

Toward discovery science of human brain function

Bharat B. Biswal^a, Maarten Mennes^b, Xi-Nian Zuo^b, Suril Gohel^a, Clare Kelly^b, Steve M. Smith^c, Christian F. Beckmann^c, Jonathan S. Adelstein^b, Randy L. Buckner^d, Stan Colcombe^e, Anne-Marie Dogonowski^f, Monique Ernst^g, Damien Fair^h, Michelle Hampsonⁱ, Matthew J. Hoptman^j, James S. Hyde^k, Vesa J. Kiviniemi^l, Rolf Kötter^m, Shi-Jiang Liⁿ, Ching-Po Lin^o, Mark J. Lowe^p, Clare Mackay^c, David J. Madden^q, Kristoffer H. Madsen^r, Daniel S. Margulies^r, Helen S. Mayberg^s, Katie McMahon^t, Christopher S. Monk^u, Stewart H. Mostofsky^v, Bonnie J. Nagel^w, James J. Pekar^x, Scott J. Peltier^y, Steven E. Petersen^z, Valentin Riedl^{aa}, Serge A. R. B. Rombouts^{bb}, Bart Rypma^{cc}, Bradley L. Schlaggar^{dd}, Sein Schmidt^{ee}, Rachael D. Seidler^{ff,uu}, Greg J. Siegle^{gg}, Christian Sorg^{hh}, Gao-Jun Tengⁱⁱ, Juha Veijola^{jj}, Arno Villringer^{ee,kk}, Martin Walter^{ll}, Lihong Wang^q, Xu-Chu Weng^{mm}, Susan Whitfield-Gabrieliⁿⁿ, Peter Williamson^{oo}, Christian Windischberger^{pp}, Yu-Feng Zang^{qq}, Hong-Ying Zhangⁱⁱ, F. Xavier Castellanos^{bj}, and Michael P. Milham^{b,1}

^aDepartment of Radiology, New Jersey Medical School, Newark, NJ 07103; ^bPhyllis Green and Randolph Cöwen Institute for Pediatric Neuroscience, New York University Child Study Center, NYU Langone Medical Center, New York, NY 10016; ^cFMRI Centre, Oxford University, Oxford OX3 9DU, UK; ^dHoward Hughes Medical Institute, Harvard University, Cambridge, MA 02138; ^eSchool of Psychology, University of Wales, Bangor, UK; ^fDanish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark; ^gMood and Anxiety Disorders Program, National Institute of Mental Health/ National Institutes of Health, Department of Health and Human Services, Bethesda, MD 20892; ^hBehavioral Neuroscience Department, Oregon Health & Science University, Portland, OR 97239; ⁱDepartment of Diagnostic Radiology, Yale University School of Medicine, New Haven, CT 06511; ^jDivision of Clinical Research, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY 10962; ^kBiophysics Research Institute, Medical College of Wisconsin, Milwaukee, WI 53226; ^lDepartment of Diagnostic Radiology, Oulu University Hospital, Oulu, Finland; ^mDonders Institute for Brain, Cognition, and Behavior, Center for Neuroscience, Radboud University Nijmegen Medical Center, 6500 HB Nijmegen, The Netherlands; ⁿBiophysics Research Institute, Medical College of Wisconsin, Milwaukee, WI 53226; ^oInstitute of Neuroscience, National Yang-Ming University, Taiwan; ^pImaging Institute, The Cleveland Clinic, Cleveland, OH 44195; ^qBrain Imaging and Analysis Center, Duke University Medical Center, Durham, NC, 27710; ^rDepartment of Cognitive Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; ^sDepartment of Psychiatry and Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322; ^tCentre for Advanced Imaging, University of Queensland, Brisbane, Australia; ^uDepartment of Psychology, University of Michigan, Ann Arbor, MI 48109; ^vLaboratory for Neurocognitive and Imaging Research, Kennedy Krieger Institute, Baltimore, MD, 21205; ^wDepartment of Psychiatry, Oregon Health & Science University, Portland, OR 97239; ^xF.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD 21205; ^yFunctional MRI Laboratory, University of Michigan, Ann Arbor, MI 48109; ^zMcDonnell Center for Higher Brain Functions, Washington University School of Medicine, St. Louis, MO 63110; ^{aa}Departments of Neurology and Neuroradiology, Klinikum Rechts der Isar, Technische Universität München, 81675 Munich, Germany; ^{bb}Institute of Psychology and Department of Radiology, Leiden University Medical Center, Leiden University, Leiden, The Netherlands; ^{cc}Center for Brain Health and School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX 75080; ^{dd}Department of Neurology, Washington University School of Medicine, St. Louis, MO 63110; ^{ee}Department of Neurology, Charité Universitätsmedizin-Berlin, 10117 Berlin, Germany; ^{ff}School of Kinesiology, University of Michigan, Ann Arbor, MI 48109; ^{gg}Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15213; ^{hh}Department of Psychiatry, Klinikum Rechts der Isar, Technische Universität München, D-81675 Munich, Germany; ⁱⁱJiangsu Key Laboratory of Molecular and Functional Imaging, Department of Radiology, Zhong-Da Hospital, Southeast University, Nanjing 210009, China; ^{jj}Department of Psychiatry, Institute of Clinical Medicine and Department of Public Health Science, Institute of Health Science, University of Oulu, Oulu 90014, Finland; ^{kk}Berlin Neuroimaging Center, 10099 Berlin, Germany; ^{ll}Department of Psychiatry, Otto-von-Guericke University of Magdeburg, Magdeburg 39106, Germany; ^{mm}Laboratory for Higher Brain Function, Institute of Psychology, Chinese Academy of Sciences, Beijing 100864, China; ⁿⁿDepartment of Brain and Cognitive Sciences, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Boston, MA 02139; ^{oo}Department of Psychiatry, University of Western Ontario, London, ON N6A3H8, Canada; ^{pp}Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria; and ^{qq}State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China

Edited* by Marcus E. Raichle, Washington University, St. Louis, MO, and approved January 20, 2010 (received for review October 14, 2009)

Although it is being successfully implemented for exploration of the genome, discovery science has eluded the functional neuroimaging community. The core challenge remains the development of common paradigms for interrogating the myriad functional systems in the brain without the constraints of a priori hypotheses. Resting-state functional MRI (R-fMRI) constitutes a candidate approach capable of addressing this challenge. Imaging the brain during rest reveals large-amplitude spontaneous low-frequency (<0.1 Hz) fluctuations in the fMRI signal that are temporally correlated across functionally related areas. Referred to as functional connectivity, these correlations yield detailed maps of complex neural systems, collectively constituting an individual's "functional connectome." Reproducibility across datasets and individuals suggests the functional connectome has a common architecture, yet each individual's functional connectome exhibits unique features, with stable, meaningful interindividual differences in connectivity patterns and strengths. Comprehensive mapping of the functional connectome, and its subsequent exploitation to discern genetic influences and brain-behavior relationships, will require multicenter collaborative datasets. Here we initiate this endeavor by gathering R-fMRI data from 1,414 volunteers collected independently at 35 international centers. We demonstrate a universal architecture of positive and negative functional connections, as well as consistent loci of inter-individual variability. Age and sex emerged as significant determinants. These results demonstrate that independent R-fMRI datasets can be aggregated and shared. High-throughput R-fMRI can provide quantitative phenotypes for molecular genetic studies and biomarkers of developmental and

pathological processes in the brain. To initiate discovery science of brain function, the 1000 Functional Connectomes Project dataset is freely accessible at www.nitrc.org/projects/fcon_1000/.

database | neuroimaging | open access | reproducibility | resting state

Much like the challenge of decoding the human genome, the complexities of mapping human brain function pose a challenge to the functional neuroimaging community. As dem-

