Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 45 (2007) 427–436
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

The pygmy hog is a unique genus: 19th century taxonomists got it right first time round

Stephan M. Funk [a], Sunil Kumar Verma [b], Greger Larson [c,1], Kasturi Prasad [b], Lalji Singh [b], Goutam Narayan [a], John E. Fa [a,*]

[a] *Durrell Wildlife Conservation Trust, Les Augrès Manor, Jersey JE3 5BP, Channel Islands, UK*
[b] *Centre for Cellular and Molecular Biology, Hyderabad 500 007, India*
[c] *Henry Wellcome Ancient Biomolecules Centre, University of Oxford, South Parks Road, OX1 3PS, UK*

Received 3 August 2006; revised 8 August 2007; accepted 8 August 2007
Available online 26 August 2007

Abstract

The pygmy hog, *Sus salvanius*, the smallest and rarest extant suid was first described as the only member of the genus *Porcula*. It is currently regarded as member of the genus *Sus* and a sister taxon of the domestic pig/Eurasian wild boar (*Sus scrofa*). Phylogenetic analyses of 2316 bp from three mtDNA loci (control-region, cytochrome *b*, 16S) by Bayesian inference and statistical testing of alternative phylogenetic hypotheses all support the original classification of the pygmy hog as a unique genus. Thus, we propose that the species name *Porcula salvania* should be resurrected. The reclassification will heighten awareness of the need for the future protection and survival of this unique species.

Keywords: Pygmy hog; *Porcula salvania*; *Sus*, alternative phylogenetic hypotheses; Cytochrome *b*; Control region; 16S

1. Introduction

The pygmy hog, *Sus salvanius*, is now probably restricted to a single population in India and is one of the most endangered mammals in the world (Oliver and Deb Roy, 1993). First described as the only member of the genus *Porcula* (Hodgson, 1847), it is currently regarded as the closest relative of the Eurasian pig *Sus scrofa* (Groves, 1981; Oliver, 1993). Hodgson (1847) and Pilgrim (1926) justified the classification of the pygmy hog as a separate genus compared to *Sus*, based on morphological differences, especially skull and dental characteristics. These differences were later interpreted as 'superficial' (Groves, 1981) and a consequence of body size miniaturization (Groves, 1981; Oliver, 1993), leading to the general, though not unanimous (Ghosh, 1988), acceptance of the pygmy hog as a *Sus* (Corbet and Hill, 1992; Groves and Grubb, 1993b; Nowak, 1999). Members of the genus *Sus* have been separated into two species groups by the presence of conspicuous facial adornments in adult males, referred to as 'warts': (1) non-warty pigs—Eurasian wild boar *S. scrofa* and *S. salvanius* and (2) warty pigs—all other *Sus* (Groves and Grubb, 1993b), or into three distinct clusters using the ratio between the width of the inferior surface and the posterior surface of the lower canine or ''canine index'': (1) the *S. scrofa* group containing *S. scrofa* and *S. salvanius*; (2) the *S. philippensis* group; (3) the *S. verrucosus* group containing the Javan warty pig *S. verrucosus*, the bearded pig *S. barbatus*, the Sulawesi pig *S. celebensis*, and the Visayan pig *S. cebifrons* (Groves, 2001).

This paper is the first to draw conclusions based upon the genetics of the pygmy hog. Here, we present data on mitochondrial DNA variation among a range of Suidae, and explore the relationship of the pygmy hog with other pigs.

* Corresponding author. Fax: +44 1534 860001.
E-mail address: jfa@durrell.org (J.E. Fa).
[1] Present address: Department of Medical Biochemistry and Microbiology, Uppsala University Biomedical Center, Box 597, Uppsala 751 24, Sweden.

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.08.007

2. Materials and methods

2.1. Sampling

To investigate the relationships between pygmy hog and other genera of the Suidae, we sequenced almost the entire mitochondrion of three wild-caught pygmy hogs (Assam, India) and a fragment of the control region (CR) of a museum specimen (Sikkim, India) collected by Hodgson (Hodgson, 1847). We restricted the analysis to the CR, cytochrome *b* (*cytb*) and 16S ribosomal RNA (16S) for which comparative sequence data for Suidae were available from GenBank.

Blood samples were collected from three of six individuals, captured in 1995 from the Manas Tiger Reserve to start a captive breeding program (Narayan and Jyoti Deka, 2002). The museum specimen (Spec #: 1879.11.21.666/1077.f) from the "Tarai of Sikim" collected by Hodgson, was obtained with permission from the Natural History Museum, London. A Dremmel drill was used to remove a small fragment of the mandible of the museum sample. DNA extraction, PCR amplification and sequencing were conducted independently for the recent and museum samples at the Centre for Cellular and Molecular Biology in Hyderabad, and the Henry Wellcome Ancient Biomolecules Centre (HWABC) in Oxford, respectively.

2.2. Laboratory procedures

Total genomic DNA was extracted from blood samples following the standard proteinase k/phenol–chloroform method (Sambrook et al., 1989). We sequenced almost the whole mitochondrion of the three wild-caught pygmy hogs using conserved primers, designed from the aligned mitochondrial gene sequences of *S. scrofa* and related species, and amplification conditions as described elsewhere (Verma and Singh, 2003). Because of the bias in availability of Suidae sequences from GenBank, we focused our analysis on control region (CR), *cytb*, and 16S. Three sets of primer combinations resulted in overlapping sequence fragments (GenBank Accession Nos. EU107788 to EU107791) for NADH5-*cytb* (positions 12,768–14,234, referenced to the whole pig mitochondrium, Ursing and Arnason, 1998), *cytb* (positions 14,328–15,015) and *cytb*-NADH1 (positions 15,193–3024).

The extraction, amplification and sequencing protocol at the HWABC followed Shapiro et al. (2004) and employed the use of stringent ancient DNA protocols (Cooper and Poinar, 2000). Combinations of 12 primers (Larson et al., 2005) were used to amplify overlapping CR fragments (positions 15,463–15,759; GenBank Accession No. EF472247).

