Rare Non-Synonymous Exonic Variants in Addiction and Behavioral Disinhibition

Supplement 1

Table S1. Descriptive Statistics for Constituent Measures of Substance Use and Behavioral Disinhibition

Variable	Mean (SD)	Range	% with 0
Nicotine Dependence			
Nicotine frequency number of days per month	12.4 (13.7) Exposed: 23.4 (11.2)	0 to 30	27.5
Nicotine quantity number of uses per day	8.1 (13.5) Exposed: 16.6 (15.6)	0 to 99	48.1
Nicotine dependence symptoms	1.2 (1.9) Exposed: 2.6 (2.1)	0 to 7	65.4 Exposed: 27.7
Alcohol Consumption			
Alcohol frequency Likert scale from never (0) to 3x day (10)	4.1 (2.5) Exposed: 4.7 (2.1)	0 to 10	14.5
Number of lifetime intoxications	59.3 (190.0) Exposed: 71.6 (206.7)	0 to 999	29.5
Maximum consumption greatest number of drinks in 24 hours	9.4 (9.1) Exposed: 11.2 (9.0)	0 to 50	14.6
Alcohol Dependence			
Alcohol dependence Social and occupational consequences	0.7 (1.2) Exposed: 0.9 (1.2)	0 to 9	56.6 Exposed: 48.8
Alcohol dependence Withdrawal and tolerance	0.3 (0.8) Exposed: 0.4 (0.8)	0 to 4	81.8 Exposed: 79.0
Alcohol dependence Compulsive use and impairment in major life activities	0.5 (1.5) Exposed: 0.6 (1.6)	0 to 13	80.5 Exposed: 78.4
Illicit Drug Use			
Number of lifetime marijuana uses	74.9 (234.7)	0 to 999	55.0
Number of different drug classes used	0.9 (1.5)	0 to 11	54.2
Drug use disorder symptoms	0.5 (1.5)	0 to 9	81.9
Behavioral Disinhibition			
Adult antisocial behavior symptoms	0.6 (1.0)	0 to 7.5	64.0

Conduct disorder symptoms	0.9 (1.4)	0 to 10.5	51.9
Dissocial behaviors	0.4 (0.6)	0 to 3	63.1
Delinquent Behavior Inventory	3.3 (3.5)	0 to 21	21.8
Aggressive undercontrol	41.5 (9.0)	20 to 77	N/A

All substance use and substance use disorder measures refer to period of heaviest use. Exposed refers to the means and standard deviations for the subsample of individuals who met the study criterion for adequate exposure (i.e., sufficient levels of use) to nicotine or alcohol. The number of lifetime intoxications and marijuana uses was capped at 999. The number of maximum drinks in 24 hours was capped at 50. The different drug classes included marijuana, amphetamines, barbiturates, cocaine, psychedelics, heroin, other opiates, inhalants, PCP, tranquilizers, and gas. Adult antisocial behavior refers to the adult criteria for antisocial personality disorder. Dissocial behaviors is the sum of the following behaviors: ever being suspended from school, ever being arrested, and early initiation of sexual intercourse (age 18 or older = 0, age 15-17 = 0.5, age 14 or younger = 1). The Delinquent Behavior Inventory is a self-report questionnaire of 21 different antisocial acts committed in childhood or adolescence. Aggressive undercontrol is a 20-item personality scale (1 to 4 response format). Items taken were the aggression, self-control (vs. impulsivity), harm avoidance (vs. thrill seeking), and traditionalism (vs. rebelliousness and nonconformity) scales of the Multidimensional Personality Questionnaire.

Table S2. Phenotypic Correlations for Family Member Pairs

Family Pairing	NIC			DRK			ALD			DRG			BD			NICe			DRKe			ALDe		
	r	**95% CI**		**r**	**95% CI**		**r**	**95% CI**		**r**	**95% CI**		**r**	**95% CI**		**r**	**95% CI**		**r**	**95% CI**		**r**	**95% CI**	
		Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi
MZ Twins	.71	(.68	.74)	.72	(.69	.75)	.75	(.73	.78)	.74	(.71	.76)	.74	(.72	.77)	.70	(.64	.74)	.69	(.65	.72)	.70	(.66	.74)
DZ/Bio Sibs	.44	(.38	.50)	.46	(.39	.51)	.41	(.34	.47)	.46	(.40	.51)	.43	(.37	.49)	.51	(.41	.60)	.45	(.37	.52)	.31	(.22	.39)
Adoptive Sibs*	-.02	(-.31	.29)	.21	(-.17	.50)	.20	(-.05	.41)	.10	(-.14	.33)	.31	(.01	.54)	.19	(-.51	.62)	.21	(-.18	.51)	.28	(.01	.50)
Mom/Bio	.22	(.18	.26)	.23	(.19	.27)	.25	(.20	.29)	.27	(.23	.31)	.30	(.26	.34)	.14	(.05	.22)	.14	(.09	.20)	.19	(.14	.25)
Mom/Adopt*	.31	(-.02	.52)	.19	(-.10	.41)	.29	(.07	.47)	.18	(.00	.35)	.19	(-.18	.45)	-.07	(-.41	.31)	.39	(.07	.57)	.42	(.16	.59)
Dad/Bio	.23	(.18	.27)	.23	(.18	.27)	.23	(.18	.27)	.26	(.21	.30)	.31	(.27	.35)	.20	(.11	.28)	.15	(.09	.20)	.19	(.14	.25)
Dad/Adopt*	.20	(-.04	.40)	.19	(-.13	.43)	.12	(-.09	.31)	.08	(-.13	.27)	.20	(-.12	.44)	.15	(-.34	.45)	.16	(-.09	.37)	.14	(-.08	.33)
Mom/Dad	.34	(.30	.38)	.40	(.36	.44)	.37	(.32	.41)	.56	(.53	.59)	.38	(.34	.42)	.20	(.10	.29)	.32	(.27	.37)	.32	(.26	.36)

ALD, alcohol dependence; BD, behavioral disinhibition; Bio sibs, biological siblings; CI, confidence interval; DRG, illicit drug use; DRK, alcohol consumption; DZ, dizygotic; MZ, monozygotic; NIC, nicotine dependence.

NICe = nicotine dependence among exposed individuals, and similarly for DRKe and ALDe.

Mom/Bio is correlation between mother and biological offspring; Dad/Bio correlation between father and biological offspring; Mom/Adopt and Dad/Adopt are analogous for adoptive offspring.

*These comparisons are based on few observations and have large confidence intervals.

Table S3. Correlations Among Residualized Phenotypes

	Nicotine	**Alcohol Consumption**	**Alcohol Dependence**	**Illicit Drugs**	**Behavioral Disinhibition**	**Nicotine (Exposed)**	**Alcohol Consumption (Exposed)**	**Alcohol Dependence (Exposed)**
Nicotine	1.0							
Alcohol Consumption	.49	1.0						
Alcohol Dependence	.60	.64	1.0					
Illicit Drugs	.51	.48	.64	1.0				
Behavioral Disinhibition	.57	.66	.72	.64	1.0			
Nicotine (Exposed)	.90	.45	.43	.38	.51	1.0		
Alcohol Consumption (Exposed)	.49	.94	.66	.45	.64	.52	1.0	
Alcohol Dependence (Exposed)	.56	.59	.99	.59	.72	.44	.68	1.0

Table S4. Covariates Significant in the Covariate Correction

	Age	Age^2	BYR	Generation	Sex	BYR × Generation	Generation × Sex	Age × Generation	PCs 1-10	R^2
Nicotine	x	x		x		x	x	x		.12
Alcohol Consumption	x	x	x	x		x	x	x		.27
Alcohol Dependence	x	x	x	x	x	x	x	x	1,8	.23
Illicit Drugs	x		x	x		x	x	x	4	.14
Behavioral Disinhibition	x	x		x	x	x	x	x	7	.18
Nicotine (Exposed)	x		x	x		x		x	1,4,7	.39
Alcohol Consumption (Exposed)	x	x	x	x	x	x	x	x	1	.30
Alcohol Dependence (Exposed)	x	x	x	x	x		x	x	1,8	.21

BYR, birth year; PC, principal components.

Covariate correction was done with multiple regression. An "x" denotes that the covariate was significantly associated with the phenotype. The PCs 1-10 column denotes the PC (if any) that was significantly associated with that phenotype. PC associations were not strongly associated in any case. R^2 is the adjusted r-squared for the full model containing all covariates, regardless of significance.