Author contributions: B.B.B., R.L.B., J.S.H., R.K., A.V., Y.Z., F.X.C., and M.P.M. designed research; B.B.B., M.M., X.N.Z., S.G., C.K., S.M.S., C.F.B., J.S.A., R.L.B., S.C., A.-M.D., M.E., D.F., M.H., M.J.H., J.S.H., V.J.K., R.K., S.J.L., C.P.L., M.J.L., C.E.M., D.M., K.H.M., D.S.M., H.S.M., K.M., C.S.M., S.M., B.J.N., J.J.P., S.J.P., S.E.P., V.R., S.A.R., B.R., B.L.S., S.S., R.D.S., G.S., C.S., G.J.T., J.M.V., A.V., M.W., L.W., X.C.W., S.W.-G., P.W., C.W., Y.Z., H.Y.Z., B.B.B., F.X.C., and M.P.M. performed research; S.M.S., C.F.B., R.L.B., S.C., A.-M.D., M.E., D.F., M.H., M.J.H., J.S.H., V.J.K., R.K., S.J.L., C.P.L., M.J.L., C.E.M., D.M., K.H.M., D.S.M., H.S.M., K.M., C.S.M., S.M., B.J.N., J.J.P., S.J.P., S.E.P., V.R., S.A.R., B.R., B.L.S., S.S., R.D.S., G.S., C.S., G.J.T., J.M.V., A.V., M.W., L.W., X.C.W., S.W.-G., P.W., C.W., Y.Z., H.Y.Z., B.B.B., F.X.C., and M.P.M. contributed new reagents/analytic tools; B.B.B., M.M., X.N.Z., S.G., C.K., F.X.C., and M.P.M. analyzed data; and B.B.B., M.M., X.N.Z., C.K., J.S.A., F.X.C., and M.P.M. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

Data deposition: All data used in this work were released on December 11, 2009 via www.nitrc.org/projects/fcon_1000/.

¹To whom correspondence should be addressed. E-mail: michael.milham@nyumc.org.

This article contains supporting information online at www.pnas.org/cgi/content/full/0911855107/DCSupplemental.

Results

We applied three distinct analytic methods commonly used in the R-fMRI literature: seed-based functional connectivity, independent component analysis (ICA), and frequency-domain analyses. Across the three approaches, we found evidence of (i) a universal intrinsic functional architecture in the human brain, (ii) center-related variation in R-fMRI measures, and (iii) consistent effects of age and sex on R-fMRI measures, detectable across centers despite the presence of center-related variability (Fig. 1). Specifically, seed-based correlational analyses revealed highly consistent patterns of functional connectivity across centers for both the “default mode” (30) and “task-positive” networks (31), supporting a universal functional architecture (Fig. S1). Similarly, a data-driven, temporal concatenation ICA approach, combined with dual regression (32–34), revealed consistent patterns of functional connectivity across centers for 20 spatially independent functional networks (Fig. 1 and Figs. S2 and S3). In addition, for each of the functional connectivity measures, within-center coefficient of variation maps showed a high degree of concordance across centers (Fig. S4). This suggests that common loci of variation exist: centers demonstrated a high degree of agreement on which connections are characterized by relative variance or invariance. Despite the high degree of concordance between centers, there were appreciable center-related variations in the strength of functional connectivity throughout the brain (8). The effect of center was especially prominent in regions exhibiting greater interregional connection strength, because these have the least within-center variability (See *SI Results* and Fig. S5 for further discussion of center-related variability.) However, even when taking this center-related variability into account, robustly reliable effects of age and sex remained appreciable (Fig. 2 and Figs. S1 and S2). (See *SI Results* and Fig. S6 for an examination of the impact of sample size on effects of age and sex.)

The detection of sex differences was particularly noteworthy, because these differences are rarely appreciated in the R-fMRI

literature (35). Sexual dimorphism in human genomic expression (36) is known to affect numerous physiological variables that can influence the fMRI signal (37, 38). For example, males and females differ in terms of hemoglobin concentrations and hematocrit (39). However, global variables such as these do not explain the regionally specific sex-related phenomenon noted in the present work. Hormonal effects (e.g., estrogen), operating both during brain development (40) and acutely (41), are known to have regional specificity (42), making them potential contributors to the differences observed. Given the discovery nature of the present work and the lack of prior coordination among centers, the specific sex differences that we observed should be interpreted with caution until replicated in an independent sample.

Along with examining patterns of functional connectivity, we measured the amplitude of low-frequency fluctuations at each voxel using two common periodogram-based measures: amplitude of low frequency fluctuation (ALFF; total power <0.1 Hz) (2, 17, 43) and fractional ALFF (fALFF; total power <0.1 Hz/total power in the measured spectrum) (44). Concordant with previous work, the dominance of low-frequency fluctuations was consistently noted within gray matter regions, but not white matter (44). As with our analyses of functional connectivity, despite clear evidence of center-related effects, we were again able to demonstrate age- and sex-related differences in the magnitude of low-frequency fluctuations in various regions, particularly medial wall structures (Fig. 2 and Fig. S7).

Beyond data pooling for statistical analyses, we demonstrate the potential to use high-throughput datasets to develop normative maps of functional systems in the brain, which is a prerequisite for clinical applications. Specifically, we exploit a key property of functional connectivity maps, the presence of well-differentiated borders between functionally distinct regions (45). The voxelwise measures of coefficients of variation for each type of functional connectivity map delineate putative functional boundaries based on the presence of marked variability in func-

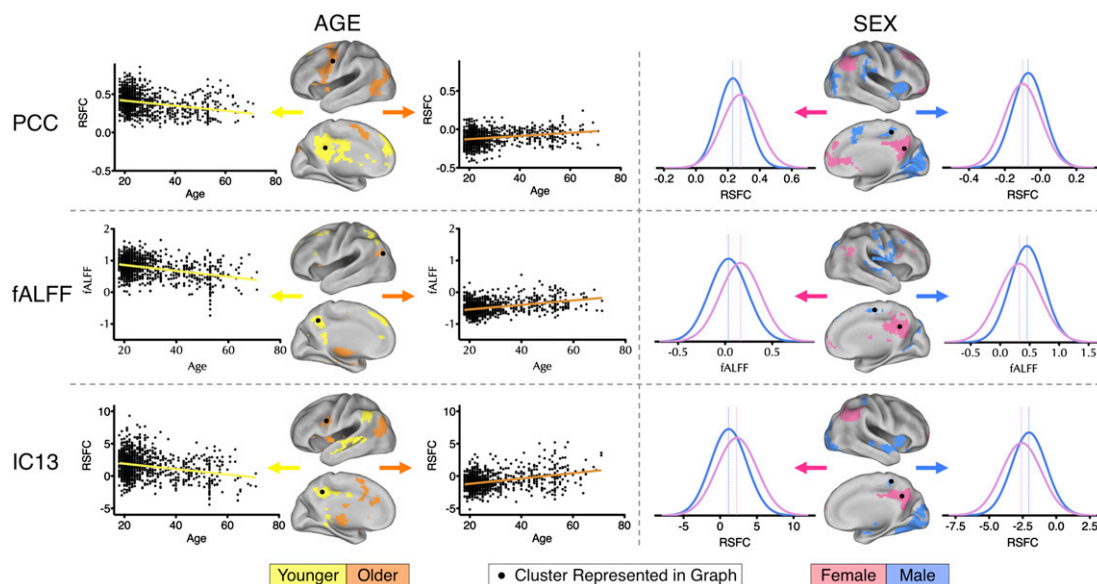


Fig. 2. Illustrative areas exhibiting age- and sex-related variation in R-fMRI properties. Significant group-level variance in functional connectivity maps was explained by age and sex (cluster-based Gaussian random-field corrected: $Z > 2.3$; $P < 0.05$). For each of three methods (seed-based, fALFF, and ICA), variance in connectivity strength explained by age (*Left*) and sex (*Right*) is illustrated both anatomically and graphically. Age-related differences are represented as scatterplots. Sex-related differences are represented as histograms depicting the distributions of resting-state functional connectivity (RSFC) values for males and females separately. Vertical lines indicate peak values. Corresponding topographical brain areas are indicated with dots. “Male” refers to significantly greater connectivity (or amplitude, i.e., fALFF) in males; similarly, “female” refers to significantly greater connectivity (or amplitude) in females. “Older” refers to significantly increasing connectivity (or amplitude) with increasing age, whereas “younger” refers to significantly increasing connectivity (or amplitude) with decreasing age.

the origins and biological significance of spontaneous low-frequency fluctuations of neuronal and hemodynamic activity, the impact of intrinsic activity on evoked responses (and vice versa), and the ideal means of acquiring, processing, and analyzing R-fMRI data. Nevertheless, the potential of discovery science is vast, from the development of objective measures of brain functional integrity to help guide clinical diagnoses and decision-making, to tracking treatment response and assessing the efficacy of treatment interventions. Finally, whereas the present work examines functional connectivity alone, future studies may combine R-fMRI with other modalities (e.g., EEG, magnetoencephalography, diffusion-tensor imaging, volumetrics) and genetics to achieve a more complete understanding of the human brain.

