

## Starch genomics and bioarchaeology - Review

Gabriel Dorado <sup>1</sup>, Francisco Javier S. Sánchez-Cañete <sup>2</sup>, Plácido Pascual <sup>3</sup>, Inmaculada Jiménez <sup>4</sup>, Fernando Luque <sup>5</sup>, Margarita Pérez-Jiménez <sup>1</sup>, Patricia Raya <sup>6</sup>, Manuel Gálvez <sup>7</sup>, Jesús Sáiz <sup>8</sup>, Adela Sánchez <sup>8</sup>, Teresa E. Rosales <sup>9</sup> Víctor F. Vásquez <sup>9</sup>, Pilar Hernández <sup>10</sup>

<sup>1</sup> Author for correspondence, 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>; <sup>2</sup> EE.PP. Sagrada Familia de Baena, Avda. Padre Villoslada 22, 14850 Baena (Córdoba); <sup>3</sup> Laboratorio Agroalimentario de Córdoba, Consejería de Agricultura y Pesca, Junta de Andalucía, 14004 Córdoba; <sup>4</sup> IES Puertas del Campo, Avda. San Juan de Dios 1, 51001 Ceuta; <sup>5</sup> Laboratorio de Producción y Sanidad Animal de Córdoba, Ctra. Madrid-Cádiz km 395, 14071 Córdoba; <sup>6</sup> Servicio de Protección Radiológica, Facultad de Medicina, Avda. Menéndez Pidal s/n, Universidad de Córdoba, 14071 Córdoba; <sup>7</sup> Dep. Radiología y Medicina Física, Unidad de Física Médica, Facultad de Medicina, Avda. Menéndez Pidal s/n, Universidad de Córdoba, 14071 Córdoba; <sup>8</sup> Dep. Farmacología, Toxicología y Medicina Legal y Forense, Facultad de Medicina, Avda. Menéndez Pidal, s/n, Universidad de Córdoba, 14071 Córdoba; <sup>9</sup> Centro de Investigaciones Arqueobiológicas y Paleoecológicas Andinas ARQUEOBIOS, Apartado Postal 595, Trujillo (Peru); <sup>10</sup> Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Alameda del Obispo s/n, 14080 Córdoba

### Abstract

Glucose is preferentially stored as starch in plants. Thus, the remains of such polysaccharide may represent an invaluable source of archaeological evidence. The genomics, proteomics and metabolomics pathways can be also exploited to ascertain key events in phylogenetic processes. This includes the reconstruction of polyploidization events in plants and the identification of taxonomic positions. These developments are particularly relevant for the study of selection, evolution and domestication events of microorganisms, plants and animals. They include the adaptation to such food resource and the study of paleodiets and their implications for ancient and modern animals, including humans.

*Key words:* archaeobotany, ancient DNA, aDNA, sequencing, dendrogram, phylogenetic tree, genes, proteins, biochemical pathways.

### Resumen

La glucosa se almacena preferentemente como almidón en plantas. Por tanto, los restos de tal polisacárido pueden representar una fuente inestimable de información arqueológica. La genómica, proteómica y rutas metabólicas pueden ser también explotadas para determinar eventos cruciales en procesos filogenéticos. Estos incluyen la reconstrucción de eventos de poliploidización en plantas y la identificación de posiciones taxonómicas. Estos desarrollos son particularmente relevantes para el estudio de procesos de selección, evolución y domesticación de microorganismos, plantas y animales. Incluyen la adaptación a tal recurso alimentario y el estudio de las paleodietas y sus implicaciones para animales ancestrales y modernos, incluyendo los humanos.

*Palabras clave:* arqueobotánica, ADN antiguo, ADNa, secuenciación, dendrograma, árbol filogenético, genes, proteínas, rutas metabólicas.

**Introduction**

Although lipids have more energy per dry weight than carbohydrates (glucids), the latter are the preferential energy-substrate for cells. This stems from the fact that the catabolism in living cells is primarily engineered to use carbohydrates. Thus, energy can be obtained in an easier and faster way from carbohydrates as substrates. More specifically, glucose is the best substrate for catabolism. Therefore, the cells have developed anabolic pathways to store glucose as a chemical-energy source. These include the biosynthesis of starch in plants and glycogen in animals. The former is indeed made of two types of molecules: 20 to 25% amylose (linear and helical); and 75 to 80% amylopectin (branched) by weight, depending on the plant species. The latter is a more-branched version of amylopectin (Figs. 1-3).

