PRELIMINARY STUDY OF MOLECULAR VARIABILITY FOR NEOLITHIC PIG (*SUS SCROFA DOMESTICUS*) FROM ROMANIA USING THE CYTOCHROME B

LUCIAN GORGAN[1*], ANDREI STEFAN[1], LUMINITA BEJENARU[1], ROMEO CAVALERIU[1], SIMINA STANC[1]

Keywords: mitochondrial DNA, cytochrome b, *Sus scrofa domesticus*, Neolithic

Abstract: The aim of the present study is to reveal molecular differences between old and actual populations of pig (*Sus scrofa domesticus*), based on the analysis of hyper variable first part of cytochrome b. Neolithic pig remains (bone fragments) are sampled from Poduri-Dealul Ghindaru *Tell* (Bacau County). The *Tell* of Poduri-Dealul Ghindaru (positioned in the Eastern part of Romania) has a complex stratigraphy and the archaeologists have carried out analyses on Chalcolithic (Precucuteni and Cucuteni) and Bronze Age levels.

INTRODUCTION

The Neolithic represents one of the most important times of Prehistory, when mankind achieved spiritual and physical progress, whose legacies can be traced up to modern times. One of the most important changes in human behaviour, during the Neolithic, is the transition from the way of predator and consumer, to the way of producer of food resources and other goods, thus making the new age mean more than just a “food revolution”. During the transition from the Mesolithic to the Neolithic, an important element in the evolution of human communities was, besides the beginnings of plant cultivation, the ability of taming wild animals, which would later be domesticated and raised. This domestication is a fundamental acquisition of the Neolithic way of life, whose premises are found in the final stage of the Upper Palaeolithic when human communities achieved an advanced form of hunting. The beginnings and the ways of domestication are still debateable because the morphological differences between the domestic and wild animals are not clear but the step forward achieved by the process of domestication is obvious (Boghian, 2003). In recent times, based on DNA studies, a phylogenetic tree of modern domestic species can be constructed. Analyses of DNA extracted from fragmentary remains of skeletons and teeth provide an important insight into the ancestry of domesticates (Moritz *et al*., 1987; Nei & Kumar, 2000; Reitz & Wing, 2008; Haile *et al*., 2010).

Our most important domesticates originated from south-western Asia in the area known as the Fertile Crescent and then spread following a northwest or northeast direction. Because of the environmental conditions, a series of modifications occurred in the conformation, variability and behaviour of domestic animals (Reitz & Wing, 2008). In the domestic fauna of Europe, cattle (*Bos taurus*) and pig (*Sus scrofa domesticus*) were predominant and were also abundant in the wild form, some researchers even assuming the possibility of local domestication and re-domestication of aurochs (*Bos primigenius*), as in the case of the wetland, forest and steppe Neolithic, and wild boar (*Sus scrofa ferus*). The eventual domestications would have contributed to the increase of local livestock or to its regeneration following battles, natural disasters and plagues (Boghian, 2003). Pigs also suffered some domestication adaptations even if some archaeozoologists have shown that they could have been raised in a state of semi-domestication. It is important to show that the domestication of animals did not stop at the Neolithic stage and, after a period of the previously acquired human knowledge consolidation, this process continued.

Cucuteni Culture appeared, evolved and spread in eastern part of Romania during a millennium, 4600-3500 Cal B.C. (Mantu, 1998), and it is divided in three chronological phases: *A*, *AB*, and *B*. Cucuteni Culture developed in the Chalcolithic period, representing the end of the Neolithic and the transition to Bronze Age. The Cucuteni culture is one of the most representative for the Romanian Chalcolithic, known especially for its decorative and figurative art. The Cucuteni culture still is an important topic in prehistory research because of the complexity of the uncovered artefacts and the spatial and social organization elements. This is one of the reasons why a synthetic archaeozoological analysis was necessary (Cavaleriu & Bejenaru, 2009).

The *Tell* of Poduri-Dealul Ghindaru (Bacau County) has the following position: 46º47’’ North latitude and 26º53’’ East longitude, with an absolute altitude of 429 m. The *Tell* is situated on a fragment of the terrace of 30 m on the right bank of the Tazlau Sarat River and it currently has a surface of around 1.2 ha. The high complexity of the stratigraphy was emphasized along over 27 archaeological excavations campaigns that took place so far. There were reported levels belonging to the Precucuteni and Cucuteni Chalcolithic cultures and to Bronze Age.

