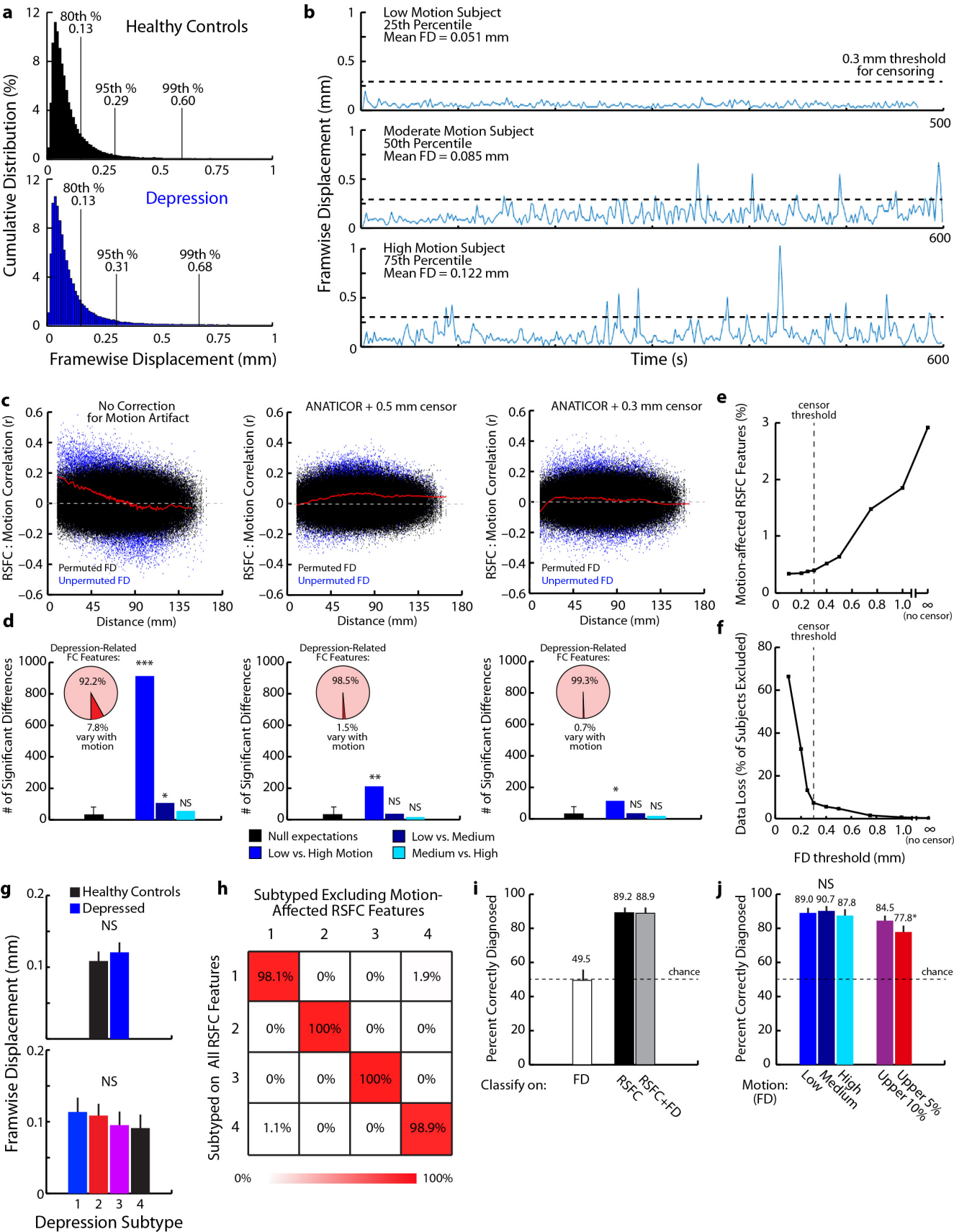<div align="center">**Supplementary Figures**</div>

Supplementary Fig. S1: Analysis and Validation of Controls for Motion-Related Artifact.

a. Cumulative distribution of head motion per volume (framewise displacement [FD] in mm) for healthy controls (upper panel) and depressed patients (lower panel) in Dataset 1. In both patients and controls, head position was stable to within 0.13 mm for the vast majority (>80%) of brain volumes.

b. Representative FD traces for low-, moderate-, and high-motion subjects at the 25th, 50th, and 75th percentiles of mean FD, respectively. Prior studies have shown that significant excursions from baseline FD values are associated with demonstrable motion artifact in the BOLD signal.[1-4] A 0.3 mm threshold effectively censored most significant excursions from baseline (FD < ~0.1 mm).

c. Prior studies have shown that motion artifacts tend to vary with neuroanatomical distance between brain nodes.[1-4] To further evaluate our controls for motion artifact, we conducted quality control analyses as described in Ref. [5], computing correlations between head motion (mean FD) and each resting state functional connectivity (RSFC) feature and plotted them as a function of neuroanatomical distance (mm) for subjects in Dataset 1 (blue data points). Smoothing curves (in red) were plotted using a moving average filter. To estimate this distribution under chance conditions (no relationship between FD and RSFC), we randomly permuted the FD values 1,000 times and performed the same analysis (black data points). To illustrate how nuisance signal regression and censoring affects these distributions, we plot them for "uncorrected" data (left panel; no nuisance signal regression, no censoring) and after nuisance signal regression using the ANATICOR method of Jo et al.[6] and censoring at a threshold of 0.5 mm (middle panel) or 0.3 mm (right panel).

d. To further quantify these effects, we divided subjects into three groups using a tertile split for mean FD and tested for group differences in each RSFC feature as in Ref. [5]. Null expectations were estimated by randomly permuting FD values 1,000 times as above (error bar = 99% confidence interval). Here, we plot the number of significant differences (as in Ref. [5], p < 0.0005 uncorrected, FDR = 0.018–0.15) for each pairwise comparison. ANATICOR nuisance signal regression with censoring at 0.3 mm (right panel) effectively corrected for most motion-related differences in RSFC values. A small number of RSFC features (N=109 or 0.3% of the 33,153 features tested) were significantly different in low vs. high motion subjects after ANATICOR regression and censoring at 0.3 mm. Insets: to illustrate how motion affected depression-related RSFC features that were used in clustering or classification, we also plot the proportion of these features that varied with motion at a more liberal threshold of p < 0.005 uncorrected. After ANATICOR regression with a 0.3 mm

<div align="center">1</div>

censor, only 0.7% of these features varied with motion even at this liberal threshold (NS vs. chance rate of 0.5%).

e.   The 0.3 mm threshold was selected by repeating the analysis above across a range of FD thresholds. Decreasing the threshold beyond 0.3 mm did not substantially reduce the proportion of motion-affected RSFC features.

f.   Subjects were excluded from further analysis if censoring resulted in an insufficient number of data points for the deconvolution step in preprocessing (~4 minutes). Prior studies have shown that estimates of RSFC coefficients are unstable in these conditions.[7] Decreasing the threshold beyond 0.3 mm resulted in substantial data loss (~35–65% for 0.1–0.2 mm) without substantial reductions in motion artifact (e).

