Supplemental Methods

Procedures

Recording and testing conditions were similar and stimulus presentation and recording equipment were identical across sites. Experimenters across sites also were trained and continually monitored to ensure comparable laboratory data collection procedures. As a result, there were no site effects that influenced group comparisons on any laboratory biomarker measure (see Clementz 2016). Individual participants with missing data were excluded modality-wise. Phenotype assessment often included multiple individual variables for each task, paradigm, or imaging modality.

All biomarker measures were adjusted for age and sex. For each measure, age and sex effects were calculated using linear regression on the healthy control group from BSNIP-1 (N=459).[24,25] Measures with significant age effects were adjusted in probands by subtracting the product of the linear regression coefficient and age for each individual, an approach we have taken in previous B-SNIP publications.[70,75,80,83,85,86,88–91]

Laboratory Tasks

The B-SNIP phenotyping battery was collected on each participant, including neurocognitive assessments, electrophysiological measures, and structural and functional brain imaging.

Cognition: The *Brief Assessment of Cognition in Schizophrenia (BACS)* battery assesses multiple cognitive functions, with psychosis-related cognitive impairment best indicated by a global score.[42,81] Age-, sex-, and race-stratified normative data were used to compute composite scores for each participant.[81]

The *Stop Signal Task* measures the efficiency and adequacy of cognitive control when response activation and generation regarding a single stimulus location are placed in conflict.[43] Individuals with psychosis and unaffected first-degree relatives show delayed response times and increased errors.[82]

The *Dot Probe Expectancy Task* assesses goal maintenance and context processing, requiring use contextual information to overcome a prepotent response.[44,45] Individuals with psychosis and their first-degree relatives show reduced targets detection and elevated false alarm rates.[83]

The *Penn Emotion Recognition Test (ER-40)* is a standardized measure of facial emotion recognition from the University of Pennsylvania Computerized Neurocognitive Test Battery that depicts 40 color photographs of faces displaying expressions for four basic emotions (happiness, sadness, anger, and fear) and no emotion (neutral).[48] Intended emotions displayed in photographs are based on those reported by healthy raters viewing the photographs.[49] Participants are asked to examine the faces and decide what emotion the person is showing.

The *Wechsler Memory Scale (WMS-III) Spatial Span* subtest assesses maintenance and manipulation aspects of spatial working memory.[75] Psychosis probands and their first-degree relatives exhibit a similar pattern of robust working memory deficits, although these deficits are largely attenuated when controlling for generalized cognitive deficits.[75]

Oculomotor: Smooth pursuit eye movement tracking assesses ability to visually track slowly moving objects; deficits can result from disturbances in neural circuitry involving motion sensitive visual area V5, parietal and frontal areas supporting sensorimotor transformation, and subcortical areas involved in motor control.[50,51,54,55] Psychosis probands show deficits on both sensorimotor and cognitive aspects of the task,[52,56–61,76,77] which have different associations with genes regulating dopamine and glutamate systems in psychotic disorders,[53,78] while only early sensorimotor function is impaired in psychosis relatives.[79]

Prosaccades assess speed of visual orienting,[63] with individuals with psychosis showing either decreased or increased response times.[64] *Antisaccades* assess inhibitory control under perceptual conflict because the visual stimulus and required response location are incompatible.[62] Individuals with psychosis have high antisaccade error rates.[64,80] These deficits are found to a lesser extent in biological relatives of individuals with psychosis[66] with performance increasing with genetic distance from the proband,[64] making

antisaccade performance one of the most robust[80] and widely replicated[65] behavioral endophenotypes for psychosis.

Electrophysiology: Electroencephalography (EEG) data were collected from 64 Ag/AgCl sensors (Quik-Cap, Compumedics Neuroscan, El Paso, TX) positioned according to the 10-10 EEG system (with the inclusion of mastoids and CP1/2 locations to provide for greater signal sampling below the cantho-meatal line), with a forehead ground and nose reference. EEG recordings were amplified (12,500×) and digitized (1000 Hz) using Neuroscan Acquire and Synamp2 (Compumedics, Charlotte, NC). Additional sensors located above and below the eyes, as well as at the outer canthi of each eye, recorded blinks and eye movements. Detailed processing and analysis procedures have been published previously.[69–71] Evoked brain responses to repetitive auditory stimuli (*paired stimuli paradigm*) and predetermined auditory targets randomly interspersed with nontarget (or "standard") auditory events (*oddball paradigm*) are deviant in individuals with psychosis.[69,70] These paradigms assess the neural dynamics of preparation for and recovery from auditory sensory activations, neural responses to stimulus salience, and neural differentiation of relevant from irrelevant auditory stimuli.[23] *Resting-state EEG* activity was examined using independent component analysis.[71] Eight independent resting-state EEG spectral components and associated spatial weights were derived with Group ICA using EEGIFT (http://icatb.sourceforge.com; also see Narayanan 2014).

