J Comp Physiol A (2004) 190: 1001–1010
DOI 10.1007/s00359-004-0556-x

ORIGINAL PAPER

**Timothy Q. Gentner · Stewart H. Hulse**
**Gregory F. Ball**

# Functional differences in forebrain auditory regions during learned vocal recognition in songbirds

Received: 19 March 2004 / Revised: 30 July 2004 / Accepted: 31 July 2004 / Published online: 21 September 2004


**Abstract** Converging evidence implicates the auditory forebrain regions caudal medial mesopallium (formerly cmHV) and caudal medial nidopallium in the perceptual processing of conspecific vocalizations in songbirds. Little is known however, about more specific processing within these regions especially during song-based perceptual behaviors. One hallmark of the caudal medial mesopallium and caudal medial nidopallium, areas analogous to mammalian secondary auditory cortical structures, is their robust expression of the immediate-early-gene *zenk* in response to conspecific songs. Using European starlings operantly trained to recognize the songs of individual conspecifics, we show that the levels and patterns of *zenk* protein expression in the caudal medial nidopallium and caudal medial mesopallium differ when song recognition demands are placed on the system. In the caudal medial mesopallium, expression is significantly elevated above basal levels during the recognition of familiar songs, the acquisition of novel associations for familiar songs, and the acquisition of novel song discriminations. In the caudal medial nidopallium, however, expression is significantly elevated above basal levels only during the acquisition of novel song discriminations. The results directly implicate the caudal medial nidopallium and caudal medial mesopallium in at least a portion of the auditory processes underlying vocal recognition. Moreover, the observed differences between these regions imply the functional localization (or at least the concentration) of different auditory processing mechanisms within the caudal medial nidopallium and the caudal medial mesopallium.

**Keywords** Birdsong · Memory · Perception · Recognition · Representation

**Abbreviations** CM: Caudal mesopallium · CLM: Caudal lateral mesopallium · CMM: Caudal medial mesopallium · NCM: Caudal medial nidopallium

T. Q. Gentner (✉)
Department of Organismal Biology and Anatomy,
University of Chicago, 1027 E. 57th Street,
Chicago, IL, 60637, USA
E-mail: tgentner@ucsd.edu

S. H. Hulse · G. F. Ball
Department of Psychological and Brain Sciences,
Johns Hopkins University,
3400 North Charles Street, Baltimore, MD 21218, USA

*Present address*: T. Q. Gentner
Department of Psychology, University of California, San Diego,
9500 Gilman Drive, La Jolla, CA, 92093-0109 USA

## Introduction

Oscine passerine birds (songbirds) exhibit acoustically rich vocal communication behaviors that are well studied, and therefore provide an excellent model system to examine the neurobiological mechanisms that mediate the perception and cognition of natural sounds. To this end, we are studying the neural mechanisms of one common song-based perceptual behavior, individual vocal recognition, in a species of songbird, the European starling (*Sturnus vulgaris*). In this study, we examine neural activation across different forebrain regions while birds are actively recognizing conspecific songs.

Male starlings tend to sing in long continuous episodes (bouts). Song bouts, in turn, are composed of much smaller acoustic units referred to as motifs (Adret-Hausberger and Jenkins 1988; Eens et al. 1991) that are composed of still smaller units called notes (Fig. 1a–c). Although several notes may occur in a given motif, the note pattern within a motif is usually highly stereotyped between successive renditions of that motif. Commonly, each motif is repeated two or more times before the next one is sung. Thus, starling song appears (acoustically) as a sequence of changing motifs, where each motif is an acoustically complex event (Fig. 1d). Although some sharing of motifs does occur among captive males (Hausberger and Cousillas 1995; Hausberger 1997), the motif repertoires of different males living in the wild are generally unique (Adret-Hausberger and Jenkins 1988;

