Genetic and morphological variation of *Halcyon senegalensis* revealing cryptic mitochondrial lineages and patterns of mitochondrial-nuclear discordance

Monica Mwale[1*], Desiré L. Dalton[1,2], Anna S. Kropff[1], Kim Labuschagne[1], Isa-Rita M. Russo[3], Samuel T. Osinubi[4]

[1]South African National Biodiversity Institute, P.O. Box 754, Pretoria, 0001, South Africa.
[2]School of Health and Life Sciences, Teesside University, Middlesbrough, TS1 3BX, United Kingdom.
[3]School of Biosciences, Cardiff University , Sir Martin Evans Building, Museum Avenue, Cardiff, CF10 3AX, UK
[4]FitzPatrick Institute of African Ornithology, University of Cape Town, South Africa.

*Corresponding author: M Mwale; m.mwale@sanbi.org.za

Abstract

The Woodland Kingfisher (*Halcyon senegalensis*) is widely distributed throughout Africa and occupies a wide variety of woodland and savannah habitat. Thus far, three subspecies have been described based on morphological variation. In the present study, using western, eastern and southern African populations, we examined the relationship between morphological and genetic divergence among two subspecies, *H. s. cyanoleuca* and *H. s. senegalensis*, using three mitochondrial (COI, CYTBand 16S) and two nuclear markers (FIB5 and RAG1). South African birds showed clear evidence for morphological divergence, with a longer wing and tail length compared to eastern and western birds. Phylogenetic analyses using Bayesian methods identified two well-characterised genetic clusters, representing the two subspecies. We determined that *H. s. senegalensis* and *H. s. cyanoleuca* are closely related subspecies that split recently (approximately 0.66 to 1.31 MYA) in the Pleistocene. Further, genetic substructure was evident within *H. s. senegalensis* with three distinct genetic clusters in each region. The separation between the lineages of *H. s. senegalensis* Ghana, Gabon and Uganda occurred approximately 0.12 to 0.57 MYA ago. Nuclear-mitochondrial discordance was however detected, where the pattern of divergence was not detected in the RAG1 and FIB5 sequences. Our results suggest that climate change, biogeographic barriers and local adaptation has played a role in the diversification of Woodland Kingfishers in Africa.

Keywords: phylogeography, nuclear-mitochondrial discordance, *H. s. cyanoleuca* and *H. s. senegalensis*

Introduction

Species that are widely distributed may display spatial phenotypic differences and patterns of genetic structuring depending on the level of connectivity between distinct populations (Avise and Ball 1990, Frankham et al. 2004). Various factors shape population genetic connectivity in mobile species including dispersal, philopatry, as well as geographic isolation by distance or barriers (Dobzhansky and Dobzhansky 1970, Greenwood 1980, Matthiopoulos et al. 2005, Taylor and Friesen 2012, Wright 1943). The level of differentiation is dependent on the balance of gene flow, genetic drift and natural selection along environmental gradients (Rice and Hostert 1993, Avise 2004, Pinho and Hey 2010). The scale of differentiation varies greatly among species and may result in speciation (Slatkin 1987). In addition, widely distributed species may be further subdivided into subspecies. Species display phenotypic distinctiveness through to reproductive incompatibility (de Quieroz 2007) whereas subspecies are reproductively compatible but display at least one heritable trait and may be associated with a specific geographic area or ecological niche (Wallin et al. 2017). Accurately determining processes involved in speciation and differentiation is challenging and complex and has not been fully explored for several avian species on the African continent. Kingfishers are an example of an understudied widely distributed species.

Kingfishers (Alcedinidae) consist of three subfamilies (Halcyoninae or Daceloninae, Alcedininae, and Cerylinae), 17 genera and 92 species, of which 18 species occur in Africa (Fry et al. 1988). Thus far, studies on the phylogenetic analysis of kingfishers are limited with disagreement regarding the basal family based on several lines of evidence (Maurer and Raikow 1981; Fry et al. 1988; Sibley and Ahlquist 1990; Johansson and Ericson 2003; Moyle 2006). The Woodland Kingfisher *Halcyon senegalensis* Linnaeus, 1766 belongs to the subfamily Halcyoninae. However, uncertainty remains regarding the relationships within the subfamily, especially at the base of the radiation (Moyle 2006). The Woodland Kingfisher is widely distributed in Africa, south of the Sahara (Moyle 2006) occupying a wide variety of woodland habitats. Although the species faces numerous threats, the International Union for Conservation of Nature (IUCN) has assessed it as least-concern, due to its extremely large distributional range (approximately 20,100 km^2) and stable population trend (Birdlife International 2021). Three subspecies are currently recognised (Figure 1), Woodland Kingfisher *Halcyon senegalensis fuscopileus* Reichenow, 1906 occurs from Sierra Leone to south Nigeria and south to Democratic Republic of Congo (DRC) and north Angola, Woodland Kingfisher *H. s. cyanoleuca* Vieillot, 1818 is found in south Angola and west Tanzania to South Africa and Woodland Kingfisher *H. s. senegalensis* is distributed in Senegal and Gambia to Ethiopia and north Tanzania. It has been reported that *H. s. fuscopileus* are found in forest habitat and are mainly or entirely resident whereas *H. s. cyanoleuca* and *H. s. senegalensis*

occur in well-developed woodland such as in tall *Acacia* stands and Mopane and are largely migratory (Fry et al. 1988). All three subspecies differ morphologically. *H. s. fuscopileus* is smaller and has a crown that is dark brownish grey with a mantle and breast greyer than other subspecies. *H. s. senegalensis* and *H. s. cyanoleuca* are similar in colour with *H. s. cyanoleuca* having a dark strip running behind the eye (Fry et al. 1988). The migratory movement patterns, drivers of migration, phenology and phylogeography of intra-African migrant species like the Woodland Kingfisher are only just being understood. Available knowledge indicates trans-equatorial migration in the *H. s. cyanoleuca*, with breeding grounds in the southern latitudes of South Africa (reportedly >23°S), though resident populations are reported in Tanzania. The *H. s. senegalensis* has resident populations on both sides of the Equator, with migrants of this subspecies making northern non-breeding movements into sub-Saharan Africa. Migrant *H. s. cyanoleuca* have been recorded in breeding ranges of *H. s. senegalensis*, however no breeding is reported to overlap, and the species is described as monogamous.

Thus far, limited phylogenetic studies have been completed for Woodland Kingfisher (Maurer and Raikow 1981; Fry et al. 1988; Sibley and Ahlquist 1990; Johansson and Ericson 2003; Moyle 2006) and no genetic studies have been conducted to elucidate the phylogenetic relationships of *H. s. cyanoleuca* and *H. s. senegalensis*. Modern molecular tools can be used to provide insight into the evolutionary history of the taxa (Jetz et al. 2012) and can be used for the discovery of previously unrecognized 'cryptic' species (Bickford et al. 2007). Thus, in the study presented here, analysis was conducted on two Woodland Kingfisher subspecies using genetic methods, namely sequencing of mitochondrial (mtDNA) and nuclear genes. Mitochondrial DNA has been preferentially used for taxonomic and phylogeographic studies due to its lack of recombination, ease of amplification and maternal inheritance (Boonseub et al. 2009). Sequence-based nuclear DNA variation can be used to complement mtDNA markers to further understand systematic relationships and genetic variation within a species (Lessa 1992). Thus, a combination of mtDNA and nuclear markers is recommended in demographic studies as there are several factors (e.g., non-neutrality, extreme rate variation and recombination) that could confound inferences provided by mitochondrial DNA.

Periodic climatic oscillations between wet and dry conditions have occurred over the last four to five million years (Zachos et al. 2001) and have led to significant habitat modification or fragmentation resulting in diversification of species (Hewitt 2003). However, the impact of climatic fluctuations varies per taxa as well as per geographic range (Stewart et al. 2010). In this study, we investigated the phylogenetic relationships and genetic history of divergence for *H. s. senegalensis* lineages from West Africa (Ghana) and East Africa (Uganda) and *H. s. cyanoleuca* from South Africa, using mitochondrial and nuclear molecular markers. We

hypothesize that the Woodland Kingfisher in Africa has been subjected to differentiation or speciation as a result of historic isolation and subsequent connectivity between populations as suitable habitat for the species expanded and contracted across Africa.

