

Una última entrevista: el seguimiento molecular a una pieza ósea de un contexto prehispánico

Rodrigo Barquera Lozano*

La metamorfosis que sufre un fragmento óseo, desde su salida del microambiente en el que se encontraba en su contexto arqueológico hasta convertirse en parte de la discusión de un artículo o libro publicado, es muchas veces ajena no sólo al público general, sino a participantes de los equipos académicos que custodian, analizan o estudian los resultados derivados de los individuos de contextos arqueológicos. Con la intención de acercar a todos los actores y espectadores a estas transformaciones de la pieza arqueológica, tanto materiales como conceptuales, el autor se dio a la tarea de realizar un seguimiento a través de todos los pasos analíticos, complementando con material fotográfico para hacer accesible la experiencia del DNA antiguo más allá de los artículos y reportes, esto con el fin de exponer el trayecto que une a la pieza ósea con su participación en la discusión científica. El DNA antiguo (aDNA) es el término con el que se designa al material genético recuperado a partir de organismos, o partes de ellos, que perecieron tiempo atrás y que posee una serie de cambios característicos en su estructura fisicoquímica que son producto de los procesos tafonómicos relacionados con su ambiente inmediato. Aunque el estudio en el campo de las moléculas antiguas es relativamente reciente —1984 marca su inicio con la primera recuperación de aDNA a partir de un espécimen de museo—, se ha reinventado en varias ocasiones y para estudiar el pasado ha echado mano de los avances tecnológicos asociados a la biología molecular, la genética de poblaciones y, más recientemente, la genómica y la proteómica. Primero en materiales de otras especies, y luego en restos humanos, los estudios de aDNA han transformado la forma de aproximarnos al estudio del pasado y también han evolucionado con el tiempo para adaptarse a las exigencias de las nuevas preguntas de investigación en el área.

* Department of Archaeogenetics, Max Planck Institute for the Science of Human History (MPI-SHH), Jena, Alemania (barquera@shh.mpg.de). Agradezco a Oana del Castillo Chávez, Diana Iraíz Hernández Zaragoza, Laura Ramírez Vázquez, Ana Ruiz, Sylvia Arnold-El Fehri, Beate Kerpen, Anke Trinkler, Vanessa Villalba, Susanna Sabin, Theas Lamnidis, James Fellows Yates, Eirini Skourtanioti, Cosimo Posth, Maria Spyrou, Stephen Clayton, Chuan-Chao Wang, Franziska Aron, Marta Burri, Cäcilia Freund, Rita Radzeviciute, Raphaela Stahl, Antje Wissgott, Guido Brandt, Stephan Schiffels, Christina Warinner, Wolfgang Haak y Johannes Krause por su invaluable apoyo en distintas partes de este proyecto. El trabajo en este material se encuentra aprobado por el Consejo de Arqueología del Instituto Nacional de Antropología e Historia (No. de oficio 401.1S.3-2017/482, oficio de exportación 401-3-6194, número de expediente de exportación AA-39-17 A). Dedicó el presente trabajo a la memoria de Jorge Balbuena.

Con la recuperación de material genético a partir de muestras de un *quagga* (*Equus quagga*), miembro extinto de la familia de los equinos (Higuchi *et al.*, 1984), se demostró por primera vez que la obtención de aDNA era posible. Sin embargo, las posibles aplicaciones en el estudio de la historia y de la prehistoria dejaron de ser anecdóticas y se volvieron prioridad en el campo con la primera extracción de material genético a partir de restos humanos momificados con una antigüedad de 2 500 años (Pääbo, 1985a; 1985b), que demostró que este tipo de análisis no necesariamente estaba restringido a ejemplares de museo y que podrían usarse eficientemente para el estudio de restos humanos del pasado distante. Han pasado 34 años desde aquellos estudios pioneros en el área y se han sucedido revoluciones tecnológicas que han impactado en los alcances y limitaciones en este campo. El advenimiento de las técnicas de secuenciación masiva paralela y el uso de estrategias de enriquecimiento selectivo han generado una mejora sustancial en la cantidad y calidad de los datos obtenidos, por lo que con técnicas menos invasivas o a partir de materiales novedosos se pueden responder preguntas que hasta hace poco hubiera sido inimaginable siquiera esbozar.

Sin embargo, entre mayor sea la complejidad de la herramienta, más restringido su acceso a la comunidad en general e incluso a la misma comunidad académica. Es con esta limitación que nace la idea del presente trabajo, en el que se expone el *seguimiento* a un fragmento óseo desde su incorporación a la línea de trabajo hasta su presentación como parte de una figura en un trabajo de investigación. Es, por decirlo de cierta forma, el seguimiento a las técnicas empleadas para realizar una última entrevista al individuo del que alguna vez formó parte, y que nos permitirá acercarnos a la historia de su vida, de sus ancestros, de sus enfermedades e incluso de su muerte, y asimismo puede relatarnos algunas cuestiones en torno a la organización social, a los rituales, a la estructura poblacional y a los procesos de migración y/o mestizaje de los que su comunidad fue partícipe.

El inicio del viaje (figura 1)

Poco podemos imaginarnos acerca de los contextos arqueológicos de donde provienen los individuos que estudiamos antes de comenzar con los análisis. ¿Quiénes eran? ¿Qué hacían ahí? ¿Cómo llegaron a ese momento y a ese lugar? ¿Por qué terminaron en esas circunstancias? No todo lo que nos preguntamos puede ser respondido, sin importar si se aborda con “técnicas clásicas” o con el uso de las llamadas “herramientas moleculares”. Pero entre las cuestiones que sí podemos responder, y tal vez las que causan más expectativa, son aquellas que tienen que ver con la causa de muerte y el estilo de vida de esos antepasados distantes. Ya decía Ernest Hemingway que “la vida de cada hombre termina de la misma manera. Son sólo los detalles de cómo vivió y cómo murió lo que distingue a un hombre de otro” (Hotchner, 1966). Y son precisamente esos detalles los que tratamos de conocer a partir de los análisis moleculares. Generalmente no se realiza el análisis a



Figura 1. El inicio del viaje, MPI-SHH, 2018. Fotografía © R. Barquera.

una única pieza, sino que, para tener una idea más completa del panorama, se trabaja con varios individuos que provienen del mismo contexto arqueológico.

En esta imagen podemos apreciar una gran cantidad de piezas óseas (54 para ser exactos), que son fragmentos del hueso temporal izquierdo a los que se encuentra unida la porción petrosa, la cual ha demostrado ser de las unidades anatómicas en las que podemos encontrar DNA en mayor abundancia y mejor estado de preservación (Pinhasi *et al.*, 2015). Cada pieza ha sido limpiada con cloro diluido y agua destilada, irradiada por 15 minutos con luz ultravioleta, empacada individualmente y registrada. Se le ha asignado ahora un número que la une inequívocamente con su identificación arqueológica y que la acompañará por el resto de los procedimientos analíticos. Las dimensiones que por ahora ocupan nuestros *entrevistados* son aproximadamente 5 litros de volumen y 686.6 gramos. Estos datos se obtienen posteriormente a la limpieza de cada pieza ósea, en la que se retiran restos de tierra y, en algunos casos, consolidantes y otros materiales.

Dimensiones del lote: 54 piezas, 5 L (V), 686.6 g (M).

Dimensiones del entrevistado: aprox. 7.4 cm, 14.98 g.

