Supplementary Materials

White Matter Abnormalities in Major Depression Biotypes Identified by Diffusion Tensor Imaging

Materials and Methods

Subjects

All participants were interviewed by an experienced psychiatrist using the Structured Clinical Interview for DSM-IV: SCID-NP for controls [1] and SCID-P for patients [2]. Patients with major depressive disorder (MDD) and healthy control individuals were well matched in terms of age ($T = 0.13$, $P = 0.899$), sex (chi-square $\chi^2 = 0.09$, $P = 0.766$) and education level ($T = -1.52$, $P = 0.129$). All patients in this study met DSM-IV criteria for MDD [3]. Exclusion criteria included any history of head trauma, neurological disorders or medical conditions that might alter cognitive function or intellectual ability. This study also excluded those who had organic brain syndrome, learning disability, substance use disorder, or psychoses secondary to medical illness. Potential healthy control participants reporting mental disorders in one or more first-degree relatives were also excluded. Patients diagnosed with depression were subject to prospective longitudinal observation over a period of 6 months to clarify the diagnosis. In this study, eighty-five out of 116 patients with MDD were drug-naïve. The other thirty-one depressed patients were relapsed but they had not taken antidepressants at least the previous three months while they were recruited into the current study. Written informed consent was obtained from all participants after the procedure had been fully

explained. Ethical approval for this study was granted by the ethics committee of the West China Hospital, Sichuan University, in accord with the Declaration of Helsinki.

Clinical Assessments

Clinical symptoms of major depression were evaluated using the 17-item Hamilton Rating Scale for Depression (HAM-D) [4]. The severity of anxiety symptoms were evaluated with the 17-item Hamilton Rating Scale for Anxiety (HAM-A) [5]. Only those patients with a HAM-D score ≥14 were eligible for the study.

Neuropsychological Assessments

For the short form of WAIS [6, 7], the Verbal IQ (VIQ) of scaled scores sum was obtained by 2 (Information + Similarities) + Arithmetic + Digit Span; Performance IQ (PIQ) sum was calculated as 2 (Picture Completion + Block Design) + Digit Symbol. Full Scale IQ (FSIQ) estimates were based on the VIQ sum + PIQ sum. According to the standard procedure and age-corrected conversion tables in the WAIS-RC manual, the estimated sums of scaled scores derived from these formulae were converted to IQ scores [8].

The computerized Cambridge Neuropsychological Test Automated Battery (CANTAB) incorporates a range of visuo-spatial cognitive paradigms developed to evaluate attention, memory and executive function that are sensitive to the dysfunction of frontal, cingulate and temporal brain regions [9, 10]. The CANTAB tests included the Rapid Visual Information Processing (RVP) and the Delayed Matching to Sample (DMS). Variables of interest across tasks included reaction time, accuracy, and errors (see Table S1).

Magnetic Resonance Imaging Acquisition

All magnetic resonance imaging scans were performed using a Philips 3T (Achieva, TX, Best, the Netherlands) scanner equipped with an eight-channel head coil. The DTI data were acquired using an echo-planar imaging (EPI) sequence with the following parameters: 32 diffusion gradient directions, b-values 0 and 1,000 mm^{-2}, repetition time (TR) 10 254 ms, echo time (TE) 92 ms, FOV (field of view, RL, AP, FH) 256 mm × 256 mm × 150 mm, acquisition matrix size 128 × 128, acquisition voxel size 2 ×2 × 2 mm^3, reconstructed voxel size 2 × 2 × 2 mm^3, EPI factor 67, slice thickness = 2.0 mm, SENSE factor 2 in the anterior-posterior direction, 75 slices throughout the whole brain. All scans were reviewed by an experienced neuroradiologist to exclude gross brain abnormalities.

