**ORIGINAL ARTICLE** 



# Densely sampled phylogenetic analyses of the Lesser Shorttoed Lark (*Alaudala rufescens*) — Sand Lark (*A. raytal*) species complex (Aves, Passeriformes) reveal cryptic diversity

Fatemeh Ghorbani<sup>1</sup> | Mansour Aliabadian<sup>1,2</sup> | Ruiying Zhang<sup>3</sup> | Martin Irestedt<sup>4</sup> | | Yan Hao<sup>3</sup> | Gombobaatar Sundev<sup>5</sup> | Fumin Lei<sup>3</sup> | Ming Ma<sup>6</sup> | Urban Olsson<sup>7,8</sup> | | Per Alström<sup>3,9</sup>

<sup>1</sup>Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran

Revised: 10 February 2020

- <sup>2</sup>Zoological Innovations Research Department, Institute of Applied Zoology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran
- <sup>3</sup>Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China
- <sup>4</sup>Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden
- <sup>5</sup>National University of Mongolia and Mongolian Ornithological Society, Ulaanbaatar, Mongolia

<sup>6</sup>Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Xinjiang, China

- <sup>7</sup>Systematics and Biodiversity, Department of Biology and Environmental Sciences, University of Gothenburg, Göteborg, Sweden
- <sup>8</sup>Gothenburg Global Biodiversity Centre, Gothenburg, Sweden
- <sup>9</sup>Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

#### Correspondence

Mansour Aliabadian, Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran. Email: aliabadi@um.ac.ir

Per Alström, Animal Ecology, Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden. Email: per.alstrom@ebc.uu.se

#### **Funding information**

Ferdowsi University of Mashhad, Grant/ Award Number: 3/29628 ; Jornvall Foundation; Swedish Research Council, Grant/Award Number: 2015-04402, 2015-04651 and 621-2014-5113

#### Abstract

The taxonomy of the Lesser/Asian Short-toed Lark Alaudala rufescens-cheleensis complex has been debated for decades, mainly because of minor morphological differentiation among the taxa within the complex, and different interpretations of the geographical pattern of morphological characters among different authors. In addition, there have been few studies based on non-morphological traits. It has recently been suggested based on a molecular study of the lark family Alaudidae that the Sand Lark A. raytal is nested within this complex. We here analysed mitochondrial cytochrome b (cyt b) from 130 individuals across the range of this complex (hereafter called Alaudala rufescens-raytal complex), representing all except two of the 18 currently recognized subspecies. We also analysed 11 nuclear markers from a subsample of these individuals, representing all of the clades found in the cyt b tree. Five primary clades were recovered, which confirmed that A. raytal is nested within this complex. Divergence time estimates among these five clades ranged from 1.76 to 3.16 million years (my; 95% highest posterior density [HPD] 1.0–4.51 my) or 1.99-2.53 my (95% HPD 0.96-4.3 my) in different analyses. Only four of the currently recognized subspecies were recovered as monophyletic in the cyt b tree. Our results call for a taxonomic revision, and we tentatively suggest that at least four

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. © 2020 The Authors. *Zoologica Scripta* published by John Wiley & Sons Ltd on behalf of Royal Swedish Academy of Sciences species should be recognized, although we stress the need for an approach integrating molecular, morphological and other data that are not yet available.

**KEYWORDS** 

cryptic species, multilocus analysis, Palearctic, phylogeny, species tree, taxonomy

# **1** | INTRODUCTION

The family Alaudidae, larks, are passerine birds within the superfamily Sylvioidea (reviewed by Alström, Olsson, & Lei, 2013). Unlike the other families within Sylvioidea (except Hirundinidae, swallows), which inhabit more or less dense habitats, such as forests, scrub and thick marshland vegetation, the larks are adapted to open habitats, and several of the species thrive in stony or sandy desert with little or no vegetation. Alaudidae comprises 92-98 species in 21 genera, which are distributed throughout Eurasia (35-39 species in 14 genera), Africa (74–78 species in 21 genera), North America (one species in one genus) and Australia (one species in one genus; some of the species and genera occur on two or three continents; Christidis, 2018; Gill & Donsker, 2019; del Hoyo & Collar, 2016). Most species of larks are cryptically coloured, with a striking correlation between their overall colouration and the substrate they live on (Donald, Alström, & Engelbrecht, 2017). The selection for crypsis in the open habitats is likely the reason for the strong plumage conservatism within the family, and many species of larks are renowned for being difficult to distinguish. The generally poor plumage differentiation and multiple cases of both morphological convergence and divergence revealed by a molecular phylogeny of more than 80% of the species (Alström, Barnes, et al., 2013) led the authors of that study to conclude that "Few groups of birds show the same level of disagreement between taxonomy based on morphology and phylogenetic relationships as inferred from DNA sequences."

The genus Alaudala comprises 3-5 species (Christidis, 2018; Clements et al., 2018; Gill & Donsker, 2019; del Hoyo & Collar, 2016), which are small (13–14 cm, c. 20–27 g), inconspicuously coloured and patterned larks showing various shades of brown and grey above and mostly whitish underparts, with variously distinct dark streaking on the upperside and breast (de Juana & Suárez, 2019). They were previously placed in the genus Calandrella, but were removed from there following phylogenetic evidence that Calandrella was non-monophyletic (Alström, Barnes, et al., 2013). The Lesser Short-toed Lark Alaudala rufescens complex is distributed across the southern Palearctic from the Canary Islands to north-east China (Christidis, 2018; Gill & Donsker, 2019; del Hoyo & Collar, 2016; de Juana & Suárez, 2019; Figure 1). The taxonomy of this complex has been much debated (e.g., Bianchi, 1905,

1906; Dementiev & Gladkov, 1954; Dickinson & Dekker, 2001; Hartert, 1904; del Hoyo & Collar, 2016; Korelov, 1958; Meinertzhagen, 1951; Sharpe, 1890; Stachanow & Spangenberg, 1931; Stepanyan, 1967, 1983, 1990; Vaurie, 1959), mainly because the different taxa are very similar morphologically, with poorly understood geographical variation, and because different authors have drawn different conclusions regarding relationships based on morphological evidence. Also, different authors have placed different faith in field studies undertaken in the former Soviet Union during the early to mid-1900s, and there have been no comprehensive studies of non-morphological data. In total, up to 16 taxa are usually recognized in the A. rufescens complex, although there is no consensus on the number of taxa and their distributions (Christidis, 2018; Gill & Donsker, 2019; del Hoyo & Collar, 2016; Peters, 1960; Table 1). These taxa are treated as a single species (e.g., Dementiev & Gladkov, 1954; Peters, 1960 [also including Somali Short-toed Lark A. somalica]; de Juana & Suárez, 2019; del Hoyo & Collar, 2016) or as two species, Lesser Short-toed Lark A. rufescens sensu stricto (s.s.) and Asian Short-toed Lark A. cheleensis (e.g., Christidis, 2018; Gill & Donsker, 2019; Roselaar, 1995; Table 1). The reasons for separating the complex into two species rest on reports of sympatry and morphological and ecological differentiation between the taxa heinei (placed in A. rufescens s.s.) and leucophaea (placed in A. cheleensis) in Central Asia (Bianchi, 1905, 1906; Korelov, 1958; Stachanow & Spangenberg, 1931; Stepanyan, 1967, 1983, 1990), although this has been questioned (Dementiev & Gladkov, 1954; del Hoyo & Collar, 2016; de Juana & Suárez, 2019). There is no consensus on the delimitation of the species when treated as two species. For example, Roselaar (1995) argued that the taxa niethammeri (*=aharonii*) and *persica* should be placed in *A. cheleensis*, whereas Christidis (2018), Clements et al. (2018) and Gill and Donsker (2019) included these in A. rufescens. Somali Short-toed Lark A. somalica and Sand Lark A. raytal have also, separately or together, been treated as conspecific with Alaudala rufescens (Meinertzhagen, 1951; Peters, 1960); Alström, Barnes, et al. (2013) found A. raytal to be nested within the A. rufescens complex and A. somalica (=*Calandrella athensis*) to be the more deeply diverged sister species to these.