Nuestro entrevistado (figura 2)

¿Qué preguntas podremos hacer, y cómo hacerlas, a nuestro pequeño viajero en el tiempo? Con base en el contexto arqueológico, los datos previamente obtenidos mediante el estudio del microambiente, de los materiales acompañantes, de la determinación de la naturaleza (primaria o secundaria) del entierro, datos sobre marcas de corte, mutilaciones u otras modificaciones *ante, peri o post mortem* ayudan a refinar el tipo de preguntas que deseamos realizar. Nuestro entrevistado tenía una edad de entre 4 y 8 años, estimación derivada a partir del estudio de huesos largos, aunque el contexto no permite trazar inequívocamente una relación entre un hueso largo particular y la porción petrosa de nuestro invitado. A falta de datos sobre su identidad, se le ha asignado un nombre: Chultún 01. Chultún 01 fue encontrado dentro de una estructura subterránea semejante a una cueva, pero de origen artificial, denominada precisamente "Chultún". De acuerdo con lo que se sabe de su contexto, Chultún 01 fue depositado ahí tiempo después de su muerte en una modalidad conocida como entierro secundario colectivo u osario (Márquez, 2010). Dado el tipo de depósito de los huesos, es extremadamente complicado encontrar relaciones entre los distintos huesos que pertenecen a cada individuo, por lo que la única estrategia para asegurar que sólo estamos *entrevistando* una vez a cada individuo es tomar una pieza ósea que pueda ser identificada como única para cada individuo. En este caso, decidimos que esa pieza sería el hueso temporal izquierdo de cada individuo, el cual es fácilmente lateralizable y por tanto nos permite garantizar que cada porción petrosa proviene de un individuo distinto. Las piezas óseas fueron empacadas individualmente, registradas, inventariadas y reportadas ante el Consejo de

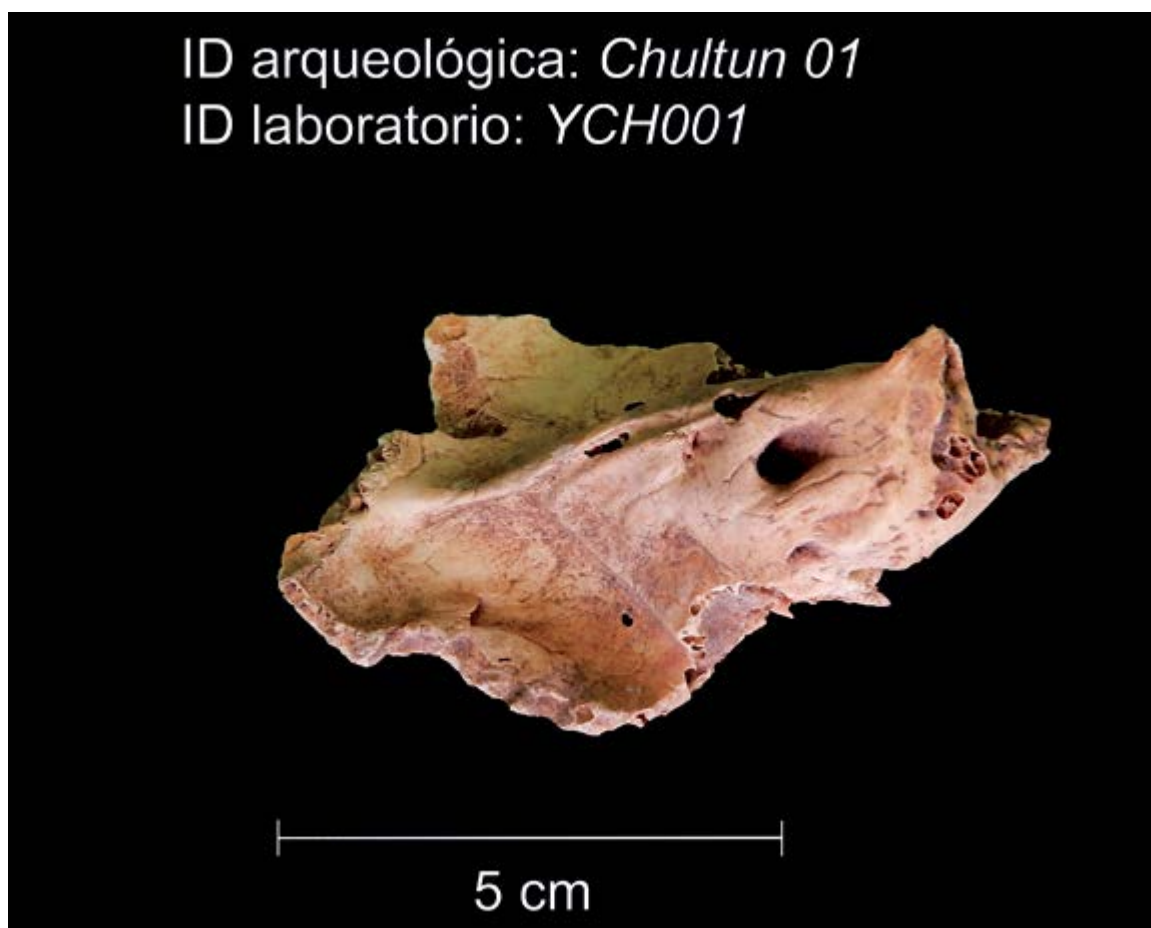


Figura 2. *Nuestro entrevistado*, MPI-SHH, 2018. Fotografía © R. Barquera.

Arqueología antes de volar a Alemania, donde el proceso de análisis molecular se llevó a cabo en las instalaciones del Instituto Max Planck de Ciencias de la Historia Humana en Jena. A su llegada, se les recibió con un distintivo especial que consiste en tres letras y un número de tres dígitos que será su identificación durante su trayecto por el instituto.

Nuestro primer *entrevistado* fue identificado como *YCH001* y a partir de este momento nos referiremos a él con este “nombre”. Tras haber tomado una serie de fotografías que permiten preservar una imagen de la pieza antes de proceder al muestreo, se introduce en una bolsa nueva y se coloca en una caja de transferencia donde la bolsa es irradiada con luz UV para inutilizar cualquier cadena de DNA externa y evitar así la contaminación con DNA moderno. La caja de transferencia se abre desde dentro de las instalaciones del laboratorio de DNA antiguo, también conocido como *cuarto limpio*, y se traslada al cuarto de muestreo. Este cuarto limpio está acondicionado para el trabajo con material antiguo y alberga una serie de implementaciones que contribuyen a minimizar la posibilidad de contaminación con DNA moderno, tales como un número determinado de horas diarias de irradiación UV en todas las instalaciones, áreas especializadas

destinadas a una única operación unitaria (cuarto de muestreo, cuarto de extracción de DNA, etcétera) y un sistema de filtración de aire y agua que impide la entrada de material genético contaminante del exterior. Dentro, se siguen rigurosos procesos para minimizar la contaminación por parte del equipo científico y para evitar la contaminación entre muestras, como el uso de sistemas de exclusas, uso de ropa exclusiva para el trabajo dentro del laboratorio, cambio de calzado entre exclusas, trajes de cuerpo entero, tres pares de guantes, lentes protectores, mascarillas para cubrir nariz y boca, cofias para el cabello y protocolos para vestirse, ingresar y transferir elementos desde el mundo exterior hasta el corazón del laboratorio.

Dimensiones del lote: 54 piezas, 5 L (V), 686.6 g (M)

Dimensiones del entrevistado: aprox. 7.4 cm, 14.98 g

Preguntas invasivas (figura 3)

Para poder realizar las preguntas de manera adecuada y obtener respuestas claras, es necesario retirar todo aquello que pueda crear confusión, falsos resultados o minimizar la cantidad de información de YCH001 con respecto al resto de DNA que podemos encontrar en la muestra ana-

