Germline-Restricted Chromosome (GRC) is Widespread among Songbirds

Anna A. Torgasheva[1,2], Lyubov P. Malinovskaya[1,2], Kira S. Zadesenets[1], Tatyana V. Karamysheva[1], Elena A. Kizilova[1,2], Ekaterina A. Akberdina[1], Inna E. Pristyazhnyuk[1], Elena P. Shnaider[3], Valeria A. Volodkina[4], Alsu F. Saifitdinova[4], Svetlana A. Galkina[4], Denis M. Larkin[1,5], Nikolai B. Rubtsov[1,2], and Pavel M. Borodin[1,2]

[1]Institute of Cytology and Genetics, Russian Academy of Sciences, Siberian Department, Novosibirsk, 630090, Russia

[2]Novosibirsk State University, Novosibirsk, 630090, Russia

[3]SibEcoCenter LLC, Novosibirsk, 630009, Russia

[4]Saint Petersburg State University, Saint Petersburg, 199034, Russia

[5]Royal Veterinary College, University of London, London NW1 0TU, UK

Corresponding author

Pavel M. Borodin

borodin@bionet.nsc.ru

18 **Abstract**

19 An unusual supernumerary chromosome has been reported for two related avian
20 species, the zebra and Bengalese finches. This large, germline-restricted chromosome (GRC),
21 is eliminated from somatic cells and spermatids and transmitted via oocytes only. Its origin,
22 distribution among avian lineages, and function were mostly unknown so far. Using
23 immunolocalization of key meiotic proteins, we found that GRCs of varying size and genetic
24 content are present in all sixteen songbird species investigated and absent from germline
25 genomes of all eight examined bird species from other avian orders. Results of fluorescent *in*
26 *situ* hybridization of microdissected GRC probes and their sequencing indicate that GRCs show
27 little homology between songbird species and contain a variety of repetitive elements and
28 unique sequences with paralogs in the somatic genome. Our data suggest that the GRC
29 evolved in the common ancestor of all songbirds and underwent significant changes in the
30 extant descendant lineages.

31

32 **Significance Statement**

33 We discovered that contrary to other bird species and most other animals, all examined
34 songbird lineages contain a different number of chromosomes in the somatic and germ line
35 genomes. Their germ cells have an additional germline-restricted chromosome (GRC). GRCs
36 contain highly duplicated genetic material represented by repetitive elements and sequences
37 homologous to unique regions of the somatic genome. Surprisingly, GRCs even in very closely
38 related species, vary drastically in size and show little homology. We hypothesize that the GRC
39 was formed as an additional parasitic microchromosome in the songbird ancestor about 35
40 MYA and subsequently underwent significant changes in size and genetic content, becoming an
41 important component of the germ line genome.

42

43 \body

44 **Introduction**

45 In addition to a standard chromosome set, which is present in all cells of an
46 organism, cells of many animal, plant and fungi species contain additional, so called B-
47 chromosomes. Their origin, evolution and adaptive significance remain obscure. B-
48 chromosomes show erratic phylogenetic distribution indicating their independent
49 occurrence in different species. The fact that B-chromosomes vary in number between
50 different individuals of the same species or even between cells of the same individual
51 suggests they are not critical for survival and may be considered as selfish genomic
52 elements (1, 2). In birds, additional chromosomes were described so far in two related
53 species of the family Estrildidae: zebra and Bengalese finches (3, 4). However, these
54 chromosomes behave differently from a typical B-chromosome.

55 In the germline cells of these two species, a large additional acrocentric chromosome is
56 found, which is absent from somatic cells. In oocytes, this germline-restricted chromosome
57 (GRC) is usually present in two copies, forming a bivalent that undergoes recombination. In
58 spermatocytes, one copy of this chromosome forms a round heterochromatic body, which is
59 eliminated from the nucleus during the first meiotic division (4, 5). Camacho and co-workers (2,
60 6) indeed suggested that the GRC is a genomic parasite, a bird variant of supernumerary B-
61 chromosome. Recent studies revealed that the zebra finch GRC contains multiple copies of
62 genes paralogous to the genes from the somatic genome (7, 8). Some of these genes are
63 amplified and importantly could be expressed at both the RNA and protein levels in the testes
64 and ovaries. Thus, GRC contains genetic material, which could be important for germ line cells,
65 but not essential for the majority of the body (somatic) cells.

66 The origin of GRC remains unclear. Itoh and co-workers (9) found that the zebra finch
67 GRC contains sequences homologous to an interval of chromosome 3 as well as repetitive
68 elements absent from the sequenced somatic genome. Phylogenetic analysis of GRC-derived
69 sequences, together with the zebra finch and chicken somatic cell counterparts, suggests that

70 the GRC was formed after the galliform–neoaves split (9). A recent study hypothesized that the
71 GRC is evolutionarily old and could be present in other birds as well (7).

