Abnormal Cortical Growth in Schizophrenia Targets Normative Modules of Synchronized Development

Supplemental Information

Smoothness of the Spline Functions

As described in the main text, let age_{ij} denote the age of the i^{th} individual at the time of his/her j^{th} scan. The thickness measured at the v^{th} vertex is modeled as a smooth function g_v of age plus a random person effect u_{iv} plus error:

$$thickness_{ijv} = g_v(age_{ij}) + u_{iv} + error_{ijv} \quad (1)$$

Rather than specifying the form of g_v (linear, quadratic, etc.), we allow it to be an essentially arbitrary smooth function, by taking it to be a linear combination of 10 piecewise cubic B-spline functions (Main Figure 2) (1). Estimating this linear combination by conventional least squares would lead to an overly bumpy estimate of g_v. Penalized spline smoothing avoids this by instead minimizing the sum of squared errors plus a 'roughness' penalty, taken here to be $\lambda \int \boldsymbol{g_v''(t)^2 dt}$. The integral serves as an index of the roughness or bumpiness of the function, and the non-negative smoothing parameter λ determines the tradeoff between the goodness of fit and the estimated function's smoothness. If λ is small, then there is only a small penalty for roughness, and g_v will tend to be less smooth; as λ approaches infinity, g_v approaches a straight line as would result from ordinary linear regression. Finding an optimal value of λ is crucial. Often, this is done by cross-validation and related procedures (2). Here, we follow the approach of choosing the smoothing parameter by restricted maximum likelihood (3-5), a well-known method for fitting linear mixed-effect models (6). One advantage of this approach is that it allows a unified treatment when, as in our case, the model incorporates random effects for longitudinal scans of the same individual.

Comparison between Polynomial and Spline Models

Polynomial models have been used by most existing work to analyze the maturational trajectories of regions of the brain. The use of polynomial models can assess whether the relationship between age and thickness is best captured by linear, quadratic or cubic models (7; 8). When there is evidence for an 'inverted U'-type curve, regions can also be differentiated based on the peak of these curves (7; 9).

However, polynomial models impose severe limits on hypothesis testing about differences in maturational trajectories between brain regions or between study populations. First, it is not clear how to quantify differences in the shapes of growth curves, especially if they have been modeled by polynomial functions of different order (e.g. quadratic vs. cubic). Second, these growth curves are highly dependent on the precise age-range sampled, which can lead to false inferences about key features such as the location of the peak of a maturational trajectory (10; 11). Finally, accurate modeling of neurodevelopment is severely limited by the assumption that regional brain growth must correspond to linear, quadratic or cubic functions, rather than allowing the shape of the model to be determined by the complexity of the biology.

The use of semiparametric penalized spline models addresses these limitations (12), allowing us to characterize differences in the trajectories of growth curves between brain regions and between clinical groups at tens of thousands of points across the cortical mantle. It has been shown that features of such models are more robust to alterations in the sampled age range, because the fit of the curve at a given age depends on local information only (scans within a limited age range), as opposed to polynomial models where scans at the edges of the age range can affect the entire curve (10; 12; 13). In addition, it is straightforward to distinguish age-constant trait differences and age-varying trajectory differences between groups.

Table S1. **Demographic information on study sample.**

	Participants with COS	Healthy Participants	Statistic (*df*)	*P*-value
Scans	261	264		
Subjects	106	102		
Age of Scan, Mean (SD)	17.4 (4.5) Range = 7-32	17.2 (5.0) Range = 8-33	*t* = 0.56 (519.3)	0.57
Sex	35 F; 71 M	41 F; 61 M	X^2 = 0.8 (1)	0.4
Race	8 A; 25 B; 9 H; 6 O; 58 W	5 A; 22 B; 7 H; 7 O; 61 W	X^2 = 1.2 (4)	0.88
IQ, Mean (SD)	76.6 (13.7)	112.2 (12.3)	*t* = 26.2 (406.2)	<0.001
Socio-economic Status, Mean (SD)	40.2 (32.2)	62.6 (20.6)	*t* = 5.9 (176.5)	<0.001
Handedness	21 L; 85 R	6 L; 96 R	X^2 = 7.7(1)	0.005

COS, childhood-onset schizophrenia; df, degrees of freedom; F, female; M, male; R, right-handed; L, non-right-handed. Race: A, Asian; B, Black/African-American; H, Hispanic; O, other/mixed-race; W, white.

