Articles in PresS. J Neurophysiol (April 5, 2006). doi:10.1152/jn.00067.2006

Title: The Midbrain Periaqueductal Gray and Vocal Patterning in a Teleost Fish.

J. Matthew Kittelberger, Bruce R. Land and Andrew H. Bass

Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853.

Running Head: Midbrain Vocal Patterning

Contact Information: J. Matthew Kittelberger
Dept. of Neurobiology and Behavior
Seeley G. Mudd Hall
Cornell University
Ithaca, NY 14853
email: mk348@cornell.edu

FINAL ACCEPTED VERSION

Abstract

Midbrain structures, including the periaqueductal gray (PAG), are essential nodes in vertebrate motor circuits controlling a broad range of behaviors, from locomotion to complex social behaviors such as vocalization. Few single unit recording studies, so far all in mammals, have investigated the PAG's role in the temporal patterning of these behaviors. Midshipman fish use vocalization to signal social intent in territorial and courtship interactions. Evidence has implicated a region of their midbrain, located in a similar position as the mammalian PAG, in call production. Here, extracellular single unit recordings of PAG neuronal activity were made during forebrain-evoked fictive vocalizations that mimic natural call types and reflect the rhythmic output of a known hindbrain-spinal pattern generator. The activity patterns of vocally active PAG neurons were mostly correlated with features related to fictive call initiation. However, spike trains in a subset of neurons predicted the duration of vocal output. Duration is the primary feature distinguishing call types used in different social contexts, and these cells may play a role in directly establishing this temporal dimension of vocalization. Reversible, lidocaine inactivation experiments demonstrated the necessity of the midshipman PAG for fictive vocalization, while tract-tracing studies revealed the PAG's connectivity to vocal motor centers in the fore- and hindbrain comparable to that in mammals. Together, these data support the hypotheses that the midbrain PAG of teleosts plays an essential role in vocalization and is convergent in both its functional and structural organization to the PAG of mammals.

Introduction

The midbrain periaqueductal gray (PAG) is proposed to play an essential role in the production of vocal communication signals across vertebrates, including humans (Esposito et al. 1999; Jurgens 1994; Kennedy 1975; Seller 1981; Wild 1997). While inputs from forebrain limbic structures, as demonstrated in mammals, support the PAG's role in coupling motivational state to vocal behavior (Bandler and Shipley 1994; Behbehani 1995; Holstege 1998; Jurgens 2002; Sewards and Sewards 2003), few studies show how PAG neuronal activity shapes the temporal parameters of vocal motor output. PAG activity is clearly necessary for call initiation (Behbehani 1995; Jurgens 2002). Single unit recordings in vocalizing macaques show PAG activity correlated with temporal features (e.g., duration) of vocalizations (Larson 1991; Larson and Kistler 1986), implying a more direct role in vocal patterning than classically assumed (Davis et al. 1996; Jordan 1998; Jurgens 2002; Jurgens 1994; Swanson 2000). However, similar recordings in squirrel monkeys do not reveal such correlations (Dusterhoft et al. 2004, 2000). While these findings may represent species differences, the interpretation is currently complicated by an incomplete understanding of the relationship between ensemble muscle activity and the acoustic properties of primate vocalizations. This makes it difficult to establish the PAG's exact role in vocal patterning. We therefore chose to examine this question in a species with a simple vocal repertoire and a more direct translation from neural activity pattern to muscle action and vocal output.

Midshipman fish (*Porichthys notatus)* depend on vocal communication for successful courtship and reproduction. Territorial males use sonic swim bladder muscles (Fig. 1*A*) to produce several call types, differing primarily in duration (Brantley and Bass 1994) (Fig. 1*B*).

The sonic muscles are innervated by a hindbrain-spinal sonic motor nucleus (SMN) that receives input from nearby pacemaker neurons (PN) (Bass and Baker 1990) (Fig. 1*C*). The rhythmic output of the PN-SMN circuit, referred to as fictive vocalization, directly establishes muscle contraction rate and, in turn, call fundamental frequency (Bass and Baker 1990). A ventral medullary nucleus (VM) links the PN-SMN circuitry to a midbrain region similar in location to the mammalian PAG (Fig. 1*C*) (Bass et al. 1994; Goodson and Bass 2002). Electrical stimulation of this midbrain region elicits vocalization (Demski and Gerald 1974, 1972; Fine 1979; Goodson and Bass 2002). The simplicity of vocal motor production, specifically a one-to-one translation from the SMN spike train to acoustic properties of the call, make this an ideal system for identifying the role of PAG neurons in vocal initiation and/or patterning.

Here, we show that the activity patterns of a subset of PAG neurons are correlated with both call initiation and duration. In addition, direct manipulation of PAG activity by electrical microstimulation of descending PAG axons resulted in changes in call duration, further supporting a role for the PAG in temporal patterning. The necessity of PAG activity for vocal initiation is shown by PAG-specific, reversible lidocaine blockade. Focal neurobiotin injections confirm that the vocal motor connectivity of the midshipman PAG compares closely to that in mammals. Thus, two groups of distantly related vertebrates, teleost fish and mammals, share many functional and structural similarities in the descending control of vocalization.

Materials and Methods

Animals

Midshipman fish (*Porichthys notatus)* have two male reproductive morphs that differ in vocal and spawning behaviors (Bass 1996). All experiments were performed on adult type I (territorial) males, which have the most dynamic vocal repertoire (Bass et al. 1999). Fish were collected from tidal pool nesting sites or by offshore trawls in northern California and Washington, shipped to Cornell University, and maintained in artificial seawater tanks at ~15° C. All experimental procedures were approved by the Cornell University Institutional Animal Care and Use Committee.

Surgeries

Surgical procedures were similar to those described previously (Bass and Baker 1990; Goodson and Bass 2000c). Fish were anesthetized by immersion in 0.025 % benzocaine (ethyl *p*-amino benzoate, Sigma, St. Louis, MO). Local anesthetic (0.2 ml of 0.25 % bupivacaine (Abbott Labs, N. Chicago, IL) with 0.01mg/ml epinephrine (International Medication Systems, So. El Monte, CA)) was then injected subdermally to the top of the head. The hypothalamus, midbrain, and hindbrain were exposed by dorsal craniotomy. Before transfer to the experimental apparatus, fish were immobilized with an intramuscular injection of pancuronium bromide (~5 mg/kg, Baxter Healthcare, Deerfield, IL). For all experiments, fish were suspended in a parafilm sling in a plexiglass tank with their head stabilized, and artificial seawater at ~15° C was perfused continuously across their gills. Exposed portions of the brain were kept covered with an inert, electrically conductive fluorocarbon (Fluorinert, 3M, St. Paul, MN). Fish were left to rest for ~1hr before starting electrophysiology experiments to allow all residual benzocaine to wash out of their system.

Extracellular recordings

Fictive vocalizations: The vocal-motor output of the hindbrain vocal pattern generator, the "fictive vocalization", was monitored with an extracellular electrode (75 µm diameter teflon coated silver wire, A-M Systems, with an exposed ball tip of 125-200 µm in diameter) placed on an occipital nerve root that carries the axons of motor neurons from the ipsilateral sonic motor nucleus to the ipsilateral sonic muscle; both sonic motor nuclei fire in phase (Bass and Baker, 1990). These nerve root recordings were amplified 1000x and band-pass filtered from 1 Hz-20 kHz with an A-M Systems differential AC amplifier (Model 1700).

PAG neurons: Extracellular recordings from PAG neurons were obtained using glass microelectrodes (6010, A-M Systems) pulled to a tip resistance of 10-15 MΩ on a Flaming/Brown micropipette puller (Model P-97, Sutter Instruments, Novato, CA) filled with 2M NaCl. The electrode tip solution also contained 5% dextran tetramethylrhodamine (10,000 MW, D-1868, Molecular Probes, Eugene, OR) to label each recording site. Preamplified signals were amplified 1000x total (Model NB-100 amplifier, Biomedical Engineering, Thornwood, NY and an A-M Systems, Model 1700 differential AC amplifier) and band-pass filtered from 0.3-10 kHz.

Forebrain stimulation of PAG neurons and fictive vocalizations

Rationale: Previous studies established the medial longitudinal fasciculus (MLF) of the midbrain as a site where low-amplitude stimulation evokes naturalistic vocalizations (Demski and Gerald 1972; Fine and Perini 1994). Thus, our preliminary experiments focused on the MLF and showed that stimulation here elicited reliable vocal responses that did not fatigue with repeated stimulation. These experiments facilitated the reliable localization and isolation of single PAG neurons with stimulus-modulated activity concurrent with stimulus-evoked fictive vocalizations. However, our ongoing anatomical studies showed the MLF to be the main descending pathway for PAG axons (see last section of Results). Thus, changes in PAG neuron firing evoked by MLF stimulation were likely antidromically mediated, and therefore would not resemble the endogenous pattern of vocal-related PAG activity during spontaneous vocalization. Neuroanatomical and brain stimulation studies suggested that the ventral tuberal hypothalamus (vT) was a likely candidate for a source of descending input to the PAG (Goodson and Bass 2000a, 2002), as we later confirmed in our own anatomical studies (Results). Stimulation at sites in and around vT evoked a less reliable vocal response, at longer and more variable latency, that tended to fatigue with repeated stimulation. Although these features made it more difficult to isolate vocal-related PAG neurons, all of the available data indicated that PAG activity evoked by stimulating vT should be more reflective of endogenous PAG vocal activity patterns. Therefore, for the purposes of this paper, we present data from 48 PAG neurons recorded during vT stimulation. The results of the MLF stimulation experiments were used to analyze how descending midbrain output, inclusive of PAG axons, affects features of the vocal response (see below).

Procedures: Surface landmarks and micromanipulator coordinates were used to guide insulated tungsten stimulating electrodes (125 μm diameter, 8° tip angle, 5 MΩ impedance; A-M Systems, Sequim, WA) to sites in or near either vT or the MLF. Brief trains of stimuli (3-15 pulses, 1msec pulse duration, 333Hz repetition rate, 50-75 μA) were delivered using a WPI stimulus isolation unit (Model 850A, New Haven, CT), with stimulus timing parameters driven by Tucker-Davis System II software and hardware (Alachua, FL). The stimulating electrode was lowered into the brain until a stable reliable vocal output was obtained. The extracellular electrode was then positioned over the PAG and advanced into the brain using a Burleigh microdrive (Model LSS-1000, Burleigh Instruments, Fishers, NY). Once the electrode was at the approximate correct depth, we began searching for units whose activity was clearly modulated by the stimulus. Since the vocal response elicited by vT stimulation tended to fatigue with repeated stimulation, we typically repeated only about 40 stimulus sweeps (1 every 2 s) at a time, followed by a 5-10 min rest interval. This stimulation pattern was followed both while searching for PAG units and while acquiring data once a stable unit was isolated. As PAG neurons tended to have either low or no spontaneous activity (see Results), units were often only detected based on the presence of stimulus-evoked action potentials. Once a stable unit was isolated, spike and vocal output data were recorded while the stimulus duration was varied, resulting in changes in the latency, duration, and probability of the vocal response. Spontaneous spike data were also recorded. After recording, the recording sites were marked by injecting pulsed positive current (+3 μA, 50 % duty cycle, ~10 min) to iontophorese the fluorescent dye out of the electrode tip. In addition, stimulation sites were confirmed after each experiment by the small electrolytic lesion left at each site.