g.   Motion (mean FD) did not differ significantly in patients vs. controls (upper panel: Wilcoxon rank sum test, $Z = 0.48$, $p = 0.63$) or by biotype (lower panel: Kruskal Wallis ANOVA, $p = 0.202$).

h.   To further evaluate whether motion artifact affected cluster assignments, we repeated the hierarchical clustering analysis depicted in Fig. 1 after excluding the 0.7% of RSFC features that varied with motion in panel d. 99.1% of all subjects were assigned to the same cluster.

i.   To rule out the possibility that multivariate classifiers may been influenced by the aggregation of subtle between-group differences in motion artifact that were undetectable by the mass univariate approach implemented in Ref. [5], we trained support vector machine (SVM) classifiers to differentiate patients and controls on the basis of FD and obtained chance accuracy rates (49.5%) in leave-one-out classification. We also tested whether including mean FD values improved the performance of the classifiers depicted in **Fig. 3g**. Including FD values did not affect performance (McNemar's test: $\chi^2 = 0.167$, $p = 0.683$).

j.   Finally, to further rule out the possibility that classification rates were enhanced by subtle, unidentified differences in head motion, we tested whether the performance of the classifiers depicted in **Fig. 3g** differed in low, medium, or high motion subjects (tertile split as in panels c-d). There were no differences between these three groups ($\chi^2 = 1.078$, $p = 0.583$). We also assessed whether classifier rates were artificially enhanced in subjects with the highest levels of motion. Classification rates were statistically indistinguishable for subjects in the highest 10% of motion (by mean FD; $\chi^2 = 1.776$, $p = 0.183$), and were slightly *lower* in the highest 5% ($\chi^2 = 5.096$, $p = 0.024$). These results indicate that subtle differences in head motion between patients and controls did not artificially enhance classifier performance.

Supplementary Fig. S1: Analysis and Validation of Controls for Motion-Related Artifact.

Supplementary Fig. S2: Canonical correlation analysis. For related data, see Fig. 1 in the main text.

a. Clinical symptom loadings on the anxiety- and anhedonia-related clinical scores are represented here as the Pearson correlations between each HAMD item and these two canonical variates, i.e. the anxiety- and anhedonia-related canonical variables for clinical symptoms. NS = not significant.

b-c. Functional connectivity feature loadings (Pearson correlations, as above) on the anxiety- (b) and anhedonia-related (c) canonical variables for RSFC features, respectively. For visualization purposes, we plot these correlations for the 25 functional nodes (ROIs) with the largest R^2 values, summed across all connectivity features associated with a given node.

4
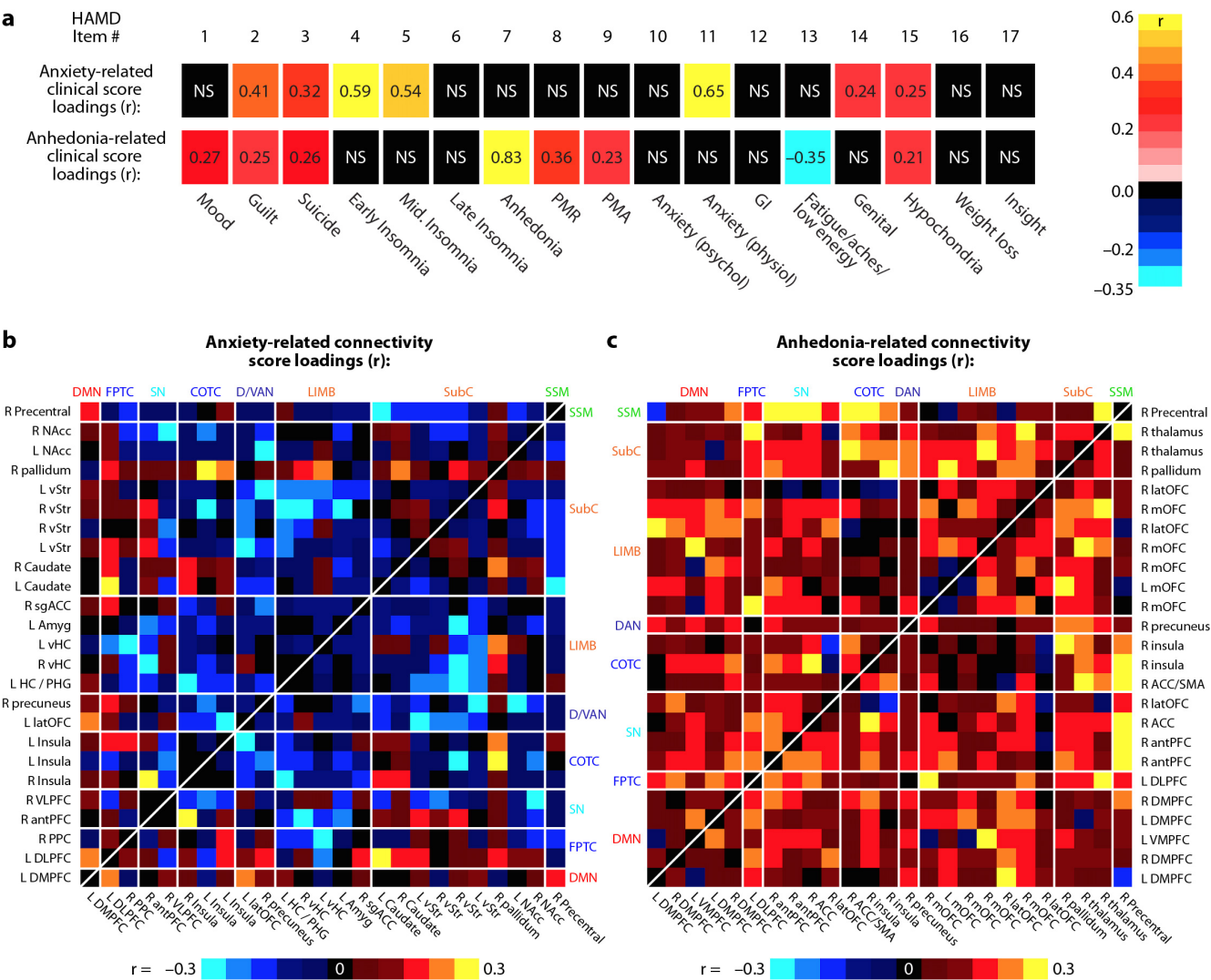
Supplementary Fig. S3: Hierarchical clustering evaluation analyses indicate an optimal four-cluster solution with stable cluster identities and stable individual cluster assignments (biotype diagnoses).

a. The variance ratio criterion (Calinski-Harabasz index),[8] defined as the ratio of between-cluster variance to within-cluster variance, is maximized at four clusters.

b. Statistical power to detect clinically meaningful differences in symptoms and treatment outcome decreases substantially for solutions with five or more clusters. Here, a 20–25% difference for an item-level response (i.e. a 1-point difference rated on a scale of 0–3 or 0–4) was considered clinically meaningful, and we plot power as a function of cluster number across the range of standard deviations (0.7–1.3) that we observed in our data, corresponding to an effect size of 0.58–0.93 (Cohen's d).