2.3. Phylogenetic analyses

A total of 1081 CR sequences, 228 *cytb* sequences and 59 16S sequences were retrieved from GenBank for Suidae and outgroups (Tayassuidae, Hippopotamidae). We used the species identification listed in each GenBank entry and associated publications. For example, Randi et al. (2002) provide convincing evidence that they sampled the rare Desert Warthog *Phacochoerus aethiopicus*. Published CR sequences vary substantially in length, ranging from short fragments of the hypervariable CRI segment to the entire locus. To ensure maximum representation of species of the family Suidae, phylogenetic analyses were restricted to the CRI region and the conserved core, corresponding to the *S. scrofa* mitochondrial genome positions 15,435–15,996 (Ursing and Arnason, 1998).

Phylogenetic analyses were conducted for the three loci separately, using only GenBank sequences spanning the whole lengths of the selected proportions of 16S, *cytb* and CR with one exception. One shorter 16S fragment from babirusa *Babyrousa babyrussa* alongside one homologous *S. scrofa* sequence of the same size were included since this is the only fragment in GenBank for the genus. Because of the bias towards *S. scrofa* and the Common Warthog *P. africanus*, groups of sequences with 99–100% match percentage, were identified using Sequencher 4.5. Only one haplotype was selected from each group. We also jointly analyzed the concatenated complete CR (CRI and conserved core) and *cytb* sequences. For those species with more than two haplotypes available from GenBank, we selected two haplotypes representing the width of the species's genetic diversity in order to avoid biases towards any species (*S. scrofa*: Swedish wild boar, Kijas and Andersson, 2001 and Chinese Min breed, Yang et al., 2003; *P. africanus*: isolates Pafr1 and Pafr3, Randi et al., 2002; *P. aethiopicus*: isolates Paet1 and Paet4, Randi et al., 2002; Collared Peccary *Pecari tajacu*: one Argentinian and one Mexican haplotype, Gongora et al., 2006). A joint 16S, *cytb* and CR analysis was not conducted because of the restricted availability of 16S sequences.

Cytb and CR sequences were aligned using ClustalX (Thompson et al., 1994). 16S sequences were aligned manually using Sequencher 4.5. The manual alignment was straightforward as it contained only a few gaps, mainly between the in-group (Suidae) and the out-group (*Pecari tajacu*). To evaluate the relationship between pygmy hog and other Suidae species, we used a Bayesian Markov Chain Monte Carlo simulation (MCMC) to estimate the most likely phylogenetic trees with Mr. Bayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Bayesian inference used two runs of 6,000,000 generations each, sampled every 600 generations, and excluded the first 10% of tree samples as initialization (Ronquist and Huelsenbeck, 2003). *Cytb* was partitioned by codon position. The joint CR and *cytb* analysis partitioned into the two loci and cytb was partitioned by codon position. Parameter estimation was unlinked between partitions. We applied model parameters identified by ModelTest 3.6 (Posada and Crandall, 1998) using the Akaike Information Criterium. For the CR analysis, trees were rooted using three peccary species (Family Tayassuidae), the

White-lipped Peccary (*Tayassu pecari*), collared peccary (*Pecari tajacu*) and Chacoan (*Catagonus wagneri*). Collared peccary was used for 16S. The common and pygmy hippo (*Hippopotamus amphibius* and *Hexaprotodon liberiensis*; Family Hippopotamidae) were used to root trees in the *cytb* analysis. Pairwise distances were estimated in PAUP 4.0 (Swofford, 2001) using the best fitting model of substitutions, identified by ModelTest 3.6 (Posada and Crandall, 1998).

We evaluated alternative phylogenetic hypothesis using the non-parametric Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa, 1999). The likelihood, $-\ln L$, of competing trees for each locus was estimated by rooted maximum likelihood (ML) heuristic searches in PAUP 4.0 with 20 random addition sequence replicates and tree bisection and reconnection (TBR) branch swapping for topological unconstrained and constrained trees depending on the hypothesis being tested. ML searches employed model parameters identified by Bayesian analysis for the most likely phylogenetic tree at each locus. Searches enforcing topological constraints were used to identify the most likely trees for competing alternative phylogenetic hypothesis. Between-topology likelihood differences were statistically evaluated by SH tests in PAUP 4.0 with 10,000 RELL bootstrap replicates (Kishino and Hasegawa, 1989).

3. Results

3.1. Sequences obtained

The aligned CR sequences from the three wild-caught pygmy hogs from Assam totaled 1349 bp, corresponding to the *S. scrofa* mitochondrial genome position 15,435–16,679 (Ursing and Arnason, 1998), and defined three haplotypes. The haplotypes diverged from each other by a single base deletion, located in the hypervariable region I (CRI), and by a differing number of repeats of the 10 bp motif GTACACGTGC, located between the conserved sequence blocks 1 and 2 (Ursing and Arnason, 1998). The alignment of overlapping CR fragments from the museum sample from Sikkim totaled 319 bp (positions 15,463–15,759) and diverged from the first three haplotypes by six transitions, one transversion and two deletions. *Cytb* was amplified only for the three wild-caught pygmy hogs using three overlapping fragments. The three individuals exhibited identical sequences for the fragments including the 3′ and the 5′ end of *cytb* (positions 14,163–14,234 and 15,193–15,302, respectively). However, DNA from two individuals resulted in 'unclean' and not analyzable sequences for the central *cytb* (positions 14,328–15,015); DNA from the third individual exhibited a frame shift in the central *cytb* caused by a two-base deletion compared with pig mtDNA (Ursing and Arnason, 1998), thus indicating the probable amplification of a nuclear pseudogene (Leister, 2005). We thus excluded this proportion of *cytb* and only analyzed the concatenated 3′ and 5′ ends. The alignment of the 16S fragment spanned the complete locus (1570 bp, positions 1100–2668) and only one haplotype was observed.