Data Analysis and Statistics of PAG Neuronal Activity

Both PAG and nerve root signals were digitized at 50 kHz, and were acquired using a Tucker-Davis System II data acquisition system with BrainWare v6.3 software (Tucker-Davis). Initial post hoc analyses of the extracellular and vocal nerve recordings were performed with Brainware v6.3 software (Tucker-Davis). The shapes of all action potentials were examined, and multiple units, when present, were defined based on differences in various parameters of spike shape (amplitude of each peak, inter-peak interval). When all spikes were plotted in parameter space, distinct clusters were readily apparent. Visual inspection of voltage versus time plots of each cluster confirmed that all spikes within a cluster were from the same unit (e.g., Fig. 1*D*). The few spikes that did not conform to cluster boundaries were discarded. Of the 143 total neurons recorded, the majority of all recording sites were single units (112 of 127); of the 15 multi-unit sites, 14 revealed two clearly distinguishable units each, and one revealed 3 clear units. The 48 units recorded during vT stimulation represent 42 separate recording sites: 36 sites with only a single unit plus 6 sites each yielding two clearly distinguishable single units. The spike times for each unit, and for the coincident nerve recording of the vocal response pulses, were exported for further analysis.

Custom Matlab scripts (version 7.0.1 (R14), The Math Works) were used to compile, annotate, and analyze the spike records. For each unit confirmed by histology to be in the PAG (see below), we calculated (Table 1): 1) the mean spontaneous firing rate (from trials with no stimulus), 2) the net mean number of stimulus-evoked spikes (over the first 800 msec post-stim, reduced by the number of expected spontaneous spikes), 3) the mode of the inter-spike interval distribution, 4) the latency to the first post-stimulus spike, and 5) the mean lag time between the first post-stimulus unit spike and the first pulse of the vocal response. In addition, for each of these fish, we determined various parameters of the vocal response, including (Table 1): 1) the total mean number of pulses (over the first 800 msec post-stimulus), 2) the mean number of pulses in the first vocal burst (a burst was defined as a series of pulses, each separated from the next by no more than 50 msec), 3) the latency to the first pulse of the vocal response, 4) the mean inter-pulse interval, and 5) the mean inter-burst interval. Post-stimulus time histograms (PSTH) for each unit and vocal response, peri-event time histograms (PETH) for each unit (centered on the time of the first vocal response pulse), and frequency distributions of inter-spike intervals (for each unit and vocal response) were plotted using Origin version 6.1 software (Origin Lab Corp., Northampton, MA).

Statistical analyses, including tests for trial-by-trial correlations between the unit spike activity and the vocal response, were performed using JMP version 5.0.1a software (SAS Institute, Cary, NC). For each unit, we tested for 3 possible correlations: 1) between the net number of stimulus-evoked unit spikes and the total number of vocal response pulses, 2) between the net unit spikes and the latency of the vocal response, and 3) between the latency of the unit response and the latency of the vocal response. The total number of vocal response pulses per trial reflects the product of the probability of a vocal response times the duration (number of pulses) of each vocal burst times the number of vocal bursts. Thus, a correlation between the number of spikes of a particular PAG neuron and the total number of vocal pulses could reflect a correlation with the presence of the vocal response, the number of vocal bursts, and/or the duration of each burst. In order to directly determine whether unit activity predicted burst duration, an additional (fourth) correlation was therefore assessed between the number of vocal pulses in the first vocal burst on each trial and the net number of unit spikes up to the end of this vocal burst. Trials with no vocal response were excluded from this analysis. Note that because the inter-pulse interval of all vocal responses was relatively invariant (see Table 1), the number of vocal pulses is highly predictive of the total time duration of vocalization. We therefore use number of vocal pulses to represent vocal duration throughout this paper.

It seemed likely that variation in the number of stimulus pulses from trial to trial could affect both the unit activity and the vocal response separately. We therefore wanted to be sure that any detected correlations between the unit activity and the vocal response were not being driven by changes in the stimulus duration. Thus, for all four correlational analyses, we ran multi-variate regressions with both stimulus pulse number and unit activity (either latency or net spikes) as independent variables. In addition, we tested for correlations between the vocal response and unit activity on the subset of trials in which stimulus duration was kept constant. Because for some neurons there was little data at particular stimulus durations, the multi-variate regression analysis was used as the criterion for whether or not there was a significant correlation between unit activity and vocal response for each cell. For clarity, however, the illustrative data presented in the figures is only for trials in which stimulus duration was held constant. Finally, for each cell, we tested whether the net number of unit spikes was different on trials on which a vocal response occurred than on trials on which the vocal response failed (while holding stimulus duration constant).

To compare how tightly each unit's spike pattern was timed relative to the stimulus onset versus the vocal onset, we calculated the width of each PSTH and PETH distribution peak at half maximal height. These distributions were noisy: the spikes per time bin often did not decay smoothly away from the peak. We calculated the width at the first time point greater and less than the peak time where the number of spikes per bin fell below half the number of spikes in the peak bin, and again at the last time points (i.e., furthest from the peak) where the spikes per bin fell below this half peak value. The former number will tend to underestimate the half-height width of the distribution, while the latter will overestimate this width. We averaged these minimum and maximum values to arrive at a final estimate of the width of both the PSTH and PETH spike distributions for each unit. A comparison of these two numbers was used to judge whether spike timing for each cell was more tightly locked to the stimulus onset or to the vocal onset.

MLF stimulation of vocal output

We wanted to directly test whether altering the ensemble output of PAG neurons influenced specific parameters of the vocal response, including temporal features such as call duration and inter-pulse interval. To do so, we analyzed the properties of fictive vocal responses elicited by stimulation of the MLF (as described above). All stimulation sites were confirmed histologically to be in the MLF, through which PAG axons descend to connect with the hindbrain vocal circuit. We systematically altered the number of stimulation pulses (stimulus duration), and analyzed four features of the vocal response that might change in correlation with stimulus duration: 1) the duration of the stimulus-evoked vocal burst (number of vocal pulses), 2) the probability of eliciting a vocal response, 3) the latency to the vocal response, and 4) the mean interval between pulses within the vocal response. All 4 features were quantified at each stimulus duration. Using SAS version 9.1.3 statistical software (SAS Institute), we ran repeated measures ANOVAs to ask whether there was a significant effect of stimulus duration on each of these 4 vocal parameters across experiments.

Lidocaine inactivation experiments.

In these experiments, we sought to determine whether reversible inactivation of the PAG affected the stimulus-evoked vocal output. Surgeries, vT stimulation, and vocal nerve recording were as described above. Once a stable and reliable vocal response was obtained, we recorded the baseline response for at least 40 min (16-24 stimulus trials every 10 min). About 6-7 min after the final baseline recording, a glass micropipette containing 4 % lidocaine hydrochloride (Sigma) and 4% fluorescent dextran conjugate (either fluorescein- or tetramethylrhodamine-, Molecular Probes) in 0.1 M phosphate-buffered saline (PBS) was stereotactically guided to PAG. Micropipettes were made as described above for extracellular recording electrodes, but the tips were broken back to ~5-10 μm inner diameter. Either lidocaine or control (vehicle plus dye only) solution was pressure-ejected from the pipette using a picospritzer (Biomedical Engineering) set to deliver 1 pulse per second, 10-50 msec duration each, at 25-30 psi, for 30 s, for a total injection volume of approximately 0.5 nl. We continued to record the stimulus-evoked vocal responses for up to an hour post-injection. PAG injections were bilateral, ipsilateral to the stimulus, or contralateral. In some fish, control injections were targeted to midbrain areas outside the PAG, including the torus and ventral tectum. In some cases post-hoc analysis revealed that injections intended for the PAG had in fact missed; these injections were also treated as controls. Vehicle (PBS) only or sham injections to the PAG were used as additional controls. Typically, 1 experimental and 1 control injection were made in each fish, with at least 60 min between the 2 injections, allowing for complete recovery from any effect of the first injection on the vocal response. The locations of the injection sites were confirmed post hoc (see below).

Post hoc analyses were similar to the extracellular single unit analysis described above. Vocal response pulses were discriminated using Brainware, and pulse times were exported for further analysis. Matlab scripts were used to compile and annotate the data, and to calculate the mean number of vocal response pulses over the first 800 msec post stimulus for each time point pre- and post-injection. Data were graphed in Origin, and statistical analysis performed using JMP. Within each experiment, an injection was determined to have had a significant effect on the vocal response if: 1) there was a significant overall effect of time on the mean vocal response over the duration of the experiment, and 2) a Tukey-Kramer test indicated that the vocal response at a minimum of at least one time point within the first 20 minutes post-injection was significantly less than at a minimum of at least two time points during the pre-injection baseline. To analyze the data by treatment type, we first normalized the post-injection vocal response data to the mean pre-injection baseline response for each experiment, and then pooled this normalized data by treatment type. Within each treatment type, we performed a repeated measures ANOVA across the experimental time points. Because the data were normalized to the pre-injection baselines, a significant effect of time indicates a consistent post-injection effect within the group. In such cases we also determined which post-injection time points were different from pre-injection using a Tukey-Kramer test, with a significance level of 0.05.

Tract tracing experiments.

To characterize the connectivity, both antero- and retrograde, of neurons in the PAG, we made focal iontophoretic injections of 5% neurobiotin (Vector Labs, Burlingame, CA). Surgeries were as described above, except that only a small craniotomy was made over the midbrain on one side. Fish were transferred to the recording rig, and glass micropipettes were stereotactically guided to the PAG. Micropipettes were the same as those used as extracellular recording electrodes. Neurobiotin (in 2M KCl) was iontophoresed for 5-10 min using +3 μA pulsed current (15 s on / 15 s off). Only a single injection was made in each fish. After a 10 h survival time to allow transport of the tracer, fish were perfused and brains processed as described below.

Histology

At the end of all experiments, fish were deeply anesthetized (0.025 % benzocaine) and perfused with ice-cold, teleost Ringer's solution with 10 units/ml of heparin (Elkins-Sinn, Cherry Hill, NJ), followed by 4 % paraformaldehyde in 0.1 M phosphate buffer (PB). Brains were removed, post-fixed overnight, and transferred to 0.1 M PB (pH 7.2) for storage. Brains were equilibrated in 30 % sucrose for 24 h and then sectioned transversely at 50 μm on a freezing microtome. For brains with fluorescent dye injections (fluorescein- or rhodamine-dextran, as in the recording and lidocaine inactivation experiments), sections were collected in 0.1M PB, mounted immediately on gelatin subbed slides, dried overnight, and coverslipped using a fluorescent mounting medium (Vectashield, Vector Labs). Neurobiotin injections, as used in the tract-tracing experiments, were visualized using a standard avidin-biotin-peroxidase protocol. Briefly, sections were collected in PBS, permeabilized for 30 min in PBS with 0.04% Triton-X100 (Sigma), incubated for 3h at room temperature in avidin-biotin-peroxidase in PBS

(Vectastain ABC Elite, Vector Labs), rinsed twice in PB, reacted with 0.05% diaminobenzidine (Sigma) and 0.0072% H_2O_2 in PB, rinsed twice more in PB, and mounted on gelatin subbed slides. Alternate sections were mounted on separate slides, dried overnight, cleared in xylene, and coverslipped. Later, one set of sections was stained with 0.5% cresyl violet (Sigma) to delineate the borders of the different cell clusters. Sections were examined and photomicrographs taken on a Nikon Eclipse E-800 microscope.

