# Transcriptional regulatory divergence underpinning species-specific learned vocalization in songbirds

[鳴禽類(ソングバード)の種特異的な歌行動に関わる遺伝子発現制御の研究]

**A DISSERTATION**

**Submitted to the Graduate School of Life Science,**

**Hokkaido University**

**In partial fulfilment of the requirements for the degree**

**DOCTOR OF LIFE SCIENCE**

**Hongdi Wang**

**March 2020**

# Contents

# General Introduction

Species-specific behaviors are shaped by evolutionary forces to adapt to environment diversity (Crews, 1997; Katz & Harris-Warrick, 1999). As an important factor for species co-occurrence, species-specific behavior contributes to a variety of inter- and intra-specific interactions including communication, sexual selection and predator avoidance (Tinbergen, 1963; Krebs et al., 1978; Lingle, 1992). Species-specific behavior exists in most animal taxa, such as alarming gaits behavior between deer species (Lingle, 1992), pair-banding behavior between vole species (Winslow *et al.*, 1993), and different courtship song behavior between *Drosophila* species (Wheeler *et al.*, 1991). Figuring out the physiological and molecular difference in the neural systems that induced species-specific behavior between closely related animal species can help to understand how the neural mechanisms regulate behavior (Crews, 1997; Katz & Harris-Warrick, 1999).

Evolutionary pressures shape organisms to produce species-specific behavior during speciation, which suggests neuronal changes exist among evolutionally closely related species (Searcy & Andersson, 1986; Crews, 1997). The nervous system incorporates the changes to generate novel species-specific behavior during speciation (Tierney, 1995; Katz & Harris-Warrick, 1999). The differences in structure and development of neural circuits among evolutionally closely-related species are the basis for species-specific behavior formation. Such structural and developmental differences can be caused by the genetic difference in closely related species (Young *et al.*, 1997; Lim *et al.*, 2004; Reaume & Sokolowski, 2011; Bendesky *et al.*, 2017; Chandra *et al.*, 2018).

With development and application in the life science of genomics, finding out the genetic differences which induced species-specific behavior between closely related species is possible to explore the mechanisms of animal behavior on the molecular level (Reaume & Sokolowski, 2011; Bengston *et al.*, 2018). Genomic researches of species-specific behavior have been performed in some animal species (Ahmadiyeh *et al.*, 2005; Weber *et al.*, 2013; Ding *et al.*, 2016; Bendesky *et al.*, 2017). For example, the sin and pulse song, two features of species-specific courtship song of *Drosophila simulans* and *D. mauritiana*, were correlated with different genomic locus (Ding *et al.*, 2016). In addition, quantitative trait locus (QTL) mapping found that species-specific nest digging behaviors of two closely related mice species, *Peromyscus polionotus* and *P. maniculatus*, were highly associated with a few genome regions (Weber *et al.*, 2013). These two studies narrow down the candidate genomic regions by using backcross hybrid offspring between closely related species. F1 hybrids between closely related species are also useful to select candidate genomic regions related to species-specific behavior. For instance, genes related with parental nest building behavior were figured out by using biased allele expressed genes in F1 hybrid comparing with QTL mapping and species differentially expressed genes (Bendesky *et al.*, 2017). Furthermore, the allelic expression ratio in the F1 hybrid and expression ratio between their parental species can be used to evaluate the transcription regulatory difference between closely related species (Wittkopp *et al.*, 2004; Signor & Nuzhdin, 2018). These studies suggested that genomic analysis utilizing hybrid offspring is powerful to explore genes/genome alleles which contributed to species-specific behavior formation.

The structural and developmental differences in the neural circuits between species are mainly caused by the expression difference of orthologous genes in conserved neural circuits (Seeholzer *et al.*, 2018; Ding *et al.*, 2019). Previous research found that different expression levels of orthologous genes between closely related animal species induced species-specific behavior. For instance, the vasopressin 1a receptor (V1aR) gene show different expression levels in the ventral pallidum which induced species difference in pair-bonding behavior between monogamous (*Microtus ochrogaster*) and polygamous (*M. pennsylvanicus*) vole species (Insel *et al.*, 1994; Young *et al.*, 1997; Young & Wang, 2004). Further research found that the different length of microsatellites at the upstream of the V1aR changes the gene expression level, which in turn generates different behavioural traits between the two vole species (Hammock & Young, 2005). These issues suggest that gene transcriptional regulatory difference which induced species different expression patterns of orthologous genes potentially contributes to the generation of species-specific behavior. Gene expression is driven by transcriptional regulatory elements which mainly contain two kinds, *cis*- and *trans*-regulatory elements (Wray *et al.*, 2003; Wittkopp, 2005; Bryois *et al.*, 2014). Transcriptional regulatory divergences of closely related species could induce the expression level difference of orthologous genes in the brain (Wittkopp *et al.*, 2004; Carroll, 2005; Signor & Nuzhdin, 2018). Orthologous genes in closely related species are inherited from their ancestral species, while *cis*- and/or *trans*- elements may change during evolution among species (**Fig. I-1A**). Such mutations can affect transcription initiation, transcription rate, and transcript stability which induced species-specific

expression patterns (Carroll, 2005; Wittkopp *et al.*, 2008; Signor & Nuzhdin, 2018). Mutations such as single nucleotide polymorphism (SNP) and insert and deletion (indels) in the *cis*- sequence could induce expression level difference of orthologous genes (Wray *et al.*, 2003; Wittkopp *et al.*, 2004). *Trans*-regulatory elements are genes that regulate the expression of distant genes through an intermolecular interaction. The different expression levels or amino acid substitutions of *trans*-regulators can drive different expression levels of their downstream genes between closely related species (Wray *et al.*, 2003; Wittkopp *et al.*, 2004). However, little is known about how transcriptional regulatory divergence contributes to the generation of species-specific behavior, especially for learned behavior, by changing the expression level of orthologous genes.

There are more than 4,000 songbird species in the world and they produce species unique songs which play important roles for mating interaction and habitat declaring (Marler & Slabbekoorn, 2004; Brenowitz & Beecher, 2005). Species-specific birdsongs are complex vocal signals acquired through vocal learning, which makes songbird species become an excellent model to explore the mechanism for vocal learning behavior (Jarvis, 2004; Mori & Wada, 2015b). The species-specific songs as an important characteristic of speciation raised along with the evolution of songbird species (Jarvis, 2004). The species-specific songs are developed and produced by conserved neural circuits called the song pathway in the brain of songbirds (**Fig. I-1B**). The song pathway consists of two major neural circuits: the anterior forebrain pathway (AFP) and the vocal motor pathway (Nottebohm *et al.*, 1976; Scharff & Nottebohm,

1991; Zeigler & Marler, 2008). There are two features of birdsong, the sound characteristics of song elements (syllables) featured as “syllable acoustic”, and the arrangement of each song syllable order featured as “syllable sequence”. The AFP contributes to song learning and vocal plasticity during song development, but it is not crucial for the production of the bird’s learned song (Bottjer *et al.*, 1984; Scharff & Nottebohm, 1991). Species-specific songs of adult songbirds are produced by the vocal motor circuit of the song pathway. The robust nucleus of arcopallium (RA) and the song nuclei HVC (as a proper name) in the vocal motor circuit control the syllable acoustic and sequence features of birdsong, respectively (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Sober *et al.*, 2008). Both acoustic and sequence of the birdsong show species-specific differences, which suggests that the structure and physiological activity of RA and HVC are different among songbird species. That further meant these two nuclei could be important for determining species-specific song traits of adult songbirds. However, how the transcriptional regulatory differences of orthologous genes in the vocal motor pathway induce species differential expression which driven species-specific songs among songbird species is still unknown. Concerning this question, I hypothesized that gene expression difference induced by *cis*- and/or *trans*- changes in HVC and RA among songbird species is one of the potential reasons for the generation of species-specific birdsong. Identification of the transcriptional regulatory divergence which induced gene expression difference in the motor pathway can enhance our understanding of the molecular mechanisms for species-specific vocal learning behavior.

Based on the above information, I predicted that orthologous genes with different expression levels are regulated by *cis*- and/or *trans*-regulatory divergences among songbird species in HVC and RA. To verify this hypothesis, I used zebra finch (ZF; *Taeniopygia guttata*), owl finch (OF; *T. bichenovii*) and their reciprocal first-generation (F1) hybrids to analyse gene expression difference and regulatory divergence.

In **Chapter I**, I examined differences in song phenotypes between ZF and OF. To test whether the song difference between ZF and OF was genetically constrained, I analysed the vocal learning behavior of ZFs and OFs which were tutored by conspecific and cross-specific songs. To identify genes with potential contribution to the generation of species-specific songs, I analysed gene expression differences between ZF and OF in both HVC and RA.

In **Chapter II**, I analyzed the expression ratio between parental species (ZFs and OFs) of all detectable genes and the allelic expression ratio in the F1 hybrids. Depending on these two sets of values, I identified *cis*- and/or *trans*-regulatory difference between ZF and OF. To explore the biological function of genes with *cis*- and/or *trans*-regulatory divergences, I performed function analysis of genes belongs to all categories of regulatory divergence.

In **Chapter III**, depending on the candidate upstream regulators of *trans*-regulated different genes in RA, I performed a correlation analysis in F1 hybrids between parameters of acoustic and sequence features and ASE, or expression level of the most significant candidate *trans*-mediator, BDNF. To test the expression level of BDNF

maintaining species-specific songs, I further performed pharmacological over-activation of BDNF receptor and RNA-seq (RNA sequencing) of the song nucleus RA of adult ZFs.

Based on the results obtained through a set of experiments, I will discuss the potential contribution of *cis*- and/or *trans*-regulatory differences for the generation of species-specific vocal learning behavior among songbird species.

# Chapter I

# Species difference in song phenotypes and gene expression in the motor song nuclei between zebra finch and owl finch

## 1.1 introduction

Different expression of orthologous genes is thought to underlie phenotypic differences between species, which allows us to study how the genomic mutation affect the evolution of closely related species. The evolution of distinct phenotypic traits between different taxa, such as anatomical, physiological, and behavioral characteristics are induced by the co-function of different expressions of orthologues genes (Rifkin *et al.*, 2003; Brawand *et al.*, 2011). Past studies reported that gene expression differences between closely related species in the nervous system caused species-specific behavior (Insel & Shapiro, 1992; Young *et al.*, 1997; Bendesky *et al.*, 2017; Tamvacakis *et al.*, 2018). Identifying such genes showing differential expression in the nervous system between closely related species allows us to figure out which genes might regulate species-specific behavior.

Like human, songbirds possess vocal learning ability to acquire vocalizations through imitation, not as instinct development (Jarvis, 2004; Marler & Slabbekoorn, 2004). The analogous neural pathway for human speech and birdsong made songbirds excellent model to explore the physiological and molecular mechanisms of vocal learning behavior (Jarvis, 2004; Pfenning *et al.*, 2014; Mori & Wada, 2015b). More than 4,000 songbird species produce species-specific songs by conserved neural circuits, called the song pathway (Marler & Slabbekoorn, 2004; Brenowitz & Beecher, 2005). Candidate genes for regulation of species-specific vocal learning behavior can be found by analyzing genes with different expression in the song pathway between closely related species. Zebra finch (ZF; *Taeniopygia guttata*) and owl finch (OF; *T. bichenovii*)

are two closely-related species of oscine songbirds which were diverged about 6.5 million years ago (Hooper & Price, 2015). These two species share overlapping habitats in the north and west of Australia (Immelmann & Cayley, 1982; Forshaw & Shephard, 2012). Songs of these two species show species-specific differences in both syllable acoustics and sequence which make them be good candidates for the research of species-specific vocal learning. A previous study explored an expression difference in muscarinic acetylcholine receptor 2 (CHRM2) in a motor song nucleus HVC of the song pathway among songbird species including ZF and OF (Asogwa *et al.*, 2018). However, no research has been performed about the gene expression difference between songbird species at global transcriptome level. In the song pathway, HVC and RA song nuclei of the motor pathway contribute to regulating learned song patterns of adult songbirds (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Sober *et al.*, 2008). Figuring out the different expression of orthologous genes in HVC and RA between ZF and OF is helpful to understand the molecular mechanism of species-specific vocal learning behavior.

In this chapter, I first compared the song phenotype between ZF and OF which were tutored with conspecific and cross-specific songs to test whether species-specific songs affected by genetically constrained or not. To explore how many genes showing different expression between ZF and OF, I performed transcriptome analysis between ZF and OF in HVC and RA.

## 1.2 Materials and Methods

### *Animals*

Zebra finches (ZF; *T. guttata*) and owl finches (OF; *T. bichenovii*) were obtained from our breeding colony at Hokkaido University and local breeders. Reciprocal F1 hybrids were bred by pairing ZF and OF in our lab at Hokkaido University (**Fig. I-2**). All birds were constantly maintained with food and water available *ad libitum* under 13/11 light/dark photoperiod. The sex of chicks was determined by previously reported method (Wada *et al.*, 2006; Soderstrom *et al.*, 2007). Chicks of ZF and F1 hybrids were fed by both parents until 10-20 dph (days post hatching). Males of parents were removed to prevent chicks listening father's song. Female of parents and chicks were moved in a sound-attenuation box and mothers were moved out until chicks fledged (around 35 dph). OF chicks were hand-raised after hatching until they could feed themselves (30–40 dph). After fledging, juveniles were subsequently housed in individual isolation boxes and then individually housed in a sound-attenuating box containing a mirror to reduce social isolation (Gallup & Capper, 1970). Animal experiments were performed following the guidelines of the Committee on Animal Experiments of Hokkaido University which based on national regulations for animal welfare in Japan (Law for the Humane Treatment and Management of Animals; after partial amendment number 105, 2011).

***Song recording and analysis***

Male juveniles were isolated in sound-attenuation boxes (cage size: 37x42x44 cm; box size: 44x53x55 cm) after fledging (36-48 dph). Bird songs were acoustically recorded and saved by microphone (SM57, Shure) connected computer with Sound Analysis Pro software (Tchernichovski *et al.*, 2000). ZF and OF chicks were tutored by conspecific or cross-specific songs depending on the purpose. F1 hybrids were tutored by combined ZF and OF songs. Tutor songs were played back 10 times a day (5/5 morning/afternoon) at 55-75 dB from a speaker (SRS-M30, SONY) passively controlled by Sound Analysis Pro software.

Acoustic and sequence features of ZF and OF songs analysis were performed. For acoustic score feature analysis, 500 syllables were randomly selected from the song of ZF and OF individuals. A total of 10 acoustic features were measured which contain syllable duration, inter-syllable gap duration, mean pitch, pitch goodness, Wiener entropy, entropy variance, mean amplitude modulation (AM), AM variance, mean frequency modulation (FM), and FM variance (Tchernichovski *et al.*, 2000). Statistical analysis was performed for these 10 acoustic features between ZF (n = 6 each, conspecific tutored; n = 4 each, cross-specific tutored) and OF (n = 6 each, conspecific tutored; n = 3 each, cross-specific tutored) songs by one-way ANOVA. For sequence feature analysis, we performed syllable similarity matrices (SSM) analysis depending on the published method (Imai *et al.*, 2016). Motif pattern and repetition pattern of ZF (n = 6 each, conspecific tutored; n = 4 each, cross-specific tutored) and OF (n = 6 each,

conspecific tutored; n = 3 each, cross-specific tutored) songs were calculated and performed statistical analysis by one-way ANOVA.

***Brain tissue sample and RNA extraction for RNA-seq***

*Whole brain sampling.* Adult ZF and OF male individuals (ZF: n=4, 234-786 dph; OF: n=4, >180 dph) were isolated in sound-attenuation boxes more than one day before sacrifice. All birds were sacrificed under silent and dark condition. Whole brains were frozen on dry ice immediately and stored at -80°C until RNA extraction. Total RNA was isolated using TRIzol Reagent according to manufacturer's protocol (Invitrogen) and were treated with RNase-free DNase.

*Laser capture microdissection (LCM) of HVC and RA.* Adult ZF, OF and F1 hybrid individuals (ZF: n=4, 138-305 dph; OF: n=4, >180 dph; ZO: n=4, 152-174 dph; OZ: n=4, 150-279 dph) were isolated in sound-attenuation boxes and sacrificed under silence and dark condition. Whole brains were removed into Tissue-Tek OCT and rapidly frozen on dry ice, then stored at -80°C until section. Brain section and LCM were performed by the previously reported method (Mori & Wada, 2015a). Briefly, the whole brains were serially cryosectioned with 14 μm thickness onto handmade membrane slides and performed Nissl staining to confirm the presence and boundaries of HVC and RA. HVC and RA were microdissected using a laser capture microscope (Arc-turusXT; Arcturus Bioscience) with setting parameters as follows: spot diameter, 100 μm; laser power, 80 mW; and laser duration, 80 ms. The captured tissues were dissolved into RLT buffer (Qiagen) with β-mercaptoethanol (Wako) and then stored at

-80 °C until RNA extraction. Total RNA was purified using a column-based method (RNeasy Micro kit; Qiagen) and treated with DNase in the column to avoid contamination of genomic DNA.

***RNA-seq library construction and sequencing***

RNA integrity number (RIN) and concentration of whole brain total RNA were measured by Bioanalyzer 2100 (Agilent Technologies) to make sure the high quality of total RNA used for library construction (RIN: 9.3-9.6). After treating with RNase-free DNase, the total RNA was used for library construction by using TruSeq DNA Sample Prep Kits (Illumina). All libraries were sequenced using Illumina Hiseq2500 platform for 150 bp paired ends sequencing.

For HVC and RA tissues, RIN and concentration were measured by Bioanalyzer 2100 to guarantee the RNA quality (RIN: 5.2-8.1). Purified total RNA (1-2 ng) of RA and HVC were used to synthesize first strand cDNA by previously described method (Sasagawa *et al.*, 2013). The PCR amplification condition is: 14 cycles of 98°C (10 s), 65°C (15 s), 68°C (5 min). Amplified cDNA samples were purified by PCR purification column (MiniElute PCR Purification Kit; Qiagen). The concentrations and smearing patterns of cDNA samples were checked by Bioanalyzer 2100 to make sure the quality of cDNA. Amplified cDNA samples were fragmented using a DNA Shearing System LE220 (Covaris) and then purified by column. After the end repair of DNA fragments, adaptors were ligated and amplified using a ligation based Illumina multiplex library preparation method (LIMprep) with a KAPA Hyper Prep Kit (Nippon genetics) and 10

PCR cycles. All libraries were sequenced by using Illiumina Hiseq2500 platform for 150 bp paired-end sequencing. The library construction and sequencing steps were performed at Suzuki laboratory in the Tokyo University as a support by the Grant-in-Aid for Scientific Research on Innovative Areas 'Genome Science'.

All NGS data of whole brain, HVC and RA samples were updated to DDBJ Sequence Read Archive (DRA005548 and DRA002970).

***Read mapping and quantification of gene expression level***

Reads with low quality and the adaptor sequence in the NGS raw data of all HVC and RA samples were removed by using Filter FASTQC (Simon, 2010). Clean reads were mapped to reconstructed ZF genome by TopHat2 (Kim *et al.*, 2013). Cufflinks software (Trapnell *et al.*, 2010) was used to evaluate expression levels of each gene by calculating RPKM (Reads Per Kilobase per Million) value of HVC and RA samples of ZF and OF. During RPKM calculation, new genome annotation GTF (Gene Transfer Format) file of ZF which improved by Hayase *et al.* (Taeniopygia_guttata taeGut3.2.4. .76.gtf) was used (Hayase *et al.*, 2018). After getting the RPKM value of ZF and OF individuals, species-differently expressed (SDE) genes between ZF and OF were identified by using R package DEseq2 (Love *et al.*, 2014) (adjusted $p$-value < 0.05, Benjamini-Hochberg's procedure).

## 1.3 Results

***Species-specific song phenotypic differences between ZF and OF were genetically constrained***

First, I compared the song features of ZF and OF reared with conspecific song tutoring in our breeding colony to confirm whether a laboratory-controlled environment could maintain species-specific song features. I compared the songs of the two species regarding syllable acoustics and sequential features (12 parameters; See the Materials and methods) at the adult stage (**Fig. I-3A**) and identified significant differences in six acoustic syllable parameters (i.e., syllable duration, inter-syllable gap duration, entropy variance, amplitude modulation [AM] variance, mean frequency modulation [FM], and FM variance) and in syllable sequence features (motif and repetition transition rates) (n = 6 birds each, $p < 0.01$, one-way ANOVA) (**Fig. I-3B, 3C, and Fig. I-4A**). The results showed that the range, but not the pattern, of each acoustic feature's distribution overlapped between ZFs and OFs (3,000 syllables from n = 6 birds each and 500 syllables/bird) (**Fig. I-4A**), thus suggesting that the species differences in the syllable acoustics were not caused by physical species-specific constraints in the peripheral vocal organs.

To examine the genetic contribution on the regulation of species-specific song features of ZF and OF, I performed cross-species song tutoring experiments for both ZF and OF juveniles. Under the cross-species song tutoring condition, juveniles heard only the counter-species songs as tutor songs. By comparing with songs from

conspecific and cross-species song tutoring conditions, I found that song tutoring effects on most of the song parameters, including syllable sequence and acoustics (entropy variance, AM variance, mean FM, and FM variance). However, excluding AM variance, other all song parameters retained species specificity (ZF n = 4, OF n = 3; one-way ANOVA, $p < 0.05$) (**Fig. I-3B, 3C**). In line with this result, I performed principal component analysis (PCA) to investigate the song feature distribution of conspecific and cross-species song tutored birds by reducing the dimensionality of the syllable acoustics and sequential features. I observed that clusters were separable by species but not by song tutoring conditions (**Fig. I-4B**). As many studies in songbirds reported (Marler & Peters, 1977; Eales, 1985; Clayton, 1989), these results also indicate that song learning of these two species is implemented based on the species-specific genetic constraint.

***Species differentially expressed (SDE) genes in both HVC and RA***

To investigate the genetic reason inducing species-specific songs, I analyzed gene expression levels in HVC and RA between ZF and OF. Clean reads (33.5 – 47.0 Mb) were mapped to the ZF reconstructed genome and Reads Per Kilobase per Million mapped reads (RPKM) was calculated to estimate the gene expression level. As a result, 11,501 and 11,487 genes in HVC and RA were identified as genes with detectable expression levels in either ZF or OF (average RPKM $\geqslant$ 1). Statistical analysis was performed to compare the expression level between ZF (n = 4) and OF (n = 4) in HVC and RA. A total of 333 and 374 genes showed significantly differential expression in

HVC and RA, respectively, between ZF and OF ($p$-value adjusted by the Benjamini-Hochberg method; $p < 0.05$) (**Fig. I-5A**). These results showed expression differences of orthologous genes existed in song nuclei of the motor pathway, which could be the potential reason for the generation of species-specific songs between ZF and OF.

Sequence and acoustic features of songs were controlled by HVC and RA, respectively (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Sober *et al.*, 2008). To investigate how many species differential expressed (SDE) genes were uniquely expressed in HVC and RA, I compared the SDE genes between HVC and RA. Only 123 (ZF expressed higher) and 51 (OF expressed higher) SDE genes were shared by HVC and RA, respectively (**Fig. I-5B**). There were more than 65% of SDE genes exist alone in HVC or RA, which showed the expression of such SDE genes were also regulated in a brain region specific manner.

## 1.4 Discussion

Species-specific animal behavior is one of important phenotypic traits between closely related species during speciation, which play roles for reproduction and habitat use (Krebs *et al.*, 1978; Searcy & Andersson, 1986). Past researches had found different expression levels and patterns of orthologous genes in neural system induced species-specific behavior between closely related species (Young *et al.*, 1997; Seeholzer *et al.*, 2018; Ding *et al.*, 2019). In this chapter, I analyzed the vocal learning behavior of two closely related songbird species, ZF and OF, and found that their songs were significantly different both in acoustic and sequence features. To explore the genetic reason of the species-specific vocal learning behavior, I explored gene expression differences in song nuclei of the motor pathway and found 333 and 374 SDE genes in HVC and RA, respectively.

ZF and OF juveniles tutored by conspecific songs showed significant differences in acoustic and sequence features of songs, as well as tutored by cross-specific songs (**Fig. I-3**). These results were in line with previous research (Marler & Peters, 1977; Eales, 1987; Clayton, 1989; Nelson, 2000; Fehér *et al.*, 2009) and suggested that vocal learning behavior of ZF and OF were genetically constrained. Male songbirds display territorial and courtship behavior by singing songs in nature (Sossinka & Böhner, 1980; Williams, 2004; Fujita *et al.*, 2011). Species-specific features of the birdsong are important signals for intra- and inter-species identification of songbird species during the territorial and courtship behavior in the wild (Krebs *et al.*, 1978; Williams, 2004; Fujita *et al.*, 2011). ZF and OF shared overlapping inhabits in the north and west of

Australia (Forshaw & Shephard, 2012), which means juveniles of ZF and OF can hear songs of adult males of both species during song development. The genetic constraint of species-specific vocal learning behavior of juvenile ZFs and OFs is benefited to learn the conspecific song during song development avoiding heterospecific song learning.