Phylogenetic analyses were restricted to 566 bp at the 5′ end of the 1349 bp long pygmy hog CR (CRI and conserved core), 180 bp of cytb (combined 3′ and 5′ ends) and 1564 bp of the 1570 bp long 16S. Pygmy hog haplotypes were compared with published sequences, retrieved from GenBank and representing groups of sequences with 99–100% match percentage. For CR analysis, haplotypes represented the family Suinae including all extant subfamilies (subfamily Babyrousinae: 1 species, 2 haplotypes; subfamily Phacochoerinae: 2 species, 4 haplotypes; subfamily Suinae: 4 species, 38 haplotypes; GenBank accession numbers listed in Fig. 1a). Two genera of the subfamily Suinae (*Hylochoerus*; 1 species; *Potamochoerus*: 2 species), and three *Sus* species (Vietnam Warty Pig *S. bucculentus*, Philippine Warty Pig *S. philippensis* and Visayan Warty Pig *S. cebifrons*) were not represented. The alignment was 584 bp in length, of which 125 characters were variable and parsimony-informative in the 47 taxa of the in-group (Suidae). *Cytb* haplotypes represented the same subfamilies, genera and species except that sequences were also available for both *Potamochoerus* species and for the *S. philippensis* and the *S. cebifrons* (GenBank accession numbers listed in Fig. 1b). The 180-bp alignment showed 38 variable and parsimony-informative characters in 23 taxa of Suidae. Sequence availability was more restricted for 16S and only three *Sus* and one *Phacochoerus* species and the Babirusa were available. The 1577-bp alignment showed 93 variable and parsimony-informative characters in seven taxa of Suidae.

The Akaike Information Criteria indicated that the Hasegawa et al. (1985) HKY model of substitution with rate heterogeneity and gamma distribution was the best fit for the CR (HKY + I + G; $-\ln L = 3392.6$, AIC = 6791.1, T_i/T_v ratio = 5.6, proportion of invariable sites $I = 0.47$, gamma distribution shape parameter $\Gamma = 0.77$), the HKY model with gamma distribution as the best fit for the combined 3′ and 5′ ends of *cytb* (HKY + G; $-\ln L = 916.0$, AIC = 1841.4, T_i/T_v ratio = 3.28, $I = 0.0$, $\Gamma = 0.39$) and the GTR model with gamma distribution for 16S (GTR + G; $-\ln L = 4000.2$, AIC = 8018.5, $I = 0.0$, $\Gamma = 0.15$).

Molecular pairwise distances (Table 1) do not suggest a close relationship between pygmy hog and *Sus* for both CR and the combined 3′ and 5′ ends of *cytb*. Mean ± SD pairwise distances between pygmy hog and *Sus* species (CR: 0.134 ± 0.025, $n = 114$ pairwise comparisons; 3′/5′ *cytb*: 0.091 ± 0.0014, $n = 12$) largely exceeded distances between *Sus* species (CR: 0.032 ± 0.011, $n = 703$; 0.028 ± 0.018, $n = 69$) and was of similar magnitude as *Sus* species compared with *P. aethiopicus*, *P. africanus* and *Babyrousa babyrussa* (CR: 0.180 ± 0.042, $n = 228$; *cytb*: 0.128 ± 0.054, $n = 84$).

The phylogenetic tree produced by the Bayesian analysis for CR (Fig. 1a) joined the three haplotypes for the recent