Using the morphometric differentiation in *S. scrofa*, 16 different subspecies have been described in the past (Groves, 1981). Groves and Grubb (1993) classified the European wild boar into a group of ‘western races’, including the subspecies *S. s. scrofa* (Central-Western Europe), *S. s. meridionalis* (Sardinia and Corsica), *S. s. attila* (Eastern Europe)

and *S. s. lybicus* (Southern Balkans). Nonetheless, multivariate morphometric of skull measurements suggested the existence of a single subspecies (Genov, 1999) with a possible north-east to south-west dimensional cline due to environmental variation and its effect on food availability and growth rate (Randi *et al*., 1989). In addition, the wild boar is the ancestor of the domestic pig, with which it shares a close genetic affinity and can hybridize. Thus, genetic introgression from the domestic pig may represent a further source of differentiation throughout the natural range of the wild boar, and may represent a major threat to its genetic integrity (Scandura *et al*., 2011).

Many researchers have sampled European wild boars, for studies with various purposes, and so, most of the data are scattered. No molecular survey has yet been carried out to evaluate the overall genetic variation throughout Europe, on the basis of wide and comprehensive sampling. Our goal is to focus on the relationship between pigs from different historical periods, by evaluating the results obtained with first part of cytochrome b as molecular marker.

MATERIALS AND METHODS

Samples' description

For this study we analysed pig remains (bone fragments) sampled in 2007 from the Cucuteni *A* level within the *Tell* in Poduri-Dealul Ghindaru, dated between 4460-4050 Cal B.C. (Mantu, 1998; Monah *et al*., 2003). With a view to comparing, we also analysed bone fragments from Middle Ages sites and muscle tissue from modern specimens (Table 1). Anatomical and taxonomical primary identification was made according to archaeozoological methodology (Udrescu *et al.*, 1999).

Mitochondrial DNA isolation and purification

The bone fragments have been previously washed with a solution of sodium hypochlorite and exposed to UV rays for 12 hours. The DNA template purification was performed using the DNA IQ Kit (Promega) for forensic analysis and quantified by agarose gel electrophoresis and by spectrophotometry.

DNA amplification and Sequencing

According to complete the fragments containing the first part of cytochrome b (approximately 500 bp) two primers L14724 (5′-CGAAGCTTGATATGAAAACCATCGTTG-3′) (Pääbo, 1990) and H15149 (5′-AACTGCAGCCCTGAATATTTGTCCTCA-3′) (Kocher *et al*., 1989) were used for this study. The PCR protocol was carried out in a 25 µl volume using 1XGoTaq Green Master Mix (Promega), 2 mM $MgCl_2$, 0.2 mM each dNTP, 10 mM of each primer, GoTaqDNA polymerase 1.25U. The amplification process was performed in an initial denaturation step at 95°C for 3 minutes, followed by 35 cycles at 95°C for 1 minute, 45°C for 1 minute and 72°C for 1 minute. The amplicons were tested by 1% agarose gel electrophoresis, purified from the gel through Wizard SV Gel and PCR Clean-up System columns (Promega) commercial kit, following the producer's protocol. Purified PCR products were sequenced with the both PCR primer, due to the fragments length, following the Methods Development Kit (Beckman Coulter) protocol and the ethanol precipitation. Samples were sequenced with 8 capillaries CEQ 8000 Beckman Coulter analyser.

Table 1. *Sus scrofa domesticus* samples' type and origin.

Sample ID	Samples' age	Archaeological site	Tissues type
01S1	2009	-	muscle
01S2	2009	-	muscle
31S1	2009	-	muscle
31S2	2009	-	muscle
SR	Middle Ages (IV-VI century A.D.)	Slava Rusa (Tulcea County)	bone fragment
2667	Neolithic (4665-4050 B.C.)	Poduri-Dealul Ghindaru (Bacau County)	bone fragment

Sequence alignment and phylogenetic analysis

For phylogenetic inference, one alignment corresponding to the concatenation of all gene sequences was created. Sequences of each fragment were automatically aligned using the Clustal W method (Thompson *et al*., 1994) and MegAlign software (Lasergene v.7; DNASTAR Inc., Madison, WI).

Phylogenetic analysis including neighbour-joining (NJ), maximum-parsimony (MP) and maximum-likelihood (ML) trees of DNA sequence alignments analysis were conducted using Paup 4.0b10 (Swofford, 2003) using PaupUp graphical interface (Calendini & Martin, 2005). Genetic distances used in NJ trees are Kimura two-parameter model distances with a transition: transversion ratio of 2:1. Non-synonymous/synonymous substitutions distances (dN/dS) ratios, as described by Nei & Gojobori (1986), were calculated using MEGA4 software (Tamura *et al*., 2007). Bootstrap analysis was made with 1000 replicates except in ML where only 100 replicates were generated. Hierarchical likelihood ratio tests were conducted using a batch file supplied with MODEL TEST 3.7 (Posada & Crandall, 1998) to provide the evolutionary models used in ML and Bayesian analysis. In total, the data set were analysed under three criteria, described below.