c. Minimum, mean, and maximum cluster size decrease with cluster number. With five or more clusters, the smallest cluster contained only 17 subjects, so power was substantially decreased as seen in panel b, relative to 2-, 3-, and 4-cluster solutions.

d. To evaluate the stability of the clustering solution depicted in Fig. 1d-e, we repeated this analysis in 10,000 randomly selected subsamples of the cluster discovery set (70% of the 220 subjects). This bivariate histogram depicts the frequency of cluster centroid locations, colored by biotype with arrows denoting the centroid locations from the analysis of the full dataset depicted in Fig. 1e. For scale, the relative frequencies are depicted in a 3D histogram in the right panel, illustrating highly stable clustering outcomes distributed closely around the labeled peaks. A secondary clustering solution (denoted by the dashed green line), highly similar to the modal outcome (denoted by the solid green line), occurred for 25–30% of samples. For evaluation of other cluster solutions and an illustration of comparatively unstable clustering on clinical symptoms, see Supp. Fig. S5.

e. To evaluate the stability of cluster assignments (biotype diagnoses) at the individual subject level, subjects left out of the cluster identification process (30% of the 220 subjects) in each of the 10,000 samples in panel **d** were assigned to clusters using linear discriminant analysis classifiers as described in the **Online Methods**. Cluster assignment stability was assessed by testing whether pairs of subjects assigned to the same cluster in the main analysis in Fig. 1 were also assigned to the same cluster in the left-out set for each sample, repeated for all pairwise subject combinations. Overall, 92.6% of all pairwise subject combinations were stable across the 10,000 samples (89–97% stability by cluster).

f. Overall cluster assignment stability as a function of cluster number.

Supplementary Fig. S3: Cluster evaluation analyses indicate an optimal four-cluster solution with stable cluster identities and stable individual cluster assignments (biotype diagnoses).
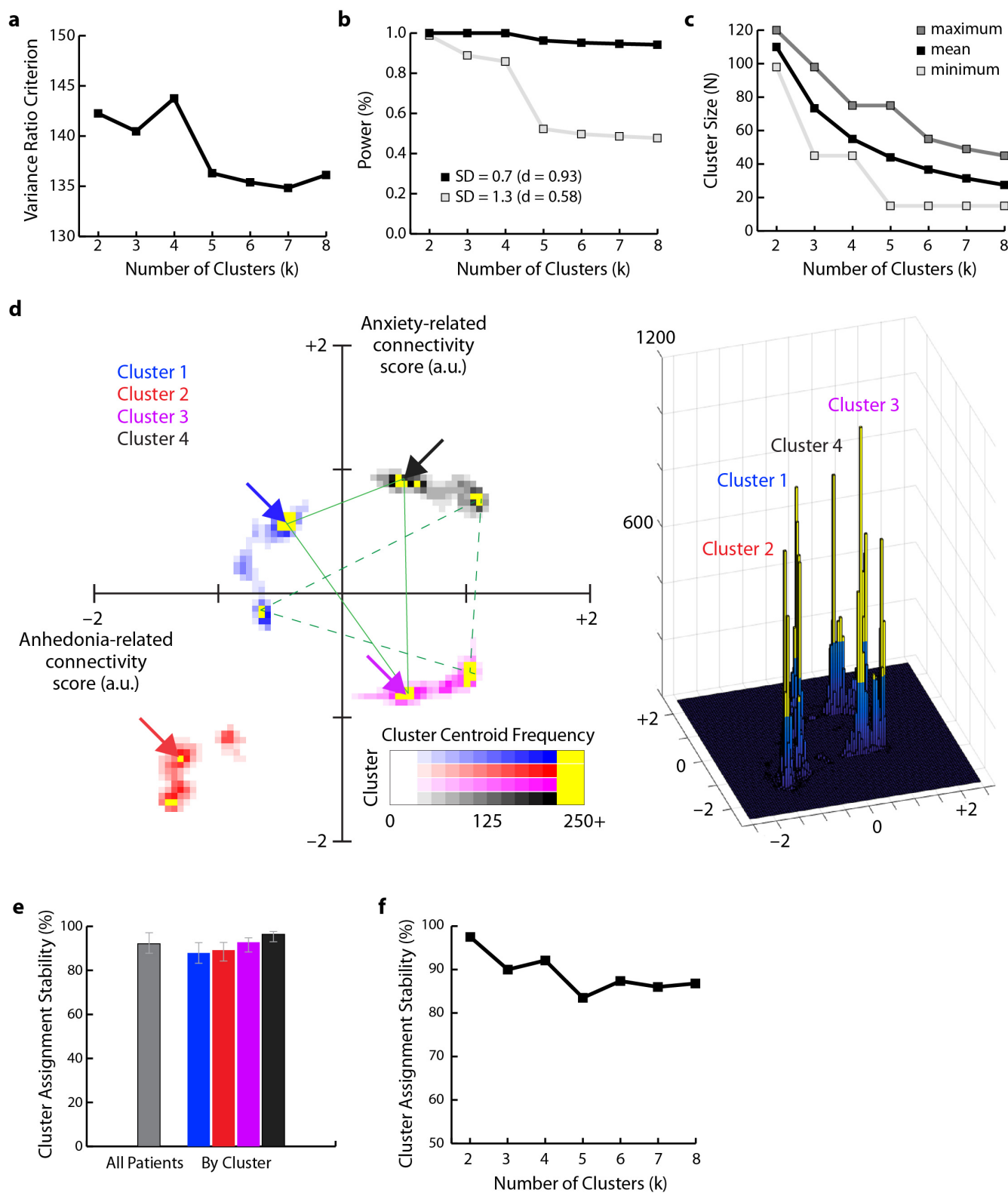
Supplementary Fig. S4: Post-hoc clinical analyses (N=220 depressed patients comprising the cluster discovery set) and replication in an independent set of subjects left out of clustering (N=92 depressed patients), the latter comprising all subjects across both datasets in Supplementary Table 3 for whom HAMD assessments were and who were not included in the 220-subject cluster discovery set.

a. Biotype differences across all 17 HAMD items in the cluster discovery set (N=220).

b. Distribution of symptom severity scores for the six depressive symptoms that varied most significantly by cluster in the cluster discovery set (N=220).

