DNA-identification of the first Dutch Siberian northern shrike *Lanius borealis sibericus* from a toepad sample

Peter de Knijff and Marije te Raa.
Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.

This document with public access URL https://doi.org/10.5281/zenodo.6982812 was prepared as supplementary information to the following article to appear in Dutch Birding:
Siberische Noordelijke Klapekster bij Nuenen in november 1909; written by Justin J F J Jansen, Pepijn Kamminga, Martin Brandsma, Reuven Yosef & Peter de Knijff.

Summary
On 25 November 1909, a grey shrike *Lanius* was purchased by Rijksmuseum voor Natuurlijke Historie (now Naturalis Biodiversity Center) at Leiden, Zuid-Holland, Netherlands (skin RMNH.AVES1663). The bird, a first-winter female, was collected at Nuenen, Noord-Brabant, Netherlands. It was at the time identified as Great Grey Shrike *L excubitor*. The bird was recently re-identified as a Northern Shrike *L borealis* (*L b borealis* or *L b sibericus*). Here we describe all relevant technical aspects of the DNA-study that was performed in order to support this identification. For this we used a toe pad sample. We obtained a mitochondrial Cytochrome B fragment of 893 bp. which gave a 100% match with *L b sibericus*. This was the first record for the Netherlands, and chronologically the second for Europe.

Introduction
Since museum-collection toepad samples are generally old and very dry, the DNA (if any) therein is normally degraded and one needs a series of short and overlapping DNA-fragments to obtain a sufficiently long-enough DNA fragment allowing a comparison with known reference sequences. In this case, we needed a total of 14 different primers by means of which 14 overlapping DNA-fragments could be amplified and sequenced (see Table 1 and Figure 1 for details).
We used this protocol on three different samples. First, two fresh feather-samples from Great Grey Shrike *L excubitor*, that were previously received from bird-banding station Castricum, were used to test the full amplification and sequencing strategy. Subsequently, the third sample was the toepad sample from the possible Northern Shrike *L borealis* (*L b borealis* or *L b sibericus*) skin present in the Naturalis collection (RMNH 1663).

The three final sequences CytB sequence fragments of 893 bp. were deposited in GenBank under numbers MW775037 (*L excubitor* , Castricum 19 October 2013), MW775038 (*L excubitor*, Castricum 31 October 2014), and MW775039 (L borealis, Nuenen November 1909, RMNH aves 1663).

<u>DNA-sequencing</u>

DNA-isolation from toepad or feather samples was performed using the QIAamp DNA mini and blood mini kit. Toepad, feathers or parts of feathers were put in a 2mL tube containing 400µL ATL buffer. 10µL ProtK and 3µL DTT (1M) were added. Samples were incubated at 56°C. After 2 hours incubation, samples were checked to see whether feather pieces were still visible. If so, another 10µL ProtK and 3µL DTT (1M) was added. When after 4 hours of incubation, no feather pieces were visible, extraction was continued. If pieces were still visible the incubation was left overnight and extraction was continued by adding 400µL AL buffer, vortexing and 10 minutes incubation at 70°C. Afterwards 400µL 100% ethanol was added. The mixture was applied to the column according to step 6 of the QIAamp DNA Mini and Blood Mini Handbook Protocol: DNA Purification from Dried Blood Spots. The extraction was completed following the remainder of the steps in this protocol except for the elution which was performed with 80µL nuclease free water with a centrifuge step of 8000rpm for 2 minutes in soft mode.

DNA-sequencing (Sanger)

PCR reaction

In order to sequence a fragment of 893 bp. of the *Lanius* mitochiondrial cytB gene, a series of overlapping Monoplex PCR's were performed with an input of 2.5-5µL of DNA-extract. See Table 1 for PCR-primer sequences and Figure 1 for a schematic layout of the various monoplexes. The PCR-mix contained Geneamp™ 10x PCR-buffer1 (Applied biosystems), 0.2mM dNTP's (GE healthcare), 0.4pmol per primer (tabel1) and 2U AmpliTAq gold DNA polymerase (Applied biosystems) in a total volume of 50µL. PCR's were run on a GeneAmp® PCR System 9700 with the following program. 94°C for 10 min, 36 cycles of 94°C for 30 seconds, 50°C for 30 seconds and 72°C for 2 minutes ending with 72°C for 10 minutes.

