

Secuenciación de ácidos nucleicos de segunda generación y bioarqueología - Revisión

Gabriel Dorado¹, Fernando Luque², Plácido Pascual³, Inmaculada Jiménez⁴, Francisco Javier S. Sánchez-Cañete⁵, Margarita Pérez-Jiménez¹, Patricia Raya⁶, Manuel Gálvez⁷, Jesús Sáiz⁸, Adela Sánchez⁸, Teresa E. Rosales⁹, Víctor F. Vásquez¹⁰, Pilar Hernández¹¹

¹ Autor para correspondencia, Dep. Bioquímica y Biología Molecular, Campus Rabanales C6-1-E17, Campus de Excelencia Internacional Agroalimentario (ceiA3), Universidad de Córdoba, 14071 Córdoba (Spain), eMail: <bb1dopeg@uco.es>; ² Laboratorio de Producción y Sanidad Animal de Córdoba, Ctra. Madrid-Cádiz km 395, 14071 Córdoba; ³ Laboratorio Agroalimentario de Córdoba, Consejería de Agricultura y Pesca, Junta de Andalucía, 14004 Córdoba; ⁴ IES Puertas del Campo, Avda. San Juan de Dios 1, 51001 Ceuta; ⁵ EE.PP. Sagrada Familia de Baena, Avda. Padre Villoslada 22, 14850 Baena (Córdoba); ⁶ Servicio de Protección Radiológica, Facultad de Medicina, Avda. Menéndez Pidal s/n, Universidad de Córdoba, 14071 Córdoba; ⁷ Dep. Radiología y Medicina Física, Unidad de Física Médica, Facultad de Medicina, Avda. Menéndez Pidal s/n, Universidad de Córdoba, 14071 Córdoba; ⁸ Dep. Farmacología, Toxicología y Medicina Legal y Forense, Facultad de Medicina, Avda. Menéndez Pidal, s/n, Universidad de Córdoba, 14071 Córdoba; ⁹ Laboratorio de Arqueobiología, Avda. Universitaria s/n, Universidad Nacional de Trujillo, Trujillo (Perú); ¹⁰ Centro de Investigaciones Arqueobiológicas y Paleoecológicas Andinas Arqueobios, Apartado Postal 595, Trujillo (Perú); ¹¹ Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Alameda del Obispo s/n, 14080 Córdoba

Resumen

La revolución de la secuenciación de ácidos nucleicos comenzó hace 31 años. El desarrollo de la secuenciación de segunda generación (SGS) ha permitido un mayor rendimiento y menor precio por base secuenciada, abriendo la posibilidad de secuenciar genomas antiguos, incluyendo epigenética. Esta revisión describe las principales plataformas de SGS: i) Roche 454 Life Sciences, basada en PCR en emulsión (emPCR) y posterior pirosecuenciación; ii) Illumina (amplificación por puente y secuenciación mediante terminadores reversibles); iii) Life Technologies SOLiD (emPCR y ligación de oligonucleótidos para interrogar ADN); y iv) Life Technologies Ion-Torrent-chip (emPCR y microchips pH-metros). Distintos genomas antiguos (incluyendo virus, microorganismos, plantas y animales) han sido secuenciados. Ello ha permitido estudiar la evolución de patógenos, domesticación de microorganismos, plantas y animales, paleodietas y paleoambientes, incluyendo cambios climáticos. Todavía quedan por superar obstáculos y desafíos, pero los avances tecnológicos continuos en aislamiento de ácidos nucleicos, secuenciación y bioinformática (junto con mayor potencia de computación) prometen un futuro brillante para la bioarqueología en general, y la paleogenómica en particular, permitiendo analizar no sólo genomas aislados, sino también abordar la genómica y evolución de poblaciones antiguas.

Palabras clave (no en título o resumen, que son siempre indexados): ADN antiguo (ADNa), secuenciación de alto rendimiento (SAR), ADN paleoambiental (ADNPAlAmb), paleogenomas, reacción en cadena de la polimerasa (RCP).

Abstract

The nucleic-acid revolution started 31 years ago. The development of the second-generation sequencing (SGS) has allowed a higher throughput and lower price per sequenced base, opening the possibility to sequence ancient genomes, including epigenetics. The main SGS platforms are described in this review: i) Roche 454 Life Sciences, based on emulsion PCR (emPCR) and further pyrosequencing; ii) Illumina (bridge-amplification and subsequent reversible-terminator sequencing); iii) Life Technologies SOLiD (emPCR coupled with oligonucleotide ligation to interrogate DNA); and iv) Life Technologies Ion-Torrent-chip (emPCR, further using microchip pH-meters). Different ancient genomes (including viruses, microorganisms, plants and animals) have been sequenced. This has allowed to study the evolution of pathogens, domestication of microorganisms, plants and animals, paleodiets and paleoenvironments, including climate changes. Some hurdles and challenges must yet be overcome, but the steady technological advances in nucleic-acid isolation, sequencing and bioinformatics (together with higher computing power) promise a bright future for bioarchaeology in general, and paleogenomics in particular, allowing to analyze not just single genomes, but also to address ancient population genomics and evolution.

Key words (not in title or abstract, which are always indexed): ancient DNA (aDNA), high-throughput sequencing (HTS), paleoenvironmental DNA (PalEnDNA), paleogenomes, polymerase chain-reaction (PCR).

Introducción

La secuenciación de ADN antiguo (ADNa) se inició hace más de tres décadas con el quagga (*Equus quagga*), una especie de cebra extinguida en 1883 (Higuchi et al, 1984), utilizando la metodología de secuenciación de primera generación (FGS). Más tarde, la secuenciación de segunda generación (SGS) de ADN (a veces descrito con la denominación algo ambigua de “próxima generación”; NGS) revolucionó las ciencias de la vida y disciplinas relacionadas, incluyendo la arqueología (bioarqueología). Las principales ventajas sobre las plataformas de primera generación fueron el más alto rendimiento y menor precio por base secuenciada. Tal tecnología de secuenciación de alto rendimiento (HTS) permitió secuenciar genomas completos de una forma asequible. Además, algunas plataformas de segunda generación han permitido lo que antes se consideraba una tarea imposible: secuenciar genomas antiguos. Por lo tanto, la mayoría de los avances en la secuenciación de ADNa se han llevado a cabo con dicha segunda generación (Dorado et al, 2007-2014; Charman et al, 2015; Hagelberg et al, 2015; Orlando et al, 2015). La SGS también permite estudiar la epigenética, habiendo sido aplicado recientemente al ADNa (Orlando et al, 2015; Pedersen et al, 2014; Smith et al, 2015; Seguin-Orlando et al, 2015).

