nature research

https://doi.org/10.1038/s41591-020-1103-1

Supplementary information

Using common genetic variation to examine phenotypic expression and risk prediction in 22q11.2 deletion syndrome

In the format provided by the authors and unedited

Supplementary Material

Table of Contents

1. Supplementary Tables 1-6

Purpose	Regression (bold = what in p, effect columns)	N	effect size	p
1 IV and mediator	mediator ~ **IV** + covars	943	-0.405	0.001
1* IV and mediator *(no schizophrenia cases)*	mediator ~ **IV** + covars	725	-0.221	0.202
2 IV and DV *(same as original)*	DV ~ **IV** + covars	755	0.239	0.025
3 mediator and DV	DV ~ **mediator** + covars	725	1.886	0.067
4 IV, DV and mediator	DV ~ **IV** + mediator + covars	725	0.250	0.021

Supplementary Table 1. Mediation analysis between subthreshold psychosis, mood disorders, and PS_SZ. IV = Independent Variable, DV = Dependent Variable.

	Schizophrenia spectrum diagnosis (SSD)	Merged controls (all ages)
$90 < PS^{ile}$	OR = 1.44 [0.88, 2.37] PPV = 0.325 [0.222, 0.428] N+ = 26, N- = 181	N+ = 54, N- = 543
$75 < PS^{ile}$	OR = 1.62 [1.14, 2.31] PPV = 0.332 [0.265, 0.398] N+ = 64, N- = 143	N+ = 129, N- = 468
$50 < PS^{ile}$	OR = 1.91 [1.38, 2.64] PPV = [0.274, 0.366] N+ = 126, N- = 81	N+ = 268, N- = 329
$PS^{ile} < 50$	OR = 0.52 [0.38, 0.72] PPV = 0.198 [0.159, 0.236] N+ = 81, N- = 126	N+ = 329, N- = 268
$PS^{ile} < 25$	OR = 0.41 [0.27, 0.62] PPV = 0.147 [0.099, 0.195] N+ = 31, N- = 176	N+ = 180, N- = 417
$PS^{ile} < 10$	OR = 0.26 [0.12, 0.55] PPV = 0.091 [0.031, 0.151] N+ = 8, N- = 199	N+ = 80, N- = 517

Supplementary Table 2. OR for schizophrenia based on polygenic score cutoffs. Results show for a given binary cutoff based on Polygenic Score percentile, how many 22q11.2DS individuals fall above or below that cutoff, stratified by having SSD, or being a control (regardless of age). ORs and PPVs are given for SSD against merged controls. Prevalence of SSD (observed) is 26% (versus controls).

FSIQ Polygenic Score cutoff (percentile)	22q11.2DS with ID	22q11.2DS without ID
$PS^{ile} < 10$	OR = 2.64 [1.59, 4.4] PPV = 0.629 [0.515, 0.742] N+ = 44, N- = 246	N+ = 26, N- = 384
$PS^{ile} < 25$	OR = 2.07 [1.47, 2.93] PPV = 0.549 [0.475, 0.622] N+ = 96, N- = 194	N+ = 79, N- = 331
$PS^{ile} < 50$	OR = 1.85 [1.37, 2.51] PPV = 0.489, [0.436, 0.541] N+ = 171, N- = 119	N+ = 179, N- = 231
$50 < PS^{ile}$	OR = 0.54 [0.4, 0.73] PPV = 0.34 [0.29, 0.39] N+ = 119, N- = 171	N+ = 231, N- = 179
$75 < PS^{ile}$	OR = 0.49 [0.34, 0.71] PPV = 0.291 [0.224, 0.359] N+ = 51, N- = 239	N+ = 124, N- = 286
$90 < PS^{ile}$	OR = 0.42 [0.24, 0.74] PPV = 0.243, [0.142, 0.343] N+ = 17, N- = 273	N+ = 53, N- = 357

Supplementary Table 3. OR and PPV for ID based on polygenic score cut-offs. Results show for a given binary cut-off based on Polygenic Score percentile, how many 22q11.2DS individuals fall above or below that cut-off, stratified by having ID or not having ID. Odds-ratios and PPVs are shown for each percentile cut-off. ID is defined as IQ < 70. Overall prevalence of ID is 41%.

