Fine mapping on chromosome 13q32-34 and brain expression analysis implicates *MYO16* in schizophrenia

Laura Rodriguez-Murillo[1], Bin Xu[1,2], J. Louw Roos[3], Gonçalo R. Abecasis[4], Joseph A. Gogos[2,5], Maria Karayiorgou[1,*]

Author affiliations:

[1]Columbia University, Department of Psychiatry, 1051 Riverside Drive, New York, NY 10032;

[2]Columbia University, Department of Physiology, 630 West 168th Street, New York, NY 10032;

[3]University of Pretoria, Department of Psychiatry and Weskoppies Hospital, Pretoria, RSA 0001;

[4]University of Michigan, Department of Biostatistics, Center for Statistical Genetics, 1420 Washington Heights, Ann Arbor, MI 48109-2029;

[5]Columbia University, Department of Neuroscience, 630 West 168th Street, New York, NY 10032

[*]**Corresponding author**

Tel.: 212-568-4189

E-mail: mk2758@columbia.edu

Running title: MYO16 fine mapping and brain expression in schizophrenia

Key words: Schizophrenia, association, MYO16, linkage, CNVs, SNPs, expression

Abstract

We previously reported linkage of schizophrenia and schizoaffective disorder to 13q32-34 in the European descent Afrikaner population from South Africa. The nature of genetic variation underlying linkage peaks in psychiatric disorders remains largely unknown and both rare and common variants may be contributing. Here, we examine the contribution of common variants located under the 13q32-34 linkage region. We employ densely spaced

SNPs to fine map the linkage peak region using both a discovery sample of 415 families and a meta-analysis incorporating two additional replication family samples. In a second phase of the study, we use one family-based dataset with 237 families and independent case-control datasets for fine mapping of the common variant association signal using HapMap SNPs. We report a significant association with a genetic variant (rs9583277) within the gene encoding for the myosin heavy chain Myr 8 (MYO16), which has been implicated in neuronal phosphoinositide 3-kinase (PI3K) signaling. Follow-up analysis of HapMap variation within *MYO16* in a second set of Afrikaner families and additional case-control datasets of European descent highlighted a region across introns 2 to 6 as the most likely region to harbor common *MYO16* risk variants. Expression analysis revealed a significant increase in the level of MYO16 expression in brains of schizophrenia patients. Our results suggest that common variation within *MYO16* may contribute to the genetic liability to schizophrenia.

Introduction

Susceptibility to schizophrenia is determined by multiple genetic and possibly environmental factors. Recent studies addressing the role of high-penetrant rare variants (Walsh *et al*, 2008; Xu *et al*, 2012; Xu *et al*, 2011; Xu *et al*, 2008; Xu *et al*, 2009) or common genetic variants with low effect (ISC, 2008; Lee *et al*, 2012; O'Donovan *et al*, 2008; Ripke S, 2011; Shi *et al*, 2009; Shi *et al*, 2011; Stefansson *et al*, 2009) suggest that patient genomes contain risk alleles at a wide range of frequencies, some driving and some merely modifying the disease risk and expression, which in concert may affect the structure and function of neural circuits (ISC, 2008; O'Donovan *et al*, 2008; Rodriguez-Murillo *et al*, 2012; Shi *et al*, 2009; Stefansson *et al*, 2009; Xu *et al*, 2011; Xu *et al*, 2008; Xu *et al*, 2009).

In complex diseases, the genetic structure of linkage signals most likely involves one or several rare alleles with strong effect on disease risk, or a combination of rare and common alleles, in the same or different genes (Bowden *et al*, 2010). Also, linkage analyses of inbred mice have shown that more than one gene can contribute to the same linkage signal for a given QTL trait (Karst *et al*, 2011). Along the same lines, association

studies coupled with targeted re-sequencing have suggested that the same genes carrying common risk variants can also show an excess of rare risk variants implicated in the disease (Cirulli and Goldstein, 2010; Di Rienzo, 2006; Manolio *et al*, 2009; Trynka *et al*, 2011).

A 9-cM genome-wide linkage scan on families from the European descent Afrikaner population from South Africa identified three linkage signals on chromosomes 1, 9 and 13 (Abecasis *et al*, 2004). Subsequently, we increased the genomic coverage to better define the linkage regions, and performed a 2-cM genome-wide linkage scan on an extended set of Afrikaner families. The results from this genome scan identified chromosome 13q32-34 as the most robustly linked locus in this population. We also addressed the contribution of rare CNVs to schizophrenia in this cohort and found that, at the level of resolution of the linkage scan, none of the linkage signals observed in these families may be caused by the presence of CNVs within these genomic intervals (Xu *et al*, 2009).

Here, we present the results of our ongoing systematic effort to elucidate the genetic structure of our 13q32-34 linkage peak obtained in our 2-cM genome scan, by analyzing the contribution of local common variants via a multistage association study. In addition to genuine contributions to the risk associated with a given linkage signal, even in cases where the linkage signal is accounted for only by rare variants, common variants may in some cases help pinpoint with more accuracy the location of rare risk variants (Dickson *et al*, 2010; Lin *et al*, 2004; Sanna *et al*, 2011). First, we genotyped 1223 individuals from 415 Afrikaner families for 723 SNPs localized within 13q32-34. Subsequently, the most significant SNPs were followed-up in two independent family-based replication samples of European origin. One SNP showed replicated association in one of the two independent samples and remained significant after meta-analysis and correction for multiple testing. This SNP is located within the *MYO16* (myosin XVI) gene (Patel *et al*, 2001; Yokoyama *et al*, 2011). Second, we performed a comprehensive fine-scale mapping of the genetic contribution of this gene with respect to common variation, by genotyping an independent set of families from the Afrikaner population for 102 SNPs within *MYO16* and by imputing the rest of the HapMap SNPs within the gene boundaries.

These analyses identified a preponderance of common variants implicated in schizophrenia within introns 2-6 of the gene *MYO16*. Furthermore, expression analysis of the *MYO16* gene in brain samples from patients and controls identified a significantly elevated level of expression in patients with schizophrenia.

Methods and Materials

We used a family-based approach studying families with at least one affected individual per family. Datasets are presented in Supplementary Table 1.

Afrikaner Cohorts: Affected families were recruited and diagnosed as part of our ongoing, large-scale genetic study of schizophrenia in the European descent Afrikaner population from South Africa, as previously described (Abecasis *et al*, 2004; Karayiorgou *et al*, 2004; Xu *et al*, 2008; Xu *et al*, 2009). Affected subjects were classified as either narrowly or broadly affected. The narrow diagnosis includes subjects with schizophrenia or schizoaffective disorder-depressive type, as previously described (Abecasis *et al*, 2004; Xu *et al*, 2009). The broad diagnosis includes all individuals classified under the narrow definition, as well as individuals with schizoaffective disorder-bipolar type (Xu *et al*, 2009).

Afrikaner set 1 (SAF1): This dataset includes the 143 families used for the linkage scan, plus an additional 272 families. The entire set comprises 474 affected individuals who meet the narrow diagnostic criteria or 741 who meet the broad diagnostic criteria.

Afrikaner set 2 (SAF2): This dataset includes 237 families, 85 of whom have family history of schizophrenia in the previous two generations. 232 individuals in these families meet the narrow diagnostic criteria, while 266 individuals meet the broad diagnostic criteria.

Rutgers families: From the entire set of families collected under the NIMH Schizophrenia Genetics Initiative, maintained by the Rutgers University Cell and DNA Repository, we selected a subset of 301 families matched according to ethnicity. Our selected Caucasian,

Table I LAMP *P*-values and risk alleles for the discovery (SAF1) and replication (Rutgers and US) family samples. Na stands for narrow schizophrenia and Bd for broad schizophrenia. Rutgers and US phenotype is broad schizophrenia. Meta-analysis was performed with SAF1(Bd).

						SAF1 (Na)			SAF1 (Bd)			Rutgers			US			
CHR	SNP	bp	GENES	A1	A2	FREQ A1	*P*	TDT OR	FREQ A1	*P*	TDT OR	FREQ A1	*P*	TDT OR	FREQ A1	*P*	TDT OR	META-*P*
13	rs1323666	108054116	*FAM155A*	G	C	0.48	0.021	0.86	0.46	0.087	0.89	0.43	0.550	1.03	0.47	0.450	0.80	0.245
13	rs716504	108061624	*FAM155A*	G	C	0.32	0.026	1.39	0.33	0.880	1.33	0.35	0.360	0.89	0.33	0.750	1.03	0.786
13	rs7994782	108393891	*FAM155A*	G	A	0.46	0.002	1.28	0.46	0.002	1.20	0.46	0.890	1.23	0.48	0.980	0.97	0.036
13	rs2940695	108567155	-	G	A	0.38	0.012	1.17	0.38	0.021	1.08	0.42	0.620	1.11	0.44	0.160	0.78	0.244
13	rs4325412	108708242	-	T	G	0.23	0.003	0.63	0.23	0.014	0.77	0.20	0.045	1.02	0.19	0.630	1.18	0.790
13	rs9583277	109333749	*MYO16*	C	A	0.26	0.002	1.35	0.25	0.001	1.40	0.36	0.002	1.18	0.33	0.600	0.88	**$2.25\ 10^{-4}$**
13	rs277828	109693885	*MYO16*	C	A	0.22	0.011	1.24	0.21	0.010	1.21	0.26	0.460	0.86	0.29	0.260	0.74	0.444
13	rs9521372	110093044	-	T	G	0.16	0.019	0.66	0.16	0.072	0.79	0.13	0.440	0.94	0.12	0.360	1.37	0.225
13	rs4773155	110971066	*COL4A2*	C	A	0.46	0.013	0.79	0.45	0.036	0.84	0.39	0.590	1.05	0.37	0.550	1.04	0.416
13	rs9515201	111040798	*COL4A2*	C	A	0.30	0.004	1.41	0.30	0.021	1.31	0.32	0.240	0.78	0.34	0.058	0.84	0.982
13	rs1927343	111053959	*COL4A2*	A	C	0.32	0.011	1.16	0.33	0.013	1.17	0.37	0.160	1.01	0.37	0.015	0.87	0.190
13	rs3742193	111280002	*FLJ10769*	T	C	0.10	0.001	1.41	0.11	0.014	1.31	0.11	0.270	1.23	0.15	0.470	0.75	0.054
13	rs4771711	111502659	-	G	A	0.42	0.009	1.27	0.41	0.003	1.26	0.44	0.054	0.88	0.48	0.960	0.95	0.395
13	rs1163830	112113080	-	G	A	0.44	$2.0\ 10^{-04}$	0.68	0.45	$7.6\ 10^{-04}$	0.65	0.46	0.250	0.76	0.45	0.990	0.97	0.004
13	rs1550192	112833723	-	T	C	0.18	0.014	1.55	0.17	0.002	1.44	0.19	0.150	1.33	0.18	0.910	1.03	0.004

*TDT ORs are calculated with respect to the A1 allele

European ancestry, subset includes a total of 1241 individuals (631 affected with schizophrenia).

US families: 210 trios (consisting of one affected individual and both biological unaffected parents for a total of 630 individuals) were included in this sample of Caucasian, European descent families recruited from the US. All probands met full diagnostic criteria for schizophrenia or schizoaffective disorder. Description of this dataset and the methods of subject selection and clinical evaluation have been previously described in Sobin et al. (Sobin *et al*, 2001; Sobin *et al*, 2003).

GAIN dataset: This study is part of the Genetic Association Information Network (GAIN) (ID phs000021.v2.p1). Details on inclusion criteria and participants are available at dbGap (Suarez *et al*, 2006). In total, 1314 cases and 1368 controls of European descent were included in the final set.

MGS_nonGAIN dataset: This study is part of the Molecular Genetics of Schizophrenia (MGS) genome wide association study (ID phs000167). Details on inclusion criteria and participants are available at dbGap. 1405 cases and 1347 controls of European descent were included in the final set.

PGC dataset: This dataset is part of the Schizophrenia Psychiatric Genome-wide association study consortium (Ripke *et al*, 2011). We included the results from stage 1 mega-analysis published in Ripke et al (Ripke *et al*, 2011) that correspond to the *MYO16* gene region overlapping SNPs genotyped or imputed in our SAF2 dataset. This dataset included 9,394 schizophrenia cases.

Genotyping, quality control and imputation

SAF1: Family members were genotyped for 723 SNPs covering 14.65 Mb under the 13q32-34 linkage peak and within candidate genes (*ZIC2, ZIC5, NALCN, FGF14, G72* and *EFNB2*) in the immediate vicinity of the linkage peak (Supplementary Table 2), on the Illumina GoldenGate platform at the Center for Inherited Disease Research (CIDR).

Rutgers and US samples: Family members were genotyped for 22 SNPs on a Taq Man Open Array Genotyping Platform (Applied Biosystems). These 22 SNPs were chosen

among the top associated SNPs resulting from the association analysis in stage 1 (SAF1) or surrogates of those (i.e., in strong LD with at least one of the top associated SNPs).

SAF2: This set of families was genotyped as part of a wider genotyping project on a Human Genome-Wide SNP Array 5.0 (Affymetrix), which contains 500,568 SNPs (manuscript in preparation). Samples were processed as previously described (Xu *et al*, 2008). Average call rate on arrays used in this study was 99.43%. All microarray experiments were performed in the Vanderbilt Microarray Shared Resource.

GAIN and *MGS*: Individual genotypes as well as phenotypic information were available to download from the dbGap website. Only individuals of European descent were included in the analysis.

For all datasets, quality control procedures per family, individual, and marker were performed with PLINK (Purcell *et al*, 2007) and PedStats (see URLs). All datasets went through quality control and we only selected samples with a call rate > 95%. We eliminated from the analysis duplicated SNPs, monomorphic SNPs, and SNPs with Hardy-Weinberg Exact Test $P < 10^{-6}$. Only SNPs with minor allele frequency over 0.01 were included in the downstream analyses. We also checked for Mendelian inheritance errors among families, and removed SNPs with more than 4 Mendelian errors in the total sample. For the case-control datasets, we corrected for population stratification with the program EIGENSTRAT, eliminating outliers from the downstream analyses.

Imputation of non-genotyped HapMap SNPs for SAF2, GAIN and MGS datasets was performed with MACH (see URLs) using 100 Markov iterations with the two-step procedure recommended in the manual. HapMap Phased Haplotypes (release 22) on CEU subjects were used in the imputation. After imputation, only SNPs with a MACH R^2 over 0.3 were further considered. This estimates the correlation between imputed and true genotypes; a value less than 0.3 flags poorly imputed SNPs (Li *et al*, 2010). In addition, Mendelian checks (for the family-based samples) and Hardy-Weinberg equilibrium tests were performed to eliminate unreliable imputation calls in order to include imputed

genotypes in downstream analyses. Imputed SNPs were then analyzed as the genotyped SNPs.

Statistical analyses

Family-based association testing for single SNPs was performed using LAMP (see URLs) (Li *et al*, 2005, 2006). We adopted a free model for the analysis that does not constrain the penetrances for the three genotypes. Haplotype-based associations were assessed by means of the Transmission Disequilibrium Test (TDT) for haplotypes implemented in PLINK. For the case-control datasets, a trend-test was performed to evaluate the SNP association. We applied Bonferroni correction in all tests to obtain an α corrected threshold. We calculated the number of independent tests in each case based on LD patterns between SNP pairs. These procedures were performed in PLINK (see URLs).

Meta-analysis of the results for the independent samples was performed with Metal (see URLs). The algorithm checks for heterogeneity and performs meta-analysis under a fixed effects model. All base pair positions are based on the current Human genome assembly (hg19) (see URLs).

To identify duplicated individuals and family relationships between individuals across datasets, we performed identity by descent (IBD) analysis of GAIN, MGS_nonGAIN and Rutgers samples merged together using PLINK. Duplicated and related individuals across datasets were removed from all but one of the datasets to avoid bias in the analysis. Specifically, GAIN and MGS_nonGAIN included 10 duplicated individuals that were removed from the larger MGS_nonGAIN dataset.