Table S2. Brain anatomical areas with any group difference between typical development and childhood-onset schizophrenia. In other words, the null hypothesis H_0, that β_v (age) in equation (2) is identically zero, is rejected. Regions listed have group differences encompassing at least 20% of their surface area, using an FDR threshold of 0.001.

	MNI Coordinate of Peak Vertex			Percent of AAL Region with Group Differences
AAL Region	**x**	**y**	**z**	
IFGtriang.R	51.6	24.6	7.6	86%
IFGtriang.L	-37.8	23.2	8	77%
PreCG.L	-27.6	-11.9	56.3	75%
SMA.L	-7.5	16.5	67.3	67%
IFGoperc.R	50.4	22.2	11.7	64%
SFGmed.R	4	27.8	53.4	62%
STG.L	-62.7	-25.1	10.6	50%
IPL.L	-55.7	-30.9	47.1	48%
SMA.R	3.8	21.9	59.8	40%
SFGdor.L	-22.4	11.1	52.6	40%
IFGoperc.L	-40.6	20.4	8.5	40%
MFG.L	-26.6	15.1	47.9	39%
PreCG.R	43.3	6.2	23.3	38%
PoCG.L	-37.3	-18.3	42.5	38%
SFGmed.L	-4	41.8	44.7	36%
SMG.L	-60.4	-43.3	40.3	36%
HES.R	53	-20.5	5.4	32%
HES.L	-60.2	-16.2	4.9	31%
ITG.L	-56	-48.2	-5.9	30%
SFGdor.R	16.9	31.4	56.8	30%
MFG.R	37.2	12.1	28.8	25%
MTG.L	-65.1	-46.9	-2	24%
PCL.L	-11.2	-9.4	71.9	23%

AAL, Automated Anatomical Labeling; FDR, false discovery rate; MNI, Montreal Neurological Institute. Regions: IFGtriang.R; right inferior frontal gyrus, triangular part; IFGtriang.l, left inferior frontal gyrus, triangular part; PreCG.L, left precentral gyrus; SMA.L, left supplementary motor area; IFGoperc.R, right inferior frontal gyrus, opercular part; SFGmed.R, right medial superior frontal gyrus; STG.L, left superior temporal gyrus; IPL.L, left inferior parietal lobule; SMA.R, right supplementary motor area; SFGdor.L, left dorsolateral superior frontal gyrus; IFGoperc.L, left inferior frontal gyrus, opercular part; MFG.L, left middle frontal gyrus; PreCG.R, right precentral gyrus; PoCG.L, left postcentral gyrus; SFGmed.L, left medial superior frontal gyrus; SMG.L, left supramarginal gyrus; HES.R, right Heschl's gyrus; HES.L, left Heschl's gyrus; ITG.L, left inferior temporal gyrus; SFGdor.R, right dorsolateral superior frontal gyrus; MFG.R, right middle frontal gyrus; MTG.L, left middle temporal gyrus; PCL.L, left paracentral lobule.

Table S3. Brain anatomical areas with trait differences between typical development and childhood-onset schizophrenia. In other words, the null hypothesis H_{0a} that $\beta_v^{(1)} = 0$ in equation (3) is rejected. Regions listed have group differences encompassing at least 20% of their surface area, using an FDR threshold of 0.001. Regions listed have group differences over at least 20% of their surface area, using an FDR threshold of 0.001.

	MNI Coordinates of Peak Vertex			Percent of
AAL Region	**x**	**y**	**z**	**Region**
IFGtriang.L	-39.2	23.4	7	97%
PreCG.L	-27.9	-12.4	58.3	83%
IFGoperc.L	-41.7	20.2	8.4	80%
STG.L	-62.7	-25.1	10.6	79%
SMA.L	-4.5	20.2	56.7	75%
SFGmed.R	4.5	26.7	52.3	71%
HES.R	51.5	-19.8	5.8	62%
MFG.L	-24.7	9	54.5	60%
SFGdor.L	-24.3	1.6	51.6	52%
SMG.L	-61.3	-43.5	38.3	52%
HES.L	-60.2	-16.2	4.9	51%
IFGtriang.R	41.8	13.2	25.8	51%
ORBinf.L	-38.1	29.4	-3.2	50%
SMA.R	3.8	21.9	59.8	47%
PreCG.R	52.9	0.7	44.5	47%
MTG.L	-65.1	-46.9	-2	45%
SFGmed.L	-4.7	40.2	45	44%
SFGdor.R	12.9	37.5	50.4	44%
ITG.L	-56	-48.2	-5.9	41%
MFG.R	47.4	3.2	45.4	36%
IPL.L	-57.1	-28.7	45.4	36%
IFGoperc.R	42.8	12.3	25.3	33%
PCL.L	-12.4	-11.5	69.1	33%
SMG.R	60.5	-33.1	45.9	27%
STG.R	53.1	-21.3	5.9	25%

