

coordinated activity of these networks in the striatal function, and, therefore, they may participate in its dysfunction in brain disorders. Our results demonstrate the existence of functional astro-neuronal networks that comprise subpopulations of astrocytes, neurons, and synapses belonging to specific brain circuits, which may differentially control specific circuit activity through selective signaling between particular astrocytes and neurons.

REFERENCES AND NOTES

1. A. Araque *et al.*, *Neuron* **81**, 728–739 (2014).
2. C. Eroglu, B. A. Barres, *Nature* **468**, 223–231 (2010).
3. M. M. Halassa, P. G. Haydon, *Annu. Rev. Physiol.* **72**, 335–355 (2010).
4. G. Perea, M. Navarrete, A. Araque, *Trends Neurosci.* **32**, 421–431 (2009).
5. A. Volterra, J. Meldolesi, *Nat. Rev. Neurosci.* **6**, 626–640 (2005).
6. G. Perea, A. Araque, *Science* **317**, 1083–1086 (2007).
7. C. Agulhon, T. A. Fiacco, K. D. McCarthy, *Science* **327**, 1250–1254 (2010).
8. T. A. Fiacco *et al.*, *Neuron* **54**, 611–626 (2007).
9. B. C. Shonesy *et al.*, *Nat. Neurosci.* **16**, 456–463 (2013).
10. A. C. Kreitzer, R. C. Malenka, *Nature* **445**, 643–647 (2007).
11. W. Shen, M. Flajolet, P. Greengard, D. J. Surmeier, *Science* **321**, 848–851 (2008).
12. M. Uchigashima *et al.*, *J. Neurosci.* **27**, 3663–3676 (2007).
13. V. Chevaleyre, K. A. Takahashi, P. E. Castillo, *Annu. Rev. Neurosci.* **29**, 37–76 (2006).
14. A. C. Kreitzer, R. C. Malenka, *Neuron* **60**, 543–554 (2008).
15. R. L. Albin, A. B. Young, J. B. Penney, *Trends Neurosci.* **12**, 366–375 (1989).
16. L. M. Suárez *et al.*, *Biol. Psychiatry* **75**, 711–722 (2014).
17. S. Ares-Santos, N. Granado, I. Espadas, R. Martínez-Murillo, R. Moratalla, *Neuropsychopharmacology* **39**, 1066–1080 (2014).
18. M. Gómez-Gonzalo *et al.*, *Cereb. Cortex* (2014).
19. M. Navarrete, A. Araque, *Neuron* **68**, 113–126 (2010).
20. R. Min, T. Nevian, *Nat. Neurosci.* **15**, 746–753 (2012).
21. M. Navarrete *et al.*, *PLoS Biol.* **10**, e1001259 (2012).
22. M. A. Di Castro *et al.*, *Nat. Neurosci.* **14**, 1276–1284 (2011).
23. J. Petrávic, T. A. Fiacco, K. D. McCarthy, *J. Neurosci.* **28**, 4967–4973 (2008).
24. T. Fellin *et al.*, *Neuron* **43**, 729–743 (2004).
25. M. Navarrete, A. Araque, *Neuron* **57**, 883–893 (2008).
26. G. Perea, A. Araque, *J. Neurosci.* **25**, 2192–2203 (2005).
27. G. Cui *et al.*, *Nature* **494**, 238–242 (2013).
28. A. V. Kravitz *et al.*, *Nature* **466**, 622–626 (2010).
29. I. Ruiz-DeDiego, B. Mellstrom, M. Vallejo, J. R. Naranjo, R. Moratalla, *Biol. Psychiatry* **77**, 95–105 (2015).
30. V. M. André *et al.*, *J. Neurosci.* **31**, 1170–1182 (2011).

ACKNOWLEDGMENTS

We thank D. Redish, W. Buño, A. Diez, and R. Gómez for helpful comments; J. Lerma and A. Valero for valuable help; B. Pro for technical assistance; and J. Chen (University of California–San Diego) and A. Zimmer (University of Bonn) for providing *IP₃R2^{-/-}* and *CB1R^{-/-}* mice, respectively. A.A. was supported by Cajal Blue Brain, the Human Frontier Science Program (grant RGP0036/2014), the Wallin family, and V. and L. Han. Support for this work was also provided by the Ministerio de Economía y Competitividad of Spain (grants BFU2013-47265-R, CSD2010-00045, and RYC-2012-12014 to G.P. and grants SAF2013-48532-R and CB06/05/0055 to R.Mo.) and by the Juan de la Cierva Program (grant JCI-2010-07693 to R.Ma.). Data described in the paper are archived at the Instituto Cajal (Madrid, Spain). The authors declare no competing financial interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6249/730/suppl/DC1
Materials and Methods
Figs. S1 to S7
References (31–39)

