

A quantitative method for analyzing species-specific vocal sequence pattern and its developmental dynamics



Raimu Imai^a, Azusa Sawai^a, Shin Hayase^a, Hiroyuki Furukawa^a,
Chinweike Norman Asogwa^a, Miguel Sanchez^a, Hongdi Wang^a, Chihiro Mori^a,
Kazuhiro Wada^{a,b,c,*}

^a Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 060-810, Japan

^b Department of Biological Sciences, Hokkaido University, Sapporo, Hokkaido 060-810, Japan

^c Faculty of Science, Hokkaido University, Sapporo, Hokkaido 060-810, Japan

HIGHLIGHTS

- New method is developed for discrimination of species specificity of vocalization.
- Prevalence of the characteristic syllable transitions is calculated.
- It quantitatively evaluates inter- and intraspecific differences of bird song.
- The method allows tracking the developmental trajectories of song patterns.

ARTICLE INFO

Article history:

Received 18 March 2016

Received in revised form 24 June 2016

Accepted 25 June 2016

Available online 30 June 2016

Keywords:

Vocal learning
Vocal development
Zebra finch
Syntax
Song learning
Species specificity
Individual variation

ABSTRACT

Background: Songbirds are a preeminent animal model for understanding the neural basis underlying the development and evolution of a complex learned behavior, bird song. However, only a few quantitative methods exist to analyze these species-specific sequential behaviors in multiple species using the same calculation method.

New method: We report a method of analysis that focuses on calculating the frequency of characteristic syllable transitions in songs. This method comprises two steps: The first step involves forming correlation matrices of syllable similarity scores, named syllable similarity matrices (SSMs); these are obtained by calculating the round-robin comparison of all the syllables in two songs, while maintaining the sequential order of syllables in the songs. In the second step, each occurrence rate of three patterns of binarized “2 rows × 2 columns” cells in the SSMs is calculated to extract information on the characteristic syllable transitions.

Results: The SSM analysis method allowed obtaining species-specific features of song patterns and intraspecies individual variability simultaneously. Furthermore, it enabled quantitative tracking of the developmental trajectory of the syllable sequence patterns.

Comparison with existing method: This method enables us to extract the species-specific song patterns and dissect the regulation of song syntax development without human-biased procedures for syllable identification. This method can be adapted to study the acoustic communication systems in several animal species, such as insects and mammals.

Conclusions: This present method provides a comprehensive qualitative approach for understanding the regulation of species specificity and its development in vocal learning.

© 2016 Elsevier B.V. All rights reserved.

* Corresponding author at: Faculty of Science, Department of Biological Sciences, Hokkaido University, Room 910, Building No. 5, North 10, West 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan.

E-mail address: wada@sci.hokudai.ac.jp (K. Wada).

1. Introduction

The acquisition of species-specific behavioral patterns and their subsequent lifelong maintenance is a crucial research theme in order to understand the diversity and evolution of complex learned behaviors. Specifically, species-specific behaviors, such as mating, foraging, nest construction, and social communication, are generally composed of a rendition of temporal sequence patterns of behavioral units. These sequential behavioral characteristics are represented with species specificity even between closely related species (Berridge, 1990; Donaldson and Young, 2008; Greenspan and Ferveur, 2000; Weber and Hoekstra, 2009; Weber et al., 2013). Therefore, the sequential behavioral pattern is a critical species-specific biosignal for the recognition of conspecifics and the discrimination of heterospecifics in nature (Leininger and Kelley, 2015; Marler and Peters, 1988; Pollack and Hoy, 1979; Woolley and Moore, 2011). However, a limited number of methods have been developed for quantitative analysis of the temporal sequence of behavioral patterns.

Over 3000 species of songbirds around the world have species-specific song patterns. These patterns comprise stereotyped acoustic units, called syllables, which are arranged in characteristic species-specific sequential orders. Similar to that of human language, songbirds learn such song features through auditory and vocal experiences in a defined sensitive period during which juvenile songbirds listen, memorize, and gradually match their own developing vocalizations to the song pattern of an adult tutor (Konishi, 1965; Tchernichovski et al., 2001; Waser and Marler, 1977). An understanding of how syllables are sequentially arranged to form songs patterns is of general interest. This is because the learning and ordering of stereotyped motor patterns are the basis of vocal communication in vertebrates including speech in human. They thus may underlie other aspects of motor control. However, there exist some limitations on the analysis of vocal phenotypes. This was especially evident in the regulation of the temporal sequential patterns of syllables in many songbird species because of the difficulty in identifying each syllable (Gardner et al., 2005; Liu and Nottebohm, 2007). Acoustic features of syllables are variable in multiple factors, such as duration, frequency modulation, amplitude, and pitch, across song renditions and even in a song bout (Tchernichovski et al., 2000). Nevertheless, for the analysis of sequence or syntax of the song, categorization of syllable types was often performed based on human visual inspection of the spectrogram shapes. Although a few studies adapted semi-automatic clustering of syllables based on their mean acoustic features (Deregnacourt et al., 2004; Ravbar et al., 2012; Wu et al., 2008), it remains difficult to set clear standards for identifying unique song syllables. This misleads the researcher given the observer-dependent bias on the identification of syllables, and precludes the extraction of precise information on the temporal sequence pattern of syllables in the song. A similar problem also occurs in the analysis of vocal patterns in mammals. For example the ultrasonic vocalizations produced by many species of rodents consist to a great extent of acoustically variable syllables (Holy and Guo, 2005). In addition, during vocal development, highly variable acoustic features of syllables are observed at the subsong and plastic song stages before the period of vocal crystallization. This makes it difficult to extract the quantitative information on the development of the temporal sequence of syllables.

Here, we adapted a correlation matrix capturing information on the similarity between syllables in two songs, termed as “syllable similarity matrix (SSM),” to extract songs’ temporal structures by quantification of specific syllable transition types rather than transitions of identified syllables. This method allows the detection of the species specificity of song temporal patterns, and quantitative

tracking of the developmental trajectories without human-biased procedures in syllable identification.

2. Materials and methods

2.1. Animals

All zebra finches (ZF, *Taeniopygia guttata*) and the Bengalese finches (BF, *Lonchura striata* var. *domestica*) were laboratory bred. Other species, such as owl finch (OF, *Taeniopygia bichenovii*), star finch (SF, *Neochmia ruficauda*), Java sparrow (JS, *Padda oryzivora*), and canary (CN, *Serinus canaria*) were obtained from local breeders. The photoperiod was constantly maintained at a 13:11 h light/dark cycle with food and water provided *ad libitum*. All bird experiments were performed according to the guidelines specified by the Committee on Animal Experiments at Hokkaido University. The guidelines are based on the national regulations for animal welfare in Japan (Law for the Humane Treatment and Management of Animals; after partial amendment No. 68, 2005).

