ARTICLE

Amino acid patterns in independent lineages of vocal learning and other birds give insights into convergent evolution

Chul Lee[1, 2, +], Seoae Cho[3], Kyu-Won Kim[4], Dongahn Yoo[1], Matthew Davenport[2], Jae Yong Han[5], Hong Jo Lee[5,6], Gregory Gedman[2], Jean-Nicolas Audet[2], Erina Hara[7], Miriam Rivas[8], Osceola Whitney[9], Andreas R. Pfenning[10,11], Heebal Kim[1,2,4]*, and Erich D. Jarvis[2,12]*

Affiliation

[1]Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Republic of Korea. [2]Laboratory of Neurogenetics of Language, The Rockefeller University, New York City, NY, 10065, USA, [3]eGenome, inc., Seoul, Republic of Korea. [4]Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea. [5]Department of Agricultural Biotechnology and Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea. USA. [6]Division of Animal Sciences, University of Missouri, Columbia, MO, USA. [7]Institute for Genomic Medicine, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA. [8]Integrated Laboratory Systems, Inc., Research Triangle Park, NC 27706, USA. [9]Department of Biology, Division of Science, The City College of New York, 160 Convent Avenue, New York, NY 10031, USA. [10]Computational Biology Department, the School of Computer Science, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA. [11]Neuroscience Institute, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA. [12]Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, MD 20815, USA.

Corresponding author: HBK: heebal@snu.ac.kr, EDJ: ejarvis@rockefeller.edu

Abstract

Vocal learning, the ability to imitate sounds and a component of spoken language, is a complex convergent trait observed in only a few independent lineages of mammals and birds. Clear gene expression convergences have been found in vocal learning brain regions among several vocal learners, but amino acid convergences remain an open question. Here, we investigated whether avian vocal learning clades have amino acid convergences that could be related to their specialized trait. We developed a tool, Convergent Sequence Variant finder (CSV finder), applied to an alignment of 48 species representing nearly all bird orders, and identified convergent single amino acid variants among vocal learners and among most polyphyletic species combinations. We discovered that the numbers of convergent variants were associated with the product of branch lengths of the most recent common ancestors of each species combination. The number of convergent variants in vocal learning clades did not exceed that of control species combinations. However, a subset of genes with the vocal learner-specific variants was uniquely enriched in the 'learning' process, under positive selection, and supported by meta-analyses for *FOXP2* targets, singing-induced regulation, and differential expression in song learning nuclei. Moreover, we confirmed convergent patterns were still enriched in 363 species densely sampled across the avian tree. We propose a hypothesis of a steady state background of amino acid and nucleotide convergence upon which selection acts for convergent traits, with the deeper in time their common ancestor the higher proportions of convergent genetic changes.

Keywords: Convergent evolution, vocal learning, convergent sequence variant finder (CSV finder), convergent single amino acid variant (ConSAV), product of original branch lengths (POB), *B3GNT2*.

Significance

Convergent evolution is a common principle of biological system, with vocal learning being one of the examples of an advanced behavior that gives rise to speech in humans and learned song in birds. Here, we developed new tools to detect amino acid convergences amongst different combinations of species, found that most combinations were correlated with the product of ancestral branch lengths, and that for avian vocal learners these convergences were in genes associated with learning. Based on our findings, we suggest that there is steady background rate of convergent amino acid changes among species that is influenced by their branch lengths, for which selection acts upon for convergent traits.

Introduction

Single amino acid variants (SAV) are one of the potential drivers of evolution for various traits. For example, the Forkhead box P2 (*FOXP2*) transcription factor has two well-known human-specific SAVs which might have been positively selected for learning behavior related to spoken language (Enard et al. 2002; Scharff and Petri 2011). Mutant mice humanized for the two SAVs of *FOXP2* showed more advanced learning abilities (Schreiweis et al. 2014) and alterations of cortico-basal ganglia circuits (Enard et al. 2009; Enard 2011; Reimers-Kipping et al. 2011), which play critical roles in spoken-language (Jarvis 2019); mice containing a heterozygous missense mutation that causes speech syllable apraxia in humans also showed syllable sequencing deficits (Castellucci et al. 2016; Chabout et al. 2016).

A crucial component of spoken-language is vocal learning, the ability to produce vocalizations through imitation, and is a convergent trait observed in only a few animals, including songbirds, parrots, and hummingbirds among birds, and bats, dolphins/whales, seals, elephants, and humans among mammals (Nottebohm 1972; Jarvis 2004; Petkov and Jarvis 2012; Nowicki and Searcy 2014; Jarvis 2019; Cahill et al. 2021). Both vocal learners and vocal non-learners share an auditory pathway that controls auditory

learning, while only the vocal learning birds and humans have been found to share a specialized convergent forebrain pathway that controls vocal learning (Jarvis 2004; Petkov and Jarvis 2012; Nowicki and Searcy 2014; Pfenning et al. 2014). Supporting the hypothesis of independent origins of vocal learning, the first genome-scale phylogenetic tree of birds showed that the three avian vocal-learner lineages are indeed not monophyletic (Jarvis et al. 2014) and the next tree with more densely sampled birds kept showing the independency (Feng et al. 2020). Even though songbirds and parrots are relatively closely related, their two closest relatives, sub-oscines and New Zealand wrens, are considered vocal non-learning lineages.

In the first genome-scale analyses for vocal learning among birds, genes with positively selected changes in zebra finch (a songbird) compared to chicken were identified (Warren et al. 2010). Some of the positively selected genes were ion channels, which are known to control neurological function, behavior and disease (Warren et al. 2010). However, the comparison by necessity at that time was narrow, between only one vocal learner (zebra finch) and one vocal non-learner (chicken), which are also very distant relatives (Jarvis et al. 2014), like a marsupial is to a placental mammal.

The first draft genome sequences of the Avian Phylogenomics Project (Bird 10K project, B10K) consisting of 48 avian species representing nearly all bird orders (Zhang et al. 2014) provided an unprecedented opportunity to investigate genetic features specific to polyphyletic vocal learning clades. These studies found convergent brain gene expression specializations in vocal-learning birds and human (Jarvis 2004; Petkov and Jarvis 2012; Pfenning et al. 2014). In preliminary analyses with these genomes, we also found mutually exclusive amino acid substitutions unique to vocal learners, using a novel method (Target-specific Amino Acid Substitutions [TAAS] analysis) (Zhang et al. 2014). However, the later study overlooked several viewpoints reported around that time for principles of molecular convergence (Castoe et al. 2009; Goldstein et al. 2015); it did not separately test for convergent (identical) versus divergent

(different) amino acid substitutions; it did not compare any control species combinations, with phylogeny similar to vocal learning birds; and it did not test for possible influence of close relatives.

Here, we overcame the above limitations. We developed a new computational method, Convergent Sequence Variant finder (CSV finder), to identify convergent and divergent amino acid substitutions unique to vocal learners or any other polyphyletic species combination, by considering phylogenetic relationships and estimating ancestral sequences. Control sets of different species combinations and their molecular convergences and divergences were compared with that of vocal learning birds. Based on these comparisons, we discovered additional principles of genetic convergence, dependent in the relative molecular clocks of each lineage. Despite vocal learning not having a higher level of convergence than expected, multiple lines of evidence identified new candidate genes with convergence in vocal learners.

Results

Convergent amino acid variants specific to avian vocal learning clades. Among the genomes of 48 avian species (Zhang et al. 2014) spanning most orders (Jarvis et al. 2014), we scanned their alignments of 8,295 orthologous protein coding genes consisting of 4,519,041 homologous amino acid sites for shared amino acid patterns in the 6 vocal learning species from the 3 vocal learning orders or suborders (songbirds: zebra finch, medium ground finch, and American crow; parrots: budgerigar, and kea; and hummingbirds: Anna's hummingbird) that were not found in any of the remaining 41 vocal non-learning birds (**Fig. 1A**). Rifleman, a close New Zealand wren relative of songbirds and sub-oscines, was initially excluded because of the uncertainty of its vocal learning ability, although assumed to be a vocal non-learner (Jarvis et al. 2014).

Based on the previous TAAS algorithm (Zhang et al. 2014), we developed a new tool that scans for coding

variants specific to vocal learning birds mutually exclusive to vocal non-learning birds with amino acid substitutions as well as single point insertions/deletions (**Methods**). Using this approach, we scanned the alignment for single amino acid variants (SAV) shared in vocal non-learners and classified them into the following four types (**Fig. 1B**): Type 1, mutually exclusive identical amino acid substitutions between vocal learners and vocal non-learners; Type 2, identical amino acid substitutions in vocal learners and different substitutions in vocal non-learners; Type 3, the inverse of Type 2, with different substitutions in vocal learners not shared with an identical substitutions in vocal non-learners; and Type 4, mutually exclusive different sets of amino acid substitutions between vocal learners and vocal non-learners. We called the Types 1 and 2 identical substitutions as "convergent SAVs (ConSAVs)" and the Types 3 and 4 as different substitutions as "divergent SAVs (DivSAVs)" in polyphyletic clades. In this manner, the broad term "convergent" traditionally refers to both ConSAVs and DivSAVs, but the specific term here applies just to ConSAVs, and divergent to DivSAVs.

Using our approach, we found 148 sites (0.0033% of total sites) in 135 genes (1.6% of total genes) with SAVs specific to avian vocal learners (AVL-SAVs, **Supplementary table S1**). Out of these 148 AVL-SAVs, 3 and 21 were classified as Type 1 and 2 SAVs (24 total AVL-ConSAVs), and 6 and 118 were classified as Type 3 and 4 SAVs (124 AVL-DivSAVs). It is logical that Types 1 was in the minority and Type 4 in the majority because Types 1 has the most stringent criterion that each group of species must have identical substitutions, whereas Type 4 has the most relaxed criterion allowing neutral substitutions. An example of Type 1 is the 253rd site of *B3GNT2* with an asparagine (N) in all avian vocal learning species sequenced and a histidine (H) in all vocal non-learning species sequenced; an example of Type 4 is the 217th site of *SMRC8* with a glutamine, valine, or leucine (Q, V, or L) in all avian vocal learners, and alanine (A) or isoleucine (I) in all vocal non-learners (**Fig. 1C**).

To confirm evolutionary directions from ancestral states to these AVL-SAVs of extant species, we

performed ancestral sequence reconstruction analysis with RAxML (Stamatakis 2014) using the avian family tree (Jarvis et al. 2014) in **Fig. 1A**, and confirmed that all 148 AVL-SAV sites had inferred evolutionary directions of either convergent identical or divergent different substitutions in vocal learners relative to their most recent common ancestors (MRCA; **Supplementary table S1**). We classified the 24 AVL-ConSAVs as 1 simple parallel path from an identical amino acid at ancestral nodes to another identical amino acid within vocal learners and 23 complex convergent paths with complex accumulative substitutions with different amino acids at ancestral nodes to another identical amino acid within vocal learners. On the other hands, the 124 AVL-DivSAVs had 4 simple and 120 complex divergent evolutionary paths. These findings indicate that with combining the initial scanning with the ancestral reconstruction, we have a robust tool, which we called Convergent Sequence Variant Finder (CSV Finder, **Supplementary fig. S1**), for identifying and classifying various types of amino acid convergences in a group of species, given a tree, and that vocal learning bird species have several or more of all types.

Convergent versus divergent variant relationships across multispecies combinations. We next tested whether avian vocal learners have a higher frequency of convergent substitutions relative to control sets of species. Considering the polyphyletic relationship of the 6 vocal learning species examined, we designed 2 types of clade-specific control sets based on 10,737,573 possible species combinations (**Supplementary fig. S2**): 1) All controls consisting of 8,238 different species combinations given the phylogeny with 6 target species from 3 independent lineages without considering any traits; 2) Of these 8,238 control combinations, 59 core controls consisting of all possible combinations with 6 target species having at least 2 vocal learning clades and 1 non-learning clade. For example, the latter included the control combination of songbirds and parrots as 2 vocal learning clades and swift as a vocal non-learning clade which is a close relative to hummingbirds. We conducted the CSV Finder analysis on each of the

8,238 control species combinations and identified their amino acid convergences and divergences.