Materials and Methods

Sample collection and ethical approval

Across western, eastern and southern African, between November 2015 and January 2019, Woodland Kingfishers were trapped using varying numbers and lengths of mist nets, as well as spring traps baited with superworms (*Zophobas moria*). The traps were deployed during morning (06h00-10h00) and evening (15h00-18h00) sessions. Subspecies field identification during sample collection was based on subspecies range as well as the diagnostic eye stripe of the *H. s. cyanoleuca* and the darker plumage of the *H. s. fuscopileus* (the *H. s. fuscopileus* was not encountered during the study). A total of 60 individual Woodland Kingfishers were trapped and sampled at least once, with 15 additional re-traps that occurred in southern Africa. In Ghana (western Africa) blood samples were collected from 10 *H. s. senegalensis* individuals in the Greater Accra Province (n = 5) and in the Central Province (n = 5), which are c. 150 km apart. In Uganda (eastern Africa), blood samples were collected from 12 *H. s. senegalensis* individuals trapped in the Central Region. In South Africa (southern Africa), blood samples were collected from 38 *H. s. cyanoleuca* individuals trapped in the Limpopo Province. Below is the estimated distance between the focal countries:

- Ghana and Uganda ≈ 5,000 km,
- Ghana and South Africa ≈ 7,000 km, and
- Uganda and South Africa ≈ 3,700 km.

Generally, samples were collected during the breeding season across the sub-regions. In Ghana samples were collected between June and September. In Uganda samples were collected between July and August. In South Africa samples were collected between November and January. Trapped birds were ringed using individually coded aluminium rings that followed the ringing scheme in each country, as well as with plastic colour rings that followed a unique combination. The metal rings ensured individuality of samples as well as the identification of individual birds if re-trapped, while the plastic colour rings facilitated the identification of the individuals when free and perched. All trapped individuals were weighed and measured before tissue samples were collected, and immediately released after sampling. Blood samples were collected using the brachial venipuncture method with 27-gauge needles and 100 µl capillary tubes, and the blood samples were stored in lysis buffer (Seutin et al. 1991). All biological materials collected were stored at the Biobank of the South African National Biodiversity Institute (SANBI).

Morphometric measurements included:

- Mass (g); measured in the first two years using a spring balance (Pesola 20100 Micro-Line Metric Spring Scale – 100 g) during which the bird is weighed in the bird bag and the bird bag is weighed afterwards to determine the mass of the bird. A digital scale (Pesola PPS200 Professional Digital Pocket Scale – 200 g) was used in subsequent years during which a small plastic container was tared before each measurement, then the bird was placed in the container and weighed.
- Wing length (mm); measured using a wing rule, placing the bend of the wing against the top of the rule, flattening the wings and feathers so that the measure is maximised, and taking the reading from the tip of the longest wing feather (the primaries).
- Tail length (mm); measured using a flat rule and taking the measurement from the base of the tail to the tip of the longest tail feather.
- Head length (mm); measured using a digital caliper (0-150 mm), and the measurement was taken from the back of the skull to the front of the skull. This measure excludes the length of the culmen from the total head length.
- Tarsus length (mm); measured using a digital caliper (0-150 mm), and the measurement was taken from the notch on the metatarsus (where it meets the tibiotarsus) to the top of the bone above the bent toes.

The project was registered and approved as P14/23 by the Research and Ethics Committee (RESC) of SANBI. Permission to do research in terms of Section 20 of the Animal Diseases Act, 1984 (Act No. 35 of 1984) was issued by the Department of Agriculture, Land Reform and Rural Development (DALRRD, approval number 12/11/1/1/18). A dispensation on Section 20 approval in terms of the Animal Diseases Act, 1984 (Act No. 35 of 1984) was also issued by DALRRD to store the samples collected for the project at the SANBI Wildlife and Conservation Biobank. The lead bird ringer on the project was licensed in accordance with the South African Bird Ringing Unit (SAFRING), and collection permits (ZA/LP/93056, ZA/LP/WMD/1257, CPM 36408) were secured from the Limpopo Provincial Department of Economic Development, Environment and Tourism during each sampling year. Blood samples were collected under South African Veterinary Council (SAVC) authorisation (AL17/15903). Additionally, support for sampling was obtained from relevant national and local authorities in Ghana and Uganda.

Multivariate analysis

Summary statistic of variation among morphometric variables were estimated in PAST (Hammer and Harper, 2005). Principal Component Analysis (PCA) was performed to assess significant variation of five morphometric characters among the localities and subspecies, for the different sexes using PAST (Hammer and Harper 2005). Four morphometric characters were direct measurements for each bird namely, head length (HL), tarsus length (TarL), wing

length (WL) and tail length (TaiL) while the Body Mass Index (BMI) was calculated as a relationship of tarsus length to mass. This was done using the following formula: BMI = (mass [g]/tarsus length [mm]) (Nesbitt et al. 2008). All morphometric measurements were taken using the same devices to reduce the effects of measurement errors (ME). Furthermore, images of the birds and the sampling process were taken for verification of field data and notes. Measurement error (ME) was tested for using one-way Analysis of Variance (ANOVA) using the Kruskal-Wallis test (Kruskal and Wallis 1952) by comparing measurements taken by the samplers for each variable. The Kruskal-Wallis test can compare variation between two or more independent samples of equal or different sample sizes To minimize size-related dissimilarity, a PCA with log-transformed morphometric variables was also performed in PAST for adult birds. Log transformation of continuous data sets removes the skewness to make data as normal as possible to ensure more valid analysis. Multivariate analysis of variance (MANOVA) was also conducted in PAST to test whether there was significant variation in morphometric variation among sexes for the three localities and two subspecies.

Molecular sexing and amplicon amplification

Samples were extracted using the Quick-DNA Miniprep Plus Kit (Zymo Research, California, USA) following the manufacturer's instructions. In order to determine concentration and purity of the extracted DNA, a Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies, Inc., Delaware, USA) was used. Amplification of the CHD1 gene in order to determine gender was conducted using two primer sets; P2/P8 (Griffiths et al., 1998) and the NP/MP (Ito et al., 2003) using a standard Polymerase Chain Reaction (PCR) protocol as described in Mucci *et al.* (2017). Both reactions were completed using the Ampliqon Red *Taq* Mastermix (Ampliqon A/S, Odense, Denmark) in a final reaction volume of 15 µl containing 6.25 µl Ampliqon Red *Taq* Mastermix (0.1 M Tris/HCl, pH 8.5, $(NH_4)_2SO_4$, 4 mM $MgCl_2$, 0.2 % Tween 20, 0.4 mM deoxynucleotides, 0.2 units μl^{-1} Taq DNA Polymerase, inert red dye and stabilizer), 0.1 µM of the forward and reverse primers (Thermo Scientific, California, USA), 5.25 µl double distilled water (ddH_2O) and 50 ng of the template DNA. The cycling conditions were as follows: one cycle at 95ºC for 5 min; 35 cycles at 95ºC for 30 sec, 55ºC for 30 sec and 72ºC for 30 sec; followed by one cycle at 72ºC for 20 mins. Polymerase Chain Reactions were performed in a Labnet™ MultiGene™ OptiMax Thermal Cycler (Labnet International, Inc.). The resulting PCR products were run on the 3130 Genetic Analyzer (Thermofisher Scientific, California, USA) and results were analysed using GeneMapper software (Thermofisher Scientific, California, USA).

Amplification was conducted for three mitochondrial genes namely Cytochrome oxidase 1 (COI), Cytochrome *b* (CYTB) and the large subunit ribosomal RNA (16S). For COI, a region

of the gene was amplified using primers (BirdF1: 5'-TTC TCC AAC CAC AAA GAC ATT GGC AC-3' and BirdR1:5'-ACG TGG GAG ATA ATT CCA AAT CC TG-3'; Hebert et al., 2004). Amplification of a region of CYTB was conducted using external primers (L14764: 5'- TGR TAC AAA AAA ATA GGM CCM GAA GG-3' and H15915A: 5'- AGT CTT CAG TCT CTG GTT TAC AAG AC-3') and internal primers (H15541: 5'- GGG TGG AAK GGR ATT TTR TC-3' and L15087: 5'-TAC TTA AAC AAA GAA ACC TGA AA-3') for sequencing (Edwards et al. 1991, Sorenson et al. 1999; Fain, et al. 2007). Lastly, a region of 16S was amplified using primers 16S-f (5'-CGC CTG TTT AAC AAA AAC AT-3') and 16S-r (5'-CCG GTC TGA ACT CAG ATC ACG T-3'; Miya and Nishida 1996). In addition, amplification was conducted for two nuclear genes namely; recombination activating gene 1 (RAG1) and nuclear b-fibrinogen gene, intron 5 (FIB5). RAG1 were amplified with the primers RAG-1-F1 (5'-GAT TCT GTC ACA ACT GTT GGA GT-3'), and RAG-1-R2 (5'-TCC CAC TTC TGT GTT AGT GGA-3'; Gardner et al., 2010). Lastly, a region of the FIB5 gene was amplified with the following primers FIB5 (5'-CGC CAT ACA GAG TAT ACT GTG ACA T-3') and FIB6 (5'-GCC ATC CTG GCG ATT CTG AA-3'; Marini and Hackett 2002). All reactions were conducted using Ampliqon Red *Taq* Mastermix in a final reaction volume of 15 µl containing 6.25 µl Ampliqon Red *Taq* Mastermix (Ampliqon A/S, Odense, Denmark) (0.1 M Tris/HCl, pH 8.5, $(NH_4)_2SO_4$, 4 mM $MgCl_2$, 0.2 % Tween 20, 0.4 mM deoxynucleotides, 0.2 units μl^{-1} Taq DNA Polymerase, inert red dye and stabilizer), 0.1 µM of the forward and reverse primers (Thermo Scientific, California, USA), 5.25 µl ddH_2O and 10 ng of the template DNA. The cycling conditions were as follows: one cycle at 95ºC for 5 min; 35 cycles at 95ºC for 30 sec, 58ºC for 30 sec and 72ºC for 30 sec; followed by one cycle at 72ºC for 20 mins. Polymerase Chain Reaction was performed in a Labnet™ MultiGene™ OptiMax Thermal Cycler (Labnet International, Inc.). The PCR product was purified by adding 0.25 µl of 10 U Exonuclease 1 and 2 µl of 2 U FastAP Thermosensitive Alkaline Phosphatase (Thermofisher Scientific, California, USA) to the PCR product. The purification reaction was run for one cycle at 37ºC for 15 mins followed by one cycle at 85ºC for 15 mins in a Labnet™ MultiGene™ OptiMax Thermal Cycler (Labnet International, Inc.). Cycle sequencing reactions were completed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermofisher Scientific, California, USA) using the Sanger chain termination method. Sequencing was conducted in a final reaction volume of 10 µl containing 0.7 µl BigDye, 2.25 µl of BigDye Terminator v3.1 5x Sequencing buffer, 0.75 µl ddH_2O, 3.2 µM of primer and 5 µl of the PCR product. The cycling conditions were as follows: one cycle at 95ºC for 2 min; 40 cycles at 85ºC for 10 sec, 55ºC for 10 sec and 60ºC for 2 min 30 sec. Reactions were performed in a Labnet™ MultiGene™ OptiMax Thermal Cycler (Labnet International, Inc.). Sequencing reactions were completed in both the forward and reverse direction. The cycle sequencing product was purified using the BigDye Xterminator Purification Kit

