter, J.C., Pääbo, S., Rubin, E.M., 2005. Genomic sequencing of Pleistocene cave bears. Science 309, 597–599. https://doi.org/10.1126/science.1113485
Oliveira, G.C.X., Vieira, I.C.G., Tremblay, R.L., 2021. Editorial: Gene Flow Among Wild

and Domesticated Plant Species in the Neotropical Region. Frontiers in Ecology and Evolution 9.
Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., Warinner, C., 2021a. Ancient DNA analysis. Nat Rev Methods Primers 1, 1–26. https://doi.org/10.1038/s43586-020-00011-0
Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., Warinner, C., 2021b. Ancient DNA analysis. Nat Rev Methods Primers 1, 1–26. https://doi.org/10.1038/s43586-020-00011-0
Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M.V., Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A.S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J.V., Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J., Willerslev, E., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature 499, 74–78. https://doi.org/10.1038/nature12323
Ortner, D., 2003. Identification of Pathological Conditions in Human Skeletal Remains. Identification of Pathological Conditions in Human Skeletal Remains 1–645.
Ortner, D.J., 2003. Identification of Pathological Conditions in Human Skeletal Remains. Academic Press.
Outram, A.K., Stear, N.A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., Thorpe, N., Evershed, R.P., 2009. The Earliest Horse Harnessing and Milking. Science 323, 1332–1335. https://doi.org/10.1126/science.1168594
Pääbo, S., 1985. Molecular cloning of Ancient Egyptian mummy DNA. Nature 314, 644–645. https://doi.org/10.1038/314644a0
Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., 2004. Genetic analyses from ancient DNA. Annu Rev Genet 38, 645–679. https://doi.org/10.1146/annurev.genet.37.110801.143214
Paijmans, J.L.A., Fickel, J., Courtiol, A., Hofreiter, M., Förster, D.W., 2016. Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. Mol Ecol Resour 16, 42–55. https://doi.org/10.1111/1755-0998.12420
Parker, C., Rohrlach, A.B., Friederich, S., Nagel, S., Meyer, M., Krause, J., Bos, K.I., Haak, W., 2020. A systematic investigation of human DNA preservation in medieval skeletons. Sci Rep 10, 18225. https://doi.org/10.1038/s41598-020-75163-w
Patterson, N., Price, A.L., Reich, D., 2006. Population Structure and Eigenanalysis. PLOS Genetics 2, e190. https://doi.org/10.1371/journal.pgen.0020190
Pedersen, J.S., Valen, E., Velazquez, A.M.V., Parker, B.J., Rasmussen, M., Lindgreen, S., Lilje, B., Tobin, D.J., Kelly, T.K., Vang, S., Andersson, R., Jones, P.A., Hoover, C.A., Tikhonov, A., Prokhortchouk, E., Rubin, E.M., Sandelin, A., Gilbert, M.T.P., Krogh, A., Willerslev, E., Orlando, L., 2014. Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. Genome Res 24, 454–466. https://doi.org/10.1101/gr.163592.113
Pedersen, M.W., Ruter, A., Schweger, C., Friebe, H., Staff, R.A., Kjeldsen, K.K., Mendoza, M.L.Z., Beaudoin, A.B., Zutter, C., Larsen, N.K., Potter, B.A., Nielsen, R., Rainville, R.A., Orlando, L., Meltzer, D.J., Kjær, K.H., Willerslev, E., 2016. Postglacial viability and colonization in North America's ice-free corridor. Nature 537, 45–49. https://doi.org/10.1038/nature19085

Peyrégne, S., Peter, B.M., 2020. AuthentiCT: a model of ancient DNA damage to estimate the proportion of present-day DNA contamination. Genome Biology 21, 246. https://doi.org/10.1186/s13059-020-02123-y
Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., Anders, A., Pietrusewsky, M., Rollefson, G., Jovanovic, M., Trinhhoang, H., Bar-Oz, G., Oxenham, M., Matsumura, H., Hofreiter, M., 2015. Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. PLOS ONE 10, e0129102. https://doi.org/10.1371/journal.pone.0129102
Piperno, D.R., 2017. Assessing elements of an extended evolutionary synthesis for plant domestication and agricultural origin research. Proceedings of the National Academy of Sciences 114, 6429–6437. https://doi.org/10.1073/pnas.1703658114
Piperno, D.R., Holst, I., Moreno, J.E., Winter, K., 2019. Experimenting with domestication: Understanding macro- and micro-phenotypes and developmental plasticity in teosinte in its ancestral pleistocene and early holocene environments. Journal of Archaeological Science 108, 104970. https://doi.org/10.1016/j.jas.2019.05.006
Piperno, D.R., Ranere, A.J., Holst, I., Iriarte, J., Dickau, R., 2009. Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. Proceedings of the National Academy of Sciences 106, 5019–5024. https://doi.org/10.1073/pnas.0812525106
Poinar, H.N., Hofreiter, M., Spaulding, W.G., Martin, P.S., Stankiewicz, B.A., Bland, H., Evershed, R.P., Possnert, G., Pääbo, S., 1998. Molecular Coproscopy: Dung and Diet of the Extinct Ground Sloth Nothrotheriops shastensis. Science 281, 402–406. https://doi.org/10.1126/science.281.5375.402
Posth, C., Nakatsuka, N., Lazaridis, I., Skoglund, P., Mallick, S., Lamnidis, T.C., Rohland, N., Nägele, K., Adamski, N., Bertolini, E., Broomandkhoshbacht, N., Cooper, A., Culleton, B.J., Ferraz, T., Ferry, M., Furtwängler, A., Haak, W., Harkins, K., Harper, T.K., Hünemeier, T., Lawson, A.M., Llamas, B., Michel, M., Nelson, E., Oppenheimer, J., Patterson, N., Schiffels, S., Sedig, J., Stewardson, K., Talamo, S., Wang, C.-C., Hublin, J.-J., Hubbe, M., Harvati, K., Nuevo Delaunay, A., Beier, J., Francken, M., Kaulicke, P., Reyes-Centeno, H., Rademaker, K., Trask, W.R., Robinson, M., Gutierrez, S.M., Prufer, K.M., Salazar-García, D.C., Chim, E.N., Müller Plumm Gomes, L., Alves, M.L., Liryo, A., Inglez, M., Oliveira, R.E., Bernardo, D.V., Barioni, A., Wesolowski, V., Scheifler, N.A., Rivera, M.A., Plens, C.R., Messineo, P.G., Figuti, L., Corach, D., Scabuzzo, C., Eggers, S., DeBlasis, P., Reindel, M., Méndez, C., Politis, G., Tomasto-Cagigao, E., Kennett, D.J., Strauss, A., Fehren-Schmitz, L., Krause, J., Reich, D., 2018. Reconstructing the Deep Population History of Central and South America. Cell 175, 1185-1197.e22. https://doi.org/10.1016/j.cell.2018.10.027
Poullet, M., Orlando, L., 2020. Assessing DNA Sequence Alignment Methods for Characterizing Ancient Genomes and Methylomes. Frontiers in Ecology and Evolution 8.
