

Phenotypic features of patients with schizophrenia carrying de novo gene mutations: A pilot study

P.J. Malherbe^{a,*}, J.L. Roos Jr.^a, R. Ehlers^b, M. Karayiorgou^{c,d}, J.L. Roos^a

^a Department of Psychiatry, University of Pretoria, Weskoppies Hospital, Private Bag X113, Pretoria 0001, South Africa ^b Department of Statistics, University of Pretoria, Pretoria, South Africa

^c Department of Psychiatry, Columbia University, New York, USA

^d Extraordinary Professor, University of Pretoria, Pretoria, South Africa

ABSTRACT

Genome-wide scans have revealed a significant role for de novo copy number variants (CNVs) and Single Nucleotide variants (SNVs) in the genetic architecture of schizophrenia. The present study attempts to parse schizophrenia based on the presence of such de novo mutations and attempts genotype–phenotype correlation. We examined phenotypic variables across three broad categories: clinical presentation, premorbid function, disease course and functional outcome and compared them in individuals with schizophrenia carrying either a de novo CNV, a de novo SNV, or no de novo mutation. Work skills were worst affected in patients carrying de novo CNVs. More learning disabilities were found in subjects carrying de novo SNVs. Patients with either mutation had older parents at birth and worse functional outcome as measured by SLOF scores. We found no relation between treatment resistance and the presence of de novo mutations. The combined consideration of the functional outcome scores and early deviant behaviours was found to have higher predictive value for underlying genetic vulnerability. Due to the rare nature of the de novo mutations the sample sizes studied here were small. Despite this, valuable phenotypic characteristics were identified in schizophrenia patients carrying de novo mutations and studying larger samples will be of interest.

Keywords:

Copy number variant
Single nucleotide polymorphism
Paternal age
Rare mutation
Early deviance
Functional outcome

1. Introduction

Schizophrenia is a leading cause of disability worldwide and is both highly heritable and highly genetically heterogeneous (Rodriguez-Murillo et al., 2012). Advances in genomics have enabled the next generation of studies into the genetics of schizophrenia. Rare gene-disrupting Copy Number Variants (CNVs), Single Nucleotide Variants (SNVs) and small insertions and deletions (indels) have been found to contribute substantially to the disorder (McClellan and King, 2010; Walsh et al., 2008; Xu et al., 2008, 2011, 2012; Stefansson et al., 2008; Gulsuner et al., 2013) and carry high pathogenicity value in their de novo form. A de novo mutation is a genetic alteration that is present for the first time in one family member as a result of a mutation in a germ cell (ovum or sperm) of one of the parents or in the embryo itself. Genes disrupted in schizophrenia and schizoaffective disorder may be revealed by de novo mutations in affected persons from otherwise healthy families (sporadic cases).

We have previously conducted genome-wide scans for de novo CNVs and SNVs in a well-characterized cohort of trios of Afrikaner

families in South Africa consisting of individuals affected with schizophrenia and their biological parents. These scans revealed a significant role of de novo CNVs and SNVs in the genetic architecture of schizophrenia (Xu et al., 2008, 2009, 2011, 2012).

In the present study, we aim to determine whether phenotypic differences could be identified among three groups of patients with schizophrenia as stratified by their de novo mutational status. We compare patients who carry a de novo CNV (Group A), to patients who carry a de novo SNV (Group B), to patients carrying no detectable de novo mutations (Group C). We examined phenotypic variables across three broad categories: clinical variables as a way to assess qualitative differences, premorbid variables as indication of a neurodevelopmental course influenced by the presence of a de novo mutation; disease course and functional outcome as a means to determine whether presence of a de novo CNV or SNV is associated with worse outcome.

2. Materials and methods

Subject recruitment: a large number of families with schizophrenia have been recruited from the Afrikaner population over a number of years for a collaborative genetic study (Karayiorgou et al., 2004). The families are of varying structure, and include a large number of trios of families. Each subject underwent a careful,

* Corresponding author. Tel.: +27 12 319 9720.

in-person diagnostic evaluation using the Diagnostic Interview for Genetic Studies (DIGS) at recruitment (Nurnberger et al., 1994).

