



# Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans

Christian J. Sumner<sup>a,1,2</sup>, Toby T. Wells<sup>a</sup>, Christopher Bergevin<sup>b,c</sup>, Joseph Sollini<sup>a,3</sup>, Heather A. Kreft<sup>d,e</sup>, Alan R. Palmer<sup>a</sup>, Andrew J. Oxenham<sup>d,e</sup>, and Christopher A. Shera<sup>f,g</sup>

<sup>a</sup>Medical Research Council Institute of Hearing Research, School of Medicine, The University of Nottingham, NG7 2RD Nottingham, United Kingdom; <sup>b</sup>Department of Physics & Astronomy, York University, Toronto, ON M3J 1P3, Canada; <sup>c</sup>Centre for Vision Research, York University, Toronto, ON M3J 1P3, Canada; <sup>d</sup>Department of Psychology, University of Minnesota, Minneapolis, MN 55455; <sup>e</sup>Department of Otolaryngology, University of Minnesota, Minneapolis, MN 55455; <sup>f</sup>Caruso Department of Otolaryngology, University of Southern California, Los Angeles, CA 90033; and <sup>g</sup>Department of Physics and Astronomy, University of Southern California, Los Angeles, CA 90089

Edited by Dale Purves, Duke University, Durham, NC, and approved September 17, 2018 (received for review June 26, 2018)

**Frequency analysis of sound by the cochlea is the most fundamental property of the auditory system. Despite its importance, the resolution of this frequency analysis in humans remains controversial. The controversy persists because the methods used to estimate tuning in humans are indirect and have not all been independently validated in other species. Some data suggest that human cochlear tuning is considerably sharper than that of laboratory animals, while others suggest little or no difference between species. We show here in a single species (ferret) that behavioral estimates of tuning bandwidths obtained using perceptual masking methods, and objective estimates obtained using otoacoustic emissions, both also employed in humans, agree closely with direct physiological measurements from single auditory-nerve fibers. Combined with human behavioral data, this outcome indicates that the frequency analysis performed by the human cochlea is of significantly higher resolution than found in common laboratory animals. This finding raises important questions about the evolutionary origins of human cochlear tuning, its role in the emergence of speech communication, and the mechanisms underlying our ability to separate and process natural sounds in complex acoustic environments.**

cochlear tuning | frequency selectivity | auditory nerve | psychoacoustics | otoacoustic emissions

The cochlea within the inner ear acts like an acoustic prism to decompose sound into its constituent frequency components, creating a frequency-to-place map along its length. This decomposition establishes the tonotopic encoding of sound frequency that remains a fundamental organizing principle of the auditory system from the cochlea to the auditory cortex (1–4). The resolution with which the cochlea performs this frequency analysis influences our ability to perceptually separate different sounds and to communicate in complex acoustic environments. The loss of cochlear frequency resolution, through damage or disease, underlies some of the most troublesome problems associated with hearing impairment, including difficulty understanding speech in noise (5).

For many years, a consensus existed that cochlear tuning was similar across a wide range of mammalian species, including humans. That conclusion was based on the relatively good correspondence between indirect behavioral estimates of human tuning (6, 7) and direct measures of cochlear tuning taken from the auditory nerve of smaller laboratory animals (8, 9). Very few physiological human data existed, and those that did were not sufficient in number or did not deviate sufficiently from animal data to suggest any fundamental differences between species (10). However, more recent studies have suggested that human cochlear tuning may be sharper, by a factor of two or more, than cochlear tuning in typical laboratory animals, such as cat and guinea pig. The latest estimates from humans combined more refined behavioral measures and new noninvasive objective measures based on otoacoustic emissions (OAEs)—sounds that

are emitted by the cochlea and can be recorded in the ear canal (11).

Knowledge of any interspecies differences in the frequency resolution of the cochlea is critical to our understanding of a diverse range of issues (12). For example, the claimed disparities in estimates between animal and human tuning are sufficiently large to substantially affect the neural coding and representation of speech and other critical natural sounds (13–15). Quantification of species differences is also important for understanding the mechanisms underlying frequency analysis. For instance, it has been claimed that the cortical representation of frequency results from neural sharpening by the central auditory system from a less sharply tuned representation in the cochlea (16). This claim hinges critically on the assumption that human cochlear tuning is similar to that of small mammals.