One interesting finding in this chapter is that more than half of SDE genes between ZF and OF were not shared by HVC and RA (**Fig. I-5**). This result suggested that the SDE genes showed brain region-specific expression in the song pathway. In the motor pathway, HVC and RA contribute to regulation of syllable sequence and acoustic of songs, respectively (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Marler & Slabbekoorn, 2004; Sober *et al.*, 2008), which may be caused by SDE genes showing brain region-specific expression. There are multiple subpopulations of neurons in the HVC (Hahnloser *et al.*, 2002; Kozhevnikov & Fee, 2007) and RA (Spiro *et al.*, 1999; Leonardo & Fee, 2005). In addition, previous research found that different cell types in the brain of mammals showed different gene expression patterns (Raff *et al.*, 1979; Zeisel *et al.*, 2015; Rosenberg *et al.*, 2018). In my experiment, however, I used LCM to sample the HVC and RA tissues, which mixed all subpopulations of the neurons together to perform RNA-seq. The technique limitation of LMC led to losing cell type information of SDE genes. In the future research, performing single-cell RNA-sequencing (scRNA-seq) and figuring out SDE genes between ZF and OF in each cell type of HVC and RA may help to explain why SDE genes show brain region-specific expression. The results of scRNA-seq will also be helpful to understand the molecular mechanism of species-specific songs on single cell level.

## 1.5 Figures

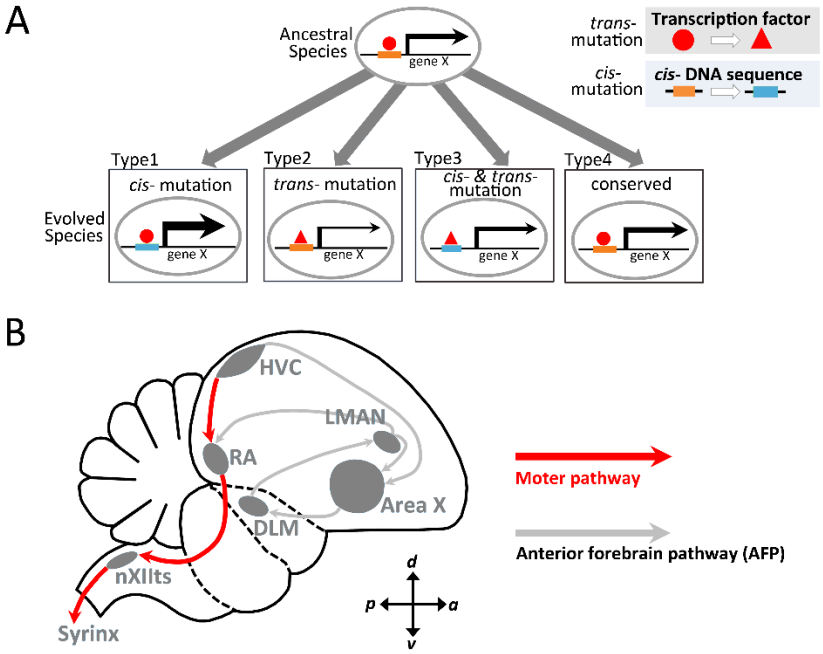



**Figure I-1: C*is*- and/or *trans*-regulatory changes during evolution in song pathway**

**(A)** During evolution, transcriptional regulatory changes between closely-related species. Red circles and triangles: orthologous transcription factors in ancestral and evolved species; Blue and orange squares: the *cis*-sequence of orthologous gene in ancestral and evolved species; Black arrows: different thickness to show different expression levels of orthologous genes in ancestral and evolved species.

**(B)** Schematic showing the song pathway for vocal learning and production in the songbird brain. The posterior motor pathway and the anterior forebrain pathway (AFP) are represented as red and gray lines, respectively. HVC (used as a proper name); RA, the robust nucleus of the arcopallium; Area X, Area X of the basal ganglia; DLM, dorsal lateral nucleus of the medial thalamus; LMAN, lateral magnocellular nucleus of the anterior nidopallium; nXIIts, tracheosyringeal part of the hypoglossal nucleus. Axes indicate brain orientation: *a* = anterior, *p* = posterior, *d* = dorsal, *v* = ventral.

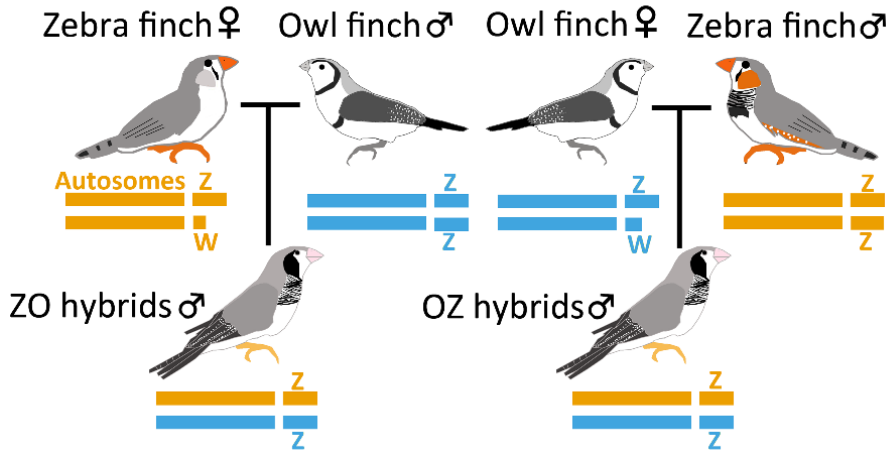



**Figure I-2**: **Reciprocal F1 hybrid between ZF and OF and genomic construction.**

F1 hybrids were bred between zebra finch (ZF) and owl finch (OF). Squares under birds represent genome composition of reciprocal F1 hybrids between ZF (orange) and OF (blue). ZO represents F1 hybrid offspring between ZF♀ x OF♂. OZ hybrids are the opposite. Male F1 hybrids share identical sets of auto- and sex-chromosomes.

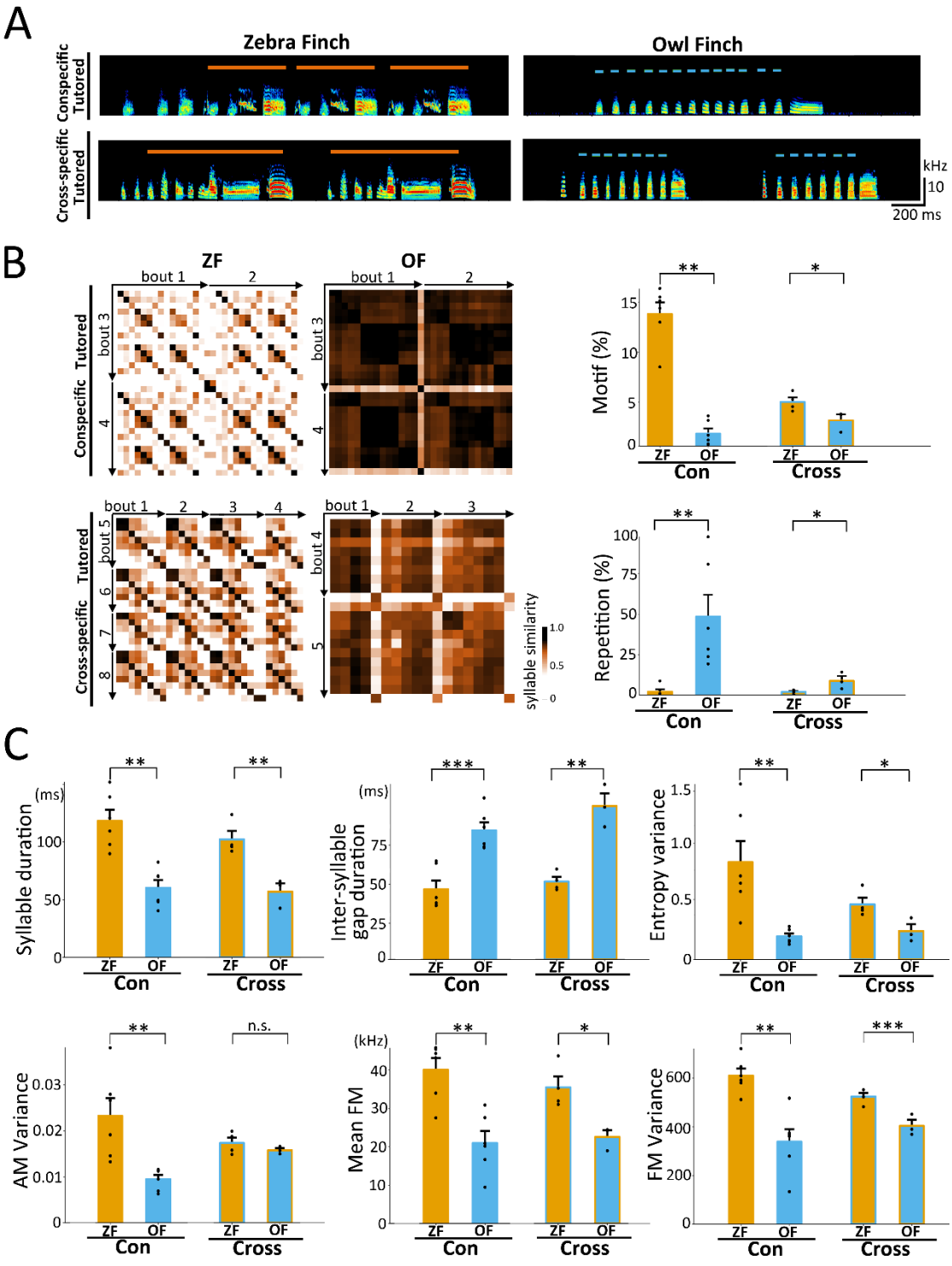

A
Zebra Finch
Owl Finch
Conspecific Tutored
Cross-specific Tutored
kHz
10
200 ms
B
ZF
OF
bout 1
2
3
4
Conspecific Tutored
Cross-specific Tutored
bout 3
bout 5
6
7
8
bout 4
5
syllable similarity
1.0
0.5
0
Motif (%)
Repetition (%)
ZF
OF
Con
Cross
C
Syllable duration
(ms)
Inter-syllable gap duration
Entropy variance
AM Variance
Mean FM
(kHz)
FM Variance
n.s.

**Figure I-3. Species difference in song phenotype between ZF and OF**

**(A)** Typical examples of songs from ZFs and OFs which were reared with conspecific song tutoring and cross-species song tutoring. Orange solid and blue dotted lines represent the motif and repetitive structure of syllables, respectively.

**(B)** Species differences in syllable sequence of ZF and OF songs. (left) Syllable similarity matrices (SSM) for songs produced by ZFs and OFs which were reared with conspecific song tutoring and cross-species song tutoring. (Right) Motif and repetition indices of ZF and OF songs (Mean ± SE; "Con": n = 6 each from conspecific song tutored ZF and OF; "Cross": n = 4 and 3 from cross-species song tutored ZF and OF, respectively; one-way ANOVA, $*p < 0.05$, $**p < 0.01$). Each dot corresponds to individual birds.

**(C)** Species differences in syllable acoustics (syllable duration, inter-syllable gap duration, entropy variance, AM variance, mean FM, and FM variance) of ZF and OF songs (Mean±SE; "Con": n = 6 each from conspecific song tutored ZF and OF; "Cross": n = 4 and 3 from cross-species song tutored ZF and OF, respectively; one-way ANOVA, $p* < 0.05$, $**p < 0.01$, $p*** < 0.001$). Each dot corresponds to individual birds.

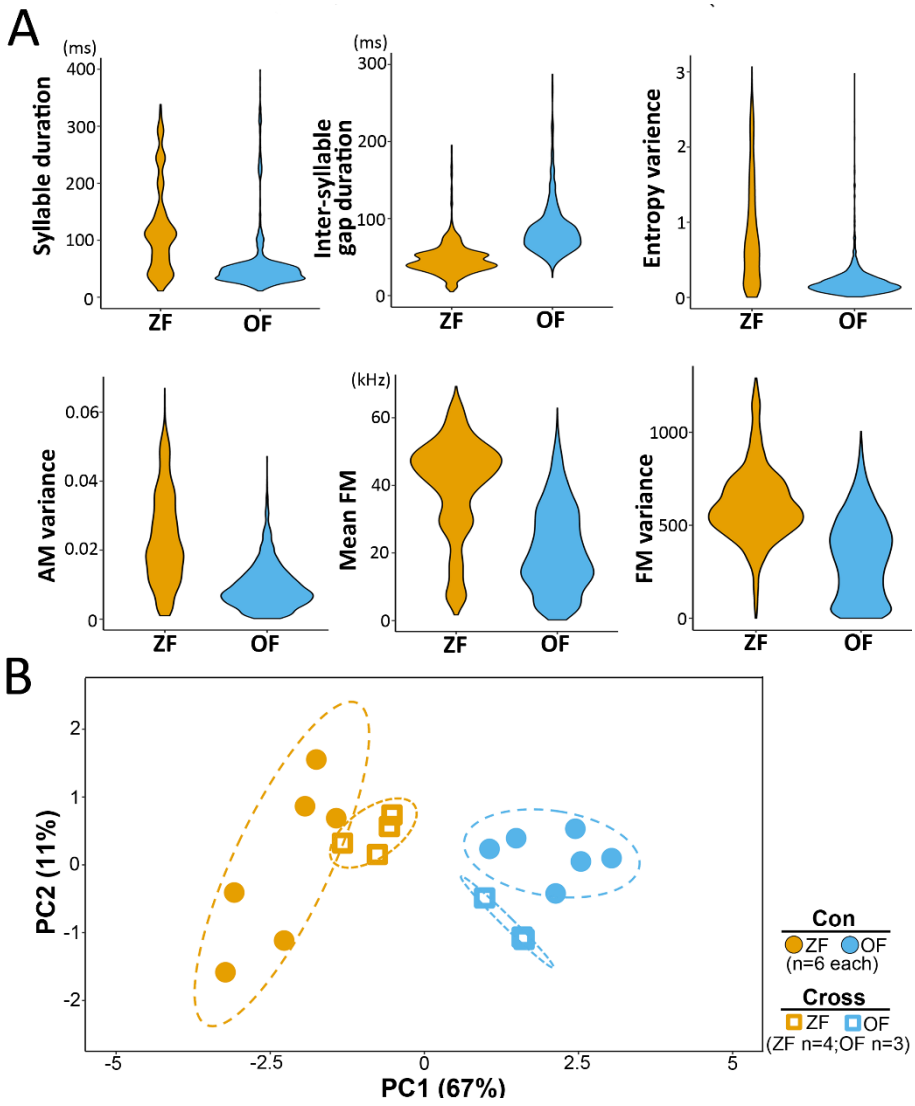



**Figure I-4: The similar distribution range of syllable acoustic traits between ZF and OF.**

**(A)** The distribution of syllable duration, inter-syllable gap duration, entropy variance, AM variance, mean FM, and FM variance from ZF and OF which were reared with conspecific song tutoring (total 3,000 syllables from n = 6 birds each and 500 syllables/bird).

**(B)** PCA of the song features of ZFs and OFs reared under conspecific and cross-species song tutoring conditions ("Con": n = 6 each from conspecific song tutored ZF and OF; "Cross": n = 4 and 3 from cross-species song tutored ZF and OF, respectively).

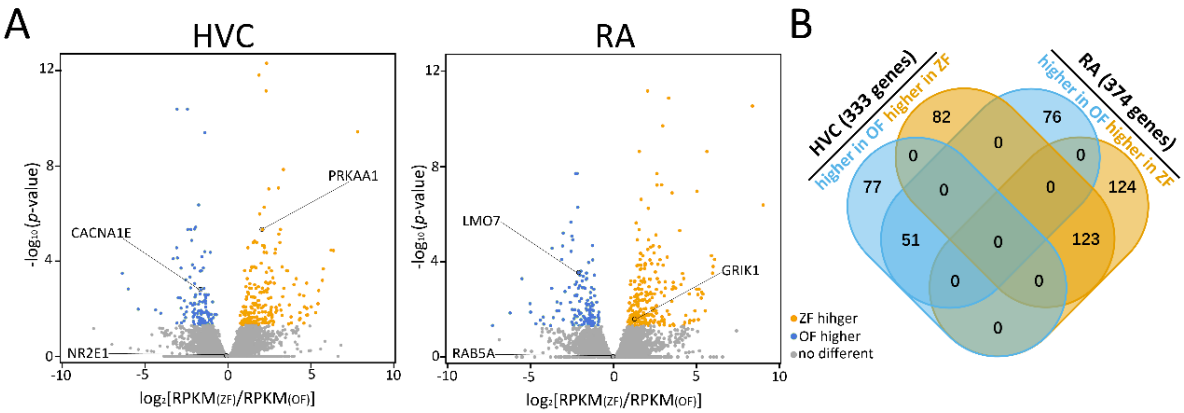



**Figure I-5. Species-differentially expressed (SDE) genes in HVC and RA**

**(A)** Species-differentially expressed (SDE) genes in HVC and RA. Orange and blue colored spots represent significantly higher expression in ZF or OF, respectively. Grey spots mean no significant expression between ZF and OF. (R DEseq2 package; corrected with Benjamini-Hochberg's method, $p < 0.05$).

**(B)** Venn diagram representing number of genes in HVC and RA, which were differently expressed between ZF or OF.

# Chapter II

# Regulatory divergence influenced gene expression differences between zebra finch and owl finch in the vocal motor pathway

## 2.1 Introduction

Species-specific birdsongs are produced by conserved song pathways in songbird species (Nottebohm *et al.*, 1976; Marler & Slabbekoorn, 2004). Species-specific songs are characterized by syllable sequence and acoustics, controlled by HVC and RA in the motor pathway, respectively (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Sober *et al.*, 2008). Previous studies have shown genes (such as forkhead box protein P2, FoxP2 and muscarinic acetylcholine receptor 2, CHRM2) with species-differential expression in song pathway of songbirds (Haesler *et al.*, 2004; Wada *et al.*, 2013b; Asogwa *et al.*, 2018). In **Chapter I**, I found that genes were differentially expressed in HVC and RA between ZF and OF as the SDE genes. However, the regulatory mechanism that induces differences in gene expression is not understood. To the best of my knowledge, there is no research to explore how transcriptional regulatory difference affects on the evolution of species-specific birdsong in songbirds.

Differences in gene expression are caused by *cis*-regulatory changes that affect transcription initiation, transcription rate, and/or transcript stability, or *trans*-regulatory changes that modify the activity or expression of factors that interact with *cis*-regulatory sequences (Wittkopp *et al.*, 2004; Carroll *et al.*, 2013). During evolution, mutations in *cis*- and *trans*-elements could have changed the expression levels of related genes in evolved species as shown in **Fig. I-1A** (Wittkopp *et al.*, 2004; Mack & Nachman, 2017). The F1 hybrids between closely related species are good models to estimate *cis*- and/or *trans*- changes during speciation, because the two alleles inherited from both parents share same transcriptional environment in the cell nuclei of F1 hybrids. Therefore, *cis*-

difference between parental species could be identified by comparing allelic expression levels of genes in F1 hybrids (**Fig. II-1**). Moreover, comparison of the allelic expression ratio in F1 hybrids and expression ratio of orthologous genes between parental species could estimate *trans*-regulatory difference (**Fig. II-1**) (Wittkopp *et al.*, 2004; Wittkopp, 2005; Mack & Nachman, 2017; Signor & Nuzhdin, 2018). Such strategy has been used in some species to explore *cis*- and/or *trans*-regulatory difference between closely related species in *Drosophila*, mice, birds, and wasps (Wittkopp *et al.*, 2004; Goncalves *et al.*, 2012; Davidson & Balakrishnan, 2016; Wang *et al.*, 2016). However, how *cis*- and/or *trans*-regulatory mutation affects species-specific behaviors, especially learned behaviors and the brain region specific regulation between closely related species are still unknown. Identification of *cis*- and/or *trans*-regulatory differences driving species-specific gene expression in the song pathway among songbird species may enhance understanding of how songbirds produce species-specific song using conserved neural circuits. In addition, elucidating *cis*- and/or *trans*-differences in the song pathway of songbirds may further illuminate the molecular mechanism of vocal learning in songbirds. Previous work of Dr. Wada lab found that reciprocal F1 hybrid songbirds between ZF and OF can be bred under laboratory conditions (**Fig. I-2**). The F1 hybrid songbird is good model to identify transcriptional regulatory divergence between ZF and OF. As the strategy introduced above, the parental expression ratio of each gene can be calculated as RPKM values as shown in **Chapter I**. Meanwhile, the allelic expression ratio of each gene in F1 hybrids can be calculated by counting the number of mRNA molecular which transcribed from ZF and OF allele. Using the two values

(parental expression ratio and allelic expression ratio in F1 hybrid) of each gene, *cis*- and/or *trans*-regulatory difference between ZF and OF can be identified.

In this chapter, I evaluated the *cis*- and/or *trans*-regulatory differences between ZF and OF in both HVC and RA using strategy in **Fig. II-1**. Firstly, the allelic expression ratio in F1 hybrids between ZF and OF were calculated utilizing species-specific single nucleotide polymorphisms (ss-SNP). By comparing parental expression ratio and allelic expression ratio of each gene, I identified *cis*- and/or *trans*-regulatory difference between ZF and OF, in both HVC and RA. Since HVC and RA playing different roles for the adult song of songbird (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Sober *et al.*, 2008), I compared the *cis*-, *trans*-, and both *cis*- and *trans*-regulatory different genes between HVC and RA. To examine the biological function of transcriptional regulatory different genes, I performed functional analysis of *cis*-, *trans*-, and both *cis*- and *trans*-regulated genes between ZF and OF in both HVC and RA.

## 2.2 Materials and methods

***Calculation of gene expression ratios between parental species***

I estimated the expression level of each gene in HVC and RA of each animal (ZF: n = 4; OF: n = 4) with the RPKM values shown in Materials and methods of **Chapter I**. The average expression level of each gene was calculated using the average RPKM values of each gene which were defined as follows.

$$A = \frac{(A1+A2+A3+A4)}{4} \qquad B = \frac{(B1+B2+B3+B4)}{4}$$

Where $A_1$, $A_2$, $A_3$, and $A_4$ and $B_1$, $B_2$, $B_3$, and $B_4$ represent the RPKM value of orthologous genes in each ZFs and OFs, respectively. Gene expression ratios between parental species were calculated as $X = log_2(A/B)$, where X is the parental expression ratio between ZF and OF; A and B are the average expression ratios described above. Examples of the parental expression ratio calculation were shown in **Fig. II-2A**.

***Identification of species-specific single nucleotide polymorphisms (ss-SNPs)***

NGS raw data of ZF and OF whole brain samples were filtered by removing adaptor and low quality reads using Trimmonatic software (Bolger *et al.*, 2014). Clean reads of ZF and OF whole brain samples were mapped to the ZF reference genome which was downloaded from Ensembl (Taeniopygia guttata.taeGut3.2.4). Ss-SNPs were discovered from the mapping result and were defined as following rules: the base variants were same in all individuals of a species but different from the base found in all individuals of another species. SNPs in individuals of same species were shield to

same base for both ZF and OF reconstructed genome DNA. The positions with ss-SNP of ZF genome (Taeniopygia guttata.taeGut3.2.4) were replaced by ZF type and OF type SNP to make ZF version and OF version reconstructed genome. MUMmer software (Delcher *et al.*, 2003) was used to identify ss-SNPs by performing alignment sequence of each chromosome of the reconstructed ZF and OF genome.

***Analysis of allelic specific expression (ASE) genes in F1 hybrid***

To determine the allelic origin of reads, mapping results of HVC and RA of F1 hybrid individuals were used for the following analysis depending on the user guide of SNPsplit software (Krueger & Andrews, 2016). Firstly, bases at the ss-SNP position of reconstructed ZF genome were replaced by character 'N' to make N-marked genome. Secondly, the clean reads of HVC and RA (ZF, OF and F1 hybrids samples) individuals were mapped to the N-marked genome by TopHat2 (Kim *et al.*, 2013). Lastly, SNPsplit software was used to distinguish reads from ZF or OF allele depending on the discovered ss-SNP information. The reads numbers of each ss-SNP were counted by using samtools (Li *et al.*, 2009) based on the results of SNPsplit.

