Supporting files for the manuscript ”Speciation and population divergence in a mutualistic seed dispersing bird”

**Merged\_VCF.filtered\_scaffname\_MAF0.1.vcf.gz**

Contains all nutcracker individuals (Fig 1)

**Caryocatactes\_filtered\_mmis0.96.vcf.gz**

Contains individuals belonging to the species *N. caryocatactes* (Northern) (Fig 3)

**Hemispila\_filtered\_mmis0.96.vcf.gz**

Contains individuals belonging to the species *N. caryocatactes* (Southern) (Fig 4)

VCF files were filtered with VCFtools v 0.1.17. (Danecek et al 2011) for biallelic positions and a minimum and maximum coverage ranging from 8X to 2.5 times mean depth of the respective sample. To account for the bias resulting from mapping to an in-group reference genome, we applied a strict filtering approach in which at maximum one individual was allowed to have missing data at each site. VCF-files were converted to plink’s PED format with VCFtools v 0.1.17 and principal component analyses was subsequently performed with Plink version 1.9 (Purcell 2007). To account for linkage disequilibrium, sites within a 50-SNP stepping window with a correlation coefficient higher than 0.1 were omitted.

**Morphometric\_data.zip** (Fig 1, 3, 4)

12 morphological traits were measured on nutcracker specimens from museum collections. Morphological traits associated with nutcracker foraging and diet, flight ability, manoeuvrability and bipedal locomotion were used (detailed description of all measures in Supplementary Table 10). Traits were measured on a total number of 118 specimens from the Zoological Research Museum Alexander Koenig, Bonn, Germany, the Natural History Museum, Berlin, Germany, Senckenberg Natural History Collections, Dresden, Germany and the Senckenberg Forschungsinstitut und Naturmuseum, Frankfurt, Germany. The final data set include the 90 specimens with the least amount of missing data.

**Fig2\_treefiles.zip** (Fig 2)

The consensus function of Bcftools v.1.2 (Li et al., 2009) was used to call concensus sequences for all individuals. Phylogenomic inference of the nuclear genome consisted of removal of heterozygous sites and N’s in the consensus sequences for all individuals using bedtools v2.28.0. Consensus sequences were subsequently aligned using custom designed scripts available at <https://github.com/mobilegenome/phylogenomics/>. Maximum likelihood inference was performed with IQ-TREE v.1.6.11, whereas species tree inference based on the multispecies coalescent (MSC) was performed with ASTRAL-III v5.6.1 (Zhang et al., 2018). Network analyses was performed using SplitsTree v.4.15.1 (Huson & Bryant 2006). The file contains 192 individual treefiles generated with IQ-TREE that were used for the phylogenetic analyses.