Results

Extracellular recording of vocal-related activity in PAG

Extracellular recordings were made from a total of 143 isolated units in the midbrains of 53 type I male midshipman fish, while stimulating different sites to elicit vocal output (see Methods). Of these, a total of 48 neurons were recorded from 23 fish while vocal responses were being elicited by microstimulation of vT, a known vocal structure providing descending afferent input to the PAG (see Methods, and last section of Results). Twenty-four of these 48 neurons, in 12 fish, were confirmed histologically post hoc to be in the PAG (e.g., Fig. 6*C*), a compact cell layer along the periventricular surface of the midbrain (Fig., 6*A, B, K*). Recording sites were often (21/24) localized to the PAG axons just ventral to the main PAG cell body layer (e.g., Fig. 6*C*), but all 24 of these sites had clearly back-filled neurons within the PAG proper. Of the 24 neurons outside the PAG whose activity was also modulated by vT stimulation, 11 were localized to the ventral tectum, 9 were in the midbrain tegmentum, and 4 were rostral of the PAG, in the dorsal thalamus; in none of these recording sites were there any back-filled neurons in the PAG. The mean and median values, including range and standard deviations, for spontaneous and stimulus-evoked firing properties of PAG neurons are summarized in Table 1. All cells fired spontaneous action potentials at either a low rate (up to 8 Hz; 17 cells) or not at all (7 cells). Ventral tuberal stimulation induced an excitatory response in all PAG neurons recorded. From cell to cell, this response ranged from a single, short-latency spike (9 cells), to either a discrete burst of spikes or a broad, longer-lasting increase in firing that decayed back to the spontaneous baseline over several hundred msec. In 18 of the 24 cells, the mean latency to the first post-stimulus spike was < 50 msec.

Stimulus-evoked unit spiking typically began well before the vocal response, and some units continued to fire during the fictive vocalization (Fig. 2*A*1, *B*1). This relationship is best visualized by comparing the post-stimulus time histograms (PSTH) of vocal and unit activity (e.g., Fig. 2*A*2 vs. *A*3, *B*2 vs. *B*3). In 23 of the 24 PAG units, the peak of the unit PSTH preceded the peak of the vocal PSTH, across all trials. In addition, we computed the mean lag time as the difference between the times of the first unit spike and of the first vocal pulse, using only trials where both occurred. The mean lag time for the population of PAG units (–165.6 +/- 110 msec) was significantly less than 0 (p < 0.0001, ANOVA). All but 2 PAG units had mean lag times less than 0, that is, unit leading vocal. However, there was a great deal of variation, both from unit to unit, and from trial to trial within each unit, in the lag time magnitude (see Table 1).

When examining the trial-to-trial raster plots of spike versus vocal activity (e.g., Fig. 2*A*1, *B*1), it was clear that much of the trial-to-trial variance in either the latency or duration of the vocal response, or in whether a vocal response even occurred, could not be explained by variance in either the number or latency of spiking of any individual PAG neuron. Indeed, vT stimulation elicited excitatory responses in PAG neurons on nearly all trials, regardless of whether such stimulation elicited a subsequent vocal response. Therefore, several statistical analyses were used to determine whether the stimulus-evoked firing of PAG neurons was significantly related to the stimulus-evoked vocal response. Neuronal activity was considered to be vocal related if: either: a) spike number (as in Fig. 2*A*4 and *B*4) or latency was significantly different on trials with vocal responses than on trials in which the vocal response failed, with stimulus parameters held constant, or b) there was a significant correlation, across trials, between the number of net unit spikes evoked by the stimulus and either the number of vocal response pulses (e.g., Fig. 2*A*5, *B*5) or the latency of the vocal response (e.g., Fig. 2*B*6) (see Methods for details of statistical tests). By these criteria, 13 of the 24 PAG neurons, including both neurons illustrated in Figure 2, showed a significant relation between their spike activity and the vocal response. There was no clear spatial specificity within the PAG between neurons classified as vocal and those classified as non-vocal using these criteria. Furthermore, none of the basic firing properties of the neurons listed in Table 1 (e.g., spontaneous firing rate) were statistically different between these two populations (Wilcoxon tests).

In 9 of the 13 "vocal" neurons, the net number of stimulus-evoked spikes (i.e., total spike count minus spontaneous activity) was significantly correlated with the total number of vocal response pulses (Table 2, Fig. 2*A*5, *B*5), suggesting that spike number is correlated with either the duration (i.e., number of vocal pulses) of each vocal burst or with the number of vocal bursts. Alternatively, these data could reflect either increased or decreased neuronal firing when a vocal response occurred, with no correlation to the duration of the response itself. This latter possibility would support the hypothesis that PAG neuronal activity contributes to initiation of the vocal response, but has no clear role in the patterning of response duration. To directly test whether spike number in any of the vocal neurons predicted the duration of the subsequent vocal burst, we re-ran the correlation analyses excluding all trials when the vocal response failed, considering only the duration of the first vocal burst on each trial, and only considering spikes occurring before the end of this first vocal burst. In 5 of the 9 neurons that initially showed significant correlations between net spike number and total vocal pulses, significant correlations (3 positive, 2 negative) persisted between net spike number and vocal burst duration (Fig. 2*A*6). In the remaining 4 neurons, this second test revealed no significant relation between spike count and vocal duration, implying that the initial correlation between spike count and total vocal pulses in these 4 neurons was due to an underlying relationship between spike count and the binary presence or absence of a vocal response or between spike count and the number of vocal bursts. Indeed, in 7 of the 13 vocal neurons, including these 4, there was a significant relationship between the net stimulus-evoked spike number and the simple presence or absence of a vocal response (Fig. 2*A*4, 2*B*4), evidence further supporting a role in initiation (but not excluding a role in patterning). Only a minority of the 13 vocal neurons revealed significant correlations either between spike number and vocal latency (3 cells, Fig. 2*B*6), or between mean unit latency and the presence of a vocal response (4 cells). Finally, in none of the 24 confirmed PAG neurons was there any significant relationship between the firing frequency of the unit spike train and the discharge frequency of the vocal response (i.e., rate of fictive sound pulses; see Table 1) or between the latency of the first unit spike and the latency of the vocal response. Histograms of inter-spike interval (ISI) distributions for each neuron never showed a clear peak at or near 10 msec, the characteristic inter-pulse interval of both stimulus-evoked fictive vocalizations (see Vocal Response Data, Table 1) and natural calls (Fig. 1*B*). In summary, approximately half of all histologically confirmed PAG neurons recorded were significantly correlated with the initiation and / or duration of the vocal output.

There were no obvious patterns across neurons in the combination of these correlations that might indicate the presence of distinct sub-populations of vocal PAG neurons. The different categories of correlations described in the preceding paragraph were not mutually exclusive, as the spike activity of some neurons predicted multiple parameters of the vocal response, such as the presence of a vocal response and vocal burst duration (the neuron in Fig. 2*A*), or the presence of a vocal response and the latency of the response (the neuron in Fig. 2*B*) (see also Table 2). On the other hand, overlap between these categories of correlations did not appear to be obligatory. For example, cells with a significant positive correlation between the number of net stimulus-evoked unit spikes and the number of vocal response pulses were neither more nor less likely than other cells to have a significant correlation between the number of unit spikes and the vocal latency. The one possible exception is one of the two cells that showed a significant negative correlation between the number of unit spikes and vocal burst duration. This cell was also one of only two cells with significantly fewer spikes on trials with a vocal response than on trials without. Thus there may be a small subset of PAG neurons with net inhibitory effects on vocal initiation and/or duration.

Another approach to examining the relationship between spike trains and vocal output was to re-plot each unit's PSTH as a peri-event time histogram (PETH) relative to the onset of the vocal response on each trial. By comparing the PSTH and PETH for each unit, it is possible to quantify how tightly each unit's spiking is distributed relative to the stimulus versus to the onset of the vocal response (Fig. 3). We quantified this difference by measuring the width of the peak of each distribution at half of the maximum height (see Methods), and comparing the widths of the two distributions for each unit. Across all PAG neurons the PETH width was, on average, 119 +/- 29 msec (SEM) greater than the PSTH width, indicating that spiking was more tightly time-locked to the stimulus than to the vocal response. However, when cells were sub-divided into "vocal" and "non-vocal" categories based on correlations with the vocal response as described above, the vocal cells had a significantly smaller difference between the PETH and PSTH distribution widths (Fig. 3, top left inset). In 7 of the 13 cells classified as vocal (and only 1 of the non-vocal cells), the PETH width was within +/- 50 msec of the PSTH width. Thus, in these cells, spiking was approximately as well time-locked to the vocal response as to the stimulus itself. While the *difference* between the PETH and PSTH half-height widths was significantly less in vocal cells than in non-vocal, the absolute values of these widths, for either PETH or PSTH distributions were no different between the two groups ($p > 0.2$ and 0.9, respectively, Wilcoxon test). This analysis validates the separation of PAG neurons into vocal and non-vocal subgroups, and shows that the spike trains of vocal neurons were not only significantly correlated with the vocal response, but were also more tightly distributed relative to the onset of the fictive vocalization.

Direct manipulation of PAG output alters fictive vocal initiation and duration.

If PAG activity does contribute to both initiation of the vocal response and establishing vocal duration, one prediction is that increasing or decreasing the duration of PAG ensemble spiking should cause changes in both the probability and duration of the resultant vocal output. Our preliminary recording experiments involved stimulation in the MLF (see Methods), which is the sole pathway by which PAG axons connect to the hindbrain vocal circuit (see below). Thus, MLF stimulation presumably elicits vocal responses via excitation of the descending PAG axons, and changes in the duration of such stimulation will produce longer or shorter trains of action potentials in these descending inputs to the hindbrain. We examined the relationship between stimulus duration and vocal response properties in 11 fish in which the stimulation site was histologically confirmed post hoc to be in the MLF and in which vocal responses had been recorded at multiple stimulus durations. Repeated measures ANOVAs revealed a significant positive relationship across experiments between stimulus duration and both the duration of each vocal burst (Fig. 4*A*) and the probability of a vocal response (Fig. 4*B*). In contrast, stimulus duration did not significantly affect either the latency of the vocal response (Fig. 4*C*), or the interval between vocal pulses within the response (Fig. 4*D*). Thus, not only is PAG neuronal activity correlated with the initiation and duration of the vocal output, but directly altering the output of these neurons also causes changes in these same vocal parameters.

Reversible inactivation of the PAG blocks vocal output

To directly test whether PAG activity was necessary for the stimulus-evoked vocalizations to occur, focal injections of the reversible sodium channel blocker lidocaine were made into the PAG and surrounding midbrain structures after recording baseline vocal responses to vT stimulation. After injection, we continued recording to monitor any effects of the injection and recovery from these effects. Fluorescent dyes included with the lidocaine allowed us to examine post hoc the location of each injection, and to estimate the relative size and spread away from the injection site.

Representative results of two lidocaine inactivation experiments are shown in Figure 5*A-D*. The vocal response was completely suppressed immediately following all lidocaine injections to the PAG ipsilateral to the site of vT stimulation (n=6) and 9 out of 10 lidocaine injections to the PAG bilaterally. In contrast, none of the control injections, including sham (n=3) and vehicle (n=4) injections to ipsilateral PAG, and lidocaine injections to either the lateral midbrain tectum (n=3) or to the underlying torus semicircularis (n=2) resulted in the complete and immediate blockade of the vocal response (see summary in Fig. 5*E*).

