

Genomes of three closely related Caribbean Amazons provide insight for species history and conservation

Sofiia Kolchanova ^{1,2,3}, Sergei Kliver ^{3,*}, Aleksei Komissarov ³, Pavel Dobrinin ³, Gaik Tamazian ³, Kirill Grigorev ^{1,4}, Walter W. Wolfsberger ^{1,5,6}, Audrey J. Majeske ¹, Jafet Velez-Valentin ⁸, Ricardo Valentin de la Rosa ⁹, Joanne R. Paul-Murphy ¹⁰, David Sanchez-Migallon Guzman ¹⁰, Michael H. Court ¹¹, Juan L. Rodriguez-Flores ⁴, Juan Carlos Martínez-Cruzado ¹ and Taras K. Oleksyk ^{1,5,6,*}

¹ Department of Biology, University of Puerto Rico at Mayaguez, Mayaguez, PR 00680, USA; sofii.kolchanova@upr.edu (S.K.); kig2007@med.cornell.edu (K.G.); wwolfsberger@oakland.edu (W.W.W.); audrey.majeske@upr.edu (A.J.M.); juancarlos.martinez@upr.edu (J.C.M.-C.)

² Department of Biology, University of Konstanz, 78464 Konstanz, Germany

³ Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, 199034 St. Petersburg, Russia; ad3002@gmail.com (A.K.); pdobrynin@gmail.com (P.D.); gaik.tamazian@gmail.com (G.T)

⁴ Department of Genetic Medicine, Weill Cornell Medical College, New York, NY 10021, USA; jur2014@med.cornell.edu

⁵ Department of Biological Sciences, Oakland University, 118 Library Drive, Rochester, MI 48309, USA

⁶ Department of Biological Sciences, Uzhhorod National University, 88000 Uzhhorod, Ukraine

⁷ Beaumont BioBank, William Beaumont Hospital, Royal Oak, MI 48073, USA;

⁸ Conservation Program of the Puerto Rican Parrot, U.S. Fish and Wildlife Service, Rio Grande, PR 00745, USA; jafet_vez@fws.gov

⁹ The Recovery Program of the Puerto Rican Parrot at the Rio Abajo State Forest, Departamento de Recursos Naturales y Ambientales de Puerto Rico, Arecibo, PR 00613, USA; el.cotorro.electrico@gmail.com

¹⁰ Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California Davis, Davis, CA 95616, USA; paulmurphy@ucdavis.edu (J.R.P.-M.); guzman@ucdavis.edu (D.S.-M.G.)

¹¹ Program in Individualized Medicine (PrIme), Pharmacogenomics Laboratory, Department of Veterinary Clinical Sciences, College of Veterinary Medicine, Washington State University, 100 Grimes Way, Pullman, WA 99164, USA; michael.court@wsu.edu

* Correspondence: mahajrod@gmail.com (S.K.); oleksyk@oakland.edu (T.K.O.); Tel.: +1-248-370-3359 (T.K.O)

