Altered ultrasonic vocalization in mice with a disruption in the Foxp2 gene

Weiguo Shua,b, Julie Y. Cho,b,c,d, Yuhui Jiangd,e, Minhua Zhangd,e, Donald Weiszf,g,h, Gregory A. Elderd, James Schmeiderd,j, Rita De Gasperid, Miguel A. Gama Sosad, Donald Rabidoul, Anthony C. Santuccij, Daniel Perl d,j,k, Edward Morrisseyd,j and Joseph D. Buxbaumt,d,e,m,n

*Molecular Cardiology Research Center, Department of Medicine, University of Pennsylvania Medical Center, 956 Biomedical Research Building II/III, Philadelphia, PA 19104; 1Laboratory of Molecular Neuropsychiatry and Departments of Psychiatry, Neuroscience, Neurosurgery, and Neurology, Otalaryngology, Pathology, Geriatrics and Adult Development, and *Biomathematical Sciences, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029; and†Department of Psychology, Manhattanville College, 2900 Purchase Street, Purchase, NY 10577

Communicated by Peter Palese, Mount Sinai School of Medicine, New York, NY, May 5, 2005 (received for review October 13, 2004)

Neurobiology of speech and language has previously been studied in the KE family, in which half of the members have severe impairment in both speech and language. The gene responsible for the phenotype was mapped to chromosome 7q31 and identified as the FOX2 gene, coding for a transcription factor containing a polyglutamine tract and a forkhead DNA-binding domain. Because of linkage studies implicating 7q31 in autism, where language impairment is a component of the disorder, and in specific language impairment, FOX2 has also been considered as a potential susceptibility locus for the language deficits in autism and/or specific language impairment. In this study, we characterized mice with a disruption in the murine Foxp2 gene. Disruption of both copies of the Foxp2 gene caused severe motor impairment, premature death, and an absence of ultrasonic vocalizations that are elicited when pups are removed from their mothers. Disruption of a single copy of the gene led to modest developmental delay but a significant alteration in ultrasonic vocalization in response to such separation. Learning and memory appear normal in the heterozygous animals. Cerebellar abnormalities were observed in mice with disruptions in Foxp2, with Purkinje cells particularly affected. Our findings support a role for Foxp2 in cerebellar development and in a developmental process that subsumes social communication functions in diverse organisms.

A n important genetic locus associated with normal development of both speech and language was identified by studies in the KE family, in which half of the members have severe abnormalities in these faculties (1). The core deficits of the disorder in KE are still unresolved, but affected members have been cited in various studies as having impaired ability to perform the coordinated movements that are required for speech (verbal dyspraxia), an impairment of speech and verbal comprehension (dysphasia) due to the inability to use the rules of grammar to denote tense, number, etc., a deficit in phonological and language-production systems, or a severe speech disorder interfering with all aspects of language (2).

In the study of affected members of the KE family, the phenotype was clearly segregating as an apparently autosomal dominant trait. The gene responsible for the phenotype was mapped to chromosome 7q31 (3) and identified as the FOX2 gene (4). The FOX2 gene codes for a transcription factor containing a polyglutamine tract and a forkhead DNA-binding domain (4). A missense mutation in the forkhead domain was found in the affected members (4).

Interestingly, Foxp2 is expressed at high levels during vocal learning in zebra finches in the striatal nucleus Area X, an area necessary for vocal learning (5). Moreover, in canaries, Foxp2 expression in Area X varies seasonally, with more expression observed at periods when the song becomes unstable (5). The expression pattern of Foxp2 in humans and songbirds is quite similar (6). The expression of Foxp2 in songbirds supports a role for Foxp2 in communication in multiple species.

Ignoring the polyglutamine tract, the human FOXP2 protein differs at only 3 amino acids with its homolog in mouse (7). The expression of the human and mouse homologs is very similar, both during development and in adulthood (2, 8–10). This finding supports the potential relevance of a mouse model to the study of the development of the neurobiological substrates for communication.

