



编辑精选: 美国人类遗传学会第 58 届年会, 宾夕法尼亚州费城, 2008 年 11 月 11-15 日

# 现代墨西哥线粒体 DNA 多样性分布概况

安智利·莱山渡<sup>1,2</sup>, 北官·武国<sup>2,3</sup>, 龚·帕米尔·艾佳<sup>3</sup>, 希达·佛乐·瑞卡<sup>4</sup>, 瑞奇·可缙<sup>3</sup>, 胡奇·罗布<sup>3</sup>, 安霍·诺曼<sup>3</sup>, 多灵·安东<sup>2</sup>, 麦思·纳达<sup>3</sup>, 梧沃·苏格<sup>3</sup>

1. 佩鲁贾大学环境和细胞生物学系, 意大利 佩鲁贾 06100; 2. 帕维亚大学遗传学与微生物学系, 意大利 帕维亚 27100; 3. 索伦森分子家谱基金会, 美国 犹他州 盐湖城 84115; 4. 墨西哥社会安全研究所东北生物医学研究研究中心遗传学组, 墨西哥 新莱昂州 蒙特雷 64720

翻译: 张一帆, 康奈狄格大学健康中心血管生物学中心, 美国 康州 堡明屯 06030

**摘要:** 拥有超过 180 个语言民族群体的墨西哥是人类学和群体研究的丰富资源。这个国家见证了包括著名的玛雅文明和阿兹台克文明在内的重大文明的兴起和衰亡。但是由于欧洲的大规模殖民涌入, 其人口概况在过去的五个世纪里发生了剧烈的变化。今天不到 30% 的现代墨西哥人自我认同为完全或者部分美洲印地安人, 而其余的人群却似乎与他们在哥伦布时代之前的祖先鲜有共同之处。但是, 在仔细研究了母系遗传成分之后, 我们发现事实并非如此。通过对来自两千多个样本的线粒体 DNA 调控区序列(包括 HVS-I、HVS-II、HVS-III)的分析, 我们发现了现代墨西哥人群中存在着极大的美洲原住民痕迹。现代墨西哥人中, 90% 的线粒体 DNA 属于四个主要的泛美洲单倍群(A2、B2、C1、D1)。这个发现不仅支持欧洲人对墨西哥基因库的贡献主要是通过男性殖民者的学说, 还证实了采用单亲传递的线粒体 DNA 来重构一个国家历史这个方法的有效性。

Editors' Pick from the American Society of Human Genetics 58th Annual Meeting. Philadelphia, Pennsylvania, Nov. 11-15, 2008

## The Mitochondrial DNA Landscape of Modern Mexico

Alessandro ACHILLI<sup>1,2</sup>, Ugo A. PEREGO<sup>2,3</sup>, J. Edgar GOMEZ-PALMIERI<sup>3</sup>, Ricardo M. CERDA-FLORES<sup>4</sup>, Katie H. RITCHIE<sup>3</sup>, Robert HUGHES<sup>3</sup>, Norman ANGERHOFER<sup>3</sup>, Antonio TORRONI<sup>2</sup>, Natalie M. MYRES<sup>3</sup>, Scott R. WOODWARD<sup>3</sup>

1. Dip. di Biologia Cellulare e Ambientale, Università di Perugia, 06100 Perugia, Italy; 2. Dip. di Genetica e Microbiologia, Università di Pavia, 27100 Pavia, Italy; 3. Sorenson Molecular Genealogy Foundation, Salt Lake City, UT84115 USA; 4. Genetics Division, Northeast Biomedical Research Center, Instituto Mexicano del Seguro Social, Monterrey, Nuevo León, 64720 Mexico

Chinese translator: ZHANG Yifan, Center for Vascular Biology, University of Connecticut Health Center, Farmington, CT06030 USA