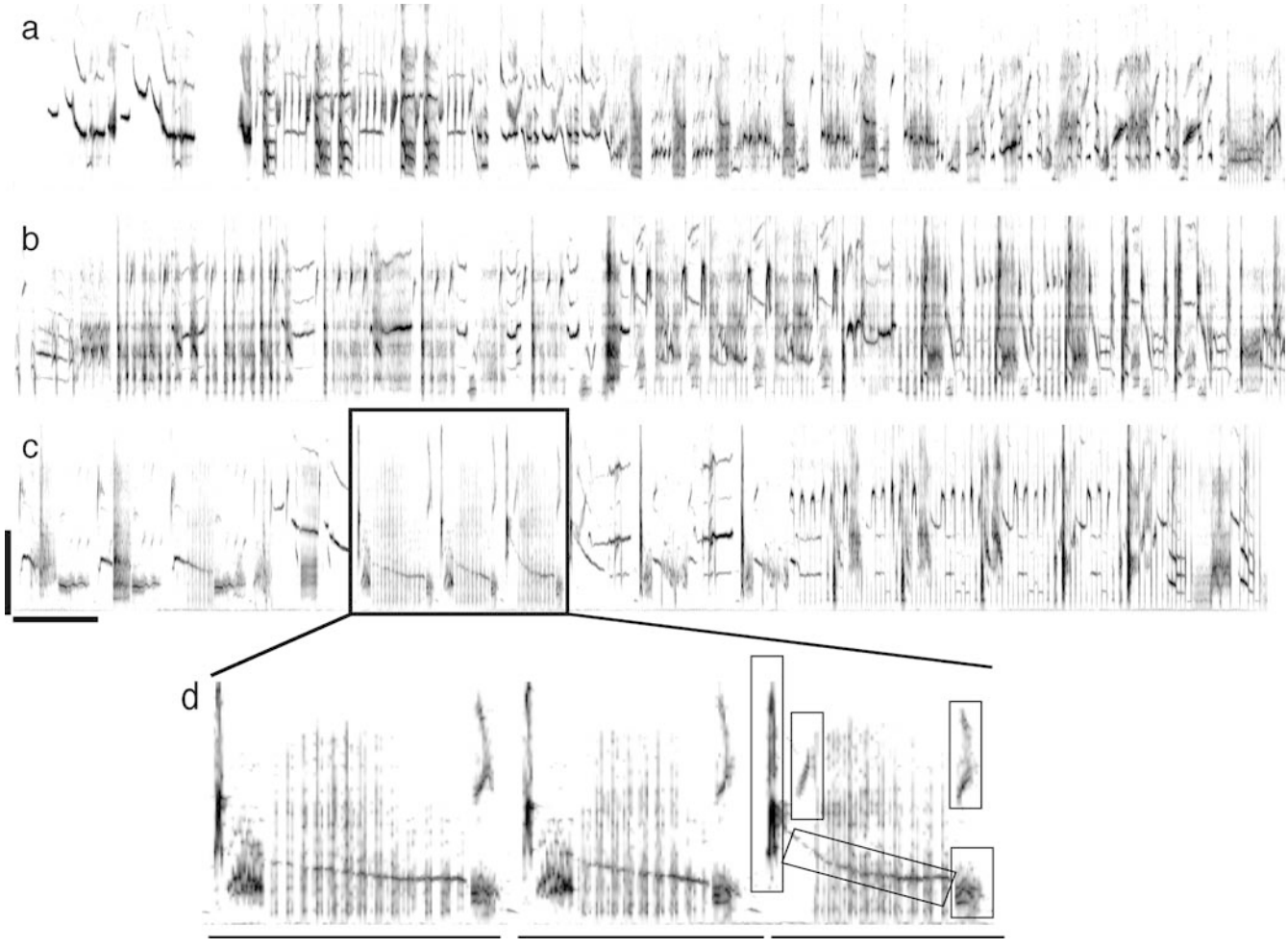



**Fig. 1a–d** Stimulus exemplars. **a–c** Sonograms of three song stimulus exemplars, showing one of the eight song stimuli that comprised each of the three stimulus sets. Each exemplar (**a–c**) is a sample of continuously recorded singing from a different male starling. *Horizontal* and *vertical* scale bars show 1 s and 10 kHz, respectively. **d** A temporally expanded view of the three motifs outlined in **c**. Each separate motif is *underlined*, and the notes in the final motif are *outlined*. Note the largely stereotyped pattern of notes within the three motifs

Eens et al. 1989, 1991; Chaiken et al. 1993; Gentner and Hulse 1998). Thus, the presence or absence of specific motifs in a given bout provides diagnostic cues for song recognition. Behavioral data from several studies support the notion that motifs form the basic units of individual song recognition in starlings (Gentner and Hulse 1998, 2000; Gentner et al. 2000).

A number of forebrain auditory regions (Fig. 2) likely play a role in the recognition of individual songs. Field L is the primary telencephalic target for auditory information from the thalamus, and interconnected subregions of field L (see Fig. 2) project to the caudal medial nidopallium (NCM) and reciprocally to the caudal lateral mesopallium (CLM; formerly clHV [1]). The NCM and CLM, in turn, share reciprocal connections with the caudal medial mesopallium (CMM; formerly cmHV[1]). Responses to conspecific song are observed throughout these forebrain regions, but song selectivity varies among them. Neurons in L1 and L3 show greater selectivity to species-specific vocalizations than those in L2 (Leppelsack and Vogt 1976; Bonke et al. 1979; Müller and Leppelsack 1985; Theunissen and Doupe 1998). Response selectivity continues to increase along the processing hierarchy into NCM and the caudal mesopallium (CM) (Müller and Leppelsack 1985), where selective responses to conspecific (compared to heterospecific) songs are observed (Grace et al. 2003). Increasing song selectivity along the processing hierarchy is coincident with an increase in the non-linear components of the neural responses (Sen et al. 2001), suggesting that 'higher' regions in the pathway, i.e., NCM and CMM, are involved in the extraction of complex acoustic features (e.g., Leppelsack 1983).

Recent data from adult starlings directly support the role of CMM in the extraction of complex song features, and suggest how the system may function in song recognition behavior. Following operant song recognition training, many cells in CMM show selective responses to acoustic features of specific song motifs, but

[1]See Reiner et al. 2004 for recent changes in avian brain nomenclature

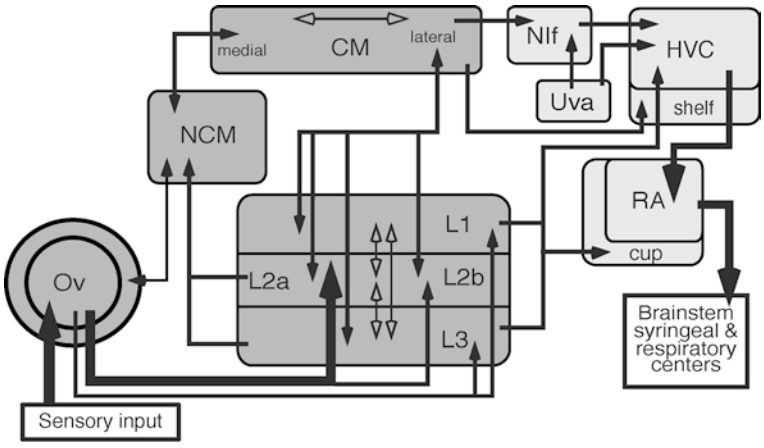



**Fig. 2** Songbird auditory system. Schematic drawing of the major forebrain auditory regions (in *dark gray*) and their pattern of connectivity. Nuclei of the classic 'song system' primary motor pathway are shown in *light grey*. The present study examines IEG protein expression in the medial CM and the NCM. *Ov* ovoidalis; *L1-3* field L complex; *NIf* nucleus interfacialis; *Uva* nucleus uvaeformis; *RA* robustus archistriatalis; *HVC* used as a proper name

only those motifs present in songs that birds have learned to recognize (Gentner and Margoliash 2003a). Moreover, the strength of these experience-dependent selective responses is related to the operant contingencies used to train the song recognition (Gentner and Margoliash 2003a).

A second line of research supporting the role of NCM and the CM in the processing of conspecific song comes from studies of stimulus driven expression of the immediate–early-gene (IEG) *zenk*. The density of *zenk* expression in both NCM and CM is higher in response to the presentation of conspecific songs than to a variety of other acoustic stimuli (Mello et al. 1992), and the *zenk* response is tuned to the acoustics of particular conspecific song syllables (Ribiero et al. 1998). Reminiscent of the song-specific habituation seen in NCM neuronal spike rates (Chew et al. 1995; Stripling et al. 1997), *zenk* expression also habituates to the repeated presentation of the same conspecific song on a similar time-scale (Mello et al. 1995). In addition, pairing the presentation of song with an aversive stimulus leads to increased expression relative to controls in which song is unpaired, suggesting that associative mechanisms may modulate at least a portion of the NCM IEG response (Jarvis et al. 1995).