We here study the phylogeny of the genus *Alaudala*, with special focus on the *A. rufescens* complex and *A. raytal* 

(hereafter collectively referred to as the *A. rufescens-raytal* complex), based on one mitochondrial and 11 nuclear loci. Our sampling covers the entire range of the genus and comprises all but two of the unanimously recognized subspecies within the *A. rufescens-raytal* complex. Our results reveal deep divergences among different groups of taxa and suggest that the taxonomy of the genus *Alaudala* needs to be revised.

#### 2 | MATERIALS AND METHODS

### 2.1 | Sampling

For taxonomic names and distributions, we follow Christidis (2018) (Table 1). We obtained 130 samples of *Alaudala*, including 124 samples from 14 of the 15 taxa in the *A. rufescens* complex (*A. r. nicolli* missing) and six samples from two of the three subspecies of *A. raytal* (*A. r. krishnakumars-inhji* missing). In addition, we obtained one *Galerida cristata* and one *Calandrella dukhunensis* to be used as outgroups. We also downloaded two cytochrome *b* sequences of *A. so-malica* from GenBank.

Sixty-one samples were collected during fieldwork and deposited at the Department of Biology and Environmental Sciences (previously Department of Zoology), University of Gothenburg, Sweden (DZUG), Institute of Zoology, Chinese Academy of Sciences, Beijing, China (IOZ) and National University of Mongolia and Mongolian Ornithological Society (Table S1), and 69 samples were borrowed from the ornithological collections at the Burke Museum, University of Washington, Seattle, USA (UWBM); American Museum of Natural History, New York, USA (AMNH); Zoological Museum Berlin, Berlin, Germany (ZMB); Swedish Museum of Natural History, Stockholm, Sweden (NRM); Natural History Museum, University of Copenhagen, Copenhagen, Denmark (ZMUC); Zoological Institute, Almaty, Kazakhstan (ZIA); University of Michigan Museum of Zoology, Ann Arbor, USA (UMMZ); Natural History Museum Wien, Vienna, Austria (NMW); Institute of Zoology, Chinese Academy of Sciences, Beijing, China (IOZ); and Natural History Museum, Tring, UK [NHMUK, formerly BMNH]; Table S1.

Tissue samples were preserved in 20% dimethylsulphoxide (DMSO) or 95% ethanol. Blood samples were



**FIGURE 1** Breeding distribution (based on de Juana & Suárez, 2019) of the *Alaudala rufescens* complex (pale grey) and *A. raytal* (beige), with samples used in this study indicated by different symbols (see legend). The numbers refer to the localities in Table S1. Note that sample number 1 in the *heinei* clade (*A. c. 'stegmanni'* Jiangsu, China ZMB 36779) was collected in winter south of the breeding range of the complex

t authors
differen
according to
study
this
Ξ.
nded
Ē
ц.
of taxa
istributions
p
' and
Taxonomy
1
$(\mathbf{T})$
Г
В
-

TABLE 1	Taxonomy and	distributions of taxa ii	ncluded in this stu	dy according to differ	rent authors					
Christidis (2018	0		Peters (1960)		del Hoyo and Collar ( (2019)	2016), de Juana and Suárez	Clements et al. (2018)		Gill and Donsker (20)	(6)
Species	Subspecies	Distribution	Species/subspecies	Distribution	Species/subspecies	Distribution	Species/subspecies	Distribution	Species/subspecies	Distribution
A. raytal (Blyth, 1845)	raytal (Blyth, 1845)	S Nepal, NC and NE India, W Bangladesh, C and S Myanmar	C. raytal raytal	N India from Rajputana and the United Provinces to Bengal, and Burma south to Thayetmyo.	A. raytal raytal	N India (E from Haryana to Arunachal Pradesh), Nepal, Bangladesh and C & S Myanmar	A. raytal raytal	Coastal SE Iran to Afghanistan, N India and S Myanmar	A. raytal raytal	NC India to S Myanmar
	krishnakumarsinhji (Vaurie & Dharmakumarsinhji, 1954)	W India (Kathiawar Pen. in Gujarat)	C. raytal krishnakumarsinhji	Bhavnagar, Saurashtra, India	A. raytal krishnakumarsinlıji	Bhavnagar, in S Gujarat (W India)	A. raytal krishnakumarsinhji	NW India (Kathiawar Peninsula)	A. raytal krishnakumarsinhji	WC India
	adamsi (A. O. Hume, 1871)	S Iran, E Afghanistan, Pakistan, NW India	C. raytal adamsi	Resident along the Makran coast (W to Bandar Abbas) and in the Punjab, Sind, and Cutch	A. raytal adamsi	SE Iran, Pakistan and NW India (Punjab, Haryana, NW Gujarat)	A. raytal adamsi	Coastal NW India	A. raytal adamsi	SE Iran, Pakistan, E Afghanistan and NW India
A. rufescens (Vieillot,	rufescens (Vieillot, 1819)	Tenerife (Canary Is.)	C. r. rufescens	Laguna, Tenerife (Canary Is.)	A. r. rufescens	Tenerife, WC Canary Is.	A. r. rufescens	Tenerife (Canary Is.)	A. r. rufescens	Tenerife (WC Canary Is.)
1819)	minor (Cabanis, 1851)	Morocco and N Algeria to NW Egypt; Sinai to SW Turkey, SW Iran and N Arabian	C. r. minor	N Sahara from Morocco to Egypt: Sinai and the Syrian Desert	A. r. minor	N Africa (from S Morocco E to NW Egyp) and from S Turkey S to N Sinai Peninsula and E to C Iraq	A. r. minor	North Africa to N Sinai and S Turkey and E Iraq	A. r. minor	Morocco to NW Egypt, S Turkey to the Sinai Pen. and E Iraq
	apetzii (A.E. Brehm, 1857)	S Portugal, S Spain	C. r. apetzii	S Spain	A. r. apetzii	S Portugal, S and E Spain	A. r. apetzii	S Iberian Peninsula	A. r. apetzii	E and S Iberian Pen.
	polatzeki (E. J. O. Hartert, 1904)	Gran Canaria, Fuerteventura, Lanzarote (Canary Is.)	C. r. polatzeki	Gran Canaria, Fuerteventura, Lanzarote (Canary Is:)	A. r. polatzeki	E Canary Is (Gran Canaria, Fuerteventura, Lanzarote)	A. r. polatzeki	E Canary Is. (Gran Canaria, Fuerteventura and Lanzarote)	A. r. polatzeki	E Canary Is.
	<i>heinei</i> (Homeyer, 1873)	S Ukraine to E Kazakhtan and NW Altai Mts.	C. r. heinei	SE Russia, N to the Volga region, S to the Aral Sea and the valley of the Syr Darya, recorded in winter from Asia Minor, Palestine, Iraq, Iran, and N Egypt	A. r. heinei	Ukraine E to E Kazakhstan (Semipalatinsk region), S to Aral Sea and Syr Darya Valley; probably winters in the Middle East and SW Asia	A.r. heinei	SE Russia N to Volga basin and SW Siberia	A. r. heinei	Ukraine to E Kazakhstan
	persica Sharpe, 1890	C and E Iran to W and S Afghanistan to Pakistan	C. r. persica	S Iraq, Iran, SE Afghanistan; possibly Bahawalpur	A. r. persica	E & S Iraq, Iran (except NW) and W & SE Afghanistan; some populations winter in lowlands of Pakistan and W India	A. r. persica	S Iraq and Iran to SE Afghanistan	A. r. persica	E and S Iraq to S and SW Afghanistan
	<i>nicolli</i> (E. J. O. Hartert, 1909)	Nile delta (Egypt)	C. r. nicolli	Northern parts of the Nile delta (Egypt)	A. r. nicolli	Nile Delta (N Egypt)	A. r. nicolli	N Egypt (northern Nile Delta)	A. r. nicolli	Nile Delta (N Egypt)
										(Continues)

