Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds

Alexander Suh1, Martin Paus1, Martin Kiefmann1, Gennady Churakov1, Franziska Anni Franke2,3, Jürgen Brosius1, Jan Ole Krieger1,4,* & Jürgen Schmitz1,*

The relationships of passerines (such as the well-studied zebra finch) with non-passerine birds is one of the great enigmas of avian phylogenetic research, because decades of extensive morphological and molecular studies yielded highly inconsistent results between and within data sets. Here we show the first application of the virtually homoplasly-free retroposon insertions to this controversy. Our study examined ~200,000 retroposon-containing loci from various avian genomes and retrieved 51 markers resolving early bird phylogeny. Among these, we obtained statistically significant evidence that parrots are the closest and falcons the second-closest relatives of passerines, together constituting the Psittacopasserae and the Eufalconimorphae, respectively. Our new and robust phylogenetic framework has substantial implications for the interpretation of various conclusions drawn from passerines as model organisms. This includes insights of relevance to human neuroscience, as vocal learning (that is, birdsong) probably evolved in the psittacopasseran ancestor, >30 million years earlier than previously assumed.
Birds are important model organisms in many fields\textsuperscript{1,2}, but even since the time of Darwin, numerous attempts to reconstruct their phylogenetic relationships have yielded at least as many controversies\textsuperscript{3-7}. In recent years, however, some morphological\textsuperscript{8,9} and most molecular studies\textsuperscript{10-13} have found congruence regarding the earliest chapters of bird evolution\textsuperscript{14}. The root of extant birds lies between Palaeognathae (ratites and tinamous) and Neognathae, the latter comprising Gallorostrales (chicken and ducks) and Neoaves (all remaining birds).

Despite immense efforts, most of the basal relationships among Neoaves remain unsolved. This includes one issue of great interdisciplinary relevance\textsuperscript{8}: the discovery of the putative sister group of passerines (>50\% of all birds species, including all songbirds), one of the most-studied groups of animals\textsuperscript{2}. Morphological studies indicated a close affinity either to woodpeckers and rollers or to cuckoos\textsuperscript{3}, whereas a more basal position among Neoaves was suggested by DNA hybridization data\textsuperscript{4}. On the other hand, nucleotide sequences of mitochondrial genomes placed passersines as the sister group to all remaining Neoaves\textsuperscript{5,6}, to a woodpecker/roller/trogon clade or to cuckoos\textsuperscript{3}, whereas nuclear sequence analyses proposed a relation to woodpeckers and rollers\textsuperscript{11} or to parrots\textsuperscript{12}, falcons and seriema\textsuperscript{11}.

A promising approach to overcome the present phylogenetic ambiguities is the use of retroposon insertions. Retroposons, jumping genetic elements that copy via RNA intermediates and insert nearly randomly anywhere in the genome (although some biases of insertion and retention have been proposed\textsuperscript{13}), provide (by inheritance) virtually homoplasy-free evidence of relatedness\textsuperscript{14} that is detectable for more than 100 million years. Because parallel insertions or exact excisions are highly unlikely\textsuperscript{15}, presence/absence patterns of retroposons at orthologous genomic loci are powerful, clear-cut phylogenetic markers capable of resolving long-standing uncertainties\textsuperscript{16-20}.

In this study, we present an improved resolution of bird evolution using retroposon insertions, a marker system that rarely undergoes uncertainties or exact excisions are highly unlikely\textsuperscript{16}; presence/absence patterns of retroposons at orthologous genomic loci are powerful, clear-cut phylogenetic markers capable of resolving long-standing uncertainties\textsuperscript{16-20}.

Results

Reconstructing the avian tree of life using retroposon insertions.

From the over 200,000 retroposed elements (REs) present in the chicken and zebra finch genomes\textsuperscript{1}, we selected the two most numerous fractions (>97\% of all REs), namely, both the chicken repeat 1 (CR1) family of long interspersed elements (LINEs) and the long terminal repeat elements (LTRs) of endogenous retroviruses. Utilizing three different search strategies (Methods), we extracted 131 CR1 and 75 LTR loci that were experimentally tested via high-throughput PCR, leading to the identification of 51 phylogenetically informative markers. For each marker, representatives of the key avian lineages\textsuperscript{13,14} were sampled, sequenced and aligned using standard procedures\textsuperscript{21}. To measure the strength of support for all recovered branches, we calculated \( P \) values using the Waddell et al.\textsuperscript{22} likelihood ratio test for retroposon data. Thus, statistically significant retroposon evidence \((P < 0.05)\) is reached with three conflict-free markers \((P = 0.0370, (3 \ 0 \ 0))\). Because of the mentioned strength and clearness of retroposon markers, our resultant maximum parsimony-based phylogenetic tree (Fig. 1, branches A–L) is effectively a maximum likelihood estimation\textsuperscript{23}.