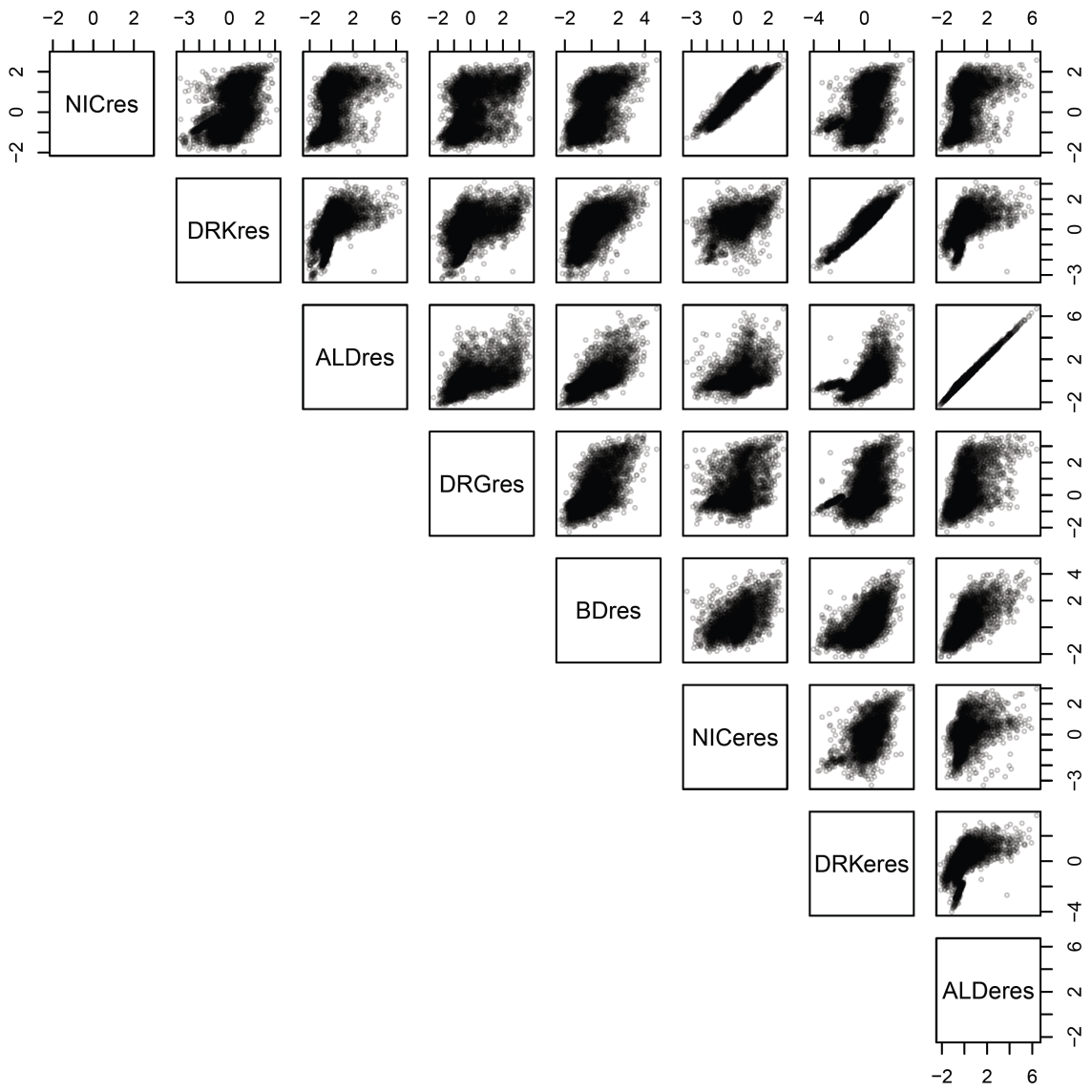

Figure S1. Phenotypic Scatterplot Prior to Inverse Normalization. NICres, DRKres, ALDres, DRGres, and BDres are the residuals calculated for the nicotine, alcohol consumption, alcohol dependence, drug use, and behavioral disinhibition phenotypes, respectively. NICeres, DRKeres, and ALDeres are the phenotypes conditioned on exposure to the relevant substance. The clumps observed for DRKeinv are due to individuals who had been sufficiently exposed to alcohol but do not currently drink. That is, their phenotypic value was zero, but the zero value was perturbed in covariate correction, leading to a “clump” in the scatterplots.

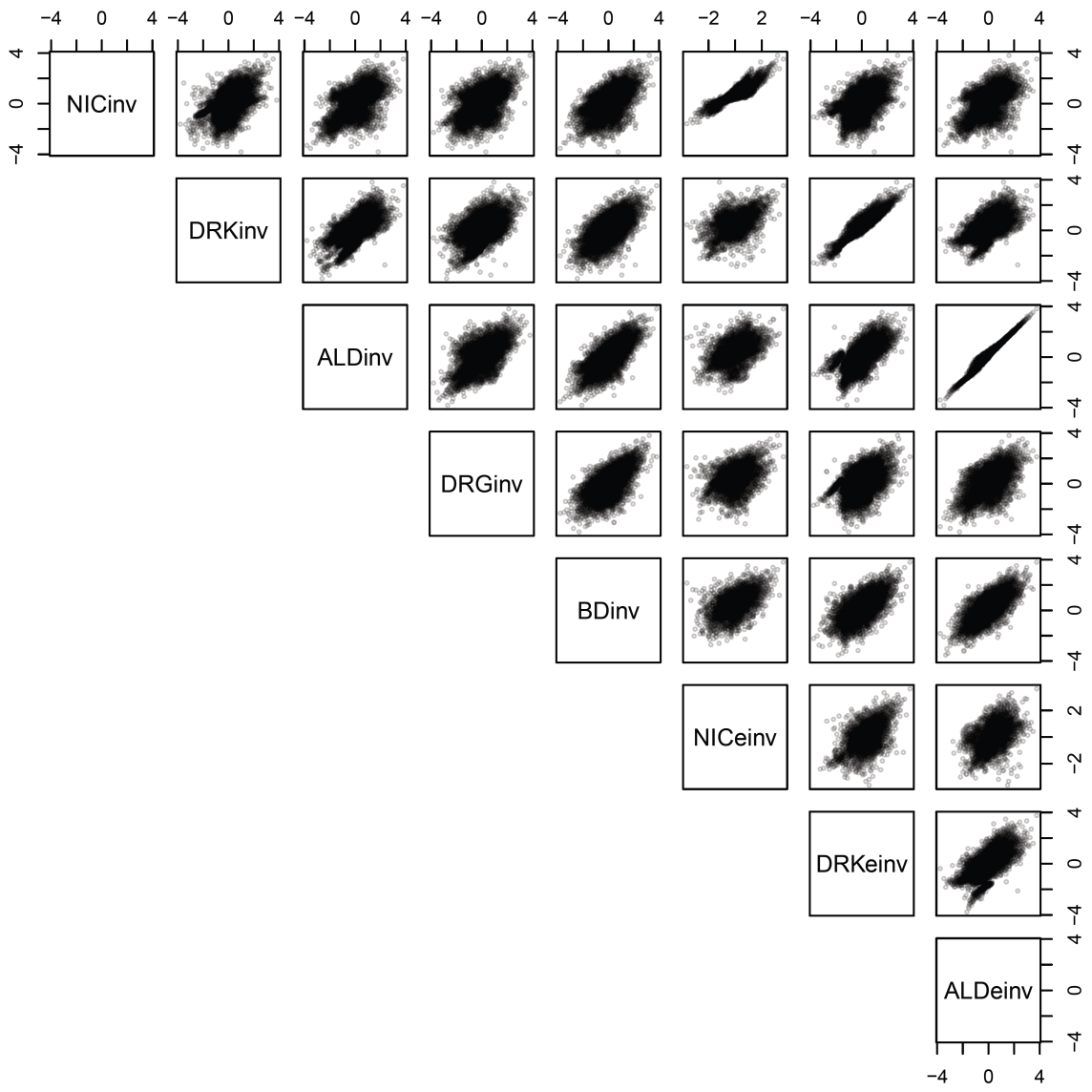

Figure S2. Phenotypic Scatterplot after Inverse Normalization. NICinv, DRKinv, ALDinv, DRGinv, and BDinv are the residuals calculated for the nicotine, alcohol consumption, alcohol dependence, drug use, and behavioral disinhibition phenotypes, respectively. NICeinv, DRKeinv, and ALDeinv are the phenotypes conditioned on exposure to the relevant substance. The clumps observed for DRKeinv are due to individuals who had been sufficiently exposed to alcohol but do not currently drink. That is, their phenotypic value was zero, but the zero value was perturbed in covariate correction, leading to a “clump” in the scatterplots.