Building off of previous studies on convergent evolution in reptile and mammalian lineages (Castoe et al. 2009; Goldstein et al. 2015; Thomas and Hahn 2015) that tested pair-wise combinations of two species, we then compared multi-wise combinations of all 8,238 control species sets and found a strong correlation between the number of convergent SAVs (ConSAVs) and divergent SAVs (DivSAVs) specific to vocal learners and control sets (**Fig. 1D**). There were four outliers (x) with higher number of ConSAVs of control sets (Ctrl-ConSAV sets) than expected given the regression line (adjusted $p > 0.05$), but none of them were vocal learners. Among the outliers, the highest residual was 32.46 in a combination that included 4 passerines (songbirds and a sub-oscine), a parrot, and a falcon, and 17.61 in a core control that included 3 songbirds, an Anna's hummingbird, and 2 land fowls (**Fig. 1D**). These outlier species combinations do not share known convergent traits as far as we are aware. These findings support that amino acid convergences (ConSAVs) are widespread, and their numbers vary in a linear fashion for different species combinations, implying an underlying principle of convergence. We next sought for underlying principles of phylogenetics and the central dogma theory.

Produce of most recent ancestor divergence times influence amino acid convergences. According to previous studies on mammalian and drosophila nuclear genomes (Zou and Zhang 2015) and vertebrate mitochondrial genomes (Goldstein et al. 2015), fewer convergent substitutions are expected with greater phylogenetic branch distances reflecting time to accumulate mutations. However, the correlations found in those studies showed high levels of variation, which make it difficult to identify principles reliably as well as identify the outliers. Here, we analyzed many other phylogenetic variables (**Fig. 2A**). We found strong and significant correlations between ConSAVs and DivSAVs with the product of the MRCA (most recent common ancestor - origin) branch length lengths (POB) for both the broad control set and core

control set of species combinations (**Fig. 2B-D**). Much weaker to no correlations were observed with the product of terminal branches (PTB), distances among terminal branches (DTB) and distances among terminal nodes (DTN); **Fig. 2**). Like the ConSAV versus DivSAV correlation analyses (**Fig. 1D**), the avian vocal learners were not a significant outlier relative to all control sets in any the phylogenetic factor analyses (**Fig. 2**). These findings suggest that POB largely explains the proportion of ConSAVs in polyphyletic species, where the longer evolutionary time on the ancestral branch of the polyphyletic species combination the greater frequencies of ConSAVs.

Convergent or complex nucleotide changes cause amino acid convergences. To trace the molecular basements producing amino acid substitutions under the central dogma theory, we performed CSV finder analysis for avian codon alignments and identified avian vocal learner-specific single codon variants (AVL-SCVs, **Fig. 3A**) and single nucleotide variants (AVL-SNVs, Fig. 3**B**). Theoretically, nonsynonymous nucleotide sources causing amino acid substitutions can be simple and complex: SNVs at a homologous codon position and complex multiple non-exclusive nucleotide variants (CNENVs) in the same or different codon positions (**Fig. 3C**). We also analyzed synonymous variants that do not lead to AVL-SAVs (**Fig. 3D**). In the 4,519,041 homologous codons from 13,557,123 homologous nucleotides among the 8,295 orthologous genes in birds, we found 600 AVL-SCVs (**Fig. 3E**), and 148 nonsynonymous AVL-SCVs out of these explained the all 148 AVL-SAVs (**Supplementary table S1**). At the nucleotide level, of these 148 nonsynonymous AVL-SCVs, 54 (36%) resulted from AVL-SNVs and 94 (64%) resulted from CNENVs at different codon positions (**Fig. 3E**). By comparison for nonsynonymous substitutions, we found 452 synonymous AVL-SCVs, of which 111 (24%) resulted from AVL-SNVs and 341 (76%) resulted from CNENVs (**Fig. 3E**).

Next, focusing on amino acid convergences, we investigated evolutionary types of codon and nucleotide

sources of 24 convergent AVL-SAVs (AVL-ConSAVs) among the 148 AVL-SAVs. At codon level, the 24 AVL-ConSAVs were caused by both types of convergent AVL-SCVs (AVL-ConSCVs, 62.5%) and divergent AVL-SCVs (AVL-DivSCVs, 37.5%) (**Fig. 3F**). On the other hands, nucleotide sources were more complex: The 24 AVL-ConSAVs were caused by convergent AVL-SNVs (AVL-ConSNVs, 70.8%), divergent AVL-SNVs (AVL-DivSNVs, 4.2%), and CNENVs (25%) (**Fig. 3G**).

An example of the simple cases with a nucleotide convergence causing an amino acid convergence is the AVL-ConSAV site of *B3GNT2* mentioned earlier, where all vocal learners have the same convergent nucleotide (A) in the first position of the codon AAT for Asparagine (N) and all vocal non-learners have CAT or CAC for Histidine (H; **Fig. 3H**). Another example of complex cases with complex nucleotide variants at different sites causing an amino acid convergence is the ConSAV site of *LRRN4*, where all of vocal learners have CAC or CAT codon for Histidine (H), while nearly all vocal non-learners have either TAT or TAC codon for Tyrosine (Y) and two have more complex changes of TCG and GCG codons (**Fig. 3H**).

We also performed similar analyses on the 8,238 control species combinations, of which 8,109 (including the core 59) had one or more single codon variants (Ctrl-SCVs) and single nucleotide variants (Ctrl-SNVs) specific to each control set, with proportional distributions similar to vocal learners (**Fig. 3E-G**). These findings suggest that the convergent amino acid evolution can originate from not only convergent nucleotide variants but also divergent and complex nucleotide variants at same or different codon positions.

Codon and nucleotide variants are also strongly correlated with the product of MRCA branch lengths. To assess if phylogenetic time factors also influence codon and nucleotide variant numbers and compare among variant types, we performed correlation tests between all nine types of single sequence

convergent variants (three evolutionary types [all, convergent, and divergent] at three molecular levels [amino acid, codon, and nucleotide]) and four types of phylogenetic features (POB, PTB, DTB, and DTN). As expected, all nine types of single sequence variants were highly correlated with each other, in all control (**Fig. 4**) and core control species combinations (**Supplementary fig. S3**). For the phylogenetic features, the POB showed the strongest correlation with all variant types, where the others (DTB, DTN, and PTB) were either weaker or not correlated at all (**Fig. 4**, **Supplementary fig. S3**). Correlations were overall stronger in the all control species combinations (**Fig. 4**, **Supplementary fig. S3**), presumably due to the higher number of species combinations in the analyses than the core control sets. Like amino acid convergences (**Fig. 2**), the number of convergent vocal learner-specific codon or single nucleotide variants still were not outliers from the POB relationships (**Fig. 4**).

Fixation and positive selection of convergent amino acid sites in vocal learning birds. Like most studies, we have one genome per species, and thus some of the variants identified could instead be due to population differences within a species. Although with 47 species, we are less likely to detect many populations variants as convergent. Nevertheless, we checked for vocal learning variants in several species in a population analyses. We scanned the dbSNP database of zebra finch (n =1,257 samples; build 139) a vocal learner and chicken (n =9,586 samples, build 145) a vocal non-learner. At the 148 AVL-SAV sites, zebra finches showed complete fixation without any non-synonymous polymorphisms. And chicken only showed one missense SNP in the *OTOA* gene (c.2581A>G, p.Thr861Ala) resulting in an amino acid change identical to that of vocal learners (**Supplementary fig. S4**). We also validated fixation of the convergent substitutions in *DRD1B* by PCR of genomic DNA and sequencing of 3 male and 3 female zebra finches and 3 male and 3 female chickens each (**Supplementary fig. S5;** n = 6 total of each species). These findings indicate that the vast majority (99.3%) of the single amino acid variants we identified in

vocal learners are the result of true species-specific variants, that are fixed and thus presumably selected upon.

To check for positive selection, we performed dN/dS analyses with the branch-site model for the 24 AVL-ConSAV genes in the avian vocal learning species. We found that ~42% of the amino acid convergences in the vocal learning birds (10 of 24 genes) showed signs of positive selection (likelihood ratio value (D) > 0, posterior probability > 0.5, dN/dS ω_2 values of ancestral branches of each vocal learning clade > 1). With a stricter statistical significance (adjusted $p < 0.05$), only 3 of 24 AVL-ConSAVs (12.5%) were under positive selection. We compared the proportions of amino acid convergences in the closest control set (songbirds, parrots, and swifts) under positive selection, and found similar proportions where 46% (12 of 26 Swift-ConSAV genes) were positively selected with the cutoff for the likelihood ratio test (D) > 0 and 6 genes (23%) with the stricter significance (adjusted $p < 0.05$). These findings suggest that a subset of genes with amino acid convergences in vocal learners have been positively selected, but like the number of convergent sites, that this positive selection rate maybe the background rate on convergent substitutions as seen in the closely related control species combination without a convergent trait.

Post hoc analyses reveal rifleman variants are shared more with vocal non-learners. Although rifleman, a presumed vocal non-learner, was excluded from the initial CSV Finder analyses, we could ask in an unbiased way whether its sequences are more similar to vocal non-learners or vocal learners. We applied principal component analysis (PCA) and phylogenetic analysis for the 148 AVL-SAV sites and the subset of 24 AVL-ConSAV sites unique to vocal learning clades. Despite being more closely related to songbirds, the pattern of these 148 sites in rifleman clustered the species among the vocal non-learners (**Supplementary fig. S6A**). For the 24 AVL-ConSAV subset, the pattern in rifleman was separate from the two groups but still closer to vocal non-learners (**Supplementary fig. S6B**). Phylogenetic analyses of

the AVL-SAV and AVL-ConSAV sets were consistent with the PCA results, where instead of branching with its closest relative, the songbirds, rifleman was on a branch outside and next to the vocal learners (**Supplementary fig. S6A, B**). These results support the assumption that rifleman is a vocal non-learner, and further suggest that despite the number of variants and their positive selection not differing from expectation, that the actual substitutions and their associated genes could be related to the convergent trait of vocal learning.

Biological functions of genes with amino acid convergences. To investigate the biological functions of genes with convergent sequence variants in vocal learners and in other species combinations, we performed gene ontology (GO) analyses for 53,058 lists of genes with 1 or more of each type of sequence variants for vocal learners and all 8,238 control species combinations. Among them, at least one significant (adjusted $p < 0.05$) GO term was found for 7,901 gene lists (14.9%). For these 7,901 lists, we found a positive correlation between the number of significant GO terms and the number of genes with convergent variants in each list (**Fig. 5A**), meaning the lower number of convergent genes the less likely to find a significant shared GO function. The vocal learners were at the lower end of this correlation, and though they did not have significant GO enrichment for their total 148 AVL-SAV gene list, they did so for their fewer AVL-ConSAV 24 gene list, which was significantly enriched for 'learning' (GO:0007612, adjusted $p = 0.042$). Four genes were responsible for this enrichment: *DRD1B* [also known as *DRD1B*], *LRRN4*, *PRKAR2B*, and *TANC1* (**Fig. 5B**). The amino acid convergences of *DRD1B*, *PRKAR2B*, and *TANC1* were caused by codon convergences (AVL-ConSCVs) in all vocal learners, while that of *LRRN4* were caused by codon divergences (AVL-DivSCVs) with complex nucleotide variants (CNENVs; **Fig. 3H, 5C**). Based on literature reviews, 2 of these 4 genes function in a cAMP signaling pathway for learning, including several in vocal learning circuits, and 2 others are indirectly function in this pathway (**Fig. 5D**).

Out of the 8,238 control species combinations, only one had 2 gene lists, with Ctrl-DivSCVs and Ctrl-DivSNVs, that showed significant enrichment for 'learning' (GO:0007612, both adjusted *p* values = 0.02); the associated set of species (**Fig. 5E**) did not include any vocal learners, but another convergent variant in *LRRN4* contributed to this functional enrichment (**Fig. 5F**). The findings indicate that a subset of genes with identical convergent variants in vocal learners are enriched for learning functions, and this convergent enrichment is rare.

Candidate genes supported by meta-analyses for vocal learning. We next tested if there was any relationship between genes with amino acid convergences specific to vocal learning birds that we detected here and previous candidate genes implicated in vocal learning behavior and brain regions. Out of 8,295 singleton orthologous genes, we analyzed 6,932 genes with same gene symbols in at least one of six meta data sets from 3 types of comparisons: 1) 786 target genes of the *FOXP2* transcription factor (Lachmann et al. 2010; Lovell et al. 2020); 2) 1,769 genes (25.5%) that are up- or downregulated in song learning nuclei of the zebra finch in response to singing (Hilliard, Miller, Fraley, et al. 2012; Hilliard, Miller, Horvath, et al. 2012; Whitney et al. 2014; Lovell et al. 2020); and 3) Last, total 3,001 differentially expressed genes (DEGs, 43.3%) that have specialized up- or down-regulated expression in the avian song nuclei or human speech regions compared to their surrounding non-vocal motor brain regions (Lovell et al. 2018; Gedman et al. 2022). Out of the total 6,932 genes, 4,353 genes were supported by at least one of meta data sets (**Supplementary table S2**).