**ABSTRACT:** With more than 180 ethnic and linguistic groups, Mexico is a rich source for anthropological and population studies. This country witnessed the rise and fall of major civilizations, including the well-known Maya and Aztec civilizations, but as a result of heavy European colonization and influx, the population landscape has dramatically changed over the past five centuries. Today less than 30% of modern Mexicans identify themselves as being fully or partly Amerindians and the remaining population seems to have very little in common with their pre-Columbian ancestors. However, this is not the case when the maternal genetic component is evaluated in detail. Analysis of the mitochondrial DNA (mtDNA) control region sequences, including HVS-I, HVS-II and HVS-III, from more than 2,000 subjects revealed an overwhelming Native American legacy in the modern Mexican population, with ~90% of mtDNAs belonging to the four major pan-American haplogroups A2, B2, C1 and D1. This finding supports a European contribution to the Mexican gene pool primarily by male settlers and confirms the effectiveness of employing the uniparentally-transmitted mtDNA as a tool to reconstruct a country's history.

Agreement was received from the authors to translate this paper into Chinese.

在现今墨西哥的河谷和山地区域, 5000 年前放牧的美洲原住民部落在居民点定居下来, 然后开始发展农业。其中的一些村落在接下来的几个世纪中经历了巨大的增长。由此产生了复杂的社会经济、宗教和政治的城市中心, 最终成为了如玛雅(Maya)和阿兹台克(Aztecs)等伟大文明的基石[1]。在十六世纪初的欧洲人占领和殖民之后, 墨西哥的人口从此就被改造了。今天, 超过 1 亿的墨西哥

人口是由不同的人种和语言学种群构成的, 虽然大多数人视自己为欧洲人和土著祖先的混血(Mestizos)[2]。最近的估计表明 60% 人口的祖先是土著, 30% 是混血, 9% 纯种白人, 还有 1% 是其他(国家统计局, 地理及信息学, INEGI)。然而, 如果以语言作为选择依据, 美洲土著人口的比例则小于 10%, 而且这个比例呈现减小的趋势。墨西哥几乎整个人口都讲官方语言西班牙语。讲西班牙语的墨西

收稿日期: 2008 年 11 月 26 日 修回日期: 2008 年 12 月 9 日 联系人: 北官武国 Ugo@SMGF.org



表 1 墨西哥的线粒体单倍群频率 Tab. 1 Frequencies of mtDNA haplogroups in Mexico

Clade 分支	N 个数	频率 %	Clade 分支	N 个数	频率 %
<b>Native 土著</b>	<b>1800</b>	<b>89.06</b>	<b>West Eurasia 欧亚西部</b>	<b>171</b>	<b>8.46</b>
A2*	719	35.58	H*	56	2.77
A2e	62	3.07	H1a	4	0.20
A2g	29	1.43	H1c	2	0.10
A2h	29	1.43	H1l	4	0.20
A2i	6	0.30	H5	1	0.05
B2*	277	13.71	H6	3	0.15
B2a	79	3.91	J1b	8	0.40
C1b	212	10.49	J1c	3	0.15
C1c	226	11.18	J2a1	2	0.10
C1d	37	1.83	J2b	1	0.05
D1	112	5.54	K1a	14	0.69
D4h3	12	0.59	K2*	13	0.64
<b>Africa 非洲</b>	<b>42</b>	<b>2.08</b>	K2a2a	6	0.30
L0a1a	2	0.10	T1b	2	0.10
L0a2a1	1	0.05	T2*	3	0.15
L1c1	2	0.10	T2c	1	0.05
L1c2	1	0.05	T2d	4	0.20
L1c3a	2	0.10	U2e	11	0.54
L2a1	8	0.40	U3a	7	0.35
L2b	1	0.05	U4	1	0.05
L2c	8	0.40	U5a	7	0.35
L2d	1	0.05	U5b	7	0.35
L2e	1	0.05	U6a	3	0.15
L3b	3	0.15	U6d	1	0.05
L3d1	2	0.10	V	5	0.25
L3e1	2	0.10	X2b	2	0.10
L3e2	2	0.10	<b>East Asia 东亚</b>	<b>8</b>	<b>0.40</b>
L3e3,5	2	0.10	C1a	3	0.15
L3f2	3	0.15	F1a1a	4	0.20
L4	1	0.05	M*	1	0.05