In large part, claims of sharper tuning in the human cochlea remain controversial (17–19) because of a lack of commensurate

## Significance

Sound consists of a dynamic stream of energy at different frequencies. Auditory processing of sound frequency is critical in determining our ability to interact and communicate in a complex acoustic world, yet fundamental gaps remain in our understanding of how this is achieved. Indeed, the resolving power of the system, how best to measure it, and the mechanisms that underlie it are all still debated. Here, we provide critical evidence demonstrating that humans can resolve the frequency components of competing sounds better than other commonly studied mammals. This finding raises important questions both for theories of auditory perception and for our understanding of the evolutionary relationships between the auditory system and acoustic communication, including speech.

Author contributions: C.J.S., A.R.P., A.J.O., and C.A.S. designed research; C.J.S., T.T.W., C.B., J.S., H.A.K., A.R.P., and A.J.O. performed research; C.J.S., T.T.W., C.B., J.S., H.A.K., A.J.O., and C.A.S. analyzed data; and C.J.S., C.B., J.S., A.R.P., A.J.O., and C.A.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Data deposition: The data have been deposited in the Nottingham Research Data Management Repository (<http://dx.doi.org/10.17639/nottingham.373>).

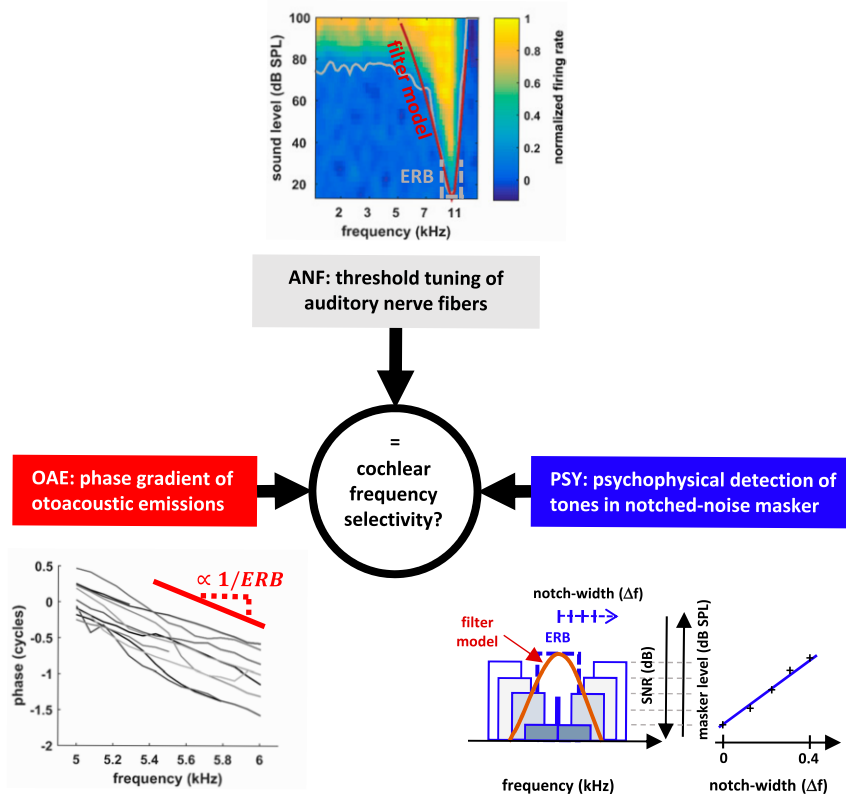
<sup>1</sup>To whom correspondence should be addressed. Email: [christian.sumner2@nottingham.ac.uk](mailto:christian.sumner2@nottingham.ac.uk).

<sup>2</sup>Present address: Department of Psychology, Nottingham Trent University, NG1 4FQ Nottingham, United Kingdom.

<sup>3</sup>Present address: University College London Ear Institute, WC1X 8EE London, United Kingdom.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1810766115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1810766115/-DCSupplemental).

Published online October 15, 2018.



**Fig. 1.** Three different ways of estimating cochlear tuning used in ferrets. ANFs, threshold levels (gray line) for a response are fit with a filter model (red line), from which the ERB (dashed gray line) is calculated. OAEs, the mean phase gradient of OAEs (red line) is used to estimate filter sharpness,  $Q_{\text{ERB}}$  ( $=f/\text{ERB}$ ), using the approximate species invariance of the tuning ratio. PSY, the behavioral detection of a pure tone in the presence of two bands of noise, separated by varying spectral distances. ERB (blue dashed line) is estimated by fitting a filter model to the detection thresholds. ANF data from ref. 27.

