Patterns of variation influencing antipsychotic treatment outcomes in South African first-episode schizophrenia patients

Aim: Many antipsychotic pharmacogenetics studies have been performed examining candidate genes or known variation; however, our understanding of the genetic factors involved in antipsychotic pharmacogenetic traits remains limited. **Materials & methods:** A well-characterized cohort of first-episode schizophrenia (FES) patients was used to identify a subset of nonresponders and responders to antipsychotic treatment for exome sequencing (n = 11). The variation observed in the responders and nonresponders was subsequently compared and a prioritization strategy was employed to identify variants for genotyping in the entire FES cohort (n = 103) as well as an additional Xhosa schizophrenia cohort (n = 222). **Results:** Examination of coding variation revealed a potential role for rare loss-of-function variants in treatment response outcomes. One variant, rs11368509, was found to be weakly associated with better treatment outcomes in the FES cohort (p = 0.057) and the Xhosa schizophrenia cohort (p = 0.016). In addition, the majority of the loss-of-function variation that was considered likely to be involved in antipsychotic treatment response was either novel or rare in Asian and European populations. **Conclusion:** This pilot study has highlighted the importance of exome sequencing for antipsychotic pharmacogenomics studies, particularly in African individuals. Furthermore, the results emphasize once again the complexity of antipsychotic pharmacogenomics and the need for future research.

Original submitted 8 July 2013; Revision submitted 24 October 2013

KEYWORDS: antipsychotic treatment ▪ exome ▪ genome ▪ rare variants ▪ schizophrenia ▪ South Africa ▪ UPP2

Britt I Drogemöller[1], Dana JH Niehaus[2], Bonginkosi Chiliza[2], Lize van der Merwe[3,4], Laila Asmal[2], Anil K Malhotra[5], Galen EB Wright[‡1], Robin Emsley[2] & Louise Warnich*[1]

[1]Department of Genetics, Stellenbosch University, Private Bag XI, Matieland, 7602, Western Cape, South Africa
[2]Department of Psychiatry, Stikland Hospital, Stellenbosch University, Western Cape, South Africa
[3]University of the Western Cape, Department of Statistics, Bellville, Western Cape, South Africa
[4]Stellenbosch University, Department of Molecular Biology & Human Genetics, Tygerberg, Western Cape, South Africa
[5]Department of Psychiatry, Zucker Hillside Hospital, North Shore-Long Island Jewish Health System, USA
[‡]Current address: South African National Bioinformatics Institute, University of the Western Cape, South Africa
**Author for correspondence:*
Tel.: +27 21 8085888
Fax: +27 21 8085833
lw@sun.ac.za

This year marks the 10-year anniversary of the release of the complete human genome sequence and many advances in human genetics have been made possible through the availability of this sequence [101]. Nonetheless, there remain a large number of diseases for which substantial genomic and pharmacogenomic information is lacking, of which schizophrenia is one [1,2]. Although it has been reported that both susceptibility to schizophrenia and response to antipsychotic treatment are heritable, the exact mechanisms involved remain poorly understood [3,4]. To aid in genomic research efforts to identify the genetic factors involved in the missing heritability, large consortiums such as the Psychiatric Genome-Wide Association Study (GWAS) Consortium (PGC) [102] have been formed. Unfortunately, even though these consortiums are utilizing thousands of patients and novel loci have been identified [5], there are still gaps in our knowledge regarding the genetics of schizophrenia.

Therefore, with the rapid decrease in sequencing costs [6], research efforts have turned to whole-genome and exome sequencing (WGES) for much needed answers. In contrast to the hundreds of candidate gene studies that have been performed [103], there are currently only four published schizophrenia studies using WGES strategies [7–10] and no WGES studies focusing on antipsychotic treatment response. The results from the four WGES schizophrenia studies have added to the growing evidence that there may be more variants, both *de novo* and rare, present in schizophrenia patients, and that these variants are more likely to have an adverse effect on the resulting protein products. This corroborates earlier findings that schizophrenia patients are more likely to harbor damaging variants in the form of rare copy number variants [11]. In addition to this, it has been reported that rare variants confer a bigger risk for schizophrenia susceptibility than common variants [4]. In contrast to GWAS, WGES is able to detect both common and rare variants, highlighting the importance of utilizing such technologies when investigating schizophrenia and antipsychotic treatment response outcomes.

Although there are no large consortiums or WGES studies that have focused on antipsychotic pharmacogenomics to date, this avenue of schizophrenia research is important.

Antipsychotics are the only known therapeutic agents that are consistently shown to be superior to placebos for the treatment of schizophrenia [12,13]. However, treatment outcomes are plagued by a lack of efficacy and many adverse drug reactions [14,15]. Reports that differences in treatment response are heritable [3] highlight the need to elucidate the genetic factors involved in influencing response to antipsychotic treatment. Unfortunately, past research has focused predominantly on candidate genes [16] and the majority of findings have not been well replicated [3]. Therefore, it seems likely that many genes and variants, some that could still be unknown, may be involved in antipsychotic treatment response outcomes, making WGES an ideal tool for antipsychotic pharmacogenomic research.

This pilot study sequenced the exomes of 11 South African first-episode schizophrenia (FES) patients to examine the patterns of variation potentially contributing to antipsychotic treatment outcomes, focusing predominantly on functional variants. These results should play a role in enhancing our understanding of antipsychotic treatment response in the South African context and guide the design of future antipsychotic pharmacogenomic studies.

Materials & methods

Patient samples

Discovery cohort

Written informed consent was provided by all patients and/or their guardians prior to this study and ethical clearance was obtained from the Committee for Human Research, Faculty of Health Sciences, Stellenbosch University (South Africa). A cohort of 130 South African FES patients was collected over 4 years. Patients between the ages of 16 and 45 years were included in this study provided they received a Diagnostic and Statistical Manual of Mental Diseases, Fourth Edition (DSM-IV) [17] diagnosis of schizophreniform disorder, schizophrenia or schizoaffective disorder, and have during their lifetime been exposed to a maximum of 4 weeks of antipsychotic medication. Subjects were excluded based on current substance abuse, previous long-acting injectable antipsychotic treatment and significant physical illness or mental retardation. From this cohort, 104 patients provided genomic DNA (gDNA) samples and all unrelated patients (n = 103) were included in the subsequent genetic association analyses. All of these patients were outpatients at recruitment and the demographic data for these patients is provided in Table 1.