Lidocaine effects were reversible: vocal responsiveness returned within 60 minutes after all ipsi- or bilateral PAG injections. Qualitatively, the effects of the various control treatments varied. Lidocaine injections outside of the PAG (including tectal and toral injections) reduced the number of vocal response pulses evoked, but this effect was always delayed by at least 10 min post-injection, and in 2 of these 5 cases, the vocal response was not completely blocked at any post-injection time point. The delayed effect of these injections was consistent with the diffusion of lidocaine from the injection site to the PAG. Indeed, in all but one of these injections, we confirmed the spread of fluorescent dye from the injection site to structures surrounding the ventricle, including the PAG (see example, Fig. 5*B*'). Vehicle injections to the PAG did, in 2 of 4 cases, cause a decrease of > 70%, but not a complete suppression, in the vocal response immediately post injection; in both cases the response returned to baseline levels within 10 minutes. While this shows an apparent influence of the vehicle alone on the results, we emphasize that these results sharply contrast with the lidocaine injections in the ipsilateral PAG that show an immediate and complete cessation of vocal activity (see above and statistical tests below). Sham injections to the PAG never resulted in any detectable change in the vocal response (see Methods for how we defined an effect of treatments within each experiment).

The effects of lidocaine injections to only the PAG contralateral to the stimulation were inconsistent. Of the 4 injections, 3 caused an immediate decrease in the vocal response, with 2 of these completely blocking the vocal response. The latter 2 injection sites both showed clear diffusion of dye across the midline to the ipsilateral PAG, but the immediacy of the effects nonetheless suggests the involvement of the contralateral PAG in vocalizations evoked by unilateral vT stimulation. That the results of these experiments were inconsistent may imply relatively weak contralateral connectivity (see anatomical experiments below for confirmation of bilateral PAG to PAG connectivity). Nevertheless, only lidocaine injections directly in the ipsilateral PAG blocked the vocal response consistently, immediately, and completely, demonstrating that neuronal activity in the PAG, but not in surrounding midbrain structures, is necessary for stimulation-evoked vocal responses.

These results were confirmed by statistical analyses of the different treatment types. Repeated measures ANOVAs confirmed that there was a significant overall effect across post-injection time of both bilateral and ipsilateral lidocaine injections to the PAG ($p < 0.0001$, both groups). Toral and tectal lidocaine injections also had significant effects on the vocal response ($p < 0.0001$ and 0.001, respectively), but the timing of these effects was different. Tukey-Kramer tests indicated that lidocaine injections to the PAG, both ipsi- and bilateral, caused a significant decrease in the vocal response, relative to baseline, at the 0, 10 and 20 min post-injection time points. In contrast, lidocaine injections to the torus and tectum caused decreases at the 10, 20 and 30, but not 0, min time points. This confirms that the effects of lidocaine injections outside the PAG were indeed delayed relative to injections that directly hit the PAG, supporting the hypothesis that activity only in the PAG, and not in neighboring midbrain structures, is essential for vocal initiation. By this analysis, there were no significant overall effects of vehicle or sham injections to the PAG, or of lidocaine injections to the contralateral PAG ($p > 0.6, 0.35, 0.06$ respectively).

Anatomical connectivity of the PAG

Two main types of neurophysiological data support the hypothesis that vT provides descending afferent input to the PAG. First, the latency of vocal responses elicited by stimulation of vT is much longer than responses evoked by MLF stimulation (~200 msec vs. ~15-30 msec). Second, PAG inactivation blocks vocal responses evoked by vT stimulation. In order to more clearly define the connectivity of the PAG, both anterograde and retrograde, we made focal, unilateral injections of neurobiotin targeted to the vocal PAG regions in 4 type I male and in 1 female midshipman (e.g., Fig. 6*A, B, K*). Our two goals were 1) to determine whether there were direct connections from our stimulation site in the anterior hypothalamus (vT) to the PAG and from the PAG to the known vocal motor structures in the hindbrain, and 2) to be able to compare the connectivity of the presumptive teleost PAG with that of the mammalian PAG. Photomicrographs of label resulting from 2 such injections are shown in Figure 6 (*A, B, K*). We found strong retrograde labeling of cell bodies in vT (Fig. 6*D*, *E*), as well as in other vocally active portions of the hypothalamus (the anterior tuberal nucleus) and preoptic area (the anterior and posterior parvocellular preoptic nuclei; not shown but see Goodson and Bass, 2002). In the hindbrain, we found labeled fibers with swellings indicative of presynaptic boutons in VM (Fig. 6*F*, *G, L, M*) and in an area ventral and medial of VM (Fig. 6*F*, *H*) where commissural axons from VM neurons cross the midline to the contralateral VM (Bass et al. 1994). The projection from the PAG to VM was highly specific, as terminals were not found in surrounding structures.

In no case did we find fibers or terminals in the PN or SMN, the two other components of the hindbrain vocal motor circuit (Fig. 1*C*). Labeled axons terminating in VM traveled caudally from the PAG within the MLF. Other labeled axons crossed the midline at the level of the PAG (Fig. 6*I*) and a small number of retrogradely labeled neurons and anterogradely labeled terminals were observed in the contralateral PAG (Fig. 6*J*). This reciprocal connectivity provides an anatomical basis for the immediate effects on the vocal response observed after some contralateral PAG lidocaine injections (see previous section of Results). In sum, these experiments confirm a descending pathway from the anterior hypothalamus (and, more specifically, from vT) to the PAG and then to VM, but neither to the motoneurons innervating the sonic muscles nor to the premotor pacemaker neurons (see Fig. 1*C* for overview).

Discussion

The data presented here demonstrate for the first time in a teleost fish: (1) the activity patterns of single midbrain PAG neurons that predict the initiation and duration of fictive vocalizations, (2) the necessity of the PAG for vocal production, and (3) the specific anatomical pathway by which descending vocal motor information from the forebrain flows through the PAG to the hindbrain-spinal vocal circuit that determines the fundamental frequency of calls (Bass and Baker, 1990).

Single neurons in the midshipman PAG, excited by forebrain hypothalamic (vT) stimulation, showed neuronal responses preceding the evoked fictive vocal response, with significant correlations between spike number and the presence, duration, and/or latency of fictive calls. Such “vocal” neurons were more tightly timed with regard to the vocal onset than were other "non-vocal" PAG neurons. However, the frequency of neuronal firing did not correlate with the discharge frequency of the fictive call. This is consistent with other studies showing that this trait, which establishes the fundamental frequency of natural calls, is primarily, if not solely, determined by the hindbrain-spinal vocal circuitry (Bass and Baker, 1990). Stimulation of the MLF, the sole descending pathway for PAG axons connecting to the hindbrain vocal circuitry, affected the probability and duration of vocal bursts, but not their latency or discharge frequency. These data strongly support the hypothesis that the correlations between individual PAG neuron spike activity and vocal output are relevant for vocal initiation and duration patterning, but not for establishing the fine temporal structure (in this case, fundamental frequency) of a vocalization. Consistent with these findings, reversible lidocaine inactivation experiments of the PAG showed that its activity is necessary for the production of vocal responses that are elicited by forebrain stimulation. Lastly, the above neurophysiological experiments, together with the demonstration of a direct anatomical pathway from vT to the PAG, and then from the PAG to VM, an integral component of the hindbrain vocal motor complex (Fig. 1*C*), provide strong support for the general hypothesis that the vT-PAG-VM circuit is the major descending pathway for vocal motor commands originating in the forebrain.

Comparisons with mammals

These findings also provide critical empirical support for proposals that this portion of the teleost brain is similar both structurally and functionally to the mammalian, and more generally tetrapod, PAG (Goodson and Bass 2002). Neurons in the vocal portion of the mammalian PAG do not project directly to the motoneuron pools involved in vocalization, but rather to the nucleus retroambiguus (NRA), a primary premotor structure in the caudal medulla (Holstege 1989; Jurgens 2002; Vanderhorst et al. 2000). The NRA then connects to the various motoneuron pools innervating the vocal musculature (e.g., muscles of the larynx, pharynx, tongue, jaw; see Jurgens, 2002). Similarly, we find that the teleost PAG projects to VM, which in turn connects to the PN–SMN circuit that includes sonic motoneurons. As in mammals, we find no evidence of direct connections from the PAG to vocal motoneurons. The mammalian PAG receives dense inputs from a variety of structures in the limbic system, including hypothalamic and preoptic nuclei (Dujardin and Jurgens 2005; Jurgens 2002), inputs which are also present in midshipman (Results and Goodson and Bass, 2002).

An important distinction between our results and those in mammals is that the mammalian PAG appears to be topographically organized. The vocal region itself occupies a discrete portion of the PAG (Holstege 1989; Larson and Kistler 1986; Vanderhorst et al. 2000). Furthermore, there is evidence of topographic mapping of call types across the vocal area of the PAG (Dujardin and Jurgens 2005; Zhang et al. 1994). In contrast, we found no obvious spatial clustering of vocal neurons within the teleost PAG.

Functionally, either lesions or pharmacological inactivation of the mammalian PAG blocks both spontaneous vocalizations as well as those elicited by stimulation of PAG afferents (Jurgens 1994; Siebert and Jurgens 2003), just as lidocaine inactivation of the teleost PAG prevented vocal responses to vT stimulation. Single unit PAG recordings in bats (Suga and Yajima 1988), squirrel monkeys (Dusterhoft et al. 2004, 2000), and macaques (Larson 1991; Larson and Kistler 1986) demonstrate that neuronal activity correlates with vocal output, as we describe here. Generally, PAG neurons in midshipman show similar patterns of vocal-related activity to those described in these other systems. Specifically, both here (see above) and in mammals: 1) Only a small subset of PAG neurons have vocal-related activity. 2) Vocal neurons usually begin firing before vocal onset. 3) Vocal responses are heterogeneous, sometimes ceasing to fire prior to vocal onset and sometimes continuing to fire during vocalization. Furthermore, different neurons are correlated with different aspects of the vocal output. 4) From trial to trial, there is significant variability in the lag time between the onset of neuronal spiking and the vocal onset. Thus, the variability in vocal-related spike activity we found in neurons in the midshipman PAG is, in fact, typical of that described previously in the mammalian PAG (Larson 1991; Larson and Kistler 1986). One minor difference between our results and those in mammals are that we found a small subset (2/24) of "vocal" PAG neurons with an apparent net inhibitory relationship to the vocal response. Such neurons have not been described in the mammalian PAG (Dusterhoft et al. 2004, 2000; Larson 1991; Larson and Kistler 1986; Suga and Yajima 1988). Despite the presence of these inhibitory neurons, the summed effect of PAG activity is still excitatory vis a vis vocalization, as demonstrated by the fact that PAG inactivation completely blocks vocalization, and that MLF stimulation excites vocal responses. Taken together, the structural and functional similarities between the mammalian and teleost PAG suggest that our findings regarding the role of PAG neurons in the initiation and temporal patterning of vocalization are likely to be generally applicable across sonic vertebrates.