Una nueva secuenciación de tercera generación (TGS) de ácidos nucleicos promete revolucionar aún más la bioarqueología en los próximos años, permitiendo la secuenciación de moléculas individuales (sin amplificación previa, evitando los sesgos que la misma podría ocasionar). Ello permite secuenciar no sólo ADNa, como se ha demostrado con el caballo (Orlando et al, 2011; Ginolhac et al, 2012), sino también ARN antiguo (ARNa), aunque dicha tecnología está en desarrollo. Por lo tanto, esta revisión trata de la utilidad de la segunda generación de secuenciación de ácidos nucleicos en bioarqueología.

Plataformas de secuenciación

Existen varias plataformas de secuenciación de segunda generación. Las más populares se describen a continuación.

a. Secuenciación Roche 454 Life Sciences

Esta tecnología revolucionó la secuenciación del ADN, impulsando la secuenciación de segunda generación. Se basa en la amplificación *in vitro* mediante la reacción en cadena de la polimerasa de emulsión (emPCR), lo que permite un alto rendimiento de reacciones paralelas en una emulsión agua-en-aceite (W/O). Además se utilizan reacciones de pirosecuenciación para leer la secuencia de ADN en picopocillos. Las lecturas son finalmente ensambladas utilizando herramientas bioinformáticas. Representa la plataforma de segunda generación con lecturas más largas, aunque son también las más caras.

El poder de esta estrategia ha sido demostrado mediante la secuenciación de genomas antiguos, como el de neandertal (Green et al, 2006; Noonan et al, 2006).

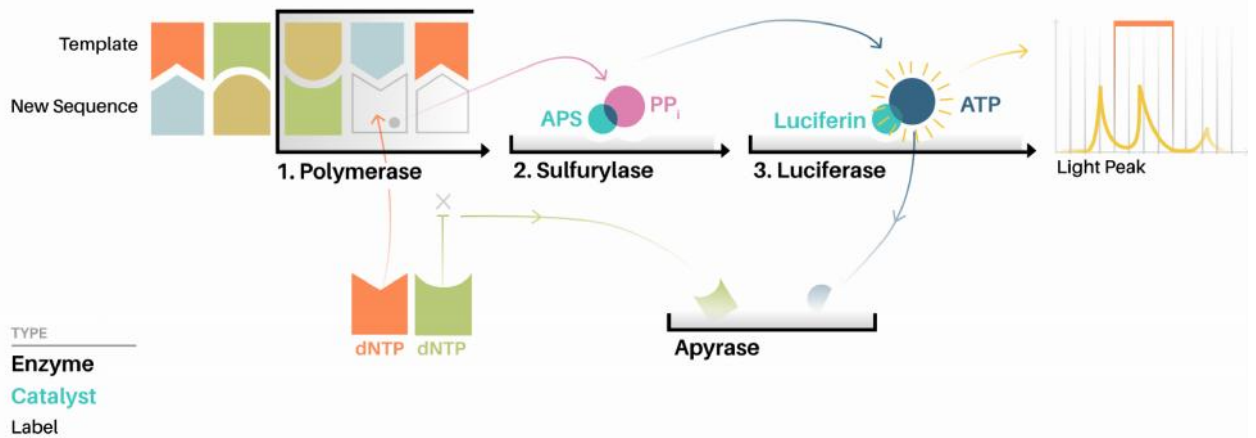


Figura 1. Secuenciación Roche 454 Life Sciences. La reacción de pirosecuenciación (junto a la polimerización del ADN) genera luz, permitiendo la secuenciación del ADN. Créditos: How pyrosequencing works. © 2014 Jacopo Pompili, Wikimedia Commons <<http://commons.wikimedia.org>> y Creative Commons <<http://creativecommons.org>>.

b. Secuenciación Illumina

Se basa en tecnología de amplificación por Reacción en Cadena de la Polimerasa (PCR) de puente, para generar colonias de ADN en soportes sólidos (protocolos sin PCR también están disponibles). Se lleva a cabo una secuenciación masiva en paralelo con terminadores-reversibles, generando luz en células de flujo. Posteriormente se utilizan sofisticados algoritmos (software) para ensamblar las lecturas y generar los fragmentos de ADN contiguos ("contigs"), cromosomas y genomas.

Esta plataforma permite una alta cobertura del genoma secuenciado. Aunque las lecturas originales eran cortas (que a veces era un problema bloqueante, ya que las herramientas bioinformáticas pueden no ser capaces de ensamblar secuencias cortas/repetitivas), la tecnología ha evolucionado, generando actualmente lecturas significativamente más largas.

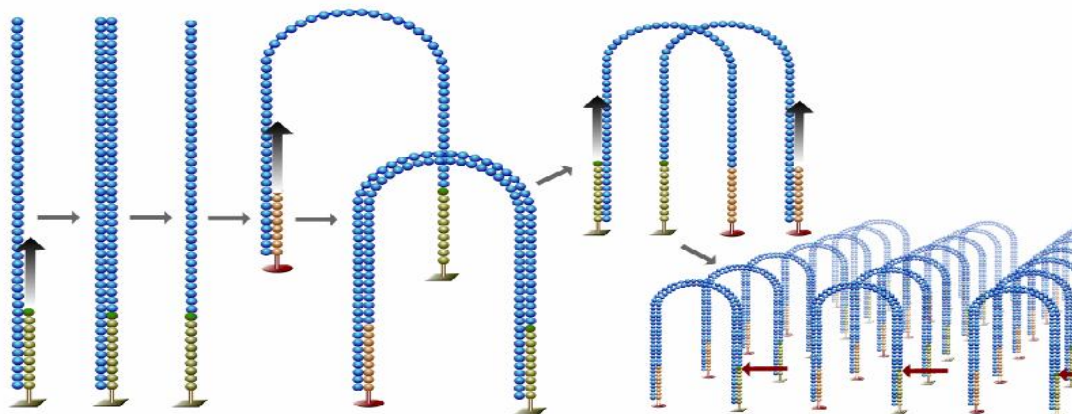
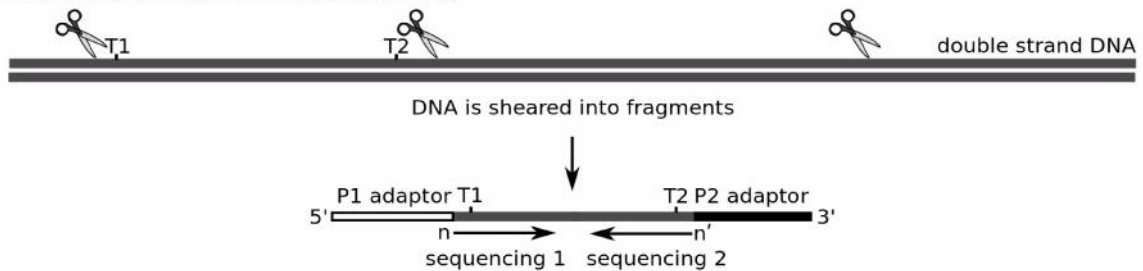


Figure 2. Secuenciación Illumina. La amplificación por puente permite un alto rendimiento. Créditos: Bridge amplification. © 2009 Abizar, Wikimedia Commons <<http://commons.wikimedia.org>> y Creative Commons <<http://creativecommons.org>>.