Expression analysis

Total RNA from frontal cortex was obtained from the Stanley Medical Research Institute (SMRI) (Bethesda, Maryland) (see URLs). The SMRI Array Collection includes 35 individual subjects in each of three groups: control, schizophrenia and bipolar disorder subjects (Torrey *et al*, 2000). qRT-PCR was performed with pre-designed TaqMan® Gene Expression assay by ABI (Applied Biosystems; ABI assay number

#Hs01031284_m1) on a 7900HT Fast Real Time PCR system (Applied Biosystems). Human glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as the endogenous control. Relative quantitation of expression comparing the three groups (schizophrenia, bipolar disorder, and controls) was tested with generalized linear models (GLM) and incorporating covariates into the model. Descriptive statistics, means comparison and GLM were analyzed with R statistics software.

Results

SNP association in a discovery set of Afrikaner families, replication, and meta-analysis

To follow-up the 13q32-34 linkage signal obtained through our 2-cM coverage linkage scan, we genotyped 723 SNPs from chromosome 13, on 1223 individuals from the 143 Afrikaner families included in the 2-cM linkage scan plus 272 additional families from the same homogeneous population (SAF1). We performed family-based association tests on these 415 families using LAMP (Li *et al*, 2005, 2006). 115 SNPs reached nominal significance at this stage (Figure 1 and Supplementary Table 2). None of these SNPs were within candidate genes (*ZIC2, ZIC5, NALCN, FGF14, G72 and EFNB2*) abutting the linkage peak. 22 associated SNPs were followed-up in two independent samples of European descent (Rutgers and US samples, see Methods). These 22 SNPs were selected from the top associated SNPs, or surrogates of these, chosen based on the LD structure, each one representing one independent LD block, and the availability of a genotyping assay in a TaqMan Open Array Genotyping platform (Applied Biosystems). Family members from the replication samples (Rutgers and US) were genotyped for these 22 SNPs (Table 1). Following quality control procedures, 7 SNPs were removed from the analysis, two due to bad calls, four due to Mendelian errors (these were not concentrated in specific families), and one due to deviations from Hardy-Weinberg equilibrium. The remaining 15 SNPs were tested for association in the replication samples by using LAMP. Table 1 shows the *P*-values and odd ratios for the association in the discovery sample (SAF1) and in the replication samples (Rutgers and US). Subsequently, meta-

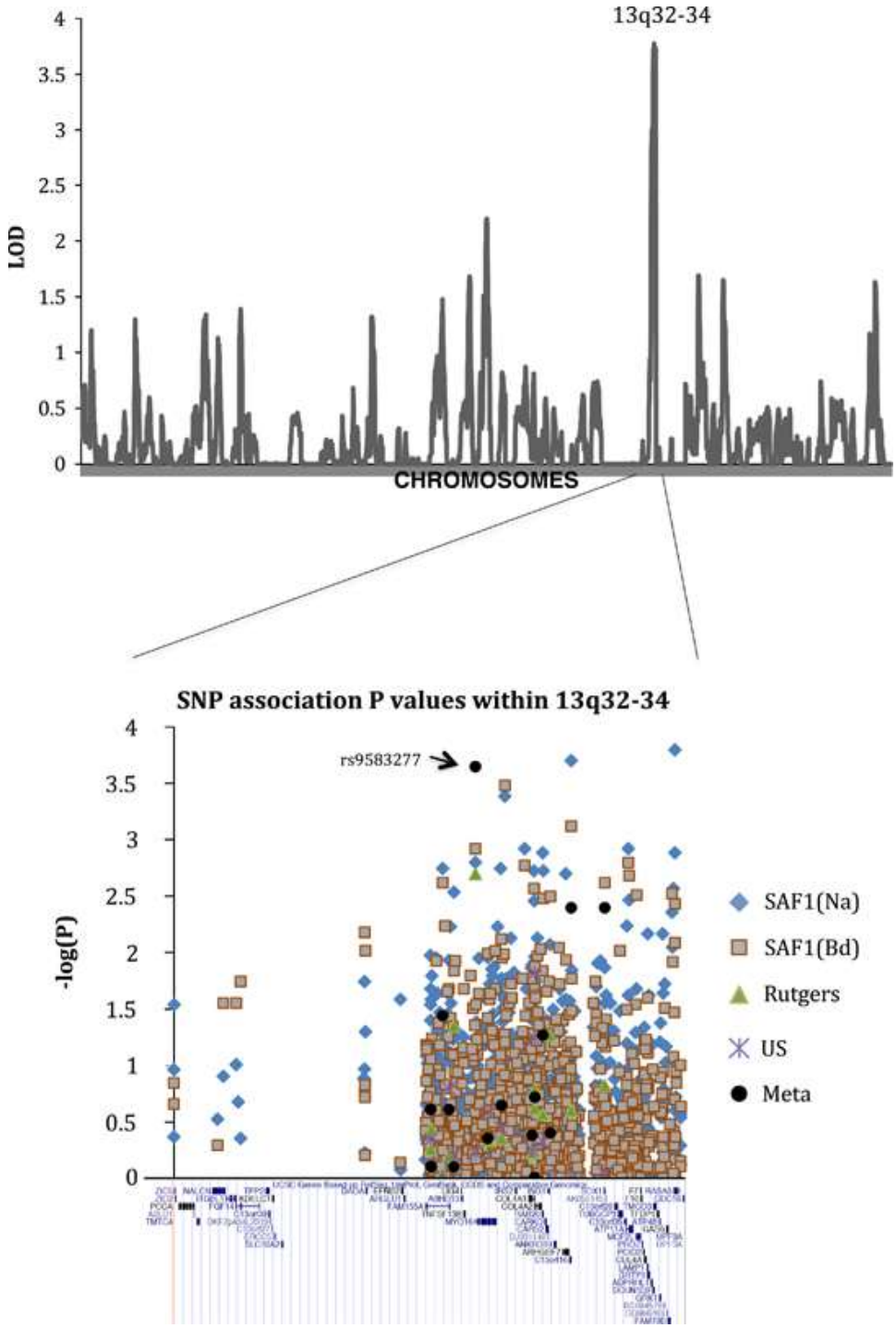


Figure 1: Genome-wide linkage and fine mapping. Bd, broad; LOD, logarithm of the odds; Na, narrow; SNP, single-nucleotide polymorphism.

analysis was performed combining *P*-values obtained from the SAF1 and both replication samples. Meta-analysis identified one SNP with combined *P*-values that survive Bonferroni in correction for multiple testing (Table 1) ($\alpha_{corrected} = 0.0023$). The top associated SNP, rs9583277, has a meta-analysis *P*-value of $1.86x10^{-4}$ and $2.25x10^{-4}$ for the combined sample SAF1-Rutgers-US families, for both narrow and broad definition of schizophrenia, respectively. It is worth noting that the Bonferroni correction we employ to declare significance reflect the number of independent tests (n) we performed (n = 22, $\alpha_{corrected} = 0.05/22 = 0.0023$) and therefore are not as stringent as thresholds employed in GWAS that reflect corrections for ~1 million tests performed. The identified variant (rs9583277) maps to 109,333,749 bp on chromosome 13q33.3, within the second intron of the *MYO16* gene. Our previous linkage analysis indicated dominant inheritance for the risk locus at 13q32-34 (Xu *et al*, 2009). Consistent with this finding, we did not detect any excess of homozygosity (an indication of recessive mode of inheritance) at rs9583277, either when the entire SAF1 dataset was considered or upon stratified analysis including only families linked to 13q32-34 (data not shown). Overall, our analysis, employing densely spaced SNPs to fine map the prior 13q32-34 linkage peak region on a discovery and two replication family samples (a total of 923 families), highlighted a potential contribution of the *MYO16* gene locus.

Fine mapping of the common variant association signal using *MYO16* HapMap SNPs

Having identified a significant association with a genetic variant (rs9583277) within the *MYO16* gene, we then performed a comprehensive fine scale mapping of the common variant association signal in an independent set of families and cases. To this end, we genotyped, or imputed when necessary, all HapMap SNPs within the *MYO16* gene boundaries, according to UCSC genome browser genome positions (hg19). First, we examined a sample of 228 Afrikaner families with an average of 3.2 individuals per family (SAF2). Following quality control procedures (see Methods), 102 genotyped and 470 imputed HapMap SNPs were available for analysis of the *MYO16* gene locus with respect to underlying common risk variants. It should be noted that there is no overlap with *MYO16* SNPs genotyped in SAF1 and that the SNP previously found associated in

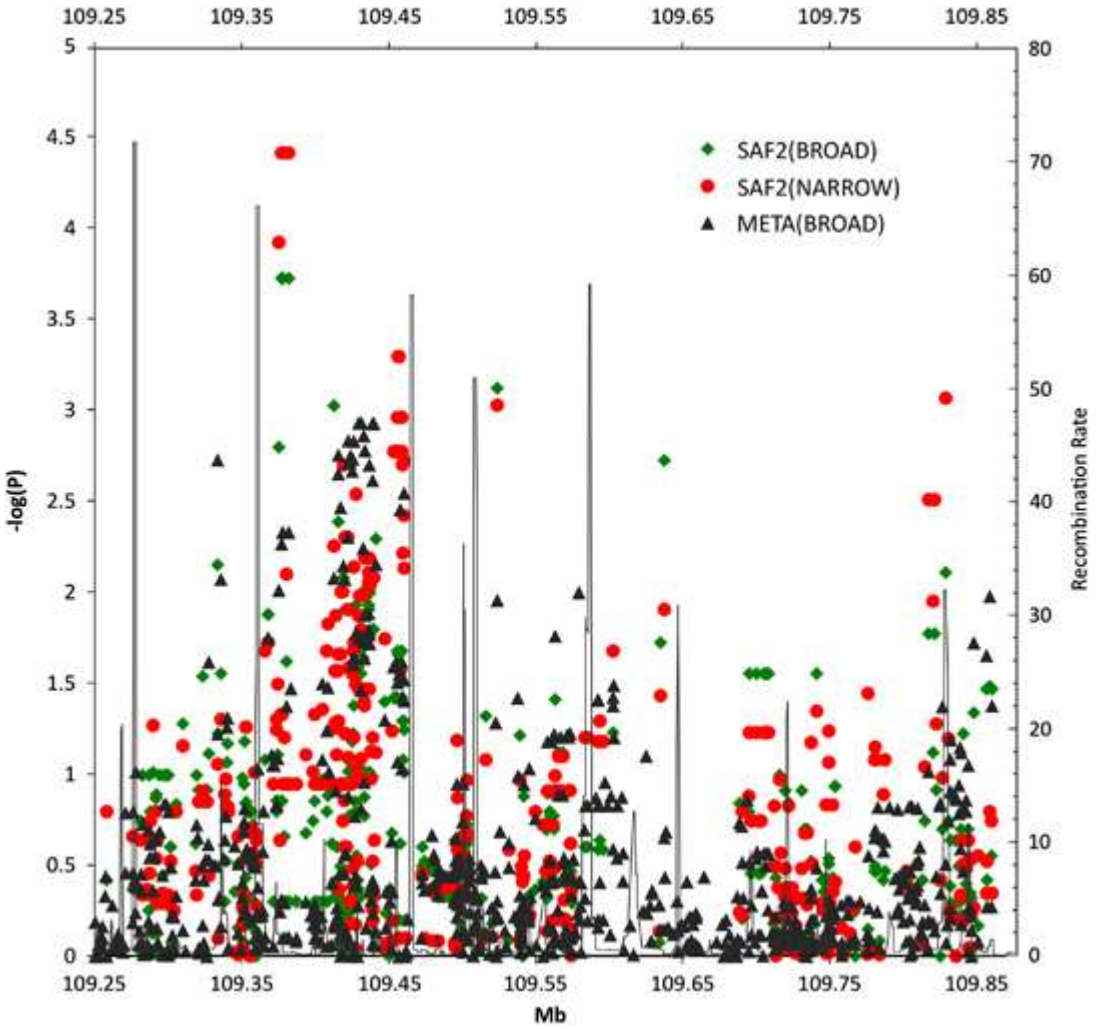


Figure 2: Plot depicts the negative logarithm of the SNP association *p* values for the Afrikaner Set 2 data set with narrow (SAF2(NARROW)) and broad (SAF2(BROAD)) schizophrenia diagnosis, as well as the meta-analysis *p*-values for the combined sample SAF2(BROAD)/GAIN/MGS data sets (META(BROAD)). The background graph represents the recombination rate throughout the region. GAIN, Genetic Association Information Network; MGS, Molecular Genetics of Schizophrenia.

Table 2 LAMP *P*-values and maximum-likelihood estimates of Penetrance, Genotype Relative Risk (GRR), Population Attributable Risk (PAR), and Odd Ratios (OR) for top associated SNPs in SAF2.

SNP	bp	*P*	A1	A2	FREQA1	Pen(1/1)	Pen(1/2)	Pen(2/2)	GRR(1/2)	PAR	TDT-OR*
Narrow											
rs9520990	109376993	$3.9\ 10^{-5}$	C	T	0.685	0.014	0.007	0.005	1.37	0.47	1.82
rs9520991	109378071	$3.9\ 10^{-5}$	G	A	0.685	0.014	0.007	0.005	1.37	0.47	1.82
rs932678	109382119	$3.9\ 10^{-5}$	A	C	0.685	0.014	0.007	0.005	1.37	0.47	1.82
rs9514889	109375338	$1.2\ 10^{-4}$	A	G	0.668	0.013	0.007	0.006	1.17	0.37	1.69
rs984298	109456170	$5.1\ 10^{-4}$	G	A	0.792	0.008	0.013	0.018	0.72	0.81	0.72
rs9514918	109457217	$5.1\ 10^{-4}$	C	G	0.792	0.008	0.013	0.018	0.72	0.81	0.72
rs7324758	109828818	$8.6\ 10^{-4}$	T	C	0.821	0.012	0.006	0.005	1.38	0.55	1.11
rs7321660	109523912	$9.4\ 10^{-4}$	T	C	0.628	0.010	0.012	0.006	2.00	0.45	2.10
Broad											
rs9520990	109376993	$1.9\ 10^{-4}$	C	T	0.687	0.013	0.008	0.006	1.42	0.44	1.69
rs9520991	109378071	$1.9\ 10^{-4}$	G	A	0.687	0.013	0.008	0.006	1.42	0.44	1.69
rs932678	109382119	$1.9\ 10^{-4}$	A	C	0.687	0.013	0.008	0.006	1.42	0.44	1.69
rs7321660	109523912	$7.6\ 10^{-4}$	T	C	0.641	0.009	0.012	0.007	1.84	0.34	1.40
rs9521010	109412639	$9.5\ 10^{-4}$	G	A	0.505	0.011	0.012	0.006	1.86	0.38	1.41
rs9521011	109412704	$9.5\ 10^{-4}$	A	G	0.505	0.011	0.012	0.006	1.86	0.38	1.41

*ORs are calculated with respect to the A1 allele.

the SAF1 sample (rs9583277) was not genotyped in the SAF2 sample since it is not a HapMap SNP, nor is it present in the common genotyping platforms. Therefore, this stage is not intended to be a replication of the previous findings, but a deeper characterization of the common variation within *MYO16* in the context of schizophrenia. Even though several SNPs genotyped at this stage are located in the general vicinity of rs9583277 within the *MYO16* gene, rs9583277 is in a region of low LD.

Figure 2 shows LAMP *P*-values for the association of *MYO16* SNPs for both narrow and broad definitions of schizophrenia, along with the recombination frequency across the region. Table 2 shows the top associated SNPs within this dataset. Notably, 4 SNPs showed significant association with narrow definition schizophrenia after correction for multiple testing ($\alpha_{corrected} = 2.17x10^{-4}$, based on a Bonferroni correction after estimating the number of independent tests to 230, taking into account the LD pattern among SNP pairs). Three of these four SNPs also showed association with the disease under its broader definition. All these four SNPs were located within intron 3 of the *MYO16* gene, within an LD block that expands from intron 2 to intron 6.