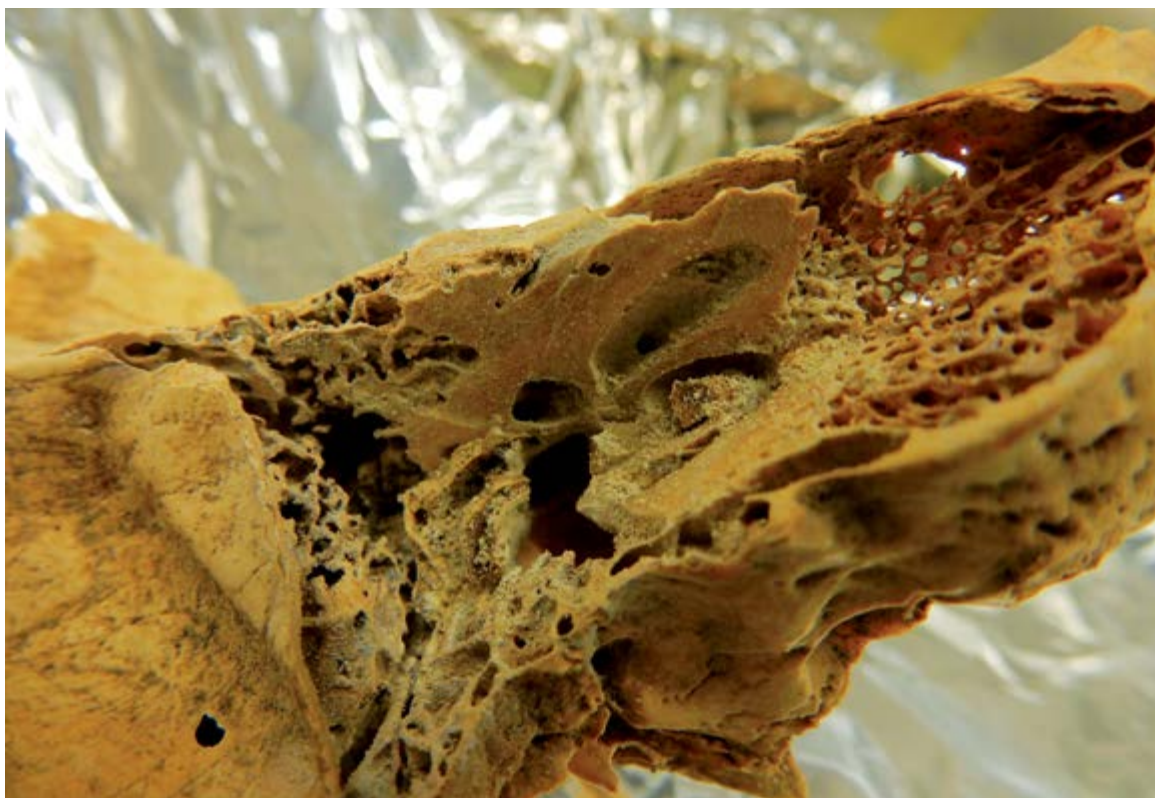


Figura 3. Preguntas invasivas, MPI-SHH, 2018. Fotografía © R. Barquera.

lizada. Tras seguir el protocolo de limpieza y descontaminación de superficies en el cuarto de muestreo, es hora de comenzar nuestra entrevista. Con ayuda de una sierra circular realizamos un corte longitudinal a través de la porción petrosa con la intención de dejar al descubierto el área más densa del hueso, que es la que contiene más material genético y también el mejor preservado. Debemos evitar las partes porosas porque no sólo no contienen tanto material genético como la región más densa, sino que pueden albergar partículas de tierra o polvo, restos vegetales o animales, y otro tipo de sustancias que, si son procesadas junto con el hueso de nuestro entrevistado, disminuirán la cantidad de DNA propio de la muestra (denominado *DNA endógeno*) y dificultará más el proceso para obtener la información genética necesaria. Todo este procedimiento se lleva a cabo dentro de una campana que minimiza la posibilidad de contaminación al mantener condiciones controladas en el ambiente inmediato a la pieza ósea.

Dimensiones del lote: 54 piezas, 5 litros (V), 686.6 gramos (M)

Dimensiones del entrevistado: aprox. 7.4 cm, 14.98 g

Hecho polvo (figura 4)

Con la ayuda de una broca con punta de diamante, se taladra la porción más densa del hueso para obtener un polvo fino que contiene nuestro material genético. Este polvo se pesa (se trabaja a partir de entre 50 y 100 mg de polvo) y se coloca en un tubo de plástico en preparación para las siguientes fases analíticas. Entre cada muestreo de distintos individuos se debe seguir un riguroso protocolo de descontaminación que asegure que no haya material genético en la campana de trabajo de muestras previas al momento de muestrear cada individuo. Si llegara a existir alguna partícula de polvo o tierra, se retira con ayuda de material de muestreo exclusivo para cada pieza (pinzas o espátulas), mismo que es lavado con cloro, enjuagado e irradiado antes de poder ser utilizado en otra muestra. El polvo de hueso está ahora listo para la extracción de su material genético y esperará en un congelador a -20°C hasta que ese paso ocurra. El resto de la pieza ósea también se almacena a esa temperatura para garantizar la mejor preservación del material genético en caso de ser necesario repetir un muestreo. Algunas veces puede elegirse una porción del material óseo para realizar otros análisis complementarios, como la datación con ^{14}C (radiocarbono), el análisis de isótopos estables o extracción de proteínas para análisis proteómicos. Cada muestra es un mundo particular, por lo que no existe un procedimiento único estándar para decidir en qué región muestrear, o qué pieza debe ser la indicada. La preservación depende del microambiente de cada pieza, el cual puede variar de manera significativa de entierro en entierro, de individuo en individuo e incluso entre distintos huesos del mismo individuo. El criterio de la persona que realiza el muestreo es fundamental para maximizar las probabilidades de obtener material genético analizable en nuestra muestra de polvo de hueso.



Figura 4. Hecho polvo, MPI-SHH, 2018. Fotografía © R. Barquera.

Dimensiones del lote: 63 piezas más controles y blancos, aprox. 12 ml (V), 6.006 gramos (M)

Dimensiones del entrevistado: aprox. 200 μ l, 90 mg

La historia de una vida en un susurro (figura 5)

A pesar de todas las precauciones y el empleo de tecnología de punta, la norma indica que la cantidad de DNA antiguo a obtener de cada muestra será baja. Por ello suelen realizarse varias extracciones de una misma pieza, o de varias, para maximizar la cantidad de DNA obtenido de cada individuo. El proceso de extracción comienza con la descalcificación de la muestra, lo que permite liberar el DNA de la matriz ósea en una solución acuosa. Esta solución acuosa contendrá todo el material genético de la muestra, sea del individuo, de algún patógeno que haya estado presente en la pieza al momento de su muerte, de microbiota acompañante, de microorganismos u otros seres vivos en el entorno inmediato al momento de la muerte o después de la misma, o de la o las personas que hayan entrado en contacto con la pieza ósea desde la excavación hasta su ingreso al laboratorio. Tras un incubado prolongado con sustancias que retiran el calcio de los huesos y enzimas que degradan



Figura 5. *La historia de una vida en un susurro*, MPI-SHH, 2018. **Fotografía** © R. Barquera.

las proteínas del hueso, el DNA se purifica con ayuda de columnas de silicato a las que el material genético recuperado se une y permite realizar lavados para retirar la mayoría de los posibles contaminantes. Al terminar los lavados, el DNA se eluye (es decir, se separa de la columna) con ayuda de una disolución que además de recuperarlo, ayudará a que no se degrade por cambios en el pH del medio. Entre mayor sea la cantidad de DNA “propio” de la muestra, mayor cantidad de preguntas moleculares podremos realizar a nuestro “entrevistado”. Algunas veces la cantidad de DNA endógeno de una muestra en particular es tan bajo que no se le pueden formular preguntas en lo absoluto.

Dimensiones del lote: 30 tubos, aprox. 3 ml (V)

Dimensiones del entrevistado: 100 μ l

Varias versiones del mismo relato (figura 6)

Como ya se ha mencionado, a veces es mejor realizar una encuesta para conocer el contexto general, en lugar de levantar una única entrevista. En el caso del entierro del chultún de Chichén