Here, we show that disruption of both copies of the Foxp2 gene in mice causes severe motor impairment, premature death, and an absence of ultrasonic vocalization in response to stressors. Even disruption of a single copy of the gene causes modest developmental delay and a significant alteration in ultrasonic vocalization. Cerebellar abnormalities are also observed, even with disruption of a single copy of the gene. Our results suggest that Foxp2 plays an instrumental role in the developmental process that subsumes social communication functions in diverse organisms.

Methods

Disruption of Foxp2. All studies followed Institutional Animal Care and Use Committee-approved procedures. The Foxp2 targeting construct was generated to replace exon 12 and exon 13 with a neomycin cassette. Linearized target vector was electroporated into embryonic stem (ES) cells, and a positive/negative selection strategy by using neomycin and gancyclovir was used to facilitate the isolation of homologous recombinants. The results were verified by Southern blot analysis, and targeted ES clones were microinjected into C57BL/6 blastocysts. Chimeric male offspring were bred to C57BL/6 females, and agouti F1 offspring were tested for transmission of the disrupted allele by Southern blot analysis. Heterozygous matings of the F1 mice were then carried out to produce homozygous F2 mutant mice.

Behavior. Animals selected from 12 distinct litters were tested in behavioral analyses, with litter sizes ranging from 6 to 13 animals. Mice were not routinely culled, because this procedure has been challenged for its scientific value (4). At postnatal day 3, each pup was permanently marked for identification, and each day starting at 1000 hours, each litter was observed for nesting behavior. Each pup was separated from the litter and weighed in a disposable weighing boat before all testing. For clarity, the description of all of the studies carried out are divided into four experiments below.

Experiment 1. Testing was carried out in mice from six distinct litters immediately after observation of nesting behavior and weighing.

www.pnas.org/cgi/doi/10.1073/pnas.0503739102

PNAS | July 5, 2005 | vol. 102 | no. 27 | 9643–9648

© 2005 by The National Academy of Sciences of the USA

Freely available online through the PNAS open access option.

Abbreviation: EGL, external granular layer.

[1]W.S. and J.Y.C. contributed equally to this work.


[3]To whom correspondence should be addressed. E-mail: joseph.buxbaum@msm.edu.
Testing was carried out in 14 wild-type animals, 15 heterozygous animals, and 8 knockout animals. Testing consisted of rooting reflex, catalepsy, negative geotaxis, righting reflex, beaker dump, hanging, mid-air righting, and toe pinch in that order. Physical development such as opening of the eyes and extroversion of the ears was observed and noted. Testing began at postnatal day 6 and was conducted at 3-day intervals until postnatal day 15. For negative geotaxis, each pup was oriented downward on a downward-sloping mesh grid, and the time before the mouse reoriented itself in an upward facing position was recorded (to a maximum time of 180 sec). To measure emotionality, fear-induced freezing was measured after a traumatic stimulus. In this study, the pup was placed in a 100-ml beaker, and the beaker was inverted, dropping the pup on a padded beaker. The righting reflex was evaluated by turning the mouse onto its back, and the delay before it could turn itself over was noted (to a maximum time of 180 sec). Further, the pups were oriented belly up 30 cm above a padded surface and were dropped to observe whether they could right themselves before landing.

Experiment 2. Animals from experiment 1 were taken after behavioral testing on postnatal day 6 and were used for automated vocalization monitoring. Each pup was separated from the litter one at a time and placed in a shallow beaker. The pup was then positioned below a Mini5 bat detector set to detect at 40–70 Hz in a soundproof chamber maintained at 21 ± 0.5°C. Six minutes of recording commenced after the pup habituated to the chamber for 30 sec. ULTRAVOX 2.0 (Noldus Information Technology, Wageningen, The Netherlands) was used to count the ultrasonic vocalizations. Click-like sounds whose duration was <4 ms were filtered out.