		Mean Age (SD)	Sex (%M)	SSD	Control	Putative ctrl	Sub-threshold	No pheno data
Included	962	24.0 (12.4)	48.6%	207 (21.5%)	215 (22.3%)	382 (39.7%)	158 (16.4%)	0
Excluded	824	18.2 (9.1)	48.5%	123 (14.9%	79 (9.6%)	468 (56.8%)	110 (13.3%)	11 (1.3%)

Supplementary Table 4. An overview of demographic differences between included and excluded individuals in the study. Note that these differences reflect the history of the IBBC recruitment strategy. In "phase 1" submission of DNA from individuals who were either (schizophrenia spectrum) case or true control (age >25) was encouraged. There are more Affymetrix data available from this "phase 1", because by the time the second wave started, the WGS effort was up and running. This "phase 2" also included individuals who did not directly qualify as either case or definitive control. The main reason for exclusion for the current study was lack of availability of Affymetrix data. Therefore, as a result of the said prioritization of phase 1 (schizophrenia cases and definitive controls), the mean age of subjects with available Affymetrix data is also higher compared to those without Affymetrix data (enriched in phase 2). The age difference occurred because the onset of schizophrenia is generally after age 18 years, and true controls were defined as only those without psychosis and older than 25 years. In addition, given the on average lower age range in the individuals with no available Affymetrix data (hence: not included in this study), it is expected that the proportion of putative controls is higher in the excluded samples.

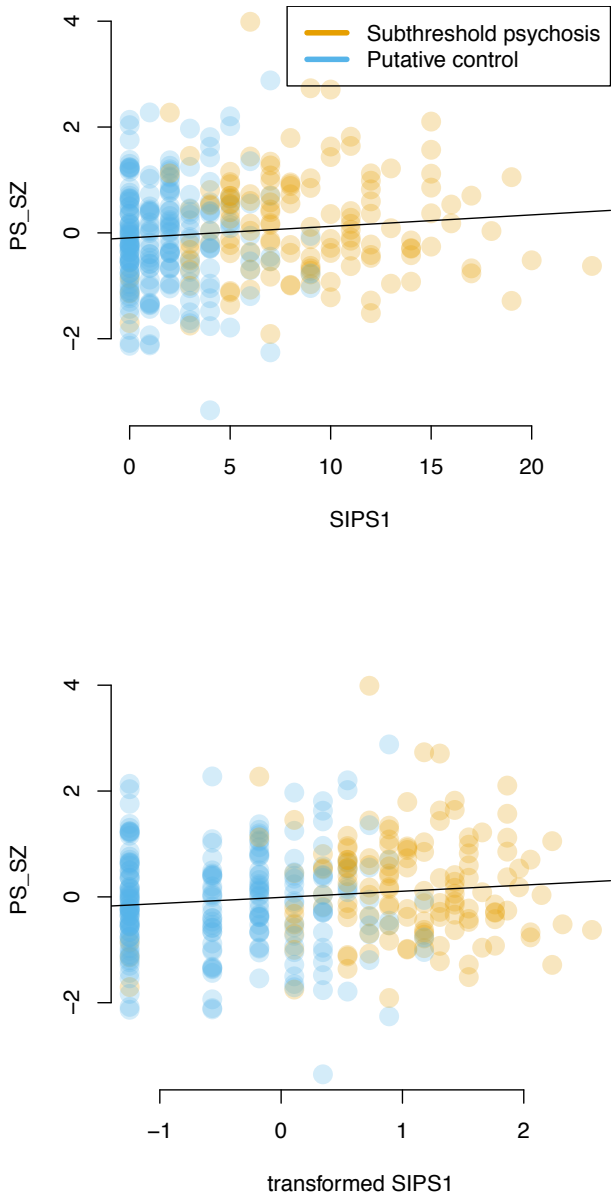
	Estimate	CI Lower Bound	CI Upper Bound
K_SZ	0.45	0.37	0.56
K_subthreshold psychosis	0.32	0.28	0.4
age_shape1	1.63	1.48	1.79
age_shape2	3.58	3.23	3.95
SZ_mean_age	23.05	19.64	27.98
SZ_sd_age	10.33	7.26	15.12
Subthreshold psychosis_mean_age	9.94	8.82	11.89
Subthreshold psychosis _sd_age	1.69	0.76	6.48

Supplementary Table 5. Parameter estimates for model that was used to inform power calculations in this study.