Prado, M., Franco, C.M., Fente, C.A., Cepeda, A., Vázquez, B.I., Barros-Velázquez, J., 2002. Comparison of extraction methods for the recovery, amplification and species-specific analysis of DNA from bone and bone meals. Electrophoresis 23, 1005–1012. https://doi.org/10.1002/1522-2683(200204)23:7/8<1005::AID-ELPS1005>3.0.CO;2-1
Price, T.D., Bar-Yosef, O., 2011. The Origins of Agriculture: New Data, New Ideas: An Introduction to Supplement 4. Current Anthropology 52, S163–S174. https://doi.org/10.1086/659964
Prüfer, K., de Filippo, C., Grote, S., Mafessoni, F., Korlević, P., Hajdinjak, M., Vernot, B., Skov, L., Hsieh, P., Peyrégne, S., Reher, D., Hopfe, C., Nagel, S., Maricic, T., Fu, Q., Theunert, C., Rogers, R., Skoglund, P., Chintalapati, M., Dannemann, M., Nelson, B.J., Key, F.M., Rudan, P., Kućan, Ž., Gušić, I., Golovanova, L.V., Doronichev, V.B., Patterson, N., Reich, D., Eichler, E.E., Slatkin, M., Schierup, M.H., Andrés, A.M., Kelso, J., Meyer, M.,

Pääbo, S., 2017. A high-coverage Neandertal genome from Vindija Cave in Croatia. Science 358, 655–658. https://doi.org/10.1126/science.aao1887
Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., Li, H., Mallick, S., Dannemann, M., Fu, Q., Kircher, M., Kuhlwilm, M., Lachmann, M., Meyer, M., Ongyerth, M., Siebauer, M., Theunert, C., Tandon, A., Moorjani, P., Pickrell, J., Mullikin, J.C., Vohr, S.H., Green, R.E., Hellmann, I., Johnson, P.L.F., Blanche, H., Cann, H., Kitzman, J.O., Shendure, J., Eichler, E.E., Lein, E.S., Bakken, T.E., Golovanova, L.V., Doronichev, V.B., Shunkov, M.V., Derevianko, A.P., Viola, B., Slatkin, M., Reich, D., Kelso, J., Pääbo, S., 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature 505, 43–49. https://doi.org/10.1038/nature12886
Qin, P., Stoneking, M., 2015. Denisovan Ancestry in East Eurasian and Native American Populations. Molecular Biology and Evolution 32, 2665–2674. https://doi.org/10.1093/molbev/msv141
Quintana-Murci, L., 2019. Human Immunology through the Lens of Evolutionary Genetics. Cell 177, 184–199. https://doi.org/10.1016/j.cell.2019.02.033
Quintana-Murci, L., Clark, A.G., 2013. Population genetic tools for dissecting innate immunity in humans. Nat Rev Immunol 13, 280–293. https://doi.org/10.1038/nri3421
Racimo, F., Marnetto, D., Huerta-Sánchez, E., 2017. Signatures of Archaic Adaptive Introgression in Present-Day Human Populations. Mol Biol Evol 34, 296–317. https://doi.org/10.1093/molbev/msw216
Ralf, A., Montiel González, D., Zhong, K., Kayser, M., 2018. Yleaf: Software for Human Y-Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. Mol Biol Evol 35, 1291–1294. https://doi.org/10.1093/molbev/msy032
Ramos-Madrigal, J., Smith, B.D., Moreno-Mayar, J.V., Gopalakrishnan, S., Ross-Ibarra, J., Gilbert, M.T.P., Wales, N., 2016. Genome Sequence of a 5,310-Year-Old Maize Cob Provides Insights into the Early Stages of Maize Domestication. Curr Biol 26, 3195–3201. https://doi.org/10.1016/j.cub.2016.09.036
Rascovan, N., Sjögren, K.-G., Kristiansen, K., Nielsen, R., Willerslev, E., Desnues, C., Rasmussen, S., 2019. Emergence and Spread of Basal Lineages of Yersinia pestis during the Neolithic Decline. Cell 176, 295-305.e10. https://doi.org/10.1016/j.cell.2018.11.005
Rasmussen, M., Guo, X., Wang, Y., Lohmueller, K.E., Rasmussen, S., Albrechtsen, A., Skotte, L., Lindgreen, S., Metspalu, M., Jombart, T., Kivisild, T., Zhai, W., Eriksson, A., Manica, A., Orlando, L., De La Vega, F.M., Tridico, S., Metspalu, E., Nielsen, K., Ávila-Arcos, M.C., Moreno-Mayar, J.V., Muller, C., Dortch, J., Gilbert, M.T.P., Lund, O., Wesolowska, A., Karmin, M., Weinert, L.A., Wang, B., Li, J., Tai, S., Xiao, F., Hanihara, T., van Driem, G., Jha, A.R., Ricaut, F.-X., de Knijff, P., Migliano, A.B., Gallego Romero, I., Kristiansen, K., Lambert, D.M., Brunak, S., Forster, P., Brinkmann, B., Nehlich, O., Bunce, M., Richards, M., Gupta, R., Bustamante, C.D., Krogh, A., Foley, R.A., Lahr, M.M., Balloux, F., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Wang, J., Willerslev, E., 2011. An Aboriginal Australian genome reveals separate human dispersals into Asia. Science 334, 94–98. https://doi.org/10.1126/science.1211177
Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., Bertalan, M., Nielsen, K., Gilbert, M.T.P., Wang, Y., Raghavan, M., Campos, P.F., Kamp, H.M., Wilson, A.S., Gledhill, A., Tridico, S., Bunce, M., Lorenzen, E.D., Binladen, J., Guo, X., Zhao, J., Zhang, X., Zhang, H., Li, Z., Chen, M., Orlando, L., Kristiansen, K., Bak, M., Tommerup, N., Bendixen, C., Pierre, T.L., Grønnow, B., Meldgaard, M., Andreasen, C., Fedorova, S.A., Osipova, L.P., Higham, T.F.G., Ramsey, C.B., Hansen, T. v O., Nielsen, F.C., Crawford, M.H., Brunak, S., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Krogh, A., Wang, J., Willerslev, E., 2010. Ancient human genome

sequence of an extinct Palaeo-Eskimo. Nature 463, 757–762. https://doi.org/10.1038/nature08835
Rasmussen, S., Allentoft, M.E., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K.-G., Pedersen, A.G., Schubert, M., Van Dam, A., Kapel, C.M.O., Nielsen, H.B., Brunak, S., Avetisyan, P., Epimakhov, A., Khalyapin, M.V., Gnuni, A., Kriiska, A., Lasak, I., Metspalu, M., Moiseyev, V., Gromov, A., Pokutta, D., Saag, L., Varul, L., Yepiskoposyan, L., Sicheritz-Pontén, T., Foley, R.A., Lahr, M.M., Nielsen, R., Kristiansen, K., Willerslev, E., 2015a. Early Divergent Strains of Yersinia pestis in Eurasia 5,000 Years Ago. Cell 163, 571–582. https://doi.org/10.1016/j.cell.2015.10.009
Rasmussen, S., Allentoft, M.E., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K.-G., Pedersen, A.G., Schubert, M., Van Dam, A., Kapel, C.M.O., Nielsen, H.B., Brunak, S., Avetisyan, P., Epimakhov, A., Khalyapin, M.V., Gnuni, A., Kriiska, A., Lasak, I., Metspalu, M., Moiseyev, V., Gromov, A., Pokutta, D., Saag, L., Varul, L., Yepiskoposyan, L., Sicheritz-Pontén, T., Foley, R.A., Lahr, M.M., Nielsen, R., Kristiansen, K., Willerslev, E., 2015b. Early Divergent Strains of Yersinia pestis in Eurasia 5,000 Years Ago. Cell 163, 571–582. https://doi.org/10.1016/j.cell.2015.10.009
Raxworthy, C.J., Smith, B.T., 2021. Mining museums for historical DNA: advances and challenges in museomics. Trends in Ecology & Evolution 36, 1049–1060. https://doi.org/10.1016/j.tree.2021.07.009
Rebourg, C., Chastanet, M., Gouesnard, B., Welcker, C., Dubreuil, P., Charcosset, A., 2003. Maize introduction into Europe: the history reviewed in the light of molecular data. Theor Appl Genet 106, 895–903. https://doi.org/10.1007/s00122-002-1140-9
Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., Maricic, T., Good, J.M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E.E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M.V., Derevianko, A.P., Hublin, J.-J., Kelso, J., Slatkin, M., Pääbo, S., 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468, 1053–1060. https://doi.org/10.1038/nature09710
Reich, D., Patterson, N., Kircher, M., Delfin, F., Nandineni, M.R., Pugach, I., Ko, A.M.-S., Ko, Y.-C., Jinam, T.A., Phipps, M.E., Saitou, N., Wollstein, A., Kayser, M., Pääbo, S., Stoneking, M., 2011. Denisova Admixture and the First Modern Human Dispersals into Southeast Asia and Oceania. Am J Hum Genet 89, 516–528. https://doi.org/10.1016/j.ajhg.2011.09.005
Renaud, G., Slon, V., Duggan, A.T., Kelso, J., 2015. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biology 16, 224. https://doi.org/10.1186/s13059-015-0776-0
Resnick, D., Niwayama, G., 1987. Diagnosis of bone and joint disorders. Second edition.
Riehm, J.M., Tserennorov, D., Kiefer, D., Stuermer, I.W., Tomaso, H., Zöller, L., Otgonbaatar, D., Scholz, H.C., 2011. Yersinia pestis in Small Rodents, Mongolia. Emerg Infect Dis 17, 1320–1322. https://doi.org/10.3201/eid1707.100740
Roberts, C.A., Manchester, K., 2005. The Archaeology of Disease. Sutton.