72 In order to answer questions about the origin, architecture, and widespread of GRCs in
73 avian lineages we performed a comprehensive comparative cytogenetic study of the germ cell
74 chromosomes from 24 avian species representing eight Orders. To further examine the degree
75 of GRC conservation between distinct species we made a sequence-based comparison of the
76 microdissected GRC probes from four passerine species.

77 **Results and discussion**

78 Using antibodies to the core proteinaceous structure of meiotic chromosomes, the
79 synaptonemal complex (SC), we found that GRCs are present in all 16 songbird species
80 examined (14 in this study and two in the previous studies). These species represent nine
81 families of Passeri (Fig. 1). In ten species, the GRCs were large acrocentric
82 macrochromosomes (macro-GRCs) absent from bone marrow cells (Fig. 2A–2B and see SI
83 Appendix, fig. S1). In oocytes, the macro-GRCs were usually present as a bivalent, containing
84 one or two terminally located recombination sites visualized by antibodies to MLH1, a mismatch
85 repair protein. In spermatocytes, the GRC usually occurred as a univalent lacking recombination
86 sites, and was diffusely labeled with centromere antibodies (Fig. 2A and see SI Appendix, fig.
87 S1). At the end of male meiotic prophase, this GRC was transformed into a dense round body
88 and ejected from the nucleus (see SI Appendix, fig. S2). A similar meiotic behavior has been
89 described for GRCs in zebra and Bengalese finches (3, 4).

90 In male germline cells of six other species, we detected micro-GRCs, which appeared as
91 a univalent surrounded by a cloud of centromere antibodies and lacking recombination sites,
92 similar to the behavior of macro-GRCs in the ten other species. In the oocytes of these species,
93 the GRCs formed a bivalent indistinguishable from the standard microchromosomes. We did not
94 observe any phylogenetic clustering for the GRCs by size. Both macro- and micro-GRCs were
95 present within the families Fringillidae and Hirundinidae (Fig. 1).

96 Every examined primary spermatocyte of the 16 songbird species contained a GRC.
97 This suggests that the GRC is an important component of the germ line genome. However, no
98 GRCs were observed (by re-analyzing our own data (14, 15) and published SC images (16–19))
99 in eight species of non-passerine birds from seven separate lineages (Fig. 1). This implies a
100 monophyletic origin of the GRC. The estimated time since songbird divergence from other avian
101 lineages is 35 MYA (20). However, basal oscines, suboscines, and Acanthisittidae species have
102 not been examined yet, so we cannot exclude the possibility that GRCs formed in the common
103 ancestor of all Passeriformes, about 60 MYA (20).

104 To estimate the sequence homology between GRCs of different species and to get
105 insight into their genetic content, we prepared DNA probes of macro-GRCs for four
106 representatives of three families: Estrildidae (zebra and Bengalese finches), Fringillidae
107 (Eurasian siskin), and Hirundinidae (pale martin). We microdissected the round dense bodies
108 (see SI Appendix, fig. S2) containing the GRC from spermatocyte spreads and carried out
109 whole-genome amplification of the dissected material. The resulting probes were used for
110 fluorescent *in situ* hybridization (FISH) and for sequencing.

111 Reverse FISH with these GRC probes produced strong specific signals on the GRCs of
112 each species, indicating that the round dense bodies are indeed the ejected GRCs (see SI
113 Appendix, fig. S2C and fig. S3). In cross-species FISH experiments, the intensity of specific
114 GRC signals was much lower. Interestingly, micro-GRCs were painted with DNA probes derived
115 from macro-GRCs of closely related species, indicating that GRCs of related species share at
116 least a part of their genetic content. In both reverse and cross species FISH, we also detected
117 GRC signals on somatic chromosomes. Some signals remained visible after suppression of
118 repeated sequences with Cot-1 DNA. This indicates that GRCs contain multiple copies of
119 sequences homologous to genomic repeats, as well as sequences homologous to unique
120 regions present in the somatic genomes.