We calculated whether there were any overlapping enrichments between gene sets using hypergeometric test. As the results, total nine significant enrichments between only 5 gene sets and 5 meta-data sets before multiple testing corrections (p value < 0.05, **Supplementary fig. S7**), while all sets did not show any statistical significance after FDR test (adjusted p value =1). First, AVL-ConSAV genes and AVL-

ConSAV genes with higher likelihoods for positive selection (AVL-ConSAV_PS_LRT) were enriched for *FOXP2* targets reported in the ZEBrA database (p = 0.03 and 0.01, respectively). Next, 3 vocal learner-specific gene sets (AVL-SAV_PS_LRT genes, AVL-ConSAV genes, and AVL-ConSAV_PS_LRT genes) showed enrichments for singing regulated genes in song nuclei of the ZEBrA database and in area X of Hilliard et. al. Last, AVL-ConSAV genes under positive selection (AVL-ConSAV_PS_FDR) and the Swift-DivSAV_PS_LRT gene set of the closest control set replacing hummingbird to swift were enriched for DEGs in nucleus XII of Lovell et al. and in RA of Gedman et al. Afterthen, we also calculated proportions of variable genes of vocal learners and the closest control set supported by meta data sets (**Fig. 6A**). The 1,463 DEGs of three song nuclei (Area X, HVC, and RA) showed the maximum proportion (4/10, 40%) for AVL-ConSAV genes under positive selection, while 2,266 DEGs of all song nuclei showed the higher proportion (10/23, 43.5%) for Ctrl-ConSAV genes of the closest control set.

These meta-analyses identified potent candidate genes supported by multiple lines of evidence for vocal learning. Out of the 24 AVL-ConSAV genes, 10 had at least one relationship to vocal learning gene expression data sets (**Fig. 6B**), and 6 of these 10 genes had higher likelihood values for positive selection on vocal learning clades (D>0): *B3GNT2, DRD1B, FNDC1, PIK3R4, PRKAR2B, and SMPD3* (**Table 1**). The top candidate gene, *B3GNT2*, was supported by all 3 major types of meta data sets and its down-regulations in HVC and RA were additionally validated with in-situ hybridization in the ZEBrA database (Lovell et al. 2020) (**Fig. 6C**). Two of these 6 genes, *DRD1B* and *PRKAR2B*, also contributed to the GO finding for learning functions (**Fig. 5C, D**). Further *DRD1B* has a specialized up-regulation specific to adult Area X compared to its surrounding striatum (Kubikova et al. 2010) (**Fig. 6D**).

Most enriched substitutions in vocal learners remain with more species. Until now, we investigated vocal learner-specific amino acid convergences by comparing 48 species from the first phase of the B10K

project. While we were completing this study, the consortium additionally released Illumina-based genome assemblies and constructed genome-wide alignments of a total of 363 avian species representing 218 out of 236 bird families, providing an opportunity to check whether the convergent patterns still exist in the more densely sampled data set of the avian family tree (Feng et al. 2020). We analyzed 8 key candidate sites with vocal learner ConSAVs supported by multiple lines of evidence (**Fig. 7**, **Supplementary table S3**). Although exclusivity to known vocal learners was reduced, the substitutions were still highly enriched to the vocal learners: depending on site, 89.6% to 100% of the 154 vocal learning species still had the substitution and 86.1% to 100% of the 209 species regarded as vocal non-learners did not have it (**Fig. 7A, Supplementary table S4**). For example, *B3GNT* kept its mutually exclusive variants convergently unique to 154 species from 3 vocal learning clades compared with the other 209 species (**Fig. 7A, B**). For *DRD1B*, 153 of the 154 vocal learning birds maintained the Alanine (A) substitutions, with the exception being the noisy scrubbird (Atrichornis clamosus), once thought to be extinct, which lost the pattern to Valine (V); only 1 to 2 species in 3 bird lineages (a suboscine, Otidimorphae, and Palaeognather) of 209 vocal non-learners had the substitution seen in vocal learners. For *LRRN4*, one of the two hummingbird species lost the substitution seen in all the remaining 153 vocal learners (**Fig. 7B**). These findings indicate although exclusive substitutions in vocal learners are rare, extremely enriched substitutions are present, which also questions whether vocal learning is exclusive to the known clades.

Discussion

As the primary structure of proteins, amino acid substitutions can contribute to various traits including human spoken language (Lai et al. 2001; Enard 2011; Berwick et al. 2013). To assess whether there are convergent amino acid substitutions to species that have a component of spoken language, the rare trait of vocal learning (Jarvis 2019), we developed new analyses tools to scan the genome alingnments of species

representing the avian family tree, in different polyphyletic species combinations. In doing so, we discovered positive correlations between the frequencies of molecular convergences at the amino acid, codon, and nucleotide levels with the product of the MRCA (=origin) branch lengths (POB), implying a natural background rate of genetic convergence among species. Although vocal learners did not have a higher preponderance of convergent variants above background levels, a subset of the convergent genes was significantly enriched for a biological process and meta-data sets associated with vocal learning. To explain our findings, we propose following hypotheses of selection on a background of convergent substitutions for convergent traits.

When searching for convergent substitutions among species, we believe our approach of a multi-wise species comparisons and the POB maybe more informative than past approaches. Previous studies found correlations between convergent and divergent variants (called convergent and divergent substitutions in those studies) between pairs of species (Castoe et al. 2009; Thomas and Hahn 2015). We find that such a relationship exists also at higher dimensional species combinations. But this type of analyses does not control for species relationships. Several other studies found that the rate or number of convergent substitutions decreases with increasing branch distance between two polyphyletic species (Goldstein et al. 2015; Zou and Zhang 2015), whereas we fail to find strong relationships of genetic convergence with such phylogenetic parameters including summations of branch lengths. However, we find much stronger associations between numbers of convergent variants and the product of the MRCA branch lengths in polyphyletic clades. This result suggest that the deeper in time their common ancestor, the more likely the evolution of higher proportions of detectable convergences at the amino acid, codon, and nucleotide levels. We believe that this finding may provide a new null hypothesis of convergent evolution according to phylogeny.

Against this phylogenetic background of convergence, our positive selection, gene functional associations, and gene expression analyses suggest that selection occurs on some of these convergent substitutions to contribute to evolving novel, convergent traits, in our case vocal learning. According to this hypothesis, it is not about how many genes show convergence, but which specific genes and specific nucleotide sites that show convergence, as the more important factor to consider.

The best example is the *DRD1B* dopamine receptor, a striatum-enriched gene, with further specialized further up-regulation in the Area X song nucleus of the striatum, suggesting further selection on regulatory genomic region changes. Often coding and regulatory genomic sources of trait evolution are pitted against in each other as alternatives (Carroll 2005), but our findings suggest that they could synergistically influence evolution of each other. Preliminary studies in our group find the promoter regulatory region of *DRD1B* has differential open chromatin activity within Area X versus the surrounding striatum, and the striatum further has differential open chromatin differences with the other cortical-pallial regions and song nuclei (Rhie et al. 2021; Kim et al. 2022). Ongoing studies are analyzing convergent nucleotide substitutions in these differential regulatory regions, that are more completely assembled with long reads by the Vertebrate Genomes Project (VGP) (Rhie et al. 2021). The assemblies we used here are from the 48 avian genomes generated from the first phase of Avian Phylogenomics Project (Jarvis et al. 2014; Zhang et al. 2014), and the 363 species from the second phase (Feng et al. 2020), but are short-read-based and we found are missing many of the GC-rich regulatory promoter regions; they do have the most of the protein coding regions of genes (Rhie et al. 2021; Kim et al. 2022). The new long read assemblies, however, do not yet have all bird or vertebrate orders represented. When they are complete, we will be able to further analyze possible convergent relationships between coding and non-coding regulatory regions.

Based on literature reviews for candidate genes related to learning process, we suggest a potential molecular pathway for vocal learning. This includes: *DRD1B*, which functions to regulate striatal activity associated with learning (da Silva et al. 2012; Wong et al. 2012); *LRRN4* that affects long lasting memory (Bando et al. 2005); *TANC1* that regulates dendritic spines and spatial memory (Han et al. 2010); and *PRKAR2B* (Protein Kinase cAMP-dependent Type II Regulatory Subunit Beta), an enzyme that activates cAMP-dependent protein kinase (PKA) inside the cell (Solberg et al. 1992). Blocking of PKA inhibits learning, including vocal learning (Abe et al. 2015), namely through the cAMP response element binding protein (*CREB1)*, a transcription factor responsive to cAMP signaling via PKA, which regulates genes that convert short-term memories into long-term memories (Kandel 2012). *DRD1B*, through its G-protein, also regulates activity of adenylyl cyclase's synthesis of cAMP in the cell membrane (Sunahara et al. 1991; Rangel-Barajas et al. 2015). Based on findings in the literature, all four of these may be part of a nexus at targeting the cAMP signaling pathway associated with learning (**Fig. 5D**).

Our finding that the substitutions of these genes were extremely enriched but no longer exclusive to vocal learners when expanding the number of species 7.5-fold (from 48 to 363) across the family tree, does not negate their association with vocal learners. Vocal learning has often been treated as a dichotomous trait, as one either has it or doesn't. But experiments in mice, suboscine birds, and other so-called vocal non-learning species (Arriaga et al. 2012; Petkov and Jarvis 2012; Liu et al. 2013; Jarvis 2019; Fischer et al. 2020; Ten Cate and Fullagar 2021) have found various degrees of component traits from the behavior to neural connectivity seen in vocal learners. Even within well-established vocal learning clades, there are differences in vocal learning abilities, with some species able to imitate human speech and other species, and other species with a limited repertoire of just one learned song for its entire life (Kroodsma 1983). This has led to the continuum hypothesis (Arriaga et al. 2012; Petkov and Jarvis 2012; Arriaga and Jarvis 2013; Jarvis 2019) and multidimensional hypothesis of vocal learning (Wirthlin et al. 2019). We suggest

that the extremely enriched convergent substitutions found in known or advanced vocal learners and also in 1 to 3 other species in certain clades, could be associated with a continuum. Another alternative is that it is the combination of these convergent substitutions occurring together in a species that is more important than the presence of any one of the substitutions alone. Our findings provide candidates for testing either of these non-mutually exclusive hypotheses.

Other ideas for future studies that our studies suggest will be whether similar or different rules apply for genetic convergence and the POB in non-coding regions regulating differential gene expressions. It will be useful to determine if the convergent rules we identified here are specific to birds, or are more widespread across life forms. Vocal learning species also share other convergent traits besides vocal learning (Naguib and Riebel 2014; Nowicki and Searcy 2014; Mason et al. 2017; Jarvis 2019), and the identified genes and their functions in the surrounding brain regions could be associated with these other traits. Overall, our study generates new hypotheses on principles of convergent evolution.

Materials and Methods

Multiple sequence alignments of singleton orthologous genes in birds

In our preliminary studies, the Avian Phylogenomics Project (now the Bird10K project) defined 8,295 singleton orthologous gene sets across 48 avian species, and constructed the phylogenetic avian family tree consisting of at least 34 orders (Jarvis et al. 2014; Zhang et al. 2014; Jarvis et al. 2015). This 1:1 orthologous gene set was identified by reciprocal best blast hits and synteny, using two species as a reference: chicken and zebra finch. They were then aligned across all species using SATé+MAFFT and SATé+Prank, for both nucleotide and amino acid sequences. Alignment frameshift errors were corrected when translating into amino acid sequence alignments. This resulted in detection of 4,519,041 amino acid and 13,557,123 nucleotide homologous sites. In our previous analyses for amino acid substitutions, we

used Gblocks (Castresana 2000) to remove poorly scored alignments with sequence divergences and columns with gaps in at least one species included. However, here we found that this was too aggressive, removing 65% of the aligned sequences. For example, vocal learner-specific amino acid substitutions of *DRD1B* (*DRD1B*) were excluded because of gaps in one of the outgroup species (Lizard) in the previous study (Zhang et al. 2014) (data is not shown). Therefore, in the current study we used alignments without the Gblocks trimming step.