(Thermofisher Scientific, California, USA) and sequences were visualized using the 3500 Genetic Analyzer (Thermofisher Scientific, California, USA).

Phylogenetic and genetic diversity analyses

Forward and reverse sequences were aligned in BioEdit (Hall 1999) to create a consensus sequence and all sequences were manually trimmed and checked for ambiguous peaks. Absence of nuclear mitochondrial DNA sequences (numts) was confirmed following comparisons with published mitochondrial genomes and via visual inspection of the chromatograms. We included one in-group reference sequence of *H. senegalensis* from Gabon (complete mtDNA genome; MN356338) and out-group reference sequences from the genus. The three mitochondrial genes were initially analysed separately and included outgroup reference sequences of White-breasted Kingfisher *H. smyrnensis* (KY940559; COI, CYTB and 16S), Black-capped Kingfisher *H. pileata* (KJ476742; COI and CYTB) and Ruddy Kingfisher *H. coromanda* (MK327578 and KT356219; COI, CYTB and 16S). For the concatenated dataset of the three mitochondrial genes (COI, CYTB and 16S), the outgroup reference sequences included a consensus of two species, the Black-capped Kingfisher *H. pileata* (KJ476742) and the White-breasted Kingfisher *H. smyrnensis* (KY940559). Inter- and intraspecific p-distances between subspecies and lineages for the concatenated mtDNA dataset were calculated using maximum likelihood (ML) genetic distance in MEGA7. Haplotype diversity (h), nucleotide diversity (π) and levels of gene-flow were all calculated using DNASP (Rozas et al. 2003), while Tajima's D statistic was estimated in Arlequin 3.5.2.2 (Schneider et al. 2000). The mismatch distribution (Rogers and Harpending 1992) of the *H. senegalensis* population was calculated using Arlequin 3.5.2.2 for both gene regions in order to test for demographic changes, The Harpending's raggedness index (Hri) (Harpending et al. 1993) and Sum of Squared deviations (SSD) estimates were calculated to determine the fit of the expected frequencies under the demographic expansion model. A small raggedness index suggests an expanded population while a large raggedness index indicates a stationary population (Harpending 1994). Analysis of Molecular Variance (AMOVA) was used as implemented in Arlequin 3.5.2.2. to estimate population differentiation by testing hypotheses about genetic variation and geographic differentiation between the two subspecies and among the three localities respectively.

The two nuclear genes were also analysed separately. For the RAG1 gene region, we included ingroup reference sequences of *H. cyanoleuca* (MK579328) and *H. senegalensis* (DQ111818) and an outgroup sequence of Rufous-collared Kingfisher *Actenoides concretus* (MG0081831). For the FIB5 gene region, we included an ingroup reference sequences of *H. senegalensis* (MG001233) and an outgroup sequences of Chocolate-backed Kingfisher *H. badia*, Striped

295 Kingfisher *H. chelicuti*, Brown-hooded Kingfisher *H. albiventris*, Grey-headed Kingfisher *H.*
296 *leucocephala* and Brown-breasted Kingfisher *H. gularis* (MG001216-MG001221). The Grey-
297 headed Kingfisher *H. leucocephala* was the only available outgroup DNA sequence for the
298 concatenated nuclear DNA genes (RAG-1 and FIB5) partition (MG001221 and MG008244).
299 All reference sequences were obtained from the National Centre for Biotechnology Information
300 (NCBI) GenBank. We determined the best fitting substitution model in MEGA7 (Kumar et al.
301 2016). Haplotype reconstruction for the nDNA sequence was done after phasing the data
302 using the algorithms provided in Stephens et al. (2001) and Wang and Xu (2003) for
303 heterozygous sites (polymorphic nucleotide positions) in DNASP.
304
305 We computed phylogenetic trees in MrBayes (Ronquist and Huelsenbeck, 2003) and MEGA7
306 by using the Bayesian phylogenetic inference (BI) and ML methods for the separate as well
307 as concatenated mitochondrial and nuclear DNA datasets. The BI analysis was conducted
308 using the General Time Reversal (GTR) model with the following parameters: the maximum
309 likelihood model employed 6 substitution types, with gamma rate variation (0.07) across sites
310 modelled using a gamma distribution for the mtDNA data partition only (mtDNA only analysis
311 and concatenated portioned dataset of all genes). The Markov Chain Monte Carlo (MCMC)
312 approach was used to search for trees using 4 chains for 5,000,000 generations, with trees
313 sampled every 50,000 generations (the first 20% were discarded as "burnin"). The resultant
314 trees were formatted in MEGAX (Kumar et al. 2018) with posterior probabilities (PB).
315 Phylogenetic relationships were also reconstructed by ML based on the Hasegawa-Kishino-
316 Yano (HKY +G) model for COI, CYTB and the concatenated MtDNA dataset, the Kimura 2-
317 parameter (K2) model for 16S and RAG1 and the Jukes–Cantor (JC) one-parameter model
318 for FIB5. Branch support values were estimated using non-parametric bootstrap with 1000
319 replicates. The consistency of the two methods (BI and ML) was evaluated by comparing
320 support values at identical nodes and by assessing the similarity of the tree topologies.
321 Phylogenetic analysis of the concatenated mtDNA gene regions identified divergence
322 between subspecies as well as demonstrated well-supported geographic lineages. The
323 degree of divergence between subspecies and lineages was thus calculated using ML genetic
324 distance in MEGA7.
325
326 *Haplotype network inference*
327 Haplotype networks were constructed separately for the concatenated mtDNA gene regions
328 (COI, CYTB and 16S) as well as for concatenated nDNA gene regions (RAG-1 and FIB5)
329 using PopART v. 1.7 (Leigh and Bryant 2015) that employs an agglomerative approach where
330 clusters are progressively combined with one or more connecting edges. This was done using

the TCS network, a parsimony-based method of analysis defined by Templeton et al. (1992) and Clement et al. (2002).

Divergence time estimation

Phylogenetic reconstruction and molecular dating calibration analysis was done using Bayesian evolutionary analysis in BEAST v1.8.4 (Drummond et al. 2012) implemented with BEAGLE (Ayres et al. 2011) using the mtDNA datasets. A configuration file generated in BEAUti v1.8.4, was run using a random starting tree under the following parameters: GTR model, 10 million generations of the Markov Chain, 10,000 tree sampling frequency with the first 10% discarded as burn-in, a strict clock model and speciation (Yule) process. The divergence times were calibrated using published date estimates for the three major clades of kingfishers (Alcedininae, Cerylinae and Halcyoninae) of the family Alcedinidae. An average node age of 16.3 Ma (95% CI 13.2–19.6) was used for the root of the tree representing the Miocene origin of the Halcyoninae (Anderson et al. 2018). Three species of its sister clade Cerylinae (*Ceryle rudis, Megaceryle lugubris* and *Chloroceryle aenea*) were included as the outgroup for this analysis. For the Halcyoninae clade, published mtDNA sequences of *Halcyon* (*coromanda, pileata,* and *smyrnensis*) and haplotypes of *H. s. senegalensis and H. s. cyanoleuca* generated by this study were included in the analysis with three species from closely related genera (*Todiramphus, Dacelo* and *Actenoides*). [Information on the additional GenBank sequences used in this phylogenetic tree is given in Supplementary Table 1]. The information contained within the sampled trees of the BEAST output, was summarised using TreeAnnotator v1.8.4, with the annotated tree being visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

Molecular sexing

Molecular gender determination analysis was successful for 59 out of the 60 samples tested. Collected samples consisted of 30 male and 29 female birds. In South Africa, 17 female and 20 male birds were identified. Samples from Ghana were identified to include six females and four males whereas from Uganda, six birds were determined to be female and six as male.