Temporal patterning of vocalization

A role for the PAG in the direct patterning of the temporal and motor structure of vocalization remains unresolved. In macaques, the activity of single PAG neurons is correlated both with the activation of single or functionally-related groups of vocal muscles and with various acoustic properties of vocalizations, including duration and fundamental frequency (Larson 1991; Larson and Kistler 1986). This data has been interpreted to imply a role for the PAG in vocal patterning. However, other data imply that the PAG initiates each call type, but that detailed patterning of the temporal and acoustic features of calls occurs in premotor and motor circuitry of the hindbrain (Jurgens 1994). Consistent with this, the activity of single PAG neurons in squirrel monkeys correlates with call type, but does not predict the gross temporal structure of amplitude and frequency modulations (Dusterhoft et al. 2004, 2000). However, analysis of the temporal details of the unit spike trains in these experiments may yet reveal temporal correlations between unit activity and acoustic features of the vocalizations. Although abundant experimental evidence exists for supra-hindbrain temporal patterning of learned vocalizations in songbirds (Ashmore et al. 2005; Hahnloser et al. 2002; Kittelberger and Mooney 2005; Leonardo and Fee 2005; Vu et al. 1994; Yu and Margoliash 1996), none of these studies address the role of the vocal midbrain.

In midshipman fish, the direct translation of SMN activity to vocal traits simplifies the analysis of whether PAG spike trains predict the fine temporal details of vocalization. Natural calls in midshipman have either a fundamental frequency (for harmonic “hums”) or pulse repetition rate (for nonharmonic “grunts”) close to 100 Hz (Fig. 1*B*). Stimulus evoked fictive vocalizations invariably have a similar frequency (Bass and Baker, 1990). We find no indication that PAG neuronal spike trains establish this characteristic frequency. However, the number of spikes scaled with the duration of vocal bursts in a number of neurons that could participate in establishing call duration, the primary feature distinguishing call types (Fig. 1).

While the activity of single PAG neurons in midshipman was significantly correlated with various parameters of the vocal output, including duration, no single neuron’s activity explained a high percentage of the variance in any one parameter (as indicated by the generally low r-values of the linear regressions such as those shown in Figure 2). This may reflect the fact that initiation and patterning of the vocal output depends on a population of neurons, both within the PAG and in other portions of the vocal circuit (see Fig. 1*C*). Thus, the activity of no single neuron within this population will likely be sufficient to account for the entire variance in any vocal parameter, such as initiation or duration, even when that neuron is functionally involved in shaping that parameter.

In support of the conclusion that PAG neuronal activity shapes vocal duration, changing the duration of electrical stimuli delivered to the MLF consistently altered fictive call duration, with concurrent effects on the probability of eliciting a vocal response, but not on discharge frequency. MLF stimulation likely excites descending axons from mid- and forebrain sources other than the PAG. Thus, taken alone, these experiments only demonstrate that call duration can be influenced by activity manipulation above the level of the hindbrain. However, several pieces of evidence suggest the most parsimonious interpretation is that the effects of MLF stimulation on vocalization are attributable to effects specifically on descending PAG axons. First, injections of biotin tracers into the hindbrain region of VM that innervates the PN-SMN circuit (Fig. 1) mainly label PAG neurons (Goodson and Bass, 2002). The tract tracing experiments reported here confirm the specificity of this projection. Consistent with this circuitry, the current study also demonstrates that PAG activity is predictive of a vocal response, while lidocaine experiments show that activity in the PAG, but not in midbrain structures surrounding the PAG, is necessary for forebrain-evoked fictive vocalizations. Previously published reports also support a role for the midshipman PAG in patterning vocal duration. First, neuropeptide injections near the PAG induce dose-dependent changes in fictive call duration but not in discharge frequency (Goodson and Bass 2000b). Furthermore, certain effects of androgenic steroids on vocal duration have been localized to the midbrain (Remage-Healey and Bass 2004); a particularly high density of androgen receptors in the PAG (Forlano et al. 2005), makes it a likely locus for these effects.

The PAG is a critical site in the initiation of a variety of "emotional" behaviors which involve vocal communication signals, from defense and escape to courtship and reproduction (Behbehani 1995; Holstege 1998). Similarly, the multiple midbrain regions comprising the mesencephalic locomotor center initiate locomotion via projections to reticulospinal neurons driving spinal motoneuron pools (Grillner 2003; Jordan 1998). In all of these contexts, midbrain neuronal activity has been thought to play little role in directly patterning the motor output. Here we show that PAG neuronal activity correlates with discrete aspects of vocal initiation. Furthermore, we demonstrate that PAG activity correlates with the duration of the vocal output, but clearly not with other fine temporal features, such as fundamental frequency. In general, the results of this study provide a compelling example of how distantly related groups of vertebrates have adopted similar mechanisms to solve common problems in vocal-acoustic communication, in this case the patterning of brainstem vocal motor output. To the extent that the specifics of these findings prove to be generally applicable across motor systems and groups of vertebrates, initiation and patterning functions may not be localized in distinct, separate nodes of the descending motor pathway, but rather may be distributed properties of the motor circuit considered in its entirety.

Acknowledgements: We would like to thank Aly Melin and Midge Marchaterre for technical assistance, and B. Arthur, J. Fetcho, B. Johnson, R. Harris-Warrick for helpful comments on the manuscript. James Booth, Professor of Biological Statistics and Computational Biology, provided helpful advice on statistical methods. These studies were funded by grants from the National Science Foundation (IBN 9987341, 0516748 to AHB) and the National Institutes of Health, National Institute for Deafness and Communication Disorders (RO1 DC00092 to AHB, NRSA F32 DC06156 to JMK).

References

Ashmore RC, Wild JM, and Schmidt MF. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J Neurosci* 25: 8543-8534, 2005.

Bandler R and Shipley MT. Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends Neurosci* 17: 379-389., 1994.

Bass AH. Shaping brain sexuality. *Am Scientist* 84: 352-363, 1996.

Bass AH and Baker R. Sexual dimorphisms in the vocal control system of a teleost fish: morphology of physiologically identified neurons. *J Neurobiol* 21: 1155-1168, 1990.

Bass AH, Bodnar D, and Marchaterre MA. Complementary explanations for existing phenotypes in an acoustic communication system. In: *The Design of Animal Communication*, edited by Hauser MD and Konishi M. Cambridge, MA: MIT Press, 1999, p. 493-514.

Bass AH, Marchaterre MA, and Baker R. Vocal-acoustic pathways in a teleost fish. *J Neurosci* 14: 4025-4039, 1994.

Behbehani MM. Functional characteristics of the midbrain periaqueductal gray. *Prog Neurobiol* 46: 575-605., 1995.

Brantley RK and Bass AH. Alternative male spawning tactics and acoustic signals in the plainfin midshipman fish *Porichthys notatus* girard (Teleostei, Batrachoididae). *Ethology* 96: 213-232, 1994.

Davis PJ, Zhang SP, and Bandler R. Midbrain and medullary regulation of respiration and vocalization. *Prog Brain Res* 107: 315-325, 1996.

Demski LS and Gerald JW. Sound production and other behavioral effects of midbrain stimulation in free-swimming toadfish, *Opsanus beta*. *Brain Behav Evol* 9: 41-59, 1974.

Demski LS and Gerald JW. Sound production evoked by electrical stimulation of the brain in toadfish (*Opsanus beta*). *Anim Behav* 20: 507-513, 1972.

Dujardin E and Jurgens U. Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey. *Brain Res* 1034: 114-131, 2005.

Dusterhoft F, Hausler U, and Jurgens U. Neuronal activity in the periaqueductal gray and bordering structures during vocal communication in the squirrel monkey. *Neurosci* 123: 53-60., 2004.

Dusterhoft F, Hausler U, and Jurgens U. On the search for the vocal pattern generator. A single-unit recording study. *NeuroReport* 11: 2031-2034., 2000.

Esposito A, Demeurisse G, Alberti B, and Fabbro F. Complete mutism after midbrain periaqueductal gray lesion. *NeuroReport* 10: 681-685, 1999.

Fine ML. Sounds evoked by brain stimulation in the oyster toadfish Opsanus tau L. *Exp Brain Res* 35: 197-212, 1979.

Fine ML and Perini MA. Sound production evoked by electrical stimulation of the forebrain in the oyster toadfish. *J Comp Phys A* 174: 173-185, 1994.

Forlano PM, Marchaterre MA, Deitcher DL, and Bass AH. Distribution of androgen receptor mRNA in vocal and nonvocal circuitry of a teleost fish. *Soc Neurosci Abstr* 2005: 1001.1006, 2005.

Goodson JL and Bass AH. Forebrain peptides modulate sexually polymorphic vocal circuitry. *Nature* 403: 769-772, 2000a.

Goodson JL and Bass AH. Rhythmic midbrain-evoked vocalization is inhibited by vasoactive intestinal polypeptide in the teleost *Porichthys notatus. Brain Res* 865: 107-111, 2000b.

Goodson JL and Bass AH. Vasotocin innervation and modulation of vocal-acoustic circuitry in the teleost *Porichthys notatus. J Comp Neurol* 422: 363-379, 2000c.

Goodson JL and Bass AH. Vocal-acoustic circuitry and descending vocal pathways in teleost fish: convergence with terrestrial vertebrates reveals conserved traits. *J Comp Neurol* 448: 298-322, 2002.

Grillner S. The motor infrastructure: from ion channels to neuronal networks. *Nat Rev Neurosci* 4: 573-586., 2003.

Hahnloser RHR, Kozhevnikov A, and Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419: 65-70, 2002.

Holstege G. Anatomical study of the final common pathway for vocalization in the cat. *J Comp Neurol* 284: 242-252, 1989.

Holstege G. The emotional motor system in relation to the supraspinal control of micturition and mating behavior. *Behav Brain Res* 92: 103-109, 1998.

Jordan LM. Initiation of locomotion in mammals. *Annals NY Acad Sci* 860: 83-93, 1998.

Jurgens U. Neural pathways underlying vocal control. *Neurosci Biobehav Rev* 26: 235-258., 2002.

Jurgens U. The role of the periaqueductal grey in vocal behaviour. *Behav Brain Res* 62: 107-117, 1994.

Kennedy MC. Vocalization elicited by electrical stimulation of the midbrain. *Brain Res* 91: 321-325, 1975.

Kittelberger JM and Mooney R. Acute injections of brain-derived neurotrophic factor in a vocal premotor nucleus reversibly disrupt adult birdsong stability and trigger syllable deletion. *J Neurobiol* 62: 406-424, 2005.

Larson CR. On the relation of PAG neurons to laryngeal and respiratory muscles during vocalization in the monkey. *Brain Res* 552: 77-86, 1991.

Larson CR and Kistler MK. The relationship of periaqueductal gray neurons to vocalization and laryngeal EMG in the behaving monkey. *Exp Brain Res* 63: 596-606, 1986.

Leonardo A and Fee MS. Ensemble coding of vocal control in birdsong. *J Neurosci* 25: 652-661, 2005.

Remage-Healey L and Bass AH. Rapid, hierarchical modulation of vocal patterning by steroid hormones. *J Neurosci* 24: 5892-5900, 2004.

Seller TJ. Midbrain vocalization centres in birds. *Trends Neurosci* 4: 301-303, 1981.

Sewards TV and Sewards MA. Representations of motivational drives in mesial cortex, medial thalamus, hypothalamus and midbrain. *Brain Res Bull* 61: 25-49., 2003.

Siebert S and Jurgens U. Vocalization after periaqueductal grey inactivation with the GABA agonist muscimol in the squirrel monkey. *Neurosci Letters* 340: 111-114, 2003.

Suga N and Yajima Y. Auditory-vocal integration in the midbrain of the mustached bat: periaqueductal gray and reticular formation. In: *The physiological control of mammalian vocalization.*, edited by Newman JD: Plenum, 1988, p. 87-107.

Swanson LW. Cerebral hemisphere regulation of motivated behavior. *Brain Res* 886: 113-164, 2000.