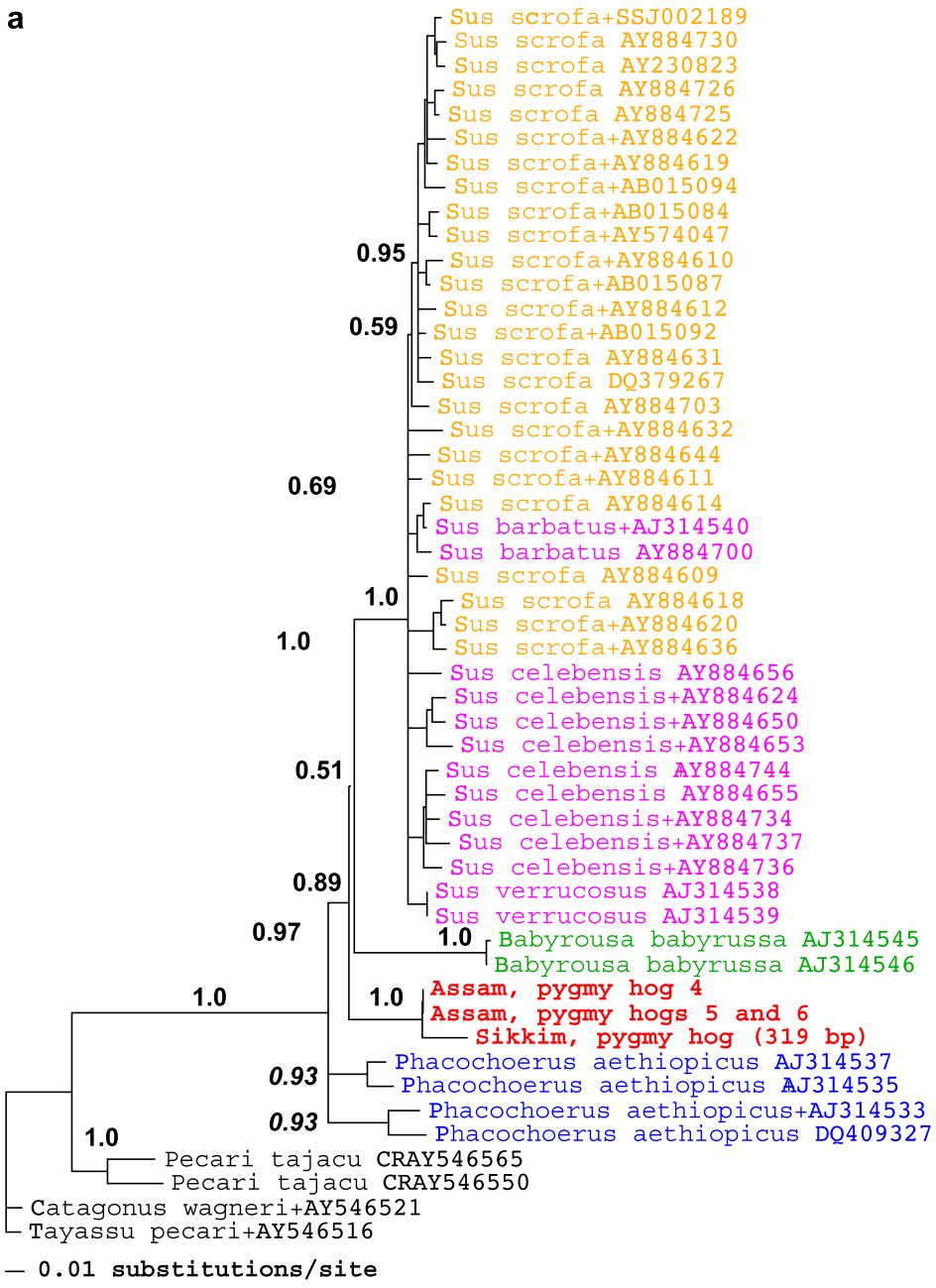

Fig. 1. Bayesian (MCMC) consensus phylograms of 584 bp mtDNA control region (a), 180 bp of combined 3′ and 5′ ends of *cytochrome b* (b), concatenated CR and cytb (c) and 1577 bp of 16S (d) of Suidae. All sequences spanned the full length of the respective fragment except the museum sample of pygmy hog for CR (a), and *Babyrousa babyrussa* and one *Sus scrofa* haplotype for 16S (b). Trees were rooted using peccary (*Tayassu pecari, Pecari tajacu* and *Catagonus wagneri*; Tayassuidae) for CR, common and pygmy hippo (*Hippopotamus amphibious* and *Hexaprotodon liberiensis*; Hippopotamidae) for *cytb* analysis, and *Pecari tajacu* for 16S, respectively. Species name and GenBank accession number identify haplotypes. Haplotypes representing groups of sequences with ⩽1% divergence are marked by '+'. Posterior probabilities (p) larger than 0.50 are included next to nodes.

and ancient samples in a single, strongly supported clade (posterior probability $p = 1.0$). No affinity of pygmy hog haplotypes with *Sus* was revealed. Three highly divergent and well supported ($p = 1.0$) clades are evident (a single clade consisting of four species of *Sus* except pygmy hog, a clade of pygmy hogs, a clade of babirusa) and a grouping of warthogs. The relationships between the *Sus*, *Babyrousa* and pygmy hog clades are unresolved with relatively low posterior probability ($p = 0.51$). Although the pairwise distances involving *Babyrousa* and *Sus* are consistently larger than those between *Phacochoerus and Sus*, *Babyrousa* is not basal to the *Phacochoerus*, *Sus* and pygmy hog clades. In contrast, warthogs are excluded from the *Sus*/*Babyrousa*/pygmy hog clade with $p = 0.89$.

The tree of combined 3′ and 5′ ends of *cytb* (Fig. 1b) showed a similar topology as the CR fragment but branch

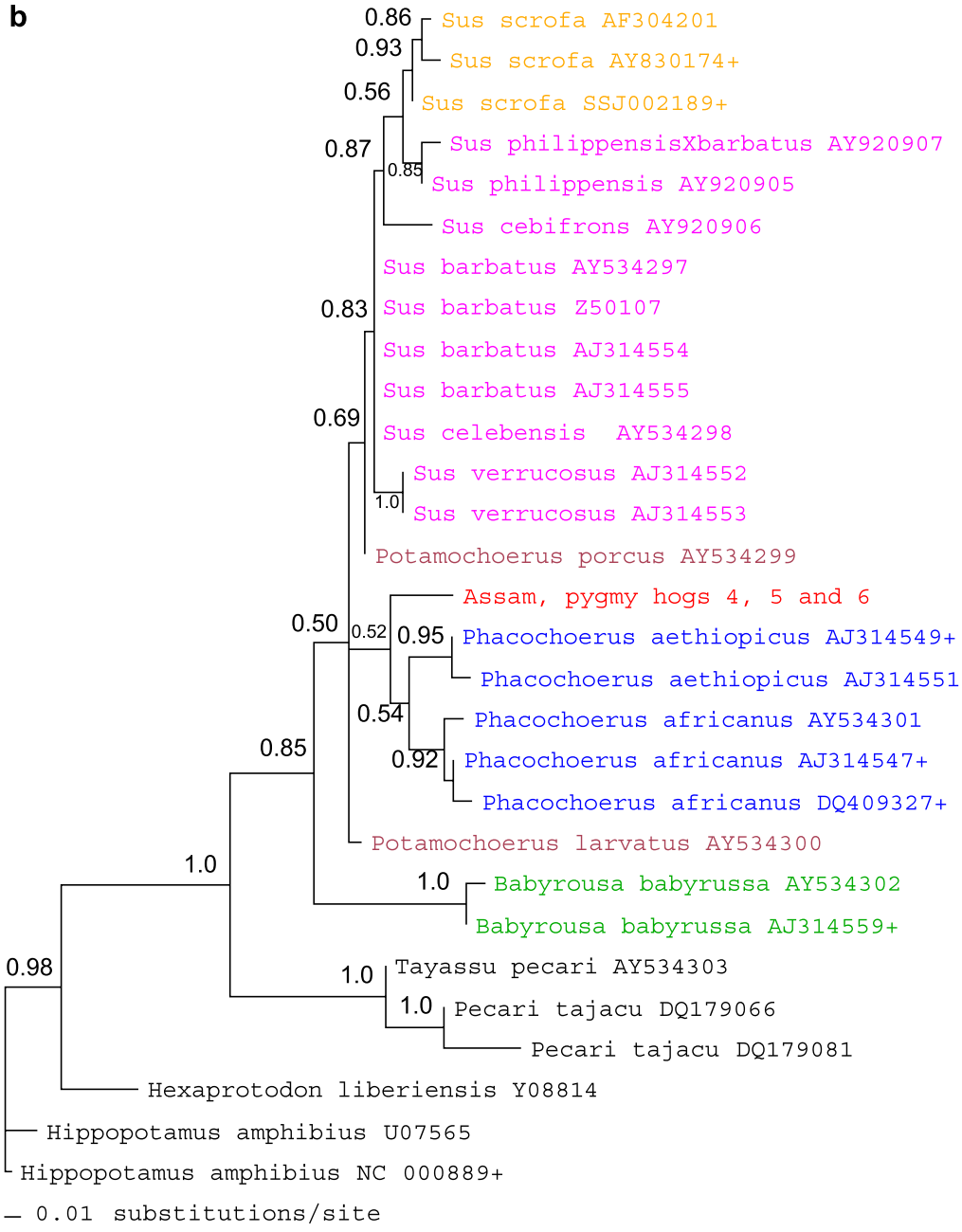

Fig. 1 (*continued*)