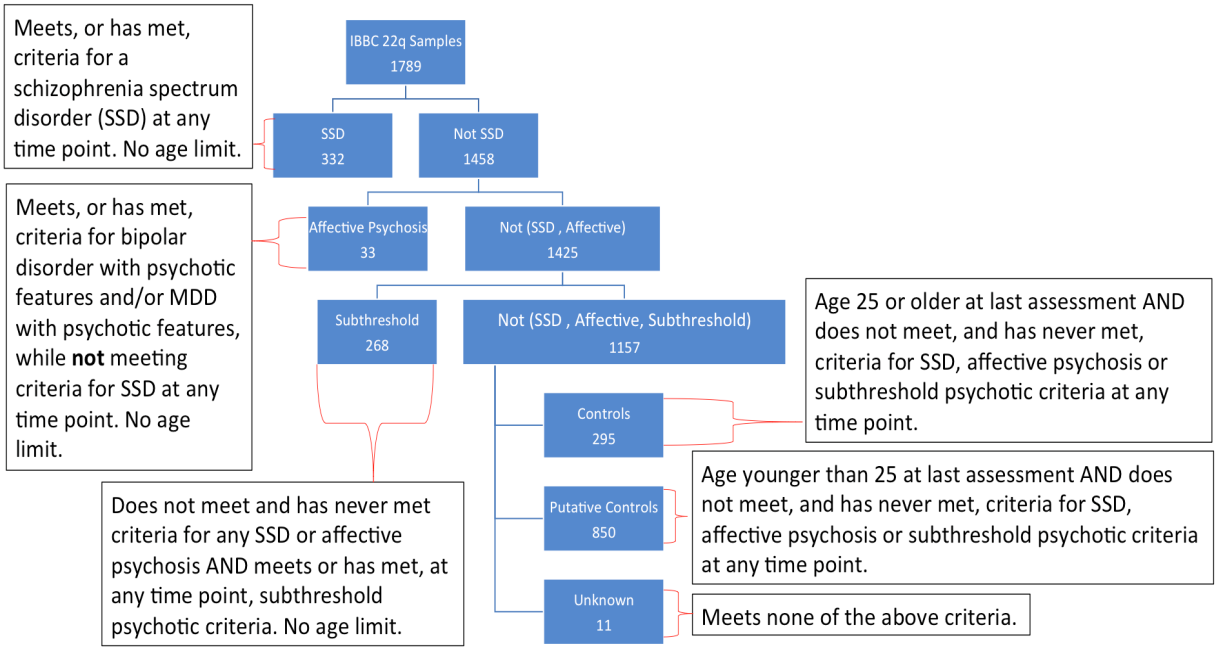
Dependent variable	IV	Power (alpha = 0.05)*	Relevant Supplementary Figure
SSD	PS_SZ	0.997 [r_g = 1]	5
Subthreshold psychosis		0.062 [r_g = 0], 0.974 [r_g = 0.95]	5
Baseline FSIQ		0.32 [r_g = 1]	6
VIQ decline		0.058 [r_g = 0], 1 [r_g = 0.8]	6
SSD	PS_IQ	0.189 [r_g = 1]	5
Subthreshold psychosis		0.048 [r_g = 0], 0.867 [r_g = 0.95]	5
Baseline FSIQ		1 [r_g = 1]	6
VIQ decline		0.048 [r_g = 0], 0.996 [r_g = 0.8]	6

Supplementary Table 6. Power analyses for primary analyses regarding genetic relationships between dependent variables (phenotypes) and independent variables (polygenic score). r_g is given between dependent variable and either schizophrenia (first four rows), or IQ (last four rows). Values of r_g between schizophrenia and IQ are fixed at -0.234, while otherwise, conditional on this, we report power for minimum and maximum possible genetic correlation between dependent variable and independent variable. IV = Independent Variable.

