



# Identification of dopamine receptors across the extant avian family tree and analysis with other clades uncovers a polyploid expansion among vertebrates

Asher Haug-Baltzell<sup>1,2</sup>, Erich D. Jarvis<sup>3,4</sup>, Fiona M. McCarthy<sup>5,6</sup> and Eric Lyons<sup>2,6,7\*</sup>

<sup>1</sup> Arizona Biological/Biomedical Sciences Program, University of Arizona, Tucson, AZ, USA, <sup>2</sup> Genetics GIDP, University of Arizona, Tucson, AZ, USA, <sup>3</sup> Department of Neurobiology, Duke University Medical Center, Durham, NC, USA, <sup>4</sup> Howard Hughes Medical Institute, Chevy Chase, MD, USA, <sup>5</sup> School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, AZ, USA, <sup>6</sup> BIO5 Institute, University of Arizona, Tucson, AZ, USA, <sup>7</sup> The School of Plant Sciences, University of Arizona, Tucson, AZ, USA

## OPEN ACCESS

### Edited by:

Igor Ponomarev,  
The University of Texas at Austin, USA

### Reviewed by:

Philippe Vernier,  
Centre National de la Recherche  
Scientifique, France  
Carolina Frank-Vilches,  
Max Planck Institute for Ornithology,  
Germany

### \*Correspondence:

Eric Lyons,  
BIO5 Institute, School of Plant  
Science, University of Arizona, 1657  
E. Helen St. Tucson, AZ 85745, USA  
ericlyons@email.arizona.edu

### Specialty section:

This article was submitted to  
Neurogenomics,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 26 July 2015

**Accepted:** 21 September 2015

**Published:** 07 October 2015

### Citation:

Haug-Baltzell A, Jarvis ED,  
McCarthy FM and Lyons E (2015)  
Identification of dopamine receptors  
across the extant avian family tree and  
analysis with other clades uncovers a  
polyploid expansion among  
vertebrates. *Front. Neurosci.* 9:361.  
doi: 10.3389/fnins.2015.00361

Dopamine is an important central nervous system transmitter that functions through two classes of receptors (D1 and D2) to influence a diverse range of biological processes in vertebrates. With roles in regulating neural activity, behavior, and gene expression, there has been great interest in understanding the function and evolution dopamine and its receptors. In this study, we use a combination of sequence analyses, microsynteny analyses, and phylogenetic relationships to identify and characterize both the D1 (DRD1A, DRD1B, DRD1C, and DRD1E) and D2 (DRD2, DRD3, and DRD4) dopamine receptor gene families in 43 recently sequenced bird genomes representing the major ordinal lineages across the avian family tree. We show that the common ancestor of all birds possessed at least seven D1 and D2 receptors, followed by subsequent independent losses in some lineages of modern birds. Through comparisons with other vertebrate and invertebrate species we show that two of the D1 receptors, DRD1A and DRD1B, and two of the D2 receptors, DRD2 and DRD3, originated from a whole genome duplication event early in the vertebrate lineage, providing the first conclusive evidence of the origin of these highly conserved receptors. Our findings provide insight into the evolutionary development of an important modulatory component of the central nervous system in vertebrates, and will help further unravel the complex evolutionary and functional relationships among dopamine receptors.

**Keywords:** dopamine receptors, comparative genomics, genome evolution, polyploidy, whole genome duplication

## Introduction

Dopamine (DA) is an important neurotransmitter that functions in the central nervous system of vertebrates. DA functions through two main classes of receptors, designated D1 (Class 1) and D2 (Class 2), to influence neural activity, behavior, and gene expression (Stoof and Keibarian, 1984; Beaulieu and Gainetdinov, 2011). All DA receptors are integral membrane proteins and belong to the rhodopsin family of G protein-coupled receptors, but the two classes are structurally,

functionally, and genetically distinct. D1 receptors consist of a single exon open reading frame with no introns, contain a short third cytoplasmic loop and long cytoplasmic C-terminal stretch, and activate adenylyl cyclase resulting in increased cAMP levels (Callier et al., 2003; Le Crom et al., 2003; Yamamoto et al., 2013). Conversely, the D2 receptors possess multiple exons and introns, have a long third cytoplasmic loop and a short C-terminus stretch that remains anchored in the membrane, and reduced adenylyl cyclase activity (Callier et al., 2003; Le Crom et al., 2003). The source of DA and the expression of its associated receptors are well conserved between birds and mammals, and are important in brain modulation of many behaviors, including reproduction, learning, vocalization, addiction, and reward, making them of particular interest in both health and agricultural applications (Schnell et al., 1999; Chaiseha et al., 2003; Sasaki et al., 2006; Sartssoongnoen et al., 2008; Kubikova et al., 2010).

Dopamine and its receptors have been studied both functionally and evolutionarily. The D1 family is currently known to include four paralogous receptors DRD1A (human DRD1), DRD1B/X (human DRD5), DRD1C/D, and DRD1E (Callier et al., 2003; Yamamoto et al., 2013). The D2 family includes three paralogous receptors DRD2, DRD3, and DRD4 (Callier et al., 2003). Not all dopamine receptors are present in all species, and lineage-specific duplications have increased receptor number in some species (Yamamoto and Vernier, 2011). For example, only two D1 receptors DRD1A and DRD1B have been identified in mammals, while in birds a combination of DRD1A, DRD1B, DRD1C, and/or DRD1E may be present (Kubikova et al., 2010; Yamamoto et al., 2013). In the teleost fishes, DRD1B has duplicated and formed the DRD1X paralog (Yamamoto et al., 2013). It is commonly accepted that the two classes of vertebrate dopamine receptors are no more closely related to each other than to other classes of monoamine receptors, suggesting that they likely converged upon binding dopamine (Callier et al., 2003; Yamamoto and Vernier, 2011; Robertson et al., 2012; Yamamoto et al., 2013). However, questions still remain regarding the complex functional and evolutionary relationships between the receptors of each class.

Here, we identified and characterized D1 (DRD1A, DRD1B, DRD1C, and DRD1E) and D2 (DRD2, DRD3, and DRD4) dopamine receptor gene families in 43 recently sequenced bird genomes spanning the avian family tree at the ordinal level (Table 1) (Zhang et al., 2014). We used phylogenetic relationships, macro- and microsynteny, and comparisons to homologous receptors in other vertebrate genomes to make significant advances in our understanding of the evolutionary origins of D1 and D2 receptors among birds and other vertebrates.

## Materials and Methods

### Sequence Data

To identify putative dopamine receptors across the avian lineage, we used the DNA and protein coding sequences of previously identified D1 and D2 receptors in Chicken, Turkey, and Zebra Finch to BLAST against 43 unmasked avian genome sequences from the Avian Phylogenomics Consortium (Jarvis et al., 2014;

Zhang et al., 2014). The 43 species were chosen based on being representative of all major orders of modern birds, and having a now well-vetted species-level phylogenetic trees (Jarvis et al., 2014). All avian BLAST query reference sequences were obtained from the Ensembl Database v81 (Flicek et al., 2014). Accession numbers are listed in Table 3. BLAST searches were performed with CoGeBlast (Lyons and Freeling, 2008), and results were filtered based on *e*-value and HSP coverage to remove false positives (Schnable and Lyons, 2012). The remaining putative DA receptor gene protein coding sequences were exported and compiled into multi-FASTA files. To serve as an outgroup, the protein coding sequence for a Ciona D1-like receptor was acquired from GenScript (GenScript)<sup>1</sup>, and appended to the multi-FASTA (Supplementary File 4).

To infer an evolutionary model for the origin of DA receptors, DA receptor protein coding sequences were obtained for a range of vertebrate and invertebrate species, including human, mouse, chicken, turkey, peregrine falcon, alligator, fugu, lamprey, ciona, and drosophila. Protein coding sequences for chicken, turkey, human, and mouse D1 and D2 receptors, and fugu D2 receptors were obtained from Ensembl. Fugu D1 sequences were identified with CoGeBLAST using the orthologous chicken gene coding sequences as queries and confirmed via microsynteny analyses (Supplemental File 3). The remaining bird sequences were identified in this study. Alligator sequences were obtained using CoGe tool SynFind (<https://genomeevolution.org/CoGe/SynFind.pl>) using bird receptors as queries to identify orthologous genes (and genomic regions) in alligator. Ciona D1-like was obtained from GenScript. To serve as the outgroup, Drosophila dopamine 1 receptor (Dop1) protein coding sequence was acquired from Flybase (Pierre et al., 2014). All coding sequences were compiled into a multi-FASTA file (Supplementary File 5) for alignment and phylogenetic analysis. Accession IDs for all sequences obtained from external resources other than CoGe may be found in Table 3.

### Phylogenetic Analysis

All phylogenetic analyses were generated from protein-coding sequences (Figures 1A,B, 7A). For each analysis, the respective multi-FASTA file(s) were uploaded to the iPlant Data Store (Goff et al., 2011). Sequences were aligned into PHYLIP Interleaved format using the iPlant Discovery Environment (DE) implementation of MUSCLE 3.8.31 (Supplementary Files 6, 7) (Edgar, 2004). To confirm the validity of programmatic (fully automated) alignment, the multiple sequence alignments used for construction of all phylogenies were visualized using Jalview (Waterhouse et al., 2009). Minor errors were observed, but manual adjustment did not create well-supported (bootstrap values) improvements to the gene-tree in regards to species-level relationships (Supplemental File 2: Figures S2A, S2B). As such, automatic alignments were used for all analyses, which also increases the reproducibility of our study. Phylogenetic analyses were performed on the aligned sequences using a

<sup>1</sup>GenScript - Make Research Easy - The leader in molecular cloning and gene synthesis, peptide synthesis, protein and antibody engineering. Available online at: <http://www.gencript.com/> (Accessed December 19, 2014).

customized iPlant implementation of RAxML 8.1.2.2 (“Protein-Based RAxML Analysis”) (Goff et al., 2011; Stamatakis, 2014). Best ML trees were calculated from 20 replicates using the PROTGAMMA Model and WAG AA substitution matrix. To generate a final tree with bootstrap support values, 100 bootstrap replicates were computed, using the same model as for best tree calculations, and the values were drawn onto the best tree using the PROTCAT model. Phylogenetic analyses can be regenerated from the Supplementary Material

multi-FASTA files (Supplementary Files 4, 5) with the iPlant DE tools “MUSCLE 3.8.31” and “Protein-Based RAxML Analysis.” Alternately, phylogenies can be regenerated directly from the multiple sequence alignments (Supplementary Files 6, 7) using only iPlant DE tool “Protein-Based RAxML Analysis” or any RAxML 8.1.2 installation. Tree visualizations with bootstrap support values were performed using Dendroscope (Huson and Scornavacca, 2012), and visualizations with branch lengths using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