Resolving early bird phylogeny.

Our retroposon markers are located on 14 different chromosomes, significantly clarifying more than the well-established\textsuperscript{14,17-19} avian relationships. We obtained six retroposon insertions that are shared among paleognaths and neognaths, corroborating the monophyly of extant birds (Fig. 1, branch A). These retroposon insertions feature a unique, diagnostic deletion present only in some avian CR1 elements (subtypes CR1-Y and CR1-Z; this deletion is absent in crocodilian and all other avian CR1 elements), and can therefore be regarded as bird-specific REs (Supplementary Fig. S1). Additionally, the root of living birds is located between the significantly supported Neognathae (Fig. 1, branch B; five REs, \( P = 0.0041, (5 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}) and Palaeognathae (Fig. 1, branch C; four REs, \( P = 0.0123, (4 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}). Significant support was also found for the monophyly of Neoaves (Fig. 1, branch D; six REs, \( P = 0.0014, (6 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}), Galloanserae (Fig. 1, branch E; four REs, \( P = 0.0123, (4 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}) and Passeriformes (Fig. 1, branch I; six REs, \( P = 0.0014, (6 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}).

Resolving the neoavian radiation. Within the hitherto largely unresolved\textsuperscript{14} radiation of Neoaves, we obtained four markers whose insertion patterns seem inconsistent with one another (Fig. 1, label F; Supplementary Fig. S2; Supplementary Table S1). As CR1 and LTR retroposons exhibit no or very short (<6bp) target site duplications, exact excisions as proposed for primates Alu short interspersed elements\textsuperscript{24} cannot have occurred\textsuperscript{25}. Of the nearly 1.2 billion\textsuperscript{26} potential insertion sites in the avian genome, parallel insertions (featuring exactly the same target site, retroposon type, orientation and truncation) should be extremely rare. Therefore, the incongruent patterns among the four retroposon insertions are most likely a result of incomplete lineage sorting (leading to hemiplasy)\textsuperscript{27} of retroposon presence/absence dimorphisms that persisted during the very beginning of the neoavian radiation and were randomly fixed (that is, one of the two alleles was lost) in each of the descendant lineages (Supplementary Fig. S2). This complex evolutionary phenomenon was previously revealed by retroposons (for example, in the rapid radiations of cichlid fishes\textsuperscript{28} and placental mammals\textsuperscript{29-30}) and is a further indication that the earliest period of the rapid radiation of Neoaves is a putative polytomy\textsuperscript{31}.

The remaining retroposon evidence within Neoaves exhibits no incongruent presence/absence patterns. We recovered the previously reported\textsuperscript{12,13} ‘landbird’ assemblage (Fig. 1, branch G; two REs), a novel clade consisting of all ‘landbirds’ to the exclusion of mousebirds (Fig. 1, branch H; two REs) and a close affinity\textsuperscript{12,13} among seriema, falcons, parrots and passerines (Fig. 1, branch I; two REs). Statistical testing\textsuperscript{32} of the support for these three branches is not applicable, as some of the above incongruent presence/absence patterns are also inconsistent to these (Supplementary Fig. S2; Supplementary Table S1).

Unexpectedly, we obtained a wealth of conflict-free retroposon markers for two branches that were previously proposed by the Hackett et al.\textsuperscript{15} study of nuclear intronic sequences, and which received relatively moderate bootstrap support in their study. Seven retroposon insertions are exclusively present in falcons, parrots and passerines, but absent in hawks, woodpeckers and other ‘landbirds’ (Fig. 1, branch J; seven REs, \( P = 0.0005, (7 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}); we therefore suggest the new name Eu Falconimorphae (true Falconimorphae) for this significantly supported monophyly. Most strikingly, the shared presence of three retroposon insertions solely in parrots and passerines (Fig. 1, branch K; three REs, \( P = 0.0370, (3 \ 0 \ 0) \), likelihood ratio test\textsuperscript{22}); see also Fig. 2 for sequence alignments) provides statistically significant evidence of parrots as the living sister group of the Passeriformes. To make this new phylogenetic resolution easily comprehensible, we propose the new name Psittacopasserae (parrots and passerines). It is worth noting that with this evidence, for the first time, passerines can be confidently placed within the avian tree of life.