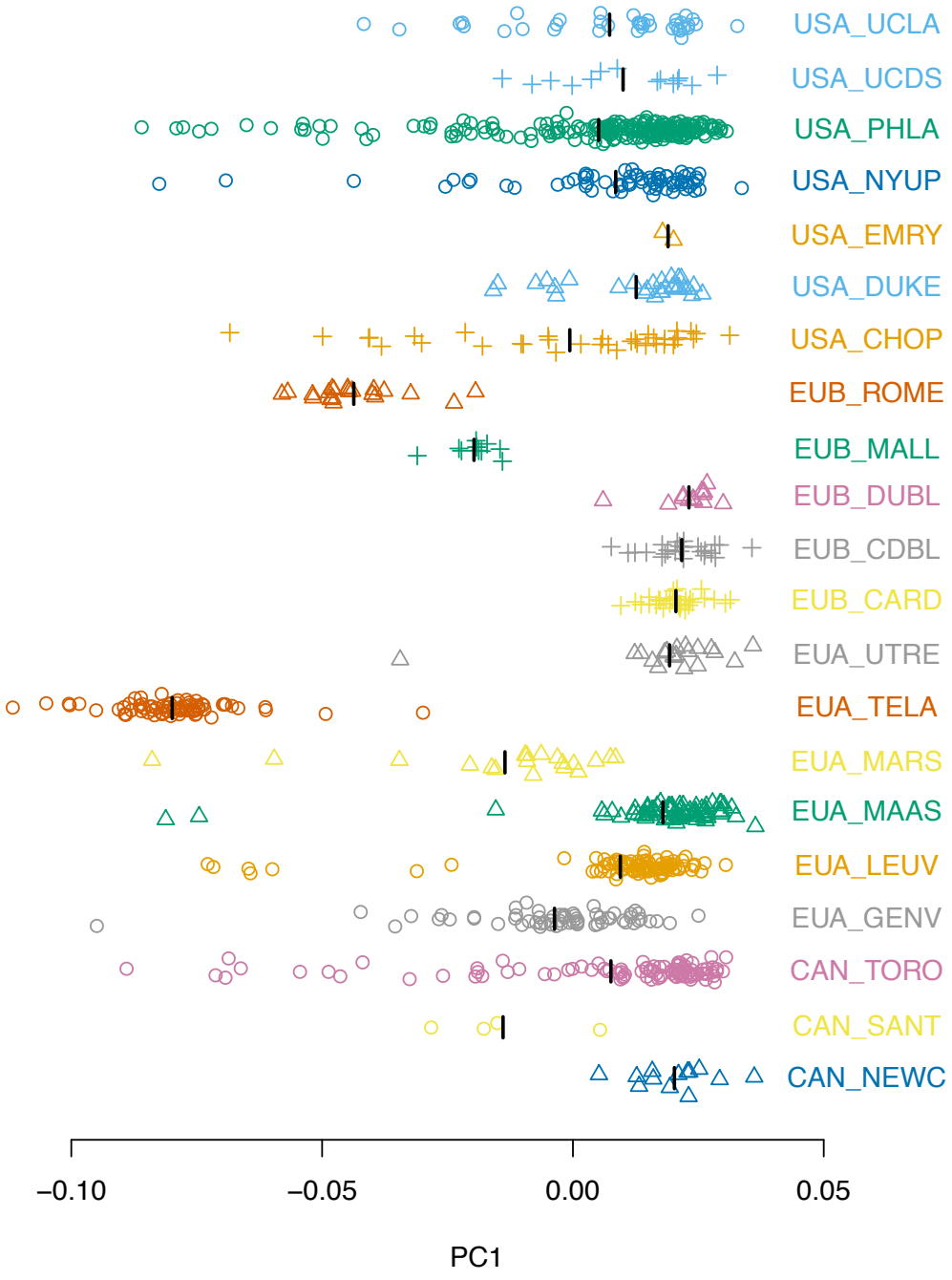
2. Supplementary Figures 1-8



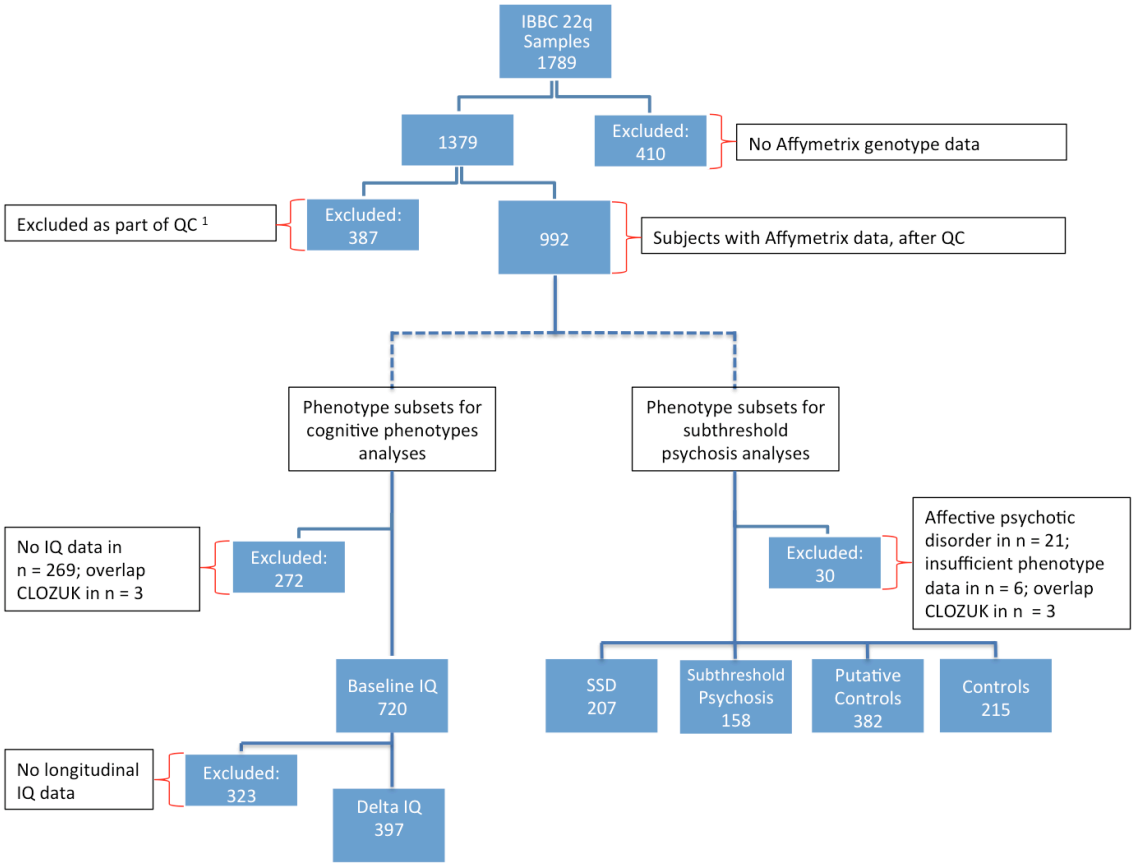
Supplementary Figure 1. Correlation plots between PS_SZ and a quantitative measure of subthreshold psychotic symptom severity. Upper panel shows untransformed SIPS values, lower panel shows transformed SIPS values. When adjusting for the previous binary indicator of subthreshold psychosis versus