



Molecular data in suid systematics and phylogeny: Utility and limitations

¹Claudia Balint, ²Tudor Păpuc, ²Cristian Coroian, ¹Daniela Bordea, ³Camelia Oroian, ³Florin D. Bora, ⁴Maria Popescu, ^{1,5,6}I. Valentin Petrescu-Mag

¹Department of Environmental Engineering and Protection, Faculty of Agriculture, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Romania; ² Faculty of Animal Science and Biotechnologies, University of Agricultural Science and Veterinary Medicine Cluj-Napoca, Cluj-Napoca, Romania; ³ Faculty of Horticulture and Business for Rural Development, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Romania; ⁴ Equine Clinic, Faculty of Veterinary Medicine, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Cluj-Napoca, Romania; ⁵ Bioflux SRL, Cluj-Napoca, Romania; ⁶ University of Oradea, Oradea, Romania.
Corresponding author: I. V. Petrescu-Mag, zoobiomag2004@yahoo.com

Abstract. Molecular data have significantly advanced the systematics and phylogenetic understanding of the Suidae family, which includes wild and domestic pigs. This paper reviews the strengths and limitations of molecular approaches, particularly mitochondrial and nuclear DNA sequences, in resolving evolutionary relationships within Suidae. Molecular analyses have clarified both deep divergences and recent splits, revealed cryptic lineages and hybridization events, and informed domestication and phylogeographic patterns. However, challenges such as mitochondrial bias, incomplete lineage sorting, hybridization, and nuclear-mitochondrial discordance complicate phylogenetic interpretations. Integrative methodologies, combining genomic, morphological, ecological, and ancient DNA data, are increasingly necessary to address taxonomic ambiguities and improve the resolution of suid evolutionary history. While molecular tools are powerful, they must be used critically and in concert with other evidence to produce robust phylogenies and informed taxonomic decisions.

Key Words: evolutionary history, domestication, hybridization, mitochondrial DNA, molecular markers, nuclear DNA, phylogenomics, phylogeny, taxonomy, suidae, systematics.

Introduction. The family Suidae (true pigs) includes several genera of even-toed ungulates (Artiodactyla) found primarily in Africa and Eurasia. The currently recognized extant genera are: *Babyrousa* – babirusas (Proorocu & Petrescu-Mag 2022a), *Hylochoerus* – giant forest hog (Gavriloaie & Petrescu-Mag 2023), *Phacochoerus* – warthogs (Gavriloaie & Burduhos 2023), *Potamochoerus* – bush pigs and red river hogs (Oroian & Proorocu 2023), *Porcula* – pygmy hog (Proorocu & Petrescu-Mag 2022b), *Sus* – domestic pigs and wild boars (Petrescu-Mag & Oroian 2018). These genera include both widely distributed and regionally restricted species. Taxonomy within Suidae remains under revision, especially with regard to subspecies and cryptic diversity (Iannucci 2022). Advances in molecular biology have provided powerful tools for investigating the evolutionary history of suids (Șuteu 2011; Xie et al 2022). Nucleotide and amino acid sequences from mitochondrial and nuclear genomes have been widely used to resolve phylogenetic relationships within this family (Xie et al 2022; Bolner et al 2024). However, as in other taxa, molecular data have both strengths and limitations. This mini-review critically evaluates how far we can rely on such sequences in suid systematics and phylogeny and where caution is warranted.

Strengths of Molecular Data in Suid Phylogeny

Clarifying deep and recent relationships. Molecular data, especially mitochondrial DNA (mtDNA) and nuclear genes, have been instrumental in clarifying both deep evolutionary splits and recent divergences within Suidae.

Mitochondrial genes (e.g., *cytochrome b*, *D-loop*, *COI*) (Niu et al 2023; Kušec & Gvozdanić 2023) have been used to resolve relationships within the *Sus* genus and among other genera such as *Phacochoerus*, *Potamochoerus*, and *Babyrousa* (Benjamin et al 2023).