Roca-Rada, X., Politis, G., Messineo, P.G., Scheifler, N., Scabuzzo, C., González, M., Harkins, K.M., Reich, D., Souilmi, Y., Teixeira, J.C., Llamas, B., Fehren-Schmitz, L., 2021. Ancient mitochondrial genomes from the Argentinian Pampas inform the early peopling of the Southern Cone of South America. iScience 24, 102553. https://doi.org/10.1016/j.isci.2021.102553
Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., Reich, D., 2015. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos Trans R Soc Lond B Biol Sci 370, 20130624. https://doi.org/10.1098/rstb.2013.0624
Rohland, N., Hofreiter, M., 2007a. Ancient DNA extraction from bones and teeth. Nat Protoc

2, 1756–1762. https://doi.org/10.1038/nprot.2007.247
Rohland, N., Hofreiter, M., 2007b. Comparison and optimization of ancient DNA extraction. BioTechniques 42, 343–352. https://doi.org/10.2144/000112383
Rohland, N., Siedel, H., Hofreiter, M., 2010. A rapid column-based ancient DNA extraction method for increased sample throughput. Mol Ecol Resour 10, 677–683. https://doi.org/10.1111/j.1755-0998.2009.02824.x
Ross, Z.P., Klunk, J., Fornaciari, G., Giuffra, V., Duchêne, S., Duggan, A.T., Poinar, D., Douglas, M.W., Eden, J.-S., Holmes, E.C., Poinar, H.N., 2018. The paradox of HBV evolution as revealed from a 16th century mummy. PLOS Pathogens 14, e1006750. https://doi.org/10.1371/journal.ppat.1006750
Rothschild, B.M., Martin, L.D., Lev, G., Bercovier, H., Bar-Gal, G.K., Greenblatt, C., Donoghue, H., Spigelman, M., Brittain, D., 2001. Mycobacterium tuberculosis Complex DNA from an Extinct Bison Dated 17,000 Years before the Present. Clinical Infectious Diseases 33, 305–311. https://doi.org/10.1086/321886
Rouf Shah, T., Prasad, K., Kumar, P., 2016. Maize—A potential source of human nutrition and health: A review. Cogent Food & Agriculture 2, 1166995. https://doi.org/10.1080/23311932.2016.1166995
Ruppert, K.M., Kline, R.J., Rahman, M.S., 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation 17, e00547. https://doi.org/10.1016/j.gecco.2019.e00547
Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., Pääbo, S., 2012. Temporal Patterns of Nucleotide Misincorporations and DNA Fragmentation in Ancient DNA. PLOS ONE 7, e34131. https://doi.org/10.1371/journal.pone.0034131
Scheib, C.L., Hui, R., D'Atanasio, E., Wohns, A.W., Inskip, S.A., Rose, A., Cessford, C., O'Connell, T.C., Robb, J.E., Evans, C., Patten, R., Kivisild, T., 2019. East Anglian early Neolithic monument burial linked to contemporary Megaliths. Annals of Human Biology 46, 145–149. https://doi.org/10.1080/03014460.2019.1623912
Schlebusch, C.M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A.R., Vicente, M., Steyn, M., Soodyall, H., Lombard, M., Jakobsson, M., 2017. Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. Science 358, 652–655. https://doi.org/10.1126/science.aao6266
Schubert, M., Ermini, L., Sarkissian, C.D., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, R., Kircher, M., McCue, M., Willerslev, E., Orlando, L., 2014. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat Protoc 9, 1056–1082. https://doi.org/10.1038/nprot.2014.063
Schuenemann, V.J., Avanzi, C., Krause-Kyora, B., Seitz, A., Herbig, A., Inskip, S., Bonazzi, M., Reiter, E., Urban, C., Pedersen, D.D., Taylor, G.M., Singh, P., Stewart, G.R., Velemínský, P., Likovsky, J., Marcsik, A., Molnár, E., Pálfi, G., Mariotti, V., Riga, A., Belcastro, M.G., Boldsen, J.L., Nebel, A., Mays, S., Donoghue, H.D., Zakrzewski, S., Benjak, A., Nieselt, K., Cole, S.T., Krause, J., 2018a. Ancient genomes reveal a high diversity of Mycobacterium leprae in medieval Europe. PLOS Pathogens 14, e1006997. https://doi.org/10.1371/journal.ppat.1006997
Schuenemann, V.J., Lankapalli, A.K., Barquera, R., Nelson, E.A., Hernández, D.I., Alonzo, V.A., Bos, K.I., Morfín, L.M., Herbig, A., Krause, J., 2018b. Historic Treponema pallidum genomes from Colonial Mexico retrieved from archaeological remains. PLOS Neglected Tropical Diseases 12, e0006447. https://doi.org/10.1371/journal.pntd.0006447
Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jäger, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J.L., Kjellström, A., Wu, H.,

Stewart, G.R., Taylor, G.M., Bauer, P., Lee, O.Y.-C., Wu, H.H.T., Minnikin, D.E., Besra, G.S., Tucker, K., Roffey, S., Sow, S.O., Cole, S.T., Nieselt, K., Krause, J., 2013. Genome-Wide Comparison of Medieval and Modern Mycobacterium leprae. Science 341, 179–183. https://doi.org/10.1126/science.1238286
Seersholm, F.V., Werndly, D.J., Grealy, A., Johnson, T., Keenan Early, E.M., Lundelius, E.L., Winsborough, B., Farr, G.E., Toomey, R., Hansen, A.J., Shapiro, B., Waters, M.R., McDonald, G., Linderholm, A., Stafford, T.W., Bunce, M., 2020. Rapid range shifts and megafaunal extinctions associated with late Pleistocene climate change. Nat Commun 11, 2770. https://doi.org/10.1038/s41467-020-16502-3
Seguin-Orlando, A., Costedoat, C., Der Sarkissian, C., Tzortzis, S., Kamel, C., Telmon, N., Dalén, L., Thèves, C., Signoli, M., Orlando, L., 2021a. No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. iScience 24, 102383. https://doi.org/10.1016/j.isci.2021.102383
Seguin-Orlando, A., Costedoat, C., Der Sarkissian, C., Tzortzis, S., Kamel, C., Telmon, N., Dalén, L., Thèves, C., Signoli, M., Orlando, L., 2021b. No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. iScience 24, 102383. https://doi.org/10.1016/j.isci.2021.102383