AAL, Automated Anatomical Labeling; FDR, false discovery rate; MNI, Montreal Neurological Institute. Regions: IFGtriang.l, left inferior frontal gyrus, triangular part; PreCG.L, left precentral gyrus; IFGoperc.L, left inferior frontal gyrus, opercular part; STG.L, left superior temporal gyrus; SMA.L, left supplementary motor area; SFGmed.R, right medial superior frontal gyrus; HES.R, right Heschl's gyrus; MFG.L, left middle frontal gyrus; SFGdor.L, left dorsolateral superior frontal gyrus; SMG.L, left supramarginal gyrus; HES.L, left Heschl's gyrus; IFGtriang.R; right inferior frontal gyrus, triangular part; ORBinf.L; left inferior frontal gyrus, orbital part; SMA.R, right supplementary motor area; PreCG.R, right precentral gyrus; MTG.L, left middle temporal gyrus;

SFGmed.L, left medial superior frontal gyrus; SFGdor.R, right dorsolateral superior frontal gyrus; ITG.L, left inferior temporal gyrus; MFG.R, right middle frontal gyrus; IPL.L, left inferior parietal lobule; IFGoperc.R, right inferior frontal gyrus, opercular part; PCL.L, left paracentral lobule; SMG.R, right supramarginal gyrus; STG.R, right superior temporal gyrus.

Table S4. Brain anatomical areas with trajectory differences between typical development and childhood-onset schizophrenia. In other words, the null hypothesis H_{0b} that $\beta_v^{(2)}(\text{age})$ in equation (3) is identically zero is rejected. Regions listed have group differences encompassing at least 20% of their surface area, using an FDR threshold of 0.05.

	MNI Coordinates of Peak Vertex			Percent of AAL Region with Group Differences
AAL Region	**x**	**y**	**z**	
IFGtriang.R	52.1	28.6	-1.4	73%
IFGoperc.R	50.3	21.1	11.5	53%
PCG.L	-8.1	-42.7	33.5	45%
ORBsupmed.R	3.1	23.8	-18.9	31%
REC.R	3.5	33.2	-21.1	30%
ORBinf.R	28.8	26.1	-22.3	27%
DCG.L	-2.3	-13.5	36.2	24%
PoCG.L	-27.3	-30.3	74.2	20%

AAL, Automated Anatomical Labeling; FDR, false discovery rate; MNI, Montreal Neurological Institute. Regions: IFGtriang.R, right inferior frontal gyrus, triangular part; IFGoperc.R, right inferior frontal gyrus, opercular part; PCG.L, left posterior cingulate gyrus; ORBsupmed.R, right medial orbital superior frontal gyrus; REC.R, right gyrus rectus; ORBinf.R, right inferior frontal gyrus, orbital part; DCG.L, left median cingulate and paracingulate gyri; PoCG.L, left postcentral gyrus.

Table S5. Brain anatomical areas with both trajectory differences and trait differences between typical development and childhood-onset schizophrenia. In other words, the null hypotheses H_{0a} and H_{0b}, that $\beta_v^{(1)} = 0$ and $\beta_v^{(2)}$(age) is identically zero in equation (3), are both rejected. Regions listed have group differences encompassing at least 20% of their surface area, using an FDR threshold of 0.05.

	MNI Coordinate of Peak Vertex			Percent of AAL Region with Group Differences
AAL Region	**x**	**y**	**z**	
IFGtriang.R	55.9	28.9	11.6	69%
IFGoperc.R	56.7	14.8	13.5	53%
PoCG.L	-30.8	-29.9	72.3	17%
REC.R	7.3	40.4	-24.5	13%

AAL, Automated Anatomical Labeling; FDR, false discovery rate; MNI, Montreal Neurological Institute. Regions: IFGtriang.R, right inferior frontal gyrus, triangular part; IFGoperc.R, right inferior frontal gyrus, opercular part; PoCG.L, left postcentral gyrus; REC.R, right gyrus rectus.