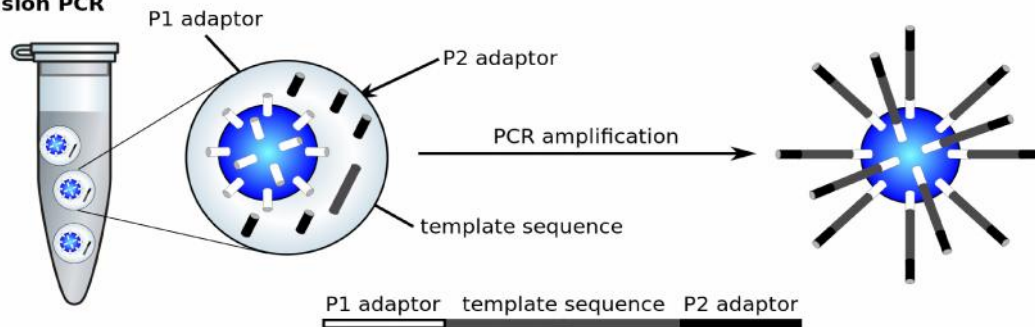
c. Secuenciación Life Technologies SOLiD

Esta metodología utiliza emPCR en un soporte sólido, junto con ligación de oligonucleótido de diferentes cebadores universales y sondas marcadas para interrogar al ADN. La fluorescencia generada por la escisión de la sonda permite leer el ADN en la reacción de secuenciación. Posteriormente, se utilizan herramientas bioinformáticas para ensamblar las lecturas.

A) Single-end and paired-end sequencing



B) Emulsion PCR



C) Mate-pair sequencing

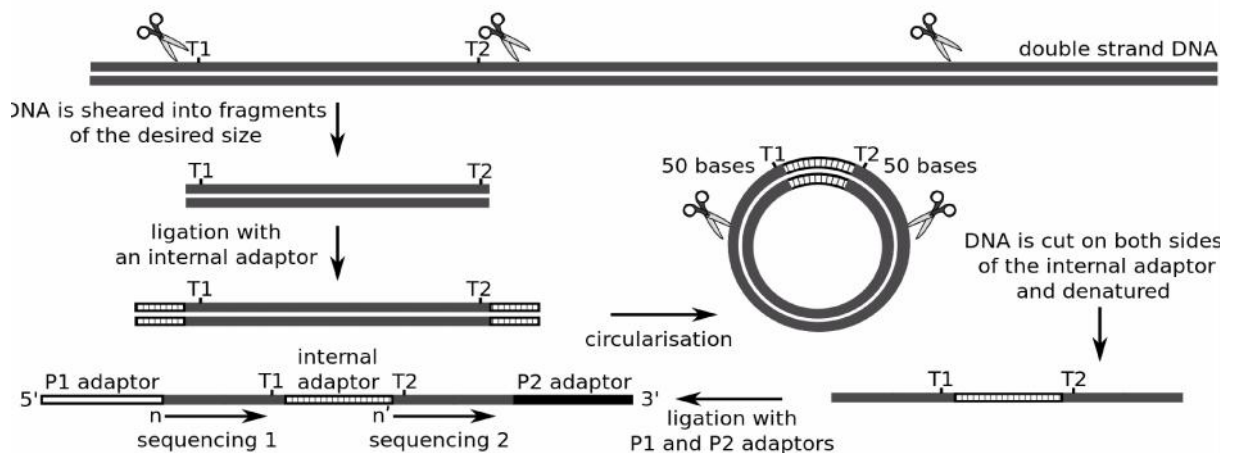


Figura 3. Secuenciación Life Technologies SOLiD. El ADN cortado se somete a emPCR y secuenciación. Créditos: Library preparation for the SOLiD platform. © 2012 Philippe Hupé, Wikimedia Commons <<http://commons.wikimedia.org>> y Creative Commons <<http://creativecommons.org>>.

d. Secuenciación Life Technologies Ion-Torrent-chip

Se genera una biblioteca de fragmentos de ADN ligados a adaptadores con biotina, que son capturados con perlas recubiertas con estreptavidina. El

ADN se desnaturaliza y los fragmentos no biotinilados son capturados con perlas recubiertas por cebadores. El ADN se amplifica por emPCR utilizando cebadores biotinilados, lo que permite aislar los productos de extensión unidos a las perlas. Ellos son capturados en picopocillos, llevándose a cabo las reacciones de secuenciación en femtopocillos que detectan los protones generados (H^+) en la polimerización del ADN, funcionando por lo tanto como microchips pH-metros. La lecturas se ensamblan usando aplicaciones bioinformáticas.