We also investigated if there was any specific configuration of alleles or haplotypes conferring susceptibility to schizophrenia for either narrow or broad definition. Only directly genotyped SNPs were used to test association on haplotypes. Haplotype-based association in Afrikaner families was assayed with the TDT. First, we estimated haplotype blocks based on the LD structure by means of the default procedure implemented in Haploview. Subsequently, each haplotype within each block was tested for association with the hap-tdt option implemented in PLINK. In this fashion, we tested 74 haplotypes, each comprising of 2 to 10 SNPs. Table 3 shows the top associated haplotypes for either schizophrenia definition. Two distinct two-SNP haplotypes show under-transmission and significant association with schizophrenia ($\alpha_{corrected} = 6.76x10^{-4}$, 74 independent tests). It is worth noting that, of the two haplotypes with significant *P*-values, the CG haplotype including SNPs rs558322 and rs4976845 is associated with the narrow definition; and the GA haplotype including rs4578513 and rs10492418 is associated with the broad definition. Notably, these two haplotypes reside within distinct haplotype blocks, suggesting that the observed association signals are independent of

Table 3 Haplotypic association

	Start (bp)	End (bp)	Kb	Haplotype	Freq	T	U	*P*	SNPs
	109379006	109380726	1.721	GA	0.313	55	96	$8.48\ 10^{-4}$	rs4578513|rs10492418
Narrow	109819961	109824928	4.968	CG	0.192	30	64	**$4.53\ 10^{-4}$***	rs558322|rs4976845
	109379006	109380726	1.721	GA	0.313	65	110	**$6.69\ 10^{-4}$**	rs4578513|rs10492418
Broad	109819961	109824928	4.968	CG	0.192	43	71	0.0087	rs558322|rs4976845

α Bonferroni corrected = 0.00068

T: Transmitted

U: Untransmitted

each other. Of note, there are other examples where independent haplotypes are associated with distinct forms of a disease (Cruz *et al*, 2008). In our study, the two independent haplotypes might be acting as modifiers of the clinical presentation or reflect two distinct patient subpopulations.

We extended our follow-up studies to two additional, independent, case-control datasets, which are part of genome-wide genotyping projects (GAIN and MGS). To this end, we extracted SNP genotypes located within the *MYO16* gene boundaries and also imputed non-genotyped HapMap SNPs from this region in order to facilitate comparison with the SNPs in our SAF2 dataset. The SNPs extracted from GAIN and MGS datasets matched the SNPs in SAF2, and so the LD patterns were equivalent. Therefore, we employed the same significance threshold for the SAF2 as well as the GAIN and MGS datasets. Following quality control procedures, 572 single SNP *P*-values from the three datasets SAF2, GAIN and MGS (a total of 3,307 cases) were meta-analyzed. The lowest combined *P*-value after meta-analysis was 1.1×10^{-3} for the combined sample SAF2 (broad status of SCZ)-GAIN-MGS for SNP rs4772996 (Table 4). Although this SNP does not survive the correction for multiple testing when considering 230 independent tests (corrected α level, $0.05/230 = 2.17 \times 10^{-4}$), it is important to note that direction of association is consistent across all three datasets for the top associated SNPs. Moreover, all top associated SNPs following meta-analysis are located within intron 4 of the gene, in complete LD with the top associated SNPs in the SAF2 dataset (D' = 1), strongly suggesting that the association signal obtained upon meta-analysis points to the same associated region within the *MYO16* gene. The fact that these SNPs do not reach significance after correction for multiple testing likely reflects the presence of heterogeneity across datasets (Table 4). Sample heterogeneity also likely explains change of ranking among top SNPs. Specifically, although the top associated SNPs in the SAF2 dataset continue to show nominally significant association in the meta-analysis they are not present among the top-ranking SNPs (Table 4). However, top-ranking SNPs from either dataset are in high LD with each other and likely represent the same association signal.

Table 4 *P*-values for the individual replication samples and *P*-values following meta-analysis (Meta-*P*).

SNP	BP	Allele 1 (reference)	*P* SAF2 (Na)	*P* SAF2 (Bd)	*P* GAIN	*P* MGS	META-*P*	Direction of the association
rs4772996	109429712	A	0.034	0.097	0.013	0.125	1.16×10^{-3}	+++
rs878536	109431493	T	0.034	0.097	0.013	0.125	1.16×10^{-3}	---
rs1019863	109438967	A	0.008	0.016	0.003	0.673	1.18×10^{-3}	+++
rs1022801	109439661	A	0.008	0.016	0.003	0.673	1.18×10^{-3}	+++
rs12862455	109433076	T	0.034	0.097	0.015	0.135	1.36×10^{-3}	---
rs17482465	109422400	T	0.013	0.084	0.007	0.209	1.45×10^{-3}	---
rs9634572	109426194	A	0.063	0.103	0.040	0.061	1.47×10^{-3}	+++
rs1118797	109433982	T	0.034	0.097	0.023	0.139	1.67×10^{-3}	+++
rs732974	109415758	A	0.051	0.004	0.021	0.781	1.78×10^{-3}	---
rs12857877	109424041	A	0.170	0.026	0.017	0.103	1.60×10^{-3}	+++

We further compared the results obtained in our SAF2 dataset with recently available results from the Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium (Ripke *et al*, 2011) The PGC study is a meta-analysis that combines various datasets, including GAIN and MGS. Therefore, meta-analysis of our dataset and the PGC dataset is not intended as a replication, but as a test of our hypothesis using a more extensive set of data. We extracted results for SNPs mapping within *MYO16* and performed a meta-analysis following the same procedure as with the SAF2-GAIN-MGS datasets. We meta-analyzed 248 SNPs overlapping across datasets, for a total of 22,640 individuals. The top associated SNP following meta-analysis is rs9284246 (Supplementary Table 3) located within intron 2 of the *MYO16* gene (109,327,788 bp). This finding further points to the region across introns 2 to 6 as the most likely region to harbor common variants implicated in schizophrenia.

Expression analysis

In seeking convergent supporting evidence we also tested the expression levels of the *MYO16* gene in brains of patients with schizophrenia. Our analysis of the SMRI Array Collection using qRT-PCR showed that mean levels of expression of MYO16 were significantly higher in the frontal cortex of schizophrenia patients as compared to controls [$F(1, 66) = 4.2$; $P = 0.044$]. The significance holds when we incorporate either sex and age at death [$F(3,64) = 3.008$; $P = 0.037$] or brain pH and post-mortem interval [$F(3,64) = 2.778$; $P = 0.048$] as covariates in our analysis. The comparison of the bipolar group to controls did not result in a significant difference, although mean levels of expression were slightly higher in the bipolar group (see Supplementary Figure 1 for a scatter plot of expression levels). Furthermore, 6 out of 11 expression studies that have profiled the SMRI Array Collection samples using array technology reported increased levels of MYO16 expression in schizophrenia patients versus controls.

We also tested 28 SNPs genotyped in the Stanley Array Collection, located at both ends of our significant SNPs, but none of these SNPs showed association with MYO16 expression levels ($P > 0.05$). It should be noted that SNP rs9583277, as well as most of the significant SNPs in SAF2, were not included in this set since they had not been genotyped in the Stanley Array Collection.

Discussion

This study employed seven patient cohorts and a dense array of SNPs to fine map the prior linkage region at the 13q32-34 locus. We provide evidence suggesting that variants within *MYO16* contribute to the genetic liability to schizophrenia conferred by the 13q32-34 locus. The *MYO16* gene stretches along 611,856 bps on chromosome 13q33. It consists of 35 exons and has several isoforms. All associated SNPs from SAF1, SAF2 and meta-analysis, and one haplotype from SAF2 are located within introns 2 to 6 of the gene. Considering that there was no significant excess of total genotyped SNPs in this region, this finding indicates that the signal related to common risk variation from this gene is likely localized in this region of the gene. It should be noted that incorporating the initial findings in the meta-analyses is necessary due to the small effect sizes of common variants and the need to increase the power of our association study, albeit the potential for introducing biases (Zeggini and Ioannidis, 2009). The effect of the associated SNPs on the function of the *MYO16* gene remains unknown. It should be noted, however, that one of the most significant SNPs in the SAF2 dataset (rs9301323, *P*-value = $1.7x10^{-3}$), which is in strong LD with the top associated SNP in the same dataset (rs9520990), is located within a splice site region in intron 6 (see URLs), and could affect the pattern of splicing of the *MYO16* gene. This position is conserved in the mouse. The Human Splicing Finder (HSF) program (Desmet *et al*, 2009) indicates that the minor allele (G) of rs9301323 disrupts a predicted branch point sequence in intron 6.

Myosin XVI appeared very recently during the evolution of mammals and is unique in both its structure and function (Thompson and Langford, 2002). Earlier evidence suggested that MYO16 is important for neuronal migration and brain development (Patel *et al*, 2001). More recently, MYO16 has been implicated in neuronal phosphoinositide 3-kinase (PI3K) signaling (Yokoyama *et al*, 2011), an extensively studied pathway involved in neuronal function and morphogenesis as well as in a number of neurological and psychiatric disorders, including schizophrenia and autism (Waite and Eickholt, 2010). MYO16 is a member of the Neuronal tyrosine-phosphorylated Adaptor for the PI 3-kinase (NYAP) family of phosphoproteins, which is comprised of NYAP1, NYAP2,

and Myo16/NYAP3. The NYAPs are expressed predominantly in developing neurons and upon stimulation with Contactin5, they are tyrosine phosphorylated by Fyn. Phosphorylated NYAPs interact with PI3K p85 and activate PI3K, Akt, and Rac1. In addition, NYAPs interact with the WAVE1 complex, thus serving as a bridge for a PI3K-WAVE1 interaction, which mediates PI 3-kinase-dependent remodeling of the actin cytoskeleton. Importantly, disruption of the *NYAP* genes in mice affects brain size and neurite elongation (Yokoyama *et al*, 2011). Notably, meta-analysis of the SAF2, GAIN and MGS datasets (a total of 2,956 cases) showed a gene-wise significant association (*P*-value of $1.8x10^{-5}$) with a SNP located within the third intron of the *NYAP2* gene (rs1897227) suggesting that variation within this gene family may be modulating the risk of schizophrenia.

Additional supporting evidence was provided by expression analysis in brain samples (frontal cortex), which revealed a significant increase in the levels of MYO16 expression in schizophrenia patients compared to controls. Finally, convergent supporting evidence could be found in the existing literature. First, according to the SCAN database (see URLs) the top associated *MYO16* SNP rs9583277 is a potential trans-acting eQTL (expression quantitative trait locus) for *MAP3K13* (mitogen-activated protein kinase 13) gene on chromosome 3q27 ($P = 8x10^{-5}$). Given a potential convergence of *MAP3K13* and PI3K pathways (Ambacher *et al*, 2012), regulation in trans of *MAP3K13* may be mediated by altered *MYO16* activity. Interestingly, MAP3K13 can phosphorylate MAP2K7 (mitogen activated kinase protein 7), which has been recently implicated in schizophrenia (Winchester *et al*, 2012). In addition, the 7 top associated SNPs identified by our meta-analysis of the SAF2-GAIN-MGS datasets (Table 4) are reported by the SCAN database to have a trans-acting effect on the expression of *PAG1* (phosphoprotein associated with glycosphingolipid microdomains 1) on chromosome 8q21.23 ($P = 2x10^{-6}$), a gene implicated in brain maturation (Lindquist *et al*, 2011). Notably, we have previously reported a non-synonymous *de novo* mutation within *PAG1* in a schizophrenia proband (Xu *et al*, 2012; Xu *et al*, 2011). Finally, a recent study (Nakayama *et al*, 2002), reported a physical interaction between the gene products of MYO16 and NRXN1, a synaptic neuronal adhesion molecule that connects presynaptic and postsynaptic neurons and has an important role in cognitive process (Sudhof, 2008). Rare and recurrent

deletions disrupting *NRXN1* have been reported in patients with schizophrenia and neurodevelopmental disorders. Furthermore, *MYO16* has been identified as a candidate risk gene in a genome-wide association study of autism where suggestive association signals were reported in two independent discovery cohorts (Wang *et al*, 2009), as well as in GWAS of alcohol response (Joslyn *et al*, 2010) and smoking cessation (Rose *et al*, 2010).

While our results suggest that common variation within *MYO16* may contribute to the genetic liability to schizophrenia we cannot exclude the possibility that common variants within *MYO16* act in combination with or as surrogates of rare alleles with strong effect in the same or different genes to generate the observed linkage signal in the 13q32-34 locus. We started addressing this question using inherited exonic variant data extracted from our recent whole exome sequencing study in 146 Afrikaner and 85 US parent-proband trios afflicted with schizophrenia or schizoaffective disorder (Xu *et al*, 2012). Trios used in the present study and the Xu et al. (2012) study overlap by ~50% (~72% if we considered just the South African sample). None of the *MYO16* variants located in exons 2-6 (Supplementary Table 4) are in LD with associated SNPs, show differential enrichment in cases versus controls or show strong allele transmission distortion in affected families. Also, no homozygous or compound heterozygous carriers were identified. Although further analysis in expanded samples and in linked families is required, these results suggest that the association observed with common variants of the *MYO16* gene is unlikely to be due to rare exonic variants.

Our results establish *MYO16* as a novel candidate gene for schizophrenia. Interpretation of our findings awaits replication in independent datasets.

Funding: This work was supported in part by National Institute of Mental Health (NIMH) grant MH061399 (to M.K.) and the Lieber Center for Schizophrenia Research at Columbia University. L.R.M. was partially supported by a Gray Matters Fellowship and B.X. was partially supported by a NARSAD Young Investigator Award.

Financial Disclosures Statement: The authors have no biomedical financial interests or potential conflicts of interest to declare.

Acknowledgments

We thank all the families who participated in this research. For our Afrikaner cohorts, we thank H. Pretorius and nursing sisters R. van Wyk, C. Botha and H. van den Berg for their assistance with subject recruitment, family history assessments and diagnostic evaluations. We thank Yan Sun for technical assistance with all DNA sample preparations. We also thank the Center for Inherited Disease Research (CIDR), and in particular J. Roberts and KF Doheny, for custom genotyping services under NIH contract number N01-HG-65403. We thank the Rutgers University Cell and DNA Repository for samples and also the Stanley Medical Research Institute for brain samples. The Rutgers samples were collected in three projects that participated in the NIMH Schizophrenia Genetics Initiative. From 1991 to 1997, the Principal Investigators and Co-Investigators were as follows: (i) Harvard University: MT Tsuang, S Faraone and J Pepple; (ii) Washington University: CR Cloninger, T Reich and D Svrakic; and (iii) Columbia University: C Kaufmann, D Malaspina and JH Friedman.

For the GAIN and MGS_nonGAIN datasets, funding was provided through U01s MH79469 and MH79470. Assistance with data cleaning was provided by the National Center for Biotechnology Information. The GAIN and MGS datasets used for the analyses included in this manuscript were obtained from the database of Genotype and Phenotype (dbGaP) found at dbGaP (see URLs) through dbGaP accession numbers phs000021.v2.p1 (GAIN) and phs000167.v1.p1 (nonGAIN). Samples and associated phenotype data for the MGS GWAS study were collected under the following grants: NIMH Schizophrenia Genetics Initiative U01s: MH46276 (CR Cloninger), MH46289 (C Kaufmann), and MH46318 (MT Tsuang); and MGS Part 1 (MGS1) and Part 2 (MGS2)

R01s: MH67257 (NG Buccola), MH59588 (BJ Mowry), MH59571 (PV Gejman), MH59565 (Robert Freedman), MH59587 (F Amin), MH60870 (WF Byerley), MH59566 (DW Black), MH59586 (JM Silverman), MH61675 (DF Levinson), and MH60879 (CR Cloninger). Further details of collection sites, individuals, and institutions may be found in data supplement Table 1 of Sanders et al. (2008; PMID: 18198266) and at the study dbGaP pages.

References

Abecasis GR, Burt RA, Hall D, Bochum S, Doheny KF, Lundy SL*, et al* (2004). Genomewide scan in families with schizophrenia from the founder population of Afrikaners reveals evidence for linkage and uniparental disomy on chromosome 1. *Am J Hum Genet* **74**(3): 403-417.

Ambacher KK, Pitzul KB, Karajgikar M, Hamilton A, Ferguson SS, Cregan SP (2012). The JNK- and AKT/GSK3beta- Signaling Pathways Converge to Regulate Puma Induction and Neuronal Apoptosis Induced by Trophic Factor Deprivation. *PLoS One* **7**(10): e46885.

Bowden DW, An SS, Palmer ND, Brown WM, Norris JM, Haffner SM*, et al* (2010). Molecular basis of a linkage peak: exome sequencing and family-based analysis identify a rare genetic variant in the ADIPOQ gene in the IRAS Family Study. *Hum Mol Genet* **19**(20): 4112-4120.

Cirulli ET, Goldstein DB (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* **11**(6): 415-425.

Cruz E, Whittington C, Krikler SH, Mascarenhas C, Lacerda R, Vieira J*, et al* (2008). A new 500 kb haplotype associated with high CD8+ T-lymphocyte numbers predicts a less severe expression of hereditary hemochromatosis. *BMC Med Genet* **9**: 97.