I performed quality control of the genes used for allelic expression ratio analysis in HVC and RA of F1 hybrids. Thresholds were set as follows. 1) For each ss-SNP site in HVC and RA of ZF, the percentage of reads transcribed from ZF allele should more than 98% (vice versa in OF HVC and RA samples) to double check the veracity of identified ss-SNP. 2). The reads number of each ss-SNP site should be more than 5. 3) Median value of RPKM of 16 individuals (ZF: n = 4; OF: n = 4; ZO: n = 4; OZ: n = 4)

should be more than 10. Genes with at least one ss-SNP which passed the above threshold were used for allelic expression ratio calculation. A total of 5,827 and 6,328 genes passed the threshold were used for the following analysis in HVC and RA, respectively. Allelic expression ratio was measured by *d*-score which was described in previous research as the following formula (Eckersley-Maslin *et al.*, 2014).

$$d = \frac{\text{Reads}_{(\text{ZF})}}{\text{Reads}_{(\text{ZF})} + \text{Reads}_{(\text{OF})}} - 0.5$$

The range of *d*-score is from -0.5 to 0.5 which means gene bias from OF to ZF, and *d*-score equal to 0 means no bias to either allele. The significant difference between reads number of ZF and OF allele was tested by using *Fisher*-test which compared ZF and OF allele reads against the average number of reads from both alleles. Genes were defined as allelic imbalance in a F1 hybrid individual with threshold |*d*-score| > 0.18 and *p*-value < $10^{-8}$ (*Chi*-square test, FDR adjust *p*-values). For eight F1 hybrids (ZO: n =4; OZ: n =4), genes with at least four F1 hybrid individuals showing allelic bias towards either ZF or OF and no individual showing bias to the other allele was considered as allelic imbalance gene.

***Imprinting gene identification between ZO and OZ***

I performed correlation analysis to test the potential of genomic imprinting in F1 hybrids using Spearman's rank correlation of gene allelic expression ratio between ZO (n = 4) and OZ (n = 4). In addition, the differences in allelic expression ratios of each

gene were compared between ZO (n = 4) and OZ (n = 4) hybrids using one-way ANOVA (adjust *p*-value by Benjamini-Hochberg's method).

***Calculation of allelic expression ratio in F1 hybrids***

Reads numbers transcribed from ZF or OF alleles of genes passed the threshold were counted in each F1 hybrid individuals. Allelic expression ratios of F1 hybrids were calculated as $Y = log_2(a/b)$, where $Y$ is the allelic expression ratio between two alleles; $a$ and $b$ are the read numbers of ZF allele and OF alleles in F1 hybrid individuals, respectively. Two gene examples of allelic expression ratios calculation were shown in **Fig. II-2B**. Since there was no imprinting genes in F1 hybrids, we treated ZO and OZ F1 hybrids as the same to perform the following analysis.

***Identification of cis- and/or trans-regulatory divergence between ZF and OF***

*Cis*- and *trans*-effects on gene expression divergence were estimated by the scheme described in **Fig. II-1**. In brief, the regulation mechanism of gene expression between ZF and OF was (1) *cis*-regulatory difference, if $X = Y$ and $Y \neq 0$; (2) *trans*-regulatory difference, if $X \neq Y$ and $B = 0$; (3) both *cis*- and *trans*-regulatory difference, if $X \neq Y$ and $Y \neq 0$; (4) no *cis*- and *trans*-regulatory difference (i.e., conserved), if $X = Y$, and $Y = 0$. The student's *t*-test was used to determine the difference between gene expression ratio in parental species and allelic expression ratio in F1 hybrids. The *SGoF* program (Carvajal-Rodriguez *et al.*, 2009) was employed to correct *p* values for multiple testing (adjusted $p \leq 0.05$). The previous standard method for estimating regulatory divergence

can lead to a negative correlation as an artefact when *cis*-estimates have any errors (Fraser, 2019; Zhang & Emerson, 2019). To avoid this bias, first, I randomly selected four individual F1 hybrids as a group to estimate *cis*-effects using their average ASE ratio while the remaining four F1 individual hybrids were used to compare the expression ratio between ZF and OF. For each gene, a total of 70 combinations were constructed by randomly selection of four of eight F1 hybrid birds (n = 4 each from ZO and OZ). Thus, *cis*- and/or *trans*-regulatory identification was done for each gene for each pair of 70 total combinations. During this cross-replicate comparison, some genes were categorized as different transcriptional regulations due to a large variance in ASE ratios among F1 individuals. Therefore, I finally determined which transcriptional divergence made the main regulatory effect on each gene by two steps of statistics as followed: (i) calculation of the difference between four categories (*cis*-, *trans*-, both *cis*- and *trans*-, and conserved) using *chi*-square test (with adjust *p*-value by FDR < 0.05) and (ii) a comparison of the difference between the first and second strongest regulatory effects using a *Fisher's* exact-test (adjust *p*-value by FDR < 0.05). If genes did not show significance at both two tests, such genes were defined as "ambiguous" regulatory genes (**Fig. II-3**).

In addition, I performed analysis of *cis*- and/or *trans*-regulatory divergence using a standard method (Wittkopp *et al.*, 2004) and compared these results with those from the above method. The difference of the standard method from my method is that the allelic expression ratios of all eight F1 hybrids (ZO = 4, OZ = 4) were used to estimate *cis*- and *trans*-regulatory effects. In brief, the parental expression ratio value *X* and the

allelic expression ratio in F1 hybrid value *Y* were calculated similarly to the new method. The average value *Y* of eight F1 hybrid individuals were used to compare with value *X* and 0, respectively, to estimate *cis*- and *trans*-effects by the scheme described in **Fig. II-1** and **Fig. II-3** (Student's *t*-test). *SGoF* program (Carvajal-Rodriguez *et al.*, 2009) was employed to perform multiple testing correction (adjust *p*-value ≤ 0.05).

***Functional analysis of gene group with cis- and/or trans-regulatory different genes***

The function of gene groups with *cis*-, *trans*- and *cis*- and *trans*-regulatory differences between ZF and OF in HVC and RA were annotated by Gene Ontology (GO) analysis using a website tool, DAVID Bioinformatics Resources 6.8 (https://david.ncifcrf.gov) (Huang *et al.*, 2009). GO enrichment analysis allows elucidating the potential biofunctional significance of gene groups depending on the pre-annotated function of each gene. Therefore, I performed GO enrichment analysis of *cis*- and/or *trans*-regulatory different genes by using Fisher's test (*p*-value was adjusted by Bonferroni method). Since *trans*-regulated genes in RA were enriched the most GO terms, I performed upstream regulation analysis for *trans*-regulatory different genes in RA. Ingenuity Pathway Analysis (IPA) was employed to perform upstream regulation enrichment analysis based on the human gene function database (Krämer *et al.*, 2013). The fold change of each gene *[Foldchange = $log_2$(ZF_RPKM/OF_RPKM)]* and *p*-value (got from species differential expressed gene analysis, **Fig. I-5**) were used as input parameters for IPA analysis.

## 2.3 Results

***Identification of ss-SNPs and allelic specific expression (ASE) genes***

Depending on the sequencing results of whole brains, 2,409,063 ss-SNPs were identified (average 2.02 SNPs/Mb) between ZF and OF. To quantify ASE genes in F1 hybrids, I figured out genes passed the threshold (ss-SNP number ⩾1; RPKM value > 10; Read number of each ss-SNP ⩾ 5) described in the Materials and methods part. As a result, a total of 5,827 and 6,328 genes passed the threshold survived in HVC and RA, respectively. To test the existence of genomic imprinting in F1 hybrid songbirds, I performed a correlation analysis of parental allelic expression ratio between ZO (n = 4) and OZ (n = 4). The results showed that parental allelic expression ratios were highly correlated between ZO and OZ in HVC and RA genes (Pearson correlation, $r = 0.527$, $p < 2.2e\text{-}16$ in HVC; $r = 0.550$, $p < 2.2e\text{-}16$ in RA, **Fig. II-4**). Furthermore, statistical analysis showed that the parental expression ratio of these genes was not significantly different between ZO and OZ (one-way ANOVA, $p$-value < 0.05 adjusted by the Benjamini-Hochberge's method). These results suggested that there was no evidence of parental imprinting gene exist in reciprocal F1 hybrids, which was in line with previous research in avian species (Fresard *et al.*, 2014). Therefore, I used ZO and OZ individuals equally perform allelic imbalance expression calculations.

I identified 504 allelic imbalance expression genes in HVC, which contain 402 OF allele biased genes (OF-biased) and 102 ZF biased allele genes (ZF-biased); while 403 genes were identified as allelic imbalance expression genes, which contain 234 OF-

biased genes and 169 ZF-biased genes in RA (**Fig. II-5B, 5C**). Two genes, 5-hydroxytryptamine receptor 1B (HTR1B) and Beta-1,3-Glucuronyltransferase 2 (B3GAT2), were shown as examples representing non-allelic imbalance and allelic imbalance gene in **Fig. II-5A, 5B**. A total of 145 genes showing OF biased expression and 50 genes showing ZF biased expression existed in both HVC and RA. However, more than half (309 of 504 genes in HVC; 208 of 403 genes) ASE genes only existed in HVC or RA (**Fig. II-5C**). These results indicate that ASE genes were brain region-specifically regulated in the motor pathway in F1 hybrid songbirds. I then explored how genes with allelic imbalance distribute on the entire genome including autosome and sex chromosomes. OF allele expression-biased genes (402 and 234 genes) exist more than ZF allele expression-biased genes (102 and 169 genes) in HVC and RA, respectively, on the whole genome level. In contrast, such patterns were opposite on chromosome Z (**Fig. II-5D**). Only 10 and 8 genes showed OF allele expression-biased, while 26 and 29 genes showed ZF allele expression-biased on chromosome Z in HVC and RA, respectively, which meant that the number of ZF allele expression-biased genes were significantly higher than OF allele expression-biased on chromosome Z in the F1 hybrids (Binomial test, $p = 0.01$, HVC; $p = 0.0008$, RA).

***Transcriptional regulatory divergences existed between ZF and OF***

To identify *cis*- and/or *trans*-regulatory difference between ZF and OF, I calculated the parental expression ratios and allelic ratios of genes passed the threshold (**Fig. II-1; Fig. II-3; Fig. II-6**). In F1 hybrids, *cis*-differences inherited from two parental species

were maintained and the expression levels of two alleles in F1 hybrid were different. For *trans*- regulation different genes, *trans*-elements were shared in the same cellular environment in the F1 hybrids. Therefore, the expression differences between the two alleles were eliminated. To identify the transcriptional regulatory divergence between ZF and OF, log2-transformed gene expression ratio between parental species [$\log_2$(A/B)] and allelic expression ratio in F1 hybrids [$\log_2$(a/b)] were calculated as described in the Materials and methods (**Fig. II-1**, **II-6)**. As a result, 158 (2.4% of total 5,827 genes) of genes showed evidence of significant *cis*-regulatory difference, 271 (4.7%) of genes showed *trans*-regulatory difference, 183 (3.1%) genes showed both *cis*- and *trans*-regulatory difference, 4,489 (77.0%) genes showed conserved between ZF and OF, and 726 (12.5%) genes were classified as ambiguous genes in HVC. In RA, 246 (3.9% of the total 6,328 genes), 383 (6.1%), 183 (2.9%), 4782 (75.6%) and 734 (11.6%) genes were classified as only *cis*-, only *trans*-, both *cis*- and *trans*-, conserved and ambiguous regulatory difference between ZF and OF, respectively (**Fig. II-6A, 6B**). These results indicated that the expression of 600–800 genes (approximately 10–15 % of the expressed genes) in the vocal motor song nuclei was modified by altered transcriptional regulation between the two species.

***Brain region-specific alternation of cis- and/or trans-regulatory divergence***

In both HVC and RA, genes with *trans*-alternation were more prevalent than *cis*-alteration. Furthermore, the majority of the genes under conserved regulation were highly expressed in both HVC and RA [3,523 genes of 4,489 (78.5%) and 4,782 (73.7%)

genes expressed in HVC and RA, respectively]. In contrast, most of the *cis*- and/or *trans*-regulated genes were not shared between HVC and RA (**Fig. II-6B**), showing a brain region-specific transcriptional regulatory alteration. This result was similar to the results obtained with the estimation method using the average of ASE of all F1 hybrids, which showed similar rates of *cis*- versus *trans*-regulation divergence (**Fig. II-6C**).

***Transcriptional regulatory changes significantly correlated with species-differentially expressed (SDE) genes in the song nuclei***

I further examined whether the SDE genes in HVC and RA were affected by the transcriptional regulatory divergences between ZF and OF. Based on the RPKM values of each gene expressed in ZF and OF, 333 and 374 genes showed significantly different expression in HVC and RA, respectively, between the two species (2.9% and 3.3% of all the genes expressed in HVC and RA, respectively) (DEseq2, *p*-value corrected by Benjamini-Hochberg's method, $p < 0.05$; n = 4 each from ZF and OF) (**Fig. I-5**). A total of 209 and 242 SDE genes in HVC and RA, respectively, passed the ss-SNPs threshold for calculating the ASE ratio in F1 hybrids. Such SDE genes were significantly enriched with a higher probability for *cis*-, *trans*-, and both *cis*- and *trans*-regulatory effects, compared with those of non-SDE genes, in both HVC and RA (*Fisher's* exact test, $p^{***} < 0.001$) (**Fig. II-7B**). These results showed a significant association of transcriptional regulatory changes with SDE genes in the song nuclei.

***Trans-regulatory divergence was with a predominant effect on neural function***

To determine the potential function of *cis*-, *trans*- and *cis*- and *trans*- genes in HVC and RA, I performed Gene Ontology (GO) enrichment analysis using these sets of genes. The results showed that more GO categories were enriched for *trans*-regulated genes compared to the other types of transcriptional regulatory divergence in both HVC and RA (**Fig. II-8**). GO categories related to neural functions associated with synapse transmission, soma to dendritic compartment, and nervous system development were significantly enriched for RA *trans*-regulated genes (Fisher's test, *p*-value corrected by Benjamini-Hochberg's method, $p < 0.05$). These results motivated me to focus on altered *trans*-regulation in RA. To predict the potential regulatory mediators driving species differences in the expression of *trans*-regulated genes in RA, I performed upstream regulatory analyses using IPA (Krämer *et al.*, 2013) and found that potential upstream regulators that were significantly enriched. In these upstream regulators, brain-derived neurotrophic factor (BDNF) was the most significant upstream mediator of genes under *trans*-regulation in RA (**Fig. II-9A**), which included genes [e.g., glutamate decarboxylase (GAD) 2, NMDA glutamate receptor (GRIN) 2A, neuropeptide Y (NPY), and collapsin response mediator protein (CRMP) 1] for neural plasticity and dendritic spine development (*Fisher's* exact test, $p = 6.44E-07$) (**Fig. II-9A, 9B**).

## 2.4 Discussion

In this chapter II, I identified *cis*- and/or *trans*-regulatory differences between ZF and OF in HVC and RA. The transcriptional regulatory differences were brain region-specifically exist, in turn, which significantly affected gene expression difference between ZF and OF. The *trans*-regulatory genes were predominant than *cis*-regulatory genes in both HVC and RA. Functional analysis revealed that *trans*-regulatory genes were associated with synaptic formation and transmission in RA. In addition, I found that some genes including BDNF were enriched as upstream mediators for the *trans*-regulatory genes in RA.

One of my important findings in this chapter was that *trans*-regulatory different genes between ZF and OF were much greater than *cis*- and both *cis*- & *trans*-regulatory genes. This result was different from the previous studies that perform regulatory divergence analysis in other species (Goncalves *et al.*, 2012; Gomes & Civetta, 2015; Davidson & Balakrishnan, 2016; Wang *et al.*, 2016; Reuveni *et al.*, 2018). In previous studies, researchers used whole organ or whole body of animal to explore the regulatory difference between closely related species (Gomes & Civetta, 2015 using whole body of fruit fly, Wang *et al*., 2016 using whole body of wasps, Davidson & Balakrishnan, 2016 using whole brain of songbird, Goncalves *et al*., 2012 & Reuveni *et al.*, 2018 using whole liver). The results of these researches showed that genes with *cis*-regulatory differences were much more than *trans*-regulatory differences, which were opposite with my results. The predominant *trans*-regulatory differences made me speculate whether the method using the cross-replicated calculation to estimate *cis*-

and/or *trans*-regulatory difference caused such discrepancies. Studies reported that the traditional standard method which was used for *cis*- and/or *trans*- identification could overestimate the *cis*-regulatory genes (Fraser, 2019; Zhang & Emerson, 2019). To verify this possibility, I also used the traditional standard method to analyze *cis*- and/or *trans*-regulatory difference between ZF and OF and obtained a similar result with the cross-replicated calculation method (**Fig. II-6C**). One of the reasons why *trans*-regulatory genes showed predominantly might be using of song nuclei tissues, instead of using the whole body or whole organ. In the future, it will be necessary to perform transcriptional regulatory analyses using whole brain or non-song nuclei tissue as a background control between ZF and OF.

Additionally, I found that HVC and RA shared smaller numbers of genes which showed *cis*-, *trans*-, and both *cis*- and *trans*-regulatory difference compared with conserved genes. These results suggest that transcriptional regulatory differences between ZF and OF were brain region-specific appearing in HVC and RA. Why did genes show brain region-specific transcriptional regulatory differences even the genome sequences between HVC and RA are the same in one species? One reason may be that HVC and RA play different roles for regulating the species-specific song traits (Nottebohm *et al.*, 1976; Hahnloser *et al.*, 2002; Sober *et al.*, 2008). First, *trans*-mediators show different expression levels between HVC and RA, such genes can induce gene expression differences of downstream genes. In line with this, previous research already found that genes including transcription factors showed brain region-specific expression in the song pathway of songbirds (Asogwa *et al.*, 2018; Hayase *et*

*al.*, 2018). Secondly, epigenetic modification, such as DNA methylation, of the *cis*-DNA sequence of singing related genes may become different between HVC and RA during song development. Memory and learning progress normally accompany epigenetic changes in the *cis*- sequences which include DNA methylation and histone post-translational modifications (Levenson & Sweatt, 2005; Peixoto & Abel, 2013). The epigenetic modification causes DNA compaction and relaxation that induce to gene transcriptional repression and activation in the neural system of animals (Korzus *et al.*, 2004; Martin & Sun, 2004; Kim & Kaang, 2017). For example, *cis*-regulatory differences could be due to the difference of methylation in the upstream DNA sequence between closely related species (Wang *et al.*, 2016). In future research, it will be essential to figure out the epigenetic difference of the whole genome between HVC and RA in both ZF and OF. For instance, we can analyze the difference in DNA methylation state between HVC and RA by performing whole genome bisulfite sequencing. Therefore, the understanding of brain-region specific transcriptional regulatory divergences is a potential chance to deeply understand the molecular mechanism of species-specific vocal learning.

To the best of my knowledge, this is the first study to use the cross-replicated method to perform *cis*- and/or *trans*-identification. I defined a group of genes as “ambiguous” genes that were not revealed by the traditional calculation method. About 10% of all expressed genes in HVC and RA were categorized as “ambiguous”. Ambiguous regulated genes could have resulted from intraspecies genomic variations, such as intraspecies SNPs and indels (insert and deletion). In Dr. Wada’s lab, we found that F1

hybrids showed a wide range of individual differences in learned song structure under the same tutoring condition. The potential contribution of ambiguous genes for the song variability of F1 hybrids may be a crucial research topic in the future.

In this chapter, I found that many genes were regulated with allelic imbalance expression in HVC and RA of F1 hybrid (**Fig. II-5**). Interestingly, the number of ZF allele expression-biased genes are significantly higher than OF allele expression-biased genes on chromosome Z in F1 hybrids (**Fig. II-5D**), which suggested *cis*-regulatory differences on the chromosome Z enhanced gene expression of ZF alleles or inhibited the gene expression of OF alleles on the chromosome Z. A potential reason might be difference in DNA methylation state between ZF- and OF-alleles (Teranishi *et al.*, 2001; Wright *et al.*, 2015). Such *cis*-regulatory difference could induce differential expression of genes on chromosome Z between ZF and OF. Previous studies found that genes located on sex chromosomes play potential roles for singing behavior in songbird species (Tomaszycki *et al.*, 2009; Itoh *et al.*, 2011). My results suggested that species differential expressed genes which were regulated by *cis*-regulatory differences on chromosome Z might be important candidate genes for species-specific songs between ZF and OF.

## 2.5 Figures

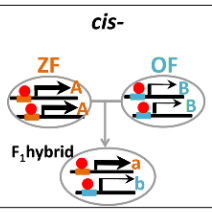
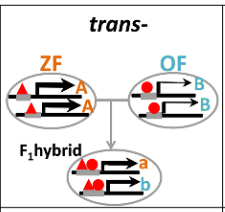
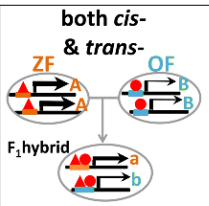
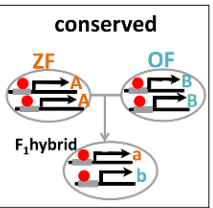
| | *cis*- | *trans*- | both *cis*- & *trans*- | conserved |
|---|---|---|---|---|
| |  |  |  |  |
| Parental expression ratio A/B | ≠ 1 | ≠ 1 | ≠ 1 | = 1 |
| $F_1$ hybrid allelic expression ratio a/b | ≠ 1 | = 1 | ≠ 1 | = 1 |
| Difference between P and $F_1$ (A/B) /(a/b) | = 1 | ≠ 1 | ≠ 1 | = 1 |

**Figure II-1: Strategy to identify *cis*- and/or *trans*-regulatory difference between ZF and OF.**

Classification of species differences in *cis*- and/or *trans*-regulations based on the comparison of relative gene expression ratio between parental species and the allelic expression ratio in their F1 hybrids. For each gene, “A” and “B” represent gene expression levels in ZF and OF, respectively. “a” and “b” are gene expression levels from ZF and OF alleles, respectively, in F1 hybrids. “A/B” and “a/b” are the expression ratios between parental species and allelic expression ratios in F1 hybrids, respectively.

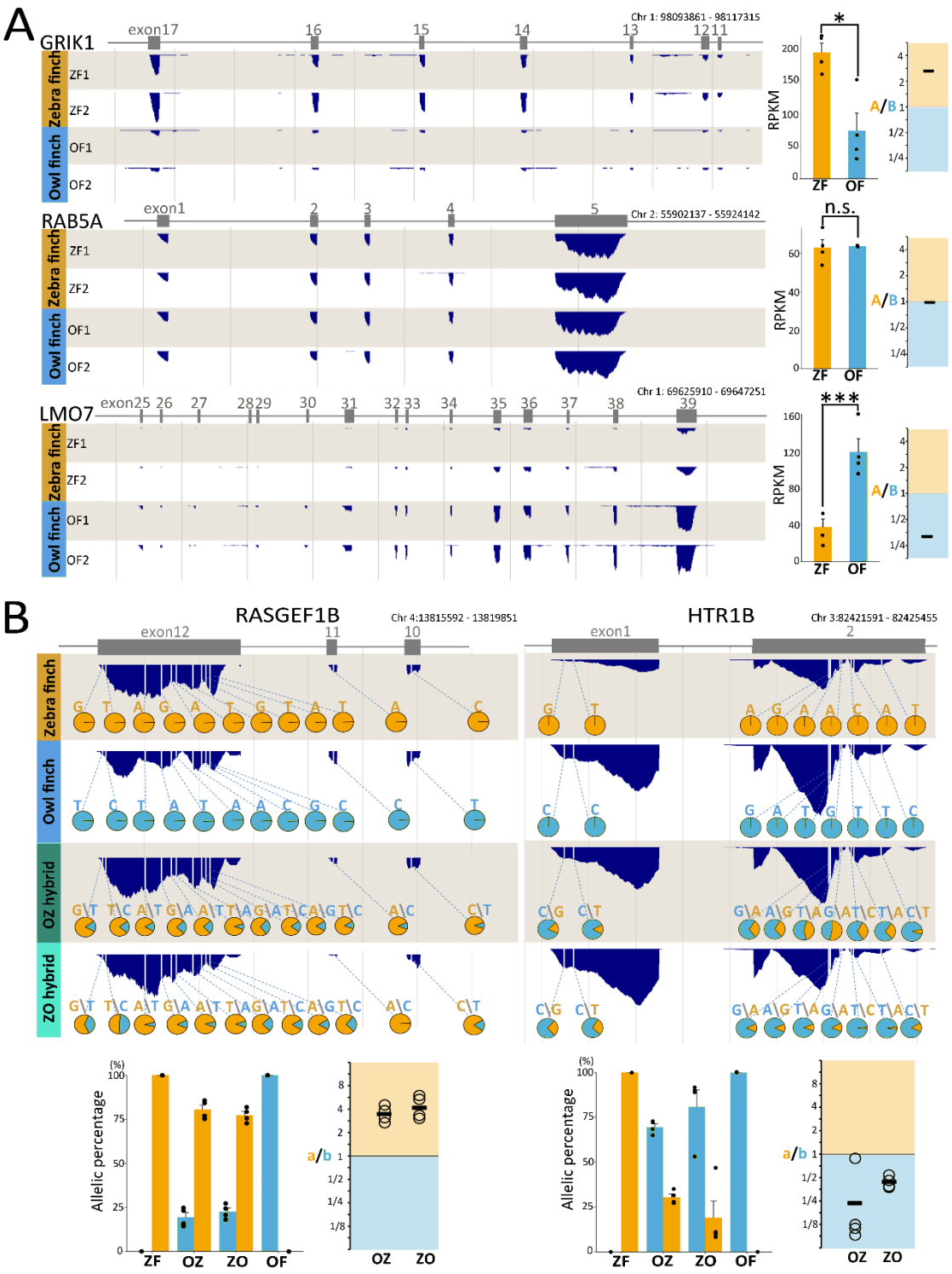

A
GRIK1
Chr 1: 98093861 - 98117315
exon17
16
15
14
13
12 11
Zebra finch
ZF1
ZF2
Owl finch
OF1
OF2
RPKM
*
ZF
OF
A/B
RAB5A
Chr 2: 55902137 - 55924142
exon1
2
3
4
5
n.s.
LMO7
Chr 1: 69625910 - 69647251
exon25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
***
B
RASGEF1B
Chr 4:13815592 - 13819851
exon12
11
10
HTR1B
Chr 3:82421591 - 82425455
exon1
2
OZ hybrid
ZO hybrid
Allelic percentage
(%)
ZF
OZ
ZO
OF
a/b
OZ
ZO

**Figure II-2: Species differences in gene expression between ZF and OF and allele specific expression in F1 hybrids.**

**(A)** Examples of species differences in gene expression between ZF and OF. (Left panels) Expression levels of Glutamate receptor, ionotropic, kainate 1 (GRIK1), Ras-related protein Rab-5A (RAB5A), and LIM domain only protein 7 (LMO7) in song nucleus RA of ZFs and OFs. Gray boxes represent the position of exons for each gene. Dark blue peaks below exons represent read density. (Right panels) Gene expression levels in ZF and OF and the average expression ratio between ZF and OF. Each dot represents RPKM value for each individual. Data are Mean ± SEM. (n = 4 birds each; one-way ANOVA, *$p < 0.05$, ***$p < 0.001$, n.s.: not significant).