Imaging: All subjects underwent a single 5-minute run of *resting-state fMRI* on a 3T MRI scanner at each site.[72] Credible resting-fMRI connectivity networks (as opposed to physiological/susceptibility artifacts) identified with Group ICA using GIFT v1.3f (http://mialab.mrn.org/software/gift; also see Meda et al., 2016). Fractional anisotropy measures of white matter integrity, assessed with *diffusion tensor imaging (DTI)*, were collected at two of the five B-SNIP1 sites.[74] Differences in white matter structure and integrity have been observed both as a global measure and in specific structures (see Skudlarski et al., 2013). Fractional anisotropy sums for 50 individual white matter tracts were considered.

Biotype Creation

Biotype creation and classification was based on BACS composite score and PCA-reduced data from pro-/anti-saccades, stop signal task, and evoked EEG responses to auditory oddball and paired stimuli paradigms. The resulting nine composite variables were subjected to *k*-means clustering, with the optimal number of subgroups to extract (3) being previously determined and verified using multiple estimation methods. The greater distinctiveness of these biological phenotypes, or "Biotypes", over clinical DSM diagnoses was verified with multiple external validators. Detailed procedures can be found in Clementz et al. (2016).

Supplemental Analyses

Specificity of Negative Symptom Associations

To assess the specificity of the currently identified biomarker associations for negative symptoms, the CCA analyses were repeated using PANSS positive symptom items rather than negative. The cognition battery showed an association with positive symptoms ($F_{91,2159.1}$ = 1.37, Wilks' λ = .706, Canonical Correlation = .35, p = .014). Specifically, more severe conceptual disorganization (P2) and hallucinations (P3) and decreased hostility (P7) correlated with decreased performance on a number of cognitive tasks. No other biomarker modality showed a significant relationship with positive symptoms.

Supplemental Tables

Supplemental Table 1. Mean PANSS Negative Symptom Factor Total (N1 + N2 + N3 + N4 + N6 + G7 + G16) scores as a function of MADRS inclusion criteria

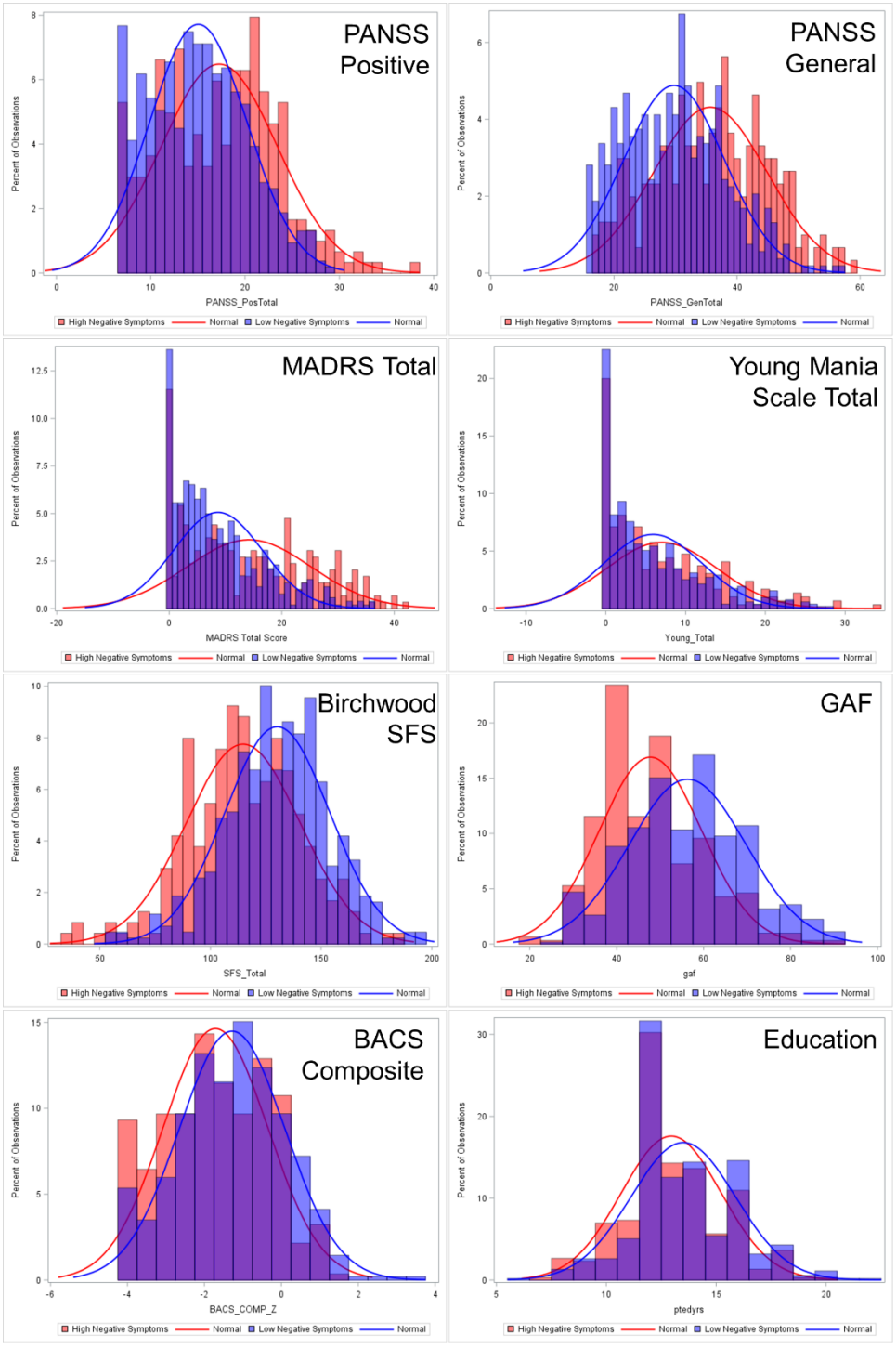
	NS			Non-NS		
	Cumulative N	Mean	(SD)	Cumulative N	Mean	(SD)
MADRS ≤ 19	211	19.3	(5.4)	454	11.0	(3.4)
MADRS ≤ 34	302	19.6	(5.3)	499	11.1	(3.4)
No MADRS criteria	322	19.7	(5.4)	515	11.1	(3.4)