After a wash-out period of up to 7 days, all patients were treated with flexible doses of flupenthixol decanoate as follows: the starting dose of flupenthixol decanoate was 10 mg every 2 weeks, with 6-weekly increments of 10 mg every 2 weeks permitted, to a maximum of 30 mg every 2 weeks. No other antipsychotics, mood stabilizers or psychostimulants were permitted. However, other psychotropic medications such as benzodiazepines, antidepressants or anticholinergics were used in parallel to the flupenthixol decanoate treatment when required.

Treatment response was measured by means of the Positive and Negative Syndrome Scale (PANSS) for 12 months. Nonresponders to flupenthixol decanoate were taken off the study and changed to another antipsychotic chosen at the clinician's discretion. Patients were monitored every 2 weeks for the first 6 weeks, and every 3 months thereafter. Although the PANSS score was the only scale used to assess treatment refractoriness, psychopathology was also evaluated using the Clinical Global Impression (CGI) scale and the Calgary Depression Scale for Schizophrenia (CDSS). Investigators were medical doctors trained in the use of the assessment instruments and underwent inter-rater reliability testing prior to commencing the study. The inter-rater reliability was >75% for all scales.

A subset of ten South African Mixed Ancestry (also referred to as Coloured in the South African census) FES patients falling at extreme ends of the treatment response phenotype were selected for exome sequencing. These patients consisted of five nonresponders and five responders to treatment. Nonresponse was defined as study discontinuation due to poor response, end point percentage change in PANSS total score <20% or end point PANSS total score >70, provided that patients met the criteria, and had completed at least 3 months of treatment and had not experienced a relapse. The five responder individuals all exhibited a greater than 40% improvement in PANSS total scores during the 12 months of treatment and were matched for age (within 5 years), gender and ethnicity. In addition, to allow for the examination of shared variation between related individuals, a sibling of one of the nonresponders was selected for exome sequencing. This sibling was also a nonresponder to flupenthixol decanoate treatment.

Replication cohort

In addition to the well-characterized FES cohort, an additional Xhosa schizophrenia cohort was available for replication analyses. In short, this

cohort consisted of 222 Xhosa schizophrenia patients who were assessed by means of the Scale for the Assessment of Negative Symptoms (SANS) and the Scale for the Assessment of Positive Symptoms (SAPS) in a cross-sectional manner. More details regarding this cohort are provided by Wright *et al.* [18].

Exome sequencing

gDNA from the 11 individuals was extracted from venous blood using the QIAamp DNA Blood Maxi Kit (Qiagen, Germany), according to the manufacturer's instructions. Thereafter, the gDNA samples were sent to the HudsonAlpha Genomics Service Laboratory (AL, USA) [104] for exome capture and sequencing. Exome capture was performed using the Agilent SureSelect Human All Exon 50 Mb kit (Agilent Technologies, CA, USA) and 50-bp paired-end sequencing was subsequently performed on the Illumina HiSeq™ 2000 (Illumina, CA, USA).

Bioinformatics pipeline for exome sequence analysis

The generated exome data for the 11 FES individuals was analyzed with the use of a bioinformatics pipeline that made use of the following programs: Burrows–Wheeler Alignment Tool [19] for sequence read alignment; SAM tools [20] for sorting and indexing of reads; the Genome Analysis Toolkit [21], including variant quality score recalibration, for variant calling; and SeattleSeq Annotation 134 [105] for variant annotation. Only variants that were assigned a 'PASS' value in the 'FILTER' field of the vcf files that were generated by Genome Analysis Toolkit were included in downstream analyses. All novel SNPs, as determined by SeattleSeq Annotation 134, were submitted to dbSNP [106] and submitted SNP numbers were assigned.

Comparison of the patterns of variation observed in the FES individuals

In order to examine the patterns of variation observed in the unrelated nonresponder and responder FES individuals, the total and average number of coding variants observed in each group was compared. More specifically, the variants were divided into different classes based on the effect that they were predicted to have on protein function. In each case the percentage of novel variants was investigated.

With regards to the nonresponder siblings, all variants that were predicted to abolish the function of the resulting protein product, that is, loss-of-function (LOF) variants (frameshift, splice-site and stop lost/gained variants) were examined. These LOF variants were then prioritized for genotyping in the entire FES and Xhosa schizophrenia cohorts using the following filtering pipeline: exclusion of all variants that were not shared between the nonresponder siblings; exclusion of all variants that were present in the responder individuals; stepwise exclusion of all

Table 1. Patient demographic data.

Patient demographic descriptor	Count, n (%)
Mean age (years)	24 ± 7
Gender	
Male	76 (74)
Female	27 (26)
Ethnicity	
African descent	13 (13)
Mixed ancestry	82 (80)
European descent	8 (8)
Highest level of education	
Elementary school	7 (7)
Secondary school	65 (63)
Matriculation	21 (20)
Tertiary education	8 (8)
Technical college	2 (2)
Marital status	
Single	87 (84)
Married	8 (8)
Divorced	5 (5)
Widowed	2 (2)
Cohabiting	1 (1)
Urban vs rural living	
Urban	101 (98)
Rural	2 (2)
Previous antipsychotic medication	
No	57 (55)
Yes	46 (45)
Family history of psychosis	
None	41 (41)
Schizophrenia	38 (38)
Bipolar	1 (1)
Depression	5 (5)
Substance abuse	7 (7)
Other	9 (9)

variants based on their presence in additional nonresponder individuals. Based on this filtering system, the most promising variant was subsequently genotyped in the entire FES and Xhosa schizophrenia cohorts using a custom TaqMan SNP genotyping assay (Life Technologies™, CA, USA), according to the manufacturer's instructions. Allele frequency data were obtained from the 1000 Genomes Project Browser [107].

Statistical analyses

Logistic regression modeling was used to compare the numbers of each type of variant detected in the sequenced unrelated responder and nonresponder samples, adjusting for the matching factors: age and gender. p-values less than 0.05 were considered statistically significant.