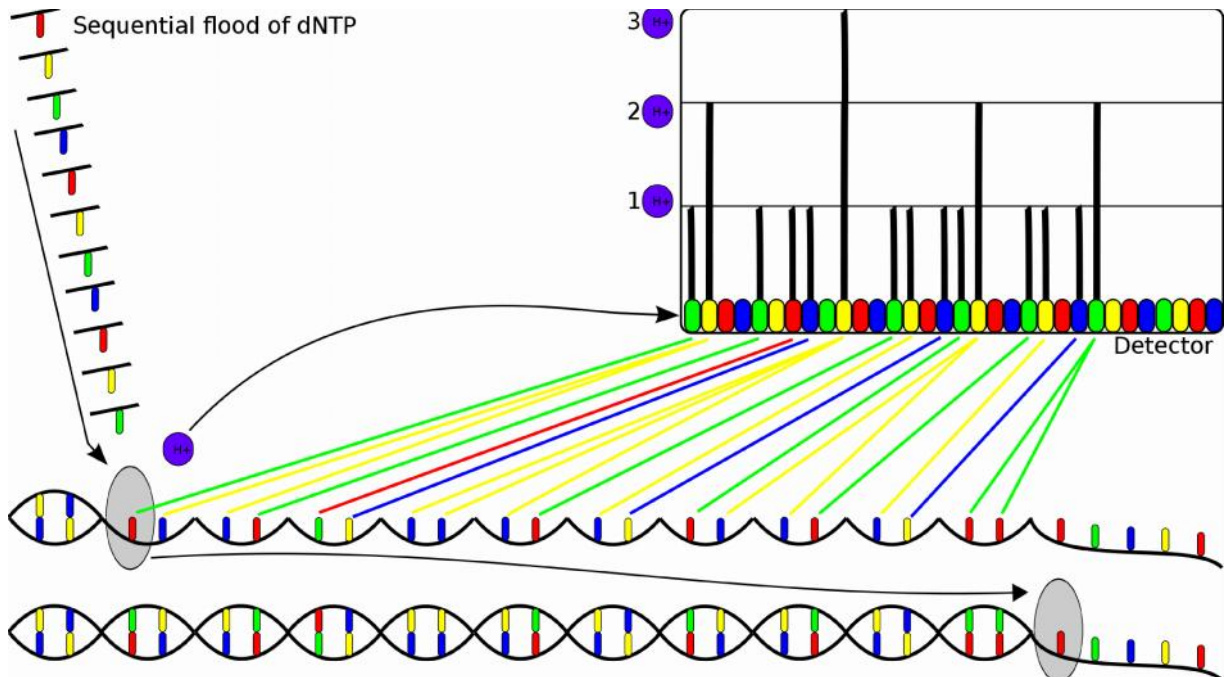


Figura 4. Secuenciación Life Technologies Ion-Torrent-chip. Detección de protones durante la polimerización del ADN. Créditos: dNTP incorporation - hydrogen magnitude. © 2011 David Tack, Wikimedia Commons <<http://commons.wikimedia.org>> y Creative Commons <<http://creativecommons.org>>.

3. Secuenciando genomas antiguos

Diferentes genomas antiguos (paleogenomas) se han secuenciado hasta la fecha. Inicialmente se llevó a cabo mediante la amplificación in vitro utilizando la PCR. La metodología ha mejorado en los últimos años, incluyendo el aislamiento de mayor cantidad/calidad de ADN, así como lecturas de mayor longitud a precios más bajos. La paleogenómica permite la secuenciación total o parcial de genomas antiguos, incluidos virus (Ng et al, 2014) y microorganismos como las cianobacterias (MartínezDeLaEscalera et al, 2014; Pal et al, 2015). Del mismo modo, orgánulos como cloroplastos y mitocondrias (Haak et al, 2010; Bon et al, 2012; Fu et al, 2012; Hung et al, 2013; Paijmans et al, 2013; Fernández et al, 2014; Meyer et al, 2014; Orlando, 2014; Shapiro y Ho, 2014; Sheng et al, 2014; Hervella et al, 2015; Immel et al, 2015; Marsolier-Kergoat et al, 2015) y los genomas nucleares de plantas como algodón (Palmer et al, 2012a, b; Brown et al, 2015) y animales salvajes como mamut (Poinar et al, 2006; Miller et al, 2008), oso (Barnes et al, 2002; Noonan et al, 2005), hiena de las cavernas (Bon et al, 2012), caballo (Orlando et al, 2013) y bisonte (Marsolier-Kergoat et al, 2015), así como ganado como ovejas (Teasdale et al, 2015) y homínidos (Hominini) como neandertales (Paabo, 2015)

y denisovanos (Meyer et al 2012; Brown y Barnes, 2015) desde el Pleistoceno medio hasta hace un millón de años (Orlando, 2014). Las nuevas tecnologías de secuenciación también permiten estudiar la genómica de poblaciones antiguas y evolución, como se ha descrito recientemente para el pingüino adalaida (*Pygoscelis adeliae*) (Parques et al, 2015) y los humanos (Allentoft et al, 2015).

La paleogenómica también se está utilizando para estudiar la evolución de agentes patógenos, incluidos los que causan enfermedades infecciosas como el tizón de la patata (*Phytophthora infestans*), la peste (*Yersinia pestis*), lepra (*Mycobacterium leprae*) y tuberculosis (*Mycobacterium tuberculosis*) (Donoghue et al, 2015). Del mismo modo, la domesticación de microorganismos (eg., fermentaciones), plantas, animales y paleodietas. Curiosamente, una secuenciación masiva de 101 genomas antiguos humanos de la Edad del Bronce en Eurasia (3000 a 1000 aC) ha puesto de manifiesto el hecho sorprendente de que la tolerancia a la lactosa en dicho periodo fue sólo del 10% en Europa (Allentoft et al, 2015), lo que demuestra que la principal selección positiva surgió más recientemente (1000 aC o posterior) de lo que se pensaba (5500 aC) (Itan et al, 2009). Por otra parte, la secuenciación de plantas y animales extintos puede arrojar nueva luz sobre el cambio climático, migraciones, adaptaciones e historia evolutiva de las especies (Haak et al, 2010; Fu et al, 2012; Fernández et al, 2014; Brandao et al, 2015; Cooper et al, 2015; Hervella et al, 2015). De esta manera, los flujos de genes pueden ser también determinados, demostrando antiguas mezclas genéticas endogámicas entre seres humanos modernos los denisovanos y los neandertales (que también se cruzaron entre sí) (Prüfer et al, 2014; DerSarkissian et al, 2015; Ermini et al, 2015; Hofreiter et al, 2015; Knapp et al, 2015; Paabo, 2015; Perry y Orlando, 2015; Vernot y Akey, 2015). Esto ha demostrado que son subespecies de la misma especie.

Por otra parte, el ADN del medio ambiente (eDNA), y en particular el ADN paleoambiental (PalEnDNA), incluye depósitos como sedimentos y suelos, y restos como coprolitos y contenido intestinal (Clack et al, 2012; MartínezDeLaEscalera et al, 2014; Ng et al, 2014; Pawlowski et al, 2014; Rawlence et al, 2014; Pal et al, 2015; Pedersen et al, 2015; Thomsen y Willerslev, 2015). Por lo tanto, la secuenciación de especies extintas se puede utilizar para reconstruir los ecosistemas antiguos, incluso en ausencia de fósiles visibles a simple vista o al microscopio.