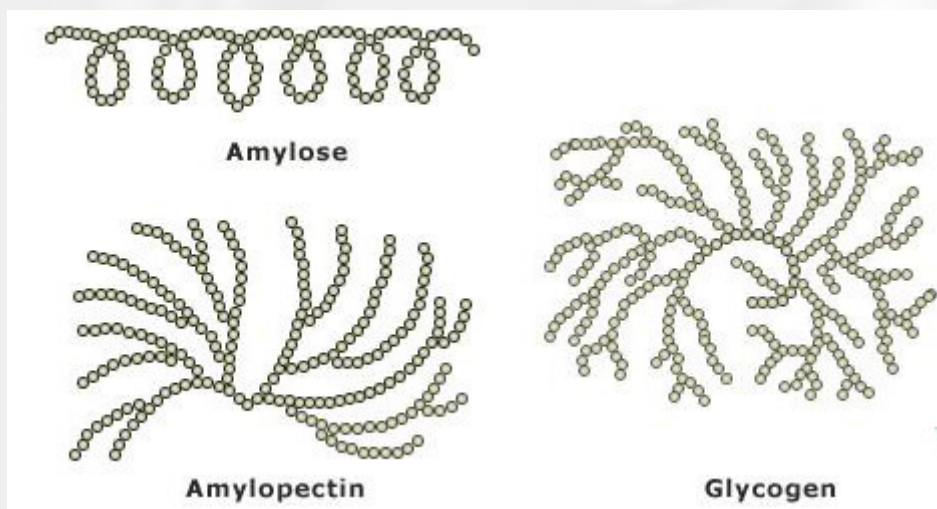


Figure 1. Ball structure of amylose, amylopectin and glycogen. Figure credit: Storage polysaccharides. © 2006 NEUROtiker, Wikimedia Commons <<http://commons.wikimedia.org>> and Creative Commons <<http://creativecommons.org>>.

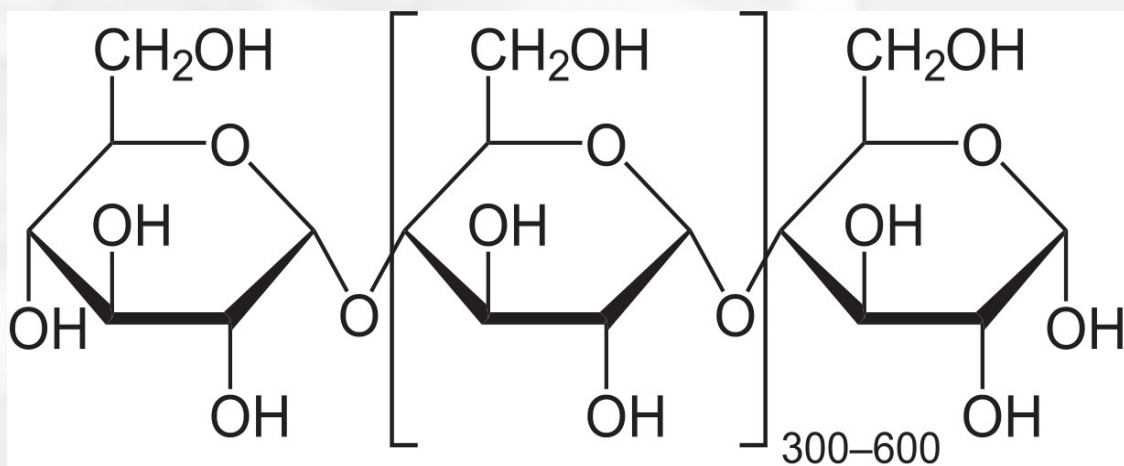


Figure 2. Chemical structure of amylose. Figure credit: Amylose. © 2007 NEUROtiker, Wikimedia Commons <<http://commons.wikimedia.org>> and Creative Commons <<http://creativecommons.org>>.