The variant genotyped in the entire FES and Xhosa schizophrenia cohorts was tested for Hardy–Weinberg equilibrium using the Tools For Population Genetic Analysis (TFPGA) Software v1.3 [108]. Logistic regression was used to compare the genotype distribution between the nonresponder and responder FES samples for this variant. With regards to the Xhosa schizophrenia cohort, logistic regression was also used to compare genotype distribution between those individuals with total SAPS equal to 0 and those individuals with total SAPS greater than 0. Linear regression was used to compare the nonzero total SAPS and genotype groups, adjusting for age and gender. All p-values and effect sizes (odds ratio, differences) with confidence intervals are derived from these models.

Results

Exome sequencing was performed successfully for all 11 FES individuals and the sequence coverage obtained for the individual exomes is displayed in Table 2. A total of 56,346 coding variants were identified in the ten FES individuals, of which 5557 were novel (9.86%). Comparison of the variation observed in the nonresponder and responder FES individuals did not reveal any significant differences, although slightly more coding variants were observed in the nonresponder individuals with reference to both the total number of variants (42,678 vs 42,333 coding variants) and the average number of variants per individual (19,362 vs 19,253 coding variants; Table 3).

Examination of the novel variants revealed that these variants were far less likely to be shared between the responders and nonresponders when compared with known variants, such that 55.63% of known coding variants were shared between the two groups and only 7.43% of novel variants were shared (Figure 1). This was particularly the case for the stop gained/lost variants, which are likely to have a large impact on the function of the protein. None of the stop gained/lost variants that were detected in both nonresponders and responders were novel. However, a large percentage of the stop gained/lost variants occurring in only the nonresponders or the responders were novel (35.71 and 31.25%, respectively). Furthermore, the percentage of novel variants was highly correlated to the predicted effect that the variant has on the protein product, such that LOF variants were more likely to be novel (Figure 2). Examination of the variants that were not shared between the nonresponders and responders revealed that a large percentage of these variants were present in only one individual, such that 98.00% of the novel variants and 82.17% of the known variants that were unique to either of the groups were present in only one individual.

When the LOF variants present in the nonresponder siblings were examined, it was observed that there were a total of 211 LOF variants that were common to both siblings. However, after removing all variants that were also present in the responder individuals, only 22 LOF variants remained (Table 4), all of which were assigned Phred-scaled quality scores above 206. The 21 genes in which these variants occurred were then submitted to Reactome [22,23] to determine if any pathways were enriched. Sixteen of the genes were not assigned a pathway, and although none of the detected pathways were significantly enriched, three genes were assigned to the metabolism pathway (Table 4). When examining all 22 LOF variants, it was observed that nine of these variants were novel, as there were no population frequency data available on the 1000 Genomes Project Browser for two variants, and one variant (rs11368509 in *UPP2*) has only been detected in the Khoisan population to date [24]. Eight of the ten remaining SNPs have allele frequencies of less than 5% in the Asian and European populations genotyped by the 1000 Genomes Project (Table 4 & Figure 3).

When applying the filtering pipeline, it was observed that seven variants were present in at least one of the additional unrelated nonresponder individuals, three variants were present in at least two additional nonresponders and one variant, rs11368509 in *UPP2*, was present in three additional nonresponders. None of the LOF variants were present in all the nonresponders and absent from the responders. Nonetheless, the most promising variant, rs11368509, was genotyped in both the FES and Xhosa schizophrenia cohorts and

Table 2. Summary of sequence coverage obtained for the targeted exomes of the 11 sequenced samples.

Sample no.	Median read depth	1x coverage (%)	4× coverage (%)	8× coverage (%)	20× coverage (%)	30× coverage (%)
1	72×	99.11	97.41	95.36	89.53	83.93
2	67×	98.94	97.14	94.88	88.42	82.00
3	82×	99.14	97.81	96.20	91.53	86.93
4	63×	99.09	97.23	94.87	87.80	80.63
5	81×	99.20	97.69	95.92	90.96	86.24
6	66×	99.21	97.49	95.31	88.68	81.99
7	82×	99.22	97.78	96.00	90.79	85.95
8	76×	99.55	98.36	96.58	90.94	85.83
9	67×	99.20	97.64	95.67	89.55	83.13
10	61×	98.91	96.91	94.45	87.13	79.63
11	63×	98.78	96.85	94.50	87.70	80.74

was in Hardy–Weinberg equilibrium in both of these cohorts. With regards to the FES cohort, as all the individuals who were homozygous for the deletion allele of rs11368509 were present in the responder individuals, they were grouped with the heterozygous individuals for the statistical analysis – resulting in a dominant inheritance model. The difference in dominant genotype distributions between the nonresponder and responder individuals trended towards significance (between 0.05 and 0.10) after adjusting for age, gender and ancestry ($OR_{deletion\text{-}allele}$ = 5.17; 95% CI: 0.95–38.45; p = 0.057; FIGURE 4). In the Xhosa schizophrenia cohort, the difference in rs11368509 genotype distributions between those with zero and those with nonzero total SAPS was not significant (p = 0.234). However, considering only the nonzero SAPS scores, those who were homozygous for the deletion allele had scores that were 2.42 (95% CI: -0.01 to 4.84; p = 0.051) higher than those who were homozygous for the insertion allele of rs11368509, and also 3.04 (95% CI: 0.58–5.50; p = 0.016) higher than those individuals who were heterozygous for the variant (FIGURE 5).

Discussion

To our knowledge, this is the first published study utilizing exome sequencing to examine the pharmacogenomics of antipsychotic response. Furthermore, to date there is only one publically available non-European South African genome sequence, namely that of Archbishop Desmond Tutu [24]. Thus, this study provides important data regarding the variation present in South African schizophrenia patients. This is of significance as African populations are underrepresented in both genomic and psychiatric research [25,26].

Table 3. Total and average number of variants observed in the five nonresponder and responder individuals.