Table S6. Brain anatomical areas with trajectory differences between typical development and childhood-onset schizophrenia that also fall into the cingulo-fronto-temporal module of typical development. Regions listed have group differences encompassing at least 20% of their surface area, using an FDR-corrected *p*-value threshold of 0.05.

	MNI Coordinate of Peak Vertex			Percent of AAL Region with Group Differences
AAL Region	**x**	**y**	**z**	
IFGtriang.R	52.1	28.6	-1.4	70%
ORBsupmed.R	3.1	23.8	-18.9	30%
REC.R	3.5	33.2	-21.1	30%
PCG.L	-8.1	-42.7	33.5	26%
IFGoperc.R	50.3	21.1	11.5	23%

AAL, Automated Anatomical Labeling; FDR, false discovery rate; MNI, Montreal Neurological Institute. Regions: IFGtriang.R, right inferior frontal gyrus, triangular part; ORBsupmed.R, right medial orbital superior frontal gyrus; REC.R, right gyrus rectus; PCG.L, left posterior cingulate gyrus; IFGoperc.R, right inferior frontal gyrus, opercular part.

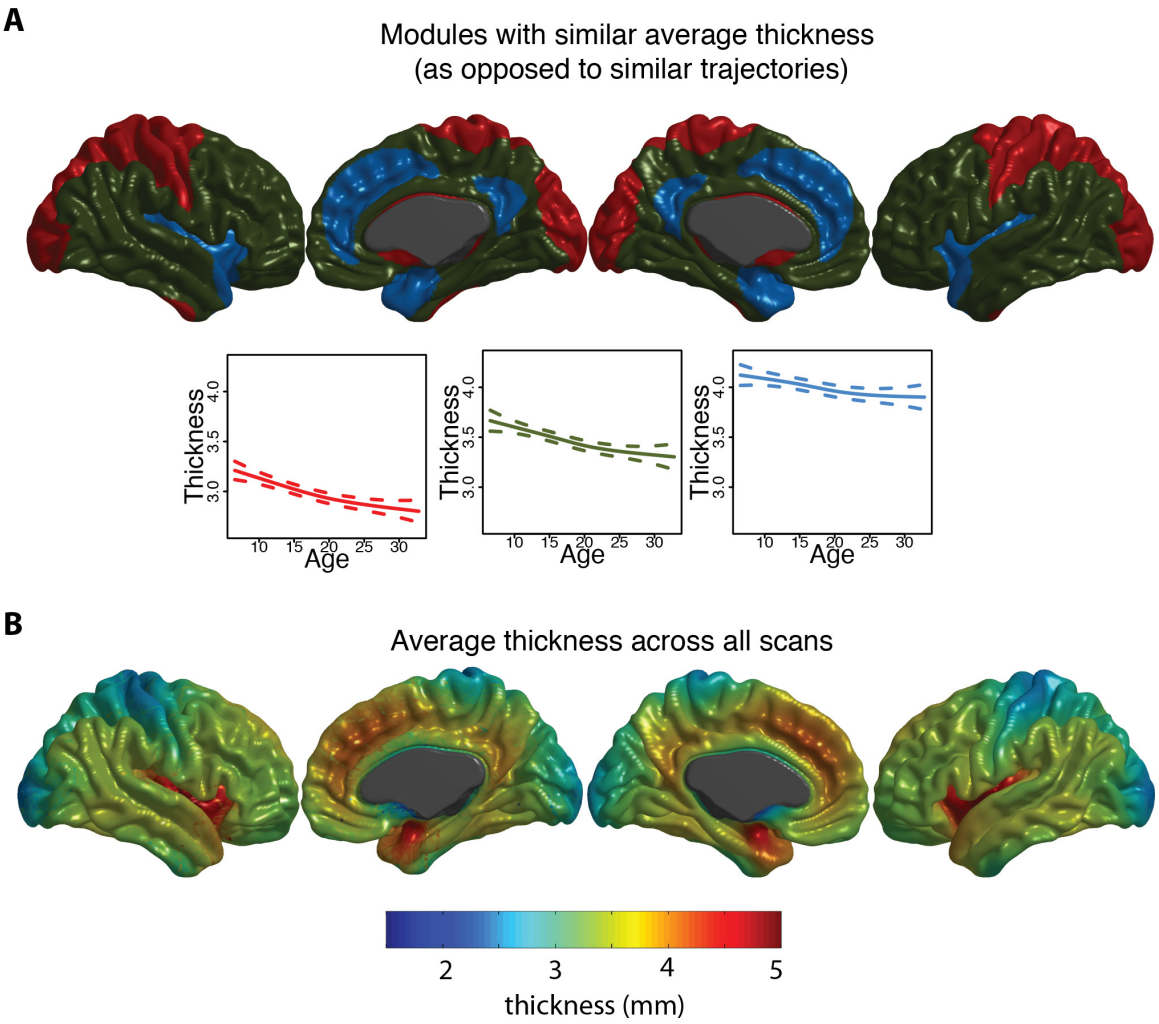

Figure S1. Modules comprised of regions with similar cortical thickness during childhood and adolescence. (**A**) A *k*-medoids algorithm was applied to the normative functional data without taking the first derivative. The illustrative case of K=3 modules is presented, as well as the growth trajectory of the average of the vertices in each module. (**B**) The average cortical thickness across all normative scans. It is evident that the modules in (**A**) simply separate the brain into regions of high, medium and low cortical thickness.