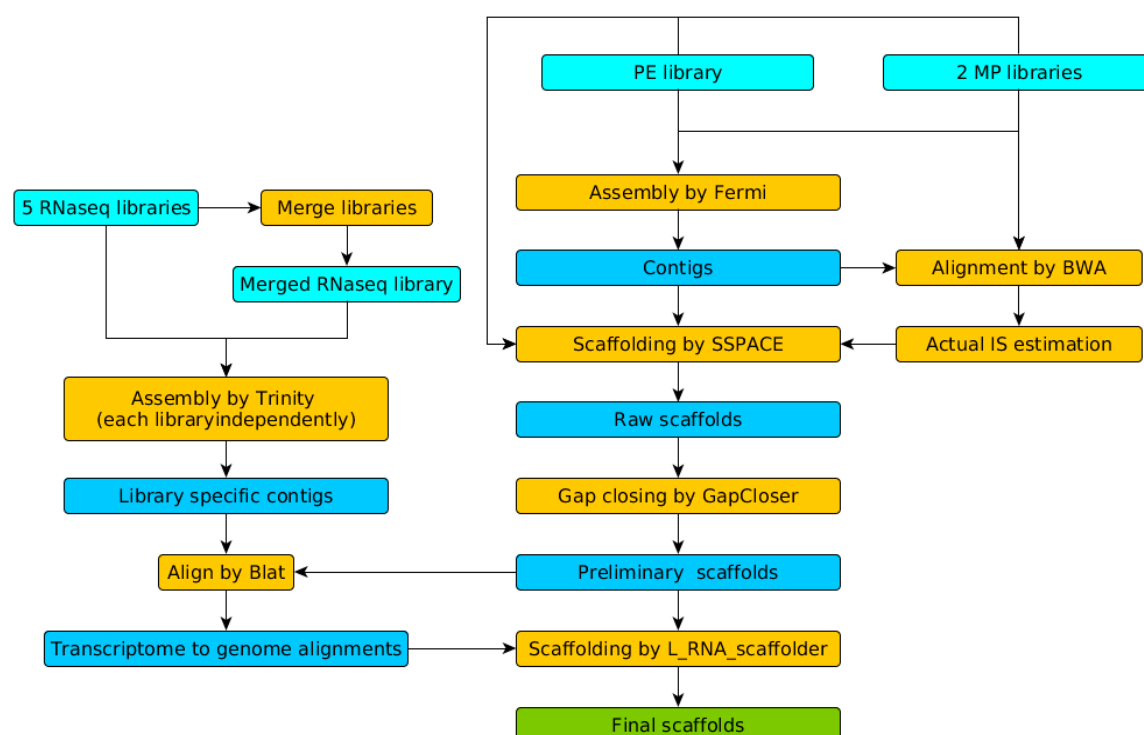


Figure S1. Pipeline used to assemble *A. vitatta* genome. One PE and two MP *A. vitatta* genome libraries were used to generate preliminary scaffolds complemented by additional scaffolding step using five *A. ventralis* RNaseq libraries.