2.2. Song recording and tutoring

Song recordings were performed as previously reported (Mori and Wada, 2015). Briefly, the birds were individually housed in a sound-attenuation box. A microphone (SHURE microphone SM57, SHURE incorporated, IL, USA) was connected to a FirePod FP10 amplifier (PreSonus Audio Electronics Inc. Florida, USA). Songs were recorded at a sampling rate of 44 kHz and 16 bit amplitude resolution by using Sound Analysis Pro version 1.04 (Tchernichovski et al., 2001). Songs were then automatically saved for 24 h per day by using the Sound Analysis Pro program. Low-frequency and high-frequency noises (<0.5 kHz and >15.8 kHz, respectively) were filtered using Avisoft SASLab pro software (Avisoft Bioacoustics, Berlin Germany). Audacity’s noise-canceling procedure (version 2.0.5) was used to further filter ventilator and fluorescent-delivered noises.

With respect to the song tutoring experiment, the sex of ZF and BF chicks was determined by polymerase chain reaction within 5 post-hatching day (phd), as described previously (Wada et al., 2006). Between 6 and 15 phd, before juveniles could start memorizing a tutor song, they were separated from their fathers and were kept with their mothers and siblings until fledging. Song tutoring was started at 15–20 phd and continued through over 100 phd. After fledging, birds were individually housed in a sound-attenuating box attached containing a mirror to reduce social isolation. Tutor songs were played five times in the morning and five times in the afternoon at 55–75 dB from a speaker (SRS-M30, SONY) passively controlled by Sound Analysis Pro (v1.04) (Tchernichovski et al., 2000).

2.3. Calculation of similarity score between syllables

In this study, a song bout was defined as a continuous production of songs separated by a silent period longer than four folds of the mean value of more than 30 randomly selected inter-syllable gaps. Syllables in the song bouts were used for making SSMs. More than 50 syllables in song bouts were used as a song rendition. Introductory notes in a song were not included in a song rendition. The songs of the ZF, OF, SF, BF, and JS usually contained less than 50 syllables in a song bout. In these cases, two or more song bouts were merged as a combined song rendition to contain more than 50 syllables. For the CN, songs with more than 200 syllables were used because their songs consisted of a considerable number of syllable repetitions when compared with that of the other species. Song spectrograms were formed as WAVE format files (.wav) by Avisoft

SASLab Pro software. Each syllable in the song bouts was identified and individually saved as a WAVE format file by automated section labeling (threshold 0.001 V, hold time 0.01 s, and margin 0 s) by using Avisoft SASLab Pro software. Then, the machine-based separated syllables were double-checked by human eyes for the precision of syllable separation. The individual syllables were next saved as their separated SON format files (.son) labeled with a unique name including information on its sequential ordering position in the song. The series of separated syllable files of two song renditions were transferred to the Avisoft CORRELATOR program for calculating the similarity scores between the syllables of two songs by the round-robin comparison. The score of syllable similarity was calculated as the peak correlation coefficient between two syllables. The two spectrograms of the separated syllables were shifted incrementally past each other along the time axis. For each offset position, the correlation coefficient is computed according to the following formula:

$$\Phi_{XY} = \frac{\sum_X \sum_Y ((a_{xy} - m_a) * (b_{xy} - m_b))}{\sqrt{\sum_X \sum_Y (a_{xy} - m_a)^2 * \sum_X \sum_Y (b_{xy} - m_b)^2}}$$

where m_a and m_b are the mean values of the spectrograms a and b , respectively. a_{xy} and b_{xy} are the intensities of the spectrogram points at the locations x and y , respectively. Syllable similarity score is a value ranging from 0 to 1. A value of 1 means that the two spectrograms are identical. A value of 0 means that there was no similarity between the spectrograms. Supplementary material 1 is shown as an example comparison of similarity scores between syllables with variability in duration and other acoustic features taken from a plastic song sung by a juvenile bird.

Similarity scores between the syllables in two song renditions were exported into cells in Microsoft Excel spreadsheet by keeping the syllable sequence order in the original songs [1st step (left) in Fig. 1A]. The spreadsheet was named with the information of the similarity scores between the syllables as a SSM. In this study, 10 SSMs per bird were prepared by the round-robin comparison of five song renditions. To qualitatively visualize the information of syllable temporal sequences in songs, each cell in the SSM was color-coded according to the value of the similarity score [1st step (right) in Fig. 1A].

2.4. Extraction of frequency of characteristic syllable transitions of interest

For the quantitative analysis of syllable temporal structures, the occurrence rate of characteristic patterns of binarized “2 rows \times 2 columns” cells in the SSMs was calculated. For the binarized patterning of 2×2 cells in the SSMs, the R software program was used to find the most similar binarized pattern for each 2×2 cell in the SSM from 12 possible patterns by comparison with the mean values 0.86 and 0.33 as similar (that was represented by black colors) and different (that was represented by white colors) syllables, respectively (R Core Team, 2013) (Supplementary material 2). The two mean values were defined as the most frequently observed values with two peaks in a total of 240,000 syllable similarity scores from six songbird species (ZF, OF, SF, BF, JS, and CN; $n = 4$ birds/each species, 1000 syllables/bird) (Supplementary material 3). As an alternative way for the binarization of cells in the SSMs, 0.55–0.60 of the similarity scores could be used as a threshold to distinguish similar (black) or different (white) syllables (Supplementary material 4). Then, for each SSM, the prevalence of each binarized 2×2 cell patterns in a SSM was calculated. In this study, particularly three specific types of 2×2 cells patterns were used as characteristic syllable transitions of interest (2nd step in Fig. 1A). Syllable

transition type I was defined as a “paired syllables transition” indicating the existence of two successive syllables that were different but with same sequential order in two songs. This can be illustrated by “song 1 [...A B.....] vs. song 2 [...A B.....]” (in this case, A and B represent two different syllables). Syllable transition type II was a case of existence of the “repetitive syllables transition” by two successive identical or very similar syllables in two songs. For instance “song 1 [.....A A..] vs. song 2 [...A A.....]”. Syllable transition type III was set as a “nonmatching syllables transition” in the case where two successive syllables were not identical within a song and across two songs. An example of this is “song 1 [...A B.....] vs. song 2 [...C D.....]” (in this case, A, B, C, and D represented different syllables). The suitability of calculating each occurrence rate of three binarized 2×2 cells patterns as a syllable transition analysis was examined using artificial models of ZF, BF, and CN song patterns. The detection of species specificity of the model songs (Fig. 1B) was then calculated. The mean of the occurrence rate of the syllable transition types I, II, and III and their coefficients of variation (CV) from 10 total SSMs as an individual animal were used for statistical analyses.