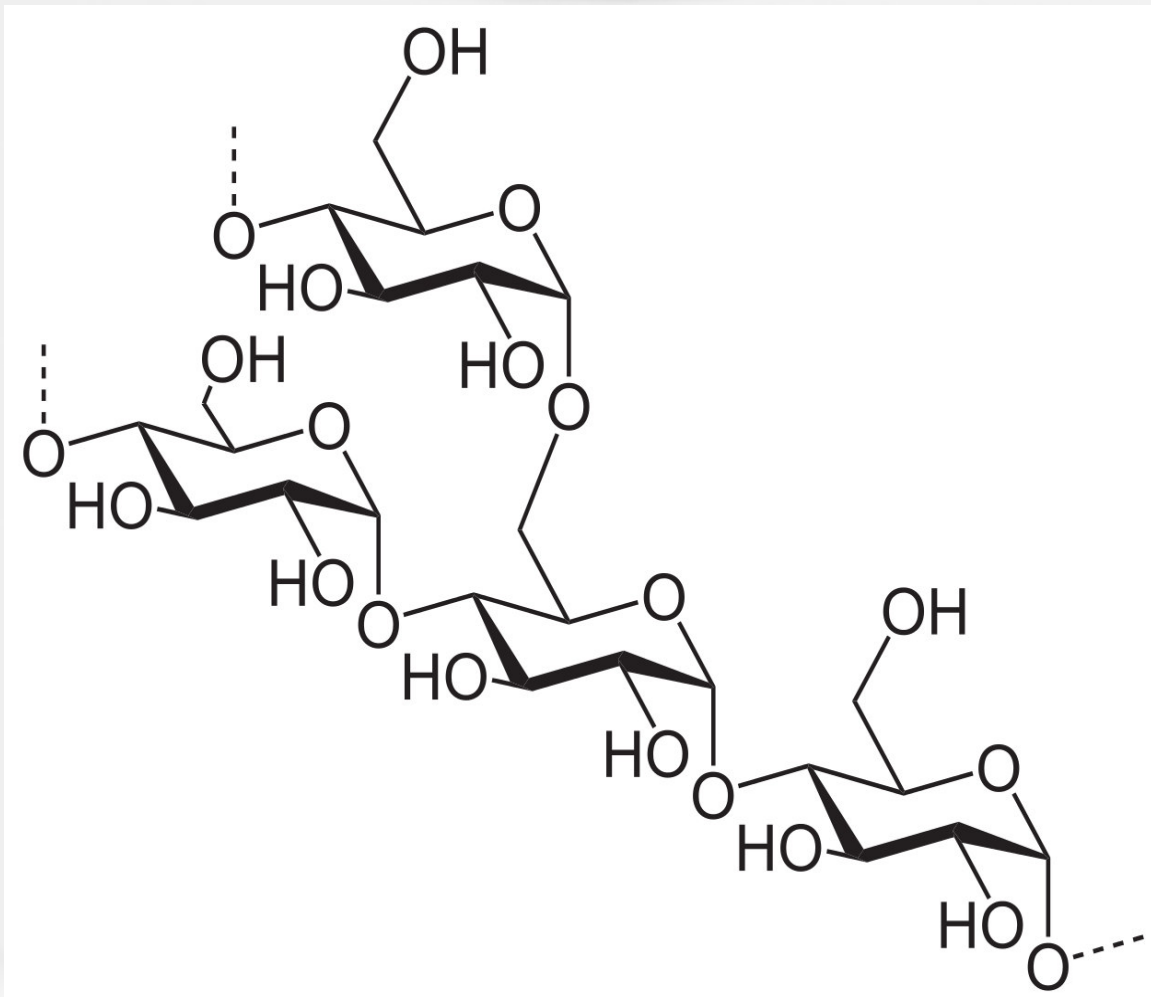


Figure 3. Chemical structure of amylopectin. Figure credit: Amylopectin. © 2008 NEUROtiker, Wikimedia Commons <<http://commons.wikimedia.org>> and Creative Commons <<http://creativecommons.org>>.

On the other hand, animals do not usually contain a high percentage of glycogen (which is mainly stored in the liver and, at a lower extent, in the muscles). Thus, they usually have other energy-rich components like proteins, which are usually less abundant in plants. The latter may accumulate large quantities of glucose as sucrose (saccharose) disaccharide (made of glucose and fructose), like in the case of the sugar beet and sugar cane. Other plants like potato (*Solanum tuberosum*) and cereals like wheat (*Triticum aestivum*), rice (*Oryza sativa*), maize (*Zea mays*), etc store glucose as starch (Fig. 4).

Curiously, fungi also store glycogen, being indeed more related to animals than plants, as demonstrated by the genome-sequencing projects in the past years. This is also in agreement with the fact that the cell-wall of fungi is not made of cellulose (as plants), but of chitin (as arthropods).