Vanderhorst VGJM, Terasawa E, Ralston HJ, and Holstege G. Monosynaptic projections from the lateral periaqueductal gray to the nucleus retroambiguus in the rhesus monkey: implications for vocalization and reproductive behavior. *J Comp Neurol* 424: 251-268, 2000.

Vu ET, Mazurek ME, and Kuo YC. Identification of a forebrain motor programming network for the learned song of zebra finches. *J Neurosci* 13: 6924-6934, 1994.

Wild JM. Functional anatomy of neural pathways contributing to the control of song production in birds. *Eur J Morph* 35: 303-325, 1997.

Yu AC and Margoliash D. Temporal hierarchical control of singing in birds. *Science* 273: 1871-1875, 1996.

Zhang SP, Davis PJ, Bandler R, and Carrive P. Brain stem integration of vocalization: role of the midbrain periaqueductal gray. *J Neurophys* 72: 1337-1356, 1994.

Figures and Legends

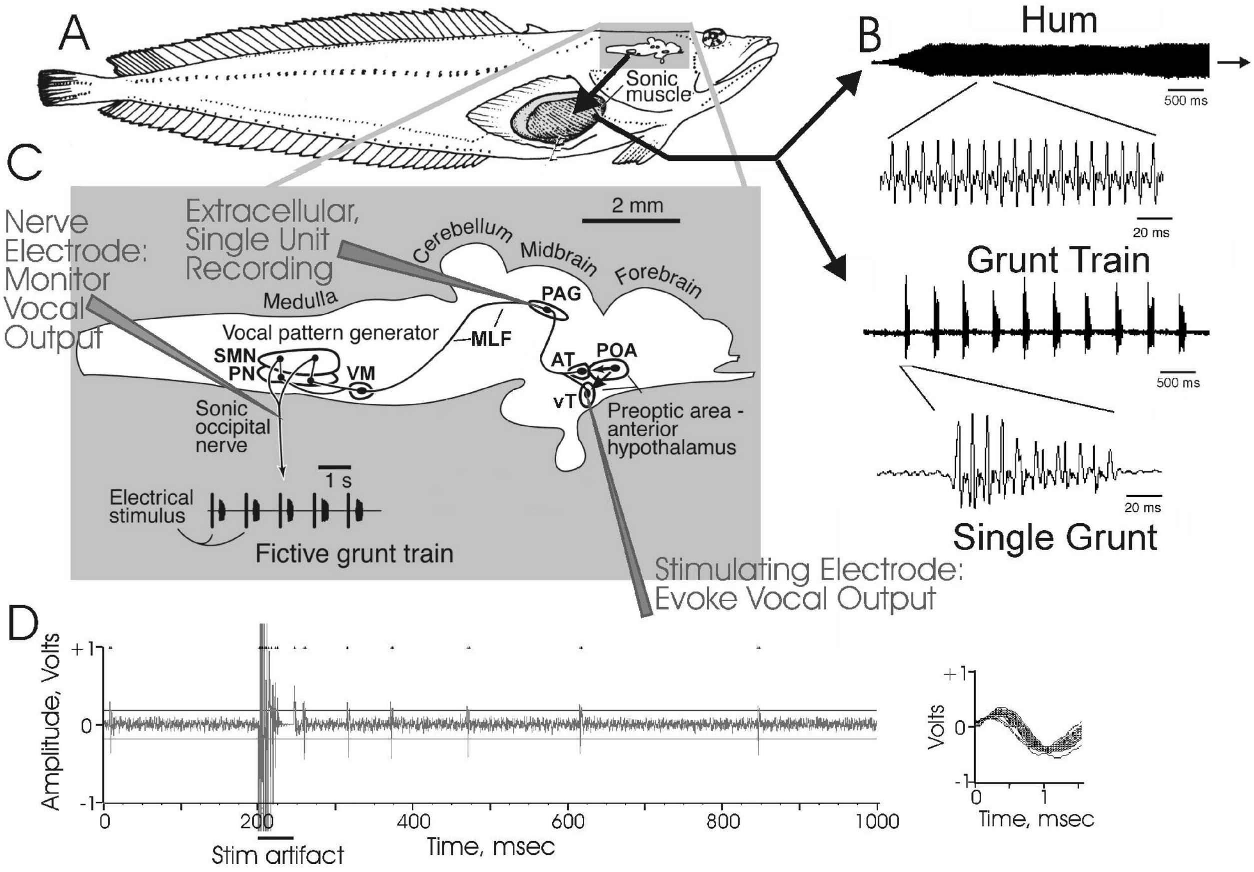

Figure 1. Vocal control system in midshipman fish.

A. Side view of a midshipman with cutaway showing the position of one of the pair of sonic muscles attached to the lateral wall of the swim bladder. These muscles contract synchronously to produce sound. *B*. Oscillograms (y-axis = amplitude, x-axis = time) of the primary call types of type I male midshipman recorded from nest sites. Long duration (min to >1h) “hums” are used to attract and court females to the nest site, while trains of brief (msec) grunts are used in agonistic encounters such as repelling competing males (Brantley and Bass, 1994). The hum’s fundamental frequency and the grunt’s pulse repetition rate (shown with expanded time axes) are close to 100 Hz. *C*. Sagittal view of the central network implicated in vocal production. The pathway from the ventral tuberal nucleus of the anterior hypothalamus (vT) to the midbrain periaqueductal gray (PAG) and then to the ventral medullary nucleus (VM) in the caudal hindbrain is delineated in this study. The synchronous firing of a pacemaker (PN) - sonic motor neuron (SMN) circuit drives the synchronous muscles contraction and sound pulses in a one to one manner. The bottom trace illustrates the rhythmic vocal motor volley, or "fictive vocalization" that mimics the temporal features (pulse repetition rate, duration) of natural grunts. Fictive vocalizations are recorded from a sonic occipital nerve root that carries SMN axons to the ipsilateral sonic muscle and are readily elicited by electrical stimulation at various forebrain and midbrain sites, including the PAG studied here (y-axis = amplitude, volts; x-axis = time). Other abbreviations: AT: anterior tuberal nucleus of the hypothalamus; MLF: medial longitudinal fasciculus; POA: preoptic area. *D*. Example of the activity recorded from a single PAG neuron after hypothalamic (vT) stimulation; the stimulus (Stim) artifact is indicated. In this example, thresholds for spike detection (horizontal lines) were set to +/- 180mV (post amplification voltage). Spike times are indicated by dots above the activity trace. The shapes of discriminated spikes (inset, right) were tightly clustered.

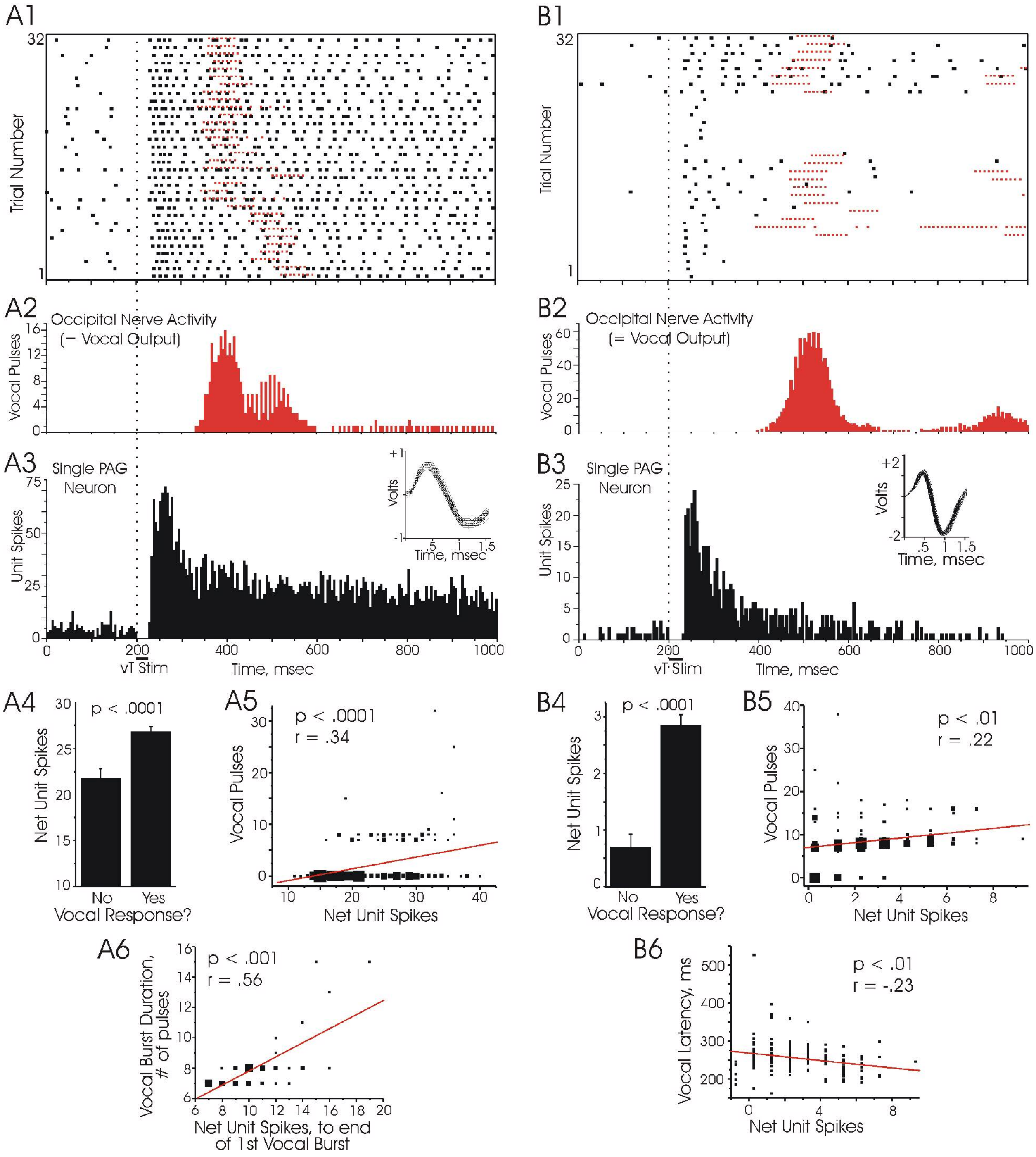

Figure 2. Midbrain PAG activity patterns correlated with vocal output.