3. Results

3.1. Visualization and extraction of species-specific temporal features of syllable sequence by the SSM analysis

Six species of songbird, namely zebra finch (ZF), owl finch (OF), star finch (SF), Bengalese finch (BF), Java sparrow (JS), and canary (CN) were selected to examine the potential of the SSM analysis method for extracting the species-specific features of the syllable temporal ordering from songs (Fig. 2). The six songbird species were represented by species-specific song patterns including motif structure, combination of syllable chunks and repetition, and repetitive phrases. For instance, SF songs represented the “motif-like” structure, which was not as strictly consistent across song renditions as the motif of ZF songs. The OF and JS produced repetitive syllables as a part of their songs. This was similar to the BF and CN. However, the numbers of repeated syllables and the number of repetitive syllables types were different across these species. Based on the observation of the species-specific features of songs in these species, the SSMs of the songs of each species were visualized as color-coded heat maps to observe the species-specific song temporal patterns qualitatively. The color-coded SSMs from each species represented distinct species-specific unique patterns (Fig. 2). The patterns of SSMs from the songs of the ZF and SF showed striped patterns that were derived from their motif structure. In contrast, the SSMs from the OF, BF, JS, and CN indicated darkened squares with variable sizes, which belonged to syllable-repetitive phrases in their songs. Furthermore, parts of the SSM from the JS songs indicated checkered-flag patterns. This suggested that there were alternative two syllable repetitions, like ABABABA (Fig. 2). Then, the prevalence of the characteristic three types of syllable transitions was calculated from binarized SSMs. These syllable transitions were named as transition type I, II, and III. Syllable transition type I included the consistent sequential production of two different syllables between two songs, as “paired syllables transition”. In contrast, transition type II referred to a repetition of closely resembling syllables between two songs, as “repetitive syllables transition”. Transition type III was defined as a type of syllable transition belonging to different syllable transitions, termed as “nonmatching syllables transition”. Similar to the scenario demonstrated using artificial song model patterns (Fig. 1B), even in the cases that used real songs from these species, the SSM pattern analyses succeeded in extracting species-specific occurrence rates of the characteristic three transition types. Then, different appearance ratios of the transition types were observed among the species (Fig. 2). For instance,

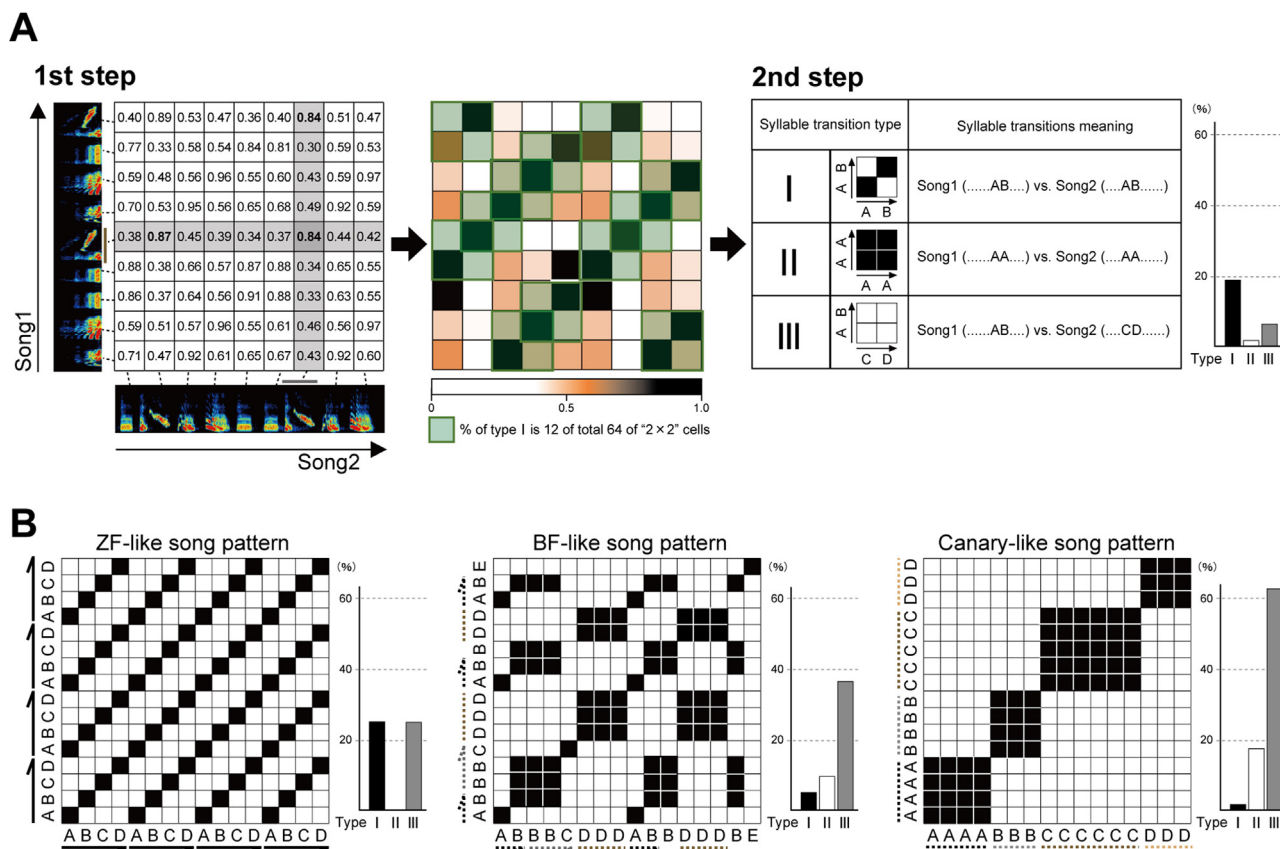


Fig. 1. Syllable similarity matrix (SSM) method for detection of syllable temporal transitions.

(A) The SSM method consisted of two steps. First, a correlation matrix including syllable similarity scores was generated by the round-robin comparison of all syllable comparisons in two songs, keeping the sequential order of the syllables in the songs. Second, the occurrence rate of three patterns of binarized “2 rows × 2 columns” cells in the SSM was calculated as a percentage of the characteristic syllable transition types I, II, and III (see Materials and methods).

(B) Test analyses of the SSM-based method were conducted by using artificial song models mimicking the songs of the zebra finch (ZF), the Bengalese finch (BF), and the canary (CN) that included motif, chunk, and repetition structures, respectively, as their typical syllable transitions. The occurrence rates of the syllable transition types I, II, and III in the three song models are represented as bar graphs.

the ZF song typically showed a higher value in the transition type I compared with those of other species. In contrast, OF, JS, and CN songs had higher values in the transition type II when compared with the other species. However, the three species possessed different values in the transition type III. This species specificity of the prevalence of the syllable transition types was clearly shown in three- or two-dimensional maps (Fig. 3A and B).

