DATA NOTE

The complete genome sequences of 22 parrot species
(Psittaciformes, Aves) [version 1; peer review: 2 approved with reservations]

Taylor Hains¹, Kathleen O'Neill², Jafet Velez³, Nancy Speed⁴, Susan Clubb⁴, Taras Oleksyk⁵, Stacy Pirro¹

¹Parrot Conservation, Terra Wildlife Genomics, Washington, DC, 20037, UK
²Biodiversity, Iridian Genomes, Bethesda, MD, 20817, USA
³Conservation Program of the Puerto Rican Parrot, U.S. Fish and Wildlife Service, Rio Grande, PR, 00745, USA
⁴Rainforest Clinic for Birds & Exotics, Hurricane Aviaries Inc, Loxahatchee, FL, 33470, USA
⁵Department of Biological Sciences, Oakland University, Rochester, MI, 48309, USA

Abstract
The parrots (Psittaciformes, Aves) are a group of colorful, intelligent, long-lived birds with a wide range of body sizes and plumage colors and patterns. One third of the parrot species is threatened with extinction due to habitat loss and the pet trade, a larger percentage than any other comparable bird order. We present the complete genome sequences of 22 species of parrots from 14 genera and 3 families: Anodorhynchus hyacinthinus, Ara ararauna, Ara chloropterus, Ara glaucogularis, Ara militaris, Aratinga solstitialis, Aratinga weddellii, Cacatua leadbeateri, Ectlectus roratus, Guaruba guarouba, Loriculus garrulus, Myiopsitta monachus, Nymphicus hollandicus, Pionus senilis, Psittacus erithacus, Psittacus timneh, Psitteutes goldiei, Pyrrhura frontalis, Pyrrhura griseiceps, Pyrrhura molinae, Pyrrhura perlata. Genomic data can be used to better understand species identity, hybridization, genetic diversity, and identification of animal products possibly derived from endangered species.

Keywords
Parrots, birds, Psittaciformes, cockatoos, genomes

This article is included in the Genome Sequencing gateway.
Introduction
The parrots (Psittaciformes, Aves) are an order of 370 extant species (Clements et al., 2019) characterized by high intelligence (Pepperberg, 2017), diverse plumage colors and patterns (Berg & Bennett, 2010), and clawed zygodactyl feet (Botelho et al., 2014). A third of all parrot species are threatened by extinction due to habitat loss and the practice of capturing young birds for the pet trade (Butchart et al., 2004).

Complete genome sequences for parrot species will benefit such diverse fields as biodiversity, evolutionary studies, species hybridization, and research into the development of brain function. Genomic data will also aid in the development of molecular markers to identify products, which may be derived from endangered species (Yang et al., 2018).

Methods
Samples from 22 parrot species were obtained from molted feathers from pet birds, and blood samples from birds from captive breeding programs (Table 1). Where blood samples were used, they were obtained by a licensed veterinarian coinciding with a standard health check. Animals were handled in a manner consistent with accepted practices (NRC, 2011). In the case of molted feathers, pulp was excised from the shaft of a large tail or primary wing feather prior to extraction.

Table 1. Species sequenced for this study.