Morphometric analyses

Morphometric measurements were collected for 52 individuals from 15 female and 20 male adult birds from South Africa, six female and four male birds from Ghana and four female and three male birds from Uganda (Table 1A). Measurements from juvenile birds were excluded (five birds). ANOVA results for effect of measurement errors were not significant ($P>0.05$) for three variables (Head length, tail length and mass), showing no differences among

measurement means taken by different samplers [*F* (2,45), critical value=3.20, α=0.05). There was however a significant difference in measurements taken for tarsus and wing length among samplers, but only due to sampling in South Africa (Table 1A). Therefore, this difference was not attributed to measurement error but actual difference in size among regions. The PCA of the five variables recovered six factors that had eigenvalues > 1, explaining 99% of the variation for log-transformed and untransformed datasets (Table 1B). Principal component 1 (PC1) described 66.27% and 51.79% of the variation while PC2 accounted for 19.86% and 26.04% of the variations for untransformed and log-transformed datasets respectively. Morphometric variables with significant factor loadings in PC1, PC2 and 3 were the mass, wing length and tail length while mass, BMI and tarsus length remained significant after log-transformations. These variables all had factor loadings more than 0.50 above the 0.30 that is considered significant. MANOVA indicated that there was no significant difference among sexes among localities for the two subspecies (Wilks´ lambda: 0.78 and P=0.07). A plot (Figure 2) of the first two PC showed that the male and female samples could not be segregated into groups, with high degree of overlap. However, there was clustering among populations with the Ghana individuals being more morphologically discrete from the other two populations (little to no overlap), with a higher degree of overlap between Uganda and South Africa.

Mitochondrial DNA phylogeographic and phylogenetic analysis

In this study, we generated sequences for 60 samples, of which 49 samples included all three gene regions (Supplementary Table 1): CYTB (GenBank accession numbers: OL602280 - OL602343), COI (GenBank accession numbers: OL518993-OL519047) and 16S (GenBank accession numbers: OL519049-OL519104). Lack of amplification of certain gene regions may be due to degradation of DNA at primer binding sites.

Analysis of subspecies: A total of 19 haplotypes were identified based on the concatenated mtDNA dataset, of which nine haplotypes were unique to *H. s. cyanoleuca* (South Africa) and ten were unique to *H. s. senegalensis* (Figure 3A). Single-locus tree topologies showed two resolved monophyletic, divergent clades for all three mitochondrial gene regions (Supplementary Figures 1A-C); thus, we produced a final tree with the concatenated dataset (mtDNA = 2281 bp, Figure 3B). Both ML (Figure 3B) and BI (Supplementary Figure 2) analysis showed identical tree topologies, with two resolved monophyletic, divergent clades corresponding to *H. s. cyanoleuca* and *H. s. senegalensis* with high bootstrap support values (100% for both ML and BI analysis). The overall data set has very high haplotype diversity (0.73) and low nucleotide diversity (0.0007) (Table 2). Tajima's D statistic estimates were not significant (*P*>0.05), however the estimate was negative (-1.16) for *H. s. cyanoleuca* suggesting that the South African population may have recently undergone a population size

405 expansion (Table 2). This result was also supported by the mismatch distribution
406 (Supplementary Figure 4) for the species, which was bimodal and significant (P<0.05) for the
407 mtDNA data suggesting either a recent population expansion or balancing selection favouring
408 the genetic variation among populations (Harpending et al. 1993). The nDNA had a more
409 ragged mismatch (to the right and broader) that was not significant (P>0.05) displaying a high
410 level of nucleotide variation with individual sequences differing at many sites, a pattern
411 expected following a bottleneck in a previously large population with possible incomplete
412 lineage sorting (Harpending et al. 1993). The low and non-significant Hri and SSD estimates,
413 suggest the presence of non-equilibrium and that the nDNA data has a relatively good fit to
414 the expansion model. Pairwise sequence divergence (F_{ST}) for concatenated mtDNA
415 sequences between the two subspecies was very high and significant (>0.94, *P*<0.05) (results
416 not shown). However, the hierarchical AMOVA analyses of populations (locality groups)
417 between the two subspecies indicated no significant variation (*P*>0.05) although 83.88% of
418 the total genetic variation was contributed by 'among groups' variation with genetic variation
419 from 'among populations within groups' accounting for 9.69% of the variation (Table 3). A
420 significant 6.43% of the variation was detected within populations. Our analysis showed that
421 the sequence divergence between species (*H. smyrnensis* and *H. senegalensis*; *H. pileata*
422 and *H. senegalensis*) was 9.3 and 8.6% respectively (Table 4). Sequence divergence between
423 outgroup species (*H. smyrnensis* and *H. pileata*) was 3.9%. Divergence between subspecies
424 (*H. s. senegalensis* and *H. s. cyanoleuca*) was 1.1% (Table 4).
425
426 *Analysis of lineages:* Within *H. s. senegalensis* unique haplotypes were distributed per locality
427 in Uganda (5), Gabon (1) and Ghana (4) (Figure 3A, Table 2). Thus, no shared haplotypes
428 were detected among the three *H. s. senegalensis* localities suggesting geographic sub-
429 structuring. The phylogenetic trees (COI, CYTB and concatenated mtDNA) also showed sub-
430 structuring in the *H. s. senegalensis* clade with two lineages comprising of samples from
431 Uganda, and the second lineage including samples from Ghana and Gabon with high
432 bootstrap support values for the concatenated mtDNA (97% ML and 100% for BI analysis)
433 suggesting isolation over the geographic range between western and eastern Africa (Figure
434 3B; Supplementary Figure 2). However, absence of sub-structuring was detected in the 16S
435 phylogenetic tree (Supplementary Figure 1C). Pairwise sequence divergence (F_{ST}) for
436 concatenated mtDNA sequences was high and significant (0.802, *P*<0.05) between the two
437 *H. s. senegalensis* geographic lineages (Table 2). Distribution of the two lineages clearly
438 corresponds to their geographic origins: Uganda and Ghana. AMOVA comparisons were done
439 for the three groups namely, 1. South Africa, 2. Uganda and 3. Ghana and Gabon and further
440 highlighted that there was higher genetic variation within populations than among assigned
441 groups (Table 3). The estimated gene flow (Nm) was 0.22, suggesting that the genetic

differentiation among these populations is significant and that gene flow is restricted. Govindajuru (1989) indicated that levels of gene flow with Nm <0.25 represent low gene flow, while as Nm >1 can be categorized as high gene flow with 0.25 to 0.99 representing intermediate gene flow. Sequence divergence between geographic lineages within *H. s. senegalensis* (Ghana, Gabon and Uganda) varied between 0.2 to 0.4% (Table 4).

Nuclear DNA phylogenetic analysis

Here, we generated 704 bp of sequence for RAG1 (GenBank accession numbers: OL602344 - OL602402), and 801 bp of sequence for FIB5 (GenBank accession numbers: OL602403 - OL602453) in 46 samples (Supplementary Table 2B) represented by 92 phased sequences. Bayesian inference (BI) of the concatenated nuclear genes (1,505 bp) showed short internal branches possibly due to a rapid radiation of lineages. The BI tree (Supplementary Figure 2) and individual ML trees (Supplementary Figures 1D and E) also showed a different topology to the mitochondrial DNA tree and some lineages were not monophyletic. The overall data set has very high haplotype diversity (0.99) and low nucleotide diversity (0.0113) (Table 2). A total of 77 unique haplotypes for this dataset were identified and none of the haplotypes were shared among the three population localities, similar to the mtDNA (Supplementary Figure 3). There was however some evidence of genetic sub-structuring among subspecies with pairwise F_{ST} comparisons between the South African (*H. s. cyanoleuca*) samples being significantly different from samples of *H. s. senegalensis* from Ghana and Uganda. AMOVA analyses conducted for the two subspecies (1. South Africa vs 2. Uganda and Ghana) further revealed that a very high and significant percentage of variation (79.2%; $P<0.05$) was attributed to within population variation. This variation was also observed in the similarly high haplotype diversity values for the three localities (all > 0.98).