control, the association between the transformed quantitative SIPS phenotype and PS_SZ was not significant (N = 347, p = 0.77, r^2 = 0.0001).



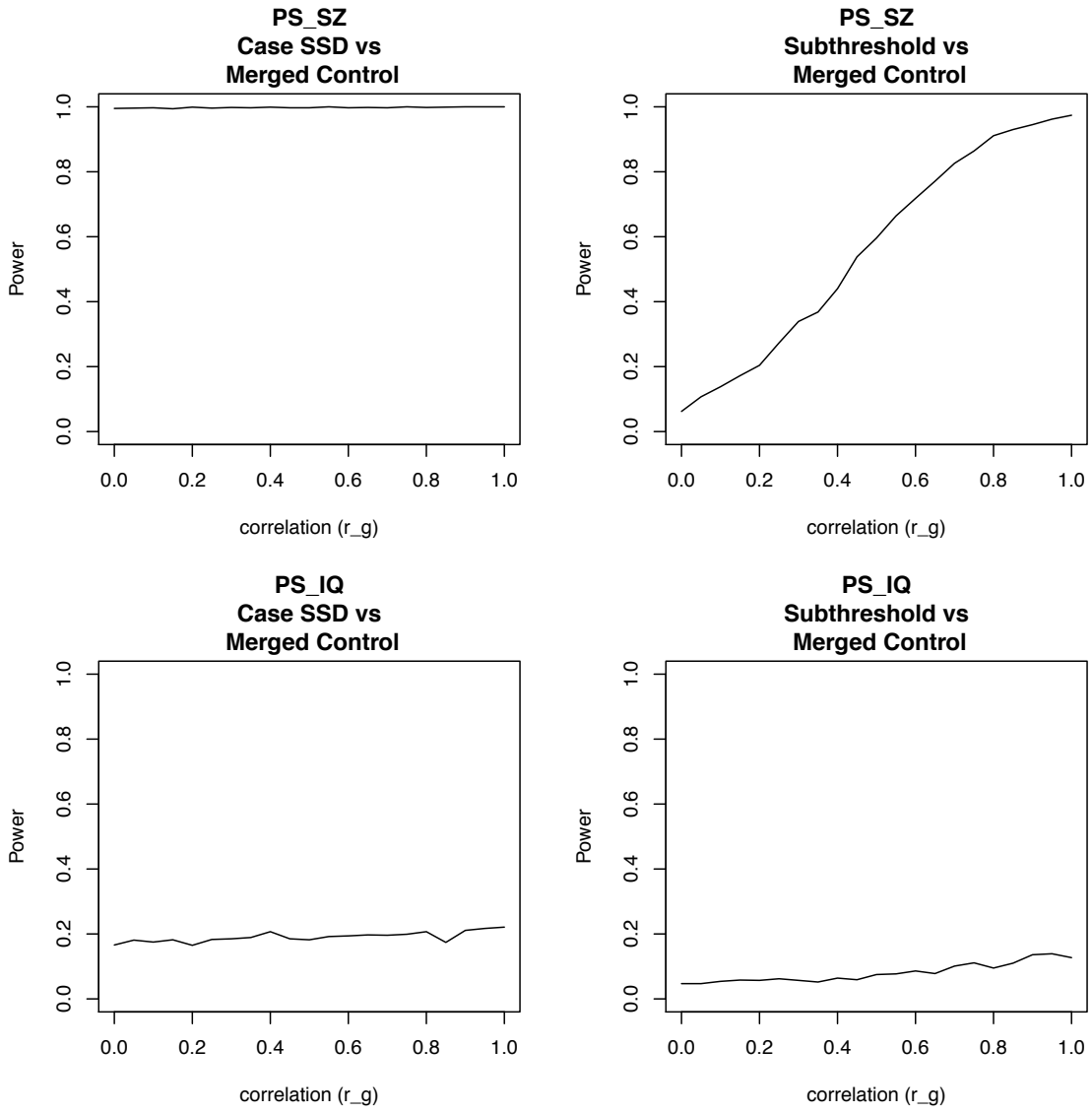
Supplementary Figure 2. Flowchart of IBBC cohort (full cohort) outlining the criteria used to assign IBBC subjects into different diagnostic classes regarding schizophrenia spectrum disorder (SSD) and related phenotypes.