Seguin-Orlando, A., Donat, R., Der Sarkissian, C., Southon, J., Thèves, C., Manen, C., Tchérémissinoff, Y., Crubézy, E., Shapiro, B., Deleuze, J.-F., Dalén, L., Guilaine, J., Orlando, L., 2021c. Heterogeneous Hunter-Gatherer and Steppe-Related Ancestries in Late Neolithic and Bell Beaker Genomes from Present-Day France. Current Biology 31, 1072-1083.e10. https://doi.org/10.1016/j.cub.2020.12.015
Seguin-Orlando, A., Gamba, C., Sarkissian, C.D., Ermini, L., Louvel, G., Boulygina, E., Sokolov, A., Nedoluzhko, A., Lorenzen, E.D., Lopez, P., McDonald, H.G., Scott, E., Tikhonov, A., Stafford, Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Shapiro, B., Willerslev, E., Prokhortchouk, E., Orlando, L., 2015. Pros and cons of methylation-based enrichment methods for ancient DNA. Sci Rep 5, 11826. https://doi.org/10.1038/srep11826
Seguin-Orlando, A., Korneliussen, T.S., Sikora, M., Malaspinas, A.-S., Manica, A., Moltke, I., Albrechtsen, A., Ko, A., Margaryan, A., Moiseyev, V., Goebel, T., Westaway, M., Lambert, D., Khartanovich, V., Wall, J.D., Nigst, P.R., Foley, R.A., Lahr, M.M., Nielsen, R., Orlando, L., Willerslev, E., 2014. Genomic structure in Europeans dating back at least 36,200 years. Science 346, 1113–1118. https://doi.org/10.1126/science.aaa0114
Seguin-Orlando, A., Schubert, M., Clary, J., Stagegaard, J., Alberdi, M.T., Prado, J.L., Prieto, A., Willerslev, E., Orlando, L., 2013. Ligation Bias in Illumina Next-Generation DNA Libraries: Implications for Sequencing Ancient Genomes. PLOS ONE 8, e78575. https://doi.org/10.1371/journal.pone.0078575
Shibata, N., Kawarai, T., Lee, J.H., Lee, H.-S., Shibata, E., Sato, C., Liang, Y., Duara, R., Mayeux, R.P., St George-Hyslop, P.H., Rogaeva, E., 2006. Association studies of cholesterol metabolism genes (CH25H, ABCA1 and CH24H) in Alzheimer's disease. Neurosci Lett 391, 142–146. https://doi.org/10.1016/j.neulet.2005.08.048
Signoli, M., 2022. History of the plague of 1720-1722, in Marseille. La Presse Médicale 104138. https://doi.org/10.1016/j.lpm.2022.104138
Signoli, M., Bello, S., Dutour, O., 1998. La rechute épidémique de la grande peste de Marseille [mai-juillet 1722] : Le charnier de l'observance.
Sikora, M., Seguin-Orlando, A., Sousa, V.C., Albrechtsen, A., Korneliussen, T., Ko, A., Rasmussen, S., Dupanloup, I., Nigst, P.R., Bosch, M.D., Renaud, G., Allentoft, M.E., Margaryan, A., Vasilyev, S.V., Veselovskaya, E.V., Borutskaya, S.B., Deviese, T., Comeskey, D., Higham, T., Manica, A., Foley, R., Meltzer, D.J., Nielsen, R., Excoffier, L., Mirazon Lahr, M., Orlando, L., Willerslev, E., 2017. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. Science 358, 659–662.

https://doi.org/10.1126/science.aao1807
Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., Jakobsson, M., 2014. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proceedings of the National Academy of Sciences 111, 2229–2234. https://doi.org/10.1073/pnas.1318934111
Skoglund, P., Storå, J., Götherström, A., Jakobsson, M., 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. Journal of Archaeological Science 40, 4477–4482. https://doi.org/10.1016/j.jas.2013.07.004
Slon, V., Hopfe, C., Weiß, C.L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., Rosas, A., Soressi, M., Knul, M.V., Miller, R., Stewart, J.R., Derevianko, A.P., Jacobs, Z., Li, B., Roberts, R.G., Shunkov, M.V., de Lumley, H., Perrenoud, C., Gušić, I., Kućan, Ž., Rudan, P., Aximu-Petri, A., Essel, E., Nagel, S., Nickel, B., Schmidt, A., Prüfer, K., Kelso, J., Burbano, H.A., Pääbo, S., Meyer, M., 2017. Neandertal and Denisovan DNA from Pleistocene sediments. Science 356, 605–608. https://doi.org/10.1126/science.aam9695
Smith, N.H., Hewinson, R.G., Kremer, K., Brosch, R., Gordon, S.V., 2009. Myths and misconceptions: the origin and evolution of Mycobacterium tuberculosis. Nat Rev Microbiol 7, 537–544. https://doi.org/10.1038/nrmicro2165
Spigelman, M., Lemma, E., 1993. The use of the polymerase chain reaction (PCR) to detect Mycobacterium tuberculosis in ancient skeletons. International Journal of Osteoarchaeology 3, 137–143. https://doi.org/10.1002/oa.1390030211
Spyrou, M.A., Bos, K.I., Herbig, A., Krause, J., 2019a. Ancient pathogen genomics as an emerging tool for infectious disease research. Nat Rev Genet 20, 323–340. https://doi.org/10.1038/s41576-019-0119-1
Spyrou, M.A., Keller, M., Tukhbatova, R.I., Scheib, C.L., Nelson, E.A., Andrades Valtueña, A., Neumann, G.U., Walker, D., Alterauge, A., Carty, N., Cessford, C., Fetz, H., Gourvennec, M., Hartle, R., Henderson, M., von Heyking, K., Inskip, S.A., Kacki, S., Key, F.M., Knox, E.L., Later, C., Maheshwari-Aplin, P., Peters, J., Robb, J.E., Schreiber, J., Kivisild, T., Castex, D., Lösch, S., Harbeck, M., Herbig, A., Bos, K.I., Krause, J., 2019b. Phylogeography of the second plague pandemic revealed through analysis of historical Yersinia pestis genomes. Nat Commun 10, 4470. https://doi.org/10.1038/s41467-019-12154-0
Spyrou, M.A., Tukhbatova, R.I., Wang, C.-C., Valtueña, A.A., Lankapalli, A.K., Kondrashin, V.V., Tsybin, V.A., Khokhlov, A., Kühnert, D., Herbig, A., Bos, K.I., Krause, J., 2018. Analysis of 3800-year-old Yersinia pestis genomes suggests Bronze Age origin for bubonic plague. Nat Commun 9, 2234. https://doi.org/10.1038/s41467-018-04550-9
Staller, J.E., Tykot, R.H., Benz, B.F. (Eds.), 2006. Histories of maize: multidisciplinary approaches to the prehistory, linguistics, biogeography, domestication, and evolution of maize. Elsevier Academic Press, Amsterdam ; Boston.