Nuclear markers (e.g., introns, exons from genes like *IRBP* or *RAG1*) provide complementary information that is less prone to maternal lineage bias. Nuclear markers, such as introns and exons from genes like *IRBP* and *RAG1*, are valuable in suid phylogeny because they provide phylogenetic signals that are independent of the maternal inheritance patterns seen in mitochondrial DNA, thus reducing lineage bias and offering a more balanced view of evolutionary relationships (Matthee et al 2001). Analyses using multiple nuclear DNA segments have demonstrated greater robustness and less homoplasy compared to mitochondrial markers, leading to more reliable reconstructions of both deep and recent divergences within Suidae and related artiodactyl groups (Matthee et al 2001). For example, studies analyzing thousands of nucleotide positions from several nuclear genes have produced phylogenies that are congruent across loci and less sensitive to taxon sampling, supporting key evolutionary splits and clarifying the relationships among suid lineages, including the diphyletic nature of Suidae and the basal position of Suiformes within Cetartiodactyla (Matthee et al 2001).

Identifying cryptic lineages and hybridization. Molecular studies have uncovered significant hidden genetic diversity within *Sus scrofa*, revealing that populations in Asia are highly structured and may represent distinct evolutionary lineages, rather than a single species. Whole-genome analyses have detected ancient hybridization events between *S. scrofa* and other wild suid species, such as the pygmy hog (*Porcula salvania*), indicating that interbreeding occurred during the wild boar's expansion across Eurasia and contributed to its rapid spread and genetic diversity (Liu et al 2019, 2020). Additionally, evidence of introgression between domestic pigs and European wild boars has been documented, complicating both the taxonomy and conservation of these populations by introducing domestic alleles into wild gene pools (Scandura et al 2011). These findings highlight the importance of advanced molecular tools in identifying cryptic lineages and understanding the complex evolutionary history and ongoing hybridization within the *S. scrofa* complex (Liu et al 2019, 2020; Scandura et al 2011).

Phylogeography and domestication studies. Mitochondrial haplotype studies have traced the origins and spread of domesticated pigs, revealing that independent domestication events occurred in Europe, East Asia, and Southeast Asia, rather than from a single center of origin (Larson et al 2005; Wu et al 2007; Zhang et al 2021; Peng et al 2022). Ancient DNA (aDNA) analysis has been instrumental in reconstructing prehistoric suid movements and gene flow, showing that domestic pigs in Europe and Asia have distinct maternal lineages and that gene flow between wild and domestic populations was common during and after domestication (Kijas & Andersson 2001; Groenen et al 2012; Frantz et al 2015; Zhang et al 2021). In East Asia, fine-scale mitochondrial analyses indicate at least two separate domestication events, particularly in the Mekong region and the middle and downstream Yangtze River, with local wild boars contributing to the genetic makeup of domestic pigs (Wu et al 2007; Zhang et al 2021; Peng et al 2022). In Europe, modern domestic pigs primarily derive from local wild boar rather than Near Eastern ancestors, with ancient DNA revealing a near-complete genomic replacement of early domestic lineages by European wild boar (Larson et al 2005; Zhang et al 2021). Genome-wide studies further demonstrate that, despite ongoing gene flow, strong selection for domestic traits created distinct genetic signatures in domestic pigs, counteracting the homogenizing effects of interbreeding with wild populations (Rubin et al 2012; Frantz et al 2015). These findings highlight the complexity of pig domestication,

involving multiple centers, recurrent gene flow, and diverse management practices across regions and time periods (Larson et al 2005; Frantz et al 2015; Price & Hongo 2020).

Key Challenges and Limitations. Despite their usefulness, molecular sequences can lead to misleading interpretations if not carefully analyzed in the suid context.

Mitochondrial DNA bias and maternal lineage. Mitochondrial DNA (mtDNA) represents only the maternal lineage and can obscure species relationships when there is hybridization and introgression, as seen between domestic pigs and wild boars or among insular suid species, and when mtDNA fails to reflect nuclear genome structure due to sex-biased dispersal, which is common in suids where males tend to disperse and females are philopatric (Randi et al 1996; Guisheng et al 2006; Frantz et al 2016). Reliance on mtDNA alone can result in phylogenies that do not accurately represent the evolutionary history of the group, especially since mtDNA is sensitive to mutational saturation and may not capture the complexity of nuclear gene flow (Randi et al 1996; Matthee et al 2001). Nuclear DNA, by contrast, provides a more comprehensive view but is often underutilized due to technical and analytical challenges (Matthee et al 2001). Additionally, high levels of homoplasmy in mtDNA and its sensitivity to taxon sampling can further complicate phylogenetic reconstructions (Matthee et al 2001). Therefore, integrating both mitochondrial and nuclear data is essential for robust phylogenetic and taxonomic conclusions in suids (Matthee et al 2001; Frantz et al 2016).