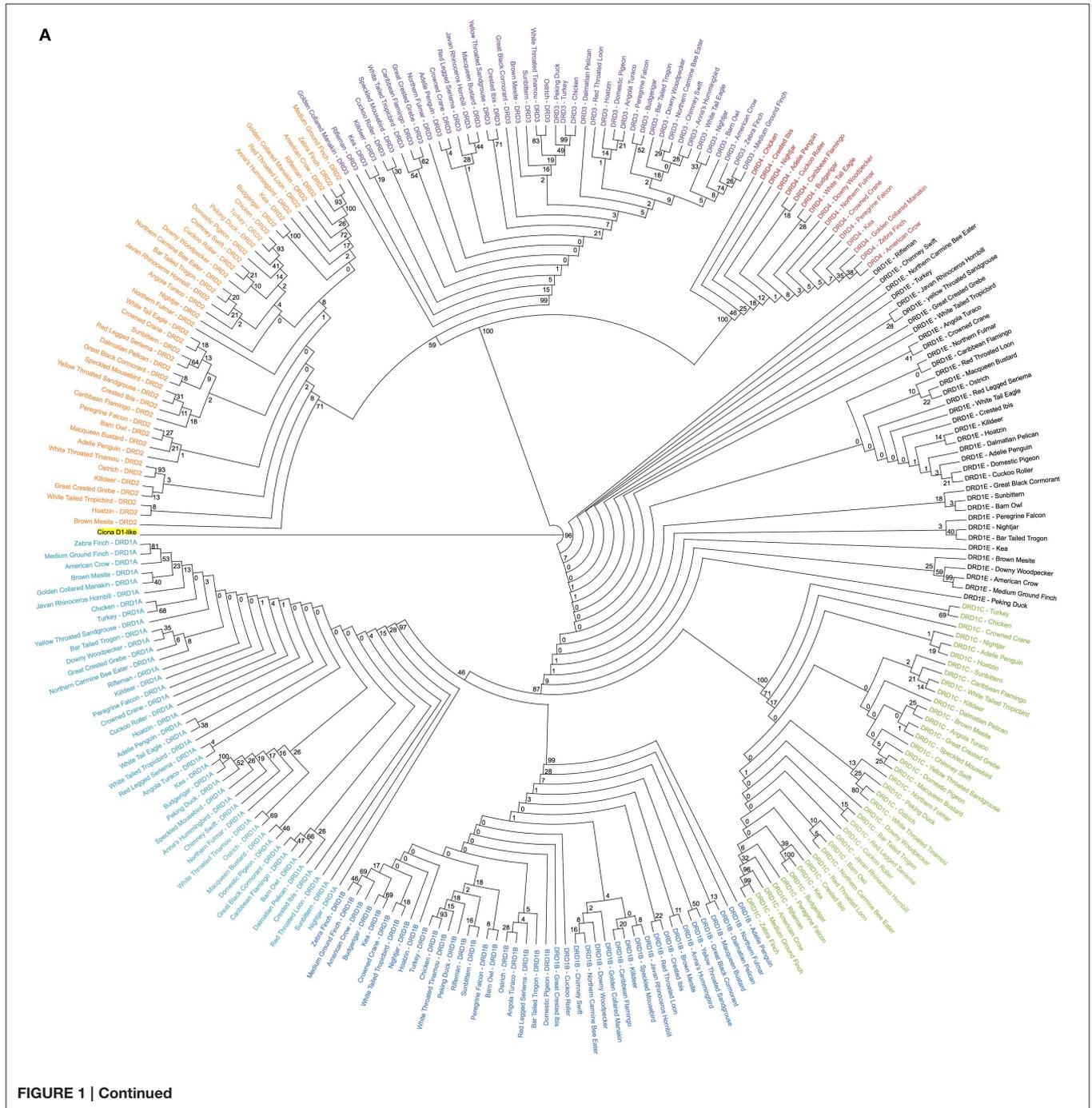
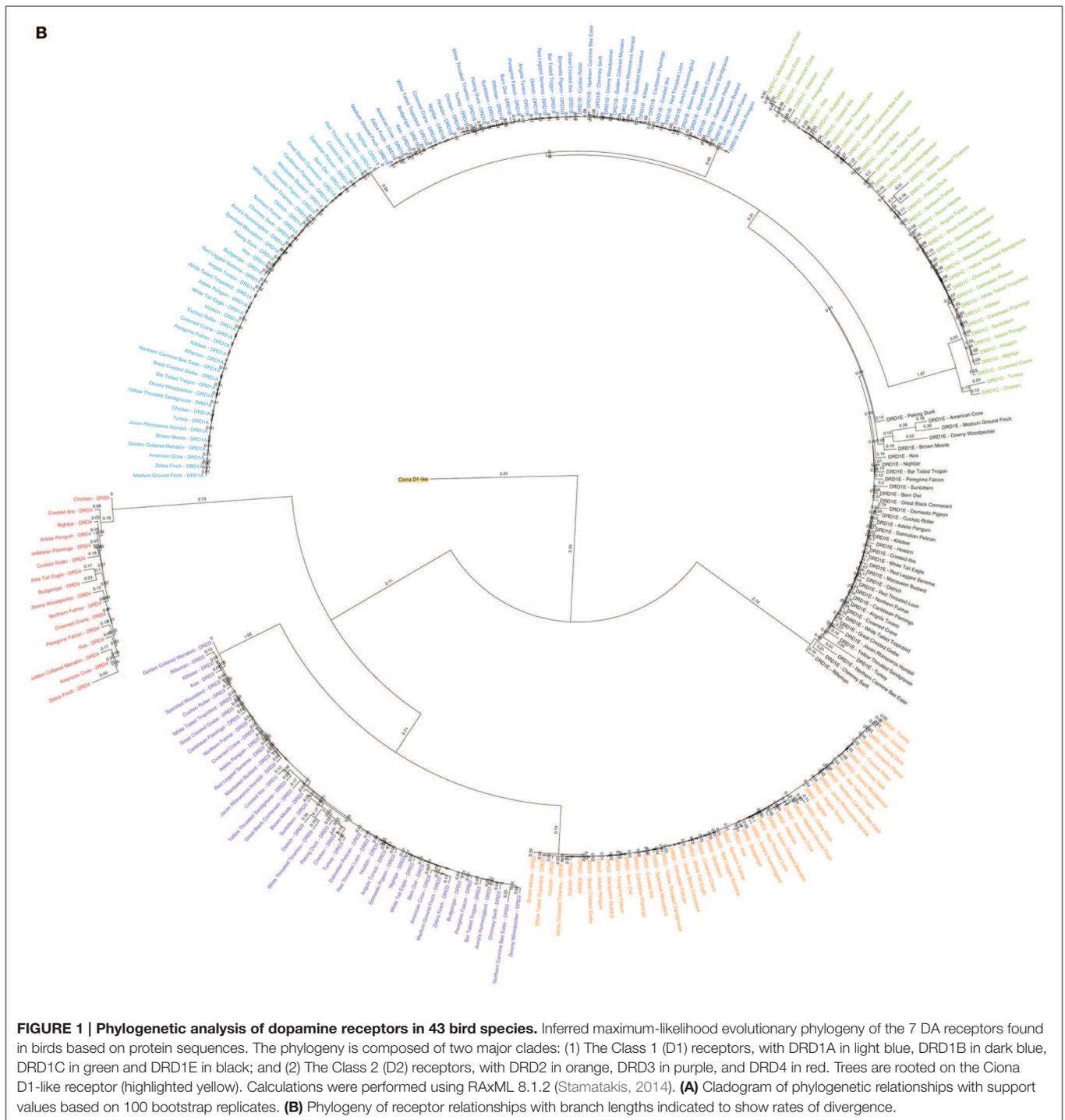


FIGURE 1 | Continued



### Synteny Analysis Microsynteny

Genomic microsynteny analyses were performed to confirm putative gene identity and to differentiate between absent and unannotated receptors. All microsynteny visualizations were generated using the GEvo tool, available through the comparative genomics platform CoGe ([www.genomevolution.org](http://www.genomevolution.org)). GEvo

plots annotated gene features along chromosome panels of user-designated lengths, and uses LAST algorithms to identify regions of high sequence similarity. Regions with high sequence similarity are physically connected using colored boxes and lines. Annotated features overlapped by these regions are colored purple, so that users can infer collinear gene order of homologous gene pairs between multiple species/chromosomes. All GEvo

microsynteny analyses have links provided in figure legends to regenerate the analysis and visualization.

### Genome Wide Synteny

To identify genome wide duplication patterns and investigate the nature of the DA receptor duplications, the CoGe tool SynMap was used to generate a full-genome syntenic dot plot between Chicken and Turkey (**Figure 5**). Synonymous mutation (Ks) values for each syntenic gene pair were calculated using CODEML (Yang, 2007). Ks values were plotted onto a histogram, and a color scheme was applied across the histogram. The corresponding colors were mapped onto each dot to visualize genome-wide duplication event patterning. Dopamine receptor genes were located on the plot to determine how their duplication history fits with the genome-wide duplication history. The SynMap analysis has a link provided in the figure legend to regenerate the analysis and visualization.

### Raw Reads BLAST Search

To confirm the gene absences indicated by microsynteny analysis, BLAST searches were performed against the raw sequencing reads for target bird genomes. Raw reads were extracted from SRA files obtained from the NCBI BioProject for each avian species (**Table 1**), and BLAST was performed using the NCBI blastn suite (Altschul et al., 1990). Program selection was optimized for highly similar sequences (megablast), with the following modifications made to default settings: “Expect threshold = 5,” “Word size = 16,” “Gap Costs = Existence: 5, Extension: 2.” For each search, the query sequence was composed of 5–6 known coding sequences for the receptor of interest. For each species investigated, a BLAST for DRD1A (always present) served as a control. All raw read BLAST search results are included in Supplemental File 1, and both BLAST and microsynteny results are summarized in **Table 1**.

### Human Ortholog Comparisons

To test for possible additional evidence into the WGD origin of the DRD1A/DRD1B receptors, maps comparing human chromosome 4 (containing DRD1, the DRD1A ortholog) and chromosome 5 (containing DRD5, the DRD1B ortholog) were generated using “Paralogons in the Human Genome” tool from the Wolfe lab ([http://wolfe.ucd.ie/cgi-bin/dup\\_528/chrom\\_plot](http://wolfe.ucd.ie/cgi-bin/dup_528/chrom_plot)) (McLysaght et al., 2002) (**Figure 6A**). Minimum threshold for paired proteins per block was set as 12, as values greater than 6 were shown as statistically likely as having originated from a whole genome duplication event (McLysaght et al., 2002), and a threshold of 12 produced only a single block overlapping the region of interest in our analysis. For DRD2/DRD3 receptors, similar maps comparing human chromosome 11 (containing DRD2) and chromosome 3 (containing DRD3) were generated using a minimum threshold for paired proteins per block of 7 (**Figure 6B**). Specific locations of the receptor genes were marked on the map by comparison with the human reference genome. Similar maps were generated for the other DA receptor pairs, but genes did not fall directly within regions shown as duplicated (not shown).

## Results

### Identification and Classification of DA Receptors in Bird Species

Sequence comparisons (BLAST) were used to identify putative DA receptor encoding genes in the 43 sequenced bird genomes (**Table 1**) (Zhang et al., 2014). Maximum-likelihood phylogenies, rooted on the outgroup D1-like receptor from *Ciona* (a tunicate invertebrate), were constructed from protein coding sequences of the putative DA receptors using RAXML (**Figure 1**) (Stamatakis, 2014). Gene identity was inferred by presence within specific receptor clades and further confirmed by syntenic analysis. Putative genes that did not have both phylogenetic and syntenic evidence were discarded.