28 January 2015; accepted 17 June 2015
10.1126/science.aaa7945

LANGUAGE DEVELOPMENT

The developmental dynamics of marmoset monkey vocal production

D. Y. Takahashi,^{1,2*} A. R. Fenley,^{1,2} Y. Teramoto,¹ D. Z. Narayanan,^{1,2} J. I. Borjon,^{1,2} P. Holmes,^{1,3} A. A. Ghazanfar^{1,2,4*}

Human vocal development occurs through two parallel interactive processes that transform infant cries into more mature vocalizations, such as cooing sounds and babbling. First, natural categories of sounds change as the vocal apparatus matures. Second, parental vocal feedback sensitizes infants to certain features of those sounds, and the sounds are modified accordingly. Paradoxically, our closest living ancestors, nonhuman primates, are thought to undergo few or no production-related acoustic changes during development, and any such changes are thought to be impervious to social feedback. Using early and dense sampling, quantitative tracking of acoustic changes, and biomechanical modeling, we showed that vocalizations in infant marmoset monkeys undergo dramatic changes that cannot be solely attributed to simple consequences of growth. Using parental interaction experiments, we found that contingent parental feedback influences the rate of vocal development. These findings overturn decades-old ideas about primate vocalizations and show that marmoset monkeys are a compelling model system for early vocal development in humans.

Human vocal development is the outcome of interactions among an infant's developing body and nervous system and his or her experience with caregivers (1, 2). Infant cries decline over the first 3 months as they transition into preverbal vocalizations (3). The rates of these transitions are influenced by social feedback: Contingent responses of caregivers spur the development of more mature vocalizations (4). In contrast, nonhuman primate vocalizations are widely viewed as undergoing little or no production-related acoustic changes during development, and any such changes are attributed solely to passive consequences of growth (5).

We tracked the vocal development of marmoset monkeys (*Callithrix jacchus*; $n = 10$)—a voluble, cooperative breeding species (6)—from the first postnatal day (P1) until they produced adultlike calls at 2 months of age. Recordings were taken at least twice weekly in two contexts: undirected (social isolation) and directed (with auditory, but not visual, contact with their mother or father). Such early and dense sampling is necessary to accurately capture developmental changes in marmosets because this species develops rapidly (7). Each recording session began with ~5 min in the undirected context followed by ~15 min in the directed context, with mothers and fathers alternating between each session. In the undirected context, infants exhibited a

dramatic change in vocal production (Fig. 1A and audio S1 to S8). At P1, vocalizations were more numerous and variable in their spectrotemporal structure than those recorded in later weeks. The number and variability of calls diminished over 2 months, approaching mature vocal output with exclusive production of whistle-like “phee” calls in this context (8).

To quantify this developmental change as a continuous process without the bias of ethological labels (9), for each of the 73,421 recorded utterances, we measured four acoustic parameters similar to those used for tracking birdsong development (10): duration, dominant frequency, amplitude modulation (AM) frequency, and Wiener entropy (a measure of spectral flatness) (Fig. 1B). Changes in all four parameters were statistically significant ($n = 301$ sessions, $P < 0.001$), showing that vocalizations underwent a transformation in the first 2 months, whereby utterances lengthened, dominant and AM frequencies decreased, and entropy decreased. This pattern of change is consistent with both human and songbird vocal development (10, 11). These changes in infant vocalizations, although not subtle, may be due solely to physical maturation (5). To test this, we used body weight as a proxy for overall growth [weight correlates well with vocal apparatus size in monkeys (12)]. Weight changes visibly contrasted with the trajectories of the acoustic parameters (Fig. 1, B and C). To quantify this difference, we used weight to predict changes in the acoustic parameters. Predicted average parameter values, given the average weight for each postnatal day, are shown in Fig. 1B. If growth completely explained the acoustic change, the residues would be uncorrelated and identically distributed across postnatal days. Using the Akaike information criterion (AIC), the best polynomial-fit order was three

¹Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA. ²Department of Psychology, Princeton University, Princeton, NJ 08544, USA.

³Department of Mechanical and Aerospace Engineering and Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA.

⁴Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA.

*Corresponding author. E-mail: takahashi@princeton.edu (D.Y.T.); asifg@princeton.edu (A.A.G.)