Incomplete Lineage Sorting (ILS). ILS is a significant challenge, especially in regions where *Sus* species have radiated rapidly, such as the islands of Southeast Asia, because ancestral genetic polymorphisms may persist across species boundaries, leading to discordant gene trees (Guisheng et al 2006; Frantz et al 2016). For example, the relationships among *Sus verrucosus*, *Sus celebensis*, and *Sus barbatus* are still debated due to gene tree incongruence caused by ILS and hybridization (Guisheng et al 2006; Frantz et al 2016). This discordance complicates efforts to resolve the evolutionary history and taxonomy of these species, as different genes may suggest conflicting relationships depending on which ancestral alleles have been retained or lost in each lineage (Frantz et al 2016). Hybridization further exacerbates this issue by introducing genetic material across species boundaries, making it difficult to distinguish between shared ancestry and recent gene flow (Frantz et al 2016). As a result, phylogenetic reconstructions based solely on a limited set of molecular markers may not accurately reflect the true evolutionary relationships within the genus *Sus* (Guisheng et al 2006; Frantz et al 2016).

Hybridization and admixture. Hybridization and admixture are notable challenges, as suid species are known to hybridize both naturally and under human influence, resulting in reticulate evolution, rather than strictly bifurcating phylogenies (Frantz et al 2016). Hybridization between *S. scrofa* and *S. celebensis* or *S. barbatus* has been reported in captivity and possibly in the wild, complicating the reconstruction of clear species boundaries (Frantz et al 2016). Furthermore, genomic segments from *S. verrucosus* have been identified in domestic pig genomes from Southeast Asia, providing evidence for past admixture events (Frantz et al 2016). These processes blur phylogenetic signals and can lead to conflicting interpretations of evolutionary relationships, especially when only a limited set of molecular markers is used (Guisheng et al 2006; Frantz et al 2016). As a result, hybridization and admixture must be carefully considered in molecular studies to avoid misrepresenting the evolutionary history and taxonomy of suids (Frantz et al 2016).

Nuclear-mitochondrial discordance. Nuclear-mitochondrial discordance is a common issue, as phylogenetic trees built from mtDNA and nuclear genes often show conflicting topologies due to their differing inheritance patterns and evolutionary histories (Matthee et al 2001; Frantz et al 2016). Nuclear markers generally evolve more slowly and may not resolve recent splits, making them less informative for distinguishing closely related

or recently diverged suid species (Matthee et al 2001). In contrast, mitochondrial genes, while more variable and useful for detecting recent divergences, may mislead due to mutational saturation or the effects of selective sweeps, which can obscure true evolutionary relationships (Randi et al 1996; Matthee et al 2001). This discordance complicates the reconstruction of accurate phylogenies and highlights the importance of integrating both nuclear and mitochondrial data for robust taxonomic and evolutionary conclusions in suids (Matthee et al 2001; Frantz et al 2016).

Taxonomic ambiguity in the face of morphological plasticity. Taxonomic ambiguity often arises in the face of morphological plasticity, as genetic divergence does not always align with morphological or ecological data (Frantz et al 2016). For instance, some island suid populations exhibit significant genetic divergence, but limited morphological differentiation, raising questions about their species status (Guisheng et al 2006; Frantz et al 2016). Conversely, strong phenotypic variation in wild boars across Eurasia is not always matched by deep genetic divergence, which complicates subspecies designations and challenges traditional taxonomy based solely on morphology (Randi et al 1996; Frantz et al 2016). This discordance between genetic and morphological data highlights the limitations of relying on a single line of evidence and underscores the need for integrative approaches in suid phylogeny and taxonomy (Frantz et al 2016).

Current Best Practices and Future Directions. To address the complexities of suid evolution, researchers now employ integrative approaches such as whole-genome sequencing (WGS) provides dense data across the genome, allowing detection of admixture, gene flow, and introgression, or multi-locus nuclear phylogenies, using concatenated or coalescent methods, help resolve discordant gene trees. Ancient DNA (e.g., from archaeological pig remains) adds historical context to phylogeographic and domestication patterns. Morphological, ecological, and behavioral data are increasingly integrated with molecular evidence to refine taxonomic decisions.