Overall topography of the receptor phylogeny was similar to previous reports, with two major monophyletic clades comprised of the two major classes of DA receptors, D1 and D2 (**Figure 1**). The D1 clade was broken into 4 major sub-clades, each containing exclusively DRD1A, DRD1B, DRD1C, or DRD1E receptors. Phylogenetic relationships between the receptors, (DRD1E,(DRD1C,(DRD1A,DRD1B))), were congruent with those reported by Yamamoto et al. (2013). Within the D1 class, DRD1E was the fastest diverging DA receptor family, followed by DRD1C, with DRD1A and DRD1B diverging slowest. Additionally, within individual receptor family clades, DRD1C and DRD1E diverged at a rate 2–4 times faster than DRD1A and DRD1C (**Figure 1B**). Receptor clades for DRD1A, DRD1B, and DRD1C clades were monophyletic while DRD1E was polyphyletic, supporting the findings of Yamamoto et al. (2013). However, our phylogeny had significantly higher bootstrap support for the nodes supporting these clades: 87 vs. 21 for the node separating DRD1C and DRD1A/1B; 46 vs. 17 for the node separating DRD1A and DRD1B (**Figure 1A**) (Yamamoto et al., 2013).

The D2 clade was broken into three major sub-clades, each monophyletic for DRD2, DRD3, and DRD4 receptors. Branching relationships were congruent with previously published relationships (DRD4,(DRD2,DRD3)), with DRD4 diverging fastest, both as a family within the D2 class and as individual receptors within the DRD4 clade (Le Crom et al., 2003). While we had similarly high bootstrap support (100 vs. 99) for the node separating DRD4 and DRD2/3, our bootstrap value for the node separating DRD2 and DRD3 was lower (59 vs. 100) (**Figure 1A**) (Le Crom et al., 2003).

Species relationships within each receptor clade were not congruent with the avian species tree phylogeny (Jarvis et al., 2014) except for closely related species on the terminal branches. This result was expected as recent findings show that: (1) gene trees almost never correlate perfectly with species trees; (2) single genes have lower resolution (bootstrap support) than thousands of genes or whole genomes; and (3) protein coding genes have lower phylogenetic signal due to high sequence conservation, including G-coupled protein receptors (Jarvis et al., 2014). Additionally, minor errors in the alignment resulting from the automated alignment did not appear to be contributing to these differences, as they were not remedied by manual corrections of the alignment (Supplemental File 2). Despite the discrepancies in

**TABLE 1 | Avian genome information and summary of gene absence confirmation by microsynteny and BLAST searches of raw sequencing reads.**

Species	Common name	NCBI bioproject accession	Missing gene	Microsynteny	BLAST coverage	Citation
<i>T. guttata</i>	Zebra Finch	PRJNA17289	DRD1E	Unidentified	N/A	Warren et al., 2010
<i>G. fortis</i>	Medium Ground Finch	PRJNA156703	DRD4	Unidentified	Good	Zhang et al., 2014
<i>C. brachyrhynchos</i>	American Crow	PRJNA212869	N/A	N/A	N/A	Zhang et al., 2014
<i>M. vitellinus</i>	Golden Collared Manakin	PRJNA212872	DRD1C DRD1E	Absent Absent	Poor (30%) Good	Zhang et al., 2014
<i>A. chloris</i>	Rifleman	PRJNA212877	DRD4	Unidentified	Good	Zhang et al., 2014
<i>M. undulatus</i>	Budgerigar	PRJEB1588	DRD1E	Absent	Poor (34%)	Ganapathy et al., 2014
<i>N. notabilis</i>	Kea	PRJNA212900	N/A	N/A	N/A	Zhang et al., 2014
<i>F. peregrinus</i>	Peregrine Falcon	PRJNA159791	N/A	N/A	N/A	Zhan et al., 2013
<i>C. cristata</i>	Red Legged Seriema	PRJNA212889	DRD4	Unidentified	Good	Zhang et al., 2014
<i>M. numbicus</i>	Northern Carmine Bee Eater	PRJNA212898	DRD4	Unidentified	Good	Zhang et al., 2014
<i>P. pubescens</i>	Downy Woodpecker	PRJNA212874	N/A	N/A	N/A	Zhang et al., 2014
<i>B. silvestris</i>	Javan Rhinoceros Hornbill	PRJNA212887	DRD4	Unidentified	Good	Zhang et al., 2014
<i>A. vittatum</i>	Bar Tailed Trogon	PRJNA212878	DRD4	Unidentified	Good	Zhang et al., 2014
<i>L. discolor</i>	Cuckoo Roller	PRJNA212897	N/A	N/A	N/A	Zhang et al., 2014
<i>C. striatus</i>	Speckled Mousebird	PRJNA212892	DRD1E DRD4	Unidentified Unidentified	Good Good	Zhang et al., 2014
<i>T. alba</i>	Barn Owl	PRJNA212909	DRD4	Unidentified	Good	Zhang et al., 2014
<i>H. albicilla</i>	White Tailed Eagle	PRJNA212896	DRD1B	Unidentified	Good	Zhang et al., 2014
<i>P. crispus</i>	Dalmatian Pelican	PRJNA212901	DRD4	Unidentified	Good	Zhang et al., 2014
<i>N. nippon</i>	Crested Ibis	PRJNA232572	N/A	N/A	N/A	Zhang et al., 2014
<i>P. carbo</i>	Great Black Cormorant	PRJNA212903	DRD1C DRD4	Absent Unidentified	Poor (34%) Good	Zhang et al., 2014
<i>F. glacialis</i>	Northern Fulmar	PRJNA212894	N/A	N/A	N/A	Zhang et al., 2014
<i>P. adeliae</i>	Adelie Penguin	PRJNA235983	N/A	N/A	Good	Zhang et al., 2014
<i>G. stellata</i>	Red Throated Loon	PRJNA212895	DRD4	Unidentified	Good	Zhang et al., 2014
<i>P. lepturus</i>	White Tailed Tropicbird	PRJNA212902	DRD4	Unidentified	Good	Zhang et al., 2014
<i>E. helias</i>	Sunbittern	PRJNA212893	DRD4	Unidentified	Good	Zhang et al., 2014
<i>C. vociferous</i>	Killdeer	PRJNA212867	DRD4	Unidentified	Good	Zhang et al., 2014
<i>O. hoazin</i>	Hoatzin	PRJNA212873	DRD4	Unidentified	Good	Zhang et al., 2014
<i>C. anna</i>	Anna's Hummingbird	PRJNA212866	DRD1C DRD1E DRD4	Absent Unidentified Unidentified	Poor (15%) Poor (35%) Good	Zhang et al., 2014
<i>C. pelagica</i>	Chimney Swift	PRJNA210808	DRD4	Unidentified	Good	Zhang et al., 2014
<i>A. carolinensis</i>	Night Jar	PRJNA212888	N/A	N/A	N/A	Zhang et al., 2014
<i>C. macqueenii</i>	Macqueen Bustard	PRJNA212891	DRD4	Unidentified	Good	Zhang et al., 2014
<i>T. erythrolophus</i>	Angola Turaco	PRJNA212908	DRD4	Unidentified	Good	Zhang et al., 2014
<i>M. unicolor</i>	Brown Mesite	PRJNA212899	DRD4	Unidentified	Good	Zhang et al., 2014
<i>P. gutturalis</i>	Yellow Throated Sandgrouse	PRJNA212906	DRD4	Unidentified	Good	Zhang et al., 2014
<i>C. livia</i>	Domestic Pigeon	PRJNA167554	DRD4	Unidentified	Poor (33%)	Shapiro et al., 2013
<i>P. ruber</i>	Caribbean Flamingo	PRJNA212904	N/A	N/A	N/A	Zhang et al., 2014
<i>P. cristatus</i>	Great Crested Grebe	PRJNA212905	DRD4	Unidentified	Good	Zhang et al., 2014
<i>G. gallus</i>	Chicken	PRJNA13342	DRD1E	Unidentified	Good	Hillier et al., 2004
<i>M. gallopavo</i>	Turkey	PRJNA42129	DRD4	Unidentified	N/A	Dalloul et al., 2010
<i>A. domestica</i>	Peking Duck	PRJNA46621	DRD4	Unidentified	Good	Huang et al., 2013
<i>T. guttatus</i>	White Throated Tinamou	PRJNA212876	DRD1E DRD4	Unidentified Unidentified	Good Good	Zhang et al., 2014
<i>S. camelus</i>	Ostrich	PRJNA212875	DRD4	Unidentified	Good	Zhang et al., 2014

All bird species, including both scientific and common names, investigated in this study. Table is organized in order of bird phylogeny (Figure 2). Each species includes a reference to the original sequencing publication, the NCBI BioProject accession number for the raw data, and a summary of evidence supporting absence or lack of identification for each unidentified receptor. For microsynteny analyses, "Absent" refers to gene appearing deleted in microsynteny analyses, while "Unidentified" indicates sequence similarity but no intact gene annotation. BLAST coverage summarizes the extent to which raw reads covered a query receptor sequence when raw sequencing data (obtained from NCBI) was BLAST searched. "Poor" coverage ( $\leq 35\%$ ) suggests receptor is absent, while "good" ( $> 35\%$ ) coverage suggests receptor may be present. In cases where a receptor is indicated as missing, but raw reads coverage is listed as "N/A," raw reads for that species were either unavailable or corrupted and thus the analysis could not be performed on that species (i.e., Zebra Finch). Raw results and additional explanation are compiled in Supplementary File 1.

species level relationships, gene-level relationships demonstrated high support and were congruent with those previously reported (Kubikova et al., 2010; Yamamoto et al., 2013).

### Independent Preferential Loss of Derived DA Receptors Among Bird Species

Not all receptors could be identified in the genome assemblies of all species of birds (Figure 2), suggesting losses could have occurred in some lineages. To test and confirm this possibility, we performed additional BLAST searches against the raw genomic sequencing reads for each missing receptor to ensure that receptor identification was not hampered by incomplete assembly of the genome in current builds (examples in Figure 3, Table 1, Supplementary File 1).

For the D1 DA receptors, DRD1A was successfully identified and classified in all investigated species (Figure 2, closed circles). DRD1B was identified in the genome builds of all species except White-Tailed Eagle. However, microsynteny alignments showed a scaffold (#35946) that contained a section of DRD1B coding sequence on its terminal end, indicative that the gene was not fully assembled following genome sequencing (Figure 2, open circle). Additionally, when DRD1B receptor coding sequences from Nightjar, Peregrine Falcon, Northern Fulmar, Medium Ground Finch, and Turkey were used as BLAST queries against the raw genomic sequencing reads from White-Tailed Eagle, matching BLAST hits covered almost the entire length of the query sequences (Table 1, Supplementary File 1). Thus, the White Tailed Eagle DRD1B is present, but lacking in annotation due to an incomplete genome assembly. DRD1C was not identified in the genome assembly of three distantly related species, golden collared manakin, great black cormorant, and Anna's hummingbird. Microsynteny analyses and BLAST searches of the raw sequencing reads suggested DRD1C was truly absent in these species (Figure 2, no circle; Table 1, Supplementary File 1). DRD1E was not found in the genome assembly builds of 7 of the 43 species, but only one (parakeet/budgerigar) appeared as a genuine loss (Figure 2, open circles; Table 1, Supplementary File 1). Overall, DRD1A and DRD1B were present in all species, while DRD1C and DRD1E were lost in some lineages independently. Notably, D1 receptor losses appear to correspond to the fastest diverging receptors, DRD1C and DRD1E (Figure 1B).