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* **37**(9): e67.

Di Rienzo A (2006). Population genetics models of common diseases. *Curr Opin Genet Dev* **16**(6): 630-636.

Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB (2010). Rare variants create synthetic genome-wide associations. *PLoS Biol* **8**(1): e1000294.

ISC (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**(7210): 237-241.

Joslyn G, Ravindranathan A, Brush G, Schuckit M, White RL (2010). Human variation in alcohol response is influenced by variation in neuronal signaling genes. *Alcohol Clin Exp Res* **34**(5): 800-812.

Karayiorgou M, Torrington M, Abecasis GR, Pretorius H, Robertson B, Kaliski S*, et al* (2004). Phenotypic characterization and genealogical tracing in an Afrikaner schizophrenia database. *Am J Med Genet B Neuropsychiatr Genet* **124B**(1): 20-28.

Karst S, Cheng R, Schmitt AO, Yang H, de Villena FP, Palmer AA*, et al* (2011). Genetic determinants for intramuscular fat content and water-holding capacity in mice selected for high muscle mass. *Mamm Genome* **22**(9-10): 530-543.

Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME*, et al* (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* **44**(3): 247-250.

Li M, Boehnke M, Abecasis GR (2005). Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. *Am J Hum Genet* **76**(6): 934-949.

Li M, Boehnke M, Abecasis GR (2006). Efficient study designs for test of genetic association using sibship data and unrelated cases and controls. *Am J Hum Genet* **78**(5): 778-792.

Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34**(8): 816-834.

Lin S, Chakravarti A, Cutler DJ (2004). Exhaustive allelic transmission disequilibrium tests as a new approach to genome-wide association studies. *Nat Genet* **36**(11): 1181-1188.

Lindquist S, Karitkina D, Langnaese K, Posevitz-Fejfar A, Schraven B, Xavier R*, et al* (2011). Phosphoprotein associated with glycosphingolipid-enriched microdomains differentially modulates SRC kinase activity in brain maturation. *PLoS One* **6**(9): e23978.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ*, et al* (2009). Finding the missing heritability of complex diseases. *Nature* **461**(7265): 747-753.

Nakayama M, Kikuno R, Ohara O (2002). Protein-protein interactions between large proteins: two-hybrid screening using a functionally classified library composed of long cDNAs. *Genome Res* **12**(11): 1773-1784.

O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V*, et al* (2008). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* **40**(9): 1053-1055.

Patel KG, Liu C, Cameron PL, Cameron RS (2001). Myr 8, a novel unconventional myosin expressed during brain development associates with the protein phosphatase catalytic subunits 1alpha and 1gamma1. *J Neurosci* **21**(20): 7954-7968.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D*, et al* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**(3): 559-575.

Ripke S, O'Dushlaine C, Chambert K, Moran J, Kähler A, Akterin S*, et al* (2011). Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* **43**(10): 969-976.

Rodriguez-Murillo L, Gogos JA, Karayiorgou M (2012). The genetic architecture of schizophrenia: new mutations and emerging paradigms. *Annu Rev Med* **63**: 63-80.

Rose JE, Behm FM, Drgon T, Johnson C, Uhl GR (2010). Personalized smoking cessation: interactions between nicotine dose, dependence and quit-success genotype score. *Mol Med* **16**(7-8): 247-253.

Sanna S, Li B, Mulas A, Sidore C, Kang HM, Jackson AU*, et al* (2011). Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. *PLoS Genet* **7**(7): e1002198.

Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I*, et al* (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* **460**(7256): 753-757.

Shi Y, Li Z, Xu Q, Wang T, Li T, Shen J*, et al* (2011). Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nat Genet* **43**(12): 1224-1227.

Sobin C, Blundell ML, Conry A, Weiller F, Gavigan C, Haiman C*, et al* (2001). Early, non-psychotic deviant behavior in schizophrenia: a possible endophenotypic marker for genetic studies. *Psychiatry Res* **101**(2): 101-113.

Sobin C, Roos JL, Pretorius H, Lundy LS, Karayiorgou M (2003). A comparison study of early non-psychotic deviant behavior in Afrikaner and US patients with schizophrenia or schizoaffective disorder. *Psychiatry Res* **117**(2): 113-125.

Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D*, et al* (2009). Common variants conferring risk of schizophrenia. *Nature* **460**(7256): 744-747.

Suarez BK, Duan J, Sanders AR, Hinrichs AL, Jin CH, Hou C*, et al* (2006). Genomewide linkage scan of 409 European-ancestry and African American families with schizophrenia: suggestive evidence of linkage at 8p23.3-p21.2 and 11p13.1-q14.1 in the combined sample. *Am J Hum Genet* **78**(2): 315-333.

Sudhof TC (2008). Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**(7215): 903-911.

Thompson RF, Langford GM (2002). Myosin superfamily evolutionary history. *Anat Rec* **268**(3): 276-289.

Torrey EF, Webster M, Knable M, Johnston N, Yolken RH (2000). The stanley foundation brain collection and neuropathology consortium. *Schizophr Res* **44**(2): 151-155.

Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, *et al* (2011). Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* **43**(12): 1193-1201.

Waite K, Eickholt BJ (2010). The neurodevelopmental implications of PI3K signaling. *Curr Top Microbiol Immunol* **346**: 245-265.

Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, *et al* (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320**(5875): 539-543.

Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, *et al* (2009). Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459**(7246): 528-533.

Winchester CL, Ohzeki H, Vouyiouklis DA, Thompson R, Penninger JM, Yamagami K, *et al* (2012). Converging evidence that sequence variations in the novel candidate gene MAP2K7 (MKK7) are functionally associated with schizophrenia. *Hum Mol Genet* **21**(22): 4910-4921.

Xu B, Ionita-Laza I, Roos JL, Boone B, Woodrick S, Sun Y, *et al* (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* **44**(12): 1365-1369.

Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, *et al* (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet* **43**(9): 864-868.

Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M (2008). Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* **40**(7): 880-885.

Xu B, Woodroffe A, Rodriguez-Murillo L, Roos JL, van Rensburg EJ, Abecasis GR, *et al* (2009). Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proc Natl Acad Sci U S A* **106**(39): 16746-16751.

Yokoyama K, Tezuka T, Kotani M, Nakazawa T, Hoshina N, Shimoda Y, *et al* (2011). NYAP: a phosphoprotein family that links PI3K to WAVE1 signalling in neurons. *EMBO J* **30**(23): 4739-4754.

Zeggini E, Ioannidis JP (2009). Meta-analysis in genome-wide association studies. *Pharmacogenomics* **10**(2): 191-201.

URLs:

dbGaP: http://www.ncbi.nlm.nih.gov/projects/gap

Plink: http://pngu.mgh.harvard.edu/~purcell/plink/

PedStats: http://www.sph.umich.edu/csg/abecasis/PedStats/

MACH: http://www.sph.umich.edu/csg/abecasis/MACH

LAMP: http://www.sph.umich.edu/csg/abecasis/LAMP

Metal: http://www.sph.umich.edu/csg/abecasis/Metal

UCSC genome browser: http://genome.ucsc.edu/

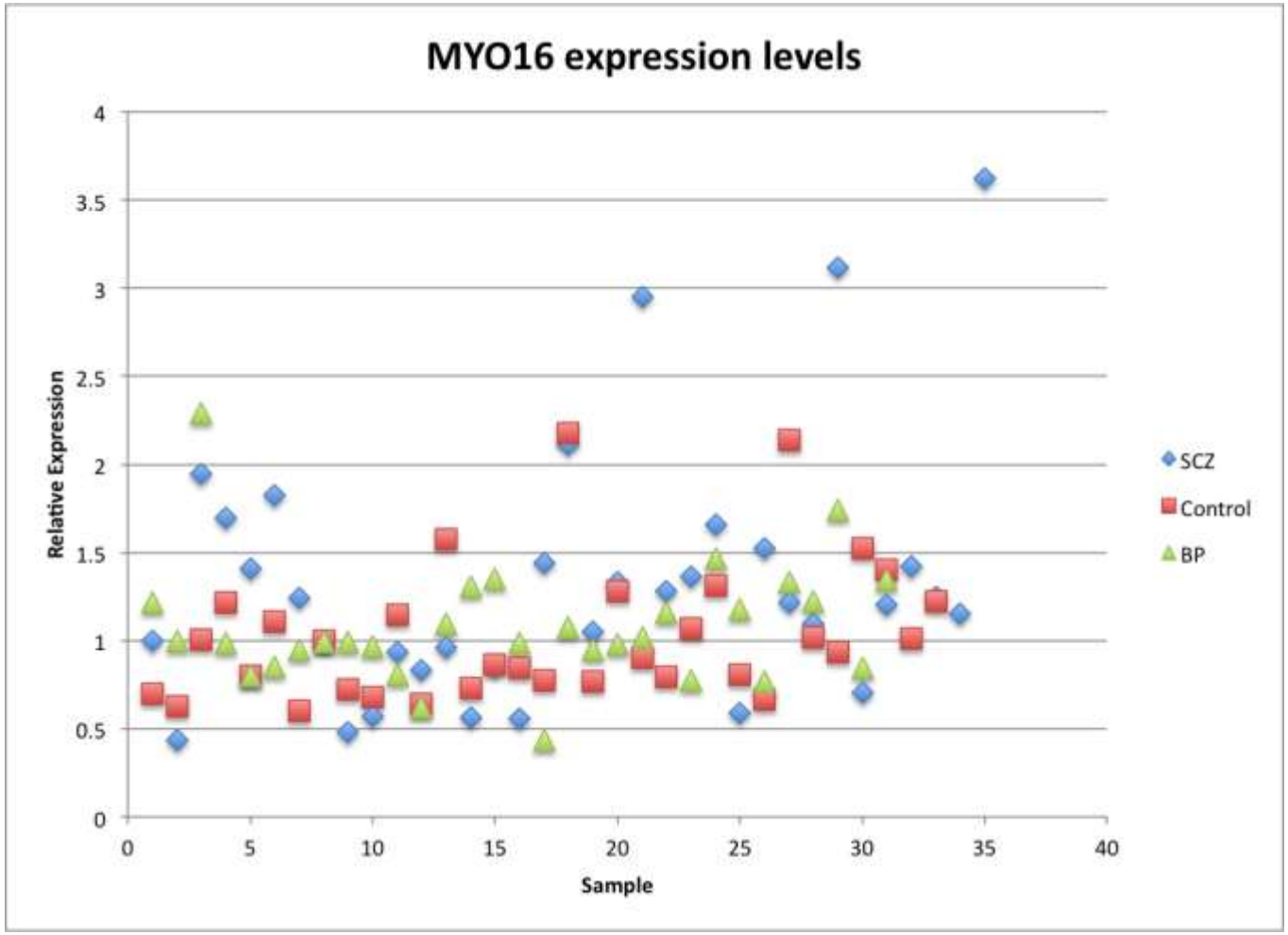
Stanley Medical Research Institute: http://stanleyresearch.org

Scan database: http://scandb.org

1000 genomes project: www.1000genomes.org, accession #: ENST00000357550

SUPPLEMENTARY INFORMATION

Supplementary figure 1 legend. Scatter plot by sample of MYO16 relative expression levels in schizophrenia (SCZ), controls (Control) and bipolar (BP) subjects.



Supplementary Table 1

	SAF1	Rutgers	US	SAF2	GAIN	MGS
# SNPs	723	22	22	Genome-Wide	Genome-Wide	Genome-Wide
# Individuals	1661	1241	630	784	2659	2752
# Families	415	301	207	237	Case/control	Case/control
# Affecteds	741 (474 narrow)	631 (587 narrow)	207 (155 narrow)	266 (232 narrow)	1217	1405
Family history	Yes	Yes	No	Yes	?	?

Unless otherwise specified, # affecteds refer to the broad definition of schizophrenia

Supplementary Table 2

CHR	SNP	POSITION	SAF1 Narrow	SAF1 Broad	alleles	gene	feature
13	rs12860901	113901892	1.60E-04	0.0082	C/T	*RASA3*	intron[NM_007368.2]
13	rs1163830	110911081	2.00E-04	7.60E-04	A/G	NA	NA
13	rs1576166	108986242	4.10E-04	3.30E-04	C/T	NA	NA
13	rs11619453	109564351	0.0012	0.0017	A/C	NA	NA
13	rs10492684	112561683	0.0012	0.0016	C/G	*ATP11A*	intron[NM_032189.3]
13	rs3742193	110078003	0.0013	0.0144	C/T	*FLJ10769*	intron[NM_018210.2]
13	rs7324447	113908097	0.0013	0.0037	A/G	*RASA3*	intron[NM_007368.2]
13	rs9583277	108131750	0.0016	0.0012	A/C	*MYO16*	intron[NM_015011.1]
13	rs7994782	107191892	0.0018	0.0024	A/G	*FAM155A*	
13	rs10492482	108862601	0.0018	0.052	A/G	NA	NA
13	rs7991436	109824441	0.0019	0.0027	A/G	*COL4A2*	intron[NM_001846.2]
13	rs440500	110098800	0.0019	0.0033	C/T	*CARS2*	intron[NM_024537.1]
13	rs942648	110749312	0.002	0.18	C/T	*ARHGEF7*	intron[NM_003899.3]
13	rs9525253	113871032	0.0027	0.084	A/G	*RASA3*	intron[NM_007368.2]
13	rs4325412	107506243	0.0029	0.0144	G/T	NA	NA
13	rs1278760	112579676	0.0034	0.0021	C/T	*ATP11A*	intron[NM_032189.3]
13	rs9515201	109838799	0.0035	0.021	A/C	*COL4A2*	intron[NM_001846.2]
13	rs7334530	113816040	0.0044	0.003	A/G	*RASA3*	intron[NM_007368.2]
13	rs2033539	112511089	0.0058	0.126	C/T	*ATP11A*	intron[NM_032189.3]
13	rs1359480	107413961	0.0059	0.038	A/C	NA	NA
13	rs9555589	108783638	0.0059	0.8	C/G	NA	NA
13	rs7989816	113118794	0.0068	0.025	C/T	NA	NA
13	rs7400121	113557366	0.0068	0.24	C/T	*GAS6*	intron[NM_000820.1]
13	rs7991409	109161160	0.0075	0.019	C/T	NA	NA
13	rs2281973	109928675	0.0075	0.0107	A/G	*COL4A2*	intron[NM_001846.2]
13	rs4771711	110300660	0.0085	0.0032	A/G	NA	NA
13	rs1928454	113831971	0.0089	0.0121	A/G	*RASA3*	intron[NM_007368.2]

13	rs9520422	106832542	0.0106	0.17	C/T	*FAM155A*	
13	rs1927343	109851960	0.0109	0.0133	A/C	*COL4A2*	intron[NM_001846.2]
13	rs277828	108491886	0.011	0.0102	A/C	*MYO16*	intron[NM_015011.1]
13	rs1018601	107008201	0.0112	0.0118	C/T	*FAM155A*	
13	rs1019863	108236968	0.0113	0.023	A/T	*MYO16*	intron[NM_015011.1]
13	rs2940695	107365156	0.0117	0.021	A/G	NA	NA
13	rs7993373	111742897	0.0118	*	A/G	NA	NA
13	rs7319311	109828579	0.0123	0.024	A/G	*COL4A2*	intron[NM_001846.2]
13	rs942336	111681733	0.0124	0.039	C/T	NA	NA
13	rs4773155	109769067	0.0129	0.036	A/C	*COL4A2*	intron[NM_001846.2]
13	rs1328837	108677100	0.0134	0.33	C/T	NA	NA
13	rs1550192	111881724	0.0136	0.0024	C/T	NA	NA
13	rs2391824	109759283	0.0137	0.017	A/G	*COL4A2*	intron[NM_001846.2]
13	rs496313	108615052	0.0141	0.049	A/G	*MYO16*	intron[NM_015011.1]
13	rs723067	110917523	0.0143	0.017	A/G	NA	NA
13	rs1414318	109244941	0.015	0.125	C/T	*LOC728767*	
13	rs1830756	106878224	0.016	0.058	A/C	*FAM155A*	
13	rs2038706	108630288	0.016	0.22	C/T	*MYO16*	intron[NM_015011.1]
13	rs2391753	108902236	0.017	0.0112	C/T	NA	NA
13	rs7330849	109034647	0.017	0.15	C/T	NA	NA
13	rs1539070	104922458	0.018	0.0066	C/G	*DAOA*	intron[NM_172370.3]
13	rs4512966	109880059	0.018	0.036	C/T	*COL4A2*	intron[NM_001846.2]
13	rs277848	108547255	0.019	0.39	A/G	*MYO16*	intron[NM_015011.1]
13	rs9521372	108891045	0.019	0.072	G/T	NA	NA
13	rs7400029	113611527	0.019	0.031	A/G	*FAM70B*	intron[NM_182614.2]
13	rs9555682	109739978	0.02	0.38	A/G	*COL4A1*	intron[NM_001845.4]
13	rs767210	110268039	0.02	0.018	C/T	*LOC100129390*	
13	rs1923735	111632186	0.02	0.035	G/T	NA	NA
13	rs1323666	106852117	0.021	0.087	C/G	*FAM155A*	
13	rs1033869	108700278	0.021	0.021	G/T	NA	NA