Dating and historical demography (mtDNA data)

The estimated time of divergence of *H. senegalensis* on the basis of the constructed Bayesian phylogenetic tree (Figure 4 and Table 5), was estimated to have occurred 1.31 Mya ago in the Pleistocene after the late Miocene diversification of the Halcyoninae. The divergence of the species *H. senegalensis* appeared to be recent and rapid around 0.97 MYA (95% HPD 1.31–0.66) after the colonisation of the genus *Halcyon* into Africa (approximately 1.8 MYA; Andersen et al., 2018). The diversification of the two subspecies, *H. s. cyanoleuca* (0.22 MYA) and *H. s. senegalensis* (0.39 MYA), was estimated to have been even more recent in the latter Pleistocene, based on this data set.

Discussion

In this study, we describe for the first time genetic and morphological variation of two broadly distributed subspecies of Woodland Kingfisher. Phylogenetic analysis based on mtDNA markers recovered *H. s. cyanoleuca* and *H. s. senegalensis* as two strongly supported, reciprocally monophyletic clades that corresponded to currently recognized subspecies designation within Woodland Kingfisher. The clear split between the two subspecies was supported by F_{ST} (>0.94) analysis. The AMOVA revealed high variance between the two subspecies (83.88%), however this was not significant which may be attributed to high within subspecies variation. Distinct nuclear DNA lineages were however not detected (RAG1 and FIB5 and nuclear haplotype analysis identified 77 haplotypes that were not shared between subspecies or localities). Thus, biogeographic discordance and phylogenetic incongruence between mitochondrial and nuclear markers was observed in this study. Several hypotheses have been put forward to explain mitochondrial-nuclear discordance including incomplete lineage sorting, hybridisation and ancestral population structure (Toews and Brelsford 2012, Linck et al. 2019). Funk and Omland (2003) reported that a high proportion of bird species (16.7%) were paraphyletic with the most common reason being identified as incomplete lineage sorting due to recent speciation (McKay and Zink 2010). The divergence between *H. s. cyanoleuca* and *H. s. senegalensis* is fairly recent (approximately 0.66 to 1.31 MYA), therefore the result of incomplete lineage sorting is not unexpected as these signals are generally detected within species that have evolved more recently.

It has been previously reported that the Halcyoninae clade comprising of the genera *Lacedo*, *Pelargopsis* and *Halcyon* originated in the Indomalayan region approximately 16.3 MYA (13.2–19.6 MYA). Andersen *et al*. (2018) placed the arrival on the African continent of the kingfisher lineage that resulted in the Woodland Kingfisher at 8 MYA, during the Miocene. The split of the Woodland Kingfisher and the Mangrove Kingfisher (*Halcyon senegaloides*), the closest congeneric of the Woodland Kingfisher, occurred around 2 MYA, during the Pleistocene. The single colonization of Halcyon into Africa was then followed by a back-colonization into Asia approximately 1.8 MYA (1.2-2.3 MYA) ago (Anderson et al. 2018). Periodic climate oscillations across the African continent have been reported with three peaks of aridification being described at approximately 2.8, 1.7 and 1.0 MYA (deMenocal 1995). During these time periods repeated oscillations in temperature and rainfall shifted between humid-warm (pluvials) and arid-cool (interpluvials) phases. During pluvials, forests expanded and during interpluvials, arid areas expanded. In this study, we suggest that the common ancestor of these two subspecies was more broadly distributed in Africa around 1 to 2 MYA. The interpluvial period that occurred 0.8 to 1.2 MYA ago resulted in forest contraction and the expansion of savannah and grassland habitat (deMenocal and Bloemendal 1995) which is unfavourable to Woodland Kingfisher. The pronounced dry period most likely closed the

515 corridor between the southern and eastern African Woodland Kingfisher populations resulting
516 in morphological/genetic divergence and differentiation. The reported arid period (0.8 to 1.2
517 MYA) matches the dating of the *H.s. senegalensis* and *H.s. cyanoleuca* split in the Pleistocene
518 (0.66 to 1.31 MYA) identified in this study. The period of isolation may have been followed by
519 secondary contact between *H.s. senegalensis* from East and West Africa and *H.s. cyanoleuca*
520 from southern Africa after the last recorded interpluvial period (0.8 to 1.2 MYA ago) due to the
521 replacement of grassland habitat with woodland. This habitat change could very well be the
522 precursor of the current migratory pattern of the *H. s. cyanoleuca*. Due to replacement of the
523 bushveld with woodlands in South Africa, Woodland Kingfishers would have gradually
524 extended their range in South Africa (McLachlan and Liversidge, 1957). Today, breeding pairs
525 of *H.s. cyanoleuca* are found as far south as the Gauteng Province due to land use changes
526 (South African Bird Atlas Project 2 [SABAP2, http://sabap2.birdmap.africa]). Here we further
527 suggest that contemporary restrictions to gene flow between Woodland Kingfisher subspecies
528 may be due to allochrony (divergence in breeding time, Servedio et al. 2011). This is
529 associated with temporal variation, availability and abundance of insects, geographic distance
530 and/or local adaptation to their habitat. South African Woodland Kingfisher populations are
531 reported to breed in November, western populations breed in June and eastern populations
532 breed in January. Thus, seasonal separation of breeding times may be an important driver
533 contributing to continued isolation of the subspecies (Taylor and Friesen 2017).
534
535 Genetic structure was also recovered within *H. s. senegalensis* suggesting that diversity within
536 Woodland Kingfisher is underestimated. Phylogenetic and network analyses of mtDNA
537 sequence data revealed significant differences at different geographic scales (Ghana, Gabon
538 and Uganda) with three distinct lineages detected within *H. s. senegalensis* suggesting a
539 significant historic isolation among these populations. The separation between the lineages of
540 *H. s. senegalensis* from western/central Africa from eastern Africa occurred between 0.22 to
541 0.57 MYA, whereas the separation between western and central African populations occurred
542 more recently between 0.12 to 0.36 MYA ago. Divergence of populations from western, central
543 and eastern Africa has been reported for several terrestrial vertebrates (lizards, mammals and
544 birds) within a similar time period. For example, chimpanzees constituted a single population
545 until approximately 0.1 MYA and were subsequently divided into three populations (southern
546 Cameroon, central Africa, and eastern Africa) (Gonder et al. 2011). Two species of
547 woodpeckers (*Campethera caroli* and *C. nivosa*) from the upper and lower Guinean forest
548 blocks were reported to each consist of populations that diverged between 0.5 to 0.8 MYA
549 (Fuchs and Bowie 2015). Further, Perktaş et al (2020) described significant genetic and
550 morphological variation within turacos and go-away-birds (Musophagidae) that were generally
551 concordant with the organization of the montane avifaunal regions of Africa which have in the

past been driven by Pliocene forest dynamics. Several potential biogeographic barriers have been identified including the Dahomey Gap, which has been described as a 200 km wide forest-savannah mosaic corridor separating the west African and central African rain forest (Moreau 1966, Hall and Moreau 1970, Crowe and Crowe 1982). The habitat and size of the Dahomey Gap has varied over time in response to large scale climate shifts (Mayr and O'Hara 1986). Additional potential barriers include the Niger Delta and the Cameroon volcanic line which are reported to be barrier to gene flow between populations from West Africa and Central Africa (Hassanin et al. 2015). Lastly, the Congo rainforest may potentially present a source of isolation between the central (Gabon) and eastern African group.

The morphological analysis supported the structuring results detected with mtDNA and our results showed morphological differentiation between the eastern, western, and southern populations. Thus far limited studies have reported on morphological variation within Woodland Kingfisher populations and have only described phenotypic differences between subspecies such as the presence of an eye stripe in *H. s. cyanoleuca* (Fry et al. 1988). In this study, birds from South Africa had longer wing and tail lengths, had lower mass and constituted a highly differentiated genetic cluster. The lower mass observed in *H. s. cyanoleuca* may be due to physiological processes that occur due to long-distance migration (Ramenosky 1990) and can vary in the course of an annual cycle. Birds from Uganda were intermediate whereas Woodland Kingfishers from Ghana had the shortest wing and tail lengths. The *H. s. senegalensis* in southern parts of Ghana are reported to be resident populations, thus shorter wings in populations that do not migrate was expected (Perez-Tris and Tellerıa 2001). However, in this study we additionally detected a continental gradient in wing length from western to the eastern populations.