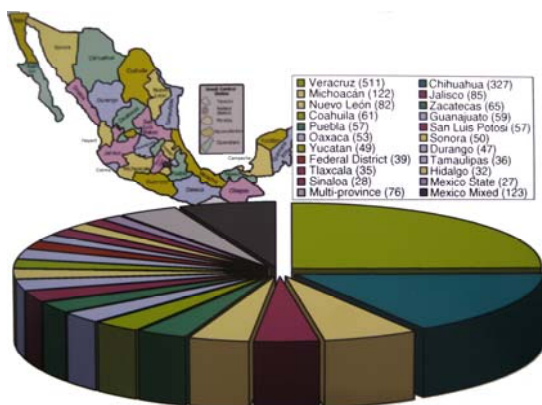


图 2 2021 个墨西哥样本的地理分布

Fig.2 Geographical distribution of 2021 Mexican samples

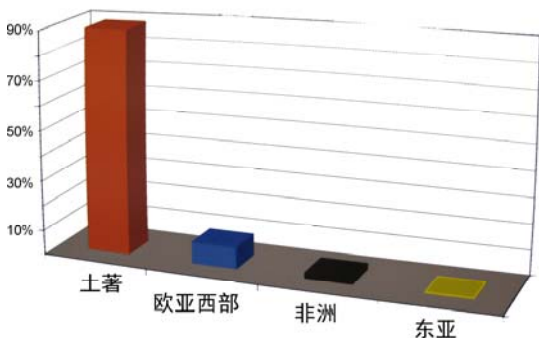


图 3 墨西哥的线粒体单倍群频率

Fig.3 Frequencies of mtDNA haplogroups in Mexico

线粒体 DNA 鉴定：线粒体 DNA 分子分析

的第一步包括 DNA 抽提、PCR 扩增、以及对整个线粒体调控区的测序。然后我们使用 SeuencerTM4.7(Gene Codes)软件对电泳带谱进行了对准、拼接和比较。最后，为利于清晰地单倍群分类，我们通过精确分析相对参考序列(rCRS [10])的突变差异来识别突变组合。我们使用 DnaSP 4.50 估测了某些对人群和/或世系之中和之间 DNA 序列变异的量度。还使用 MEGA 4.1 软件进行树状图构建，评估了线粒体 DNA 调控区变异的进化关联。最后，我们通过使用 Surfer 6.04(Gloden Software, Inc)执行 Kriging 程序获取了频率分布图，还通过整个数据组来推测在每个地图网格点。

二、结果和讨论

我们采集了大约 3000 个样本来评估现代墨西哥人的特征以及线粒体 DNA 变异的程度。通过对谱系数据精确地分析，我们发现其中 2577 个样本的母系祖先源头发生于墨西哥。现在，我们对核酸位置(nucleotide position, np)16000 到 np 580 的 1179 对碱基(base pairs, bps)进行了测序。这项测序覆盖了整个调控区，还包括了三个高变区 (HVS-1:



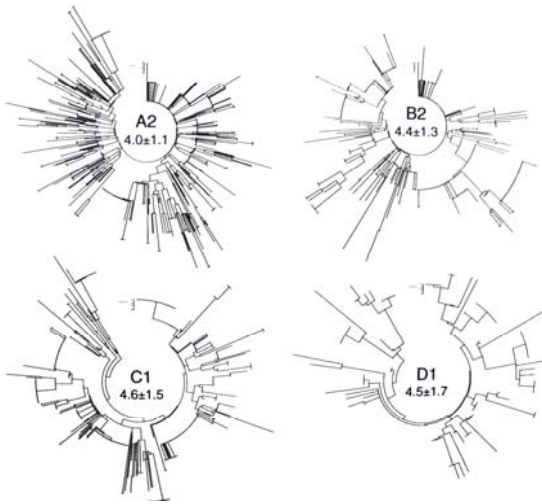


图 4 泛美洲单倍群系统树示意图

Fig.4 Schematic phylogenies of pan-American haplogroups

nps 16024-16383、HVS-II: nps 057-372、HVS-III: nps 438-576)。我们选取的样本代表了墨西哥联邦除金塔纳罗奥州(Quintana Roo)之外的 31 个州和特区(图 2)。