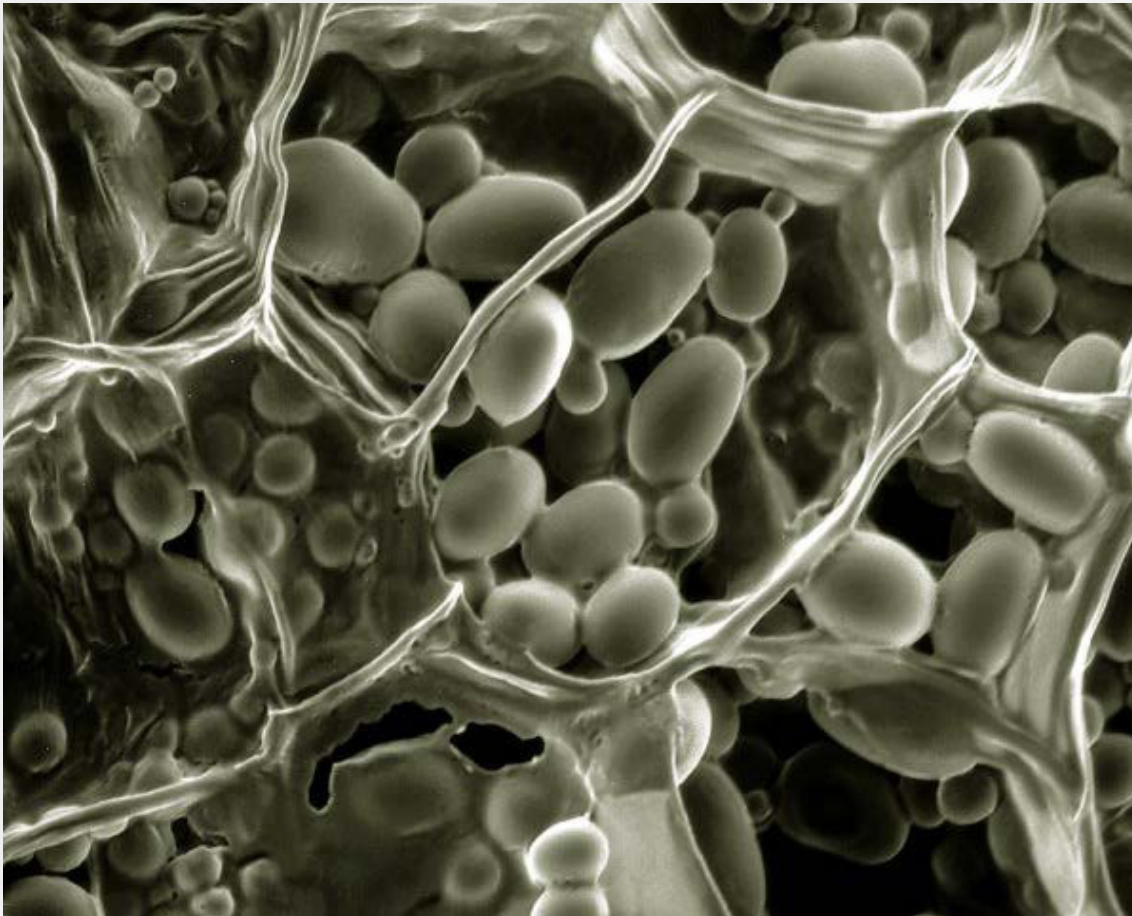


Figure 4. Scanning electron micrograph of starch-containing potato cells. Figure credit: Potato cells. © 1998 Philippa Uwins, Wikimedia Commons <<http://commons.wikimedia.org>> and Creative Commons <<http://creativecommons.org>>.

As a consequence, animals in general and humans in particular have developed metabolic pathways to use glucose, sucrose, glycogen and starch as energy sources. This has meant the adaptation towards the available food-resources by means of the evolution of the corresponding genomics, proteomics and metabolomics pathways. It has also meant the selection and domestication of microorganisms, plants and animals best adapted for the use of such food sources.

It is difficult to find glycogen in archaeological remains, for obvious reasons. But, fortunately, the starch remains are relatively common. Thus, they have a significant importance in bioarchaeology, being used to infer and ascertain paleodiets, as has been done with stable isotopes (Dorado et al, 2012). The starch-biosynthesis genes are also used for phylogenetic studies, as described below.