A subset of probands from this sample was re-contacted for participation in the current study by the principal clinical investigator of the collaborative study. Follow-up evaluations were performed by two senior registrars (psychiatrists in training) under the supervision of the principal clinical investigator:

- A new diagnostic interview using the DIGS was conducted in order to confirm diagnosis stability since initial recruitment (Nurnberger et al., 1994).
- A Specific Level of Functioning assessment scale (SLOF) was completed by a caregiver (Schneider and Streuning, 1983).
- A checklist on early deviant behaviour in the first 10 years of life was completed (Sobin et al., 2003).
- Other relevant data were collected, including present medication and parents' age at birth. Information was obtained by family members where the patient could not give details.

The senior registrars were blind to the original recruitment DIGS summary reports as well as the genetic status and grouping of the patients. The principal clinical investigator was either present at the follow up interviews or reviewed and discussed the findings of their interviews with the registrars. He was also blind to the genetic status and grouping of the patients.

Ethical approval for this study was obtained from the Ethics Committee of the Faculty of Health Sciences at the University of Pretoria.

A total of 24 patients were identified for participation in this study (7 from Group A, 8 from Group B, and 9 from Group C). Two patients from Group B were lost to follow up. The sample sizes in this study were determined by the number of subjects previously identified to carry de novo CNVs (Group A in this study). In our Xu et al. (2008) study, of 152 subjects that were included in our genome-wide microarray scan, a total of 15 subjects were identified to carry one or more de novo CNV. Of this 15, two were excluded in the present study because they carried at least 2 de novo CNVs; another three were excluded because they carried deletions in chromosome 22q11.2, which occurs recurrently and is being studied separately. Of the remaining 10 subjects carrying de novo CNVs, we were able to recontact 7 probands successfully and secure their participation in the present study (Group A). The remaining groups (B and C) were formed to match Group A in size.

A detailed Table is provided in Appendix A listing the exact chromosomal location of each CNV and SNV, as well as the specific genes disrupted. In addition, age of onset, as well as the number of all de novo CNVs or SNVs per subject is provided.

2.1. The Specific levels of Functioning (SLOF) Scale

The SLOF Scale was the best rated scale by the Validation of Everyday Real-World Outcomes (VALERO) study, (Harvey et al., 2011).

SLOF is a 43-item multidimensional behavioural survey administered in person to the caseworker or caregiver of a schizophrenic patient. The scale assesses the patient's current functioning and behaviour across 6 domains: (1) – physical functioning; (2) – personal care skills; (3) – interpersonal relationships; (4) – social acceptability; (5) – activities of community living; and (6) – work skills. Each of the questions is rated on a 5-point Likert scale and total scores range from 43 to 215. The higher total score, the better the overall functioning of the patient (Schneider and Streuning, 1983).

2.2. Early deviant behaviour checklist

The early childhood behaviour questionnaire probes seven areas of possible deviance including social dysfunction (avoidance of other children, inability to have friends, isolated play), extreme odd behaviours (unprovoked screaming fits, disorganised or irrational behaviour, inappropriate affect), unprovoked aggression, extreme anxiety, chronic sadness, attentional impairment and learning disabilities (Sobin et al., 2003).

2.3. Statistical methodology

Due to the nature of the study, limited patients were available to include in the sample. Since the sample size is small, the assumption of normality necessary for parametric tests could not be verified and is probably violated. Hence statistical data analysis was performed by utilising permutation tests. These tests are not based on any underlying assumptions of the distribution of the data (Edgington and Onghena, 2007). The advantage of permutation tests over the more conventional distribution-free tests is that all the original information in the data are used compared to using only the ranks. Another disadvantage of small sample sizes is that the power of the test to detect significant differences is very low. Because of this and the novel nature of this research, it was decided to not only report results that are significant at the conventional 5% level, but also results that are only moderately significant (p -value < 0.10) (Albright et al., 2005). This approach will assist to gain insight and to direct further research in this field of study.