supports are relatively low in accordance with the short sequence length of only 180 bp: the relationships between the *Sus*, *Phacochoerus* and *Babyrousa* clades remain unresolved and the pygmy hog was not included in the *Sus* clade. Joint CR/*cytb* and 16S analyses show the same patterns (Fig. 1c and d), but branch supports for the clades are all very high with exception for the clade joining the two warthog species ($p = 0.86$). In neither case was the basal position of *Babyrousa* supported by high posterior probabilities. None of the analyses support the split of *Sus* into two clusters as predicted from Groves and Grubb, 1993b designation of two subgenera of non-warty pigs (*S. scrofa* and pygmy hog) and "warty" pigs (all other *Sus* species).

For each locus, the ML analyses without topological constraints resulted in tree topologies identical to the Bayesian approach, placing the pygmy hog outside the *Sus* clade. SH tests between these trees and ML topologies, enforcing pygmy hog as the sister species of *S. scrofa*, rejected the null hypothesis that all tested trees are equally good explanations of the CR data (tree-score $-\ln L$ of unconstrained tree: 3382.0; $-\ln L$ of constrained tree: 3421.9, $p = 0.0147$) and 16S ($-\ln L = 4005.0$ and $-\ln L = 4042.6$, respectively; $p = 0.0043$). The SH test was also significant for *cytb* ($-\ln L = 933.7$ and $-\ln L = 949.2$, respectively; $p = 0.0186$) despite the short sequence length, the overall lower posterior probabilities (Fig. 1b) and the

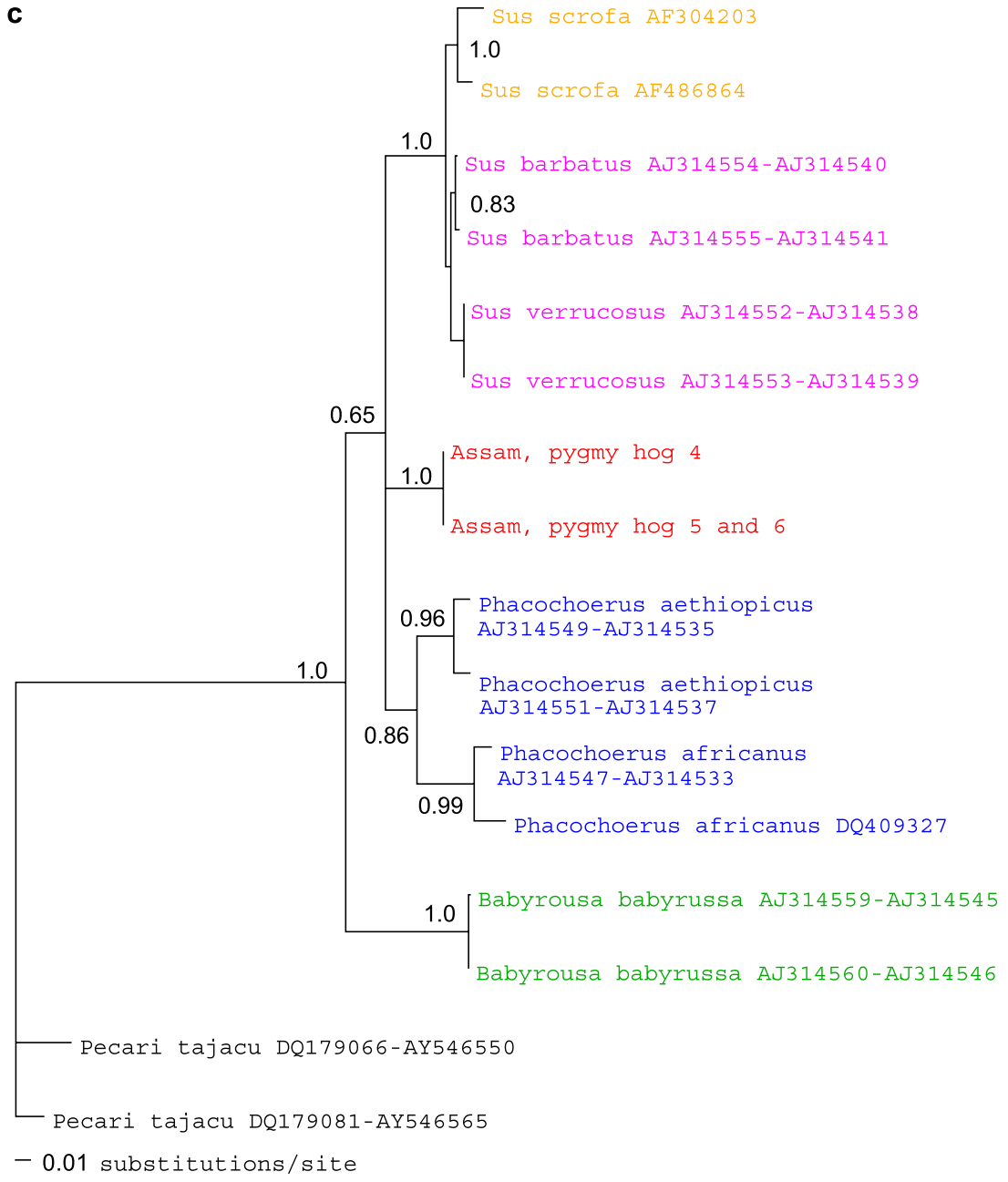

Fig. 1 (*continued*)

non-significance of the SH test for other phylogenetic relationships, in particular the relationship between the two *Potamochoerus* species (tree-score $-\ln L$ of unconstrained tree, i.e. non-monophyly: 933.7; $-\ln L$ of constrained tree, i.e. monophyly: 940.6, $p = 0.11$).