Detection of convergent variants

We initially developed an algorithm to find amino acid substitutions specific to a group of species, called Target-specific Amino Acid Substitution (TAAS) analysis (Zhang et al. 2014). It could not detect insertion/deletions (indels) specific to a group of species. In this study, we improved the algorithm to detect indels as well as applied ancestral sequence reconstructions to find convergent variants at amino acid, codon, and nucleotide levels, and named it as convergent variant finder (ConVarFinder; **Supplementary fig. S1**). ConVarFinder focuses on identifying molecular convergences specific to multiple species from polyphyletic lineages, while TAAS ignored phylogenetic relationships.

First, ConVarFinder identifies mutually exclusive variants at amino acid, codon, and nucleotide levels between a target group of species relative to all other species tested. To focus on point mutations, we excluded continuous variants potentially regarded as structural variants. Examples of single amino acid and codon variants were summarized and visualized by using WebLogo (v2.8.2) (Schneider and Stephens 1990). Second, ConVarFinder classifies the mutually exclusive variants into 4 types based on equality or inequality of sequence information in each group: Type 1, mutually exclusive identical amino acid substitutions between group A and group B species; Type 2, identical amino acid substitution in group A and different substitutions in group B; Type 3, the inverse of Type 2, with different substitutions in group

A, not shared with an identical substitution in group B; and Type 4, mutually exclusive different sets of amino acid substitutions in group A and group B. Third, it infers the evolutionary histories of these substitution variants from their common ancestors to terminal taxa using a given phylogenetic tree. The ancestral sequences were estimated by RAxML (version 8.2.12) (Stamatakis 2014) for codon substitutions with '-f A -m GTRCAT -p 12345' options and for indels converted as binary sequences with '-f A -m BINCAT -p 12345' options. RAxML usually removes the codon sites consisting of all gaps ('---' or 'NNN') in all species, so we trimmed the reduced sequences when we merged the codon and indel sequences by using a custom python script. Based on the 'RAxML_marginalAncestralStates' and 'RAxML_nodeLabelledRootedTree' outputs, we checked the substitutions on the most recent common ancestral (MRCA=origin) branches of each clade of group A species and classified their evolutionary directions as convergences or divergences. The source codes of ConVarFinder and estimated ancestral sequences are accessible at the following link (https://github.com/chulbioinfo/ConVarFinder).

Control sets of species combinations from three independent lineages

Considering that we have 6 vocal learning species, we calculated all 6 species combinations of 47 birds in the avian family tree excluding Rifleman, which was 10,737,573 combinations (**Supplementary fig. S2**). Of these, 8,239 combinations of 6 species originated from 3 independent lineages including 3 vocal learning lineages (songbirds, parrots, hummingbirds). From these convergent combinations, we designed two main types of control sets: all control set from the 8,238 set of 6 species with 3 independent origins; and a core control set consisting of 59 possible convergent combinations of species that have a similar phylogenetic history to vocal learners, but contained 6 species originated from 2 clades out of 3 vocal learning clades and 1 vocal non-learning clade.

Correlation tests

To check statistical significances of correlations between various features we discovered in this study, such as, convergences and divergences at the amino acid level (ConSAVs VS DivSAVs), we calculated Spearman rank correlation coefficient as:

$$rho = \frac{\Sigma(x' - m_{x'}) * (y' - m_{y'})}{\sqrt{\Sigma(x' - m_{x'})^2 * \Sigma(y' - m_{y'})^2}}$$

where x' and y' are each rank of x and y, respectively; and $m_{x'}$ and $m_{y'}$ correspond to the means of rank(x) and rank(y), respectively. By using 'cor.test' function with the option method = "spearman" in R package (ver. 3.5.1), we tested correlations between ConSAVs and DivSAVs in the multiple combinations of species (e.g. a set of avian vocal learners, 8,238 all control sets, and 59 core control sets). After then, we performed linear regression analysis for modeling the relationship between ConSAVs and DivSAVs based on 'lm' function, and visualized it with 'plot', 'points', and 'abline' function in R package (ver. 3.5.1) (R Core Team, R 2013). We also performed Bonferroni Outlier Test to check whether the number of convergent variants of vocal learners or other species combinations are outliers, as determined by residuals from the regression model with the 'outlierTest' function in R package (ver. 3.5.1) (Fox et al. 2012; R Core Team, R 2013); option for limitation of the max number of outliers as 3: 'n.max=3'. The source code and dataset to perform above analyses are accessible at the following link (https://github.com/chulbioinfo/ConVarFinder).

Phylogenetic features related to the number of molecular convergences

We performed multiple clade-wise comparisons of at least 3 polyphyletic clades to find relationships between convergent variants and various phylogenetic features. Using the branch lengths of the avian total evidence phylogenetic tree from Jarvis et al (Jarvis et al. 2014) (**Fig. 1A**), we calculated four types of phylogenetic branch measures for convergent groups of species: product of origin branch lengths (POB);

product of terminal branch lengths (PTB); distance between terminal branches (DTB); and distance between terminal nodes (DTN; **Fig. 2A-D**). POB was calculated by multiplying lengths of most recent common ancestral (MRCA=origin) branches of each target clade and PTB as branch lengths of terminal taxa. DTB was calculated as a summation of lengths of all branches between the MRCA node of the 47 birds and each terminal taxon, whereas the DTN was calculated as the summation between the MRCA node and the most recent ancestral nodes of each terminal taxon (**Fig. 2A**). The source code to calculate each phylogenetic feature is accessible at the following link (Https://github.com/chulbioinfo/ConVarFinder).

PCA and ML tree analyses for Rifleman

Principle component analysis (PCA) was performed using the method as implemented in JalView (Waterhouse et al. 2009). Focusing on the 148 AVL-SAV sites (AVL-ConSAVs + AVL-DivSAVs) and 24 AVL-ConSAV sites alone, pairwise scores between bird species was computed by summing the substitution scores from BLOSUM62. Then, we performed spectral decomposition of the score matrix to obtain principal component (PC) vectors and eigenvalues of the respective vectors. The PCA biplot was computed using PC1 and PC2 vectors. For the maximum likelihood (ML) tree, we constructed it using MEGA (Kumar et al. 2018), and selected the JTT model, on the part of the amino acid sequence alignment of all 148 AVL-SAV or 24 AVL-ConSAV sites.

Gene ontology functional annotations and gene network analyses

We conducted Gene Ontology (GO) analysis by using g:Profiler (v 0.3.5.) (Raudvere et al. 2019) with the default option and ClueGO (ver. 2.3.3.) in Cytoscape (Shannon et al. 2003) with the following options: GO BiologicalProcess-GOA (released in 08.04.2016); all of GO tree interval; all of GO Term/Pathway

selection; multiple testing correction by Bonferroni analyses (adjusted p-value < 0.05); and default options of others. We then tested whether the number of genes is correlated with the number of significant GO terms, by applying regression analyses using 'lm' function. We visualized the results with 'plot', 'points', and 'abline' functions in the R package (ver. 3.5.1) (R Core Team, R 2013).

After then, we focused on two GO lists enriched for learning process: AVL-ConSAV gene list of vocal learning birds and Ctrl-DivSCV and Ctrl-DivSNV gene lists of a control set (different codon convergences specific to Dalmatian pelican, little egret, houbara bustard, red-crested turaco, white-throated tinamou, and ostrich). We searched for networks between the enriched genes for learning by analyzing protein-protein interactions among convergent genes, using CluePedia ver. 1.3.3. (Bindea et al. 2009) in Cytoscape (Shannon et al. 2003), selecting the following databases: STRING-ACTIONS_v10.0 (released in 07.05.2015); activation v10.0; binding v10.0; catalysis v10.0; expression v10.0; inhibition v10.0; ptmod v10.0, and reaction v10.0. Sequences of the convergent variants of gene lists of vocal learners and a control set associated with learning were visualized by WebLogo (v2.8.2) (Schneider and Stephens 1990; Crooks et al. 2004).

Fixed differences of vocal learner-specific amino acid variants within populations of zebra finch and chicken

More than 20 million (20,739,045) and 1.6 million (1,661,545) variants have been reported in chicken (n=9,586) and zebra finch (n=1,257), respectively, according to Ensembl database release 84 (Yates et al. 2016; Zerbino et al. 2018). Hence, we performed additional analysis to check if the AVL-ConSAV and AVL-DivSAV sequences we identified in vocal learners were not due to within species variation. Local alignment was conducted for the CDS sequences containing these AVL-ConSAVs using BLAST (ver. 2.8.1) (Madden 2003) to find the position of SAV on the chromosome sequence of chicken (Galgal4) and

zebra finch (taeGut3.2.4) according to Ensembl database release 84. Fixation of sequence in a species was assessed by comparing the nucleotide site in question in an alignment from Ensembl dbSNP build 145 and 139 of chicken and zebra finch, respectively (Sherry et al. 2001).

We also performed additional fixation analyses on several genes amplified by PCR from red blood cells of zebra finch (n = 3 males and 3 females) and chicken (n = 3 males and 3 females). The *DRD1B* gene was cloned from genomic DNA by using zebra finch specific primers (forward 5'-GCC CTG CGT CAG TGA GAC CA-3' and reverse 5'-CCG CCA GCC CCC TGT ATG AC-3') and white-leghorn chicken specific primers (forward 5'-CAG ATC TCC CCC GAC CCC GA-3' and reverse 5'-GGC AAC AAT GCC GCC TGG AG-3'). The PCR reaction was conducted in a total volume of 20 ul containing 100 ng genomic DNA, 10x PCR buffer, 0.4 µl dNTP (10 mM each), 10 pmol of each primer, and 0.5 U Taq polymerase (BioFACT) in the following thermocycling conditions: 2 min at 95°C, followed by 35 cycles of 20 s at 95°C, 40 s at 60°C, 2 min at 72°C, and, finally, 5 min at 72°C. The PCR products were cloned into the pGEM-T easy vector (Promega) and sequenced using an ABI Prism 3730 XL DNA Analyzer (Thermo Fisher–Applied Bio- systems).

Positive selection on MRCA branches of vocal learners and the closest control set

The *dN* (the rate of non-synonymous substitution), *dS* (the rate of synonymous substitution) and $\omega = dN/dS$ were estimated along each branch of the phylogenetic tree and across sites by using the branch-site model A, implemented in codeml within PAML ver. 4.6 (Yang 2007) with F3X4 codon frequencies. We assumed the vocal learning trait independently originated from the most recent common ancestral branch of each vocal learning lineage (Cahill et al. 2021). Log likelihood ratio test (LRT, *D* value) was performed to compare the null hypothesis with a fixed ω (model 2) for neutral and negative selection ($\omega \leq 1$) and an alternative hypothesis with an estimated ω (model 2) for positive selection ($\omega > 1$). Orthologs with ω2

Foreground > 1 and number of accelerated sites (BEB > 0.5) > 0 were retained (branches tested for positive selection are referred to as "foreground" branches and all other are referred to as "background" branches). The data set of codon sequences of each gene list, including alignment gaps in species, was analyzed with a codeml option (cleandata = 0) and robust cutoff of adjusted *p*-value (<0.05; false discovery rate, FDR). FDR was calculated in R (ver.3.0.1). Genes with three types of single amino acid variants (all, convergence, and divergence) were analyzed with consideration on likelihoods and statistical significances of their alternative models for positive selection (D>1 and adjusted p value < 0.05, respectively) specific to vocal learner set and the closest control set (3 songbirds, 2 parrots, and the closest vocal non-learning relative species of hummingbirds, swift).

Meta-analyses for vocal learning candidate genes

We collected 6 published meta data sets associated with vocal learning for three major types of comparisons: *FOXP2* targets, singing induced genes in song nuclei, and differentially expressed genes in song nuclei relative to their surrounding non-vocal motor brain regions. Out of the 8,295 singleton orthologous gene set, 83.6% (6,932 genes) was listed up at least one of meta data sets with same gene symbols.

1) *FOXP2* targets: total 786 target genes of *FOXP2*, the foremost language gene, in the ZEBrA database (Lovell et al. 2020) and ChEA database (Lachmann et al. 2010).

2) Singing induced genes: A data set for 165 singing induced genes in the ZEBrA database (Lovell et al. 2020). A data set for 2,013 genes induced by singing behaviors in Area X (Hilliard, Miller, Fraley, et al. 2012). A data set for 852 singing induced genes in Area X compared to the striato-pallidum ventral to area X (VSP) which were significantly enriched in functional annotations (Hilliard, Miller, Horvath, et al.