For D2 DA receptors, DRD2 and DRD3 were present in the genome builds of all avian species, but we were unable to find the DRD4 receptor in the genome builds over half of the species. However, microsynteny analysis and BLAST searches of the raw sequencing reads showed fragments of DRD4 were present in all the remaining half of the species, either in unannotated assembled regions or in unassembled regions (Figures 2, 3C, Supplementary File 1). These absences were not limited to low coverage genomes, and included high-coverage, well-assembled genomes with scaffold N50s > 1Mb [medium ground finch, Dalmatian pelican, hoatzin, Anna's hummingbird, Chimney swift, pigeon, duck, and ostrich (Zhang et al., 2014)]. However, coverage and genome N50 appears to contribute to these findings, as 17 of 21 (81%) of low coverage/N50 < 1 Mb genomes had unassembled DRD4 reads, whereas only 7 of 22

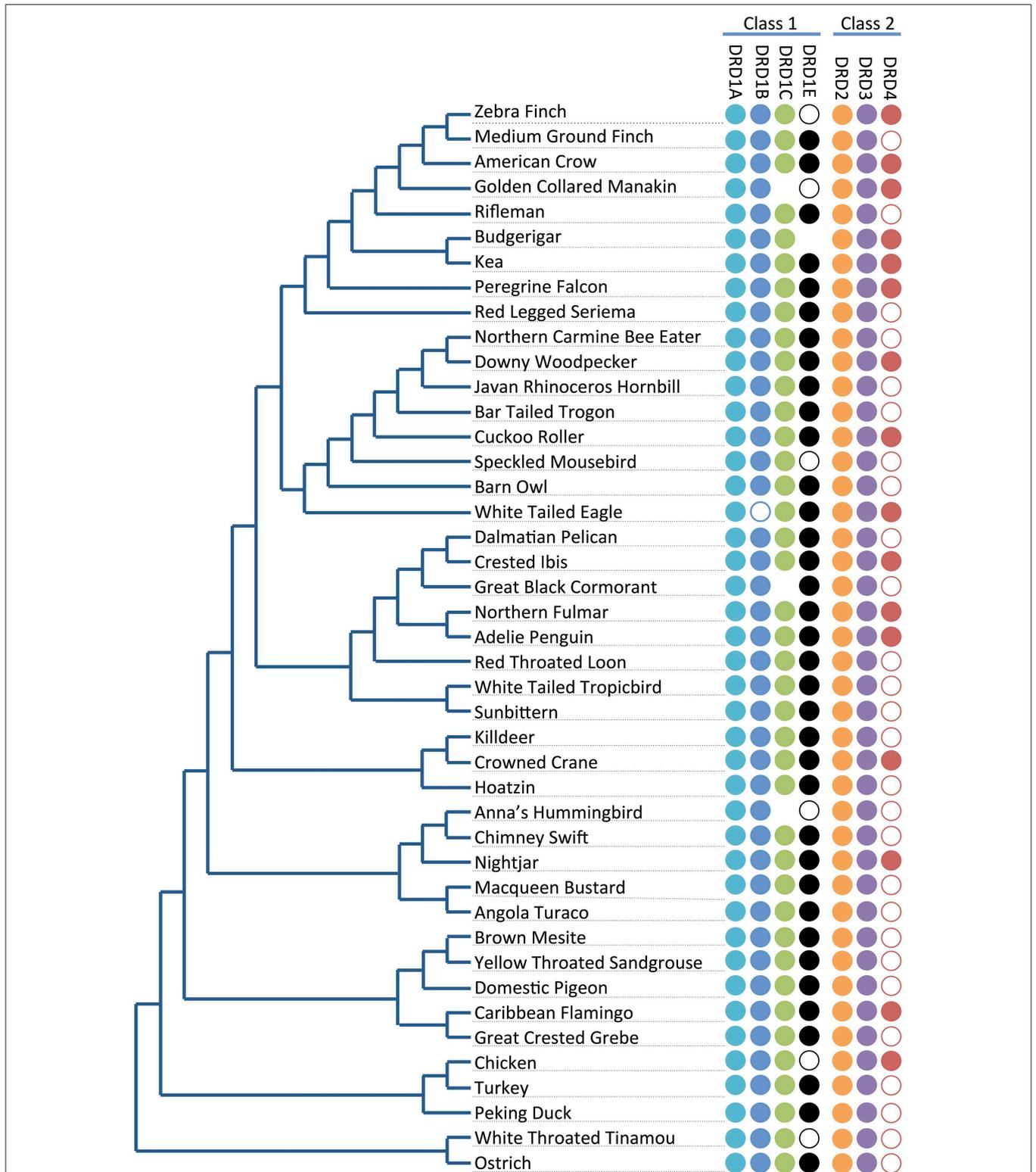
(32%) of high coverage/N50 > 1 Mb genomes had unassembled reads. Similar to DRD1C and DRD1E in the D1 receptors, DRD4 has the fastest diverging sequence of the D2 receptors, indicative of relaxed selective pressure on the gene. Combined, this evidence suggests the DRD4 receptor has a genomic sequence that is difficult to assemble and structurally annotate, and that it is likely that the gene is being pseudogenized and independently lost in many of the bird species.

### D1 and D2 Vertebrate Dopamine Receptor Families Expanded through Whole Genome Duplication

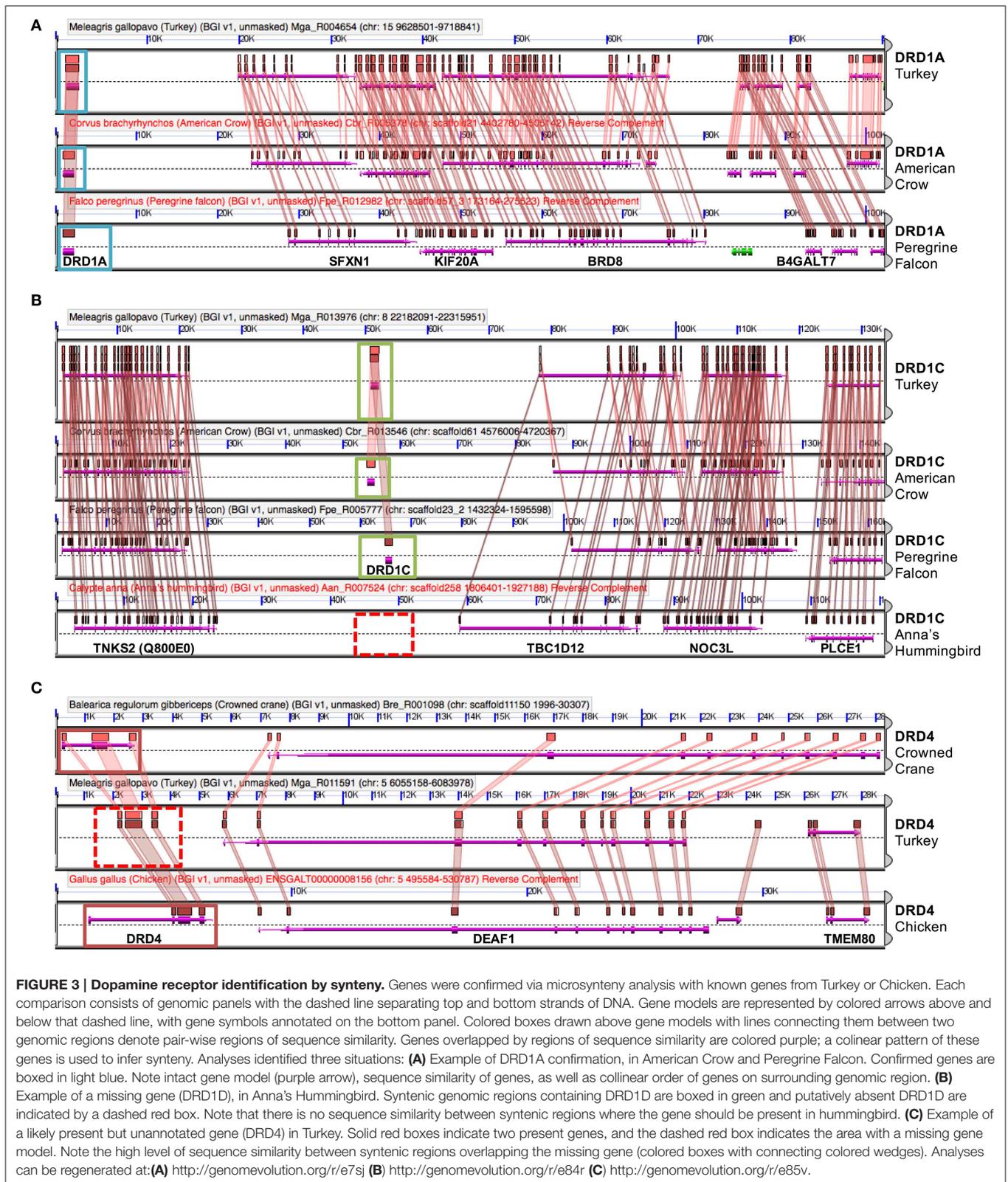
As expected, genomic regions containing each receptor showed strong synteny with regions from other species that contained the orthologous receptor (e.g., Figures 3A, 4A). Interestingly, the genomic regions containing the D1 receptors also showed weak but persistent synteny signals with the genomic regions that contained the other D1 receptors (strongest between DRD1A/DRD1B), both within and across species (Figure 4, Table 2). A similar pattern of intragenomic synteny evidence was observed between chromosome regions containing DRD2 and DRD3 (Figure 4B, Table 2). This synteny signal indicated that these genomic regions, and thus these receptor sets, might have originated from a common ancestor by duplication of large chromosomal regions (e.g., polyploidy). This pattern was not observed between any Class 1 and Class 2 receptors, or between the other D2 receptor pairs (Figure 4B).

To uncover more information on the nature of these possible duplications, we used SynMap (Lyons and Freeling, 2008) to generate a syntenic dot plot of the Chicken and Turkey genomes, with syntenic regions identified where 4 colinear genes fell within a 40 gene window. Synonymous mutation values (Ks) were calculated for each identified syntenic gene pair, and a color scheme was used to illustrate the values on the plot (Figure 5A). Ks values are frequently used as a relative molecular clock, since higher values are indicative of more neutral mutations having occurred and hence more time since divergence from a common ancestor (Kreitman and Akashi, 1995). The majority of syntenic regions occurred within three distinct clusters of synonymous mutation values (Figure 5B). The largest cluster (\*) contained the orthologous syntenic gene pairs that derived from the divergence of two species. These genes share an average  $\log_{10}$  Ks value of  $\sim -0.8$  ( $\sim 0.16$  synonymous mutations per synonymous site). The smallest cluster (\*\*\*) appeared to have also originated contemporaneously. The high substitution values ( $\sim 10$  synonymous mutations/synonymous site) of this second cluster, along with the genome-wide scattering of the orthologous regions (Figure 5A) follow a pattern consistent with an ancient whole genome duplication (WGD) event (Blanc et al., 2003; Blanc and Wolfe, 2004). The remaining cluster (\*\*) had a  $\log_{10}$  Ks value =  $\sim 1.9$  ( $\sim 80$  synonymous mutations/synonymous site), which is consistent with algorithmic noise (Tang and Lyons, 2012).