Steinbock, R.T., 1978. Paleopathological diagnosis and interpretation. Bone disease in ancient human populations. Medical History 22, 220–220. https://doi.org/10.1017/S0025727300032610
Straube, N., Lyra, M.L., Paijmans, J.L.A., Preick, M., Basler, N., Penner, J., Rödel, M.-O., Westbury, M.V., Haddad, C.F.B., Barlow, A., Hofreiter, M., 2021. Successful application of ancient DNA extraction and library construction protocols to museum wet collection specimens. Molecular Ecology Resources 21, 2299–2315. https://doi.org/10.1111/1755-0998.13433
Suchan, T., Chauvey, L., Poullet, M., Tonasso-Calvière, L., Schiavinato, S., Clavel, P., Clavel, B., Lepetz, S., Seguin-Orlando, A., Orlando, L., 2022a. Assessing the impact of USER-treatment on hyRAD capture applied to ancient DNA. Molecular Ecology Resources 22, 2262–2274. https://doi.org/10.1111/1755-0998.13619
Suchan, T., Kusliy, M.A., Khan, N., Chauvey, L., Tonasso-Calvière, L., Schiavinato, S.,

Southon, J., Keller, M., Kitagawa, K., Krause, J., Bessudnov, A.N., Bessudnov, A.A., Graphodatsky, A.S., Valenzuela-Lamas, S., Wilczyński, J., Pospuła, S., Tunia, K., Nowak, M., Moskal-delHoyo, M., Tishkin, A.A., Pryor, A.J.E., Outram, A.K., Orlando, L., 2022b. Performance and automation of ancient DNA capture with RNA hyRAD probes. Molecular Ecology Resources 22, 891–907. https://doi.org/10.1111/1755-0998.13518
Suchan, T., Pitteloud, C., Gerasimova, N.S., Kostikova, A., Schmid, S., Arrigo, N., Pajkovic, M., Ronikier, M., Alvarez, N., 2016. Hybridization Capture Using RAD Probes (hyRAD), a New Tool for Performing Genomic Analyses on Collection Specimens. PLOS ONE 11, e0151651. https://doi.org/10.1371/journal.pone.0151651
Sun, Y.-C., Jarrett, C.O., Bosio, C.F., Hinnebusch, B.J., 2014. Retracing the Evolutionary Path that Led to Flea-Borne Transmission of Yersinia pestis. Cell Host & Microbe 15, 578–586. https://doi.org/10.1016/j.chom.2014.04.003
Taubenberger, J.K., Reid, A.H., Lourens, R.M., Wang, R., Jin, G., Fanning, T.G., 2005. Characterization of the 1918 influenza virus polymerase genes. Nature 437, 889–893. https://doi.org/10.1038/nature04230
Treille, G.-F., Yersin, A., 1894. La peste bubonique à Hong Kong. Presented at the VIIIe Congrès international d'hygiène et de démographie, Pesti konyvnyomda-részvénytársaság, p. 310.
Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solórzano, A., Swayne, D.E., Cox, N.J., Katz, J.M., Taubenberger, J.K., Palese, P., García-Sastre, A., 2005. Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus. Science 310, 77–80. https://doi.org/10.1126/science.1119392
Tzortzis, S., Signoli, M., 2016. Characterization of the Funeral Groups Associated with Plague Epidemics, in: Paleomicrobiology of Humans. John Wiley & Sons, Ltd, pp. 13–20. https://doi.org/10.1128/9781555819170.ch2
Vågene, Å.J., Herbig, A., Campana, M.G., Robles García, N.M., Warinner, C., Sabin, S., Spyrou, M.A., Andrades Valtueña, A., Huson, D., Tuross, N., Bos, K.I., Krause, J., 2018. Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. Nat Ecol Evol 2, 520–528. https://doi.org/10.1038/s41559-017-0446-6
Vallebueno-Estrada, M., Rodríguez-Arévalo, I., Rougon-Cardoso, A., Martínez González, J., García Cook, A., Montiel, R., Vielle-Calzada, J.-P., 2016. The earliest maize from San Marcos Tehuacán is a partial domesticate with genomic evidence of inbreeding. Proceedings of the National Academy of Sciences 113, 14151–14156. https://doi.org/10.1073/pnas.1609701113
van der Valk, T., Pečnerová, P., Díez-del-Molino, D., Bergström, A., Oppenheimer, J., Hartmann, S., Xenikoudakis, G., Thomas, J.A., Dehasque, M., Sağlıcan, E., Fidan, F.R., Barnes, I., Liu, S., Somel, M., Heintzman, P.D., Nikolskiy, P., Shapiro, B., Skoglund, P., Hofreiter, M., Lister, A.M., Götherström, A., Dalén, L., 2021. Million-year-old DNA sheds light on the genomic history of mammoths. Nature 591, 265–269. https://doi.org/10.1038/s41586-021-03224-9
van Dijk, E.L., Jaszczyszyn, Y., Naquin, D., Thermes, C., 2018. The Third Revolution in Sequencing Technology. Trends in Genetics 34, 666–681. https://doi.org/10.1016/j.tig.2018.05.008
van Heerwaarden, J., Doebley, J., Briggs, W.H., Glaubitz, J.C., Goodman, M.M., de Jesus Sanchez Gonzalez, J., Ross-Ibarra, J., 2011. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. Proceedings of the National Academy of Sciences 108, 1088–1092. https://doi.org/10.1073/pnas.1013011108
Vilgalys, T.P., Klunk, J., Demeure, C.E., Cheng, X., Shiratori, M., Madej, J., Beau, R., Elli, D., Patino, M.I., Redfern, R., DeWitte, S.N., Gamble, J.A., Boldsen, J.L., Carmichael, A., Varlik, N., Eaton, K., Grenier, J.-C., Golding, G.B., Devault, A., Rouillard, J.-M., Yotova, V.,

