

## REVIEW

## Ancient nucleic acids from maize - A review

Gabriel Dorado <sup>1</sup>, Teresa Esperanza Rosales <sup>2</sup>, Fernando Luque <sup>4</sup>, Francisco Javier S. Sánchez-Cañete <sup>4</sup>, Isabel Rey <sup>5</sup>, Inmaculada Jiménez <sup>6</sup>, Arturo Morales <sup>7</sup>, Manuel Gálvez <sup>8</sup>, Jesús Sáiz <sup>9</sup>, Adela Sánchez <sup>9</sup>, Víctor F. Vásquez <sup>10</sup>, Pilar Hernández <sup>11</sup>

<sup>1</sup> Author for correspondence, Dep. Bioquímica y Biología Molecular, Campus Rabanales C6-1-E17, Universidad de Córdoba, 14071 Córdoba (Spain), eMail: <bb1dopeg@uco.es>; <sup>2</sup> Laboratorio de Arqueobiología, Avda. Universitaria s/n, Universidad Nacional de Trujillo, Trujillo (Peru); <sup>3</sup> Laboratorio de Producción y Sanidad Animal de Córdoba, Ctra. Madrid-Cádiz km 395, 14071 Córdoba; <sup>4</sup> EE.PP. Sagrada Familia de Baena, Avda. Padre Villoslada 22, 14850 Baena (Córdoba); <sup>5</sup> Colección de Tejidos y ADN, Museo Natural de Ciencias Naturales, 28006 Madrid; <sup>6</sup> IES Puertas del Campo, Avda. San Juan de Dios 1, 51001 Ceuta; <sup>7</sup> Dep. Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Cantoblanco (Madrid); <sup>8</sup> Dep. Radiología y Medicina Física, Unidad de Física Médica, Facultad de Medicina, Avda. Menéndez Pidal s/n, Universidad de Córdoba, 14071 Córdoba; <sup>9</sup> Dep. Farmacología, Toxicología y Medicina Legal y Forense, Facultad de Medicina, Avda. Menéndez Pidal, s/n, Universidad de Córdoba, 14071 Córdoba; <sup>10</sup> Centro de Investigaciones Arqueobiológicas y Paleoecológicas Andinas ARQUEOBIOS, Apartado Postal 595, Trujillo (Peru); <sup>11</sup> Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Alameda del Obispo s/n, 14080 Córdoba

## Abstract

The maize (*Zea mays*) is a widely cultivated American cereal grass, being domesticated from the teosinte in southwestern Mexico about 9,000 years ago. The maize was a cornerstone on the development and preponderance of the Maya and Aztec civilizations, as well as other ancient American cultures. Formerly, the maize remains were analyzed from a morphological point of view. Fortunately, the development of molecular biology methodologies has allowed the study of ancient DNA (aDNA). Reports of uniformity and lack of polymorphism of maize aDNA could be due to a biased selection process favoring homozygosity. The available results suggest that the current maize gene pool is millions of years old, and that the current domesticated varieties are derived from several wild ancestral populations. On the other hand, the new second- and third-generation nucleic acid sequencing platforms allow to sequence full genomes in a cost-efficient way, and most importantly from an archaeological point of view, they even allow the sequencing of ancient genomes. The possibility to analyze the ancient RNA (aRNA) is also exciting. The study ancient nucleic acids from maize using the new sequencing technologies opens the door to understand not only the evolution and domestication of such cereal grass, but also to decipher the mysteries of some American cultures from a multi disciplinary point of view.

*Key words:* Indian maize, corn, ears, cobs, kernels, molecular markers, microsatellites, simple sequence repeats, SSR, aRNA, amRNA, nuclear DNA, nuDNA, chloroplastial DNA, cpDNA, mitochondrial DNA, mtDNA.