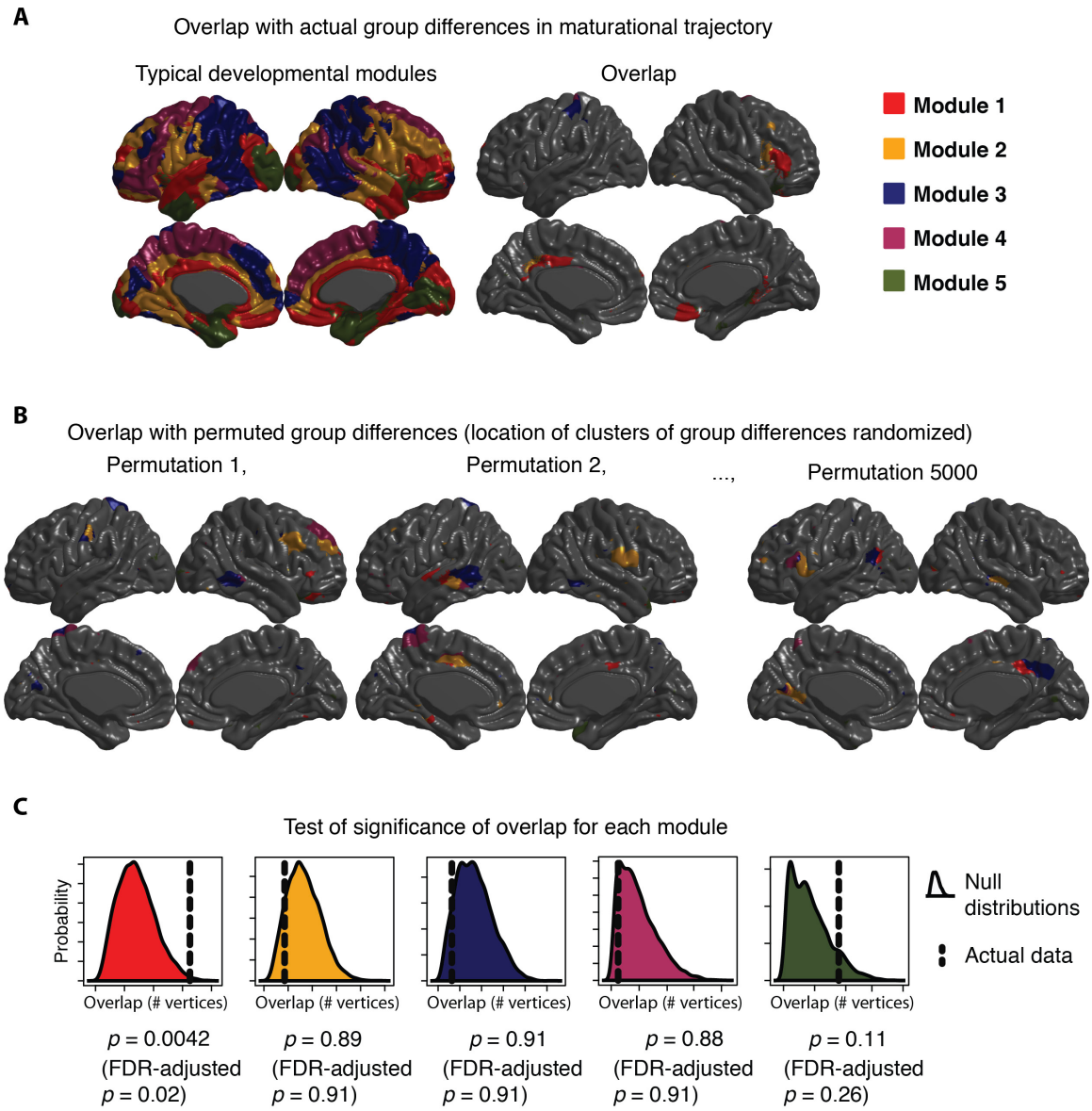

Figure S2. Randomization test of the significance of the overlap between typical developmental modules and childhood-onset schizophrenia-related alterations in the shape of maturational trajectories. (**A**) As illustrated in main text Figure 3, the overlap between developmental modules and cortical areas with significant clinical differences in the shape of their maturational trajectories (FDR < .05). (**B**) Five thousand surrogate sets of spatially contiguous brain areas, of the same size as the actual brain of group differences, but at random locations across the cortex, were simulated. The overlap between three of these simulated patterns of group differences and the typical developmental modules is shown. (**C**) Null distributions of the overlap with each module were estimated using the randomly simulated data. *P*-values were calculated as the proportion of simulated overlap values that exceeded the actual overlap. FDR, false discovery rate.