All data and analytic tools used in the present work will be made available at www.nitrc.org/projects/fcon_1000/. We anticipate that the open availability of the 1000 Functional Connectomes dataset will recruit the broad participation and collaboration among the scientific community necessary for successful implementation of discovery-based science of human brain function. In addition, we hope that it will further advance the ethos of data sharing and collaboration initiated by such efforts as fMRIDC (www.fmriddc.org), FBIRN (www.birncommunity.org), OASIS (www.oasis-brains.org), BrainScape (www.brainscape.org), and BrainMap (www.brainmap.org).

Methods

Resting-state fMRI scans were aggregated from 35 community-based datasets ($n = 1,414$). The present analysis was restricted to 24 centers ($n = 1,093$; 21 published, 3 unpublished; mean age <60 years; only participants over age 18; one scan per participant; duration: 2.2–20 min; $n = 970$ at 3 T, $n = 123$ at 1.5 T; voxel size, 1.5–5 mm within plane; slice thickness, 3–8 mm). Each contributor's respective ethics committee approved submission of deidentified data. The institutional review boards of NYU Langone Medical Center and New Jersey Medical School approved the receipt and dissemination of the data.

For functional connectivity, we used seed-based correlation analysis, based on six previously identified seed regions (31), and model-free ICA, using temporal

concatenation to generate group-level components and dual regression to generate individual participant maps. For amplitude measures at each voxel, we used the FFT-based ALFF (2, 17, 43) and its normalized variant, fALFF (44).

Standard image preprocessing was performed (i.e., motion correction, spatial filtering with FWHM = 6 mm, 12-dof affine transformation to MNI152 stereotactic space). For seed-based correlation approaches and dual regression following ICA analysis, nuisance signals (e.g., global signal, WM, CSF, motion parameters) were regressed out. Temporal filtering was tailored for each analytic approach (29, 31, 32, 44).

ICA components for dual regression analyses were determined by (i) low-dimensional (20 components) temporal concatenation ICA carried out 25 times (each with 18 participants randomly selected from each of 17 centers with minimum of 165 time points) and (ii) low-dimensional (20 components) meta-ICA, a second concatenation-based ICA using the component sets produced by the 25 runs (see [SI Results](#) for a description of an alternative method). For each participant, dual regression (32–34) was performed using the 20 components identified by the meta-ICA ([Fig. S3](#)), yielding a connectivity map for each component.

Aggregate statistical analyses of center, sex, and age effects were based on a generalized linear model implementation of one-way ANOVA (factor: center; covariates: age and sex). To identify functional boundaries, we calculated voxelwise coefficients of variation across all 1,093 participants, and ranked each voxel based on the absolute value of its coefficient of variation.

ACKNOWLEDGMENTS. We thank David Kennedy and www.nitrc.org for supporting the 1000 Functional Connectomes Project data release, Avi Snyder for providing helpful insights and advice concerning project goals, and Cameron Craddock for helpful advice on this study. Financial support for the 1000 Functional Connectomes project was provided by grants from the National Institutes of Mental Health (R01MH083246 and R01MH081218 to F.X.C. and M.P.M.), National Institute on Drug Abuse (R03DA024775, to C.K.; R01DA016979, to F.X.C.), Autism Speaks, National Institute of Neurological Disorders and Stroke (R01NS049176, to B.B.), and the Howard Hughes Medical Institute (to J.S.A. and R.L.B.), as well as gifts to the NYU Child Study Center from the Stavros Niarchos Foundation, Leon Levy Foundation, Joseph P. Healy, Linda and Richard Schaps, and Jill and Bob Smith and an endowment provided by Phyllis Green and Randolph Cowen. NITRC is funded by the National Institutes of Health's Blueprint for Neurosciences Research (neuroscienceblueprint.nih.gov) (Contract N02-EB-6-4281, to TCG, Inc.).

- Wise J (2008) Consortium hopes to sequence genome of 1000 volunteers. *BMJ* 336: 237.
- Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34: 537–541.
- Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700–711.
- Margulies DS, et al. (2007) Mapping the functional connectivity of anterior cingulate cortex. *Neuroimage* 37:579–588.
- Smith SM, et al. (2009) Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci USA* 106:13040–13045.
- Van Dijk KR, et al. (2010) Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization. *J Neurophysiol* 103:297–321.
- Sporrs O, Tononi G, Kötter R (2005) The human connectome: A structural description of the human brain. *PLOS Comput Biol* 1:e42.
- Shehzad Z, et al. (2009) The resting brain: Unconstrained yet reliable. *Cereb Cortex* 19: 2209–2229.
- Zuo XN, et al. (2010) Reliable intrinsic connectivity networks: test-retest evaluation using ICA and dual regression approach. *Neuroimage* 49:2163–2177.
- Zuo XN, et al. (2010) The oscillating brain: Complex and reliable. *Neuroimage* 49: 1432–1445.
- He Y, et al. (2009) Uncovering intrinsic modular organization of spontaneous brain activity in humans. *PLoS One* 4:e5226.
- Buckner RL, et al. (2009) Cortical hubs revealed by intrinsic functional connectivity: Mapping, assessment of stability, and relation to Alzheimer's disease. *J Neurosci* 29: 1860–1873.
- Mormann AM, Fletcher PC (2007) Does the brain have a baseline? Why we should be resisting a rest. *Neuroimage* 37:1073–1082.
- Damoiseaux JS, et al. (2006) Consistent resting-state networks across healthy subjects. *Proc Natl Acad Sci USA* 103:13848–13853.
- Castellanos FX, et al. (2008) Cingulate precuneus interactions: A new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biol Psychiatry* 63: 332–337.
- Church JA, et al. (2009) Control networks in paediatric Tourette syndrome show immature and anomalous patterns of functional connectivity. *Brain* 132:225–238.
- Zang YF, et al. (2007) Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev* 29:83–91.
- Rombouts SA, et al. (2009) Model-free group analysis shows altered BOLD FMRI networks in dementia. *Hum Brain Mapp* 30:256–266.
- Kelly AMC, et al. (2009) Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cereb Cortex* 19:640–657.
- Supekar K, Musen M, Menon V (2009) Development of large-scale functional brain networks in children. *PLoS Biol* 7:e1000157.
- Fair DA, et al. (2009) Functional brain networks develop from a "local to distributed" organization. *PLoS Comput Biol* 5:e1000381.
- Andrews-Hanna JR, et al. (2007) Disruption of large-scale brain systems in advanced aging. *Neuron* 56:924–935.
- Hagmann P, et al. (2008) Mapping the structural core of human cerebral cortex. *PLoS Biol* 6:e159.
- Vincent JL, et al. (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447:83–86.
- Margulies DS, et al. (2009) Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc Natl Acad Sci USA* 106:20069–20074.
- Fukunaga M, et al. (2006) Large-amplitude, spatially correlated fluctuations in BOLD fMRI signals during extended rest and early sleep stages. *Magn Reson Imaging* 24: 979–992.
- Greicius MD, et al. (2008) Persistent default-mode network connectivity during light sedation. *Hum Brain Mapp* 29:839–847.
- Achard S, Bullmore E (2007) Efficiency and cost of economical brain functional networks. *PLoS Comput Biol* 3:e17.
- Kelly C, et al. (2009) L-dopa modulates functional connectivity in striatal cognitive and motor networks: A double-blind, placebo-controlled study. *J Neurosci* 29:7364–7378.
- Gusnard DA, Raichle ME, Raichle ME (2001) Searching for a baseline: Functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685–694.
- Fox MD, et al. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
- Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. *Philos Trans R Soc Lond B Biol Sci* 360:1001–1013.
- Filippini N, et al. (2009) Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. *Proc Natl Acad Sci USA* 106:7209–7214.
- Calhoun VD, Pekar JJ, Pearson GD (2004) Alcohol intoxication effects on simulated driving: Exploring alcohol-dose effects on brain activation using functional MRI. *Neuropsychopharmacology* 29:2997–3017.
- Liu H, Stufflebeam SM, Sepulcre J, Hedden T, Buckner RL (2009) Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. *Proc Natl Acad Sci USA* 106:20499–20503.
- Ellegren H, Parsch J (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* 8:689–698.