(6)	Distribution	C Turkey	MN	E Turkey, Transcaucasia and N Iran	SC Siberia, NE Mongolia and NE China	NW China	S Mongolia and NC China	MN	Kazakhstan to Turkmenistan	WC China	NW Mongolia and S Russia	MN
Gill and Donsker (201	Species/subspecies	A. r. aharonii	synonym of A. r. aharonii	A. r. pseudobaetica	A. c. cheleensis	A. C. xeebohmi	A. c. beicki	NM	A. c. leucophaea	A. c. kukunoorensis	A. c. tuvinica	MN
r (2016), de Juana and Suárez Clements et al. (2018)	Distribution	1	C Turkey (Anatolia high plateau)	Highlands of Armenia and SW shores of Caspian Sea	E Transbaikalia and N Mongolia to NE China	NW Mongolia and W China (Xinjiang E to Khotan River)	W China (S Gobi Desert to NW Gansu)	Arid W China (NE Gansu)	Turkestan (S of Kyzyl Kum Desert)	SW Mongolia (Lake Kokonor basin)	MN	Mts. of W China (S Kokonor) to NE Tibet
	Species/subspecies	synonym of A. r. persica	A. r. niethammeri	A. r. pseudobaetica	A.c. cheleensis	A. c. seebohmi	A. c. beicki	A. c. stegmanni	A. c. leucophaca	A. c. kukunoorensis	synonym of A. c. seebohmi	A. c. tangutica
	Distribution	C Turkey (Anatolian High Plateau); winters S to C Syria	NM	E Turkey, Armenia, Azerbaijan and NW Iran	Transbaikalia, NE Mongolia and NE China (S Heilongjiang and Inner Mongolia S to S Shaanxi and N Jiangsu); N populations migrate farther S	S Russia (Tuva), W Mongolia and extreme Wor China (Xinjiang, from Tien Shan S to Kashgaria, E to R Hotan); N populations migrate farther S	SW Mongolia and bordering areas of N China (Gansu, Ningxia)	NC China (NW Gansu)	Turkmenistan, Uzbekistan (S of Kyzylkum Desert) and W & S Kazakhstan E to L Balkhash and Fergana Basin; winters in SW Asia	WC China (Qaidam Basin, in N Qinghai)	MN	S Qinghai and NE Tibet
del Hoyo and Collar ( (2019)	Species/subspecies	A. r. aharonii	synonym of A. r. aharonii	A. r. pseudobaetica	A. r. cheleensis	A. r. seeboltmi	A. r. beicki	A. r. stegmanni	A. r. leucophaea	A. r. kukunoorensis	synonym of <i>A. r.</i> seebohmi but mentioned that it may be a valid race	A. r. tangutica
Peters (1960)	Distribution	Asia Minor (Konya); winters in the Syrian Desert (Karyatein) and Jordan	WN	Highlands of Armenia and SW Caspian Sea	Eastern Transbaikalia and N Mongolia from Ulan Bator E to Manchuria, and NE China in provinces of Hopeh, Shantung, and Kiangsu	NW Mongolia, and Sinkiang E at least to the Khotan River.	S Gobi south to W Kansu and W Ningsia along the course of the Edsin Gol; Koko Nor.	MN	Russian Turkestan, S Kyzyl Kum; on migration or in winter to Asia Minor (accidental?) and E Iran	SW Mongolia S to Tsaidam	MN	
	Species/subspecies	C. r. aharonii	MN	C. r. pseudobaetica	C. cheleensis	C. r. seebohmi	C. r. beicki	synonym of C. r. beicki	C. r. leucophaea	C. r. kukunoorensis	MN	synonym of <i>C. r.</i> kukunoorensis
	Distribution	C Asia Minor	MN	E Turkey, Transcaucasia, NW and N Iran	Transbaikalia, C and E Mongolia, NE to C and E China	Tarim Basin (NW China)	S Mongolia, N China (NE Qinghai to N Gansu and W Nei Mongol)	MN	W and S Kazakhstan to NW China (NW Xinjiang)	SW Mongolia to E Xizang and N Qinghai	Tuva Republic and NW Mongolia	MN
	Subspecies	aharonii (E. J. O. Hartert, 1910)	niethammeri* Kumerloeve, 1963	pseudobaetica (Stegmann, 1932)	cheleensis Swinhoe, 1871	seebolmi Sharpe, 1890	beicki (Meise, 1933)	<i>stegmanni</i> * (Meise, 1937)	leucophaea (Severtsov,1873)	kukunoorensis Przevalski, 1876	<i>tuvinica</i> Stepanyan,1975	tangutica*(E. J. O. Hartert & F. Steinbacher, 1933)
Christidis (2018)	Species				A. cheleensis Swinhoe, 1871							

Note: "This taxon was not mentioned by Christidis (2018). Taxa not mentioned or postdated by a particular study are marked "NM".

TABLE 1 (Continued)

mixed immediately in 95% ethanol or a blood storage buffer (0.1 M Tris–HCl, 0.04 M EDTA.Na2 or 1.0 M NaCl, 0.5% SDS). Additionally, 22 toepad samples from museum specimens were obtained for this study. See Tables S1 and S2 for details of sampling localities, institutions and GenBank accession numbers and Figure 1 for sampling localities.