DTI Data Preprocessing, and Feature Extraction

DTI data were processed using FSL software (FMRIB Software Library, FMRIB, Oxford, UK) [11]. Quality control of DTI data was carried out using DTIPrep (translation < 2 mm, rotation < 0.5 mm) [12]. Participants who met the criteria were included. The preprocessing procedures included motion and eddy current correction (old edition), brain extraction, tensor model fitting and MNI normalization of tract-based spatial statistics (TBSS) [13]. The tractography-based tract atlases could provide unsurpassed efficiency for understanding the tract anatomy as they delineate convoluted trajectories and relationships with other brain structures [14]. Fractional anisotropy (FA) was then calculated for 20 fiber tracts defined by the JHU white-matter tractography atlas [15]. The identified 20 fiber tracts were forceps major (Fmaj), the frontal projection of the corpus callosum (the forceps minor – Fmin) as well as the following bilateral bundles: anterior thalamic radiation (ATR), gyrus part of the cingulum cingulate (CGC), hippocampal part of the cingulum (CGH), corticospinal tract

(CST), inferior fronto-occipital fasciculus (IFO), temporal part of the superior longitudinal fasciculus (SLFt), uncinate fasciculus (UNC), inferior longitudinal fasciculus (ILF), superior longitudinal fasciculus (SLF).

Cluster Validation - Gap Value

To aid in determining the optimal cluster number, the gap value was calculated in this study. The gap criterion formalizes the common graphical approach by estimating the "elbow" location as the number of clusters with the highest gap value [16]. The optimal number of clusters defines at the solution with the largest local or global gap value within a tolerance range.

The gap value is defined as

$$Gap_n(k) = E_n^*\{\log(W_k)\} - \log(W_k),$$

where n is the sample size, k is the number of clusters being evaluated, and W_k is the pooled within-cluster dispersion measurement

$$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} D_r,$$

where n_r is the number of data points in cluster r, and D_r is the sum of the pairwise distances for all points in cluster r.

The expected value $E_n^*\{\log(W_k)\}$ is determined by Monte Carlo sampling from a reference distribution, and is $\log(W_k)$ computed from the sample data.

Tract-Based Spatial Statistics (TBSS) between Groups

Between-group diffusion tensor imaging comparisons were conducted using TBSS [13]. Fractional anisotropic maps from all participants were aligned to the standard space, using non-linear image registration tool (FNIRT) in FSL. A mean voxel-wise fractional anisotropic image was generated, skeletonized, and set the threshold value of 0.2. In the *randomize* step, we performed a nonparametric 2-sample *t*-test between each subgroup and the healthy controls. With FSL Randomise, 5 000 permutations per contrast were conducted to investigate the differences between groups with age, sex and years of education as covariates [17]. The threshold-free cluster enhancement was used to generate voxel-wise probability values for multiple testing with family-wise error correction (significant at $P < 0.05$).

Supplementary Tables

Table S1. Neurocognitive tests and measurements

Cognitive Tests	Measurements	Evaluation
WAIS-RC	Verbal IQ, Performance IQ, IQ	General intelligence
CANTAB		
Rapid Visual Information Processing - RVP	RVP_TH (total hits), RVP_TM (total miss), RVP_TFA (total false alarms), RVP_TCR (total correct rejections), RVP_PFA (probability of false alarm); RVP_ML (mean latency)	Sustained attention and inhibition

	DMS_MCL (mean correct latency),
	DMS_MCLA (mean correct latency in all
	delay), DMS_MCLS (mean correct latency in
Delayed Matching to Sample	simultaneous); DMS_TC (total correct),
- DMS	DMS_TCA (total correct in all delays),
	DMS_TCS (total correct in all simultaneous),
	DMS_PEGC (Probability error given correct)

Immediate and delayed visual memory

Reaction time is in milliseconds (ms).

Table S2. Comparison of mean FA between control group and patient subgroups across 20 fiber tracts