Although our exhaustive zebra finch-based retroposon screening did not detect any evidence for incomplete lineage sorting within...
Eufalconimorphae, we cannot completely exclude the possibility of its occurrence in this part of the neoavian tree. Considering this, we expect that, once the genome sequence of a parrot or a falcon is available, parrot- or falcon-based retroposon screenings will permit an even stronger resolution of this issue and a reevaluation of the conflict-free support for Psittacopasserae reported here.

Reconstructing the chronology of Mesozoic retroposon activity. In addition to resolving phylogenetic controversies, our markers enabled us to reconstruct the temporal retroposon impact on the avian genome during early bird phylogeny via the comparison of these experimentally verified insertion events with computational estimates of retroposon activity. To determine a computational chronology of retroposon activities, 995 nested retroposons (retroposon subtypes were active during relatively short periods, it is possible that this period of extensive and accelerated speciation events was enabled us to reconstruct the temporal retroposon impact on the avian genome during early bird phylogeny via the comparison of these experimentally verified insertion events with computational estimates of retroposon activity. To determine a computational chronology of retroposon activities, 995 nested retroposons (retroposon subtypes were active during relatively short periods, it is possible that this period of extensive and accelerated speciation events was

Eufalconimorphae | DoI: 10.1038/ncomms1448 | www.nature.com/naturecommunications

ARTICLE

Figure 1 | Retroposon evidence for the early branching events in the avian tree of life. The tree topology is derived from our presence/absence matrix (Supplementary Table S1; Supplementary Software) utilizing maximum parsimony and considering representatives of the key avian lineages\textsuperscript{5,14}. Black filled circles (branch A) are bird-specific retroposon insertions (exhibiting a 6-nt deletion that is found in some avian CR1s, but absent in all other avian and reptilian CR1s, and dark grey balls represent retroposon presence/absence markers that are congruent with one another. Light grey balls on grey gradient (label F) are retroposon markers that were probably inserted at the very beginning of the neoavian radiation and were subjected to incomplete lineage sorting of retroposon dimorphisms, as they exhibit presence/absence patterns that are incongruent with one another and with some of the retroposon markers on the dashed branches. Nodes without retroposon support are not collapsed (but highlighted by asterisks) if they received very strong support in nucleotide sequence analyses\textsuperscript{5,14}. Higher-ranking taxa are in red letters (English terms in orange letters and parentheses), including the new taxa Eufalconimorphae (falcons + parrots + passerines) and Psittacopasserae (parrots + passerines), and some recently introduced superordinal groupings\textsuperscript{14,51,54}. Bird names in bold letters belong to the nearest bird icon.
chronology of retroposon activity during the Mesozoic Era of bird evolution. We identified retroposons that were inserted at the very beginning of the neoavian radiation and were probably subjected to incomplete lineage sorting, a phenomenon that likely accounts for some of the incongruent results from sequence-based phylogenies. Retroposons constitute unique tools for understanding such complex and otherwise irresolvable evolutionary scenarios. Furthermore, we have determined a statistically significant resolution of a later part of the neoavian radiation, namely, the sister group relationship of passerines and parrots (Psittacopasserae) and their mutual affinity to falcons (Falconiformes). Our retroposon evidence can serve as a robust prior hypothesis for future studies focusing on these bird taxa. As such, parrots and passerines not only share the ability to learn vocalization, but also have a direct common ancestor. Although hummingbirds are also vocal learners, our phylogeny indicates that they are only distantly related to Psittacopasserae; therefore it is most parsimonious to assume that their vocal learning capability evolved after the divergence of hummingbirds and swifts (Fig. 4). Nevertheless, the phylogenetic resolution of Psittacopasserae raises the question as to what extent the striking neuroanatomical and gene expression parallels (for example, the anterior-medial vocal pathway R5d) between parrots and oscine passerines (songbirds) are homologous and thus evolved in their shared ancestor (Fig. 4). Behavioural and neuroanatomical data on 'non-osce' passerines (Suboscines and Acanthisiidae) is scarce and, to our knowledge, limited to New World Suboscines, suggesting that some representatives do not learn vocalizations (that is, Tyrannidae8-13), whereas others possibly do (that is, the