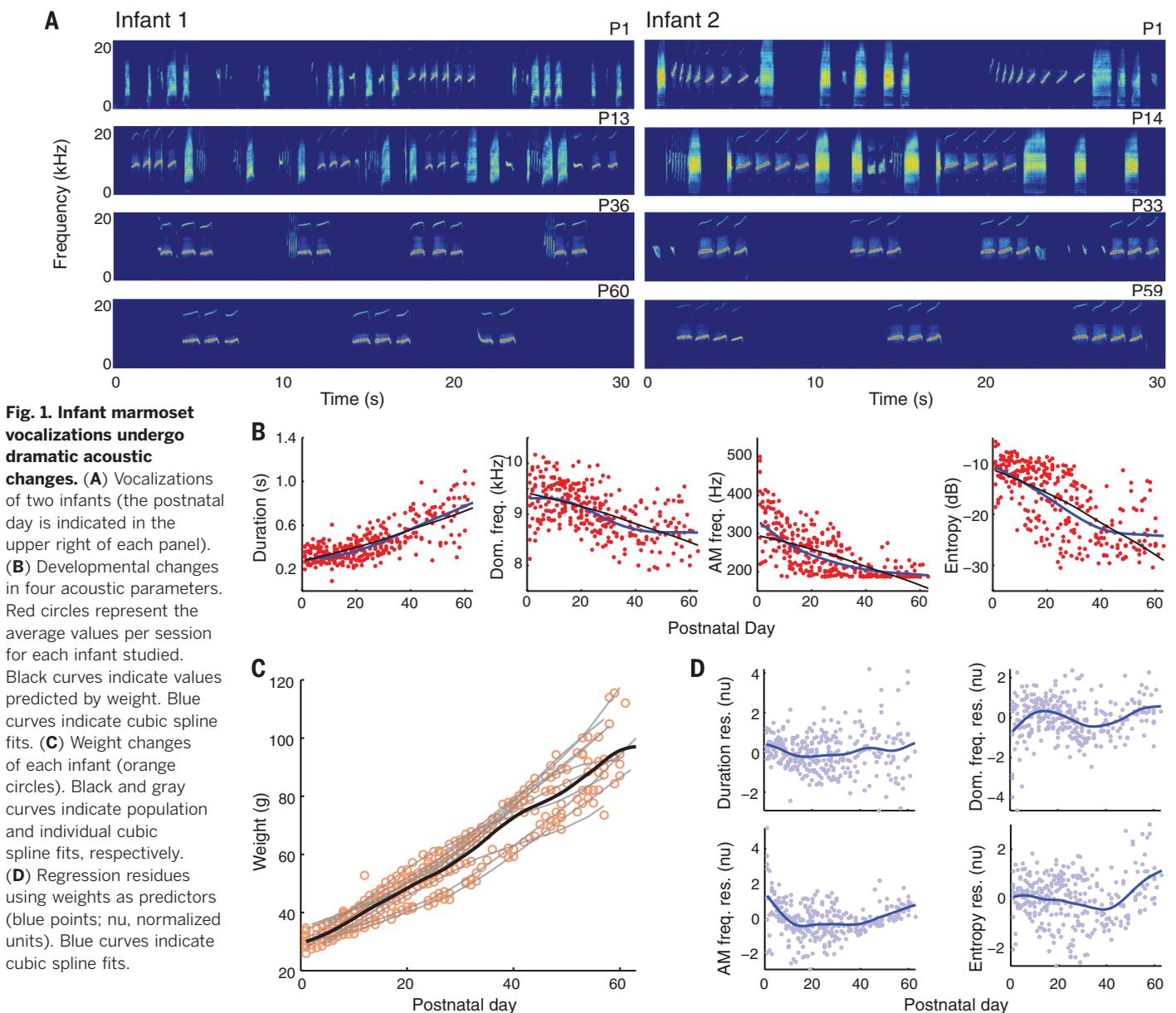
for all residues related to the acoustic parameters (Fig. 1D). To account for possible nonlinear relationships between growth and acoustic parameters, we log-transformed the weight and acoustic parameters. The log-transformed weight did not predict the log-transformed acoustic parameters (fig. S1, A to C). Thus, simple patterns of growth (linear or nonlinear) do not accurately predict acoustic changes in infant marmoset vocalizations.

A subset of the early vocalizations of humans and songbirds are incorporated into the adult repertoire, whereas others are transient, serving as scaffolding for later vocalizations (3, 10). To test whether infant marmosets follow a similar trajectory, we first measured the extent to which their calls were distinct. Two parameters, duration and entropy, identified disjoint clusters

in syllable sequences (Fig. 2A). With development, the clusters became more distinct and less numerous. Using all four parameters, we computed optimal cluster numbers for each marmoset in each session (Fig. 2B). On average, the number of clusters decreased from around four to one or two ($P < 0.001$). The clusters represent distinct ethologically based syllable types (13, 14) (Fig. 2C). Phee syllables increased to over 95% of all vocalizations by 2 months ($P < 0.001$); all other calls decreased ($P = 0.005$ for trills; $P < 0.001$ for all other syllables) (Fig. 2D). The changes in syllable proportions potentially represent two independent processes: change in usage (15) and transformation of immature calls into mature versions (Fig. 2E). Twitters and trills are produced frequently by marmosets of all ages (13, 14), but in adults, they are typically produced when

in visual contact with conspecifics and not in the undirected context. Thus, twitters and trills undergo a change in usage in the first 2 months. In contrast, cries, phee-cries, and subharmonic phees are only produced by infants; mature phees are produced almost exclusively during vocal exchanges that occur when out of visual contact with conspecifics (8).