4. Discussion

4.1. Phylogenetic position of the pygmy hog and potential loss of evolutionary uniqueness

The phylogenetic exclusion of the pygmy hog from the *Sus* clade indicates that this taxon, like the babirusa and warthogs, deserves recognition as a distinct genus. This is in line with the originally described morphological uniqueness of the pygmy hog, which argued for the species to be recognised as a member of a separate genus (Hodgson, 1847). This designation was changed in a taxonomic review that reclassified the pygmy hog as a species within *Sus* and referred to as the closest relative of *S. scrofa,* belonging to the subgenus of non-warty pigs (all other *Sus* species are presumed to form the subgenus of "warty" pigs). We suggest that the original genus status should be resurrected and the species name *Porcula salvania* adopted. The conservation status of the pygmy hog (Oliver, 1993) is critical, and its extinction would result in the loss of a unique evolutionary branch of pigs.

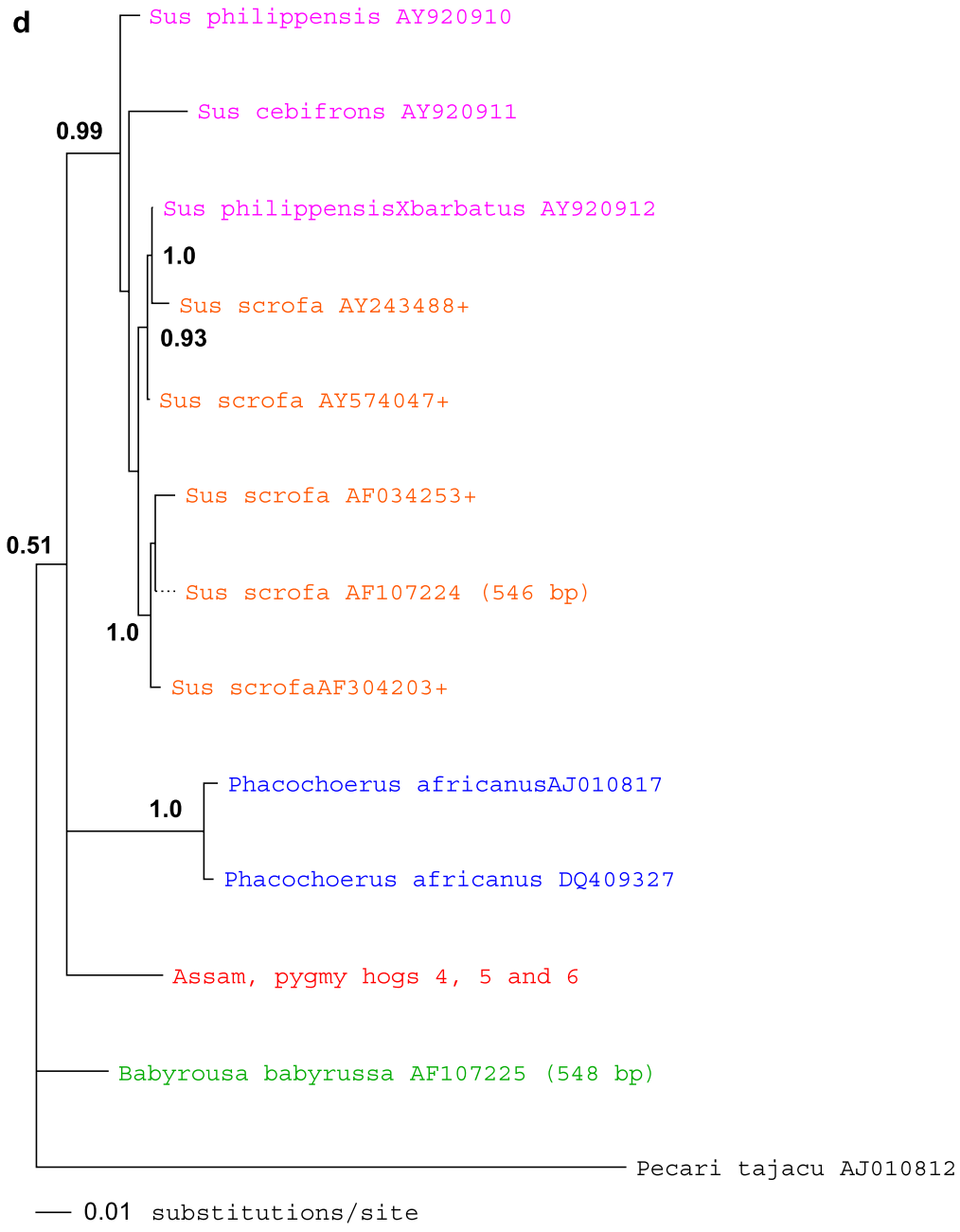

Fig. 1 (*continued*)

The evolutionary relationships between the pig genera *Sus*, *Porcula*, *Potamochoerus*, *Babyrousa* and *Phocochoerus* remain unresolved using the dataset of three mitochondrial loci. All associations between genera other than polytomies are characterized by low statistical support. This applies especially to the surprising associations between pygmy hog and warthogs (*cytb*; $p = 0.52$) and the separate clustering of *Potamochoerus* (*cytb*). The most surprising result relates to the issue of which Suid genus is most basal. Morphological comparisons have traditionally identified the babirusa as the most basal genus (Groves and Grubb, 1993a), but no published molecular dataset has rigorously tested this hypothesis. Though babirusa rests in a basal position in three of the four phylograms presented here (*cytb*, 16S, and concatenated CR/*cytb*), the posterior probabilities supporting this position are not strong (0.50, 0.51 and 0.65, respectively). Ironically, the strongest support for any basal grouping is for warthogs on the CR phylogram (0.89). Understanding evolutionary relationships between the genera of Suidae will require further sampling and sequencing, especially for *Babyrousa* and *Potamochoerus*, and the inclusion of additional, especially nuclear, loci.