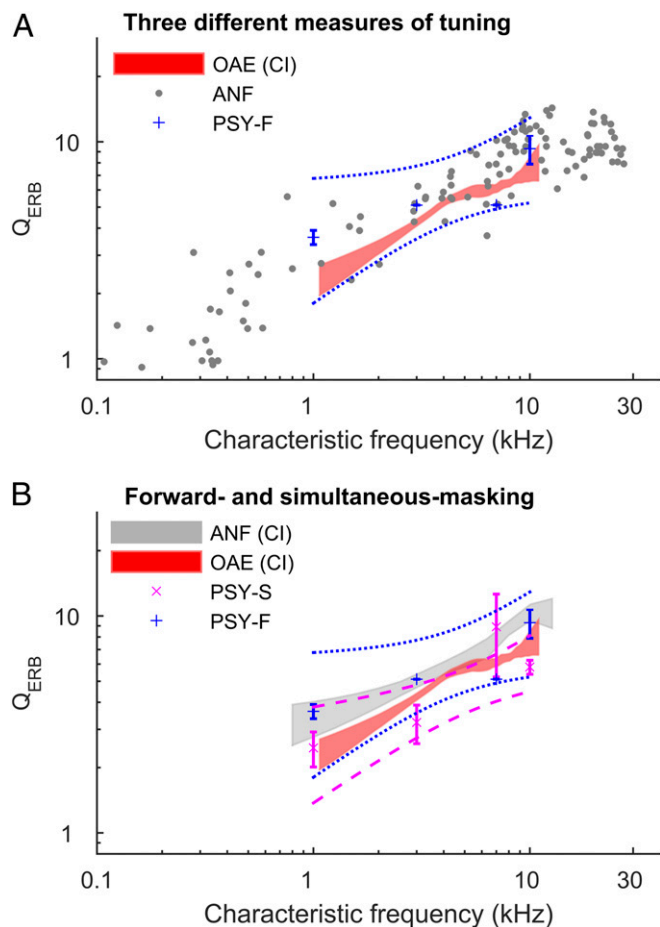
measures across species. Direct measures of tuning from single-unit recordings in the auditory nerve [auditory nerve fiber (ANF) in Fig. 1] have been obtained in laboratory animals but are too invasive to be performed in humans. Conversely, the more recent psychophysical methods (PSY) used in humans, involving the masking of a probe tone by spectrally notched noise under PSY forward masking (PSY-F; Fig. 1), have not yet been tested in animals. Estimates based on OAE measurements have been obtained in both humans and smaller mammals and are consistent with the claim of sharper tuning in humans (11, 18). However, uncertainty surrounding the mechanisms by which OAEs are generated, and their relationship to cochlear tuning, leave room for doubt (20, 21). In summary, three types of measure have been used to estimate cochlear tuning—behavioral, otoacoustic, and neural—but have never all been measured and compared in the same species. To resolve this problem, we used ferrets to examine all three measures within the same species. We reasoned that if the two indirect measures (OAE and PSY) provide accurate estimates of cochlear tuning, then they should both agree with the direct neural (ANF) measures. By employing all three methods in the same species, our experiments provide the strongest test to date of the validity of the indirect measures used to assess cochlear frequency tuning in humans.

## Results

We estimated ferret frequency tuning perceptually using a psychophysical notched-noise masking paradigm (Fig. 1, PSY; *SI Appendix, Fig. S1*). This paradigm measures the effectiveness of noises with various spectral shapes at masking a narrowband signal, such as a pure tone. By varying the frequency extent of a spectral notch in the masking noise, the shape and bandwidth of

the effective auditory filter can be derived (*SI Appendix, Experimental Methods*). We applied this method in ferrets performing behavioral detection tasks, and from the results derived the equivalent rectangular bandwidths (ERBs) of the filters, along with a corresponding dimensionless measure of tuning sharpness,  $Q_{\text{ERB}}$  (center frequency/ERB). For any filter shape, the ERB is the bandwidth of the rectangular filter with the same peak height that passes the same total power.