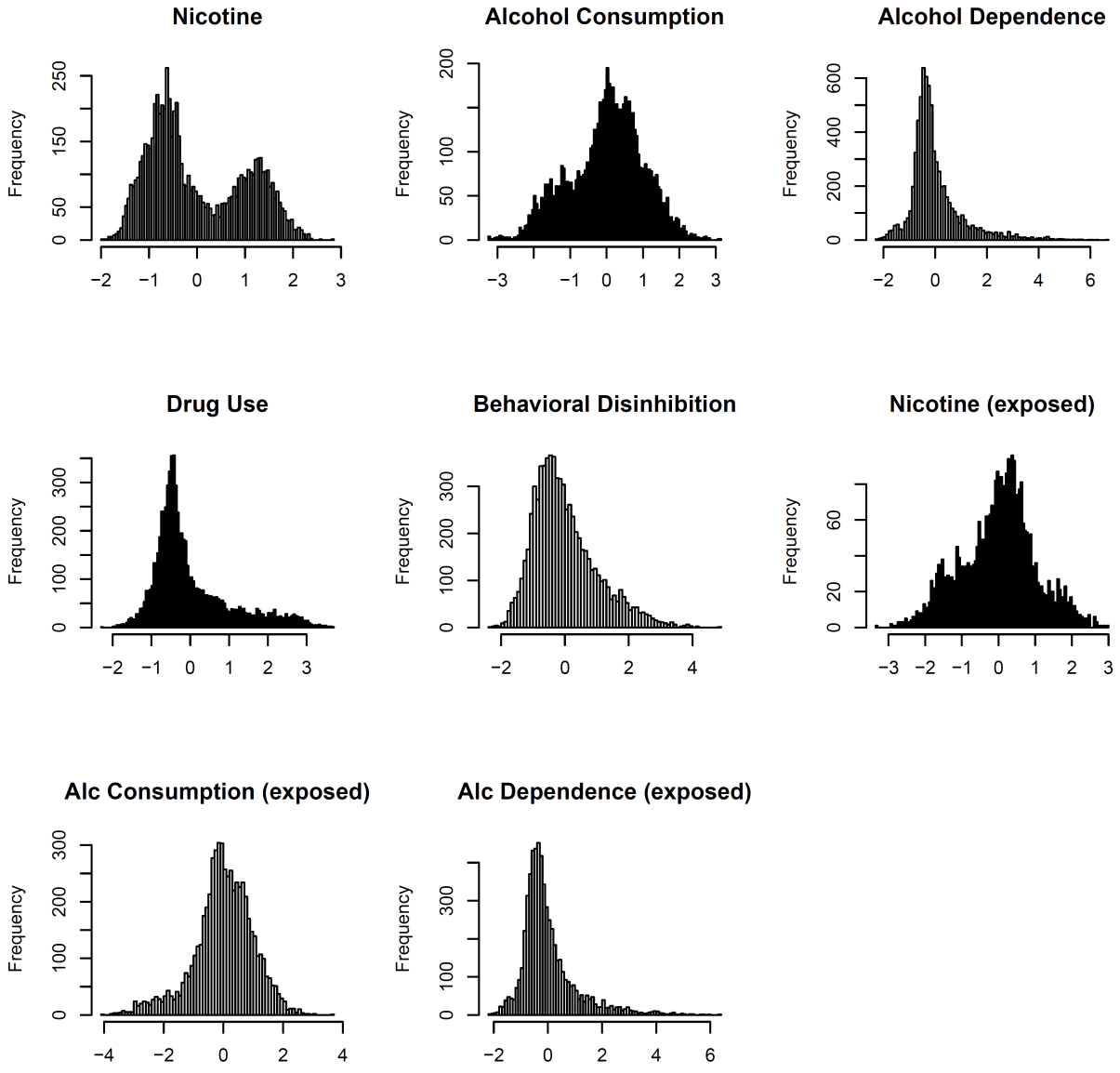

Figure S3. Phenotype Histograms after Covariate Correction, Prior to Inverse Normalization. Alc, alcohol.

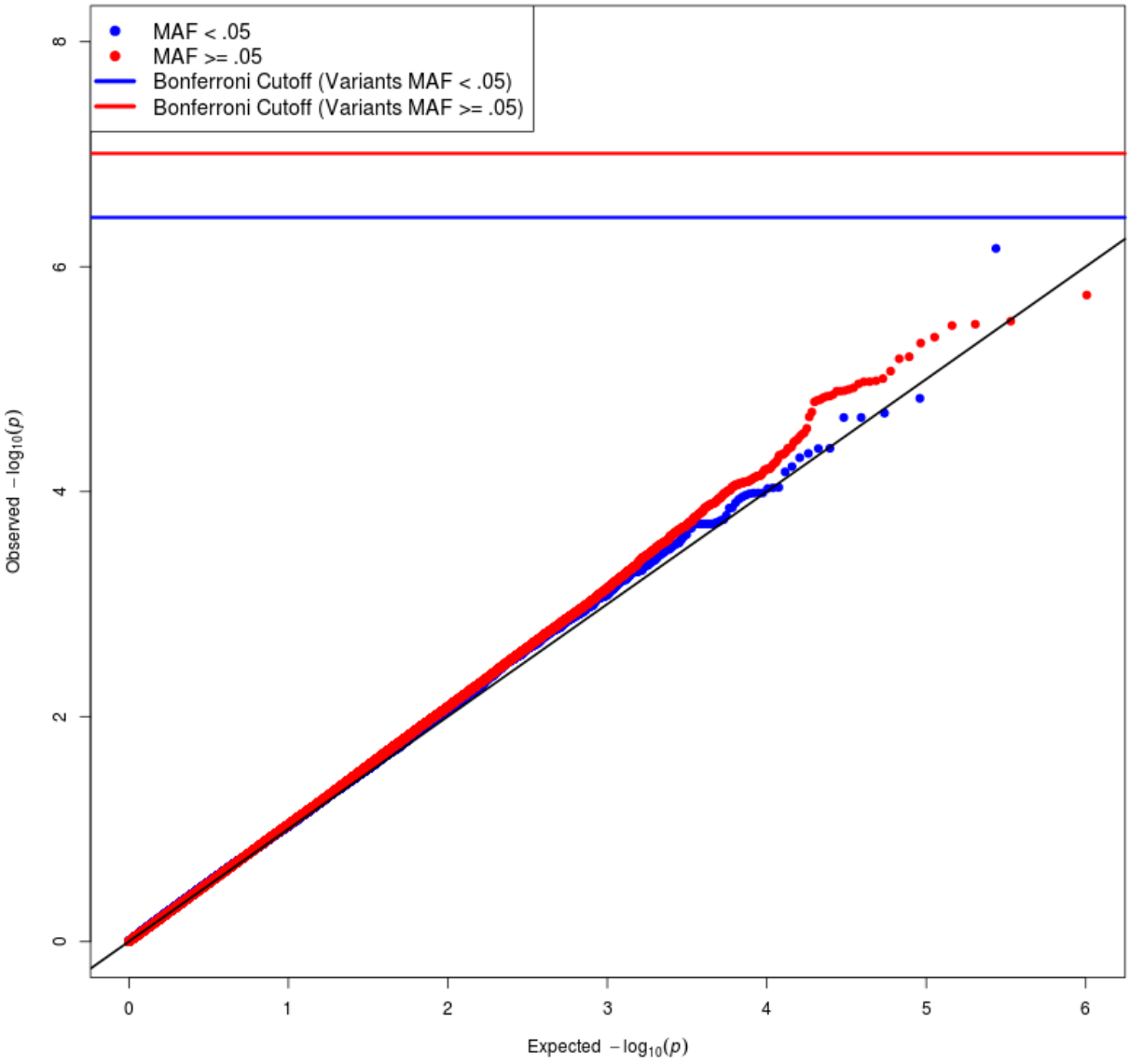

Figure S4. Single Variant QQplot of Nicotine. MAF, minor allele frequency.

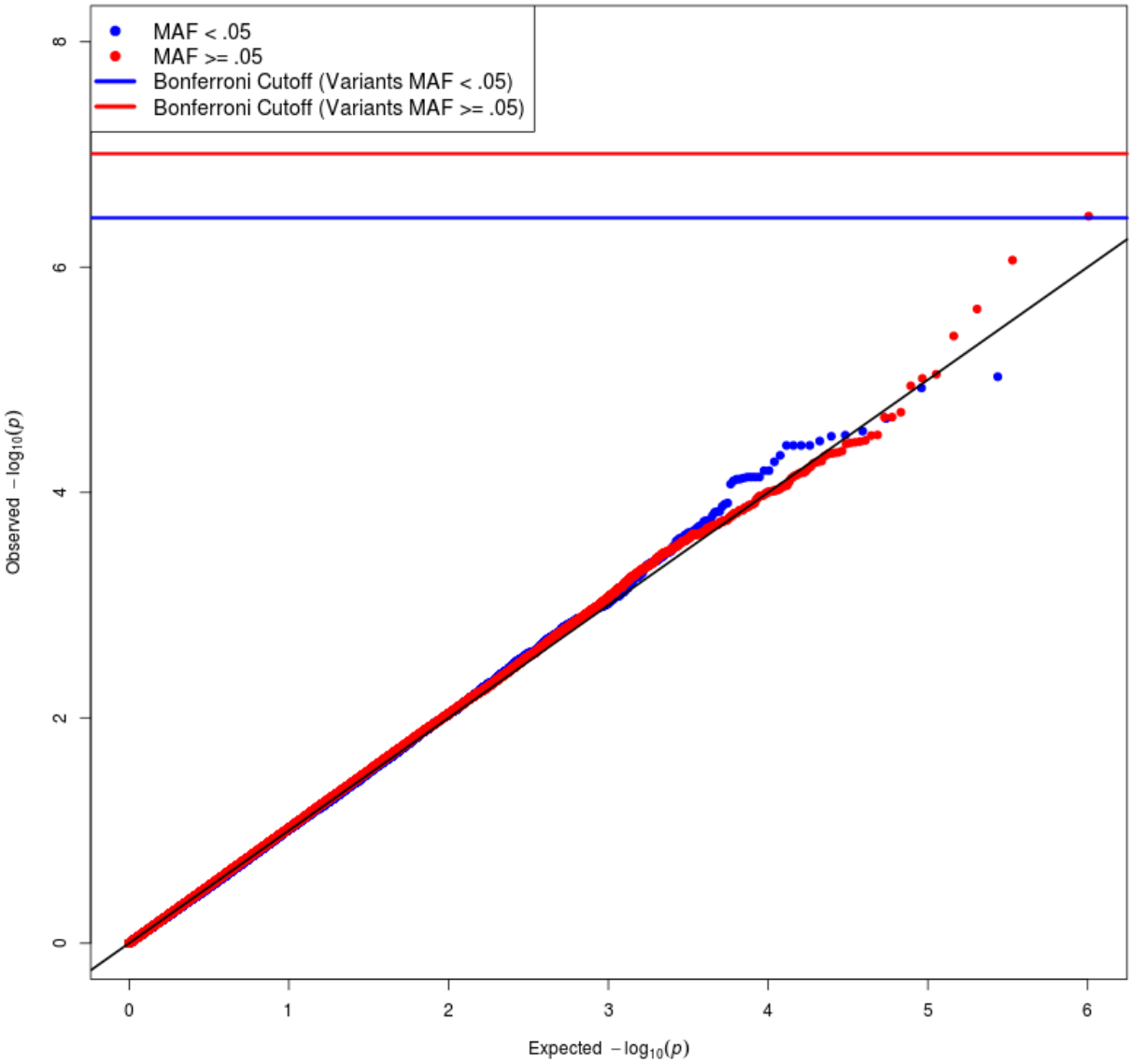

Figure S5. Single Variant QQplot of Alcohol Consumption. MAF, minor allele frequency.