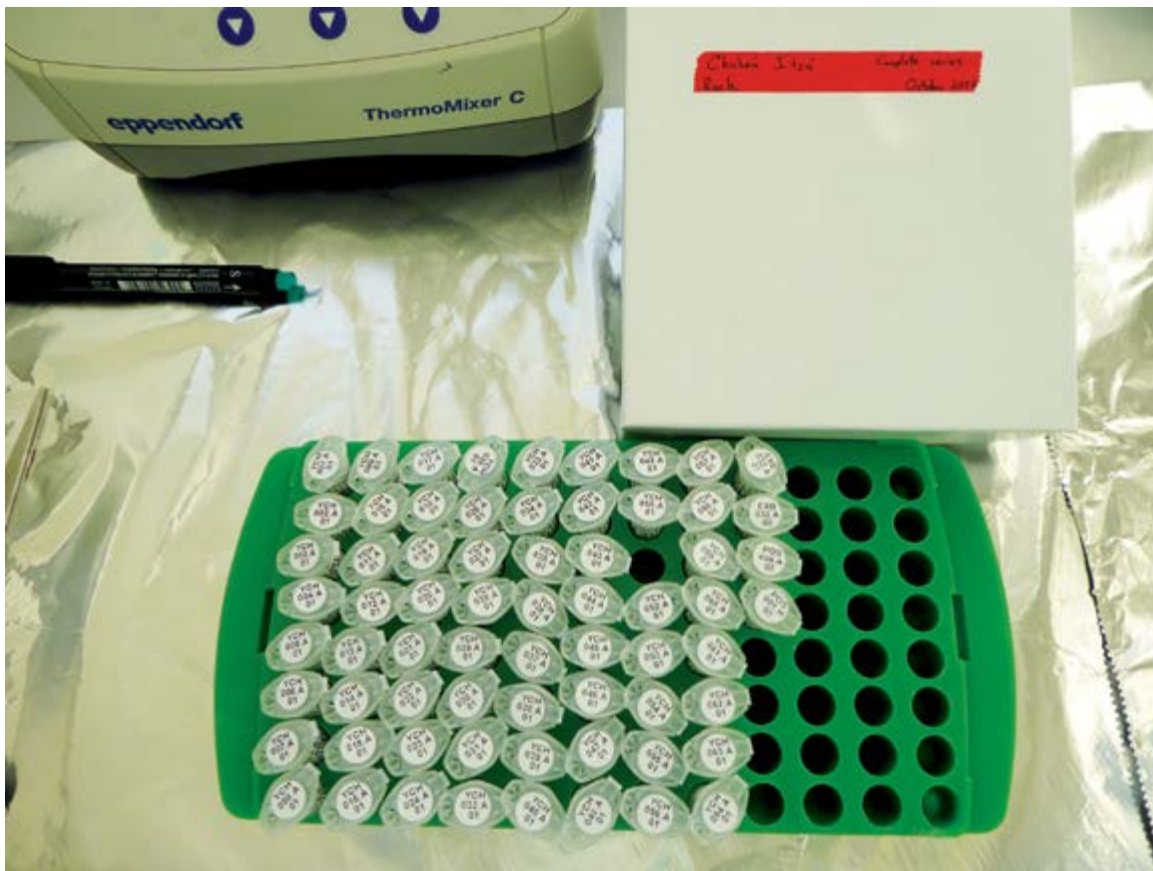


Figura 6. Varias versiones del mismo relato, MPI-SHH, 2018. Fotografía © R. Barquera.

Itzá, ésta es la estrategia empleada. Además de YCH001, “entrevistamos” a otros 63 individuos acompañantes que nos permitirán conocer con más detalle lo ocurrido en ese contexto y, posiblemente, responder algunas preguntas de interés arqueológico y/o antropológico. En la imagen podemos apreciar la colección entera, cada muestra en su tubo, así como blancos de extracción y controles positivos. Los blancos de extracción son tubos en los que se ha seguido el mismo proceso que con cualquiera de las muestras, excepto que no se agregó ningún tipo de muestra. El blanco sirve para evaluar si nuestros métodos de muestreo, extracción y purificación fueron adecuados, pues si en los blancos encontráramos DNA significaría que se ha transferido material genético de alguna fuente a estos tubos y, por lo tanto, corremos el riesgo de que todo nuestro lote de muestras esté igualmente contaminado. El control positivo, por otro lado, es polvo de un hueso largo de oso de las cavernas (*Ursus spelaeus*) del que se sabe que contiene DNA. Al procesarse de la misma manera que las muestras, debemos obtener DNA en concentraciones consistentes mediante los distintos lotes de extracción. Si en un determinado lote ninguna muestra posee DNA detectable, pero nuestro control positivo sí, entonces quiere decir que esa colección de muestras no posee material genético analizable. Si el control positivo tampoco presenta DNA, quiere decir

que hemos hecho algo mal durante el proceso y por tanto debemos buscar la causa del problema y repetir el análisis del lote entero.

Dimensiones del lote: 30 tubos, aprox. 3 ml (V)

Dimensiones del entrevistado: 100 μ l

Edición (figura 7)

El material genético antiguo auténtico posee ciertas características que derivan del daño esperado ocurrido a lo largo de las decenas, centenas o millares de años en que el DNA estuvo expuesto a condiciones ambientales poco decuadas para su preservación. Incluso en los mejores escenarios, se espera la presencia de daño en las moléculas de DNA producto de reacciones químicas características, tales como la ruptura de las cadenas de DNA, la pérdida de ciertas partes de la molécula —como los grupos amino ($-\text{NH}_2$) en las bases nitrogenadas— y daños oxidativos. Todo esto hace que la molécula típica de DNA antiguo tenga un tamaño más bien corto, con alta presencia de transiciones

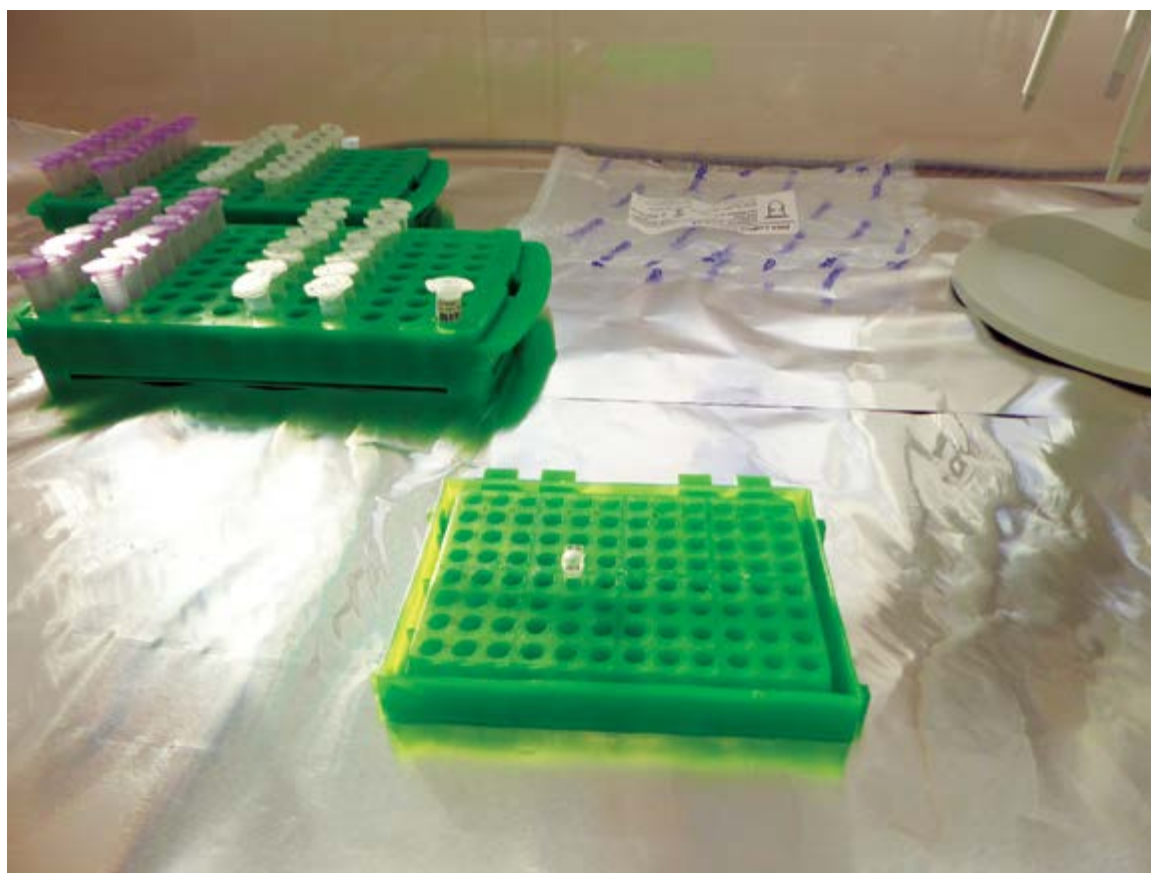


Figura 7. Edición, MPI-SHH, 2018. Fotografía © R. Barquera.

(cambio de citosinas por uracilo o hidroxipuracilo, por ejemplo), lo que genera sustituciones aparentes de citosinas por timinas (Hofreiter *et al.*, 2001) y daños en los extremos de la cadena. Aunque pudiera parecer algo negativo, en realidad estas características nos permiten distinguir al DNA antiguo del DNA moderno y, por lo tanto, contar con criterios de autenticidad al momento de analizar los resultados obtenidos. Sin embargo, para poder analizar las moléculas de DNA de nuestros entrevistados debemos “repararlas”. La reparación del DNA se realiza como parte de la generación de las denominadas bibliotecas genómicas, que no son otra cosa sino la totalidad del material genético obtenido de nuestra muestra, aún sin distinguir si pertenece de forma exclusiva al individuo analizado o no. Los pasos incluyen la remoción enzimática de uracilos de la cadena, la restauración de cadenas sencillas en cadenas dobles y la adición de adaptadores, necesarios para los procesos de secuenciación masiva de las bibliotecas genómicas en una serie de incubaciones sucesivas acompañadas de pasos de purificación para remover las enzimas empleadas en el paso anterior (Briggs *et al.*, 2010).