Supplemental Table 2. Standardized Factor Loadings for the Two-Factor Model of Negative Symptoms.

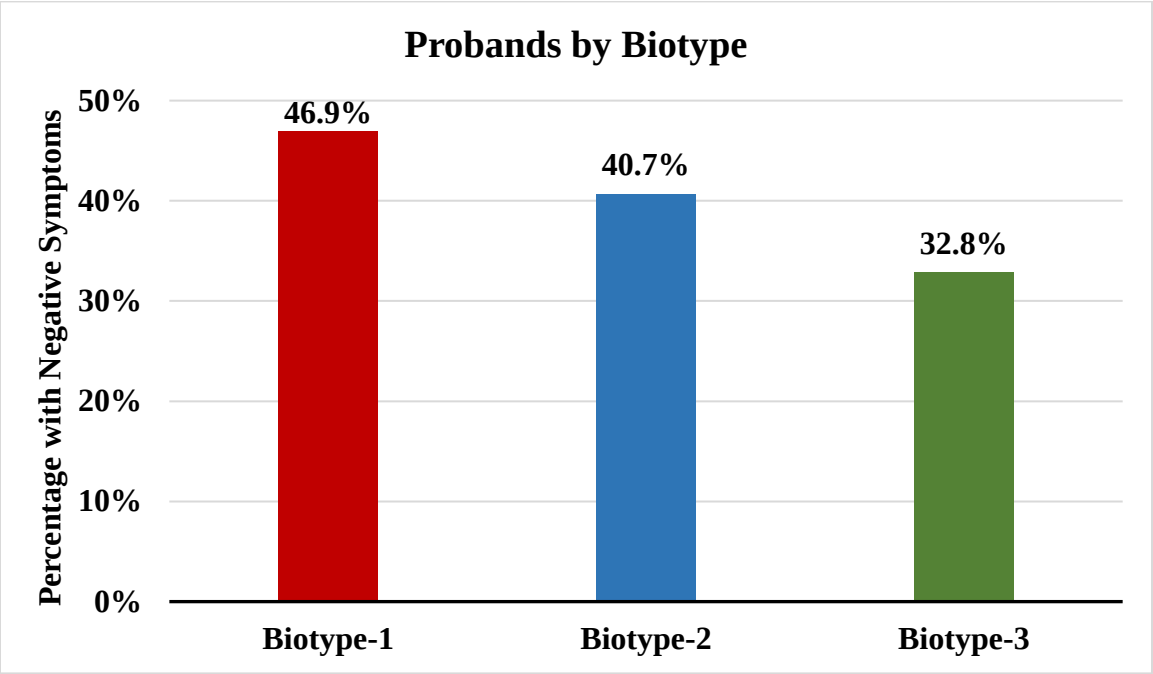
	Diminished Expression	**Avolition/Apathy**
N1 Blunted affect	0.76	–
N3 Poor rapport	0.75	–
N6 Lack of spontaneity	0.80	–
G7 Motor Retardation	0.68	–
N2 Emotional withdrawal	–	0.85
N4 Passive social withdrawal	–	0.76
G16 Active social avoidance	–	0.73