2012). A data set for 2,740 transcripts responded for singing behaviors in song nuclei (Whitney et al. 2014).

3) Differentially expressed genes: A data set for 2,640 marker genes of three song nuclei including Area X, HVC, and RA comparably regulated in their surrounding non-vocal brain regions (Lovell et al. 2018). A data set measuring 21,617 genes, of which 5,473 were differentially expressed between 4 song nuclei including Area X, HVC, LMAN, and RA relative to their surrounding non-vocal motor brain regions (Gedman et al. 2022).

Based on the above meta data sets of precedented candidate genes for vocal learning, we calculated their proportions in lists of genes with single amino acid evolution that we detected in this study.

Amino acid patterns of candidate convergent sites of vocal learners in 363 avian genome alignments

For 6 key candidate genes (*B3GNT2, DRD1B, FNDC1, PIK3R4, PRKAR2B,* and *SMPD3*) and 2 more leaning-related genes (*LRRN4* and *TANC1*), we collected and analyzed the reference-free genome-wide CACTUS alignments of 363 birds generated by the 2nd phase of Avian phylogenomics consortium (B10K) (Feng et al. 2020). To find genomic positions of AVL-ConSAVs of candidate genes, we parsed zebra finch's coding sequences of codons producing the amino acid convergences with 10bp nucleotides at their up/down streams of in coding gene alignments of 1:1 orthologs. Next, matching the version of zebra finch genome assembly (taeGut2) in the CACTUS alignments, we manually scanned the genomic positions of parsed sequences matched in its genome sequence (**Supplementary table S3**) by using the UCSC genome browser (Nassar et al. 2023). Based on the zebra finch's genomic locations of AVL-ConSAVs of candidate genes, we extracted chromosome-wide MAF files by setting zebra finch as the reference species from the published HAL file (https://cgl.gi.ucsc.edu/data/cactus/363-avian-2020.hal) (Feng et al. 2020) by using hal2maf in the HAL tool (Hickey et al. 2013). After then, we extracted alignment blocks with AVL-

ConSAVs, changed the species order following the avian family tree (https://cgl.gi.ucsc.edu/data/cactus/363-avian-2020-phast.nh) (Feng et al. 2020), and finally extracted codon sequences of all species at the AVL-ConSAV sites of candidate genes in the alignment blocks by using the in-house code, MafScan.py (https://github.com/chulbioinfo/CSAVanalysis).

At each AVL-ConSAV site of 8 candidate genes, vocal learning-type amino acid patterns were defined from amino acid information of zebra finch. We tested and visualized amino acid enrichments of vocal learner-type amino acids in vocal learning clades (n=154) and the other vocal non-learning clades (n=209) by using sequence Logo (Crooks et al. 2004) and calculated and visualized the proportions of vocal learner-type amino acid patterns in total 22 clades by using Microsoft Excel and Adobe Illustrator.

Institutional Review for animal cares and experiments

The care and experimental use of animals (zebra finch or chicks) were approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-150827-1) and the Rockefeller University IACUC. The experimental animals were maintained according to a standard management program at the University Animal Farm, Seoul National University or the Rockefeller University. The procedures for animal management adhered to the standard operating protocols of the laboratory at Seoul National University, Korea or the at the Rockefeller University.

Table legends

Table 1. Candidate genes with convergent amino acid changes under positive selection on vocal learning clades and supported by multiple lines of evidence for vocal learning. 'Pos_AA' Column shows amino acid position in gene-wide multiple sequence alignments. 'AA_avian vocal learners (AVLs)' and 'AA_avian vocal non-learners (AVNLs)' column shows amino acid profiles of vocal learning and non-learning group, respectively. 'dN/dS (ω) FG' column shows dN/ dS (ω) values of ancestral branches of vocal learning clades assigned as foreground branches assuming positive selection (ω>1). 'Likelihood ratio (*D*)' column shows likelihood ratios between null and alternative hypothesis assuming neutral and negative selection ($D \leq 0$) and positive selection ($D > 0$), 'Adj. *p* (FDR)' column shows adjusted p-value assuming positive selection (ω>1) applied with false discovery rate (FDR) correction (Adj. $p < 0.05$). 'Posterior probability (BEB)' column shows posterior probabilities calculated from bayes empirical bayes posterior approach to estimate probabilities assuming the amino acid sites under positive selection. The other columns show meta data sets for 3 major types of precedented candidates for vocal learning: target genes of the famous language gene, *FOXP2*, in the ZEBrA database (Lovell et al. 2020); singing induced genes in 'Area X (ZEBrA)' from the ZEBrA database (Lovell et al. 2020), 'Singing_AreaX (Hilliard_2012)' (Hilliard, Miller, Fraley, et al. 2012), 'Singing_AreaX_vs_VSP (Hilliard_2012)' (Hilliard, Miller, Horvath, et al. 2012), and 'Singing (Whitney_2014)' (Whitney et al. 2014), respectively; and differentially expressed genes of song nuclei compared to their surrounding sub-brain regions of zebra finch collected from 'DEG_songnucleiVSsurrounding (Lovell_2018)' (Lovell et al. 2018)and 'DEG_songnucleiVSsurrounding (Gedman_2022)' (Gedman et al. 2022) (see more details in **Materials and Methods**).

Table 1

		AA profiles		Positive selection				FOXP2_target (ZEBrA)	Singing				DEG_song nuclei _VS_surrounding	
Symbol	Pos_AA	AA_avian vocal learners (AVLs)	AA_avian vocal non-learners (AVNLs)	dN/dS (ω) FG	Likelihood ratio (*D*)	Adj. p (FDR)	Posterior probability (BEB)		Area X (ZEBrA)	Area X (Hilliard_2012)	Area X_vs_VSP (Hilliard_2012)	All songnuclei (Whitney_2014)	AX, HVC, RA (Lovell_2018)	All songnuclei (Gedman_2022)
B3GNT2	253	N	H	3.3	1.1	2.8.E-01	0.999	+	+	+		AXdown	HVCdown	HVCdown /RAdown
DRD1B (DRD1B)	416	A	I,V	3.7	1.2	2.7.E-01	0.5		+				RAup	LMANdown
FNDC1	1034	S	G	6.7	4.3	6.4.E-02	0.981				+		RAup	
PIK3R4	671	C	R	10.4	4.9	4.9.E-02	0.997						RAup	
PRKAR2B	32	V	I,-	295.8	25.4	3.6.E-06	0.999				+			
SMPD3	307	C	Y,-	14.3	6.0	3.4.E-02	0.994							AXup

Figure legends

Fig. 1. Convergent amino acid variants in vocal learning birds and other species combinations. (**A**) Avian family tree and genomes analyzed. The branch lengths are estimated from the RAxML tree (Jarvis et al. 2014). Red text, vocal learning lineages. Bold red lines, most recent common ancestral branches (origin branch) of each vocal learning clade. (**B**) Example illustration of the four types of convergent single amino acid variants (sky blue-colored boxes), in vocal learning birds versus vocal non-learning birds as an example. (**C**) Example amino acid sequence logo cases of a Type 1 ConSAV site in *B3GNT2* and a Type 3 DivSAV site in *SMRC8*. (**D**) Correlation plots between amino acid convergences (ConSAVs; y-axis) and divergences (DivSAV; x-axis) of control species combinations consisting of 6 species from 3 independent lineages relative to 42 other species. x = outliers, and colors denote different types of species combinations. Two correlations are shown, for the core control set of 59 species combinations (black) and the broader control set of 8,238 species combinations (grey).

Fig. 2. Phylogenetic features correlated with amino acid convergences. (**A**) Four types of phylogenetic tree features measured: product of origin branch lengths (POB); product of terminal branch lengths (PTB); distance among terminal branches (DTB); and distance among terminal nodes (DTN). In the example trees, red lines show the branches used for the calculations and red texts show the species clades that have a convergent trait. (**B-D**) Regression analyses of three categories for convergent single amino acid variants (SAVs) - convergent SAVs (ConSAVs), divergent SAVs (DivSAVs), and total convergent and divergent SAVs (ConSAVs + DivSAVs) in the vocal learning set and control sets of avian species - with each type of phylogenetic features. Color coding is the same as in **Fig 1D**.

Fig. 3. Molecular sources causing amino acid convergences. (**A**) Example illustration of four types of single codon variants (SCV). (**B**) Example illustration of four types of single nucleotide variants (SNV). (**C**) Concept of single amino acid variants (SAVs) explained by non-synonymous SCVs and SNVs. Left case, ConSAVs caused by non-synonymous ConSCVs with ConSNVs at a homologous nucleotide site. Right case, ConSAVs caused by non-synonymous DivSCVs with complex non-exclusive nucleotide variants (CNENVs) at different sites in the codon of different species. (**D**) Concept of non-exclusive amino acid variants explained by synonymous substitutions between species, which do not cause amino acid changes. Left case, same amino acids caused by synonymous ConSCVs with ConSNVs at a homologous nucleotide site. Right case, same amino acids caused by synonymous DivSCVs with CNENVs. (**E**) Venn diagrams of the different subsets of nonsynonymous and synonymous SCVs and SNVs outlined in (**C**) and (**D**), in avian vocal learners (n=1), control sets with at least one SCVs (n=8,109), and the core control sets (n=59). (**F, H**) Proportions of types of codon and nucleotide sources causing amino acid convergences. (**H**) Examples of amino acid convergences among vocal learners (ConSAVs) originating from different types of nucleotide variants at the same site (in *B3GNT2*) or different sites (in *LRRN4*). Red

text, avian vocal learners. Sky blue boxes, sites with SAVs; Light sky-blue boxes, SCVs; Dark sky-blue box, SNVs.

Fig. 4. Phylogenetic parameters proportional to various types of genetic variants at the amino acid, codon, and nucleotide levels in all control species combinations. Regression plots between frequencies of convergences among genetic variants and phylogenetic features are visualized in lower diagonal matrices. *p* values and Adjusted R^2 of correlations are visualized at upper diagonal matrix ($p<0.05$*, $p<0.01$**, and $p<0.001$***). Histograms of frequencies of each convergent variant and values of each phylogenetic feature are visualized on the diagonal matrix. Grey, orange, and red spots indicate all control sets (n=8,237), the closest control set of vocal learners (n=1), and the set of avian vocal learners (n=1), respectively. Black lines and black 'X' marks indicate regression lines and outliers, respectively. POB = product of origin branch lengths, PTB = product of terminal branch lengths, DTB = distance between terminal branches, DTN = distance between terminal nodes, SAV = convergent + divergent single amino acid variants, ConSAV = convergent SAV, DivSAV = divergent SAV, SCV = single codon variants, ConSCV = convergent SCV, DivSCV = divergent SCV, SNV = convergent + divergent single nucleotide variants, ConSNV = convergent SNV, DivSNV = divergent SNV. Correlations of core control sets are shown in **Supplementary fig. S4**.

Fig. 5. Functional ontology of genes with amino acid convergences. (**A**) Correlation plot between the number of significantly enriched GO terms and the number of genes with 1 or more convergent variants in each species data set with 9 convergent sequence types (SAVs, ConSAVs, DivSAVs, SCVs, ConSCVs, DivSCVs, ConSNVs+DivSNVs, ConSNVs, and DivSNVs) (n=53,058 returned GO terms). (**B**) GO and network results for four learning associated genes with amino acid convergences (ConSAVs) in vocal

learners (*adj.* $p < 0.05$). (**C**) Codon and amino acid sequence logos of AVL-ConSAV genes associated with learning. (**D**) cAMP-related learning molecular pathways. Red hexagons indicate the four AVL-ConSAV genes that function in learning. Transparent red hexagons indicate muti-cellular location of the candidate genes. Out of 4 learning associated genes, *DRD1B* (*DRD1B)* and *PRKAR2B* interact with cAMP, and/or its target *CREB1*, a transcription factor well known for its role in helping to convert short-term memories into long-term memories. Blue rectangles indicate traits related to learning. Black arrows and dashed blue arrows indicate trait-gene relationships and trait-trait relationships, respectively. (**E**) GO and network analysis for learning associated genes with codon and nucleotide divergences (DivSCVs and DivSNVs) of a control set (*adj.* $p < 0.05$). (**F**) Codon and amino acid sequence logos of control DivSCV and DivSNV genes associated with learning.