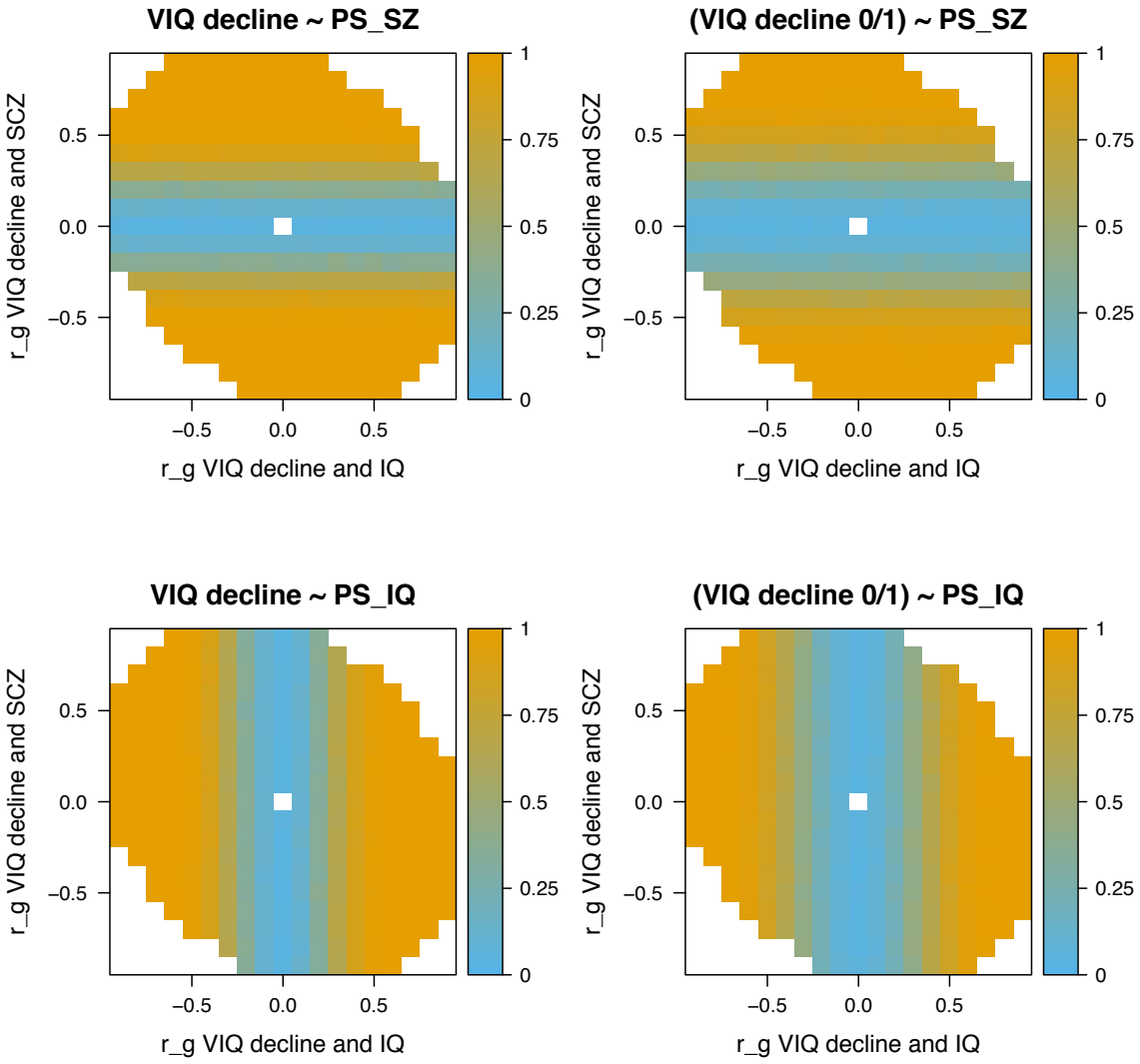
Supplementary Figure 3. Principal component 1 as function of study site. X-axis denotes value per-individual on PC1, while Y-axis is arbitrary to separate study sites plus jitter. Different sites are separated vertically and are grouped together by colour and plot icon. Black vertical bar indicates per study site average.