	Subgroup1	Subgroup2	Subgroup3	HC	Post hoc (P)		
					Sub1 vs. HC	Sub2 vs. HC	Sub3 vs. HC
ATR_L	0.458 (0.008)	0.467 (0.008)	0.482 (0.011)	0.476 (0.022)	<0.0001**	0.024*	0.372
ATR_R	0.456 (0.009)	0.466 (0.010)	0.477 (0.009)	0.471 (0.023)	<0.0001**	0.448	0.363
CST_L	0.553 (0.011)	0.562 (0.014)	0.571 (0.014)	0.579 (0.026)	<0.0001**	<0.0001**	0.348
CST_R	0.538 (0.014)	0.550 (0.012)	0.560 (0.012)	0.567 (0.027)	<0.0001**	<0.0001**	0.316
CGC_L	0.489 (0.147)	0.494 (0.012)	0. 506(0.012)	0.517 (0.230)	<0.0001**	<0.0001**	0.026*
CGC_R	0.471 (0.015)	0.489 (0.012)	0.495 (0.012)	0.507 (0.027)	<0.0001**	0.313	0.031*
CGH_L	0.396 (0.020)	0.419 (0.026)	0.431 (0.025)	0.441 (0.035)	<0.0001**	0.087	0.441

|							
CGH_R	0.424 (0.0230	0.449 (0.020)	0.463 (0.018)	0.451 (0.032)	<0.0001**	0.998	0.106
Fmaj	0.532 (0.011)	0.541 (0.009)	0.556 (0.007)	0.554 (0.020)	<0.0001**	<0.0001**	0.987
Fmin	0.518 (0.010)	0.525 (0.010)	0.544 (0.011)	0.536 (0.022)	<0.0001**	<0.0001**	0.16
IFO_L	0.452 (0.012)	0.466 (0.010)	0.480 (0.010)	0.474 (0.021)	<0.0001**	0.025*	0.225
IFO_R	0.455 (0.010)	0.474 (0.009)	0.485 (0.009)	0.478 (0.020)	<0.0001**	0.687	0.088
ILF_L	0.419 (0.010)	0.434 (0.012)	0.450 (0.009)	0.445 (0.024)	<0.0001**	0.005*	0.581
ILF_R	0.395 (0.010)	0.413 (0.009)	0.426 (0.008)	0.419 (0.022)	<0.0001**	0.195	0.142
SLF_L	0.440 (0.008)	0.454 (0.009)	0.466 (0.009)	0.464 (0.021)	<0.0001**	<0.0001**	0.995
SLF_R	0.439 (0.010)	0.457 (0.008)	0.466 (0.008)	0.459 (0.021)	<0.0001**	0.942	0.156
UNC_L	0.419 (0.016)	0.430 (0.014)	0.443 (0.013)	0.439 (0.025)	<0.0001**	0.072	0.855
UNC_R	0.421 (0.013)	0.433 (0.013)	0.447 (0.014)	0.444 (0.026)	<0.0001**	0.008*	0.988
SLFt_L	0.429 (0.020)	0.442 (0.018)	0.454 (0.018)	0.456 (0.029)	<0.0001**	0.034*	0.988
SLFt_R	0.473 (0.014)	0.489 (0.014)	0.492 (0.015)	0.503 (0.024)	<0.0001**	0.929	0.03*

Fmaj, forceps major; Fmin, forceps minor; ATR, anterior thalamic radiation; CGC, gyrus part of the cingulum cingulate; CGH, hippocampal part of the cingulum; CST, corticospinal tract; IFO, inferior fronto-occipital fasciculus; SLFt, temporal part of the superior longitudinal fasciculus; UNC, uncinate fasciculus; ILF, inferior longitudinal fasciculus; SLF, superior longitudinal fasciculus. L, left. R, right.

