

INFORME FINAL
ESTUDIO PALEOGENÓMICO DE LA COEVOLUCIÓN DEL
FRIJOL Y SUS SIMBIONTES DURANTE LA
DOMESTICACIÓN

Informe final
Exportación de muestras arqueológicas con fines de investigación
científica

Presentado ante la Dirección de Calificación de Intervenciones Arqueológicas
Ministerio de Cultura

Por
PhD. Jazmín Ramos Madrigal
Elsa Peters Ruiz De Chávez
Mag. George Edward Chauca Iparraguirre

Lima, 9 de septiembre de 2026

Contenido

1.- Antecedentes	2
2.- Problemática	3
3.- Finalidad de los análisis	4
4.- Lugar de procedencia de las muestras.....	4
5.- Datos cuantitativos y cualitativos de las muestras.....	13
6.- Metodología	17
6.1.- Extracción de ADN	17
6.2.- Preparación de bibliotecas de secuenciación y secuenciación	19
6.3.- Procesamiento bioinformático de los datos de secuenciación.....	20
6.4.- Preparación del panel de genomas modernos de referencia	21
6.5.- Llamado de variantes	22
6.6.- Análisis de ancestría de los genomas antiguos	23
7.- Resultados	23
7.1.- Preservación de ADN antiguo en las semillas arqueológicas.....	23
7.2.- Secuenciación de las semillas arqueológicas	26
7.3.- Ancestría genómica de las semillas antiguas.....	28
8.- Conclusiones y perspectivas	29
9.- Referencias	30

Estudio paleogenómico de la coevolución del pajarito/frijol y sus simbiosis durante la domesticación

1.- Antecedentes

El frijol es una de las plantas domesticadas más importantes a nivel mundial, tanto por su valor nutricional, debido a su alto contenido de proteínas, como por su relevancia ecológica y cultural. En América, varias especies del género *Phaseolus* fueron domesticadas (Bitocchi et al. 2017), siendo *P. vulgaris* y *P. lunatus* las más representativas y más comúnmente cultivadas. Estas dos especies fueron domesticadas independientemente en Mesoamérica y los Andes hace aproximadamente 8000 años (Bitocchi et al. 2013) y actualmente se cultivan en varias regiones del mundo.

Durante la domesticación, el frijol sufrió cambios morfológicos y fisiológicos como consecuencia de la selección humana y natural. El frijol domesticado carece de sensibilidad al fotoperiodo y dormancia de las semillas. Además, posee semillas más grandes en comparación con su contraparte silvestre (Koinange et al. 1996) con una gran diversidad de color y tamaño en las semillas entre las diferentes variedades de frijol domesticado. Estudios previos han comenzado a identificar algunas de las variantes genéticas responsables de estos cambios (Rendón-Anaya et al. 2017). Demostrando que cerca del 16% del genoma del frijol posee señales de selección asociadas al proceso de domesticación (Bellucci et al. 2014).

El frijol representa una de las fuentes de proteínas más importantes para la población humana (Castro-Guerrero et al. 2016). Su alto contenido proteico se debe a su capacidad de asociarse con bacterias fijadoras de nitrógeno del suelo conocidas como rizobios (Martínez-Romero 2003). Esta relación simbiótica es de gran importancia comercial, pues reduce la dependencia de la planta de fertilizantes, pero también evolutiva pues le permite al frijol crecer en ambientes con limitación de nutrientes, lo que lo hace ideal para procesos de reforestación (Foyer et al. 2019). Estudiar la domesticación del frijol nos permite entender mejor las prácticas agrícolas ancestrales y es fundamental para el desarrollo de variedades mejoradas frente a los desafíos actuales del cambio climático y la seguridad alimentaria.

2.- Problemática

Aunque estudios recientes han comenzado a describir el proceso de domesticación del frijol, un área que permanece poco explorada es la coevolución entre esta planta y sus bacterias simbióticas. El frijol depende en gran medida de su relación simbiótica con los rizobios en su raíz para adaptarse a entornos cambiantes (Hayat et al. 2010), por lo que es posible que esta interacción haya tenido un papel importante durante el proceso de domesticación. Sin embargo, estudiar dicha coevolución es complejo, debido a procesos como el flujo génico o el reemplazo de simbioses, los cuales pueden oscurecer la historia evolutiva de estas asociaciones.

La paleogenómica, o el estudio del ADN antiguo, ha transformado la genómica evolutiva al permitir recuperar información genética de materiales arqueológicos

y paleontológicos (Orlando et al. 2021). Estos avances permiten recuperar no sólo el ADN del organismo al que estos materiales pertenecen, sino también datos sobre su dieta, patógenos y microbiomas asociados (Jensen et al. 2019; Wibowo et al. 2021). En este contexto, las semillas arqueológicas de frijol ofrecen una oportunidad única, ya que son frecuentes en contextos arqueológicos, presentan buena preservación de ADN (Trucchi et al. 2021) y, como se ha demostrado en semillas modernas, contienen rizobios endofíticos que colonizan las raíces al germinar (López-López et al. 2010). Por tanto, permiten estudiar tanto la evolución del frijol como sus asociaciones simbióticas en el pasado durante la domesticación.

Por medio de la generación de datos de secuenciación genómica de semillas arqueológicas de frijol, estudiaremos la coevolución del frijol y sus bacterias simbióticas durante la domesticación.

3.- Finalidad de los análisis

Los análisis tienen como finalidad investigar la coevolución del frijol (“frejol” y “pallar”) y sus simbiontes bacterianos, por medio del estudio de genomas generados a partir de preparar librerías de secuenciación y la extracción de ADN de muestras (semillas y vainas) antiguas.

4.- Lugar de procedencia de las muestras

Las 30 muestras objeto del análisis científico realizados fueron seleccionadas y tomadas durante el desarrollo del proyecto de investigación de colecciones y fondos museográficos administrados por el Ministerio de Cultura “Estudio

paleogenómico de la coevolución del frijol y sus simbioses durante la domesticación”, autorizado mediante la Resolución Directoral N°000040-2025-DGM-VMPCIC/MC, emitido el 09 de junio de 2025.

Las muestras formaban parte del acervo cultural del Museo Regional de Ica “Adolfo Bermúdez Jenkins”.

Antes de formar parte de la colección del museo regional, las muestras fueron recolectadas en los siguientes sitios arqueológicos del departamento de Ica: La Yerba III, Pernil Alto, Ánimas Altas/Ánimas Bajas, Cerro Córdova, Tajahuana, Cahuachi, Maymi, Casa Vieja, Huaca del Loro, H-8, Huaca Guadalupe y Tacaraca A.

La Yerba III está ubicado en la margen derecha del río Ica y cerca del litoral. La Yerba III es un sitio precerámico (6485-5893 cal. a. P.), conjuga los restos de “viviendas” circulares semisubterráneas de 4.5 de diámetro, las cuales están cubiertas con estratos (capas) compuestos de restos de alimentos marinos y vegetales, incluyendo cultivos como pallares (Beresford-Jones et al., 2015; 2017; 2021).

La Yerba III se encuentra al interior del distrito de Santiago, provincia y departamento de Ica. Las muestras fueron colectadas por el Proyecto de Investigación Arqueológica Samaca durante la temporada 2015. La muestra 13 proviene de la unidad de excavación Pozo 3, contexto UE 9021: se trata de la acumulación de restos botánico y animales mezclados con arena. La muestra 13 proviene de la unidad de excavación Pozo 2, contexto 9766: se trata de una capa compuesta de restos botánicos y animales mezcladas con sedimento.

Pernil Alto es un sitio arqueológico del valle medio de Río Grande, a 6 kilómetros al noreste de la ciudad de Palpa. El sitio conserva evidencias de dos momentos de ocupación: 1.- las viviendas semisubterráneas y entierros de los primeros agricultores del periodo precerámico (3600 a 3000 a. C.) y 2.- las construcciones de barro, artefactos y ecofactos de pobladores sedentarios del Periodo Inicial (1500-850 a. C.) (Gorbahn, 2013; Reindel e Isla, 2009). Pernil Alto se encuentra dentro de los límites del distrito Río Grande, provincia de Palpa, departamento de Ica.

Los bienes culturales de interés para la presente solicitud fueron colectados por el Proyecto Arqueológico Palpa Lucanas 2008 y Proyecto Nasca Palpa 2009. De la colección de Pernil Alto tomamos siete muestras: muestra 23, muestra 24, muestra 25, muestra 26, muestra 27, muestra 28 y muestra 29.

La muestra 23 fue colectada en un sedimento de grano fino mezclado con restos de plantas y piedras, este sedimento cubre un “piso”. La muestra 24 proviene de la Unidad 5, capa F, matriz 9. La muestra 25 fue colectada en la Unidad 5, capa G (restos de encendido de fuego). La muestra 26 fue colectada en la Unidad 5, capa G (restos de encendido de fuego). La muestra 27, la muestra 28 y la muestra 30 provienen de la Unidad 5, capa F.