We found evidence of a co-amplification of mitochondrial *cytb* and a pseudogene in case of the sequence fragment amplified from primers for the central *cytb*. First, the 'unclean' and not analyzable sequences for the central

Table 1
Molecular pairwise distance matrix for control region (below diagonal) and cytochrome *b* (above diagonal) between suid species

	N	*S. sc*	*S. ba*	*S. cel*	*S. ceb*	*S. ph*	*S. ve*	Pygmy hog	*P. la*	*P. po*	*P. ae*	*P. af*	*B. ba*
		3	4	1	1	1	2	1	1	1	2	3	2
S.sc	25	0.015 ± 0.003 0.027 ± 0.01	0.03 ± 0.005	0.06 ± 0.007	0.03 ± 0.006	0.028 ± 0.01	0.048 ± 0.008	0.1 ± 0.013	0.059 ± 0.007	0.037 ± 0.007	0.102 ± 0.01	0.113 ± 0.012	0.229 ± 0.022
S.ba	2	0.028 ± 0.009	0 ± 0 0.013	0.039 ± 0	0 ± 0	0.03 ± 0	0.015 ± 0.003	0.081 ± 0	0.024 ± 0	0.006 ± 0	0.081 ± 0	0.087 ± 0.012	0.186 ± 0.017
S.cel	9	0.04 ± 0.008	0.033 ± 0.005	— 0.034 ± 0.013	0.039	0.046	0.051 ± 0.006	0.122	0.071	0.047	0.121 ± 0	0.12 ± 0.017	0.224 ± 0.026
S.ceb	0	—	—	—	—	0.03	0.015 ± 0.004	0.081	0.024	0.006	0.081 ± 0	0.087 ± 0.014	0.186 ± 0.022
S.ph	0	—	—	—	—	—	0.048 ± 0.005	0.089	0.059	0.037	0.088 ± 0	0.1 ± 0.008	0.207 ± 0.023
S.ve	2	0.027 ± 0.009	0.012 ± 0.001	0.031 ± 0.005	—	—	0.0 0.0	0.089 ± 0.007	0.043 ± 0.005	0.022 ± 0.005	0.104 ± 0.006	0.094 ± 0.014	0.222 ± 0.022
Pygmy hog	3	0.134 ± 0.027	0.125 ± 0.026	0.139 ± 0.019	—	—	0.083 ± 0.013	— 0.017 ± 0.014	0.081	0.072	0.088 ± 0.011	0.097 ± 0.008	0.217 ± 0.024
P.la	0	—	—	—	—	—	—	—	— —	0.018	0.081 ± 0	0.065 ± 0.007	0.151 ± 0.001
P.po	0	—	—	—	—	—	—	—	—	— —	0.071 ± 0	0.065 ± 0.007	0.202 ± 0.023
P.ae	2	0.13 ± 0.009	0.124 ± 0.009	0.128 ± 0.01	—	—	0.117 ± 0.008	0.114 ± 0.006	—	—	0.011 0.025	0.078 ± 0.013	0.202 ± 0.023
P.af	2	0.176 ± 0.011	0.167 ± 0.005	0.175 ± 0.013	—	—	0.161 ± 0.01	0.138 ± 0.012	—	—	0.11 ± 0.014	0.07 ± 0.008 0.04	0.197 ± 0.016
B.ba	2	0.228 ± 0.016	0.211 ± 0.006	0.233 ± 0.01	—	—	0.216 ± 0.007	0.22 ± 0.021	—	—	0.235 ± 0.009	0.313 ± 0.012	0.012 0.002

Species abbreviations: *Sus scrofa* (*S. Sc*), *Sus barbatus* (*S. ba*), *Sus celebensis* (*S. cel*), *Sus cebifrons* (*S. ceb*), *Sus philippensis* (*S. ph*), *Potamochoerus larvatus* (*P. la*), *Potamochoerus porcus* (*P. po*), *Phacochoerus aethiopicus* (*P. ae*), *Phacochoerus africanus* (*P. af*) and *Babyrousa babyrussa* (*B. ba*). GenBank accession numbers of the sequences used are given in Fig. 1. Distances between haplotypes within species are shown on the diagonal (bottom: control region; top: *cytochrome b*). Listed are means ± SD distances applying the HKY + *G* + *I* and HKY + *G* models of DNA substitutions for control region and cytochrome *b*, respectively.

cytb of two individuals indicate that two loci were co-amplified. Second, the sequence for the third individual had a two-base deletion compared to Suid mtDNA, thus suggesting the amplification of a pseudogene. Phylogenetic analysis (not shown) placed this pygmy hog sequence basal to all Suids including babirusa (posterior probability $p = 1.0$), a phylogenetic position concordant with the lower mutation rate typical for nuclear transpositions. However, an erroneous phylogenetic exclusion of the pygmy hog from the *Sus* clade as a result of amplification of nuclear transposition is highly unlikely for the CR, 16S and concatenated 5′/3′ ends of *cytb* analyzed here. None of the sequences indicated the co-amplification of two loci, indicating that either the target mtDNA locus or a pseudogene was amplified. No pseudogenes were reported for large-scale CR and *cytb* screening in pigs and warthogs (Gongora et al., 2006; Larson et al., 2005; Muwanika et al., 2003) , indicating that the primers used are mtDNA specific. The small sequence divergence between the three pygmy hog control region haplotypes derived from recent samples and the sequence of the historic sample (amplified independently in different laboratories with different sets of primers, of which one set is the same as used in Larson et al., 2005) supports amplification of mitochondrial products only.

4.2. The pygmy hog and pig domestication

Our data also have implications for the domestication of pigs. A previous mtDNA study revealed multiple regions across Eurasia where domestication of *S. scrofa* occurred (Larson et al., 2005). One centre of domestication lies within the distribution of the pygmy hog in India (Larson et al., 2005), and though it has been suggested that pygmy hogs may have been a wild progenitor of some Asian pig breeds (Groves, 1981; Mohr, 1960), the lack of affinity of pygmy hog with any *Sus* species rules out a significant maternal genetic input to any of the numerous sampled breeds of modern domestic pigs, and suggests pygmy hogs have not been independently domesticated.