Because these infant-only calls share some features with the mature phee call (e.g., a common duration), we hypothesized that they represent immature phees, consistent with vocal transformations observed in preverbal human infants (11) and songbirds (10) but contrasting with previous reports on developing primates (5). It is possible that these transitional forms are related to growth but sound distinct because of nonlinearities in the vocal production



system (16). This would suggest that a single biomechanical mechanism generates cries, phees, and the transitional forms and that the transitional calls result from smooth changes through a parameter space. To test this idea, we developed a model based on one that successfully reproduces syllable types in zebra finch song but that is nonspecific with regard to songbird versus mammalian vocal anatomy (Fig. 3A) (17). Our simulations verified that the model can reproduce the marmoset call types described above (Fig. 3, B to E). The simulations also revealed the underlying biomechanics corresponding to different calls at different levels of pressure (respiratory power) and laryngeal muscle tension (Fig. 3F). Broadband cries were produced at low pressure and muscle tension, where small variations cause large changes in spectral content because of nonlinear vocal fold dynamics. Phees occurred at higher pressures and tensions, and subharmonic phees occurred in an intermediate region, supporting their classification as

transitional calls. Rapid switching between high and low pressure and tension states produced the phee-cries. Throughout, linear changes in pressure and tension produced nonlinear acoustic effects.

To test the model's overall validity and the prediction that respiration during cries is less stable than during phees, we measured respiratory activity via electromyography (EMG) in five P1 infants. We investigated whether different respiratory patterns underlie cries and phees with similar intersyllable intervals (Fig. 3, G and I). The EMG signals were more uniform across phees than across cries (Fig. 3, H and J), as quantified by the cost of dynamic time warping (DTW) (18). For each infant, the mean DTW costs for phees were smaller than they were for cries ($P < 0.001$) (Fig. 3K). Therefore, phee syllables at least partly result from more stable respiration; immature respiratory control leads mainly to cries early in life, consistent with the model prediction. Overall, these data support our hypothesis that cries are immature phees.

Thus, although vocal acoustic changes were dramatic, physiological growth could explain the transition from cries to phees, as improved respiratory and/or laryngeal control modulates spectral parameters (Fig. 1B), reducing the entropy. However, if the cries-to-phees transition was solely driven by physical maturation, it would be impervious to social feedback. Yet, consistent with a role for vocal feedback in guiding development, marmoset monkeys exhibit a developmental pattern of *FoxP2* expression in their thalamocortical-basal ganglia circuit (19) that is analogous to that of songbirds and humans (20). This suggests that marmoset infants may use this circuit to guide their phee-call development through reward-based parental feedback, as birds and humans do (21). To assess the effect of parent-infant vocal interactions in marmosets, we quantified their vocal exchanges in the directed context, where infants and their mother or father were in auditory, but not visual, contact.

Infant and parent vocalizations were parsed into whole multisyllabic calls according to the