PCR purification

The PCR-products were visualized using the QIAxcel. Afterwards a purification step using the QIAquick® PCR purification kit (QIAGEN) was performed according to the protocol from this kit. Elution was performed in 20-70µL Aquabraun depending on the amount of PCR-product visible on the Qiaxcel (Figure 2).

Sequencing PCR

Forward and reverse sequencing PCR's were performed using an input of 1-4µL of purified PCR-product (samples eluted in 50-70µL water had an input of 1µL in the sequencing PCR. Samples eluted in 20-30µL had an input of 4µL in the sequencing PCR). PCR-mix contained: 5x sequencing buffer big dye terminator V1.1 v3.1 (Life Technologies), 0.6pmol sequencing primer (tabel2) and 2µL BigDye® Terminator v3.1 ready reaction mix (Life Technologies) in a total volume of 10µL. Sequencing PCR's were run on a pre-heated (96°C) GeneAmp® PCR System 9700 with the following program. 96°C for 1 min, 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes.

Sequencing PCR purification

12µL water was added to the sequencing PCR-product. Then the product was purified using the DyeEx® 2.0 Spin kit (QIAGEN), protocol for Dye-Terminator Removal.

Sequencing

Purified sequencing PCR-product was run on an AB3100 Genetic Analyzer.

Table1: Forward and reverse primer sequences for the CytB sequencing of *Lanius excubitor/borealis* samples. Lowercase sequences tgtaaaacgacggccagt and caggaaacagctatgacc are M13 tails.

Primer	Reverse primer sequence
Primer 1F	tgtaaaacgacggccagtTTCAGCTCCGTAGCCCATA
Primer 2F	tgtaaaacgacggccagtATCGGCCGAGGRCTCTACTA
Primer 2R	caggaaacagctatgaccTAGTAGAGYCCTCGGCCGAT
Primer 3F	tgtaaaacgacggccagtTGAGGGGCTACAGTTATCACT
Primer 3F	caggaaacagctatgaccAGTGATAACTGTAGCCCCTCA
Primer 4F	tgtaaaacgacggccagtGACAACCCAACWCTCACTCG
Primer 4R	caggaaacagctatgaccCGAGTGAGWGTTGGGTTGTC
Primer 5F	tgtaaaacgacggccagtCATGCTAGGAGACCCAGAAA
Primer 5R	caggaaacagctatgaccTTTCTGGGTCTCCTAGCATG
Primer 6F	tgtaaaacgacggccagtCCCCAACAAACTAGGAGGRGT
Primer 6R	caggaaacagctatgaccACYCCTCCTAGTTTGTTGGGG
Primer 7F	tgtaaaacgacggccagtCTGAGCYCTAGTAGCCAACA
Primer 7R	caggaaacagctatgaccTGTTGGCTACTAGRGCTCAG
Primer 8R	caggaaacagctatgaccCTAGCACRCTTGCGATAGG

Table2: Overview of the forward and reverse primer sequences for the sequencing PCR