121 To identify these sequences we aligned the GRC NGS reads to the repeat-masked
122 zebra finch reference genome (Taeniopygia_guttata-3.2.4) using BLAT (21) with a 90% identity
123 setting. Average genome coverage estimated in 10 kb windows was 0.15 ± (S.D.) 4.60, 0.12 ±

124 3.29, 0.03 ± 1.16, and 0.01 ± 0.25 for reads of zebra finch, Bengalese finch, Eurasian siskin,
125 and pale martin GRC libraries, respectively. The coverage was highly uneven. GRCs of different
126 species show homology to different regions of the reference genome. Using the four GRC
127 libraries, we characterized 27 regions longer than 10 kb, covered in at least 30% of their length
128 and with an excess of two S.D. from the genome average (see SI Appendix, table S1). In some
129 regions, where the GRC of one species showed a high coverage, GRCs of other species
130 showed lower, but still above average, coverage. This may indicate that the unique sequences
131 located in these regions have been copied from the ancestral somatic genome into the
132 ancestral GRC and have then subsequently become diverged at the sequence level and/or in
133 copy number.

134 The longest of such excessively covered genomic regions were also detected by FISH at
135 the SCs of the corresponding species. Some regions partially overlapped sequences of zebra
136 finch genes (22) or sequences homologous to non-zebra finch RefSeq genes (23) (see SI
137 Appendix, table S1). For example, the zebra finch GRC probe gave a strong hybridization signal
138 on the short arm of the zebra finch SC3 (corresponding to TGU1) and on one of the largest SCs
139 of other species examined (Fig. 2C and see SI Appendix, fig. S3). In the corresponding region
140 of TGU1, we found a 2.5 Mb long cluster of several regions with ~70 fold coverage excess (see
141 SI Appendix, table S1). This cluster overlapped with two genes: completely with ROBO1, a
142 gene involved in vocal learning (24); and partially with GBE1, a gene encoding 1,4-alpha-glucan
143 branching enzyme 1. The homology between the zebra finch GRC and a part of the genomic
144 interval on TGU1 has been detected earlier by the RAPD-PCR technique (9) and recently
145 confirmed by Kinsella et al. (7).

146 Besides functional genes, GRCs also contain multiple repeated sequences. We
147 estimated their representation in the GRC reads and in the somatic genome of zebra finch using
148 RepeatMasker (25) with the RepBase avian library (26) (see SI Appendix, table S2). This
149 revealed both simple and low complexity repeats. The fraction of transposable elements (TEs)
150 in the GRCs was typical for avian genomes (27). The majority were LTRs and LINEs, while
151 SINEs and DNA TEs were represented in lower fractions than in the somatic genome. Overall

152 abundance of LTRs and LINEs and their ratio varied between different species' GRCs,
153 reflecting their different evolutionary trajectories. It has been shown that although activity of TEs
154 in avian genomes was rather low and ancient (especially for SINEs), avian species differed for
155 the timing of TE family activities. Interestingly, the zebra finch genome shows a peak of LTR
156 activity from 5 to 20 MYA (27). This is a likely reason why LTRs are more abundant in zebra
157 finch GRC than in other GRCs. On the other hand, SINEs are rare in avian genomes and they
158 did not show any activity during last 30 MY, yet they are present in the GRCs of all four
159 examined species, likely being inherited from the GRC ancestor. This provides further evidence
160 for the formation of GRCs in the songbird genome rather than in older avian ancestors, because
161 GRCs had a chance to accumulate at higher rate LTRs active in songbirds but not older SINEs.
162 Therefore, a few SINEs found in the GRC likely represent copies transferred from the somatic
163 genome and amplified in the GRCs rather than those inserted during the actual activity of
164 SINEs.

165 To examine the general pattern of GRC transcription in oogenesis, we analyzed
166 lampbrush GRCs isolated from zebra finch oocytes at the previtellogenic growth phase. The
167 lampbrush GRC exhibited a typical chromomere-loop pattern, with several pairs of
168 transcriptionally active lateral loops extending from all chromomeres except for those located in
169 a prominent DAPI-positive region. Antibodies against RNA-polymerase II labeled the whole
170 GRC except for this region (see SI Appendix, fig. S4). Thus, lampbrush GRCs display a pattern
171 of transcription typical for somatic chromosomes (28).

172 Indeed, recent studies demonstrated that many GRC-linked genes are transcribed (7, 8)
173 and at least some of them are translated in the zebra finch germ line (7). The evolutionary
174 history of some of these genes points to the songbird origin of the proto-GRC. This is an
175 excellent complementary confirmation of our own findings which drove us to the same
176 conclusion based on a direct (cytogenetic) observation and indirect (transposable element
177 composition) analysis of our sequenced GRC libraries. Our study, however, also points to the
178 existence of both micro- and macro-chromosome versions of GRCs in avian lineages,
179 suggesting that this chromosome is highly dynamic in songbird evolution.