Individual vocal recognition of same-species members, in which songs are associated with individuals, or groups of individual conspecifics, is widespread among songbirds (Falls 1982; Stoddard 1996). Recognition of this form can be considered as a classification problem in associative learning and memory, in which target sounds (i.e., songs) are classified or associated with some specific referent (e.g., a mate, intruder, motivational state, location in space, etc), and the association retained until the sound's next occurrence. Thus, interaction between stimulus-driven neuronal activation in NCM/CMM and response components linked to reward, attention or other so-called 'top–down' mechanisms is not unexpected given the putative role of these high-level auditory structures in individual song recognition. Nonetheless, the respective contributions of NCM and CMM to the representational and associative processing of behaviorally relevant songs are not known. For a variety of technical reasons, electrophysiological recordings in songbirds that are actively engaged in song recognition are difficult (cf. Gentner and Margoliash 2003b). In the present study, we use IEG *zenk* protein expression to examine population level activation patterns across NCM and CMM while birds are actively engaged in different song recognition tasks. The results show that these two regions contribute in significantly different ways to the recognition of familiar songs, and the learning of novel song discriminations.

## Materials and methods

### Subjects

Seventeen adult male European starlings (*Sturnus vulgaris*) served as subjects for this experiment. Subjects were wild caught as adults, and housed in gang cages on a natural-light photoperiod prior to behavioral training and testing.

### Stimuli

Recordings of the songs from six male European starlings were used to generate all the stimuli for this experiment. Song recording procedures have been detailed elsewhere (Gentner and Hulse 1998). Briefly, we recorded a minimum of 0.5 h of song from each male in a large sound-attenuating booth. During recording, each male had visual and auditory access to the same single female starling. We selected four song bouts from all the song bouts recorded from each male, and from each selected bout drew one, roughly 15 s, sample of continuous singing (Fig. 1a–c). This gave us four song samples from each of six males. Similar song bout samples have been used extensively to study vocal recognition behavior in this species (Gentner and Hulse 1998, 2000; Gentner et al. 2000; Gentner and Margoliash 2003a). We dived the song samples into two sets of

"baseline training" songs, and one set of "novel transfer" songs. Each of these three sets contained four song samples from two different males (eight songs total). Half the subjects were trained using one baseline set, the other half were trained with the other baseline set (see below). For training on a given baseline set, the four songs from one male were associated with one operant response and the four songs from the second male were associated with another operant response (see below). The novel transfer stimuli were used with a subset of the birds during the testing phase. Any familiarity with the song stimuli prior to baseline training is unlikely, as zthe birds serving as subjects and those used to generate the song stimuli were captured at field sites over 100 miles apart.

## Behavioral training

Subjects learned to work an operant apparatus mounted inside a sound-attenuating chamber (Fig. 3a; see Gentner and Hulse 1998 for complete description) by pecking at three small buttons for food reward. The baseline training procedure taught each subject to discriminate between two sets of song stimuli (see above) using a well-established discrimination procedure (Hulse 1995; Macmillan and Creelman 1991). The subject started each baseline training trial by pecking the center button, which initiated the playback of a randomly selected song stimulus. Once the song stimulus finished, the subject could peck either the left or right response button to receive reinforcement. For each subject, four songs (all sung by one male) were paired with the left response and the four other songs in the set (all sung by a second male; see above) were paired with the right response. If the subject responded correctly, by pecking the response button with which a given song was paired, it was rewarded with 2.5 s access to the food hopper. If the subject responded incorrectly, by pecking the opposite response button, e.g., the left response button following a stimulus associated with the right response (or vice versa), it was punished with a 6-s timeout during which the house light was extinguished and no food access was allowed. Assignment of the songs within a baseline set to either the left or right response button was counterbalanced across subjects. Each subject was trained on only one of the two baseline training stimulus sets (see above).

In addition to causing a timeout, an incorrect response initiated a correction-trial sequence in which the same exemplar repeated on all subsequent trials until the bird responded either correctly or not at all. In the event that a bird failed to respond within 5 s following the completed presentation of a given exemplar, the trial ended without a timeout, and the computer waited for a center key peck to begin the next trial. Prior to testing, the maximum sound level of the acoustic stimuli inside the test chamber was set at $70 \pm 2$ dB (A) SPL. During training and testing animals were maintained on a closed economy with a single 11-h daily session, starting at 8:00 a.m.

Once each subject reached asymptotic performance on the baseline training procedure (set as four consecutive days at 85% correct or better), we transferred the animal to a 'pretest' condition. Trials in the pretest condition used the same stimuli as the baseline trials, but did not require a peck to the center button to initiate a song presentation. Instead, trials began every 30 s, so that during the pretest sessions each animal was exposed to the same total number of stimuli per unit time. Once the subject performed at or above 90% correct for 3 consecutive days under pretest conditions, the animal was assigned to one of four test conditions.

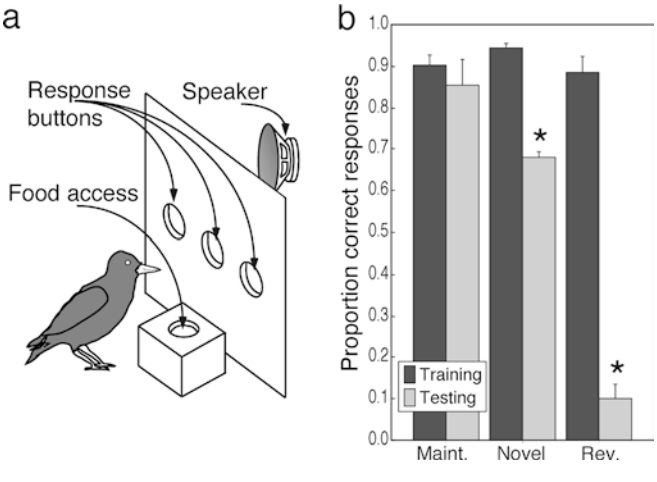



**Fig. 3a, b** Song recognition behavior. **a** Diagram of the operant apparatus used for recognition training and testing. Subjects learned to work the buttons on the panel in response to different song stimuli to obtain food reward. **b** Mean ($\pm$SEM) song recognition performance during the last three blocks (100 trials/block) of baseline training, and during the testing session (60 trials) for each group of subjects. *Significant difference between training and testing ($P < 0.0003$, paired $t$-test)