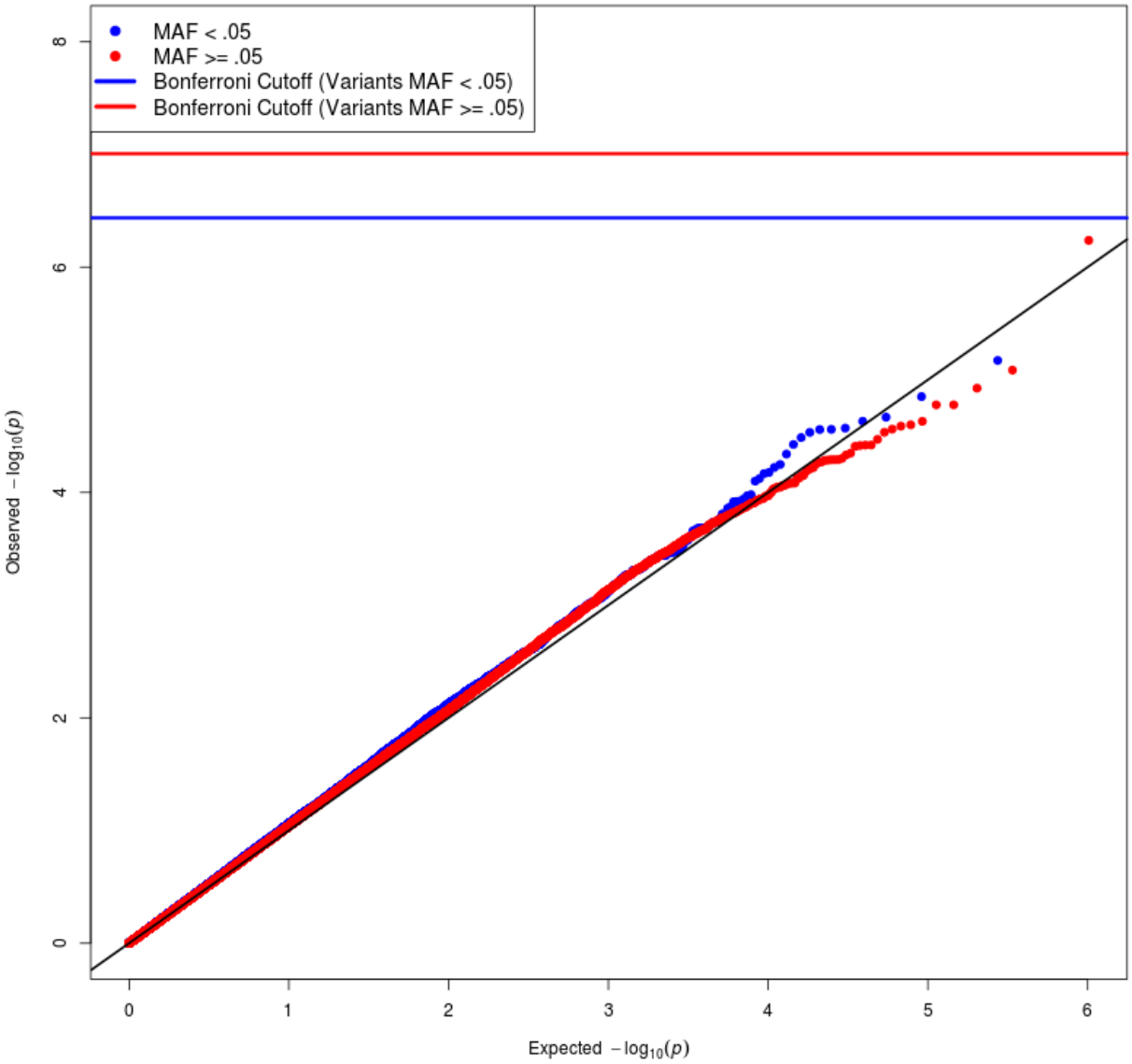

Figure S6. Single Variant QQplot of Alcohol Dependence. MAF, minor allele frequency.

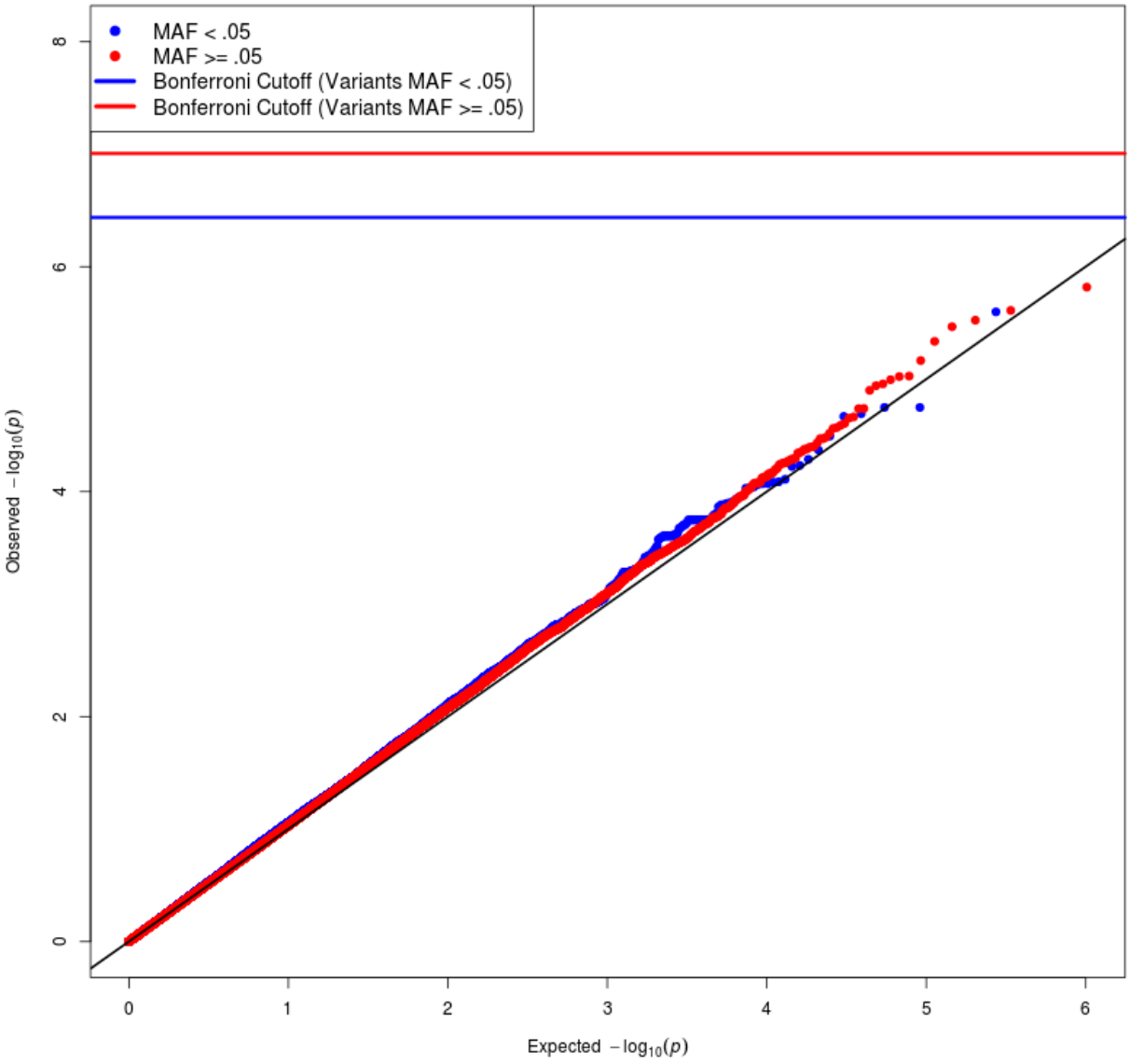

Figure S7. Single Variant QQplot of Illicit Drug Use. MAF, minor allele frequency.