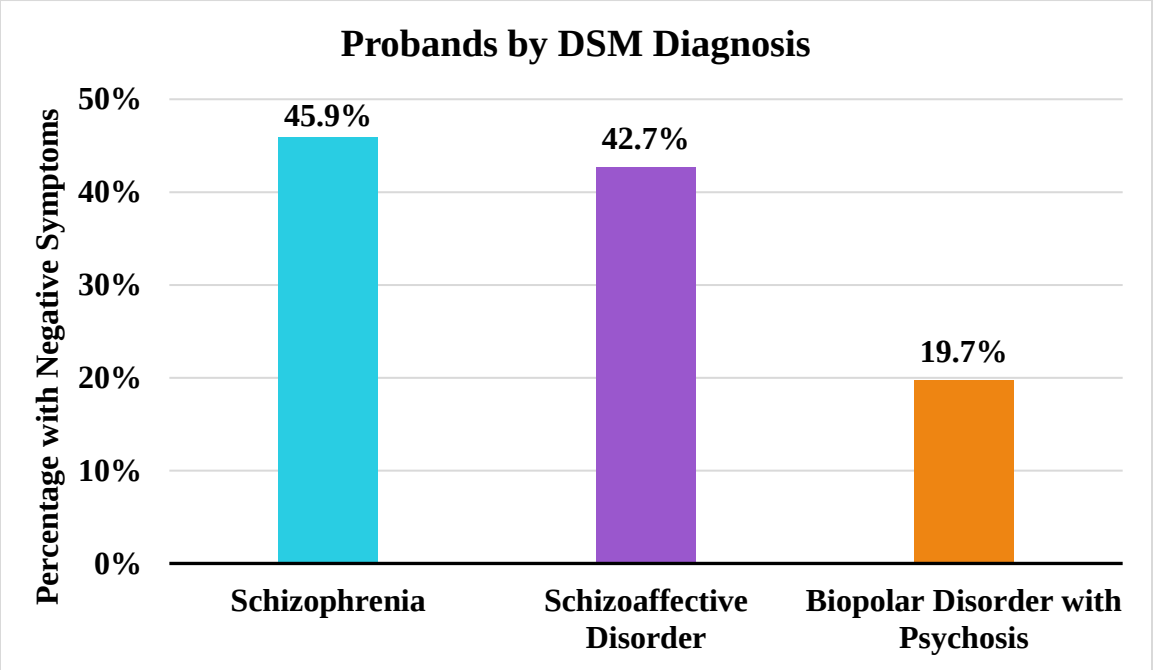
Supplemental Figures

Supplemental Figure 1. Histograms of socio-demographic and clinical characteristics for probands with predominant negative symptoms versus those without.