37. Goldstein JM, et al. (2001) Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb Cortex* 11:490–497.
38. Gur RC, et al. (1995) Sex differences in regional cerebral glucose metabolism during a resting state. *Science* 267:528–531.
39. Zeng SM, Yankowitz J, Widness JA, Strauss RG (2001) Etiology of differences in hematocrit between males and females: Sequence-based polymorphisms in erythropoietin and its receptor. *J Gend Specif Med* 4:35–40.
40. Lenroot RK, et al. (2007) Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage* 36:1065–1073.
41. Protopopescu X, et al. (2008) Hippocampal structural changes across the menstrual cycle. *Hippocampus* 18:985–988.
42. McEwen BS (1999) Permanence of brain sex differences and structural plasticity of the adult brain. *Proc Natl Acad Sci USA* 96:7128–7130.
43. Cordes D, et al. (2001) Frequencies contributing to functional connectivity in the cerebral cortex in “resting-state” data. *AJNR Am J Neuroradiol* 22:1326–1333.
44. Zou QH, et al. (2008) An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *J Neurosci Methods* 172:137–141.
45. Cohen AL, et al. (2008) Defining functional areas in individual human brains using resting functional connectivity MRI. *Neuroimage* 41:45–57.
46. Roy AK, et al. (2009) Functional connectivity of the human amygdala using resting-state fMRI. *Neuroimage* 45:614–626.
47. Albert NB, Robertson EM, Miall RC (2009) The resting human brain and motor learning. *Curr Biol* 19:1023–1027.
48. Harrison BJ, et al. (2008) Modulation of brain resting-state networks by sad mood induction. *PLoS ONE* 3:e1794.
49. Lettre G (2009) Genetic regulation of adult stature. *Curr Opin Pediatr* 21:515–522.
50. Manolio TA, et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753.

Supporting Information

Biswal et al. 10.1073/pnas.0911855107

SI Methods

Image Preprocessing. Overview. All available resting-state scans were preprocessed using both AFNI (1) and FSL (www.fmrib.ox.ac.uk). Specific commands can be found in the preprocessing scripts that will be released at www.nitrc.org/projects/fcon_1000/ on publication of this paper. After the first five time points of every scan were discarded, to remove possible T1 stabilization effects, the data were corrected for motion by aligning each volume to the mean image volume using Fourier interpolation in AFNI. Then the data were spatially smoothed using a 6-mm FWHM Gaussian kernel. Mean-based intensity normalization was done by scaling all volumes by the same factor (10,000).

Seed-based correlation analyses. The data were temporally filtered using both a high-pass (Gaussian-weighted least squares straight-line fitting, with $\sigma = 100.0$ s) and low-pass (Gaussian low-pass temporal filtering, with a HWHM of 2.8 s) filter, followed by linear detrending to remove any residual drift.

Independent component analysis. Temporal concatenation group analysis. Consistent with common practice, temporal filtering for ICA analyses was limited to high-pass filtering (Gaussian-weighted least squares straight-line fitting, with $\sigma = 100.0$ s).

Dual regression. This step used the same preprocessed data as used in the seed-based correlation analyses.

ALFF/fALFF. No temporal filtering was carried out, because the data were examined in the frequency domain within select bands (2, 3). Temporal despiking with a hyperbolic tangent squashing function was performed, however, to limit extreme values. Linear trends were then removed from the data.

Registration and normalization. After the skull was removed using AFNI, registration of each individual's high-resolution anatomic image to a common stereotactic space [the Montreal Neurological Institute's 152-brain template (MNI152); 3 mm isotropic voxel size] was done using a 12-degrees of freedom linear affine transformation (FLIRT) (4, 5). The resulting transformation was then applied to each individual's functional dataset. We did not further optimize the normalization with a nonlinear algorithm, because of concerns about image quality and limited coverage in some datasets.

Functional Connectivity: Seed-Based Correlation Analysis. Nuisance signal regression. Consistent with common practice in the fMRI literature, nuisance signals were removed from the data via multiple regression before functional connectivity analyses were performed. This step is designed to control for the effects of physiological processes, such as fluctuations related to motion and cardiac and respiratory cycles. Specifically, each individual's 4D time series data were regressed on nine predictors: white matter (WM), cerebrospinal fluid (CSF), the global signal, and six motion parameters. The global signal regressor was generated by averaging across the time series of all voxels in the brain. The WM and CSF covariates were generated by segmenting each individual's high-resolution structural image (using FAST in FSL). The resulting segmented WM and CSF images were thresholded to ensure 80% tissue type probability. These thresholded masks were then applied to each individual's time series, and a mean time series was calculated by averaging across time series of all voxels within each mask. The six motion parameters were calculated in the motion-correction step during preprocessing. Movement in each of the three cardinal directions (X, Y, and Z) and rotational movement around three axes (pitch, yaw, and roll) were included for each individual.

Seed selection. Six 7.5-mm-radius seed regions of interest (ROIs) (containing 33 voxels) centered on the coordinates previously used

by Fox et al. (6) were created to examine functional connectivity for each of six regions, three regions within the "task-positive" network and three within the "default mode" network. The ROIs within the task-positive network were located in the IPS (-25, -57, 46), the middle temporal region (MT+; -45, -69, -2), and the right frontal eye field (FEF) region of the precentral sulcus (25, -13, 50). The default mode network seed ROIs were located in the left lateral parietal cortex (LP; -45, -67, 36), medial prefrontal cortex (MPF; -1, 47, -4), and PCC (-5, -49, 40).

Individual seed-based functional connectivity analysis. First, each individual's residual 4D time series data were spatially normalized by applying the previously computed transformation to the MNI152 standard space. Then the time series for each seed was extracted from these data. Time series were averaged across all voxels in each seed's ROI. For each individual dataset, the correlation between the time series of the seed ROI and that of each voxel in the brain was determined. This analysis was implemented using 3dfm+ (AFNI) to produce individual-level correlation maps of all voxels that were positively or negatively correlated with the seed's time series. Finally, these individual-level correlation maps were converted to Z-value maps using Fisher's *r*-to-*z* transformation.

Functional Connectivity: Independent Component Analysis. Overview. Temporal-concatenation group ICA (TC-GICA) was used to generate group-level components for the dataset (7) using MELODIC (FSL). Given computational resource limitations (e.g., 32 GB of physical memory), as well as a number of centers with a small number of time points due to repetition times >2.0 s, each TC-GICA run was applied to a dataset consisting of 18 participants/center from the 17 centers that collected a minimum of 165 functional volumes per scan. This approach also ensured that a single center's data would not drive the ICA components detected. Consistent with recent work on low-dimensional ICA (8), the number of components was fixed at 20. Given the potential for such factors as initial random values and subject sampling to affect ICA results, 25 TC-GICA analyses were performed, each using a unique resampling from each of the 17 centers. A meta-ICA analysis was then carried out across the 25 runs to extract the 20 spatially independent components consistently identified across the 25 runs. An alternative hierarchical clustering approach based on ICASSO (9) is described below. The two approaches yielded similar results. Dual regression (10, 11) was then carried out using the 20 resulting components as templates, to produce individual participant maps for each of the 20 components.