3. Results

3.1. Stability of the diagnoses

The lifetime diagnoses originally assigned to the subjects were remarkably stable across all 3 groups. The initial study diagnosis was made by a best-estimate process using medical records and collateral information. The average number of years to the follow up assessment since the initial recruitment was 10, 12, and 13 years in Groups A, B, and C, respectively. The diagnoses remained the same in all but 2 cases from Group A, 1 case from Group B, and 1 case from Group C. The stability of the diagnoses were confirmed by re-administering the DIGS and collecting other relevant data.

In Group A, one male patient had a dual diagnosis of Asperger Syndrome and schizophrenia at initial assessment. After follow up assessment the diagnosis of Asperger Syndrome was discarded.

The modification of diagnosis from schizophrenia to schizoaffective disorder in one patient in each of the 3 groups was done because the longitudinal course of the illness was taken into account and a more accurate picture of the mood syndrome was available at the follow up evaluation. The reliability coefficients for schizoaffective disorder are lower than for other diagnoses made in the DIGS (Nurnberger et al., 1994). It remains difficult to assess

Table 1
Specific level of functioning (SLOF) scores.

	Group A (n=7)		Group B (n=6)		Group C (n=9)		p-Value
	Mean	± S.D.	Mean	± S.D.	Mean	± S.D.	
Self-maintenance							
(a) Physical functioning	24.7	0.8	24.3	1.2	24.4	0.7	0.769
(b) Personal care skills	32.7	3.5	32.3	3.8	32.8	2.8	0.968
Sum of (a) and (b)	57.4	3.5	56.7	3.9	57.2	2.8	0.922
Social functioning							
(c) Interpersonal relationships	19.1	6.6	21.3	6.6	22.8	6.3	0.541
(d) Social acceptability	32.0	2.1	30.8	3.8	32.0	4.5	0.826
Sum of (c) and (d)	51.1	6.0	52.2	8.4	54.8	9.4	0.658
Community living skills							
(e) Activities	40.7	12.6	45.3	9.8	48.7	7.4	0.311
(f) Work skills	13.9	4.0	19.0	5.3	20.1	6.8	0.102
Sum of (e) and (f)	54.6	14.0	64.3	13.3	68.8	13.7	0.143
Total SLOF	163.1	18.2	173.2	23.3	180.8	24.2	0.310

reliably the mood syndrome criteria in the DSM IV of schizoaffective disorder.

Importantly, in all 22 patients, psychosis remained central to their clinical presentation supporting the reliability of the final best-estimate process initially employed in making a lifetime diagnosis.

3.2. Functional outcome

The SLOF assessment scale was used to quantify life skills. Six domains across three broad categories of self-maintenance, social functioning and community living skills were interrogated. The results for the comparison of different SLOF scores for the three de novo mutational groups are given in Table 1. There were no significant differences between any of the mean SLOF scores for the three groups. The total mean SLOF scores were: Group A=163.1; Group B=173.2; Group C=180.8. Although not statistically significant, the trend observed is that the higher the genetic loading, the lower the functional outcome. It would be interesting to see if this finding replicates in larger sample sizes.

In the domain 'work skills', the mean SLOF score for group A (patients carrying de novo CNVs) was 13.9 – lower than the mean of 20.1 for Group C (with no detectable de novo events) indicating

a worse adaptability to life demands and less capacity to develop independent skills in the group carrying de novo CNVs.

3.3. Early deviant behaviours

The presence of one or more of seven early, premorbid behaviours that might have been present before the age of 10 were evaluated. The age of 10 was determined as a cut-off to avoid interference from behaviours that coincide with puberty onset. The 7-point scale published earlier by Sobin et al. (2003) has been used in other studies with interesting results. When examining probands with one of these early deviances the groups failed to differentiate during statistical analysis, however when two or more of these deviances were present a subset of patients could be identified carrying de novo mutations with early developmental significance indicating that they capture early premorbid neurodevelopmental pathology.