Dimensiones del lote: 63 tubos, aprox. 2.4 ml (V)

Dimensiones del entrevistado: 38 μ l en un tubito de 200 μ l

Etiquetas (figura 8)

Para poder continuar con el proceso, cada palabra, es decir, cada fragmento de DNA, debe ser identificado como parte de esa única entrevista. Para ello se colocan índices, que son pequeños fragmentos de DNA sintético con una secuencia única que de manera particular identificará a cada biblioteca genómica. Todos los fragmentos de DNA obtenidos a partir de la muestra de hueso de YCH001 contienen las mismas secuencias de identificación en sus extremos, sin importar su origen. De esta manera, la biblioteca genómica ahora posee dos características: resulta en extremo difícil contaminarla, y es *inmortal*. Contaminarla es complicado porque necesitaríamos más DNA con esa combinación de índices de una fuente distinta para lograrlo. Es inmortal porque a partir de este punto pueden realizarse tantas copias *in vitro* como sean necesarias por medio de la reacción en cadena de la polimerasa, o PCR por sus siglas en inglés. A partir de este punto no sólo no son necesarias las instalaciones del cuarto limpio, sino que no se debe realizar la síntesis de DNA *in vitro* dentro de éste, pues cualquier contaminación con un número elevado de moléculas dentro de estas instalaciones sería difícil de contener. Cabe mencionar que esta síntesis *in vitro* de DNA se conoce también como *amplificación*. Por ello se sacan los tubos con las reacciones listas para correr y en tubos perfectamente sellados, pero el proceso —que se lleva a cabo en un termociclador— ocurre fuera del cuarto limpio en instalaciones físicamente separadas. En ellas se introducirán los índices por medio de una reacción enzimática mediada por ciclos de temperaturas (Meyer y Kircher, 2010). Dado que cada par de índices se incorpora a una molécula de DNA, la



Figura 8. Etiquetas, MPI-SHH, 2018. Fotografía © R. Barquera.

cantidad de índices que debemos colocar en la reacción es directamente proporcional a la cantidad de DNA presente en la biblioteca genómica. Empero, incrementar el volumen para permitir la incorporación de suficientes índices por reacción sólo haría menos eficiente la misma. En cambio, la estrategia que se sigue es partir la biblioteca en un número razonable de reacciones, cada una con una cantidad similar de índices. Esto puede llevar a números teóricos elevados de reacciones. En la fotografía podemos apreciar el número de reacciones a las que se elevaron nuestras bibliotecas para este lote: ¡son más de 800!

Dimensiones del lote: 828 tubos, aprox. 82.8 ml (V)

Dimensiones del entrevistado: 800 μ l en ocho tubos de 200 μ l

Amplificar los detalles (figura 9)

Una vez indizadas las bibliotecas genómicas, es necesario purificar de nuevo las bibliotecas y volver a colocar todas las reacciones en su contenedor respectivo. Es decir, reunir todas las re-

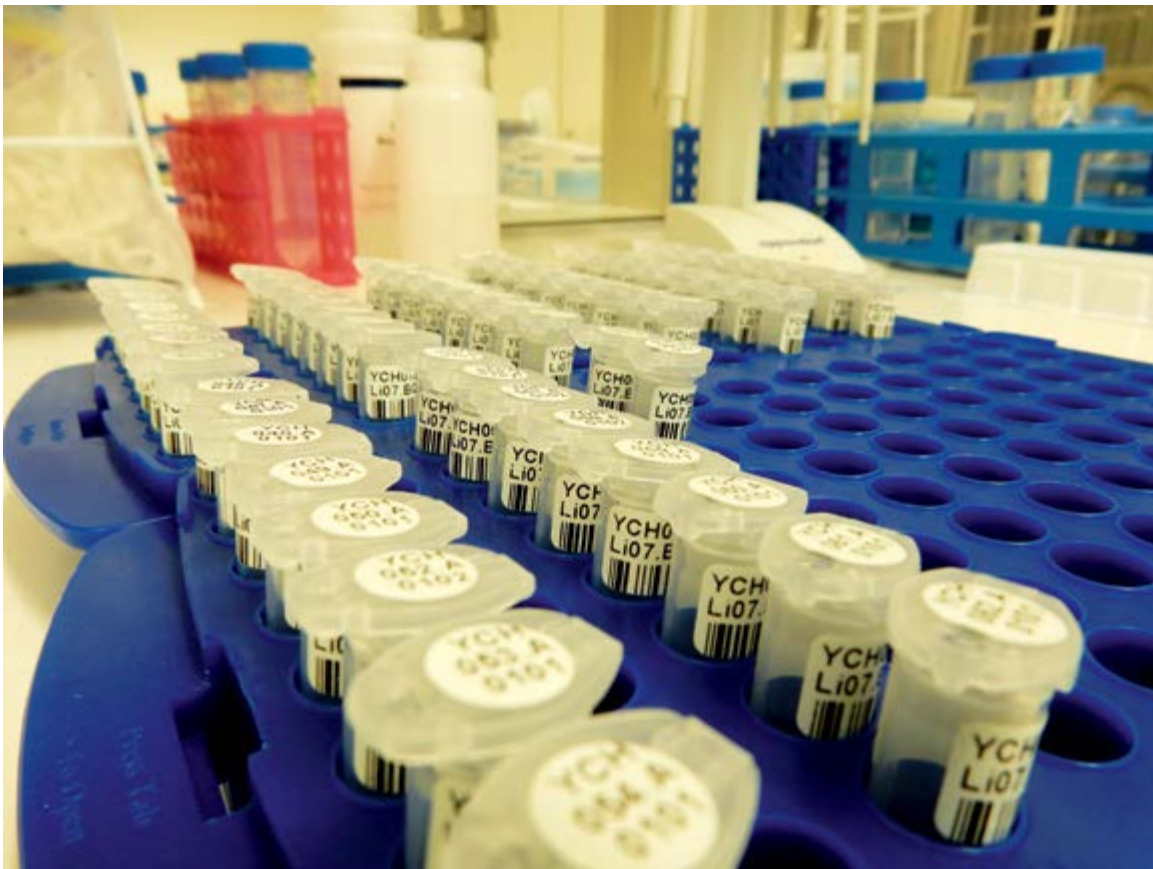


Figura 9. Amplificar los detalles, MPI-SHH, 2018. Fotografía © R. Barquera.

acciones que corresponden a una biblioteca en un solo tubo. Una vez reunidas, se realiza una cuantificación y se determina la eficiencia de la reacción; generalmente se incrementó la cantidad de DNA en la reacción, en un rango de decenas, a cientos de veces. Estas bibliotecas genómicas ahora están listas para sus últimos ajustes antes de entrar al secuenciador. Para poder secuenciarse juntas (una ventaja característica de la secuenciación de siguiente generación o NGS por sus siglas en inglés) es necesario ajustar todas las muestras a la misma concentración. Gracias a la lectura anterior sabemos la cantidad de DNA presente en cada biblioteca y podemos calcular el número de ciclos necesarios para alcanzar diez billones (1×10^{13}) de copias de DNA indizado por biblioteca. Esto lo hacemos nuevamente con el uso de la PCR, y volvemos a purificar y reunir las bibliotecas en sus respectivos tubos. Estas bibliotecas *reamplificadas* son el material a partir del cual trabajaremos de aquí en adelante, tanto para la secuenciación como para las demás técnicas analíticas.