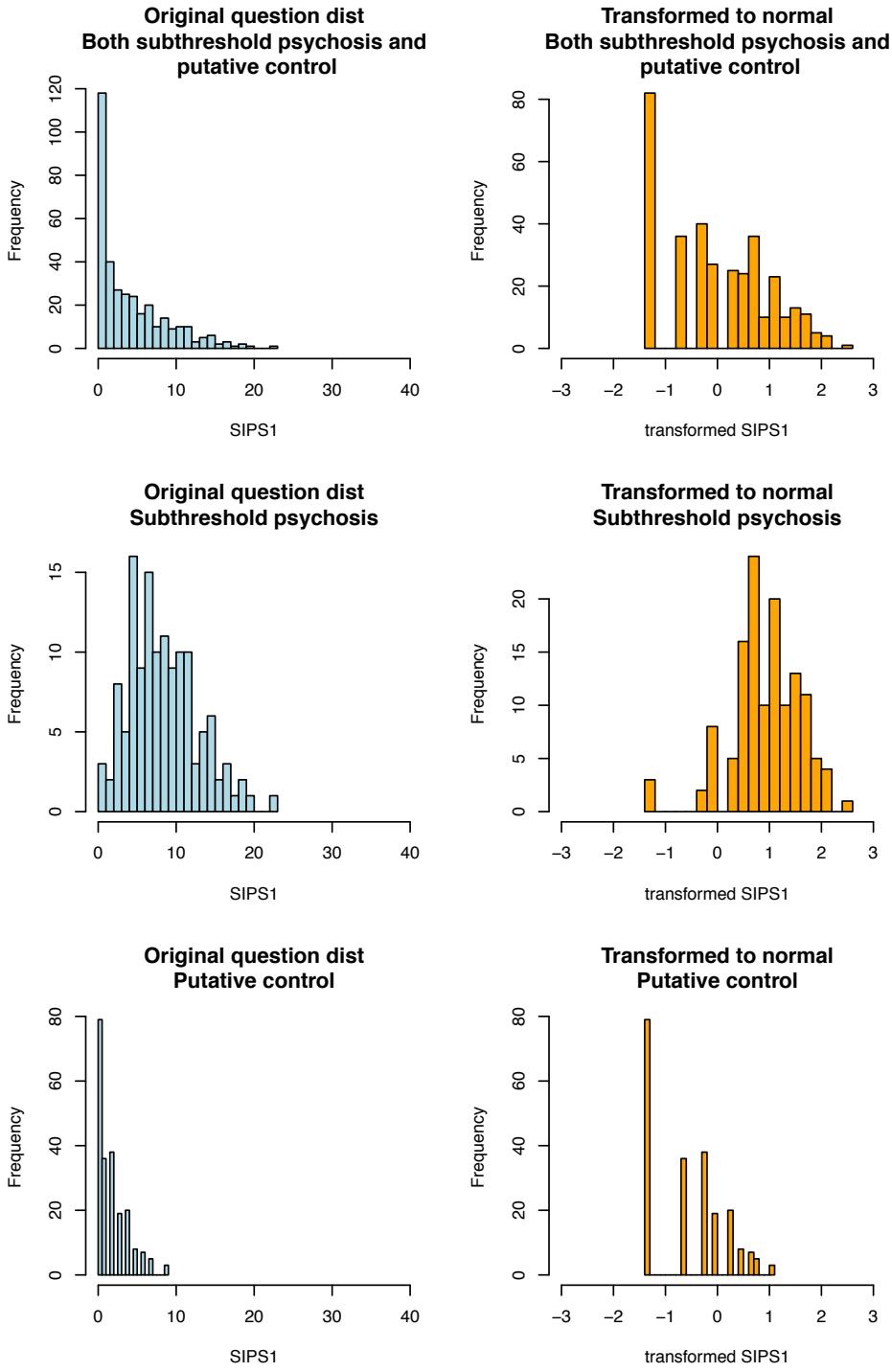
Supplementary Figure 4. Flowchart of subjects of IBBC cohort, outlining the different phenotypic subsets for the current study. [1.] Arrays excluded for sex coding reasons (5); missingness (84); IBD analysis (174); PCA (124).