Class of variation	Nonresponders (n = 5)		Responders (n = 5)		p-value
	Total	*Average*	*Total*	*Average*	
Synonymous	22,805	10,509	22,549	10,393	0.569
Missense†	18,917	8337	18,815	8336	0.889
Frameshift‡	652	354	660	363	0.064
Splice site‡	204	95	211	95	0.760
Stop gained/lost‡	171	66	165	66	0.076
Total coding	42,749	19,361	42,400	19,253	

p-values are for the difference between responders and nonresponders, adjusting for the matching variables age and gender.
†*Changes the protein product.*
‡*Abolishes the function of the protein product.*

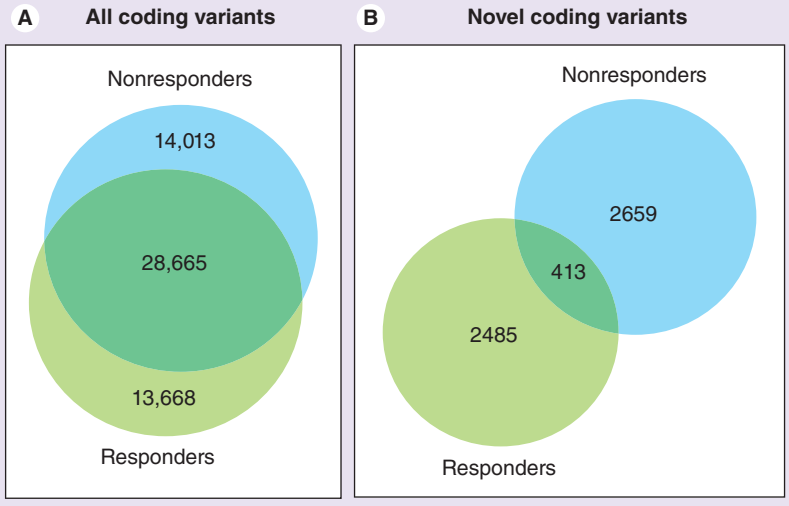

Figure 1. Number of variants shared between the nonresponder and responder individuals. (A) The majority of the coding variants that were observed in the ten first-episode schizoprenia individuals were shared by the nonresponders and responders. There were, however, slightly more coding variants observed in the nonresponder individuals. **(B)** A small portion of the novel coding variants were shared by the two groups, with the majority observed in only the nonresponder or responder groups.

When comparing the data generated by this project to the 'healthy' high coverage trio genomes sequenced by the 1000 Genomes Project [23], it was observed that the number of coding variants observed in the FES individuals fell within the individual range reported by the 1000 Genomes Project. Comparison of the nonresponders and responders revealed that there were slightly more variants present in the nonresponder individuals, although these differences were not significant (Table 3). Of greater significance for future studies was the identification of LOF variants that were shared between the sibling nonresponders but were not present in the responder individuals. These analyses identified a total of 22 variants in 21 genes, of which none were present in all of the nonresponder individuals. The inability of this study to identify a single variant that could completely explain nonresponse, even with the use of well-characterized patients, provides further evidence that antipsychotic treatment response is a complex trait.

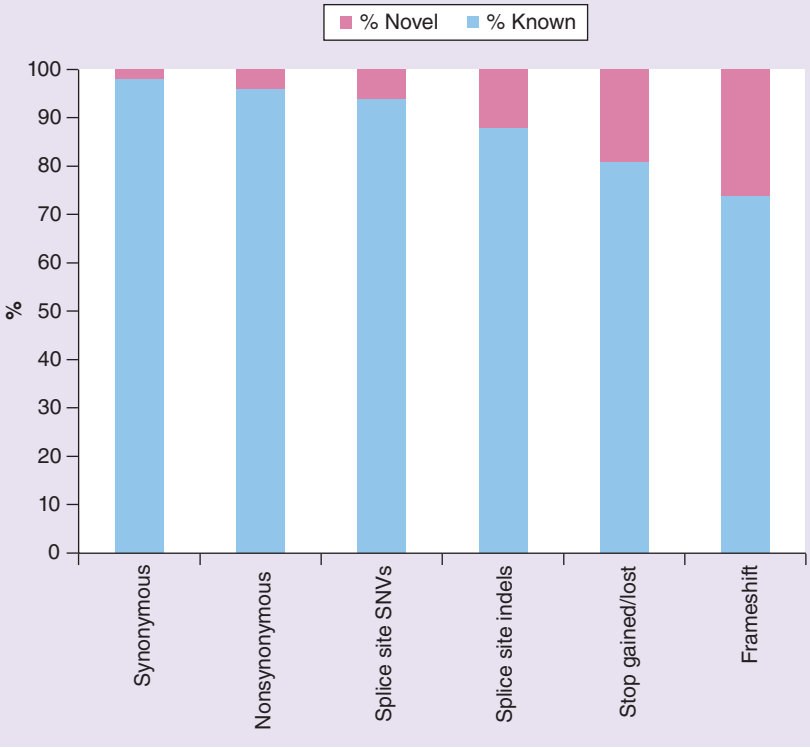

Figure 2. Average percentage novel variation per individual observed for each class of variant.
SNV: Single nucleotide variant.

This study was however able to identify one variant, rs11368509, that was present in all but one of the nonresponder individuals and absent from all the responder individuals (Figure 3). This frameshift variant, occurring in *UPP2*, was genotyped in the entire FES cohort as well as in a further Xhosa schizophrenia cohort. Statistical analyses revealed that this variant was weakly associated with treatment response outcomes in both the FES and Xhosa schizophrenia cohorts. Investigation of the function of UPP2 revealed that this enzyme plays a role in the metabolism of the anticancer drug fluorouracil [27], highlighting the relevance that this gene may have for pharmacogenetic applications. In addition, UPP2 is responsible for metabolizing uridine to uracil and ribose-1-phosphate [28], therefore reducing the amount of uridine available. As uridine has been shown to enhance the antagonism of the dopamine system by the antipsychotic haloperidol [29], a decrease in uridine metabolism may enhance the action of antipsychotic agents. This would explain why individuals who are heterozygous (in the case of the FES cohort) or homozygous (in the case of the Xhosa schizophrenia cohort) for the nonfunctional allele (i.e., the insertion allele of rs11368509) appear to respond better to antipsychotic treatment. Interestingly, this association was in the opposite direction to what was initially observed in the exome data, highlighting the importance of validating the initial findings from the exome data in larger cohorts. Furthermore, as this variant appears to show different patterns of inheritance in the two schizophrenia cohorts, these results should be interpreted with caution and further research into the role that this variant may play in treatment response is required.

Table 4. Descriptive data for the 22 loss-of-function variants present in the nonresponder siblings, but absent from the responder individuals.