* represents subgroup compared to HC, P <0.05

** represents subgroup compared to HC, P <0.005

Supplementary Figures

Fig. S1 Cluster validation - gap value
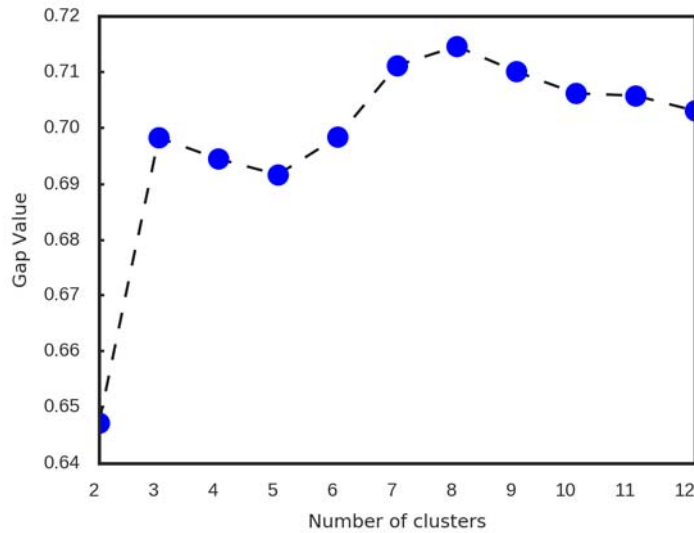
The X-axis represents the number of clusters. The Y-axis represents gap value. The plot of validity has the number of clusters from 2 to 12. Our cluster result has the "elbow" with gap value 0.6983 when cluster number equals to 3, suggesting the optimal cluster number that best represent the data structure is 3.
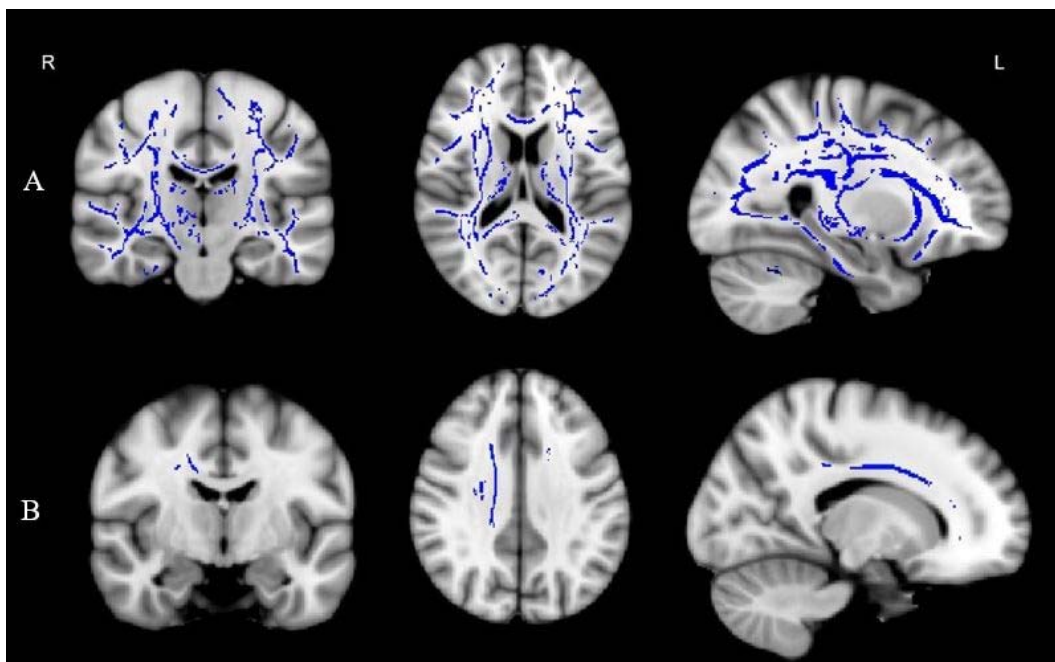
Fig. S2 Patterns of white matter alternations between patient subgroup vs. control group
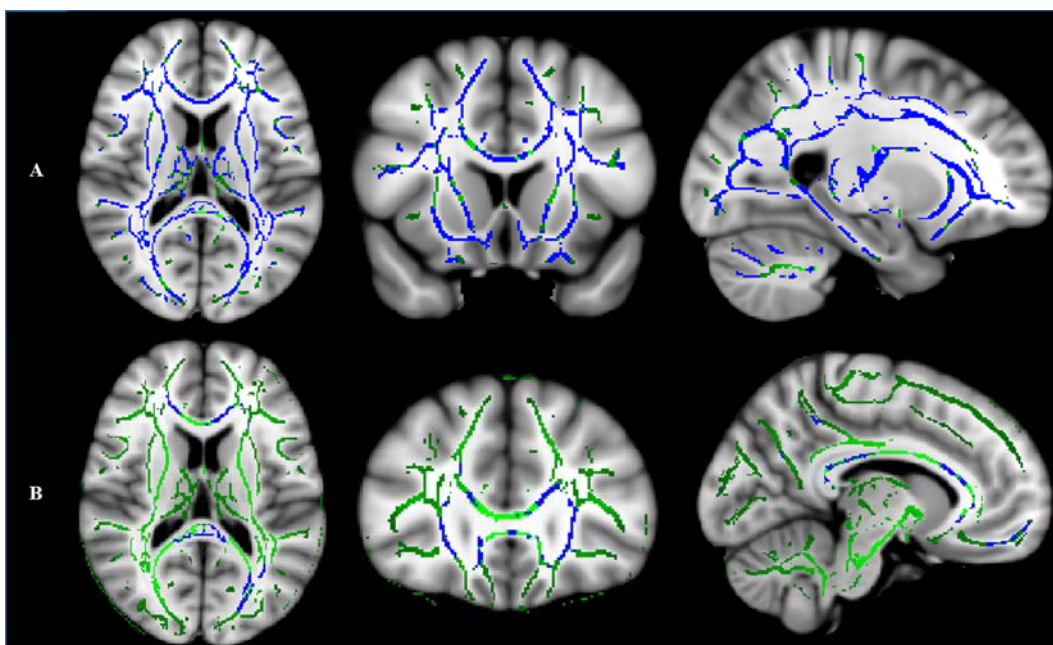
A Subgroup 1 compared to control group had widespread white matter disruptions. **B** Subgroup 2 compared to control group had regional white matter alternations, mainly in portion of corpus callosum and part of left cingulum cingulate.