## Starch biosynthesis, genomics and plant evolution

The starch biosynthesis and accumulation are important for the plant survival, being considered a circadian output process (Serrano et al, 2009). Additionally, the study of ancient DNA and its comparison with modern genomes has become an invaluable tool in modern archaeology, phylogenetic and evolutionary studies (Dorado et al, 2007-2013). Thus, the analyses of several genes, including the ones involved in starch production, have allowed to reveal that current maize alleles were already present 4,400 years ago, albeit allelic selection may be still working (Jaenicke-Despres et al, 2003; Dorado et al, 2011).

Other interesting results involve the Shrunken 2 (*Sh2*) gene, which encodes the large subunit of the ADP-glucose pyrophosphorylase (AGPase; major endosperm-starch biosynthesis enzyme). Surprisingly, the polymorphism of such a gene in teosinte (maize ancestor) and 50 maize accessions is very low (with highly conserved genes and protein sequences), as compared to other genes and proteins. This indicates a purifying selection, predating domestication. An ancient selective-sweep in such a gene was also discovered when comparing *Zea mays* subspecies and outgroups. Thus, the *Sh2* alleles were maintained by natural selection before the domestication of such species, without further involvement in the domestication from teosinte into the cultivated maize (Manicacci et al, 2007).

Additionally, the waxy (*Wx*) gene encodes a granule-bound starch synthase (GBSS) enzyme, being involved in the amylose synthesis of plants. Interestingly, the bioinformatics analyses have allowed to propose an ancient *Gbss* gene transfer from maize (*Zea mays*) into *Dimeria lawsonii*. Thus, the polymorphic region for exons 8-9 was amplified by the polymerase chain-reaction (PCR), sequenced and aligned, generating the corresponding dendrogram (phylogenetic tree) (Baranov et al, 2014).

On the other hand, most eukaryotes are diploid, thus having two paired (homologous) sets of chromosomes (one inherited from each parent). The duplication of a whole chromosome set is known as polyploidization. This generates extra sets of chromosomes, being especially common in vascular plants (Pecinka et al, 2011). The polyploidization may contribute to the species adaptation and thus evolutionary success. Yet, the study and reconstruction of polyploidization may be complicated, due to the complex gene and genome histories in plants. Fortunately, the phylogenetic studies of the genome, including paralogs of the waxy gene have allowed to dissect such process (Yuan et al, 2006; Okuyama et al, 2012). In this respect, the sequencing of the *Gbss* gene in the Rosaceae family has revealed the presence of several loci, with differences in the presence or absence of introns and mutations in splicing sites. Such phylogenetic studies have shed new light on the evolution and speciation of such species (Evans et al, 2000).

Additionally, the retroposons have been proposed as promising molecular markers to analyze the plant polyploid evolution. Thus, the sequencing of the *Gbss* gene from diploid, tetraploid and hexaploid species in

the Geinae group (Colurieae: Rosaceae) suggest that the hexaploid species were generated via two consecutive allopolyploidization events. Besides, the allopolyploid speciation events were relatively ancient (Geinae should be ~17 million years old) (Smedmark et al, 2003). Likewise, the analyses of retroposon insertion events into the *Gbss* gene of the polyploid species belladonna (*Atropa belladonna*), have allowed to reveal its taxonomic position (Hyoscyameae tribe; Solanaceae family) and ancient hybrid-history. This has been confirmed by phylogenetic analyses of this medically important plant (Yuan et al, 2006).

Likewise, the water lily known as sacred lotus (*Nelumbo nucifera*) is also relevant for phylogenetic studies of flowering plants, as a basal eudicot. It is used for ornamental and medical purposes, besides being a food source. Its genome has been recently sequenced, allowing confirmation of its evolutionary position, additionally suggesting an interesting whole-genome duplication event. Furthermore, an expansion of starch synthase genes was also found. This was likely related to higher rates of starch synthesis and storage into the rhizomes (Wang et al, 2013).

### **Starch resources and species domestication**

The starch-grain remains have been used to analyze plant domestication in archaeobotanical studies (Piperno, 2012). Different paleoecological, archaeobotanical and genetic-molecular studies involving ancient DNA and starch have been used to ascertain the origin of plant domestication and thus agriculture (Zizumbo-Villarreal and Colunga-GarciaMarin, 2010). It was initially considered that agriculture allowed the development of ancient settlements and complex-societies, eventually giving rise to religious rituals. Yet, the shocking findings at Gobekli Tepe (Turkey) have reversed such assumption, being the oldest religious temple to date. Thus, as Klaus Schmidt pointed out, “first came the temple, then the city”. Indeed, such a discovery has changed everything in the sense that it shows that the building of monumental temple complexes was carried out by hunter-gatherers, before the establishment of sedentary farming-societies and cities with their further civilization developments (Dietrich et al, 2012).