Pero diferentes desafíos se deben superar para secuenciar con éxito el ADN (Kircher, 2012; Shapiro y Hofreiter, 2012), incluyendo su recuperación con la cantidad/calidad suficiente y requerida integridad química y física (Overballe-Petersen et al, 2012; Parques y Lamber, 2015), sin estar contaminado con ADN moderno. Además de la fragmentación física, otras posibles alteraciones químicas del ADN incluyen tautomerización, desaminación, pérdida de bases (principalmente despurinación; sobre todo en guanosinas), oxidación e hidrólisis (mayormente en bases purínicas). Los artefactos de secuenciación pueden ser también específicos de la plataforma utilizada (Seguin-Orlando et al, 2013). Así, se han identificado recientemente artefactos de secuencias palindrómicas (Estrella et al, 2014). De hecho, se ha

encontrado que aunque la edad de la muestra puede ser, obviamente, relevante, otros aspectos como la historia tafonómica de los restos pueden ser más importantes para determinar la cantidad y la calidad del ADN recuperable de los restos arqueológicos. Por lo tanto, el frío (eg., permafrost en el Ártico y regiones antárticas) y la sequedad (eg., ambientes desérticos y salinos) producen generalmente las muestras de ADN mejor conservadas.

Se han desarrollado aplicaciones bioinformáticas para determinar daños del ADN (Jonsson et al, 2013). Recientemente se ha aconsejado usar muestras de ADN con limitada fragmentación y desaminación para evitar resultados sesgados en los estudios epigenómicos (Seguin-Orlando et al, 2015). Se han evaluado diversas estrategias de reparación enzimática para disminuir los daños en el ADN y aumentar su cantidad (Mouttham et al, 2015). Además, algunos tratamientos de la muestra, como digestiones enzimáticas en presencia de ácido etilendiamino-tetraacético (EDTA) pueden aumentar significativamente el ADN recuperado (Damgaard et al, 2015). Además, y como es lógico, puede haber también una cantidad/calidad diferencial del ADN aislado de las diferentes partes de las muestras, como los dientes y los huesos (Damgaard et al, 2015; Pinhasi et al, 2015). El enriquecimiento de fragmentos de ADN por captura mediante hibridación con sonda, o estrategias de captura sin sondas, como afinidad de dominios de unión metilados (MBD) para dinucleótidos CpG metilados (mCpG), así como el aislamiento de ADN de cadena sencilla (ssDNA) a partir de muestras antiguas, han representado hitos en el progreso de la secuenciación de ADN en los últimos años (Carpenter et al, 2013; Gansauge y Meyer, 2013; Enk et al, 2014; Ávila-Arcos et al, 2015; Brown y Barnes, 2015; Hofreiter et al, 2015).

Además, la gran cantidad de datos generados con la secuenciación de segunda generación exige nuevas estrategias de desarrollo bioinformático, tanto para hardware como software. Así, se está utilizando procesamiento paralelo con microprocesadores de múltiples núcleos, junto con herramientas bioinformáticas para ensamblar lecturas cortas. Consecuentemente, estos avances están impulsando también un cambio conceptual para abordar adecuadamente los desafíos experimentales, analizar e interpretar los resultados (Schubert et al, 2012; Rawlence et al, 2014; Hofreiter et al, 2015).

4. Perspectivas futuras y conclusiones

El futuro es prometedor para los estudios de ADN en general y la paleogenómica en particular, sobre todo debido al desarrollo de nuevas plataformas de secuenciación que permiten secuenciar genomas antiguos rápidamente usando menos material de partida, con lecturas mucho más largas, mayor precisión y rendimiento, a un costo menor. Así, los estudios de paleogenómica de poblaciones serán más rentables con la secuenciación de tercera generación de ácidos nucleicos, que permite leer directamente las secuencias de ácidos nucleicos individuales sin previa amplificación in vivo (eg., clonación molecular dentro de las bacterias) o in vitro (eg., PCR). Además, estas nuevas plataformas también podrían permitir a la secuencia de ARN. Pero para llegar a esas metas, nuevos desarrollos y mejoras en potencia de computación (principalmente mediante ejecuciones paralelas en

chips de muchos núcleos) y algoritmos bioinformáticos serán también necesarios para manejar la creciente complejidad de los conjuntos de datos generados. Por supuesto, se debe tener especial cuidado para mantener la integridad de los ácidos nucleicos, aumentar el rendimiento y evitar contaminaciones cruzadas, como ha demostrado la experiencia de investigación en ADNa en los últimos 31 años.

Agradecimientos. Financiado por Ministerio de Economía y Competitividad (proyectos MINECO AGL2010-17316 y BIO2011-15237-E) e Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (MINECO e INIA RF2012-00002-C02-02); Consejería de Agricultura y Pesca (041/C/2007, 75/C/2009 y 56/C/2010), Consejería de Economía, Innovación y Ciencia (P11-AGR-7322 y P12-AGR-0482) y Grupo PAI (AGR-248) de Junta de Andalucía; y Universidad de Córdoba (Ayuda a Grupos), Spain.

Bibliografía

Allentoft ME, Sikora M, Sjogren KG, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlstrom T, Vinner L, Malaspinas AS, Margaryan A, Higham T, Chivall D, Lynnerup N, Harvig L, Baron J, DellaCasa P, Dabrowski P, Duffy PR, Ebel AV, Epimakhov A, Frei K, Furmanek M, Gralak T, Gromov A, Gronkiewicz S, Grupe G, Hajdu T, Jarysz R, Khartanovich V, Khokhlov A, Kiss V, Kolar J, Kriiska A, Lasak I, Longhi C, McGlynn G, Merkevicius A, Merkyte I, Metspalu M, Mkrtychyan R, Moiseyev V, Paja L, Palfi G, Pokutta D, Pospieszny L, Price TD, Saag L, Sablin M, Shishlina N, Smrcka V, Soenov VI, Szeverényi V, Toth G, Trifanova SV, Varul L, Vicze M, Yepiskoposyan L, Zhitenev V, Orlando L, Sicheritz-Ponten T, Brunak S, Nielsen R, Kristiansen K, Willerslev E (2015): Population genomics of Bronze Age Eurasia. *Nature* 522: 167-172.

Avila-Arcos MC, Sandoval-Velasco M, Schroeder H, Carpenter ML, Malaspinas AS, Wales N, Penaloza F, Bustamante CD, Gilbert MTP (2015): Comparative performance of two whole-genome capture methodologies on ancient DNA Illumina libraries. *Methods Ecol Evol* 6: 725-734.

Barnes I, Matheus P, Shapiro B, Jensen D, Cooper A (2002): Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science* 295:2267-2270.

Bon C, Berthonaud V, Maksud F, Labadie K, Poulain J, Artiguenave F, Wincker P, Aury JM, Elalouf JM (2012): Coprolites as a source of information on the genome and diet of the cave hyena. *Proc Biol Sci* 279: 2825-2830.

Brandao MM, Spoladore L, Faria LC, Rocha AS, Silva-Filho MC, Palazzo R (2015): Ancient DNA sequence revealed by error-correcting codes. *Sci Rep* 5: 12051 (9 pp).

