Molecular evidence suggests that the enigmatic Sulawesi endemic *Geomalia heinrichi* belongs in the genus *Zoothera* (Turdidae, Aves)

Urban OLSSON [1,✉], **Per ALSTRÖM** [2,3]

[1] Department of Biology and Environmental Sciences, University of Gothenburg, Box 463, SE 405 30 Göteborg, Sweden

[2] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China

[3] Swedish Species Information Centre, Swedish University of Agricultural Sciences, Box 7007, SE-750 07 Uppsala, Sweden

Abstract The taxonomic status of the Sulawesi endemic *Geomalia heinrichi* has long been debated, and it has variously been treated as a babbler (Timaliidae) or a turdid (Turdidae). We estimated the phylogeny of 43 taxa in the family Turdidae based on the mitochondrial cytochrome *b* gene and the nuclear myoglobin intron 2 and ornithine decarboxylase introns 6–7. *Geomalia heinrichi* was shown to be part of the *Zoothera* clade with high support. We propose that *Geomalia* is transferred to *Zoothera* under the name *Zoothera heinrichi*.

Introduction

The Sulawesi endemic Geomalia (*Geomalia heinrichi*) is a medium-sized terrestrial bird living in mossy forest at 1700–3500 m a.s.l., mainly 1800–2200 m (Collar, 2005). In some respects it resembles both thrushes (Turdidae) and babblers (Timaliidae/Pellorneidae/Leiotrichidae), and no consensus as to its taxonomic position has been reached to date. The species was discovered by Gerd Heinrich in 1930 in the Latimodjong mountains in central Sulawesi. He made his first observation already on his first day in the field, and initially mistook it for an Island Thrush (*Turdus poliocephalus*). Throughout the description of his first encounters with the species, he referred to it as a thrush (quoted in Stresemann, 1940). Stresemann (1931) provided a detailed descrip-

tion, pointing out morphological features of which some seemed to indicate that *Geomalia* was closer to thrushes, and others suggesting that it belonged to the babblers. The early attempts to classify *Geomalia* by Stresemann (1931, 1940) followed a tradition in which specific characters judged to be more important than others were relied on to determine affinity, rather than synapomorphies. His initial conclusion (Stresemann, 1931) was that he thought that *Geomalia* was related to the shortwings, particularly *Heinrichia calligyna*, alternatively best placed in its own monotypic subfamily, Geomaliinae. Later, Stresemann (1940) reconsidered this and placed *Geomalia* in Timaliinae.

Following Stresemann's initial descriptions and taxonomic conclusions (Stresemann, 1931, 1940), very little primary data have been added. The taxonomic position of this monotypic genus has remained controversial, with opinion swinging between babbler or thrush affinity. Ripley (1952) was the first to disagree with Stresemann (1940), by placing *Geomalia* in Turdidae, immediately before *Zoothera* in a linear taxonomy, although

without providing any reasons. Subsequent authors have followed either of these two positions. Most major world lists or handbooks have placed *Geomalia* before *Zoothera/Geokichla* in Turdidae (Sibley and Monroe, 1990; Dickinson, 2003; Clement, 2000; Clements et al., 2012; Gill and Donsker, 2013). Collar (2005) emphasized the uncertain phylogenetic position of *Geomalia* by placing it among the first species in Turdinae. Collar (2004, 2005) noted that *Geomalia* has a distinct juvenile plumage, with spotted underparts, unlike any babbler.

White and Bruce (1986) placed *Geomalia* in Timaliidae, but mixed up their reference to Stresemann by stating that he placed *Geomalia* in Timaliidae in 1931, which in fact he did in 1940. They further stated that Stresemann (1931) did not mention a thrush-like juvenile plumage, which is correct in the sense that Stresemann (1931) did not refer to the juvenile plumage as "thrush-like". However, his description, although somewhat vague, stated that the feathers of the breast are edged grey-black in juveniles, not uniformly colored as in the adult (Stresemann, 1931). Holmes and Philipps (1996) and Coates and Bishop (1997) followed White and Bruce (1986) in placing *Geomalia* in Timaliidae, in the latter case reiterating the incorrect reference to Stresemann (1931).