Ánimas Altas/Ánimas Bajas es un complejo arqueológico que conjuga 100 montículos dentro de un área mayor a 90 hectáreas. El complejo se ubica en la cuenca de Callango del valle bajo de Ica y, en su interior, se ha podido identificar arquitectura pública ceremonial (algunas con frisos), áreas de residencias, espacios de producción, cementerios y plazas. El estudio estilístico de la

cerámica ha permitido plantear que el lugar fue ocupado entre el 400 al 100 a. C. (Bachir Bacha y Llanos, 2013).

El sitio se ubica en el distrito de Ocucaje, provincia de Ica, departamento de Ica. Las muestras de interés fueron colectadas por el Programa Arqueológico Animas Altas -Ica Perú, durante el desarrollo de su temporada 2011, en el Montículo 26. La muestra 1 fue colectada durante la excavación de un relleno constructivo, el cual fue denominado U.E. 46, de igual forma, la muestra 2 proviene de un relleno constructivo (U.E. 113c).

Cerro Córdova, también conocido como Cerro Max Uhle, es un complejo arqueológico cuyos componentes se distribuyen sobre la superficie de dos elevaciones rocosas mayores a 30 metros, y el área que los separa, conocidas como Cerro Córdova. Actualmente el sitio se encuentra segmentado en dos bloques debido a la construcción de una pista asfaltada y seriamente deteriorado por la constante acción de huaqueros. A pesar de la alteración, es posible observar restos de muros de adobes y quinchá, así como rellenos constructivos expuestos. Tiestos de vasijas de cerámica de los estilos Paracas y Nasca se encuentran con frecuencia sobre la superficie (Llanos, 2017).

El sitio se ubica en el distrito de Ocucaje, provincia de Ica, departamento de Ica. La muestra de interés fue colectada por el Programa Arqueológico Animas Altas -Ica Perú, durante el desarrollo de su temporada 2014. La muestra 3 fue colectada en la denominada Zona III del bloque oeste en el que está dividido el sitio y dentro del relleno UE 188.

Tajahuana (circa 370-200 a. C.) está ubicado en el valle medio de Ica. Los componentes del sitio se distribuyen sobre la cima, quebradas, laderas y faldas del cerro Tajahuana. En la cima resaltan tres sectores. El Sector Norte es el más conocido y en él destaca la concentración de estructuras de planta rectangular que fueron identificadas como viviendas, varios cientos de metros de muros altos, largos y anchos (también llamados murallas) y un espacio abierto delimitado por cinco plataformas. En el Sector Oeste se encuentra un geoglifo trapezoidal y cinco colinas; cuatro de ellas irradian geoglifos en forma de líneas desde la cumbre. En el Sector Sur destacan dos monumentos de planta rectangular alargada (Massey, 2022; Williams y Pazos, 1974).

Tajahuana se ubica en el distrito de Santiago, provincia de Ica, departamento de Ica. Los bienes culturales de interés fueron colectados el Proyecto de Investigación Arqueológica Tajahuana durante el desarrollo de sus temporadas 2019. La muestra 4, 5 y 30 provienen del mismo contexto: UE 2101, el cual se trata de un relleno constructivo. El contexto UE 2101 fue registrado dentro de la Unidad 2, la cual fue intervenida a fin de conocer las características arquitectónicas de las murallas que divide la zona de geoglifos con el área de las edificaciones del Sector Norte.

Cahuachi (200 a 450 d. C.) fue el centro ceremonial de carácter regional de la época Nasca. El complejo arqueológico se ubica en la cuenca de Río Grande, a 18 kilómetros al oeste de la ciudad de Nasca. La arquitectura del complejo abarca 24 hectáreas y el sector central está compuesto por varias pirámides truncas edificadas con adobes. En esta zona, además, es posible advertir patios abiertos, amplias terrazas con recintos techados y murallas entre los edificios

monumentales. En la parte superior de las pirámides se conservan las bases de grandes recintos cuyas coberturas fueron sostenidas por decenas de columnas (Lasaponara, Masini y Orefici, 2016). Cahuachi se ubica en el distrito de Nasca, provincia de Nasca, departamento de Ica.

Las tres muestras de interés para el presente proyecto fueron colectadas por Proyecto De Investigación Arqueológica Proyecto Nasca 2017-2018. La muestra 19 fue colectada en Sector Y1/ Unidad de excavación Exp. 184/ Cuadrícula de excavación Q20/ Capa/Nivel B/2, Contexto Ampliación 1/Lado oeste. La muestra 20 fue recolectada en el Sector Y1/ Unidad de excavación Exp. 185/ Cuadrícula de excavación Q17/ Capa/Nivel A/1-2, Contexto Relleno al norte de muro 10. Y la muestra 21 en el Sector Y1/ Unidad de excavación Exp. 184/ Cuadrícula de excavación Q18/ Capa/Nivel A/1, Contexto 5D.

Maymi (*circa* 550-800 d. C.) se encuentra en la margen norte del río Pisco, a 13 kilómetros del mar. Los restos arqueológicos están distribuidos sobre una terraza. En el lugar, gracias a excavaciones y reconocimiento de superficie, se han hallado evidencias relacionadas a dos momentos de ocupación: Nasca Tardía y Wari. Las evidencias arqueológicas incluyen, recintos edificados con adobes, muros de quincha, contextos funerarios, así como residuos de alimentos y utensilios relacionados a la producción de cerámica relacionada a los estilos del Estado Wari (Anders, 1990).

Maymi se ubica en el distrito de San Clemente, provincia de Pisco, departamento de Ica. Los tres bienes culturales de interés para esta investigación fueron colectados por el Proyecto Arqueológico Maymi. La muestra 15 fue colectada en

el sector IV, Sub-sector C, Unidad 1 y al interior de la Capa C. La muestra 16 fue colectada en el Sector IV, Sub-sector C, Unidad 5 y dentro de la Capa C. Y la muestra 22 fue recolectada en el Sector IV, Sub-sector C, Unidad I-O y dentro de la Capa D.

Casa Vieja (*circa* 550-800 d. C.) se encuentra en la margen este del río Ica, en la cuenca de Callango. El sitio mide 9 hectáreas y en su interior se encuentra vestigios de la época Nasca tardía y Horizonte Medio 2. De acuerdo con Cook (1994, p. 80), el lugar conjuga tres montículos. Expuestos en superficie se observó muros de quinchas y restos relacionadas a la vida cotidiana, actos rituales y producción. El sitio se ubica en el distrito de Ocucaje, provincia de Ica, departamento de Ica. Las muestras fueron colectadas por Proyecto Arqueológico Valle Bajo de Ica 1999. Las dos muestras provienen de la Zona D. La muestra 17 proviene del Locus 100b y la muestra 18 del Locus 129, ambos locus forman parte de acumulaciones de desperdicios.

Huaca del Loro (*circa* 550-800 d. C.) es un sitio relevante en la discusión sobre la expansión del Estado Wari hacia el territorio Nasca. El establecimiento se encuentra en el valle de Las Trancas y abarca cerca de 20 hectáreas. Los recientes trabajos de Conlee y colegas han permitido identificar cinco sectores. En el Sector I se compone de edificio de plantar ortogonal. El Sector I conjuga la estructura monumental de planta en D y edificios monumentales rectangulares. El Sector III es el cementerio muy deteriorado por huaqueo. Por último, el Sector IV y Sector V contiene abundantes restos de residuos sobre la superficie, pero no se aprecia arquitectura expuesta (Conlee et al., 2021).

Huaca del Loro se encuentra al interior del distrito de Vista Alegre, provincia de Nasca, departamento de Ica. Las muestras de interés fueron colectadas por Proyecto de Investigación Arqueológica Huaca del Loro durante el desarrollo de sus temporadas 2019 y 2022. Las dos muestras (muestra 11 y muestra 12) provienen de la capa B, este junto con otros (A, C y D) cubrieron un piso de una estructura edificada con muros de piedra.

H-8 es un asentamiento de 0.8 hectáreas y está emplazado en el valle bajo de Ica, conserva restos de habitaciones y recintos edificados con muros de tapia, adobes y piedras. En la arquitectura conservada es posible advertir corredores, patios y unidades domésticas. Las investigaciones realizadas por el PIA Samaca, en el transcurso de los últimos diez años, han mostrado que fue ocupado de manera continua por más de 600 años. Este lapso temporal transcurre desde el Periodo Horizonte Medio hasta el inicio de la época colonial (Cook, 1990; Santana et al., 2022). Las muestras de interés: muestras 8 y muestra 9, fueron colectadas por el Proyecto de Investigación Arqueológica Samaca durante su temporada 2024. Ambas muestras provienen del mismo contexto: UE2401, el cual se trata de una acumulación de restos de alimentos y otros desechos, ocurrido al interior de un recinto (unidad 2).