## Resumen

El maíz (*Zea mays*) es un cereal herbáceo americano extensamente cultivado, que fue domesticado a partir del teocinte suroeste de Méjico hace unos 9.000 años. El maíz fue una piedra angular para el desarrollo y preponderancia de la civilización Maya y Azteca, así como otras culturas antiguas americanas. Inicialmente, los restos del maíz fueron analizados desde un punto de vista morfológico. Afortunadamente, el desarrollo de métodos de biología molecular ha permitido el estudio del ADN antiguo (ADNa). Diversos estudios acerca de la uniformidad y la falta de polimorfismo del aDNA del maíz indican que podría deberse a un proceso de selección sesgada que favorece la homocigosis. Los resultados disponibles sugieren que el actual acervo genético del maíz tiene millones de años de antigüedad, y que las actuales variedades domesticadas son el resultado de diversas poblaciones ancestrales silvestres. Por otro lado, los nuevos programas de secuenciación de ácidos nucleicos de segunda y tercera generación permiten secuenciar genomas completos de manera eficaz, y lo más importante desde el punto de vista arqueológico, incluso permiten la secuenciación de genomas antiguos. La posibilidad de analizar el ARN antiguo (ARNa) es también apasionante. El estudio de los ácidos nucleicos antiguos del maíz usando las nuevas técnicas de secuenciación abre la puerta para entender no sólo la evolución y domesticación de la mayoría de los cereales herbáceos, sino también para descifrar los misterios de algunas de las culturas americanas desde un punto de vista multidisciplinar.

*Palabras clave:* maíz indio, mazorca, marcadores moleculares, microsátelites, polimorfismo de nucleótido sencillo, SSR, ARNa, ARNma, microARN, ADN nuclear, ADNnu, ADN cloroplastidial, ADNcp, ADN mitocondrial, ADNmt.

## Introduction

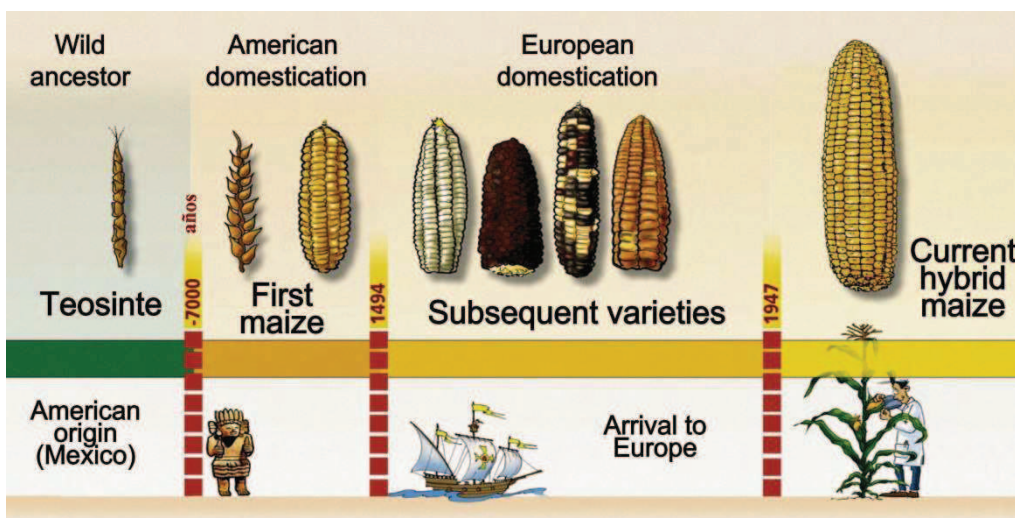
The *Zea mays* plant is known as maize, Indian maize or corn. It is a widely cultivated American cereal grass, bearing seeds on elongated ears (cobs). The words maize and corn are also used to designate the cobs and its edible seeds (kernels). From a biochemical point of view, the maize is a C4 carbon-fixation plant, being therefore more efficient than C3 plants under certain conditions, as drought and high temperatures. Thus, the maize uses more efficiently precious resources like water, maintaining in this way the soil moisture more efficiently, and thus being better capable of growing in arid environments. In fact, its shallow root system makes such plant dependent on soil moisture. On the other hand, maize is intolerant to cold. These facts have a significant agronomical impact for maize breeding.

From an archaeological and historical point of view, the maize cultivation has had an extraordinary relevance. Previous to the industrial revolution and international trading as a source of commerce and thus currency and wealth, ancient cultures flourished only when a stable and abundant food supply was available, once agricultural practices started about 15,000 years Before Present (BP) in the Middle East. Cereal grasses like wheat (Middle East), rice (Asia) and maize (America) have played a key role on the emergence and further

development of ancient cultures. Thus, the maize had a vital influence on the Maya and Aztec civilizations, as well as other ancient American cultures. Actually, the archaeological and genetic evidence suggests that about 9,000 years BP there was a single domestication event from teosinte by the Aztec, in the Rio Balsas drainage area of southwestern Mexico (Benz, 2001; Matsuoka et al, 2002; Lia et al, 2007; Zizumbo-Villarreal and Colunga-GarciaMarin, 2010) (Figs. 1 & 2).