Sindeaux, R., Ye, C.J., Bikaran, M., Dumaine, A., Brinkworth, J.F., Missiakas, D., Rouleau, G.A., Steinrücken, M., Pizarro-Cerdá, J., Poinar, H.N., Barreiro, L.B., 2023. Reply to Barton et al: signatures of natural selection during the Black Death. https://doi.org/10.1101/2023.04.06.535944
Wagner, D.M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J.W., Enk, J., Birdsell, D.N., Kuch, M., Lumibao, C., Poinar, D., Pearson, T., Fourment, M., Golding, B., Riehm, J.M., Earn, D.J.D., Dewitte, S., Rouillard, J.-M., Grupe, G., Wiechmann, I., Bliska, J.B., Keim, P.S., Scholz, H.C., Holmes, E.C., Poinar, H., 2014. Yersinia pestis and the plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect Dis 14, 319–326. https://doi.org/10.1016/S1473-3099(13)70323-2
Wagner, S., Lagane, F., Seguin-Orlando, A., Schubert, M., Leroy, T., Guichoux, E., Chancerel, E., Bech-Hebelstrup, I., Bernard, V., Billard, C., Billaud, Y., Bolliger, M., Croutsch, C., Čufar, K., Eynaud, F., Heussner, K.U., Köninger, J., Langenegger, F., Leroy, F., Lima, C., Martinelli, N., Momber, G., Billamboz, A., Nelle, O., Palomo, A., Piqué, R., Ramstein, M., Schweichel, R., Stäuble, H., Tegel, W., Terradas, X., Verdin, F., Plomion, C., Kremer, A., Orlando, L., 2018. High-Throughput DNA sequencing of ancient wood. Molecular Ecology 27, 1138–1154. https://doi.org/10.1111/mec.14514
Wang, H., Avican, K., Fahlgren, A., Erttmann, S.F., Nuss, A.M., Dersch, P., Fallman, M., Edgren, T., Wolf-Watz, H., 2016. Increased plasmid copy number is essential for Yersinia T3SS function and virulence. Science 353, 492–495. https://doi.org/10.1126/science.aaf7501
Wang, K., Goldstein, S., Bleasdale, M., Clist, B., Bostoen, K., Bakwa-Lufu, P., Buck, L.T., Crowther, A., Dème, A., McIntosh, R.J., Mercader, J., Ogola, C., Power, R.C., Sawchuk, E., Robertshaw, P., Wilmsen, E.N., Petraglia, M., Ndiema, E., Manthi, F.K., Krause, J., Roberts, P., Boivin, N., Schiffels, S., 2020. Ancient genomes reveal complex patterns of population movement, interaction, and replacement in sub-Saharan Africa. Science Advances 6, eaaz0183. https://doi.org/10.1126/sciadv.aaz0183
Wang, L., Beissinger, T.M., Lorant, A., Ross-Ibarra, C., Ross-Ibarra, J., Hufford, M.B., 2017. The interplay of demography and selection during maize domestication and expansion. Genome Biol 18, 215. https://doi.org/10.1186/s13059-017-1346-4
Wang, Y., Pedersen, M.W., Alsos, I.G., De Sanctis, B., Racimo, F., Prohaska, A., Coissac, E., Owens, H.L., Merkel, M.K.F., Fernandez-Guerra, A., Rouillard, A., Lammers, Y., Alberti, A., Denoeud, F., Money, D., Ruter, A.H., McColl, H., Larsen, N.K., Cherezova, A.A., Edwards, M.E., Fedorov, G.B., Haile, J., Orlando, L., Vinner, L., Korneliussen, T.S., Beilman, D.W., Bjørk, A.A., Cao, J., Dockter, C., Esdale, J., Gusarova, G., Kjeldsen, K.K., Mangerud, J., Rasic, J.T., Skadhauge, B., Svendsen, J.I., Tikhonov, A., Wincker, P., Xing, Y., Zhang, Y., Froese, D.G., Rahbek, C., Bravo, D.N., Holden, P.B., Edwards, N.R., Durbin, R., Meltzer, D.J., Kjær, K.H., Möller, P., Willerslev, E., 2021. Late Quaternary dynamics of Arctic biota from ancient environmental genomics. Nature 600, 86–92. https://doi.org/10.1038/s41586-021-04016-x
Warinner, C., García, N.R., Spores, R., Tuross, N., 2012. Disease, Demography, and Diet in Early Colonial New Spain: Investigation of a Sixteenth-Century Mixtec Cemetery at Teposcolula Yucundaa. Latin American Antiquity 23, 467–489. https://doi.org/10.7183/1045-6635.23.4.467
Warinner, C., Herbig, A., Mann, A., Fellows Yates, J.A., Weiß, C.L., Burbano, H.A., Orlando, L., Krause, J., 2017. A Robust Framework for Microbial Archaeology. Annu Rev Genomics Hum Genet 18, 321–356. https://doi.org/10.1146/annurev-genom-091416-035526
Warinner, C., Rodrigues, J.F.M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R.Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H.U., Salazar-García, D.C., Eppler, E., Seiler, R., Hansen, L.H., Castruita, J.A.S., Barkow-Oesterreicher, S., Teoh, K.Y., Kelstrup, C.D., Olsen, J.V., Nanni, P., Kawai, T., Willerslev,

E., von Mering, C., Lewis, C.M., Collins, M.J., Gilbert, M.T.P., Rühli, F., Cappellini, E., 2014. Pathogens and host immunity in the ancient human oral cavity. Nat Genet 46, 336–344. https://doi.org/10.1038/ng.2906
Warner, T.G., Dambach, L.M., Shin, J.H., O'Brien, J.S., 1980. Separation and characterization of the acid lipase and neutral esterases from human liver. Am J Hum Genet 32, 869–879.