180 Thus, GRCs are present in all the songbirds studied, but are absent from germ lines of
181 birds from other orders. These chromosomes vary drastically in size and show a low sequence
182 similarity between different species. GRCs contain various highly duplicated regions
183 represented in the somatic genome by both unique and repetitive sequences. The spectrum of
184 transposable elements found in our sequenced GRC libraries suggests that the GRC was more
185 likely formed in the ancestral songbird lineage followed by an extensive sequence divergence in
186 the descendent species genomes rather than to appear in the avian ancestor and then being
187 lost in the non-songbirds.

188 Therefore, we propose that the GRC has formed as an additional 'parasitic B-like'
189 microchromosome in the ancestral songbird genome likely due to a whole-chromosome
190 duplication (Fig. 3). If this proto-GRC already contained some copies of somatic genes
191 contributing to reproductive and developmental processes it could become beneficial due to
192 providing a higher dosage of these genes and therefore escape purifying selection pressure in
193 the germ line. Its presence in the germ line only could also relax selection for the functional
194 integrity of the GRC's genetic content.

195 This in turn could make the GRC a target for selfish genetic elements active during its
196 evolutionary history. Additional copies of unique sequences (e.g. genes) from the somatic
197 genome could also populate the GRC through non-allelic recombination process, using its own
198 and somatic genome transposable elements as templates. Suppression of recombination along
199 GRCs (except for their termini in female meiosis) could facilitate their divergence and the
200 degradation of their original genetic content via Muller's ratchet (29). This could lead to a rapid
201 and massive loss of homology between various species' GRCs.

202 However, as the contemporary GRCs contain expressed and transcribed genes and
203 persist in the germ line of all songbirds studied, it likely has changed its original 'parasitic' state
204 to a more 'symbiotic' one, providing evolutionary benefits to the representatives of the most
205 speciose group of birds. We believe that a detailed comparison of micro- and macro-GRCs,
206 phylogenetic studies of shared and lineage-specific GRC sequences, and detailed analysis of

207 their stratification within each GRC will shed further light on the origin and evolution of this
208 highly dynamic and surprising chromosome.
209

210 **Materials and methods**

211 ***Experimental model and subject details***

212 Adult males of pale martin, great tit, barn swallow, European pied flycatcher, Blyth's reed
213 warbler and black tern were captured at the beginning of breeding season. Nestling females of
214 sand martin, pale martin, barn swallow, great tit, and European pied flycatcher were collected
215 from nests ~3-6 days after hatching.

216 Adult male zebra finch, Gouldian finch, Bengalese finch, Eurasian siskin, European
217 goldfinch, Eurasian skylark, pine bunting, Eurasian bullfinch, common canary and budgerigar
218 were purchased from a commercial breeder. Sexually mature zebra finch females were
219 provided by the Leningrad Zoo (Saint Petersburg, Russia). An adult male rook with fatal
220 accident trauma was provided by the Bird Rehabilitation Centre of Novosibirsk and euthanized
221 in our laboratory.

222 Capture, handling and euthanasia of birds followed protocols approved by the Animal
223 Care and Use Committee of the Institute of Cytology and Genetics SD RAS (protocol #35 from
224 26.10.2016) and by the Saint Petersburg State University Ethics Committee (statement #131-
225 03-2). Experiments described in this manuscript were carried out in accordance with the
226 approved national guidelines for the care and use of animals. No additional permits are required
227 for research on non-listed species in Russia.

228 ***Mitotic metaphase chromosomes***

229 Mitotic chromosome preparations were obtained from short-term bone marrow cell
230 cultures incubated for 2 h at 37°C with 10 µg/ml colchicine in culture Dulbecco's Modified
231 Eagle's medium with UltraGlutamine. Hypotonic treatment was performed with 0.56% KCl
232 solution for 15 min at 37°C and followed by centrifugation for 5 min at 500×*g*. Fresh cold fixative
233 solution (methanol : glacial acetic acid, 3:1) was changed three times. Cell suspension was
234 dropped on cold, wet slides (76 mm x 26 mm, 1 mm thick). The slides were dried for 2 hours at
235 65°C and stained for 4 min with 1 µg/ml solution of DAPI in 2xSSC. Then slides were washed in

236 deionized water, dried at room temperature and mounted in Vectashield antifade mounting
237 medium (Vector Laboratories, USA) to reduce fluorescence fading.