Additionally, these dimensional maps represented intraspecies individual variability of the temporal patterns. For example, in the ZF song, the motif structure was observed as a typical species-specific feature. Nevertheless, the number of syllable types composed in the structure ranging with 2–5 syllables was different across individuals. Such individual differences of the motif structure were shown as different values of the transition type I. That is, the motif structure formed with three syllable types showed higher values of type I than the motif structure formed with four syllable types. In the group data of each species, the value and ratio of mean occurrence rates of the three transition types and its CV were represented differently across the six species. This indicated species-specific traits of the syllable transition patterns (Fig. 3C). For example, it is known that some ZFs produce repetitive call-like syllable transitions at the end of their song motifs. Such variability in syllable repetition was detected as a high CV value of syllable transition type II in the ZF (Fig. 3C). These results implied that the analysis of the SSM-based method could be suitable for discriminating the species-specific features of syllable sequence transition in the crystallized songs at an adult stage in multiple songbird species.

3.2. Detection of the development of the species specificity of syllable transitions in a song

To elucidate the development of species-specific song patterns in songbirds, the SSM analysis method was further adapted for monitoring the developmental trajectory of the syllable transition of songs through the various song acquisition stages. It was a challenge to analyze the development of species-specific song patterns across multiple species by using the same platform of the analysis algorithm. This was due to the difficulty in quantitatively calculating the amorphous vocal patterns called subsong and plastic song. These were observed as intermediate states until song crystallization. The analysis using the SSM method did not require identification or alphabetization of syllables in songs. Therefore, the quantitative development of species-specific song patterns was analyzed regardless of the existence of highly variable acoustic features of syllables during subsong and plastic song stages.

For this purpose, song files from ZFs and BFs were used. These were recorded through development under playback tutoring conditions with conspecific songs (each $n = 4$) (Fig. 4). The results of the SSM analyses for the development of the species specificity of song patterns revealed that the songs of both ZF and BF juveniles did not indicate any characteristic patterns in the SSMs at 40–50 phd. They indicated low occurrence rates of both paired- (transition type I) and repetitive- (transition type II) syllable transitions but with high scores of transition type III. Then, at 60–70 phd, the differences of the species specificity of song patterns appeared in the song

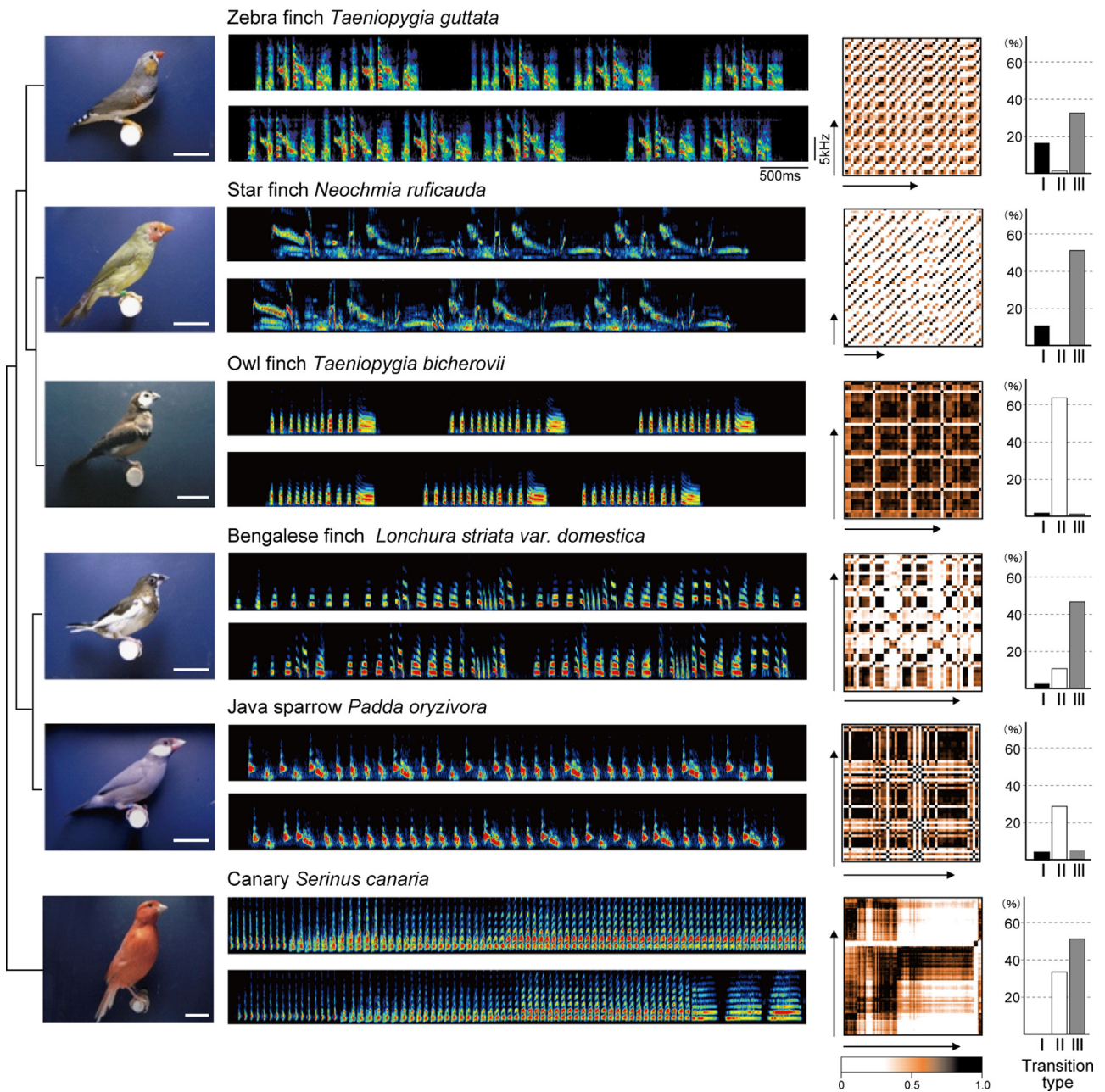


Fig. 2. Visualization and extraction of the species-specificity of song temporal patterns by the SSM method.

