SUPPLEMENTARY MATERIAL

Methods

DNA and Protein Analyses

Blood samples were collected from all participants with FraX to confirm the genetic diagnosis and analyzed by Kimball Genetics (Denver, CO). The diagnosis of FraX was confirmed by standard DNA analysis and peripheral (lymphocyte) FMRP levels.[1, 2] All females with FraX carried the full mutation. Of the 45 males, 6 (13.3%) displayed a mosaic (full mutation/permutation) genotype.

Cognitive Testing

Children under the age of 4 years underwent the Mullen Scales of Early Learning[3] and subjects age 4 years and above underwent the age appropriate Wechsler Intelligence test.[4-6]

Behavioral Assessments

The presence and severity of autistic symptoms were measured for individuals with FraX by the Autism Behavior Checklist (ABC).[7] The ABC contains 57 items measuring a wide range of symptoms of autism within five domains: language, relating, body and object use, sensory stimulation, social and self-help. The total score of the ABC has demonstrated good internal consistency and high discriminant validity[7] and has also been used for measuring autistic symptoms in subjects with FraX.[8] Maladaptive behavioral symptoms commonly associated with developmental disability and mental retardation were measured by the Aberrant Behavior Checklist-Community (AbBC).[9, 10] The AbBC

resolves into five statistically derived subscales: (1) Irritability-Agitation, (2) Lethargy-Social Withdrawal, (3) Stereotypic Behavior, (4) Hyperactivity-Noncompliance, and (5) Inappropriate Speech. The subjects' parents completed both behavioral scales.

Volumetric Measures

The processing procedures included removal of non-brain tissues from the images, correction of equipment related image artifacts, and separation (segmentation) of tissue components (grey, white, cerebrospinal fluid (CSF)). Each brain was positionally normalized with a 6-parameter transformation to make it parallel to the plane defined by the anterior and the posterior commissures. Brain grey matter (GM) was subdivided into four cerebral lobes, subcortex and brainstem on the basis of a stereotaxic atlas template.[11] The cerebellum was manually segmented.

Additional regions of interest (ROIs) were obtained using previously described methods.[11-17] Raters were blind to subject diagnosis, gender and other subject characteristics; inter-rater reliabilities for volumes of manually drawn ROIs were above 0.90 as calculated by the intraclass correlation coefficient calculated on the same gold standard set of ROIs.

ROIs: the prefrontal cortex was defined as all frontal cortical GM lying anterior to a coronal plane intersecting the most anterior point of the genu of the corpus callosum.[12] The prefrontal cortex was than subdivided by three axial planes parallel to a line passing through the anterior and posterior commissure. Each axial plane was determined at the same location for each participant after fitting a proportional (Talairach) grid system to the normalized brain.[18]

Trained research assistants following detailed protocols delineated additional ROI volumes manually. The ROI variables included the volumes of the CN and lenticular nuclei,[13, 14] hippocampus, amygdala[11], thalamus[15], and cerebellar vermis[17] and the GM volume of the superior temporal gyrus.[16] The CN ROI was further divided into anterior (head) and posterior (body and tail) components by a coronal plane perpendicular to the AC-PC axis at the level of the anterior commissure.[13] The cerebellar vermis was divided on the midsagittal image to anterior (lobules I to V) and posterior (lobules VI to X) parts.[17]

QROC: For this analysis we entered the brain volumes (adjusted for differences in total brain tissue) that were significantly different between FraX and Controls as independent variables and the binary outcome measure as FraX vs. Controls. The QROC searched all independent variables, values representing adjusted brain volumes and their associated cutpoints, and identified the one with the optimal balance between sensitivity and specificity for distinguishing between FraX and controls (maximal weighted kappa, putting equal weight on false positives and false negatives). Once the optimal variable and associated cutpoint was identified, the cutpoint was tested using a 2x2 chi-square with 1% significance level. If that variable passed this "test", the sample was divided into two subgroups, above and below the cutpoint. The QROC analysis was then restarted, separately, for each of these two subgroups. The procedure was reiterated until a stopping rule activated, which occurred when: (1) the sample size was too small, i.e., fewer than 10 subjects in any of the cells of the 2x2 table or (2) $p > 0.01$.[19-22]

Voxel-Based Morphomtery (VBM)