Figure 2 | Alignment of presence/absence regions of three monophyly markers for Psittacopasserae. Potential target site duplications (direct repeats) are in black boxes, 5’ and 3’ ends of the retroposon insertions are shown in lower case letters in grey boxes.
earlier-branching\textsuperscript{38} Cotingidae\textsuperscript{39} and Pipridae\textsuperscript{39}). Thus, to assume that vocal learning evolved in the psittacopassarian ancestor (with a secondary loss in at least one lineage of suboscine passerines) seems more parsimonious than hypothesizing four independent evolutions of vocal learning within Psittacopasserae. Accordingly, the emergence of vocal learning of songbirds would have happened at least 30 million years\textsuperscript{39} earlier than evident from the previous assumption of the independent evolution of cerebral vocal nuclei\textsuperscript{10} in parrots and in (oscine) passerines. Thorough reevaluation of this issue will impact various conclusions drawn from passerines and might thereby change our current understanding of the evolution of vocal learning in general.

**Methods**

**General approach.** We used three different search strategies to computationally screen over 200,000 REs present in the chicken and zebra finch genomes (see Supplementary Table S1 for information on the contribution of each strategy to the phylogenetically informative markers). On the basis of their suitability for cross-species PCR amplification (that is, only retrotransposition events situated in well-conserved intronic or intergenic regions smaller than 1.5 kb were considered), we identified 131 CR1 and 75 LTR candidate RE-containing loci. These loci were then experimentally screened in a reduced taxon sampling (comprising Aves, B; Neognathae; D, Neoaves; F, incongruent markers; G, ‘landbirds’; H, ‘landbirds’ without mousebirds; I, Eufalconiformes + sieriemas; J, Eulidaformes, K, Psittacopasserae; L, Passeriformes). CR1 retrotransposons are highlighted in blue and LTR retrotransposons are shown in red. The dashed bracket consists of retrotransposon markers that were inserted during the neovian radiation; the grey dashed vertical line indicates the estimated end of the Mesozoic Era at the Cretaceous/Tertiary boundary\textsuperscript{35}. We note that the exceptionally long TinT activity range of the CR1-E_Pass element (no. 4) is most probably an overestimation because of CR1 subfamily misidentification, as only a few diagnostic nucleotides distinguish this retrotransposon from other CR1 subfamilies.

**In silico screening.** Initially, (first strategy; a) genomic three-way alignments (comprising emu, chicken, and zebra finch) were compiled by MAFFT\textsuperscript{35} (FFT-NS-2, version 6, http://mafft.cbrc.jp/alignment/server/index.html) using –2.55 million bp of emu genomic contigs available in GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) and the corresponding regions in the chicken and zebra finch genomes (assemblies galGal3 and taGut1 in Genome Browser\textsuperscript{40}, http://genome.ucsc.edu/cgi-bin/hgBlat). REs were annotated using CENSOR (http://www.girinst.org/censor/index.php), and retrotransposition loci situated in well-conserved intronic or intergenic regions were chosen for primer generation. To identify additional candidate loci (first strategy; b), all avian sequences available in GenBank were screened for REs and (if a retrotransposition was present) aligned to the corresponding regions in the chicken and zebra finch genomes using MAFFT (E-INS-I, version 6). Second strategy: based on the insights gained by strategy I into the phylogenetic informativeness of representatives of certain CR1 and LTR subfamilies for our phylogenetic questions of interest, whole-genome in silico screenings for selected retrotransposons were conducted. This was done by extracting retrotransposition events including their flanking sequences (13 kb of each flank) from chicken or zebra finch genomes and BLAST screening these against chicken annotated unique exonic sequences to obtain well-conserved loci (<1.5 kb). Alternatively, retrotranscription consensus sequences from Refbase (http://www.girinst.org/repbase/index.html) were BLAST\textsuperscript{41} screened against chicken or zebra finch genomes and well-conserved loci in introns (of any size) or intergenic regions were chosen for primer generation. Third strategy: a CR1-enriched retroposon library of emu genomic DNA was constructed via a protocol utilizing digestion and circularization of genomic DNA and subsequent inverse PCR\textsuperscript{42}. A total of 242 clones were sequenced and BLAST screened against chicken and zebra finch genomes to find CR1 insertions (situated in well-conserved regions) specific to the lineage leading to the emu and suitable for experimental presence/absence screening.