Of the D1 receptors, gene pairs DRD1A/DRD1B were located in regions identified by SynMap with evidence of originating through a WGD, DRD1A/DRD1C, and DRD1B/DRD1C were located near, but not within such regions, and WGD evidence was not present in regions containing DRD1E (Figure 5, Table 2).

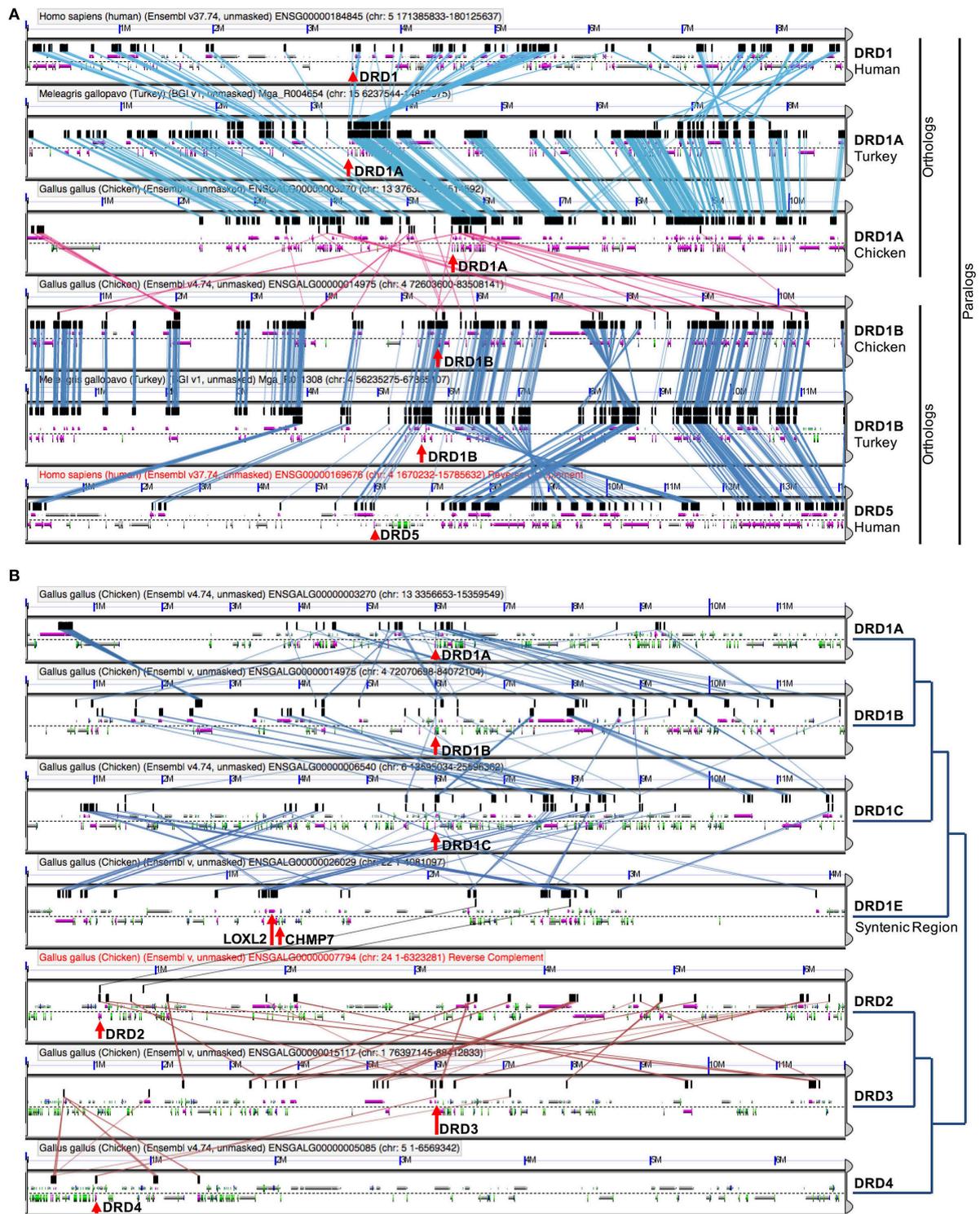


**FIGURE 2 | Receptor presence/absence by species.** Presence and absence of dopamine receptors (right) across the phylogeny of the 43 investigated bird species (left; tree adapted from Jarvis et al., 2014). Full circles indicate receptor genes are present and identified. Empty circles indicate receptor genes not annotated in the genome, but with at least significant portions of the sequences identified by microsynteny analysis (e.g., **Figure 3B**) or BLAST searches against raw sequencing reads (Supplementary File 1). No circles represent receptor genes are likely absent in that species due to not being found through any analyses. Losses do not follow a phylogenetic pattern, and thus appear to be lost independently in different lineages.



This pattern of conservation is consistent with the fractionation of homeologous (partially homologous) gene content that follows polyploidy, as has been shown in plant systems (Thomas et al.,

2006; Schnable et al., 2011; Wang et al., 2011; Cheng et al., 2012; Sankoff and Zheng, 2012). For D2 receptors, no gene pairs were located in regions identified by SynMap with evidence to



**FIGURE 4 | DA receptor microsynteny analysis. (A)** Visualization of shared microsynteny between human, chicken, and turkey chromosome regions containing DRD1A (human DRD1) and DRD1B (human DRD5) receptors. Strong synteny can be seen between Turkey/Chicken DRD1A receptors and Turkey/Chicken DRD1B receptors. Human chromosome regions containing DRD1 and DRD5 receptors show clear synteny with respective orthologs containing chicken and turkey chromosome regions. Chicken DRD1A and DRD1B regions show synteny with one another, although the signal is weaker than between orthologous receptors in different species. **(B)** Microsynteny visualizations (left) and phylogenetic relationships (right) of all DA receptors in Chicken. Microsynteny can be observed between D1 receptors, with the strongest between DRD1A and DRD1B. No pattern of microsynteny is seen between the two classes of receptors, with weak synteny seen between DRD2 and DRD3. Details on how to read these microsynteny analyses are described in **Figure 3**. Results can be regenerated at **(A)** <https://genomevolution.org/r/ho6d> and **(B)** <http://genomevolution.org/r/hjp0>.

TABLE 2 | Duplication evidence summary.

	Microsynteny		SynFind		SynMap	Human Paralogy
	2 MB	12 MB	Colinear	Density		
<b>DRD1A/DRD1B</b>	Y	Y	Y	Y	Y	Y
<b>DRD1A/DRD1C</b>	Y	Y	Y	Y	Y	N
<b>DRD1A/DRD1E</b>	Y	Y	Y	Y	N	N
<b>DRD1B/DRD1C</b>	Y	Y	N	Y	Y	N
<b>DRD1B/DRD1E</b>	Y	Y	N	Y	N	N
<b>DRD1C/DRD1E</b>	Y	Y	N	N	N	N
<b>DRD2/DRD3</b>	N	Y	N	Y	N	Y
<b>DRD2/DRD4</b>	N	N	N	N	N	N
<b>DRD3/DRD4</b>	N	N	N	N	N	N

Evidence supporting a WGD origin of each DA receptor pair. For the D1 receptors, evidence follows a pattern consistent with fractionation loss across all D1 receptors, but is only conclusive for DRD1A/DRD1B. For D2 receptors, evidence suggests DRD2/DRD3 arose from a WGD, but the evidence is weaker than for DRD1A/DRD1B. In no cases was evidence for duplication between classes D1 and D2 receptors found (not shown). Microsynteny analyses were performed with GGeo for two genomic windows: 2 MB and 12 MB. SynFind analyses were performed using both collinear and density based scoring algorithms, with a gene window size of 40 and minimum number of four syntenic genes. SynMap analyses aimed to identify whether gene pairs were located in chromosome regions with evidence of a WGD. GGeo, SynFind, and SynMap are all CoGe tools, are available online at: <http://www.genomeevolution.org> (Lyons and Freeling, 2008). Human paralogy analyses (e.g., **Figure 7**) are based on McLysaght et al. (2002) and aimed to identify whether the human orthologs of gene pairs fell in regions shown as statistically likely as having duplicated from a WGD event. In all analyses, Y indicates evidence was found, and N indicates a lack of evidence.

support a WGD. However, when a “density” based algorithm (4 syntenic genes within a 40 gene window, order disregarded) was used instead of the colinear algorithm to determine syntenic gene regions, genomic regions containing DRD2 and DRD3 showed a signal of synteny (**Table 2**). This suggest that while the colinear arrangement of syntenic genes was lost around DRD2 and DRD3, the overall regions are indeed syntenic (Wolfe, 2001). The DRD4 containing region did not show evidence of synteny with the other D2 receptor regions and no pair of D1/D2 genes showed synteny evidence, under either colinear or density based algorithms.