Fisher's exact test was performed to determine whether there is a significant relationship between de novo mutational status and occurrence of specific deviant behaviour in the first ten years of life. The results are displayed in Table 2. The relationship between de novo mutational status and early occurrence of learning disabilities is moderately significant (p -value=0.097). The percentage of patients with learning disabilities is higher than expected in Group B and in Group C the percentage is lower than expected under the assumption of independence. Also, although not statistically significant, an enrichment of early deviant social behaviour, labelled "Social dysfunction" in the scale, was noted among the patients carrying de novo CNVs (Group A) (86% compared to 50% in Group B and 44% in Group C).

Table 3 shows results when considering the subset of patients with two or more early deviant behaviours. The mean ages of parents at the time of the patients' birth as well as SLOF scores were again compared for the three groups. Significant differences were detected between the mean ages of both fathers and mothers for the three groups (p -values of 0.015 and 0.013 respectively). Post-hoc pairwise comparisons, using permutation tests and a Bonferroni correction for multiple comparisons, showed that the mean age of fathers in Group B (36.3) was significantly different from the mean age of fathers in Group C (27.2) (p -value =0.012). For mothers there was only moderate evidence (p -value=0.024)

Table 2

Early deviant behaviours in the first ten years of life.

	Group A (n=7)		Group B (n=6)		Group C (n=9)		p-Value
	n (yes)	%	n (yes)	%	n (yes)	%	
One or more early deviant behaviours	6	86	5	83	9	100	0.494
Two or more early deviant behaviours	4	57	3	50	6	67	0.867
Social dysfunction	6	86	3	50	4	44	0.226
Extreme odd behaviours	1	14	0	0	2	22	0.755
Unprovoked aggression	1	14	0	0	3	33	0.406
Extreme anxiety	0	0	1	17	3	33	0.252
Chronic sadness	2	29	2	33	2	22	1.000
Attentional impairment	3	43	3	50	6	67	0.659
Learning disabilities	2	29	4	67	1	11	0.097

Table 3

Comparison of patients with two or more early deviant behaviours.

	Group A (n=4)		Group B (n=3)		Group C (n=6)		p-Value
	Mean	± S.D.	Mean	± S.D.	Mean	± S.D.	
Ages of parents at birth of patient							
Father	31.8	4.9	36.3	0.6	27.2	3.4	0.015
Mother	30.5	2.1	34.3	2.1	26.7	3.6	0.013
Specific level of functioning (SLOF) results							
Self-maintenance							
(a) Physical functioning	24.5	1.0	23.7	1.5	24.7	0.5	0.416
(b) Personal care skills	31.3	4.2	32.3	3.8	33.0	2.9	0.756
Sum of (a) and (b)	55.8	3.9	56.0	4.0	57.7	3.1	0.675
Social functioning							
(c) Interpersonal relationships	16.0	1.6	19.0	7.9	21.7	7.3	0.429
(d) Social acceptability	32.3	2.6	33.0	2.6	31.8	5.4	0.979
Sum of (c) and (d)	48.3	2.1	52.0	9.5	53.5	10.8	0.659
Community living skills							
(e) Activities	36.3	15.6	46.3	12.4	47.2	8.8	0.390
(f) Work skills	14.5	4.7	16.0	2.0	18.7	7.6	0.573
Sum of (e) and (f)	50.8	18.5	62.3	14.2	65.8	15.7	0.390
Total SLOF	154.8	20.2	170.3	25.0	177.0	28.5	0.425

that the mean age in Group B (34.4) is different from the mean age in Group C (26.7). Patients with two or more early deviant behaviours therefore tend to have older parents, however an overall paternal age trend (mentioned in Section 3.4), may account for the difference seen in this subgroup.

It is interesting to note that in this subset, as in the entire sample set, the total average SLOF scores maintain the same gradient, with Group A having more dysfunction than Group B, than Group C. Also, of interest is that the total scores in this subset (154.8/170.3/177.0 across the three Groups) are lower than they were for the entire sample set (163.1/173.2/180.8, respectively, see Table 1) consistent with worse functional outcomes for patients with 2 or more early deviant behaviours. This is particularly evident in the interpersonal relationships, where the scores drop from 19.1/21.3/22.8 across the 3 Groups in the entire sample to 16.0/19.0/21.7 in the subset. The community living skills follow a similar trend dropping from 54.6/64.3/68.8 in the entire sample to 50.8/62.3/65.8 in the subset. Circumspect interpretation of these data should take into consideration the small sample size.