Variant	Gene	LOF variant	AFR	ASN	EUR	Novel	Reactome pathway
rs112899189	*CLLU1OS*	Frameshift	**0.03**	**0.00**	**0.01**	No	NA
rs10666583	*GRIN3B*	Frameshift	0.07	0.08	0.28	No	NA
rs73439094	*C6orf52*	Splice-5	0.26	**0.03**	**0.00**	No	NA
rs8065203	*CYTH1*	Splice-3	**0.03**	**0.00**	**0.00**	No	NA
rs57118523	*HK1*	Splice-3	0.07	**0.00**	**0.00**	No	Metabolism
rs149764161	*MYL5*	Splice-3	**0.01**	**0.00**	**0.00**	No	NA
rs17104991	*SLC25A21*	Splice-5	0.12	**0.00**	**0.00**	No	Metabolism, transmembrane transport
rs74141230	*TRIM17*	Splice-3	0.15	**0.01**	**0.00**	No	NA
rs28759013	*TXNDC16*	Splice-3	0.11	**0.00**	**0.00**	No	NA
rs80220955	*OR5AC2*	Stop-lost	0.08	**0.01**	0.23	No	Signal transduction
ss678319342	*C6*	Frameshift	NA	NA	NA	Yes	Immune system
ss678371274	*DNMBP*	Frameshift	NA	NA	NA	Yes	NA
ss678329213	*HIST1H2BM*	Frameshift	NA	NA	NA	Yes	NA
ss678372253	*NRAP*	Frameshift	NA	NA	NA	Yes	NA
ss678431152	*SLC25A41*	Frameshift	NA	NA	NA	Yes	NA
ss678374868	*STK33*	Frameshift	NA	NA	NA	Yes	NA
ss678374869	*STK33*	Frameshift	NA	NA	NA	Yes	NA
ss678424277	*TMEM235*	Splice-3	NA	NA	NA	Yes	NA
ss678333610	*ZBTB24*	Splice-3	NA	NA	NA	Yes	NA
rs71710115	*HOMEZ*	Frameshift	NA	NA	NA	No	NA
rs35744335	*PKD1L3*	Frameshift	NA	NA	NA	No	NA
rs11368509	*UPP2*	Frameshift	NA	NA	NA	No	Metabolism

Minor allele frequencies <0.05 are shown in bold.
AFR: African 1000 genomes population; ASN: Asian 1000 genomes population; EUR: European 1000 genomes population; LOF: Loss of function; NA: Not applicable.

The results from the analyses of rs11368509 emphasize the likelihood that several different genes and variants may be involved in treatment response outcomes. Therefore, the functions of all 21 genes that were affected by the 22 prioritized LOF variants were also examined. To do this, a literature search was performed in PubMed [109] utilizing each of the gene names and the term 'antipsychotic'. The results from this search revealed that only one of the genes, *GRIN3B*, had a previous connection to antipsychotic response [30]. To further assess the functions of these genes, a pathway analysis was performed in Reactome, which revealed that four genes were involved in processes that may be relevant to the mechanism of action of drugs (i.e., metabolism, signal transduction and transport of small molecules; Table 4). These pathway analyses were, however, unable to assign pathways to 16 of the genes. The lack of information pertaining to the function of these genes as determined by both the pathway analyses and NCBI's gene resource [110] highlights a need for future studies to examine the function of these genes.

With reference to the specific variants, of the 22 prioritized LOF variants, 20 were novel or had minor allele frequencies <0.05 in the 1000 Genomes Project Asian and European populations. Although, to our knowledge, none of the 22 variants have been reported in past antipsychotic research, all but two occur at very low frequencies in the Asian and European populations. These populations have been the focus of the majority of such studies, leaving African populations under-represented in antipsychotic pharmacogenomic studies [31]. Thus, variants affecting African individuals may differ to those affecting non-African individuals. Due to the fact that this study was a pilot study, we investigated only the most promising variant identified from the exome sequence analyses. However, these exome data have detected additional variants that may be of

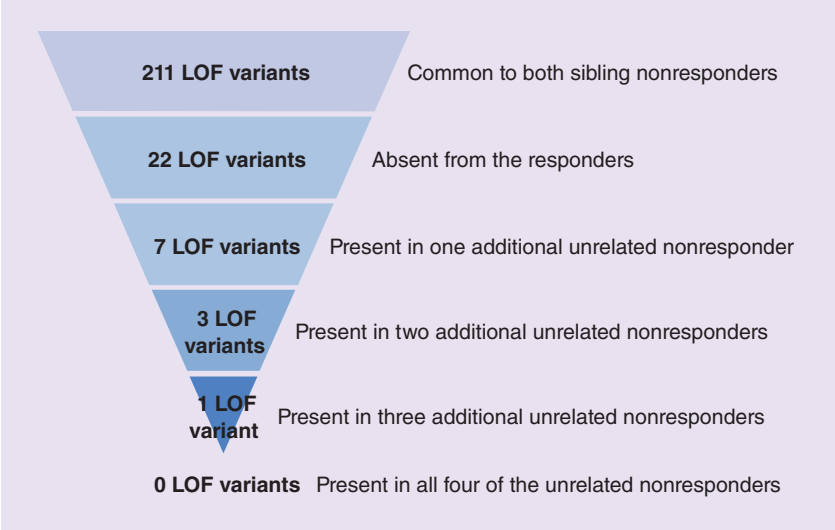

Figure 3. Filtration pipeline used to prioritize variants to be genotyped in the entire cohort.
LOF: Loss of function.

particular relevance to African individuals. Thus this study has provided information that may be of value to future studies examining antipsychotic treatment outcomes in the context of Africa.