**Figure 1. Maize domestication.** Graphical representation of the maize domestication from the small teosinte ear to the large cob of the modern hybrid corn. Credit: original picture from Prof. Dr. José Ignacio Cubero; first published as the front cover of his book “Introducción a la Mejora Genética Vegetal” [“Introduction to Plant Breeding”]; © Mundi-Prensa (1999) <<http://www.mundiprensa.com>>, now Ediciones Paraninfo <<http://www.paraninfo.es>> (Grupo Paraninfo) <<http://www.grupoparaninfo.com>>.



**Figure 2. Maize domestication timeline.** Chronology of the maize domestication; a remarkable breeding accomplishment that took thousands of years. Credit: Groupement National Interprofessionnel des Semences et plants (GNIS); © Gnis, 2006 <<http://www.gnis.fr>>.

## Maize archaeology, ancient DNA and genomics

Due to its strategic relevance to maintain human civilizations, the maize ears and the corn seeds have been a typical element on archaeological remains of the Aztec and other American civilizations. Thus seeds and even full ears have been preserved to our days. This has been favored by the presence of dry climates, as is the case of some sites located on the Pacific coast of America, as is the case of Peru. Formerly, the maize remains were analyzed from a morphological point of view, using paleoecological and archaeobotanical data. Fortunately, the development of molecular biology methodologies has allowed the study of ancient DNA (aDNA), which brings such powerful technology to the archaeology (Rollo et al, 1988, 1994; Goloubinoff et al, 1993, 1994; Bird et al, 1995; Jaenicke-Despres, 2003; Dorado et al, 2007-2009; Lia et al, 2007; Schlumbaum et al, 2008). Such methodologies may involve the use of the Polymerase Chain Reaction (PCR) *in vitro* amplification, as well as modern second- and third-generation sequencing approaches, as described below. It is important to remark here that besides the scarcity of material and the fact that the DNA may be highly fragmented and chemically degraded (eg., tautomeric nitrogenous base changes), special care should be taken to avoid cross-contamination of aDNA with modern DNA (mDNA) (Goloubinoff et al, 1993, 1994).

The analyses of plant aDNA in general (Schlumbaum et al, 2008) and maize aDNA in particular opens the door for genomic comparisons with modern maize (synteny). This allows to study the taxonomy of ancient species and varieties, their evolutionary and domestication pathways, phylogeography and biogeography, as well as their adaptation to both abiotic and biotic stresses. One of the goals of these studies is to determine the formation of the maize-bean-squash multicrop milpa system and agro-food system in America, which started about 10,000 years ago and allowed the development of important ancient cultures in America (Zizumbo-Villarreal and Colunga-GarciaMarin, 2010). It is also important to note that the problem of dealing with scarce and incomplete data from the ancient samples to estimate the nitrogenous base substitution rates has been addressed by different statistical approaches, including the Bayesian Markov Chain Monte Carlo (MCMC) (Navascues and Emerson, 2009; Ho et al, 1011).

Concerning the molecular markers, the Variable Number of Tandem Repeats (VNTR) can be classified as minisatellites and microsatellites. The later are also known as Simple Sequence Repeats (SSR) in plants and Short Tandem Repeats (STR) in animals. They are very popular for DNA fingerprinting (genotyping) among plant and animal breeders, since they are codominant (the heterozygote can be differentiated from the two homozygotes), simple to analyze, amenable to automation, fast and cost-effective. Thus, they have also been used with aDNA in general and maize aDNA in particular. Reports of uniformity and lack of polymorphism of maize SSR from nuclear DNA (nuDNA) have been published from samples dated 300 to 2,000 (Lia et al, 2007), which could be due to a biased selection process favoring homozygosity. Indeed, even current rural populations around the world follow an agricultural process that involves selection of the best ears for the next season

sowings (instead of random seeds or the best seeds from the total pool of seeds). Such practice may drastically reduce the maize polymorphism, mainly if there are not germplasm introgressions from other places. Some other loci have also been analyzed on maize aDNA samples dated 4,400 to 4,700 years ago, including genes related to the gene alcohol dehydrogenase 2 (*adh2*) gene, as well as genes related to the starch biosynthesis, storage proteins and control of the plant architecture, finding the same alleles than the modern maize has. These results suggest that the current maize gene pool is millions of years old, and that the current domesticated varieties are derived from several wild ancestral populations (Goloubinoff et al, 1993, 1994; Jaenicke-Despres, 2003).