Experiment 3. Additional vocalization measurements were made at postnatal day 10 with a fresh set of animals chosen from six distinct litters. Testing was carried out in 8 wild-type animals, 30 heterozygous animals, and 8 knockout animals. An Ultrasonic microphone type 2670 (Bruel & Kjaer Instruments, Marlborough, MA) was used in place of the bat detector. Animals were placed into a Petri dish that was contained within a ventilated styrofoam box (interior dimension of 14 cm × 12 cm × 12 cm). The microphone was placed ~6 cm above the Petri dish. The styrofoam box was placed inside of a custom-made sound-attenuating chamber (83 cm × 33 cm × 71 cm). After an initial 1-min habituation, vocalizations were recorded over a 6-min period. A Nexus conditioning amplifier (Bruel & Kjaer) amplified the vocalizations. Sound files were recorded by using SPIKE 2.4.18 (Cambridge Electronic Design, Cambridge, U.K.) at a sampling frequency of 333 kHz. Recorded waves were filtered (high pass, 10,000 Hz with a 3,000 Hz transition gap) before analysis. The duration of each call was measured between two movable cursors at the beginning and the end of the call. The power spectrum was obtained for each call by using a Fast Fourier Transformation (with 256 bins). Bandwidth was measured between two movable cursors at the beginning and the end of the spectrum. Peak frequency (the strongest sound frequency) was measured by the highest peak of the spectrum. Twenty successive clicks of each pup and up to 50 successive whistles, if the pup vocalized >50, were analyzed for duration, bandwidth, and peak frequency.

Experiment 4. Wild-type and heterozygous animals (n = 11) from experiments 1 and 3 were used to test for shock sensitivity as well as learning and memory at 3–4 months of age. Knockout animals did not survive to adulthood and, hence, were not included in these tests. Shock sensitivity was assessed by exposing animals to electric foot shock with a maximum intensity of 0.2 mA. Shock intensity was increased in gradations of 10% of maximum and responses (paw withdrawal, running, jumping, and vocalization) were noted. The Morris water maze place-learning paradigm was performed as described in ref. 12 by using a combined mass and spaced regime, such that on each day, each animal was given four training trials in a row, and this regiment continued for 4 days.
antiserum (1:500, Chemicon, Temucula, CA) and a rat monoclonal anti-GFAP antibody (1:500, the gift of Dr. Virginia Lee, University of Pennsylvania). DAPI Nuclear staining was performed by adding 1 μg/ml of 4',6-diamino 2-phenylindole dihydrochloride to the next-to-last wash solution. TUNEL staining was performed by using a commercially available kit (In Situ Cell Death Detection Kit, Roche Applied Sciences).

**Statistics.** There are well known issues regarding litter effects with multiparous species such as mice. To control for these issues, we used mice from at least six distinct litters in each experiment. In addition, we used mice of differing genotypes from within each litter, so that any comparison of genotypes was between mice of the same litter. Finally, to assess litter effects and litter by genotype interactions, we included multiple mice with the same combination of litter and genotype. If there was a significant litter by genotype interaction, this result was interpreted to indicate that the genotype differences in the litters were significantly discrepant from one another. Therefore, if there was a significant interaction, so that the litters did not exhibit a consistent genotype difference, statistical analysis was not performed for that outcome variable.

Some outcome variables differed substantially from a normal distribution by having outliers. For this reason, except for weight and mid-air righting, all variables were logarithmically transformed before analyses.

For each data set, a preliminary analysis was performed by using sex of the mice. In all cases, sex of the mice was not related to the outcomes as a main effect or an interaction with genotype or litter, so sex of the mice was not included as an additional covariate in any subsequent analyses.

Mice of each genotype were compared pairwise in all litters by analysis of covariance, with the litter as a covariate (because litter is a categorical rather than linear variable, dummy coding was used). Comparisons were made between wild-type and both heterozygous and knockout animals. For weight and for other longitudinal data as indicated, repeated measures analysis of covariance is reported, examining differences between groups in means averaged over all times and in linear trends. For weight, only repeated measures analyses were carried out.