Brown TA, Barnes IM (2015): The current and future applications of ancient DNA in Quaternary science. *J Quat Sci* 30: 144-153.

Brown TA, Cappellini E, Kistler L, Lister DL, Oliveira HR, Wales N, Schlumbaum A (2015): Recent advances in ancient DNA research and their implications for archaeobotany. *Vegetation History and Archaeobotany* 24: 207-214.

Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H, Allentoft ME, Sikora M, Rasmussen M, Gravel S, Guillén S, Nekhrizov G, Leshtakov K, Dimitrova D, Theodossiev N, Pettener D, Luiselli D, Sandoval K, Moreno-Estrada A, Li Y, Wang J, Gilbert MT, Willerslev E, Greenleaf WJ, Bustamante CD (2013): Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *Am J Hum Genet* 93: 852-864.

Charman DJ, Duller GAT, Long AJ, Schreve DC, Scourse JD (2015): Editorial: Quaternary revolutions. *J Quaternary Sci* 30: 101-103.

Clack AA, MacPhee RD, Poinar HN (2012): *Myiodon darwinii* DNA sequences from ancient fecal hair shafts. *Ann Anat* 194: 26-30.

Cooper A, Turney C, Hughen KA, Brook BW, McDonald HG, Bradshaw CJ (2015): Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover. *Science* 349: 602-606.

Damgaard PB, Margaryan A, Schroeder H, Orlando L, Willerslev E, Allentoft ME (2015): Improving access to endogenous DNA in ancient bones and teeth. *Sci Rep* 5: 11184 (pp).

DerSarkissian C, Allentoft ME, Ávila-Arcos MC, Barnett R, Campos PF, Cappellini E, Ermini L, Fernández R, DaFonseca R, Ginolhac A, Hansen AJ, Jonsson H, Korneliusson T, Margaryan A, Martin MD, Moreno-Mayar JV, Raghavan M, Rasmussen M, Velasco MS, Schroeder H, Schubert M, Seguin-Orlando A, Wales N, Gilbert MT, Willerslev E, Orlando L (2015): Ancient genomics. *Philos Trans R Soc Lond B Biol Sci* 370: 20130387 (12 pp).

Donoghue HD, Spigelman M, O'Grady J, Szikossy I, Pap I, Lee OY, Wu HH, Besra GS, Minnikin DE (2015): Ancient DNA analysis - An established technique in charting the evolution of tuberculosis and leprosy. *Tuberculosis* 95 Suppl 1: S140-S144.

Dorado G, Jiménez I, Rey I, Sánchez-Cañete FJS, Luque F, Morales A, Gálvez M, Sáiz J, Sánchez A, Rosales TE, Vásquez VF, Hernández P (2013): Genomics and proteomics in bioarchaeology - Review. *Archaeobios* 7: 47-63.

Dorado G, Rey I, Rosales TE, Sánchez-Cañete FJS, Luque F, Jiménez I, Gálvez M, Sáiz J, Sánchez A, Vásquez VF (2009): Ancient DNA to decipher the domestication of dog (REVIEW). *Archaeobios* 3: 127-132.

Dorado G, Rey I, Rosales TE, Sánchez-Cañete FJS, Luque F, Jiménez I, Morales A, Gálvez M, Sáiz J, Sánchez A, Hernández P, Vásquez VF (2010): Biological mass extinctions on planet Earth (REVIEW). *Archaeobios* 4: 53-64.

Dorado G, Rosales TE, Luque F, Sánchez-Cañete FJS, Rey I, Jiménez I, Morales A, Gálvez M, Sáiz J, Sánchez A, Vásquez VF, Hernández P (2011): Ancient nucleic acids from maize - A review. *Archaeobios* 5: 21-28.

Dorado G, Rosales TE, Luque F, Sánchez-Cañete FJS, Rey I, Jiménez I, Morales A, Gálvez M, Sáiz J, Sánchez A, Vásquez VF, Hernández P (2012): Isotopes in bioarchaeology - Review. *Archaeobios* 6: 79-91.

Dorado G, Sánchez-Cañete FJS, Pascual P, Jiménez I, Luque F, Pérez-Jiménez M, Raya P, Gálvez M, Sáiz J, Sánchez A, Rosales TE, Vásquez VF, Hernández P (2014): Starch genomics and bioarchaeology - Review. *Archaeobios* 8: 41-50.

Dorado G, Vásquez V, Rey I, Luque F, Jiménez I, Morales A, Gálvez M, Sáiz J, Sánchez A, Hernández P (2008): Sequencing ancient and modern genomes (REVIEW). *Archaeobios* 2: 75-80.

Dorado G, Vásquez V, Rey I, Vega JL (2007): Archaeology meets Molecular Biology (REVIEW). *Archaeobios* 1: 1-2.

Enk JM, Devault AM, Kuch M, Murgha YE, Rouillard JM, Poinar HN (2014): Ancient whole genome enrichment using baits built from modern DNA. *Mol Biol Evol* 31: 1292-1294.

Ermini L, DerSarkissian C, Willerslev E, Orlando L (2015): Major transitions in human evolution revisited: a tribute to ancient DNA. *J Hum Evol* 79: 4-20.

Fernández E, Pérez-Pérez A, Gamba C, Prats E, Cuesta P, Anfruns J, Molist M, Arroyo-Pardo E, Turbón D (2014): Ancient DNA analysis of 8000 B.C. near eastern farmers supports an early Neolithic pioneer maritime colonization of Mainland Europe through Cyprus and the Aegean Islands. *PLoS Genet* 10: e1004401 (16 pp).

Fu Q, Rudan P, Paabo S, Krause J (2012): Complete mitochondrial genomes reveal Neolithic expansion into Europe. *PLoS One* 7: e32473 (6 pp).

Gansauge MT, Meyer M (2013): Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat Protoc* 8: 737-748.

Ginolhac A, Vilstrup J, Stenderup J, Rasmussen M, Stiller M, Shapiro B, Zazula G, Froese D, Steinmann KE, Thompson JF, Al-Rasheid KA, Gilbert TM, Willerslev E, Orlando L (2012): Improving the performance of true single molecule sequencing for ancient DNA. *BMC Genomics* 13: 177 (14 pp).

Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, Du L, Egholm M, Rothberg JM, Paunovic M, Paabo S (2006): Analysis of one million base pairs of Neanderthal DNA. *Nature* 444: 330-336.

Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, Der Sarkissian CS, Brandt G, Schwarz C, Nicklisch N, Dresely V, Fritsch B, Balanovska E, Vilems R, Meller H, Alt KW, Cooper A; Members of the

Genographic Consortium (2010): Ancient DNA from European early Neolithic farmers reveals their near eastern affinities. PLoS Biol 8: e1000536 (16 pp).