Examples of song spectrograms, color-coding SSMs, and mean occurrence rates of the syllable transition types I, II, and III in the SSMs of six species of the songbird. Their phylogenetic relationship was represented as a dendritic tree. Two arrow lines on the sides of the SSMs represent parts of songs shown as two spectrograms. Color-coding of the SSMs indicates the value of the syllable similarity scores (darker brown color means higher similarity score between compared syllables). Scale bars (in the images of birds) = 2 cm.

developmental trajectory in the two-dimensional maps of transitions I and II between ZF and BF (Fig. 5). Interestingly, each juvenile from the two species showed an individually unique trajectory of song development with distinct combination of drastic changes and slumber states of both transition types I and II during 10 day intervals until 100 phd. For instance, bird ZF3 showed the largest change of the mean occurrence rate in transition type I between 50 and 60 phd. The bird then slowly crystallized its song pattern by the iteration of minor changes (ZF3 in Fig. 4C). Furthermore, individual variability in song learning strategy was also observed among birds keeping the same species-specific developmental constraints (Fig. 4C and D). Although both BF1 and BF4 represented unique individual trajectories of song development similar to the tutor song,

BF1 increased only in the repetition transition shown as an accumulation of the value of the transition type II. However, BF4 increased in both paired (transition type I) and repetition (transition type II) transitions (BF1 and BF4 in Fig. 4D). In the species group data, the song development trajectory of the ZFs and BFs revealed a gradual increase in the species-specific prevalence of syllable transitions I and II, respectively, over the course of development ($p = 0.00423$ and $p = 3.88E-06$, respectively, two-way ANOVA). In contrast, the prevalence of the syllable transition type III, “nonmatching syllables transition”, continuously decreased during song development in the ZFs and BFs (Fig. 5, upper panels). However, there was a trend toward differences in the mean number of Type III transitions in ZF vs. BF starting from an early developmental stage, with more tran-

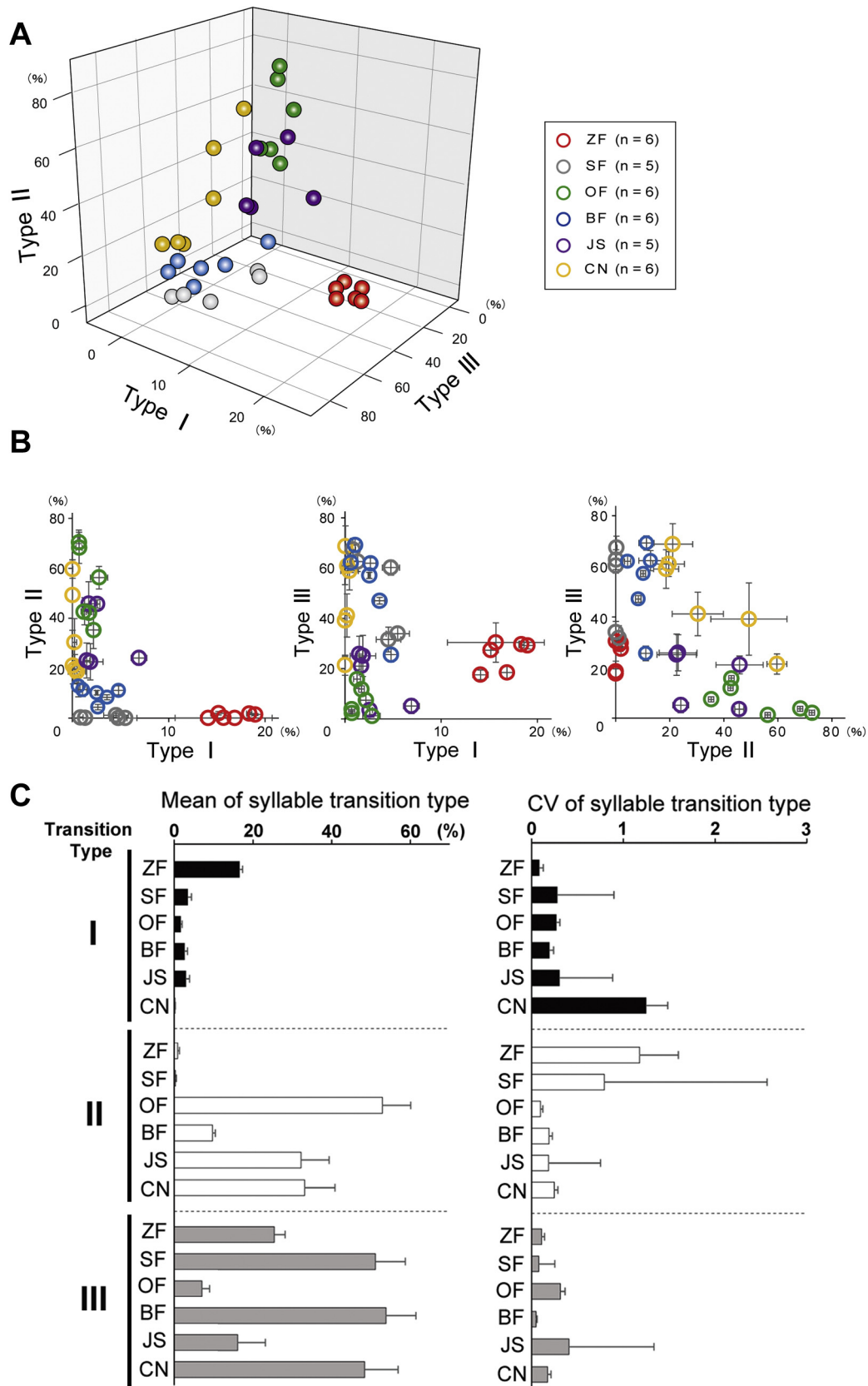


Fig. 3. Species-specific features of the syllable transitions represented by the SSM analyses.

(A) Three-dimensional plot showing the prevalence of the three transition types (transition patterns I, II, and III) of syllables in the six species of the songbirds (red: ZF; $n=6$, gray: SF; $n=5$, green: OF; $n=6$, blue: BF; $n=6$, purple: JS; $n=5$, and orange: CN; $n=6$).

(B) Two-dimensional plots for the prevalence of the syllable transition types shown in (A). Each dot represents mean \pm S.D. of 10 SSMs in a bird.

(C) Mean of the occurrence rate of the three types of the syllable transitions and its CV in six songbird species. Bar graphs represent mean \pm S.E.M.