## Testing

Each subject was exposed to one of four test conditions. In three of the four test conditions, subjects worked the apparatus in response to song stimuli for 30 min (60 trials), with trials presented at the same constant rate as they had been during the pretest training. Subjects in the first group, the *maintenance* test condition ($n = 4$), ran 60 trials with exactly the same song stimuli and response contingencies used during their baseline and pretest training. Subjects in the second group, the *reversal* test condition ($n = 5$), ran 60 trials with the same song stimuli as in the baseline and pretest training, but with the response pairings reversed. That is, songs that had originally been associated with a response to the left button were now associated with a response to the right button, and vice versa. Subjects in the third group, the *novel stimulus* test condition ($n = 5$), ran 60 trials with a

set of unfamiliar conspecific song stimuli (described above). Responses during the novel test condition were reinforced, but the subjects began the session without prior knowledge of the stimulus-response associations, and thus had to learn (as best as possible during the session) both the new stimuli and their associations. Subjects in the 'no-stimulus' (i.e., control) condition ($n = 3$) worked the operant apparatus for food on a 30-s fixed-interval schedule during the first 30 min of the 60-min session, but heard no song stimuli. To increase appetitive motivation and maximize the likelihood that subjects would respond to most of the test trials, pretest training on the day prior to testing was halted at 12 noon. In the event that an animal responded to fewer than half the trials during the testing phase, that session was aborted and rescheduled for a later day. In practice, this rescheduling was only necessary twice. The mean ($\pm$SEM) number of no-response trials during the test session was $8.1 \pm 0.6$. Immediately following the completion of the 60-min test session the animal was sacrificed and the brain tissue prepared for *zenk* protein (ZENK) immunocytochemistry (see below). All test sessions began between 8:00 and 11:00 a.m.

### Histology

Sixty minutes after the start of the test session, and 30 min after the completion of the last trial, the subject was administered a lethal dose (80 mg $kg^{-1}$) of secobarbitol (50 mg $ml^{-1}$ i.m.). Once deeply anesthetized we perfused the brain tissue, via the carotid artery, with approximately 50 ml of heparinized (150 IU/10 ml) 0.9% saline followed by 300–400 ml of 4% paraformaldehyde in a 0.1 mol $l^{-1}$ phosphate buffer (pH 8.5). The brain was then extracted, post-fixed in 4% paraformaldehyde for 24 h, cryoprotected in 30% sucrose solution (in 0.1 mol $l^{-1}$ PBS (pH 7.5) for 2–4 days, and then stored at $-70$°C until immunocytochemistry (ICC) for *Egr-1* protein (ZENK) was performed. Forty-μm sagittal brain tissue sections were collected and processed as free-floating tissue, using a previously published protocol (Gentner et al. 2001). We incubated the tissue in a commercially available polyclonal *Egr-1* antibody (Santa Cruz Biotechnology, catalog no. sc-189; see Mello and Ribeiro 1998) at 1:10,000 dilution in 0.1 mol $l^{-1}$ PBS with 0.3% Triton X-100 (Sigma). Reaction time for DAB visualization was held constant across all the tissue in a given ICC run. The distributions of subjects from each group were balanced within and across ICC runs. No systematic variation across the separate ICC runs was observed.

### Data analyses

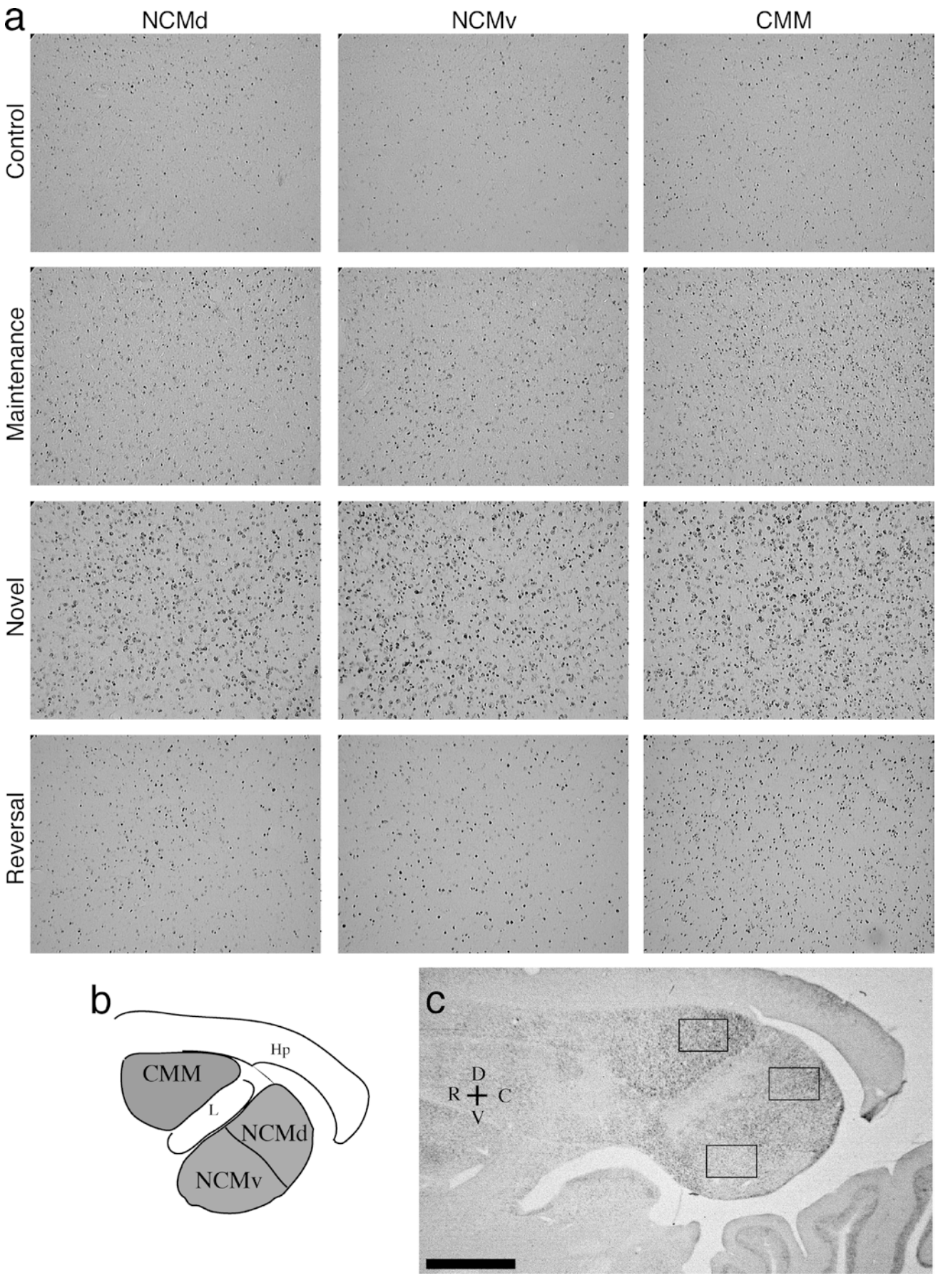
We quantified the level of ZENK expression in three regions of the auditory forebrain: the CMM, the dorsal portion of the caudo-medial nidopallium (NCMd) and the ventral portion of the NCM (NCMv). For each subject we took 96 samples, 16 samples per hemisphere per region, within 700 μm of the midline. Each sample was a 640×480 pixel gray-scale image of the tissue in each region at 200× magnification (992×770 μm actual size; Fig. 4a). We counted the number of ZENK protein positive cells in each sample using an automated image analysis routine that normalized for background optical density variation (see Gentner et al. 2001 for complete description). The sample windows in the CMM were bounded dorso-caudally by the lateral ventricle and the lamina hyperstriatica (LH; Fig. 4b). The rostral border of NCM was defined in medial sections by the LH and in lateral sections by the band of low ZENK expression characteristic of field L2a (Mello et al. 1992). We arbitrarily divided the NCM into dorsal and ventral halves, and took one sample from each (Fig. 4c). All tissue sampling was done blind to the independent variables.