No cladistic analysis of morphological characters including *Geomalia* has been undertaken. We here present a hypothesis of the phylogenetic position of *Geomalia heinrichi* based on DNA sequences of mitochondrial and nuclear sequences obtained from one of Heinrich's specimens. To the best of our knowledge, this is the first molecular study of the phylogenetic position of *Geomalia heinrichi*.

Materials and methods

DNA sampling and sequencing

DNA was extracted from feathers and a claw from a juvenile *Geomalia heinrichi heinrichi* housed in the Museum für Naturkunde der Humboldt-Universität, Berlin, collected in the Latimodjong Mountains, southern Sulawesi on 16 July 1939 by Gerd Heinrich (ZMB 34.160; field number 1018), and from tissue samples of a few additional taxa (Supplementary Table S1).

DNA was extracted using QIA Quick DNEasy Kit (Qiagen, Inc) according to the manufacturer's instructions, but with 30 μL DTT added to the initial incubation step for the extraction from feathers, toepads and

claws. For the amplification and sequencing of the fresh samples, we sequenced the mitochondrial cytochrome *b* (cytb) gene and the entire nuclear myoglobin intron 2 (myo) following the protocols described in Olsson et al. (2005), and the nuclear ornithine decarboxylase exon 6 (partial), intron 6, exon 7, intron 7 and exon 8 (partial) (ODC) following the protocols described in Friesen et al. (1999) and Allen and Omland (2003).

We sequenced a few additional taxa, and downloaded some more sequences from GenBank (Supplementary Table S1). Both the cytb gene and the nuclear introns from the museum toepad samples were sequenced in small fragments, using a range of specific primer combinations as described by Irestedt et al. (2006), Dario Zuccon (personal communication) and some primers designed specifically for this study (Supplementary Table S2). All sequences have been submitted to GenBank (Supplementary Table S1).

Phylogenetic analyses

Sequences were aligned using MegAlign 4.03 in the DNAstar package (DNAstar Inc.).

Phylogenies were estimated by Bayesian inference (BI) using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). All analyses were run under the best-fit models according to the Bayesian Information Criterion (BIC), calculated in jModeltest 0.1.1 (Posada, 2008a, 2008b). The preferred model for the cytb data was the general time-reversible (GTR) model (Lanave et al., 1984; Tavaré, 1986; Rodríguez et al., 1990), assuming rate variation across sites according to a discrete gamma distribution with four rate categories (Γ; Yang, 1994), and an estimated proportion of invariant sites (I; Gu et al., 1995) (GTR + Γ + I). For ODC the preferred model was GTR + I, and for myo the HKY (Hasegawa et al., 1985) model + Γ.

The three loci were analysed both separately and concatenated. In the combined BI analysis, the data were divided into locus-specific partitions. In each analysis, four Metropolis-coupled MCMC chains were run for 30 million generations and sampled every 1000 generations. The posterior distributions were summarized as a majority-rule consensus tree. The first 7.5 million generations of each run, well after the chains reached stationarity, were discarded as burn-in.

The concatenated data were also analysed by maximum likelihood bootstrapping (MLBS) and parsimony bootstrapping (MPBS). MLBS (1000 replicates) was

conducted with RAxML-HPC2 version 7.3.2 (Stamatakis, 2006; Stamatakis et al., 2008) on the Cipres portal (Miller et al., 2010). The data were partitioned by locus, and as per default GTRCAT was used for the bootstrapping phase, and GTRGAMMA for the final tree inference. The MPBS was performed in PAUP* (Swofford, 2001): heuristic search strategy, 10000 replicates, starting trees obtained by stepwise addition (random addition sequence, 10 replicates), TBR branch swapping. *Myophonus caeruleus,* which is part of Muscicapidae, the sister group of Turdidae (Sangster et al., 2010), was used as outgroup in all analyses.