Fig. 6. Meta-analysis for overlapping genes sets in vocal learning species. (**A**) Proportions of candidate genes for vocal learning in meta data sets (color lines) per genes with amino acid evolution specific to vocal learning birds (white background) and control sets (grey background). Three types of single amino acid variants (SAVs, ConSAVs, and DivSAVs) under positive selection with different stringent statistical levels: Higher likelihoods ($D > 1$) and statical significance (adjusted p value < 0.05) of the alternative model assuming positive selection. Meta data sets include 3 major types of published candidates for target genes of the famous language gene, *FOXP2*, singing induced genes in song nuclei of zebra finch, and differentially expressed genes of song nuclei compared to their surrounding sub-brain regions of zebra finch collected from published sources mentioned in **Supplementary fig. S7.** (see more details in **Materials and Methods**). (**B**) Songbird brain diagram showing the song learning system. Yellow, forebrain song learning brain regions that had their gene expression profiled. Grey, other song learning nuclei. White arrows, connections between the song nuclei. Red-up arrow and blue-down arrow, numbers

of the subset of ConSAV genes, supported by two independent data sources (Lovell et al. 2018; Gedman et al. 2022). (**C**) *B3GNT2* mRNA expression patterns in zebra finch HVC and RA. Image modified from the ZEBrA database (Lovell et al. 2020). (**D**) *DRD1B* mRNA expression pattern in zebra finch Area X and surrounding striatum (St) at three different developmental time points, with specialized expression (white arrows) appearing by adulthood. Image used with permission from Kubikova et al. (Kubikova et al. 2010).

Fig. 7. Enrichments of vocal learner type amino acid patterns in 363 bird species. (**A**) Sequence logos of amino acid profiles at 8 convergent sites in 154 vocal learing species and 209 species considered vocal non-learners. Color codes applied for positively charged residues (K, R, H) as blue, negatively charged residues (D, E) as red, hydrophobic residues (A, F, G, I, L, P, V, W, Y) as green, and the other residues as black. (**B**) Heatmap of vocal learner-type amino acid enrichments for 363 bird species grouped into 22 avian ordinal or higher classification lineages. The color scale shows the proportions of the vocal learner-type amino acid patterns assigned from the representative vocal learner, zebra finch, in each lineage. Numbers in parentheses of each lineage indicate the numbers of species. On the left cladogram, red bold branches indicate the most recent ancestral branches of vocal learning clades.