The fact that many of these LOF variants were either novel or occurred at low frequencies

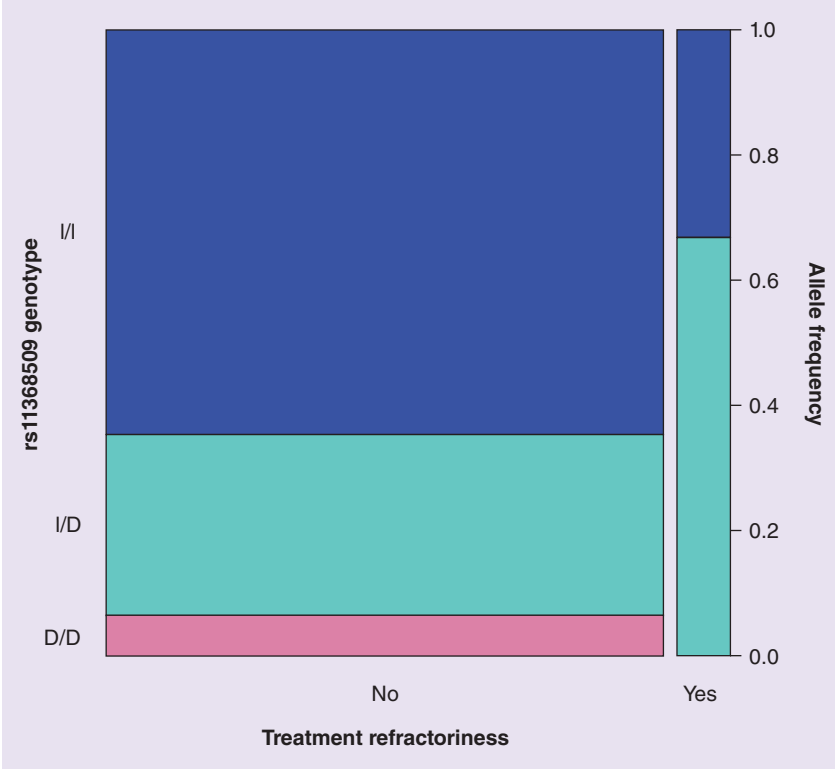

Figure 4. Comparison of genotype distributions according to treatment refractoriness in the first-episode schizoprenia cohort. A higher proportion of deletion alleles was observed in the 'yes' group (treatment-refractory group). The six individuals who were homozygous for the deletion allele were all 'no' (nontreatment-refractory group).

also highlights the likely role that rare or novel variants may play in the antipsychotic treatment response phenotypes. Due to the fact that examination of known common variation through the use of GWAS has yielded incomplete information regarding the contribution of genetics to schizophrenia and antipsychotic treatment response [5], it remains likely that some of the answers to the missing heritability lie in rare or novel variation, which can only be detected through sequencing. Further evidence for this theory is provided by the fact that a recent publication examining autism has reported a two-fold increase in rare LOF variants (minor allele frequency <0.05) in cases when compared with controls [32]. As the genetics of schizophrenia and autism have been shown to overlap [33], such findings may be transferable to schizophrenia and related phenotypes.

Sequencing studies examining schizophrenia have also emphasized the potential role that rare and *de novo* variants play in the development of schizophrenia [7–10,34,35]. As these variants are unlikely to be shared between different populations [36], it may be necessary to examine many different populations to determine the exact variation contributing to poor treatment outcomes and the development of schizophrenia. This is especially relevant for southern African populations, which exhibit high levels of variation, but have been under-represented in genomic research to date [24–26]. Therefore, the results obtained from this study have provided valuable information regarding the variation present in South African schizophrenia patients, which may be used to guide future studies examining antipsychotic treatment response in the context of Africa.

Conclusion

This study utilized well-characterized patients at extreme ends of the antipsychotic treatment spectrum, as well as the addition of sibling nonresponders, to prioritize variation that may be involved in antipsychotic treatment. These analyses identified 22 LOF variants that may be involved in antipsychotic nonresponse, the effects of which should be investigated in a larger cohort. It should also be noted that the use of schizophrenia patients who have been treated for several years and have never responded to treatment, rather than FES patients, may have increased the power of this study. Nonetheless, using a filtering strategy, this study was able to identify a novel candidate gene, *UPP2*, that may play a role in antipsychotic treatment response. These results will,

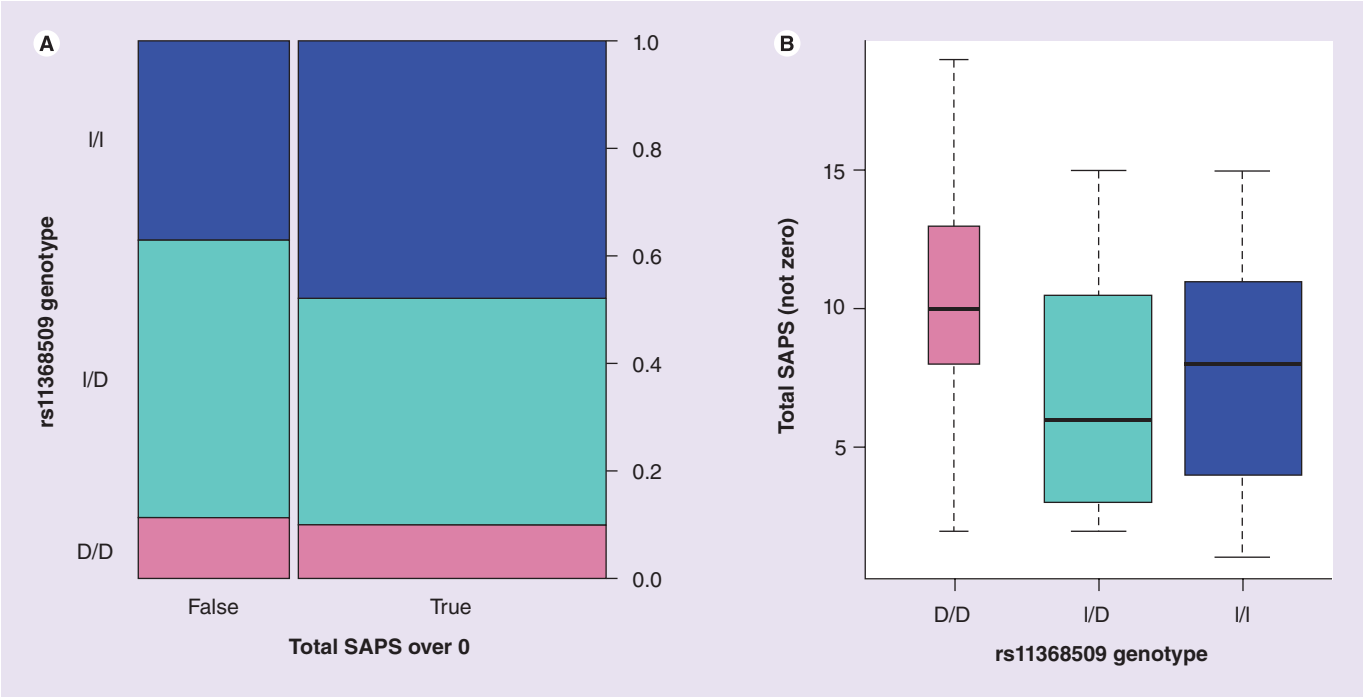

Figure 5. Comparison of genotype distributions according to total SAPS in the Xhosa schizoprenia cohort. (A) Comparison of genotype distributions according to zero or nonzero SAPS. **(B)** Comparison of genotype distributions according to nonzero SAPS. SAPS: Scale for the Assessment of Positive Symptoms.