Acknowledgments

We are grateful to the British Museum Natural History (London) for permission to take samples from the specimen in their collection, and to Dr. Richard Sabin for his support and assistance. All blood samples from Assamese pygmy hogs were collected under permission of the Forestry Department of the State of Assam. We gratefully acknowledge Prof. Mike Bruford and Mr. William Oliver for suggestions on a first draft of the paper, and two anonymous reviewers for their helpful comments.

References

Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139.

Corbet, G.B., Hill, J.E., 1992. The Mammals of the Indomalayan Region: A Systematic Review. Oxford University Press, Oxford.

Ghosh, M., 1988. The craniology and dentition in the pigmy hog, with a note on the genetic status of *Porcula* Hodgson, 1847. Records of the Zoological Survey of India: A Journal of Indian Zoology 85, 245–266.

Gongora, J., Morales, S., Bernal, J.E., Moran, C., 2006. Phylogenetic divisions among Collared peccaries (*Pecari tajacu*) detected using mitochondrial and nuclear sequences. Molecular Phylogenetics and Evolution 41, 1–11.

Groves, C., 1981. Ancestors for the Pigs: Taxonomy and Phylogeny of the Genus *Sus*. Technical Bulletin 3, Department of Prehistory, Research School of Pacific Studies, Australian National University, pp. 96.

Groves, C., 2001. Taxonomy of wild pigs of Southeast Asia. Asian Wild Pig News 1, 3–4.

Groves, C.P., Grubb, P., 1993a. The suborder Suiformes – Chapter 1. In: Oliver, W.L.R. (Ed.), Pigs, Peccaries and Hippos Status Survey and Action Plan. IUCN/SSC Pigs and Peccaries Specialist Group & IUCN/SSC Hippo Specialist Group.

Groves, C.P., Grubb, P., 1993b. The Eurasian Suids *Sus* and Babyrousa, Taxonomy and description – Chapter 5.1. In: Oliver, W.L.R. (Ed.), Pigs, Peccaries and Hippos Status Survey and Action Plan. IUCN/SSC Pigs and Peccaries Specialist Group & IUCN/SSC Hippo Specialist Group.

Hasegawa, M., Kishino, K., Yano, T., 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22, 160–174.

Hodgson, B.H., 1847. On a new form of the hog kind or Suidae. Journal of the Asiatic Society of Bengal 16, 423–428.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Kijas, J.M.H., 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.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29, 170–179.

Larson, G., Dobney, K., Albarella, U., Fang, M.Y., Matisoo-Smith, E., Robins, J., Lowden, S., Finlayson, H., Brand, T., Willerslev, E., Rowley-Conwy, P., Andersson, L., Cooper, A., 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307, 1618–1621.

Leister, D., 2005. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. Trends in Genetics 21, 655–663.

Mohr, E., 1960. Wilde Schweine (reprint 2004, Neue Brehm-Bücherei no. 247). Springer, Wittenberg.

Muwanika, V.B., Nyakaana, S., Siegismund, H.R., Arctander, P., 2003. Phylogeography and population structure of the common warthog (*Phacochoerus africanus*) inferred from variation in mitochondrial DNA sequences and microsatellite loc. Heredity 91, 361–372.

Narayan, G., Jyoti Deka, P., 2002. Pygmy hog conservation programme in Assam, India. Asian Wild Pig News 2, 5–7.

Nowak, R.M., 1999Walker's Mammals of the World, vol. II. John Hopkins University Press, Baltimore, MD.

Oliver, W.L.R. (Ed.), 1993. Pigs, Peccaries and Hippos Status Survey and Action Plan. IUCN/SSC Pigs and Peccaries Specialist Group & IUCN/SSC Hippo Specialist Group.

Oliver, W.L.R., Deb Roy, S., 1993. The Pigmy Hog (*Sus salvanius*) – Chapter 5.3. In: Oliver, W.L.R. (Ed.), Pigs, Peccaries and Hippos Status Survey and Action Plan. IUCN/SSC Pigs and Peccaries Specialist Group & IUCN/SSC Hippo Specialist Group.

Pilgrim, G.E., 1926. The fossil Suidae of India. Palaeontologica Indica n.s. 8, 1–65.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Randi, E., D'Huart, J.P., Lucchini, V., Aman, R., 2002. Evidence of two genetically deeply divergent species of warthog, *Phacochoerus africanus*

and *P-aethiopicus* (Artiodactyla : Suiformes) in East Africa. Mammalian Biology 67, 91–96.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sambrook, J., Fritsh, E.F., Maniatis, T., 1989, 2nd ed.Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory Press, New York.

Shapiro, B., Drummond, A.J., Rambaut, A., Wilson, M.C., Matheus, P.E., Sher, A.V., Pybus, O.G., Gilbert, M.T.P., Barnes, I., Binladen, J., Willerslev, E., Hansen, A.J., Baryshnikov, G.F., Burns, J.A., Davydov, S., Driver, J.C., Froese, D.G., Harington, C.R., Keddie, G., Kosintsev, P., Kunz, M.L., Martin, L.D., Stephenson, R.O., Storer, J., Tedford, R., Zimov, S., Cooper, A., 2004. Rise and fall of the Beringian Steppe Bison. Science 306, 1561–1565.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16, 1114–1116.

Swofford, D.L., 2001. PAUP: Phylogenetic Analysis Using Parsimony. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

Ursing, B.M., Arnason, U., 1998. The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). Journal of Molecular Evolution 47, 302–306.

Verma, S.K., Singh, L., 2003. Novel universal primers establish identity of an enormous number of animal species for forensic application. Molecular Ecology Notes 3, 28–31.

Yang, J., Wang, J., Kijas, J., Liu, B., Han, H., Yu, M., Yang, H., Zhao, S., Li, K., 2003. Genetic diversity present within the near-complete mtDNA genome of 17 breeds of indigenous Chinese pigs. Journal of Heredity 94, 381–385.