Wilbur, A.K., Bouwman, A.S., Stone, A.C., Roberts, C.A., Pfister, L.-A., Buikstra, J.E., Brown, T.A., 2009. Deficiencies and challenges in the study of ancient tuberculosis DNA. Journal of Archaeological Science 36, 1990–1997. https://doi.org/10.1016/j.jas.2009.05.020
Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M.B., Brand, T.B., Hofreiter, M., Bunce, M., Poinar, H.N., Dahl-Jensen, D., Johnsen, S., Steffensen, J.P., Bennike, O., Schwenninger, J.-L., Nathan, R., Armitage, S., de Hoog, C.-J., Alfimov, V., Christl, M., Beer, J., Muscheler, R., Barker, J., Sharp, M., Penkman, K.E.H., Haile, J., Taberlet, P., Gilbert, M.T.P., Casoli, A., Campani, E., Collins, M.J., 2007. Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. Science 317, 111–114. https://doi.org/10.1126/science.1141758
Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M.E., Lorenzen, E.D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L.S., Pearman, P.B., Cheddadi, R., Murray, D., Bråthen, K.A., Yoccoz, N., Binney, H., Cruaud, C., Wincker, P., Goslar, T., Alsos, I.G., Bellemain, E., Brysting, A.K., Elven, R., Sønstebø, J.H., Murton, J., Sher, A., Rasmussen, M., Rønn, R., Mourier, T., Cooper, A., Austin, J., Möller, P., Froese, D., Zazula, G., Pompanon, F., Rioux, D., Niderkorn, V., Tikhonov, A., Savvinov, G., Roberts, R.G., MacPhee, R.D.E., Gilbert, M.T.P., Kjær, K.H., Orlando, L., Brochmann, C., Taberlet, P., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51. https://doi.org/10.1038/nature12921
Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., Cooper, A., 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science 300, 791–795. https://doi.org/10.1126/science.1084114
Wilson, P.H., 2009. Europe's Tragedy: A New History of the Thirty Years War. Penguin UK.
Winther Pedersen, M., Overballe-Petersen, S., Ermini, L., Sarkissian, C.D., Haile, J., Hellstrom, M., Spens, J., Thomsen, P.F., Bohmann, K., Cappellini, E., Schnell, I.B., Wales, N.A., Carøe, C., Campos, P.F., Schmidt, A.M.Z., Gilbert, M.T.P., Hansen, A.J., Orlando, L., Willerslev, E., 2015. Ancient and modern environmental DNA. Philos Trans R Soc Lond B Biol Sci 370, 20130383. https://doi.org/10.1098/rstb.2013.0383
Wood, J.W., Milner, G.R., Harpending, H.C., Weiss, K.M., Cohen, M.N., Eisenberg, L.E., Hutchinson, D.L., Jankauskas, R., Cesnys, G., Česnys, G., Katzenberg, M.A., Lukacs, J.R., McGrath, J.W., Roth, E.A., Ubelaker, D.H., Wilkinson, R.G., 1992. The Osteological Paradox: Problems of Inferring Prehistoric Health from Skeletal Samples [and Comments and Reply]. Current Anthropology 33, 343–370.
Worobey, M., Watts, T.D., McKay, R.A., Suchard, M.A., Granade, T., Teuwen, D.E., Koblin, B.A., Heneine, W., Lemey, P., Jaffe, H.W., 2016. 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. Nature 539, 98–101. https://doi.org/10.1038/nature19827
Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., Saunders, S.R., 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. Am J Phys Anthropol 105, 539–543. https://doi.org/10.1002/(SICI)1096-8644(199804)105:4<539::AID-AJPA10>3.0.CO;2-1
Yang, M.A., Gao, X., Theunert, C., Tong, H., Aximu-Petri, A., Nickel, B., Slatkin, M., Meyer, M., Pääbo, S., Kelso, J., Fu, Q., 2017. 40,000-Year-Old Individual from Asia Provides

Insight into Early Population Structure in Eurasia. Current Biology 27, 3202-3208.e9. https://doi.org/10.1016/j.cub.2017.09.030
Yang, R., Atkinson, S., Chen, Z., Cui, Y., Du, Z., Han, Y., Sebbane, F., Slavin, P., Song, Y., Yan, Y., Wu, Y., Xu, L., Zhang, C., Zhang, Y., Hinnebusch, B.J., Stenseth, N.C., Motin, V.L., 2023. *Yersinia pestis* and Plague: Some Knowns and Unknowns. Zoonoses 3. https://doi.org/10.15212/ZOONOSES-2022-0040
Yu, T., Keto-Timonen, R., Jiang, X., Virtanen, J.-P., Korkeala, H., 2019. Insights into the Phylogeny and Evolution of Cold Shock Proteins: From Enteropathogenic Yersinia and Escherichia coli to Eubacteria. International Journal of Molecular Sciences 20, 4059. https://doi.org/10.3390/ijms20164059
Zhou, Z., Lundstrøm, I., Tran-Dien, A., Duchêne, S., Alikhan, N.-F., Sergeant, M.J., Langridge, G., Fotakis, A.K., Nair, S., Stenøien, H.K., Hamre, S.S., Casjens, S., Christophersen, A., Quince, C., Thomson, N.R., Weill, F.-X., Ho, S.Y.W., Gilbert, M.T.P., Achtman, M., 2018. Pan-genome Analysis of Ancient and Modern Salmonella enterica Demonstrates Genomic Stability of the Invasive Para C Lineage for Millennia. Current Biology 28, 2420-2428.e10. https://doi.org/10.1016/j.cub.2018.05.058
Ziesemer, K.A., Mann, A.E., Sankaranarayanan, K., Schroeder, H., Ozga, A.T., Brandt, B.W., Zaura, E., Waters-Rist, A., Hoogland, M., Salazar-García, D.C., Aldenderfer, M., Speller, C., Hendy, J., Weston, D.A., MacDonald, S.J., Thomas, G.H., Collins, M.J., Lewis, C.M., Hofman, C., Warinner, C., 2015. Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. Sci Rep 5, 16498. https://doi.org/10.1038/srep16498
Zink, A., HAAS, C.J., REISCHL, U., SZEIMIES, U., NERLICH, A.G., 2001. Molecular analysis of skeletal tuberculosis in an ancient Egyptian population. Journal of Medical Microbiology 50, 355–366. https://doi.org/10.1099/0022-1317-50-4-355