In Maximum parsimony (MP) analysis characters were unordered and equally weighted and gaps were treated as missing data (“?”). Analyses started with a stepwise addition tree with taxa randomly added in 5000 replicates. During searches branches were collapsed when minimum branch length was zero. Trees of MP analyses resulting in more than one most parsimonious tree were summarized in strict consensus trees. Branch support values for combined data sets were estimated with nonparametric bootstrap values (BP; Felsenstein, 1985) (5000 replicates).

For Maximum likelihood (ML) we used the Akaike Information Criterion (AIC) as implemented in Model Test (Posada & Crandall, 1998) to choose substitution models for the entire partition. ML analysis for the data set was run using HKY (Hasegawa *et al*. 1985) model.

RESULTS AND DISCUSSIONS

The first part of mitochondrial cytochrome b (cyt b) was amplified by PCR and sequenced in both directions in all tested individuals. The amplicons were tested by agarose gel electrophoresis (Figure 1). The electrophoregram analysis reveals fragments of approximately 500 base pairs (bp) for all analysed individuals. The amplicons resulted from the PCR reaction after purification process, presents a sequencing optimal amount of DNA only for 6 individuals: 01S1, 01S2, 31S1, 31S2, 2667 and SR. All others individuals have a very low amount of total DNA due to fragmentation process present in old bones fragments, unusable for sequencing.

The 300bp sequences were analysed and aligned for haplotypes identification. Based on the alignment process 6 different haplotypes were analysed and characterized. The haplotypes were compared and the similarity and divergence matrix was achieved (Table 2). Comparing the similarity percentage for the pig individuals they vary between 99.7% for present specimens, 69.9% comparing present specimens with the Middle Ages individuals and 67.2% for the Neolithic pig compared with present ones.

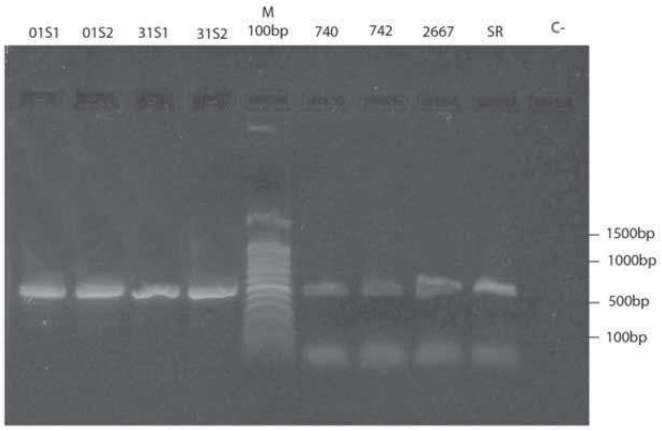

Figure 1. The first part of cytochrome b amplicons.

Table 2. The similarity and divergence percentage for the analysed haplotypes.

	Similarity Percentage								
Divergence Percentage		01S1	01S2	31S1	31S2	HM1047	2667	SR	
	01S1		100.0	99.7	99.7	100.0	76.7	69.9	01S1
	01S2	0.0		99.7	99.7	100.0	76.7	69.9	01S2
	31S1	0.3	0.3		99.3	99.7	77.0	69.6	31S1
	31S2	0.3	0.3	0.7		99.7	76.3	69.6	31S2
	HM1047	0.0	0.0	0.3	0.3		76.7	69.9	HM1047
	2667	28.1	28.1	27.6	28.7	28.1		67.2	2667
	SR	39.1	39.1	39.7	39.7	39.1	43.9		SR
		01S1	01S2	31S1	31S2	HM1047	2667	SR	

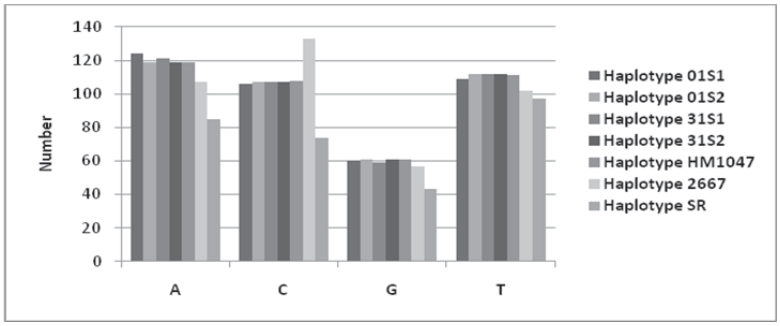

Figure 2. The nucleotides frequencies for the analysed haplotypes.