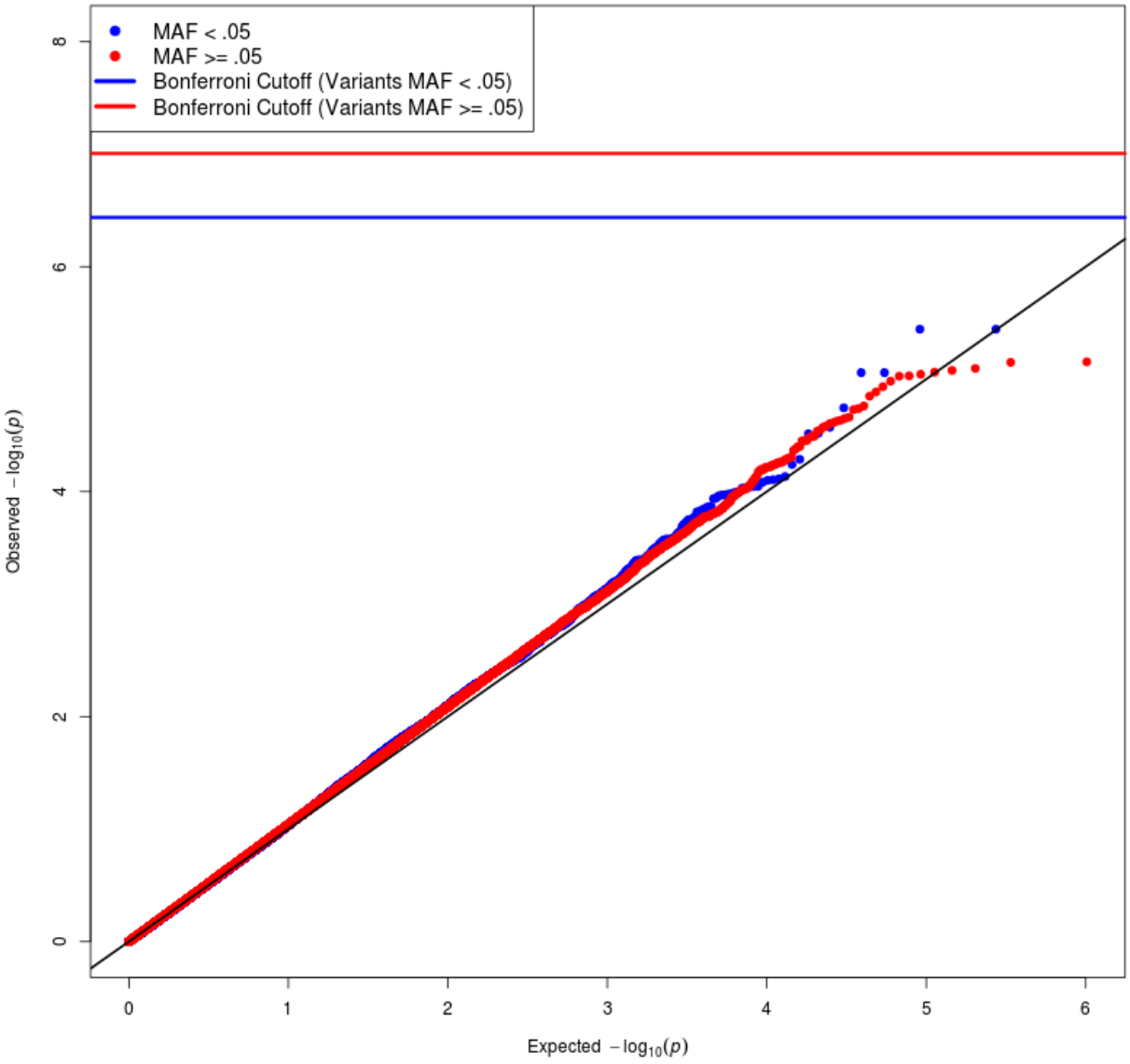

Figure S8. Single Variant QQplot of Behavioral Disinhibition. MAF, minor allele frequency.

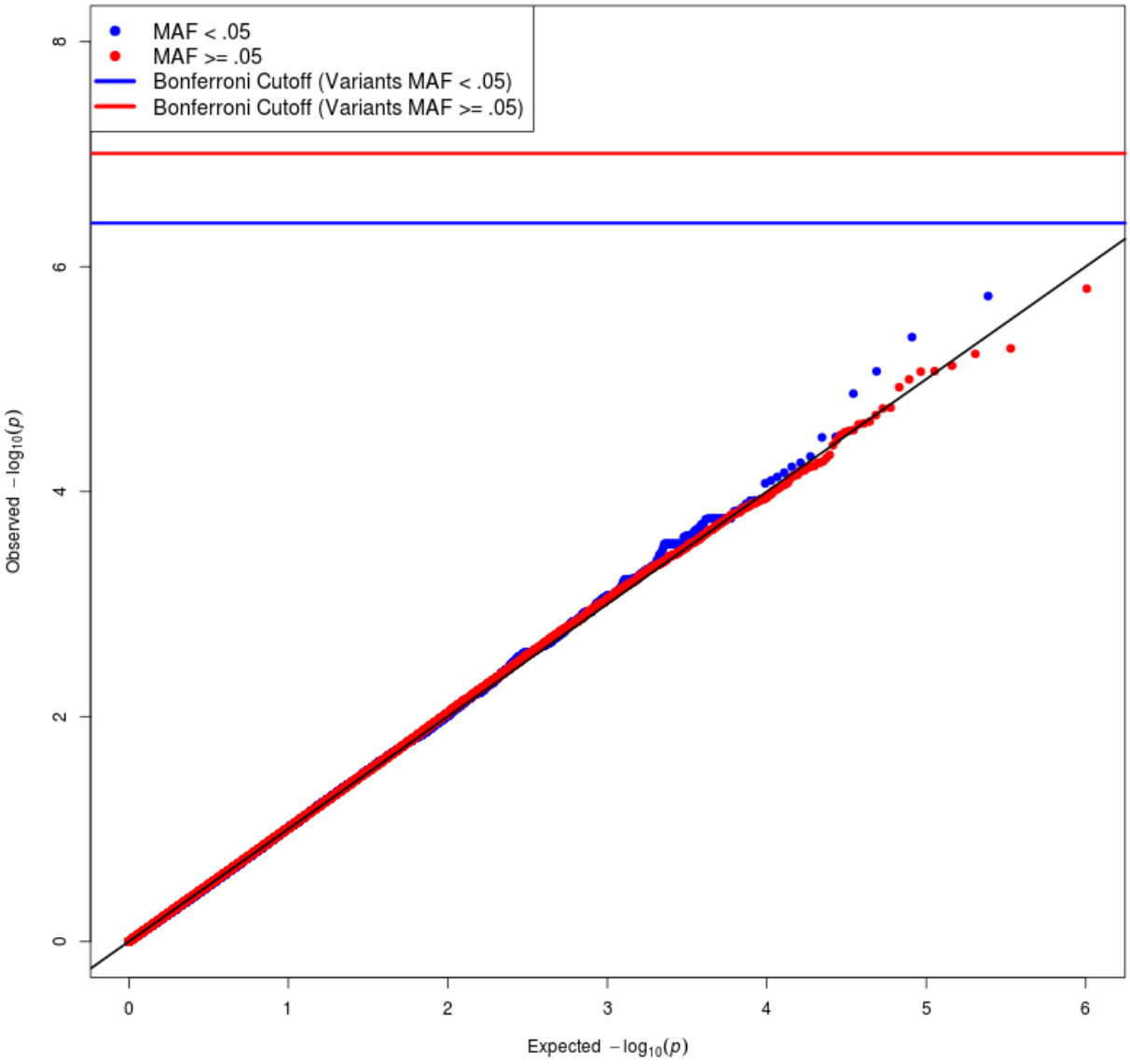

Figure S9. Single Variant QQplot of Nicotine in Individuals Exposed to Nicotine. MAF, minor allele frequency.