however, need to be investigated in greater detail in the future to determine their validity. This study has highlighted once again the complexity of antipsychotic treatment response, as well as the importance of rare and novel LOF variation in this phenotype. This reiterates the need for WGES to detect the entire spectrum of variation, particularly in the under-represented African populations. Furthermore, these genotyping strategies should be accompanied by statistical methods to account for variation in entire pathways, the importance of which has been highlighted by Kiezun *et al.* [37].

Future perspective

In the next 5–10 years, WGES will become even more affordable and accessible. Current psychiatric research focuses on common variants; however, it seems likely that it may become more important to focus on rare variants. It appears that the variants involved in antipsychotic treatment response may differ between individuals; however, the pathways involved may be common. Thus, in future, statistical methods that are able to examine rare variants in multiple genes will be necessary. The implementation of these strategies may identify novel drug targets and aid in optimal antipsychotic treatment regimens. In addition, due to the known heterogeneity of antipsychotic treatment response in schizophrenia patients, in the future large consortiums such as the PGC should begin to implement carefully designed pharmacogenomic studies. Such studies should aid in the elucidation of the genetic variants involved in antipsychotic treatment response.

Disclaimer

The opinions expressed and conclusions arrived at, are those of the authors and are not necessarily attributed to the funding sources.

Financial & competing interests disclosure

The work reported here was supported by grants to the following authors: BI Drögemöller is the recipient of a National Research Foundation (NRF) research bursary and the L'Oréal-UNESCO for Women in Science in sub-Saharan Africa Fellowship; DJH Niehaus is the recipient of a South African Medical Research Council (MRC) operating research grant; B Chiliza is the recipient of a South African MRC and Hamilton Naki Clinical Research Fellowship; AK Malhotra acts as a consultant for Genomind Inc. and has received a grant from AbbVie; R Emsley has participated in speakers/advisory boards and received honoraria from AstraZeneca, Bristol-Myers Squibb, Janssen, Lilly, Lundbeck, Organon, Pfizer, Servier, Otsuka and Wyeth. He has also received research funding from Janssen, Lundbeck and AstraZeneca; L Warnich is the recipient of a NRF operating research grant. This study was supported by a New Partnership for Africa's Development (NEPAD) grant, through the Department of Science and Technology

of South Africa. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

Background

- Although schizophrenia and antipsychotic treatment response are heritable, the genetic factors involved remain poorly understood.
- There is growing evidence that rare functional variants may be involved in the development of schizophrenia and antipsychotic response.
- Whole-genome and exome sequencing is well-suited to antipsychotic pharmacogenomic studies due to its ability to detect common and rare variation in all genes.

Materials & methods

- The exomes of 11 FES patients, falling on extreme ends of the antipsychotic treatment response spectrum, were sequenced. The 11 individuals were as follows:
  - Five nonresponders to antipsychotic treatment.
  - Five responders to antipsychotic treatment.
  - A sibling of one of the nonresponder individuals, who was also a nonresponder to treatment.
- The variation in five unrelated antipsychotic nonresponders and five unrelated responders was compared.
- The loss-of-function (LOF) variants present in both nonresponder siblings were examined.
- A filtering strategy was applied to identify variants that were most likely to be involved in treatment response outcomes.
- The prioritized variant was genotyped in two schizophrenia cohorts and association analyses were performed.

Results

- No significant differences in the number of variants in the nonresponders and responders were identified.
- Only a small portion of the novel variation detected was shared between the nonresponder and responder individuals.
- Novel variation was more likely to abolish the function of the protein.
- Twenty-two LOF variants were detected in both nonresponder siblings, which were absent in the responder individuals.
- Twenty of these LOF variants were novel or had minor allele frequency <0.05 in the Asian and European 1000 Genomes Project populations.
- One variant survived the filtering pipeline and was found to be weakly associated with better antipsychotic treatment outcomes.

Discussion

- This is the first study to utilize exome sequencing for antipsychotic pharmacogenomics.
- Although one variant was found to be weakly associated with antipsychotic treatment response, this study was unable to identify a single variant that could explain the nonresponse phenotype, highlighting the complexity of this phenotype.
- A panel of variants that could potentially impact on treatment nonresponse were identified, the majority of which were rare or novel.
- Whole-genome and exome sequencing is an important tool for future antipsychotic pharmacogenomic studies, particularly in the context of Africa, due to its ability to detect rare and novel variation in many genes.

References

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1 Manolio TA, Collins FS, Cox NJ *et al.* Finding the missing heritability of complex diseases. *Nature* 461(7265), 747–753 (2009).

2 Girard SL, Dion PA, Rouleau GA. Schizophrenia genetics: putting all the pieces together. *Curr. Neurol. Neurosci. Rep.* 12(3), 261–266 (2012).

3 Arranz MJ, Kapur S. Pharmacogenetics in psychiatry: are we ready for widespread clinical use? *Schizophr. Bull.* 34(6), 1130–1144 (2008).

▪▪ **Good review of past antipsychotic pharmacogenomic studies.**

4 Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet.* 13(8), 537–551 (2012).

▪ **Excellent summary of the psychiatric research performed to date.**

5 Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat. Genet.* 43(10), 969–976 (2011).

▪ **Highlights the work done by the Psychiatric Genome-Wide Association Study Consortium, which is a frontrunner in psychiatric research.**

6 Lander ES. QnAs with Eric S. Lander. Interview by Prashant Nair. *Proc. Natl Acad. Sci. USA* 108(28), 11319 (2011).

7 Xu B, Roos JL, Dexheimer P *et al.* Exome sequencing supports a *de novo* mutational paradigm for schizophrenia. *Nat. Genet.* 43(9), 864–868 (2011).