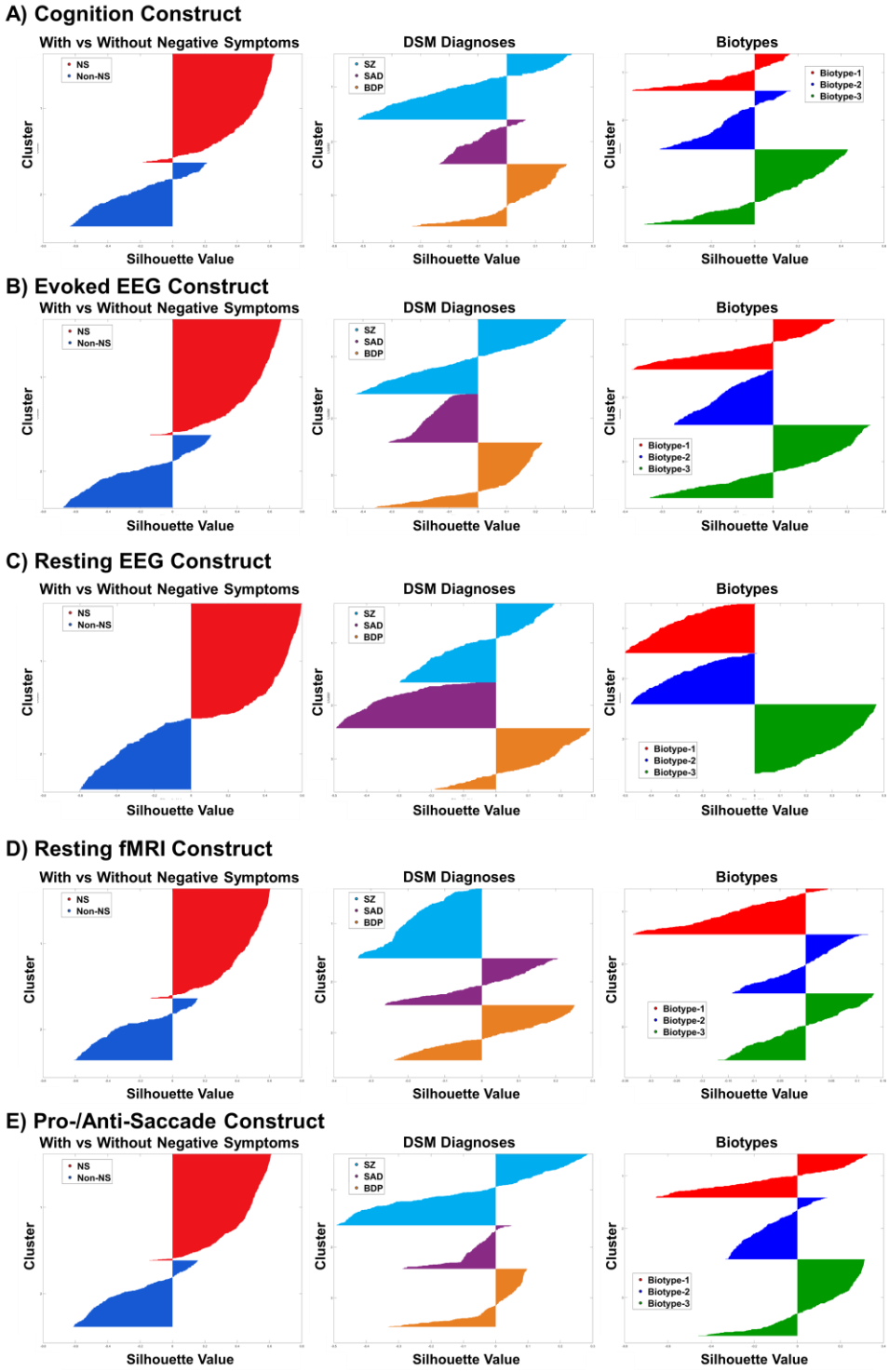


Supplemental Figure 2. Percentage of probands with predominant negative symptoms, by Biotype and DSM diagnosis

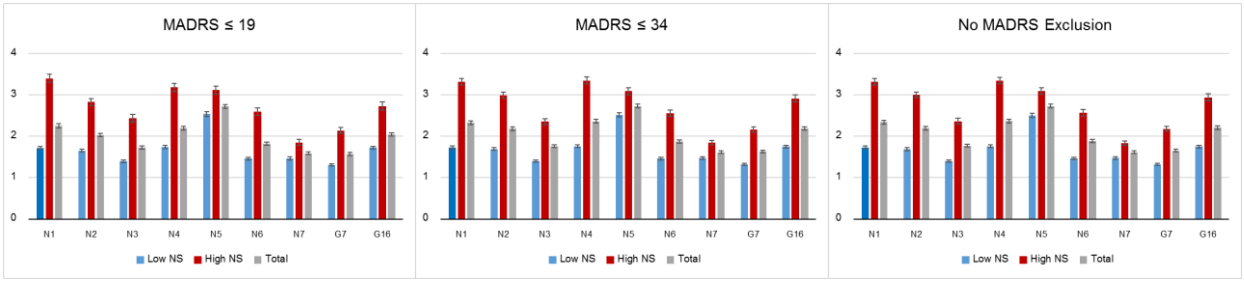




Supplemental Figure 3. Silhouette plots of canonical variate scores color coded by probands with vs without predominant negative symptoms (left), clinical diagnosis (middle), and Biotype (right) indicate largely overlapping subgroups. Scores represent the sum of the standardized data weighted by the loading strength of individual negative symptom and biomarker variables on their respective latent variates for each canonical correlation analysis.



Supplemental Figure 4. PANSS Negative Symptom Factor scores as a function of MADRS inclusion criteria



Supplemental Figure 5. Scatterplots of Brief Assessment of Cognition in Schizophrenia composite scores vs negative symptom items

Each "+" is an individual patient. Each NS item shows a significant negative correlation with BACS score, but the strength of each correlation is quite low.