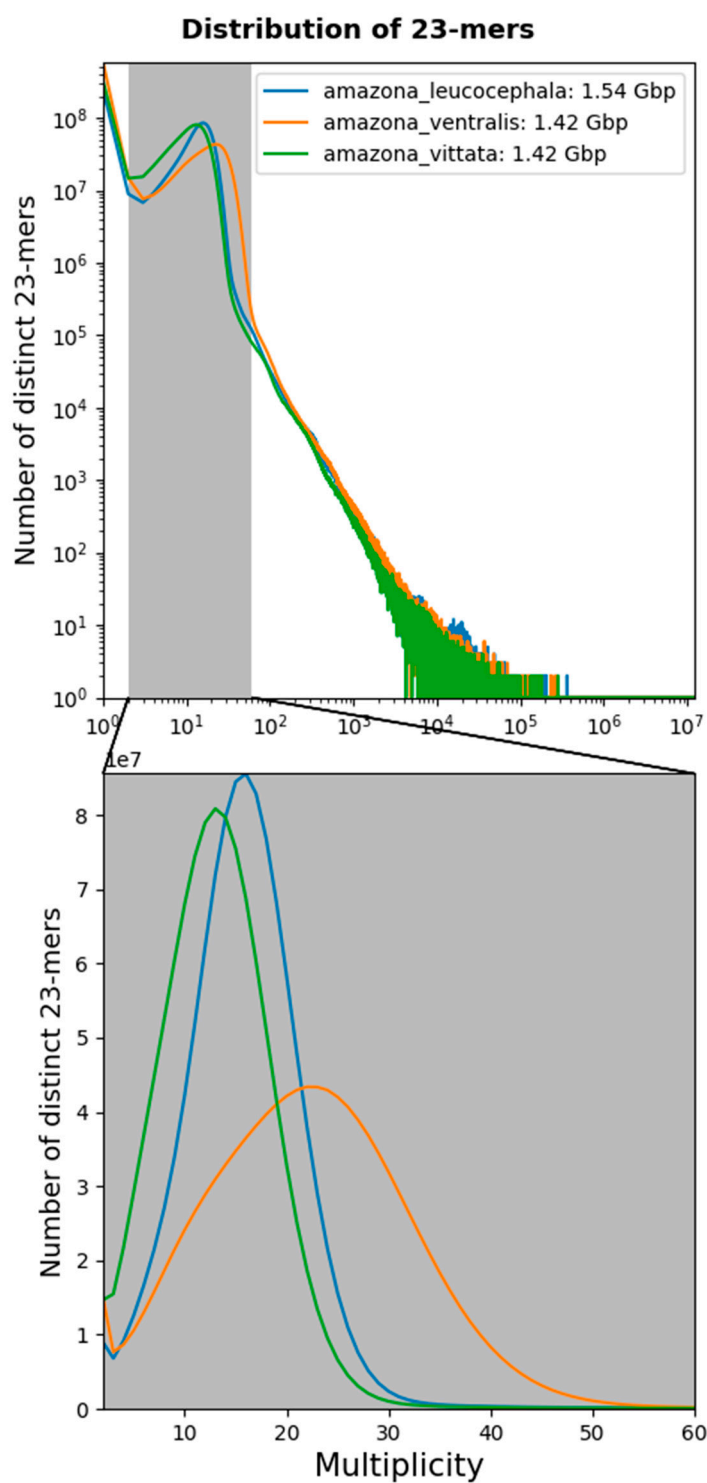


Figure S2. Distributions of 23-mers from PE libraries of three parrot species. Corresponding genome size estimation present in figure and also in Table 1. Only one major peak was detected for each genome, suggesting low heterozygosity level in all three species.

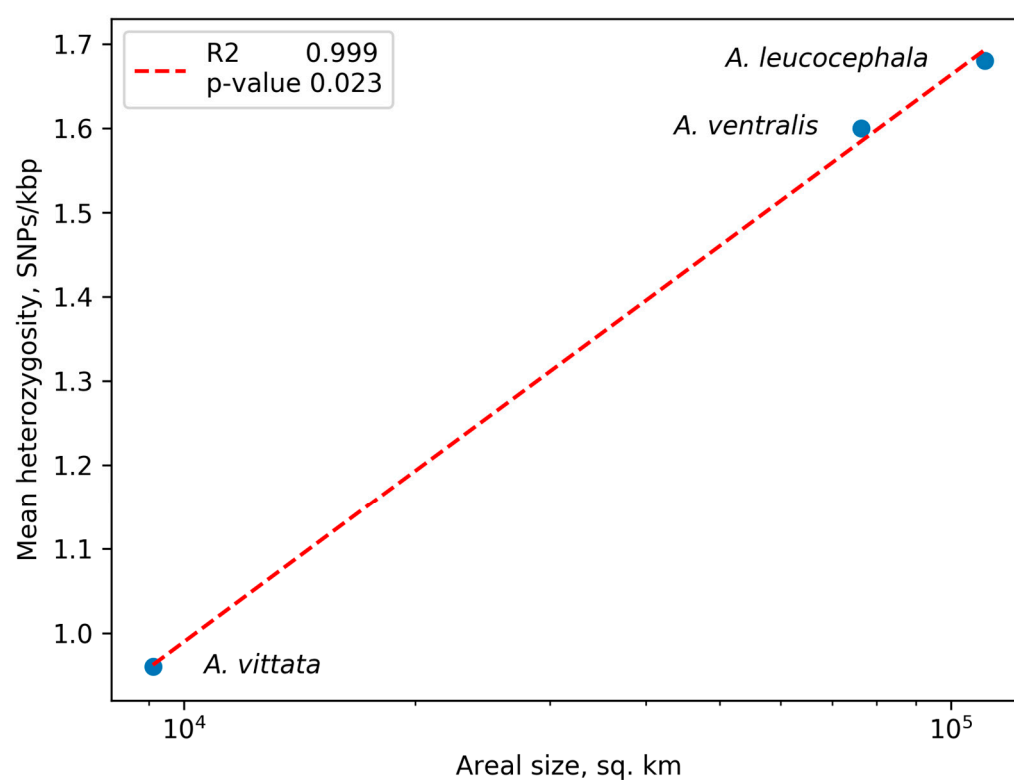


Figure S3. Connection between heterozygosity and areal size of the corresponding islands for the three *Amazona* species: *A. vittata* (Puerto Rico), *A. ventralis* (Hispaniola) and *A. leucocephala* (Cuba). The X axis is on the log₁₀ scale, and a logarithmic trendline is displayed ($r^2 = 0.999$, p -value = 0.023)