On the other hand, the study of ancient DNA has also demonstrated the domestication of wolves (*Canis lupus*) to generate dogs (*Canis lupus familiaris*) (Dorado et al, 2009). Curiously, the genome sequencing of such animals has revealed the evolution of the starch metabolism in the domestication process. Thus, a total of 10 genes encoding proteins with key roles in starch metabolism were identified. The mutations in such genes should have contributed to the switching from a carnivorous diet (wolf) to a starch-rich one (dog), thus facilitating its adaptation and domestication by human communities (Axelsson et al, 2013).

Of course, domestication does not only involve plants and animals, but also microorganisms. Thus, *Aspergillus flavus* is a widely-distributed saprotrophic and pathogenic filamentous-fungus. Interestingly it is considered that its ancient domestication by means of starch fermentation into simple sugars gave rise to *Aspergillus oryzae*. Therefore, the empiric selection by

humans to increase the starch-degrading-enzyme secretion generated to the new species, which is widely used nowadays (Hunter et al, 2011).

### Starch and paleodiets

Ancient civilizations considered the consumption of ethanol (fermented sugars) and other drugs as a way to contact the spiritual world. Thus, the fermentation of starch to generate beer was an important technology in different civilizations, as in the ancient Egypt, being also consumed as staple food. Some archaeological residues containing starch have shown a well-preserved starch microstructure, being subjected to different analyses, including scanning-electron microscopy (SEM). This has revealed interesting details of the brewing process, like the use of dates instead of bread as the standard ingredient (Samuel, 1996).

Some archaeological remains of starch grains from potato (Andean plant) date back to 6,000 years ago. They have been also found in pre-ceramic sites in the northern coast of Peru, revealing that it was not an exclusive Andean crop. Furthermore, the microscopic analyses of sediments adhering ceramic-fragments has allowed the visualization and isolation of starch grains in Andean Prehispanic cultures. Indeed, such microbotanic fragments have allowed to update the archaeobotanical record. Thus, it was possible to reveal the presence of starch remains of such plant in the domestic pottery of the Moche period (200 to 700 AD) (Vásquez and Rosales, 2014).

On the other hand, the fossilized feces (coprolites) may represent a useful source of archaeological evidence. They may contain starch among other debris, thus revealing paleodiets. This way, it has been found that the ancient diets of the inhabitants of some southwest American regions were diverse and rich in fiber, including plants containing inulin instead of starch. Curiously, this may be related to the fact that the current descendants of such populations are more prone to obesity and type II diabetes when eating starch-rich Western diets (Bryant and Reinhard, 2012).

Additionally, the ancient human skeletons, and in particular the dental remains, may also be useful sources of archaeological data. This is specially relevant for edible parts of plants, which may be otherwise difficult to preserve in archaeological remains. Thus, the presence of starch remains in teeth calculi has revealed the consumption of cultivated plants. This has allowed to infer ancient diets and agricultural activities (Piperno and Dillehay, 2008).

Finally, the Neandertals belong to an extinct subspecies of *Homo sapiens* (*Homo sapiens neanderthalensis*), sharing genome regions with modern man (*Homo sapiens sapiens*), as demonstrated by ancient DNA (aDNA) sequencing (Dorado et al, 2008, 2013). Some have proposed that Neandertals did not consume plant foods, thus facilitating their extinction. Yet, the study of dental calculi of Neanderthal skeletons has revealed the presence of starch grains. Indeed, their diet included dates, legumes and grass seeds. Curiously, many starches from grass seeds exhibited damages typical of cooking, suggesting a sophisticated diet (Henry et al, 2011).

## Concluding remarks

Starch is an ubiquitous food resource in nature as an efficient glucose-storage in plants. Thus, plant and animal remains containing such polysaccharide may represent invaluable sources of archaeological evidence, as in the case of paleodiets. This way, the starch genomics, proteomics and metabolomics pathways are also useful tools to ascertain the phylogeny and evolution of microorganisms, plants and animals, including domestication events.