c. Convergent findings in clinical data from sites not included in clustering (N=92). For all panels, ** = Significant effect of biotype by Kruskal Wallis ANOVA at $p < 0.01$. * = Significantly greater than mean symptom severity rating for all patients ($Z = 0$) at $p < 0.05$.
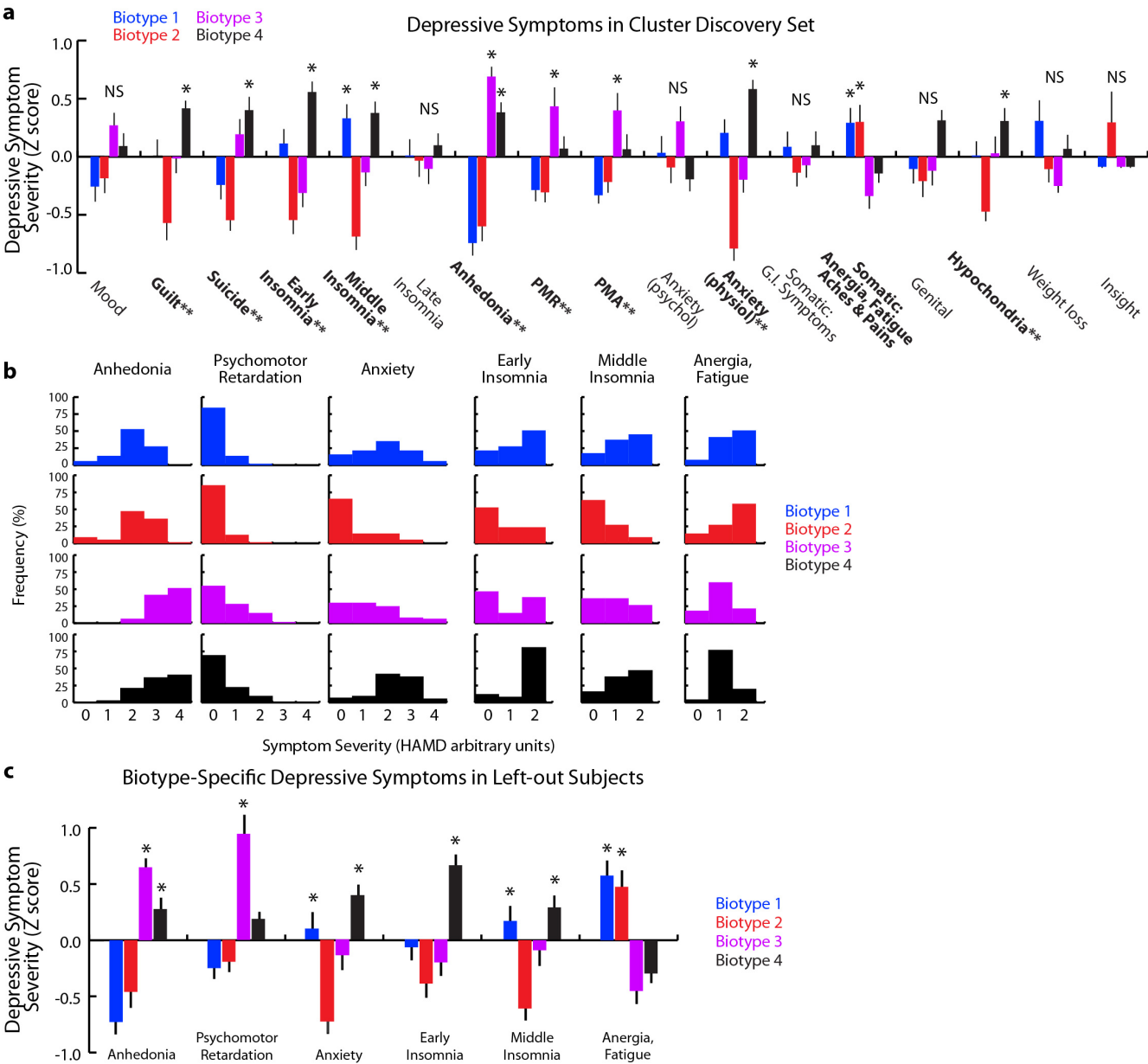
Supplementary Fig. S5: Clustering on clinical symptoms alone yields unstable solutions. Compared to clusters based on RSFC features, clustering based on clinical symptoms yields unstable outcomes, which account for relatively little variance across subjects, are longitudinally unstable, and fail to predict treatment response.

a. Dendrogram for hierarchical clustering on clinical symptoms, with evidence of relatively weak clustering into two groups. HAMD item numbers are labeled at left.

b. Variance ratio criterion is maximized at K=2 for clustering based on clinical symptoms, but it never exceeds 30. For comparison, the variance ratios for RSFC clusters are plotted in red.

c. Bivariate histograms for cluster centroid locations over 10,000 patient subsamples (70% of the full 220-subject cluster discovery set). The results indicate unstable clustering outcomes when clustering on clinical symptoms, especially for 3 or more clusters (upper panels). Comparatively stable cluster centroids for clustering on RSFC features (lower panels). Saturation denotes relative frequency and color (blue, red, violet, black) denotes cluster identity. Green denotes indeterminate cluster identity. Note that for viewing purposes, the centroid locations for clinical symptom clusters are depicted in a two-dimensional space based on the first two principal components (PC1, PC2) accounting for 39.2% of the variance across subjects. Similarly, RSFC clusters are represented in a two-dimensional space defined by the two canonical variates (CV1, CV2) described in Fig. 1.

d. Quantification of the results depicted in panel c. The variability in cluster centroids was quantified by calculating the distance between the location of each cluster centroid for a given patient subsample and the corresponding cluster centroid for the full 220-subject cluster discovery set, and then averaging across all 10,000 samples. To facilitate comparison between the RSFC and clinical symptom clusters, which were defined in distinct feature spaces, the results were normalized with respect to the average distance between each subject and every other subject in the RSFC- or clinical symptom-spaces, respectively. Clustering based on clinical symptoms yielded unstable outcomes compared to clustering based on RSFC features.

e. Longitudinal stability of subtype diagnoses based on clinical symptoms (upper panels, blue) or RSFC features (lower panels, red) in 50 subjects assessed on two occasions, 4-5 weeks apart. For additional details, see main text and Fig. 3h. Although a majority of patients (81.3–87.5%) assigned to one of the clinical symptom clusters was assigned to the same cluster 4 weeks later, all other cluster diagnoses were longitudinally unstable. By contrast, subtype diagnoses based on RSFC features were relatively stable (93-95% for 2- and 3-cluster solutions, 86–93% for 4-cluster solutions).

f. Summary of the results depicted in panel e, indicating the percent of subjects receiving stable subtype diagnoses based on clinical symptom clusters vs. RSFC features. In both cases, longitudinal stability tends to decrease with the number of clusters but is consistently higher for clustering based on RSFC features.

g-h. Clustering on clinical symptoms does not yield clusters that predict treatment outcome as indexed by the percent of subjects showing a full treatment response (defined conventionally as a 50% reduction in symptoms) and by the percent reduction in the severity of depressive symptoms (total HAMD17 score).