In this study, molecular analysis confirmed subspecies designation of *H. s. cyanoleuca* and *H. s. senegalensis.* In addition, we further identified distinct lineages within *H. s. senegalensis*. Here, we provide evidence that climate change leading to expansion and contraction of geographic range has played an important role in shaping populations of Woodland Kingfisher s. We further suggest that limited contemporary gene flow in these populations due to distance, differences in breeding behaviour and/or local adaptation to their habitat. Our results were found to be concordant with other vertebrate species that have identified unique populations in central, eastern, western, and southern Africa. However, additional data needs to be collected throughout the distribution range of Woodland Kingfishers to develop a comprehensive picture of intra- and interspecific variation within *H. s. senegalensis* and how this is distributed geographically with the other co-occurring sub-species. The relationships of *H. senegalensis* including the *H. s. fuscopileus* subspecies that we were not able to sample in

589 our study, clearly awaits molecular analysis. Although Woodland Kingfishers are currently
590 abundant and have been assessed as least-concern, the identification of unique populations
591 in a continually transforming habitat may require future conservation management.
592
593 **Acknowledgements**
594 This research was supported by the International Foundation for Science, Stockholm, Sweden
595 (IFS; J/4611-3), with additional support from the African Bird Club (ABC; Expedition Award)
596 and the British Ecological Society (BES; EA17/1146).
597
598 **References**
599 Andersen MJ, McCullough JM, Mauck III WM, Smith BT, Moyle RG, 2018. A phylogeny of
600 kingfishers reveals an Indomalayan origin and elevated rates of diversification on oceanic
601 islands. *Journal of biogeography* 45: 269-281.
602 Avise JC. 2004. *Molecular Markers, Natural History, and Evolution*. Sinauer: Sunderland.
603 Avise JC, Ball RM. 1990. Principles of genealogical concordance in species concepts and
604 biological taxonomy. *Oxford Surveys in Evolutionary Biology*, 7: 45–67.
605 Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, Huelsenbeck JP, Ronquist F,
606 Swofford DL, Cummings MP, Rambaut A. 2012. BEAGLE: an application programming
607 interface and high-performance computing library for statistical phylogenetics. *Systematic*
608 *biology* 61: 170-3.
609 Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic
610 species as a window on diversity and conservation. *Trends in ecology and evolution* 22:
611 148-55.
612 BirdLife International. 2021. *Species factsheet: Halcyon senegalensis*. Downloaded from
613 http://www.birdlife.org on 18/01/2021.
614 Boonseub S, Tobe SS, Linacre AM. 2009. The use of mitochondrial DNA genes to identify
615 closely related avian species. *Forensic Science International: Genetics Supplement Series*
616 2: 275-7.
617 Catanach TA, Johnson KP, Marks BD, Moyle RG, Valim MP, Weckstein JD. 2019. Two
618 lineages of kingfisher feather lice exhibit differing degrees of cospeciation with their hosts.
619 *Parasitology* 146:1083-95.
620 Clement M, Snell Q, Walke P, Posada D, Crandall K. 2002. TCS: estimating gene
621 genealogies. *Proceedings of the 16th International Parallel and Distributed Processing*
622 *Symposium* 2:184.
623 Crowe TM, Crowe AA. 1982. Patterns of distribution, diversity and endemism in Afrotropical
624 birds. *Journal of Zoology* 198: 417-42.
625 DeMenocal PB. 1995. Plio-pleistocene African climate. *Science* 270: 53-9.

626 DeMenocal PB, Bloemendal J. 1995. Plio-Pleistocene climatic variability in subtropical Africa
627 and the paleoenvironment of hominid evolution: A combined data-model approach, in: E.
628 Vrba, G. Denton, L. Burckle, T. Partridge (Eds.), *Paleoclimate and Evolution With Emphasis*
629 *on Human Origins*, Yale University Press, New Haven, 262-288.
630 De Queiroz K. 2007. Species concepts and species delimitation. *Systematic biology* 56: 879-
631 86.
632 Dobzhansky T, Dobzhansky TG. 1970. *Genetics of the evolutionary process*, vol. 139. New
633 York: Columbia University Press.
634 Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti
635 and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973.
636 Edwards SV, Arctander P, Wilson AC. 1991. Mitochondrial resolution of a deep branch in the
637 genealogical tree for perching birds. *Proceedings of the Royal Society of London. Series*
638 *B: Biological Sciences* 243: 99-107.
639 Fain MG, Krajewski C, Houde P. 2007. Phylogeny of “core Gruiformes”(Aves: Grues) and
640 resolution of the Limpkin–Sungrebe problem. *Molecular phylogenetics and evolution* 43:
641 515-29.
642 Feng S, Stiller J, Deng Y, Armstrong J, Fang Q, Xie D, Chen G, Paten B, Zhang G. (in press).
643 Densely sampling genomes across the diversity of birds increases power of comparative
644 genomics analyses. Nature.
645 Frankham R, Ballou JD, Briscoe DA. 2004. *A primer of conservation genetics*. Cambridge
646 University Press.
647 Fry CH, Keith S, Urban EK. 1988. *The Birds of Africa*, Vol. 3. Academic Press, London.
648 Fuchs J, Bowie RC. 2015. Concordant genetic structure in two species of woodpecker
649 distributed across the primary West African biogeographic barriers. *Molecular*
650 *phylogenetics and evolution* 88: 64-74.
651 Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and
652 consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology,*
653 *Evolution, and Systematics* 34: 397-423.
654 Gardner JL, Trueman JW, Ebert D, Joseph L, Magrath RD. 2010. Phylogeny and evolution of
655 the Meliphagoidea, the largest radiation of Australasian songbirds. *Molecular*
656 *Phylogenetics and Evolution* 55: 1087-102.
657 Gonder MK, Locatelli S, Ghobrial L, Mitchell MW, Kujawski JT, Lankester FJ, Stewart CB,
658 Tishkoff SA. 2011. Evidence from Cameroon reveals differences in the genetic structure
659 and histories of chimpanzee populations. *Proceedings of the National Academy of*
660 *Sciences* 108: 4766-71.
661 Govindajuru RD 1989. Variation in gene flow levels among predominantly self-pollinated
662 plants. *Journal of Evolutionary Biology* 2: 173–181.

663 Greenwood PJ. 1980. Mating systems, philopatry and dispersal in birds and mammals. *Animal*
664 *Behaviour* 28: 1140–1162.
665 Griffiths R, Double MC, Orr K, Dawson RJ. 1998. A DNA test to sex most birds. *Molecular*
666 *Ecology* 7: 1071-5.
667 Hall T. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis
668 program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.
669 Hall BP, Moreau RE. 1970. *Speciation in African Passerine Birds*. British Museum of Natural
670 History, London.
671 Hammer, Ø. and Harper D.A.T. 2005: PAST – PAlaeontological STatistics, ver. 1.34.
672 http://folk.uio.no/ohammer/past.
673 Hassanin A, Khouider S, Gembu GC, Goodman SM, Kadjo B, Nesi N, Pourrut X, Nakoune E,
674 Bonillo C. 2015. The comparative phylogeography of fruit bats of the tribe Scotonycterini
675 (*Chiroptera, Pteropodidae*) reveals cryptic species diversity related to African Pleistocene
676 forest refugia. *Comptes Rendus Biologies* 338: 197-211.
677 Harpending HC, Sherry ST, Rogers AR, Stoneking M. 1993. The Genetic Structure of Ancient
678 Human Populations. *Current Anthropology* 34(4), 483–496.
679 Harpending HC. 1994. Signature of ancient population growth in a low-resolution
680 mitochondrial DNA mismatch distribution. *Human Biology* 66(4):591-600. PMID: 8088750.
681 Hebert PD, Stoeckle MY, Zemlak TS, Francis CM, Godfray C. 2004. Identification of birds
682 through DNA barcodes. *PLoS biology* 2: e312.
683 Hewitt G. 2003. *Ice ages: species distributions, and evolution*. In Evolution on planet Earth.
684 Academic press, 339-361.
685 Ito H, Sudo-Yamaji A, Abe M, Murase T, Tsubota T. 2003. Sex identification by alternative
686 polymerase chain reaction methods in Falconiformes. *Zoological Science* 20: 339-44.
687 Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. 2012. The global diversity of birds in
688 space and time. *Nature* 491: 444-8.
689 Johansson US, Ericson PG. 2003. Molecular support for a sister group relationship between
690 Pici and Galbulae (*Piciformes sensu* Wetmore 1960). *Journal of Avian Biology* 34: 185-97.
691 Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis
692 version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33: 1870-1874.
693 Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary
694 genetics analysis across computing platforms. *Molecular biology and evolution* 35: 1547-
695 1549.
696 Kruskal W H, Wallis WA. 1952. Use of ranks in one-criterion variance analysis. *Journal of*
697 *the American Statistical Association* 47(260), 583–621.
698 https://doi.org/10.1080/01621459.1952.10483441