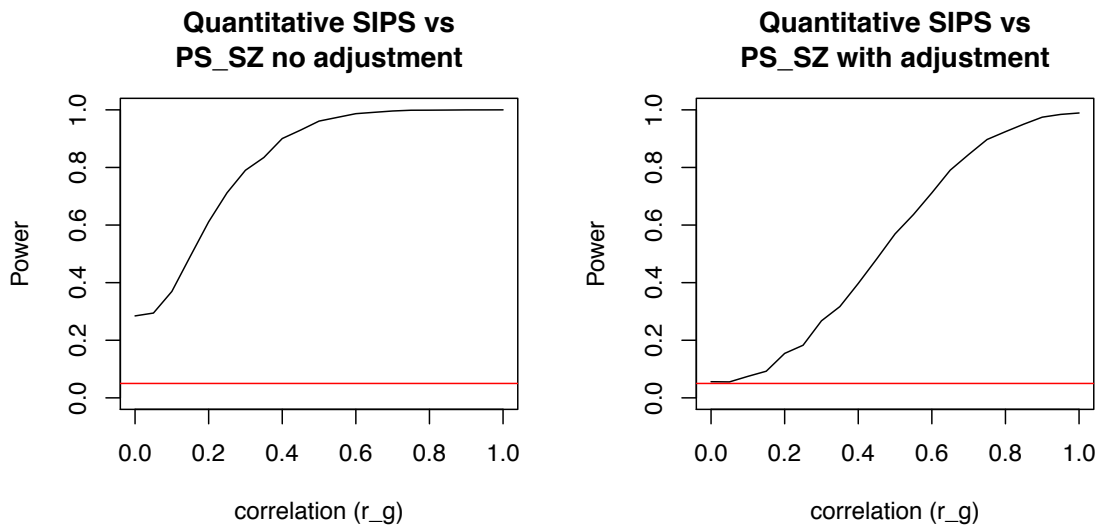
Supplementary Figure 5. Power to differentiate SSD status given genetic correlation. Shown are power at alpha=0.05 when comparing groups as specified in the plot sub-titles for their difference in polygene score as specified in the title, given genetic correlation between subthreshold psychosis and SSD.



Supplementary Figure 6. Power to differentiate VIQ decline given genetic correlations. Shown are power at alpha=0.05 when regressing continuous or binary VIQ decline against PS_SZ or PS_IQ, shown as a function of both the genetic correlation between VIQ decline and IQ, as well as between VIQ decline and SSD.



Supplementary Figure 7. Histogram of pre and post transformed SIPS measure. Transformation is defined by "qnorm(pexp(q = x + 0.5, rate = 0.2238))".



Supplementary Figure 8. Power analysis for quantitative subthreshold psychosis measure based on SIPS either without an adjustment for binary subthreshold psychosis (left) or with (right). Results were generated using simulation including only those simulations where a significant (i.e. alpha < 0.05) observation was made between subthreshold psychosis and PS_SZ. Note that the plot on the right, with the binary conditioning, is unbiased, unlike the plot on the left.