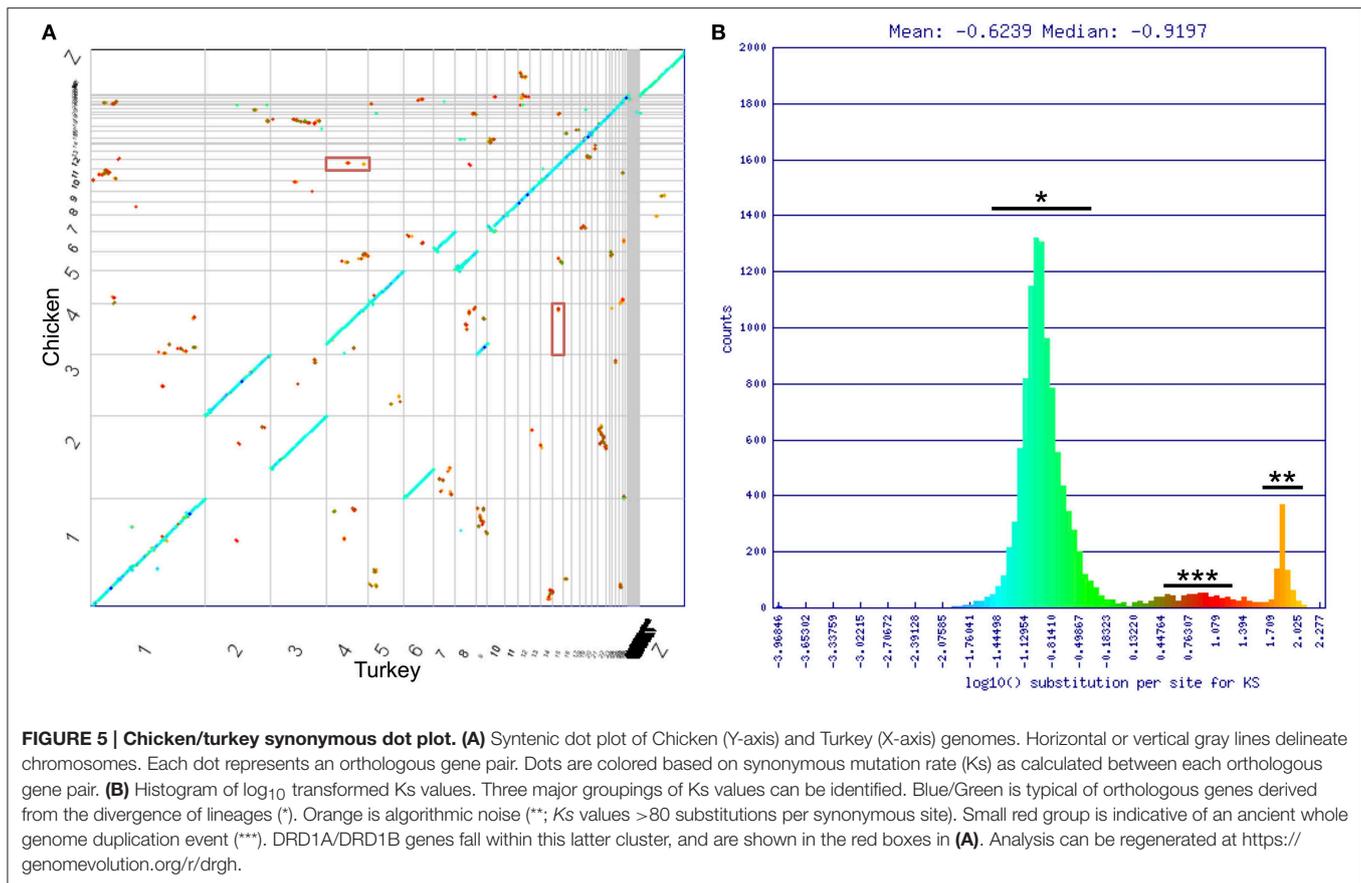
Additional evidence for the WGD origin of the DRD1A/DRD1B and DRD2/DRD3 receptors was sought through comparative analysis with the human genome, where we investigated whether the human orthologs of the avian genes were located in chromosome regions previously identified as likely derived from an ancient vertebrate WGD (McLysaght et al., 2002). To confirm previous assertions that human DRD1 (hDRD1) and hDRD5 are direct orthologs to the avian DRD1A and DRD1B receptors, respectively, a microsynteny analysis of genomic regions containing these genes between Human, Turkey, and Chicken was performed (**Figure 4A**). In this analysis, it was clear that hDRD1 and avian DRD1A are orthologous, and hDRD5 and avian DRD1B are also orthologous. Similar analyses revealed that the three D2 receptors in Chicken and Human were all orthologous between species (not shown). hDRD1, hDRD5, hDRD2, hDRD3, and hDRD4 were all mapped onto human chromosome plots marked with regions

TABLE 3 | Reference receptor accessions.

Species	Gene name	Accession ID
Chicken	DRD1A	ENSEMBL: ENSGALG00000003270
Chicken	DRD1B	ENSEMBL: ENSGALG00000014975
Chicken	DRD2	ENSEMBL: ENSGALG00000007794
Chicken	DRD3	ENSEMBL: ENSGALG00000015117
Chicken	DRD4	ENSEMBL: ENSGALG00000005085
Turkey	DRD1A	ENSEMBL: ENSMGAG00000015442
Turkey	DRD1B	ENSEMBL: ENSMGAG00000012700
Turkey	DRD2	ENSEMBL: ENSMGAG00000004223
Turkey	DRD3	ENSEMBL: ENSMGAG00000014321
Turkey	DRD4	ENSEMBL: ENSMGAG00000003742
Zebra Finch	DRD1A	ENSEMBL: ENSTGUG00000000340
Zebra Finch	DRD1B	ENSEMBL: ENSTGUG00000009908
Zebra Finch	DRD2	ENSEMBL: ENSTGUG00000000255
Zebra Finch	DRD3	ENSEMBL: ENSTGUG00000013405
Zebra Finch	DRD4	ENSEMBL: ENSTGUG00000007173
Human	DRD1	ENSEMBL: ENSG00000184845
Human	DRD2	ENSEMBL: ENSG00000149295
Human	DRD3	ENSEMBL: ENSG00000151577
Human	DRD4	ENSEMBL: ENSG00000069696
Human	DRD5	ENSEMBL: ENSG00000169676
Mouse	DRD1A	ENSEMBL: ENSMUSG000000021478
Mouse	DRD1B	ENSEMBL: ENSMUSG000000039358
Mouse	DRD2	ENSEMBL: ENSMUSG000000032259
Mouse	DRD3	ENSEMBL: ENSMUSG000000022705
Mouse	DRD4	ENSEMBL: ENSMUSG000000025496
Fugu	DRD2	ENSEMBL: ENSTRUG00000014690
Fugu	DRD3	ENSEMBL: ENSTRUG00000000584
Fugu	DRD4	ENSEMBL: ENSTRUG000000012466
Lamprey	DRD1A/B	ENSEMBL: ENSPMAG00000010421
Ciona	D1-Like	GenScript: XM_004226475
Drosophila	Dop1 (D1-like)	Flybase:FBgn0011582

Accession numbers for receptors sequences obtained from external resources. Coding sequence for each receptor was obtained from the listed databases.

demonstrated by McLysaght et al. (2002) as statistically likely to have originated from an early vertebrate WGD (**Figure 6**). hDRD1/hDRD5, as well as hDRD2/hDRD3, fell within these regions. Since the human orthologs of avian DRD1C and DRD1E have been lost in the mammalian lineage and could not be directly compared, we aligned human chromosome regions syntenic to the avian regions containing DRD1C and DRD1E, identified where the genes were most likely located prior to loss, and plotted these locations onto human chromosome maps. Gene regions containing DRD1, DRD5, and those syntenic to DRD1C and DRD1E fell in four regions believed to have been a linkage group quadruplicated in stem chordates, (Pébusque et al., 1998; Putnam et al., 2008; Yamamoto et al., 2013), but did not fall directly within regions shown by McLysaght et al. (2002) to be statistically likely as having originated from a WGD (**Table 2**). No pair of human D1/D2 receptors fell within these regions, and hDRD4 did not fall in a region with the other human D2 receptors.



## Evolutionary Analysis of DRD1A/DRD1B Duplication in the Vertebrate Lineage

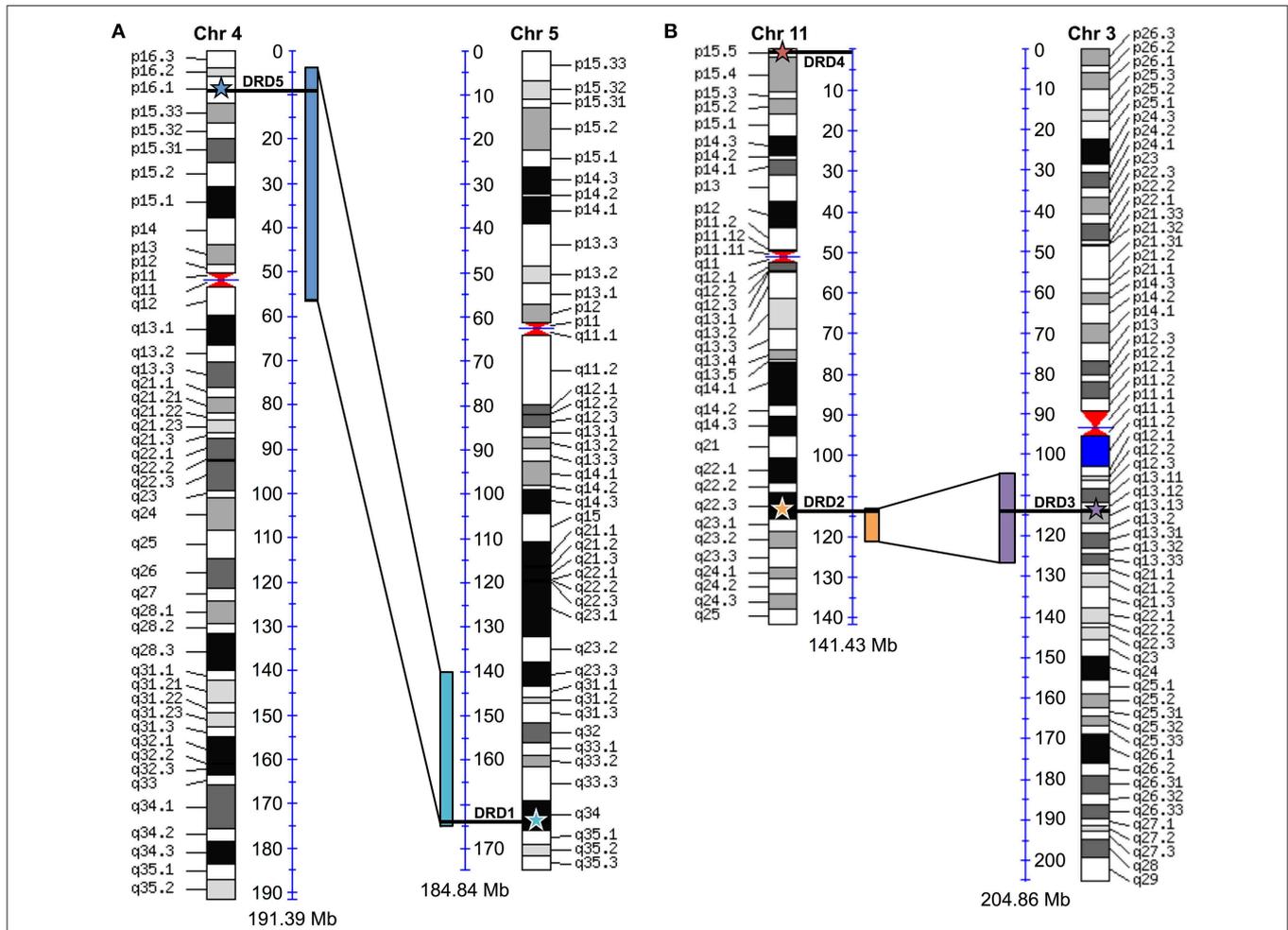
To determine if the WGD that gave rise to the DA receptor expansion was specific to the vertebrate lineage, we constructed a phylogeny of DA receptors across a range of vertebrate and invertebrate species rooted on an arthropod, drosophila (Figure 7A). As expected, each family of receptors formed distinct clades, each monophyletic for that family. Using our methods for gene identification, duplicates of each D1 receptors DRD1A, DRD1B, and DRD1C were identified in fugu (Figure 7A, Supplemental File 3). These duplications are likely explained by an additional WGD that occurred in the fugu lineage (Panopoulou and Poustka, 2005). With regards to DRD1A/DRD1B duplication, the phylogeny showed a clear transition in the vertebrate lineage, with a WGD event near the origin of vertebrates approximately 450 MYA (Figure 7A) (Dehal and Boore, 2005).