	Forward	Reverse
Sequencing PCR primers	TGTAAAACGACGGCCAGT	CAGGAAACAGCTATGACC

Figure 1: PCR and sequence primer strategy

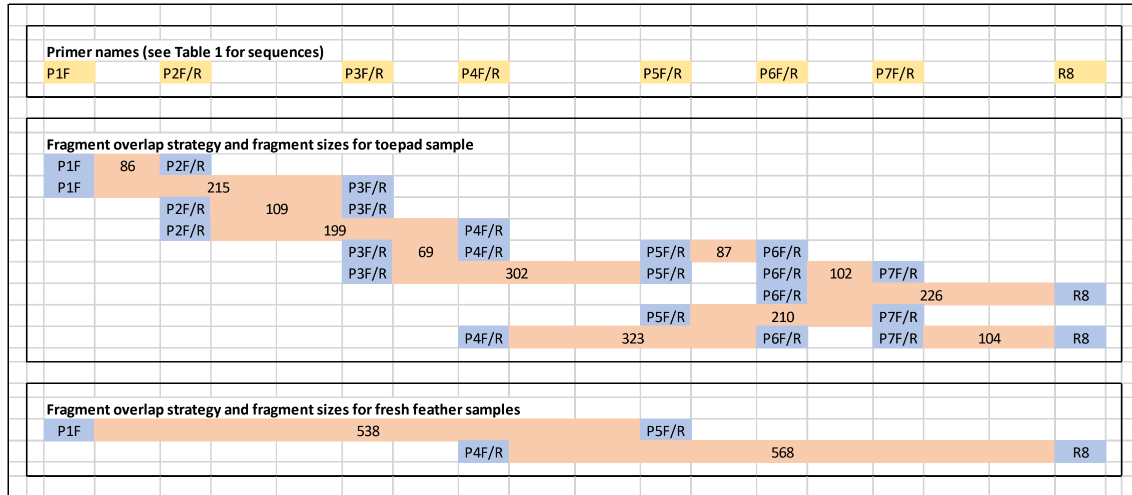
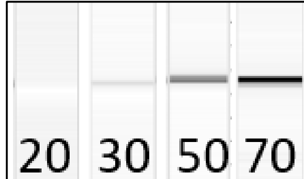


Figure 2: example of elution volumes (µL) after purification based on Qiaxcel results.



Network preparation and analyses

As a first step, we combined the following CytB sequences, obtained from GenBank and our own study (the last three sequences) into a single fasta file, aligned the sequences and clipped them to an overlapping size of 893 bp:

GU253498.1 Lanius excubitor excubitor voucher UWBM:57238
GU253499.1 Lanius excubitor excubitor voucher UWBM:74439
GU253500.1 Lanius excubitor excubitor voucher UWBM:59641
GU253501.1 Lanius excubitor excubitor voucher NRM:20076397
GU253502.1 Lanius excubitor excubitor voucher NRM:20016213
GU253503.1 Lanius excubitor excubitor voucher NRM:20036012
GU253504.1 Lanius excubitor excubitor voucher NRM:996541
GU253505.1 Lanius excubitor excubitor
GU253508.1 Lanius excubitor homeyeri voucher NRM:569603
GU253531.1 Lanius excubitor mollis voucher NHM 1898.9.20.1090
GU253541.1 Lanius excubitor sibiricus voucher NRM:568489
GU253542.1 Lanius excubitor sibiricus voucher NRM:556595
GU253543.1 Lanius excubitor sibiricus voucher NRM:556596
GU253545.1 Lanius excubitor sibiricus voucher NHM 1898.9.20.508
GU253480.1 Lanius excubitor bianchii voucher UWBM:47014
GU253481.1 Lanius excubitor borealis voucher UWBM:60536
GU253482.1 Lanius excubitor borealis voucher UWBM:62203
GU253483.1 Lanius excubitor borealis voucher BMNH X7458
GU253484.1 Lanius excubitor borealis voucher BMNH X7672
GU253485.1 Lanius excubitor borealis voucher BMNH AWJ143
GU253486.1 Lanius excubitor borealis voucher BMNH JK95028

GU253487.1 Lanius excubitor borealis voucher BMNH JDW0035
GU253488.1 Lanius excubitor borealis voucher BMNH X7530
GU253489.1 Lanius excubitor borealis voucher BMNH X7756
GU253490.1 Lanius excubitor borealis voucher BMNH X7780
MW775037 Lanius excubitor FLDO-KE01, Castricum 19-October-2013
MW775038 Lanius excubitor FLDO-KE02, Castricum 31 October-2014
MW775039 Lanius borealis sibericus] Specimen_voucher=RMNH:AVES:1663] FLDO-KE03, Nuenen November 1909