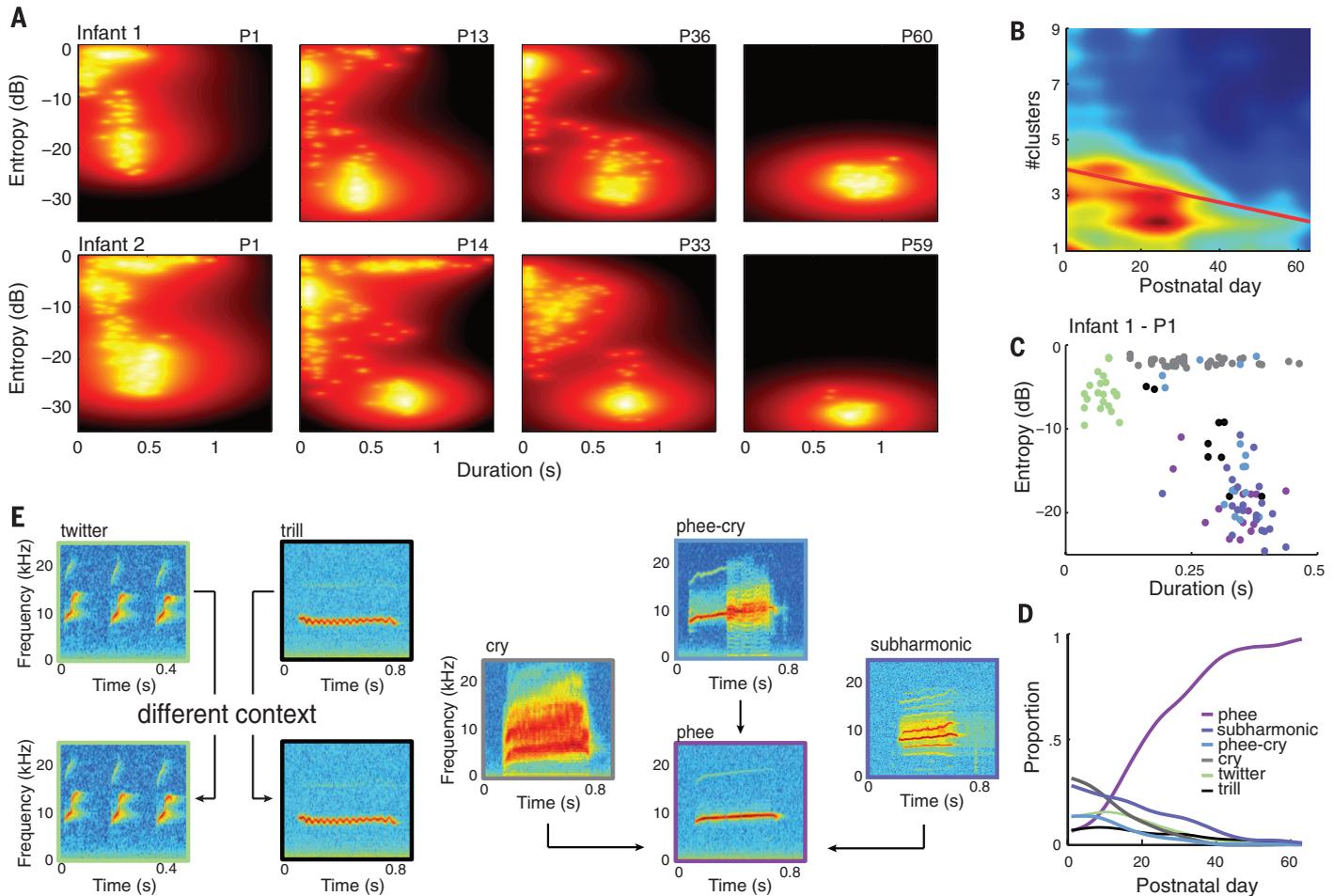


Fig. 2. A subset of infant marmoset calls transform into adultlike phee calls. (A) Probability maps of duration and entropy for the infants in Fig. 1A. Lighter colors indicate higher probabilities. (B) Distribution of cluster numbers by developmental age. Warmer colors indicate more frequent occurrences of cluster numbers. The red line is the regression fit. (C) Correspondence between clusters and different syllable types for a marmoset at P1. (D) Developmental changes in the proportion of call types. Colors in (C) indicate the same call types as in (D). (E) Twitters and trills change in usage, whereas cries, phee-cries, and subharmonic phees are hypothesized to transition to phee calls.

bimodal distribution of their intersyllable intervals (8). We recorded 8800 infant phees, 11,798 infant cries, and 6567 adult phees, of which 2512 were contingent responses to infant phees [those falling within a turn-taking interval as seen in adults (8)]. Parents produced mostly phee calls (>98%). Typical examples of infant phee and cry production during interactions over the first 2 months and the phee/cry ratio across days are shown in Fig. 4, A and B. As in the undirected context (Fig. 2D), cries gave way to phees, but the transition occurred rapidly. For each infant, we used the point where the phee/cry ratio first crossed zero to mark the transition day (Fig.

4C). Transitions were typically sharp, but their timing varied substantially across infants (~10 to 40 postnatal days). If physiological growth completely explained the cries-to-phees transition, the weight-change rate and the timing of the transition (zero-crossing) day would be correlated. However, we found no significant correlation ($n = 10$ infants, t test, $P = 0.684$) (Fig. 4D); growth alone cannot explain the timing of the cries-to-phees transition.

We then investigated whether parental responses to infant vocalizations affect the timing of the cries-to-phees transition. This would explain, at least partially, its variability across infants.