13	rs12017058	109621341	0.021	0.051	A/G	*COL4A1*	intron[NM_001845.4]
13	rs9555773	110571109	0.021	0.019	A/G	*ARHGEF7*	intron[NM_145735.2]
13	rs555212	112804541	0.021	0.0031	A/G	NA	NA
13	rs4517649	107045669	0.023	0.98	NA	NA	NA
13	rs1509095	107336856	0.023	0.056	G/T	NA	NA
13	rs1927674	108874741	0.023	0.0076	C/T	NA	NA
13	rs7399860	113567277	0.023	0.114	A/C	*LOC100128430*	
13	rs192505	111923787	0.024	0.039	A/G	NA	NA
13	rs9549575	112560332	0.024	0.3	C/T	*ATP11A*	intron[NM_032189.3]
13	rs3814254	113160288	0.024	0.52	G/T	*DCUN1D2*	
13	rs12021271	108312802	0.025	0.055	G/T	*MYO16*	intron[NM_015011.1]
13	rs7983579	105957447	0.026	0.72	G/T	*EFNB2*	intron[NM_004093.2]
13	rs716504	106859625	0.026	0.88	C/G	*FAM155A*	
13	rs9515076	108966763	0.026	0.0103	C/T	NA	NA
13	rs4773340	110747254	0.026	0.0116	A/G	*ARHGEF7*	intron[NM_003899.3]
13	rs2147686	110665502	0.027	0.04	G/T	*ARHGEF7*	intron[NM_003899.3]
13	rs1859756	110867758	0.027	0.126	A/G	NA	NA
13	rs1325372	107431712	0.028	0.079	C/T	NA	NA
13	rs913746	109814125	0.028	0.052	A/C	*COL4A2*	intron[NM_001846.2]
13	rs475518	111640795	0.028	0.049	A/T	NA	NA
13	rs9577645	111893886	0.028	0.068	A/T	NA	NA
13	rs185792	112048584	0.028	0.26	A/G	NA	NA
13	rs1334586	99433607	0.029	*	A/G	*ZIC2*	intron[NM_007129.2]
13	rs7982576	106844787	0.029	0.28	C/T	*FAM155A*	
13	rs2031837	108930376	0.031	0.056	C/T	NA	NA
13	rs701580	109040392	0.031	0.032	A/G	NA	NA
13	rs2391882	110377552	0.032	0.16	C/G	NA	NA
13	rs6492204	108972142	0.033	0.053	C/T	NA	NA
13	rs953386	109741693	0.033	0.066	A/G	*COL4A1*	intron[NM_001845.4]
13	rs1163838	111001762	0.033	0.075	A/G	NA	NA

13	rs7987644	112322196	0.033	0.0096	A/G	NA	NA
13	rs7984847	107096864	0.034	0.041	C/T	*FAM155A*	
13	rs996969	109040933	0.034	0.21	A/G	NA	NA
13	rs7325670	109122216	0.034	0.028	C/T	NA	NA
13	rs4771685	109966952	0.034	0.136	A/G	NA	NA
13	rs2136267	107332783	0.035	0.022	A/G	NA	NA
13	rs9515008	108648676	0.035	0.29	C/T	*MYO16*	intron[NM_015011.1]
13	rs4771748	110810722	0.035	0.066	C/T	NA	NA
13	rs913745	109683981	0.037	0.028	A/T	*COL4A1*	intron[NM_001845.4]
13	rs2296353	110731099	0.037	0.085	C/G	*ARHGEF7*	intron[NM_003899.3]
13	rs387298	109957629	0.038	0.086	A/G	*COL4A2*	intron[NM_001846.2]
13	rs9521694	109775574	0.039	0.49	A/T	*COL4A2*	intron[NM_001846.2]
13	rs944899	111798962	0.039	0.81	C/T	NA	NA
13	rs4597175	106846695	0.04	0.2	C/T	*FAM155A*	
13	rs2077891	110197898	0.04	0.022	A/T	NA	NA
13	rs195249	108170191	0.041	0.28	A/G	*MYO16*	intron[NM_015011.1]
13	rs1328247	108917530	0.041	0.28	C/T	NA	NA
13	rs157008	108556379	0.042	0.026	A/G	*MYO16*	intron[NM_015011.1]
13	rs7394	110090045	0.042	0.104	C/T	*FLJ10769*	utr-3[NM_018210.2]
13	rs754599	110152789	0.042	0.043	A/C	*CARS2*	intron[NM_024537.1]
13	rs6602897	113501804	0.042	0.128	C/T	*FLJ44054*	
13	rs277796	108435624	0.046	*	C/T	*MYO16*	intron[NM_015011.1]
13	rs4293272	111553204	0.046	0.028	A/G	NA	NA
13	rs280795	111914113	0.046	0.49	A/G	NA	NA
13	rs2270393	113023895	0.046	0.061	C/T	*LAMP1*	
13	rs9559068	107032518	0.048	0.107	A/G	*FAM155A*	
13	rs3906815	110203578	0.048	0.065	G/T	NA	NA
13	rs1536678	112143948	0.049	0.06	A/G	NA	NA
13	rs954580	104949855	0.05	0.0096	C/T	NA	NA
13	rs1935133	107126055	0.05	*	A/T	*FAM155A*	

13	rs942653	110632032	0.05	0.2	C/T	*ARHGEF7*	intron[NM_003899.3]
13	rs9549653	112767835	0.051	0.029	C/T	*MCF2L*	intron[NM_024979.3]
13	rs7328800	107489197	0.052	0.045	C/G	NA	NA
13	rs1473792	107675630	0.052	0.091	C/T	*ABHD13*	intron[NM_032859.2]
13	rs483838	111660805	0.052	0.2	C/T	NA	NA
13	rs920008	109191651	0.053	0.025	C/T	NA	NA
13	rs4107301	107322224	0.054	0.16	A/G	NA	NA
13	rs4773173	109823119	0.054	0.062	A/G	*COL4A2*	intron[NM_001846.2]
13	rs1536621	110006337	0.054	0.049	G/T	*RAB20*	intron[NM_017817.1]
13	rs2391610	107468407	0.055	0.28	C/T	NA	NA
13	rs277854	108564557	0.055	0.023	C/T	*MYO16*	intron[NM_015011.1]
13	rs815176	111906520	0.055	0.53	C/G	NA	NA
13	rs7999900	107795019	0.056	0.021	C/T	NA	NA
13	rs622911	108666181	0.056	0.24	C/T	NA	NA
13	rs11069816	109388612	0.057	0.065	C/T	NA	NA
13	rs1926542	108118919	0.058	0.044	A/G	*MYO16*	intron[NM_015011.1]
13	rs157023	108575332	0.058	0.06	G/T	*MYO16*	intron[NM_015011.1]
13	rs1325395	107495302	0.059	0.08	A/G	NA	NA
13	rs7325927	107273514	0.06	0.146	C/T	*FAM155A*	
13	rs12016920	113474918	0.062	0.26	A/G	NA	NA
13	rs492560	109717977	0.063	*	A/G	*COL4A1*	intron[NM_001845.4]
13	rs2094700	109808269	0.064	0.071	C/T	*COL4A2*	intron[NM_001846.2]
13	rs393192	112507640	0.064	0.032	A/T	*ATP11A*	intron[NM_032189.3]
13	rs3858816	108441465	0.065	0.4	C/T	*MYO16*	intron[NM_015011.1]
13	rs1046518	113252370	0.065	0.42	A/G	*TMCO3*	utr-3[NM_017905.4]
13	rs2068813	108341285	0.066	0.124	C/T	*MYO16*	intron[NM_015011.1]
13	rs188126	111925441	0.066	0.94	A/G	NA	NA
13	rs196141	108333318	0.067	0.058	C/T	*MYO16*	intron[NM_015011.1]
13	rs3910315	109156880	0.067	0.141	C/T	NA	NA
13	rs512610	110787471	0.067	0.087	A/G	*C13orf16*	intron[NM_152324.1]

13	rs6422414	113646243	0.068	0.18	C/T	*FAM70B*	intron[NM_182614.2]
13	rs1328240	108822845	0.069	0.094	A/G	NA	NA
13	rs2993342	112655072	0.07	0.148	G/T	NA	NA
13	rs2011035	108688891	0.071	0.34	A/G	NA	NA
13	rs1771138	107314854	0.072	0.44	C/T	*FAM155A*	
13	rs474128	106941677	0.074	*	C/T	*FAM155A*	
13	rs9549655	112773368	0.074	0.16	A/G	*MCF2L*	intron[NM_024979.3]
13	rs6042	112818069	0.074	0.36	C/T	*F7 F7 F7*	
13	rs4773139	109730194	0.075	0.126	A/G	*COL4A1*	intron[NM_001845.4]
13	rs2391933	110823483	0.076	0.083	A/C	NA	NA
13	rs3742181	110751904	0.077	0.129	A/G	*ARHGEF7*	intron[NM_003899.3]
13	rs1931356	107587302	0.079	0.142	C/T	NA	NA
13	rs2391812	109554665	0.08	0.34	C/T	NA	NA
13	rs4772864	106710934	0.082	0.066	C/T	*FAM155A*	
13	rs1556122	109783608	0.083	0.148	C/T	*COL4A2*	intron[NM_001846.2]
13	rs2479429	109975673	0.083	0.044	A/G	*RAB20*	intron[NM_017817.1]
13	rs1571862	110463849	0.083	0.039	C/T	NA	NA
13	rs9577826	111632380	0.083	0.042	A/T	NA	NA
13	rs1151451	110936930	0.084	0.21	C/T	NA	NA
13	rs3858799	107133874	0.086	0.34	A/G	*FAM155A*	
13	rs1408911	107423256	0.086	0.22	C/T	NA	NA
13	rs10492418	108178727	0.086	0.071	G/T	*MYO16*	intron[NM_015011.1]
13	rs9559900	110325330	0.087	0.093	A/G	NA	NA
13	rs1924349	107905294	0.088	0.44	A/G	NA	NA
13	rs1022963	110296002	0.088	0.068	A/G	NA	NA
13	rs2391891	110430395	0.089	0.17	A/G	NA	NA
13	rs1359476	107404941	0.09	0.51	C/T	NA	NA
13	rs7998446	112658843	0.09	0.079	C/T	NA	NA
13	rs927606	108654322	0.091	*	A/G	*MYO16*	intron[NM_015011.1]
13	rs1023168	109029934	0.091	0.036	A/G	NA	NA

13	rs4771678	109874941	0.091	0.025	C/T	*COL4A2*	intron[NM_001846.2]
13	rs9588436	110881278	0.091	0.062	A/C	NA	NA
13	rs7317784	109753075	0.093	0.084	C/T	*COL4A1*	intron[NM_001845.4]
13	rs1853861	110112905	0.093	0.32	A/G	*CARS2*	intron[NM_024537.1]
13	rs2391887	110420594	0.093	0.33	C/T	NA	NA
13	rs1074866	111810283	0.093	0.063	A/G	NA	NA
13	rs9314891	113891307	0.093	0.034	A/G	*RASA3*	intron[NM_007368.2]
13	rs9634589	109359256	0.094	0.127	C/G	NA	NA
13	rs378953	112026872	0.094	0.63	G/T	NA	NA
13	rs7489746	113637032	0.094	0.18	A/G	*FAM70B*	intron[NM_182614.2]
13	rs9558996	106740433	0.095	0.28	C/T	*FAM155A*	
13	rs7326145	109845351	0.096	0.066	A/C	*COL4A2*	intron[NM_001846.2]
13	rs4907475	112657291	0.097	0.68	A/G	NA	NA
13	rs1887697	101214784	0.098	0.028	C/T	*FGF14*	intron[NM_004115.2]
13	rs3915580	107180494	0.098	0.39	A/G	*FAM155A*	
13	rs8002899	108201934	0.099	0.08	C/G	*MYO16*	intron[NM_015011.1]
13	rs7994403	110523933	0.099	0.143	G/T	NA	NA
13	rs992529	108075875	0.1	0.091	A/T	*MYO16*	intron[NM_015011.1]
13	rs275944	108597821	0.1	0.097	A/T	*MYO16*	intron[NM_015011.1]
13	rs872484	109506369	0.1	0.05	A/G	NA	NA
13	rs912943	110001696	0.1	0.069	G/T	*RAB20*	intron[NM_017817.1]
13	rs912947	109946660	0.101	0.034	C/G	*COL4A2*	intron[NM_001846.2]
13	rs4578513	108177007	0.104	0.059	A/G	*MYO16*	intron[NM_015011.1]
13	rs4238277	110034434	0.105	0.0093	A/G	NA	NA
13	rs1164133	110847931	0.105	0.035	C/T	NA	NA
13	rs7330570	111660345	0.106	0.44	C/T	NA	NA
13	rs9549845	111813664	0.106	0.047	A/G	NA	NA
13	rs1925888	112294946	0.106	0.114	A/G	NA	NA
13	rs3916968	104925530	0.108	0.17	A/G	*DAOA*	intron[NM_172370.3]
13	rs943900	110831429	0.108	0.31	C/T	NA	NA

13	rs9585307	99419777	0.109	0.142	C/G	*ZIC5*	intron[NM_033132.3]
13	rs715738	108990492	0.109	0.062	A/G	NA	NA
13	rs766974	110015423	0.109	0.12	C/T	NA	NA
13	rs9604456	112508272	0.109	0.15	C/G	*ATP11A*	intron[NM_032189.3]
13	rs947361	107123121	0.11	*	C/T	*FAM155A*	
13	rs958378	110542939	0.11	0.049	G/T	NA	NA
13	rs4444189	106823669	0.111	0.34	A/G	*FAM155A*	
13	rs878538	108228609	0.111	0.122	A/T	*MYO16*	intron[NM_015011.1]
13	rs4771635	109013314	0.112	0.19	A/G	NA	NA
13	rs961762	106770873	0.113	0.108	C/T	*FAM155A*	
13	rs7981705	109229892	0.115	0.101	C/T	*IRS2*	intron[NM_003749.2]
13	rs7331116	113103906	0.116	0.141	A/G	NA	NA
13	rs9583484	109764350	0.117	0.071	G/T	*COL4A2*	intron[NM_001846.2]
13	rs3803229	109932781	0.119	0.17	C/T	*COL4A2*	intron[NM_001846.2]
13	rs8000897	107630926	0.121	0.37	C/T	NA	NA
13	rs943767	108828552	0.121	*	A/C	NA	NA
13	rs9521623	109607383	0.123	0.23	A/C	*COL4A1*	intron[NM_001845.4]
13	rs928543	109396965	0.124	0.103	A/G	NA	NA
13	rs8000376	100848840	0.125	0.028	C/T	*NALCN*	intron[NM_052867.2]
13	rs977387	108062451	0.126	0.031	C/T	*MYO16*	intron[NM_015011.1]
13	rs9514997	108609512	0.126	0.34	A/G	*MYO16*	intron[NM_015011.1]
13	rs1163857	110957774	0.126	0.24	A/G	NA	NA
13	rs1411628	110997322	0.126	0.31	A/G	NA	NA
13	rs1926503	108148775	0.127	0.029	C/T	*MYO16*	intron[NM_015011.1]
13	rs958952	109540198	0.127	0.68	C/T	NA	NA
13	rs9520479	106984579	0.128	0.2	C/G	*FAM155A*	
13	rs1224166	107765412	0.128	0.076	C/T	NA	NA
13	rs3923530	108790155	0.128	0.5	G/T	NA	NA
13	rs4773291	110398933	0.128	0.083	A/G	NA	NA
13	rs9520928	107990778	0.129	0.016	G/T	NA	NA