## Discussion

Understanding DA and its associated receptors is critical for both health and agricultural sciences due to the significant roles they play in behavior, learning, and general neuromodulation. In this study we aimed to expand our understanding of both the recent evolution of DA receptors across the avian lineage as well as the ancient evolution of DA receptors throughout the

larger vertebrate lineage. Using a combination of phylogenetic and sequence analyses, we identified and classified the DA receptors in 43 recently sequenced bird species spanning the diversity of extant avian species. Our analyses show that the ancestral bird genome contained at least 7 DA receptors: four D1 receptors (DRD1A, DRD1B, DRD1C, and DRD1E) and three D2 receptors (DRD2, DRD3, and DRD4). Two of the D1 receptors (DRD1A, DRD1B) and two of the D2 receptors (DRD2 and DRD3) were maintained across all extant species. The other two D1 receptors, DRD1C and DRD1E, have been lost in some lineages independently, although they are still widely conserved across the lineage. The remaining D2 receptor DRD4 has been either pseudogenized and thus functionally lost in many species, or has increased sequence complexity making it difficult to computationally identify, assemble, and annotate.

When receptor losses are compared to the species phylogeny, no apparent species-level pattern of gene loss can be identified. With regards to the Class 1 receptors, these findings strongly support earlier claims that DRD1C and DRD1E receptors can be more easily lost in multiple vertebrate lineages independently (Yamamoto et al., 2013). One interesting observation is that these two receptors that have undergone loss are diverging 2–4 times faster than the retained receptors DRD1A and DRD1B. Similarly, the fastest diverging D2 receptor, DRD4, appears to have been lost independently from species spanning the avian species tree. This pattern of loss of the more-rapidly

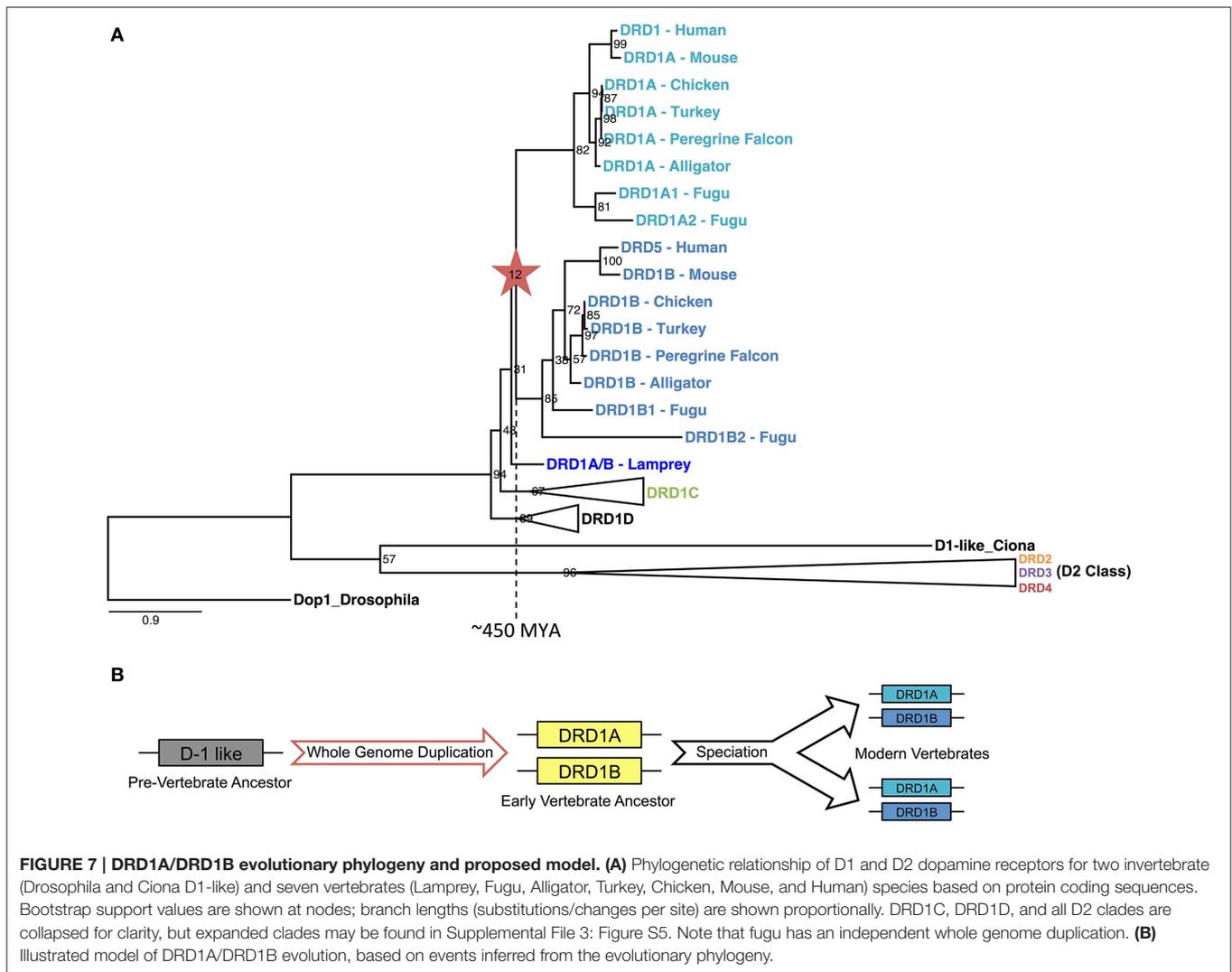


**FIGURE 6 | Human paralagon maps.** Human chromosome maps showing regions previously classified as likely having originated from a whole genome duplication event in vertebrate history. **(A)** Human DRD1 (light blue star) and DRD5 (dark blue star) are marked on their respective chromosomes, and can be seen to fall within these regions. **(B)** Human DRD2 (orange star), DRD3 (purple star), and DRD4 (red star) mapped on their respective chromosomes. DRD2/DRD3 fall within these regions. Maps were generated using “Paralogs in the Human Genome” tool from the Wolfe lab, and are based on a minimum threshold for paired proteins per block greater than 12 **(A)** or 7 **(B)**. Paired proteins per block greater than 6 are statistically probable as having occurred from a WGD event rather than by chance (McLysaght et al., 2002). Chromosome maps can be regenerated and viewed at [http://wolfe.ucd.ie/cgi-bin/dup\\_528/chrom\\_plot](http://wolfe.ucd.ie/cgi-bin/dup_528/chrom_plot).

diverging receptors may suggest relaxed functional constraints due to overlapping function, allowing some receptors to undergo modification (e.g., neofunctionalization) and/or loss without negative impacts in those species. The possible pseudogenization of DRD4 in many bird species is intriguing, considering that variations in the coding sequence of the gene are correlated with creativity and other personality differences within a species, including in birds and humans (Fidler et al., 2007; Ben-Israel et al., 2015). If future studies confirm that DRD4 is indeed being pseudogenized and lost in many lineages, it would be interesting to perform additional experiments in species that do not have a DRD4 receptor to determine possible effects on personality traits, and whether other dopamine receptors functionally replace it. Detailed expression studies of the location and timing of these genes should shed further light into this possibility.

In terms of evolutionary history of DA receptors, while there has been long-standing speculation that D1 receptors DRD1A and DRD1B could have originated from a WGD event, conclusive evidence has not been provided (Holland, 1999; Le Crom et al., 2003). Through microsynteny, genome-wide synteny, synonymous substitution metrics, phylogenetic, and comparative genomic analyses of many genomes, we provide structural and temporal evidence that DRD1A and DRD1B, as well as DRD2 and DRD3 receptors originated through WGD.

We hypothesize that the weaker colinear signal of homologous genes from the genomic regions containing DRD2 and DRD3 could be due to the diploidization process following polyploidy, where many duplicated homeologous genes are lost through a process known as fractionation, relatively quickly reducing the gene content of a genome to a state more similar to the pre-duplicated ancestral genome (Woodhouse et al., 2010).



**FIGURE 7 | DRD1A/DRD1B evolutionary phylogeny and proposed model. (A)** Phylogenetic relationship of D1 and D2 dopamine receptors for two invertebrate (*Drosophila* and *Ciona* D1-like) and seven vertebrates (*Lamprey*, *Fugu*, *Alligator*, *Turkey*, *Chicken*, *Mouse*, and *Human*) species based on protein coding sequences. Bootstrap support values are shown at nodes; branch lengths (substitutions/changes per site) are shown proportionally. *DRD1C*, *DRD1D*, and all D2 clades are collapsed for clarity, but expanded clades may be found in Supplemental File 3: Figure S5. Note that *fugu* has an independent whole genome duplication. **(B)** Illustrated model of *DRD1A/DRD1B* evolution, based on events inferred from the evolutionary phylogeny.

In addition to fractionation, larger structural rearrangements, including chromosome fissions, fusions, translocations, and local inversions may also occur causing further obfuscation of the colinear signal of retained duplicated genes, which has been seen in vertebrate lineages (Smith et al., 2002). This difference in synteny signal strength is interestingly correlated with a single exon D1 receptor (i.e., low recombination) and multi-exon D2 receptors. However, despite the surrounding gene losses and local rearrangements in D2 receptor chromosome regions, we find that many duplicated genes around the receptors are retained, and a signal of synteny can still be identified.

Yamamoto et al. (2013) proposed an evolutionary model for the origin of the *DRD1* receptors where an ancient duplication gave rise to the ancestor of *DRD1A/B* and *DRD1C/E*. While our data clarifies the mechanism and origin of *DRD1A* and *DRD1B*, it is not conclusive on the origin of *DRD1C* and *DRD1E* sharing an ancestor to the exclusion of *DRD1A/B*. Similar to the results they present, our phylogeny of D1 receptors (Figure 1) shows a polytomy for *DRD1E* even though there is syntenic evidence for their orthologous relationship across the sampled avian taxa,

which is likely a result of this family's rapid divergence rate. In addition, syntenic analysis of *DRD1C* and *DRD1E* shows weak support for being derived from WGD, but not to a great extent when compared pairwise with *DRD1A* and *DRD1B* (Table 2). Nevertheless, we believe the data and methods we present, when combined with additional data and comparative analyses, will aid future studies to resolve the exact relationships and origins of *DRD1C* and *DRD1E* receptors.

In summary, our evidence suggests a large scale, retained expansion of the DA receptor family through an early vertebrate WGD event, followed by subsequent losses of the more rapidly diverging receptors in some species lineages independently, suggesting relaxed functional constraints on some receptors in the DA receptor superfamily. We propose that the early vertebrate ancestor possessed only one D1 receptor. An ancient WGD, likely one of the vertebrate-lineage-specific 1R or 2R events (Dehal and Boore, 2005; Panopoulou and Poustka, 2005), gave rise to the *DRD1A* and *DRD1B* receptors approximately 450 MYA (Figure 7B). An additional round of WGD or translocation duplication events may have given rise to the additional D1

receptors DRD1C/D and DRD1E, but fractionation complexities of layered polyploidy events mask conclusive evidence (Le Crom et al., 2003; Thomas et al., 2006; Sankoff and Zheng, 2012). In regards to the D2 receptors, our evidence suggests that DRD2 and DRD3 also originated from an early WGD, but DRD4 originated from a smaller scale duplication event. Our evidence provides the first conclusive support for a long-standing hypothesis of dopamine receptor evolution, and also suggests that the expansion of dopamine receptors through WGD was greater than previously expected. These findings shed light into the evolutionary relationships of a gene family important in brain function and will hopefully help unravel complex functional relationships between similar but distinct neuroregulatory receptors.