Results

Sequence characteristics

The aligned cytb of 44 taxa comprised 1038 characters, of which 374 (36%) were parsimony informative. For 41 sequences the complete target stretch was available, while 3 included missing data and varied between 976 and 1011 bp in length. We found no unexpected stop codons, indels or distinct double peaks in the chromatograms that would indicate the presence of nuclear pseudogenes (e.g. Sorenson and Quinn, 1998).

The length of the nuclear introns varied depending on multiple indels. The aligned myo sequences of 43 taxa comprised 758 characters, of which 94 (12%) were parsimony informative. For 36 sequences the complete target stretch of 695–713 bp in length was obtained. Nine sequences included missing data. Eight of these varied between 674 and 705 bp in length, one sequence was only 181 bp, and for *Ridgwayia pinicola* no sequence from this locus was available. Seven indels were shared by two or more taxa, and one of these, a 2 bp deletion, supported the position of *Geomalia* in the phylogeny (see below; Fig. 1).

The aligned ODC of 43 taxa comprised 733 characters, of which 104 (14%) were parsimony informative. For 29 sequences the complete target stretch of 690–706 bp in length was available. Fourteen sequences included missing data, the majority varying between 499 and 691 bp in length. The *Geomalia* sequence was 191 bp, and *Ridgwayia pinicola* 156 bp. For *Zoothera everetti* sequence data from this locus was missing.

Single-locus analyses

The trees based on single-locus analyses (hereafter SLAs) varied in resolution. All three SLAs recovered *Geomalia* as part of the *Zoothera* clade, with good support (PP 0.92 in the cytb tree; PP 0.97 in the ODC tree and PP 0.99 in the myo tree) (Supplementary Figs. S1–S3).

Concatenated multilocus analyses

The tree based on the concatenated sequences (Fig. 1) was overall congruent with various trees presented by Cibois and Cracraft (2004), Klicka et al. (2005), Nylander et al. (2008), Voelker et al. (2007) and Voelker and Klicka (2008). In agreement with previous studies, the exact positions of some problematic taxa, such as *Geokichla sibirica*, *Ixoreus naevius*, *Ridgwayia pinicola*, *Hylocichla mustelinus* and *Entomodestes leucotis* were uncertain to some degree. *Geomalia* was part of the *Zoothera* clade with PP 1.00, 95% MLBS and 62% MPBS, although its exact position within the clade was unresolved.