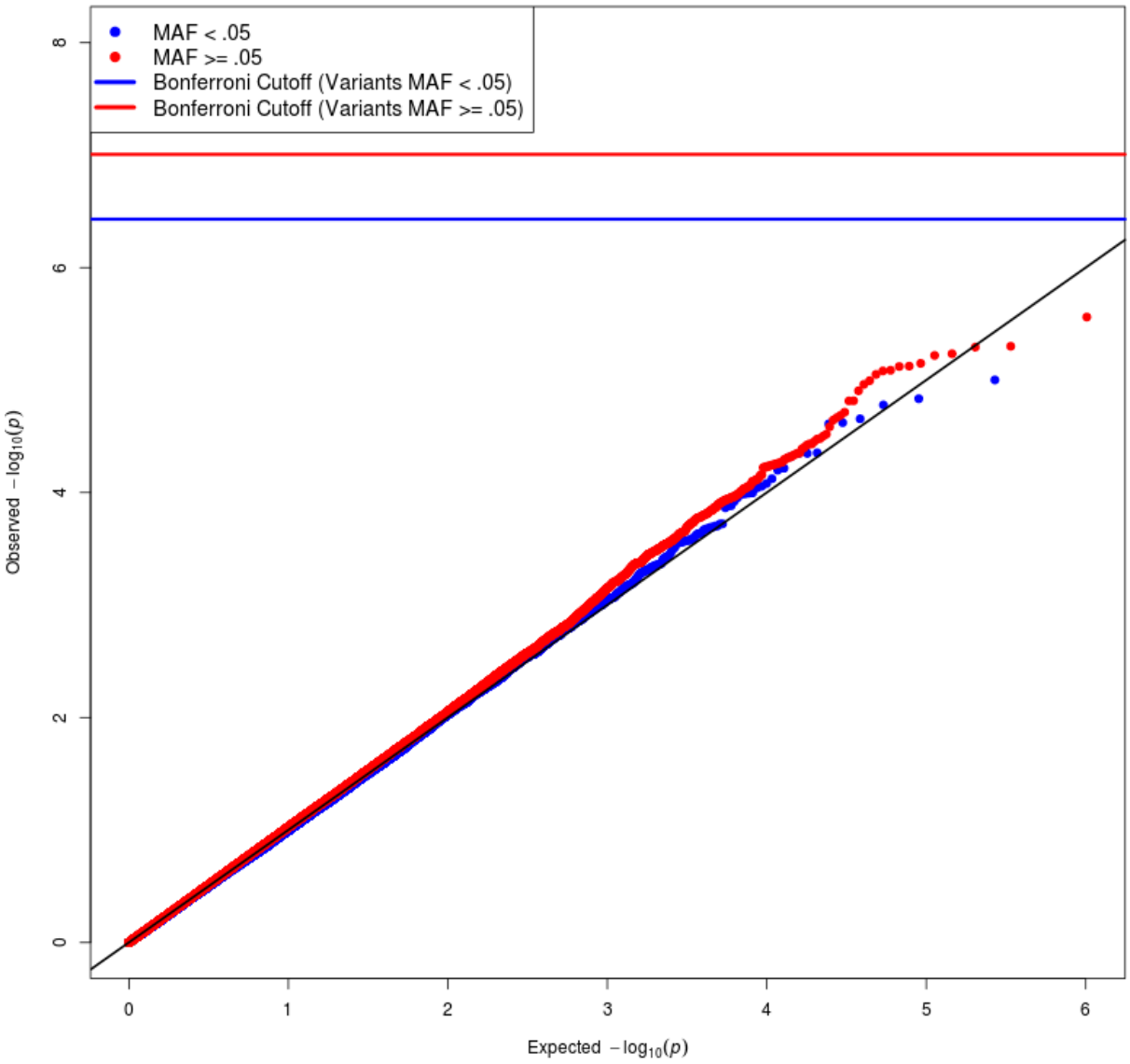

Figure S10. Single Variant QQplot of Alcohol Consumption in Individuals Exposed to Alcohol. MAF, minor allele frequency.

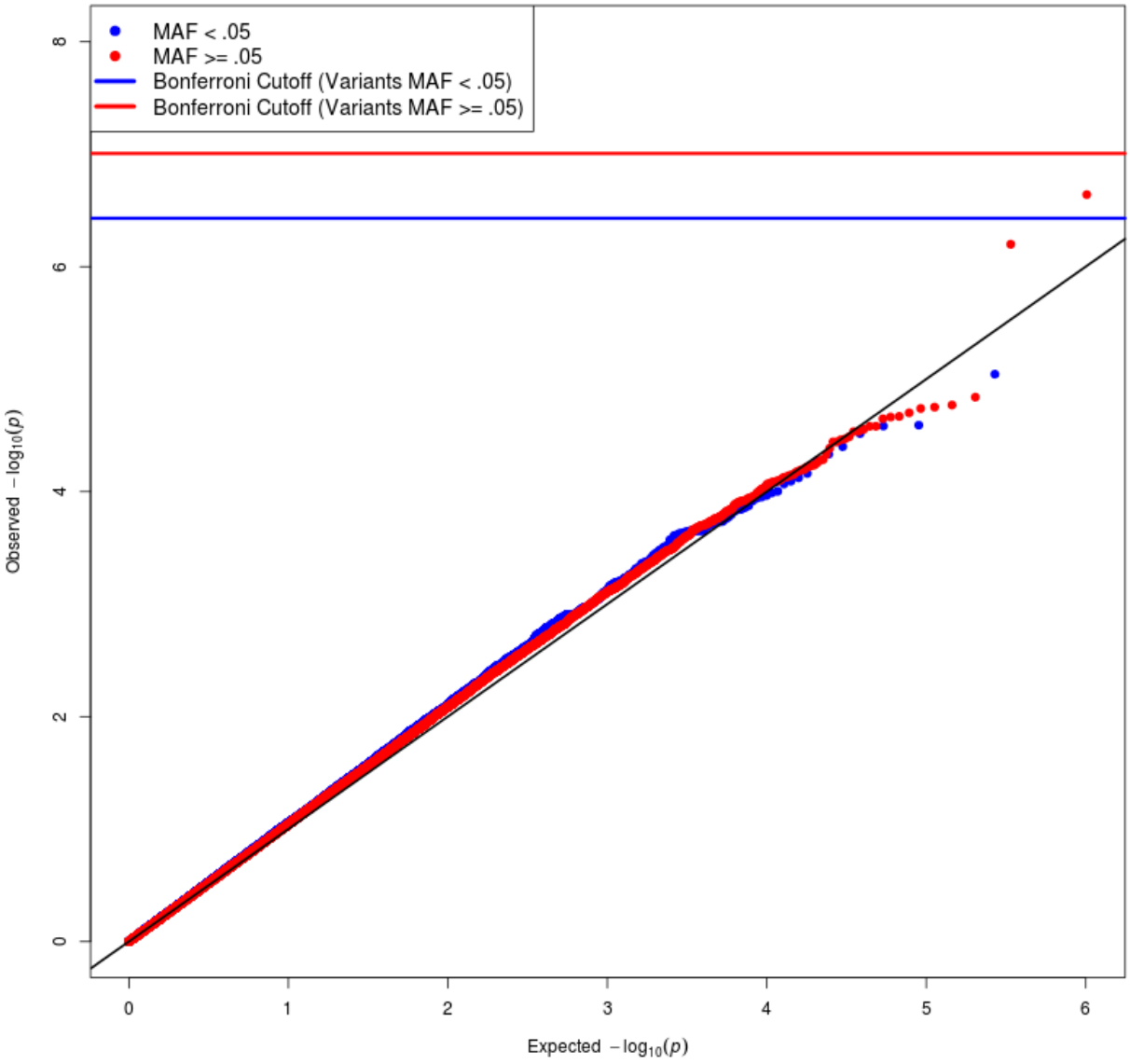

Figure S11. Single Variant QQplot of Alcohol Dependence in Individuals Exposed to Alcohol. MAF, minor allele frequency.