Huaca Guadalupe (circa 1000-1532 d. C.) es un asentamiento ubicado en el valle medio de Ica, en la margen derecha. El sitio conserva dos sectores. En ambos lugares es posible observar arquitectura alzada con adobes y mortero de asiento. Los edificios tienen los muros cubiertos con enlucido y diseño de planta ortogonal (Mallco, 2019). Las muestras de interés (muestras 6 y 7) fueron colectadas por el Proyecto Arqueológico Ica-Investigación en el Sitio

Arqueológico de Huaca Guadalupe, distrito de Salas, Provincia de Ica, Departamento Ica 2016-2017. Ambas muestras provienen del mismo contexto: Capa B/Nivel 2.

Tacaraca, sector A (*circa* 1000-1532 d. C.) es un montículo ubicado en el valle medio de Ica. Se trata de un edificio monumental de planta ortogonal que por acción de fuerzas erosivas naturales y culturales ha perdido su configuración original. El montículo, junto con otros, forman parte del sitio arqueológico Ica Vieja, este habría sido el centro administrativo religioso del valle durante el Periodo Intermedio Tardío (*circa* 1000-1470 d. C.) (Menzel 1976). Las excavaciones realizadas en el monumento han podido registrar la superposición de recintos construidos con adobes del periodo tardío debajo de habitaciones de la época Inca (Mallco, 2015). El monumento pertenece al distrito de Pueblo Nuevo, provincia de Ica, departamento de Ica.

El bien cultural de interés fue colectado por el Proyecto de Investigación Arqueológica y Puesta en Valor del Sitio Arqueológico de Tacaraca-Ica 2012. La muestra 10 proviene de la Unidad 1, Unidad arquitectónica N°2 y del contexto Capa C/Nivel 4, el cual se trata del relleno de tierra que cubrió un conjunto de soguillas, todo cubierto por la construcción de un piso (Capa C).

5.- Datos cuantitativos y cualitativos de las muestras

Muestra	Tipo de material	Cantidad	Peso gr	Descripción	Procedencia
1	Botánico	1	1.68	Vaina incompleta con semilla de pallar	Ánimas Altas/Ánimas Bajas
2	Botánico	1	1.24	Semilla de pallar	Ánimas Altas/Ánimas Bajas
3	Botánico	1	1.03	Semilla de pallar	Cerro Córdova
4	Botánico	1	1.00	Semilla de pallar	Tajahuana
5	Botánico	1	2.63	Vaina de pallar	Tajahuana
6	Botánico	1	0.82	Semilla de pallar	Huaca Guadalupe
7	Botánico	1	0.94	Semilla de frejol	Huaca Guadalupe
8	Botánico	1	1.35	Vaina de pallar	H-8
9	Botánico	1	0.7	Semilla de frejol	H-8
10	Botánico	1	0.87	Semilla de frejol	Tacaraca
11	Botánico	1	0.57	Semilla de pallar	Huaca del Loro
12	Botánico	1	0.39	Semilla de frejol	Huaca del Loro
13	Botánico	1	0.73	Semilla de pallar	La Yerba III
14	Botánico	1	0.82	Semilla de pallar	La Yerba III
15	Botánico	1	1.05	Semilla de pallar	Maymi
16	Botánico	1	1.07	Semilla de pallar	Maymi
17	Botánico	1	0.41	Semilla de frejol	Casa Vieja
18	Botánico	1	1.29	Semilla de pallar	Casa Vieja
19	Botánico	1	1.30	Semilla de pallar	Cahuachi
20	Botánico	1	0.32	Semilla de frejol	Cahuachi
21	Botánico	1	1.20	Semilla de pallar	Cahuachi
22	Botánico	1	0.85	Semilla de pallar	Maymi
23	Botánico	1	0.68	Semilla de pallar	Pernil Alto
24	Botánico	1	0.80	Semilla de pallar	Pernil Alto
25	Botánico	1	1.10	Semilla de frejol	Pernil Alto
26	Botánico	1	0.57	Semilla de pallar	Pernil Alto
27	Botánico	1	0.52	Semilla de frejol	Pernil Alto
28	Botánico	1	0.92	Semilla de frejol	Pernil Alto
29	Botánico	1	1.31	Semilla de pallar	Pernil Alto
30	Botánico	1	0.65	Semilla de pallar	Tajahuana

Fotos de cada muestra

	
Muestra 1	Muestra 2
	
Muestra 3	Muestra 4
	
Muestra 5	Muestra 6
	
Muestra 7	Muestra 8
	
Muestra 9	Muestra 10



Muestra 11



Muestra 12



Muestra 13



Muestra 14



Muestra 15



Muestra 16



Muestra 17



Muestra 18



Muestra 19



Muestra 20



Muestra 21



Muestra 22



Muestra 23



Muestra 24



Muestra 25



Muestra 26



Muestra 27



Muestra 28



Muestra 29



Muestra 30

6.- Metodología

6.1.- Extracción de ADN

Las semillas arqueológicas de frijol (n=30) fueron procesadas en los laboratorios de ADN antiguo del Instituto Globe de la Universidad de Copenhague, Dinamarca, utilizando protocolos especiales para minimizar la contaminación. Estos laboratorios cuentan con presión positiva de aire, filtros de aire de partículas de alta eficiencia (HEPA), un sistema de descontaminación con luz ultravioleta diario y protocolos de limpieza estrictos.

Cada una de las semillas fue fotografiada, pesada y pulverizada en condiciones estériles. El material pulverizado fue transferido a un tubo de eppendorf para su posterior extracción. La extracción de ADN fue hecha utilizando métodos previamente estandarizados para muestras arqueológicas de plantas (Wales et al., 2014; Wales et al., 2015). En breve, el polvo se incubó en un rotador a 55 °C durante toda la noche en 1 ml de la siguiente solución: 10 mM de hidrocloreuro de tris(hidroximetil) aminometano (Tris-HCl, pH 8,0), 10 mM de NaCl, 2 % p/v de dodecilsulfato de sodio (SDS), 5 mM de CaCl₂, 2,5 mM de ácido etilendiaminotetraacético (EDTA, pH 8,0), 40 mM de ditioneitol (DTT) y 10 % de proteinasa K. El ADN se extrajo utilizando un protocolo de extracción de fenol-cloroformo con dos rondas de limpieza con fenol. Posteriormente, las muestras se purificaron utilizando un kit de purificación de PCR MinElute con las modificaciones descritas en (Dabney et al. 2013). En breve, después de centrifugar, se recolectó la capa acuosa superior, se añadió a 13 volúmenes de una solución de PB, se mezcló bien y se transfirió a una columna del High Pure

Viral Nucleic Acid Large Volume Kit (Roche Life Science). Las muestras se centrifugaron durante 4 minutos a $1500 \times g$, y las columnas se lavaron dos veces con 700 μ l de la solución PE y se transfirieron a un tubo Eppendorf LoBind de 1,5 ml. Posteriormente, las columnas se incubaron a 37 °C durante 15 minutos en 30 μ l de EB y el ADN se eluyó mediante centrifugación a $16\ 100 \times g$ durante un minuto. Finalmente, para eliminar posibles compuestos inhibidores que podrían afectar la construcción de las bibliotecas de ADN, utilizamos el OneStep PCR Inhibitor Removal Kit de Zymo Research. Las muestras fueron extraídas en grupos de 9, más un blanco, para minimizar las posibilidades de contaminación cruzada, y el color de la solución después de la digestión, extracción y limpieza con las columnas de Zymo fue registrado.

Los extractos de ADN se midieron utilizando un fluorómetro Qubit y el kit de ensayo de alta sensibilidad de ADN bicatenario (Life Technologies) siguiendo el protocolo del fabricante (Tabla 1). Todos los blancos resultaron en concentraciones de ADN no detectables demostrando que no hubo contaminación durante el procesamiento de las muestras.

Tabla 1. Semillas arqueológicas utilizadas para extracción de ADN, concentración de ADN y color registrado después de la digestión, extracción y limpieza con el kit de Zymo.