<table>
<thead>
<tr>
<th>Scientific name</th>
<th>Common name</th>
<th>Accession number</th>
<th>Genome size</th>
<th>Sample type</th>
<th>Family name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anodorhynchus hyacinthinus</td>
<td>Hyacinth Macaw</td>
<td>GCA_009936445</td>
<td>1,119,461,287</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Ara ararauna</td>
<td>Blue-And-Yellow Macaw</td>
<td>GCA_010014805</td>
<td>1,128,178,780</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Ara chloropterus</td>
<td>Red-And-Green Macaw</td>
<td>GCA_01001472</td>
<td>1,117,846,189</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Ara glaucogularis</td>
<td>Blue-Throated Macaw</td>
<td>GCA_010015425</td>
<td>1,135,105,966</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Ara militaris</td>
<td>Military Macaw</td>
<td>GCA_01001596</td>
<td>1,177,230,593</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Aratinga solstitialis</td>
<td>Sun Parakeet</td>
<td>GCA_009867195</td>
<td>1,214,690,015</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Aratinga weddellii</td>
<td>Dusky-headed Parakeet</td>
<td>GCA_009867205</td>
<td>1,114,053,963</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Cacatua leadbeateri</td>
<td>Major Mitchell's Cockatoo</td>
<td>GCA_010015045</td>
<td>1,069,984,771</td>
<td>feather</td>
<td>Cacatuidae</td>
</tr>
<tr>
<td>Eclectus roratus</td>
<td>Eclectus Parrot</td>
<td>GCA_01176353</td>
<td>1,101,365,900</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Eupsittula pertinax</td>
<td>Brown-Throated Parakeet</td>
<td>GCA_011317185</td>
<td>1,171,654,275</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Guaruba guarouba</td>
<td>Golden Parakeet</td>
<td>GCA_011800095</td>
<td>1,151,548,123</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Lories garrulus</td>
<td>Chattering Lory</td>
<td>GCA_010014925</td>
<td>1,091,728,784</td>
<td>feather</td>
<td>Psittaculidae</td>
</tr>
<tr>
<td>Myiopsitta monachus</td>
<td>Monk Parakeet</td>
<td>GCA_013036005</td>
<td>1,118,816,385</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Nymphaicus hollandicus</td>
<td>Cockatiel</td>
<td>GCA_009823435</td>
<td>1,205,077,083</td>
<td>feather</td>
<td>Cacatuidae</td>
</tr>
<tr>
<td>Pionus senilis</td>
<td>White-Capped Parrot</td>
<td>GCA_011762725</td>
<td>1,067,577,099</td>
<td>feather</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Psittacus erithacus</td>
<td>Grey Parrot</td>
<td>GCA_009867235</td>
<td>1,118,034,822</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Psittacus timneh</td>
<td>Timneh Parrot</td>
<td>GCA_009867315</td>
<td>1,103,923,287</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Psitteuteles goldiei</td>
<td>Goldie's Lorikeet</td>
<td>GCA_010014875</td>
<td>1,125,670,205</td>
<td>feather</td>
<td>Psittaculidae</td>
</tr>
<tr>
<td>Pyrrhura frontalis</td>
<td>Maroon-Bellied Parakeet</td>
<td>GCA_010014865</td>
<td>1,112,821,352</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Pyrrhura griseicapctus</td>
<td>Gray-Breasted Parakeet</td>
<td>GCA_010014965</td>
<td>1,129,186,538</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Pyrrhura molinae</td>
<td>Green-Cheeked Parakeet</td>
<td>GCA_011763355</td>
<td>1,122,251,672</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
<tr>
<td>Pyrrhura perlata</td>
<td>Crimson-bellied Parakeet</td>
<td>GCA_011763455</td>
<td>1,111,465,322</td>
<td>blood</td>
<td>Psittacidae</td>
</tr>
</tbody>
</table>
DNA extraction was performed on blood and feather pulp using the Qiagen DNAeasy genomic extraction kit, using the manufacturer’s process. A paired-end sequencing library was constructed using the Illumina TruSeq kit, according to the manufacturer’s instructions. The library was commercially sequenced at Genewiz (New Jersey, USA) on an Illumina Hi-Seq platform in paired-end, 2 x 150bp format.

The resulting fastq files were trimmed of adapter/primer sequence and low-quality regions with Trimmomatic (v0.33) (Bolger et al., 2014). The trimmed sequence was assembled by SPAdes (v2.5) (Bankevich et al., 2012). This was followed by a finishing step using RagTag (v1.0.0) (Alonge, 2020) to make additional contig joins based on conserved regions in other parrot species: Melopsittacus undulatus (GCA_012275295), Amazona guildingii (GCA_013399615), Agapornis roseicollis (GCA_002631895), and other species in this study.

**Results**
The genome sizes and NCBI Assembly accessions for each species are shown in Table 1.

Genome sizes ranged from 1.07 – 1.21 G (scaffolded), with no clear delineation between the three represented families (Psittacidae, Psittaculidae, Cacatuidae).