Discussion

The phylogeny is overall relatively well supported and in agreement with previous publications (e.g. Cibois and Cracraft, 2004; Klicka et al., 2005; Nylander et al., 2008; Voelker et al., 2007; Voelker and Klicka, 2008) regarding reference taxa. We are not concerned here with minor differences within clades; all major clades found in earlier works are recovered here, and the position of *Geomalia* as part of the *Zoothera* clade is corroborated both by the concatenated data and by all of the SLAs. Within the *Zoothera* clade, the position of *Geomalia* is unclear. The present data indicate that it stems from a basal node, suggesting that it may have become isolated on Sulawesi early in the *Zoothera* radiation. The juvenile plumage of *Geomalia* shows a dark vertical bar on the rear ear-coverts and a dark malar stripe, and the breast feathers have dark scallops (photos available on www.orientalbirdimages.org), all traits recalling several *Zoothera* species. It is worth noting that the Borneo endemic *Z. everetti* resembles *Geomalia* in having warm orange underparts (albeit with whitish belly), lacking the dark spotting or scalloping typical of the genus *Zoothera,* and in showing a facial pattern reminiscent of juvenile *Geomalia*. Juvenile *Z. everetti* has more heavily scalloped underparts than juvenile *Geomalia*. It thus appears that the typical *Zoothera* pattern is retained in juveniles, whereas it has become lost in adults in both
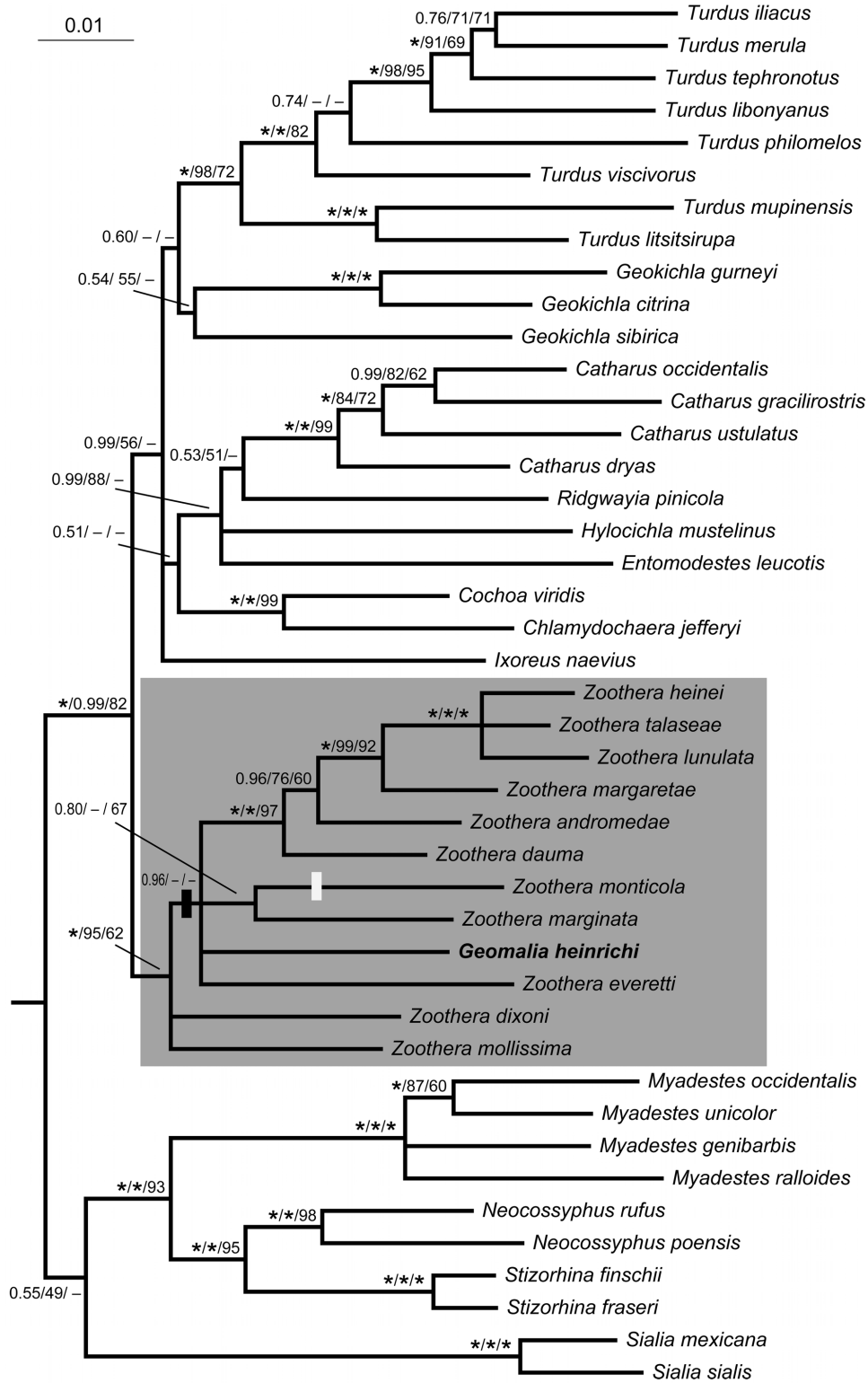
Fig. 1 Majority rule (50%) consensus tree based on concatenated nuclear intron 2 of myoglobin and introns 6–7 of ODC and mitochondrial cytochrome *b*, inferred by Bayesian inference, analysed in three partitions, one per locus. Support values are indicated at the nodes, in the order posterior probability/maximum likelihood bootstrap/parsimony bootstrap; an asterisk represents support 1.00/100%. The shaded rectangle indicates the genus *Zoothera*. One 2 bp deletion in the myoglobin alignment supporting monophyly of the *Zoothera* clade including *Geomalia heinrichi* is indicated by a black bar; a white bar denotes an inferred reversal of this state.