Index	ID	ADN (ng/μl)	Color (después de la digestión)	Color (después de la extracción)	Color (después de la columna Zymo)
1	AA_1	0.202	Café oscuro/ambar	Amarillo claro	Transparente
2	AA_2	0.272	Café oscuro/ambar	Amarillo claro	Transparente
3	CC_3	0.243	Café oscuro	Amarillo claro	Transparente
4	Taj_4	0.357	Café oscuro	Amarillo claro	Transparente
5	Taj_5	0.383	Café oscuro/ambar	Narana	Amarillo claro
6	HG_6	2.37	Ambar	Naranja	Amarillo claro
7	HG_7	1.88	Ambar	Amarillo claro	Transparente
Blan	Blank 26/01/26	Too low	Transparente	Transparente	Transparente
8	H8_8	12.6	Ambar	Amarillo claro	Transparente
9	H8_9	7.98	Ambar	Amarillo claro	Transparente
10	Tar_10	3.84	Ambar	Amarillo claro	Transparente
11	HL_11	0.489	Ambar	Amarillo claro	Transparente
12	HL_12	8.06	Ambar	Amarillo claro	Amarillo claro
13	LY_13	0.148	Café oscuro	Amarillo claro	Transparente
14	LY_14	0.256	Café oscuro	Amarillo claro	Transparente
15	May_15	0.469	Café oscuro	Naranja	Amarillo claro
16	May_16	0.415	Café oscuro	Ambar	Amarillo claro
Blank	Blank 23/02/26	Too low	Transparente	Transparente	Transparente
17	CV_17	2.31	Café oscuro	Amarillo claro	Transparente
18	CV_18	0.697	Café oscuro	Amarillo claro	Transparente
19	Cah_19	0.62	Café oscuro	Amarillo claro	Transparente
20	Cah_20	0.447	Café oscuro	Amarillo claro	Transparente
21	Cah_21	0.248	Café oscuro	Amarillo claro	Transparente
22	May_22	0.371	Café oscuro	Café	Amarillo claro
23	PA_23	0.099	Café oscuro	Amarillo claro	Transparente
24	PA_24	0.448	Café oscuro	Amarillo claro	Transparente
25	PA_25	0.339	Café oscuro	Amarillo claro	Transparente
Blank	Blank 26/03/26	Too low	Transparente	Transparente	Transparente
26	PA_26	0.318	Naranja	Amarillo claro	Transparente
27	PA_27	0.13	Ambar	Amarillo claro	Transparente
28	PA_28	0.274	Naranja	Amarillo claro	Transparente
29	PA_29	0.145	Naranja	Amarillo claro	Transparente
30	Taj_30	0.302	Café oscuro	Amarillo claro	Transparente

6.2.- Preparación de bibliotecas de secuenciación y secuenciación

Las muestras con concentración de ADN mayor a 1 ng/μl (con la excepción de la muestra Taj_5, que se procesó a pesar de tener una concentración de 0.383

ng/μl) se utilizaron para preparar bibliotecas de Illumina mediante un protocolo optimizado en un solo tubo (Carøe et al. 2018). La reparación de extremos se realizó en reacciones de 40 μl que contenían 32 μl de extracto de ADN, con una incubación de 30 min a 20 °C seguida de 30 min a 65 °C. La ligación de adaptadores se llevó a cabo con 2 μl de adaptadores y ligasa T4 en reacciones de 50 μl, incubadas durante 30 min a 20 °C y 10 min a 65 °C. La reacción de relleno (fill-in) se realizó añadiendo 10 μl de mezcla maestra, seguida de una incubación de 15 min a 65 °C y 15 min a 80 °C. Las bibliotecas se purificaron utilizando perlas magnéticas SPRI, se eluyeron en 60 μl de buffer EB, se amplificaron, se indexaron con índices duales (Kircher et al. 2012), se cuantificaron, se agruparon y se secuenciaron en modo paired-end, con 50 ciclos, en un secuenciador NovaSeqXPlus en el centro de secuenciación del Instituto Globe de la Universidad de Copenhague. Para cada biblioteca generamos entre 150 y 250 millones de lecturas y entre 5-10 millones de lecturas para los blanks.

6.3.- Procesamiento bioinformático de los datos de secuenciación

Utilizamos AdapterRemoval2 (Schubert et al. 2016) para cortar los adaptadores de lecturas de secuenciación, además, regiones de baja calidad y poly-Ns situadas al inicio o al final de las lecturas se cortaron utilizando el mismo programa. Las lecturas de menos de 30 pb fueron descartadas. Posteriormente, las lecturas fueron mapeadas al genoma de referencia de *Phaseolus vulgaris* (v2.0) (Schmutz et al. 2014) utilizando BWA aln 0.7.12 (Li and Durbin 2009). La semilla de BWA se desactivó (-l 1000) con el fin de evitar sesgos en el alineamiento debidos a sustituciones en el extremo 5' causadas por daños en el

ADN, y se descartaron las lecturas con una calidad de mapeo inferior a 30. Los duplicados de PCR se identificaron y removieron de los alineamientos resultantes utilizando Picard 1.130 (<http://picard.sourceforge.net>). Finalmente, las lecturas se realinearon contra el genoma de referencia utilizando Genome Analysis Toolkit (GATK 3.3), y la etiqueta MD se re-calculó mediante samtools 1.2 (Li et al. 2009). Los resultados de secuenciación se resumen en la Tabla 2. Finalmente, utilizamos mapDamage 2.0 (Jónsson et al. 2013) para evaluar la autenticidad de las lecturas de secuenciación.

Tabla 2. Estadísticas de secuenciación.

Nombre de muestra	Nombre de la biblioteca	Lecturas generadas	Lecturas sin adaptador	Lecturas mapeadas	Lecturas mapeadas y filtradas	% ADN endógeno	Profundidad de cobertura genómica
H8-8	H8-8-E1L1A1	300205968	158257659	56652221	17761821	11.22	1.84
H8-9	H8-9-E1L1A1	258889544	135306248	47038773	16068667	11.88	1.66
HG-6	HG-6-E1L1A1	477702790	251906731	40448259	11364632	4.51	1.08
HG-7	HG-7-E1L1A1	520396742	262665077	91349990	23243094	8.85	2.36
HL-11	HL-11-E1L1A1	656157730	239918302	46490350	5004027	2.09	0.41
HL-12	HL-12-E1L1A1	293045090	158314361	152565217	62184755	39.28	6.65
Taj-5	Taj-5-E1L1A1	460933754	69368968	17311829	1560282	2.25	0.13
Tar-10	Tar-10-E1L1A1	378476728	186270478	60275923	16801368	9.02	1.60
Blanco	Ext170625-L1A2	121111166	1928064	206	87	0.00	0.00
Blanco	Ext230925-L1A2	9170926	1912884	100	9	0.00	0.00
Blanco	Ext240226-L1A1	52294392	1886123	395	54	0.00	0.00
Blanco	Ext260126-L1A1	68627008	2156322	125	38	0.00	0.00
Blanco	Lib171125-A2	46835912	2408051	20	7	0.00	0.00
Blanco	Lib150426-A1	37140438	1012960	51	15	0.00	0.00
Blanco	Lib2171125-A2	59353800	3652489	741	313	0.01	0.00

6.4.- Preparación del panel de genomas modernos de referencia

Preparamos un panel de datos con 297 genomas modernos de varias especies de *Phaseolus* (Tabla 3) para comparar con nuestros genomas antiguos. Para cada uno de estos genomas, descargamos las lecturas de secuenciación del “European Nucleotide Archive” (ENA). Las lecturas fueron posteriormente mapeadas al genoma de referencia de *Phaseolus vulgaris* (v2.0) (Schmutz et al.

2014) con los mismos parámetros descritos en la sección anterior, pero sin desactivar la semilla al momento del mapeo con BWA aln.

Tabla 3. Genomas de referencia de *Phaseolus* utilizadas para el análisis genómico.

Especie	Número de genomas	Publicaciones
<i>Phaseolus vulgaris</i>	153	(Bellucci et al. 2023) Nat Communications
<i>Phaseolus lunatus</i>	63	(Frascarelli et al. 2025) Theoretical and Applied Genetics
<i>Phaseolus acutifolius</i>	63	
<i>Phaseolus coccineus</i>	4	(Kates et al. 2024) Nature Communications
<i>Phaseolus dumosus</i>	2	(Lobaton et al. 2018) The plant genome
<i>Phaseolus amblyosepalus</i>	1	
<i>Phaseolus carteri</i>	1	(Lozano-Arce et al. 2023) Communications Biology
<i>Phaseolus costaricensis</i>	1	
<i>Phaseolus filiformis</i>	1	(Moghaddam et al. 2021) Nature Communications
<i>Phaseolus hintonii</i>	1	
<i>Phaseolus jaliscanus</i>	1	(Rendón-Anaya et al. 2017) Genome Biology
<i>Phaseolus leptostachyus</i>	1	(Trucchi et al. 2021) Nature Plants
<i>Phaseolus macrolepis</i>	1	
<i>Phaseolus maculatus</i>	1	
<i>Phaseolus micranthus</i>	1	
<i>Phaseolus microcarpus</i>	1	
<i>Phaseolus perplexus</i>	1	
<i>Phaseolus polystachios</i>	1	

6.5.- Llamado de variantes

Para identificar sitios polimórficos seleccionamos un máximo de 10 genomas de cada especie de los genomas modernos y utilizamos ANGSD (-GL 2) (Korneliussen et al. 2014) para el llamado de variantes. Bases con calidad menor a 20 y sitios con cobertura en menos del 80% de los genomas fueron descartados. Para evitar incorporar el daño debido a la degradación de ADN antiguo en las muestras arqueológicas las transiciones (C-T, G-A) fueron descartadas. Con este procedimiento, obtuvimos un total de 176,820 polimorfismos de un sólo nucleótido (SNP). Dado que la profundidad de

cobertura de la mayoría de muestras antiguas (0.4-6.6 X) no es ideal para estimar genotipos diploides, por cada SNP y cada genoma antiguo y moderno, seleccionamos una base aleatoria de los datos de secuenciación, una estrategia comúnmente utilizada en estudios de ADN antiguo (ej. (Ramos-Madrigal et al. 2016)).