Because of cochlear nonlinearities, the exact stimulus conditions employed can influence the measured bandwidths. These include whether the masking noise is presented simultaneously with the signal (PSY simultaneous, PSY-S) or directly precedes the signal (PSY-F), thereby avoiding physical interactions between the stimuli within the cochlea (22–24). The estimated bandwidths can also depend on whether the intensity of the signal is kept constant and the threshold is found by varying the intensity of the masker, or vice versa. We estimated filter bandwidths in ferrets using all of these variants. Consistent with results in humans (22–24), we observed that PSY-F produces significantly sharper estimates of tuning than simultaneous masking [ $Q_{\text{ERB}}(\text{PSY-S}) = 0.72 \times Q_{\text{ERB}}(\text{PSY-F})$ ;  $P = 0.04$ ; see Figs. 2B and 3 and *SI Appendix, Fig. S3*]. We found no significant effect of whether thresholds are derived by varying the level of the masker or target tone ( $P = 0.2$ ), contrary to expectations (19, 25, 26). The absence of a significant effect may be partly due to our use of low stimulus levels (<40 dB sound pressure level), which are generally below the onset level of the compressive cochlear nonlinearity in ferrets (27), and partly due to the relatively small number of estimates in each condition ( $n = 5$  for the fixed signal and  $n = 3$  for the fixed masker), providing limited statistical power to detect a difference. Therefore, we only



**Fig. 2.** Three measures of frequency selectivity agree. (A) Filter sharpness from PSY-F agrees closely with ANF and OAE measurements. Tuning in individual nerve fibers (gray points), psychophysical forward masking (blue points), and a loess trend and its bootstrapped 95% CI for the otoacoustic emissions measurements. Dashed lines indicate bootstrapped 95% CIs for the data. (B) Forward masking (PSY-F; blue points,  $n = 8$ ) yields a better match to auditory nerve tuning than simultaneous masking (PSY-S; magenta points,  $n = 22$ ). In B, auditory nerve data are shown as the area within the loess (SI Appendix, Experimental Methods) trend 95% CI. ANF data from ref. 27.

distinguish between forward and simultaneous masking in our further comparisons.

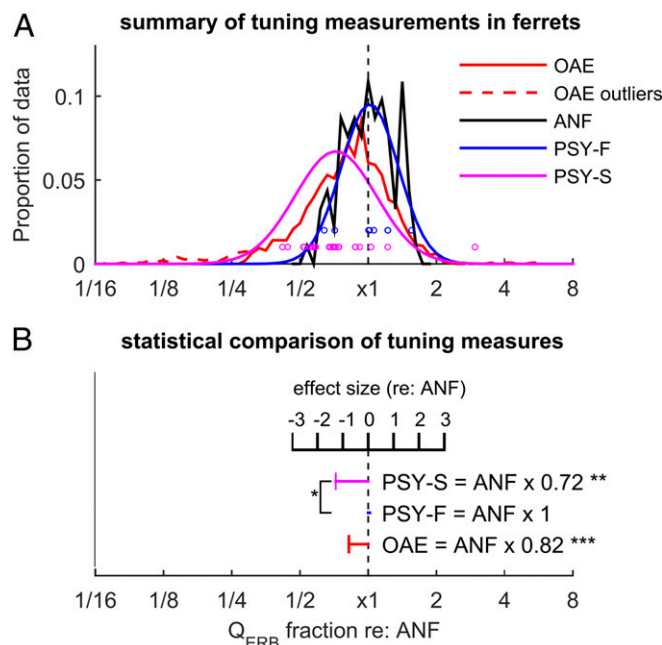
Next, we recorded stimulus-frequency otoacoustic emissions (SFOAEs) from the ears of sedated ferrets and inferred cochlear bandwidths using the emission group delay (Fig. 1; OAE, SI Appendix, Experimental Methods and Fig. S2). The OAE-based method estimates the sharpness ( $Q_{\text{ERB}}$ ) of the cochlear filters using the assumption of approximate species invariance of the “tuning ratio.” The tuning ratio is the empirical relationship between emission-delay and auditory nerve-fiber tuning trends obtained from independent measurements in other species. To estimate the ferret  $Q_{\text{ERB}}$  trend from the SFOAE delays, we followed Joris et al. (28) and used a tuning ratio obtained by averaging those previously derived for cats, guinea pigs, and chinchillas—species whose tuning ratios are all similar (18). Fig. 2A shows the trend of auditory filter sharpness inferred from the emission delays (data points are shown in SI Appendix, Fig. S2).