Dimensiones del lote: 63 tubos, aprox. 1.4 ml (V)

Dimensiones del entrevistado: 22 μ l, en un tubo de 1.5 ml

Todos juntos (figura 10)

Una última cuantificación de las bibliotecas genómicas reamplificadas basta para asegurarse que todas se encuentran en el mismo rango de concentraciones. A partir de esta medición se estima la cantidad de cada biblioteca que deberá ir en el concentrado final, o *pool*, de muestras. Este *pool* contiene una fracción de cada biblioteca genómica de nuestro lote de muestras procesadas. Un



Figura 10. Todos juntos, MPI-SHH, 2018. Fotografía © R. Barquera.

podremos ver la secuencia de cada uno de los fragmentos de DNA analizados durante la corrida de secuenciación para cada biblioteca, aun cuando es sólo una parte de la biblioteca genómica de cada individuo. Éstos son algunos de los segmentos de nuestro entrevistado YCH001. En el caso del primero, se trata de un fragmento de DNA mitocondrial, concretamente el inicio de la secuencia del genoma mitocondrial con respecto a la secuencia de referencia (NC_012920.1). La información de la primera línea, que empieza con una @, es el encabezado de la lectura, que contiene información acerca de las celdas y la corrida. Después viene la secuencia de esa lectura en el característico formato A, T, C, G.¹ Finalmente, tenemos los *scores* de calidad de cada base en formato ASCII, que nos dice la calidad de cada base en la secuencia. Esto se repite tantas veces como lecturas tengamos. Y este archivo lo podemos emplear para generar genotipos para cada individuo, lo que nos permitirá obtener nuestras respuestas: qué variantes genéticas tenía, si tenía alguna propensión a padecer alguna enfermedad, si tenía algunas características fenotípicas, y en caso de haberlas, qué patógenos se encontraban en su organismo al momento de fallecer. ¿YCH001, estás ahí? ¡Sí! ¡Oh! ¡Qué bien! ¿Y qué eres? *Un varón, señor.*

Dimensiones del lote: 9772 archivos, 26.62 GB

Dimensiones del entrevistado: 156 archivos, 378.6 MB

La autenticidad del relato (figura 12)

Pero ¿cómo saber si lo que obtenemos es en realidad la respuesta a la pregunta que nos hacemos? Una de las maneras de averiguarlo es con el uso de gráficos de daño de bases. Con éstos podemos ver la frecuencia con la que aparecen los daños característicos de las moléculas de DNA antiguo en nuestras lecturas. Tras mapear nuestras lecturas contra los genomas de referencia adecuados, podemos autenticar que nuestro material es “antiguo” si las lecturas presentan mayor cantidad de daño hacia los extremos de la molécula. Esto genera los denominados “gráficos de sonrisa” o *smiley plots* en inglés. En el caso de las moléculas tratadas enzimáticamente, la frecuencia del daño disminuye en la porción interna de cada lectura pero prevalece en los nucleótidos terminales, lo que da patrones más abruptos, pero aun así retienen la tendencia antes mencionada. Si, por el contrario, analizáramos de la misma manera moléculas modernas, el gráfico sería totalmente lineal, pues no presenta daños y, por lo tanto, los cambios en las bases nitrogenadas con respecto a la secuencia de referencia serían poco frecuentes.

Dimensiones del lote: 63 archivos PDF, 873.1 KB

Dimensiones del entrevistado: 1 archivo PDF, 13.7 KB

1. Las letras hacen referencia a las iniciales de las moléculas adenina (A), timina (T), citosina (C) y guanina (G), las bases nitrogenadas de cuya secuencia depende el material genético para preservar las instrucciones codificadas en el DNA.

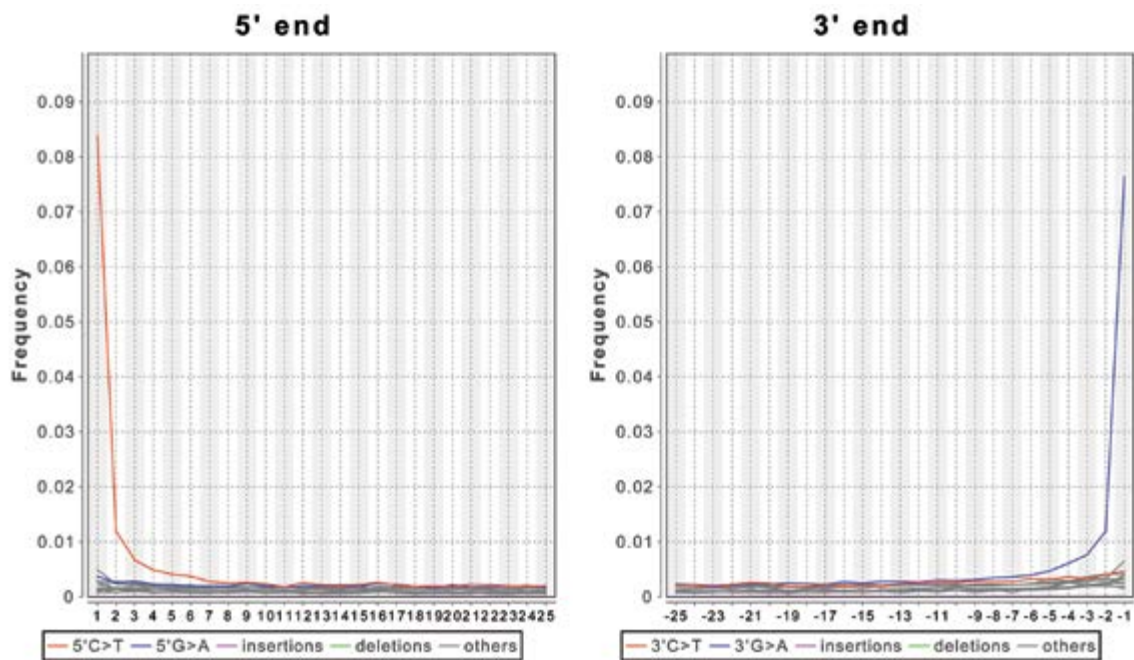


Figura 12. La autenticidad del relato, “gráficos de sonrisa” usados para verificar la autenticidad de la antigüedad del material genético analizado © MPI-SHH (Jena, Alemania).

Enfocarse en lo importante (figura 13)

A pesar de todo el esmero puesto en nuestro trabajo, difícilmente encontraremos muestras que posean altos contenidos de DNA endógeno. La mayor parte de las ocasiones nos toparemos con muestras con proporciones minúsculas de DNA auténtico procedente de la muestra, lo cual es de esperarse pues su origen estuvo enterrado por cientos o miles de años. El DNA extraído a partir de dientes y huesos humanos generalmente es una mezcla de DNA de diversos organismos (la mayoría de ellos ni siquiera antiguos, pues proceden del medio ambiente inmediato al entierro) en la que el DNA humano rara vez representa una proporción mayoritaria y la mayoría de las veces no representa más de 5% del total de las moléculas aisladas. Una estrategia posible sería secuenciar a mayor profundidad (es decir, incrementar el número de lecturas recuperadas en la corrida) pero por lo general esto no es económico ni práctico, y tal aproximación dejaría fuera a la inmensa mayoría de nuestros potenciales individuos (Allentoft *et al.*, 2015). Con la intención de incrementar la abundancia relativa de moléculas de interés se han diseñado colecciones de sondas² que se alinean y “atrapan” las cadenas complementarias a sus secuencias en las bibliotecas genómicas. Esto incrementa el número de segmentos de DNA de interés al mismo tiempo que disminuye cantidad de DNA exógeno. A continuación, las bibliotecas “capturadas” son secuenciadas como suele