3.4. Parental age

Finally, we also examined whether the parental age at birth differs among the three Groups. Not surprising, and in agreement with recent literature, we find that parental age at the time of proband's birth is significantly higher in Group B (subjects carry-

ing de novo SNVs) than Group C or A. (Kong et al., 2012) Specifically, as can be seen in Table 4, the mean ages of fathers are significantly different on a 5% level (p -value=0.042). In the post-hoc pairwise permutation tests a Bonferroni correction was applied to correct for multiple comparisons when comparing the mean ages of the fathers in Groups A, B and C. It was found that the mean age of fathers in Group C (27.1) is significantly different on the 5% level (p -value=0.007) from the mean age of fathers in Group B (34.2). No other significant differences were detected. This is in agreement with recent literature where older fathers accumulate more mutations in their germline and are therefore more likely to pass them on as de novo SNVs in their offspring.

3.5. Other variables

Table 5 presents details on the medication that was being used during the original recruitment and during the follow up evaluation ~10 years later. We categorised the current medication into 1st and 2nd generations and kept clozapine separately since it usually indicates treatment resistance.

Judging from these results, clozapine use and thus treatment resistance cannot be used as a phenotypic marker to distinguish the three groups.

4. Discussion

Technological advances in genomics, like deep sequencing is paving the way for genotype–phenotype relationship studies, which can be of significant translational value. Parsing the heterogeneous phenotype of schizophrenia to gene-specific subtypes will allow tremendous progress in our understanding of disease pathophysiology, as well as in development of biomarkers and treatments. Although there are already some gene-specific subtypes that have been identified in schizophrenia (i.e., 22q11.2 microdeletion, VIPR2 duplication, etc.), for the most part the

Table 4
Parental age at the time of the patients' birth.

	Group A (n=7)		Group B (n=6)		Group C (n=9)		p-Value
	Mean	± S.D.	Mean	± S.D.	Mean	± S.D.	
Father	32.3	7.4	34.2	5.0	27.1	2.9	0.042
Mother	29.6	5.6	30.2	5.5	26.3	2.9	0.236

Table 5
Medication used during recruitment and FU evaluation.

GROUP	Recruitment	Follow up	Group	Recruitment	Follow up	Group	Recruitment	Follow up
A 071	2° AP ⊗ BZD ⊗	1° AP ⊗ SSri ⊗ AC ⊗ BZD ⊗	B 007	1° AP ⊗ SSri ⊗ BZD ⊗	2° AP ⊗ MS ⊗ AD ⊗ BZD ⊗	C 016	1° AP ⊗ SSri ⊗ AC ⊗	1° AP ⊗ SSri ⊗ AC ⊗
A 085	2° AP ⊗	2° AP ⊗ BZD ⊗	B 058	1° AP ⊗ AC ⊗	2° AP ⊗ 1° AP ⊗	C 029	2° AP ⊗	CLZP ⊗ SSri ⊗
A 137	2° AP ⊗	2° AP ⊗ ⊗	B 080	2° AP ⊗	2° AP ⊗	C 046	2° AP ⊗ AC ⊗	2° AP ⊗
A 251	2° AP ⊗	2° AP ⊗ AC ⊗	B 169	2° AP ⊗	AD ⊗ 1° AP ⊗ AC ⊗	C 086	2° AP ⊗	2° AP ⊗
A 264	1° AP ⊗	1° AP ⊗ SSri ⊗	B 294	2° AP ⊗	2° AP ⊗ AD ⊗	C 116	2° AP ⊗	2° AP ⊗ SSri ⊗
A 287	CLZP ⊗	CLZP ⊗ 2° AP ⊗ 1° AP ⊗	B 613	2° AP ⊗ CLZP ⊗	2° AP ⊗ 2° AP ⊗ CLZP ⊗	C 120	CLZP ⊗ BZD ⊗ AC ⊗	2° AP ⊗ AC ⊗ SSri ⊗ BZD ⊗
A 333	2° AP ⊗	CLZP ⊗				C 207	1° AP ⊗ AC ⊗	1° AP ⊗ CLZP ⊗ MS ⊗ BZD ⊗
						C 245		MS ⊗ SSri ⊗
						C 342	2° AP ⊗ SSri ⊗	2° AP ⊗ MS ⊗ SSri ⊗