As a practical example of its usefulness, the ancient DNA has helped to reveal the routes taken by agriculture as it spread away from its origin. Such is the case of indigenous landraces of maize and archaeological specimens from Brazil, which have been used to study the origins of South American maize. They are based on the Single Nucleotide Polymorphisms (SNP) and different SSR found on the maize *adh2* gene. Thus, the simpler  $GA_n$  SSR has been found through the Andean region. A more complex  $GA_nTA$  SSR was found along the lowlands of the northeast coast of Brazil, which probably reached such areas through the river systems. The most complex  $GA_1AA_1GA_n$  SSR was found in Brazil and Paraguay. These results suggest two routes for the maize cultivation spreading into South America, as well as some cultural contact between Chile and Paraguay (Freitas et al, 2003).

On the other hand, the organelle DNA can be a good target for aDNA studies, since each cell may contain many chloroplasts and mitochondria (which in turn may contain many organellar genomes inside), instead of the two copies of the nuclear chromosomes found on diploid cells. Thus, plants may have about 5,000 to 10,000 chloroplast genomes (chloroplastial DNA; cpDNA) per cell and about 100 to 200 mitochondrial genomes (mitochondrial DNA; mtDNA) per cell. Besides, the organelles evolved from ancestral parasites that became endosymbionts millions of years ago, thus being protected by a double membrane envelope (the inner membrane derived from them and the outer one from the host cell). Furthermore, the organellar DNA is circular, and thus resistant to exonucleases (Besnard et al, 2011). Thus, the cytoplasmic DNA may be particularly relevant for aDNA studies, which usually deal with scarce and degraded DNA, as previously indicated.

Concerning genomics, the maize genome has 2.3 Gigabase pairs (Gbp) and about 33,000 genes (for comparison, the human genome has 3 Gbp and about 26,000 genes) as shown by the Maize Genome Sequencing Project (MGSP) <<http://www.maizesequence.org>>. Surprisingly, nearly 85% of the maize genome is made of transposons. Interestingly, most part of the wheat genome has been duplicated and reorganized by helitrons, which are a group of transposons that replicate by a rolling-circle mechanism (Schnable et al, 2009). Fortunately, the new second- and third-generation nucleic acid sequencing platforms allow to sequence full genomes in a cost-efficient way, and most importantly from an archaeological point of view, they even allow the sequencing of ancient genomes (Dorado et al, 2008), as demonstrated with several Neandertal samples collected at different sites, dating 38,000 to 70,000

years (Green et al, 2010). The sequencing of the modern maize genomes and the ones from archaeological samples should bring archaeology to unprecedented levels to decipher the evolution and domestication of his important crop.

The possibility to analyze the ancient RNA (aRNA), and in particular the messenger RNA (amRNA) and microRNA (miRNA) from archaeological samples is also exciting (Rollo et al, 1994). The microRNA are small RNA molecules (about 22 nucleotides long), found in most eukaryotic cells. They bind to complementary sequences on target mRNA, usually resulting in degradation of the target or translational inhibition, and thus gene silencing. This kind of regulation could date back more than 400 million years. An important group of genes related with the maize domestication are transcriptional regulators. Some studies have shown that maize MicroRNA target transcription factors and other genes involved in key cellular roles (Floyd and Bowman, 2004; Doebley et al, 2006; Zhang et al, 2006). Such molecules could be analyzed on ancient samples using third-generation nucleic acid platforms capable of reading single molecules, without requiring previous retrotranscription or amplification steps that could fail and bias the results.

In conclusion, the maize was a cornerstone on the development and preponderance of the Aztec and other old American cultures. The study of the aDNA and even aRNA from maize using the new sequencing technologies opens the door to understand not only the evolution and domestication of such cereal grass, but also to decipher the mysteries of some American cultures from a multidisciplinary point of view related to archaeology, biology, agronomy, geography, history and sociology.

*Acknowledgements.* Supported by “Ministerio de Ciencia e Innovación” (MICINN grants AGL2010-17316); “Consejería de Agricultura y Pesca” of “Junta de Andalucía” (041/C/2007, 75/C/2009 and 56/C/2010); “Grupo PAI” (AGR-248); and “Universidad de Córdoba” (“Ayuda a Grupos”), Spain. We are indebted to both Prof. Dr. José Ignacio Cubero and Mundi-Prensa <<http://www.mundiprensa.com>>, now Ediciones Paraninfo <<http://www.paraninfo.es>> (Grupo Paraninfo) <<http://www.grupoparaninfo.com>> for providing the picture shown on Fig. 1 with reproduction permission for this article of Archaeobios. We are also very grateful to the “Groupement National Interprofessionnel des Semences et plants” (GNIS) for providing the Fig. 2 and the reproduction permission for this article.