13	rs1341403	104914808	0.13	*	A/C	*DAOA*	
13	rs7992722	112157732	0.131	0.38	A/G	NA	NA
13	rs1886228	107067797	0.132	0.35	A/G	*FAM155A*	
13	rs4771636	109021687	0.132	0.24	C/T	NA	NA
13	rs4773124	109557712	0.132	0.33	G/T	NA	NA
13	rs1886227	107063516	0.133	0.66	C/T	*FAM155A*	
13	rs4771644	109211818	0.134	0.2	A/G	*IRS2*	intron[NM_003749.2]
13	rs7139958	104934733	0.135	0.148	A/T	*DAOA*	intron[NM_172370.3]
13	rs4773278	110371439	0.135	0.43	C/G	NA	NA
13	rs966804	108207339	0.136	0.58	C/T	*MYO16*	intron[NM_015011.1]
13	rs415756	110259218	0.136	0.41	A/G	NA	NA
13	rs6577046	113140550	0.136	0.16	G/T	*ADPRHL1*	intron[NM_199162.1]
13	rs831165	110590297	0.137	0.24	A/G	*ARHGEF7*	intron[NM_145735.2]
13	rs776897	112843672	0.138	0.46	C/T	*F10*	intron[NM_000504.3]
13	rs1323682	106909000	0.139	0.38	C/T	*FAM155A*	
13	rs7326528	108423121	0.14	0.12	C/T	*MYO16*	intron[NM_015011.1]
13	rs7991325	113127674	0.141	0.077	A/C	*ADPRHL1*	intron[NM_199162.1]
13	rs1223978	107573163	0.142	0.17	C/T	NA	NA
13	rs1041466	109042323	0.144	0.23	C/T	NA	NA
13	rs4773092	109233954	0.144	*	A/G	*IRS2 LOC728767 LOC728767*	
13	rs7335928	111053199	0.145	0.076	A/G	NA	NA
13	rs1320517	112636505	0.145	0.22	C/T	NA	NA
13	rs1164132	110853125	0.146	0.42	A/G	NA	NA
13	rs9577503	113013177	0.147	0.114	A/C	*LAMP1*	
13	rs2057504	106895244	0.15	0.89	C/T	*FAM155A*	
13	rs9521184	108580569	0.15	0.42	C/T	*MYO16*	intron[NM_015011.1]
13	rs745324	109263298	0.15	0.45	C/T	NA	NA
13	rs9555646	109364001	0.15	0.089	C/T	NA	NA
13	rs4773106	109448627	0.15	0.68	A/G	NA	NA

13	rs9521666	109695445	0.15	0.38	A/G	*COL4A1*	intron[NM_001845.4]
13	rs867790	110249709	0.15	0.36	C/T	NA	NA
13	rs11619712	111611762	0.15	0.018	A/G	NA	NA
13	rs11618595	111741481	0.15	0.106	C/T	NA	NA
13	rs9549656	112773759	0.15	0.22	A/T	*MCF2L*	intron[NM_024979.3]
13	rs9520447	106902104	0.16	0.27	A/G	*FAM155A*	
13	rs11839232	108496987	0.16	0.038	A/G	*MYO16*	intron[NM_015011.1]
13	rs7324846	108608164	0.16	*	A/G	*MYO16*	intron[NM_015011.1]
13	rs1536675	108861771	0.16	0.106	G/T	NA	NA
13	rs336230	109268342	0.16	0.079	A/G	NA	NA
13	rs927793	109998017	0.16	0.21	A/G	*RAB20*	intron[NM_017817.1]
13	rs12856863	113050384	0.16	0.072	A/G	*GRTP1*	intron[NM_024719.2]
13	rs9604573	113571085	0.16	0.53	C/T	*GAS6*	intron[NM_000820.1]
13	rs9520396	106791742	0.17	0.58	C/T	*FAM155A*	
13	rs9301247	107071857	0.17	0.042	C/T	*FAM155A*	
13	rs719185	107079922	0.17	0.31	C/T	*FAM155A*	
13	rs10492417	108181642	0.17	0.35	G/T	*MYO16*	intron[NM_015011.1]
13	rs592398	108677223	0.17	0.066	C/T	NA	NA
13	rs9301388	108934565	0.17	0.3	C/T	NA	NA
13	rs4773218	110028734	0.17	0.133	C/G	NA	NA
13	rs912937	110048882	0.17	0.091	C/T	NA	NA
13	rs2296354	110668038	0.17	0.89	A/G	*ARHGEF7 ARHGEF7 ARHGEF7*	
13	rs753178	111665322	0.17	0.68	C/T	NA	NA
13	rs7325678	112183509	0.17	0.81	A/G	NA	NA
13	rs383353	112527549	0.17	0.35	A/G	*ATP11A*	intron[NM_032189.3]
13	rs915047	107648692	0.18	0.94	A/G	NA	NA
13	rs726449	107949490	0.18	0.81	A/G	NA	NA
13	rs1926510	108153556	0.18	0.33	G/T	*MYO16*	intron[NM_015011.1]
13	rs9514995	108602549	0.18	0.18	A/C	*MYO16*	intron[NM_015011.1]

13	rs4462453	109049329	0.18	0.5	A/G	NA	NA
13	rs627527	109713136	0.18	0.37	C/T	*COL4A1*	intron[NM_001845.4]
13	rs2391876	110354344	0.18	0.38	A/T	*ANKRD10*	intron[NM_017664.2]
13	rs947170	111039063	0.18	0.31	A/G	NA	NA
13	rs1923740	111594867	0.18	0.19	A/C	NA	NA
13	rs927855	108080073	0.19	0.082	A/C	*MYO16*	intron[NM_015011.1]
13	rs4386002	108219137	0.19	*	NA	NA	NA
13	rs984299	108254079	0.19	0.79	A/G	*MYO16*	intron[NM_015011.1]
13	rs1328250	108856632	0.19	0.16	C/T	NA	NA
13	rs1410424	108976282	0.19	0.84	C/T	NA	NA
13	rs885339	110487236	0.19	0.24	A/G	NA	NA
13	rs7331778	110681068	0.19	0.09	A/G	*ARHGEF7*	intron[NM_003899.3]
13	rs884708	110813104	0.19	0.31	C/T	NA	NA
13	rs1278775	112595200	0.19	0.56	A/G	NA	NA
13	rs474810	112828911	0.19	0.65	C/T	*F10*	intron[NM_000504.3]
13	rs7320143	113046455	0.19	0.12	C/T	*GRTP1*	intron[NM_024719.2]
13	rs9562080	113941879	0.19	0.113	C/G	NA	NA
13	rs9521369	108880688	0.2	0.136	A/G	NA	NA
13	rs2391777	109056244	0.2	0.125	A/G	NA	NA
13	rs1411553	109108212	0.2	0.38	G/T	NA	NA
13	rs2119480	110141889	0.2	0.29	A/G	*CARS2*	intron[NM_024537.1]
13	rs9522149	110625168	0.2	0.5	C/T	*ARHGEF7*	intron[NM_145735.2]
13	rs9588435	110880978	0.2	0.24	A/G	NA	NA
13	rs928194	111902203	0.2	0.18	A/C	NA	NA
13	rs1041385	112354883	0.2	0.49	C/T	*C13orf35*	intron[NM_207440.1]
13	rs9604566	113533961	0.2	0.18	G/T	NA	NA
13	rs2181506	101272738	0.21	*	A/G	*FGF14*	intron[NM_004115.2]
13	rs1341373	107230495	0.21	0.49	C/T	*FAM155A*	
13	rs4772955	107597307	0.21	0.55	A/C	NA	NA
13	rs12428930	107737706	0.21	0.6	A/C	*TNFSF13B*	intron[NM_006573.3]

13	rs1360051	108709716	0.21	0.0097	A/G	NA	NA
13	rs162563	109095564	0.21	0.3	A/G	NA	NA
13	rs9301412	109256477	0.21	0.19	A/G	*LOC728767*	
13	rs7337597	109454492	0.21	0.24	C/T	NA	NA
13	rs529041	109658220	0.21	0.137	A/G	*COL4A1*	intron[NM_001845.4]
13	rs1923742	111595306	0.21	0.055	A/G	NA	NA
13	rs7318319	113882946	0.21	0.38	A/G	*RASA3*	intron[NM_007368.2]
13	rs7317198	106723017	0.22	0.137	C/T	*FAM155A*	
13	rs12561491	107602960	0.22	0.3	A/C	NA	NA
13	rs766106	107787699	0.22	0.07	A/G	NA	NA
13	rs4344597	108000287	0.22	0.22	C/T	NA	NA
13	rs406685	108425587	0.22	0.53	G/T	*MYO16*	intron[NM_015011.1]
13	rs157000	108551465	0.22	*	A/G	*MYO16*	intron[NM_015011.1]
13	rs7981385	109105063	0.22	*	C/T	NA	NA
13	rs9559849	110041737	0.22	0.36	A/G	NA	NA
13	rs9559981	110548126	0.22	0.009	A/T	NA	NA
13	rs11841484	110771589	0.22	0.084	C/T	*C13orf16*	intron[NM_152324.1]
13	rs4907561	112615675	0.22	0.28	A/G	NA	NA
13	rs4907623	113071402	0.22	0.059	A/G	NA	NA
13	rs9514649	106742081	0.23	0.068	C/T	*FAM155A*	
13	rs3858806	107264370	0.23	0.0058	C/T	*FAM155A*	
13	rs9514840	107816155	0.23	0.39	C/G	NA	NA
13	rs196159	108308282	0.23	0.41	A/G	*MYO16*	intron[NM_015011.1]
13	rs7328960	112376717	0.23	0.83	C/T	*C13orf35*	intron[NM_207440.1]
13	rs2993305	112726831	0.23	0.096	C/T	*MCF2L*	intron[NM_024979.3]
13	rs1323672	106870516	0.24	*	C/T	*FAM155A*	
13	rs7321280	107297419	0.24	0.56	A/G	*FAM155A*	
13	rs9520836	107755064	0.24	0.81	A/G	*TNFSF13B*	intron[NM_006573.3]
13	rs2181766	107855110	0.24	0.99	C/T	NA	NA
13	rs12585282	108857114	0.24	0.29	C/T	NA	NA

13	rs6492232	109383171	0.24	0.16	C/T	NA	NA
13	rs7320618	109695293	0.24	0.41	C/T	*COL4A1*	intron[NM_001845.4]
13	rs1933204	112629826	0.24	0.3	C/T	NA	NA
13	rs984300	108253836	0.25	0.67	C/T	*MYO16*	intron[NM_015011.1]
13	rs9521112	108414403	0.25	0.39	A/G	*MYO16*	intron[NM_015011.1]
13	rs9522032	110424256	0.25	0.062	A/G	NA	NA
13	rs7319901	111151026	0.25	0.91	A/G	NA	NA
13	rs1887649	112122229	0.25	0.3	A/C	*C13orf28*	intron[NM_145248.3]
13	rs3024731	112866709	0.25	0.53	A/T	*PROZ*	intron[NM_003891.1]
13	rs9525275	113905078	0.25	0.2	C/T	*RASA3*	intron[NM_007368.2]
13	rs2893357	109461227	0.26	0.25	C/T	NA	NA
13	rs4773116	109511453	0.26	0.088	C/T	NA	NA
13	rs7140030	109862160	0.26	0.33	A/G	*COL4A2*	intron[NM_001846.2]
13	rs1467652	109985751	0.26	0.18	A/G	*RAB20*	intron[NM_017817.1]
13	rs2243928	110779708	0.26	0.21	C/G	*C13orf16*	intron[NM_152324.1]
13	rs876505	110922638	0.26	0.033	A/G	NA	NA
13	rs4343137	111113593	0.26	0.47	C/T	NA	NA
13	rs530085	111649700	0.26	0.16	A/G	NA	NA
13	rs11618091	113524235	0.26	0.27	C/T	*FLJ44054*	
13	rs1408563	106775852	0.27	0.28	C/T	*FAM155A*	
13	rs1581031	107447597	0.27	0.27	A/G	NA	NA
13	rs7321660	108321913	0.27	0.22	C/T	*MYO16*	intron[NM_015011.1]
13	rs9515395	110670274	0.27	0.65	A/C	*ARHGEF7*	intron[NM_003899.3]
13	rs9515446	111015109	0.27	0.56	A/G	NA	NA
13	rs1046793	112587895	0.27	0.19	C/T	*ATP11A*	utr-3[NM_032189.3]
13	rs7989319	113132512	0.27	0.24	A/G	*ADPRHL1*	intron[NM_199162.1]
13	rs7327124	114071257	0.27	0.23	C/T	*UPF3A*	intron[NM_080687.1]
13	rs1431273	108486591	0.28	0.23	C/T	*MYO16*	intron[NM_015011.1]
13	rs3803233	109917147	0.28	0.138	A/T	*COL4A2*	intron[NM_001846.2]
13	rs1541110	110500205	0.28	0.26	A/T	NA	NA

13	rs9559738	109582412	0.29	0.036	C/T	NA	NA
13	rs2318058	111622992	0.29	0.31	C/T	NA	NA
13	rs571564	111773792	0.29	0.88	C/T	*SOX1*	utr-3[NM_005986.2]
13	rs4907674	111900674	0.29	0.59	G/T	NA	NA
13	rs7338610	112381685	0.29	0.23	C/T	*C13orf35*	utr-5[NM_207440.1]
13	rs7329468	113514092	0.29	0.26	A/G	*FLJ44054*	
13	rs9590410	114065802	0.29	0.099	A/T	*UPF3A*	intron[NM_080687.1]
13	rs3916906	100679778	0.3	0.51	C/T	*NALCN*	
13	rs4771568	106697442	0.3	0.32	A/G	*FAM155A*	
13	rs9520569	107200552	0.3	0.4	C/T	*FAM155A*	
13	rs7139848	107259958	0.3	0.43	C/G	*FAM155A*	
13	rs1224096	107701073	0.3	0.78	C/T	NA	NA
13	rs869913	108721817	0.3	0.26	C/T	NA	NA
13	rs167952	109072232	0.3	0.19	A/G	NA	NA
13	rs754730	109151514	0.3	0.99	C/T	NA	NA
13	rs1888845	110520432	0.3	0.094	A/G	NA	NA
13	rs7332266	111768387	0.3	0.26	A/G	*SOX1*	
13	rs3813739	112291289	0.3	0.38	A/C	NA	NA
13	rs9550176	112317128	0.3	0.52	C/T	NA	NA
13	rs2224904	106746123	0.31	0.19	C/T	*FAM155A*	
13	rs816960	107320522	0.31	0.5	A/G	NA	NA
13	rs7985095	107479131	0.31	0.27	C/T	NA	NA
13	rs4772972	107908415	0.31	*	C/T	NA	NA
13	rs1579520	108524944	0.31	0.2	C/T	*MYO16*	intron[NM_015011.1]
13	rs912196	112781336	0.31	0.2	A/G	*MCF2L*	intron[NM_024979.3]
13	rs9994	112932302	0.31	0.99	C/T	*CUL4A*	
13	rs11069734	107800359	0.32	0.095	C/T	NA	NA
13	rs997702	110934168	0.32	0.73	C/T	NA	NA
13	rs4129052	111562394	0.32	0.28	A/C/G/T	NA	NA
13	rs4907715	112155252	0.32	0.59	A/G	NA	NA