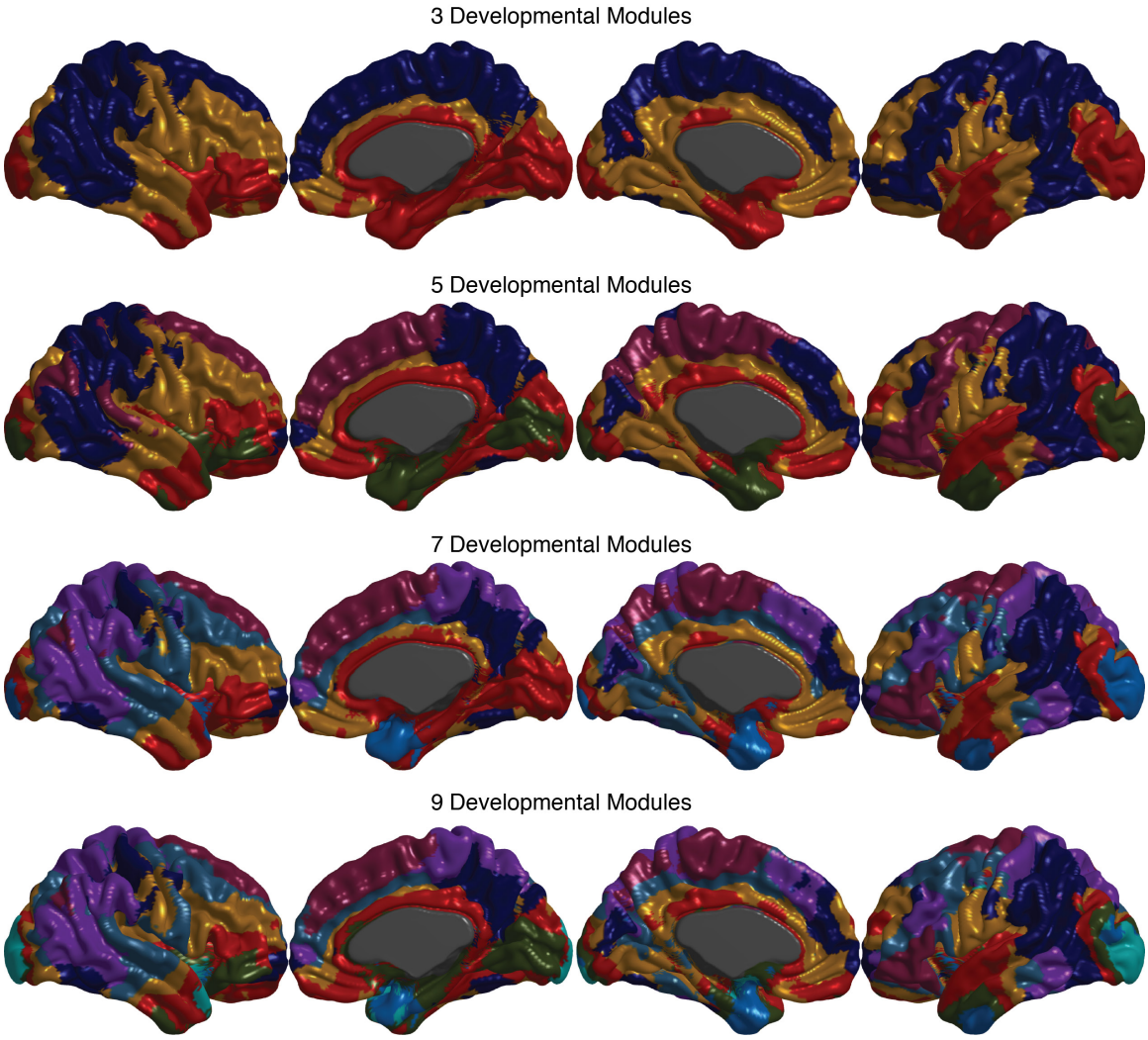

Figure S3. Typical developmental modules at different spatial resolutions. As the number of modules increases, the average size of the modules decreases accordingly. The colors of the modules across resolutions were matched algorithmically. First, the normalized mutual information (NMI) between all modular partitions was calculated (14). The partition with the second highest average NMI was matched to the partition with the highest average NMI by solving the linear assignment problem via the Hungarian algorithm (15). A soft (or fuzzy) partition based on the labels of both of these partitions was estimated, and the partition with the third highest average NMI was matched to this soft partition, also using the Hungarian algorithm, and so on for all of the partitions across resolutions. This process provides an approximate solution to the NP-hard multivariable assignment problem. It was implemented with the *R* package *clue* (16) using scripts available at http://brainnetworks.sourceforge.net/.