8

Supplementary Fig. S5: Clustering on clinical symptoms alone yields unstable solutions.
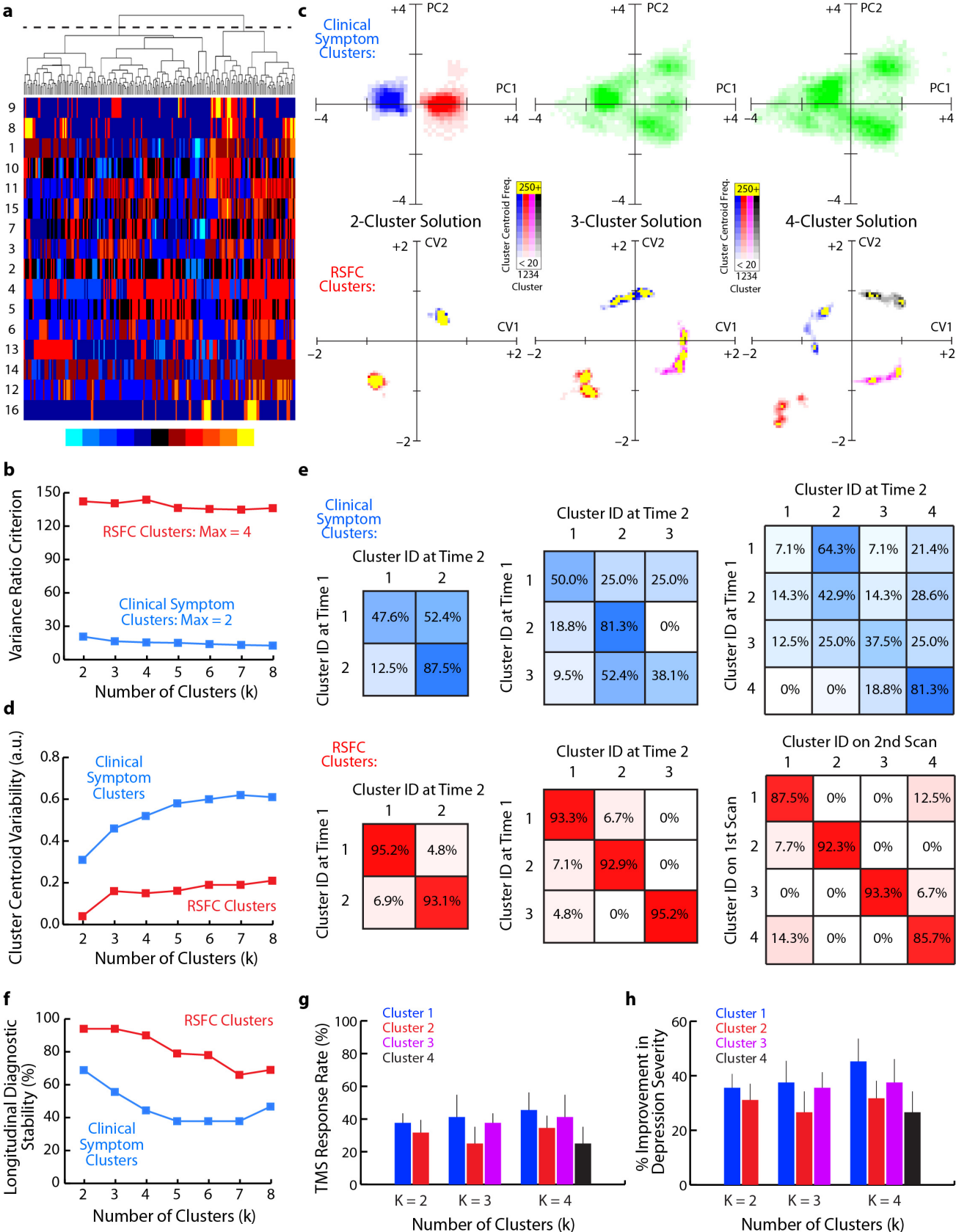
Supplementary Fig. S6: Schematic for classifier optimization. *n.b. The test subject was left out of all aspects of the training process, including clustering. For iterations when the test subject was also a member of the cluster discovery set, only 220 – 1 = 219 subjects were used in canonical correlation analysis. For additional details, see **Online Methods**, "Clustering" and "Classification" sections.
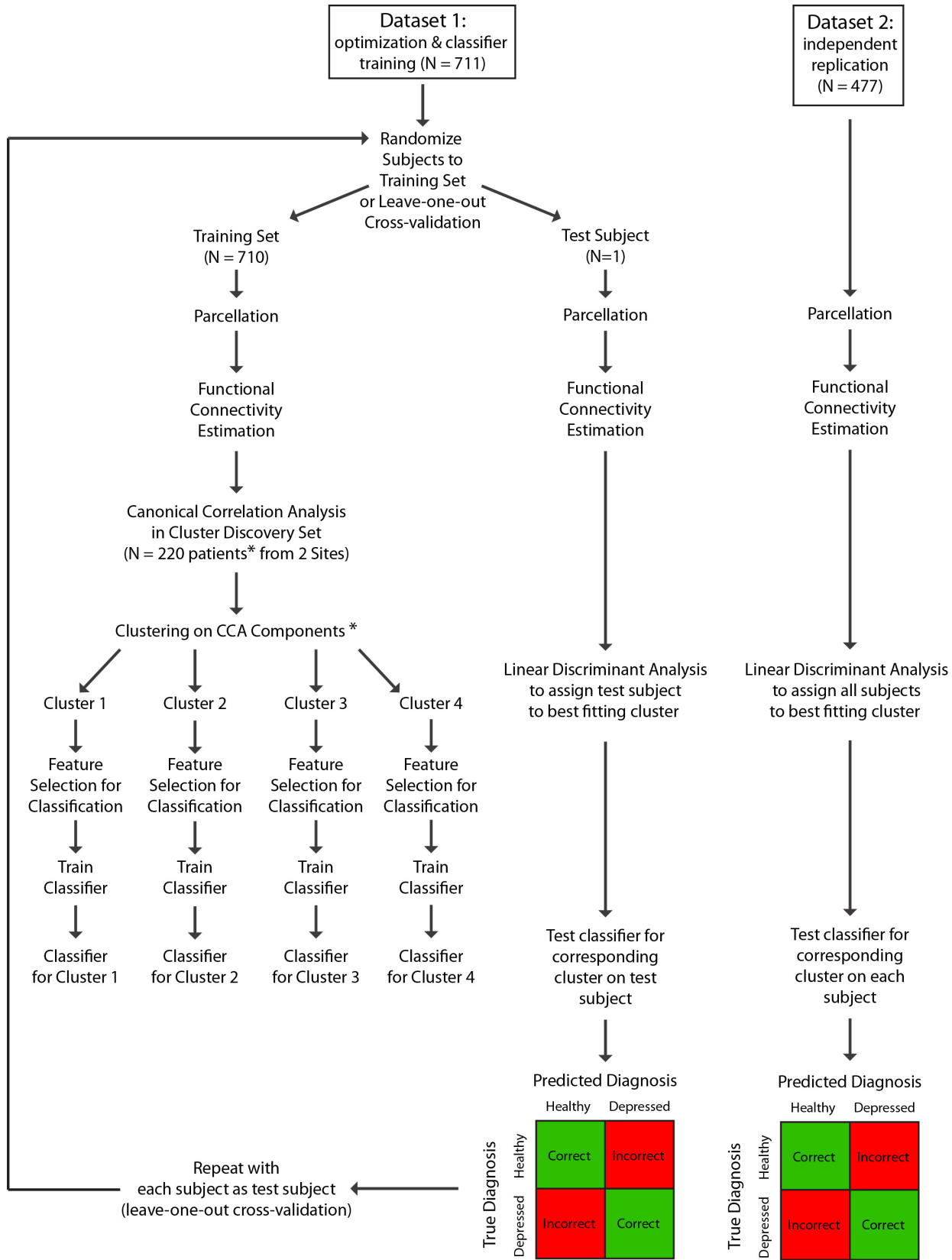
Supplementary Fig. S7:

a.   There were no significant differences in medication usage across the four clusters ($\chi^2 < 0.603$, p > 0.896). NS = not significant.