除去缺失和模糊的位点,我们在调控区域序列中发现了 283 个多态位点,核酸多样性  $\pi$ (Nei 1997)为  $0.00665 \pm 0.00193$ 。两个随机选择的序列之间的核酸差异(k, Tajima 1983)的平均值是  $7.062 \pm 2.356$ 。不同的突变组合构成了 583 个不同的单倍型。这显示了极高的多样性指数(diversity index)— $Hd=0.9874 \pm 0.0009$ (Nei 1987)。我们通过对突变诊断组合(mutational diagnostic motifs)的精确统计,把线粒体 DNA 归入 58 个单倍群、亚单倍群和大类(paragroup)(表 1)。大多数单倍群为土著世系[11-14]。正如图 3 中报道的一样,本研究检测出显著的欧亚西部(Western Eurasian)的线粒体 DNA 比例为 8.64%,而非洲撒哈拉以南的线粒体 DNA 比例为 2.08%[15-18]。然而只有 3 个东亚进化枝(0.40%)被识别出来[19]。我们比较了三大不同的世系类群之中单倍群多样性指数。其中最高的值( $Hd=0.9880 \pm 0.0008$ )是土著世系 505 个不同的单倍型。所有 4 个常见泛美洲单倍群(A2、B2、C1 和 D1)都高频出现于墨西哥人口中,其中最普遍的 A2 线粒体 DNA(占总体的 48%)可以被细分成 281 个单倍型( $Hd=0.9823 \pm 0.0018$ )和 5 个不同的亚单倍型 A2\*、A2e、A2h、A2i。本研究使用邻接法(Neighbor-Joining)推断出了属于泛美洲世系

的 1800 个调控区序列的进化关联(图 4)。排除了缺失和模糊数据以后,我们使用最大似然法(Maximum Likelihood)计算出进化距离,以碱基替换数目表示。我们发现所有的泛美洲单倍群 A2、B2、C1、D1 都有着大规模重合的序列分化时间,罕见的土著类型 D4h3( $4.1 \pm 2.3$ )也一样。本项发现表明这些墨西哥的单倍群是并行或者几乎同时到达和分散的[11,20]。而这个发现也与古印第安人(Paleo-Indian)并行散入美洲的理论一致。

有趣的是,这个时间上的相似性并没有在空间分布上面反映出来(图 5)。在这个国家中,美洲土著单倍型的亲缘地理学惊人地不同。其中 A2 在南部最频繁, C1 在西北部,而 B2 和 D1 在中央区域。不常见的 D4h3 的分布更具地理局限性,它的绝大部分分布在西北海岸。而另一方面,欧亚西部单倍群在北大西洋海岸呈现出峰顶频率。事实上,在墨西哥的东北部西方世系被观察到的频率明显高于土著世系(P 值 $<0.01$ )。为了评估不同州的线粒体 DNA 遗传结构,我们通过主成分分析(PCA, 图 6)比较了各单倍群的频率。主成分分析是一种以各单倍群为离散变量并允许把这样的数据集归入主成分(PCs)的方法[21]。在把单倍群归纳成主成分之后,我们观察到第一主成分主要是西南-东北走向,索诺拉州-下加州-南下加州(Sonora-Baja-Baja Sur)群体在一极,而恰帕斯州-塔巴斯科州-坎佩切州-尤卡坦州(Chiapas-Tabsco-Campeche-Yucatan)在另一极。如果我们观察变量对第一和第二主成分的贡献,我们会发现一个清晰的单倍群结构。具体来说,第二主成分把土著世系(A2、B2、C1、D1、D4h3)从西方进化枝(H、J、T、U、K)中分离出来,使新莱昂州-塔毛利帕斯州(Nuevo Leon-Tamaulipas)的东北类群分布在聚类外围。

### 三、结论

在本研究中,我们第一次提供了对墨西哥线粒体基因库的广泛概览。迄今为止,不同的土著人群都曾被使用不同的细胞核遗传标记来研究。结果都显示了他们是土著人、欧洲人和非洲人的混血[2,5,22]。然而,这些调查是对有限的样本进行的,而且没有覆盖