Infants could be influenced by contingent responses only or by the total number of adult vocalizations that they hear. The fraction of infant phees that elicited contingent parental phee responses before the zero-crossing day correlated significantly with the timing of the zero-crossing day ($n = 10$ infants, t test, $P = 0.005$) (Fig. 4E). Proportions of noncontingent parental calls (91.5% of all calls on average) were not significantly correlated with this timing ($n = 10$ infants, t test, $P = 0.558$) (Fig. 4F). Therefore, contingent vocal responses from parents influence the timing of the cries-to-phees transition by reinforcing the production of phee calls.

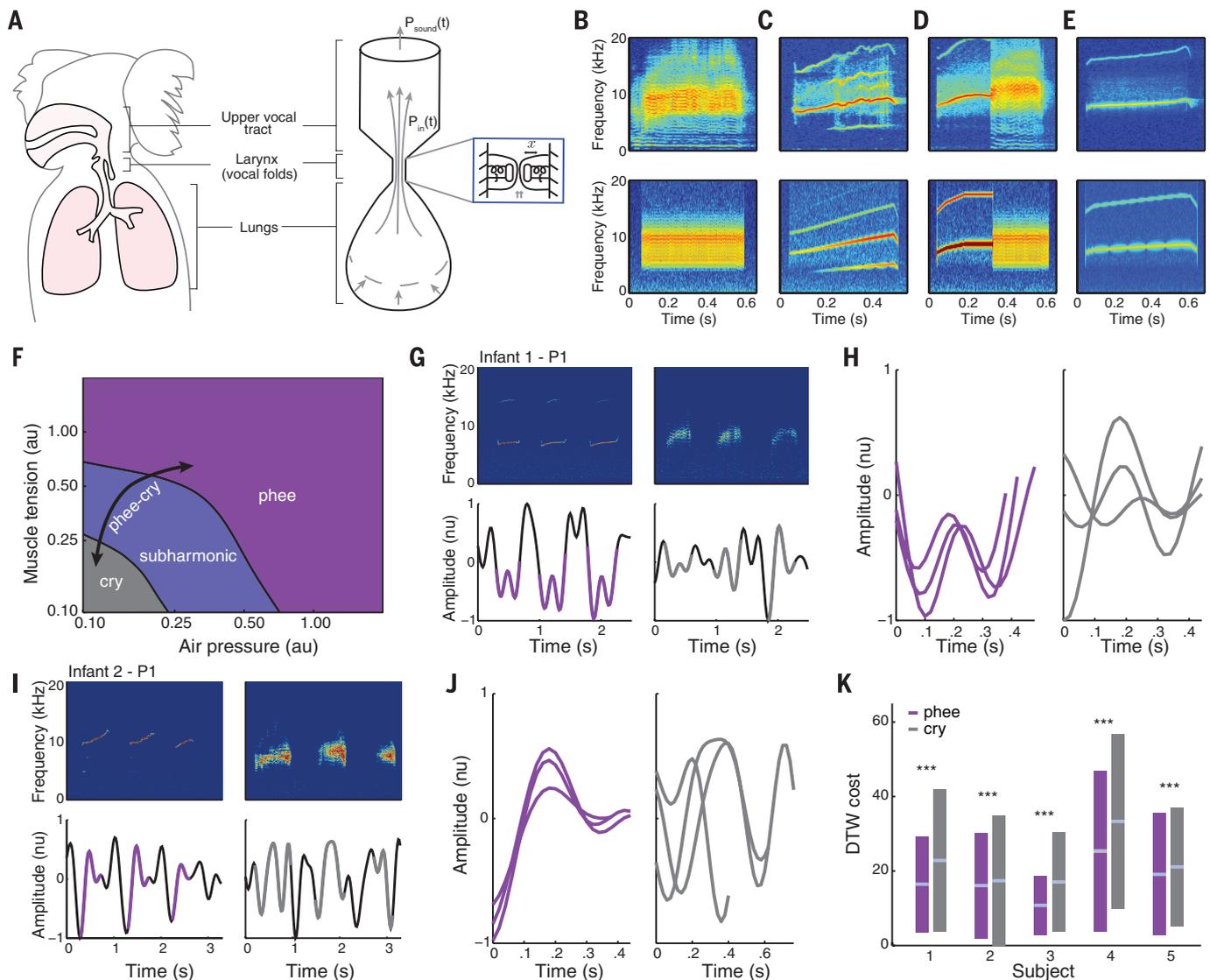


Fig. 3. Biomechanical model of vocal folds can reproduce infant marmoset calls. (A) Main elements of the vocal tract and respective model representations, where $P_{in}(t)$ is pressure at the inlet to the vocal tract, $P_{sound}(t)$ is pressure upon exit from the mouth, and x is the displacement of the vocal fold. Gray arrows show airflow. (B to E) Top: Representative recordings of a cry (B), subharmonic phee (C), phee-cry (D), and phee (E); bottom: corresponding model simulations. (F) Changing air pressure and muscle tension produces

different calls (au, arbitrary units). (G and I) Top: Three-syllable phees (left) and cries (right) for two marmosets; bottom: their corresponding EMG activities. (H and J) Respiratory EMG activity during call production. Curves are the purple and gray segments from (G) and (I), aligned at the syllable onset of phees and cries, respectively. (K) Average DTW cost on P1 for five marmosets. Central light-colored bars show means, rectangles show standard deviations, and asterisks indicate significant differences ($P < 0.001$).