For mid-air righting, an ordinal scale was used (a score of 1 reflecting no righting, 2 reflecting partial righting, and 3 reflecting complete righting).

**Results**

We disrupted the Foxp2 gene in mice by replacing exon 12 and 13 with a neomycin-resistance gene (Fig. 1a). Targeted disruption was determined by Southern analysis (Fig. 1b), and the loss of Foxp2 expression was demonstrated by Northern (data not shown) and immunoblot analysis (Fig. 1c). Offspring genotypes from heterozygous matings approached Mendelian frequencies, but mice with a disruption in the Foxp2 gene grew at slower rates (Fig. 2a).

Significant differences in mean weight of both heterozygous (P < 0.022) and knockout (P < 0.0005) animals was observed when compared with wild-type littermates. In addition, the rate of weight gain was significantly different for both heterozygous (P = 0.009) and knockout (P < 0.0005) animals when compared with wild-type littermates. The homozygous knockout animals showed severe motor abnormalities and typically died by postnatal day 21. Death did not appear to be due to the lack of maternal care or feeding. Pups showed normal nesting, huddling in clusters with no evidence that mice of a specific genotype were outside the cluster, and all pups also showed the presence of a milk band.
To define the phenotype of the Foxp2 knockouts, we carried out behavioral assessments in knockouts, heterozygotes, and wild-type littermates. Knockout animals showed a decrease in spontaneous activity (data not shown) and were delayed in righting ability when placed on their backs (Fig. 2b). Mean righting during a fall, a reflex that develops between day 6 and day 15 (Fig. 2c) in wild-type animals, was reduced in knockout animals when compared with wild-type animals ($P = 0.002$). In addition, the knockout animals typically failed to orient themselves up an inclined plane when placed facing down the plane (negative geotaxis) (Fig. 2d). Repeated measures analysis demonstrated differences in both mean time to reorient ($P < 0.0005$ for wild-type versus knockout and $P = 0.011$ for wild-type versus heterozygotes) and in the linear trend ($P = 0.014$ for wild-type versus knockout and $P = 0.027$ for wild-type versus heterozygotes). By day 9 and later, homozygous knockout animals performed indistinguishably from wild-type animals when time spent freezing after being dropped on a thick pad from 30 cm was measured, this test being taken as a simple measure of emotionality (data not shown). At day 6, knockouts appeared to freeze for extended periods when compared with controls. However, the interpretation of these results was confounded by the obvious motor deficits in the knockout mice and could not therefore be attributed to altered emotionality at this early date. The deficiencies in motor development were accompanied by a delay in eye opening and ear extroversion (Fig. 2e and f). In all of these studies, homozygous animals performed similarly to the wild-type littermates by day 15 or sooner.

Brain sections from the heterozygous and knockout mice were compared with those from the wild-type mice in a blinded manner by a neuropathologist with no data about the behavioral phenotype. No overt abnormalities were detected in the histologic appearance of the cerebral hemispheres and the subcortical structures, including the midbrain and pons. However, the knockout mouse demonstrated the presence of a 3- to 4-cell-thick external granular layer (EGL) at postnatal days 15–17, well after the normal resolution of the EGL (Fig. 3 a–f). The heterozygous animals retained a one-cell-thick EGL at this age, whereas the wild-type mice were free of this early developmental feature. By adulthood, the EGL was absent in heterozygous animals (data not shown).

The molecular layer in the knockout animals was approximately half the thickness of that of wild-type and heterozygous animals, suggesting that there might be abnormalities in the Purkinje cells. We therefore performed calbindin immunostaining on postnatal day 17 wild-type and mutant mice (Fig. 3 g–i). In contrast to the wild-type mice, calbindin-stained Purkinje cells in the mutant mice frequently failed to align in a continuous row, often forming a layer that was several cells thick with ectopically placed cells in the granule cell layer. In addition, the dendritic arbors in the mutant mice were less elaborate and often aligned at oblique angles compared with the wild-type animals, thus explaining the thinned molecular layer.