Densities of ZENK positive cells (cells $mm^{-2}$) were natural log transformed prior to statistical analysis to correct any deviation from normality in the underlying sample distributions. Accept where noted we used a mixed ANOVA model ($\alpha = 0.05$) to analyze the data with brain region, hemisphere and tissue section as nested within-subjects variables (repeated measures), and test condition as a single between-subject (factorial) variable. Planned comparisons using separate ANOVAs were conducted on each region; and where appropriate, post-hoc differences between group means were compared using the Student-Newman-Kuels (SNK) test with the harmonic means. Analyses identical to those reported here, but using the raw cell counts, yielded qualitatively similar results. For ease of interpretation, the data are shown as untransformed expression densities.

## Results

All subjects learned to discriminate accurately among the song stimulus exemplars in their respective training sets, and performed at roughly equivalent levels prior to the testing session (Fig. 3b). There were no significant differences between groups in their acquisition rates over the course of baseline training [$F_{(58,290)} = 0.58$, $P = 0.99$, repeated measures ANOVA group×training interaction]. Likewise, there were no significant differences between any of the groups in the accuracy of song recognition immediately prior to testing [$F_{(3, 11)} = 1.01$, $P = 0.43$, one-way ANOVA; $P > 0.14$; Fig. 3b].

Each group of subjects handled the song stimuli presented during the test session very differently (Fig. 3b). As expected, subjects in the 'maintenance' group showed accurate song recognition during testing. The mean ($\pm$SEM) percentage correct for the maintenance subjects ($0.856 \pm 0.061$) was significantly above


a
NCMd
NCMv
CMM
Control
Maintenance
Novel
Reversal
b
Hp
CMM
L
NCMd
NCMv
c
D
R
C
V

◄

**Fig. 4a–c** The ZENK protein expression. **a** Sample images (992×770 μm) showing immunohistochemistry visualized ZENK protein in dorsal NCM, ventral NCM and CMM for all four groups of subjects. **b** Schematic of the sagittal tissue section shown in (**c**) with the auditory regions of interest *shaded grey* (see Fig. 2). **c** Low power image showing ZENK protein expression in the sagittal plane approximately 300 μm from the midline. Approximate positions of the sampling windows within each region are depicted by the *rectangles*. Scale bar = 1 mm

chance ($P < 0.05$, *t*-test, two-tailed), and there was no significant difference in mean percent correct scores between the pretest and test conditions ($p > > 0.05$, paired *t*-test). In contrast, subjects in the 'reversal' test condition, for whom the training songs were presented with reversed response contingencies, showed a significant drop in performance from pretest levels ($P < 0.0003$, paired *t*-test) with the mean proportion of correct responses ($0.10 \pm 0.03$) falling significantly *below* chance ($P < 0.0005$, *t*-test; Fig. 3b). Subjects in the 'novel' test condition also showed a significant drop in their performance between pre-test and test conditions ($P < 0.0001$, paired *t*-test). For the novel group, however, the mean proportion of correct responses during the test session ($0.68 \pm 0.02$) was significantly *above* chance ($P < 0.0005$, *t*-test), indicating rapid (though not complete) learning of the novel stimuli and the correct response associated with each.

The behavioral results establish four classes of subjects that differ in their perceptual relation to the song stimuli: (1) those maintaining accurate recognition of familiar songs (i.e., the 'maintenance' group); (2) those discriminating among familiar songs and relearning associative contingencies ('reversal' group); (3) those learning new song discriminations and new associations (the 'novel' group); (4) and those working the operant apparatus for food without hearing songs (the 'silent control' group).

In all animals that heard song during testing, we observed robust staining of ZENK protein positive cells throughout the CMM and dorsal and ventral regions of the NCM (see Fig. 4a, c). There was significant variation in the mean expression densities observed between the three brain regions [$F_{(2,\ 26)} = 17.110$, $P < 0.0001$] with CMM showing significantly higher expression levels than NCMd and NCMv (Fishers PLSD, $P < 0.0001$, both comparisons). Expression levels in the NCMd and NCMv showed no statistically significant differences between these sub regions. Subsequent analyses used pooled NCM data.

The behavioral requirements associated with the different test conditions had significant effects on the overall levels of ZENK protein expression [$F_{(3,\ 13)} = 8.999$, $P < 0.005$; main effect of group], and these effects varied significantly by brain region [$F_{(6,\ 26)} = 3.061$, $P < 0.05$; interaction between brain region and test group]. To examine the relationship between expression density and recognition behavior more closely, we conducted separate analyses on the data from CMM, and the pooled data from NCMv and NCMd. ZENK protein expression densities in NCM varied significantly among the different test groups [$F_{(3,\ 13)} = 9.668$, $P < 0.005$]. The mean expression density for the subjects tested under the novel song condition ($5{,}583 \pm 784$ cells mm$^{-2}$) was significantly greater than that for all other groups ($P < 0.05$ for all comparisons to 'novel', SNK; Fig. 5b, c). The mean NCM expression densities (in cells mm$^{-2}$) for subjects in the maintenance test condition ($2{,}606 \pm 103$) and the reversal conditions ($2{,}613 \pm 406$) did not differ significantly from the mean expression levels in the silent controls ($1{,}872 \pm 374$), or from one another (SNK; $P \geq 0.5$ all comparisons; see Fig. 5b, c).