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

## References

- Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar A, Lindblad-Toh K (2013): The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495: 360-364.
- Baranov YO, Slischuk GI, Volkova NE, Sivolap YM (2014): Bioinformatic analysis of maize granule-bound starch synthase gene. *Cytol Genet* 48: 150-154.
- Bryant VM, Reinhard KJ (2012): Coprolites and archaeology: the missing links in understanding human health. *Bulletin of the New Mexico Museum of Natural History and Science* 57: 379-387.
- Dietrich O, Heun M, Notroff J, Schmidt K, Zarnkow M (2012): The role of cult and feasting in the emergence of Neolithic communities. New evidence from Gobekli Tepe, south-eastern Turkey. *Antiquity* 86: 674-695.
- 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, 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.
- Evans RC, Alice LA, Campbell CS, Kellogg EA, Dickinson TA (2000): The granule-bound starch synthase (GBSSI) gene in the Rosaceae: multiple loci and phylogenetic utility. *Mol Phylogenet Evol* 17: 388-400.
- Henry AG, Brooks AS, Piperno DR (2011): Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). *Proc Natl Acad Sci USA* 108: 486-491.
- Hunter AJ, Jin B, Kelly JM (2011): Independent duplications of alpha-amylase in different strains of *Aspergillus oryzae*. *Fungal Genet Biol* 48: 438-444.
- Jaenicke-Despres V, Buckler ES, Smith BD, Gilbert MT, Cooper A, Doebley J, Paabo S (2003): Early allelic selection in maize as revealed by ancient DNA. *Science* 302: 1206-1208.
- Manicacci D, Falque M, Le Guillou S, Piegou B, Henry AM, Le Guilloux M, Damerval C, De Vienne D (2007): Maize *Sh2* gene is constrained by natural selection but escaped domestication. *J Evol Biol* 20: 503-516.
- Okuyama Y, Tanabe AS, Kato M (2012): Entangling ancient allotetraploidization in Asian *Mitella*: an integrated approach for multilocus combinations. *Mol Biol Evol* 29: 429-439.
- Pecinka A, Fang W, Rehmsmeier M, Levy AA, Mittelsten Scheid O (2011): Polyploidization increases meiotic recombination frequency in *Arabidopsis*. *BMC Biol* 9: 24 (7 pp).
- Piperno DR (2012): New archaeobotanical information on early cultivation and plant domestication involving microplant (phytolith and starch grain) remains. In Gepts P, Famula TR, Bettinger RL, Brush SB, Damania AB, McGuire PE, Qualset CO (Eds): "Biodiversity in Agriculture: Domestication, Evolution and Sustainability". Cambridge University Press (Cambridge, UK): 136-159.
- Piperno DR, Dillehay TD (2008): Starch grains on human teeth reveal early broad crop diet in northern Peru. *Proc Natl Acad Sci USA* 105: 19622-19627.
- Samuel D (1996): Archaeology of ancient Egyptian beer. *J Am Soc Brew Chem* 54: 3-12.

Serrano G, Herrera-Palau R, Romero JM, Serrano A, Coupland G, Valverde F (2009): *Chlamydomonas* CONSTANS and the evolution of plant photoperiodic signaling. *Curr Biol* 19: 359-368.

Smedmark JE, Eriksson T, Evans RC, Campbell CS (2003): Ancient allopolyploid speciation in Geinae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. *Syst Biol* 52: 374-385.

Vásquez VF, Rosales TE (2014): Origen del maíz andino: arqueobotánica y genética molecular. *Revista Arqueológica SIAN* 22:2-19.

Wang Y, Fan G, Liu Y, Sun F, Shi C, Liu X, Peng J, Chen W, Huang X, Cheng S, Liu Y, Liang X, Zhu H, Bian C, Zhong L, Lv T, Dong H, Liu W, Zhong X, Chen J, Quan Z, Wang Z, Tan B, Lin C, Mu F, Xu X, Ding Y, Guo AY, Wang J, Ke W (2013): The sacred lotus genome provides insights into the evolution of flowering plants. *Plant J* 76: 557-67.

Yuan YW, Zhang ZY, Chen ZD, Olmstead RG (2006): Tracking ancient polyploids: a retroposon insertion reveals an extinct diploid ancestor in the polyploid origin of belladonna. *Mol Biol Evol* 23: 2263-2267.

Zizumbo-Villarreal D, Colunga-GarcíaMarín P (2010): Origin of agriculture and plant domestication in West Mesoamerica. *Genet Res Crop Evol* 57: 813-825.