TC-GICA. Specifically, TC-GICA comprised five fundamental steps:

1. Each individual's preprocessed data were first truncated to the same number of time points (i.e., 165 EPI volumes).
2. A bootstrapping dataset was generated by randomly choosing 18 individual datasets per center, resulting in 306 individual functional datasets.
3. All 306 individual functional datasets were spatially averaged in MNI152 standard space and then used to estimate the mean covariance matrix.
4. The number of components was set at 20, and all individual functional data were projected into a subspace spanned by the first 20 eigenvectors of the mean covariance matrix, resulting in reduced individual fMRI data (in a common subspace).

5. All 306 reduced individual datasets were temporally concatenated, reduced via principal component analysis to 20 dimensions, and fed into the probabilistic ICA algorithm with a random initial value (12).

This procedure produced 20 group-level components for each TC-GICA run. Finally, 500 (20×25) group-level components were generated from the 25 TC-GICA runs.

Generation of component templates for dual regression (meta-ICA). To provide more accurate and robust ICA component templates, we carried out another low-dimensional (20 components) TC-GICA. Here we concatenated the 500 components produced by the 25 TC-GICA runs as the input data of a single-session ICA in MELODIC. The resultant maps were used as final component templates for the dual regression procedure. Of note, this method was selected as the primary approach over the alternative approach described because it guarantees the spatial independence of the 20 components, whereas the alternative approach does not.

Generation of component templates for dual regression model (alternative approach). To emphasize the robustness of the findings of the meta-ICA, here we describe an alternative approach that yields nearly identical components to the meta-ICA. The findings of the two approaches differed notably for only one of the 20 components, for which the meta-ICA finding was more plausible. Given the high degree of similarity between the two methods, we present only the findings from the meta-ICA in the present work. In the alternative approach, we used the hierarchical clustering algorithm implemented in the ICASSO toolbox (9). ICASSO was designed for validating the robustness of ICA with respect to random initial values (of the ICA mixing matrix) and the ICA cost function optimization search strategy. However, due to limitations in computational resources (e.g., 32 GB of memory in the present work), TC-GICA cannot be carried out on the full datasets. Thus, we used the bootstrapping approach described above with 25 ICA analyses, in which initial values and the specific participants selected from each center varied from one ICA analysis to the next. Here the 500 group-level components (20 components per run \times 25 runs) were sorted using hierarchical clustering. The number of clusters (20) was selected to match the number of components. The similarity between components was measured by the combination of both spatial R_s and temporal R_t correlations in Eq. (1) and the distance between components as defined in Eq. (2) (13):

$$S(i,j) = \lambda * R_s(i,j) + (1 - \lambda) * R_t(i,j) \quad (1)$$

$$D(i,j) = \sqrt{1 - S(i,j)}, \quad 1 \leq i, j \leq 500 \quad (2)$$

Considering the spatial ICA, $\lambda = 0.8$ was chosen in our clustering procedure. Finally, the median value at each voxel for each of the 20 clusters was calculated to determine the final component templates for the dual regression procedure.

Individual component reconstruction via the dual regression model. To reconstruct component maps for each participant, the recently developed dual regression procedure (11, 14) was applied to each of the 1,093 individual participants' datasets. Specifically, in the present work, dual regression consisted of two linear regressions carried out independently for each of the 20 component maps identified in temporal concatenation ICA. For each component template, the first regression model used the template as a spatial predictor for the participant's 4D data, producing a set of individual regression weights in the time domain (i.e., a time series for each spatial map). Using this time series as a temporal predictor for the 4D BOLD data, the second regression equation estimated the individual regression weights in the spatial domain (i.e., the participant-level individual spatial map). Both regressions used the same data set used for the seed-based con-

nectivity approaches, that is, each participant's 4D dataset after removal of the nine nuisance covariates. Component time series were demeaned in both regressions, but no variance normalization was used. The dual regression procedure was carried out for all 1,093 participants included across 24 centers, not just those used for the generation of ICA-based templates. For each component, these individual spatial maps were then used to evaluate group-level statistics.

Amplitude of spontaneous low-frequency fluctuations. To examine the potentially meaningful information contained within the ALFF, two fast-Fourier transformation (FFT)-based indices, ALFF and fALFF, were used to compute the amplitude of low-frequency fluctuations in the frequency domain (2, 3, 15). For each individual, ALFF and fALFF were computed to identify those voxels with significantly detectable low-frequency fluctuation amplitudes. Specifically, at each voxel, ALFF is calculated as the sum of amplitudes within a specific low frequency range (0.01–0.1 Hz). fALFF is the normalized ALFF, calculated by dividing the ALFF value by the total sum of amplitudes across the entire frequency range measured in a given time series. Voxelwise ALFF and fALFF maps were calculated for each participant in native space, and then transformed into the MNI152 standard brain space with 3-mm isotropic voxel size. Before statistical analyses, each individual ALFF or fALFF map was Z-transformed (i.e., by subtracting the mean voxelwise ALFF or fALFF obtained for the individual's entire brain, and then dividing by the corresponding SD) to improve its suitability for group-level parametric analyses. The individual Z-transformed ALFF or fALFF maps were used in subsequent group- and center-level analyses.

Unified group-level statistical model. For all three types of R-fMRI measures (seed-based correlations, ICA, and ALFF/fALFF), a unified general linear model frame was developed for center-level statistical analyses. The unified statistical model is a one-way ANOVA, treating centers as the between factor. F-contrasts were used to measure the effect of centers. Overall group mean contrasts across all centers were modeled as well. Specifically, a one-factor 24-level ANOVA (factor: center; 1,093 participants), with age and sex as covariates, was used to examine the effects of age, sex, and center on the three R-fMRI measures. Multiple comparisons were corrected at the cluster level using Gaussian random field theory (min $Z > 2.3$; cluster significance: $P < 0.05$, corrected).

SI Results

Center-Related Variability. The results presented in Fig. 1 show that the effects ascribable to center can be either interpreted as negligible, as indicated by the high between-center Kendall's W (row 2), or substantial, accounting for much of the variance (row 3). These opposite interpretations are not mutually exclusive. The high between-center Kendall's W indicates that the resting-state measures (i.e., functional connectivity, fluctuation amplitude) obtained from different centers have a high degree of similarity. Nevertheless, systematic differences exist between centers, and these are easily quantified by ANOVA. In Fig. S5 for each center, the mean functional connectivity across 40 peak voxels derived from the center effect map for the PCC seed is depicted. As the figure shows, there are between-center differences in the height of the functional connectivity values. Some centers have overall higher functional connectivity values than others; these differences in the height of functional connectivity values drive the significant between-center effects. The variability in functional connectivity values could be related to a number of factors (e.g., the specific scanner used, scanner sequence, sample characteristics, specific instructions to participants, degree of variability in participant wakefulness). Because there was no previous coordination among centers regarding scanning parameters, each of these parameters could contribute to across-center differences. Specific examination of these factors is beyond the scope of the present work, but we anticipate that it will be