In CMM, ZENK protein expression densities also varied significantly among the different groups ($F_{(3,13)} = 7.133$, $P < 0.005$). In contrast to NCM, however, mean expression densities (cells mm$^{-2}$) for subjects in all test conditions (maintenance: $3{,}937 \pm 643$; reversal: $4{,}893 \pm 1{,}033$; novel: $7{,}579 \pm 1{,}339$) were significantly elevated above mean levels in silent controls ($1{,}753 \pm 309$; SNK, $P < 0.05$ for comparisons between control and all other groups; Fig. 5a, c).

We also observed a difference in the expression levels between right and left hemispheres that varied between CMM and NCM. In CMM, mean ZENK protein expression densities (cells mm$^{-2}$) in the right hemisphere ($5{,}300 \pm 200$) were significantly higher than in the left hemisphere ($4{,}517 \pm 182$; $F_{(1,\ 13)} = 7.063$, $p < 0.05$; data not shown). In NCM, there were no significant differences between mean ZENK expression densities in the right ($3{,}473 \pm 480$) and left hemispheres ($3{,}236 \pm 442$). It is not immediately clear how to interpret these hemispheric differences, but they may reflect some degree of lateralization in the perceptual mechanisms attributable to these two regions.

## Discussion

Stimulus elicited IEG responses in the avian auditory telencephalon and their 'tuning' to conspecific song is well established. The strength of the IEG response is modulated by stimulus acoustics (Mello et al. 1992; Ribeiro et al. 1998), behavioral experience (Mello et al. 1995; Jarvis et al. 1995; Sockman et al. 2002), the behavioral relevance of specific songs (Gentner et al. 2001; Maney et al. 2003), and each bird's song learning history (Bolhuis et al. 2000; Terpstra et al. 2004). The present study examines IEG expression during song recognition, a natural song-based behavior, in which song acoustics and their associated behavioral contingencies are varied. The results suggest that there are significant differences between the functional processes mediated by activity in CMM and NCM during song recognition.

The ZENK protein expression in NCM was driven primarily by variability in conspecific song acoustics (i.e.

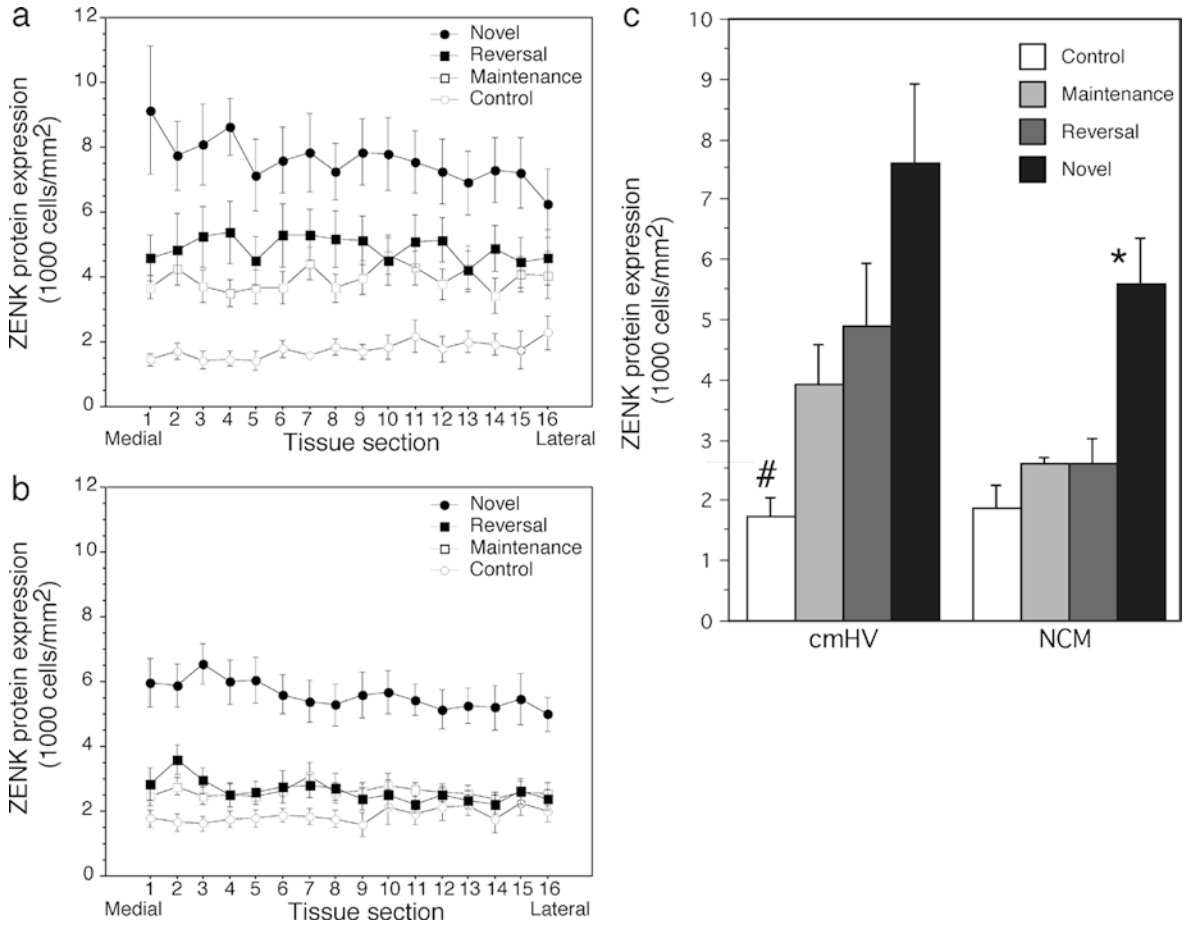



**Fig. 5a, b** Expression quantification results. **a** Mean (±SEM) ZENK protein expression density in CMM, in the 16 sections closest to the midline. The section labeled '1' is the most medial. Plotted data are pooled across hemisphere and across all subjects in each group. **b** As in **a** but for the expression densities for NCMd and NCMv pooled. **c** Mean (±SEM) ZENK protein expression densities by region and test condition. *Significantly greater than all other test conditions in that region ($P < 0.05$, Fisher's PLSD); #significantly less than all other test conditions in that region ($P < 0.05$, Fisher's PLSD)

by song novelty), whereas expression in CMM was driven by both song novelty and the ongoing maintenance of existing representations. The density of ZENK protein expression in NCM was elevated *only* in the birds required to classify the novel songs. NCM expression in the birds recognizing the baseline songs (maintenance group) or the response-reversed songs (reversal group) was not reliably different from the control birds that heard no songs (Fig. 5b, c). The pattern of expression in CMM, however, was much different, with all three groups trained to recognize songs showing significantly elevated ZENK protein expression relative to the silent controls, and birds in the novel song condition showing still higher levels relative to the reversal and maintenance birds (Fig. 5a, c). Thus, in addition to the effect of song novelty, ZENK protein expression in CMM was coincident with ongoing recognition of familiar songs.