Finally, we compared the estimates from the two indirect measures with our previously published responses of single auditory-nerve fibers in anesthetized ferrets to short (50 ms) tone pips varying in frequency and sound level (27). The spike counts

in response to these tones allowed us to map out the receptive field of each fiber (Fig. 1; ANF) (i.e., the range of stimulus conditions over which the nerve fibers responded). From the lowest (threshold; Fig. 1, ANF; gray line) sound level that produced a response at each frequency, we modeled the shape of the auditory filter in each nerve fiber by fitting a rounded-exponential function (Fig. 1; ANF; brown line, ref. 29), and derived its  $Q_{\text{ERB}}$ , in the same manner as was done with the behavioral estimates.

Fig. 2 shows that all three measures of  $Q_{\text{ERB}}$ —those derived from auditory-nerve responses (ANF), OAEs, and PSY-F—are in good agreement. The agreement includes both the overall sharpness of tuning as well as its approximate power-law dependence on frequency. The agreement is especially remarkable given the very different natures of the three measures employed.

To compare the measurements quantitatively, we fitted the data (log-transformed frequency and  $Q_{\text{ERB}}$ ) with a linear model. With respect to overall tuning sharpness, the agreement among the different measures is most apparent when the data are expressed relative to the mean auditory-nerve tuning at the same frequency (i.e., residuals of the linear model; Fig. 3). Although the mean OAE-based estimates of  $Q_{\text{ERB}}$  are similar to those obtained directly from auditory-nerve tuning curves, their ratio is less than unity [ $Q_{\text{ERB}}(\text{OAE}) = 0.82 \times Q_{\text{ERB}}(\text{ANF})$ ; Fig. 3], and this difference is statistically significant (sandwich-test,  $P < 0.001$ ; see SI Appendix, Experimental Methods), in part due to the very large sample size of the OAE data ( $n \sim 1,500$ ). The difference in means implies that the tuning ratio in ferrets derived from these data are somewhat larger than the average of those previously obtained for cat, guinea pig, and chinchilla. For comparison, the variation among the tuning ratios for these three species is shown in figure 9B of ref. 16; the approximate



**Fig. 3.** Comparing different measures of frequency resolution in the ferret, independently of the effect of signal frequency. (A) The different tuning measurements as a fraction of the mean ANF tuning at a given frequency. Dashed red lines show excluded OAE outliers (see SI Appendix, Experimental Methods). (B) Statistical comparison of the different measures of tuning. Horizontal bars show the mean of each measure as a fraction of auditory nerve tuning and also as effect size (relative to ANF tuning). Asterisks next to data points indicate significant differences compared with auditory nerve tuning. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

“invariance” of the tuning ratio typically holds to within 5–15%, with the largest variations occurring in the apical regions of the cochlea.

Consistent with findings in humans, psychophysical estimates of tuning using simultaneous masking (PSY-S) are significantly broader [ $Q_{\text{ERB}}(\text{PSY-S}) = 0.72 \times Q_{\text{ERB}}(\text{PSY-F})$ ; Fig. 3] than the tuning estimates derived from both auditory-nerve fiber responses and OAEs (sandwich test,  $P < 0.01$ ; Cohen's  $d \sim 1$ ). To adapt the behavioral experiments to animal use, we necessarily modified some procedures used in previous human experiments. To explore the possible effects of these modifications, we tested a new set of human listeners using methods (stimuli and task) directly comparable to those used in our ferret experiments, with forward masking and a fixed target level (*SI Appendix, Experimental Methods*). The estimated  $Q_{\text{ERB}}$  at 4 kHz obtained using these ferret-based procedures with humans is similar to that found in earlier human studies (22) and is more than a factor of 2 sharper than the behavioral estimates from ferrets ( $P < 0.001$ ; *SI Appendix, Fig. S3*).

## Discussion

Disparate methods for measuring cochlea tuning were employed in a single animal model. Both psychophysical and otoacoustic methods provided reliable and quantitatively accurate estimates of cochlear frequency selectivity. These direct and indirect measures combined with new human behavioral data, collected using the same methods, provide strong support for the claim that frequency resolution is sharper in humans than in common laboratory mammals (summarized in *SI Appendix, Fig. S3*).