Hagelberg E, Hofreiter M, Keyser C (2015): Ancient DNA: the first three decades. Philos Trans R Soc Lond B Biol Sci 370: 20130371 (6 pp).

Hervella M, Rotea M, Izagirre N, Constantinescu M, Alonso S, Ioana M, Lazar C, Ridiche F, Soficaru AD, Netea MG, DeLaRua C (2015): Ancient DNA from South-East Europe reveals different events during Early and Middle Neolithic influencing the European genetic heritage. PLoS One 10: e0128810 (20 pp).

Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC (1984): DNA sequences from the quagga, an extinct member of the horse family. Nature 312: 282-284.

Hofreiter M, Paijmans JL, Goodchild H, Speller CF, Barlow A, Fortes GG, Thomas JA, Ludwig A, Collins MJ (2015): The future of ancient DNA: Technical advances and conceptual shifts. Bioessays 37: 284-293.

Hung CM, Lin RC, Chu JH, Yeh CF, Yao CJ, Li SH (2013): The de novo assembly of mitochondrial genomes of the extinct passenger pigeon (*Ectopistes migratorius*) with next generation sequencing. PLoS One 8: e56301 (9 pp).

Immel A, Drucker DG, Bonazzi M, Jahnke TK, Munzel SC, Schuenemann VJ, Herbig A, Kind CJ, Krause J (2015): Mitochondrial genomes of giant deers suggest their late survival in Central Europe. Sci Rep 5: 10853 (9 pp).

Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG (2009): The origins of lactase persistence in Europe. PLoS Comput Biol 5: e1000491 (13 pp).

Jonsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L (2013): mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29: 1682-1684.

Kircher M (2012): Analysis of high-throughput ancient DNA sequencing data. Methods Mol Biol 840: 197-228.

Knapp M, Lalueza-Fox C, Hofreiter M (2015): Re-inventing ancient human DNA. Investig Genet 6: 4 (11 pp).

Marsolier-Kergoat MC, Palacio P, Berthonaud V, Maksud F, Stafford T, Begouen R, Elalouf JM (2015): Hunting the extinct steppe bison (*Bison priscus*) mitochondrial genome in the Trois-Freres Paleolithic painted cave. PLoS One 10: e0128267 (16 pp).

MartínezDeLaEscalera G, Antoniadis D, Bonilla S, Piccini C (2014): Application of ancient DNA to the reconstruction of past microbial assemblages and for the detection of toxic cyanobacteria in subtropical freshwater ecosystems. Mol Ecol 23: 5791-5802.

Meyer M, Fu Q, Aximu-Petri A, Glocke I, Nickel B, Arsuaga JL, Martínez I, Gracia A, DeCastro JM, Carbonell E, Paabo S (2014): A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* 505: 403-406.

Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prufer K, de Filippo C, Sudmant PH, Alkan C, Fu Q, Do R, Rohland N, Tandon A, Siebauer M, Green RE, Bryc K, Briggs AW, Stenzel U, Dabney J, Shendure J, Kitzman J, Hammer MF, Shunkov MV, Derevianko AP, Patterson N, Andrés AM, Eichler EE, Slatkin M, Reich D, Kelso J, Paabo S (2012): A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338:222-226.

Miller W, Drautz DI, Ratan A, Pusey B, Qi J, Lesk AM, Tomsho LP, Packard MD, Zhao F, Sher A, Tikhonov A, Raney B, Patterson N, Lindblad-Toh K, Lander ES, Knight JR, Irzyk GP, Fredrikson KM, Harkins TT, Sheridan S, Pringle T, Schuster SC (2008): Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* 456: 387-390.

Mouttham N, Klunk J, Kuch M, Fournay R, Poinar H (2015): Surveying the repair of ancient DNA from bones via high-throughput sequencing. *Biotechniques* 59: 19-25.

Ng TFF, Chen LF, Zhou YC, Shapiro B, Stiller M, Heintzman PD, Varsani A, Kondov NO, Wong W, Deng XT, Andrews TD, Moorman BJ, Meulendyk T, MacKay G, Gilbertson RL, Delwart E (2014): Preservation of viral genomes in 700-y-old caribou feces from a subarctic ice patch. *Proc Natl Acad Sci USA* 111: 16842-16847.

Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, Chen F, Platt D, Paabo S, Pritchard JK, Rubin EM (2006): Sequencing and analysis of Neanderthal genomic DNA. *Science* 314: 1113-1118.

Noonan JP, Hofreiter M, Smith D, Priest JR, Rohland N, Rabeder G, Krause J, Dettler JC, Paabo S, Rubin EM (2005): Genomic sequencing of Pleistocene cave bears. *Science* 309: 597-599.

Orlando L (2014): A 400,000-year-old mitochondrial genome questions phylogenetic relationships amongst archaic hominins: using the latest advances in ancient genomics, the mitochondrial genome sequence of a 400,000-year-old hominin has been deciphered. *Bioessays* 36: 598-605.

Orlando L, Gilbert MTP, Willerslev E (2015): Applications of next-generation sequencing reconstructing ancient genomes and epigenomes. *Nat Rev Genet* 16: 395-408.

Orlando L, Ginolhac A, Raghavan M, Vilstrup J, Rasmussen M, Magnussen K, Steinmann KE, Kapranov P, Thompson JF, Zazula G, Froese D, Moltke I, Shapiro B, Hofreiter M, Al-Rasheid KA, Gilbert MT, Willerslev E (2011): True single-molecule DNA sequencing of a Pleistocene horse bone. *Genome Res* 21: 1705-1719.

Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, Johnson PL, Fumagalli M, Vilstrup JT, Raghavan M, Korneliussen T, Malaspinas AS, Vogt J, Szklarczyk D, Kelstrup CD, Vinther J, Dolocan A, Stenderup J, Velazquez AM, Cahill J, Rasmussen M, Wang X, Min J, Zazula GD, Seguin-Orlando A, Mortensen C, Magnussen K, Thompson JF, Weinstock J, Gregersen K, Roed KH, Eisenmann V, Rubin CJ, Miller DC, Antczak DF, Bertelsen MF, Brunak S, Al-Rasheid KA, Ryder O, Andersson L, Mundy J, Krogh A, Gilbert MT, Kjaer K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wang J, Willerslev E (2013): Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499: 74-78.

Overballe-Petersen S, Orlando L, Willerslev E (2012): Next-generation sequencing offers new insights into DNA degradation. *Trends Biotechnol* 30: 364-368.

Paabo S (2015): "Neanderthal Man: In Search of Lost Genomes". Basic Books (New York, NY, USA).