Granule cell progenitors in the EGL migrate to their final position in the granule cell layer along the radial fibers of the Bergmann glia (14). To determine whether the persistence of an EGL in the knockout animal might be at least, in part, explained by a failure of radial glial development, we stained radial glial fibers in cerebellar sections from postnatal day 17 animals with GFAP immunostaining (Fig. 3 j–l). Radial glial fibers were visible in all genotypes. However, in contrast to the regularly aligned radial glial fibers in the wild-type animals, fibers in the heterozygous animals were in some areas thinner and less well aligned. In the knockout animals, there were often gaps in the radial glial network as well as areas where fibers appeared to be clumped into aggregates.

In Weaver mice, failed migration of granule cell progenitors is associated with widespread apoptosis in the EGL in the first three postnatal weeks (15). To determine whether apoptosis was occurring in the EGL remnant in Foxp2 mutant mice, we performed TUNEL staining on postnatal day 17 cerebellum. Only rare TUNEL-positive cells were observed in either heterozygous or wild-type littermates by day 15 or sooner.
knockout animals (data not shown), indicating that apoptosis in the EGL is not characteristic of the Foxp2 mutant phenotype.

Because FOXP2 has been directly implicated in speech and articulation, we examined the incidence of ultrasonic vocalizations in pups removed from their mothers. Ultrasonic calls are important for mother–infant social interaction (16) and represent important markers for neurobehavioral development (17). At postnatal day 6, the incidence of vocalization over time was dramatically reduced in both heterozygous and knockout animals as measured by automated vocalization monitoring (Fig. 4a). Repeated measures analysis demonstrated differences in mean number of vocalizations ($P < 0.0005$ for wild-type versus knockout and $P = 0.008$ for wild-type versus heterozygotes).

Based on these results, we performed a spectrographic analysis of an independent group of animals at postnatal day 10 (e.g., Fig. 4b). There was a profound decrease in the number of ultrasonic vocalizations in heterozygous and homozygous knockout animals (Fig. 4c). The duration, peak frequency, and bandwidth of these vocalizations in the heterozygous animals were indistinguishable from wild-type animals (data not shown). In the course of these analyses, we also examined broad-spectrum clicks made by the mice. These clicks are of unknown function (18), and the information content of them has not been studied. Heterozygous and homozygous knockout animals were able to produce clicks, but the homozygous knockout animals produced clicks at a reduced incidence (Fig. 4d). The duration, peak frequency, and bandwidth of these vocalizations in the heterozygous and homozygous knockout animals were indistinguishable from wild-type animals.

The altered ultrasonic vocalizations appear to be specific but may also reflect a reduced response to stress. In an attempt to examine this possibility, sensitivity of animals to shock, as measured by two stereotyped behavioral responses, paw withdrawal, and running, were compared (Fig. 4e). No differences were observed in either sensitivity to shock by these measures or in frequency of sonic vocalizations. Although this finding may suggest that the heterozygous animals are not less sensitive to stress, it is also possible that these animals have a specific, dampened response to the stress of separation from the dam.

Finally, learning and memory were compared between wild-type and heterozygous animals (Fig. 4f). Acquisition was not significantly different between the two genotypes, reaching criteria (defined as a latency of $\leq 10$ sec on three of four consecutive trials) by the fourth day of testing. Furthermore, performance in a reversal task was also not different between the two genotypes.

**Discussion**

FOXP2 is expressed in multiple regions within the developing brain including the cortical plate, basal ganglia, thalamus, inferior olive, and cerebellum. The motor deficits observed in the Foxp2 knockout are consistent with a cerebellar abnormality, an abnormality supported by the observed alteration in cerebellar development. Based on our initial histological analyses, these findings suggest that an interference in Foxp2 function influences neuronal migration and/or maturation in the development of the cerebellum. Careful screening of other brain regions, including the basal ganglia, did not detect any obvious abnormalities. However, quantitative measures will be required to fully characterize any neuropathology in the heterozygous and knockout animals.