We attribute the close correspondence in tuning measures in large part to the refined methods employed in this study and their application within a single species. However, some modest discrepancies remain that are important to address. Tuning estimates obtained here using simultaneous masking are broader than those from ANF and forward-masked methods, consistent with studies in humans (22) and macaque (28, 30). However, other published data suggest either a closer correspondence of simultaneous masking and auditory nerve tuning (31) or even little difference compared with humans (32). Our data also fail to reveal the expected difference in frequency selectivity depending on whether the signal or masker were varied to determine thresholds (19, 25, 26). These inconsistencies may point to species differences other than tuning bandwidth, such as differences in the nature and extent of cochlear nonlinearities or cognition (33). However, the sizes of any differences are not large in comparison with the variability of the data (for example, of individual nerve fibers or of individual animals). A comprehensive assessment in nonhuman mammals of the effects of iso-level (fixed-masker) vs. iso-response (fixed-signal) measurements, forward vs. simultaneous masking, and overall sound level, with larger numbers of measurements, is required to resolve these issues.

The agreement of the three tuning measures provides compelling evidence that the limits of perceptual frequency resolution (as measured in our paradigm) are determined primarily in the cochlea, in contrast to previous suggestions (16). This conclusion therefore warrants a fresh evaluation of spectral decomposition in the central auditory system. In some cases, this agreement could obviate the need to postulate additional neural sharpening mechanisms, located between the cochlea and the cortex, to explain previously presumed discrepancies between sharp cortical tuning found in humans and the broad cochlear tuning found in laboratory animals (16) or from earlier estimates in humans using simultaneous masking (6). The tuning bandwidths estimated in human cortical neurons ( $\sim 1/12$  octave) are in fact remarkably similar to the estimates of human cochlear tuning that we have validated here ( $\sim 1/13$  octave, ref. 11), indicating that further central processing may not be necessary to account for narrow cortical tuning. Our results also provide data

to inform a classical debate in auditory neuroscience on whether the auditory system extracts spectral information from sounds in the form of a rate-place code or a code based on spike timing information, or a combination of the two (34). Proposals involving timing codes have been partly motivated by the poor rate-place coding found in animal studies (13, 14). Indeed, ferret cochlear bandwidths are barely sufficient to resolve adjacent formants [e.g., in the 2- to 3-kHz region the second and third formants can be around 1/3 octave apart (35), close to the bandwidth of ferret auditory filters in this region]. According to the narrower human bandwidths validated here, however, rate-place coding schemes would have considerably more success at representing the formant peaks of human speech in the human auditory system than in other species.

Although we have confirmed sharp human cochlear tuning using low-intensity sounds similar to those used to measure auditory-nerve tuning curves in other species, tuning is known to change with sound intensity, becoming broader at high intensities. Behavioral measures in humans have also revealed broader tuning at high sound intensities (36), in line with expectations. In addition, the saturation of firing rate in the auditory nerve at higher intensities also leads to effectively broader tuning and poorer resolution in the majority of ANFs at sound levels where human speech recognition remains robust (13). It is possible that tuning under more complex acoustic conditions is sharpened by central auditory processing, beyond what can be explained by firing rate in the auditory nerve, especially at high levels. Such sharpening might occur through mechanisms involving stimulus-driven spike timing, or phase locking, and lateral inhibition based on the rapid phase transitions produced by the basilar-membrane traveling wave (37). The extent to which putative sharpening mechanisms are required to explain behavioral performance at high sound intensities remains to be explored in light of our understanding of human cochlear tuning at low intensities.

It is tempting to relate sharp human cochlear tuning to our ability to perceive the subtleties of speech (particularly those involving prosody and pitch) in complex backgrounds, and thus our ability to solve the “cocktail party problem” (38). However, there is evidence for intermediate cochlear tuning in nonhuman primates (28), and one study reported cortical tuning in a nonhuman primate that approached that observed in humans (39). In addition, studies of otoacoustic emissions in another large mammal—the tiger—have also suggested that tuning may approach that found in humans (40). These findings imply that the physical size of the cochlea and its associated tonotopic map play a more important role than any human-specific evolution of cochlear tuning (41). Although sharp cochlear tuning may not be a sufficient condition for the emergence of speech as an effective communication mode (42), it may nevertheless have played an important and perhaps necessary role in its development. Given the complexity of this and the other issues discussed, the development of cochlear models that produce realistic sharp tuning and the nonlinear characteristics that impart dependence on stimulus paradigms, will provide an important step toward evaluating such claims and consolidating our understanding of frequency selectivity, the cochlea, and their relation to perception.