2° AP: second generation antipsychotic, 1° AP: first generation antipsychotic, CLZP: clozapine, SSri: selective serotonin reuptake inhibitor, AC: anticholinergic, BZD: benzodiazepine, MS: mood stabiliser, and AD: other antidepressant.

specific genetic causes remain unknown. The collective significance of certain types of genetic mutations, such as de novo CNVs and de novo SNVs are of interest. This study attempts to parse schizophrenia based on presence of such mutations and attempts genotype–phenotype correlation studies.

Recruiting research participants based on genetic information is a potentially powerful way to study the functional significance of genetic variation, but presents ethical challenges. (Beskow et al., 2012) At recruitment ~10 years ago, permission to contact participants for further research was obtained and when the principal clinical investigator of the original research team contacted the follow up participants, all agreed to participate. Two participants were lost to follow-up – a 92% recall rate. This is remarkably high and it is due, in part, to the close knit family structure of the Afrikaner population, as well as the strong rapport between the study participants and the principal clinical investigator.

The re-administration of the DIGS, 10–13 years after original clinical assessment revealed a remarkable stability of diagnoses and offered a rare opportunity to evaluate the longitudinal course of the disorder. With the exception of 3 cases, where the mood syndrome manifested better over time resulting in a diagnosis of schizoaffective disorder, all other diagnoses remained the same. The even distribution of the 3 cases across the sample set did not correlate with the mutational status. Worth mentioning is the one patient from Group A, where the dual diagnosis of Asperger's was dismissed. It was felt that the autism features were consistent with features of early onset schizophrenia.

Deficits in the performance of critical everyday functional skills are particularly salient in schizophrenia (Wiersma et al., 2000). The subgroups of the SLOF scale attempt to capture deficits in several distinct functional domains. Work skills were the worst affected in Group A (de novo CNVs), whilst Groups B and C were very much equal. This suggests a greater level of dysfunction associated with the presence of de novo CNVs, which is not surprising as CNVs tend to affect a large number of genes leading to greater disability. Our results suggest that there is merit in evaluating functioning for particular domains rather than globally and the SLOF is a good instrument for this.

Literature suggests that schizophrenia is a neurodevelopmental disorder with roots in early childhood (Roos et al., 2009). Early deviant behaviour patterns in children have been consistently (Cannon and Jones, 1996) reported in high-risk, prospective and retrospective behavioural studies of schizophrenia.

Early behavioural deviance appears to occur in a subgroup of patients (Sobin et al., 2001). Early behavioural deviance and the various neurodevelopmental anomalies may in fact be both a subtype and disease-onset marker. These findings have important implications (Sobin et al., 2001). For those patients whose histories include early deviant behaviours, the behaviours themselves may provide important clues about brain development and, by extension, possible neurobiologic underpinnings of schizophrenia (Sobin et al., 2001).

In our sample, we find an increased percentage of patients with learning disabilities (67%) in Group B (de novo SNVs). Importantly, subjects with a history of two or more early deviant behaviours in Group B had significantly older mothers and fathers at birth compared to Group C (*p*-values of 0.013 and 0.015 respectively). This, together with the trend toward lower total SLOF scores observed in Groups A and B indicate that the combined use of the SLOF and the early deviant behaviour checklist has higher predictive value for underlying genetic vulnerability.

Treatment with clozapine serves as a marker for treatment resistance, as it is the only drug indicated for the management of treatment resistant schizophrenia and has few other indications (Elkis, 2007; Kane et al., 1988). The tertiary hospital where most of

the patients in this study were treated adheres to a strict protocol for clozapine use. We did not find any relationship between clozapine prescription and, therefore, treatment resistance and mutational status.