**(B)** Examples of allele specific expression in F1 hybrids. (Upper panels) Allelic expression ratios in F1 hybrids at species-specific SNPs (ss-SNPs) of RASGEF1B and HTR1B in the song nucleus RA. Dark blue peaks below exons represent read density. White bars in the dark blue-colored peaks represent ss-SNP positions. Pie charts of each ss-SNP represent the percentage of transcribed read numbers from ZF (orange) and OF (blue) alleles. (Bottom panels) the percentage and ratio of parental species-allelic expression of RasGEF domain family, member 1B (RASGEF1B) and HTR1B in OZ and ZO F1 hybrids. Each dot represents average allelic expression ratios of all species-specific SNPs in one individual (n = 4 birds each, mean). Orange and blue-colored bars represent the values from ZF and OF alleles, respectively. Data are Mean ± SEM (n = 4 birds each).

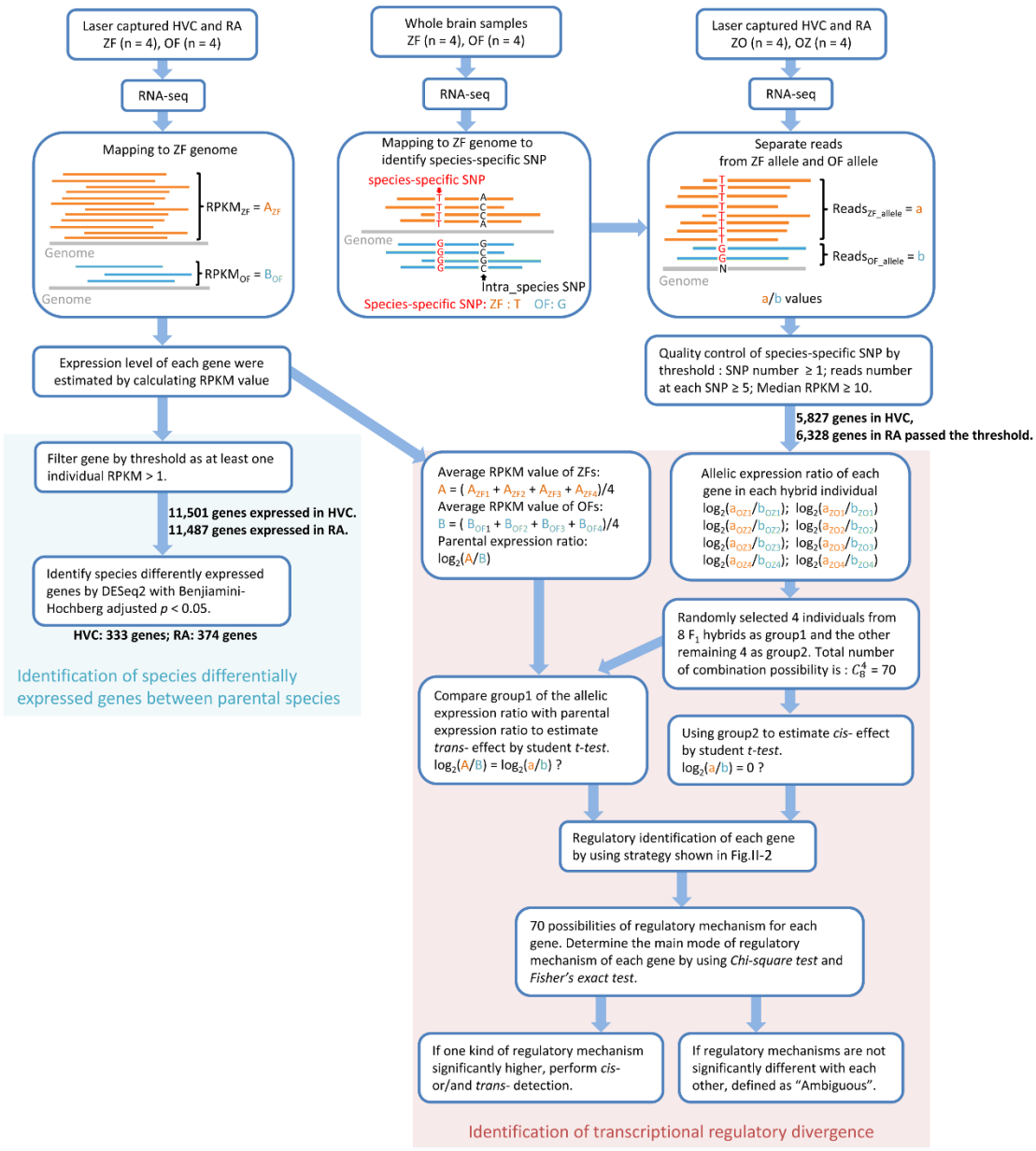



**Figure II-3: Workflow of dada analysis for SDE genes and transcriptional regulatory divergence between ZF and OF.**

Experimental workflow for the calculation of species-differently expressed genes and characterization of transcriptional regulatory divergence

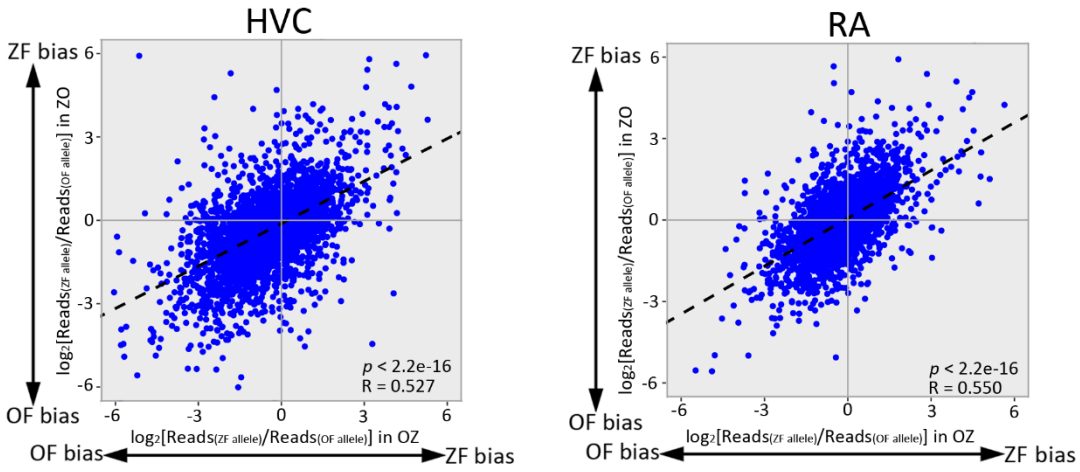



**Figure II-4. No genomic imprinting genes in reciprocal $F_1$ hybrids of ZF and OF**

Correlation analysis of genes between OZ and ZO. Scatter plots of allelic expression ratios of 5,827 and 6,328 genes in HVC and RA, respectively, of OZ and ZO hybrids (Pearson correlation coefficient).

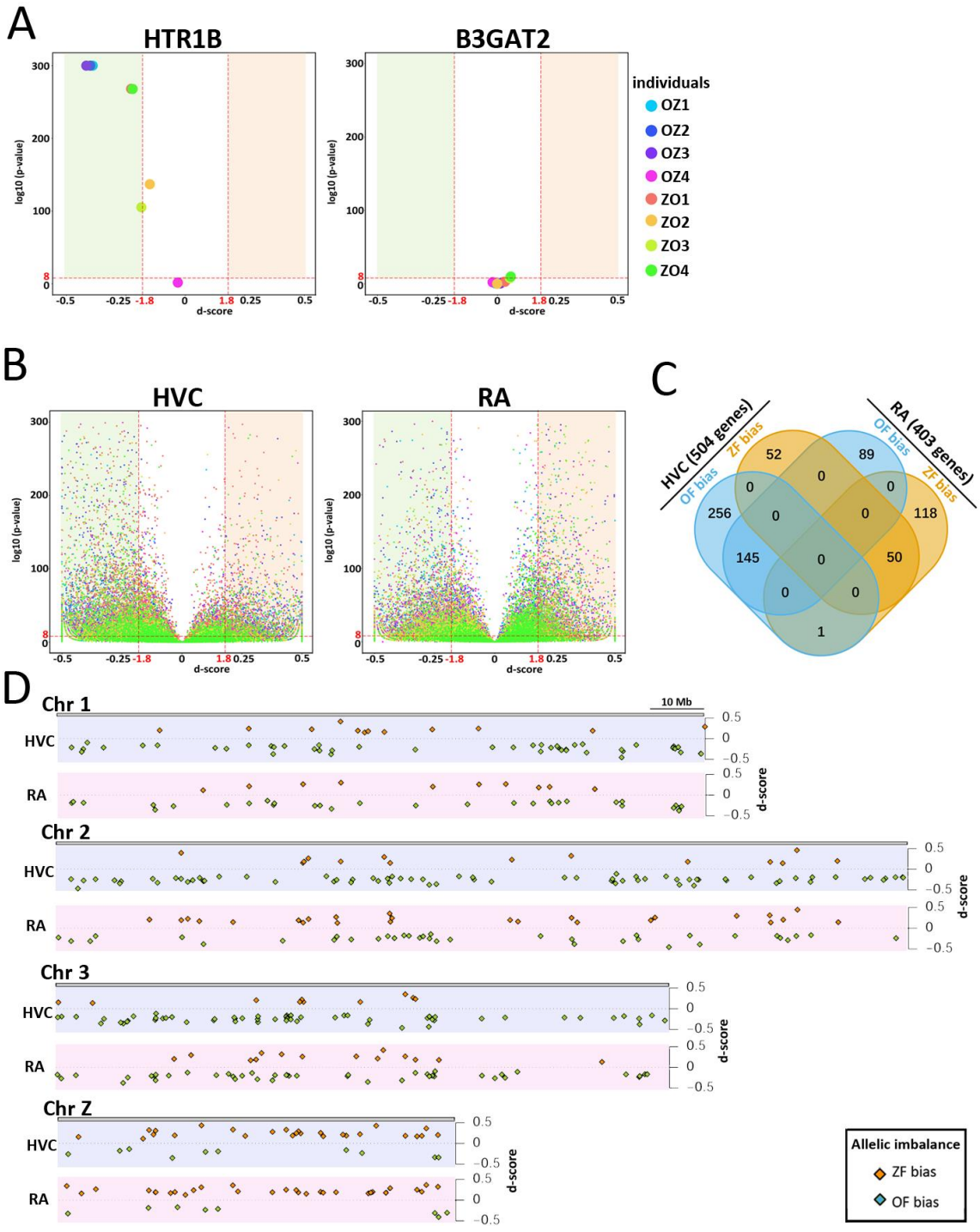

A
HTR1B
B3GAT2
log10 (p-value)
d-score
individuals
OZ1
OZ2
OZ3
OZ4
ZO1
ZO2
ZO3
ZO4
B
HVC
RA
C
HVC (504 genes)
RA (403 genes)
ZF bias
OF bias
52
89
0
0
0
0
256
0
118
145
0
50
0
0
1
D
Chr 1
Chr 2
Chr 3
Chr Z
10 Mb
HVC
RA
d-score
Allelic imbalance
ZF bias
OF bias

**Figure II-5: Allelic imbalance genes in HVC and RA in F1 hybrids.**

**(A)** Two gene examples in RA showed how to define allelic imbalanced expression genes (HTR1B) and allelic expression unbiased gene (B3GAT2).

**(B)** Scatter plot of allelic expression in each animal's HVC and RA. Different colors of dots represent different F1 hybrid individuals. Plots located in the area with light green or orange background mean genes significantly biased to OF or ZF respectively.

**(C)** Venn diagram showed gene number which was ZF-biased and OF-biased genes in both HVC and RA, respectively.

**(D)** Distribution of allelic imbalance expressed genes on three auto-chromosomes and chromosome Z in HVC and RA of F1 hybrids. X-axis represents the gene position of each chromosome. Y-axis represents average of *d*-scores (ZO: n =4; OZ: n =4) of each gene. Orange and blue squared plots represent ZF-biased and OF-biased genes, respectively.

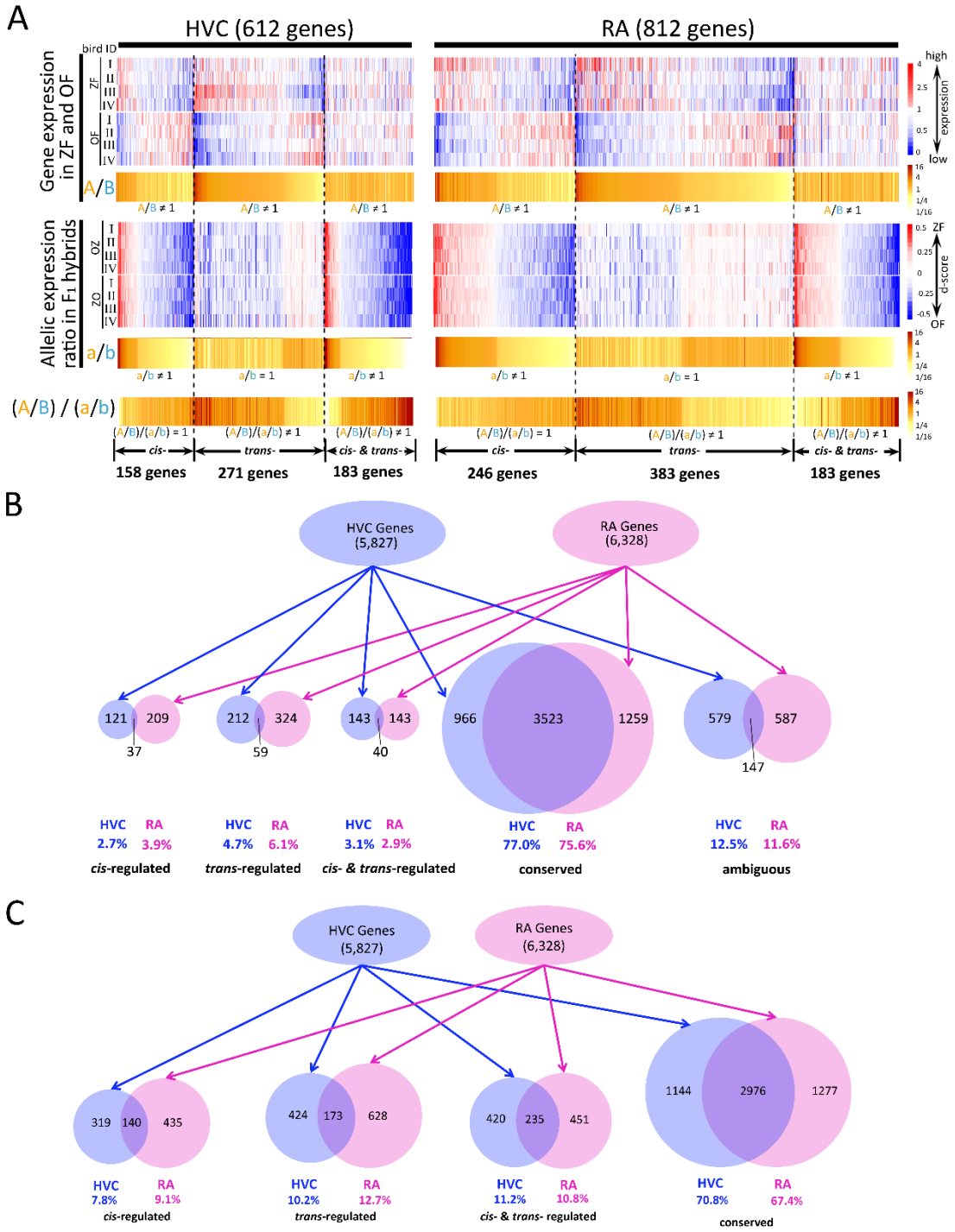

A
HVC (612 genes)
RA (812 genes)
bird ID
Gene expression in ZF and OF
ZF
OF
A/B
high
low
expression
Allelic expression ratio in F1 hybrids
OZ
ZO
a/b
ZF
OF
d-score
(A/B) / (a/b)
cis-
trans-
cis- & trans-
158 genes
271 genes
183 genes
246 genes
383 genes
183 genes
B
HVC Genes (5,827)
RA Genes (6,328)
121
209
37
212
324
59
143
143
40
966
3523
1259
579
587
147
HVC 2.7% RA 3.9%
cis-regulated
HVC 4.7% RA 6.1%
trans-regulated
HVC 3.1% RA 2.9%
cis- & trans-regulated
HVC 77.0% RA 75.6%
conserved
HVC 12.5% RA 11.6%
ambiguous
C
HVC Genes (5,827)
RA Genes (6,328)
319
140
435
424
173
628
420
235
451
1144
2976
1277
HVC 7.8% RA 9.1%
cis-regulated
HVC 10.2% RA 12.7%
trans-regulated
HVC 11.2% RA 10.8%
cis- & trans- regulated
HVC 70.8% RA 67.4%
conserved

**Figure II-6. Classification of transcriptional regulation divergence between ZF and OF.**

**(A)** Heatmaps of gene expression in ZFs and OFs, and allelic expression ratios in F1 hybrids for *cis*-, *trans*-, and both *cis*- and *trans*-regulated genes in song nuclei HVC and RA (blue-red colored). Comparison between species-different gene expression (A/B) and allelic expression ratios in F1 hybrids (a/b) in heatmaps (Dark brown-light yellow colored). "A" and "B" represent $RPKM_{(ZF\ average)}$ and $RPKM_{(OF\ average)}$, respectively. "a" and "b" represent $Reads_{(ZF\ allele)}$ and $Reads_{(OF\ allele),}$ respectively.

**(B)** Numbers of gene obtained by the new method classified as *cis*-, *trans*-, both *cis*- and *trans*-, conserved, and ambiguous regulation in HVC and RA.

**(C)** Numbers of genes obtained by the traditional standard method classified as *cis*-, *trans*-, both *cis*- and *trans*-, conserved regulation in HVC and RA.

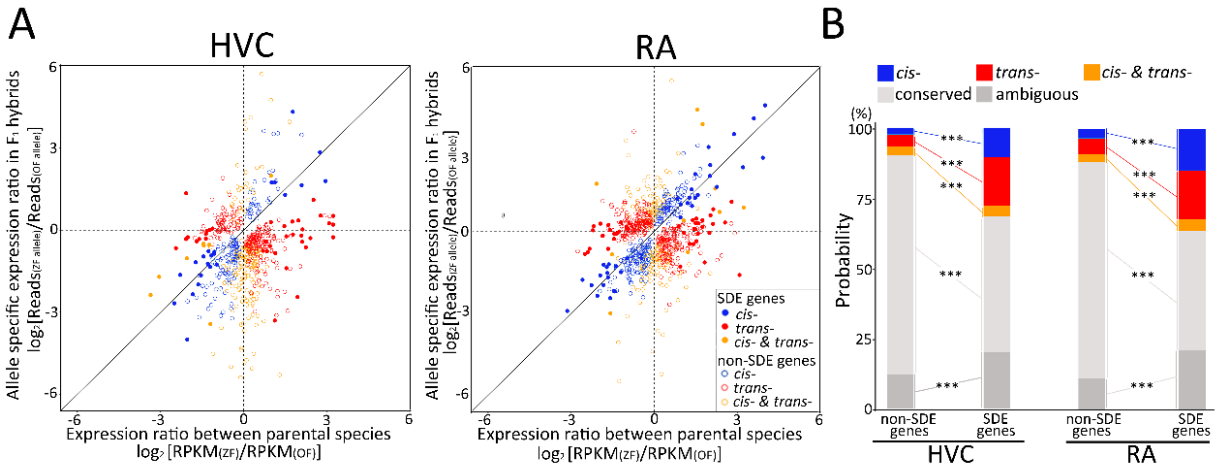



**Figure II-7: *Cis*- and *trans*-regulatory effects on species-differential expression.**

**(A)** Scatter plots of expression ratios between ZF and OF (X-axis) and allelic expression ratios in $F_1$ hybrids (Y-axis) for genes showing differential expression between species. Blue-, red-, and orange-colored spots: *cis*-, *trans*-, both *cis*- and *trans*-regulated genes, respectively. Filled spots correspond to species-differentially expressed (SDE) genes.

**(B)** *Cis*- and *trans*-effects on the expression of species-differentially regulated genes. The percentage of *cis*-, *trans*-, both *cis*- and *trans*-, conserved, and ambiguous transcriptional regulatory genes in the SDE and non-SDE genes (Fisher's exact test, ***$p < 0.001$).

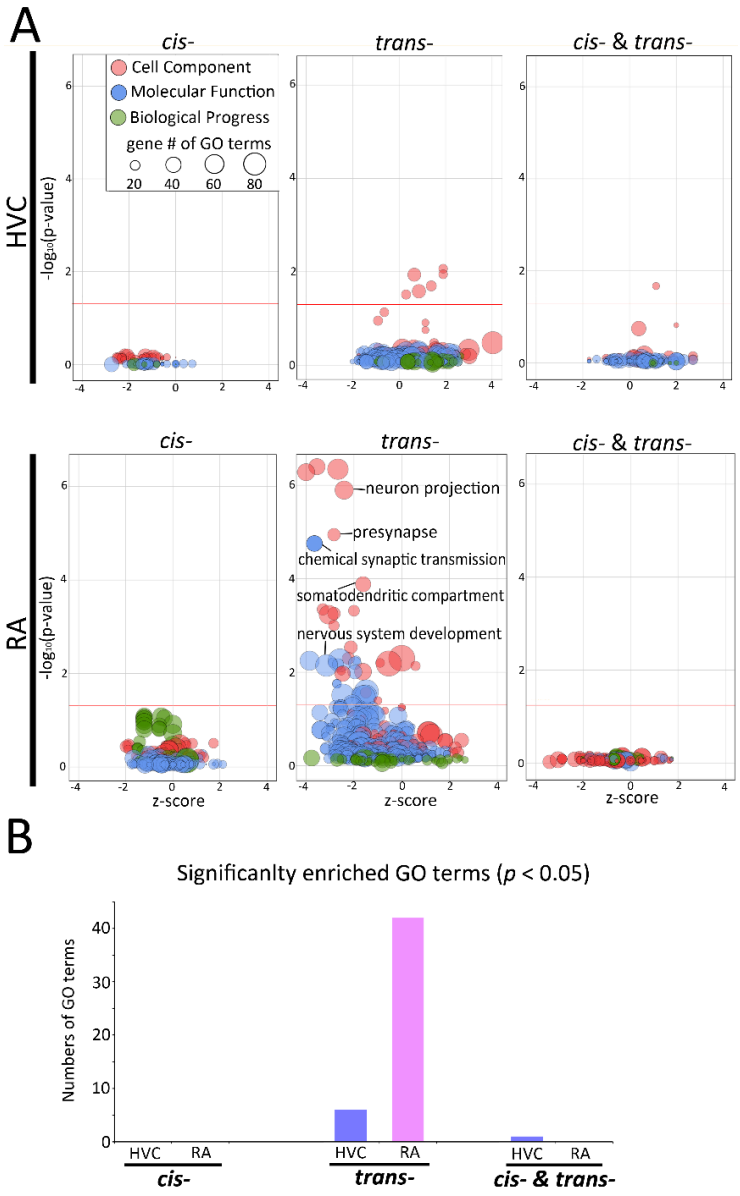



**Figure II-8: GO enrichment analysis of regulatory different genes.**

**(A)** Gene ontology (GO) enrichment analysis of the *cis*-, *trans*-, and both *cis*- and *trans*-regulated genes in HVC and RA. The size of circle represents the number of genes assigned to each GO term. Red lines represent *p*-value for significant enrichment (Fisher's exact test, *p*-value adjusted by Benjamini-Hochberg's method, $p < 0.05$).