238 ***SC spreading and immunostaining***

239 Chromosome spreads for SC analysis were prepared from spermatocytes or juvenile
240 oocytes according to Peters et al. (30). Immunostaining was performed according to the
241 protocol described by Anderson et al. (31) using rabbit polyclonal anti-SYCP3 (1:500; Abcam),
242 mouse monoclonal anti-MLH1 (1:50; Abcam), and human anticentromere (ACA) (1:100;
243 Antibodies Inc) primary antibodies. The secondary antibodies used were Cy3-conjugated goat
244 anti-rabbit (1:500; Jackson ImmunoResearch), FITC-conjugated goat anti-mouse (1:50;
245 Jackson ImmunoResearch), and AMCA-conjugated donkey anti-human (1:100; Jackson
246 ImmunoResearch). Antibodies were diluted in PBT (3 % bovine serum albumin and 0.05 %
247 Tween 20 in phosphate-buffered saline). A solution of 10% PBT was used for blocking. Primary
248 antibody incubations were performed overnight in a humid chamber at 37°C; and secondary
249 antibody incubations, for 1 h at 37°C. Slides were mounted in Vectashield antifade mounting
250 medium (Vector Laboratories, USA) to reduce fluorescence fading.

251 ***Lampbrush chromosome preparations***

252 Zebra finch lampbrush chromosomes were manually dissected from previtellogenic or
253 early vitellogenic oocytes using the standard avian lampbrush technique described in
254 Saifitdinova et al. (32). After centrifugation, preparations were fixed in 2% paraformaldehyde,
255 then in 50% and in 70% ethanol, air-dried and kept at room temperature until used for FISH. For
256 immunostaining experiments lampbrush chromosome preparations were kept in 70% ethanol at
257 +4ºC.

258 ***Preparation of the hybridization probe and FISH***

259 In order to generate a DNA probe for the GRCs of the pale martin, zebra finch,
260 Bengalese finch, and Eurasian siskin testicular cells of adult males were treated with hypotonic
261 solution (0.88% KCl) at 37^0 for 3h and then with Carnoy's solution (methanol : glacial acetic
262 acid, 3:1). The cell suspension was dropped onto clean, cold, wet cover slips (60 mm x 24 mm,
263 0.17 mm thick), dried, and stained with 0.1% Giemsa solution (Sigma) for 3-5 min at room

264 temperature. GRCs were identified as positive round bodies located near the spermatocytes I.
265 Microdissection of GRC and amplification of DNA isolated from this chromosome were carried
266 out with the GenomePlex Single Cell Whole Genome Amplification Kit (WGA4) (Sigma-Aldrich)
267 (33). Microdissected DNA probes were generated from 15 copies of GRC for each studied
268 species. The obtained PCR products were labeled with Flu-dUTP (Genetyx, Novosibirsk,
269 Russia) in additional PCR cycles or with biotin-11-dUTP (Sileks, Moscow, Russia).

270 FISH experiments with DNA probes on SC spreads of the studied avian species were
271 performed as described earlier (34) with salmon sperm DNA (Ambion, USA) as a DNA carrier.
272 In case of suppression FISH, Cot-1 DNA (DNA enriched for repetitive DNA sequences) was
273 added to the DNA probe to suppress the repetitive DNA hybridization. Chromosomes were
274 counterstained with DAPI dissolved in Vectashield antifade solution (Vector Laboratories, USA).
275 Zebra finch GRC at the lampbrush stage was identified by FISH using biotin-labelled zebra finch
276 microdissected probe with a 50-fold excess of *E. coli* tRNA as a carrier. FISH was performed
277 according to the DNA/DNA+RNA hybridization protocol omitting any chromosome pre-
278 treatment, as described previously (35). To detect biotin-labelled probe, we used avidin-
279 Alexa488 and biotinylated goat antibody against avidin (both from Thermo Fisher Scientific,
280 USA). Lampbrush chromosomes were counterstained with DAPI in an antifade solution,
281 containing 50% glycerol.

282 ***Immunostaining of the zebra finch lampbrush chromosomes***

283 Immunostaining was carried out with mouse antibodies V22 (kindly donated by
284 U. Scheer) against the phosphorylated C-terminal domain (CTD) of RNA polymerase II.
285 Lampbrush chromosome spreads, fixed in 2% paraformaldehyde, were blocked in 0.5%
286 blocking reagent (Sigma-Aldrich, USA) in PBS for 1 h at +37°C. Then preparations were
287 incubated with primary antibodies, diluted 1:200, overnight at room temperature. Slides were
288 washed in PBS with 0.05% Tween-20 and incubated with Alexa-488-conjugated goat anti-
289 mouse IgG+IgM secondary antibody (Jackson ImmunoResearch Lab). After washing in
290 PBS+0.05% Tween-20, slides were mounted in antifade solution containing DAPI.