b.   The four clusters did not vary significantly in age (Kruskal Wallis ANOVA: $\chi^2 = 0.29$, p = 0.962).

c.   Head motion did not differ by cluster (Kruskal Wallis ANOVA: $\chi^2 = 4.62$, p = 0.202).
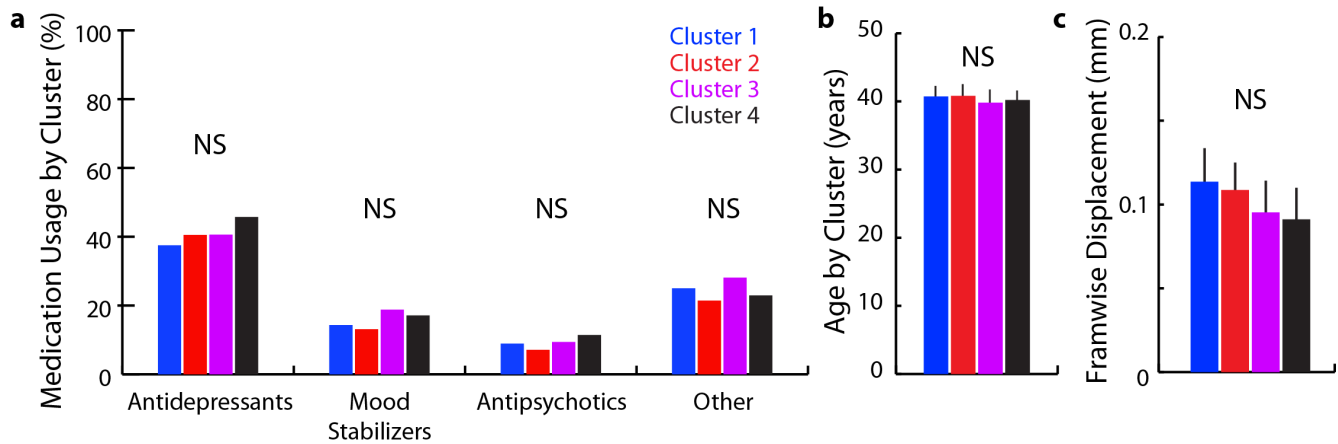
Supplementary Table 1: Demographic and clinical characteristics and MRI acquisition parameters for the 2-site cluster discovery set. To ensure that cluster discovery was not confounded by site-related differences in subject recruitment criteria or other variables, the cluster discovery analysis was restricted to this 2-site subset of Dataset 1. These two sites had identical inclusion and exclusion criteria and were matched for age ($p = 0.41$), sex ($p = 0.87$), and depression severity (HAMD17 total score; $p = 0.11$). * Other psychiatric co-morbidities = ADHD, OCD, Asperger's Syndrome, and Tourette's Syndrome. Other psychiatric medications included benzodiazepines, stimulants, thyroid hormone, and non-benzodiazepine sedative-hypnotics.

	Cornell 1	Toronto
N	96	124
Age	42.1	40.4
Sex	58.3% F	57.3% F
HAMD17 total	19.3	20.4
Inclusion Criteria	Current unipolar major depressive episode; failure to respond to at least two previous antidepressant trials	Current unipolar major depressive episode; failure to respond to at least two previous antidepressant trials
Exclusion Criteria	Bipolar disorder, psychotic disorder, active substance use disorder, active suicidal ideation, contraindications to MRI, history of seizures, current pregnancy	Bipolar disorder, psychotic disorder, active substance use disorder, active suicidal ideation, contraindications to MRI, history of seizures, current pregnancy
Psychiatric Co-morbidities		
GAD	5.2%	4.8%
PTSD	4.2%	6.5%
SAD	4.2%	4.8%
Panic D/O	3.1%	2.4%
Other*	3.1%	4.0%
Psychiatric Medications		
Antidepressant	57.3%	59.7%
Mood Stabilizer	17.7%	16.9%
Antipsychotic	15.6%	17.7%
Other*	42.7%	45.2%
rsfMRI Parameters		
Scanner	GE Signa 3T	GE Signa 3T
TR	2	2
Volumes	180	300
FOV	240	220
Slices	28	32
XY Resolution	3.75	3.44
Z Resolution	5	5

Supplementary Table 2: Inclusion criteria, exclusion criteria, psychotropic medication use, and psychiatric co-morbidities by site.

Site	Psychiatric Medications				Inclusion Criteria	Exclusion Criteria	Psychiatric Co-Morbidities
	AD	MS	AP	Other			
Cornell 1: PIs Dubin, Liston	57.3	17.7	15.6	42.7	current unipolar major depressive episode; failure to respond to at least two previous antidepressant trials	bipolar disorder, psychotic disorder, active substance use disorder, active suicidal ideation, contraindications to MRI, history of seizures, currently pregnant	GAD (5.2%), PTSD (4.2%), social anxiety disorder (4.2%), panic disorder (3.1%), other (3.1%)
Toronto: PI Downar	59.7	16.9	17.7	45.2	current unipolar major depressive episode; failure to respond to at least two previous antidepressant trials	bipolar disorder, psychotic disorder, active substance use disorder, active suicidal ideation, contraindications to MRI, history of seizures, currently pregnant	GAD (4.8%), PTSD (6.5%), social anxiety disorder (4.8%), panic disorder (2.4%), other (4.0%)
Emory 1: PI Mayberg	0	0	0	0	current unipolar major depressive episode, non-psychotic, moderate-to-severe, HAMD17 >= 18; treatment-naïve adults ages 18-65	current or past use of psychotropic medications; current or past neurologic conditions; psychotic symptoms	none
Stanford 1: PI Etkin	0	0	0	0	current unipolar major depressive episode; age >= 18	substance abuse or PTSD; history of neurologic disorder or severe mental illness (psychosis or bipolar disorder); regular use of benzodiazepines, opiates, or thyroid medications; antidepressant-free for > 6 weeks	generalized anxiety disorder (63.9%), social anxiety disorder (11.1%), panic disorder (2.8%), obsessive compulsive disorder (2.8%), bulimia (5.6%)
Stanford 2: PI Schatzberg	65.3	0	26.4	20.8	current major depressive episode with or without psychotic symptoms; HAMD21 > 18; endogenomorphic depression subscale > 6; age >= 18	illicit drug use or alcohol abuse during week before study; ECT within 4 months of study; suicidal ideation; substance abuse within 6 months of study	major depressive disorder with psychotic features (36.1%), panic disorder (6.6%), agoraphobia (6.6%), social phobia (13.1%), specific phobia (11.5%), obsessive compulsive disorder (6.6%), post traumatic stress disorder (19.7%), generalized anxiety disorder (9.8%)

Supplementary Table 3: Demographic variables and functional imaging parameters by site.