6.6.- Análisis de ancestría de los genomas antiguos

Comparamos nuestros genomas antiguos con las especies modernas en el panel de referencia por medio de un análisis de escalamiento multidimensional (MDS). Primero estimamos distancias pareadas entre todos los pares posibles de genomas modernos y antiguos utilizando Plink2.0 (--distance-matrix) (Chang et al. 2015). Después utilizamos la matriz de distancias obtenida y la función *cmdscale* de R para estimar el MDS. Los resultados fueron graficados con R (Figura 4 A).

Para incrementar la resolución de la estima de ancestría utilizamos ADMIXTURE (Alexander et al. 2009). Corrimos 50 réplicas de ADMIXTURE comenzando en diferentes semillas, incluyendo todos los genomas modernos y antiguos en el panel de SNPs y asumiendo K=10. La réplica con mejor likelihood se muestra en la figura 4 B.

7.- Resultados

7.1.- Preservación de ADN antiguo en las semillas arqueológicas

Extrajimos de forma exitosa ADN para un total de 30 muestras pertenecientes a 12 sitios arqueológicos distintos periodos: Horizonte Temprano (n=6), Horizonte

Medio (n=10), Horizonte Tardío (n=5), Periodo Inicial (n=7) y Precerámico medio (n=2) (Figura 1A). La mayoría de las extracciones de ADN presentaron un color abarcando del amarillo al café oscuro, que es indicador de la presencia de compuestos orgánicos que pueden inhibir reacciones enzimáticas y obstruir la cuantificación de ADN. Sin embargo, después de una limpieza con el kit de Zymo para eliminar inhibidores, todas las extracciones de ADN mostraron un color transparente, indicando que la mayoría de los inhibidores fueron removidos. La concentración de ADN obtenida varió en cada muestra, con un máximo de 12.6 ng/μl (H-8-8, Ica) y un mínimo de 0.099 ng/μl para la muestra (Pernil Alto, Palpa - 23) (Figura 1B). Para identificar los sitios arqueológicos que presentan mejor preservación de ADN, evaluamos la concentración por sitio. Las semillas de sitio H-8 (Ica) presentaron las concentraciones de ADN más altas, seguidas de las semillas de Huaca Loro (Nazca), Taraca A (Ica) y Huaca Guadalupe (Ica). En contraste, los sitios que resultaron con las concentraciones más bajas de ADN en promedio fueron Pernil Alto (Palpa) y La Yerba III (Ica).

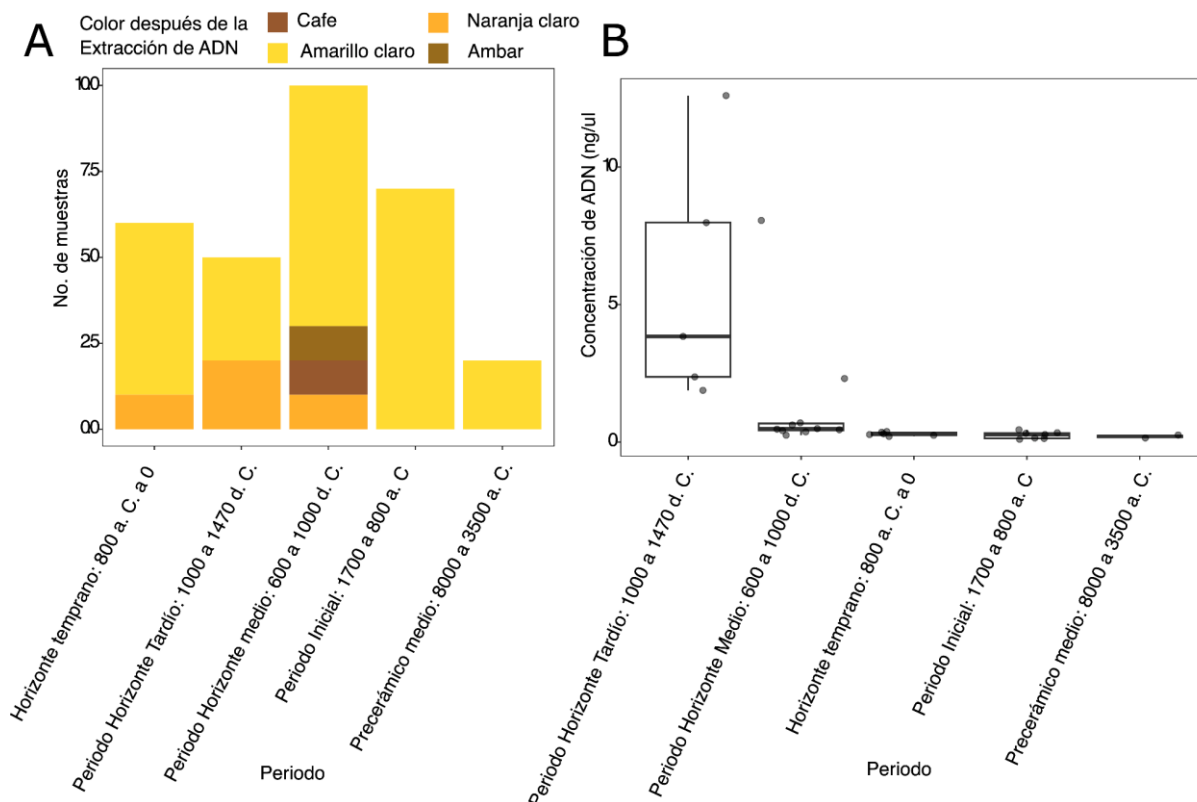


Figura 1. Preservación de ADN antiguo en semillas antiguas de Phaseolus. (A) Gráfica de barras mostrando el número de muestras procesadas por periodo. En el eje de las x se muestran los distintos periodos de las muestras analizadas, en el eje de las y el número total de muestras procesadas. El color indica el color del extracto después de seguir el protocolo de extracción. (B) Concentración de ADN por muestra. En el eje de las x se muestra el nombre del sitio arqueológico y el eje de la y indica la concentración de ADN total en ng/ul.

Puesto que el ADN se degrada conforme el paso del tiempo, evaluamos la correlación entre la edad esperada de las muestras y la concentración de ADN (Figura 2 A). Consistente con observaciones de estudios previos (e.g. (Kistler et al., 2017)), las semillas de sitios más antiguos mostraron concentraciones de ADN menores que los sitios más recientes. El sitio H-8 (Ica) del periodo Horizonte Tardío (1470 a 1532 d. C.) demostró la mejor preservación de ADN.

Compuestos orgánicos fenólicos en las muestras arqueológicas pueden afectar la eficacia de la extracción de ADN. Investigamos si las semillas con extractos más o menos teñidos resultaban en concentraciones de ADN mayores o

menores (Figura 2 B). No encontramos una correlación clara entre el color del extracto y la cantidad obtenida de ADN, sin embargo, las muestras con concentración de ADN más alta todas presentaron extractos menos teñidos (categoría “Amarillo claro”).