291 ***Microscopic analysis***

292 Images of fluorescently stained metaphase chromosomes and/or SC spreads were
293 captured using a CCD-camera installed on an Axioplan 2 compound microscope (Carl Zeiss,
294 Germany) equipped with filtercubes #49, #10, and #15 (ZEISS, Germany) using ISIS4
295 (METASystems GmbH, Germany) at the Center for Microscopic Analysis of Biological Objects
296 of SB RAS (Novosibirsk, Russia). For further image analysis we used Corel PaintShop Pro X6
297 (Corel). The location of each imaged immunolabeled spread was recorded so that it could be
298 relocated on the slide after FISH. Zebra finch lampbrush chromosome preparations were
299 examined using a Leica DM4000B fluorescence microscope installed at the "Chromas"
300 Resource Centre, Saint-Petersburg State University Scientific Park (Saint Petersburg, Russia).
301 The microscope was equipped with a black and white DFC350FX camera and filters A and I3.
302 LAS AF (Leica) software was used to capture and process color images; Adobe Photoshop CS5
303 (Adobe Systems) was used for figure assembling. The length of the SC of each chromosome
304 arm was measured in micrometers and the positions of centromeres were recorded using
305 MicroMeasure 3.3 (36). We identified individual SCs by their relative lengths and centromeric
306 indexes.

307 ***Preparation of amplified DNA and library construction***

308 DNA amplification of microdissected GRC chromosomal material was performed with the
309 GenomePlex Single Cell Whole Genome Amplification kit (WGA4) (Sigma-Aldrich) according to
310 the manufacturer's protocol. DNA library for NGS sequencing was prepared based on the
311 microdissected GRC DNA libraries using the NEBNext Ultra DNA Library Prep kit (New England
312 Biolabs).

313 ***High throughput sequencing and error correction***

314 NEBNext Ultra library was sequenced on an Illumina NextSeq 550 system with single-
315 end reads at the "Genomics" core facility of the ICG SB RAS (Novosibirsk, Russia). Read
316 lengths were 150 bp, the total number of reads obtained were 1,730,845, 1,596,722, 2,821,862
317 and 1,265,105 for zebra finch, Bengalese finch, Eurasian siskin and pale martin GRC

318 correspondingly. DNA data were quality assessed using FastQC (37) and quality trimmed using
319 Trimmomatic (38).

320 ***Estimating the homology to somatic genome and repeat content***

321 Reads from the zebra finch, Bengalese finch, Eurasian siskin, and pale martin GRC
322 sequences were aligned to the assembly of the zebra finch genome (Taeniopygia_guttata-3.2.4)
323 using BLAT (21). A custom python script was used to estimate the coverage of the zebra finch
324 genome in 10 kb windows. Overlapping of regions with high coverage with zebra finch Ensembl
325 gene predictions and non-zebra finch RefSeq (23) genes was revealed with the Ensembl
326 genome browser. Repeat content of the GRC libraries and the zebra finch genome was
327 assessed with RepeatMasker (25) by using the avian RepBase database (26)мм.

328 **Acknowledgments**

329 We thank M.I. Rodionova for her help in chromosome preparation, A. Maslov, D.
330 Taranenko, I. Korobitsyn, M. Scherbakova for help in bird collecting, J.B. Parker for English
331 editing, the “Chromas” Resource Centre of the Saint-Petersburg State University and the
332 Microscopy Center of the Siberian Department of the Russian Academy of Sciences for granting
333 access to microscopy equipment.

334 This work was supported by the Russian Foundation for Basic Research (Grant # 18-04-
335 00924) and the Ministry of Science and High Education via the Institute of Cytology and
336 Genetics (Grant # 0324-2019-0042). The funding bodies play no role in the design of the study
337 and collection, analysis, and interpretation of data and in writing the manuscript.

338

339 **References**

340 1. Houben A, Banaei-Moghaddam AM, Klemme S, Timmis JN (2014) Evolution and biology
341 of supernumerary B chromosomes. *Cell Mol Life Sci*. 71:467–478.

342 2. Camacho JPM (2005) B Chromosomes. *The Evolution of the Genome*, ed TR G (Elsevier
343 Academic Press, Amsterdam), pp 223–286.

344 3. Pigozzi MI, Solari AJ (1998) Germ cell restriction and regular transmission of an
345 accessory chromosome that mimics a sex body in the zebra finch, Taeniopygia guttata.
346 *Chromosom Res* 6(2):105–113.

347 4. del Priore L, Pigozzi MI (2014) Histone modifications related to chromosome silencing
348 and elimination during male meiosis in Bengalese finch. *Chromosoma* 123(3):293–302.

349 5. Schoenmakers S, Wassenaar E, Laven JSE, Grootegoed JA, Baarends WM (2010)
350 Meiotic silencing and fragmentation of the male germline restricted chromosome in zebra
351 finch. *Chromosoma* 119(3):311–324.