**(B)** The bar graph shows the numbers of enriched GO terms of each category of regulatory different genes.

A

| Rank | Upstream molecule | Molecule Type | Downstream genes numbers | *p*-value |
|---|---|---|---|---|
| 1 | BDNF | Nerve growth factor | 21 | 6.44E-07 |
| 2 | HTT | Unclear | 26 | 1.06E-05 |
| 3 | POU3F1 | Transcription factor | 3 | 1.33E-04 |
| 4 | MAPT | Microtubule-associated protein | 15 | 1.55E-04 |
| 5 | MKNK1 | MAP Kinase-interacting Ser/Thr Kinase | 9 | 4.25E-04 |
| 6 | PSEN1 | γ-secretase | 18 | 7.85E-04 |
| 7 | HDAC4 | Histone deacetylase | 7 | 8.34E-04 |

B

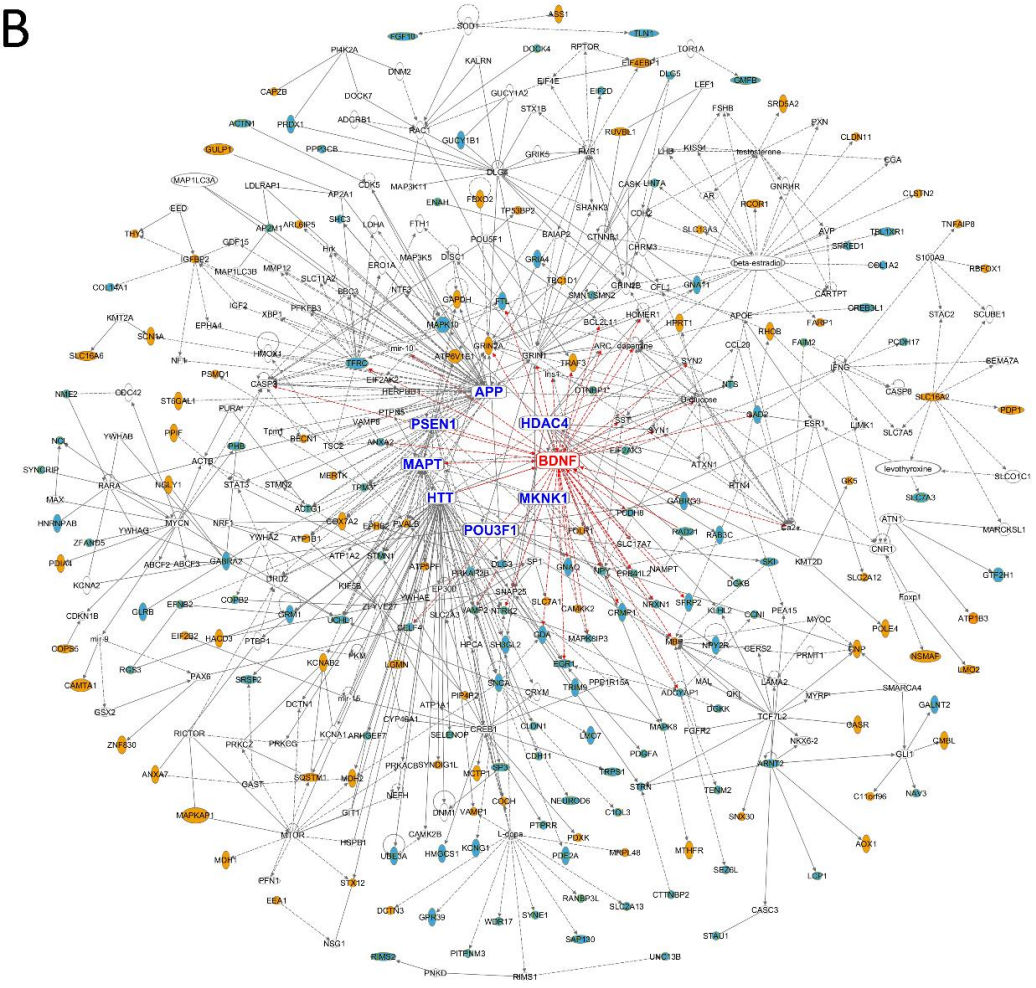



**Figure II-9: Upstream regulatory molecular enrichment analysis of *trans*-regulatory genes in RA.**

**(A)** Top 7 candidate upstream mediators for *trans*-regulated genes in RA.

**(B)** Gene-gene connections for the top 7 candidate upstream genes. Colored genes are *trans*-regulated genes in RA. Solid and dotted lines represent directed and undirected regulation, respectively, between connected genes.

# Chapter III

# A potential contribution of BDNF expression on the generation of species-specific birdsongs

## 3.1 Introduction

More than 4,000 species of songbird produce species-specific unique songs as a function of the conserved song pathway (Nottebohm *et al.*, 1976; Marler & Slabbekoorn, 2004; Brenowitz & Beecher, 2005). There are many genes that are regulated by singing activation or development of song in song nuclei of the song pathway. (Jarvis *et al.*, 1998; Haesler *et al.*, 2004; Wada *et al.*, 2006; Hayase & Wada, 2018; Hayase *et al.*, 2018). However, little is known about how the genetic difference between songbird species in the song pathway contributed to species-specific vocal learning behavior via species-specific gene expression level/pattern in the song pathway among songbird species.

The post-genomic era focuses on elucidating the function of genes (Eisenberg *et al.*, 2000; Pandey & Mann, 2000; Husi & Grant, 2001). Identifying, narrowing down, and verifying the importance of candidate genes in animal behavior is one of the key research directions of neuroscience and neuroethology in the era. In **Chapter II**, I found that the existence of RA *trans*-regulatory genes between ZF and OF correlated with neural functions and identified some candidates as upstream mediators for the *trans*-regulatory genes in RA (**Fig. II-9**). These upstream mediators may contribute to differently regulate their downstream genes to produce species-specific vocal learning behavior between ZF and OF. Furthermore, I identified the brain-derived neurotrophic factor (BDNF) as the most significant upstream mediator that could be responsible for *trans*-regulatory difference in RA between ZF and OF (**Fig. II-9A**). BDNF, a member of the neurotrophin family, plays an important role in memory and cognition by binding

to tropomyosin receptor kinase B (TrkB) receptor (Lu, 2003; Kowiański *et al.*, 2018). The expression level of BDNF is regulated by neural activity and the protein is transported and secreted to the downstream neural pathway (Altar *et al.*, 1997). In recent years, studies have elucidated the potential function of BDNF related to vocal learning behavior in songbirds. The expression level of BDNF mRNA in the song pathway was increased by singing (Li *et al.*, 2000; Wada *et al.*, 2006). Furthermore, the expression level of BDNF in the song nuclei, such as Area X, HVC, and RA, dynamically changed during song development (Akutagawa & Konishi, 1998). Based on this information, I hypothesized that the difference of BDNF could differently regulate the downstream genes expression in RA between ZF and OF as a *trans*-regulatory mediator, which play important roles in the regulation of species-specific songs.

There are two possible mechanisms by which BDNF could induce *trans*-regulatory differences in RA between ZF and OF. First, gene mutation might have caused amino acid substitutions in the BDNF protein between ZF and OF and consequently affected its ability to bind on the TrkB receptor. Secondly, expression differences in the BDNF mRNA in HVC neurons projecting to RA, or in RA neurons themselves could have affected the probability to bind to the TrkB receptor. I first tested whether amino acid substitution or expression level differences of the BDNF gene existed or not between ZF and OF. For this purpose, I cloned BDNF genes from ZF and OF and other species to explore whether amino acid substitutions generally exist among songbird species. Secondly, I examined the potential species-different expression level of the BDNF

mRNA between ZF and OF under silent and singing conditions. To test the contribution of amino acid substitution and expression level difference of the BDNF gene to the species-specific songs, I performed correlation analyses between singing phenotypes and allelic imbalance ratio or expression level of the BDNF gene by utilizing F1 hybrids. Furthermore, to verify the potential BDNF function for regulation of species-specific songs as a *trans*-regulatory mediator, I performed microdialysis with a selective agonist of BDNF receptor TrkB, 7,8-Dihydroxyflavone (7,8-DHF) into the RA of adult ZFs.

## 3.2 Materials and methods

***Clone of the BDNF gene from different songbird species***

To compare the amino acid sequence of BDNF protein in multiple songbird species, I cloned the BDNF gene from 8 species of songbirds. Total RNA was extracted from whole brain tissues of ZF (*T. guttata*, n = 14), OF (*T. bicherovii*, n = 14), star finch (*Neochmia ruficauda*, n = 3), cherry finch (*N. modesta*, n = 3), Bengalese finches (*Lonchura striata domestica*, n = 3), spice finch (*L. punctulata*, n = 3), Java sparrow (*L. oryzivora*, n = 3), and canary (*Serinus canaria*, n = 3). Total RNA was transcribed to cDNA using Superscript Reverse Transcriptase (Invitrogen) with oligo dT primers. The cDNAs were amplified by PCR using oligo DNA primers directed to the start and stop regions of ZF BDNF. The sequences of primers were as follows.

Forward: 5'-ATGACCATCCTTTTCTTTACTA-3'

Reverse: 5'-CTATCTTCCCCTTTTAATGGT-3'

PCR products were ligated into the pGEM-T Easy plasmid (Promega). The cloned sequences were searched using NCBI BLAST/BLASTX to compare with homologous genes in other avian and mammalian species.

***RI In-situ hybridization and quantification of BDNF mRNA expression level***

$^{35}$S-labeled UTP was used to synthesize riboprobes from the SP6 (Roche) transcription promoter site of pGEM-T Easy plasmid with ZF BDNF gene insert. RI *in-situ* hybridization procedures were used as previous research (Wada *et al.*, 2013a). ZFs

and OFs were sacrificed after singing 3 hours. Whole brains were frozen in OCT compound on dry ice immediately and stored at -80°C. The brains frozen in OCT compound were sectioned with 12 μm thickness. Brain sections were fixed in 3% paraformaldehyde/phosphate-buffered saline (PBS), washed 3 times in 1× PBS, acetylated, washed in 2× SSPE, dehydrated in ascending ethanol concentrations (50, 70, 90, and 100%), and then air-dried. Riboprobe ($10^6$ CPM) was mixed with 120 μl hybridization solution to cover the brain sections. Sliders were put into an oil bath for 14 hours at 65°C. After hybridization, hybridization solution was rinsed by 2× SSPE, 50% formamide in 2× SSPE, 0.1×SSPE twice and dehydrated in increasing ethanol concentrations (50, 70, 90, and 100%), air-dried. BioMax MR film (Kodak, USA) was covered on the brain section for 1 week, and then development. Photos of developed X-ray films were taken under a microscope (Z16 Apo, Leica) which connected to CCD camera (DCF490, Leica) with Leica Application Suite, v3.3.0 (Leica). Light and camera settings were the same for all images to ensure unbiased comparison. Photos were converted to grayscale, and mRNA expression levels were quantified as mean pixel intensities using Adobe Photoshop CS2 (Adobe Systems, USA).

***Correlation analysis between song phenotypes and expression level or allelic expression ratio of BDNF in F1 hybrids***

RPKM values of the BDNF gene in RA of each F1 hybrid individual were calculated depending on the read mapping result. Allelic expression ratios of BDNF in each F1 hybrid were got from the result of allelic imbalance analysis result in **Chapter II**. Since

there were 5 acoustic (syllable duration, inter-syllable gap duration, entropy variance, mean frequency modulation [FM], and FM variance) and 2 sequence (motif and repetition transition rates) parameters showed significant difference both in con-specific and cross-specific tutored ZF and OF (**Fig. I-3B, 3C**), I performed correlation analysis between expression level or allelic expression ratio and the 7 song parameters (Pearson's correlation).

***Pharmacological manipulation in RA of adult ZFs***

Custom microdialysis probes were built using a microdialysis membrane (SpectralPor, in vivo microdialysis hollow fiber, Outside diameter = 216 μm, total weight < 0.035g) attached to a drug reservoir, as previously described (Andalman & Fee, 2009). I implanted probes bilaterally at positions adjacent to RA using stereotaxic coordinates. Before setting the probe, spontaneous neural activity was measured to verify the location of RA. Following surgery, the reservoir was filled every morning with saline until the bird began to sing consistently and its phonological and syntactical features were confirmed not to be damaged by probe implantation. To verify the position of microdialysis probes, tetrodotoxin (TTX, 6–12 μM) was infused into one side of RA and a hemi-RA inactivation-induced song change was confirmed. The 7, 8-DHF (10 μg/μL in 0.9% NaCl, pH 7.4–7.6, Santa Cruz; n = 5 birds), or saline (n = 3 birds) were then continuously infused during the day. The dosage of 7,8-DHF was chosen on the basis of previous *in vivo* studies (Blugeot *et al.*, 2011; Devi & Ohno, 2012). The manipulated birds were allowed to move freely in a sound attenuation

chamber and the song of each individual was recorded over 10 days after the initiation of drug infusion. Probe position was evaluated post mortem by histological staining of tissue sections.

***Golgi staining and determination of dendritic spine density***

To investigate whether 7,8-DHF infusion affected the morphology of neurons, I performed Golgi staining for the 7,8-DHF (n = 3) and saline (n = 3) infused birds, as described by sliceGolgi Kit (Bioenno Tech). Shortly to say, birds were anesthetized under the silent and dark condition and perfused with 1× PBS. Brains were sampled and sectioned at 100 μm thickness. Sections were fixed in a fixative solution for 3 hours and then immersed section into impregnation solution for 5 days. Thereafter, sections were stained and washed. Subsequently, sections were dehydrated in ethanol and xylene and covered slip using the Permout® mounting medium. Photos of stained sections were taken by using BZ X710 Microscope (Keyence) at 100× magnification. Z stacked images were formed using 100 sections with 20 μm depth. 6 dendrites from 3 ~ 4 neurons were selected from each bird to count the number of dendritic spines. Dendritic spines were classified as previous research (Spacek & Harris, 1997; Tyler & Pozzo-Miller, 2003) as follows: type A (Dendritic spine with longer spine neck, mushroom shape); type B (dendritic spine without spine neck, stubby shape). The density of each type of dendritic spine was calculated as the number of spines per micrometer.

***Sampling of RA tissue and next generation sequencing***

To prepare the library for next generation sequencing, I sampled the RA tissue of saline or 7,8-DHF infused birds and extracted total RNA. Whole brains were frozen in OCT compound on dry ice immediately. Brains were serially cryosectioned at 20 μm thickness onto handmade membrane slides and stained in Nissl to confirm the presence and boundaries of RA. RA tissues were microdissected using a laser capture microscope (Arc-turusXT; Arcturus Bioscience) settings as follows: spot diameter, 100 μm; laser power, 80 mW; and laser duration, 80 ms. Captured tissues were dissolved in RLT buffer (Qiagen) with *β*-mercaptoethanol (Wako). Total RNA was purified based on the manufacturer's instruction using Agencourt AMPure XP (BECKMAN COULTER). RIN value and RNA concentration were measured using a Bioanalyzer 2100 (Agilent Technologies).

Quartz-seq was performed to amplify the total quantity of nucleate by following the previous research (Sasagawa *et al.*, 2013). Extracted total RNA (1 ng) from RA (7,8-DHF, n = 3; Saline, n = 3) were used to synthesize first strand cDNA. The PCR amplification condition was: 18 cycles of 98°C (10s), 65°C (15s), 68°C (5min). Amplified cDNA samples were purified by using PCR purification column (MiniElute PCR Purification Kit; Qiagen). The concentrations and RIN value of cDNA samples were checked using Bioanalyzer 2100 (Agilent Technologies) to check the quality of amplified cDNA. Amplified cDNA samples were fragmented using a DNA Shearing System LE220 (Covaris) and then purified using column. Following end repair of DNA fragments, adaptors were ligated and amplified using a ligation-based Illumina

multiplex library preparation method (LIMprep). All libraries were sequenced using the Illumina Hiseq2500 platform for 150 bp paired-ends. The library construction and sequencing steps were performed at Novogene, Bejing.

***Gene expression quantification of RA***

After sequencing, raw data were filtered to remove low quality reads. Filtered clean reads were mapped to the reconstructed zebra finch genome (see material and method in **Chapter I**) using TopHat2 software (Kim *et al.*, 2013). RPKM values of all genes in each individual were calculated to estimate the gene expression level using Cufflinks software (Trapnell *et al.*, 2010) based on the new GTF file (see in material and method in **Chapter I**). Differentially expressed genes between the control and 7,8-DHF infused birds were identified by using R package DEseq2 (Love *et al.*, 2014). Significantly different expressed genes between the control and 7,8-DHF infused birds were defined as genes with a foldchange of more than 4 and *p*-value less than 0.05.

## 3.3 Results

***Amino acid substitutions exist in the BDNF protein among songbird species***

To figure out the candidate mediator for *trans*-regulatory genes in RA, I performed upstream mediator enrichment analysis in **Chapter II**. I found that BDNF was the most significant enriched mediator (**Fig. II-9A**) which might regulate 21 *trans*-regulatory different genes in RA (**Fig. III-1A**). To test the possibility of existence of the amino acid substitution in BDNF between ZF and OF, I performed sequence assembly of the BDNF gene from the RNA-seq results to compare the amino acid sequences between ZF and OF (n = 4, each). There were two amino acid differences (Ser45Arg in the prodomain and Thr143Met in the NGF domain) in BDNF protein between ZF and OF (**Fig. III-1B**). In addition, to test whether such amino acid mutation commonly existed among songbird species or not, I cloned the BDNF coding sequence from a total of 8 species of songbirds (zebra finch, owl finch, star finch, cherry finch, Bengalese finch, spice finch, Java sparrow, and canary). After sequencing, I transformed the mRNA sequences into protein sequences and aligned the protein sequences among the 8 species. A total of 6 amino acid substitutions (five amino acid changes located in the predomain and one in the NGF domain) were found in the BDNF protein among the 8 songbird species, showed that amino acid substitutions in BDNF have repeatedly occurred with song species through evolution (**Fig. III-2**).

***Different expression level of BDNF mRNA between ZF and OF in the motor pathway***

The different expression levels of the BDNF gene between ZF and OF might be a reason that induced BDNF to be a trans-regulatory mediator. To exam whether the expression level of BDNF between ZF and OF were different or not, I evaluated the expression level of BDNF mRNA in HVC and RA between ZF and OF using the expression data from RNA-seq in **Chapter I**. As the result, expression levels of BDNF mRNA were significantly different in HVC between ZF and OF under silent condition (student's *t*-test, $p < 0.05$) **(Fig. III-1C)**. Even though expression level of BDNF in the whole brain and RA did not show significant differences between ZF and OF, the expression levels of BDNF showed same trends that higher expression levels in OF than ZF.

Previous research showed that singing behavior increased the expression level of the BDNF mRNA (Li *et al.*, 2000; Wada *et al.*, 2006). To examine the potential species difference in BDNF expression level induced by singing behavior, I performed *in-situ* hybridization of BDNF mRNA for ZF and OF brains under both silent and 3 hours singing conditions. The expression levels of the BDNF mRNA in both HVC and RA of OFs were significantly higher than ZFs under both silent and 3 hours singing conditions (Student's *t*-test, $p < 0.05$) (**Fig. III-3**). In contrast, such different expression levels were not found in the surrounding non-song nuclei area of HVC and RA (e.g., caudal nidopallium and archopallium, respectively) (**Fig. III-3**). These results showed that expression levels of BDNF mRNA were significantly different between ZF and OF both in HVC and RA under both silent and singing conditions, which further support the idea

that expression level of BDNF may play an important role as *trans*-regulatory mediator for maintaining species-specific features of the adult song.

***Song phenotypes had a higher correlation with the expression level of the BDNF mRNA than amino acid substitutions in F1 hybrids***

To examine whether the putative *trans*-regulatory effects are mediated by amino acid substitution or difference in the expression level of BDNF, I used F1 hybrids to perform correlational analyses between song phenotypes and ASE ratios or RPKM values of BDNF. ASE ratios were calculated depending on the expression levels of ZF and OF allele, which reflected the percentage of BDNF protein with ZF or OF amino acid sequence in F1 hybrids. In addition, the RPKM values quantified the mRNA expression level of the BDNF gene. The transcriptome analysis in F1 hybrids showed that F1 hybrids possessed a wide range of individual differences in the ASE ratio and expression level of BDNF mRNA in HVC and RA (**Fig. III-4A**). Furthermore, F1 hybrids acquired individually-unique songs with a wide range of ZF- and OF-biased features, even though they were tutored with both ZF and OF songs (**Fig. III-4B**). I performed correlation analyses between ASE ratios or expression levels of BDNF and 5 acoustic (syllable duration, inter-syllable gap duration, entropy variance, mean frequency modulation [FM], and FM variance) and 2 sequential (motif and repetition transition rates) parameters of the song, which were identified as species different song parameters (**Fig. I-3**). I found that only one parameter, entropy variance of syllables, showing significantly correlated with the ASE ratio of BDNF in RA ($r$ = 0.800, $p$ =

0.017, Pearson's correlation) (**Fig. III-4C**). In contrast, the expression level of BDNF mRNA in HVC significantly correlated with two acoustic and two sequential song parameters in F1 hybrids [acoustics: syllable duration ($r$ = −0.862, $p$ =0.006) and entropy variance ($r$ = −0.822, $p$ = 0.012); sequence: motif ($r$ = −0.762, $p$ = 0.028) and repetition ($r$ = 0.729, $p$ =0.040), Pearson's correlation] (**Fig. III-4C, D**). These results suggest that the expression level of BDNF mRNA in HVC, not amino acid substitutions, being the most likely RA *trans*-acting mechanism that regulated expression difference in the downstream genes between ZF and OF.

***Pharmacological activation of BDNF receptors in RA altered species-specific song features in male ZF adults***

The correlation analyses in F1 hybrids suggested that differences in the expression level of BDNF in HVC and RA should contribute to generating the *trans*-regulatory function for species-specific songs between ZF and OF. To test this idea, I performed local infusion of a BDNF receptor tropomyosin receptor kinase B (TrkB) agonist 7,8-DHF, into RA of adult ZFs (**Fig. III-5**), using microdialysis probes. Then, I evaluated the effects of receptor activation on song structure. Similar to a previous report (Kittelberger & Mooney, 2005), ZF songs changed after over-activation of BDNF receptor, with a lower consistency in syllable transition during the early stage (~ 5 days after drug infusion). However, following continuous infusion for up to 2 weeks, adult structured songs gradually changed phonologically and sequentially leading to the loss of ZF species-specific features and getting close to OF song features (**Fig. III-6**).

Although syllable sequence (i.e., motif and repetitive indexes) and some acoustic parameters (inter-syllable gap duration, entropy variance, and FM variance) were changed to be similar with OF song features by the infusion of 7,8-DHF, other acoustic parameters (syllable duration and mean FM) were still maintained with ZF-specific traits (**Fig. III-6C,E**), indicating that manipulating the activation of BDNF receptors did not simply cause a completely atypical song structure, but rather changed some species-specific song features. These results suggested that a precise amount of BDNF contributes to the maintenance of species specificity in the ZF song, supporting our earlier finding that BDNF is a potential regulatory mediator of the RA *trans*-regulated genes associated with the generation of species-specific song.