**Taxon sampling.** Our whole taxon sampling (voucher numbers of the samples in the LWL-DNA- and Gewebearchiv of the Museum für Naturkunde Münster are specified) consisted of representatives of the key lineages\textsuperscript{43} within Palaeognathae (Struthio camelus (LWL00446), Pterocnemia pennata (LWL00447), Eudromia elegans (LWL00448), Dromaius novaehollandiae (LWL00449), Galloanserae (Dendrocygna viduata (LWL00450), Anas crecca (LWL00451), Aleleula lathami (LWL00452), Gallus gallus (LWL00453) and Neoeos (Chrysolampis mosquitus (LWL00458), Aplius aput (LWL00459), Oppiochomus hoazin (LWL00457), Phoenicurus ruber roose (LWL00454), Tachyis bupitus raficollis (LWL00455)/Podiceps cristatus (LWL00456), Columba palumbus (LWL00408), Corvus cornix (LWL00460)/Caculus canorus (LWL00461), Galaxia pavo (LWL00462),
Cloning Vector (Qiagen) and electroporation into TOP10 cells (Invitrogen). All nucleotide sequences were deposited in GenBank (accession numbers JF918959–JF916445). To complete our taxon sampling, we also used previously published sequences available in Genome Browser (assemblies gaGal37 and taeGut1) and GenBank (accession numbers AB112956, AB235826, AB235829, AC153776, AC158282, AC158286, AC160232, AF255979, AF255980, DP000685, DP000802, FJ279549–FJ279555, FJ279558–FJ279573 and FJ279576–FJ279590). Some of the sequence data (44) (emz RAC sequences AC153776, AC158282–AC158286, AC158288–AC158294, Alcedo atthis, and Diphasia troglodytes RAC sequences DP000795 and DP000976) were generated by the National Institutes of Health Intramural Sequencing Center (http://www.nisc.nih.gov). The lizard genome sequence (assembly anaCari1 in Genome Browser) was generated by the Broad Institute (http://www.broadinstitute.org).

All sequences of each marker were first automatically aligned using MAFFT (v6, version 6) and then manually realigned (see Supplementary Fig. S5 for 51 full sequence alignments). Each alignment was carefully inspected and the retroposon insertion considered a phylogenetically informative marker if, in all species sharing this RE, it featured an identical orthologous genomic insertion point (target site), identical RE orientation, identical RE subtype, identical target site duplications (direct repeats, if present) and a clear absence in other species. Candidate markers exhibiting low complexity or having sequences were excluded from the analysis to minimize the possibility of inconsistencies caused by precise RE excision as reported by van de Lagemaat et al. (45).

In the case of CR1 retroposon insertions shared among all the investigated bird lineages (markers A-1 to A-6), we initially aligned the avian retroposon flanks to the corresponding BAC sequences of the alligator available in GenBank (DP000795 and DP000976). Because of the ~220 million years of bird/crocodilian sequence divergence (46), a classical presence/absence situation could not be ascertained. Although CR1 elements are also found in the genomes of other non-mammalian amniotes (47–49), we consider these retroposon insertions to be suitable markers for the monophyly of birds, because each of them exhibits a diagnostic 6-nt deletion that is only present in a few bird-based CR1 subtypes (that is, CR1-Y and CR1-Z) but not in CR1 elements of other amniotes (that is, all BLAST and BLAT search hits of avian CR1 against available genome or BAC sequences of alligator, lizard, turtle, platypus and human were inspected by eye; see Supplementary Fig. S1 for a structural comparison of the well-conserved terminal regions (50) of amniote CR1 retroposons including lineage-specific diagnostic insertions or deletions). The majority-rule consensus sequences of the previously unrecognized CR1 subtypes ‘ALL-LINEa’, ‘ALL-LINEb’ and ‘ANO-LINE’ were derived from 17, 25 and 10 BLAST hits, respectively.

On the basis of the presence/absence matrix of our 51 phylogenetically informative markers (Supplementary Table S1), our phylogenetic tree was drawn by hand considering maximum parsimony and independently verified by a maximum parsimony analysis of a 1/0-coded version of our presence/absence matrix (Supplementary Software) in PAUP* (version 4.0b10; using the ‘tree’ option of character transformation, heuristic search with 1000 random sequence additions, and TBR branch swapping). This yielded one strict consensus parsimony tree (Fig. 1, consistency index = 0.895 and tree length = 57) derived from 57 equally parsimonious trees.