352 6. Pedro J, Camacho M, Sharbel TF, Beukeboom LW (2000) B-chromosome evolution. *Phil*
353 *Trans Roy Soc L B-Biological Sci* 355:163–178.

354 7. Kinsella CM, et al. (2018) Programmed DNA elimination of germline development genes
355 in songbirds. *bioRxiv*:444364.

356 8. Biederman MK, et al. (2018) Discovery of the First Germline-Restricted Gene by
357 Subtractive Transcriptomic Analysis in the Zebra Finch, Taeniopygia guttata. *Curr Biol*
358 28(10):1620–1627.

359 9. Itoh Y, et al. (2009) Molecular cloning and characterization of the germline-restricted
360 chromosome sequence in the zebra finch. *Chromosoma* 118(4):527–536.

361 10. Reddy S, et al. (2017) Why do phylogenomic data sets yield conflicting trees? Data type
362 influences the avian tree of life more than taxon sampling. *Syst Biol* 66(5):857–879.

363 11. Prum RO, et al. (2015) A comprehensive phylogeny of birds (Aves) using targeted next-
364 generation DNA sequencing. *Nature* 526(7574):569–573.

365 12. Roquet C, Lavergne S, Thuiller W (2014) One tree to link them all: a phylogenetic dataset
366 for the European tetrapoda. *PLOS Curr Tree Life* 6:1–16.

367 13. Hooper DM, Price TD (2015) Rates of karyotypic evolution in Estrildid finches differ
368 between island and continental clades. *Evolution (N Y)* 69(4):890–903.

369 14. Torgasheva AA, Borodin PM (2017) Immunocytological Analysis of Meiotic
370 Recombination in the Gray Goose (Anser anser). *Cytogenet Genome Res* 151(1):27–35.

371 15. Malinovskaya L, Shnaider E, Borodin P, Torgasheva A (2018) Karyotypes and
372 recombination patterns of the Common Swift (Apus apus Linnaeus, 1758) and Eurasian
373 Hobby (Falco subbuteo Linnaeus, 1758). *Avian Res* 9(1):4.

374 16. Calderon PL, Pigozzi MI (2006) MLH1-focus mapping in birds shows equal recombination
375 between sexes and diversity of crossover patterns. *Chromosome Res* 14(6):605–612.

376 17. Pigozzi MI (2001) Distribution of MLH1 foci on the synaptonemal complexes of chicken
377 oocytes. *Cytogenet Cell Genet* 95(3–4):129–133.

378 18. Pigozzi MI, Solari AJ (1999) Equal frequencies of recombination nodules in both sexes of
379 the pigeon suggest a basic difference with eutherian mammals. *Genome* 42(2):315–321.

380 19. Pigozzi MI, Solari AJ (1999) Recombination nodule mapping and chiasma distribution in
381 spermatocytes of the pigeon, Columba livia. *Genome* 42(2):308–314.

382 20. Selvatti AP, Gonzaga LP, Russo CA (2015) A Paleogene origin for crown passerines and
383 the diversification of the Oscines in the New World. *Mol Phylogenet Evol* 88:1–15.

384 21. Kent WJ (2002) BLAT - The BLAST-like alignment tool. *Genome Res* 12(4):656–64.

385 22. Zerbino DR, et al. (2018) Ensembl 2018. *Nucleic Acids Res* 46(D1):D754–D761.

386 23. O'Leary NA, et al. (2016) Reference sequence (RefSeq) database at NCBI: Current
387 status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44(D1):D733-
388 45.

389 24. Wang R, et al. (2015) Convergent differential regulation of SLIT-ROBO axon guidance
390 genes in the brains of vocal learners. *J Comp Neurol* 523(6):892–906.

391 25. Smit A, Hubley R, Green P (2013) RepeatMasker Open-4.0. 2013-2015 .
392 *http://www.repeatmasker.org*.

393 26. Jurka J, et al. (2005) Repbase Update, a database of eukaryotic repetitive elements.
394 *Cytogenet Genome Res* 110(1–4):462–467.

395 27. Gao B, et al. (2017) Low diversity, activity, and density of transposable elements in five
396 avian genomes. *Funct Integr Genomics* 17(4):427–439.

397 28. Morgan GT (2018) Imaging the dynamics of transcription loops in living chromosomes.
398 *Chromosoma* 127(3):361–374.

399 29. Gabriel W, Lynch M, Burger R (1993) Muller' s Ratchet and Mutational Meltdowns.
400 *Evolution (N Y)* 47(6):1744–1757.