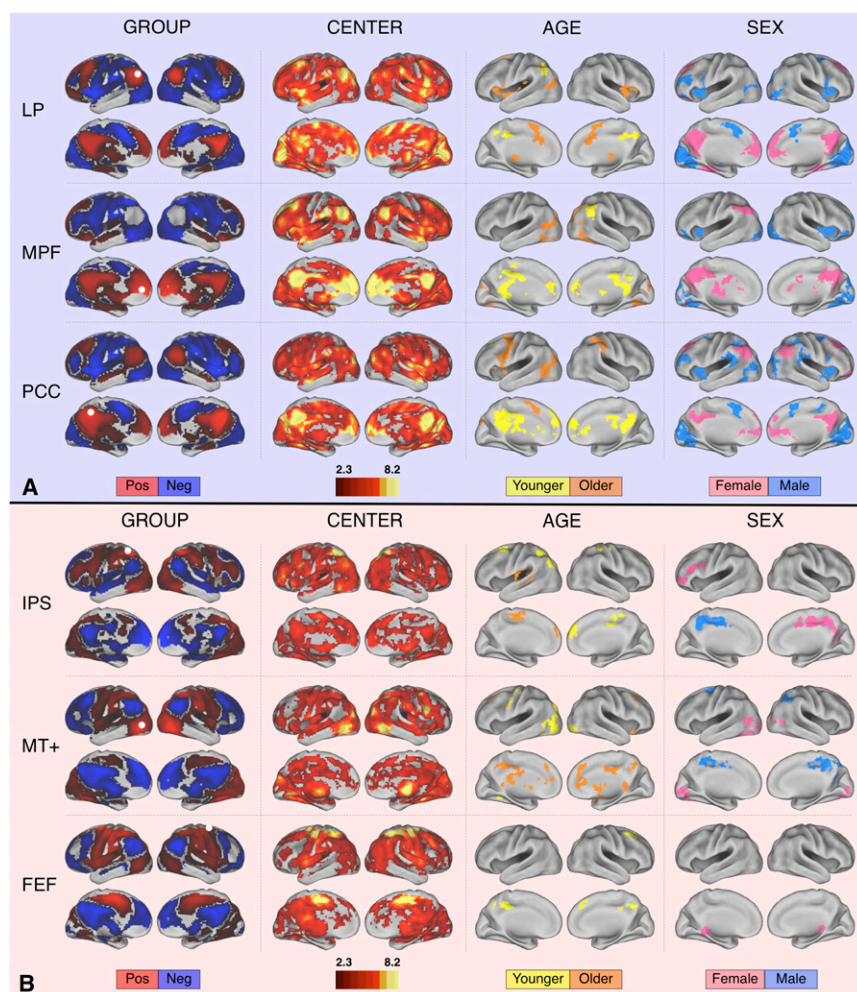


Fig. S1. Center-, age-, and sex-related variations detected in R-fMRI measures of functional connectivity using seed-based correlation analyses. The first column depicts group-level functional connectivity maps for three representative “default mode” seeds (A) and three “task-positive” network seeds (B). The seed ROIs are shown as white circles. The second column depicts voxels exhibiting significant effects of center, as detected by one-way ANOVA (across 24 centers, including 1,093 participants). Columns 3 and 4 depict voxels exhibiting age- and sex-related variations (modeled as covariates). Center, sex, and age findings were corrected for multiple comparisons ($Z > 2.3$; $P < 0.05$, corrected). All supplementary cortical surface maps are arrayed as shown in Fig. 1, with lateral views in upper rows, medial views in lower rows, left hemisphere on the left, and right hemisphere on the right. “Male” refers to significantly greater connectivity in males; similarly, “female” refers to significantly greater connectivity in females. “Older” refers to significantly increasing connectivity with increasing age, whereas “younger” refers to significantly increasing connectivity with decreasing age. “Pos”, positive functional connectivity; “neg”, negative functional connectivity.

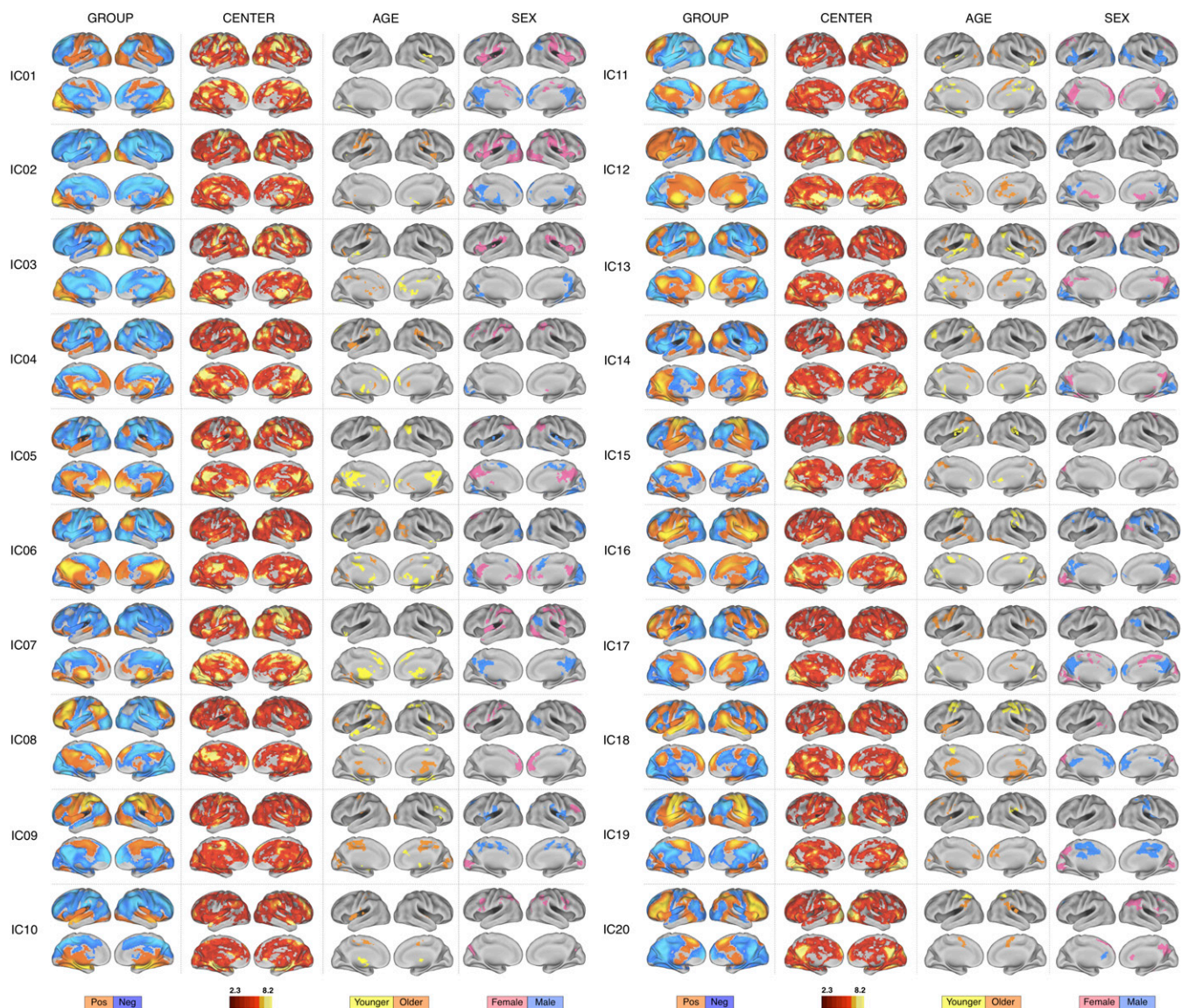


Fig. S2. Center-, age-, and sex-related differences detected in R-fMRI measures of functional connectivity combining independent component and dual regression analyses. The first column depicts group-level maps for 20 functional connectivity ICs. For each component, the second column depicts voxels exhibiting significant effects of center, as detected by one-way ANOVA (across 24 centers, including 1,093 participants). Columns 3 and 4 depict voxels exhibiting age- and sex-related variations. Center, age and sex findings were corrected for multiple comparisons ($Z > 2.3$; $P < 0.05$, corrected). "Male" refers to significantly greater connectivity in males; similarly, "female" refers to significantly greater connectivity in females. "Older" refers to significantly increasing connectivity with increasing age, whereas "younger" refers to significantly increasing connectivity with decreasing age. "Pos," positive group effect; "neg," negative group effect.