## Experimental Methods

Full details of experimental methods are given in *SI Appendix*. Briefly, we trained ferrets to detect (43) or lateralize (44) brief tones or narrowband noise, in the presence of masking noise, in a positive reinforcement procedure. Using these behavioral methods in ferrets, we measured perceptual thresholds using different variants of notched-noise maskers (6, 22). We also made measurements using similar stimulus paradigms in humans. We also recorded, in lightly anesthetized ferrets, the otoacoustic emissions elicited by pure-tone stimuli, using the SFOAE method (45). Estimates of frequency selectivity derived from these data were compared with previous recordings from the auditory nerve of anesthetized ferrets (27). In the human studies, all participants provided

written informed consent before participating, and all procedures were approved by the Institutional Review Board of the University of Minnesota. All procedures with ferrets were carried out under license from the UK Home Office, in accordance with the Animals (Scientific Procedures) Act 1986.

**ACKNOWLEDGMENTS.** We thank technicians at Nottingham for assistance with collecting behavioral data, Angie Killoran for formatting assistance,

and Brian Moore and an additional reviewer for thorough and constructive comments. Experimental work in ferrets was supported by Medical Research Council intramural funding MC\_UU\_00010/1 and U135097127. Behavioral experiments in humans were supported by NIH Grant R01DC012262 (to A.J.O. and H.A.K.). C.A.S. was supported by NIH Grant R01 DC003687. C.B. was supported by Natural Sciences and Engineering Research Council of Canada Grant RGPIN-430761-2013.