TinT analysis. To determine a chronology of retroposon activity periods, we used the web-based TinT analysis (http://www.comp.gen.uni-muenster.de/tools/tint/). As input data, the precomputed RepeatMasker files (hosted on the server) from chicken or zebra finch were selected. Only the retroposon subtypes present in the respective figures (see Fig. 2 for the zebra finch TinT of 995 nested REs or Supplementary Fig. S3 for the chicken TinT of 2355 nested REs) were included in the analysis (but note that, in the case of the zebra finch TinT, the retroposons CR1-YB1, Tgu and TgULTRs were added to the analysis but excluded from Fig. 3) using default parameters. The resultant graph of normal distributions of retroposon activity (ovaries represent 75%, vertical lines 95% and horizontal lines 99% of the probable activity period) was plotted on a simplified chronogram (44) using the experimentally verified retroposon insertions of Figure 1 as calibration points (for example, the succession of TguULTRs to TgULTRs activity in the zebra finch ancestor’s genome after the divergence of Gallus gallus and Neoraves). For this purpose, we considered the chronogram by Pereira and Baker (51) to be most suitable, as it includes molecular divergence times for the Crocodylia/Aves split, the Palaeognathae/Neognathae split, the neovaration, and the Acanthisitta/oscine Passeriformes split (other analyses of molecular divergence times (52–54) have only investigated a few of these dates).

References

1. Warren, W. C. et al. The genome of a songbird. Nature 461, 757–762 (2010).
2. Clayton, D. F., Balakrishnan, C. N. & London, S. E. Integrating genomes, brain and behavior in the study of songbirds. Curr. Biol. 19, R865–R873 (2009).
3. Cracraft, J. The major clades of birds: in The Phylogeny and Classification of the Tetrapods, Vol. 2: Aves, Suborders, Reptiles, Birds (ed. Benton, M. J.) 339–361 (Clarendon Press, Oxford, 1988).
4. Lively, B. C. & Zusi, R. L. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool. J. Linn. Soc. 114, 1–95 (2007).
5. Tazuke, G., Manegold, A. & Johansson, U. S. Monophyletic groups within ‘higher land birds’—comparison of morphological and molecular data. J. Zool. Syst. Evol. Res. 41, 233–248 (2003).

Figure 4 | Evolution of vocal learning in birds. Schematic brain drawings (adapted from Jarvis et al. (23) depict the hearing- and vocalizing-induced ZENK transcription factor expressions in hummingbirds, budgerigars (parrot representatives) and songbirds (passerine representatives). Our robust phylogenetic framework implies that some traits associated with vocal learning (for example, the anterior–medial vocal pathway (red)) are potentially homologous and thus evolved as an autopomorphitic trait (black square) in the psittacopasseran ancestor, but also independently in the distantly related ancestor of hummingbirds (dashed lines indicate that several neovarian lines are not shown). Neuroanatomical studies on early branching passerines (that is, Suboscines and Acanthisittidae) are necessary to infer that the posterior–lateral vocal pathway (green, located at different brain regions in parrots and oscine passerines) is either homologous or evolved independently in the two lineages. The caudal auditory pathway (blue) is a plesiomorphic trait (white square) and was probably inherited from a common avian ancestor (24). The putative location of the auditory pathway in falcons and swifts is not shown, as ZENK expression patterns have, to our knowledge, not yet been investigated in these birds. Scale bar, 2 mm.
20. Nishihara, H., Maruyama, S. & Okada, N. Retroposon analysis and recent... 1370–1740 (2006).

23. Jarvis, E. D. et al. Behaviourally driven gene expression reveals song nuclei in hummingbird brain. Nature 406, 628–632 (2000).

24. Jarvis, E. D. & Mello, C. V. Molecular mapping of brain areas involved in parrot hummingbird brain. J. Comp. Neurol. 426, 182–196 (2000).

25. Tello, J. G., Moyle, R. G., Marchese, D. J. & Cracraft, J. Phylogeny and phylogenetic classification of the tyrant flycatchers, cotigas, manakins, and their allies (Aves: Tyrannidae). Cladistics 25, 1–39 (2009).