Site	Patients			Healthy Controls			rsfMRI Parameters				
	N	Age	Sex	N	Age	Sex	TR	vol	FOV	xy	z
DATASET 1	333	40.6	59.2%	378	38.0	57.7%					
aka "Training Set"											
Cornell 1	96	42.1	58.3%	28	31.1	53.6%	2	180	240	3.75	5
Emory 1	20	45.8	50.0%	20	42.5	50.0%	2.9	150	220	3.44	4
Stanford 1	35	30.9	65.7%	33	34.4	72.7%	2	240	220	3.44	4.5
Stanford 2	58	42.5	63.8%	32	37.6	68.8%	2	148	200	3.125	4.5
Toronto	124	40.4	57.3%	37	35.2	56.7%	2	300	220	3.44	5
NKI				47	44.8	65.9%	2.5	260	216	3	3
Atlanta				16	29.9	62.5%	2	205	220	3.438	4
Cambridge				26	24.7	65.4%	3	119	216	3	3
Cleveland				25	41.4	64.0%	2.8	127	256	2	4
ICBM				58	48.2	48.3%	2	128	256	4	4
New York				42	33.8	50.0%	2	192	192	3	3
COBRE				14	37.5	21.4%	2	150	192	3	4
DATASET 2	125	49.8	58.4%	352	32.1	57.7%					
aka "Replication Set"											
Cornell 1	35	45.1	62.9%	35	25.8	42.9%	2	180	240	3.75	5
Cornell 2	27	71.3	55.6%	8	69.8	62.5%	2	170	224	3.5	5
Stanford 1	1	54.0	0.0%	5	31.4	60.0%	2	240	220	3.44	4.5
Stanford 2	13	48.9	61.5%	5	20.0	40.0%	2	148	200	3.125	4.5
NKI	3	33.3	33.3%	68	38.7	45.6%	2.5	260	216	3	3
Toronto	30	43.5	60.0%	11	31.4	63.6%	2	300	220	3.44	5
Harvard	16	39.5	56.3%	18	39.4	61.1%	3.2	124	240	3.75	3
Cambridge				71	20.7	63.4%	3	119	216	3	3
ICBM				12	21.3	75.0%	2	128	256	4	4
Beijing				44	23.9	63.6%	2	225	200	3.125	3.6
Milwaukee				38	55.2	71.1%	2	175	240	3.75	4
Leipzig				24	28.5	58.3%	2.3	195	192	3	4
New York				13	25.6	46.2%	2	192	192	3	3

Supplementary Table 4: MNI coordinates (mm) of ROIs in Figs. 1c-d. Abbreviations as described in main text.

ROI	Network	MNI Coordinates		
		X	Y	Z

Fig 1c: canonical correlation analysis: anhedonia-related connectivity, top 25 ROIs

ROI	Network	X	Y	Z
L DMPFC	DMN	-16	29	53
R DMPFC	DMN	13	30	59
L VMPFC	DMN	-7	51	-1
L DMPFC/ACC	DMN	-3	42	16
R DMPFC/ACC	DMN	12	36	20
L DLPFC	FPTC	-42	38	21
R antPFC	SN	31	56	14
R antPFC	SN	26	50	27
R ACC	SN	5	23	37
R VLPFC/insula	SN	34	16	-8
R SMA/ACC	COTC	7	8	51
R insula	COTC	49	8	-1
R insula	COTC	36	10	1
R precuneus	DAN	22	-65	48
R mOFC	LIMB/SubC	8	48	-15
L mOFC	LIMB/SubC	-18	63	-9
R VMPFC/mOFC	LIMB/SubC	8	42	-5
R VLPFC/latOFC	LIMB/SubC	49	35	-12
R VLPFC/latOFC	LIMB/SubC	43	49	-2
R mOFC	LIMB/SubC	24	32	-18
R VLPFC/latOFC	LIMB/SubC	27	16	-17
R globus pallidus	LIMB/SubC	15	5	7
R thalamus	LIMB/SubC	6	-24	0
R thalamus	LIMB/SubC	12	-17	8
R postcentral	AV	59	-17	29

Fig 1d: canonical correlation analysis: anxiety-related connectivity, top 25 ROIs

ROI	Network	X	Y	Z
L DMPFC	DMN	-20	45	39
L premotor	FPTC	-23	11	64
R PPC	FPTC	37	-65	40
R antPFC	SN	31	56	14
R VLPFC	SN	48	22	10
R insula	COTC	49	8	-1
L insula	COTC	-45	0	9
L insula	COTC	-34	3	4
L VLPFC/latOFC	VAN	-49	25	-1
R precuneus	DAN	10	-62	61
L HC/PHG	LIMB/SubC	-21	-22	-20
R vHC	LIMB/SubC	27	-15	-18
L vHC	LIMB/SubC	-27	-15	-18
L Amyg	LIMB/SubC	-19	-2	-21
R sgACC	LIMB/SubC	4	15	-11
L Caudate	LIMB/SubC	-12	18	-3
R Caudate	LIMB/SubC	12	18	-3
L ventral striatum	LIMB/SubC	-22	7	-5
R ventral striatum	LIMB/SubC	23	10	1
R ventral striatum	LIMB/SubC	29	1	4
L ventral striatum	LIMB/SubC	-31	-11	0
R thalamus	LIMB/SubC	9	-4	6
L nucleus accumbens	LIMB/SubC	-12	8	-8
R nucleus accumbens	LIMB/SubC	12	8	-8
R precentral	SSM	51	-6	32

Supplementary Table 5. MNI coordinates (mm) for the top 25 ROIs with the most abnormal connectivity features shared by all four biotypes in Fig. 2a (left), and for the top 50 ROIs that varied by subtype in Fig. 2d. Additional ROIs depicted in the heat maps in Figs. 2b and 2e are listed at the end. Abbreviations are as described in the main text.

Figs. 2a-b		MNI Coordinates		
ROI	Network	X	Y	Z
L PCC	DMN	-11	-56	16
L fusiform	DMN	-34	-38	-16
L MTG	DMN	-56	-13	-10
L pHC	DMN	-26	-40	-8
R pHC	DMN	27	-37	-13
L VLPFC/insula	SN	-35	20	0
R Insula	COTC	37	1	-4
R insula	COTC	36	10	1
L STG	VAN	-55	-40	14
R VMPFC/mOFC	LIMB/SubC	8	42	-5
R OFC/VMPFC	LIMB/SubC	24	45	-15
R OFC	LIMB/SubC	24	32	-18
R vHC	LIMB/SubC	27	-15	-18
L amyg	LIMB/SubC	-19	-2	-21
R amyg	LIMB/SubC	19	-2	-21
R sgACC	LIMB/SubC	4	15	-11
L ventral striatum	LIMB/SubC	-31	-11	0
R thalamus/GP	LIMB/SubC	9	-4	6
L thalamus	LIMB/SubC	-2	-13	12
L thalamus	LIMB/SubC	-10	-18	7
R postcentral/S1	SSM	50	-20	42
R precentral/M1	SSM	13	-33	75
R insula	AV	32	-26	13
R STG	AV	58	-16	7
L fusiform	AV	-47	-51	-21
L VMPFC/ACC	DMN	-11	45	8
R Precuneus	DMN	15	-63	26
L MTG	DMN	-53	3	-27
L MTG	DMN	-49	-42	1
L DLPFC	FPTC	-47	11	23
L PPC	FPTC	-28	-58	48
R Insula	COTC	49	8	-1
L temporal pole	COTC	-51	8	-2
L VLPFC	VAN	-49	25	-1
R STG/PPC	VAN	54	-43	22
L MTG	VAN	-56	-50	10
R PPC	DAN	25	-58	60
R mOFC	LIMB/SubC	8	48	-15
L VMPFC/mOFC	LIMB/SubC	-3	44	-9
L HC/PHG	LIMB/SubC	-21	-22	-20
L vHC	LIMB/SubC	-27	-15	-18
L VLPFC/latOFC	LIMB/SubC	-31	19	-19
L sgACC	LIMB/SubC	-4	15	-11
L premotor	SSM	-7	-21	65
R premotor/SMA	SSM	2	-28	60
R insula	SSM	36	-9	14
L STG	AV	-49	-26	5
L STG	AV	-60	-25	14
L lingual	AV	-15	-72	-8
L lingual	AV	-16	-52	-1