**Data availability**
**Underlying data**
Accession numbers for all genome sequences in GenBank can be found in Table 1.

**References**

Publisher Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

Publisher Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

Reference Source

Reference Source

PubMed Abstract | Publisher Full Text

PubMed Abstract | Publisher Full Text | Free Full Text
Open Peer Review

Current Peer Review Status:  ?  ?

Version 1

Reviewer Report 08 April 2021

https://doi.org/10.5256/f1000research.28208.r82331

© 2021 McCormack J et al. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Devon DeRaad
Department of Ecology & Evolutionary Biology, Kansas University, Kansas, USA

John McCormack
Moore Laboratory of Zoology, Occidental College, Los Angeles, CA, USA

These are potentially very valuable data for avian genomics, as they increase the number of parrot genomes many times. These data will provide valuable resources for evolution and conservation genomics study.

This study has a few limitations, detailed below:

1. The biggest limitation is the inability to link the genomic information with certainty to a taxonomic name because of the lack of a voucher specimen and uncertainty about the possible hybrid status of pet parrots. The samples come from either pet birds or birds in captive breeding programs. While birds in captive breeding programs are more likely to be genetically pure, pet parrots could come from anywhere (even the illegal pet trade). Parrot species are known to hybridize in introduced habitats, in the wild, and in captivity—and can even produce viable offspring between genera (Hernandez-Brito et al. 2021 in Ibis\(^1\)). More details on provenance might alleviate some concerns, but it will be hard to get away from some uncertainty about whether these genomes are accurate representations of their taxonomic identifiers.

2. The genomes are not complete, since the assemblies come from Illumina short reads and likely do not cover certain, more difficult regions of the genome. Information on depth of sequencing and contiguity (e.g., N50) is available on NCBI, but should also appear in the table to allow readers to assess genome quality more directly.

3. The introduction would be enriched by a little more information on the rationale for the study. Specifically, what is currently missing from parrots studies that these genomes could provide? Describe existing genomic resources and then narrow to a knowledge gap that the authors hope to fill with these genomes. Was there a rationale for the specific species chosen?
Likewise, the Discussion would benefit from more elaboration on the broad utility of these datasets, e.g., mapping whole genome re-sequence data or reduced representation sequence data for population genetics studies, etc. Specifically, which uses might not be diminished by some uncertainty surrounding taxonomic assignment?

References

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
No

Are the datasets clearly presented in a useable and accessible format?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Phylogenetics, genomics, birds

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Reviewer Report 19 February 2021

https://doi.org/10.5256/f1000research.28208.r77387

© 2021 Mackiewicz P. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Paweł Mackiewicz
Department of Genomics, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

The manuscript was submitted to the section Data Note. It presents very briefly the description of nuclear genomes from 22 parrot species. Although the importance of these genomes is crucial for various analyses, the presentation should be improved.

The title indicates that the genomes are complete, whereas the obtained sequences are not
assembled and are still organized in scaffolds. Therefore, the title should be changed and the authors should describe in more detail the data set including information about incompleteness and gaps in the sequences. The limitations of the datasets should necessary be presented. The authors should state which type of analyses can be performed on such incomplete data set and which further steps should be performed to assembly the sequences. Could you answer for the following questions:

Which problems prevented the obtaining the complete genomes? Are the authors going to assembly and study these genomes in the future? Moreover, the authors should present in detail sources for each studied sample, e.g. zoo, natural environment with detailed localization, individual breeders, reservation etc. It is essential. Describe in detail the captive breeding programs in the context of the samples, which you mentioned. Ethics policies should be also included. Were the samples obtained and transferred according to the species protection law? It should be clearly stated. 

Although my review may seem harsh, actually I think that the obtained genomic sequences are valuable for various studies and the data set is interesting for scientific community. I am very willing to accept the manuscript after the inclusion of information and explanations that I mentioned in my review.

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: bioinformatics, genomics, phylogenetics, molecular evolution

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com