699 Lessa EP. 1992. Rapid surveying of DNA sequence variation in natural populations. *Molecular*
700 *Biology and Evolution* 9: 323-30.
701 Leigh JW, Bryant D. 2015. PopART: Full-feature software for haplotype network construction.
702 *Methods in Ecology and Evolution* 6: 1110–1116.
703 Linck E, Epperly K, Van Els P, Spellman GM, Bryson Jr RW, McCormack JE, Canales-del-
704 Castillo R, Klicka J. 2019. Dense geographic and genomic sampling reveals paraphyly and
705 a cryptic lineage in a classic sibling species complex. *Systematic biology* 68: 956-66.
706 Linnaeus C. 1766. Systema naturae: per regna tria natura, secundum classes, ordines,
707 genera, species, cum characteribus, differentiis, synonymis, locis (in Latin). Volume 1, Part
708 1 (12th ed.). Holmiae (Stockholm): Laurentii Salvii. 180.
709 Marini MÂ, Hackett SJ. 2002. A multifaceted approach to the characterization of an
710 intergeneric hybrid manakin (*Pipridae*) from Brazil. *The Auk*. 119: 1114-20.
711 Matthiopoulos J, Harwood J, Thomas L. 2005. Metapopulation consequences of site fidelity
712 for colonially breeding mammals and birds. *Journal of Animal Ecology* 74: 716–727.
713 Maurer D, Raikow RJ. 1981. Appendicular myology, phylogeny, and classification of the avian
714 order Coraciiformes (including Trogoniformes). *Annals of Carnegie Museum Natural*
715 *History* 50: 417-434.
716 Mayr E, O'Hara RJ. 1986. The biogeographic evidence supporting the Pleistocene forest
717 refuge hypothesis. *Evolution* 40: 55-67.
718 McKay BD, Zink RM. 2010. The causes of mitochondrial DNA gene tree paraphyly in birds.
719 *Molecular Phylogenetics and Evolution* 54: 647-50.
720 McLachlan GR, and Liversidge R. 1957. Roberts birds of South Africa. Cape Times Limited,
721 Cape Town.
722 Miya M, Nishida M. 1996. Molecular phylogenetic perspective on the evolution of the deep-
723 sea fish genus Cyclothone (*Stomiiformes: Gonostomatidae*). *Ichthyological Research* 43:
724 375-98.
725 Moreau RE. 1966. The Bird Faunas of Africa and Its Islands. Academic Press, London.
726 Moyle RG. 2006. A molecular phylogeny of kingfishers (Alcedinidae) with insights into early
727 biogeographic history. *The Auk* 123: 487-99.
728 Mucci N, Mengoni C, Randi E. 2017. Discrimination of PCR products by colour and size
729 improves the accuracy of sex-typing in avian species. *Conservation genetics resources* 9:
730 73-7.
731 Nesbitt SA, Spalding MG, Candelora KL, Kubilis PS, Schwikert ST. Body mass index (BMI) of
732 normal sandhill cranes. *Proceedings of the North American Crane Workshop*, 10: 133-137.
733 Pérez-Tris J, Tellería JL. 2001. Age-related variation in wing shape of migratory and sedentary
734 Blackcaps Sylvia atricapilla. *Journal of Avian Biology* 32: 207-13.

735 Perktaş U, Groth JG, Barrowclough GF. 2020. Phylogeography, Species Limits, Phylogeny,
736 and Classification of the Turacos (Aves: Musophagidae) Based on Mitochondrial and
737 Nuclear DNA Sequences. *American Museum Novitates* 3949: 1-61.
738 Pinho C, Hey J. 2010. Divergence with gene flow: models and data. *Annual Reviews of*
739 *Ecology, Evolution and Systematics* 41: 215–230.
740 Ramenofsky M. 1990. Fat storage and fat metabolism in relation to migration. In: Gwinner E
741 (ed) Bird *migration: physiology and ecophysiology*. Springer, Berlin Heidelberg New York,
742 214–231.
743 Rice WR, Hostert EE. 1993. Laboratory experiments and speciation: what have we learned in
744 40 years? *Evolution* 47: 1637–1653.
745 Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of
746 pairwise genetic differences. *Molecular Biology and Evolution* 9(3):552-69. doi:
747 10.1093/oxfordjournals.molbev.a040727.
748 Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under
749 mixed models. *Bioinformatics* 19:1572-1574.
750 Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism
751 analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-7.
752 Schneider S, Roessli D, Excoffier L. 2000. *Arlequin ver 2.000: a software package for*
753 *population genetics data analysis* [Users manual]. Geneva: University of Geneva, Genetics
754 and Biometry Laboratory. http://lgb.unige.ch/arlequin/software/2.000/manual/Arlequin.pdf.
755 Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P. 2011. Magic traits in
756 speciation:'magic'but not rare?. *Trends in ecology and evolution* 26: 389-97.
757 Seutin G, White BN, Boag PT. 1991. Preservation of avian blood and tissue samples for DNA
758 analyses. *Canadian journal of zoology* 69: 82-90.
759 Sibley CG, Ahlquist JE. 1990. *Phylogeny and classification of birds: A study in molecular*
760 *evolution.* Yale University Press, New Haven.
761 Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP. 1999. Primers for a PCR-based
762 approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular*
763 *phylogenetics and evolution* 12: 105-14.
764 Stephens M, Smith N, Donnelly P. 2001. A new statistical method for haplotype reconstruction
765 from population data. *American Journal of Human Genetics* 68: 978-989.
766 Stewart JR, Lister AM, Barnes I, Dalén L. 2010. Refugia revisited: individualistic responses of
767 species in space and time. *Proceedings of the Royal Society B: Biological Sciences* 277:
768 661-71.
769 Taylor SA, Friesen VL. 2012. Use of molecular genetics for understanding seabird evolution,
770 ecology and conservation. *Marine Ecology Progress Series* 451: 285–304.

771 Taylor RS, Friesen VL. 2017. The role of allochrony in speciation. *Molecular Ecology* 26: 3330-
772 42.
773 Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations
774 with haplotypes inferred from restriction endonuclease mapping and DNA sequence data.
775 III. Cladogram estimation. *Genetics* 132: 619-33.
776 Toews DP, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in
777 animals. *Molecular ecology* 21: 3907-30.
778 Wallin H, Kvamme T, Bergsten J. 2017. To be or not to be a subspecies: description of
779 *Saperda populnea lapponica* ssp. n. (Coleoptera, Cerambycidae) developing in downy
780 willow (*Salix lapponum L.*). *ZooKeys*. 691: 103.
781 Wang LS, Xu Y. 2003. Haplotype inference by maximum parsimony. *Bioinformatics* 19: 1773–
782 1780.
783 Wright S. 1943. Isolation by distance. *Genetics*: 28, 114–138.
784 Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations
785 in global climate 65 Ma to present. *Science* 292: 686-93.
786

787 **Tables:**

788 **Table 1:** Descriptive statistics and ANOVA analysis results for evaluating measurement errors
789 (A) and principal components analysis results (B) of morphological variation for the two
790 subspecies for untransformed and log-transformed data. Indicated are eigen values,
791 percentage of variance and percentage of cumulative variance for each PC as well as factor
792 loading for each analysed variable. Significant F ratio values (critical value=3.20, α=0.05) and
793 PCA factor loadings are indicated in bold font.

794 A

Measurement	Head	N	Min	Max	Mean	Std. error	Stand. Dev	F ratio
Head length	All samples	52	23.55	37.24	31.98	0.32	2.31	0.56
	Sampler 1	35	23.55	35.60	32.08	0.38	2.26	
	Sampler 2	5	29.70	31.63	30.82	0.35	0.78	
	Sampler 3	7	27.93	37.24	32.13	1.04	2.74	
Tarsus length	All samples	52	13.30	21.49	17.78	0.22	1.62	**8.19**
	Sampler 1	35	13.98	19.26	17.70	0.18	1.04	
	Sampler 2	5	13.30	19.13	15.30	1.09	2.44	
	Sampler 3	7	17.78	21.49	19.58	1.50	1.34	
Wing length	All samples	52	96.00	120.00	109.89	0.92	6.65	**9.37**
	Sampler 1	35	96.00	119.00	108.83	1.06	6.27	
	Sampler 2	5	98.00	120.00	108.20	3.83	8.56	
	Sampler 3	7	111.10	120.00	116.16	1.29	3.41	
Tail length	All samples	52	60.00	83.00	66.08	0.60	4.34	2.73
	Sampler 1	35	61.00	69.00	64.71	0.42	2.49	
	Sampler 2	5	64.00	71.00	67.40	1.29	2.88	
	Sampler 3	7	64.00	79.00	67.86	1.99	5.27	
Mass	All samples	52	49.00	83.00	62.18	1.14	8.19	1.64
	Sampler 1	35	49.00	88.00	61.56	1.39	8.21	
	Sampler 2	5	50.90	65.40	59.06	2.83	6.33	
	Sampler 3	7	58.00	87.00	67.43	3.52	9.31	

795

796 **B**

Variable/ Estimate	Untransformed data						Log transformed data					
	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
Head length	0.05	-0.01	0.15	0.99	-0.05	0.00	0.09	0.17	**0.96**	0.01	0.17	0.08
Tarsus length	0.11	0.05	-0.03	0.05	0.97	0.20	0.25	**0.63**	-0.23	-0.05	0.66	-0.23
Wing length	**0.58**	**0.61**	-0.53	0.05	-0.11	0.00	0.20	**0.24**	-0.14	0.34	-0.01	**0.88**
Tail length	0.13	**0.57**	**0.80**	-0.12	-0.01	0.00	0.05	0.13	0.04	**0.88**	-0.21	-0.39
Mass	**0.80**	**-0.54**	0.24	-0.08	-0.04	-0.05	**0.71**	0.25	0.00	-0.28	-0.58	-0.13
BMI	0.02	-0.04	0.02	-0.02	-0.21	0.98	**0.61**	-0.67	0.00	0.15	0.40	-0.02
% variance	66.27	19.86	9.19	3.59	1.08	0.00	51.79	26.04	9.64	6.89	3.25	2.38

797

Table 2: Molecular diversity indices for the concatenated mitochondrial DNA (2277 bp) and nuclear DNA (1505 bp) data sets for *Halcyon senegalensis senegalensis* (Gabon, Ghana and Uganda) and *H. s. cyanoleuca* (South Africa). Results that were not significant (P>0.05) are indicated as *ns,* while *nHap* refers to the number of haplotypes.