1. Merzenich MM, Roth GL, Anderson RA, Knight PL, Colwell SA (1977) Some basic features of the organization of the central auditory nervous system. *Psychophysics and Physiology of Hearing*, eds Evans EF, Wilson JP (Academic, London), pp 485–496.
2. Malmierca MS, et al. (2008) A discontinuous tonotopic organization in the inferior colliculus of the rat. *J Neurosci* 28:4767–4776.
3. Langers DR, van Dijk P (2012) Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. *Cereb Cortex* 22:2024–2038.
4. Shera CA (2015) The spiral staircase: Tonotopic microstructure and cochlear tuning. *J Neurosci* 35:4683–4690.
5. Oxenham AJ (2018) How we hear: The perception and neural coding of sound. *Annu Rev Psychol* 69:27–50.
6. Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. *Hear Res* 47:103–138.
7. Zwicker E (1961) Subdivision of the audible frequency range into critical bands. *J Acoust Soc Am* 33:248.
8. Rhode WS (1971) Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *J Acoust Soc Am* 49:1218.
9. Liberman MC (1978) Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am* 63:442–455.
10. Harrison RV, Aran JM, Erre JP (1981) AP tuning curves from normal and pathological human and guinea pig cochleas. *J Acoust Soc Am* 69:1374–1385.
11. Shera CA, Guinan JJ, Jr, Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc Natl Acad Sci USA* 99:3318–3323.
12. Bergevin C, Manley GA, Köppl C (2015) Salient features of otoacoustic emissions are common across tetrapod groups and suggest shared properties of generation mechanisms. *Proc Natl Acad Sci USA* 112:3362–3367.
13. Sachs MB, Young ED (1979) Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate. *J Acoust Soc Am* 66:470–479.
14. Delgutte B (1984) Speech coding in the auditory nerve: II. Processing schemes for vowel-like sounds. *J Acoust Soc Am* 75:879–886.
15. Baer T, Moore BCJ (1994) Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J Acoust Soc Am* 95:2277–2280.
16. Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I (2008) Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature* 451:197–201.
17. Ruggero MA, Temchin AN (2005) Unexceptional sharpness of frequency tuning in the human cochlea. *Proc Natl Acad Sci USA* 102:18614–18619.
18. Shera CA, Guinan JJ, Jr, Oxenham AJ (2010) Otoacoustic estimation of cochlear tuning: Validation in the chinchilla. *J Assoc Res Otolaryngol* 11:343–365.
19. Lopez-Poveda EA, Eustaquio-Martin A (2013) On the controversy about the sharpness of human cochlear tuning. *J Assoc Res Otolaryngol* 14:673–686.
20. Siegel JH, et al. (2005) Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering. *J Acoust Soc Am* 118:2434–2443.
21. Charaziak KK, Siegel JH (2015) Tuning of SFOEs evoked by low-frequency tones is not compatible with localized emission generation. *J Assoc Res Otolaryngol* 16:317–329.
22. Oxenham AJ, Shera CA (2003) Estimates of human cochlear tuning at low levels using forward and simultaneous masking. *J Assoc Res Otolaryngol* 4:541–554.
23. Moore BCJ, Glasberg BR (1981) Auditory filter shapes derived in simultaneous and forward masking. *J Acoust Soc Am* 70:1003–1014.
24. Houtgast T (1974) Lateral suppression in hearing. PhD thesis (Univ of Amsterdam, Amsterdam).
25. Glasberg BR, Moore BCJ (2000) Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise. *J Acoust Soc Am* 108:2318–2328.
26. Rosen S, Baker RJ, Darling A (1998) Auditory filter nonlinearity at 2 kHz in normal hearing listeners. *J Acoust Soc Am* 103:2539–2550.
27. Sumner CJ, Palmer AR (2012) Auditory nerve fibre responses in the ferret. *Eur J Neurosci* 36:2428–2439.
28. Joris PX, et al. (2011) Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans. *Proc Natl Acad Sci USA* 108:17516–17520.
29. Patterson RD, Nimmo-Smith I, Weber DL, Milroy R (1982) The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. *J Acoust Soc Am* 72:1788–1803.
30. Burton JA, Dylla ME, Ramachandran R (2018) Frequency selectivity in macaque monkeys measured using a notched-noise method. *Hear Res* 357:73–80.
31. Evans EF (2001) Latest comparisons between physiological and behavioural frequency selectivity. *Physiological and Psychological Bases of Auditory Function*, eds Breebaart DJ, Houtsma AJM, Kohlrausch AG, Pries VF, Schoonhoven R (Shaker Publishing, Maastricht, The Netherlands), pp 382–387.
32. May BJ, Kimar S, Prosen CA (2006) Auditory filter shapes of CBA/CaJ mice: Behavioral assessments. *J Acoust Soc Am* 120:321–330.
33. Yost WA, Shofner WP (2009) Critical bands and critical ratios in animal psychoacoustics: An example using chinchilla data. *J Acoust Soc Am* 125:315–323.
34. Young ED, Sachs MB (1979) Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. *J Acoust Soc Am* 66:1381–1403.
35. Catford JC (1988) *A Practical Introduction to Phonetics* (Oxford Univ Press, Oxford).
36. Nelson DA, Freyman RL (1984) Broadened forward-masked tuning curves from intense masking tones: Delay-time and probe-level manipulations. *J Acoust Soc Am* 75:1570–1577.
37. Shamma SA (1985) Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve. *J Acoust Soc Am* 78:1612–1621.
38. Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* 25:975–979.
39. Bartlett EL, Sadagopan S, Wang X (2011) Fine frequency tuning in monkey auditory cortex and thalamus. *J Neurophysiol* 106:849–859.
40. Bergevin C, Walsh EJ, McGee J, Shera CA (2012) Probing cochlear tuning and tonotopy in the tiger using otoacoustic emissions. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 198:617–624.
41. Shera CA, Charaziak KK (July 23, 2018) Cochlear frequency tuning and otoacoustic emissions. *Cold Spring Harb Perspect Med*, 10.1101/cshperspect.a033498.
42. Fitch WT (2000) The evolution of speech: A comparative review. *Trends Cogn Sci* 4:258–267.
43. Alves-Pinto A, Sollini J, Sumner CJ (2012) Signal detection in animal psychoacoustics: Analysis and simulation of sensory and decision-related influences. *Neuroscience* 220:215–227.
44. Sollini J, Alves-Pinto A, Sumner CJ (2016) Relating approach-to-target and detection tasks in animal psychoacoustics. *Behav Neurosci* 130:393–405.
45. Shera CA, Bergevin C (2012) Obtaining reliable phase-gradient delays from otoacoustic emission data. *J Acoust Soc Am* 132:927–943.