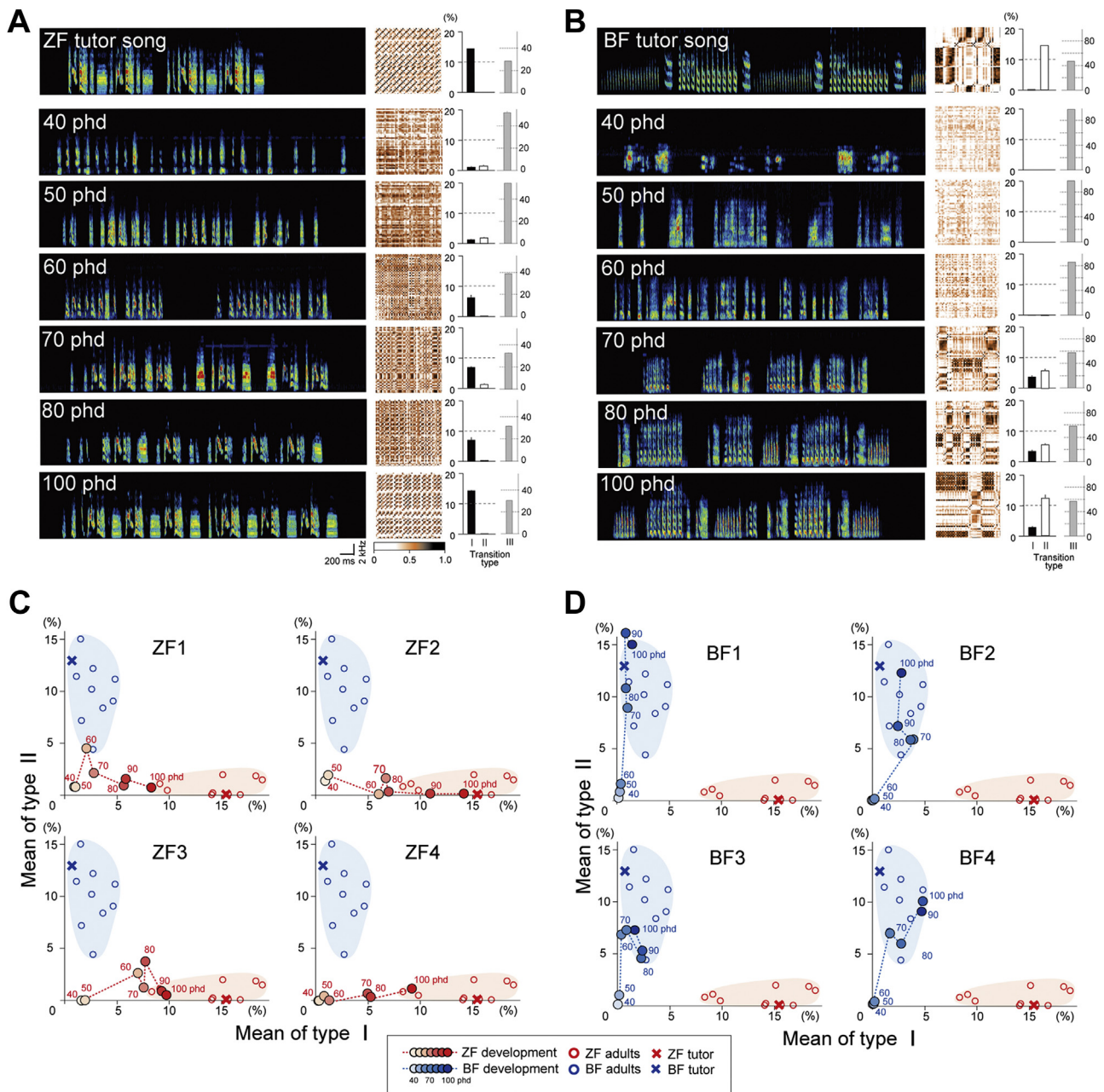


Fig. 4. Developmental dynamics of the species specificity of the song patterns in ZF and BF.

(A), (B) An example of developmental changes in the song patterns, color-coding SSMs, and prevalence of the syllable transition types I and II in ZF (A) and BF (B), which were tutored with a playback conspecific song.

(C), (D) Individual developmental trajectories of the song patterns represented with the syllable transition types I and II in ZFs (C) and BFs (D) (each n=4). Open circles denoted songs of adult ZFs (red) and BFs (blue). The red and blue colored cross marks presented tutored ZF and BF songs, respectively.

sitions observed in BF than in ZF (Fig. 5). Higher values of Type III transitions stems from songs with more variable and unique syllables. Therefore, this difference in Type III transitions suggests that BF juveniles start and continue to sing with more unstable syllables than ZF juveniles do over the course of song development. The CV of syllable transition types I, II, and III did not show significant differences between the ZFs and BFs. This suggests that both species had similar fluctuation rates in the prevalence of the syllable transitions (Fig. 5, lower panels). These results indicated the potential of the SSM analysis for studying the species-specific vocal development and monitoring individual unique traits in a variety of species.

4. Discussion

The SSM analysis method enabled the extraction of the species-specific features in syllable sequence transitions. It also allowed the examination of the developmental process of the features during song learning in songbirds. The SSM analysis of song patterns is based on obtaining and processing the prevalence of selected characteristic types of the syllable transitions (Fig. 1A). Irrespective of its simplicity, the analysis revealed the species-specific distribution of syllable transition types among the species and showed intraspecies individual variation (Figs. 3 and 4). The CV of the prevalence of characteristic transition types is also useful for detecting