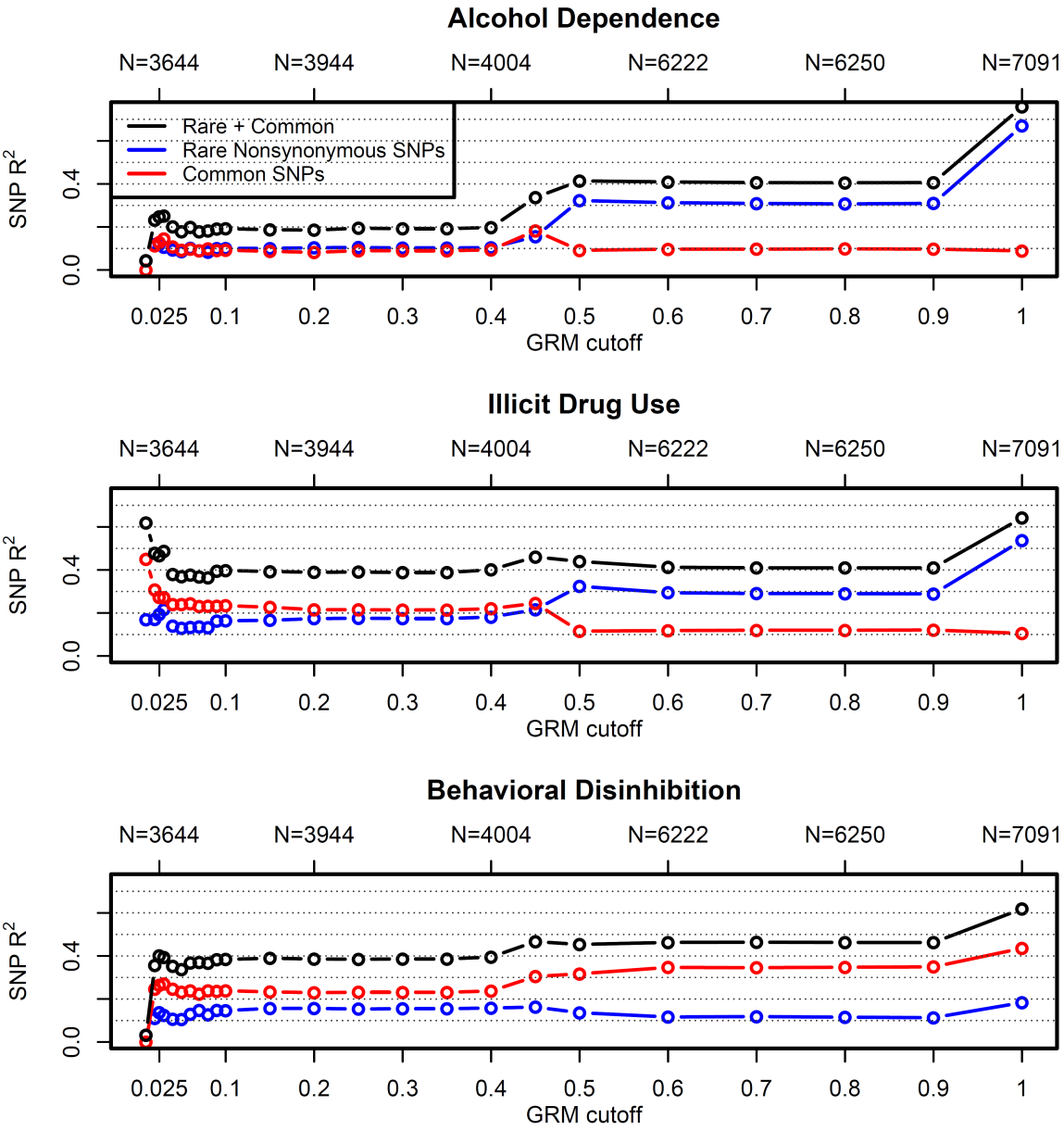

Figure S12. SNP-based Heritability as a Function of GRM Cutoff for Alcohol Dependence, Illicit Drug Use, and Behavioral Disinhibition. Note the x-axis is GRM cutoff, which is a measure of relatedness calculated from SNPs, similar to identity by state. A cutoff greater than .025 is typically not recommended for estimates of SNP-based heritability. In our family sample, cutoffs greater than .4 begin including first-degree relatives, such as siblings along with parents. At very high cutoffs such as 1 (i.e., no cutoff) the entire sample is included, including the monozygotic twin pairs. At high cutoffs the familial relatedness is being captured by the SNPs, such that the two independent sets of SNPs (rare and common) become proxies for known familial relationships, and are essentially redundant in accounting for familial relationships in the phenotype. This explains why the rare SNPs become prominent for high GRM cutoffs in alcohol dependence and illicit drug use, but not for behavioral disinhibition. GRM, genetic relationship matrix; SNP, single nucleotide polymorphism.

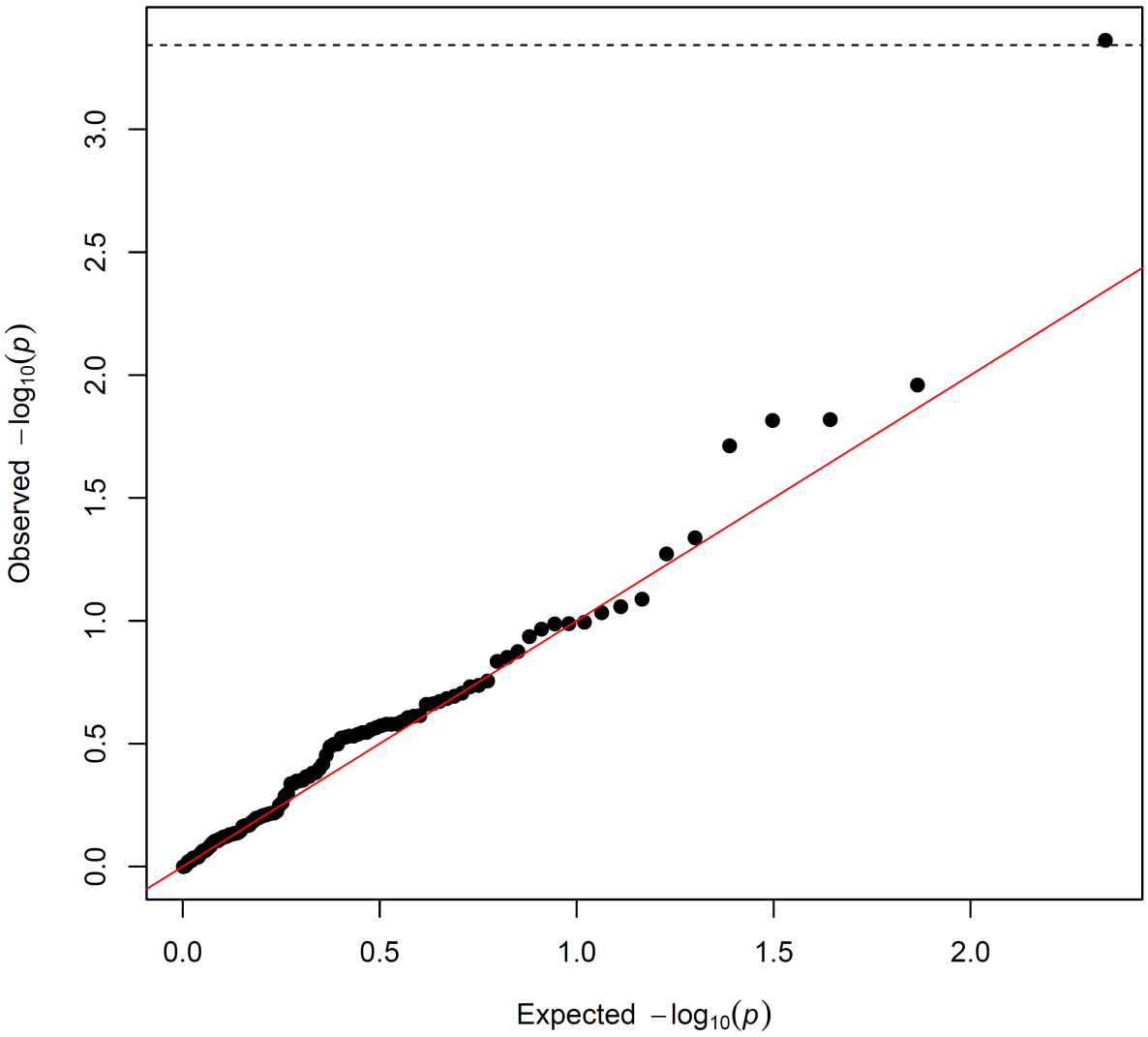

Figure S13. Nonsynonymous Variants in Genes Associated with Smoking/Drinking by GWAS Meta-Analysis. Dotted line is Bonferroni cutoff for 116 tests. The significant variant is 9:136521726 in *DBH*.

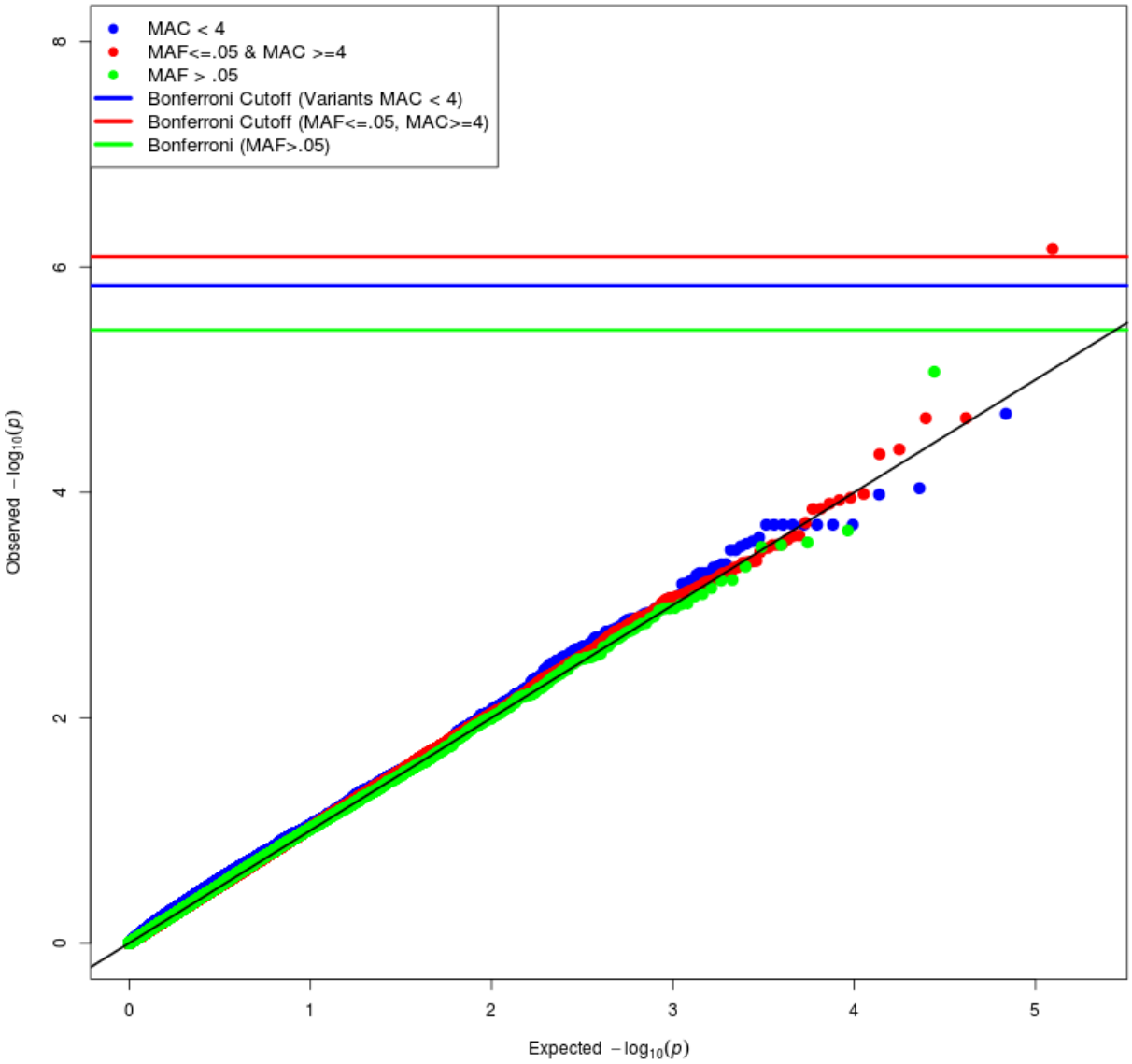

Figure S14. Single Variant QQplot of Nicotine, Considering Only Nonsynonymous Variants. MAC, minor allele count; MAF, minor allele frequency.