Table S1. The sequencing outputs for each genome used in the current assemblies. Abbreviations: paired-end (PE) mate pairs (MP)

Species	Library Type	Actual Insert Size (bp)	Number of reads (M)	Total length (Gb)	Coverage (x)
<i>A. vittata</i>	PE	281bp	124.6	24.34	13
	MP	2,172	76.7	15.46	
	MP	6,935	292.5	38.81	
<i>A. leucocephala</i>	PE	281bp	159.9	31.94	16
<i>A. ventralis</i>	PE	281bp	204	40.82	22

Table S2. BUSCO scores for all steps of assembly of *A. vittata* genome. Assembly evaluation was performed using BUSCO v3 and *Avian* dataset (Simão et al. 2015).

	Complete	Complete and single-copy	Complete and duplicated	Fragmented	Missed
<i>Fermi + SSPACE</i>	82.6	82	0.6	8.4	9
<i>Fermi + SSPACE + GapCloser</i>	84.2	83.4	0.8	8.4	7.4
<i>Fermi + SSPACE + GapCloser + L_RNA_scaffolder</i>	87.4	86.5	0.9	6.4	6.2

Table S3. Available fossil-based calibrations for speciation time dating within *Psittaciformes* and *Passeriformes* clades. All ages are in millions of years (MYA)

Split	<i>Psittaciformes</i> / <i>Passeriformes</i>	<i>Manacus</i> / <i>Taeniopygia</i> / <i>Geospiza</i>	<i>Taeniopygia</i> / <i>Geospiza</i>
Node min age	53.5 MYA	13.6 MYA	7.2 MYA
Node max age	65 MYA	16.3 MYA	11.6 MYA
Evidence	fossil, biostratigraphy	fossil (distal humerus, proximal ulna, distal tarsometatarsus)	fossil (nearly complete postcranial skeleton)
Reference	(Ksepka and Clarke 2015)	(Jarvis et al. 2014)	(Jarvis et al. 2014)
Oldest fossil taxon	<i>Pulchrapollia</i> <i>gracilis</i>	<i>Miocitta galbreathi</i>	<i>Corvus larteri</i>
Description	Oldest known species of <i>Psittaciformes</i> , K- T boundary	Internal Branch “ <i>Manacus</i> + <i>Corvus</i> + + <i>Taeniopygia</i> + <i>Geospiza</i> ” - MRCA of suboscines and oscines	Terminal Branch “ <i>Corvus</i> ” = MRCA “ <i>Corvus</i> +(<i>Taeniopygia</i> , <i>Geospiza</i>)”, <i>Corvidae</i>

Table S4. Mean heterozygosity in the genomes of three *Amazona* parrot species compared with that reported earlier for other species of birds. The table only contains species with fragmentation of genome assembly (N50) comparable to *A. vittata*.

Common name	Order	Conservation status	Whole Genome Heterozygosity
White-tailed Eagle	Accipitriiformes	EV (Once VU)	0.00040
Dalmatian pelican	Pelecaniformes	EV (VU)	0.00060
Great black cormorant	Pelecaniformes	LC	0.00139
Turkey vulture	Accipitriiformes	LC	0.00118
Kea	Psittaciformes	EV (VU)	0.00091
Puerto Rican Parrot	Psittaciformes §	EV	0.00096
Hispaniolan parrot	Psittaciformes §	NT	0.00160
Cuban parrot	Psittaciformes §	VU	0.00168

* [81]

§ this study