Our vocalization studies show that the frequency of occurrence of ultrasonic vocalizations is selectively impaired in the knockout and heterozygous mice. At the same time, analyses of the vocalization patterns and bandwidths suggest that the apparatus necessary for the production of vocalizations, including the neural control, in the vocal tract, and brainstem, is normal. There are subtle cerebellar deficits in the heterozygous animals during development, perhaps suggesting that the cerebellar abnormalities may also underlie the deficits in vocalization observed in these animals as speech is a motor activity. It was interesting that some

---

**Fig. 4.** Vocalization and behavior in mice with a disruption of the Foxp2 locus. (a) Number of ultrasonic vocalizations in 6 day old mice after separation from the dam. (b) Sample vocalization pattern showing both clicks (the two vertical traces covering both the sonic and ultrasonic region of the trace) and whistles (the four more vertical traces in the ultrasonic region of the trace). (c and d) Number of ultrasonic vocalizations (whistles) and sonic/ultrasonic vocalizations (clicks) in 10-day-old mice, after separation from the dam. (e) Shock sensitivity of adult mice. (f) Morris water maze performance of adult mice. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$. **
of the deficit in the knockouts appeared to be transient and reflected developmental delay.

There is increasing evidence for an autism susceptibility locus at or near the FOXP2 locus (19). Because of the impairment of language abilities in Foxp2 mutants, the FOXP2 gene has been considered as a potential autism susceptibility gene, specifically as associated with the language deficits in this disorder (20). The majority of association studies with FOXP2 and autism have been negative (21–23; N. Ramoz and J.D.B., unpublished data), with two positive associations observed in Chinese (24) and Japanese (25) subjects. The same region of chromosome 7 has also been implicated in specific language impairment (26, 27), although FOXP2 was not shown to be associated in one study (23). The FOXP2 gene is quite large (>600 Mb), and novel exons are still being identified (28), so FOXP2 cannot be considered excluded for either autism or specific language impairment. Although it is increasingly unlikely that FOXP2 is an autism susceptibility gene, understanding the genes and pathways that are regulated by FOXP2 may lead to candidate genes for autism and specific language impairment.

FOXP2/Foxp2 is expressed in the cerebellum of all species studied (8–10). Within the cerebellum, significant FOXP2 expression is observed in Purkinje cells and deep nuclei of the cerebellum. Within the cerebellar nuclei, expression of mouse Foxp2 and human FOXP2 are also similar during development, and there is intense expression in the alar plate during development in both the mouse and human fetal brains (8). It is therefore not surprising that there are cerebellar deficits in the heterozygous and homozygous Foxp2 knockout mice. Perhaps more surprising is the potentially associated deficits in vocalization in both heterozygous and homozygous Foxp2 knockout mice. Interestingly, voxel-based morphometry in affected KE family members looking for bilateral brain abnormalities has found reductions in gray matter in the ventral cerebellum (29). In addition, only deficits in nonword repetition and the volitional control of speech is similar to that of other motor functions and includes a critical frontocerebellar loop (2). Within this loop, abnormalities of this factor in neural development and in social communication.

Note Added in Proof. Since our submission, a significant finding on FOXP2 has been published. Specifically, FOXP2 coding exons were sequenced in 49 probands with a primary diagnoses of verbal dyspraxia, and three coding variants were identified. One variant was a nonsense mutation that cosegregates with speech and language difficulties in the proband, an affected sibling, and their mother. This paper is essential to understanding the role and functioning of FOXP2, with which our paper is chiefly concerned (30).

We thank Dr. Vahram Haroutunian for his advice and guidance. This work was supported by grants from the Seaver Autism Research Foundation (to J.D.B.), the National Institute of Health (MH066753 to J.D.B. and HD071589 to E.M.), and Cure Autism Now (to J.D.B.).