Population	***n***	***nHap***	***Tajima's D***	**Haplotype diversity (h)**	**Nucleotide diversity (*π*)**	**P distance (intra) (%)**	**Population Pairwise Comparisons (F_{ST})**	
							Ghana	**Uganda**
Mitochondrial DNA (Cyt b, COI and 16S)								
1. Gabon	1	1	N/A	N/A	N/A	N/A	N/A	N/A
2. Ghana	10	4	0.62 ns	0.7333±0.1199	0.0007±0.0005	1.6444±1.0553	-	
3. Uganda	8	5	1.09 ns	0.8571±0.1083	0.0010±0.0007	2.3928±1.4504	0.802	-
4. South Africa	31	9	-1.16 ns	0.5849±0.1032	0.0005±0.0004	1.2301±0.8025	0.949	0.940
Total	**50**	**19**	**N/A**	**0.7251±0.1105**	**0.0007±0.0005**	**N/A**	**-**	**-**
Nuclear DNA (FIB5 and RAG1)								
1. Ghana	20	19	1.64 ns	0.9947±0.0178	0.0111±0.0058	16.72±7.77		
2. Uganda	16	15	0.95 ns	0.9917±0.0254	0.0108±0.006	16.25±7.65	0.028 *ns*	
3. South Africa	56	43	0.56 ns	0.9867±0.0112	0.0087±0.0044	13.08±5.99	0.263	0.209
Total	**92**	**77**	**N/A**	**0.9910±0.0181**	**0.0102±0.0054**	**N/A**	**-**	**-**

Table 3: Hierarchical analysis of molecular variance (AMOVA) among concatenated mitochondrial (CYTB, COI and 16S) and nuclear DNA (RAG1 and FIB5) sequences. AMOVA of mtDNA sequences was conducted for variation among different groupings including 1. subspecies and 2. geographic region. Results that were not significant (P>0.05) are indicated as *ns*.

Structure tested	Source of variation	Degrees of freedom	Sum of squares	Variance components (P<0.05)	Percentage of variation (%)
Mitochondrial DNA					
1. Subspecies	Among populations	1	256.26	8.16	64.83 *ns*
	Among populations within groups	2	39.49	3.69	29.26 *ns*
	Within populations	46	34.23	0.74	5.91
	Total	**49**	**329.96**	**12.59**	**100**
2. Geographic region (South Africa, Uganda, Gabon+Ghana)	Among populations	2	293.52	9.70	83.88 *ns*
	Among populations within groups	1	2.78	1.12	9.69 *ns*
	Within populations	46	34.23	0.74	6.43
	Total	**49**	**329.96**	**11.56**	**100**
Nuclear DNA					
1. Subspecies	Among populations	1	95.01	1.91	20.22 *ns*
	Among populations within groups	1	12.43	0.29	3.07 *ns*
	Within populations	81	588.20	7.26	76.71
	Total	**83**	**695.63**	**9.46**	**100**

Table 4: Estimates of percentage sequence divergence between species, subspecies and geographic lineages. Standard error estimate(s) are shown above the diagonal.

	H. s. cyanoleuca South Africa	*H. s. senegalensis* Gabon	*H. s. senegalensis* Ghana	*H. s. senegalensis* Uganda	*H. smyrnensis*	*H. pileata*
H. s. cyanoleuca South Africa		0.003	0.003	0.002	0.007	0.006
H. s. senegalensis Gabon	1.1%		0.0000	0.001	0.007	0.006
H. s. senegalensis Ghana	1.2%	0.2%		0.001	0.007	0.007
H. s. senegalensis Uganda	1.1%	0.4%	0.4%		0.007	0.007
H. smyrnensis	9.3%	9.3%	9.3%	9.1%		0.004
H. pileata	8.4%	8.7%	8.8%	8.7%	3.9%	

Table 5: Details of estimate nodes ages for the tree for the Halcyoninae clade. The confidence interval (95% HPD interval) represents the highest-posterior-density interval containing 95% of the posterior distribution.

Nodes	Approx. node age (MYA)	Height 95% HPD		Approx. node age range (MYA)	
Halcyoninae	11.7881	0.0571	0.0761	13.42	10.27
	10.8281	0.0688	0.1120	12.19	9.52
	9.4034	0.0530	0.1044	10.92	7.95
Halcyon	8.1092	0.0490	0.0863	9.31	6.96
	7.0546	0.0429	0.0759	8.21	6.02
	2.5606	0.0121	0.0335	3.20	1.96
H. senegalensis	0.9696	0.0042	0.0152	1.31	0.66
H. s. senegalensis	0.3868	0.0010	0.0070	0.57	0.22
H. s. cyanoleuca	0.2218	0.0006	0.0057	0.36	0.12

Figures:

Figure 1: Geographical distribution of Woodland Kingfisher (*Halcyon senegalensis*) in Africa. Geographic distribution range of *H. s. senegalensis* is indicated by a solid black line, *H. s. fuscopileus* is indicated by a dotted black line and *H. s. cyanoleuca* is indicated with a dashed black line (adapted from Woodall, 2001, in del Hoyo *et al*. eds., 2014).

Figure 2: Scatterplot of the scores from PC1 and PC2 for (A) untransformed and (B) log-transformed morphometric characters of male and female *H. s. senegalensis* (Ghana and Uganda) and *H. s. cyanoleuca* (South Africa). Circles indicate females and squares indicate males. *H. s. cyanoleuca* is shown with black diamonds while *H. s. senegalensis* (Ghana) is indicated with a white triangle and *H. s. senegalensis* (Uganda) is demonstrated with a grey triangle.

Figure 3: Phylogeny of the concatenated mtDNA sequences for *Halcyon senegalensis senegalensis* (Gabon, Ghana and Uganda) and *H. s. cyanoleuca* (South Africa). A. Haplotype network. Different haplogroups are shown in different colours, where circle size corresponds to the haplotype frequency and the number of dashes indicated number of mutation steps between haplotypes. B. Maximum likelihood tree of concatenated mitochondrial sequences based on the Hasegawa-Kishino-Yano (HKY+G) model conducted in MEGA7, showing the two subspecies and their populations, South Africa (red), Gabon (yellow), Ghana (green) and Uganda (blue).

Figure 4: Divergence time estimation with BEAST based on concatenated mtDNA sequence data for Woodland Kingfisher.

Supplementary:

Supplementary Table 1: DNA sequences of all species included in phylogenetic and genetic diversity analyses as well as haplotype network inference per region for (A) mitochondrial DNA and (B) nuclear DNA regions. Cytochrome oxidase 1 = COI, Cytochrome *b* = CYTB, large subunit ribosomal RNA = 16S, recombination activating gene 1 = RAG1 and nuclear b-fibrinogen gene, intron 5 = FIB5.

Supplementary Figure 1: Maximum likelihood tree of (A) cytochrome oxidase 1 (COI) and (B) cytochrome *b* (CYTB) based on the Hasegawa-Kishino-Yano (HKY +G) model, (C) large subunit ribosomal RNA (16S) and (D) recombination activating gene 1 (RAG1) based on the Kimura 2-parameter (K2) model and (E) nuclear b-fibrinogen gene, intron 5 (FIB5) based on

the Jukes–Cantor (JC) one-parameter model. All trees were constructed in MEGA7 and numbers below branches indicate bootstrap values.

Supplementary Figure 2: Bayesian inference (BI) analysis of a (A) three-locus mitochondrial (COI, CYTB and 16S) dataset and a (B) two-locus (RAG1 and FIB5) dataset. Bayesian posterior probabilities are indicated above the branches.

Supplementary Figure 3: Haplotype network of concatenated nuclear DNA gene regions (RAG1 and FIB5). Different haplogroups are shown in different shades of grey, where circle size corresponds to the haplotype frequency and the number of dashes indicated number of mutation steps between haplotypes.

Supplementary Figure 4: Mismatch distributions of the *H. senegalensis* based on concatenated mitochondrial DNA (a) and nuclear DNA (b) regions under a model of sudden population expansion as estimated in Arlequin.