Representative examples of activity elicited by vT stimulation in two PAG neurons (*A1-A6, B1-B6*). *A1, B1*: Raster plots of PAG neuronal action potential times (black dots) and vocal pulses (red dots) recorded simultaneously across 32 separate trials (y-axis). *A2, B2*: Post-stimulus time histograms (PSTHs) of vocal pulses across all trials (including those shown in the raster plots) for these two recordings. In both cases, the peak of the vocal response distribution occurred 200-300 msec after vT stimulation. *A3, B3*: PSTHs of spike times from each neuron across all stimulus trials. In both cases, neuronal firing peaked before the vocal peak, by roughly 125 msec in the neuron on the left, and 250 msec in the neuron on the right (compare to *A2, B2*). Insets show the tight clustering of spike shapes, confirming that each recording was from a single unit. Spike shapes are shown for 8 stimulus trials in each case (226 and 53 total spikes, respectively for *A3* and *B3*). *A4, B4*: Net unit spikes (expected spontaneous spikes subtracted) evoked by vT stimulation were significantly greater on trials with a vocal response than on trials on which the vocal response failed (Wilcoxon tests). That is, in both of these neurons, vocal responses were more likely to occur on trials with a greater number of neuronal spikes than on trials with fewer spikes. *A5, B5*: Net unit spikes plotted versus the number of vocal pulses on each trial. For these two units, these variables were significantly correlated (ANOVA, p-, r-values as shown), though the distributions were not well fit by linear regression. Such correlations may reflect a relationship between spike number and vocal burst duration, but may also or instead reflect a relationship between spike number and either vocal response probability or number of vocal bursts (see Results). *A6*: In this neuron, there was a significant correlation, from trial to trial, between the number of vocal pulses and the net number of unit spikes through the end of the first vocal burst. Spike number thus statistically predicted vocal burst duration. This phenomenon was only found in a subset of the “vocal” PAG neurons. Spike number for the cell illustrated in *B*, for example, did not correlate with vocal burst duration. *B6*: In this neuron, there was a significant negative correlation between the net number of stimulus-evoked unit spikes and the latency of the vocal response. All data shown for each cell are from trials on which stimulus parameters were held constant. In *A5*, *A6*, and *B5*, the size of each data point has been scaled to represent the number of trials on which that pair of (x,y) values occurred. Note that for neuron "*A*," in order to more clearly illustrate the effect of spike number on vocal burst duration, the raster plot (*A1*) only shows trials on which a vocal response occurred.

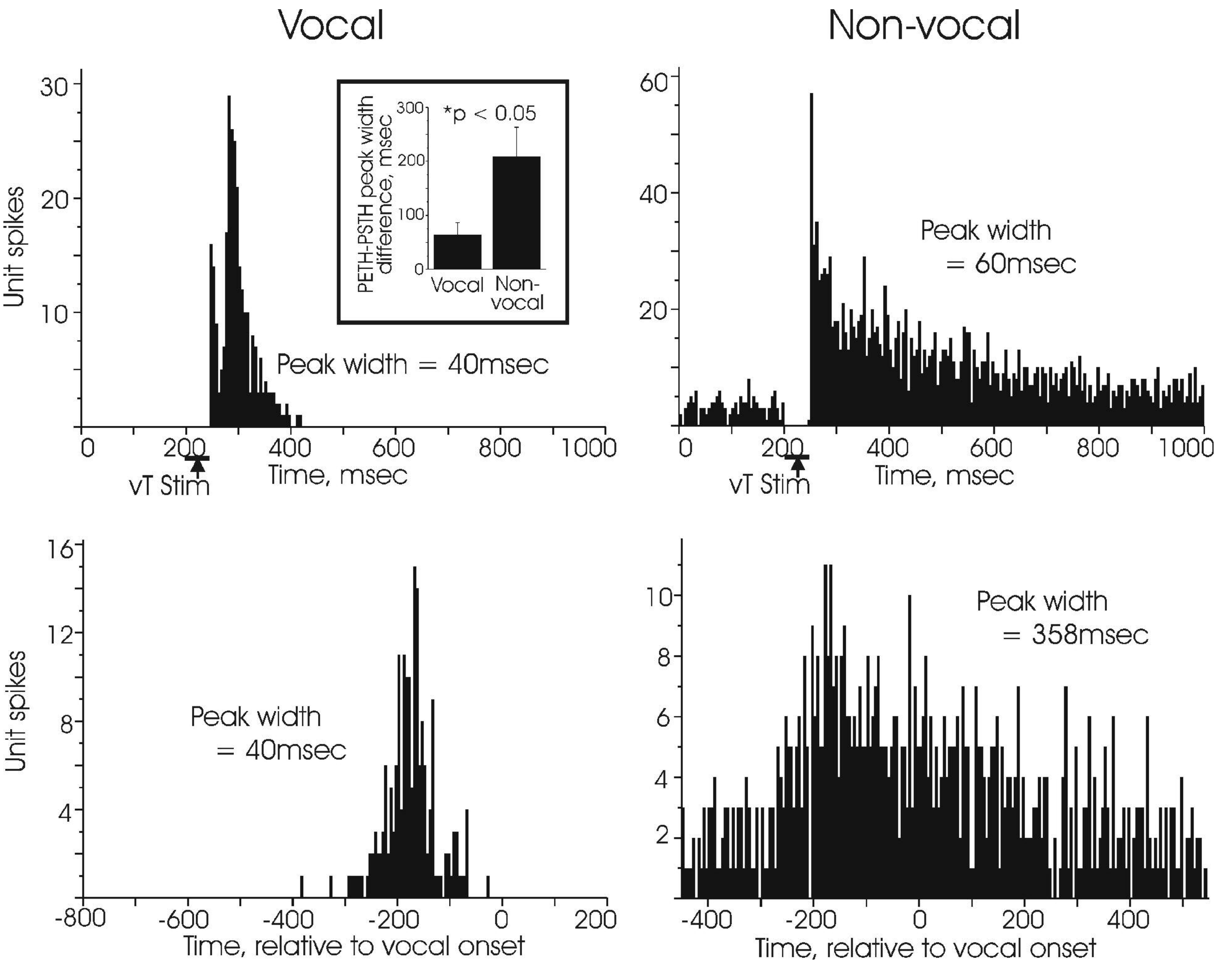

Figure 3. Spike timing in vocal PAG neurons.

Spike timing in vocal PAG neurons was more tightly distributed relative to the onset of vocalization than in non-vocal PAG neurons. Post-stimulus time histograms (PSTHs) are shown in the top row and peri-event time histograms (PETHs, relative to the time of vocal onset) are shown in the bottom row. Shown here are PSTHs and PETHs of spikes in a neuron classified as vocal (left column) because its net spike activity was significantly negatively correlated with vocal latency ($p < 0.03$, ANOVA) and in another neuron classified as non-vocal (right column) because it had no significant correlations between its spike train and the vocal response. In the vocal neuron, the spike timing was just as tightly distributed relative to the onset of the vocal response as it was to the stimulus itself, as determined by comparing the width, at half-maximal height, of the peaks of the PSTH and PETH spike distributions (see Methods). In contrast, in the non-vocal neuron, spikes were much more closely time-locked to the stimulus than to the vocal response. This effect was consistent across the populations of PAG neurons. That is, the difference between the PSTH and PETH peak width was significantly smaller in vocal neurons than in non-vocal neurons (by Wilcoxon test; inset, top left).

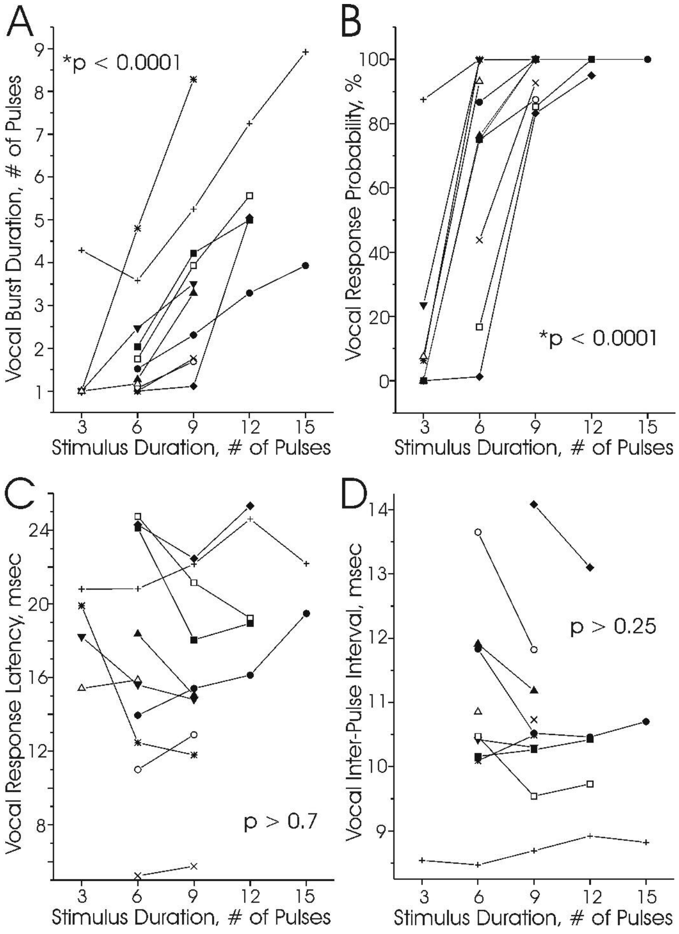

Figure 4. MLF stimulation alters both vocal initiation and burst duration.

Increasing the duration of trains of electrical stimuli delivered to the MLF, in which PAG axons descend to connect with hindbrain vocal structures, caused increases in the duration of the evoked fictive vocal response (*A*) and in the probability of evoking a response (*B*). In contrast, stimulus duration influenced neither the latency of the response (*C*) nor the inter-pulse interval of the response (*D*). Each symbol represents a different experimental animal. Each data point is the mean value for that animal across all trials (a minimum of 8) at the indicated stimulus duration. The p-values shown result from repeated measures ANOVAs. Note that the properties of the vocal responses evoked by MLF stimulation, as shown here, differ from those evoked by vT stimulation as described in Table 1. At equivalent stimulus durations, responses to MLF stimulation occurred at higher probability and shorter latencies, and were shorter in duration, than responses to vT stimulation (Wilcoxon tests).

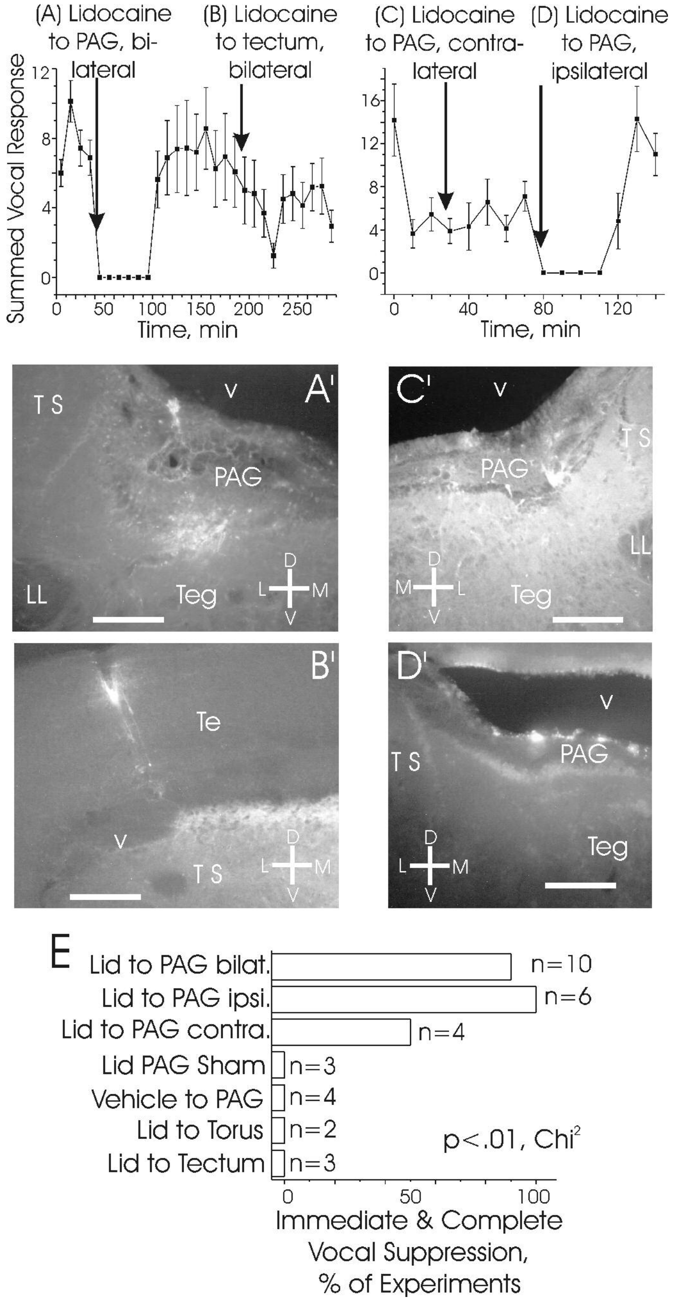

Figure 5. Lidocaine inactivation of the PAG reversibly blocked fictive vocalization.