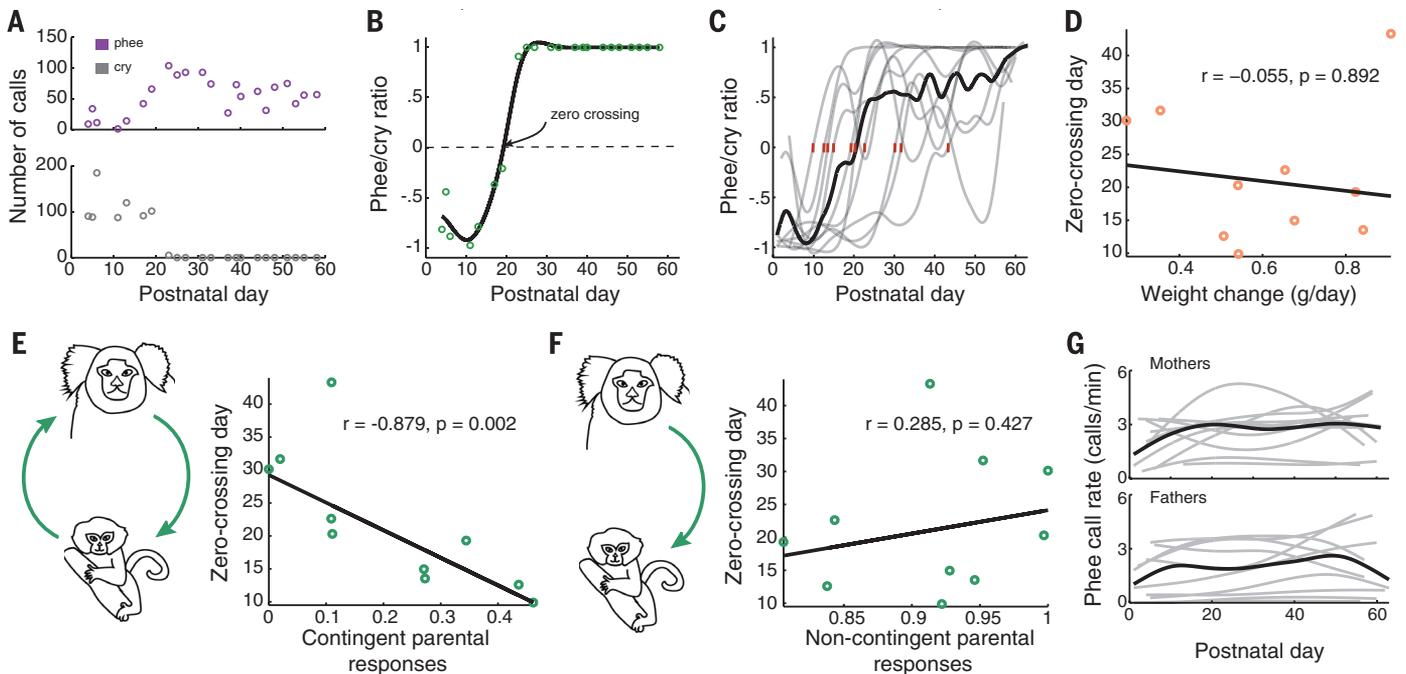


Fig. 4. Transition from cry to phee is influenced by contingent parental calls. (A) Numbers of cries and phees over 2 months for a single infant. (B) Phee/cry ratio for the infant in (A) across days. (C) Phee/cry ratios (gray curves) and zero-crossing days (red ticks) for each infant and for the population (black curve). Black and gray curves in (B) and (C) are cubic spline fits. (D) Correlation between the weight-change rate and the zero-crossing day among infants. (E and F) Correlations between the zero-crossing day and the proportion of contingent and noncontingent parental responses, respectively. (G) Rates of individual parental phee-call production during infant development (gray) and the population average (black).