Fig. 2d-e		MNI Coordinates		
ROI	Network	X	Y	Z
R DMPFC	DMN	23	33	48
L DMPFC	DMN	-20	45	39
R DMPFC	DMN	6	54	16
R DMPFC	DMN	6	64	22
L DMPFC	DMN	-8	48	23
R PCC	DMN	8	-48	31
L PCC	DMN	-3	-49	13
R MTG	DMN	65	-31	-9
R MTG	DMN	-53	3	-27
R paraHC	DMN	27	-37	-13
L premotor	FPTC	-23	11	64
L DLPFC	FPTC	-34	55	4
R DLPFC	FPTC	38	43	15
R PPC	FPTC	44	-53	47
R MTG	FPTC	58	-53	-14
R antPFC	SN	26	50	27
L ACC	SN	-11	26	25
R ACC	SN	10	22	27
L ACC	COTC	-10	-2	42
L ACC/SMA	COTC	-3	2	53
R SMA/ACC	COTC	7	8	51
L insula	COTC	-45	0	9
L insula	COTC	-34	3	4
R MTG	VAN	51	-29	-4
R PPC	DAN	25	-58	60
L mOFC	LIMB/SubC	-18	63	-9
R VLPFC/latOFC	LIMB/SubC	-46	31	-13
R VLPFC/latOFC	LIMB/SubC	49	35	-12
R VMPFC/OFC	LIMB/SubC	24	45	-15
R VLPFC/latOFC	LIMB/SubC	43	49	-2
R OFC	LIMB/SubC	24	32	-18
L HC/PHG	LIMB/SubC	-21	-22	-20
sgACC R	LIMB/SubC	4	15	-11
R thalamus/GP	LIMB/SubC	9	-4	6
R thalamus	LIMB/SubC	6	-24	0
L thalamus	LIMB/SubC	-10	-18	7
R thalamus	LIMB/SubC	12	-17	8
raphe	LIMB/SubC	6	23	-16
L postcentral/S1	SM	-45	-32	47
R precentral/M1	SM	42	-20	55
R postcentral/S1	SM	10	-46	73
R precentral/M1	SM	44	-8	57
R SMA	SM	2	-28	60
R postcentral	AV	59	-17	29
L lingual	AV	-12	-95	-13
R fusiform	AV	43	-78	-12
L inferior occipital	AV	-40	-88	-6
R middle occipital	AV	37	-84	13
R middle occipital	AV	37	-81	1
L middle occipital	AV	-42	-74	0
R PPC	FPTC	37	-65	40
L PPC	FPTC	-42	-55	45
R VLPFC	SN	37	32	-2
L VLPFC	VAN	-49	25	-1
R DLPFC	DAN	29	-5	54
R precuneus	DAN	22	-65	48
L precuneus	DAN	-27	-71	37
L PPC	DAN	-33	-46	47
R vHC	LIMB/SubC	27	-15	-18
L vHC	LIMB/SubC	-27	-15	-18
L amyg	LIMB/SubC	-19	-2	-21
R amyg	LIMB/SubC	19	-2	-21
L ventral striatum	LIMB/SubC	-15	4	8
R ventral striatum	LIMB/SubC	31	-14	2

Supplementary Table 6: Diagnostic classifier accuracy by site in the independent replication dataset

Site	Patients		Healthy Controls	
	N	Accuracy	N	Accuracy
Cornell 1	35	88.6%	35	91.4%
Cornell 2	27	81.5%	8	87.5%
Stanford 1	1	100.0%	5	60.0%
Stanford 2	13	92.3%	5	80.0%
NKI	3	66.7%	68	82.4%
Toronto	30	86.7%	11	81.8%
Harvard	16	68.8%	18	77.8%
Cambridge			71	88.7%
ICBM			12	100.0%
Beijing			44	84.1%
Milwaukee			38	89.5%
Leipzig			24	91.7%
New York			13	100.0%
Total	125	84.0%	352	86.9%

Supplementary Table 7: MNI coordinates (mm) for the top 25 ROIs differentiating rTMS responders and non-responders in Fig. 4d. Additional ROIs depicted in the heat maps in Fig. 4e are listed at the end. Abbreviations are as described in the main text.

ROI	Network	MNI Coordinates		
		X	Y	Z
R PCC	DMN	8	-48	31
L DLPFC	FPTC	-47	11	23
R DLPFC	FPTC	48	25	27
R DLPFC	FPTC	38	43	15
L PPC	FPTC	-42	-55	45
B/L ACC	SN	0	30	27
L DLPFC/Premotor	DAN	-32	-1	54
R PPC	DAN	25	-58	60
R mOFC	LIMB/SubC	6	67	-4
L mOFC	LIMB/SubC	-18	63	-9
L VLPFC/latOFC	LIMB/SubC	-42	45	-2
R VLPFC/latOFC	LIMB/SubC	43	49	-2
R VLPFC/latOFC	LIMB/SubC	27	16	-17
L amyg	LIMB/SubC	-19	-2	-21
L VLPFC/latOFC	LIMB/SubC	-31	19	-19
R globus pallidus	LIMB/SubC	15	5	7
R thalamus/GP	LIMB/SubC	9	-4	6
L nucleus accumbens	LIMB/SubC	-12	8	-8
L thalamus	LIMB/SubC	-5	-28	-4
VTA	LIMB/SubC	6	-15	-15
L postcentral/S1	SSM	-29	-43	61
L precentral/M1	SSM	-40	-19	54
L postcentral/S1	SSM	-21	-31	61
R lingual	AV	17	-91	-14
L middle occipital	AV	-26	-90	3
R DMPFC/SFG	DMN	23	33	48
L DMPFC/SFG	DMN	-10	39	52
L DMPFC/SFG	DMN	-16	29	53
R DMPFC/SMdG	DMN	13	30	59
L DMPFC/SMdG	DMN	-2	38	36
R DMPFC/SMdG	DMN	6	54	16
R DMPFC/SMdG	DMN	6	64	22
L DMPFC/SMdG	DMN	-8	48	23
L DMPFC/ACC	DMN	-3	42	16
L ACC/SMdG	FPTC	-3	26	44
L ACC	SN	-11	26	25
L ACC	SN	0	30	27
R nucleus accumbens	LIMB/SubC	12	8	-8