# 2.2 | DNA extraction

Total genomic DNA extraction was conducted using the DNeasy Blood and Tissue Kit (Qiagen) following the recommended protocol for tissue and blood samples. Toepads from old museum samples were extracted basically according to the protocol based on Irestedt, Ohlson, Zuccon, Källersjö, and Ericson (2006). QIAamp DNA Micro Kit (Qiagen) was used following the manufacturer's recommendations, with a few modifications. During lysis, 20  $\mu$ l DTT (1 M) was added, the lysis was prolonged (18–24 hr), and extra proteinase K (10  $\mu$ l) was added once. The amount of elution buffer during the final stage of the extraction was decreased to give a total volume of 80–100  $\mu$ l extract. Extractions from toepads were always performed in a room dedicated to working with old degraded DNA material, with appropriate facilities such as a UV-bench used for sterilizing equipment.

## 2.3 | PCR amplification and sequencing

### 2.3.1 | Mitochondrial DNA

For all fresh samples, the mitochondrial cytochrome b gene (cyt b) was amplified and sequenced using the primers and protocol described in Olsson, Alström, Ericson, and Sundberg (2005). The whole cyt b sequence (1,140 bp) was acquired in a single PCR and sequenced using three primers (Table S2). For the toepad samples, cyt b was amplified in short fragments (Irestedt et al., 2006) in a touchdown PCR protocol of 1 min at 95°C, three cycles each consisting of 10 s 95°C, 10 s 60°C, 15 s 72°C and then three cycles of 10 s 95°C, 10 s 58°C, 15 s 72°C, followed by 32 cycles of 10 s 95°C, 10 s 56°C, 15 s 72°C, with a final elongation step at 30 s 72°C, using specifically designed primers (Table S3). Purification of PCR product was accomplished using 0.5 µl ExoTAP (Exonuclease I and FastAP Thermosensitive Alkaline Phosphatase; Werle, Schneider, Renner, Volker, & Fiehn, 1994). Cytochrome b was sequenced for all 130 samples. All sequences have been deposited in GenBank (Table S2).

#### 2.3.2 | Nuclear DNA

Based on the cyt b tree (see Section 3), 3–8 specimens from each of the main clades of the A. rufescens-raytal

complex, as well as G. cristata and C. dukhunensis (outgroups) were selected to be analysed using 11 nuclear loci (Table S1). These nuclear loci were amplified and sequenced following the protocols described in Olsson et al. (2005) for myoglobin intron 2 (myo); Allen and Omland (2003) for ornithine decarboxylase intron 7 and exon 8 (partial; ODC); Fjeldså, Zuccon, Irestedt, Johansson, and Ericson (2003) for glyceraldehyde 3-phosphate dehydrogenase intron 11 (GAPDH); Primmer, Borge, Lindell, and Sætre (2002) for transforming growth factor beta intron 5 (TGFB2); Backström et al. (2008) for ATP6AP2 intron 7, PRADC1 intron 3, ornithine decarboxylase antizyme inhibitor (OAZI) intron 3 and POLDIP2 intron 5; Stervander (2015) for WDR12 intron 9; Friesen, Congdon, Kidd, and Birt (1999) for lactate dehydrogenase B intron 3 (LDHB); and Lovette and Rubenstein (2007) for β-fibrinogen intron 5 (Fib5; see Table S4 for details). Re-amplification was applied if necessary, with 0.5 µl template DNA. PCR was performed on a Master gradient thermocycler (Eppendorf) in 12.5 µl. PCR products were purified using ExoTAP and sequenced using PCR primers. All sequences have been deposited in GenBank (Table S2).

# 2.4 | Phylogenetic analyses of mitochondrial DNA

Sequences were aligned and assembled using MegAlign 4.03 in the DNASTAR package (DNAstar, Inc.). We performed phylogenetic analysis using 130 complete cyt b sequences from the A. rufescens-raytal complex, with and without outgroup (Galerida cristata, Calandrella dukhunensis). The best-fit models were determined based on the Bayesian information criterion (BIC) implemented in jModelTest (Darriba, Taboada, Doallo, & Posada, 2012; Table S4). Molecular phylogenetic trees were estimated using the Bayesian inference (BI) in BEAST 1.8.4 (Drummond, Suchard, Xie, & Rambaut, 2012). Input files were prepared in BEAUti version 1.8.4 (Drummond et al., 2012). Posterior probabilities (PP) were calculated under the GTR+G model and a relaxed log-normal clock. The "birth-death incomplete sampling" tree prior was selected, and default prior distributions were chosen for other parameters. The analysis was run with 500 million generations and sampled every 10,000 trees. Convergence to the stationary distribution of the single chains was inspected in Tracer v.1.6.0 (Rambaut, Suchard, Xie, & Drummond, 2014) using a minimum threshold for the effective sample size (>200). The joint likelihood and other parameter values reported large effective sample sizes (>1,000), and the trace plot had the shape of a "dense, straight, furry caterpillar". Good mixing of the MCMC and reproducibility were established by multiple runs from independent starting points. Topological convergence was examined by eye.

The first 25% sampled trees were discarded as burn-in, well after stationarity had been reached, and the posterior probabilities were calculated from the remaining samples. Trees were summarized using TreeAnnotator v.2.2.1 (Rambaut & Drummond, 2015), choosing 'Maximum clade credibility tree' and 'Mean heights' and displayed in FigTree 1.4.3 (Rambaut, 2016).

A maximum-likelihood tree was reconstructed using RAXML v.7.4.2 (Stamatakis, 2014) under the same model as in the BEAST analysis. One thousand bootstrap replicates, random starting trees were used; all other parameters were by default. The analysis was run on the CIPRES Science Gateway (Miller, Pfeiffer, & Schwartz, 2010).

## 2.5 | Phylogenetic analyses of nuclear loci

Sequences were aligned and assembled using MegAlign 4.03 in the DNASTAR package (DNAstar, Inc.); some manual adjustment was done where the program obviously mis-aligned some sequences in an indel. Heterozygous sites were coded as ambiguous. The best-fit model was selected for each locus as for the mitochondrial locus (Table S4). The data were analysed by Bayesian inference (BI) in BEAST, separately (single-locus analyses) and concatenated, partitioned by locus (11 partitions). Default priors were used, and 100 million generations were run and sampled every 10,000 trees. The MCMC output was analysed in Tracer; the first 25% sampled trees were discarded as burn-in; and trees were summarized using TreeAnnotator and displayed in FigTree (see above for details). A maximum-like-lihood tree was reconstructed with RAXML as for cyt *b* (see above), but divided into 11 locus-specific partitions.

# 2.6 | Phylogenetic analyses of concatenated cyt *b* and nuclear sequences

All loci (cyt b + 11 nuclear loci) for 31 samples from the *A*. *rufescens–raytal* complex plus one *G*. *cristata* and one *C*. *dukhunensis* were concatenated and analysed in 12 locus-specific partitions in BEAST under a relaxed log-normal clock and a "birth-death incomplete sampling" prior. The Markov chain Monte Carlo (MCMC) was run for 100 million generations. Convergence was checked, and trees were summarized and displayed as in the previous analyses (see above). A maximum-likelihood tree was reconstructed with RAXML as for the analysis of concatenated nuclear loci (see above).