As a second step, we used DnaSP v6 (from http://www.ub.edu/dnasp/), to extract and export a polymorphic sites only file in RDF-format from this fasta file. This .rdf file was used to prepare a median joining network using Network 10.2 (from https://www.fluxus-engineering.com/sharenet.htm). The resulting network was adjusted using Network Publisher (purchased via https://www.fluxus-engineering.com/sharenet.htm), and subsequently exported as an .emf file. This .emf file was edited in PowerPoint which resulted in Figure 3 below.

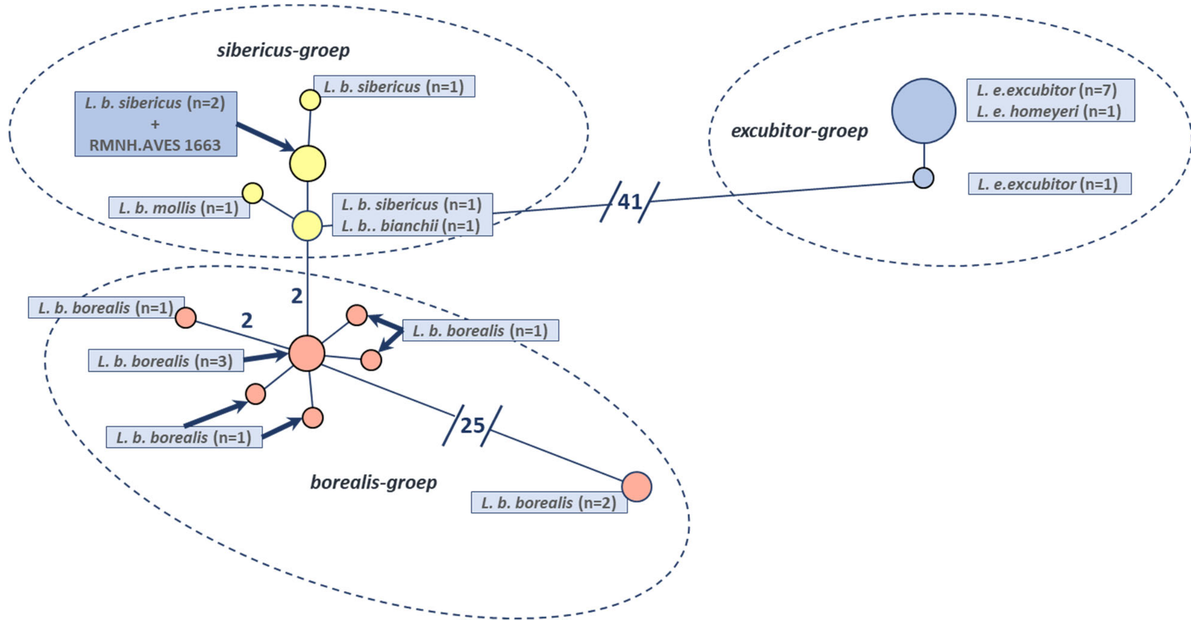


Figure 3: Network of variation in mtDNA cytochrome B gene fragment (cytB) of 26 grey shrike Lanius taxa. For this network, 25 reference sequences (Olsson et al 2010) and skin RMNH.AVES1663 were compared. Each circle represents unique cytB sequence of 893 base pairs. Label at each circle gives information on taxon and number of times that this specific sequence was observed. Relative diameter of each circle is indication of its frequency in total dataset. Short lines between circles without number mark differences on one position (of 893 positions) only, other longer lines mark two or more (as indicated by numbers along lines) different positions. Sequence of RMNH.AVES.1663 matches exactly with two analysed L borealis sibericus.