Native space images were segmented into GM, white matter (WM) and CSF. GM images were spatially normalized to a pediatric GM template created from children aged 5-18 years (Cincinnati Children's Hospital Medical Center, CCHMC, ver2, 9/2002). Thus, for the VBM analysis we included only subjects between ages 5-18 years (FraX, N=62, controls, N=63); Stanford: FraX : controls = 40:33, Hopkins: 22:30. Spatial normalization was accomplished using a combined 12 parameter affine and non-linear transformation (7x8x7 non-linear basis functions, 16 iterations). GM normalization parameters were then applied to normalize the original native space coronal images. These optimally normalized images were then segmented into GM, WM and CSF. Next, the intensity of each GM image was modulated and was smoothed with a kernel of 12 and of 4mm full-width half-maximum. TBM was performed using the Jacobian determinant obtained from the spatial normalization parameters to determine if there were local shape differences between the diagnosis groups and whether these differences contributed to any GM volume differences. While we used a customized template made using the CCHMC data, coordinates of activation are reported in both MNI space and also in Talairach space (converted using the mni2tal function http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml) as a guide.

Surface-based Anatomical Modeling

Caudate ROI data were selected randomly from 30 subjects in each group for this analysis. The average age and gender distribution of these subgroups were well matched and comparable to the overall groups (mean ± standard deviation (SD), males: females,

FraX 13.7±2.8, 15:15; controls 12.1±3.7, 16:14). Anatomical mesh modeling methods matched equivalent surface points in the CN across subjects and groups.[23]

The manually derived contours were made uniform by modeling them as a 3-dimensional variable surface mesh allowing measurements to be made at corresponding surface locations in each subject. This procedure also allows the averaging of caudate surface morphological features across all individuals belonging to a group and records the amount of variation between corresponding surface points relative to the group averages. To measure local differences, a medial 3-dimensional curve was derived from each individual's caudate threading down the central axis. The distance of each surface point from this centerline measures the radial size of the caudate.[24]

Compatibility of MR Images across Sites

Though no specific inter-site data were collected due to the long-term nature of data acquisition for this study (1993-97 at Johns Hopkins and 1997-2003 at Stanford), both academic sites used the ACR MRI phantom and followed advanced quality control procedures as related to a number of performance factors including magnetic field homogeneity, geometric distortion, image artifact, spatial resolution, slice thickness and position, image contrast and RF coil performance (e.g., signal to noise ratio-SNR). Image quality was optimized for each subject using the auto-shimming methods of the MRI system. We also calculated SNR from 10 scans per site (randomly selected by an investigator blind to the results). SNR was sampled from these scans twice and averaged.[25] No significant differences in SNR were observed across sites ($p=0.49$ and 0.45).

Volumetric

In addition to the analyses in which data were aggregated across sites (Table 2), separate (post hoc) analyses were performed for each scan site for the main variables of interest (caudate, amygdala, STG GM and PV). ANOVA models identical to those employed in the cross-site analyses were utilized. Caudate volume was significantly enlarged in the FraX group at each site (*p's* < 0.0001) and the Group x Gender interaction also was significant (*p's* < 0.05). PV size was reduced in the FraX group at each site (*p's* < 0.0001). STG GM volume was reduced in FraX as well. However, consistent with the finding that between-group STG differences across sites was not as robust as that observed for the caudate and PV, differences were marginally significant at individual sites (*p*=0.06 at Stanford and 0.04 at Hopkins). Finally, amygdala volume was reduced in the FraX group at both sites. However, this difference was significant for subjects at the Hopkins site only (*p*=0.002; Stanford site *p*=0.27). Differences in amygdala volumes across scan sites is likely to be associated with the findings that (1) the size of this structure is known to be sexually dimorphic (M>F)[26] and (2) the fact that the Hopkins site had a significantly larger male to female ratio relative to Stanford. It is important to note that in no case did the Group x Site interaction significantly contribute to any of the final statistical models listed in Table 2 – indicating that scan site did not influence brain volumes differentially in the two groups.

VBM

As a first step in assessing image comparability across the two institutions, we regressed out scan site (Hopkins/Stanford) from the analysis and observed similar results. We also analyzed the data for each site separately and found similar results (left CN: Stanford Z=4.65, Hopkins Z=4.75, right CN: Stanford Z=4.55, Hopkins Z=4.88, right fusiform gyrus: Stanford Z=3.54, Hopkins Z=3.79)(see Main Text Table 3).