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

9 Girard SL, Gauthier J, Noreau A *et al.* Increased exonic *de novo* mutation rate in

individuals with schizophrenia. *Nat. Genet.* 43(9), 860–863 (2011).

10 Need AC, McEvoy JP, Gennarelli M *et al.* Exome sequencing followed by large-scale genotyping suggests a limited role for moderately rare risk factors of strong effect in schizophrenia. *Am. J. Hum. Genet.* 91(2), 303–312 (2012).

11 Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* 148(6), 1223–1241 (2012).

12 Cole J. Evaluation of drug treatments in psychiatry. *Proc. Annu. Meet. Am. Psychopathol. Assoc.* 52, 14–31 (1964).

13 Leucht S, Barnes TR, Kissling W, Engel RR, Correll C, Kane JM. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. *Am. J. Psychiatry* 160(7), 1209–1222 (2003).

14 Kahn RS, Fleischhacker WW, Boter H *et al.* Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *Lancet* 371(9618), 1085–1097 (2008).

15 Mas S, Llerena A, Saíz J, Bernardo M, Lafuente A. Strengths and weaknesses of pharmacogenetic studies of antipsychotic drugs: the potential value of the PEPs study. *Pharmacogenomics* 13(15), 1773–1782 (2012).

16 Arranz MJ, Rivera M, Munro JC. Pharmacogenetics of response to antipsychotics in patients with schizophrenia. *CNS Drugs* 25(11), 933–969 (2011).

17 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR).* American Psychiatric Association, Washington, DC, USA (1994).

18 Wright GEB, Niehaus DJH, van der Merwe L *et al.* Association of MB-COMT polymorphisms with schizophrenia-susceptibility and symptom severity in an African cohort. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 39(1), 163–169 (2012).

19 Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25(14), 1754–1760 (2009).

20 Li H, Handsaker B, Wysoker A *et al.* The sequence alignment/map format and SAM tools. *Bioinformatics* 25(16), 2078–2079 (2009).

21 DePristo MA, Banks E, Poplin R *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43(5), 491–498 (2011).

22 Milacic M, Haw R, Rothfels K *et al.* Annotating cancer variants and anti-cancer therapeutics in reactome. *Cancers* 4(4), 1180–1211 (2012).

23 Croft D, O'Kelly G, Wu G *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39(Database issue), D691–D697 (2011).

24 Schuster SC, Miller W, Ratan A *et al.* Complete Khoisan and Bantu genomes from southern Africa. *Nature* 463(7283), 943–947 (2010).

25 Wright GE, Niehaus DJ, Koen L, Drogemoller BI, Warnich L. Psychiatric genetics in South Africa: cutting a rough diamond. *Afr. J. Psychiatry (Johannesburg)* 14(5), 355–366 (2011).

26 Drogemoller BI, Wright GEB, Niehaus DJH, Emsley RA, Warnich L. Whole-genome resequencing in pharmacogenomics: moving away from past disparities to globally representative applications. *Pharmacogenomics* 12(12), 1717–1728 (2011).

27 Cao D, Russell RL, Zhang D, Leffert JJ, Pizzorno G. Uridine phosphorylase (-/-) murine embryonic stem cells clarify the key role of this enzyme in the regulation of the pyrimidine salvage pathway and in the activation of fluoropyrimidines. *Cancer Res.* 62(8), 2313–2317 (2002).

28 Roosild TP, Castronovo S, Villoso A, Ziemba A, Pizzorno G. A novel structural mechanism for redox regulation of uridine phosphorylase 2 activity. *J. Struct. Biol.* 176(2), 229–237 (2011).

29 Myers CS, Fisher H, Wagner GC. Uridine potentiates haloperidol's disruption of conditioned avoidance responding. *Brain Res.* 651(1–2), 194–198 (1994).

30 Putnam DK, Sun J, Zhao Z. Exploring schizophrenia drug–gene interactions through molecular network and pathway modeling. *AMIA Annu. Symp. Proc.* 2011, 1127–1133 (2011).

31 Zandi PP, Judy JT. The promise and reality of pharmacogenetics in psychiatry. *Psychiatr. Clin. North Am.* 33(1), 181–224 (2010).

▪▪ **Refers to the lack of antipsychotic pharmacogenetic studies in African individuals.**

32 Lim ET, Raychaudhuri S, Sanders SJ *et al.* Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. *Neuron* 77(2), 235–242 (2013).

▪▪ **Describes an increased rate of rare loss-of-function variants in autism patients.**

33 Smoller JW, Craddock N, Kendler K *et al.* Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 381(9875), 1371–1379 (2013).

34 Jouan L, Girard SL, Dobrzeniecka S *et al.* Investigation of rare variants in *LRP1*, *KPNA1*, *ALS2CL* and *ZNF480* genes in schizophrenia patients reflects genetic heterogeneity of the disease. *Behav. Brain Funct.* 9, 9 (2013).

35 Gratten J, Visscher PM, Mowry BJ, Wray NR. Interpreting the role of *de novo* protein-coding mutations in neuropsychiatric disease. *Nat. Genet.* 45(3), 234–238 (2013).

36 Abecasis GR, Auton A, Brooks LD *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422), 56–65 (2012).

37 Kiezun A, Garimella K, Do R *et al.* Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* 44(6), 623–630 (2012).

▪▪ **Excellent perspective on exome sequencing in complex diseases.**

Websites

101 Illustration: Quantitative Advances Since the Human Genome Project (HGP). www.genome.gov/images/illustrations/hgp_measures.pdf

102 PGC – Psychiatric Genomics Consortium. https://pgc.unc.edu/

103 SZGene. www.szgene.org

104 Genomic Services Laboratory at HudsonAlpha. www.hudsonalpha.org/gsl/index

105 SeattleSeq Variation Annotation. http://snp.gs.washington.edu/SeattleSeqAnnotation134

106 Home – SNP – NCBI. www.ncbi.nlm.nih.gov/snp

107 Genomes Project Browser. http://browser.1000genomes.org/index.html

108 Miller M: Tools for population genetic analyses (TFPGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author (1997). www.marksgeneticsoftware.net/_vti_bin/shtml.exe/tfpga.htm

109 Home – PubMed – NCBI. www.ncbi.nlm.nih.gov/pubmed

110 Home – Gene – NCBI. www.ncbi.nlm.nih.gov/gene