## References Cited

- Benz BF (2001): Archaeological evidence of teosinte domestication from Guilá Naquitz, Oaxaca. *Proc Natl Acad Sci USA* 98: 2104-2106.
- Besnard G, Hernández P, Khadari B, Dorado G, Savolainen V (2011): Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology* 11: 80-92.
- Bird R, Bohorova N, GonzalezDeLeon D, Hoisington D (1995): Old maize and ancient DNA. *Maize Gen Coop News* 69: 99-100.
- Doebley JF, Gaut BS, Smith BD (2006): The molecular genetics of crop domestication. *Cell* 127: 1309-1321.

- 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, Vásquez V, Rey I, Luque F, Jiménez I, Morales A, Gálvez M, Sáiz J, Sánchez A, Hernández P (2008): Sequencing ancient and modern genomes. *Archaeobios* 2: 75-80.
- Dorado G, Vásquez V, Rey I, Vega JL (2007): Archaeology meets Molecular Biology. *Archaeobios* 1: 1-2.
- Floyd SK, Bowman JL (2004): Gene regulation: ancient microRNA target sequences in plants. *Nature* 428: 485-486.
- Freitas FO, Bandel G, Allaby RG, Brown TA (2003): DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *Journal of Archaeological Science* 30: 901-908.
- Goloubinoff P, Paabo S, Wilson AC (1993): Evolution of maize inferred from sequence diversity of an *adh2* gene segment from archaeological specimens. *Proc Natl Acad Sci USA* 90: 1997-2001.
- Goloubinoff P, Paabo S, Wilson AC (1994): Molecular characterization of ancient maize: potentials and pitfalls. In Johannessen S, Hastorf CA: "Corn and culture in the prehistoric New World". Westview Press (Boulder): 113-125.
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspina AS, Jensen JD, Marques-Bonet T, Alkan C, Prufer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Hober B, Hoffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Z, Gusic I, Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PL, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, Paabo S (2010): A draft sequence of the Neandertal genome. *Science* 328: 710-722.
- Ho, SYW, Lanfear, R, Phillips, MJ, Matthew J, Barnes I, Thomas JA, Kolokotronis SO, Shapiro B (2011): Bayesian Estimation of Substitution Rates from Ancient DNA Sequences with Low Information Content. *Sys Biol* 60: 366-374.
- Jaenicke-Despres V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, Doebley J, Paabo S (2003): Early allelic selection in maize as revealed by ancient DNA. *Science* 302: 1206-1208.
- Lia VV, Confalonieri VA, Ratto N, Hernandez JAC, Alzogaray AMM, Poggio L, Brown TA (2007): Microsatellite typing of ancient maize: insights into the history of agriculture in southern South America. *Proc Roy Soc B-Biol Sci* 274: 545-554.
- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez G J, Buckler E, Doebley J (2002): A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99: 6080-6084.
- MGSP (2011): Maize Genome Sequencing Project <<http://www.maizesequence.org>>.
- Navascues M, Emerson BC (2009): Elevated substitution rate estimates from ancient DNA: model violation and bias of Bayesian methods. *Mol Ecol* 18: 4390-4397.

- Rollo F, Amici A, Salvi R, Garbuglia A (1988): Short but faithful pieces of ancient DNA. *Nature* 335: 774-774.
- Rollo F, Venanzi FM, Amici A (1994): Dried samples: Hard tissues: DNA and RNA from ancient plant seeds. In Herrmann B, Hummel S (eds): "Ancient DNA: Recovery and Analysis Of Genetic Material from Paleontological, Archaeological, Museum, Medical, and Forensic Specimens". Springer-Verlag (New York): 218-236.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddelloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009): The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112-1115.
- Schlumbaum A, Tensen M, Jaenicke-Despres V (2008): Ancient plant DNA in archaeobotany. *Veg His Archaeobot* 17: 233-244.
- Zhang B, Pan X, Anderson TA (2006): Identification of 188 conserved maize microRNAs and their targets. *FEBS Letters* 580: 3753-3762.
- Zizumbo-Villarreal D, Colunga-GarciaMarin P (2010): Origin of agriculture and plant domestication in West Mesoamerica. *Gen Res Crop Evol* 57: 813-825.