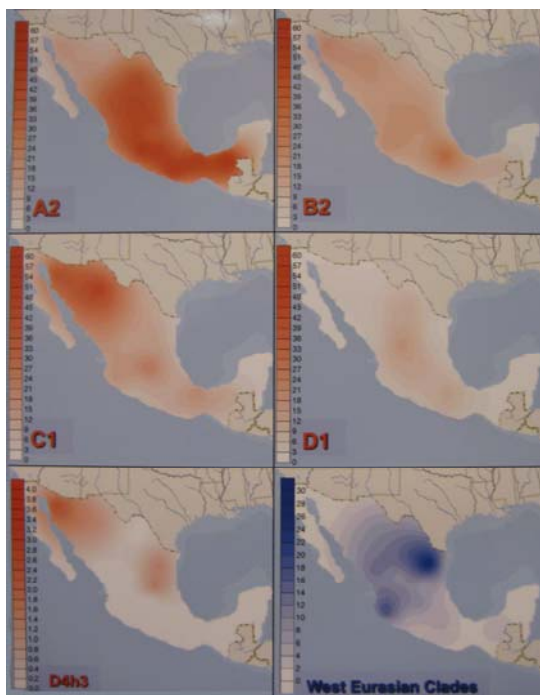


图 5 土著单倍群和欧亚西部世系的空间频率分布

Fig.5 Spatial frequency (%) distributions of Native haplogroups and Western Eurasian lineages

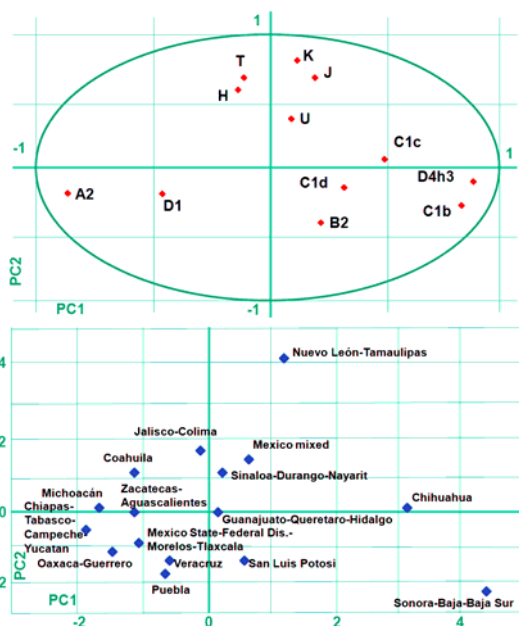


图 6 墨西哥各州的线粒体单倍群组成主成分分析

Fig.6 PCA of mtDNA haplogroup profiles in Mexico states

整个墨西哥的地理区域。

我们的发现证实了在现代墨西哥人口中存在着压倒性的美洲土著遗留。因为 90% 的线粒体 DNA 属于土著世系 A2、B2、C1、D1 和 D4h3。在被测序的 1180 bps 区段内，以 4.33 的平均碱基替换值估测，这些单倍群的分子分化时间非常相近，因此提示它们是同

时到达并扩散到这个国度的。

另一个方面，本研究鉴别出土著和欧洲世系的不均匀空间分布。欧洲单倍群并不常见而且局限于东北濒临大西洋的州。而不同的土著世系在这个国家中展示出奇异而独特的分布图景。总的说来，这些发现表明了欧洲人(主要是男性)的到来和接下来的 500 年中的人口迁移没有在本质上改变在前哥伦布时代的线粒体 DNA 变异在墨西哥的分布图谱。

**致谢** We are grateful to all the donors for providing biological specimen and to everyone at the Sorenson Molecular Genealogy Foundation for their assistance with the preliminary data. This received support from Progetti Ricerca Interesse Nazionale 2007 (Italian Ministry of the University) (to AT), the Sorenson Molecular Genealogy Foundation (to SRW), Compagnia San Paolo (to AT) and Fondazione Cariplo (to AT).