The statistical analysis of all identified haplotypes based on Clustal W alignment report was performed for the conserved, variable and parsim-informative sites (Table 3).

Table 3. Sequences statistics.

	Conserved sites	Variable sites	Parsim-info sites	Singleton sites	0-fold degenerate sites	2-fold degenerate sites	4-fold degenerate sites
Number	37	262	55	207	180	10	14
Frequency	12.3	87.6	18.3	69.2	60.2	3.3	4.6
Total	299						

Table 4. Estimates of codon-based evolutionary divergence between sequences.

	01S1	01S2	31S1	31S2	HM1047	2667	SR
01S1							
01S2	0.0						
31S1	0.0	0.0					
31S2	0.0	0.0	0.0				
HM1047	0.0	0.0	0.0	0.0			
2667	45.6	45.6	45.6	45.6	45.6		
SR	38.5	38.5	38.5	38.5	38.5	46.0	

The number of synonymous differences per sequence from analysis between sequences is shown (Table 4). All results are based on the pair wise analysis of 7 sequences. Analyses were conducted using the Nei-Gojobori method in MEGA4. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 84 positions in the final dataset.

Also, from the haplotypes comparison, the nucleotides number and frequencies have been analysed (Figure 2). The adenine is the most frequent in all present haplotypes, followed by thymine and cytosine. For the Neolithic specimens, the cytosine has highest frequency followed by adenine and thymine. The similarity dendrogram (Figure 3) was build using the MegAlign module of Lasergene software and conclude about the origin of present pig (*Sus scrofa domesticus*), in Neolithic forms of the same genera.

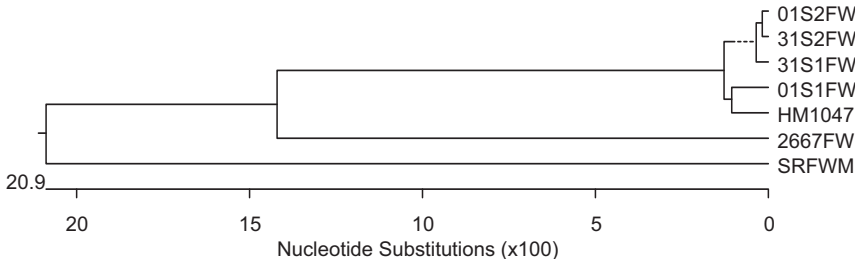

Figure 3. Phylogenetic tree of the analysed haplotypes

The evolutionary history was inferred using the Maximum Parsimony method (Figure 4) (Eck *et al*., 1966). Tree #1 out of 25 most parsimonious trees (length = 346) is shown. The consistency index is (0.982456), the retention index is (0.982143), and the composite index is 0.979304 (0.964912) for all sites and parsimony-informative sites (in parentheses). The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (5000 replicates) are shown next to the branches (Felsestein, 1985). The MP tree was obtained using the Close-Neighbor-Interchange algorithm (Nei & Kumar, 2000) with search level 3 in which the initial trees were obtained with the random addition of sequences (10 replicates). The codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). There were a total of 299 positions in the final dataset, out of which 55 were parsimony informative. Phylogenetic analyses were conducted in MEGA4 (Tamura *et al*., 2007).

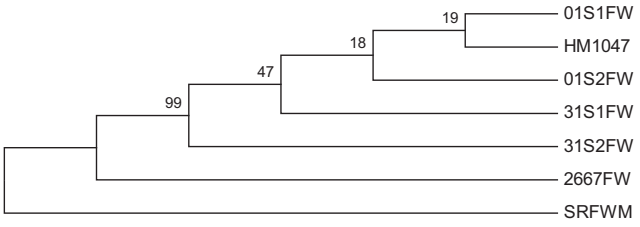

Figure 4. Evolutionary relationships of 7 taxa.

CONCLUSIONS

The results confirm the potential and the applicability of molecular techniques for inferring the phylogenies based on mitochondrial DNA sequences even from ancient bone fragments and the ability to identify the major lineages and the evaluation of homology degree between different species and subspecies of the same genera.

The sequence alignment and the phylogeny show high differentiation between individuals of *Sus scrofa* species from different historical periods of time.