We address two possible caveats to this conclusion. First, it is possible that, through shared genetics, fast-transitioning infants are born to more vocally interactive parents. To test this, we correlated the frequency of contingent parental calls and the zero-crossing day for six full siblings born from the same parents. If shared genetics were driving the result, then there would be no correlation between contingent parental responses and the zero-crossing day. We found, however, that there remained a statistically significant correlation ($n = 6$ infants, $P = 0.046$) (fig. S2). Moreover, we found no difference between the slopes of the regressions for the full-siblings and all-infants data (test for equality, $P = 0.953$).

Second, it is possible that changing patterns of infant calling are due to changes in parental call output. The phee-call production rates of each infant's parents during development are shown in Fig. 4G; neither parent changed their production rates (mother, $P = 0.132$; father, $P = 0.235$). Based on these analyses, we conclude that the cries-to-phees transition is influenced by contingent responses from parents, not by shared genetics or changes in parental vocal output.

Our findings demonstrate that infant marmoset calls undergo dramatic changes during the first 2 months of life, transforming from cries into mature, adultlike phee calls. The timing of this transition is partly attributable to maturation but is also influenced by contingent parental vocal feedback. This is consistent with preverbal vocal development in humans, whereby

(i) natural categories of sounds change as respiratory, laryngeal, and facial components mature, and (ii) in parallel, vocal feedback sensitizes infants to certain features of those sounds, and the sounds are modified accordingly. Our findings contrast with previous reports that nonhuman primate vocalizations undergo little or no postnatal change and are impervious to social feedback (5). The complex and socially dependent vocal development we observed in marmoset monkeys may be a necessary condition of the vocal learning observed in humans.

REFERENCES AND NOTES

- E. Thelen, in *Biological and Behavioral Determinants of Language Development*, N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch, M. Studdert-Kennedy, Eds. (Lawrence Erlbaum Associates, Hillsdale, NJ, 1991), pp. 339–362.
- L. Byrge, O. Sporns, L. B. Smith, *Trends Cogn. Sci.* **18**, 395–403 (2014).
- D. K. Oller, *The Emergence of the Speech Capacity* (Lawrence Erlbaum Associates, Mahwah, NJ, 2000).
- M. H. Goldstein, A. P. King, M. J. West, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8030–8035 (2003).
- S. E. R. Egnor, M. D. Hauser, *Trends Neurosci.* **27**, 649–654 (2004).
- J. I. Borjón, A. A. Ghazanfar, *Brain Behav. Evol.* **84**, 93–102 (2014).
- A. de Castro Leão, A. Duarte Dória Neto, M. B. C. de Sousa, *Comput. Biol. Med.* **39**, 853–859 (2009).
- D. Y. Takahashi, D. Z. Narayanan, A. A. Ghazanfar, *Curr. Biol.* **23**, 2162–2168 (2013).
- D. Lipkind, O. Tchernichovski, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15572–15579 (2011).
- O. Tchernichovski, P. P. Mitra, T. Lints, F. Nottebohm, *Science* **291**, 2564–2569 (2001).

- E. Scheiner, K. Hammerschmidt, U. Jurgens, P. Zwirner, *J. Voice* **16**, 509–529 (2002).
- W. T. Fitch, *J. Acoust. Soc. Am.* **102**, 1213–1222 (1997).
- B. M. Bezerra, A. Souto, *Int. J. Primatol.* **29**, 671–701 (2008).
- A. L. Pistorio, B. Vintch, X. Wang, *J. Acoust. Soc. Am.* **120**, 1655–1670 (2006).
- A. M. Elowson, C. T. Snowdon, C. Lazarro-Perea, *Behaviour* **135**, 643–664 (1998).
- W. T. Fitch, J. Neubauer, H. Herzog, *Anim. Behav.* **63**, 407–418 (2002).
- A. Amador, Y. S. Perí, G. B. Mindlin, D. Margoliash, *Nature* **495**, 59–64 (2013).
- J. A. Kogan, D. Margoliash, *J. Acoust. Soc. Am.* **103**, 2185–2196 (1998).
- M. Kato et al., *Brain Lang.* **133**, 26–38 (2014).
- I. Teramitsu, L. C. Kudo, S. E. London, D. H. Geschwind, S. A. White, *J. Neurosci.* **24**, 3152–3163 (2004).
- S. Syal, B. L. Finlay, *Dev. Sci.* **14**, 417–430 (2011).

ACKNOWLEDGMENTS

We thank L. Kelly for providing comments on an earlier draft. This work was supported by a Pew Latin American Fellowship (D.Y.T.), a Brazilian Science Without Borders Fellowship (D.Y.T.), an NSF Graduate Fellowship (J.I.B.), and the James S. McDonnell Scholar Award (A.A.G.). Data for each figure are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6249/734/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 and S2
References (22–29)
Audio S1 to S8
Supplementary Data

13 March 2015; accepted 12 June 2015
10.1126/science.aab1058