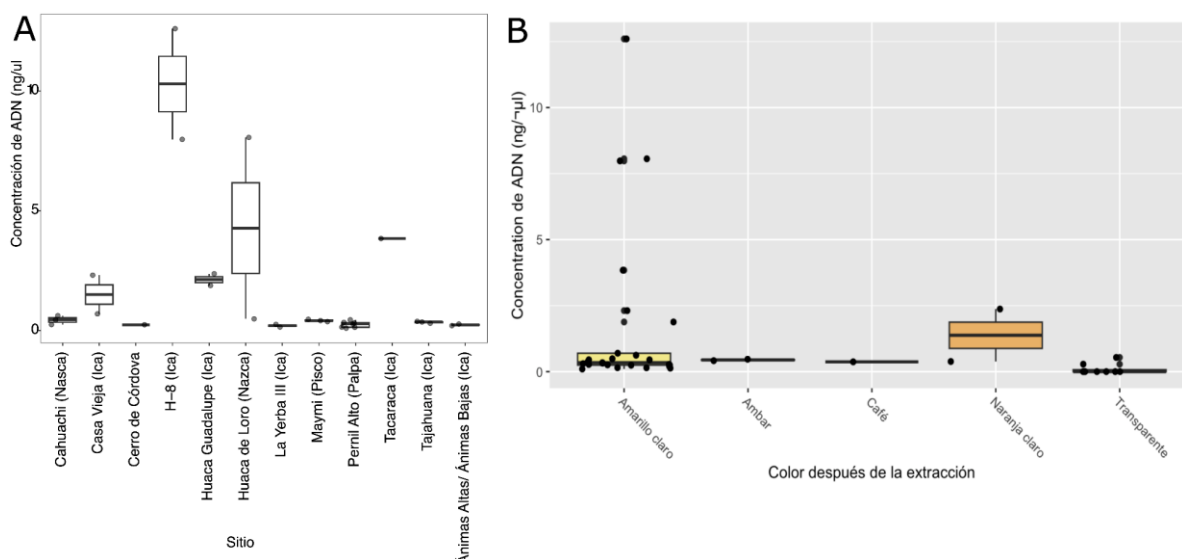


Figura 2. Preservación de ADN antiguo por sitio y periodo. (A) Concentración de ADN por periodo de tiempo. (B) Concentración de ADN de acuerdo al color del extracto post-extracción. En el eje de las x se muestran los diferentes colores presentados en los extractos post-extracción mientras que en el eje de las y la concentración total de ADN. Los puntos indican las muestras y los rectángulos agrupan a la mayoría de éstas. Las líneas en el centro de los rectángulos indican la mediana de cada grupo.

7.2.- Secuenciación de las semillas arqueológicas

Bibliotecas de secuenciación fueron construidas y secuenciadas exitosamente para 7 de las muestras con concentraciones de ADN mayor a 1 ng/μl y 1 muestra con concentración de 0.383 ng/μl. Las lecturas de secuenciación obtenidas fueron mapeadas al genoma de referencia de *P. vulgaris* (Schmutz et al. 2014). La proporción de ADN endógeno varió entre el 2 y el 39.2%, correspondiendo a una profundidad de cobertura promedio de 1.9X. La muestra con mejor

profundidad de cobertura alcanzó $6.6\times$ (HL-12), mientras que la más baja fue de $0.13\times$ (Tar-5).

Para investigar la autenticidad de los datos de secuenciación estimamos la fracción de sustituciones C-T y G-A en los extremos de las lecturas utilizando mapDamage (Jónsson et al. 2013) (Figura 3). Todas las muestras demostraron un incremento de sustituciones del tipo C-T y G-A en los extremos, consistente con lo esperado de datos provenientes de muestras antiguas (fracción de daño entre 12 y 25 %). Adicionalmente, todas las muestras mostraron lecturas cortas (longitud promedio de 40 bp; Figura 3) consistente con datos provenientes de ADN fragmentado.

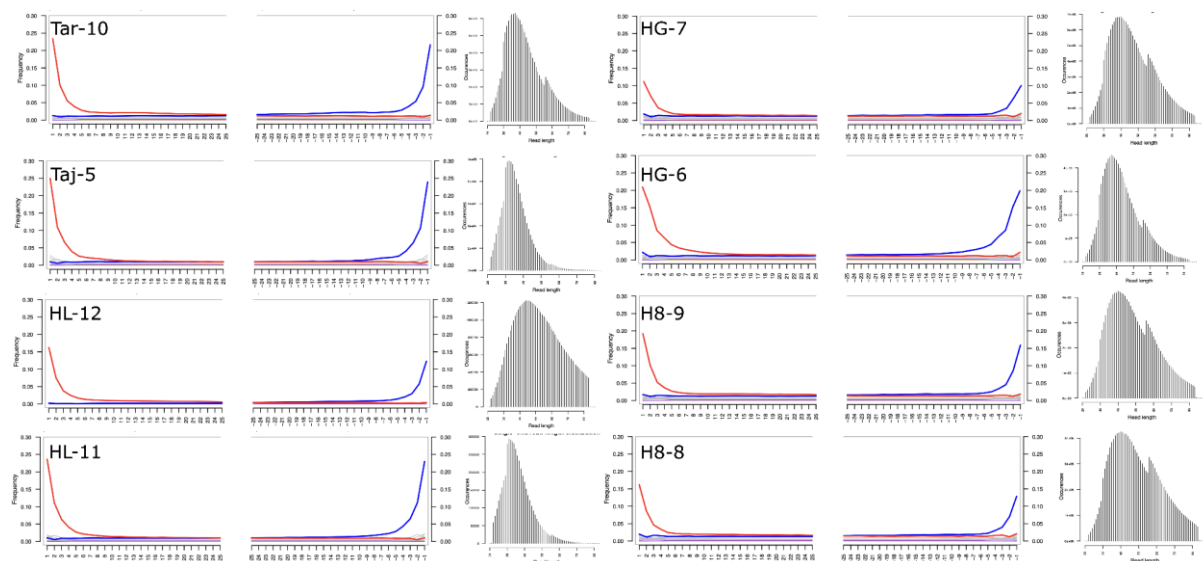


Figura 3. Patrones de daño de los genomas antiguos. Las gráficas de la izquierda muestran el incremento de sustituciones C->T y G->A en las lecturas de secuenciación en los extremos 5'y 3' respectivamente. El histograma de la derecha muestra la distribución de longitud de las lecturas de secuenciación.

7.3.- Ancestría genómica de las semillas antiguas

Para investigar la ancestría de las semillas antiguas, combinamos los genomas antiguos con genomas de especies relevantes de *Phaseolus* (Tabla 3) y realizamos un análisis MDS (Malaspinas et al. 2014) (Figura 4 A). El MDS separó los genomas por especie, donde la dimensión 1 separó *P. lunatus* del resto de las especies, y la dimensión 2 separó *P. acutifolius* de *P. vulgaris*. Lo anterior es consistente con las relaciones filogenéticas descritas previamente (Delgado-Salinas et al. 2006), donde *P. lunatus* representa el grupo más divergente entre las especies representadas en el MDS. Siete de las ocho semillas secuenciadas se agrupan en el MDS con genomas modernos de *P. lunatus* con excepción de la semilla (HL-12) del sitio Huaca de Loro (Nazca) que se agrupa con *P. vulgaris*.

Para profundizar en el análisis de la ancestría de los genomas antiguos, utilizamos ADMIXTURE (Alexander et al. 2009) para estimar ancestrías entre todos los genomas utilizados en el MDS (Figura 4 B). Las separaciones obtenidas en ADMIXTURE fueron consistentes con los grupos obtenidos en el MDS, con cada ancestría (representada como diferentes colores en la gráfica) representando una especie diferente. Las siete semillas antiguas identificadas como *P. lunatus* se agruparon junto con el grupo de *P. lunatus* de los Andes. Similarmente, la semilla identificada como *P. vulgaris*, se agrupó con *P. vulgaris* de los Andes. Cabe mencionar que ésta última comparte ancestría específicamente con el grupo de *P. vulgaris* domesticado, lo que sugiere que proviene de una planta en proceso de domesticación.