Paijmans JL, Gilbert MT, Hofreiter M (2013): Mitogenomic analyses from ancient DNA. *Mol Phylogenet Evol* 69:404-416.

Pal S, Gregory-Eaves I, Pick FR (2015): Temporal trends in cyanobacteria revealed through DNA and pigment analyses of temperate lake sediment cores. *Journal of Paleolimnology* 54: 87-101.

Palmer SA, Clapham AJ, Rose P, Freitas FO, Owen BD, Beresford-Jones D, Moore JD, Kitchen JL, Allaby RG (2012a): Archaeogenomic evidence of punctuated genome evolution in *Gossypium*. *Mol Biol Evol* 29: 2031-2038.

Palmer SA, Smith O, Allaby RG (2012b): The blossoming of plant archaeogenetics. *Ann Anat* 194: 146-156.

Parks M, Lambert D (2015): Impacts of low coverage depths and post-mortem DNA damage on variant calling: a simulation study. *BMC Genomics* 16: 19.

Parks M, Subramanian S, Baroni C, Salvatore MC, Zhang G, Millar CD, Lambert DM (2015): Ancient population genomics and the study of evolution. *Philos Trans R Soc Lond B Biol Sci* 370: 20130381 (X pp).

Pawlowski J, Lejzerowicz F, Esling P (2014): Next-generation environmental diversity surveys of foraminifera: preparing the future. *Biol Bull* 227: 93-106.

Pedersen JS, Valen E, Velazquez AM, Parker BJ, Rasmussen M, Lindgreen S, Lilje B, Tobin DJ, Kelly TK, Vang S, Andersson R, Jones PA, Hoover CA, Tikhonov A, Prokhortchouk E, Rubin EM, Sandelin A, Gilbert MT, Krogh A, Willerslev E, Orlando L (2014): Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Res* 24: 454-466.

Pedersen MW, Overballe-Petersen S, Ermini L, Sarkissian CD, Haile J, Hellstrom M, Spens J, Thomsen PF, Bohmann K, Cappellini E, Schnell IB, Wales NA, Caroe C, Campos PF, Schmidt AM, Gilbert MT, Hansen AJ, Orlando L, Willerslev E (2015): Ancient and modern environmental DNA. *Philos Trans R Soc Lond B Biol Sci* 370: 20130383 (X pp).

Perry GH, Orlando L (2015): Ancient DNA and human evolution. *J Hum Evol* 79: 1-3.

Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S, Gerritsen F, Moiseyev V, Gromov A, Raczky P, Anders A, Pietruszewsky M, Rollefson G, Jovanovic M, Trinhhoang H, Bar-Oz G, Oxenham M, Matsumura H, Hofreiter M (2015): Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One* 10: e0129102 (13 pp).

Poinar HN, Schwarz C, Qi J, Shapiro B, Macphee RD, Buigues B, Tikhonov A, Huson DH, Tomsho LP, Auch A, Rampp M, Miller W, Schuster SC (2006): Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311: 392-394.

Prufer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, de Filippo C, Li H, Mallick S, Dannemann M, Fu Q, Kircher M, Kuhlwilm M, Lachmann M, Meyer M, Ongyerth M, Siebauer M, Theunert C, Tandon A, Moorjani P, Pickrell J, Mullikin JC, Vohr SH, Green RE, Hellmann I, Johnson PL, Blanche H, Cann H, Kitzman JO, Shendure J, Eichler EE, Lein ES, Bakken TE, Golovanova LV, Doronichev VB, Shunkov MV, Derevianko AP, Viola B, Slatkin M, Reich D, Kelso J, Paabo S (2014): The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505: 43-49.

Rawlence NJ, Lowe DJ, Wood JR, Young JM, Churchman GJ, Huang YT, Cooper A (2014): Using palaeoenvironmental DNA to reconstruct past environments: progress and prospects. *J Quat Sci* 29: 610-626.

Schubert M, Ginolhac A, Lindgreen S, Thompson JF, Al-Rasheid KA, Willerslev E, Krogh A, Orlando L (2012): Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* 13: 178 (15 pp).

Seguin-Orlando A, Gamba C, Sarkissian CD, Ermini L, Louvel G, Boulygina E, Sokolov A, Nedoluzhko A, Lorenzen ED, Lopez P, McDonald HG, Scott E, Tikhonov A, Stafford TWJr, Alfarhan AH, Alquraishi SA, Al-Rasheid KA, Shapiro B, Willerslev E, Prokhortchouk E, Orlando L (2015): Pros and cons of methylation-based enrichment methods for ancient DNA. *Sci Rep* 5: 11826 (15 pp).

Seguin-Orlando A, Schubert M, Clary J, Stagegaard J, Alberdi MT, Prado JL, Prieto A, Willerslev E, Orlando L (2013): Ligation bias in Illumina next-generation DNA libraries: implications for sequencing ancient genomes. *PLoS One* 8: e78575 (11 pp).

Shapiro B, Ho SY (2014): Ancient hyaenas highlight the old problem of estimating evolutionary rates. *Mol Ecol* 23: 499-501.

Shapiro B, Hofreiter M (Eds) (2012): "Ancient DNA: Methods and Protocols". Humana Press - Springer (New York, NY, USA).

Sheng GL, Soubrier J, Liu JY, Werdelin L, Llamas B, Thomson VA, Tuke J, Wu LJ, Hou XD, Chen QJ, Lai XL, Cooper A (2014): Pleistocene Chinese cave hyenas and the recent Eurasian history of the spotted hyena, *Crocuta crocuta*. *Mol Ecol* 23: 522-533.

Smith RW, Monroe C, Bolnick DA (2015): Detection of Cytosine methylation in ancient DNA from five native American populations using bisulfite sequencing. *PLoS One* 10: e0125344 (23 pp).

Star B, Nederbragt AJ, Hansen MH, Skage M, Gilfillan GD, Bradbury IR, Pampoulie C, Stenseth NC, Jakobsen KS, Jentoft S (2014): Palindromic sequence artifacts generated during next generation sequencing library preparation from historic and ancient DNA. *PLoS One* 9: e89676 (8 pp).

Teasdale MD, van Doorn NL, Fiddyment S, Webb CC, O'Connor T, Hofreiter M, Collins MJ, Bradley DG (2015): Paging through history: parchment as a reservoir of ancient DNA for next generation sequencing. *Philos Trans R Soc Lond B Biol Sci* 370: 20130379 (7 pp).

Thomsen PF, Willerslev E (2015): Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation* 183: 4-18.

Vernot B, Akey JM (2015): Complex history of admixture between modern humans and Neandertals. *Am J Hum Genet* 96: 448-453.