Conclusions. Nucleotide and protein sequences have profoundly advanced our understanding of suid systematics and phylogeny. They are essential tools for resolving relationships, detecting cryptic diversity, and unraveling domestication histories. However, their utility is constrained by biological realities such as incomplete lineage sorting, hybridization, and mitochondrial-nuclear discordance. In Suidae, where complex evolutionary histories involve gene flow across species boundaries and island radiations with parallel evolution, molecular data must be interpreted with caution and in conjunction with other lines of evidence. Only through integrative taxonomy and phylogenomics can we achieve a truly accurate and nuanced understanding of suid diversity and evolution.

Conflict of Interest. The authors declare that there is no conflict of interest.

References

- Benjamin N. R., Crooijmans R. P., Jordan L. R., Bolt C. R., Schook L. B., Schachtschneider K. M., et al, 2023 Swine global genomic resources: insights into wild and domesticated populations. *Mammalian Genome* 34(4):520-530.
- Bolner M., Bovo S., Ballan M., Schiavo G., Taurisano V., Ribani A., et al, 2024 A comprehensive atlas of nuclear sequences of mitochondrial origin (NUMT) inserted into the pig genome. *Genetics Selection Evolution* 56(1):64.
- Frantz L., Meijaard E., Gongora J., Haile J., Groenen M., Larson G., 2016 The evolution of Suidae. *Annual Review of Animal Biosciences* 4:61-85.
- Frantz L., Schraiber J., Madsen O., Megens H., Cagan A., Bosse M., et al, 2015 Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nature Genetics* 47:1141-1148.
- Gavriloaie C., Burduhos P., 2023 A short description of *Phacochoerus aethiopicus* (Pallas, 1766). *Porc Res* 13(1):27-30.

- Gavriloaie C., Petrescu-Mag I. V., 2023 Portrait of *Hylochoerus meinertzhageni* Thomas, 1904, the giant forest hog. *Porc Res* 13(1):22-26.
- Groenen M., Archibald A., Uenishi H., Tuggle C., Takeuchi Y., Rothschild M. F., et al, 2012 Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* 491:393-398.
- Guisheng W., Jun-Feng P., Ya-Ping Z., 2006 Molecular phylogeny and phylogeography of Suidae. *Zoological Research* 27:197-201.
- Iannucci A., 2022 On the authorship, availability, and improper use of *Sus scrofa ferus* for referring to wild pigs. *Taxonomy* 2(1):91-98.
- Kijas J., Andersson L., 2001 A phylogenetic study of the origin of the domestic pig estimated from the near-complete mtDNA genome. *Journal of Molecular Evolution* 52:302-308.
- Kušec I. D., Gvozdanović K., 2023 Tracing the domestic pig using the omics technologies. In: *Tracing the Domestic Pig*. IntechOpen.
- Larson G., Dobney K., Albarella U., Fang M., Matisoo-Smith E., Robins J., et al, 2005 Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* 307:1618-1621.
- Liu L., Bosse M., Megens H., Frantz L., Lee Y., Irving-Pease E., et al, 2019 Genomic analysis on pygmy hog reveals extensive interbreeding during wild boar expansion. *Nature Communications* 10(1):1992.
- Liu L., Bosse M., Megens H., Frantz L., Lee Y., Irving-Pease E., et al, 2020 Addendum: Genomic analysis on pygmy hog reveals extensive interbreeding during wild boar expansion. *Nature Communications* 11(1):6306.
- Matthee C., Burzlaff J., Taylor J., Davis S., 2001 Mining the mammalian genome for artiodactyl systematics. *Systematic Biology* 50(3):367-390.
- Niu L., Xie J., Shi K., Zhong T., 2023 Genetic diversity of mitochondrial DNA D-loop in wild and domestic pigs (*Sus scrofa*) in East Asia. *Folia Biologica (Kraków)* 71(1):37-44.
- Oroian I., Proorocu M., 2023 Overview of the bushpig, *Potamochoerus larvatus* (Cuvier, 1822). *Porc Res* 13(1):31-36.
- Peng Y., Cai X., Wang Y., Liu Z., Zhao Y., 2022 Genome-wide analysis suggests multiple domestication events of Chinese local pigs. *Animal Genetics* 53(3):293-306.
- Petrescu-Mag I. V., Oroian C. F., 2018 Stocli breed and its taxonomic status: *Sus domesticus* or *Sus scrofa*? *Porc Res* 8(1):29-32.
- Price M., Hongo H., 2020 The archaeology of pig domestication in Eurasia. *Journal of Archaeological Research* 28:557-615.
- Proorocu M., Petrescu-Mag I. V., 2022a On the genus *Babyrousa*: biology, ecology, conservation. *Porc Res* 12(1):12-19.
- Proorocu M., Petrescu-Mag I. V., 2022b The pygmy hog, *Porcula salvania* Hodgson, 1847. *Porc Res* 12(1):20-24.
- Randi E., Lucchini V., Diong C., 1996 Evolutionary genetics of the suiformes as reconstructed using mtDNA sequencing. *Journal of Mammalian Evolution* 3:163-194.
- Rubin C., Megens H., Barrio Á., Maqbool K., Sayyab S., Schwochow D., et al, 2012 Inaugural article by a recently elected academy member: Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the United States of America* 109:19529.
- 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* 41:125-137.
- Suteu M., 2011 Porcine milk protein polymorphisms. *Porcine Research* 1(1):1-112.
- Wu G., Yao Y., Qu K., Ding Z., Li H., Palanichamy M., Duan Z., Li N., Chen Y., Zhang Y., 2007 Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. *Genome Biology* 8:R245-R245.