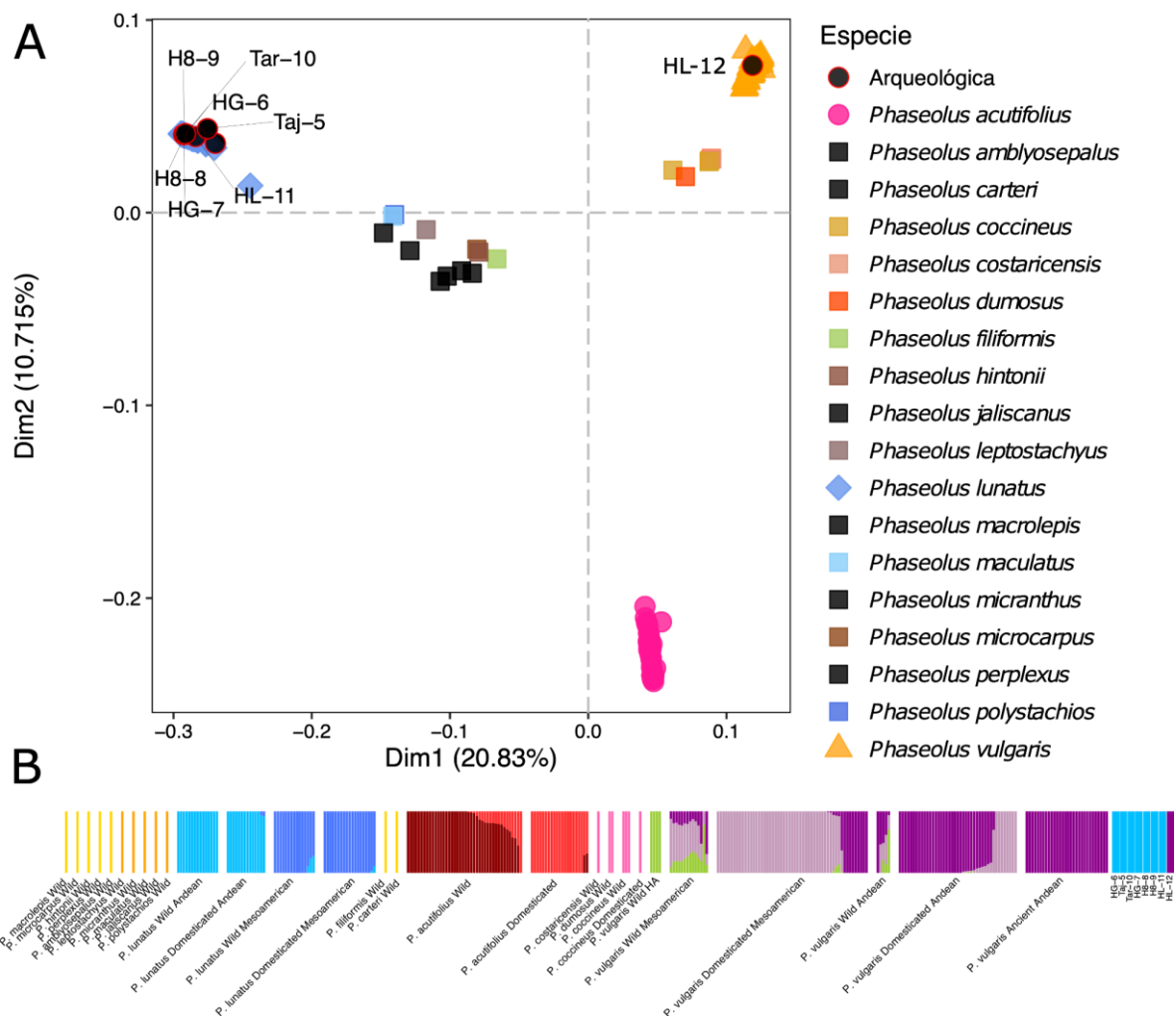


Figura 4. Análisis de MDS para los genomas modernos y antiguos de *Phaseolus*. Las muestras arqueológicas se muestran como círculos negros con borde rojo y se indican los nombres de cada muestra. Las especies modernas se muestran con diferentes colores como indica la leyenda de la figura.

8.- Conclusiones y perspectivas

De las 30 muestras procesadas 8 resultaron en concentraciones altas de ADN y fueron utilizadas para construir bibliotecas de secuenciación. Mientras que las concentraciones de ADN son bajas, éstas son consistentes con lo observado en restos macrobotánicos de otros sitios arqueológicos de períodos temporales similares (Ramos-Madrigal et al. 2024; Ramos-Madrigal 2026). Encontramos una correlación entre la edad de la muestra y la concentración de ADN. Los sitios arqueológicos H-8 (Ica) y Huaca Loro (Nazca)

dieron los mejores resultados para la extracción de ADN. No hay una correlación clara con la presencia de compuestos orgánicos visibles en el extracto y la concentración de ADN, sin embargo, queda por investigar si estos compuestos tienen algún efecto en la construcción de las bibliotecas de ADN.

Tres de las ocho semillas antiguas secuenciadas mostraron una proporción de ADN endógeno mayor al 10%, y para una de ellas fue posible generar un genoma con una profundidad de cobertura de 6X. Fue posible demostrar la autenticidad de ADN antiguo en todas las muestras. De las muestras secuenciadas, 7 fueron identificadas como *P. lunatus* de los Andes y 1 fue identificada como *P. vulgaris* domesticado de los Andes.

Una vez finalizado el análisis de los datos, en particular los análisis para investigar la presencia de rizobios en los datos antiguos, los resultados serán publicados en una revista revisada por pares de alto impacto.

9.- Referencias

- Alexander, David H., John Novembre, and Kenneth Lange (2009). Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Research* 19 (9): 1655–1664.
- Anders, M. (1990). Maymi: Un sitio del Horizonte Medio en el Valle de Pisco. *Gaceta Arqueológica*, 5(17), 27-39.
- Bachir, A. y Llanos, D. (2013). ¿Hacia un urbanismo paracas en Ánimas Altas/Ánimas bajas (valle de Ica)? *Boletín de Arqueología PUCP*, 17, 169-204.
- Bellucci, Elisa, Elena Bitocchi, Alberto Ferrarini, et al. (2014). Decreased Nucleotide and Expression Diversity and Modified Coexpression Patterns Characterize Domestication in the Common Bean. *The Plant Cell* 26 (5): 1901–1912.
- Beresford-Jones, D. G., Pullen, A. G., Whaley, O. Q., Moat, J., Chauca, G. E., Cadwallader, L., Arce, E. S., Orellana, A., Alarcón, C., Gorriti, M., Maita, P., Sturt, F., Dupeyron, A., Huaman, O., Lane, K. J., y French, C. (2015). Re-

- evaluating the resource potential of Lomas Fog Oasis environments for preceramic hunter-gatherers under past ENSO modes on the south coast of Peru. *Quaternary Science Reviews* 129, 196-215.
- Beresford-Jones, D. G., Pullen, Chauca, G. E., Cadwallader, L., García, M., Salvatierra, I., Whaley, O., Vásquez, v., Arce, S., Lane, K. y French, C. (2017). Refining the Maritime Foundations of Andean Civilization: How Plant Fiber Technology Drove Social Complexity During the Preceramic Period. *Journal Archaeological Method Theory*. DOI 10.1007/s10816-017-9341-3
- Beresford-Jones, D. G, Pomeroy, E., Alday, C., Benfer, R., Quilter, J., O'Connell, T., Lightfoot, E. (2021). Diet and Lifestyle in the First Villages of the Middle Preceramic: Insights from Stable Isotope and Osteological Analyses of Human Remains from Paloma, Chilca I, La Yerba III, and Morro I. *Latin American Antiquity*, 32(4), 741-759.
- Bitocchi, E. et al. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E788–96.
- Bitocchi, Elena, Elisa Bellucci, Alessandro Giardini, et al. (2013). Molecular Analysis of the Parallel Domestication of the Common Bean (*Phaseolus Vulgaris*) in Mesoamerica and the Andes. *The New Phytologist* 197 (1): 300–313.
- Bitocchi, Elena, Domenico Rau, Elisa Bellucci, et al. (2017). Beans (*Phaseolus* Ssp.) as a Model for Understanding Crop Evolution. *Frontiers in Plant Science* 8 (May): 722.
- Briggs, A. W. et al. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14616–14621.
- Carøe, Christian, Shyam Gopalakrishnan, Lasse Vinner, et al. (2018). Single-tube Library Preparation for Degraded DNA. *Methods in Ecology and Evolution* 9 (2): 410–419.
- Castro-Guerrero, Norma A., Mariel C. Isidra-Arellano, David G. Mendoza-Cozatl, and Oswaldo Valdés-López. (2016). Common Bean: A Legume Model on the Rise for Unraveling Responses and Adaptations to Iron, Zinc, and Phosphate

- Deficiencias. *Frontiers in Plant Science* 7 (May): 600.
- Chacón-Sánchez, M. I. & Martínez-Castillo, J. (2017). Testing Domestication Scenarios of Lima Bean (*Phaseolus lunatus* L.) in Mesoamerica: Insights from Genome-Wide Genetic Markers. *Front. Plant Sci.* **8**, 1551.
- Chang, Christopher C., Carson C. Chow, Laurent Cam Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. (2015). Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets. *GigaScience* 4 (February): 7.
- Conlee, C., Kellner, C., Walker, C. y Noriega, A. (2021). Early imperialism in the Andes: Wari colonisation of Nasca. *Antiquity*, 95 (384), 1527-1546. DOI: <https://doi.org/10.15184/aqy.2021.118>
- Cook, A. 1994. Informe Final de las Investigaciones de reconocimiento en la parte baja del valle de Ica, 1988-1990. Lima, Instituto Nacional de Cultura.
- Dabney, Jesse, Michael Knapp, Isabelle Glocke, et al. (2013). Complete Mitochondrial Genome Sequence of a Middle Pleistocene Cave Bear Reconstructed from Ultrashort DNA Fragments. *Proceedings of the National Academy of Sciences of the United States of America* 110 (39): 15758–15763.
- Delgado-Salinas, A., R. Bibler, and M. Lavin. (2006). Phylogeny of the Genus *Phaseolus* (Leguminosae): A Recent Diversification in an Ancient Landscape. *Systematic Botany* 31 (4): 779–791.
- Foyer, C. H., Nguyen, H. & Lam, H.-M. (2019). Legumes-The art and science of environmentally sustainable agriculture. *Plant Cell Environ.* **42**, 1–5.
- Frascarelli, Giulia, Teresa R. Galise, Nunzio D'Agostino, et al. (2025). The Evolutionary History of the Common Bean (*Phaseolus Vulgaris*) Revealed by Chloroplast and Nuclear Genomes Analysis. *Theoretical and Applied Genetics* 138 (2): 47.
- Gepts, P. (2010). Crop domestication as a long-term selection experiment. in *Plant Breeding Reviews* 1–44 (John Wiley & Sons, Inc., Oxford, UK).