A-D: Representative experiments from two fish. Summed vocal response to vT stimulation was recorded at 10 min intervals. Photomicrographs of coronal brain sections showing the spread of fluorescent dye from the 4 injection sites shown in the top row for experiments *A-D* are shown immediately below in *A'-D'*. Immediately following focal lidocaine injections to the PAG either bilaterally (as in injection *A*, top left), or ipsilateral to the stimulation site (injection *D*, top right), the vocal response was completely suppressed. The effect was reversible, as the vocal response returned within an average of 42 +/- 7min (SEM). Control injections of lidocaine were made to other midbrain structures, including the tectum (injection *B*) and the torus semicircularis. In contrast to the PAG injections, which immediately and completely blocked the vocal response, these control site injections had either no detectable effects on the vocal response or reduced the response, but only either partially or with a delay of at least 10 min (as per injection *B*). Injections to the contralateral PAG sometimes had no detectable effect on the vocal response (as per injection *C*), but in other instances did block the vocal response immediately (see text for details). Note the spread of dye into the ventricle from the tectal injection *B'*. Dye spreads medially towards the PAG. Note also that the contralateral PAG injection (*C'*), which had no effect on the vocal response, is well matched in both size and position with the ipsilateral PAG injection (*D'*), which immediately and completely blocked the vocal response. *E*: Summary of all experimental injections showing a significant difference across experimental groups in whether the vocal response was immediately and completely blocked.

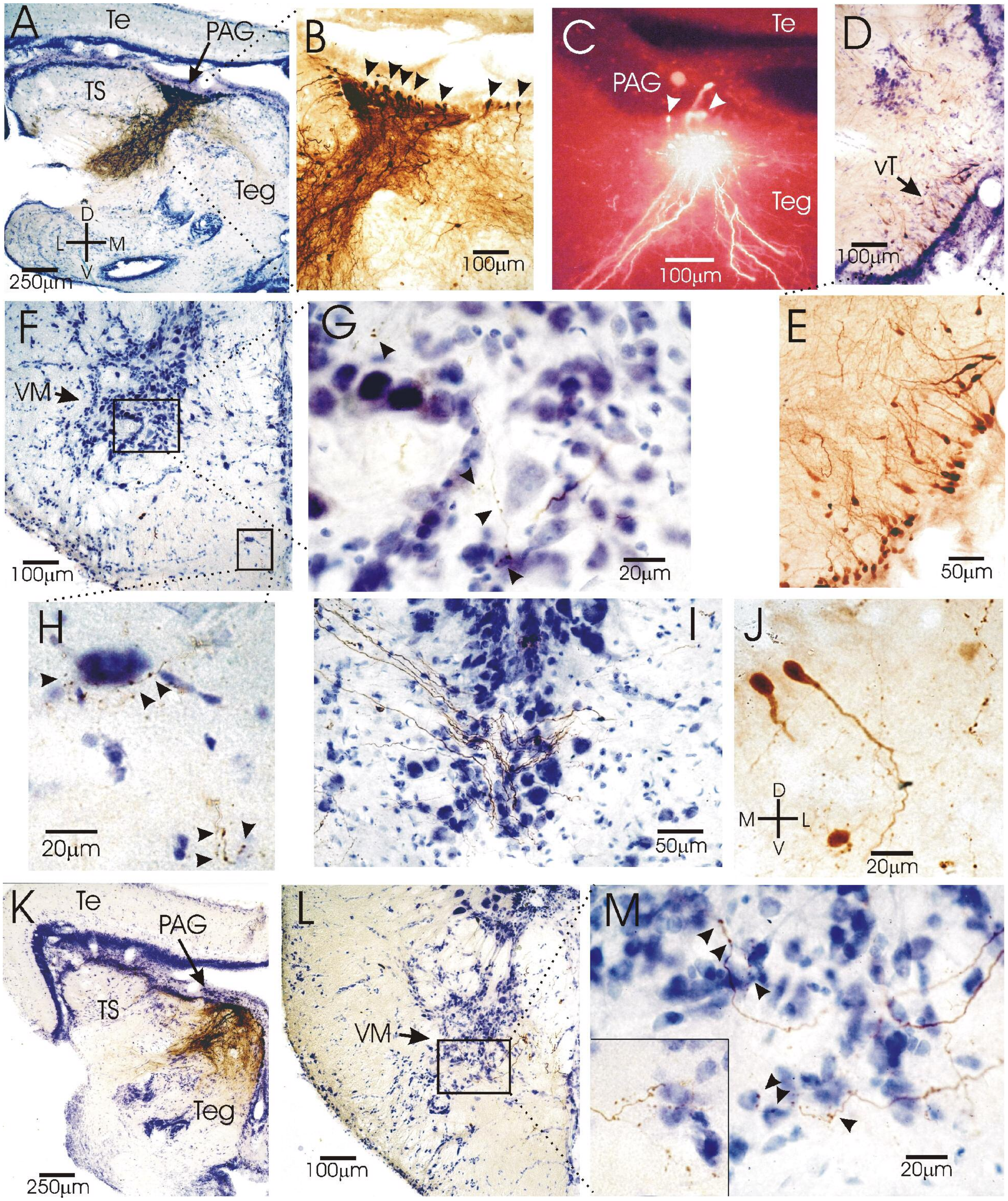

Figure 6. Vocal motor connectivity of the PAG.

Neurobiotin injections into the lateral (*A, B, D – J*) and medial (*K – M*) PAG in type I male midshipman fish. *A*. Low power view of midbrain cytoarchitecture showing position of the periventricular region designated as the PAG. Illustrated here is an injection site in the lateral PAG showing axons extending ventrolaterally through the tegmentum. *B*. Higher magnification view of the same injection site, prior to Nissl stain. The injection was restricted to the PAG, without extending laterally into the torus semicircularis (TS). Examples of labeled PAG neuronal somata are denoted with arrowheads. *C*. Photomicrograph showing the fluorescent dye injected at a representative physiological recording site in the bundle of axon fibers just ventral to the PAG. Backfilled somata in the PAG are indicated with arrowheads. Recording sites (this one is from the neuron whose activity is illustrated in Figure 2B) were highly comparable to the neurobiotin injection sites, implying that both physiological and anatomical data reflect the same population of PAG neurons. *D*. Dense cluster of backfilled neurons in the ventral tuberal hypothalamus (vT) resulting from the injection site shown in *A*. vT is the stimulation site used to elicit vocal responses in all recording and lidocaine inactivation experiments in this study. *E*. Higher magnification view of the section shown in *D*, prior to Nissl stain. *F*. Nissl stained section in the caudal medulla of the same brain, at the level of the ventral medullary nucleus (VM), a component of the hindbrain vocal pattern generator (Fig. 1*C*; see Bass et al., 1994 for detailed mapping). *G*. High magnification view of the indicated portion of VM in *F*, showing axon fibers with distinct swellings indicative of presynaptic boutons (examples denoted with arrowheads). *H*. High magnification view of the indicated area in *F* that is ventral and medial of VM, another area where axon fibers with presumptive presynaptic boutons (arrowheads) was found. This relatively cell-free area is where a dense commissural bundle of axons connects VM-VM bilaterally (Bass et al., 1994). *I*. Labeled fibers from the same injection site crossing the midline (from left to right) at the level of the PAG. Some of these fibers terminated in the medial contralateral tegmentum, others extended to the contralateral PAG. *J*. High magnification view of backfilled neurons and fibers with presumptive presynaptic boutons in the PAG contralateral to the injection site in *A*. *K*. Neurobiotin injection in another fish into the PAG medial and rostral to the injection in *A*. *L*. Nissl stained section at the level of VM. *M*. High magnification view of the indicated area at the medial edge of VM in *L*. Labeled axons, after exiting the MLF, enter VM from the right; presumptive boutons (arrowheads) are only found within the boundaries of VM. Inset shows a dense cluster of boutons at the medial edge of VM in an adjacent section. All sections coronal, scale bars as indicated. All sections, except *J,* are oriented as in section *A*. Other abbreviations: Te: midbrain tectum; Teg: midbrain tegmentum.

Table 1.

PAG Single Unit Data	**Mean**	**StdDev**	**Range**	**Median**
Spontaneous Activity Rate, Hz	1.4	2.1	0-8.2	0.2
Stim.-Evoked Spikes, net	3.7	6.0	0.3-28.8	1.5
Inter-spike interval, mode, msec	44	44	7-145	21
Latency to 1st spike, msec	43	45	5-191	24
Lag, 1st spike to voc. resp., msec	-166	110	-350-85	-161
Vocal Response Data				
% Trials w/ vocal response	62	29	17-98	65
% Vocal resp. w/ > 1 burst	18	22	0-62	8
# Pulses, 1st vocal burst	7.1	2.0	3.7-10.9	7.6
Duration, 1st vocal burst, msec	62	19	27-106	61
Latency to vocal onset, msec	201	76	92-317	216
Inter-pulse interval, msec	10.3	0.9	9.0-12.0	10.1
Inter-burst interval, msec	349	65	287-498	341

Table 1. Description of the firing properties of PAG neurons, both spontaneous and stimulus-evoked, and of the fictive vocal response. Most PAG neurons fired spontaneously at only a very low rate. Most of the single unit data distributions were skewed toward the lower end of the range, as indicated by the differences between the medians and means. Note that the average modal inter-spike interval of PAG neurons post stimulus was much higher (~4x) than the

interpulse interval of the vocal response. The properties of the fictive vocal response (pulses per burst, duration, inter-pulse interval, and interburst interval) were very similar to those of natural grunt trains (see Bass et al., 1999). All data presented here result from constant duration (12 pulse) stimulation of vT.

Table 2.

	Net Unit Spikes vs. Total Vocal Resp. Pulses (vocal failure trials included)	**Net Unit Spikes vs. Vocal Resp. Pulses, to end of 1st Vocal Burst (vocal failure trials excluded)**	**Net Unit Spikes vs. Vocal Resp. Latency**	**Net Unit Spikes: Trials w/ vs. Trials w/o Vocal Resp.**	**Unit Spike Latency: Trials w/ vs. Trials w/o Vocal Resp.**
"Vocal":	8 +	3 +	1 +	5 +	2 +
(n = 13)	1 -	2 -	2 -	2 -	2 -
"Non-Vocal":	0	0	0	0	0
(n = 11)					

Table 2. Heterogeneity of correlations between unit activity and vocal response. Vocal neurons were those with at least one of the indicated statistically significant correlations; most exhibited more than one significant correlation. For each type of correlation, the number of units with a significant positive (+) or negative (-) correlation is indicated. The last two columns indicate the number of neurons in which the net number of unit spikes or the spike latency was

significantly greater (or less) on trials with a vocal response (resp.) than on trials without (w/o) a vocal response.