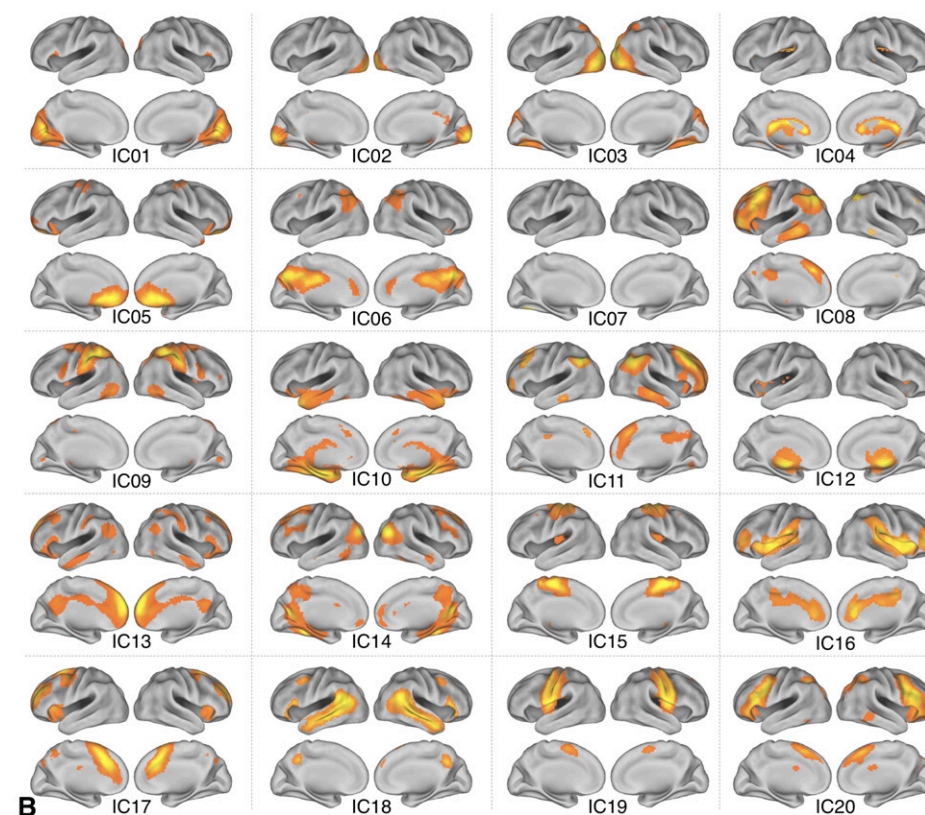


Fig. S3. IC templates used for dual regression analyses. Independent component maps resulting from the meta-ICA analysis shown on standard brain views (A) and surface maps (B). Component maps were thresholded at $P > 0.05$ using spatial mixture modeling. Peak coordinates of each IC's activity are displayed in the lower right corner of each grid panel (MNI152 standard space).

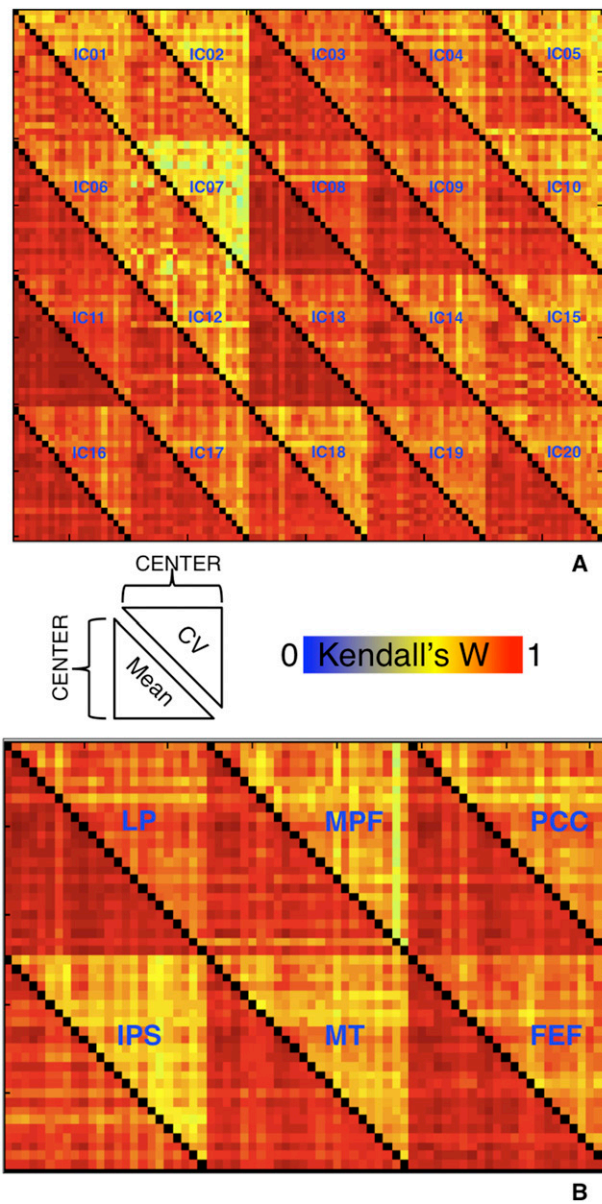


Fig. S4. Consistency of R-fMRI measures across centers: ICA combined with dual regression (A), and seed-based correlation (B). For each center, the voxelwise mean and coefficient of variation was calculated for each R-fMRI measure. The Kendall's *W* concordance of the mean or coefficient of variation maps between any two centers was calculated. The coefficient of variation is depicted above the diagonal, the mean below.



Fig. S5. Functional connectivity values observed at peak locations of between-center differences. For each center we calculated the mean across a 3 mm radius sphere centered at each of the 40 most significant voxels indexing the effect of center for the PCC seed ROI (Fig. 1, column 1, row 3). Connectivity values indexed the functional connectivity between the 3 mm radius sphere and the PCC seed ROI. All centers included in the analyses are shown ($n = 24$). Although the strength of functional connectivity values observed across centers clearly varies, the within-center variability is relatively low. This indicates that the differences in functional connectivity strength among centers are relatively stable across the brain. A center that shows higher functional connectivity in one area of the brain compared with another center most likely also shows higher functional connectivity in other areas of the brain.

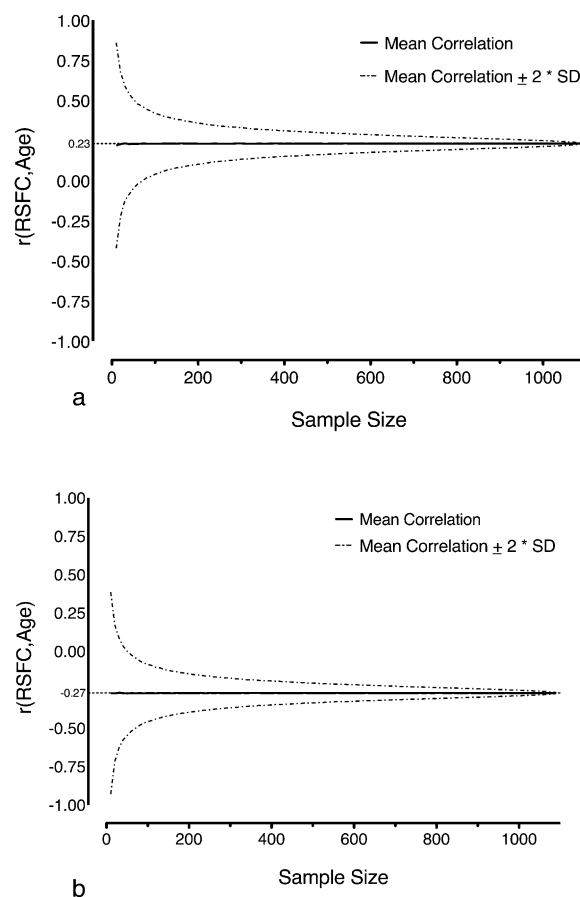


Fig. S6. Effect of sample size on the correlation between age and RSFC. Shown is the mean correlation \pm 2 SD across 10,000 calculations of the correlation between age and functional connectivity strength as a function of sample size. For each of the two regions illustrating the effect of age for the PCC seed ROI in Fig. 2, we calculated the correlation between age and RSFC as a function of sample size. We randomly sampled subgroups, ranging in size from 10 to 1,090 participants, from the total of 1,093 participants. We then calculated the correlation between age and RSFC for each of the subgroups. This procedure was iterated 10,000 times to optimize randomization. (A) Mean correlation \pm 2 times the SD across 10,000 iterations for the region illustrated in Fig. 2 that showed a positive correlation between age and RSFC with the PCC seed. (B) Mean correlation \pm 2 times the SD across the 10,000 iterations for the region illustrated in Fig. 2 that showed a negative correlation between age and RSFC with the PCC seed. In each figure, the actually observed correlation is indicated on the y-axis in a smaller font.