- Xie H. B., Yan C., Adeola A. C., Wang K., Huang C. P., Xu M. M., 2022 African suid genomes provide insights into the local adaptation to diverse African environments. *Molecular Biology and Evolution* 39(12):msac256.
- Zhang M., Liu Y., Li Z., Lü P., Gardner J., Ye M., et al, 2021 Ancient DNA reveals the maternal genetic history of East Asian domestic pigs. *Journal of Genetics and Genomics* 49(6):537-546.

Received: 03 November 2024. Accepted: 29 November 2024. Published online: 30 December 2024.

Authors:

Claudia Balint, Department of Environmental Engineering and Protection, Faculty of Agriculture, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Mănăştur Street, 400372 Cluj-Napoca, Romania, e-mail: claudia.balint@usamvcluj.ro

Tudor Păpuc, Faculty of Animal Science and Biotechnologies, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Mănăştur Street, 400372 Cluj-Napoca, Cluj, Romania, e-mail: tudor.papuc@usamvcluj.ro

Cristian Coroian, Faculty of Animal Science and Biotechnologies, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Mănăştur Street, 400372 Cluj-Napoca, Cluj, Romania, e-mail: cristian.coroian@usamvcluj.ro

Daniela Bordea, Department of Environmental Engineering and Protection, Faculty of Agriculture, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Mănăştur Street, 400372 Cluj-Napoca, Romania, e-mail: daniela.bordea@usamvcluj.ro

Camelia Oroian, Faculty of Horticulture and Business in Rural Development, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Manastur, 400372 Cluj-Napoca, Romania, e-mail: camelia.oroian@usamvcluj.ro

Florin Dumitru Bora, Viticulture and Oenology Department, Advanced Horticultural Research Institute of Transylvania, Faculty of Horticulture and Business in Rural Development, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Mănăştur Street, 400372 Cluj-Napoca, Romania, e-mail: florin-dumitru.bora@usamvcluj.ro

Maria Popescu, Equine Clinic, Faculty of Veterinary Medicine, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Mănăştur Street, 400372 Cluj-Napoca, Romania; maria.popescu@usamvcluj.ro

Ioan Valentin Petrescu-Mag, Department of Environmental Engineering and Protection, Faculty of Agriculture, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Calea Mănăştur Street, 400372 Cluj-Napoca, Romania, e-mail: zoobiomag2004@yahoo.com This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

How to cite this article:

Balint C., Păpuc T., Coroian C., Bordea D., Oroian C., Bora F. D., Popescu M., Petrescu-Mag I. V., 2024 Molecular data in suid systematics and phylogeny: Utility and limitations. *Porc Res* 14(1):56-61.