The ZENK response in NCM was similar to the habituation/dishabituation (i.e. novelty) responses first described for this region by Mello et al. (1995), and thus may be explained by non-associative mechanisms of stimulus-specific habituation. The ZENK response in CMM appears to be more complex, however, in that it is sensitive to song novelty and *resistant* to habituation for familiar song stimuli. The pattern of ZENK expression in CMM during familiar song recognition may result from a heightened sensitivity to particular stimulus acoustics, or may represent the presence of an associative signal paired with each familiar song (or both). It is difficult to isolate the separate contributions of these two processing mechanisms. Changing the associative pairings of the songs in the reversal condition yielded no measurable increase in expression levels compared to the maintenance of baseline recognition, suggesting that the response is driven entirely by stimulus acoustics. However, complete removal of the associative context during song presentation leads to simple habituation (Mello et al. 1995), suggesting an effect on expression due to associative processes, even if it is not further increased during response reversal. Similar compound effects of acoustic and associative signals may be present in the CMM (and perhaps the NCM) response during novel song recognition, where both the acoustics and the explicit associative context of the song are unknown to

the animal. It is tempting to speculate that the pairing of the initial song presentation with an appropriate reward signal (and similar pairings thereafter) help maintain the CMM response over time. In any case, the data suggest that the functional processing of songs in NCM and CMM is not homogeneous. Whereas the ZENK response in NCM may be fully described by non-associative stimulus response mechanisms, the response in CMM resists this same description.

It is important point out that the qualitative appearance of the observed differences between CMM and NCM may not correspond explicitly to the anatomical borders of these two regions. Although the power of our analysis was high ($\beta = 0.912$, ANOVA interaction between regions), the NCM data do trend in the same direction as those in CMM. Thus, the proposed shift from non-associative to associative processing mechanisms between NCM and CMM, may represent strong quantitative rather than qualitative differences in how the relevant aspects of song stimuli are coded in these regions.

The presence of an associative component in CMM activity during learned vocal recognition is consistent with electrophysiological data showing CMM tuning to acoustic features in songs that birds have learned to recognize. Interestingly, variation in the CMM neural spiking activity tied to the learned associative properties of specific songs is apparent in the magnitude of each cell's response, but not in the presence or absence of a response altogether (Gentner and Margoliash 2003a). The technique used to quantify IEG expression in the present study counts a cell as ZENK-positive if the expression density exceeded a threshold value (see Methods). Because our method does not provide a quantitative measure of protein density within each cell, it may underestimate the contribution of associative signals.

It stands to reason that associative processes play a direct role in modulating the valence of most behaviorally relevant sensory representations. Stimulus relevance is not an a priori property of most sensory objects, particularly for signals whose adaptive value is not fixed. The manner in which relevance information is introduced into high-level sensory representations, however, is not understood. The interplay of non-associative and associative learning mechanisms that we propose occurs between NCM and CMM may represent a form of "relevance-filtering" along the stimulus processing hierarchy. In this way, information about the associative value of a particular stimulus could help stabilize the experience-dependant representation of the appropriate stimulus acoustics over time. In this scenario, novel information (i.e. song acoustics) may be admitted into both CMM and NCM at all times, but only held or consolidated in CMM if coupled with an associative signal. The source of such an associative signal need not be co-localized with the acoustic representation structures. Thus, the anatomical separation of regions driven by non-associative and associative learning mechanisms, posited here for NCM and CMM, respectively, is theoretically tenable.

The hypothesized roles of NCM and CMM in the selection, representation, and consolidation of song memories, and their ZENK protein expression dovetail nicely with prevailing notions of *zenk* function. IEG *zenk* expression has proven very useful in localizing some of the songbird auditory structures important in processing conspecific vocalizations. However, emerging knowledge of *zenk*'s role in cellular signaling pathways (vis-a-vis its mammalian homologues) suggests a more explicit functional interpretation. Current models of the cellular basis for memory consolidation and retrieval involve a synapse-to-nucleus signaling pathway in which gene transcription is required to build and maintain synaptically coupled networks. IEGs such as *zenk* are thought to play a key role in this pathway. In mammals, *zif-268* (the homologue to *zenk*) has been implicated in synaptic plasticity and cellular mechanisms of memory consolidation. Induction of LTP (Cole et al. 1989) and experience in other learning paradigms (Guzowski et al. 2001) are associated with increased expression of the *zif-268* in the hippocampus. Inactivation of *zif-268* prevents certain forms of synaptic plasticity (i.e. LTP) in the dentate gyrus and long-term memory consolidation (Jones et al. 2001), and the reconsolidation of recognition memories following retrieval (Bozon et al. 2003). These data strongly suggest that *zenk*-mediated transcriptional regulation (along with that of other IEGs) is a mechanism for memory storage and retrieval. If true, then the expression patterns observed here point to differences in the memory storage and retrieval functions served by CMM and NCM during learned song recognition.

**Acknowledgements** This research was supported by the NIH R0135467 to G.F.B., and DC00389 to T.Q.G. All experiments reported here comply with the *Principles of animal care*, publication No. 86-23, revised 1985 of the National Institutes of Health, and with the current applicable laws of the United States of America.