2. Fragmentos de DNA sintéticos de secuencia conocida.

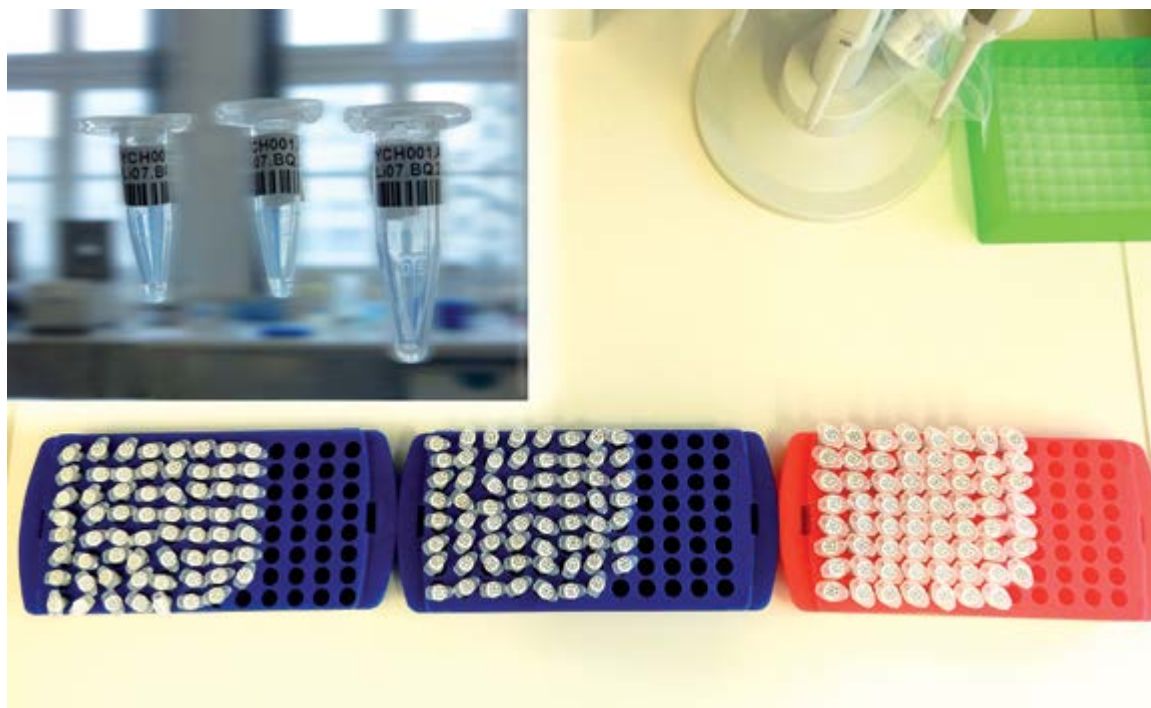


Figura 13. *Enfocarse en lo importante*, MPI-SHH (Jena, Alemania), 2018. Fotografía © R. Barquera.

hacerse, con la diferencia de que en esta ocasión contaremos con una gran cantidad de fragmentos específicos de nuestra muestra con las mismas condiciones de secuenciación. Arreglos específicos de estas sondas de captura han sido diseñados para atrapar DNA humano a nivel de genoma, enriquecido para ciertos genes o regiones genómicas de interés, patógenos o DNA mitocondrial (mtDNA). En este caso, la biblioteca genómica de YCH001 fue capturada para mtDNA, genes de respuesta inmune y una colección de 1 240 000 sitios genómicos.

Dimensiones del lote (pre captura): 189 tubos, aprox. 1.6 ml (V)

Dimensiones del lote (post captura): 29101 archivos, 95.07 GB

Dimensiones del entrevistado (pre captura): 3 tubos, 75 μ l

Dimensiones del entrevistado (post captura): 461 archivos, 1.51 GB

La historia de mamá (figura 14)

¿De dónde es tu mamá, YCH001? *De esta misma tierra, señor ¿Cuánto tiempo llevan viviendo aquí? Pues... unos 14 mil-15 mil años en este continente* (De Saint Pierre *et al.*, 2012; Brandini *et al.*, 2018). *No me acuerdo bien*. El subtipo mitocondrial presente en YCH001 pertenece al subclado mitocondrial B4b, parte a su vez, de la diversidad correspondiente al haplogrupo mitocondrial B. Este haplogrupo surge quizá en Asia cerca de 50 000 años atrás, y se diversificó en el este o sureste

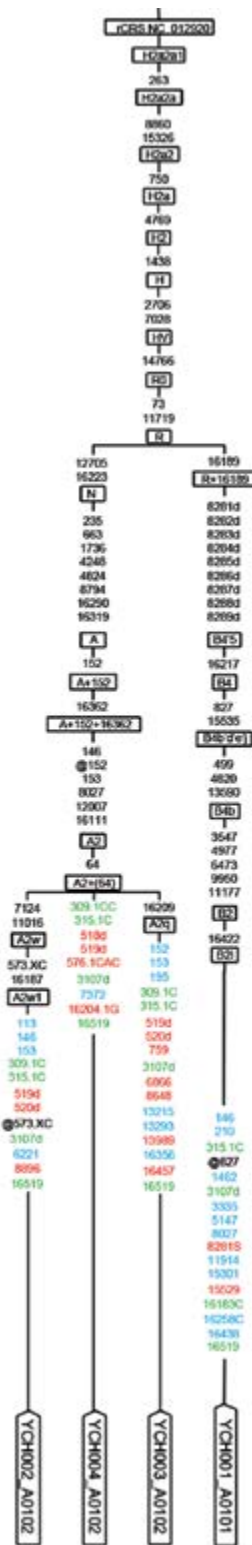


Figura 14. La historia de mamá, representación secuencial de las mutaciones diagnósticas de los linajes mitocondriales de cuatro individuos del contexto de Chichén Itzá © MPI-SHH .

asiáticos. B2 deriva de esta diversidad y más tarde, hace cerca de 15 mil años, se escinde en un subgrupo específico que es B2l, presente en poblaciones de México y del norte de Sudamérica. Cerca de la mitad de la figura, verticalmente, podemos ver a B4b. Hacia abajo están las posiciones en las que hay mutaciones características que definen al siguiente linaje, B2. Una única mutación adicional es requerida para definir al linaje B2l, al cual pertenecen YCH001, su madre, su abuela y tal vez varias generaciones de ascendencia femenina. Hacia atrás eventualmente llegamos al haplogrupo R, distribuido hoy en día sobre todo en Europa, sur y sureste de Asia y el norte de África. Este haplogrupo R es parte, a su vez, de un macrohaplogrupo conocido como N, distribuido por todo el mundo y descendiente del clado L3, el cual representa el linaje materno más común entre los pobladores fuera de África, aunque también se pueden encontrar ahí.

Dimensiones del lote: 63 archivos .bam.fq, 719.4 MB

Dimensiones del entrevistado: un archivo .bam.fq, 18.3 MB

Líneas de defensa (figura 15)

Las mutaciones que dan origen a los linajes de DNA mitocondrial son neutras, es decir, no afectan de ninguna forma el fenotipo del portador de dichos linajes, por lo cual lo único que muestran son las historias relacionadas con el movimiento de poblaciones y de personas, no su resistencia, susceptibilidad o adaptación al ambiente. Algunas de las preguntas que podemos hacer van relacionadas con cómo se adaptaban los individuos a su entorno, pero esas preguntas no pueden responderse con mutaciones neutras. Las respuestas yacen en las regiones codificantes de los genes, que en su conjunto (a nivel genómico) se les conoce como *exoma*. Las variantes génicas que YCH001 portaba transitaron desde nuestra cuna (a nivel de especie) en África, atravesaron desiertos, llanuras, estepas, tundras, bosques tropicales y se asentaron en la península de Yucatán. Esas variantes sobrevivieron a un número inimaginable de patógenos, desde virus causantes de fiebres hemorrágicas hasta parásitos intestinales, incluidas bacterias que causan infecciones agudas severas como las del género *Salmonella*, hasta las que coexisten por años sin manifestarse o cuyo cuadro clínico es una infección crónica como el complejo *Mycobacterium tuberculosis*. Son estos genes, por tanto, ganadores en la carrera de la coevolución entre los agentes infecciosos y *Homo sapiens*. Ganadores temporales. Ganadores para ese tiempo, ese lugar y ese contexto. Las variantes de YCH001 se pueden encontrar hoy en día en varias poblaciones nativas americanas, particularmente en la región del sureste de México y Guatemala, donde representantes modernos de las culturas mayas habitan hoy en día.