***Over-activation of BDNF receptor changed the expression level of putative trans-regulated downstream genes in RA***

A total of 21 *trans*-regulatory genes were predicted as downstream genes of BDNF by IPA analysis (**Fig. III-1A**). To verify the reliability of predicted downstream genes, I performed RNA-seq using RA tissues of control and 7,8-DHF infused birds whose songs were degraded at two weeks of the agonist infusion. The RPKM values of each gene were calculated to estimate its expression levels. I found that a total of 11,655 genes expressed in RA (RPKM >1). Among these genes, 570 genes were identified as differentially regulated with more than 4-fold changes between the control and 7,8-DHF infused birds (DEseq2, $p < 0.05$). Of the differentially expressed 564 genes, 6 of 21 putative downstream *trans*-regulated genes of BDNF had significantly altered

expression by over-activation of BDNF receptor (**Fig. III-7**). This result further supported the finding that the expression level of BDNF was a potential regulatory mediator of the RA *trans*-regulated genes.

***Over-activation of BDNF receptor changed the morphology of dendritic spines in the RA neurons***

Since over-activation of the BDNF receptor altered the species-specific songs of zebra finch adults, I then investigated whether over-activation of the BDNF receptor affected the dendritic spine morphology of neurons in RA by Golgi staining of control and 7,8-DHF infused birds. Following on previous studies (Spacek & Harris, 1997; Tyler & Pozzo-Miller, 2003), I defined two types of the dendritic spine: type A with longer spine neck and type B with short or no spine neck (**Fig. III-8A**). I found that the density of type A dendritic spine decreased after the over-activating of the BDNF receptor. On the other hand, the density of type B was increased after the over-activation of the BDNF receptor (**Fig. III-8B**). These results suggest that over-activation of the BDNF receptor altered the song phenotypes by affecting the synapse morphology or transmission of RA neurons which interconnected with HVC projected neurons in zebra finch adults.

## 3.4 Discussion

In this chapter III, I tested whether expression level difference or amino acid substitution of BDNF gene could be the key factor for generating different *trans*-regulatory effects between ZF and OF. By focusing on F1 hybrids, I found that the difference in BDNF expression level between ZF and OF had a greater association with species-specific song phenotypes than amino acid substitutions. BDNF is a secretory cell signalling protein that regulates neural differentiation and neural plasticity in vertebrates (Barde *et al.*, 1982; Leibrock *et al.*, 1989; Lu, 2003). Previous studies showed that singing behavior enhances the expression level of BDNF in the HVC neurons projecting to RA in songbirds (Li *et al.*, 2000; Wada *et al.*, 2006), suggesting that BDNF protein produced in HVC neurons may be transported and secreted to RA to regulate synaptic plasticity of RA neurons. Short-term injection of BDNF protein itself into RA of adult zebra finch resulted in juvenile-like songs with variability in sequential and acoustic features (Kittelberger & Mooney, 2005). In my study, local infusion of BDNF receptor agonist 7,8-DHF into RA of adult ZF induced severe degradation of species-specific song features at both acoustic and sequence levels (**Fig. III-6**). The RNA-seq results confirmed that the expression level of six predicted downstream genes of BDNF changed after 7,8-DHF infusion (**Fig. III-7**), suggesting that BDNF is a *trans*-regulatory mediator whose expression level may be crucial for maintaining species-specific adult songs. These results suggest the expression level of BDNF in HVC neurons projecting to RA, were critical for maintaining the dendritic spine morphology of neurons in RA. Previous research found that RA receives a direct

projection from two song nuclei, HVC and LMAN (Nottebohm *et al.*, 1982; Doupe & Konishi, 1991). In my study, I could not provide direct evidence that the expression level of BDNF in HVC neurons projecting to RA but not in LMAN neurons projecting to RA play roles for the species-specific song. In future research, it is necessary to over-express BDNF mRNA in HVC and LMAN separately and verifying how it effects the song of adult ZF is needed.

Over-activation of the BDNF receptor not only degraded the song of adult ZF but also changed the song features to be similar with OF songs (**Fig. III-6 C, E**). These results suggested that the molecular or morphological characters of neurons in the RA of ZF adults were changed and tend to OF patterns after the over-activation of the BDNF receptor. Consistent with this result, the expression level of around 80% predicted downstream genes of BDNF (foldchange > 2) in the RA of ZF tend to OF expression levels after 7,8-DHF perfusion (**Fig. III-7**). Similar to the RNA-seq results, it can be predicted that the morphological characteristic of RA neurons might change to OF like after over-activation of the BDNF receptor even I don't know the result of Golgi staining of RA in OF. The Golgi staining results showed that the density of type B dendritic spines was significantly increased after over-activation of the BDNF receptor (**Fig.III-8**), which was in line with previous research (Tyler & Pozzo-Miller, 2003). Previous research showed that short and stubby spines (type B) are with higher synaptic strength than spines with long and narrow necks (type A) (Segal *et al.*, 2000; Nimchinsky *et al.*, 2002). Meanwhile, the synapses with type B dendritic spines are more mature and stable than type A which play critical roles for learning and memory

(Bourne & Harris, 2007; Harris & Spacek, 2016; Gipson & Olive, 2017). That means interactivities of RA neurons were increased and then changed the adult song structures of ZF by over-activation of the BDNF receptor. Similar to the molecular changes, morphological characteristics of RA neurons in ZF may change to OF-like by BDNF receptor over-activation, which suggested that the percentage of type B spines is higher in RA neurons of OF adults than ZF adults, which may be one of the important reasons inducing species-specific songs. However, there are multiple cell types of neurons in the RA of songbirds (Spiro *et al.*, 1999; Leonardo & Fee, 2005; Ölveczky *et al.*, 2011). Therefore, it is necessary to identify TrkB-specific expressed neurons in RA since only such neurons are stimulated by BDNF releasing. In the future, performing single-cell RNA sequencing to identify TrkB-specific expressed neurons in RA and comparing the morphological difference of such neurons between ZF and OF would be essential to further understand how the expression level of BDNF contribute to species-specific songs.

Although differences in the concentration of BDNF protein in RA was important for species-specific songs between ZF and OF (**Fig. III-6**), the possibility that amino acid mutations in BDNF protein should not be ruled out for the evolution of species-specific vocal learning behavior. Interestingly, BDNF polymorphisms (Val66Met; rs6265) affect the intracellular trafficking and reduces the activity-dependent secretion of mature BDNF (Chen *et al.*, 2004). Furthermore, BDNF polymorphisms cause dendritic spine density alternation, memory formation, and extinction (Egan *et al.*, 2003; Giza *et al.*, 2018). In this chapter, I found 6 amino acid polymorphisms among 8 species of

songbird (**Fig. III-2**). In the future, it is necessary to verify the contribution of polymorphisms in BDNF protein to species-specific vocal learning behavior by gene manipulation. For example, perform overexpression of BDNF protein with OF amino acid sequence in the HVC of ZF adults by virus injection and investigate how the song of ZF adults was affected.

## 3.5 Figures

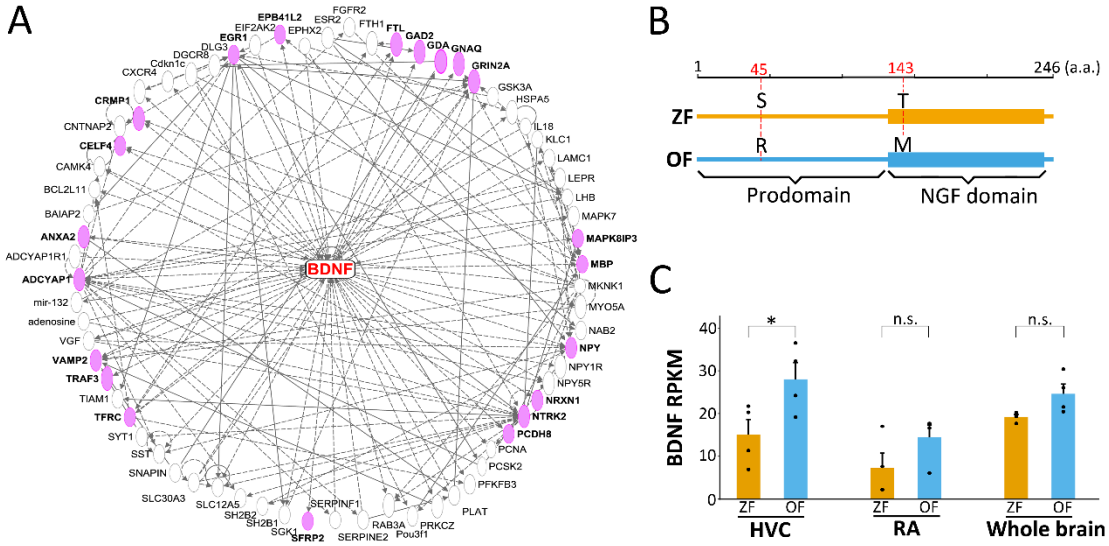



**Figure III-1: BDNF is candidate as an upstream mediator for *trans*-regulated genes**

**(A)** Gene-gene connections for downstream genes of BDNF. Pink-coloured genes are *trans*-regulated genes in RA. Solid and dotted lines represent directed and undirected regulation between connected genes, respectively.

**(B)** Alignment of the amino acid sequences between ZF and OF. Numbers at the top are the amino acid positions. Red numbers represent mutation positions.

**(C)** Differences in the expression level of BDNF in HVC, RA, and whole brain between ZF and OF. Coloured bars represent average RPKM values of BDNF in HVC, RA and whole brain of ZF and OF; each dot is the RPKM value for each individual. (Data: Mean ± SEM; n = 4 each; student's t-test, *p < 0.05, n.s: no significant)

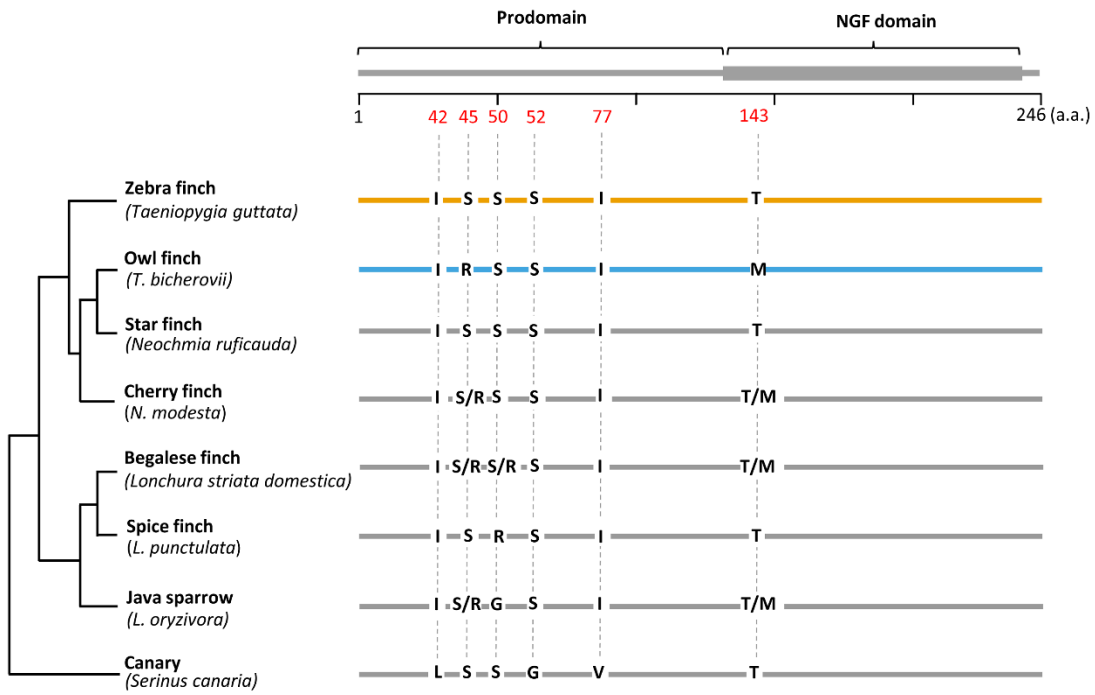



**Figure III-2: Species differences in the amino acid sequences of BDNF.**

Right: The phylogenetic tree was adapted from a previous research (Imai *et al.*, 2016). Lines for each species represent the amino acid residues which is conserved among species. Amino acid residues on the lines meant position with mutations. One mutation position with diagonal separated two amino acid residues meant animal individuals were heterozygote at this position.

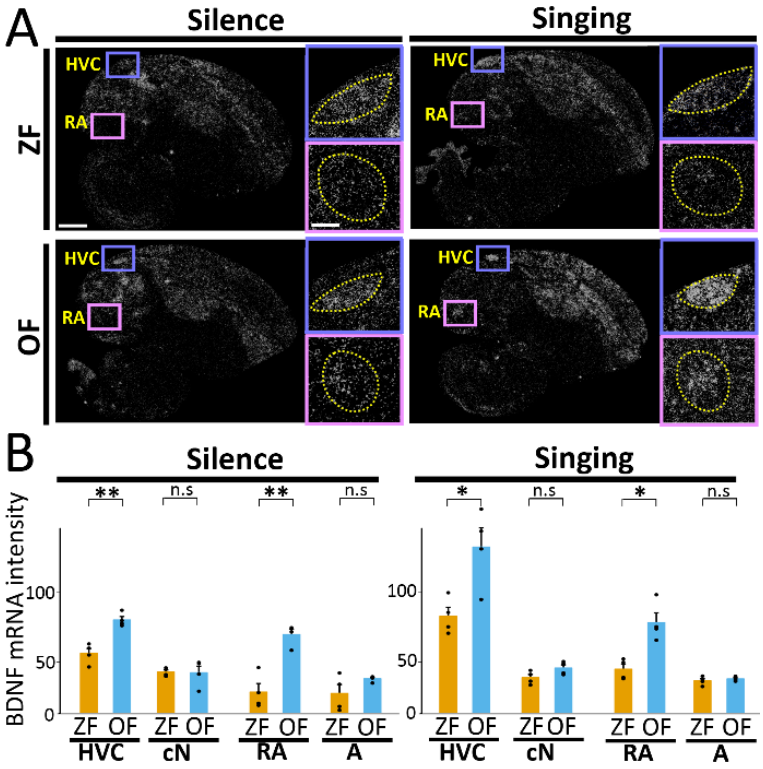



**Figure III-3: Expression level of BDNF in HVC and RA under singing condition.**

**(A)** Examples of whole brain, HVC, and RA images, showing expression of BDNF in ZF and OF at the condition of silence and 3 singing hours. Scale bars are 1 mm (left panes) and 0.2mm (right panel). White signals: BDNF mRNA.

**(B)** Expression level of BDNF mRNA in HVC, RA, and their respective surrounding areas [caudal nidopallium (cN) and archopallium (A), respectively) of ZF and OF at the condition of silence and 3 hours undirected singing (n = 4 each). Coloured bars are average values of BDNF mRNA expression in HVC, RA and surrounding areas. Each dot is the BDNF mRNA expression level of each individual. (Data: Mean ± SEM; n = 4 each; Student's *t*-test, $^{**}p < 0.01$, $^{*}p < 0.05$, n.s: not significant)

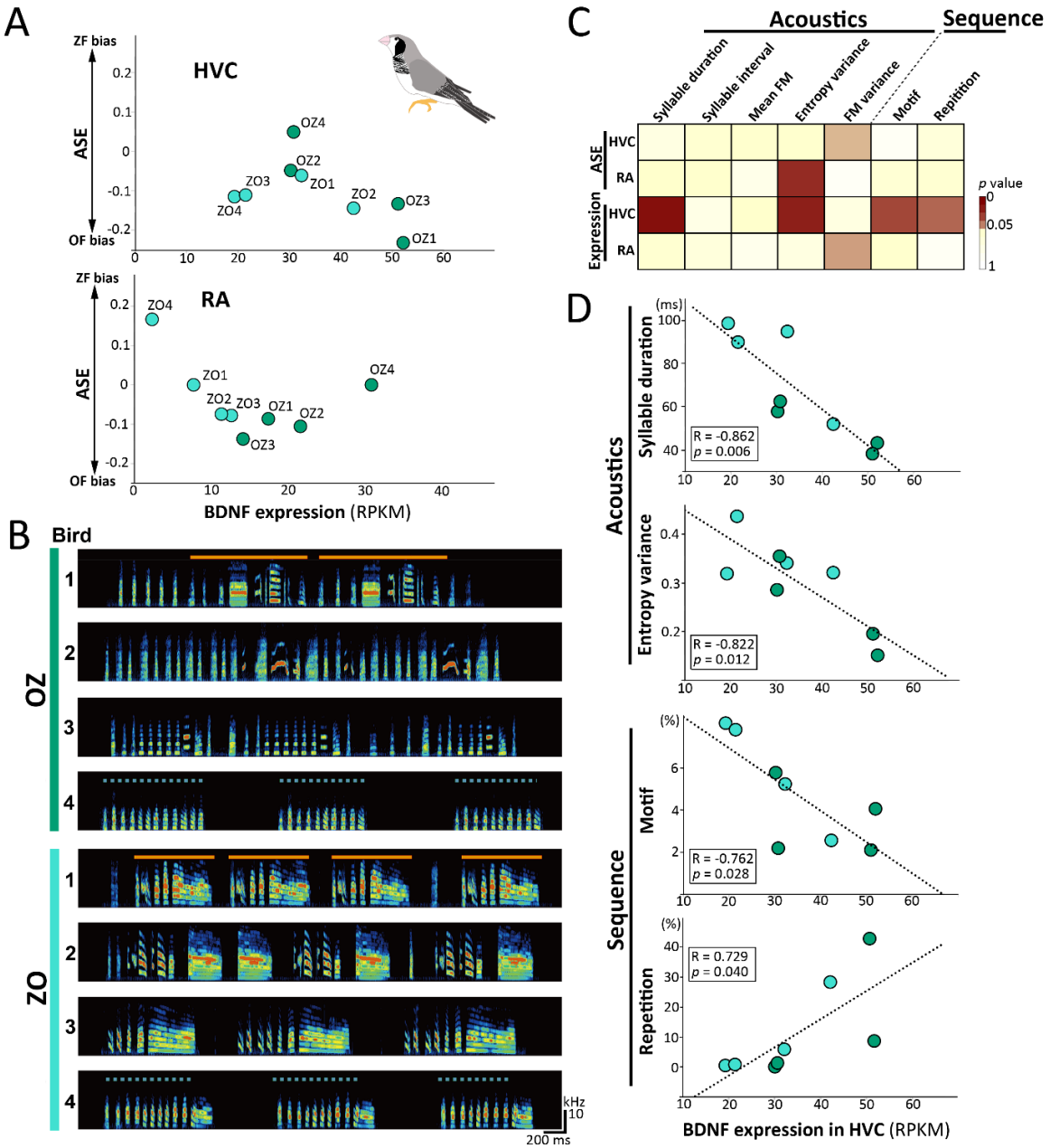

A
HVC
ZF bias
OF bias
ASE
OZ4
OZ2
ZO3
ZO1
ZO4
ZO2
OZ3
OZ1
RA
BDNF expression (RPKM)
B
Bird
OZ
ZO
kHz
10
200 ms
C
Acoustics
Sequence
Syllable duration
Syllable interval
Mean FM
Entropy variance
FM variance
Motif
Repitition
ASE
Expression
HVC
RA
p value
0
0.05
1
D
Acoustics
Syllable duration
R = -0.862
p = 0.006
Entropy variance
R = -0.822
p = 0.012
Sequence
Motif
R = -0.762
p = 0.028
Repetition
R = 0.729
p = 0.040
BDNF expression in HVC (RPKM)

**Figure III-4: Correlation between individual variation in the expression level of BDNF and species-biased song features in F1 hybrids**

**(A)** Individual variation in BDNF mRNA expression level and ASE ratios between F1 hybrids individuals.

**(B)** Individual variation in learned songs in F1 hybrids tutored with ZF and OF songs. Orange solid and blue dotted lines represent the motif and repetitive structures of song, respectively.

**(C)** Heat maps showing correlation *p*-values between BDNF expression level or ASE ratios and species-biased song phenotypes in F1 hybrids.

**(D)** Correlations between BDNF mRNA expression in HVC and species-biased song features (syllable duration, entropy variance, motif, and repetition) among F1 hybrid individuals.

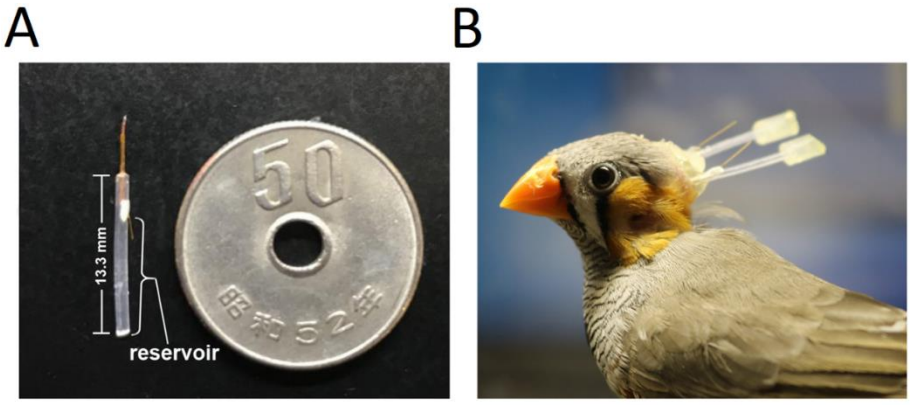



**Figure III-5: Untethered microdialysis probe for pharmacological manipulation of BDNF receptors in RA**

**(A)** Photograph of homemade microdialysis probe.

**(B)** A zebra finch with microdialysis probes bilaterally implanted in RA.

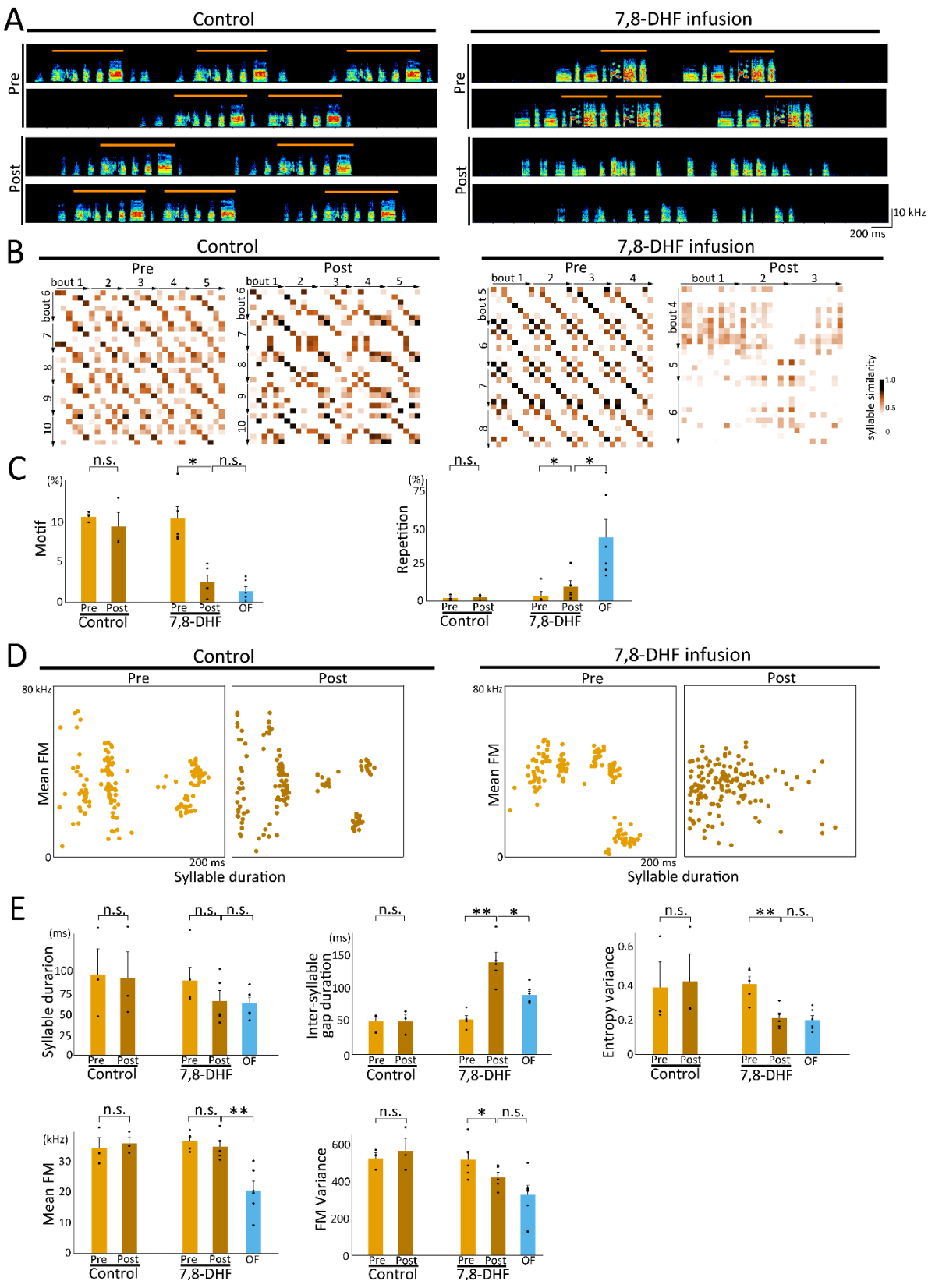

A
Control
7,8-DHF infusion
Pre
Post
10 kHz
200 ms
B
Control
7,8-DHF infusion
Pre
Post
bout 1
syllable similarity
C
Motif
Repetition
(%)
n.s.
Pre
Post
OF
Control
7,8-DHF
D
Control
7,8-DHF infusion
Mean FM
80 kHz
200 ms
Syllable duration
E
Syllable duration
Inter-syllable gap duration
Entropy variance
Mean FM
FM Variance
(ms)
(kHz)

**Figure III-6: Effects of BDNF receptor agonist infusion into RA on species-specific song features.**

**(A)** Songs before and after infusing BDNF receptor agonist, 7,8-DHF. Typical examples of songs from control and 7,8-DHF infused birds. Orange solid lines represent the motif structure of ZF songs.