Fig. S3 Patterns of white matter alternations between patient subgroups

A Subgroup 1 compared to Subgroup 2 had widespread white matter disruptions. **B** Subgroup 2 compared to Subgroup 3 tended to have regional white matter alternations in portion of right corpus callosum.

References

1. First M, Spitzer R, Williams J, Gibbon M. Structured clinical interview for DSM-IV-non-patient edition (SCID-NP, version 1.0). American Psychiatric, Washington, DC 1995.

2. First M, Spitzer R, Williams J, Gibbons M. Structured clinical interview for DSM-IV-Patient version. New York: Biometrics Research Department, New York State Psychiatric Institute 1995.

3. First MB. Diagnostic and statistical manual of mental disorders. DSM IV-4th edition. APA. 1994: 97–327.

4. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960, 23: 56–62.

5. Hamilton M. Hamilton anxiety rating scale (HAM-A). Br J Med Psychol 1959, 32: 50–55.

6. Ward LC. Prediction of verbal, performance, and full scale IQs from seven subtests of the wais-r. J Clin Psychol 1990, 46: 436–440.

7. Ryan JJ, Weilage ME, Spaulding WD. Accuracy of the seven subtest WAIS-R short form in chronic schizophrenia. Schizophr Res1999, 39: 79–83.

8. Gong Y. Manual of modified wechsler adult intelligence scale (WAIS-RC). Changsha, China: Hunan Med College 1982: 45–48.

9. Sahakian B, Owen A. Computerized assessment in neuropsychiatry using CANTAB: discussion paper. JRSM Open 1992, 85: 399.

10. Robbins TW, James M, Owen AM, Sahakian BJ, Lawrence AD, McInnes L, *et al.* A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging. Cambridge Neuropsychological Test Automated Battery. J Int Neuropsychol Soc 1998, 4: 474–490.

11. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. Neuroimage 2012, 62: 782–790.

12. Oguz I, Farzinfar M, Matsui J, Budin F, Liu Z, Gerig G*, et al.* DTIPrep: quality control of diffusion-weighted images. Front Neuroinform 2014, 8: 4.

13. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE*, et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 2006, 31: 1487–1505.

14. Mori S, Oishi K, Faria AV. White matter atlases based on diffusion tensor imaging. Curr Opin Neurol 2009, 22: 362–369.

15. Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL*, et al.* Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 2007, 36: 630–644.

16. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. J R Stat Soc Series B Stat Methodol 2001, 63: 411–423.

17. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. Neuroimage 2014, 92: 381–397.