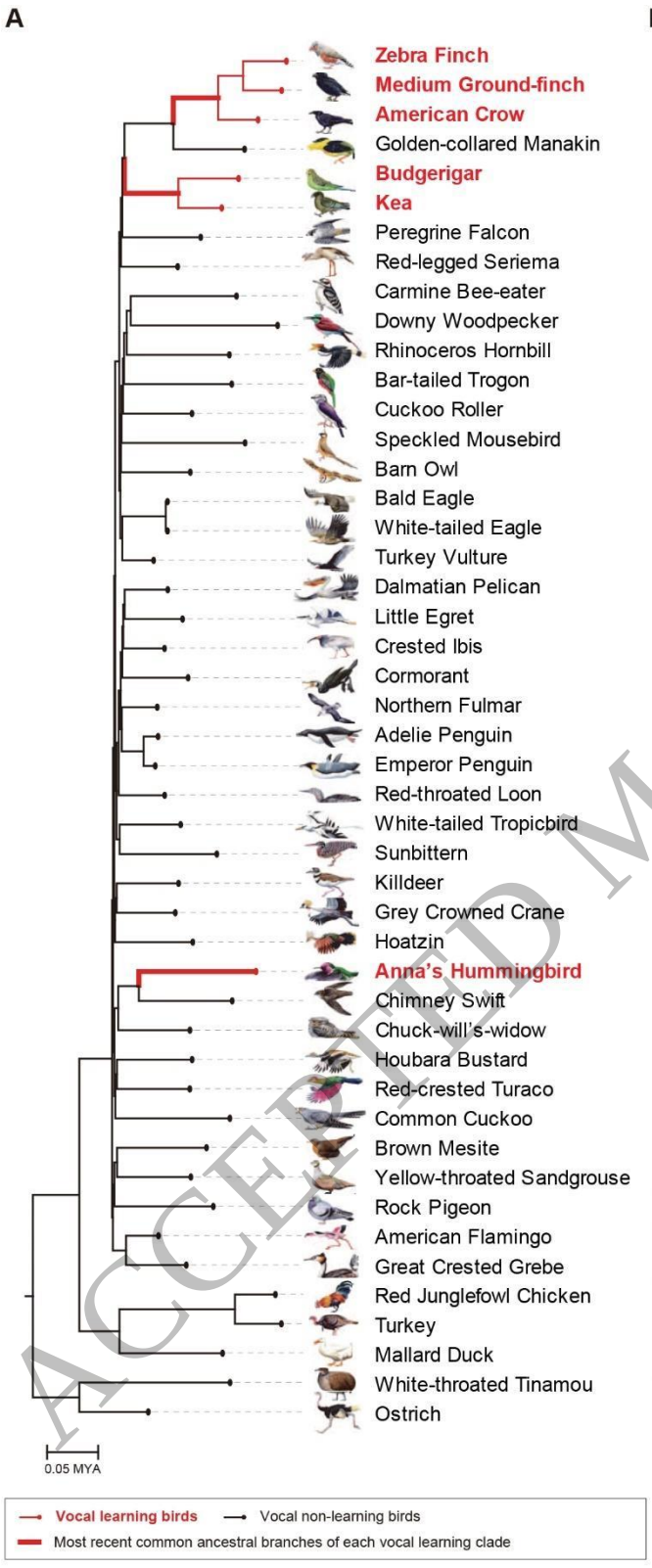


Fig. 1

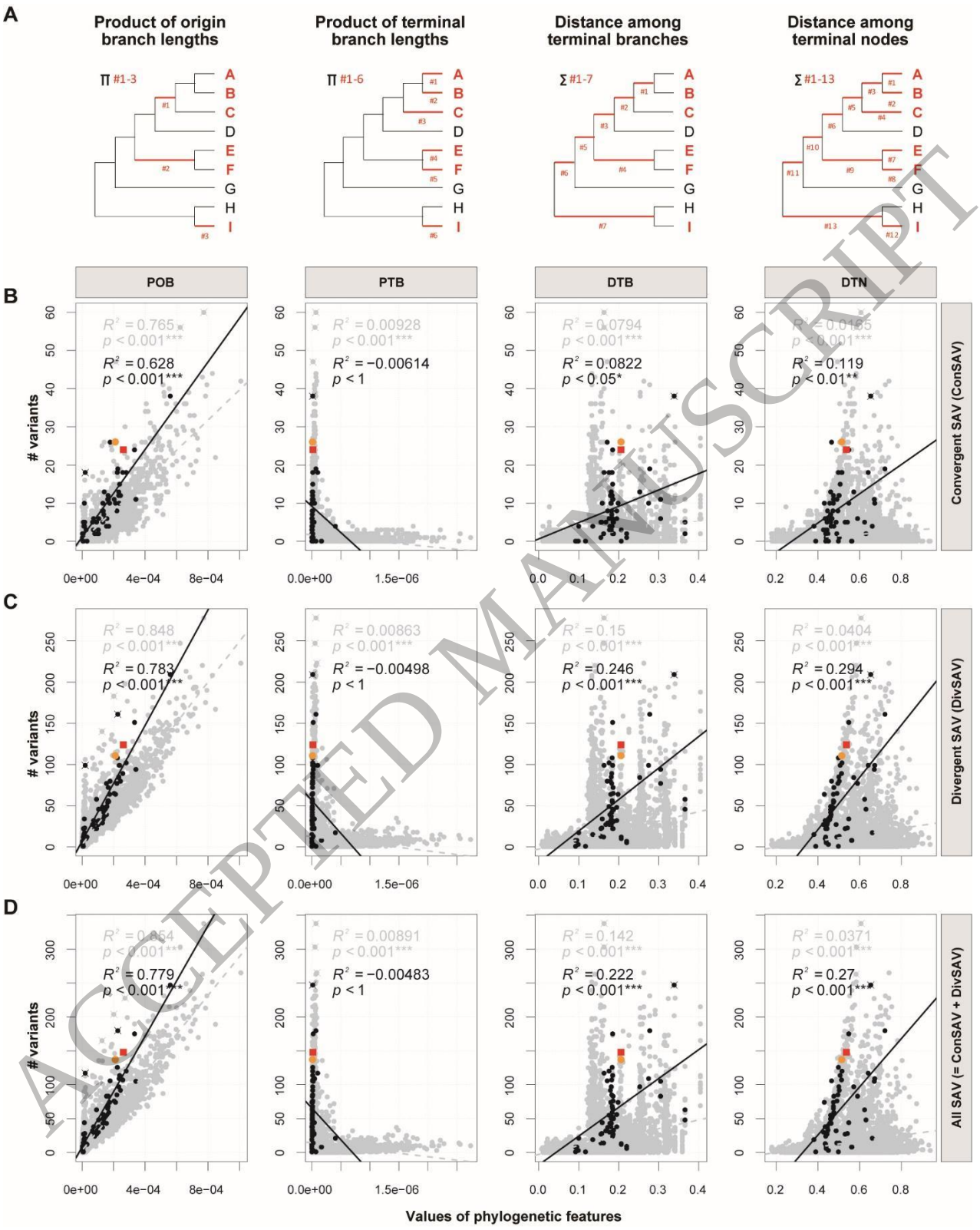


Fig. 2

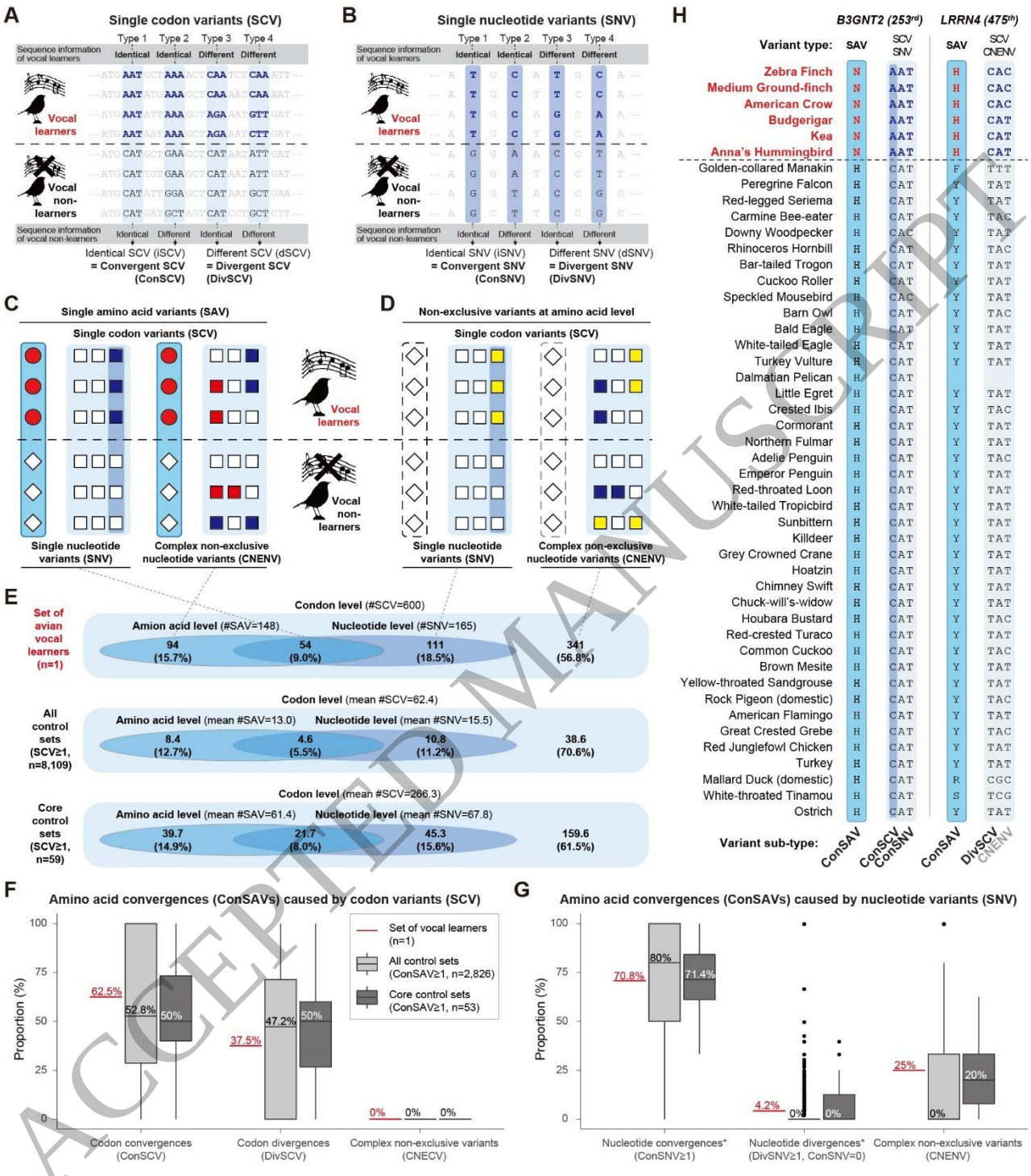


Fig. 3

Fig. 4

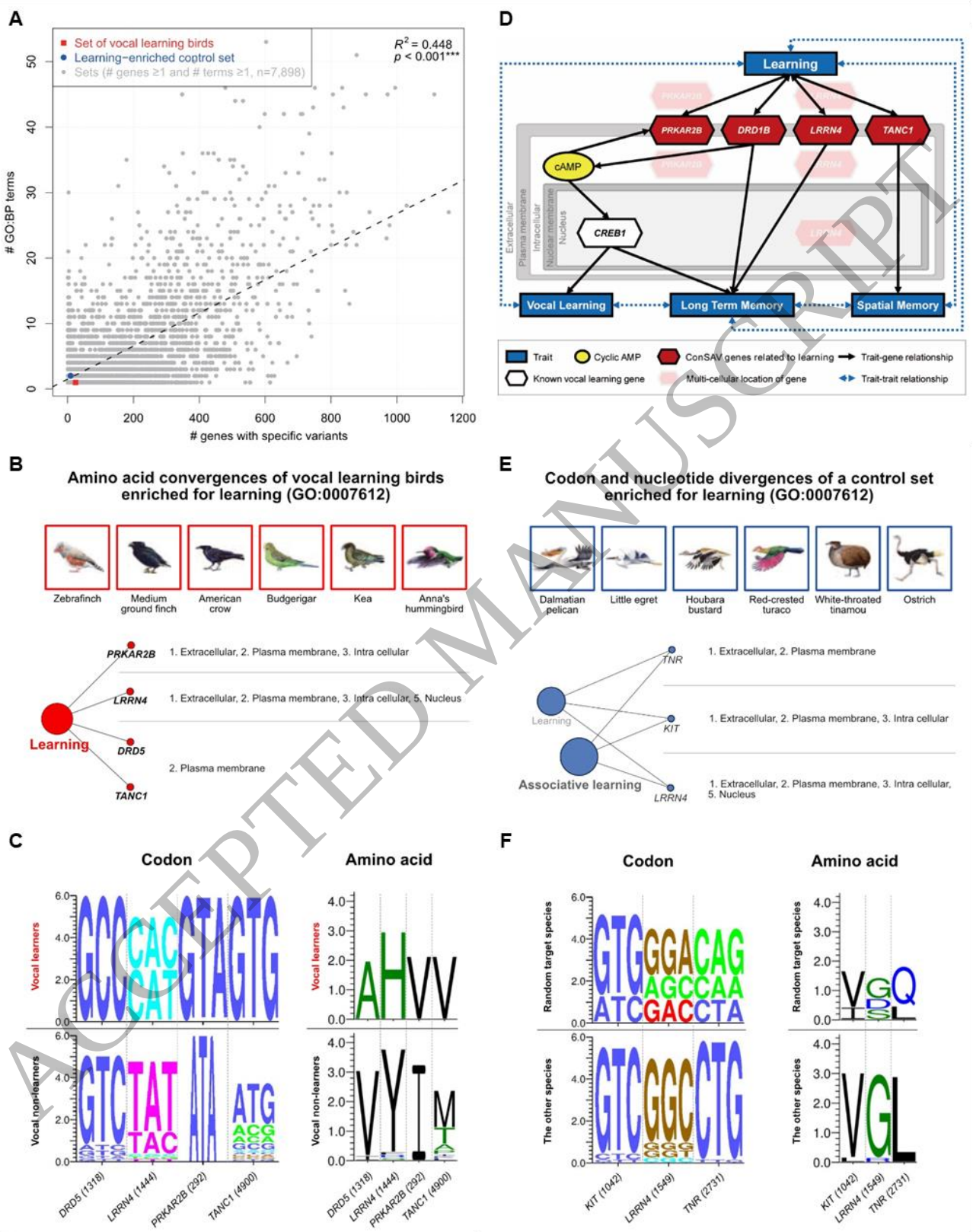


Fig. 5

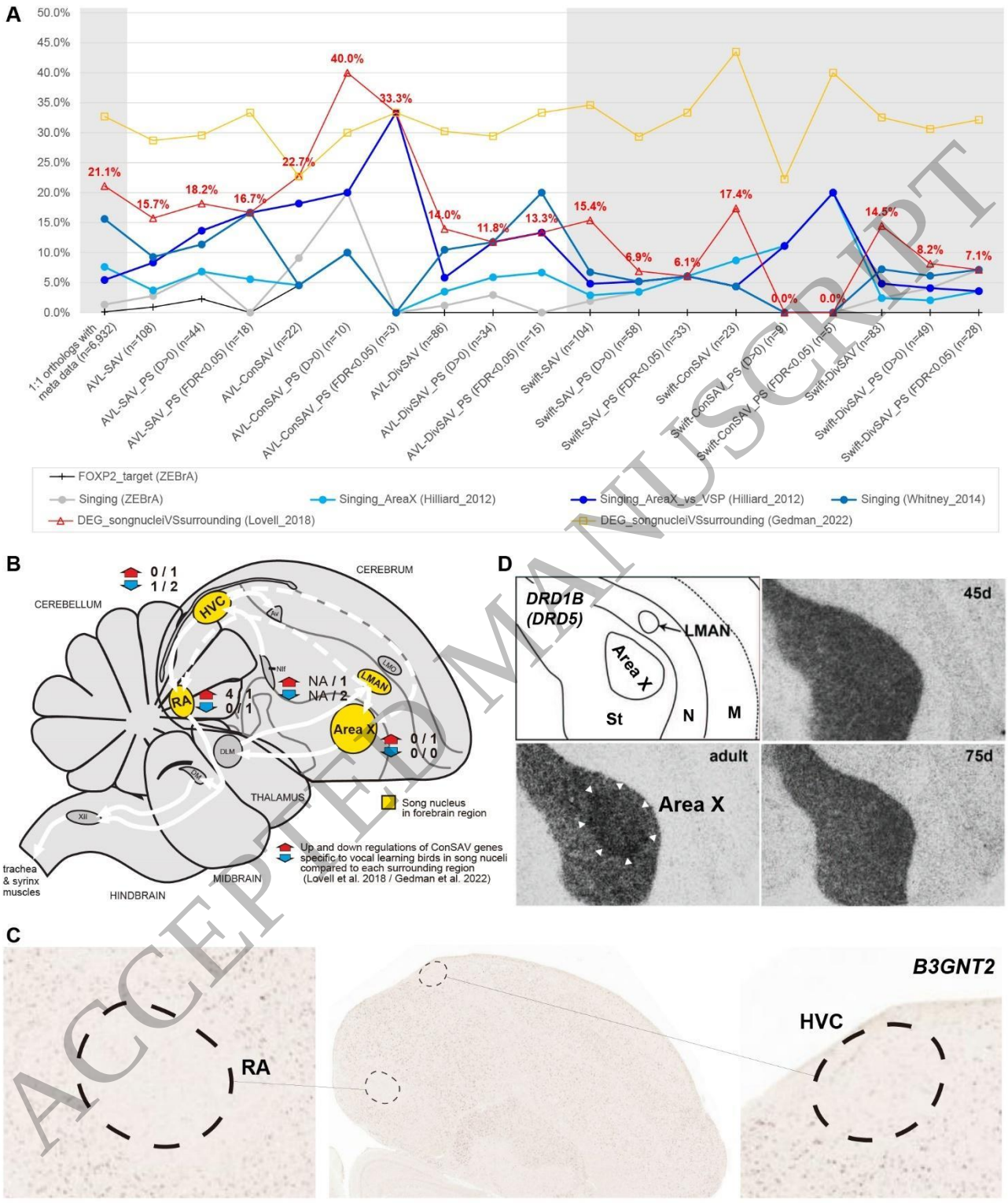
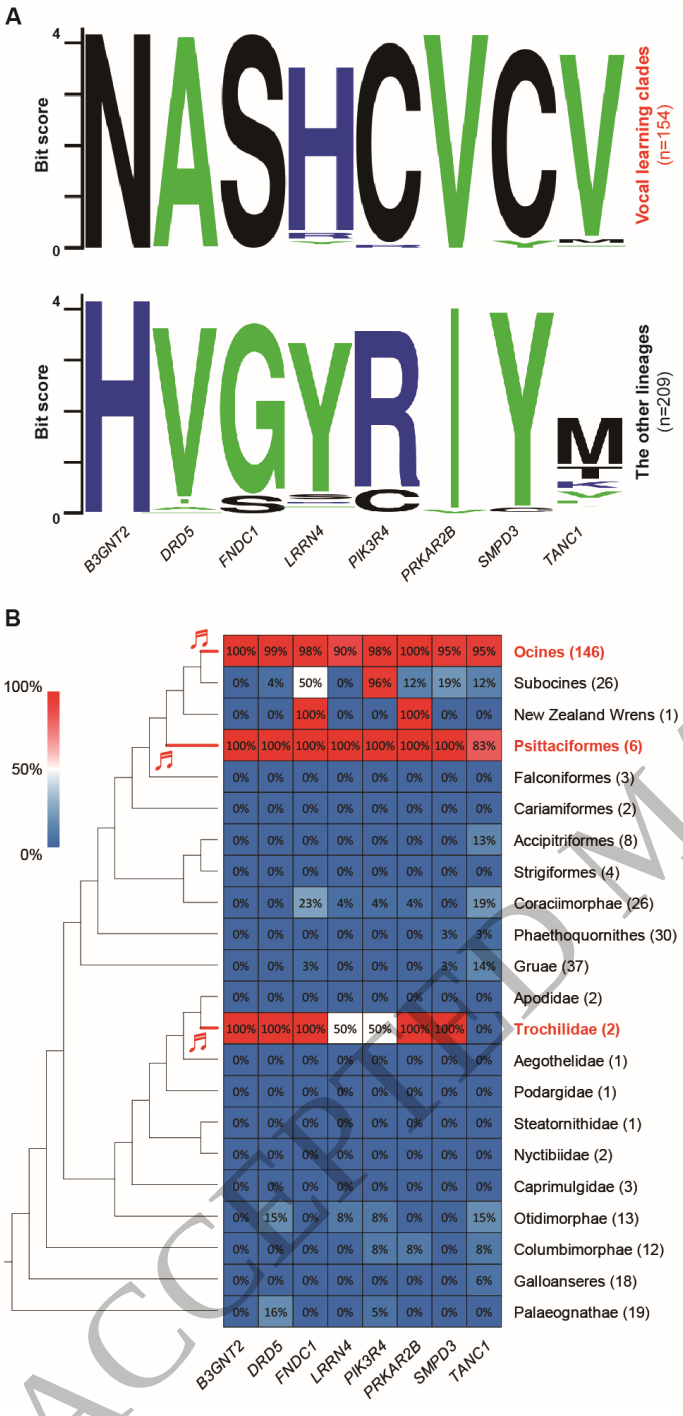


Fig. 6



1

2 **Fig. 7**

References

Abe K, Matsui S, Watanabe D. 2015. Transgenic songbirds with suppressed or enhanced activity of CREB transcription factor. *Proceedings of the National Academy of Sciences* 112:7599–7604.

Arriaga G, Jarvis ED. 2013. Mouse vocal communication system: are ultrasounds learned or innate? *Brain Lang.* [Internet] 124:96–116. Available from: http://dx.doi.org/10.1016/j.bandl.2012.10.002

Arriaga G, Zhou EP, Jarvis ED. 2012. Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds.

Bando T, Sekine K, Kobayashi S, Watabe AM, Rump A, Tanaka M, Suda Y, Kato S, Morikawa Y, Manabe T. 2005. Neuronal leucine-rich repeat protein 4 functions in hippocampus-dependent long-lasting memory. *Mol. Cell. Biol.* 25:4166–4175.