**(B)** Examples of changes in syllable sequence between pre- and post-infusion. Syllable similarity matrices (SSM) for a pair of songs produced by control and 7,8-DHF infused birds.

**(C)** Changes in the frequency of motif and repetition in songs at pre- and post-infusion stage [control ZF n = 3, ZF with 7,8-DFH infusion (7-10 days) n = 5, normal OF n = 6; paired *t*-test for pre- and post-infusion ZF; Unpaired *t*-test for post-infusion ZF and OF, $p^* < 0.05$]. Each dot corresponds to individual birds.

**(D)** Examples of changes in syllable acoustics at pre- and post-infusion. Scatter plots indicate the distribution of 150 syllables (mean frequency versus syllable duration) for control and 7,8-DHF infused birds.

**(E)** Changes in syllable acoustics (syllable duration, inter-syllable gap duration, entropy variance, mean FM, and FM variance) of songs at pre- and post-infusion stage [control ZF n = 3, ZF with 7,8-DFH infusion (7-10 days) n = 5, normal OF n = 6; paired *t*-test for pre- and post-infusion ZF; Unpaired *t*-test for post-infusion ZF and OF, $p^{**} < 0.01$, $p^* < 0.05$, n.s.: no significant]. Each dot corresponds to individual birds.

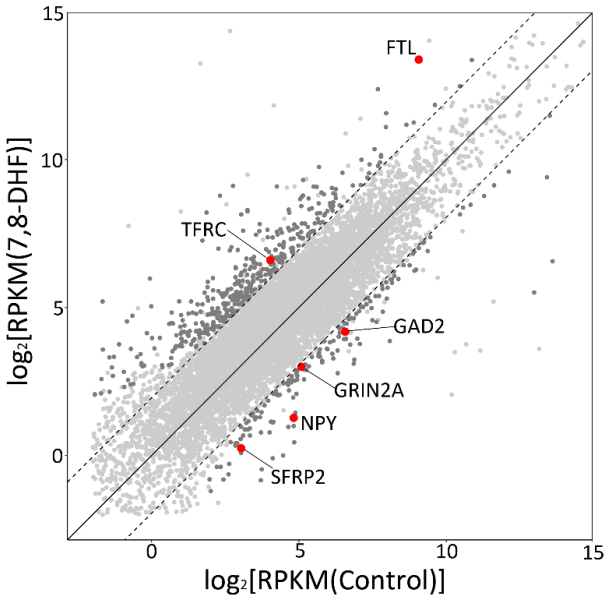



**Figure III-7: Differentially expressed genes in RA after 7,8-DHF infusion.**

Scatter plot indicating gene expression (RPKM value with log2-transformed) in RA of control (X-axis) and 7,8-DHF infused birds (Y-axis). Dashed lines represent the boundary of 4-fold expression difference. Gray dots represent not significantly different expression between control and 7,8-DHF infused birds. Darker gray-colored dots are significantly different expressed genes with more than 4-fold change between control and 7,8-DHF infused birds. Red-colored dots are *trans*-regulated downstream genes of BDNF (**Fig. III-1A**).

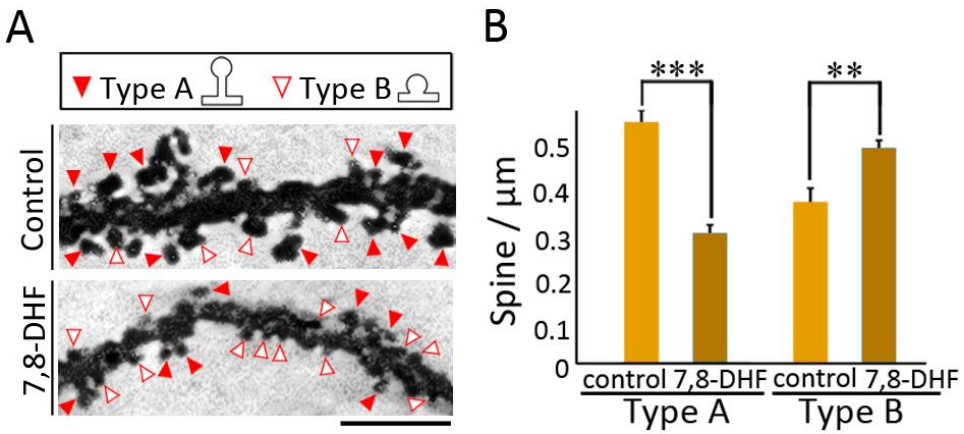



**Figure III-8: Density of dendritic spines in RA neurons of control and 7,8-DHF infused birds.**

**(A)** Photograph of two types of dendrites. Upper panel represent the types of spines: spines with mushroom shape were defined as type A and stubby shape as type B, which used filled arrowhead and empty arrowhead in the down panel; Scale bar = 5 µm

**(B)** Bar graph showing differences in the density of dendritic spines between control and 7,8-DHF infused birds. (Data: Mean ± SEM; student's *t*-test, $p^{**} < 0.01$, $p^{***} < 0.001$].

# General discussion

Like human speech, birdsong is a motor skill that is acquired through vocal learning (Marler & Slabbekoorn, 2004; Brenowitz & Beecher, 2005; Mori & Wada, 2015b). The parallel neural pathways of vocal learning with human, made songbirds good models to explore the mechanisms of vocal learning (Jarvis, 2004). Songbirds produce species-specific songs using a conserved song pathway (Nottebohm *et al.*, 1976; Doupe & Konishi, 1991), suggesting that vocal learning among songbirds species is regulated under genetic constraint. However, it is largely unclear how the difference in gene expression, especially how differences in gene transcriptional regulation induced species-specific songs among songbird species. In my PhD study, I used zebra finch, owl finch, and their reciprocal F1 hybrids to analyze the gene expression and transcriptional regulatory differences in HVC and RA of the motor pathway. In **Chapter I**, I revealed that singing behaviors were significantly different between ZF and OF under conspecific and cross-specific song tutoring. 333 and 374 genes were differentially expressed between ZF and OF in HVC and RA, respectively, of the vocal motor pathway. More than 65% of the species differential expressed (SDE) genes only existed in either HVC or RA, which showed brain region-specific regulation. In **Chapter II**, I revealed the existence of transcriptional regulatory differences in HVC and RA between ZF and OF. Results showed that 2.4% and 3.9% of genes were *cis*-regulatory different, 4.7% and 6.1% of genes were *trans*-regulatory different, 3.1% and 2.9% genes were both *cis*- and *trans*-regulatory different between ZF and OF in HVC and RA, respectively. The *trans*-regulatory genes were more frequent than other forms

of transcriptional regulatory divergences, such as *cis*- and *cis*- and *trans*- regulation, in both HVC and RA. In addition, the results of functional analysis showed that RA *trans*-regulated genes were enriched in neural functions. BDNF, one of candidate upstream mediators of the RA *trans*-regulated genes, had the highest possibility to mediate the *trans*-regulated genes which were responsible for species-specific songs. **In Chapter III**, I verified the function of BDNF as an RA *trans*-mediator for the generation of species-specific songs. First, difference in the expression level of BDNF, but not amino acid substitutions in the protein, was responsible for the *trans*-regulatory differences between ZF and OF. I performed over-activation of BDNF receptors in RA of adult ZF, resulting in an altered species-specific song feature. These results suggested that transcriptional regulatory divergences inducing species differential expression were crucial for the formation of species-specific songs among songbirds.

***Potential possibility of transcriptional regulatory divergences at different development stages of birdsong among species***

Birdsong is acquired through vocal learning with three stages, subsong, plastic song, and crystalized song, during development after hatching (Marler & Slabbekoorn, 2004; Brenowitz & Beecher, 2005). Previous studies have shown that gene expression levels in the song system changed dynamically through song development (Akutagawa & Konishi, 1998; Haesler *et al.*, 2004; Asogwa *et al.*, 2018; Hayase & Wada, 2018; Hayase *et al.*, 2018). For example, the previous study found muscarinic acetylcholine receptor 4 (CHRM4) gene show different expression levels in HVC during song

development (Asogwa *et al.*, 2018). In addition, expression levels of some clusters of genes were affected by age and singing experience in both HVC and RA (Hayase *et al.*, 2018). These researches suggested that the importance of the potential transcriptional regulatory difference uniquely existing at subsong and plastic song stages which contribute to the generation of the species-specific songs should not be ignored. However, in this study, I analyzed the transcriptional regulatory divergence in the motor pathway only at the adult song stage after song crystallization between adult ZF and OF. It is still worthwhile to identify transcriptional regulatory differences between ZF and OF at subsong and plastic song stages since these developmental stages were crucial for the generation of species-specific songs during song development. In future research, it is essential to identify new *cis*- and/or *trans*-regulatory elements by performing transcriptional regulatory divergence analysis in the subsong and plastic song stages.

***Transcriptional regulatory divergence in the anterior forebrain pathway (AFP)***

The anterior forebrain pathway (AFP) is essential for song learning at the juvenile stage and generation of vocal fluctuation at the adult stage, but not for the production of acquired song (Bottjer *et al.*, 1984; Scharff & Nottebohm, 1991; Doupe *et al.*, 2005). There is no direct evidence showing that the AFP plays a role in the acquisition of species-specific songs of songbirds. However, the potential contribution of the AFP to species-specific songs should be considered. This is because the lesion of Area X, a basal ganglia nucleus in the AFP, resulted in the development of unstable species-

specific motif structure in the ZFs (Scharff & Nottebohm, 1991). Previous research found that different expression levels of orthologous genes in certain brain region induced species-specific behaviors (Insel & Shapiro, 1992; Young *et al.*, 1997). In songbirds, some genes show different expression levels in the AFP, such as FoxP2 gene in Area X of Bengalese finch strains (Haesler *et al.*, 2004; Wada *et al.*, 2013b) and 5HT1B in LMAN among songbird species (Wada unpublished data). Transcriptional regulatory differences in the AFP which induce the different expression of orthologous genes among songbird species should play an important role in the generation of species-specific songs. In future research, exploring the transcriptional regulatory divergence in AFP among songbird species is essential since the AFP is important for song learning and plasticity.

***Transcriptional regulatory divergence among multiple songbird species***

I analyzed transcriptional regulatory divergence between two closely related songbird species, zebra finch and owl finch. These two species diverged about 6.5 million years ago (Forshaw & Shephard, 2012; Hooper & Price, 2015) and can get reciprocal F1 hybrids making them be suitable models to analyze the transcriptional regulatory divergence that may underlie species differences in learned behavior. There are more than 4,000 songbird species in the world (Marler & Slabbekoorn, 2004; Brenowitz & Beecher, 2005), and it is possible to get F1 hybrid offspring between many songbird species (Immelmann & Cayley, 1982). Therefore, it is possible to analyze conserved or unique transcriptional regulatory changes among other songbird species

by utilizing multiple F1 hybrids by adaptation of the same strategy described in this thesis (**Fig. II-1**). The sequencing and annotation of the genome of a crowd of avian species including songbird species have been conducted or recently ongoing (Zhang *et al.*, 2014; Koepfli *et al.*, 2015). Additionally, the genome information of two songbird species, zebra finch and Bengalese finch, are already published and can be used for the genomic analysis (Warren *et al.*, 2010; Colquitt *et al.*, 2018). Thus, the F1 hybrids between zebra finch, Bengalese finch, and their closely related species can be used to explore the transcriptional regulatory divergences among songbird species (**Fig. IV-1A**). By comparing the *cis*- and/or *trans*-regulated genes among multiple songbird species, it is possible to elucidate sets of genes with consistent transcriptional regulatory differences (**Fig. IV-1B**). The *cis*- and/or *trans*-regulatory element of these genes may play crucial roles in the evolution of species-specific songs. It remains unclear how species-specific songs structure evolved among songbirds species during evolution (Jarvis, 2004; Marler & Slabbekoorn, 2004). Unraveling the crucial mutations in *cis*- and/or *trans*- element by transcriptional regulatory divergence analyses among multiple songbird species may help to understand how genomic mutation contributed to the evolution of species-specific songs.

# Figure of general discussion

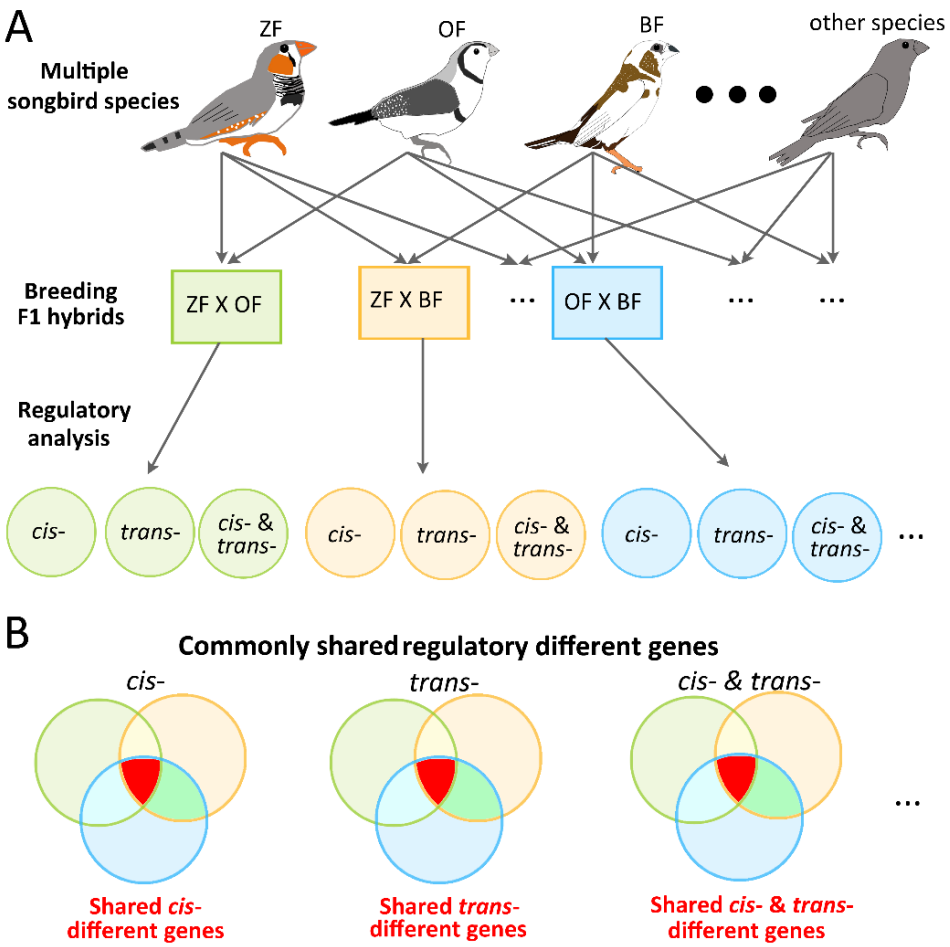



**Figure IV-1: Research plan of transcriptional regulatory divergence among multiple songbird species.**

**(A)** F1 hybrid can be bred between multiple songbird species under laboratory condition. *Cis*- and/or *trans*-regulatory difference can be identified between multiple songbird species.

**(B)** Genes with consistent transcriptional regulatory differences can be identified by comparing results of *cis*- and/or *trans*-regulatory difference analysis between multiple songbird species.

# Acknowledgements

The time I passed in the Graduate School of Life Science of Hokkaido University has taught me so much about science. Throughout the doctoral course and writing this thesis I have received a great deal of support and assistance.

Firstly, I would like to express my gratitude to my supervisor Dr. Kazuhiro Wada for mentoring me. Wada Sensei taught me not only in knowledge, concept, and philosophy but also showed me the attitude to science. I learned so much from Wada Sensei about how to develop scientific questions, solve problem and polish up scientific concepts during research. Such knowledge and spirits will accompany me in my whole science career.

I thank the members of my thesis committee, Dr. Makoto Mizunami and Dr. Asato Kuroiwa for their comments and suggestion. Thanks for their patience as my thesis committee members.

I appreciate members of laboratories of Animal Behavior and Intelligence, Drs. Toshiya Matsushima, Hiroto Ogawa, Masayo Soma, Nobuaki Tanaka, and Nina Patzke for stimulating discussions and sympathetic support for helping me to pursue my thesis research.

I thank Dr. David Wheatcroft at the Department of Ecology and Genetics of Uppsala University. You really took the time and effort to give us suggestions and English checking during the paper publish which related to this thesis.

I thank Dr. Jun Sese at the National Institute of Advanced Industrial Science and Technology, Tokyo, Japan who help to analyze RNA-seq data.

Many thanks for the current and previous members of Wada lab. Studying and working with you was really nice experience. Thanks to Dr. Chinweike Norman

Asogwa, you helped me checking the English of my paper and thesis. In addition, the friendship between us gave me much encouragement during my doctoral course. Thanks for Azusa Sawai, Dr. Noriyuki Toji, Dr. Shin Hayase for helping experiment and data analysis. I thank Miguel Sanchez for checking English of my paper. I thank Yu Ji, Rintarou Sugioka, Yukino Shibata and Yuika Suzuki for helping me microdialysis and sectioning brain samples. I thank hardly work of Keiko Sumida breeding birds which were used as research animals in this thesis. I will never forget Ippei Kojima, Jun Ishigoka, Xia Wu, Yuto Nakagawa, Yumeno Suzuki, Nasiba Afrin and Chengru Shao, I passed a nice time with you in Wada lab during my doctoral course.

Thanks for the financial support given me by the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT). I cannot finish the doctoral course without your support. Thanks to all Japanese people who carry open mind for international students, this made me love Japan and Japanese culture much more than before. In the future, I wish to do some positive efforts for promoting communication of science and culture between Japan and China.

Thanks for the financial support from the Chinese Scholarship Council (CSC). Under your support, I got the chance to study aboard in Japan.

I thank Mrs. Mari Igarashi and other members of OIAS, Faculty of Science. You helped me so much containing many aspects of my life in Japan.

Thanks to my parents, my father Mr. Zongsheng Wang and my mother Mrs. Hongying Qin. Your encouragement, whenever and wherever, is the most important factor for me to do my best.

No words can describe my gratitude for my wife, Mrs. Chengze Li. You gave the best care and support to me during these several years. Thank you for your love.

# References


Ahmadiyeh, N., Churchill, G.A., Solberg, L.C., Baum, A.E., Shimonura, K., Takahashi, J.S. & Redei, E.E. (2005) Lineage is an epigenetic modifier of QTL influencing behavioral coping with stress. *Behavior genetics*, **35**, 189-198.

Akutagawa, E. & Konishi, M. (1998) Transient expression and transport of brain-derived neurotrophic factor in the male zebra finch's song system during vocal development. *Proceedings of the National Academy of Sciences*, **95**, 11429-11434.

Altar, C.A., Cai, N., Bliven, T., Juhasz, M., Conner, J.M., Acheson, A.L., Lindsay, R.M. & Wiegand, S.J. (1997) Anterograde transport of brain-derived neurotrophic factor and its role in the brain. *Nature*, **389**, 856-860.

Andalman, A.S. & Fee, M.S. (2009) A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences*, **106**, 12518-12523.

Asogwa, N.C., Mori, C., Sánchez-Valpuesta, M., Hayase, S. & Wada, K. (2018) Inter-and intra-specific differences in muscarinic acetylcholine receptor expression in the neural pathways for vocal learning in songbirds. *Journal of Comparative Neurology*, **526**, 2856-2869.

Barde, Y.-A., Edgar, D. & Thoenen, H. (1982) Purification of a new neurotrophic factor from mammalian brain. *The EMBO journal*, **1**, 549-553.

Bendesky, A., Kwon, Y.-M., Lassance, J.-M., Lewarch, C.L., Yao, S., Peterson, B.K., He, M.X., Dulac, C. & Hoekstra, H.E. (2017) The genetic basis of parental care evolution in monogamous mice. *Nature*, **544**, 434-439.

Bengston, S.E., Dahan, R.A., Donaldson, Z., Phelps, S.M., van Oers, K., Sih, A. & Bell, A.M. (2018) Genomic tools for behavioural ecologists to understand repeatable individual differences in behaviour. *Nature Ecology & Evolution*, **2**, 944-955.

Blugeot, A., Rivat, C., Bouvier, E., Molet, J., Mouchard, A., Zeau, B., Bernard, C., Benoliel, J.-J. & Becker, C. (2011) Vulnerability to depression: from brain neuroplasticity to identification of biomarkers. *Journal of Neuroscience*, **31**, 12889-12899.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.

Bottjer, S.W., Miesner, E.A. & Arnold, A.P. (1984) Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science*, **224**, 901-903.

Bourne, J. & Harris, K.M. (2007) Do thin spines learn to be mushroom spines that remember?

*Current opinion in neurobiology*, **17**, 381-386.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A. & Kircher, M. (2011) The evolution of gene expression levels in mammalian organs. *Nature*, **478**, 343.

Brenowitz, E.A. & Beecher, M.D. (2005) Song learning in birds: diversity and plasticity, opportunities and challenges. *Trends Neurosci*, **28**, 127-132.

Bryois, J., Buil, A., Evans, D.M., Kemp, J.P., Montgomery, S.B., Conrad, D.F., Ho, K.M., Ring, S., Hurles, M., Deloukas, P., Davey Smith, G. & Dermitzakis, E.T. (2014) Cis and trans effects of human genomic variants on gene expression. *PLoS Genet*, **10**, e1004461.

Carroll, S.B. (2005) Evolution at two levels: on genes and form. *PLoS biology*, **3**, e245.

Carroll, S.B., Grenier, J.K. & Weatherbee, S.D. (2013) *From DNA to diversity: molecular genetics and the evolution of animal design*. John Wiley & Sons.

Carvajal-Rodriguez, A., de Una-Alvarez, J. & Rolan-Alvarez, E. (2009) A new multitest correction (SGoF) that increases its statistical power when increasing the number of tests. *BMC bioinformatics*, **10**, 209.

Chandra, V., Fetter-Pruneda, I., Oxley, P.R., Ritger, A.L., McKenzie, S.K., Libbrecht, R. & Kronauer, D.J. (2018) Social regulation of insulin signaling and the evolution of eusociality in ants. *Science*, **361**, 398-402.

Chen, Z.-Y., Patel, P.D., Sant, G., Meng, C.-X., Teng, K.K., Hempstead, B.L. & Lee, F.S. (2004) Variant brain-derived neurotrophic factor (BDNF)(Met66) alters the intracellular trafficking and activity-dependent secretion of wild-type BDNF in neurosecretory cells and cortical neurons. *Journal of Neuroscience*, **24**, 4401-4411.

Clayton, N.S. (1989) The effects of cross-fostering on selective song learning in estrildid finches. *Behaviour*, **109**, 163-174.

Colquitt, B.M., Mets, D.G. & Brainard, M.S. (2018) Draft genome assembly of the Bengalese finch, Lonchura striata domestica, a model for motor skill variability and learning. *GigaScience*, **7**.

Crews, D. (1997) Species diversity and the evolution of behavioral controlling mechanisms. *Annals of the New York Academy of Sciences*, **807**, 1-21.

Davidson, J.H. & Balakrishnan, C.N. (2016) Gene regulatory evolution during speciation in a songbird. *G3: Genes, Genomes, Genetics*, **6**, 1357-1364.