- Gorbahn, H. (2013). El sitio de Pernil Alto del Arcaico Medio en el sur del Perú: Comienzo de Horticultura y Sedentarismo en condiciones del Holoceno Medio. *Dialogo Andino*, 41, 61-82.
- Hansen, H. B. *et al.* (2017). Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS One* **12**, e0170940.
- Hayat, Rifat, Safdar Ali, Ummay Amara, Rabia Khalid, and Iftikhar Ahmed. (2010). Soil Beneficial Bacteria and Their Role in Plant Growth Promotion: A Review. *Annals of Microbiology* 60 (4): 579–598.
- Jensen, Theis Z. T., Jonas Niemann, Katrine Højholt Iversen, *et al.* (2019). A 5700 Year-Old Human Genome and Oral Microbiome from Chewed Birch Pitch. *Nature Communications* 10 (1): 5520.
- Jónsson, Hákon, Aurélien Ginolhac, Mikkel Schubert, Philip L. F. Johnson, and Ludovic Orlando. (2013). mapDamage2.0: Fast Approximate Bayesian Estimates of Ancient DNA Damage Parameters. *Bioinformatics* 29 (13): 1682–1684.
- Kates, Heather R., Brian C. O'Meara, Raphael LaFrance, *et al.* (2024). Shifts in Evolutionary Lability Underlie Independent Gains and Losses of Root-Nodule Symbiosis in a Single Clade of Plants. *Nature Communications* 15 (1): 4262.
- Kircher, Martin, Susanna Sawyer, and Matthias Meyer. (2012). Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform. *Nucleic Acids Research* 40 (1): e3.
- Koinange, E. M. K., S. P. Singh, and P. Gepts. (1996). Genetic Control of the Domestication Syndrome in Common Bean. *Crop Science* 36 (4): 1037–1045.
- Korneliussen, Thorfinn Sand, Anders Albrechtsen, and Rasmus Nielsen. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15 (November): 356.
- Lasaponara, R., Masini, N. y Orefici, G. (2016). *The Ancient Nasca World. New Insights from Science and Archaeology*. Springer.
- Li, Heng, and Richard Durbin. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* 25 (14): 1754–1760.
- Li, Heng, Bob Handsaker, Alec Wysoker, *et al.* (2009). The Sequence

- Alignment/Map Format and SAMtools. *Bioinformatics* 25 (16): 2078–2079.
- Llanos, D. (2017). La readaptación cultural Paracas a la entidad Nasca: un enfoque desde el asentamiento de Cerro Córdova. *Boletín de Arqueología PUCP*, 22, 159-190.
- Lobaton, Juan David, Tamara Miller, Juanita Gil, et al. (2018). Resequencing of Common Bean Identifies Regions of Inter-Gene Pool Introgression and Provides Comprehensive Resources for Molecular Breeding. *The Plant Genome* 11 (2): 170068.
- López-López, A., Rogel, M. A., Ormeño-Orrillo, E., Martínez-Romero, J. & Martínez-Romero, E. (2010). Phaseolus vulgaris seed-borne endophytic community with novel bacterial species such as Rhizobium endophyticum sp. nov. *Syst. Appl. Microbiol.* **33**, 322–327.
- Lozano-Arce, Daniela, Tatiana García, Laura Natalia Gonzalez-Garcia, Romain Guyot, Maria Isabel Chacón-Sánchez, and Jorge Duitama. (2023). Selection Signatures and Population Dynamics of Transposable Elements in Lima Bean. *Communications Biology* 6 (1): 803.
- Malaspinas, A.-S. et al. (2014). bammds: a tool for assessing the ancestry of low-depth whole-genome data using multidimensional scaling (MDS). *Bioinformatics* **30**, 2962–2964.
- Mallco, R. (2015). Investigaciones arqueológicas en Tacaraca, un sitio inca de la costa sur del Perú. *Arqueología y Sociedad*, 29, 71-106.
- Mallco (2019). La ocupación inka en el valle de Ica, perspectivas e interpretaciones en base a investigaciones realizadas en Tacaraca y Huaca Guadalupe. *ArqPerCOARPE*, 1, 275-291.
- Martínez-Romero, Esperanza. (2003). Diversity of Rhizobium-Phaseolus Vulgaris Symbiosis: Overview and Perspectives. *Plant and Soil* 252 (1): 11–23.
- Massey, S. (2022). Tajahuana: nuevas perspectivas sobre un sitio paracas en el valle medio de Ica. En Actas del VII Congreso Nacional de Arqueología (pp. 295-307). Lima: Ministerio de Cultura.
- Menzel, D. (1976). *Pottery Style and Society in Ancient Peru*. University of California Press Berkeley, Los Angeles.

- Moghaddam, Samira Mafi, Atena Oladzad, Chushin Koh, et al. (2021). The Tepary Bean Genome Provides Insight into Evolution and Domestication under Heat Stress. *Nature Communications* 12 (1): 2638.
- Mora, Y. et al. (2014). Nitrogen-fixing rhizobial strains isolated from common bean seeds: phylogeny, physiology, and genome analysis. *Appl. Environ. Microbiol.* **80**, 5644–5654.
- Motta-Aldana, J. R. et al. (2010). Multiple origins of Lima bean landraces in the Americas: Evidence from chloroplast and nuclear DNA polymorphisms. *Crop Sci.* **50**, 1773–1787.
- Nielsen, S. H. et al. (2021) 31,600-year-old human virus genomes support a Pleistocene origin for common childhood infections. *bioRxiv* 2021.06.28.450199 doi:10.1101/2021.06.28.450199.
- Orlando, Ludovic, Robin Allaby, Pontus Skoglund, et al. 2021. “Ancient DNA Analysis.” *Nature Reviews. Methods Primers* 1 (1): 1–26.
- Pickrell, J. K. & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967.
- Ramos-Madrigal, Jazmín. (2026). The Potential of Plant Palaeogenomic Research. *Current Opinion in Plant Biology* 90 (102856): 102856.
- Ramos-Madrigal, Jazmín, Gayle J. Fritz, Bryon Schroeder, et al. (2024). The Genomic Origin of Early Maize in Eastern North America. *Cell*, ahead of print, November 26. <https://doi.org/10.1016/j.cell.2024.11.003>.
- Ramos-Madrigal, J. et al. (2019). Palaeogenomic insights into the origins of French grapevine diversity. *Nat Plants* **5**, 595–603.
- Ramos-Madrigal, J. et al. (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.
- Reindel, M e Isla, J. (2009). El Periodo Inicial en Pernil Alto, Palpa, costa sur del Perú. *Boletín de Arqueología PUCP*, **13**, 259-288.
- Rendón-Anaya, Martha, Josaphat M. Montero-Vargas, Soledad Saburido-Álvarez, et al. (2017). Genomic History of the Origin and Domestication of Common Bean Unveils Its Closest Sister Species. *Genome Biology* 18 (1): 60.

- Santana, L., French, C., Bresford-Jones, D., y Rengifo, C. (2022). Geoarchaeological interpretation of the formation and environment of Samaca H-8, Ica, Valley, south coast Peru (c. AD 900-1550). *Boletín de la Sociedad Geológica Mexicana*, 74(3),
- Schmutz, J. *et al.* (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* **46**, 707–713.
- Schubert, Mikkel, Stinus Lindgreen, and Ludovic Orlando. 2016. “AdapterRemoval v2: Rapid Adapter Trimming, Identification, and Read Merging.” *BMC Research Notes* 9 (February): 88.
- Trucchi, E. *et al.* (2021). Ancient genomes reveal early Andean farmers selected common beans while preserving diversity. *Nat Plants* **7**, 123–128.
- Wales, N., Andersen, K., Cappellini, E., Avila-Arcos, M. C. & Gilbert, M. T. P. (2014). Optimization of DNA recovery and amplification from non-carbonized archaeobotanical remains. *PLoS One* **9**, e86827.
- Wales, N. *et al.* (2015). New insights on single-stranded versus double-stranded DNA library preparation for ancient DNA. *Biotechniques* **59**, 368–371.
- Williams, C. y Pazos, M. (1974). Inventario, catastro y delimitación del patrimonio arqueológico del valle de Ica. Instituto Nacional de Cultura, Centro de Investigación y Restauración de Bienes Monumentales. Lima.