13	rs816999	107308417	0.33	0.83	A/T	*FAM155A*	
13	rs231604	107524007	0.33	0.113	C/T	NA	NA
13	rs1019865	108237056	0.33	0.26	G/T	*MYO16*	intron[NM_015011.1]
13	rs4773039	108749019	0.33	0.22	A/G	NA	NA
13	rs2076914	109139017	0.33	0.76	C/T	NA	NA
13	rs496916	109649015	0.33	0.85	C/G	*COL4A1*	intron[NM_001845.4]
13	rs4771754	110907300	0.33	0.37	A/G	NA	NA
13	rs7317997	113781019	0.33	0.45	A/G	*RASA3*	intron[NM_007368.2]
13	rs9521321	108795406	0.34	0.42	C/T	NA	NA
13	rs1550042	109189696	0.34	0.17	A/T	NA	NA
13	rs2148079	109989414	0.34	0.59	A/G	*RAB20*	intron[NM_017817.1]
13	rs1224147	107754759	0.35	0.145	C/T	*TNFSF13B*	intron[NM_006573.3]
13	rs1224174	107777044	0.35	0.73	C/T	NA	NA
13	rs1328238	108866662	0.35	0.106	C/T	NA	NA
13	rs7996888	108978756	0.35	0.121	C/T	NA	NA
13	rs4773301	110492072	0.35	0.43	C/T	NA	NA
13	rs9560063	110886536	0.35	0.87	A/T	NA	NA
13	rs7320365	111848671	0.35	0.39	C/G	NA	NA
13	rs7998746	111948132	0.35	0.31	C/T	NA	NA
13	rs2476773	106931175	0.36	0.93	C/T	*FAM155A*	
13	rs2296845	109898948	0.36	0.116	A/C	*COL4A2*	intron[NM_001846.2]
13	rs1106028	111078373	0.36	0.16	A/G	NA	NA
13	rs2152929	112642865	0.36	0.54	C/G	NA	NA
13	rs14067	113158661	0.36	0.52	C/T	*DCUN1D2*	
13	rs3905075	107261220	0.37	0.65	C/T	*FAM155A*	
13	rs732974	108213759	0.37	0.25	C/T	*MYO16*	intron[NM_015011.1]
13	rs8000882	108940576	0.37	0.19	C/G	NA	NA
13	rs1411766	109050161	0.37	0.81	C/T	NA	NA
13	rs9301405	109144217	0.37	0.68	A/G	NA	NA
13	rs1888247	109164787	0.37	0.26	C/T	NA	NA

13	rs336209	109346846	0.37	0.095	G/T	NA	NA
13	rs1192198	109609579	0.37	0.97	C/T	*COL4A1*	intron[NM_001845.4]
13	rs9515299	110283570	0.37	0.101	C/T	NA	NA
13	rs1028974	110873083	0.37	0.57	C/T	NA	NA
13	rs942339	111734909	0.37	0.91	A/G	NA	NA
13	rs1320526	112608380	0.37	0.54	C/T	NA	NA
13	rs4630437	112998037	0.37	0.72	C/T	*LAMP1*	
13	rs4772985	108080882	0.38	0.84	C/T	*MYO16*	intron[NM_015011.1]
13	rs9284246	108125789	0.38	0.26	C/T	*MYO16*	intron[NM_015011.1]
13	rs189878	108250430	0.38	0.22	A/G	*MYO16*	intron[NM_015011.1]
13	rs9515341	110450160	0.38	0.17	A/C	NA	NA
13	rs1025662	111862700	0.38	0.85	C/T	NA	NA
13	rs4468469	106816536	0.39	*	C/T	*FAM155A*	
13	rs9301356	108644425	0.39	0.91	C/T	*MYO16*	intron[NM_015011.1]
13	rs3098728	110208803	0.39	0.45	C/T	NA	NA
13	rs9324305	112180463	0.39	*	C/T	NA	NA
13	rs12871648	113018663	0.39	0.59	A/C	*LAMP1*	intron[NM_005561.3]
13	rs9559125	107108857	0.4	0.121	A/T	*FAM155A*	
13	rs9555639	109296335	0.4	0.67	A/G	NA	NA
13	rs9555673	109587487	0.4	0.95	A/T	NA	NA
13	rs1018643	109894926	0.4	0.127	C/T	*COL4A2*	intron[NM_001846.2]
13	rs7320173	112207188	0.4	0.53	A/G	*TUBGCP3*	intron[NM_006322.4]
13	rs9514828	107719374	0.41	0.97	C/T	*TNFSF13B*	
13	rs954335	109516892	0.41	0.98	C/T	NA	NA
13	rs11616523	110234671	0.41	0.47	A/G	NA	NA
13	rs2391898	110454440	0.41	0.53	A/G	NA	NA
13	rs7323757	110631622	0.41	0.87	G/T	*ARHGEF7*	intron[NM_003899.3]
13	rs1163852	110962872	0.41	0.57	C/T	NA	NA
13	rs7989477	111137392	0.41	0.63	A/G	NA	NA
13	rs7995181	112014794	0.41	0.66	A/C	NA	NA

13	rs7323586	112526655	0.41	0.86	C/G	*ATP11A*	intron[NM_032189.3]
13	rs4907617	113061932	0.41	0.3	A/G	*GRTP1*	intron[NM_024719.2]
13	rs157024	108575504	0.42	0.74	A/G	*MYO16*	
13	rs7323041	109911609	0.42	0.06	G/T	*COL4A2*	intron[NM_001846.2]
13	rs2025905	109982257	0.42	0.44	C/G	*RAB20*	intron[NM_017817.1]
13	rs4771726	110518979	0.42	0.096	A/G	NA	NA
13	rs421595	112494223	0.42	0.52	C/T	*ATP11A*	intron[NM_032189.3]
13	rs1890202	112603051	0.42	0.32	A/G	NA	NA
13	rs7991547	99417463	0.43	0.22	A/G	*ZIC5*	intron[NM_033132.3]
13	rs11069721	107625853	0.43	0.114	C/T	NA	NA
13	rs1924322	107859630	0.43	0.44	A/G	NA	NA
13	rs1924345	107987033	0.43	0.65	C/T	NA	NA
13	rs4771591	108034018	0.43	0.072	C/T	NA	NA
13	rs7139390	109433529	0.43	*	A/G	NA	NA
13	rs680484	109679086	0.43	0.133	A/C	*COL4A1*	intron[NM_001845.4]
13	rs2183850	110513987	0.43	0.64	A/G	NA	NA
13	rs1183184	110839550	0.43	0.42	A/G	NA	NA
13	rs1933199	112630131	0.43	0.64	C/T	NA	NA
13	rs2993282	112681797	0.43	0.21	A/G	NA	NA
13	rs9577556	113164163	0.43	0.83	A/G	*DCUN1D2*	
13	rs9557751	101341004	0.44	0.018	A/G	*FGF14*	intron[NM_004115.2]
13	rs10508190	107242528	0.44	*	C/T	*FAM155A*	
13	rs9514827	107717404	0.44	0.83	C/T	NA	NA
13	rs8181791	107732046	0.44	0.51	A/G	*TNFSF13B*	intron[NM_006573.3]
13	rs1924326	107848056	0.44	0.71	C/T	NA	NA
13	rs630943	109675047	0.44	0.32	C/T	*COL4A1*	intron[NM_001845.4]
13	rs4447275	109780420	0.44	0.84	A/G	*COL4A2*	intron[NM_001846.2]
13	rs9549433	111571966	0.44	0.49	A/C	NA	NA
13	rs7995838	112309830	0.44	0.78	A/G	NA	NA
13	rs2183246	113919290	0.44	0.92	C/T	NA	NA

13	rs980044	106702668	0.45	*	C/T	*FAM155A*	
13	rs1931355	107579727	0.45	0.35	A/C	NA	NA
13	rs4284503	109922256	0.45	0.65	A/G	*COL4A2*	intron[NM_001846.2]
13	rs328825	110278453	0.45	0.24	C/T	NA	NA
13	rs1183680	110855749	0.45	0.75	A/G	NA	NA
13	rs744134	106922322	0.46	0.86	A/G	*FAM155A*	
13	rs2146952	106952456	0.46	*	A/G	*FAM155A*	
13	rs868284	107652214	0.46	0.75	C/T	NA	NA
13	rs10492412	108298004	0.46	0.52	G/T	*MYO16*	intron[NM_015011.1]
13	rs1034104	109203128	0.46	0.21	C/G	*IRS2*	
13	rs418543	109977797	0.46	0.74	C/T	*RAB20*	intron[NM_017817.1]
13	rs1007260	112056109	0.46	0.82	C/G	NA	NA
13	rs1001843	108477359	0.47	0.48	A/G	*MYO16*	intron[NM_015011.1]
13	rs1952109	112304134	0.47	0.38	A/G	NA	NA
13	rs12867405	113331064	0.47	0.82	A/G	*TFDP1*	intron[NM_007111.3]
13	rs9520758	107516512	0.48	0.56	A/G	NA	NA
13	rs4907563	112621373	0.48	0.29	C/T	NA	NA
13	rs6577026	112761672	0.48	0.28	G/T	*MCF2L*	intron[NM_024979.3]
13	rs755993	113053814	0.48	0.71	A/C	*GRTP1*	intron[NM_024719.2]
13	rs990181	106726884	0.49	0.072	A/C	*FAM155A*	
13	rs4238266	108779132	0.49	*	A/T	NA	NA
13	rs2083567	110223844	0.49	0.29	C/T	NA	NA
13	rs9588495	111127395	0.49	0.28	A/G	NA	NA
13	rs1571621	112367707	0.49	0.62	C/T	*C13orf35*	intron[NM_207440.1]
13	rs9549757	113076752	0.49	0.16	A/G	NA	NA
13	rs9577874	113595154	0.49	0.64	C/T	NA	NA
13	rs12584299	107612825	0.5	0.55	A/G	NA	NA
13	rs831150	110571317	0.5	0.99	C/G	*ARHGEF7*	intron[NM_145735.2]
13	rs7337905	112376526	0.5	0.66	A/G	*C13orf35*	intron[NM_207440.1]
13	rs871388	112572551	0.5	0.48	A/G	*ATP11A*	intron[NM_032189.3]

13	rs9577283	113274387	0.5	0.108	NA	NA	NA
13	rs7321414	107805127	0.51	0.85	A/G	NA	NA
13	rs913947	109154732	0.51	0.29	C/T	NA	NA
13	rs3742207	109616599	0.51	0.82	A/C	*COL4A1*	
13	rs2196579	111816835	0.51	0.54	A/G	NA	NA
13	rs4883652	113913667	0.51	0.68	A/G	*RASA3*	intron[NM_007368.2]
13	rs7994151	114056981	0.51	0.7	A/G	NA	NA
13	rs4483719	106828307	0.52	0.42	A/G	*FAM155A*	
13	rs677532	106994363	0.52	0.55	G/T	*FAM155A*	
13	rs1235133	107550437	0.52	0.0118	A/G	NA	NA
13	rs1925391	107555920	0.52	0.97	C/T	NA	NA
13	rs1408725	110703433	0.52	0.86	C/G	*ARHGEF7*	intron[NM_003899.3]
13	rs2257442	113176783	0.52	0.45	C/T	*DCUN1D2*	
13	rs3825491	108459360	0.53	0.43	C/G	*MYO16*	
13	rs7998604	110401364	0.53	0.73	A/C	NA	NA
13	rs9515326	110414704	0.53	0.63	A/G	NA	NA
13	rs2182921	112137851	0.53	0.85	A/C	NA	NA
13	rs282620	112533327	0.53	0.71	C/T	*ATP11A*	intron[NM_032189.3]
13	rs4074317	113765289	0.53	0.33	C/G	*RASA3*	
13	rs7331571	113898451	0.53	0.51	C/T	*RASA3*	intron[NM_007368.2]
13	rs4771627	108726456	0.54	0.32	A/G	NA	NA
13	rs4287436	108738864	0.54	0.6	C/T	NA	NA
13	rs10220229	109803934	0.54	0.32	C/G	*COL4A2*	intron[NM_001846.2]
13	rs9583500	109919621	0.54	0.27	C/T	*COL4A2*	
13	rs523042	111672186	0.54	0.67	A/G	NA	NA
13	rs1926717	110242642	0.55	0.39	C/G	NA	NA
13	rs1555754	110692672	0.55	0.092	A/C	*ARHGEF7*	intron[NM_145735.2]
13	rs942341	111736532	0.55	0.56	C/T	NA	NA
13	rs1888297	112470411	0.55	0.54	A/G	*ATP11A*	intron[NM_032189.3]
13	rs10508195	107270556	0.56	*	C/T	*FAM155A*	

13	rs7322498	107670522	0.56	0.4	C/T	*ABHD13*	intron[NM_032859.2]
13	rs7317745	110289758	0.56	0.24	C/T	NA	NA
13	rs1184475	110891543	0.56	0.106	C/T	NA	NA
13	rs445307	113948457	0.56	0.143	A/G	NA	NA
13	rs367910	108290417	0.57	0.17	C/T	*MYO16*	intron[NM_015011.1]
13	rs1467761	108517439	0.57	0.29	A/G	*MYO16*	intron[NM_015011.1]
13	rs2274545	109943311	0.57	0.95	G/T	*COL4A2*	intron[NM_001846.2]
13	rs4771730	110563847	0.57	0.138	C/T	*ARHGEF7*	
13	rs726455	111699501	0.57	0.17	C/T	NA	NA
13	rs2183443	111918207	0.57	0.41	A/G	NA	NA
13	rs7321084	107971966	0.58	0.2	A/G	NA	NA
13	rs701567	104939996	0.59	0.62	A/G	*DAOA*	intron[NM_172370.3]
13	rs7324250	110966878	0.59	0.67	A/C	NA	NA
13	rs7997328	113554521	0.59	0.98	C/T	*GAS6*	intron[NM_000820.1]
13	rs2296851	109936256	0.6	0.56	C/T	*COL4A2*	intron[NM_001846.2]
13	rs4771699	110059268	0.6	0.73	C/T	NA	NA
13	rs2304767	110138343	0.6	0.61	C/T	*CARS2*	
13	rs914037	111033309	0.6	0.3	C/T	NA	NA
13	rs387934	110275369	0.61	0.82	A/T	NA	NA
13	rs1151449	110931714	0.61	0.29	G/T	NA	NA
13	rs13378888	111867657	0.61	0.75	A/G	NA	NA
13	rs559054	112848623	0.61	0.61	C/T	*F10*	intron[NM_000504.3]
13	rs942335	111751180	0.62	0.42	C/T	NA	NA
13	rs2382725	113927626	0.62	0.43	A/G	NA	NA
13	rs1151403	107656374	0.63	0.92	C/T	*LIG4*	
13	rs2873579	111837270	0.63	0.19	A/G	NA	NA
13	rs957788	107055221	0.64	0.42	A/G	*FAM155A*	
13	rs1924338	107868993	0.64	0.67	C/T	NA	NA
13	rs7983084	108819564	0.64	*	C/T	NA	NA
13	rs9301376	108821974	0.64	*	C/T	NA	NA

13	rs7982209	109531103	0.64	0.77	A/G	NA	NA
13	rs7322849	111907830	0.64	0.23	C/T	NA	NA
13	rs9603917	112332192	0.64	0.72	C/T	NA	NA
13	rs3211770	112841850	0.64	0.28	A/G	*F10*	intron[NM_000504.3]
13	rs473270	110776266	0.65	0.97	A/G	*C13orf16*	intron[NM_152324.1]
13	rs3905069	107206398	0.66	0.72	C/T	*FAM155A*	
13	rs9520974	108162318	0.66	0.36	A/G	*MYO16*	intron[NM_015011.1]
13	rs1544085	108857993	0.66	0.08	A/G	NA	NA
13	rs1106723	111049123	0.66	0.89	A/G	NA	NA
13	rs2146752	112797476	0.66	0.58	C/T	NA	NA
13	rs1151457	110941191	0.67	0.34	A/G	NA	NA
13	rs1925887	112341502	0.67	*	C/T	NA	NA
13	rs4578540	112752333	0.67	0.52	A/G	*MCF2L*	intron[NM_024979.3]
13	rs12429529	106688601	0.68	0.38	C/G	*FAM155A*	
13	rs6492116	107951372	0.68	0.84	C/T	NA	NA
13	rs157014	108560008	0.68	0.52	C/T	*MYO16*	intron[NM_015011.1]
13	rs1024306	109326859	0.68	0.5	G/T	NA	NA
13	rs3211764	112840391	0.68	0.78	C/G	*F10*	intron[NM_000504.3]
13	rs6602908	113561904	0.68	0.43	A/G	*GAS6*	intron[NM_000820.1]
13	rs915017	109427779	0.69	0.98	A/C	NA	NA
13	rs568315	110797463	0.69	0.93	A/G	NA	NA
13	rs1536760	113931779	0.69	*	C/T	NA	NA
13	rs1324668	106784743	0.7	0.5	A/G	*FAM155A*	
13	rs7325027	107608084	0.71	0.83	C/T	NA	NA
13	rs191796	108338223	0.71	0.33	A/G	*MYO16*	intron[NM_015011.1]
13	rs336239	109293324	0.71	0.98	G/T	NA	NA
13	rs3803230	109917397	0.71	0.86	C/G	*COL4A2*	
13	rs9604408	112420942	0.71	0.52	A/G	*ATP11A*	intron[NM_032189.3]
13	rs4907582	112768781	0.71	0.7	C/G	*MCF2L*	intron[NM_024979.3]
13	rs157027	108578747	0.72	*	C/T	*MYO16*	intron[NM_015011.1]