Delcher, A.L., Salzberg, S.L. & Phillippy, A.M. (2003) Using MUMmer to identify similar regions in large sequence sets. *Current Protocols in Bioinformatics*, 10.13. 11-10.13. 18.

Devi, L. & Ohno, M. (2012) 7, 8-dihydroxyflavone, a small-molecule TrkB agonist, reverses memory deficits and BACE1 elevation in a mouse model of Alzheimer's disease. *Neuropsychopharmacology*, **37**, 434.

Ding, Y., Berrocal, A., Morita, T., Longden, K.D. & Stern, D.L. (2016) Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature*, **536**, 329-332.

Ding, Y., Lillvis, J.L., Cande, J., Berman, G.J., Arthur, B.J., Long, X., Xu, M., Dickson, B.J. & Stern, D.L. (2019) Neural Evolution of Context-Dependent Fly Song. *Current Biology*.

Doupe, A.J. & Konishi, M. (1991) Song-selective auditory circuits in the vocal control system of the zebra finch. *Proceedings of the National Academy of Sciences*, **88**, 11339-11343.

Doupe, A.J., Perkel, D.J., Reiner, A. & Stern, E.A. (2005) Birdbrains could teach basal ganglia research a new song. *Trends in Neurosciences*, **28**, 353-363.

Eales, L.A. (1985) Song learning in zebra finches: some effects of song model availability on what is learnt and when. *Animal Behaviour*, **33**, 1293-1300.

Eales, L.A. (1987) Do zebra finch males that have been raised by another species still tend to select a conspecific song tutor? *Animal Behaviour*, **35**, 1347-1355.

Eckersley-Maslin, Mélanie A., Thybert, D., Bergmann, Jan H., Marioni, John C., Flicek, P. & Spector, David L. (2014) Random Monoallelic Gene Expression Increases upon Embryonic Stem Cell Differentiation. *Developmental Cell*, **28**, 351-365.

Egan, M.F., Kojima, M., Callicott, J.H., Goldberg, T.E., Kolachana, B.S., Bertolino, A., Zaitsev, E., Gold, B., Goldman, D. & Dean, M. (2003) The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*, **112**, 257-269.

Eisenberg, D., Marcotte, E.M., Xenarios, I. & Yeates, T.O. (2000) Protein function in the post-genomic era. *Nature*, **405**, 823.

Fehér, O., Wang, H., Saar, S., Mitra, P.P. & Tchernichovski, O. (2009) De novo establishment of wild-type song culture in the zebra finch. *Nature*, **459**, 564.

Forshaw, J.M. & Shephard, M. (2012) *Grassfinches in Australia*. CSIRO PUBLISHING.

Fraser, H.B. (2019) Improving Estimates of Compensatory cis–trans Regulatory Divergence. *Trends in Genetics*, **35**, 3-5.

Fresard, L., Leroux, S., Servin, B., Gourichon, D., Dehais, P., Cristobal, M.S., Marsaud, N., Vignoles, F., Bed'Hom, B. & Coville, J.-L. (2014) Transcriptome-wide investigation of genomic imprinting in chicken. *Nucleic acids research*, **42**, 3768-3782.

Fujita, Y., Mutoh, A. & Kato, S. (Year) Birdsong acquisition model by sexual selection focused on habitat density. Australasian Joint Conference on Artificial Intelligence. Springer, City. p. 687-696.

Gallup, G.G. & Capper, S.A. (1970) Preference for mirror-image stimulation in finches (Passer domesticus domesticus) and parakeets (Melopsittacus undulatus). *Animal Behaviour*, **18**, 621-624.

Gipson, C.D. & Olive, M.F. (2017) Structural and functional plasticity of dendritic spines–root or result of behavior? *Genes, Brain and Behavior*, **16**, 101-117.

Giza, J.I., Kim, J., Meyer, H.C., Anastasia, A., Dincheva, I., Zheng, C.I., Lopez, K., Bains, H., Yang, J. & Bracken, C. (2018) The BDNF Val66Met prodomain disassembles dendritic spines altering fear extinction circuitry and behavior. *Neuron*, **99**, 163-178. e166.

Gomes, S. & Civetta, A. (2015) Hybrid male sterility and genome-wide misexpression of male reproductive proteases. *Sci Rep*, **5**, 11976.

Goncalves, A., Leigh-Brown, S., Thybert, D., Stefflova, K., Turro, E., Flicek, P., Brazma, A., Odom, D.T. & Marioni, J.C. (2012) Extensive compensatory cis-trans regulation in the evolution of mouse gene expression. *Genome research*, **22**, 2376-2384.

Haesler, S., Wada, K., Nshdejan, A., Morrisey, E.E., Lints, T., Jarvis, E.D. & Scharff, C. (2004) FoxP2 expression in avian vocal learners and non-learners. *Journal of Neuroscience*, **24**, 3164-3175.

Hahnloser, R.H., Kozhevnikov, A.A. & Fee, M.S. (2002) An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature*, **419**, 65.

Hammock, E.A. & Young, L.J. (2005) Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science*, **308**, 1630-1634.

Harris, K.M. & Spacek, J. (2016) *Dendrite structure*. Oxford University Press, Oxford.

Hayase, S. & Wada, K. (2018) Singing activity-driven Arc expression associated with vocal acoustic plasticity in juvenile songbird. *European journal of neuroscience*, **48**, 1728-1742.

Hayase, S., Wang, H., Ohgushi, E., Kobayashi, M., Mori, C., Horita, H., Mineta, K., Liu, W.-c. & Wada, K. (2018) Vocal practice regulates singing activity–dependent genes underlying age-

independent vocal learning in songbirds. *PLoS biology*, **16**, e2006537.

Hooper, D.M. & Price, T.D. (2015) Rates of karyotypic evolution in Estrildid finches differ between island and continental clades. *Evolution; international journal of organic evolution*, **69**, 890-903.

Huang, D.W., Sherman, B.T. & Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, **4**, 44.

Husi, H. & Grant, S.G.N. (2001) Proteomics of the nervous system. *Trends in Neurosciences*, **24**, 259-266.

Imai, R., Sawai, A., Hayase, S., Furukawa, H., Asogwa, C.N., Sanchez, M., Wang, H., Mori, C. & Wada, K. (2016) A quantitative method for analyzing species-specific vocal sequence pattern and its developmental dynamics. *J Neurosci Methods*, **271**, 25-33.

Immelmann, K. & Cayley, N.W. (1982) *Australian finches in bush and aviary*. Angus & Robertson.

Insel, T.R. & Shapiro, L.E. (1992) Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proceedings of the National Academy of Sciences*, **89**, 5981-5985.

Insel, T.R., Wang, Z.-X. & Ferris, C.F. (1994) Patterns of brain vasopressin receptor distribution associated with social organization in microtine rodents. *Journal of Neuroscience*, **14**, 5381-5392.

Itoh, Y., Kampf, K., Balakrishnan, C.N. & Arnold, A.P. (2011) Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma*, **120**, 255-264.

Jarvis, E.D. (2004) Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, **1016**, 749-777.

Jarvis, E.D., Scharff, C., Grossman, M.R., Ramos, J.A. & Nottebohm, F. (1998) For whom the bird sings: context-dependent gene expression. *Neuron*, **21**, 775-788.

Katz, P.S. & Harris-Warrick, R.M. (1999) The evolution of neuronal circuits underlying species-specific behavior. *Current opinion in neurobiology*, **9**, 628-633.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. & Salzberg, S. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*, **14**.

Kim, S. & Kaang, B.-K. (2017) Epigenetic regulation and chromatin remodeling in learning and memory. *Experimental & molecular medicine*, **49**, e281.

Kittelberger, J.M. & Mooney, R. (2005) Acute injections of brain-derived neurotrophic factor in a vocal premotor nucleus reversibly disrupt adult birdsong stability and trigger syllable deletion. *Journal of neurobiology*, **62**, 406-424.

Koepfli, K.-P., Paten, B., Scientists, G.K.C.o. & O'Brien, S.J. (2015) The Genome 10K Project: a way forward. *Annu. Rev. Anim. Biosci.*, **3**, 57-111.

Korzus, E., Rosenfeld, M.G. & Mayford, M. (2004) CBP Histone Acetyltransferase Activity Is a Critical Component of Memory Consolidation. *Neuron*, **42**, 961-972.

Kowiański, P., Lietzau, G., Czuba, E., Waśkow, M., Steliga, A. & Moryś, J. (2018) BDNF: A Key Factor with Multipotent Impact on Brain Signaling and Synaptic Plasticity. *Cellular and Molecular Neurobiology*, **38**, 579-593.

Kozhevnikov, A.A. & Fee, M.S. (2007) Singing-related activity of identified HVC neurons in the zebra finch. *Journal of neurophysiology*, **97**, 4271-4283.

Krämer, A., Green, J., Pollard Jr, J. & Tugendreich, S. (2013) Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*, **30**, 523-530.

Krebs, J., Ashcroft, R. & Webber, M. (1978) Song repertoires and territory defence in the great tit. *Nature*, **271**, 539.

Krueger, F. & Andrews, S.R. (2016) SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Research*, **5**.

Leibrock, J., Lottspeich, F., Hohn, A., Hofer, M., Hengerer, B., Masiakowski, P., Thoenen, H. & Barde, Y.-A. (1989) Molecular cloning and expression of brain-derived neurotrophic factor. *Nature*, **341**, 149.

Leonardo, A. & Fee, M.S. (2005) Ensemble coding of vocal control in birdsong. *Journal of Neuroscience*, **25**, 652-661.

Levenson, J.M. & Sweatt, J.D. (2005) Epigenetic mechanisms in memory formation. *Nature Reviews Neuroscience*, **6**, 108.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. (2009) 692 (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2693.

Li, X.-C., Jarvis, E.D., Alvarez-Borda, B., Lim, D.A. & Nottebohm, F. (2000) A relationship between behavior, neurotrophin expression, and new neuron survival. *Proceedings of the National Academy of Sciences*, **97**, 8584-8589.

Lim, M.M., Wang, Z., Olazábal, D.E., Ren, X., Terwilliger, E.F. & Young, L.J. (2004) Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature*, **429**, 754.

Lingle, S. (1992) Escape gaits of white-tailed deer, mule deer and their hybrids: gaits observed and patterns of limb coordination. *Behaviour*, **122**, 153-181.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, **15**, 550.

Lu, B. (2003) BDNF and activity-dependent synaptic modulation. *Learning & memory*, **10**, 86-98.

Mack, K.L. & Nachman, M.W. (2017) Gene regulation and speciation. *Trends in Genetics*, **33**, 68-80.

Marler, P. & Peters, S. (1977) Selective vocal learning in a sparrow. *Science*, **198**, 519-521.

Marler, P.R. & Slabbekoorn, H. (2004) *Nature's music: the science of birdsong*. Elsevier.

Martin, K.C. & Sun, Y.E. (2004) To learn better, keep the HAT on. *Neuron*, **42**, 879-881.

Mori, C. & Wada, K. (2015a) Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **35**, 878-889.

Mori, C. & Wada, K. (2015b) Songbird: a unique animal model for studying the molecular basis of disorders of vocal development and communication. *Experimental animals*, **64**, 221-230.

Nelson, D.A. (2000) A preference for own-subspecies' song guides vocal learning in a song bird. *Proceedings of the National Academy of Sciences*, **97**, 13348-13353.

Nimchinsky, E.A., Sabatini, B.L. & Svoboda, K. (2002) Structure and function of dendritic spines. *Annual review of physiology*, **64**, 313-353.

Nottebohm, F., Paton, J.A. & Kelley, D.B. (1982) Connections of vocal control nuclei in the canary telencephalon. *Journal of Comparative Neurology*, **207**, 344-357.

Nottebohm, F., Stokes, T.M. & Leonard, C.M. (1976) Central control of song in the canary, Serinus canarius. *Journal of Comparative Neurology*, **165**, 457-486.

Ölveczky, B.P., Otchy, T.M., Goldberg, J.H., Aronov, D. & Fee, M.S. (2011) Changes in the neural control of a complex motor sequence during learning. *Journal of neurophysiology*, **106**, 386-397.

Pandey, A. & Mann, M. (2000) Proteomics to study genes and genomes. *Nature*, **405**, 837.

Peixoto, L. & Abel, T. (2013) The role of histone acetylation in memory formation and cognitive impairments. *Neuropsychopharmacology*, **38**, 62.

Pfenning, A.R., Hara, E., Whitney, O., Rivas, M.V., Wang, R., Roulhac, P.L., Howard, J.T., Wirthlin, M., Lovell, P.V., Ganapathy, G., Mountcastle, J., Moseley, M.A., Thompson, J.W., Soderblom, E.J., Iriki, A., Kato, M., Gilbert, M.T.P., Zhang, G., Bakken, T., Bongaarts, A., Bernard, A., Lein, E., Mello, C.V., Hartemink, A.J. & Jarvis, E.D. (2014) Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science*, **346**, 1256846.

Raff, M.C., Fields, K.L., Hakomori, S.-I., Mirsky, R., Pruss, R.M. & Winter, J. (1979) Cell-type-specific markers for distinguishing and studying neurons and the major classes of glial cells in culture. *Brain research*, **174**, 283-308.

Reaume, C.J. & Sokolowski, M.B. (2011) Conservation of gene function in behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2100-2110.

Reuveni, E., Getselter, D., Oron, O. & Elliott, E. (2018) Differential contribution of cis and trans gene transcription regulatory mechanisms in amygdala and prefrontal cortex and modulation by social stress. *Scientific reports*, **8**, 6339.

Rifkin, S.A., Kim, J. & White, K.P. (2003) Evolution of gene expression in the Drosophila melanogaster subgroup. *Nature genetics*, **33**, 138.

Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L.T., Peeler, D.J., Mukherjee, S. & Chen, W. (2018) Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*, **360**, 176-182.

Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K.D., Imai, T. & Ueda, H.R. (2013) Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome biology*, **14**, 3097.

Scharff, C. & Nottebohm, F. (1991) A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *Journal of Neuroscience*, **11**, 2896-2913.

Searcy, W.A. & Andersson, M. (1986) Sexual selection and the evolution of song. *Annual Review of Ecology and Systematics*, **17**, 507-533.

Seeholzer, L.F., Seppo, M., Stern, D.L. & Ruta, V. (2018) Evolution of a central neural circuit underlies Drosophila mate preferences. *Nature*, **559**, 564-569.

Segal, M., Korkotian, E. & Murphy, D.D. (2000) Dendritic spine formation and pruning: common cellular mechanisms? *Trends in neurosciences*, **23**, 53-57.

Signor, S.A. & Nuzhdin, S.V. (2018) The Evolution of Gene Expression in cis and trans. *Trends in genetics : TIG*.

Simon, A. (2010) FastQC: a quality control tool for high throughput sequence data. *Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc*.

Sober, S.J., Wohlgemuth, M.J. & Brainard, M.S. (2008) Central contributions to acoustic variation in birdsong. *Journal of Neuroscience*, **28**, 10370-10379.

Soderstrom, K., Qin, W. & Leggett, M.H. (2007) A minimally invasive procedure for sexing young zebra finches. *Journal of Neuroscience Methods*, **164**, 116-119.

Sossinka, R. & Böhner, J. (1980) Song types in the zebra finch Poephila guttata castanotis 1. *Zeitschrift für Tierpsychologie*, **53**, 123-132.

Spacek, J. & Harris, K.M. (1997) Three-dimensional organization of smooth endoplasmic reticulum in hippocampal CA1 dendrites and dendritic spines of the immature and mature rat. *Journal of Neuroscience*, **17**, 190-203.

Spiro, J.E., Dalva, M.B. & Mooney, R. (1999) Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *Journal of neurophysiology*, **81**, 3007-3020.

Tamvacakis, A., Senatore, A. & Katz, P. (2018) Single neuron serotonin receptor subtype gene expression correlates with behaviour within and across three molluscan species. *Proceedings of the Royal Society B: Biological Sciences*, **285**, 20180791.

Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B. & Mitra, P.P. (2000) A procedure for an automated measurement of song similarity. *Animal Behaviour*, **59**, 1167-1176.

Teranishi, M., Shimada, Y., Hori, T., Nakabayashi, O., Kikuchi, T., Macleod, T., Pym, R., Sheldon, B., Solovei, I. & Macgregor, H. (2001) Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Research*, **9**, 147-165.

Tierney, A. (1995) Evolutionary implications of neural circuit structure and function. *Behavioural processes*, **35**, 173-182.

Tinbergen, N. (1963) On aims and methods of ethology. *Zeitschrift für tierpsychologie*, **20**, 410-433.

Tomaszycki, M.L., Peabody, C., Replogle, K., Clayton, D.F., Tempelman, R.J. & Wade, J. (2009) Sexual differentiation of the zebra finch song system: potential roles for sex chromosome genes. *BMC Neuroscience*, **10**, 24.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. & Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology*, **28**, 511-515.

Tyler, W.J. & Pozzo-Miller, L. (2003) Miniature synaptic transmission and BDNF modulate dendritic spine growth and form in rat CA1 neurones. *The Journal of physiology*, **553**, 497-509.

Wada, K., Chen, C.-C. & Jarvis, E.D. (2013a) Detecting Neural Activity-Dependent Immediate Early Gene Expression in the Brain. In Ogawa, H., Oka, K. (eds) *Methods in Neuroethological Research*. Springer Japan, Tokyo, pp. 133-149.

Wada, K., Hayase, S., Imai, R., Mori, C., Kobayashi, M., Liu, W.c., Takahasi, M. & Okanoya, K. (2013b) Differential androgen receptor expression and DNA methylation state in striatum song nucleus Area X between wild and domesticated songbird strains. *European Journal of Neuroscience*, **38**, 2600-2610.

Wada, K., Howard, J.T., McConnell, P., Whitney, O., Lints, T., Rivas, M.V., Horita, H., Patterson, M.A., White, S.A. & Scharff, C. (2006) A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes. *Proceedings of the National Academy of Sciences*, **103**, 15212-15217.

Wang, X., Werren, J.H. & Clark, A.G. (2016) Allele-specific transcriptome and methylome analysis reveals stable inheritance and cis-regulation of DNA methylation in Nasonia. *PLoS biology*, **14**, e1002500.

Warren, W.C., Clayton, D.F., Ellegren, H., Arnold, A.P., Hillier, L.W., Künstner, A., Searle, S., White, S., Vilella, A.J. & Fairley, S. (2010) The genome of a songbird. *Nature*, **464**, 757.

Weber, J.N., Peterson, B.K. & Hoekstra, H.E. (2013) Discrete genetic modules are responsible for complex burrow evolution in Peromyscus mice. *Nature*, **493**, 402.

Wheeler, D.A., Kyriacou, C.P., Greenacre, M.L., Yu, Q., Rutila, J.E., Rosbash, M. & Hall, J.C. (1991) Molecular transfer of a species-specific behavior from Drosophila simulans to Drosophila melanogaster. *Science*, **251**, 1082-1085.

Williams, H. (2004) Birdsong and singing behavior. *ANNALS-NEW YORK ACADEMY OF SCIENCES*, 1-30.

Winslow, J.T., Hastings, N., Carter, C.S., Harbaugh, C.R. & Insel, T.R. (1993) A role for central

vasopressin in pair bonding in monogamous prairie voles. *Nature*, **365**, 545.

Wittkopp, P.J. (2005) Genomic sources of regulatory variation in cis and in trans. *Cellular and molecular life sciences : CMLS*, **62**, 1779-1783.

Wittkopp, P.J., Haerum, B.K. & Clark, A.G. (2004) Evolutionary changes in cis and trans gene regulation. *Nature*, **430**, 85.

Wittkopp, P.J., Haerum, B.K. & Clark, A.G. (2008) Regulatory changes underlying expression differences within and between Drosophila species. *Nature Genetics*, **40**, 346-350.

Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V. & Romano, L.A. (2003) The evolution of transcriptional regulation in eukaryotes. *Molecular biology and evolution*, **20**, 1377-1419.

Wright, A.E., Zimmer, F., Harrison, P.W. & Mank, J.E. (2015) Conservation of regional variation in sex-specific sex chromosome regulation. *Genetics*, **201**, 587-598.

Young, L.J. & Wang, Z. (2004) The neurobiology of pair bonding. *Nature neuroscience*, **7**, 1048.

Young, L.J., Winslow, J.T., Nilsen, R. & Insel, T.R. (1997) Species differences in V1a receptor gene expression in monogamous and nonmonogamous voles: Behavioral consequences. *Behavioral neuroscience*, **111**, 599.

Zeigler, H.P. & Marler, P. (2008) *Neuroscience of birdsong*. Cambridge University Press, Cambridge.

Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L. & Betsholtz, C. (2015) Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, **347**, 1138-1142.

Zhang, G. & Li, C. & Li, Q. & Li, B. & Larkin, D.M. & Lee, C. & Storz, J.F. & Antunes, A. & Greenwold, M.J. & Meredith, R.W. & Ödeen, A. & Cui, J. & Zhou, Q. & Xu, L. & Pan, H. & Wang, Z. & Jin, L. & Zhang, P. & Hu, H. & Yang, W. & Hu, J. & Xiao, J. & Yang, Z. & Liu, Y. & Xie, Q. & Yu, H. & Lian, J. & Wen, P. & Zhang, F. & Li, H. & Zeng, Y. & Xiong, Z. & Liu, S. & Zhou, L. & Huang, Z. & An, N. & Wang, J. & Zheng, Q. & Xiong, Y. & Wang, G. & Wang, B. & Wang, J. & Fan, Y. & da Fonseca, R.R. & Alfaro-Núñez, A. & Schubert, M. & Orlando, L. & Mourier, T. & Howard, J.T. & Ganapathy, G. & Pfenning, A. & Whitney, O. & Rivas, M.V. & Hara, E. & Smith, J. & Farré, M. & Narayan, J. & Slavov, G. & Romanov, M.N. & Borges, R. & Machado, J.P. & Khan, I. & Springer, M.S. & Gatesy, J. & Hoffmann, F.G. & Opazo, J.C. & Håstad, O. & Sawyer, R.H. & Kim, H. & Kim, K.-W. & Kim, H.J. & Cho, S. & Li, N. & Huang, Y. & Bruford, M.W. & Zhan, X. & Dixon, A. & Bertelsen, M.F. & Derryberry, E. & Warren, W. & Wilson, R.K. & Li, S. & Ray, D.A. & Green, R.E. & O'Brien, S.J. & Griffin, D. & Johnson, W.E. & Haussler, D. & Ryder, O.A. & Willerslev, E. & Graves, G.R. & Alström, P. & Fjeldså, J. & Mindell, D.P. & Edwards, S.V. & Braun, E.L. & Rahbek, C. & Burt, D.W. & Houde, P. &

Zhang, Y. & Yang, H. & Wang, J. & Jarvis, E.D. & Gilbert, M.T.P. & Wang, J. (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, **346**, 1311-1320.

Zhang, X. & Emerson, J. (2019) Inferring Compensatory Evolution of cis-and trans-Regulatory Variation. *Trends in Genetics*, **35**, 1-3.

# Main Publication

**Wang H.**, Sawai A., Toji N., Sugioka R., Shibata Y., Suzuki Y., Ji Y., Hayase, S., Akama S., Sese J., Wada K. Transcriptional regulatory divergence underpinning species-specific learned vocalization in songbirds. *PLOS Biology*, 2019 Nov13; 17(11): e3000476

# List of other publications

Hayase S., **Wang H.**, Ohgushi E., Kobayashi M., Mori C., Horita H., Mineta K., Liu WC, Wada K. Vocal practice regulates singing activity-dependent genes underlying age-independent vocal learning in songbirds. *PLoS Biol*. 2018 Sep12;16(9):e2006537

Imai R., Sawai A., Hayase S., Furukawa H., Asogwa CN, Sanchez M., **Wang H.**, Mori C., Wada K. A quantitative method for analyzing species-specific vocal sequence pattern and its developmental dynamics. J Neurosci Methods. 2016 Sep 15; 271:25-33.

## Oral presentation

**Wang H.** (2018). Regulatory divergence of gene transcription for species-specific song production in songbirds. Japan Songbird Workshop at the University of Tokyo, Japan. July 30, 2018.

## Poster presentations

**Wang H**, Sawai A, Hayase S. and Wada K. (2017). Interspecific divergence of *cis*- and *trans*-regulation associated with species-specific learned vocalization. Society for Neuroscience, Washington, DC, USA. November 11 – 15, 2017.

**Wang H**, Sawai A, Toji N, Sugioka R, Shibata Y, Suzuki Y, Ji Y, Hayase S, Akama S, Sese J, Wada K. (2019). Transcriptional regulatory divergence underpinning species-specific learned vocalization in songbirds. Society for Neuroscience, Chicago, USA. October 19 – 23, 2019.