Berwick RC, Friederici AD, Chomsky N, Bolhuis JJ. 2013. Evolution, brain, and the nature of language. *Trends Cogn. Sci.* 17:89–98.

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman W-H, Pagès F, Trajanoski Z, Galon J. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* [Internet] 25:1091–1093. Available from: http://dx.doi.org/10.1093/bioinformatics/btp101

Cahill JA, Armstrong J, Deran A, Khoury CJ, Paten B, Haussler D, Jarvis ED. 2021. Positive selection in noncoding genomic regions of vocal learning birds is associated with genes implicated in vocal learning and speech functions in humans. *Genome Res.* 31:2035–2049.

Carroll SB. 2005. Evolution at two levels: on genes and form. *PLoS Biol.* 3:e245.

Castellucci GA, McGinley MJ, McCormick DA. 2016. Knockout of Foxp2 disrupts vocal development in mice. *Sci. Rep.* 6:23305.

Castoe TA, de Koning APJ, Kim H-M, Gu W, Noonan BP, Naylor G, Jiang ZJ, Parkinson CL, Pollock DD. 2009. Evidence for an ancient adaptive episode of convergent molecular evolution. *Proceedings of the National Academy of Sciences* 106:8986–8991.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540–552.

Chabout J, Sarkar A, Patel SR, Radden T, Dunson DB, Fisher SE, Jarvis ED. 2016. A Foxp2 mutation implicated in human speech deficits alters sequencing of ultrasonic vocalizations in adult male mice. *Front. Behav. Neurosci.* 10:197.

Crooks GE, Hon G, Chandonia J-M, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res.* 14:1188–1190.

Enard W. 2011. FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution.

Curr. Opin. Neurobiol. 21:415–424.

Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R. 2009. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell* 137:961–971.

Enard W, Przeworski M, Fisher SE, Lai CSL, Wiebe V, Kitano T, Monaco AP, Pääbo S. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418:869–872.

Feng S, Stiller J, Deng Y, Armstrong J, Fang QI, Reeve AH, Xie D, Chen G, Guo C, Faircloth BC. 2020. Dense sampling of bird diversity increases power of comparative genomics. *Nature* 587:252–257.

Fischer J, Wegdell F, Trede F, Dal Pesco F, Hammerschmidt K. 2020. Vocal convergence in a multi-level primate society: insights into the evolution of vocal learning. *Proc. Biol. Sci.* [Internet] 287:20202531. Available from: http://dx.doi.org/10.1098/rspb.2020.2531

Fox J, Weisberg S, Adler D, Bates D, Baud-Bovy G, Ellison S, Firth D, Friendly M, Gorjanc G, Graves S. 2012. Package "car." *Vienna: R Foundation for Statistical Computing* 16.

Gedman GL, Biegler MT, Haase B, Wirthlin ME, Fedrigo O, Pfenning AR, Jarvis ED. 2022. Convergent gene expression highlights shared vocal motor microcircuitry in songbirds and humans. *bioRxiv*:2022.07. 01.498177.

Goldstein RA, Pollard ST, Shah SD, Pollock DD. 2015. Nonadaptive amino acid convergence rates decrease over time. *Mol. Biol. Evol.* 32:1373–1381.

Han S, Nam J, Li Y, Kim S, Cho S-H, Cho YS, Choi S-Y, Choi J, Han K, Kim Y. 2010. Regulation of dendritic spines, spatial memory, and embryonic development by the TANC family of PSD-95-interacting proteins. *Journal of Neuroscience* 30:15102–15112.

Hickey G, Paten B, Earl D, Zerbino D, Haussler D. 2013. HAL: a hierarchical format for storing and analyzing multiple genome alignments. *Bioinformatics* 29:1341–1342.

Hilliard AT, Miller JE, Fraley ER, Horvath S, White SA. 2012. Molecular microcircuitry underlies functional specification in a basal ganglia circuit dedicated to vocal learning. *Neuron* 73:537–552.

Hilliard AT, Miller JE, Horvath S, White SA. 2012. Distinct neurogenomic states in basal ganglia subregions relate differently to singing behavior in songbirds. *PLoS Comput. Biol.* 8:e1002773.

Jarvis ED. 2004. Learned birdsong and the neurobiology of human language. *Ann. N. Y. Acad. Sci.* 1016:749–777.

Jarvis ED. 2019. Evolution of vocal learning and spoken language. *Science* 366:50–54.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346:1320–1331.

Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT. 2015. Phylogenomic analyses data of the avian phylogenomics project. *Gigascience* 4:s13742–014 – 0038–1.

Kandel ER. 2012. The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Mol. Brain* 5:1–12.

Kim J, Lee C, Ko BJ, Yoo DA, Won S, Phillippy AM, Fedrigo O, Zhang G, Howe K, Wood J. 2022. False gene and chromosome losses in genome assemblies caused by GC content variation and repeats. *Genome Biol.* 23:204.

Kroodsma DE. 1983. The ecology of avian vocal learning. *Bioscience* [Internet]. Available from: https://academic.oup.com/bioscience/article-abstract/33/3/165/281523

Kubikova L, Wada K, Jarvis ED. 2010. Dopamine receptors in a songbird brain. *J. Comp. Neurol.* 518:741–769.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35:1547.

Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. 2010. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* [Internet] 26:2438–2444. Available from: https://academic.oup.com/bioinformatics/article/26/19/2438/230298

Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–523.

Liu W-C, Wada K, Jarvis ED, Nottebohm F. 2013. Rudimentary substrates for vocal learning in a suboscine. *Nat. Commun.* [Internet] 4:2082. Available from: http://dx.doi.org/10.1038/ncomms3082

Lovell PV, Huizinga NA, Friedrich SR, Wirthlin M, Mello CV. 2018. The constitutive differential transcriptome of a brain circuit for vocal learning. *BMC Genomics* 19:1–27.

Lovell PV, Wirthlin M, Kaser T, Buckner AA, Carleton JB, Snider BR, McHugh AK, Tolpygo A, Mitra PP, Mello CV. 2020. ZEBrA: Zebra finch Expression Brain Atlas—A resource for comparative molecular neuroanatomy and brain evolution studies. *J. Comp. Neurol.* 528:2099–2131.

Madden T. 2003. The BLAST sequence analysis tool. *The NCBI handbook*.

Mason NA, Burns KJ, Tobias JA, Claramunt S, Seddon N, Derryberry EP. 2017. Song evolution, speciation, and vocal learning in passerine birds. *Evolution* 71:786–796.

Naguib M, Riebel K. 2014. Singing in space and time: the biology of birdsong. *Biocommunication of animals*:233–247.

Nassar LR, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee BT. 2023. The UCSC genome browser database: 2023 update. *Nucleic Acids Res.* 51:D1188–D1195.

Nottebohm F. 1972. The origins of vocal learning. *Am. Nat.* 106:116–140.

Nowicki S, Searcy WA. 2014. The evolution of vocal learning. *Curr. Opin. Neurobiol.* 28:48–53.

Petkov CI, Jarvis ED. 2012. Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 4:12.

Pfenning AR, Hara E, Whitney O, Rivas MV, Wang R, Roulhac PL, Howard JT, Wirthlin M, Lovell PV, Ganapathy G. 2014. Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* 346:1256846.

Rangel-Barajas C, Coronel I, Florán B. 2015. Dopamine receptors and neurodegeneration. *Aging Dis.* 6:349.

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47:W191–W198.

R Core Team, R. 2013. R: A language and environment for statistical computing.

Reimers-Kipping S, Hevers W, Pääbo S, Enard W. 2011. Humanized Foxp2 specifically affects cortico-basal ganglia circuits. *Neuroscience* 175:75–84.

Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J. 2021. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592:737–746.

Scharff C, Petri J. 2011. Evo-devo, deep homology and FoxP2: implications for the evolution of speech and language. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366:2124–2140.

Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18:6097–6100.

Schreiweis C, Bornschein U, Burguière E, Kerimoglu C, Schreiter S, Dannemann M, Goyal S, Rea E, French CA, Puliyadi R. 2014. Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences* 111:14253–14258.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13:2498–2504.

Sherry ST, Ward M-H, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 29:308–311.

da Silva WCN, Köhler CC, Radiske A, Cammarota M. 2012. D1/D5 dopamine receptors modulate spatial memory formation. *Neurobiol. Learn. Mem.* 97:271–275.

Solberg R, Sistonen P, Träskelin A-L, Bérubé D, Simard J, Krajči P, Jahnsen T, de la Chapelle A. 1992. Mapping of the regulatory subunits RIβ and RIIβ of cAMP-Dependent protein kinase genes on human chromosome 7. *Genomics* 14:63–69.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.

Sunahara RK, Guan H-C, O'Dowd BF, Seeman P, Laurier LG, Ng G, George SR, Torchia J, Van Tol HHM, Niznik HB. 1991. Cloning of the gene for a human dopamine D5 receptor with higher affinity for dopamine than D1. *Nature* 350:614–619.

Ten Cate C, Fullagar PJ. 2021. Vocal imitations and production learning by Australian musk ducks (Biziura lobata). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* [Internet] 376:20200243. Available from: http://dx.doi.org/10.1098/rstb.2020.0243

Thomas GWC, Hahn MW. 2015. Determining the null model for detecting adaptive convergence from genomic data: a case study using echolocating mammals. *Mol. Biol. Evol.* 32:1232–1236.

Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Künstner A, Searle S, White S, Vilella AJ, Fairley S. 2010. The genome of a songbird. *Nature* 464:757–762.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191.

Whitney O, Pfenning AR, Howard JT, Blatti CA, Liu F, Ward JM, Wang R, Audet J-N, Kellis M, Mukherjee S. 2014. Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science* 346:1256780.

Wirthlin M, Chang EF, Knörnschild M, Krubitzer LA, Mello CV, Miller CT, Pfenning AR, Vernes SC, Tchernichovski O, Yartsev MM. 2019. A Modular Approach to Vocal Learning: Disentangling the Diversity of a Complex Behavioral Trait. *Neuron* [Internet] 104:87–99. Available from: http://dx.doi.org/10.1016/j.neuron.2019.09.036

Wong PCM, Morgan-Short K, Ettlinger M, Zheng J. 2012. Linking neurogenetics and individual differences in language learning: The dopamine hypothesis. *Cortex* 48:1091–1102.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586–1591.

Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L. 2016. Ensembl 2016. *Nucleic Acids Res.* 44:D710–D716.

Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Girón CG. 2018. Ensembl 2018. *Nucleic Acids Res.* 46:D754–D761.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346:1311–1320.

Zou Z, Zhang J. 2015. Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? *Mol. Biol. Evol.* 32:2085–2096.

Acknowledgements

This study was supported by the Marine Biotechnology Program of the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Ocean and Fisheries (MOF) (No. 20180430), Republic of Korea to CL and HK, and by HHMI to EDJ, USA.

Author contributions

CL, SC, KK, HK, and EDJ designed the study. CL and EDJ collected all of raw data for this study. CL, KK, HK, and EDJ developed algorithms and programs to identify convergent variants. CL, HK, and EDJ designed control sets by considering phylogenetic relationships among vocal learning birds, and developed algorithms and programs to investigate phylogenetic factors and molecular sources of amino acid convergences. CL performed most of analyses. CL and DY performed PCA and phylogenetic analysis to estimate the vocal learning ability of Rifleman, and conducted fixed difference analyses for public SNP resources of zebra finch and chicken. JH and HL validated fixed differences of convergent amino acid variants in zebra finch and chicken. CL, MD, GG, JA, EH, MR, OW, ARP, and EDJ obtained and analyzed meta data sets for previous candidate genes related to vocal learning. CL and EDJ wrote the draft paper and all of authors reviewed it. HK and EDJ supervised this study.

Data Availability statement

The genome assemblies and corresponding gene annotations used in this study are available from Zhang et al. (2014), GigaScience, https://doi.org/10.1186/2047-217X-3-26. The multiple sequence alignments

and phylogenetic trees used for comparative analysis are available from Jarvis et al. (2014), GigaScience, https://doi.org/10.1186/s13742-014-0038-1.

The code used to perform the analyses in this study is publicly available in the GitHub repository: https://github.com/chulbioinfo/ConVarFinder. The repository includes detailed documentation and step-by-step instructions to ensure reproducibility. In particular, the input alignments used in this study are available at: https://github.com/chulbioinfo/ConVarFinder/tree/master/paper_analysis/0.rawdata/MSA.