参考文献

1. Cotterell AF (1986) A dictionary of world mythology. Oxford, UK: Oxford University Press.
2. Lisker R, Ramirez E, Briceño RP, Granados J, Babinsky V (1990) Gene frequencies and admixture estimates in four Mexican urban centers. Hum Biol 62:791-801.
3. Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, Wallace DC (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613-623.
4. Torroni A, Chen YS, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, Wallace DC (1994) mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54:303-318.
5. Buentello-Malo L, Peñaloza-Espinosa RI, Loeza F, Salamanca-Gomez F, Cerda-Flores RM (2003) Genetic structure of seven Mexican indigenous populations based on five polymarker loci. Am J Hum Biol 15:23-28.
6. Buentello-Malo L, Peñaloza-Espinosa RI, Salamanca-Gómez F, Cerda-Flores RM (2008) Genetic admixture of eight Mexican indigenous populations: based on five polymarker, HLA-DQA1, ABO, and RH loci. Am J Hum Biol 20:647-650.
7. Cerda-Flores RM, Budowle B, Jin L, Barton SA, Deka R, Chakraborty R (2002) Maximum likelihood estimates of admixture in Northeastern Mexico using 13 short tandem repeat loci. Am J Hum Biol 14:429-439.
8. Cerda-Flores RM, Villalobos-Torres MC, Barrera-Saldaña HA, Cortés-Prieto LM, Barajas LO, Rivas F, Carracedo A, Zhong Y, Barton SA, Chakraborty R (2002) Genetic admixture in three Mexican Mestizo populations based on D1S80 and HLA-DQA1 loci. Am J Hum Biol 14:257-263.
9. Peñaloza-Espinosa RI, Arenas-Aranda D, Cerda-Flores RM, Buentello-Malo L, González-Valencia G, Torres J, Alvarez B, Mendoza I, Flores M, Sandoval L, Loeza F, Ramos I, Muñoz L, Salamanca F (2007) Characterization of mtDNA haplogroups in 14 Mexican indigenous populations. Hum Biol 79:313-320.
10. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the

- Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147.
11. Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR, Salas A, Torroni A, Bandelt HJ (2008) The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. *PLoS ONE* 3:e1764.
  12. Brown MD, Hosseini SH, Torroni A, Bandelt HJ, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC (1998) mtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? *Am J Hum Genet* 63:1852-1861.
  13. Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravi CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK, Fedorova SA, Golubenko MV, Stepanov VA, Gubina MA, Zhadanov SI, Ossipova LP, Damba L, Voevoda MI, Dipierri JE, Villems R, Malhi RS (2007) Beringian standstill and spread of Native American founders. *PLoS ONE* 2:e829.
  14. Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet* 53:563-590.
  15. Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144:1835-1850.
  16. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Gölge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Di Rienzo A, Novelletto A, Oppenheim A, Norby S, Al-Zaheri N, Santachiara-Benerecetti S, Scozzari R, Torroni A, Bandelt HJ (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* 67:1251-1276.
  17. Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F, Wang CY, Chaudhuri TK, Palla V, Zhang YP (2004) Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Am J Hum Genet* 75:966-978.
  18. Behar DM, VILLEMS R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D, Bertranpetit J, Quintana-Murci L, Tyler-Smith C, Wells RS, Rosset S; Genographic Consortium (2008) The dawn of human matrilineal diversity. *Am J Hum Genet* 82:1130-1140.
  19. Kong QP, Bandelt HJ, Sun C, Yao YG, Salas A, Achilli A, Wang CY, Zhong L, Zhu CL, Wu SF, Torroni A, Zhang YP (2006) Updating the East Asian mtDNA phylogeny: a prerequisite for the identification of pathogenic mutations. *Hum Mol Genet* 15:2076-2086.
  20. Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogo MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos AK, Santos SE, Petzl-Erler ML, Bonatto SL (2008) Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am J Hum Genet* 82:583-592.
  21. Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.
  22. Peñaloza R, García-carrancá A, Ceras T, Alvarez C, Berumen J, Zavala C, Salamanca F (1995) Frequency of haplotypes in the beta globin gene cluster in a selected sample of the Mexican population. *Am J Hum Biol* 7:45-49.