#### 2.7 | Dating

In order to evaluate divergence times, we applied the cyt b molecular clock rate 2.1% divergence per million years (my)

based on Weir and Schluter (2008) (mean rate of 0.0105 substitutions/site/lineage/my and normally distributed clock prior with standard deviation 0.001). We analysed two datasets in BEAST 1.8.4: (1) cyt b on its own and (2) the 12-locus concatenated, locus-partitioned dataset. In the latter analysis, we set the rate on cyt b and allowed the other loci to be estimated. The same models as in the previous analyses were used (see above). We ran several analyses, with and without outgroup, and using different models (relaxed log-normal and strict clock). The posteriors from these analyses were compared (not shown) to select the one with the highest posterior for the final analysis. The final analysis was run under an uncorrelated log-normal relaxed clock model (Drummond, Ho, Phillips, & Rambaut, 2006) and a "birth-death incomplete sampling" tree prior. See above for further details. Xml files for all analyses and tree file in Newick format for the trees in Figures 2 and 3 are available as Appendix A.

#### 2.8 | Species tree reconstruction

A species tree was also estimated for 31 samples from the *A. rufescens–raytal* complex based on the combined mitochondrial and nuclear markers in \*BEAST (Heled & Drummond, 2010) under the birth-death model and a relaxed log-normal clock, and the same substitution models as in the previous analyses (see above). Deeply diverged clades found in the cyt *b* and concatenation trees were set as independent populations in \*BEAST, and *G. cristata* and *C. dukhunensis* were selected as outgroups. Convergence was checked and trees were summarized and displayed as in the previous analyses (see above). All data generated or analysed during this study are included in this published article (and its supplementary files and Appendix A).

# 3 | RESULTS

### 3.1 | Analysis of cytochrome *b*

The BEAST and RAxML trees are in good agreement. The BEAST tree containing 130 sequences from the *A. rufescens-raytal* complex is shown in Figure 2 (and the RAxML trees in Figures S1a,b). It shows five well-supported, geographically coherent clades (posterior probability [PP] 1.00, ML bootstrap [MLBS] 100%): A (the 'heinei clade'), comprising *A. rufescens heinei*, *A. r. aharonii*, *A. r. psedobaetica*, *A. r. persica*, *A. cheleensis beicki*, and *A. c. 'stegmanni'* (as originally labelled in ZMB), from Turkey, Iran, Afghanistan, south-east Kazakhstan and south central Mongolia, with a single geographical 'outlier' from coastal east China (*A. c. 'stegmanni'*, collected in December, south of the breeding

range); B (the '*raytal* clade'), A. *raytal raytal* (including population sometimes treated as A. r. vauriei) and A. r. adamsi, from India; C (the '*rufescens* clade'), A. *rufescens rufescens*, A. r. polatzeki, A. r. apetzii and A. r. minor, from the Canary Islands to Saudi Arabia; D (the '*cheleensis* clade'), A. *cheleensis cheleensis*, A. c. *tuvinica* and A. c. *beicki*, from northern Mongolia and neighbouring parts of Russia, and the eastern half of China, and E (the '*leucophaea* clade'), A. *cheleensis leucophaea*, A. c. *seebohmi* and A. c. *kukunoorensis*, from south Kazakhstan and western half of China. Only one taxon, A. *rufescens beicki*, was found in more than one of these clades (A and D). Clade A overlapped geographically with both clades D and E, whereas the other clades were allopatric.

The cyt *b* tree is split into two main clades, one comprising clades A–C (PP 0.95, MLBS 91%) and the other one comprising clades D–E (PP 0.97, MLBS 100%). Clades A and B are sisters with insufficient PP (0.90) but high MLBS (90%). There is little structure within the individual clades A–E. However, within clade B, the single sample of *A. r. adamsi* from west Punjab, India, is deeply diverged from the four samples of *A. a. raytal* from Delhi to Assam, India (PP 1.00, MLBS 100%; see below), and within clade E, the four samples of *A. c. leucophaea* from Kazakhstan form a clade (PP 0.85, MLBS 89%) separate from the samples from China (PP 1.00, MLBS 94%).

The two samples of *A. somalica* form a deeply diverged sister clade to the *A. rufescens-raytal* complex (Figures S1a,b).

#### 3.2 | Analyses of 11 nuclear loci

Single-locus analyses of the nuclear loci are poorly supported, and none of them fully recover clades A–E (Figure S2). However, the tree based on the concatenated 11 nuclear loci recovers clades A, B, C and D/E with strong support (PP 1.00, MLBS 100%); clades D and E are unresolved (Figure S3). Moreover, clades A and C are sisters, with low support (PP 0.82, MLBS 52%).

# 3.3 | Analyses of concatenated cytochrome *b* and 11 nuclear loci

In the BEAST analysis of the combined dataset comprising mitochondrial cyt *b* and 11 nuclear markers for 31 samples from the *Alaudala rufescens–raytal* complex, five well-supported (PP  $\ge 0.97$ , MLBS  $\ge 78\%$ ) primary clades (A–E) were

recovered (Figure 3a). These clades correspond to the five primary clades in the cyt *b* tree (Figure 2). However, in the multilocus tree, clades A and C are sisters with strong support (PP 0.98, MLBS 83%), whereas in the cyt *b* tree clades A and B are sisters with lower support (PP 0.90, MLBS 90%).

#### **3.4** | Multilocus species tree

The species tree based on the 12-locus dataset for 31 samples from the *Alaudala rufescens–raytal* complex (Figure 3b) is topologically identical to the BEAST tree based on the same concatenated dataset, although the support values are lower on average in the species tree than in the concatenation tree.

#### **3.5** | Divergence times

Divergence time estimates vary among analyses. In the analysis of the cyt *b* data (Figure 2), they ranged from 3.18 million years ago (mya; 95% highest posterior density [HPD] 2.01–4.49 mya) for the deepest split (between clades A/B/C and D/E) to 1.77 mya (95% HPD 1.03–2.60 mya) for the shallowest split among the five primary clades (between clades A and B). In the analysis based on the combined data (Figure 3), the deepest split was estimated at 2.64 mya (95% HPD 1.17–4.3 mya), whereas the shallowest split among the primary clades (between clades D and E in this case) was 1.56 mya (95% HPD 0.68–2.53 mya).

#### 4 | DISCUSSION

Larks have long presented taxonomic challenges. Strong selection for crypsis in the diverse open habitats where larks occur has resulted in poor plumage differentiation across taxa, as well as pronounced local and individual variation within taxa (Alström, Barnes, et al., 2013; Alström, Olsson, et al., 2013; Donald et al., 2017; Meinertzhagen, 1951; Vaurie, 1959). Moreover, most taxonomic work has been based exclusively on morphology, with only a small number of studies utilizing vocal, behavioural and genetic data (review in Alström, Olsson, & Lei, 2013; Drovetski, Rakovic, Semenov, Fadeev, & Red'kin, 2014; Ghorbani et al., 2020; Stervander et al., 2016). Prior molecular work based on a small number of taxa and loci has suggested that A. raytal is nested within the A. rufescens complex (Alström, Barnes, et al., 2013). Our study, based on mitochondrial and nuclear DNA combined with dense sampling across the range of the

**FIGURE 2** Cytochrome *b* tree for the *Alaudala rufescens–raytal* complex, with a timescale based on a molecular clock rate of 2.1%/million years, estimated in BEAST. Blue node bars show the 95% confidence intervals for the divergence times. Values at the nodes represent, from left to right, Bayesian posterior probabilities and maximum-likelihood bootstrap values