5. Limitations of the study

Sample sizes were small. Due to the rare nature of the mutations studied, and long term follow-up nature of this research a larger sample size could not be obtained. Larger sample sizes for future studies will be of interest. We also recognise the importance of significant findings in small study groups and would like to see this research contribute to evidence based medicine as part of future collaborations and/or meta analysis (Haidich, 2010).

Information regarding early deviant behaviours was mostly obtained from the patients themselves. The validity of this method may be questioned.

6. Conclusions

Due to the nature of the study the sample size is small and findings should be interpreted with due circumspection prior to potential replication in subsequent studies.

Follow up lifetime diagnoses remained stable over a period of more than 10 years across all three Groups of subjects. The mean age of the fathers at the time of the subjects' birth was older in subjects carrying de novo SNVs, in accordance with existing literature on an increasing number of mutations in the germline of advanced age fathers.

The SLOF scale was used as a measurement of real-world functional outcomes. The total SLOF scores followed a gradient of severity of dysfunction, with subjects carrying de novo CNVs (Group A) being the most dysfunctional followed by subjects carrying de novo SNVs (Group B) and subjects with no genetic findings (Group C). Interpersonal relationships and living skills were the most prominently affected.

Early behavioural deviance appears to occur in only a subgroup of patients. Learning disability as one of the seven areas of deviance was significantly affected in patients carrying de novo SNVs. This finding concurs with literature suggesting that schizophrenia is a neurodevelopmental disorder with roots in childhood.

Treatment resistance was not prominent in this study.

In spite of the small numbers of patients included in this study, valuable phenotypic characteristics were identified in schizophrenia patients with de novo CNVs and SNVs as identified in a subset of Afrikaner patients.

Conflict of interest

The authors do not have any actual or potential conflict of interest to disclose

Contributors

JLR, PM, JLR(Jnr), MK designed the study and wrote the protocol.

JLR(Jnr), PM performed the literature search.

JLR, PM, JLR(Jnr) performed the analysis.

RE undertook the statistical analysis.

JLR wrote the first draft of the manuscript.

All authors contributed to and have approved the final manuscript.

Appendix A

ID	Chromosomal location	Start	End	Size (kb)	#SNPs	Genes
Group A						
Subj #1	10q21.1	57665722	57693353	27,631	0	
Subj #2	19q13.12	42014819	42193363	178,544	1	ZNF790, ZNF345, DKFZp7790175, ZNF568
Subj #3	14q32.13–14q32.2	93831174	96634809	2803,635	646	SERPINA6, SERPINA1, SERPINA11, SERPINA9, SERPINA12, SERPINA4, SERPINA5, SERPINA3, SERPINA13, GSC, DICER1, CLMN, C14orf49, GLRX5, TCL6, TCL1B, TCL1A, BDKRB2, BDKRB1, C14orf103, C14orf129, AK7, PAPOLA, VRK1
	1p34.2	40063283	40677344	614,061	2	TRIT1, MYCL1, MFSD2, CAP1, PPT1, RLF, TMCO2, ZMPSTE24, COL9A2, SMAP1L
Subj #4	16p13.2	6888176	6918847	30,671	4	A2BP1
Subj #5	3q22.2	135765543	137788166	2022,623	99	CEP63, KY, EPHB1, PPP2R3A, MSL2L1, PCCB, STAG1
Subj #6	3p26.1	8126457	8193633	67,176	29	
Subj #7	12q24.23	118689478	118714328	24,85	0	CIT
	9p23	11617789	11877785	259,996	8	