## Acknowledgments

We are thankful for the insightful suggestions made by one of the reviewers that substantially strengthened this

work. AH is funded by the UA Arizona Biological and Biomedical Sciences Program. The 43 bird genomes used in this study were kindly made available pre-release by the Avian Phylogenomics Consortium (<http://avian.genomics.cn/en/>); we thank co-coordinators of that project, Guojie Zhang of BGI and Tom Gilbert of the University of Copenhagen, along with co-author EJ, for early access to these data. CoGe is supported by the Gordon and Betty Moore Foundation (#3383) and the USDA (NIFA: GRANT11301623), and is Powered by iPlant (NSF DBI – 1265383). Phylogenetic analyses were done using iPlant resources (NSF DBI – 1265383). EHL is supported by NSF (IOS – 1339156). EDJ is supported by funding from HHMI.

## Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnins.2015.00361>

## References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Beaulieu, J.-M., and Gainetdinov, R. R. (2011). The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacol. Rev.* 63, 182–217. doi: 10.1124/pr.110.002642
- Ben-Israel, S., Uzefovsky, F., Ebstein, R. P., and Knafo-Noam, A. (2015). Dopamine D4 receptor polymorphism and sex interact to predict children's affective knowledge. *Front. Psychol.* 6:846. doi: 10.3389/fpsyg.2015.00846
- Blanc, G., Hokamp, K., and Wolfe, K. H. (2003). A recent polyploidy superimposed on older large-scale duplications in the arabidopsis genome. *Genome Res.* 13, 137–144. doi: 10.1101/gr.751803
- Blanc, G., and Wolfe, K. H. (2004). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16, 1667–1678. doi: 10.1105/tpc.021345
- Callier, S., Snappyan, M., Le Crom, S., Prou, D., Vincent, J.-D., and Vernier, P. (2003). Evolution and cell biology of dopamine receptors in vertebrates. *Biol. Cell Auspices Eur. Cell Biol. Organ.* 95, 489–502. doi: 10.1016/S0248-4900(03)00089-3
- Chaiseha, Y., Youngren, O., Al-Zailaie, K., and El Halawani, M. (2003). Expression of D1 and D2 dopamine receptors in the hypothalamus and pituitary during the turkey reproductive cycle: colocalization with vasoactive intestinal peptide. *Neuroendocrinology* 77, 105–118. doi: 10.1159/000068649
- Cheng, F., Wu, J., Fang, L., Sun, S., Liu, B., Lin, K., et al. (2012). Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLoS ONE* 7:e36442. doi: 10.1371/journal.pone.0036442
- Dalloul, R. A., Long, J. A., Zimin, A. V., Aslam, L., Beal, K., Ann Blomberg, L., et al. (2010). Multi-platform next-generation sequencing of the domestic Turkey (*Meleagris gallopavo*): genome assembly and analysis. *PLoS Biol.* 8:e1000475. doi: 10.1371/journal.pbio.1000475
- Dehal, P., and Boore, J. L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3:e314. doi: 10.1371/journal.pbio.0030314
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Fidler, A. E., van Oers, K., Drent, P. J., Kuhn, S., Mueller, J. C., and Kempenaers, B. (2007). Drd4 gene polymorphisms are associated with personality variation in a passerine bird. *Proc. R. Soc. Lond. B. Biol. Sci.* 274, 1685–1691. doi: 10.1098/rspb.2007.0337
- Flicek, P., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., et al. (2014). Ensembl 2014. *Nucleic Acids Res.* 42, D749–D755. doi: 10.1093/nar/gkt1196
- Ganapathy, G., Howard, J. T., Ward, J. M., Li, J., Li, B., Li, Y., et al. (2014). High-coverage sequencing and annotated assemblies of the budgerigar genome. *GigaScience* 3:11. doi: 10.1186/2047-217X-3-11
- Goff, S. A., Vaughn, M., McKay, S., Lyons, E., Stapleton, A. E., Gessler, D., et al. (2011). The iPlant collaborative: cyberinfrastructure for plant biology. *Plant Genet. Genomics* 2:34. doi: 10.3389/fpls.2011.00034
- Hillier, L. W., Miller, W., Birney, E., Warren, W., Hardison, R. C., Ponting, C. P., et al. (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695–716. doi: 10.1038/nature03154
- Holland, P. W. H. (1999). Gene duplication: past, present and future. *Semin. Cell Dev. Biol.* 10, 541–547. doi: 10.1006/scdb.1999.0335
- Huang, Y., Li, Y., Burt, D. W., Chen, H., Zhang, Y., Qian, W., et al. (2013). The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat. Genet.* 45, 776–783. doi: 10.1038/ng.2657
- Huson, D. H., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 61, 1061–1067. doi: 10.1093/sysbio/sys062
- Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346, 1320–1331. doi: 10.1126/science.1253451
- Kreitman, M., and Akashi, H. (1995). Molecular evidence for natural selection. *Annu. Rev. Ecol. Syst.* 26, 403–422. doi: 10.1146/annurev.es.26.110195.002155
- Kubikova, L., Wada, K., and Jarvis, E. D. (2010). Dopamine receptors in a songbird brain. *J. Comp. Neurol.* 518, 741–769. doi: 10.1002/cne.22255
- Le Crom, S., Kapsimali, M., Barôme, P.-O., and Vernier, P. (2003). Dopamine receptors for every species: gene duplications and functional diversification in craniates. *J. Struct. Funct. Genomics* 3, 161–176. doi: 10.1023/A:1022686622752
- Lyons, E., and Freeling, M. (2008). How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J.* 53, 661–673. doi: 10.1111/j.1365-313X.2007.03326.x
- McLysaght, A., Hokamp, K., and Wolfe, K. H. (2002). Extensive genomic duplication during early chordate evolution. *Nat. Genet.* 31, 200–204. doi: 10.1038/ng884
- Panopoulou, G., and Poustka, A. J. (2005). Timing and mechanism of ancient vertebrate genome duplications – the adventure of a hypothesis. *Trends Genet.* 21, 559–567. doi: 10.1016/j.tig.2005.08.004
- Pébusque, M. J., Coulier, F., Birnbaum, D., and Pontarotti, P. (1998). Ancient large-scale genome duplications: phylogenetic and linkage analyses shed

- light on chordate genome evolution. *Mol. Biol. Evol.* 15, 1145–1159. doi: 10.1093/oxfordjournals.molbev.a026022
- Pierre, S. E. S., Ponting, L., Stefancsik, R., and McQuilton, P. (2014). FlyBase 102—advanced approaches to interrogating FlyBase. *Nucleic Acids Res.* 42, D780–D788. doi: 10.1093/nar/gkt1092
- Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064–1071. doi: 10.1038/nature06967
- Robertson, B., Huerta-Ocampo, I., Ericsson, J., Stephenson-Jones, M., Pérez-Fernández, J., Bolam, J. P., et al. (2012). The dopamine D2 receptor gene in lamprey, its expression in the striatum and cellular effects of D2 receptor activation. *PLoS ONE* 7:e35642. doi: 10.1371/journal.pone.0035642
- Sankoff, D., and Zheng, C. (2012). Fractionation, rearrangement and subgenome dominance. *Bioinformatics* 28, i402–i408. doi: 10.1093/bioinformatics/bts392
- Sartsoongnoen, N., Kosonsiriluk, S., Prakobsaeng, N., Songserm, T., Rozenboim, I., Halawani, M. E., et al. (2008). The dopaminergic system in the brain of the native Thai chicken, *Gallus domesticus*: localization and differential expression across the reproductive cycle. *Gen. Comp. Endocrinol.* 159, 107–115. doi: 10.1016/j.ygcen.2008.08.002
- Sasaki, A., Sotnikova, T. D., Gainetdinov, R. R., and Jarvis, E. D. (2006). Social context-dependent singing-regulated dopamine. *J. Neurosci.* 26, 9010–9014. doi: 10.1523/JNEUROSCI.1335-06.2006
- Schnable, J. C., and Lyons, E. (2012). *Comparative Genomics with Maize and other Grasses: From Genes to Genomes! Maydica*, Vol. 56. Available online at: <http://cra-journals.cineca.it/index.php/maydica/article/view/690> (Accessed December 19, 2014).
- Schnable, J. C., Springer, N. M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4069–4074. doi: 10.1073/pnas.1101368108
- Schnell, S. A., You, S., and El Halawani, M. E. (1999). D1 and D2 dopamine receptor messenger ribonucleic acid in brain and pituitary during the reproductive cycle of the turkey hen. *Biol. Reprod.* 60, 1378–1383. doi: 10.1095/biolreprod60.6.1378
- Shapiro, M. D., Kronenberg, Z., Li, C., Domyan, E. T., Pan, H., Campbell, M., et al. (2013). Genomic diversity and evolution of the head crest in the rock pigeon. *Science* 339, 1063–1067. doi: 10.1126/science.1230422
- Smith, S. F., Snell, P., Gruetzner, F., Bench, A. J., Haaf, T., Metcalfe, J. A., et al. (2002). Analyses of the extent of shared synteny and conserved gene orders between the genome of fugu rubripes and human 20q. *Genome Res.* 12, 776–784. doi: 10.1101/gr.221802
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stoof, J. C., and Keabian, J. W. (1984). Two dopamine receptors: biochemistry, physiology and pharmacology. *Life Sci.* 35, 2281–2296. doi: 10.1016/0024-3205(84)90519-8
- Tang, H., and Lyons, E. (2012). Unleashing the genome of *Brassica Rapa*. *Front. Plant Sci.* 3:172. doi: 10.3389/fpls.2012.00172
- Thomas, B. C., Pedersen, B., and Freeling, M. (2006). Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Genome Res.* 16, 934–946. doi: 10.1101/gr.4708406
- Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919
- Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Küstner, A., et al. (2010). The genome of a songbird. *Nature* 464, 757–762. doi: 10.1038/nature08819
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
- Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nat. Rev. Genet.* 2, 333–341. doi: 10.1038/35072009
- Woodhouse, M. R., Schnable, J. C., Pedersen, B. S., Lyons, E., Lisch, D., Subramaniam, S., et al. (2010). Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homeologs. *PLoS Biol.* 8:e1000409. doi: 10.1371/journal.pbio.1000409
- Yamamoto, K., Mirabeau, O., Bureau, C., Blin, M., Michon-Coudouel, S., Demarque, M., et al. (2013). Evolution of dopamine receptor genes of the D1 class in vertebrates. *Mol. Biol. Evol.* 30, 833–843. doi: 10.1093/molbev/mss268
- Yamamoto, K., and Vernier, P. (2011). The evolution of dopamine systems in chordates. *Front. Neuroanat.* 5:21. doi: 10.3389/fnana.2011.00021
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Zhan, X., Pan, S., Wang, J., Dixon, A., He, J., Muller, M. G., et al. (2013). Peregrine and saker falcon genome sequences provide insights into evolution of a predatory lifestyle. *Nat. Genet.* 45, 563–566. doi: 10.1038/ng.2588
- Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., et al. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346, 1311–1320. doi: 10.1126/science.1251385

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Haug-Baltzell, Jarvis, McCarthy and Lyons. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.