**FIGURE 3** (a) Phylogenetic tree for the *Alaudala rufescens–raytal* complex based on concatenated mitochondrial cytochrome *b* and 11 nuclear markers, estimated in BEAST. (b) \*BEAST species tree based on the same data as in (a). Blue node bars show the 95% confidence intervals for the divergence times. Values at the nodes represent, from left to right, Bayesian posterior probabilities and maximum-likelihood bootstrap values

A. rufescens-raytal complex, substantially improves our knowledge of this complex. Our analyses of all loci combined, as well as of cyt b on its own, strongly support five primary clades: the heinei clade (A), the raytal clade (B), the rufescens clade (C), the cheleensis clade (D) and the *leucophaea* clade (E). The cyt b and concatenated 12-locus trees support a deep split between the cheleensis and leu*cophaea* clades (cyt *b*: 2.14 mya, 95% HPD 1.07–3.26 mya; 12-loci: 1.56 mya, 95% HPD 0.68-2.53 mya). In contrast, the cheleensis (D) and leucophaea (E) clades are incompletely sorted in the analysis of the concatenated nuclear loci, although they form a strongly supported combined clade (DE). A close relationship between *cheleensis* and *leucophaea* has been suggested before based on similarities in wing formula (Stepanyan, 1967). The relationships among the heinei, raytal and rufescens clades differ between the analyses of cyt b and all of the multilocus analyses, although they are best supported in the multilocus analyses, where the heinei and rufescens clades are sister clades and the raytal clade is in a sister position to these. All of the analyses confirm that A. raytal is nested within the A. rufescens complex. With the exception of Meinertzhagen (1951), all subsequent authors have treated A. raytal as a separate species from the A. rufescens complex.

This reflects the faster rate of divergence in size, structure and plumage of *A. rytal* compared to the other taxa (Alström, 2020; Ganpule, 2019), which in turn is likely to be the result of adaptation to its unique habitat, sandy river banks and sandy sea coasts (Alström, 2020).

The two species, A. rufescens and A. cheleensis as circumscribed by various authors (e.g., Dickinson & Dekker, 2001; Gill & Donsker, 2019; Peters, 1960; Roselaar, 1995; Stepanyan, 1967), are not monophyletic, and accordingly, none of these different taxonomic proposals are supported by our data. Although none of our samples from different primary clades were collected in sympatry from exactly the same localities, samples from the heinei clade and the cheleensis clade were collected during the breeding season in Ömnögovi in south Mongolia only c. 70 km apart (both A. c. beicki based on distributions as given in, e.g., Christidis, 2018), and these birds were found to differ in multiple non-genetic aspects, such as morphology, vocalizations, sexual display and habitat choice (P. Alström, G. Sundev, unpublished), strongly indicating that they belong to separate species. Moreover, our samples from the heinei clade from Kazakhstan are sampled within the broad distribution of the leucophaea clade, supporting previous statements that these

taxa belong to different species, based on sympatry and differences in plumage, structure, habitat choice, breeding period, number of broods and migratory behaviour (Bianchi, 1905, 1906; Korelov, 1958; Stachanow & Spangenberg, 1931; Stepanyan, 1967, 1983, 1990).

Only four of the currently recognized subspecies are recovered as monophyletic in the cyt b analysis of 130 samples: A. rufescens aharonii (though only two samples, from the same locality), A. raytal raytal (five samples), A. rufescens rufescens (only two samples with identical haplotypes) and A. c. leucophaea (four samples). The other subspecies, that is A. r. apetzii, A. r. polatzeki, A. r. minor, A. r. heinei, A. r. pseudobaetica, A. r. persica, A. c. cheleensis, A. c. tuvinica, A. c. beicki, A. c. seebohmi and A. c. kukunoorensis, are not reciprocally monophyletic (not applicable to A. r. adamsi, for which only a single sample is available). Alaudala c. beicki is in particularly poor agreement with the current taxonomy, as samples ascribed to this taxon based on distributional data in the literature (Table 1) are placed in both the cheleensis and heinei clades. As noted above, the samples of A. c. beicki from south Mongolia apparently refer to two different species. It is hardly surprising that this has been overlooked considering the close morphological similarity between these.

The A. rufescens-raytal complex is a striking example of a group of birds where morphological conservatism has concealed deep genetic divergences, in other words, a group containing cryptic diversity. There are several other examples of passerine birds with partly similar distributions as the A. rufescens-raytal complex that have been found through molecular analyses to contain unexpectedly deep divergences, for example the Great Grey Shrike Lanius excubitor complex (Olsson, Alström, Svensson, Aliabadian, & Sundberg, 2009), the Lesser Whitethroat Sylvia curruca complex (Olsson et al., 2013), the Pale Martin Riparia diluta complex (Schweizer et al., 2018) and the Greater Short-toed Lark Calandrella brachydactyla complex (Alström, Barnes, et al., 2013; Stervander et al., 2016). Within the latter complex, the easternmost taxon C. b. dukhunensis from east Mongolia and neighbouring areas has been revealed to be more closely related to Hume's Lark C. acutirostris than to the other taxa in the C. brachydactyla complex and separated from the western taxa in the C. brachydactyla complex by 6.0 mya (95% HPD 4.6-7.5 mya; Alström, Barnes, et al., 2013; Stervander et al., 2016). Divergence times are considerably younger in the Riparia diluta complex (maximum 1.7 mya; 95% HPD 0.69-3.06 mya), whereas no time estimates are yet available for the Lanius excubitor or Sylvia curruca complexes.

Although our study suggests that the taxonomy of the *A*. *rufescens–raytal* complex is in need of revision, we refrain from doing this at this stage, as we see the need for a revision that also includes additional data, such as vocalizations, sexual behaviours and ecology. However, based on our molecular Zoologica Scripta 🚳 🕅 🐼 — WILEY

data, it seems likely that the *A. rufescens-raytal* complex should be separated into at least four species, representing the *heinei* clade (*A. heinei* by priority), the *raytal* clade (*A. raytal*), the *rufescens* clade (*A. rufescens* sensu *stricto*) and the *cheleensis* + *leucophaea* clade (*A. cheleensis* by priority). The mitochondrial data also suggest recognition of the *leucophaea* clade as a separate species (*A. leucophaea* by priority), although this is not supported by the nuclear data. The deep cyt *b* divergence between *A. r. raytal* and *A. r. adamsi* requires further study (only one sample of the latter taxon).

The cyt *b* and 12-locus trees differ somewhat with respect to estimated divergence times. For example, the age of the deepest split (between clades A/B/C and D/E) was estimated to be 0.54 my older in the cyt *b* tree than in the 12-locus tree. However, both estimates have wide confidence intervals, which are widely overlapping (2.01–4.49 mya and 1.17–4.3 mya, respectively).

### 5 | CONCLUSIONS

The Lesser Short-toed Lark *Alaudala rufescens*–Sand Lark *A. raytal* complex was found to comprise at least four distinct lineages separated c. 1.6–3.2 million years ago (95% highest posterior density c. 1.0–4.5 mya). These deep divergences have been masked by the slight plumage differentiation among the taxa in this complex. The *A. rufescens–raytal* complex should probably be treated as four or, possibly, five species, but independent data are wanted before any formal taxonomic revision is undertaken.