REFERENCES

Boghian, D. (2003): *Începuturile istoriei omenirii*, Editura Suceava.
Calendini, F., Martin J.-F. (2005): *Paup UP v1.0.3.1 A free graphical frontend for Paup* Dos software.*
Cavaleriu, R., Bejenaru, L. (2009): *Cercetări arheozoologice privind Cultura Cucuteni, faza A*, Editura Universităţii Al. I. Cuza Iaşi.
Eck, R.V., Dayhoff, M. O. (1966): *Atlas of Protein Sequence and Structure.* National Biomedical Research Foundation, Silver Springs, Maryland.
Felsenstein, J. (1985): *Confidence limits on phylogenies: An approach using the bootstrap*. Evolution, 39, 783-791.
Genov P. V. (1999): *A review of the cranial characteristics of the wild boar (Sus scrofa Linnaeus 1758), with systematic conclusions*. Mammal Review 29: 205–238.
Groves C. P., Grubb P. (1993): *The Eurasian suids: Sus and Babyrousa*. In: Oliver WLR (ed.) Pigs, Peccariesand Hippos – Status Survey and Conservation Action Plan, 107–111. IUCN/SCC, Gland, Switzerland.
Groves C. (1981): *Ancestors for the Pigs: Taxonomy and Phylogeny of the Genus Sus*. Research Schoolof Pacific Studies, Australian National University, Canberra, Australia.
Haile, J., Larson, G., Owens, K., Dobney, K., Shapiro, B. (2010): *Ancient DNA typing of archaeological pig remains corroborates historical records*. Journal of Archaeological Science, 37, 174-177.
Hasegawa, M, Kishino, K, Yano, T. (1985): *Dating the human-ape splitting by a molecular clock of mitochondrial DNA*. J. Mol. Evol., 22,160-174.
Kocher, T.D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S. (1989): *Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers.* Proc. Nati., Acad. Sci. USA, 86, 6196-6200.
Mantu, C.M. (1998): *CulturaCucuteni: evolutie, cronologie, legaturi*, Bibliotheca MemoriaeAntiquitatis, Piatra-Neamt.
Monah, D., Dumitroaia, Gh.,Monah, F., Preoteasa, C., Munteanu, R., Nicola, D. (2003): *Poduri-DealulGhindaru. O TroieînSubcarpatiiMoldovei*, Bibliotheca Memoriae Antiquitatis, Piatra-Neamt.
Moritz, C., Dowling, T., Brown, W. (1987): *Evolution of animal mithocondrial DNA: relevance for population biology and systematics*. Annu. Rev. Ecol. Syst., 18, 269-292.
Nei, M., Gojobori, T. (1986): *Simple methods for estimating the numbers of synonymous and non synonymous nucleotide substitutions.* Molecular Biology and Evolution, 3, 418-426.
Nei, M., Kumar, S. (2000): *Molecular evolution and phylogenetics*, Oxford University Press, p.128.
Pääbo, S. (1990): *Amplifying ancient DNA*. In PCR protocols: A guide to methods and applications, Eds. M. A. Tunes, D.H. Gelfand, J. J. Sninskyşi T. J. White, Academic Press San Diego, 159 – 166.
Posada, D., Crandall, K. A. (1998): *Model test: testing the model of DNA substitution.* Bioinformatics, 14 (9), 817-818.
Randi, E., Apollonio, M., Toso, S. (1989): *The systematics of some Italian populations of wild boar (Sus scrofa L.): a craniometric and electrophoretic analysis*. Zeitschriftfür Säugetierkunde 54, 40–56.
Reitz, E.J., Wing, E.S. (2008): *Zooarchaeology*, Cambridge University Press, Second edition.
Scandura, M., Iacolina, L., Apollonio, M. (2011): *Genetic diversity in the European wild boar Sus scrofa: phylogeography, population structure and wild x domestic hybridization*. Mammal Review 10, 2-5.
Swofford, D. L. (2003): *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods*). Version 4. Sinauer Associates, Sunderland, Massachusetts.
Tamura, K., Dudley, J., Nei, M., Kumar, S. (2007): *MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.* Molecular Biology and Evolution, 24, 1596-1599.
Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994): *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice*. Nucleic Acids Res. 22, 4673-4680.
Udrescu, M., Bejenaru, L., Hriscu, C. (1999): *Introducere în arheozoologie*, Editura Corson, Iaşi.

1 Alexandru Ioan Cuza University, Iasi, Carol I 20A,
* lucian.gorgan@uaic.ro

Acknowledgments
This study was made as part of *ARHEOINVEST* Platform, and was supported by the research program CNCSIS, PN II Idei_2116/2008.