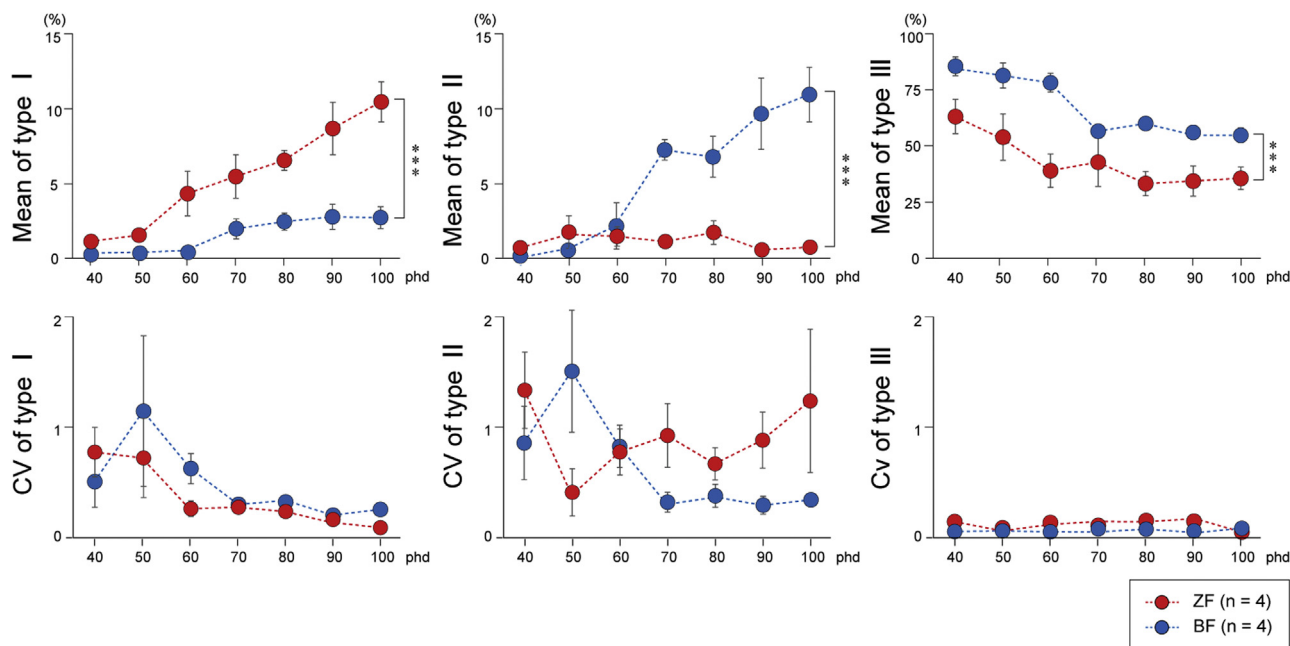


Fig. 5. Developmental dynamics of the species specificity of the song patterns in ZF and BF.

Developmental changes of the mean of occurrence rate of the syllable transition types I, II, and III and its CV in the ZF and BF (each $n=4$, mean of type I; species \times phd $F_{1,6}=3.797$, $*p<0.01$, mean of type II; species \times phd $F_{1,6}=8.724$, $**p<0.0001$, two-way ANOVA). Error bars = S.E.M.

species-specific variability of each syllable transition types (Fig. 3C). In addition, CV values of transition types indicates behavioral fluctuations occurring during the song development, as shown by higher CV values of the syllable transition I and II in the ZF and BF, respectively, at juvenile vs. adult stages (Fig. 5).

To gain a precise understanding of song development, two vocal parameters are crucial in analyzing song phenotype, namely how a juvenile develops “syllable acoustics” as the behavioral units and “syllable sequence” as the control of the behavioral unit ordering. Recent computer-aided sound recordings and analyses led to significant improvements in quantitative song analysis, especially with respect to the study of developmental changes of the syllable acoustic phenotypes (Deregnacourt et al., 2005; Tchernichovski et al., 2001). Although a few studies have adapted semi-automatic clustering of syllables based on their mean acoustic features, studying the development of the temporal sequence of syllables in songs remains a significant challenge (Daou et al., 2012; Lipkind et al., 2013; Ravbar et al., 2012; Wu et al., 2008). Many recent analyses of the syllable sequence are based on the songs with identified/alphabetized syllables (Horita et al., 2008; Okanoya and Yamaguchi, 1997; Scharff and Nottebohm, 1991), as it has been difficult to extract features of the syllable transitions with acoustically ambiguous syllables. Therefore, the analysis of syllable sequences was constrained and used a limited number of songbird species which produced consistently stable syllables across song renditions (Gardner et al., 2005; Liu et al., 2004; Liu and Nottebohm, 2007; Marler and Slabbekoorn, 2004; Rose et al., 2004). A similar problem was observed in the analysis of the development of syllable sequences in the song. This was because of a difficulty in identifying acoustically unstable syllables/notes during song learning.

Unlike the previous analysis methods based on syllable categorization, the SSM method quantifies the occurrence rate of characteristic types of the syllable transitions, not transitions of identified specific syllables. Therefore, the SSM method quantitatively detects features of the temporal sequence of songs, including species-specific differences and developmental dynamics. For instance, a majority of syllables at around 40–50 phd of ZF and BF did not have any clear transition rules and were identified as

Type III transitions (Fig. 4A and B). In contrast, some of other syllables in the same song renditions were detected as mixed population of Type I and II transitions, meaning of the initiation of the species-specificity. Detection of the onset of the species-specificity of vocal patterns is a unique strength of SSM method. Therefore, as a new quantitative method for analyzing species-specific vocal sequence patterns and their developmental dynamics, the SSM method can complement existing methods already adapted for quantification of specific syllable transitions. However, the SSM-based analysis method still has room for improvements to enable a more precise and fine detection of species-specific features in at least two points, the calculation of syllable similarity, and the pattern recognition of the SSM. With respect to the calculation of syllable similarity, parametric methods such as pitch and frequency modulation of syllables could provide more precise similarity scores than the methods based on the comparison of the image patterns of spectrograms. This could be especially applicable in a case where syllables with complicated spectrogram structures are compared (Tchernichovski et al., 2000). As for the pattern recognition of the SSM, the threshold setting for cell binarization could be arbitrarily chosen by researchers, depending on their research aims. Two different threshold settings were compared for cell binarization: “the mean value of 0.86 and 0.33 (as used in this study)” and “the similarity score 0.60 threshold (as boundary similar and different).” Consistency of the occurrence rates of the syllable transition types I, II, and III was observed using these two different thresholds for individuals of six species (Supplementary material 4). However, any specific threshold set for cell binarization is problematic for the subsong of juvenile birds and for adults in a species with variable syllable performance. To solve this potential problem, instead of using binarized 2×2 cells patterns for the fine detection of the species-specific features from the SSM pattern, machine-learning algorithms, such as supervised and/or anomaly learning methods, could be powerful tools for precise pattern recognition from SSM without cell binarization. Additionally, our SSM analysis did not include information on the inter-syllable silent gap. This is a crucial factor for regulating the song tempo (Saar and Mitra, 2008; Sasahara et al., 2015). However, even the simplest method of focus-

ing on a limited number of transition types of syllables in the SSMs, allowed the detection of species-specific song features and vocal development by monitoring individual unique traits in songbird species.

Notably, the experiment with the song tutoring in the ZF and BF by using the SSM analyses revealed that all juveniles learned a certain degree of the tutored song features. This indicated a developmental drifting approach to the tutored songs in the two-dimensional plot with the syllable transition types I and II (Fig. 4C and D). However, the developmental trajectory of the prevalence of the syllable transitions was unique among individuals. This represented individual variability in learning strategy during song development (Liu et al., 2004). Therefore, these results indicated that SSM-based analysis is a useful method for studying the developmental strategy of acquisition of the sequential behavior and its individual variability. Furthermore, the SSM analysis could be used to assess the strength of vocal learning in comparing the song of tutees and tutors. In this comparison, the SSM analysis simultaneously extracts similarity information related to both syllable transition (sequence) and acoustics (spectral) learning between the tutor's model and tutee's developing songs. In addition, the SSM method could be potentially adapted for evaluating the effects of experimental manipulation on the regulation of behavioral patterns. The changes in syllable transition regulation caused by deafening procedures and the developmental abnormality of specific syllable transitions in socially isolated birds were successfully detected to-date. It will be reported as a separate study in the near future. Also, the SSM-based analysis for sequential behaviors could be adapted to analyze other vocal behaviors that are species-specifically regulated, such as mating songs in insects and ultrasonic vocal communications in rodents, as well as the development of speech acquisition in human babies.