#### ACKNOWLEDGEMENTS

This study was financially supported by the Faculty of Science, Ferdowsi University of Mashhad (FUM; grant No. 3/29628 to FG). P.A. gratefully acknowledges support from Jornvall Foundation, Mark and Mo Constantine and the Swedish Research Council (grant No. 2015-04402); U.O. and M.I. were also supported by the Swedish Research Council (No. 2015-04651 and 621-2014-5113, respectively). The first author is indebted to Svante Martinsson and Mårten Klint for their kind help in the laboratory. For loan of samples, we are indebted to Sharon Birks/Burke Museum, University of Washington, Paul Sweet and Thomas J. Trombone/American Museum of Natural History, Sylke Frahnert and Pascal Eckhoff/Zoological Museum Berlin, Ulf Johansson and Peter Nilsson/Swedish Museum of Natural History, Jon Fjeldså and Jan Bolding Kristensen/ Natural History Museum, University of Copenhagen, the late Edward Gavrilov/Zoological Institute, Almaty, Janet Hinshaw/University of Michigan Museum of Zoology, the late Anita Gamauf/Natural History Museum Wien, Peng He/ Institute of Zoology, Chinese Academy of Sciences, Mark Adams and Hein van Grouw/The Natural History Museum, 12

WILEY—Zoologica Scripta

Tring, Oleg Belyalov, José Antonio Cañizares Mata, Pierre-André Crochet, Alban Guillaumet, Andrew Lassey, Lars Svensson, William Velmala, and the late Guillermo Delgado Castro. Gesa von Hirschheydt and Eduardo Ottimofiore assisted with analyses of distributional data. We are grateful to the National University of Mongolia and Mongolian Ornithological Society for their encouragement and support of all necessary permits and field sampling in Mongolia. Two anonymous reviewers are acknowledged for their valuable comments.

#### ORCID

Fatemeh Ghorbani D https://orcid.org/0000-0001-9091-9632 Mansour Aliabadian D https://orcid.

org/0000-0002-3200-4853

*Martin Irestedt* https://orcid.org/0000-0003-1680-6861 *Fumin Lei* https://orcid.org/0000-0001-9920-8167 *Ming Ma* https://orcid.org/0000-0003-3080-0562 *Urban Olsson* https://orcid.org/0000-0003-1435-5411 *Per Alström* https://orcid.org/0000-0001-7182-2763

#### REFERENCES

- Allen, E. S., & Omland, K. E. (2003). Novel intron phylogeny supports plumage convergence in orioles (*Icterus*). *The Auk*, *120*, 961–969. https://doi.org/10.1093/auk/120.4.961
- Alström, P. (2020). Sand Lark (*Alaudala raytal*). In J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, & E. de Juana (Eds.), *Handbook of the birds of the world alive*. Barcelona, Spain: Lynx Edicions. Retrieved from https://www.hbw.com/node/57663
- Alström, P., Barnes, K. N., Olsson, U., Barker, F. K., Bloomer, P., Khan, A. A., ... Ryan, P. G. (2013). Multilocus phylogeny of the avian family Alaudidae (larks) reveals complex morphological evolution, non-monophyletic genera and hidden species diversity. *Molecular Phylogenetics Evolution*, 69, 1043–1056. https://doi.org/10.1016/j. ympev.2013.06.005
- Alström, P., Olsson, U., & Lei, F. (2013). A review of the recent advances in the systematics of the avian superfamily Sylvioidea. *Chinese Birds*, 4, 99–131. https://doi.org/10.5122/cbirds.2013.0016
- Backström, N., Karaiskou, N., Leder, E. H., Gustafsson, L., Primmer, C. R., Qvarnström, A., & Ellegren, H. (2008). A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of avian evolution. *Genetics*, 179, 1479–1495. https://doi. org/10.1534/genetics.108.088195
- Bianchi, V. L. (1905). Birds. Scientific results of N.M. Przewalski' travel in Central Asia. Edition of the Academy of Sciences, Zoological Section. Saint Petersburg, Part II (4), pp. 193–360. [In Russian].
- Bianchi, V. L. (1906). Additional notes on Palaearctic larks (Alaudidae). *Izvestiya Akademii Nauk, Saint Petersburg*, 23, 205–240. [In Russian].
- Christidis, L. (2018). The Howard and Moore complete checklist of the birds of the world, version 4.1 (Downloadable checklist). Retrieved from https://www.howardandmoore.org
- Clements, J. F., Schulenberg, T. S., Iliff, M. J., Roberson, D., Fredericks, T. A., Sullivan, B. L., & Wood, C. L. (2018). *The eBird/Clements*

checklist of birds of the world: v2018. Retrieved from http://www. birds.cornell.edu/clementschecklist/download/

- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, 9, 772. https://doi.org/10.1038/nmeth.2109
- de Juana, E., & Suárez, F. (2019). Lesser Short-toed Lark (Alaudala rufescens). In J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, & E. de Juana (Eds.), *Handbook of the birds of the world alive*. Barcelona, Spain: Lynx Edicions. Retrieved from https://www.hbw. com/node/57662
- del Hoyo, J., & Collar, N. J. (2016). HBW and BirdLife International illustrated checklist of the birds of the world. Volume 2: Passerines. Barcelona, Spain: Lynx Edicions.
- Dementiev, G. P., & Gladkov, N. A. (1954). Birds of the Soviet Union (vol. 5, pp. 1–803). Moscow, Russia: Soviet Science [in Russian].
- Dickinson, E. C., & Dekker, R. W. R. J. (2001). Systematic notes on Asian birds. 11. A preliminary review of the Alaudidae. *Zoologische Verhandelingen Leiden*, 335, 61–84.
- Donald, P., Alström, P., & Engelbrecht, D. (2017). Possible mechanisms of substrate colour-matching in larks (Alaudidae) and their taxonomic implications. *Ibis*, 159, 699–702. https://doi.org/10.1111/ ibi.12487
- Drovetski, S. V., Rakovic, M., Semenov, G., Fadeev, I. V., & Red'kin, Y. A. (2014). Limited Phylogeographic signal in sex-linked and autosomal loci despite geographically, ecologically, and phenotypically concordant structure of mtDNA variation in the Holarctic avian genus *Eremophila*. *PLoS ONE*, *9*, 1–10.
- Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, 4(5), e88. https://doi.org/10.1371/journal.pbio.0040088
- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29, 1969–1973. https://doi.org/10.1093/ molbev/mss075
- Fjeldså, J., Zuccon, D., Irestedt, M., Johansson, U. S., & Ericson, P. G. P. (2003). Sapayoa aenigma: A new world representative of 'Old World suboscines'. Royal Society of London. Series B: Biological Sciences, 270, 238–241. https://doi.org/10.1098/ rsbl.2003.0075
- Friesen, V. L., Congdon, B. C., Kidd, M. G., & Birt, T. P. (1999). Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Molecular Ecology*, 8, 2141– 2152. https://doi.org/10.1046/j.1365-294x.1999.00802-4.x
- Ganpule, P. (2019). Field identification of Sand Lark Alaudala raytal and Lesser/Asian Short-toed Lark Alaudala rufescens/cheleensis: An unacknowledged pitfall. Indian Birds, 15, 97–111.
- Ghorbani, F., Aliabadian, M., Olsson, U., Donald, P. F., Khan, A. A., & Alström, P. (2020). Mitochondrial phylogeography of the genus *Eremophila* confirms underestimated species diversity in the Palearctic. *Journal of Ornithology*, *161*, 297–312. https://doi. org/10.1007/s10336-019-01714-2
- Gill, F., & Donsker, D. (Eds.) (2019). *IOC World Bird List (v 9.2)*. Retrieved from https://www.worldbirdnames.org/; https://doi. org/10.14344/IOC.ML.9.2
- Hartert, E. (1904). Die Vögel der paläarktischen Fauna (vol. 2 (Heft 2), pp. 113–240). Berlin: R. Friedländer & Sohn.
- Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27, 570–580. https://doi.org/10.1093/molbev/msp274