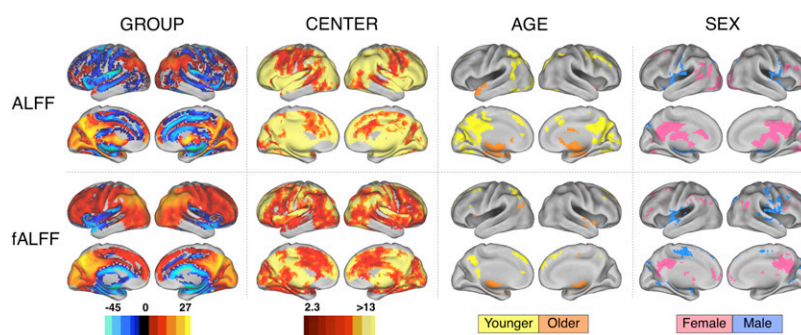


Fig. S7. Center-, age-, and sex-related variations in R-fMRI amplitude measures. The first column depicts group-level maps for voxelwise measures of ALFF (Upper) and fALFF (Lower). Before group-level analyses, each participant's ALFF/fALFF map is Z-transformed, such that positive voxels reflect greater low-frequency fluctuation amplitudes than baseline (whole brain mean) and negative voxels reflect low-frequency fluctuation amplitudes below baseline. The second column depicts voxels exhibiting significant effects of center, as detected by one-way ANOVA (across 24 centers, including 1,093 participants). Columns 3 and 4 depict voxels exhibiting age- and sex-related variations. Center, age, and sex findings were corrected for multiple comparisons ($Z > 2.3$; $P < 0.05$, corrected). "Male" refers to significantly greater connectivity in males; similarly, "female" refers to significantly greater connectivity in females. "Older" refers to significantly increasing connectivity with increasing age, whereas "younger" refers to significantly increasing connectivity with decreasing age.

Table S1. Data currently included in the 1,000 Functional Connectomes Project

	Center	PI	N	n*	Age years, mean (SD)	Age range years	Male sex %
1.	Baltimore, MD	J. J. Pekar/S. H. Mostofsky	23		29.26 (5.46)	20–40	35%
2.	Bangor, UK	S. Colcombe	20		23.4 (5.32)	19–38	100%
3.	Beijing, China	YF. Zang	198	193	21.16 (1.83)	18–26	39%
4.	Beijing, China	XC. Weng	28	27	20.41 (1.39)	18–24	27%
5.	Berlin, Germany	D. Margulies	26		29.77 (5.21)	23–44	50%
6.	Bethesda, MD	M. Ernst	18		33.00 (13.31)	18–53	22%
7.	Cambridge, MA	R. L. Buckner	198		21.03 (2.31)	18–30	38%
8.	Cambridge, MA	S. Whitfield-Gabrieli	39	35	25.09 (3.53)	20–32	49%
9.	Cleveland, OH	M. J. Lowe	31		43.55 (11.14)	24–60	35%
10.	Dallas, TX	B. Rypma	24		42.63 (20.07)	20–71	50%
11.	Hvidovre, Denmark	A.-M. Dogonowski/K. Madsen	28		41.75 (10.7)	21–68	50%
12.	Leiden, The Netherlands	S. A. R. B. Rombouts	31		22.19 (2.57)	18–28	74%
13.	Leipzig, Germany	A. Villringer	37		26.22 (5)	20–42	43%
14.	Magdeburg, Germany	M. Walter	29	28	30.43 (5.71)	22–43	93%
15.	Milwaukee, WI	SJ. Li	64		53.59 (5.79)	44–65	64%
16.	New Haven, CT	M. Hampson	19	18	31.61 (10.27)	18–48	56%
17.	New York, NY [†]	M. Milham/F. X. Castellanos	59		32.78 (8.83)	20–49	68%
18.	New York, NY [†]	M. Milham/F. X. Castellanos	20		29.75 (9.94)	18–46	40%
19.	Newark, NJ	B. B. Biswal	19		24.11 (3.91)	21–39	47%
20.	Orangeburg, NY [‡]	M. J. Hoptman	21	20	40.65 (11.03)	20–55	75%
21.	Oulu, Finland [‡]	V. J. Kiviniemi/J. Veijola	103		21.52 (0.57)	20–23	36%
22.	Oxford, UK	S. M. Smith/C. Mackay	22		29 (3.79)	20–35	55%
23.	Queensland, Australia	K. McMahon	19	18	26.28 (3.71)	20–34	61%
24.	St. Louis, MO	B. L. Schlaggar/S. E. Petersen	31		25.1 (2.31)	21–29	45%

Data from the following centers will be included in the 1000 Functional Connectomes data release but are not included in the current analyses: Ann Arbor, MI: C. S. Monk/R. D. Seidler/S. J. Peltier; Atlanta, GA: H. S. Mayberg; Berlin, Germany: S. Schmidt; Durham, NC: D. J. Madden; Durham, NC: L. Wang; London, Ontario, Canada: P. Williamson; Munich, Germany, C. Sorg/V. Riedl; Nanjing, China: G.J. Teng/HY. Zhang; Pittsburgh, PA: G.J. Siegle; Portland, OR: D. Fair/B. J. Nagel; Taipei, Taiwan: C.P. Lin; Vienna, Austria: C. Windischberger.

*Actual number of participants included in the analysis, if different from *N*.

[†]Data from the same magnet, different sequence.

[‡]1.5-T magnet.

Funding sources for each contributor (numbered by site): 1: R01 MH085328, R01 MH078160, HD-24061 (Intellectual Disabilities Research Center), M01 RR00052 (Johns Hopkins General Clinical Research Center) and P41 RR15241 (National Center for Research Resources); 3: NSFC (No.30621130074); 4: Chinese Ministry of Science and Technology (No. 2007CB512300); 5: Berlin School of Mind and Brain (DFG); 7: Howard Hughes Medical Institute; 9: National Multiple Sclerosis Society; 13: Competence Net Stroke (BMBF) and Berlin School of Mind and Brain (DFG); 17-18: NIDA (RO1DA016979), NIMH (RO1MH083246), Stavros Niarchos Foundation; 19: NINDS (RO1NS049176); 20: R01 MH064783, R21 MH084031, R01 MH0663674; 21: Academy of Finland (Grant codes 124257, 212181, 214273); 23: Australian Research Council (ARC) Discovery grant (DP0452264); 24: NIH NS053425; Atlanta, G.A.: H.S. Mayberg: URC Grant, Emory University; Durham, NC: D.J. Madden: NIH/NIA R01 AG011622; Durham, NC: L. Wang: Paul B. Beeson Career Developmental Awards (K23-AG028982); Portland, OR: D. Fair/B. Nagel: Oregon Clinical and Translational Research Institute, Medical Research Foundation, UNCF-Merck, Ford Foundation, Dana Foundation; Taipei, Taiwan: C-P. Lin: National Health Research Institute grant (NHRI-EX98-9813EC), Taiwan; Vienna, Austria: C. Windischberger: OeNB-P11468 and OeNB-P12982.