The extraction of occurrence rate of the characteristic syllable transitions from the SSM patterns represents an effective approach to evaluate species-specific song patterns and to track the developmental trajectory of the temporal regulation of syllable sequence, taking into account the intraspecies individual variability.

Conflict of interest

There are no conflicts of interest.

Acknowledgments

We thank Drs. W.-C. Liu, S. Kojima, M. Soma, T. Matsushima, and M. Mizunami for providing critical comments. This work was supported by Grant-in-Aid for Japan Society for the Promotion of Science Fellows (to R.I. and C.M.), the MEXT scholarship (to C.N.A. and H.W.), and the Asahi Glass Foundation and Grants-in-Aid for Scientific Research 225640097 and 25290063 (to K.W.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jneumeth.2016.06.023>.

References

- Berridge, K.C., 1990. Comparative fine structure of action: rules of form and sequence in the grooming patterns of six rodent species. *Behaviour*, 21–56.
- Daou, A., Johnson, F., Wu, W., Bertram, R., 2012. A computational tool for automated large-scale analysis and measurement of bird-song syntax. *J. Neurosci. Methods* 210, 147–160.
- Deregnacourt, S., Mitra, P.P., Feher, O., Maul, K.K., Lints, T.J., Tchernichovski, O., 2004. Song development: in search of the error-signal. *Ann. N. Y. Acad. Sci.* 1016, 364–376.
- Deregnacourt, S., Mitra, P.P., Feher, O., Pytte, C., Tchernichovski, O., 2005. How sleep affects the developmental learning of bird song. *Nature* 433, 710–716.
- Donaldson, Z.R., Young, L.J., 2008. Oxytocin, vasopressin, and the neurogenetics of sociality. *Science* 322, 900–904.
- Gardner, T.J., Naef, F., Nottebohm, F., 2005. Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308, 1046–1049.
- Greenspan, R.J., Ferveur, J.-F., 2000. Courtship in *Drosophila*. *Annu. Rev. Genet.* 34, 205–232.
- Holy, T.E., Guo, Z., 2005. Ultrasonic songs of male mice. *PLoS Biol.* 3, e386.
- Horita, H., Wada, K., Jarvis, E.D., 2008. Early onset of deafening-induced song deterioration and differential requirements of the pallial-basal ganglia vocal pathway. *Eur. J. Neurosci.* 28, 2519–2532.
- Konishi, M., 1965. The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z. Tierpsychol.* 22, 770–783.
- Leininger, E.C., Kelley, D.B., 2015. Evolution of courtship songs in *Xenopus*: vocal pattern generation and sound production. *Cytogenet. Genome Res.* 145, 302–314.
- Lipkind, D., Marcus, G.F., Bemis, D.K., Sasahara, K., Jacoby, N., Takahasi, M., Suzuki, K., Feher, O., Ravbar, P., Okanoya, K., Tchernichovski, O., 2013. Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*.
- Liu, W.C., Nottebohm, F., 2007. A learning program that ensures prompt and versatile vocal imitation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 20398–20403.
- Liu, W.C., Gardner, T.J., Nottebohm, F., 2004. Juvenile zebra finches can use multiple strategies to learn the same song. *Proc. Natl. Acad. Sci. U. S. A.* 101, 18177–18182.
- Marler, P., Peters, S., 1988. The role of song phonology and syntax in vocal learning preferences in song sparrow, *Melospiza melodia*. *Ethology* 77, 125–149.
- Marler, P., Slabbekoorn, H., 2004. *Nature's Music: The Science of Birdsong*. Elsevier Academic Press.
- Mori, C., Wada, K., 2015. Audition-independent vocal crystallization associated with intrinsic developmental gene expression dynamics. *J. Neurosci.* 35, 878–889.
- Okanoya, K., Yamaguchi, A., 1997. Adult Bengalese finches (*Lonchura striata* var: *domestica*) require real-time auditory feedback to produce normal song syntax. *J. Neurobiol.* 33, 343–356.
- Pollack, G.S., Hoy, R.R., 1979. Temporal pattern as a cue for species-specific calling song recognition in crickets. *Science* 204, 429–432.
- R Core Team, 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 <http://www.R-project.org/>.
- Ravbar, P., Lipkind, D., Parra, L.C., Tchernichovski, O., 2012. Vocal exploration is locally regulated during song learning. *J. Neurosci.* 32, 3422–3432.
- Rose, G.J., Goller, F., Gritton, H.J., Plamondon, S.L., Baugh, A.T., Cooper, B.G., 2004. Species-typical songs in white-crowned sparrows tutored with only phrase pairs. *Nature* 432, 753–758.
- Saar, S., Mitra, P.P., 2008. A technique for characterizing the development of rhythms in bird song. *PLoS One* 3, e1461.
- Sasahara, K., Tchernichovski, O., Takahasi, M., Suzuki, K., Okanoya, K., 2015. A rhythm landscape approach to the developmental dynamics of birdsong. *J. R. Soc. Interface* 12.
- 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. *J. Neurosci.* 11, 2896–2913.
- Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B., Mitra, P.P., 2000. A procedure for an automated measurement of song similarity. *Anim. Behav.* 59, 1167–1176.
- Tchernichovski, O., Mitra, P.P., Lints, T., Nottebohm, F., 2001. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291, 2564–2569.
- Wada, K., Howard, J.T., McConnell, P., Whitney, O., Lints, T., Rivas, M.V., Horita, H., Patterson, M.A., White, S.A., Scharff, C., Haesler, S., Zhao, S., Sakaguchi, H., Hagiwara, M., Shiraki, T., Hirozane-Kishikawa, T., Skene, P., Hayashizaki, Y., Carninci, P., Jarvis, E.D., 2006. A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes. *Proc. Natl. Acad. Sci. U. S. A.* 103, 15212–15217.
- Waser, M., Marler, P., 1977. Song learning in canaries. *J. Comp. Physiol. Psychol.* 91, 1–7.
- Weber, J.N., Hoekstra, H.E., 2009. The evolution of burrowing behaviour in deer mice (*genus Peromyscus*). *Anim. Behav.* 77, 603–609.
- Weber, J.N., Peterson, B.K., Hoekstra, H.E., 2013. Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. *Nature* 493, 402–405.
- Woolley, S.M., Moore, J.M., 2011. Coevolution in communication senders and receivers: vocal behavior and auditory processing in multiple songbird species. *Ann. N. Y. Acad. Sci.* 1225, 155–165.
- Wu, W., Thompson, J.A., Bertram, R., Johnson, F., 2008. A statistical method for quantifying songbird phonology and syntax. *J. Neurosci. Methods* 174, 147–154.