- Irestedt, M., Ohlson, J. I., Zuccon, D., Källersjö, M., & Ericson, P. G. P. (2006). Nuclear DNA from old collections of avian study skins reveals the evolutionary history of the Old World suboscines (Aves, Passeriformes). *Zoologica Scripta*, 35, 567–580. https://doi. org/10.1111/j.1463-6409.2006.00249.x
- Korelov, M. N. (1958). The specific separation of the eastern or Mongolian short-toed lark. Uchen. Zap. Moskovskoi Univ., *Ornitologiya*, 197, 223–230. [In Russian].
- Lovette, I. J., & Rubenstein, D. R. (2007). A comprehensive molecular phylogeny of the starlings (Aves: Sturnidae) and mockingbirds (Aves: Mimidae): Congruent mtDNA and nuclear trees for a cosmopolitan avian radiation. *Molecular Phylogenetics and Evolution*, 44, 1031–1056. https://doi.org/10.1016/j.ympev.2007.03.017
- Meinertzhagen, R. (1951). A review of the Alaudidae. Proceedings of the Zoological Society of London, 121, 81–132.
- Miller, M. A., 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) (pp. 1–8). New Orleans, LA.
- Olsson, U., Alström, P., Ericson, P. G. P., & Sundberg, P. (2005). Nonmonophyletic taxa and cryptic species – evidence from a molecular phylogeny of leaf-warblers (Phylloscopus, Aves). *Molecular Phylogenetics and Evolution*, 36, 261–276. https://doi.org/10.1016/j. ympev.2005.01.012
- Olsson, U., Alström, P., Svensson, L., Aliabadian, M., & Sundberg, P. (2009). The *Lanius excubitor* (Aves, Passeriformes) conundrum – Taxonomic dilemma when molecular and non-molecular data tell different stories. *Molecular Phylogenetics and Evolution*, 55, 347– 357. https://doi.org/10.1016/j.ympev.2009.11.010
- Olsson, U., Leader, P. J., Carey, G. J., Khan, A. A., Svensson, L., & Alström, P. (2013). New insights into the intricate taxonomy and phylogeny of the Sylvia curruca complex. *Molecular Phylogenetics and Evolution*, 67, 72–85. https://doi.org/10.1016/j.ympev.2012.12.023
- Peters, J. L. (1960). Family Alaudidae. In E. Mayr, & J. C. Greenway Jr. (Eds.), *Check-List of the Birds of the World* (vol. 9, pp. 30–80). Cambridge, MA: Museum of Comparative Zoology.
- Primmer, C. R., Borge, T., Lindell, J., & Sætre, G. P. (2002). Singlenucleotide polymorphism characterization in species with limited available sequence information: High nucleotide diversity revealed in the avian genome. *Molecular Ecology*, 11, 603–612. https://doi. org/10.1046/j.0962-1083.2001.01452.x
- Rambaut, A. (2016). *FigTree v1.4.3*. Retrieved from https://tree.bio. ed.ac.uk/software/figtree/
- Rambaut, A., & Drummond, A. J. (2015). *TreeAnnotator. Version 2.2.1*. Retrieved from http://beast.bio.ed.ac.uk/
- Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). *Tracer v1.7*. Retrieved from http://beast.bio.ed.ac.uk/Tracer
- Roselaar, C. S. (1995). Taxonomy, morphology and distribution of the songbirds of Turkey: An atlas of biodiversity of Turkish passerine birds. London, UK: Pica Press.
- Schweizer, M., Liu, Y., Olsson, U., Shirihai, H., Huang, Q., Leader, P. J., ... Svensson, L. (2018). Contrasting patterns of diversification in two sister species of martins (Aves: Hirundinidae): The Sand Martin *Riparia riparia* and the Pale Martin *R. diluta. Molecular Phylogenetics and Evolution*, *125*, 116–126.

- Sharpe, R. B. (1890). Catalogue of the birds in the British Museum (pp. 1–702). London, UK: British Museum.
- Stachanow, W. S., & Spangenberg, E. P. (1931). Über einige Momente der Bildung der Vögelspezies. 1. Über die Speziesbildung bei *Pseudalaudula* Bianchi 1905. Kócsag, 4, 59–65.
- Stamatakis, A. (2014). RAXML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033
- Stepanyan, L. S. (1967). Calandrella cheleensis Swinhoe a valid species. Acta Ornithologica [Warsaw], 10, 97–107.
- Stepanyan, L. S. (1983). Superspecies and sibling species in the avifauna of the USSR (pp. 1–296). Moscow, Russia: Nauka Press.
- Stepanyan, L. S. (1990). Conspectus of the ornithological fauna of the USSR (pp. 1–727). Moscow, Russia: Nauka Press. [in Russian].
- Stervander, M. (2015). On speciation in birds Genomic signatures across space and time. PhD thesis, Lund University.
- Stervander, M., Alström, P., Olsson, U., Ottosson, U., Hansson, B., & Bensch, S. (2016) Multiple instances of paraphyletic species and cryptic taxa revealed by mitochondrial and nuclear RAD data for *Calandrella* larks (Aves: Alaudidae). *Molecular Phylogenetics* and Evolution, 102, 233-245. http://dx.doi.org/10.1016/j. ympev.2016.05.032
- Vaurie, C. (1959). The birds of the palearctic fauna. Passeriformes. London, UK: Witherby.
- Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, 17, 2321–2328. https://doi. org/10.1111/j.1365-294X.2008.03742.x
- Werle, E., Schneider, C., Renner, M., Volker, M., & Fiehn, W. (1994). Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res.* 22(20), 4354–4355. https:// doi.org/10.1093/nar/22.20.4354

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ghorbani F, Aliabadian M, Zhang R, et al. Densely sampled phylogenetic analyses of the Lesser Short-toed Lark (*Alaudala rufescens*) -Sand Lark (*A. raytal*) species complex (Aves, Passeriformes) reveal cryptic diversity. *Zool Scr.* 2020;00:1–13. <u>https://doi.org/10.1111/</u> zsc.12422

#### APPENDIX A

Sources (XML files and Newick format) used to reconstruct divergence times and species trees for the *Alaudala rufes-cens-raytal* complex based on mitochondrial cytochrome *b* and 11 nuclear markers.