401 30. Peters AH, Plug AW, van Vugt MJ, de Boer P (1997) A drying-down technique for the
402 spreading of mammalian meiocytes from the male and female germline. *Chromosome*
403 *Res* 5(1):66–68.

404 31. Anderson LK, Reeves A, Webb LM, Ashley T (1999) Distribution of crossing over on
405 mouse synaptonemal complexes using immunofluorescent localization of MLH1 protein.
406 *Genetics* 151(4):1569–1579.

407 32. Saifitdinova A, Galkina S, Volodkina V, Gaginskaya E (2017) Preparation of lampbrush
408 chromosomes dissected from avian and reptilian growing oocytes. *Biol Commun*
409 62(3):165–168.

410 33. Zadesenets KS, Schärer L, Rubtsov NB (2017) New insights into the karyotype evolution
411 of the free-living flatworm Macrostomum lignano (Platyhelminthes, Turbellaria). *Sci Rep*
412 7(1):6066.

413 34. Zadesenets KS, Katokhin A V., Mordvinov VA, Rubtsov NB (2012) Telomeric DNA in
414 chromosomes of five opisthorchid species. *Parasitol Int* 61(1):81–83.

415 35. Galkina S, et al. (2006) FISH on avian lampbrush chromosomes produces higher
416 resolution gene mapping. *Genetica* 128:241–251.

417 36. Reeves A (2001) MicroMeasure: a new computer program for the collection and analysis
418 of cytogenetic data. *Genome* 44(3):439–443.

419 37. Andrews S (2010) FastQC: A quality control tool for high throughput sequence data.
420 Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

421 38. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina
422 sequence data. *Bioinformatics* 30:2114–2120.

423

424 **Figure legends**

425 **Fig. 1. Topology of the bird species examined.** Our sampling covers all major 426 passerine groups (except for the basal oscines, suboscines, and acanthisittides) as well as 427 Galloanserae, Columbaves, Apodiformes, Charadriformes, Ealconiformes, and Psittaciformes. 428 Black circles indicate species with a macro-GRC, white circles – species with a micro-GRC. 429 Numbers after the species' names indicate references for SC studies, asterisks indicate new 430 data. A consensus topology of bird orders is based on the cladogram from Reddy et al. (10). 431 Position of the common swift is defined according to Prum et al. (11). Topology of passerine 432 birds is shown according to Roquet et al. (12). Positions of species within the Estrildidae lineage 433 is established according to Hooper et al. (13).

434

435 **Fig 2. Discovery of GRCs in bird species.** (A) SC spreads of four oscine species 436 immunolabelled with antibodies against SYCP3, the main protein of the lateral element of SC 437 (red), centromere proteins (blue) and MLH1, mismatch repair protein marking recombination 438 sites (green). Arrowheads point to the largest chromosomes ordered according their size ranks, 439 ZZ (identified by its size and arm ratio), ZW (identified by heteromorphic SC and misaligned 440 centromeres), and GRCs. Arrows in the inserts point to MLH1 foci in GRCs. Micro-GRC 441 bivalents in female barn swallow and European pied flycatcher are indistinguishable from the 442 microchromosomes of the somatic chromosome set. (B) DAPI stained bone marrow cells. (C) 443 Reverse and cross-species FISH of GRC DNA probes (green) derived from Bengalese finch 444 (LST), zebra finch (TGU), Eurasian siskin (SSP), and pale martin (RDI) with SC spreads, 445 immunolabelled with antibodies against SYCP3 (red). Centromeres are labeled with antibodies 446 against centromere proteins (blue). Arrowheads point to GRCs and regions on the somatic 447 chromosome set intensely painted with GRC probes in cross-species FISH. Inserts show GRCs. 448 The Bengalese finch GRC-specific DNA probe gives a strong signal on the Bengalese finch 449 GRC and slightly paints some regions of the somatic chromosome set. The zebra finch GRC 450 probe paints the distal area of the Bengalese finch GRC and a region of the short arm of SC3.

451 The Eurasian siskin GRC probe paints a micro-GRC of European goldfinch, a region on the
452 long arm of SC3 and some pericentromeric regions. The pale martin GRC probe gives a
453 dispersed signal on the great tit GRC, the ZW bivalent and on SC5. Bar – 5 μm.

454

455 **Fig. 3. Scenario of GRC origin and evolution.** A proto-GRC forms due to duplication (Ctl-D) of
456 a microchromosome likely containing genes involved in germ cell development. Copies of
457 unique somatic cell sequences and repetitive elements invade the GRC (Ctl-C+Ctl-V).
458 Divergence of GRCs in different songbird lineages occurs due to amplification and deletion
459 (Del) of its sequences.