Table S5. Summary of Burden Tests for 151 Candidate Addiction Genes

See Supplement 2 (Excel file) for Table S5.

Table S6. Single Variant Statistics for High Priority Candidate Addiction Genes.

Gene	Phenotype	CHR	BP	MAF	MAC	Ref	Alt	Alt Eff Size	*p*-value
CHRNA5	Nicotine (Exposed)	15	78873272	0.00026	1	T	G	-0.089	0.9138
			78880752	0.01431	98	G	A	0.211	0.0458
			78882446	0.00052	5	C	T	0.557	0.2841
			78882726	0.00026	2	C	T	-0.136	0.8623
			78882821	0.00026	2	T	A	0.611	0.4288
			78882925	0.34079	2432	G	A	0.064	0.0151
			78885574	0.00026	3	T	A	0.200	0.7839
CHRNA3	Nicotine (Exposed)	15	78885574	0.00026	3	T	A	0.200	0.7839
			78893741	0.00026	1	A	C	1.109	0.2631
			78894335	0.00130	7	G	A	-0.316	0.4165
			78909414	0.00026	1	C	T	0.247	0.7643
			78911230	0.04865	344	C	T	0.073	0.2025
CHRNA6	Nicotine (Exposed)	8	42608381	0.00182	11	C	T	-0.067	0.8350
			42611120	0.00182	0	-	-	-	-
			42611665	0.00026	3	A	G	0.765	0.2128
			42611723	0.00026	0	-	-	-	-
			42611791	0.00026	1	G	T	0.699	0.4595
			42620264	0.00078	5	C	T	-0.095	0.8501
			42620320	0.00104	6	C	G	0.469	0.2895
CHRNB4	Nicotine (Exposed)	15	78917573	0.00026	1	T	C	-1.279	0.1966
			78921323	0.00052	3	C	T	0.175	0.7706
			78921343	0.00598	35	G	A	0.289	0.1030
			78921398	0.00026	3	A	G	0.755	0.2439
			78921716	0.00078	5	C	T	0.155	0.7327
			78922229	0.00702	47	T	C	-0.124	0.4304
			78922240	0.00884	54	C	T	-0.014	0.9213
			78923505	0.04110	288	G	A	0.069	0.2669
CYP2A6	Nicotine (Exposed)	19	41351935	0.00078	3	A	G	0.040	0.9435
CYP2B6	Nicotine (Exposed)	19	41512841	0.23049	1581	G	T	-0.033	0.2630
			41515943	0.23049	0	-	-	-	-
			41518204	0.00078	5	G	C	0.226	0.6487
			41518221	0.00026	3	T	C	-0.250	0.6810
			41518244	0.00026	1	C	T	0.454	0.6309
			41518265	0.00052	2	C	T	0.351	0.6064
EGLN2	Nicotine (Exposed)	19	41306568	0.00026	0	-	-	-	-
			41306650	0.01431	102	C	T	-0.047	0.6595
			41307024	0.00052	3	G	A	0.478	0.3987
DBH	Nicotine (Exposed)	9	136501569	0.00832	54	G	A	0.075	0.6054
			136501599	0.00026	2	G	A	-0.017	0.9820
			136501728	0.00546	31	C	T	-0.002	0.9933
			136501746	0.00026	1	G	A	0.247	0.8029
			136501756	0.00494	37	G	C	0.131	0.4582
			136501768	0.00494	0	-	-	-	-
			136501834	0.00234	15	T	C	0.617	0.0194
			136507332	0.00104	6	G	A	0.135	0.7316
			136507425	0.00026	3	G	A	-0.204	0.7377

Gene	Phenotype	CHR	BP	MAF	MAC	Ref	Alt	Alt Eff Size	*p*-value
			136507444	0.01977	142	A	G	-0.145	0.1025
			136507473	0.04683	305	G	A	0.000	0.9995
			136507474	0.00078	3	C	T	-0.055	0.9224
			136507528	0.00078	0	-	-	-	-
			136507559	0.00052	4	G	C	0.733	0.2072
			136508640	0.00026	1	G	A	1.274	0.1756
			136508658	0.00182	8	G	A	-0.076	0.8235
			136508691	0.00052	2	G	A	-0.373	0.5947
			136509370	0.08793	576	G	T	-0.110	0.0153
			136513028	0.00442	29	C	T	-0.079	0.6814
			136521654	0.00312	23	G	A	0.373	0.0873
			136521726	0.00078	6	G	A	1.398	0.0004
			136523487	0.00182	11	A	T	0.525	0.0926
			136523534	0.00052	6	C	T	-0.168	0.6780
			112221070	0.00038	5	G	C	0.537	0.2427
ALDH2	Drinking	12	112230413	0.00025	3	G	A	-0.103	0.8618
			112241766	0.00013	1	G	A	-1.052	0.2625
			100229016	0.00370	58	C	T	0.093	0.5162
			100229017	0.00089	14	G	A	-0.409	0.1460
			100232704	0.00166	25	C	T	0.106	0.6183
ADH1B	Drinking	4	100232773	0.00013	2	C	T	0.237	0.7479
			100239268	0.00229	41	A	G	-0.172	0.3252
			100239284	0.00892	108	T	A	0.131	0.1830
			100239319	0.00077	7	C	T	-0.317	0.3825
			69364311	0.00014	2	C	T	0.840	0.2571
			69583176	0.00014	3	C	T	-0.651	0.2849
			69583206	0.00352	40	A	G	0.174	0.2942
AUTS2	Drinking (Exposed)	7	70228020	0.03475	407	G	T	0.025	0.6356
			70229770	0.00028	5	A	G	-0.029	0.9541
			70255639	0.01548	186	G	T	0.200	0.0110
			70255956	0.00070	6	A	G	-0.064	0.8757
			70255978	0.00042	5	G	A	-0.920	0.0533
			112221070	0.00042	5	G	C	0.277	0.5492
ALDH2	Drinking (Exposed)	12	112230413	0.00028	3	G	A	-0.349	0.5640
			112241766	0.00014	1	G	A	-1.178	0.2174
			100229016	0.00394	47	C	T	0.056	0.7141
			100229017	0.00084	10	G	A	-0.483	0.1408
			100232704	0.00141	18	C	T	0.187	0.4477
ADH1B	Drinking (Exposed)	4	100232773	0.00014	1	C	T	0.736	0.4464
			100239268	0.00225	30	A	G	-0.197	0.3177
			100239284	0.00914	96	T	A	0.110	0.2938
			100239319	0.00056	5	C	T	-0.536	0.2185
			112221070	0.00038	5	G	C	0.240	0.6141
ALDH2	Alcohol Dependence	12	112230413	0.00025	3	G	A	0.634	0.2998
			112241766	0.00013	1	G	A	-0.641	0.5053
			100229016	0.00370	58	C	T	-0.049	0.7411
ADH1B	Alcohol Dependence	4	100229017	0.00089	14	G	A	-0.507	0.0814
			100232704	0.00166	25	C	T	0.243	0.2721

Gene	Phenotype	CHR	BP	MAF	MAC	Ref	Alt	Alt Eff Size	*p*-value
			100232773	0.00013	2	C	T	-0.268	0.7264
			100239268	0.00229	41	A	G	-0.183	0.3175
			100239284	0.00892	108	T	A	0.135	0.1855
			100239319	0.00077	7	C	T	-0.025	0.9470
AUTS2	Alcohol Dependence (Exposed)	7	69364311	0.00014	2	C	T	-0.083	0.9131
			69583176	0.00014	3	C	T	-0.720	0.2474
			69583206	0.00352	40	A	G	-0.023	0.8926
			70228020	0.03475	407	G	T	0.045	0.4147
			70229770	0.00028	5	A	G	-0.250	0.6245
			70255639	0.01548	186	G	T	0.130	0.1081
			70255956	0.00070	6	A	G	0.213	0.6092
			70255978	0.00042	5	G	A	-0.734	0.1332
ALDH2	Alcohol Dependence (Exposed)	12	112221070	0.00042	5	G	C	0.148	0.7545
			112230413	0.00028	3	G	A	0.580	0.3511
			112241766	0.00014	1	G	A	-0.740	0.4441
ADH1B	Alcohol Dependence (Exposed)	4	100229016	0.00394	47	C	T	-0.063	0.6875
			100229017	0.00084	10	G	A	-0.551	0.1013
			100232704	0.00141	18	C	T	0.275	0.2764
			100232773	0.00014	1	C	T	-0.304	0.7569
			100239268	0.00225	30	A	G	-0.319	0.1158
			100239284	0.00914	96	T	A	0.111	0.2975
			100239319	0.00056	5	C	T	0.118	0.7892

Alt, alternate; BP, base pair; CHR, chromosome; Eff, effect; LD, linkage disequilibrium; MAC, minor allele count; MAF, minor allele frequency; Ref, reference.

Variants with 0 MAC are in perfect LD with neighboring variant, and are set to missing to avoid double-counting.