13	rs9521585	109496262	0.72	0.017	G/T	NA	NA
13	rs9549507	112044119	0.72	0.56	A/G	NA	NA
13	rs204218	112079844	0.73	0.57	C/T	*C13orf28*	intron[NM_145248.3]
13	rs1556124	109789190	0.74	0.88	A/G	*COL4A2*	intron[NM_001846.2]
13	rs9514719	107108078	0.75	0.9	C/T	*FAM155A*	
13	rs2391690	108223447	0.75	0.79	A/G	*MYO16*	intron[NM_015011.1]
13	rs648705	109654154	0.75	0.94	A/C	*COL4A1*	intron[NM_001845.4]
13	rs188166	112403499	0.75	1	C/T	*ATP11A*	intron[NM_032189.3]
13	rs2993310	112730789	0.75	0.56	C/G	*MCF2L*	intron[NM_024979.3]
13	rs9562187	114059427	0.75	0.78	G/T	NA	NA
13	rs10508198	107740789	0.76	0.61	C/G	*TNFSF13B*	intron[NM_006573.3]
13	rs2993334	112618996	0.76	0.91	C/T	NA	NA
13	rs2015775	107640156	0.77	0.47	A/C	NA	NA
13	rs419244	109974394	0.77	0.62	A/G	*RAB20*	
13	rs11617870	113289228	0.78	0.23	A/G	*TFDP1*	intron[NM_007111.3]
13	rs10508199	107879693	0.79	0.31	C/T	NA	NA
13	rs1330540	109172131	0.79	0.31	A/G	NA	NA
13	rs626444	109662397	0.79	0.86	C/T	*COL4A1*	intron[NM_001845.4]
13	rs7986656	113157383	0.79	0.48	C/T	*DCUN1D2*	
13	rs7139897	107879625	0.8	0.85	A/G	NA	NA
13	rs9577173	111654211	0.8	0.84	C/T	NA	NA
13	rs3751411	112902561	0.8	0.75	A/G	*PCID2*	intron[NM_018386.1]
13	rs538540	107003052	0.81	0.25	A/C	*FAM155A*	
13	rs2391677	108377845	0.81	0.53	G/T	*MYO16*	intron[NM_015011.1]
13	rs7320567	112172000	0.81	0.51	C/T	NA	NA
13	rs7984269	112633899	0.81	0.63	A/G	NA	NA
13	rs2281968	109952161	0.82	0.62	C/T	*COL4A2*	intron[NM_001846.2]
13	rs4772995	108208934	0.83	0.58	C/T	*MYO16*	intron[NM_015011.1]
13	rs4255644	109887574	0.83	0.8	A/G	*COL4A2*	intron[NM_001846.2]
13	rs831168	110594974	0.83	0.83	A/G	*ARHGEF7*	intron[NM_145735.2]

13	rs4907532	112177246	0.83	0.93	A/G	NA	NA
13	rs1278114	112193789	0.83	0.81	C/T	*TUBGCP3*	intron[NM_006322.4]
13	rs5960	112849738	0.83	0.77	C/T	*F10*	
13	rs2391340	105971754	0.84	*	G/T	*EFNB2*	intron[NM_004093.2]
13	rs7320494	108020975	0.84	0.56	A/G	NA	NA
13	rs1033871	108629670	0.84	0.49	C/T	*MYO16*	intron[NM_015011.1]
13	rs2391883	110386968	0.84	0.95	C/T	NA	NA
13	rs7994242	111048395	0.84	0.66	C/T	NA	NA
13	rs4390461	108757611	0.85	0.96	C/T	NA	NA
13	rs3858807	107267996	0.86	0.93	A/G	*FAM155A*	
13	rs1317507	112679781	0.86	0.66	G/T	NA	NA
13	rs1163654	111044768	0.87	0.9	C/G	NA	NA
13	rs192530	112436304	0.87	0.29	C/T	*ATP11A*	intron[NM_032189.3]
13	rs816976	107368910	0.88	0.76	C/T	NA	NA
13	rs1410421	108905828	0.88	0.79	C/T	NA	NA
13	rs1417782	107103265	0.89	0.066	A/G	*FAM155A*	
13	rs2391643	107883579	0.89	0.56	G/T	NA	NA
13	rs3926864	109306820	0.89	0.97	A/T	NA	NA
13	rs3809340	110603526	0.89	0.99	A/T	*ARHGEF7*	intron[NM_145735.2]
13	rs9555776	110608828	0.89	0.132	A/C	*ARHGEF7*	intron[NM_003899.3]
13	rs7998896	111599869	0.89	0.36	A/G	NA	NA
13	rs12020931	113654541	0.89	0.79	A/C	NA	NA
13	rs9515212	109885564	0.9	0.27	A/G	*COL4A2*	intron[NM_001846.2]
13	rs6577102	112358548	0.9	0.73	G/T	*C13orf35*	intron[NM_207440.1]
13	rs4772974	107926181	0.91	0.57	C/T	NA	NA
13	rs443079	108303141	0.91	0.15	C/T	*MYO16*	intron[NM_015011.1]
13	rs753623	109112770	0.91	0.53	C/G	NA	NA
13	rs7982180	109451667	0.91	0.8	A/G	NA	NA
13	rs750991	112141635	0.91	0.46	A/G	NA	NA
13	rs9520594	107280986	0.92	0.69	A/G	*FAM155A*	

13	rs7332707	107898025	0.92	0.68	C/T	NA	NA
13	rs1547178	110329886	0.92	0.69	G/T	*ANKRD10*	utr-3[NM_017664.2]
13	rs1022876	106785383	0.93	0.61	C/T	*FAM155A*	
13	rs9550221	112525346	0.93	0.95	C/T	*ATP11A*	intron[NM_032189.3]
13	rs1417907	107144507	0.94	0.78	A/G	*FAM155A*	
13	rs4238264	108060380	0.94	0.61	C/T	*MYO16*	intron[NM_015011.1]
13	rs2391713	108656651	0.94	0.32	C/T	*MYO16*	intron[NM_015011.1]
13	rs7322495	109789286	0.94	0.74	C/T	*COL4A2*	intron[NM_001846.2]
13	rs1151430	110901155	0.94	0.99	A/G	NA	NA
13	rs4611350	112065161	0.94	0.71	C/T	NA	NA
13	rs8001806	113111630	0.94	0.79	A/G	NA	NA
13	rs4773042	108801494	0.95	0.2	C/T	NA	NA
13	rs1970331	111876993	0.95	*	A/G	NA	NA
13	rs7985192	112076742	0.95	0.73	A/G	*C13orf28*	
13	rs953260	110971686	0.96	*	C/T	NA	NA
13	rs4907726	112267159	0.96	0.84	A/G	*TUBGCP3*	intron[NM_006322.4]
13	rs7338868	113479759	0.96	0.81	A/C	NA	NA
13	rs4883676	113861908	0.96	0.9	C/T	*RASA3*	intron[NM_007368.2]
13	rs4505177	107875386	0.97	0.93	C/G	NA	NA
13	rs3933329	109337568	0.97	0.2	A/G	NA	NA
13	rs7996853	113488470	0.97	0.8	A/G	*FLJ44054*	
13	rs686746	109643506	0.98	0.85	A/G	*COL4A1*	intron[NM_001845.4]
13	rs1887124	108159193	0.99	0.67	A/G	*MYO16*	intron[NM_015011.1]
13	rs4145072	109697956	0.99	0.2	C/T	*COL4A1*	intron[NM_001845.4]
13	rs9550193	112372325	0.99	0.98	G/T	*C13orf35*	intron[NM_207440.1]
13	rs778294	104940236	*	0.19	A/G	*DAOA*	intron[NM_172370.3]
13	rs4397971	106838608	*	0.81	A/T	*FAM155A*	
13	rs6492073	107185360	*	0.56	C/T	*FAM155A*	
13	rs2211312	107203928	*	*	A/C	*FAM155A*	
13	rs9514806	107482525	*	0.33	A/G	NA	NA

13	rs9559211	107486706	*	0.31	A/C	NA	NA
13	rs1325382	107545226	*	0.34	C/T	NA	NA
13	rs719737	109148350	*	0.68	C/T	NA	NA
13	rs4771645	109225129	*	*	A/C	*IRS2*	intron[NM_003749.2]
13	rs1414320	109333093	*	0.47	C/T	NA	NA
13	rs1335808	109486141	*	0.54	A/T	NA	NA
13	rs1133219	109611710	*	0.87	C/T	*COL4A1*	
13	rs1111859	111092169	*	0.8	A/G	NA	NA

Highlighted are the SNPs that were followed up in Rutgers and US samples

Supplementary Table 3

MarkerName	Allele1	Allele2	Weight	P-value	Direction	P-Het
rs9284246	a	g	8784	1.9E-03	++	8.5E-01
rs7989802	t	c	8784	8.0E-03	+-	8.7E-02
rs1475142	t	c	8784	8.6E-03	+-	1.0E-01
rs732974	a	g	8784	9.0E-03	--	7.5E-02
rs9521010	a	g	8784	1.1E-02	--	2.1E-02
rs7986961	c	g	8784	1.2E-02	-+	1.1E-01
rs7330740	a	t	8784	1.3E-02	-+	1.2E-01
rs9301328	t	c	8784	1.3E-02	-+	1.2E-01
rs927856	a	g	8784	1.4E-02	--	5.2E-02
rs9521011	t	c	8784	1.5E-02	++	1.9E-02
rs9521072	t	c	8784	1.5E-02	-+	1.2E-01
rs932678	a	c	8784	1.6E-02	++	1.6E-02
rs7984431	t	c	8784	1.7E-02	++	1.8E-01
rs7327834	t	c	8784	1.8E-02	++	5.0E-02
rs7328053	a	c	8784	1.9E-02	++	5.0E-02
rs6492149	a	g	8784	2.1E-02	++	4.3E-01
rs9520990	t	c	8784	2.6E-02	--	1.4E-02
rs9587632	a	t	8784	3.4E-02	++	4.0E-02
rs16972918	a	c	8784	3.6E-02	--	9.6E-01
rs1926542	a	g	8784	3.7E-02	++	2.1E-01
rs4291792	t	c	8784	3.8E-02	++	2.3E-01
rs8002348	t	c	8784	4.0E-02	--	2.0E-01
rs9520981	t	c	8784	4.3E-02	--	8.9E-02
rs1933220	t	c	8784	4.4E-02	++	7.1E-02
rs10492420	a	t	8784	4.9E-02	++	3.4E-01

Supplementary Table 4

CHR	BP	Variant ID	REF	ALT	AF	EXON_ID	T_SA	U_SA	OR	P1	T_SA+US	U_SA+US	OR	P2	parent_count	case_count	cntrl_count	Prediction	SNPEFF_FUNCTIONAL_CLASS
13	109318370	rs117770145	G	A	0.006	2	1	1	1.00	1.000	1	1	1.00	1.000	2	1	0		SILENT
13	109379844	.	C	T	0.010	4	2	2	1.00	1.000	2	2	1.00	1.000	4	2	0		SILENT
13	109438084	rs911973	T	A	0.031	5	5	7	0.71	0.564	13	11	1.18	0.683	32	13	1	TOLERATED	MISSENSE
13	109472733	rs41308564	G	A	0.020	8	3	3	1.00	1.000	8	10	0.80	0.637	20	7	2	TOLERATED	MISSENSE
13	109475600	rs61742194	C	T	0.010	4					1	0	NA	0.317	1	1	0		SILENT
13	109496694	.	T	C	0.011	5	1	0	NA	0.317	1	0	NA	0.317	1	1	0		SILENT
13	109496813	rs16973313	T	C	0.118	5	31	33	0.94	0.803	43	53	0.81	0.307	116	44	9	TOLERATED	MISSENSE
13	109518578	rs9559428	T	C	0.122	7	37	22	1.68	0.051	53	39	1.36	0.144	115	53	13		SILENT
13	109550367	rs76952704	G	A	0.018	10	5	6	0.83	0.763	5	7	0.71	0.564	17	5	2	TOLERATED	MISSENSE
13	109550439	.	T	G	0.014	10	1	0	NA	0.317	1	0	NA	0.317	1	1	0	TOLERATED	MISSENSE
13	109562437	rs61743216	A	C	0.012	4	5	1	5.00	0.103	6	6	1.00	1.000	13	6	0	TOLERATED	MISSENSE
13	109613970	rs112439513	C	T	0.010	6					1	0	NA	0.317	1	1	0		SILENT
13	109661359	rs3825491	C	G	0.366	10	67	60	1.12	0.535	100	102	0.98	0.888	307	132	24	TOLERATED	MISSENSE
13	109672243	.	T	C	0.007	11	0	2	0.00	0.157	0	2	0.00	0.157	4	1	0	TOLERATED	MISSENSE
13	109704760	rs2303965	C	T	0.016	13	0	3	0.00	0.083	4	4	1.00	1.000	9	4	1		SILENT
13	109704761	.	A	T	0.012	13	2	1	2.00	0.564	2	1	2.00	0.564	3	2	0	DAMAGING	MISSENSE
13	109707457	.	A	G	0.011	26	0	1	0.00	0.317	2	1	2.00	0.564	3	2	0	TOLERATED	MISSENSE
13	109707827	rs9521141	T	C	0.138	27	31	36	0.86	0.541	53	59	0.90	0.571	139	59	10		SILENT
13	109772767	.	A	T	0.010	29	1	0	NA	0.317	1	0	NA	0.317	1	1	0	DAMAGING	MISSENSE
13	109777503	rs157024	A	G	0.154	30	39	38	1.03	0.909	58	61	0.95	0.783	152	68	12	TOLERATED	MISSENSE
13	109779906	rs3815983	C	T	0.382	19	53	53	1.00	1.000	81	93	0.87	0.363	309	126	23		SILENT
13	109779910	.	G	A	0.009	19	2	1	2.00	0.564	2	1	2.00	0.564	3	2	0	NA	MISSENSE
13	109793009	.	C	T	0.053	32	0	9	0.00	0.003	3	16	0.19	0.003	30	8	2		SILENT
13	109793177	.	G	A	0.026	32	3	5	0.60	0.480	8	10	0.80	0.637	21	8	3		SILENT
13	109793246	rs80260507	G	C	0.084	32	19	20	0.95	0.873	36	38	0.95	0.816	86	38	1		SILENT
13	109793424	rs113648411	G	T	0.087	32	20	16	1.25	0.505	36	30	1.20	0.460	82	42	7	TOLERATED	MISSENSE
13	109793474	rs28705148	T	C	0.998	32	1	0	NA	0.317	1	1	1.00	1.000	527	229	36		SILENT
13	109793635	.	G	A	0.012	32									1	0	1	DAMAGING	MISSENSE
13	109793720	rs7986595	A	C	0.994	32	3	2	1.50	0.655	4	2	2.00	0.414	530	229	36		SILENT
13	109817258	.	C	T	0.014	33	1	0	NA	0.317	1	0	NA	0.317	1	1	0	DAMAGING	MISSENSE
13	109859071	.	G	A	0.007	35	1	1	1.00	1.000	1	1	1.00	1.000	2	1	0	TOLERATED	MISSENSE
13	109859130	.	G	A	0.009	35					1	0	NA	0.317	1	1	0		SILENT

Note: SA case trios sample size=143. Control trios sample size=34. US case trios sample size=88

CHR: Chromosome

BP: Base Pair position

VariantID: rs number when applicable

REF: Reference allele

ALT: Alternative allele

AF: Allele Frequency

EXON_ID: Exon ID by SNPEFF

T_SA: Number of alleles transmitted to the affected proband in South African trios

U_SA: Number of alleles untransmitted to the affected proband in South African trios

P1: P value of the test transmitted/untransmitted alleles in SA

T_SA+US: Number of alleles transmitted to the affected proband in South African and US trios

U_SA+US: Number of alleles untransmitted to the affected proband in South African and US trios

OR: Odds ratios
P2: P value of the test transmitted/untransmitted alleles in SA+US
parent_count: Number of parents with the ALT allele
case_count: Number of cases with the ALT allele
cntrl_count: Number of controls with the ALT allele
Prediction
SNPEFF_FUNCTIONAL_CLASS: Functional class predicted by SNPEFF