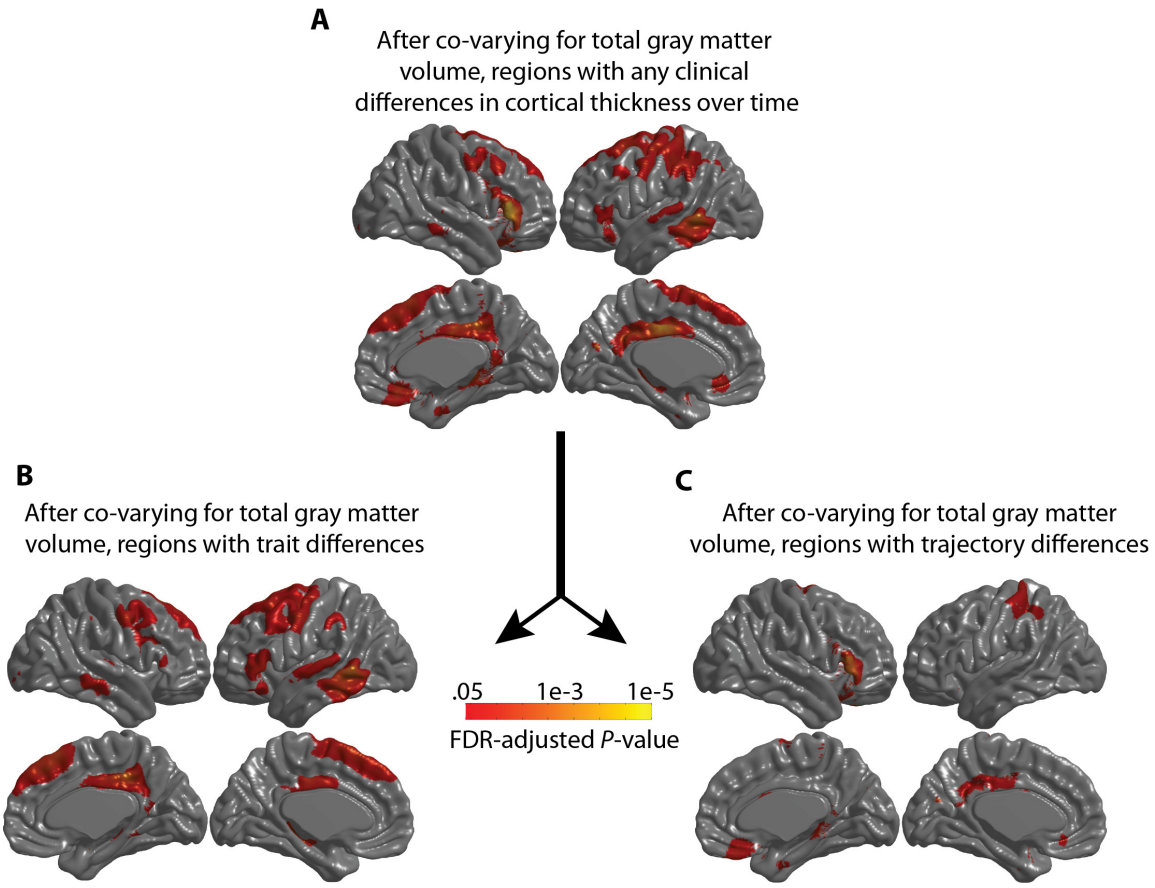

Figure S4. Abnormalities of growth curves in childhood-onset schizophrenia (COS) for each of the ~80,000 cortical vertices, using penalized spline models and FDR-adjusted *p*-values, after adding each subject's total gray matter volume as a covariate in the statistical model. (**A**) Cortical regions with any difference in the maturational trajectory in COS, either a constant trait difference or an age-varying trajectory difference. In other words, the null hypothesis H_0, that β_v (age) in equation (2) is identically zero, is rejected. (**B**) Regions with significant group differences in trait between groups, i.e., the null hypothesis H_{0a} that $\beta_v^{(1)} = 0$ in equation (3) is rejected. (**C**) Regions with significant group differences in trajectory, i.e., the null hypothesis H_{0b} that $\beta_v^{(2)}$(age) in equation (3) is identically zero is rejected. In comparison with main text Figure 3, which shows equivalent figures without covarying for total gray matter volume, the trajectory differences between the two groups are essentially identical. Relatively fewer brain regions show significant trait differences after co-varying for total gray matter volume. For example, there are no residual trait group differences in the occipital lobe, after co-varying for total gray matter volume. FDR, false discovery rate.

Supplemental References

1. Wand MP, Ormerod JT (2008): On semiparametric regression wtih O'Sullivan penalized splines. *Aust NZ J Stat* 50:179–198.
2. Craven P, Wahba G (1978): Smoothing noisy data with spline functions. *Numer Math* 31: 377–403.
3. Ruppert D, Wand MP, Carroll RJ (2003): *Semiparametric Regression*. New York: Cambridge University Press.