39. Trainer, J. M., McDonald, D. B. & Learn, W. A. The development of coordinated singing in cooperatively displaying long-tailed manakins. Behav. Ecol. 13, 65–69 (2002).

40. Wada, K., Sakaguchi, H., Jarvis, E. D. & Hagiwara, M. Differential expression of glutamate receptors in avian neural pathways for learned vocalization. J. Comp. Neuro. 476, 44–64 (2004).

41. Kotah, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 9, 286–298 (2008).

42. Fujita, P. A. et al. The UCSC Genome Browser database: update 2011. Nucleic Acids Res. 39, D876–D882 (2011).

43. Kent, W. J. BLAT—the BLAST-like alignment tool. Genome Res. 12, 656–664 (2002).

44. Wang, W. & Kirschnes, E. F. Short interspersed elements (SINEs) are a major source of canine genomic diversity. Genome Res. 15, 1798–1808 (2005).

45. Bush, K. L. DNA extraction technique for plucked body contour feathers, molted body contour feathers, and plucked down using the QIAamp DNA Micro Kit, http://www.qiagenomics.com/DNA_Extraction_Protocols/plucked_contour_feathers.html (2005).

46. Malagò, W. Jr., Franco, H. M., Matheucci, E. Jr., Medaglia, A. & Henrique-Silva, F. Large scale sex typing of ostriches using DNA extracted from feathers. BMC Biotechnol. 2, 19 (2002).

47. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432, 695–716 (2004).

48. Shedlock, A. M. Phylogenomic investigation of CR1 LINE diversity in reptiles. Syst. Biol. 55, 902–911 (2006).

49. Shedlock, A. M. et al. Phylogenomics of non-avian reptiles and the structure and evolution of the ancestral amniote genome. Proc. Natl Acad. Sci. USA 104, 2767–2772 (2007).

50. Chapus, C. & Edwards, S. V. Genome evolution in Rutilia: in silico chicken mapping of 12,000 BAC-end sequences from two reptiles and a basal bird. BMC Genomics 10, 58 (2009).

51. Shedlock, A. M. & Edwards, S. V. Amniotes (Amniota): in The Timetree of Life (eds Hedges, S.B. & Kumar, S.) 375–379 (Oxford University Press, 2009).

52. Haas, N. B. et al. Subfamilies of CR1 non-LTR retrotransposons have different 5’UTR sequences but are otherwise conserved. Genes 265, 175–183 (2001).

53. Sangster, G. A name for the flamingo–grebe clade. Ibis 147, 612–615 (2005).

54. Mayr, G. Phylogenetic relationships of the paraphragyl ‘caprimulgiform’ birds (nightjars and allies). J. Zool. Syst. Evol. Res. 48, 126–137 (2010).

Acknowledgements
We thank Judith Brooknesh for help with in vitro experiments and Werner Beckmann, Timm Spreitke (Zoo Halle), Stephanie Hodges, Sharon Birks (Burke Museum), Sandra Silmsik (Zoo Münster), Renate van den Elzen (Museum Koenig), Ommo Huppol, Nils Anthes, Michael Wink, the LWL-DNA- and Gewebearbeic, Joes Custer, Holger Schiezelt, Georg Mayer (Senckenberg Museum), Elisabeth Suh, Christoph Bleidorn and Andrew Fidler for providing feather, blood and tissue samples. Jón Baldur Hlíberg provided the bird paintings and Marsha Bundman helped with editing. We thank Gerald Mayer for valuable comments. This research was funded by the Deutsche Forschungsgemeinschaft (KR3639 to I.O.K. and J.S.) and the Medizinische Fakultät der Westfälischen Wilhelms-Universität Münster.

Author contributions
J.O.K. and I.S. conceived the project. A.S. designed the study. A.S. and G.C. performed in silico experiments. A.S., M.P. and M.K. performed in vitro experiments. A.S. analysed the data and wrote the manuscript. E.A.F., J.B., J.O.K. and J.S. contributed reagents, materials and analysis tools. G.C., J.B., I.O.K. and J.S. discussed and commented on the data and the manuscript.

Additional information
Accession codes: The nucleotide sequences have been deposited in GenBank database under accession numbers JF915895–JF916445.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Suh, A. et al. Mesotonic retrotransposons reveal parrots as the closest living relatives of passerine birds. Nat. Commun. 2:443 doi: 10.1038/ncomms1448 (2011).

License: This work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/