## References

Adret-Hausberger M, Jenkins PF (1988) Complex organization of the warbling song in starlings. Behaviour 107:138–156

Bolhuis JJ, Zijlstra GG, den Boer-Visser AM, Van Der Zee EA (2000) Localized neuronal activation in the zebra finch brain is related to the strength of song learning. Proc Natl Acad Sci U S A 97:2282–2285

Bonke D, Scheich H, Langner G (1979) Responsiveness of units in the auditory neostriatum of the Guinea fowl (*Numida meleagris*) to species-specific calls and synthetic stimuli. I. Tonotopy and functional zones. J Comp Physiol A 132:243–255

Bozon B, Davis S, Laroche S (2003) A requirement for the immediate early gene zif268 in reconsolidation of recognition memory after retrieval. Neuron 40:695–701

Chaiken M, Böhner J, Marler P (1993) Song acquisition in European starlings, *Sturnus vulgaris*: a comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. Anim Behav 46:1079–1090

Chew SJ, Mello CV, Nottebohm F, Jarvis E, Vicario DS (1995) Decrements in auditory responses to a repeated conspecific song are long lasting and require two periods of protein synthesis in the songbird forebrain. Proc Natl Acad Sci U S A 92:3406–3410

Cole AJ, Saffen DW, Baraban JM, Worley PF (1989) Rapid increase of an immediate early gene messenger RNA in hippocampal neurons by synaptic NMDA receptor activation. Nature 340:474–476

Eens M, Pinxten M, Verheyen RF (1989) Temporal and sequential organization of song bouts in the European starling. Ardea 77:75–86

Eens M, Pinxten M, Verheyen RF (1991) Organization of song in the European starling: species-specificity and individual differences. Belg J Zool 121:257–278

Falls (1982) Individual recognition by sounds in birds. In: Kroodsma DE, Miller EH (eds) Acoustic communication in birds, vol 2. Academic Press, New York, pp 237–278

Gentner TQ, Hulse SH (1998) Perceptual mechanisms for individual recognition in European starlings (*Sturnus vulgaris*). Anim Behav 56:579–594

Gentner TQ, Hulse SH (2000) Perceptual classification based on the component structure of song in European starlings. J Acoust Soc Am 107:3369–3381

Gentner TQ, Margoliash D (2003a) Neuronal populations and single cells representing learned auditory objects. Nature 424:669–674

Gentner TQ, Margoliash D (2003b) Song recognition neurons in awake birds. Soc Neurosci Abs 33:942.13

Gentner TQ, Hulse SH, Bentley GE, Ball GF (2000) Individual vocal recognition and the effect of partial lesions to HVc on discrimination, learning, and categorization of conspecific song in adult songbirds. J Neurobiol 42:117–133

Gentner TQ, Hulse SH, Duffy D, Ball GF (2001) Response biases in auditory forebrain regions of female songbirds following exposure to sexually relevant variation in male song. J Neurobiol 46:48–58

Grace JA, Amin N, Singh NC, Theunissen FE (2003) Selectivity for conspecific song in the zebra finch auditory forebrain. J Neurophysiol 89:472–487

Guzowski JF, Setlow B, Wagner EK, McGaugh L (2001) Experience-dependent gene expression in the rat hippocampus after spatial learning: a comparison of the immediate-early genes *Arc*, *c-fos*, and *zif268*. J Neurosci 21:5089–5098

Hausberger M (1997) Social influences on song acquisition and sharing in the European starling (*Sturnus vulgaris*). In: Snowden C, Hausberger M (eds) Social influences on vocal development. Cambridge University Press, Cambridge, pp 128–156

Hausberger M, Cousillas H (1995) Categorization in birdsong: from behavioural to neuronal responses. Behav Processes 35:83–91

Hulse SH (1995) The discrimination-transfer procedure for studying auditory perception and perceptual invariance in animals. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC (eds) Methods in comparative psychoacoustics, vol 10. Birkhäuser, Basel, Switzerland, pp 319–330

Jarvis ED, Mello CV, Nottebohm F (1995) Associative learning and stimulus novelty influence the song-induced expression of an immediate early gene in the canary forebrain. Learn Mem 2:62–80

Jones MW, Errington ML, French PJ, Fine A, Bliss TV, Garel S, Charnay P, Bozon B, Laroche S, Davis S (2001) A requirement for the immediate early gene Zif-268 in the expression of late LTP and long-term memories. Nat Neurosci 4:289–296

Leppelsack HJ (1983) Analysis of song in the auditory pathway of songbirds. In: Evert JP, Capranica BR, Ingle DJ (eds) Advances in vertebrate neuroethology. Plenum Press, New York, pp 783–799

Leppelsack HJ, Vogt M (1976) Responses of auditory neurons in forebrain of a songbird to stimulation with species-specific sounds. J Comp Physiol A 107:263–274

Macmillan NA, Creelman CD (1991) Detection theory: a user's guide. Cambridge University Press, Cambridge

Maney DL, MacDougall-Shackleton EA, MacDougall-Shackleton SA, Ball GF, Hahn TP (2003) Immediate early gene response to hearing song correlates with receptive behavior and depends on dialect in a female songbird. J Comp Physiol A 189:667–674

Mello CV, Ribeiro S (1998) ZENK protein regulation by song in the brain of songbirds. J Comp Neurol 393:426–438

Mello CV, Vicario DS, Clayton DF (1992) Song presentation induces gene expression in the songbird forebrain. Proc Natl Acad Sci U S A 89:6818–6822

Mello CV, Nottebohm F, Clayton D (1995) Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. J Neurosci 15:6919–6925

Müller CM, Leppelsack HJ (1985) Feature extraction and tonotopic organization in the avian forebrain. Exp Brain Res 59:587–599

Reiner A et al (2004) Revised nomenclature for avian telencephalon and some related brainstem nuclei. J Comp Neurol 473:377–414

Ribeiro S, Cecchi GA, Magnasco MO, Mello CV (1998) Toward a song code: evidence for a syllabic representation in the canary brain. Neuron 21:359–371

Sen K, Theunissen FE, Doupe AJ (2001) Feature analysis of natural sounds in the songbird auditory forebrain. J Neurophysiol 86:1445–1458

Sockman KW, Gentner TQ, Ball GF (2002) Recent experience modulates forebrain gene-expression in response to mate-choice cues in European starlings. Proc R Soc Lond B Biol Sci 269:2479–2485

Stoddard PK (1996) Vocal recognition of neighbors by territorial passerines. In: Kroodsma DE, Miller EH (eds) Ecology and evolution of acoustic communication in birds. Cornell University Press, Ithaca, pp 356-374

Stripling R, Volman SF, Clayton DF (1997) Response modulation in the zebra finch neostriatum: relationship to nuclear gene expression. J Neurosci 17:3883–3893

Terpstra NJ, Bolhuis JJ, den Boer-Visser AM (2004) An analysis of the neural representation of birdsong memory. J Neurosci 24:4971–4977

Theunissen FE, Doupe AJ (1998) Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches. J Neurosci 18:3786–3802