these species. The characters that have suggested to some earlier workers that *Geomalia* is a babbler may be adaptions to its apparently obligate terrestrial life-style. As *Geomalia* is firmly placed in *Zoothera*, we propose that it should be assigned to that genus, under the name *Zoothera heinrichi*.

References

Allen ES, Omland KE. 2003. Novel intron phylogeny supports plumage convergence in orioles (*Icterus*). Auk, 120:961–969.

Cibois A, Cracraft J. 2004. Assessing the passerine "tapestry": phylogenetic relationships of the Muscicapoidea inferred from nuclear DNA sequences. Mol Phylogenet Evol, 32:264–273.

Clement P. 2000. Thrushes. Princeton University Press, Princeton, N.J.

Clements JF, Schulenberg TS, Iliff MJ, Sullivan BL, Wood CL, Roberson D. 2012. The eBird/Clements checklist of birds of the world: Version 6.7. http://www.birds.cornell.edu/clements-checklist/downloadable-clements-checklist

Coates BJ, Bishop KD. 1997. A Guide to the Birds of Wallacea. Dove Publ., Alderley, Australia.

Collar NJ. 2004. Four odd "thrushes". BirdingAsia, 1:18–21.

Collar NJ. 2005. Family Turdidae (Thrushes), *Geomalia*. In: del Hoyo J, Elliott A, Christie DA (eds) Handbook of the Birds of the World (Vol. 10). Lynx Edicions, Barcelona, Spain, p 633.

Dickinson E (ed.). 2003. The Howard and Moore Complete Checklist of the Birds of the World. 3rd ed. Helm, London.

Friesen VL, Congdon BC, Kidd MG, Birt TP. 1999. Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. Mol Ecol, 8:2147–2149.

Gill F, Donsker D. 2013. IOC World Bird List (v. 3.3). http://www.worldbirdnames.org. Accessed 2 March 2013.

Gu X, Fu Y-X, Li W-H. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol Biol Evol, 12:546–557.

Hasegawa M, Kishino Y, Yano T. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol, 22:160–174.

Helbig A J, Seibold I, Martens J, Wink M. 1995 Genetic differentiation and phylogenetic relationships of Bonelli's Warbler *Phylloscopus bonelli* and Green Warbler *P. nitidus*. J Avian Biol, 26:139–153.

Holmes DA, Philipps K. 1996. The Birds of Sulawesi. Oxford University Press, Kuala Lumpur.

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics, 17:754–755.

Irestedt M, Ohlson JI, Zuccon D, Källersjö M, Ericson PGP. 2006. Nuclear DNA from old collections of avian study skins reveals the evolutionary history of the Old World suboscines (Aves, Passeriformes). Zool Scr, 35:567–580.

Klicka J, Voelker G, Spellman GM. 2005. A molecular systematic revision of the true thrushes (Turdinae). Mol Phylogenet Evol, 34:486–500.

Lanave C, Preparata C, Saccone C, Serio G. 1984. A new method for calculating evolutionary substitution rates. J Mol Evol, 20:86–93.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA, pp 1–8.

Nylander JAA, Olsson U, Alström P, Sanmartín I. 2008. Accounting for Phylogenetic Uncertainty in Biogeography: A Bayesian Approach to Dispersal-Vicariance Analysis of the Thrushes (Aves: Turdus). Syst Biol, 57:257–268.