4. Reiss PT, Ogden RT (2007): Functional principal component regression and functional partial least squares. *J Am Stat Assoc* 102: 984–996.
5. Wood SN (2010): Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc B* 73: 3–36.
6. Laird NM, Ware JH (1982): Random-effects models for longitudinal data. *Biometrics* 38: 963–974.
7. Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, *et al.* (2008): Neurodevelopmental trajectories of the human cerebral cortex. *J Neurosci* 28: 3586–3594.
8. Sugiura N (1978): Further analysis of the data by Akaike's information criterion and the finite corrections. *Comm Stat Theor Meth* 7: 13–26.
9. Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, *et al.* (1999): Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci* 2: 861–863.
10. Fjell AM, Walhovd KB, Westlye LT, Østby Y, Tamnes CK, Jernigan TL, *et al.* (2010): When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. *Neuroimage* 50: 1376–1383.
11. Reiss PT, Huang L, Chen YH, Huo L, Tarpey T, Mennes M (2014): Massively parallel nonparametric regression, with an application to developmental brain mapping. *J Comp Graph Stat*.
12. Wood SN (2006): *Generalized Additive Models*. Boca Raton, FL: Chapman & Hall.
13. Green PJ, Silverman BW (1994): *Nonparametric Regression and Generalized Linear Models*. Boca Raton: Chapman & Hall.
14. Kuncheva LI, Hadjitodorov ST (2004): Using diversity in cluster ensembles. In: *2004 IEEE International Conference on Systems, Man and Cybernetics.* Presented at the 2004 IEEE International Conference on Systems, Man and Cybernetics, IEEE, pp 1214–1219.
15. Papadimitriou HC, Steiglitz K (1982): *Combinatorial Optimization: algorithms and complexity*. New Jersey: Prentice Hall, pp 218–266.
16. Hornik K (2012, July 7): clue: Cluster Ensembles. *CRANRprojectorg*. Retrieved July 7, 2013, from http://CRAN.Rproject.org/package=clue.