Dimensiones del lote: 63 archivos PDF, 55.25 MB

Dimensiones del entrevistado: 1 archivo PDF, 898 KB

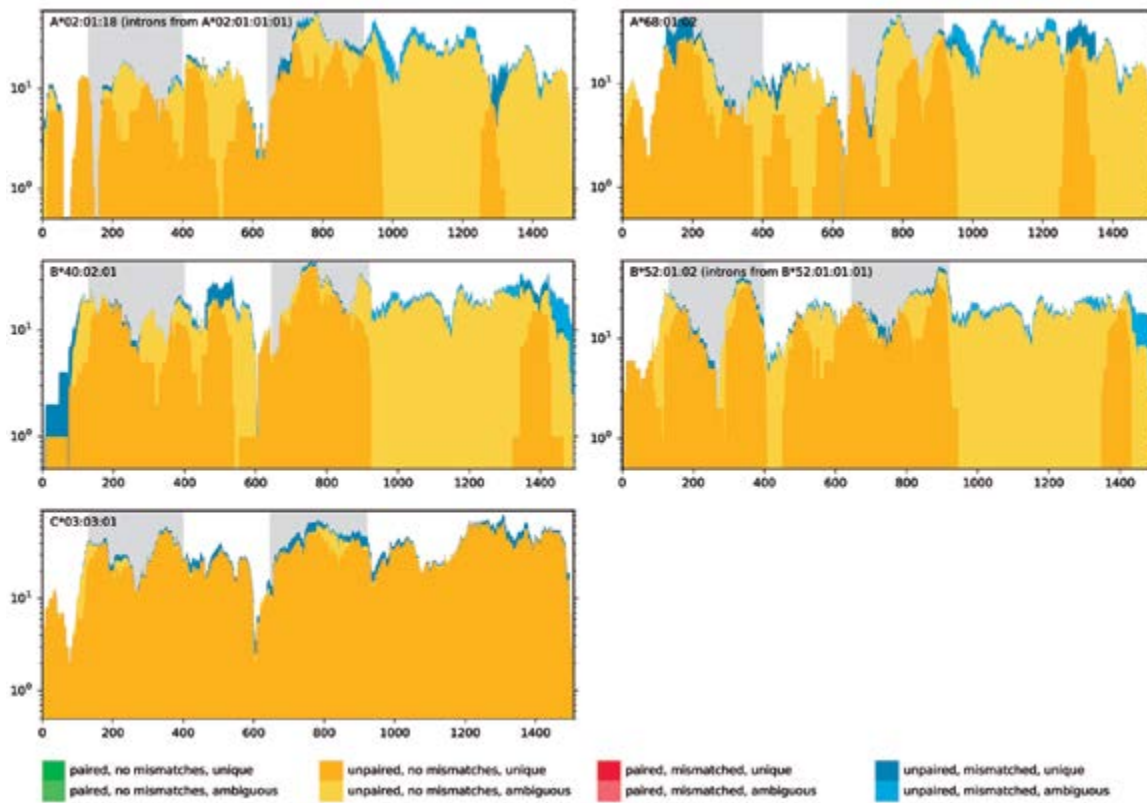


Figura 15. Líneas de defensa, gráfico de cobertura genómica de la región conocida como HLA de clase I proveniente del genoma del individuo YCH001 © MPI-SHH.

Un lugar en el universo (figura 16)

El análisis de componentes principales, o PCA por sus siglas en inglés, es una estrategia que, basada en la presencia o ausencia de colección de mutaciones dispersas por todo el genoma, permite asignar gráficamente las relaciones entre distintos grupos humanos y entre los individuos de estos grupos humanos. Entre más variantes tengamos presentes en la biblioteca genómica analizada, más precisa será la posición que obtendremos para el individuo de quien obtuvimos esa biblioteca genómica. En este caso, todos los demás individuos analizados del contexto del chultún de Chichén Itzá aparecen en el mismo gráfico, indicados con una estrella naranja. La estrella de nuestro seguimiento, YCH001, aparece indicada con la misma figura, pero en color rojo. Este gráfico nos indica, de manera visual, que los grupos mayas actuales poseen una diversidad genética comparable a la que exhibían los niños del contexto del chultún. Asimismo, se puede observar que otros grupos nativos americanos, como los pimas, zapotecos y mixtecos no se alejan mucho de la misma región, lo que indica cercanía genética entre estos grupos, al menos cuando se estudian con los marcadores empleados en este arreglo.

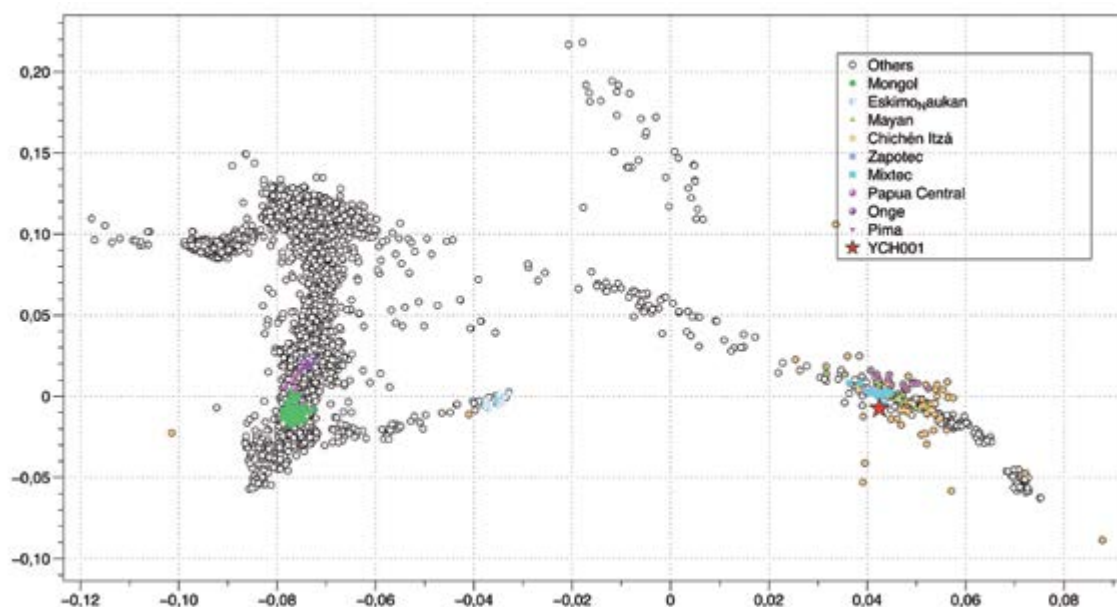


Figura 16. *Un lugar en el universo*, gráfico del análisis de componentes principales donde se muestra, con una estrella, la posición de YCH001 con respecto a otros individuos del mismo contexto y un conjunto de poblaciones mundiales © MPI-SHH.

Una despedida, por ahora

Este conjunto de análisis básicos permite dar una idea de lo que implica trabajar con DNA antiguo. Como podemos ver, aún faltan varias preguntas y respuestas, que requieren de otro tipo de análisis estadísticos y bioinformáticos, como: ¿estaban relacionados genéticamente los niños de este entierro?, ¿provenían del mismo lugar, o eran foráneos?, ¿tenían alguna característica particular a nivel de fenotipo que compartieran entre sí?, ¿qué tan distintos (o similares) son de los mayas actuales?

Pronto tendremos respuestas a estas preguntas y quizá nuevas preguntas por responder. Por ahora dejaremos descansar a YCH001, quien nos ha compartido datos acerca de su origen, su filiación genética y qué tan distinto o similar es a los habitantes mayas contemporáneos.

Dimensiones del lote: 1 archivo DataGraph, 824 KB

Dimensiones del entrevistado: 1 archivo DataGraph, 824 KB

Bibliografía

- Allentof, Morten E. *et al.* (2015). "Population genomics of Bronze Age Eurasia". *Nature*, 522, pp. 167-172.
- Brandini, Stefania *et al.* (2018). "The Paleo-Indian entry into South America according to mitogenomes". *Molecular Biology Evolution*, 35, pp. 299-311.
- Briggs, Adrian W. *et al.* (2010). "Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA". *Nucleic Acids Research*, 38, p: e87.
- Higuchi, Russell *et al.* (1984). "DNA sequences from the quagga, an extinct member of the horse family". *Nature*, 312, pp. 282-4.
- Hofreiter, Michael *et al.* (2001). "DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA". *Nucleic Acids Research*, 29, pp. 4793-4799.
- Hotchner, Aaron (1966). *Papa Hemingway: A personal memoir*. Nueva York: Random House.
- Márquez Morfín, Lourdes (2010). "Morir por los dioses... y uno que otro humano. Sacrificio de niños en Chichen Itzá o práctica funeraria". En Lourdes Márquez Morfín. *Los niños, actores sociales ignorados. Levantando el velo, una mirada al pasado* (pp. 253-282). México: ENAH-INAH / Conaculta / PROMEP.
- Meyer, Matthias, y Kircher, Martin (2010). "Illumina sequencing library preparation for highly multiplexed target capture and sequencing". *Cold Spring Harbor Protocols*, 2010 (6): prot5448. doi:10.1101/pdb.prot5448.
- Pääbo, Svante (1985a). "Preservation of DNA in ancient Egyptian mummies". *J Arch Sci*, 12, pp. 411-417.
- ____ (1985b). "Molecular cloning of Ancient Egyptian mummy DNA". *Nature*, 314, pp. 644-645.
- Pinhasi, Ron *et al.* (2015), "Optimal ancient DNA yields from the inner ear part of the human petrous bone". *PLoS One*, 10, p. e0129102.
- Saint Pierre, Michelle de *et al.* (2012). "An alternative model for the early peopling of southern South America revealed by analyses of three mitochondrial DNA haplogroups". *PLoS One*, 7, p. e43486.