Olsson U, Alström P, Ericson PG, Sundberg P. 2005. Non-monophyletic taxa and cryptic species —evidence from a molecular phylogeny of leaf-warblers (*Phylloscopus*, Aves). Mol Phylogenet Evol, 36:261–276.

Posada D. 2008a. jModelTest: phylogenetic model averaging. Mol Biol Evol, 25:1253–1256.

Posada D. 2008b. jModeltest. version 0.1.1. https://code.google.com/p/jmodeltest2. Accessed 17 March 2013.

Ripley SD. 1952. The Thrushes: a Taxonomic Study. Postilla 13.

Rodríguez J, Oliver L, Marín A, Medina R. 1990. The general stochastic model of nucleotide substitution. J Theor Biol, 142:485–501.

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

Sangster G, Alström P, Forsmark E, Olsson U. 2010. Multilocus phylogenetic analysis of Old World chats and flycatchers reveals extensive paraphyly at family, subfamily and genus level (Aves: Muscicapidae). Mol Phylogenet Evol, 57:380–392.

Sibley CG, Ahlquist JE. 1990. Phylogeny and Classification of Birds: a Study of Molecular Evolution. Yale University Press, New Haven, CT.

Sorenson MD, Quinn TW. 1998. Numts: a challenge for avian

systematics and population biology. Auk, 115:214–221.

Stamatakis A, Hoover P, Rougemont J. 2008. A fast bootstrapping algorithm for the RAxML web-servers. Syst Biol, 57:758–771.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22:2688–2690.

Stresemann E. 1931. Zur ornithologie des Latimodjong-Gebirges in südlichen Central-Celebes. Orn Monatsb, 39:7–14.

Stresemann E. 1940. Die Vögel von Celebes. Teil III. Systematik und Biologie. J Ornithol, 88:139.

Swofford DL. 2001. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4.08b. Sinauer Associates, Sunderland, MA.

Tavaré S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lec Math Life Sci, 17:57–86.

Voelker G, Klicka J. 2008. Systematics of *Zoothera* thrushes, and a synthesis of true thrush molecular systematic relationships.

Mol Phylogenet Evol, 49:377–381.

Voelker G, Rohwer S, Bowie RCK, Outlaw DC. 2007. Molecular systematics of a speciose, cosmopolitan songbird genus: defining the limits of, and relationships among, the Turdus thrushes. Mol Phylogenet Evol, 42:422–434.

White CMN, Bruce MD. 1986. The Birds of Wallacea. B.O.U. Check-list No. 7. B.O.U., London.

Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol, 39:306–314.

Supplementary material

Available on Chinese Birds website, links at http://www.chinese-birds.net, including

Tables S1–S2

Figures S1–S3

分子证据表明苏拉威西岛特有鸟种苏拉山鸫（*Geomalia heinrichi*）属于地鸫属（*Zoothera*）（鸫科，鸟纲）

Urban OLSSON [1]，Per ALSTRÖM [2,3]

（1 瑞典歌德堡大学生物与环境科学系；2 中国科学院动物研究所动物进化与系统学重点实验室，北京，100101；3 瑞典农业大学瑞典物种信息中心）

摘要：苏拉威西地区特有鸟种苏拉山鸫（*Geomalia heinrichi*）的分类地位，是属于画眉科还是鸫科，长期以来一直备受争议。基于线粒体细胞色素 *b* 基因、细胞核肌红蛋白内含子 2 以及鸟氨酸脱羧酶内含子 6–7 基因，我们估计了鸫科 43 个分类群的系统发育关系。结果表明，苏拉山鸫是地鸫属的一个分支，并具有较高的支持率。我们建议将 *Geomalia* 并入地鸫属（*Zoothera*），并命名为 *Zoothera heinrichi*。

关键词：贝叶斯推断，细胞色素 *b*，细胞核肌红蛋白内含子 2（myo），鸟氨酸脱羧酶（ODC）内含子 6–7，分类学