References

1. Oberle I, Rousseau F, Heitz D et al. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science. 1991;252:1097-1102
2. Willemsen R, Oosterwijk JC, Los FJ et al. Prenatal diagnosis of fragile X syndrome. Lancet. 1996;348:967-968
3. Mullen E. Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Service Inc., 1995
4. Wechsler D. Wechsler Adult Intelligence Scale - Third Edition. Administration and Scoring Manual. San Antonio, TX: The Psychological Corporation, 1997
5. Wechsler D. Wechsler Intelligence Scale for Children - Third Edition Manual. San Antonio, TX: The Psychological Corporation, 1991
6. Wechsler D. Manual for the Wechsler Preschool and Primary Scale of Intelligence-Revised. New York, NY: Psychological Corp, 1990
7. Wadden NP, Bryson SE, Rodger RS. A closer look at the Autism Behavior Checklist: discriminant validity and factor structure. J Autism Dev Disord. 1991;21:529-541
8. Hagerman RJ, Chudley AE, Knoll JH et al. Autism in fragile X females. Am J Med Genet. 1986;23:375-380
9. Aman MG, Singh NN, Stewart AW, Field CJ. Psychometric characteristics of the aberrant behavior checklist. Am J Ment Defic. 1985;89:492-502
10. Baumgardner TL, Reiss AL, Freund LS, Abrams MT. Specification of the neurobehavioral phenotype in males with fragile X syndrome. Pediatrics. 1995;95:744-752

11. Kates WR, Abrams MT, Kaufmann WE et al. Reliability and validity of MRI measurement of the amygdala and hippocampus in children with fragile X syndrome. Psychiatry Res. 1997;75:31-48

12. Schoenemann PT, Sheehan MJ, Glotzer LD. Prefrontal white matter volume is disproportionately larger in humans than in other primates. Nat Neurosci. 2005;8:242-252

13. Eliez S, Barnea-Goraly N, Schmitt JE et al. Increased basal ganglia volumes in velo-cardio-facial syndrome (deletion 22q11.2). Biol Psychiatry. 2002;52:68-70

14. Reiss AL, Abrams MT, Greenlaw R et al. Neurodevelopmental effects of the FMR-1 full mutation in humans. Nat Med. 1995;1:159-167

15. Chang K, Karchemskiy A, Barnea-Goraly N et al. Reduced amygdalar gray matter volume in familial pediatric bipolar disorder. J Am Acad Child Adolesc Psychiatry. 2005;44:565-573

16. Kesler SR, Blasey CM, Brown WE et al. Effects of X-monosomy and X-linked imprinting on superior temporal gyrus morphology in Turner syndrome. Biol Psychiatry. 2003;54:636-646

17. Mostofsky SH, Mazzocco MM, Aakalu G et al. Decreased cerebellar posterior vermis size in fragile X syndrome: correlation with neurocognitive performance. Neurology. 1998;50:121-130

18. Reiss AL, Eckert MA, Rose FE et al. An experiment of nature: brain anatomy parallels cognition and behavior in Williams syndrome. J Neurosci. 2004;24:5009-5015

19. Kiernan M, Kraemer HC, Winkleby MA et al. Do logistic regression and signal detection identify different subgroups at risk? Implications for the design of tailored interventions. Psychol Methods. 2001;6:35-48

20. Kraemer HC. Evaluating medical tests: objective and quantitative guidelines. Newbury Park, CA: Sage Publications, 1992

21. O'Hara R, Thompson JM, Kraemer HC et al. Which Alzheimer patients are at risk for rapid cognitive decline? J Geriatr Psychiatry Neurol. 2002;15:233-238

22. Salmond CH, Ashburner J, Vargha-Khadem F et al. Distributional assumptions in voxel-based morphometry. Neuroimage. 2002;17:1027-1030

23. Thompson PM, Hayashi KM, De Zubicaray GI et al. Mapping hippocampal and ventricular change in Alzheimer disease. Neuroimage. 2004;22:1754-1766

24. Frisoni GB, Sabattoli F, Lee AD et al. In vivo neuropathology of the hippocampal formation in AD: A radial mapping MR-based study. Neuroimage. 2006

25. Henkelman RM. Measurement of signal intensities in the presence of noise in MR images. Med Phys. 1985;12:232-233

26. Durston S, Hulshoff Pol HE, Casey BJ et al. Anatomical MRI of the developing human brain: what have we learned? J Am Acad Child Adolesc Psychiatry. 2001;40:1012-1020