ID	Chr	Start	End	Nucleotide change	cDNA	RNA/aa change	Mut. type	Property	LOF	Gene symbol	Temp. exp. patt.	AOO
Group B												
Subj #1	1	229783350	229783350	c.4000G > T	AG-tAG	p.E1334*	NV	Nonsynonymous	Y	RB2	Pr	22
	17	37374188	37374188	c.329T > C	CTG-CcG	p.L110P	NV	Nonsynonymous	N	STAC2	Po	
	19	46087891	46087891	c.132G > A	CCG-CCa	p.P44P	SNV	Synonymous	N	OPA3	Po	
Subj #2	1	43891777	43891777	c.2998delCinsCCA		p.	indel	Frameshift	Y	KIAA0467	Po	24
						Gln1001HisfsX47		(insertion)				
Subj #3	6	26157173	26157174	c.555_557delCAAinsC		p.Lys186GlufsX9	indel	Frameshift (deletion)	Y	HIST1H1E	Pr	19
	8	81905378	81905378	c.85G > A	GTC-aTC	p.V29I	SNV	Nonsynonymous	N	PAG1	Pr	
Subj #4	8	53568705	53568706	c.3682_3684delGAAinsA		p.Glu1228ThrfsX7	indel	Frameshift (deletion)	Y	RB1CC1	Pr	37
	5	75982186	75982186	c.1220C > T	CA-CtA	p.P407L	SNV	Nonsynonymous	N	CSPG4	Un	
Subj #5	11	124626163	124626163	c.166delGinsCGCTG		p.Val56ArgfsX48	indel	Frameshift	Y	ESAM	Pr	22
						(insertion)						
Subj #6	6	129835668	129835674	c.9139_9146delGAAGCCCAinsG		p.	indel	Frameshift (deletion)	Y	LAMA2	Po	19
						Ser3050ThrfsX27						
Subj #7	6	33414351	33414351	c.3583-1		r.spl?	S/c	Splice-acceptor_-1	Y	SYNGAP1	Po	22
	17	21319185	21319185	c.531T > A	TT-ATa	p.I177I	SNV	Synonymous	N	KCNJ12	Po	
Subj #8	8	19688027	9688027	c.1251G > C	AGG-AGc	p.R417S	SNV	Nonsynonymous	N	INTS10	Un	7
	8	38107312	38107316	c.1335_1340delAAACTCinsA		p.Ser447GlyfsX3	indel	Frameshift (deletion)	Y	DDHD2	Po	

Chr=chromosome, Mut. type=mutation type, LOF=loss of function, Temp. exp. patt.=temporal expression pattern, AOO=age of onset, S/c=splice/canonical, Pr=prenatally biased, Po=postnatally biased, and Un=unbiased.

References

- Albright, S.C., Winston, L.W., Zappe, C.J., 2005. *Data Analysis and Decision Making with Microsoft Excel*, 3rd ed South-Western Cengage Learning, Mason p. 503.
- Beskow, L.M., Fullerton, S.M., Namey, E.E., Nelson, D.K., Davis, A.M., Wilfond, B.S., 2012. Recommendations for ethical approaches to genotype-driven research recruitment. *Human Genetics* 131, 1423–1431.
- Cannon, M., Jones, P.M., 1996. Schizophrenia. *Journal of Neurology, Neurosurgery and Psychiatry* 60, 604–612.
- Edgington, E.S., Ongheena, P., 2007. *Randomization Tests*, 4th ed. Chapman & Hall/ CRC Press, Boca Raton, London, New York (pp. 55–67, 77–79).
- Elkis, H.J., 2007. Treatment-resistant schizophrenia. *Psychiatric Clinics of North America* 30, 511–533.
- Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Consortium on the Genetics of Schizophrenia (COGS), PAARTNERS Study Group Nimgaonkar, V.L., Go, R.C., Savage, R.M., Swerdlow, N.R., Gur, R.E., Braff, D.L., King, M.C., McClellan, J.M., 2013. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* 154, 518–529.
- Haidich, A.B., 2010. Meta-analysis in medical research. *Hippokratia* 14 (S1), S29–S37.
- Harvey, P.D., Raykov, T., Twamley, E.W., Vella, L., Heton, R.K., Patterson, T.L., 2011. Validating the measurement of real-world functional outcomes: phase 1 results of the VALERO study. *American Journal of Psychiatry* 168, 1195–1201.
- Kane, J., Honingveld, G., Singer, J., Meltzer, H., 1988. Clozapine for the treatment resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives of General Psychiatry* 45, 789–796.
- Karayiorgou, M., Rottington, M., Abecasis, G.R., Pretorius, H.W., Robertson, B., Kaliski, S., Lay, S., Sobin, C., Möller, N., Lundy, S.L., Blundell, M.L., Gogos, J.A., Roos, J.L., 2004. Phenotypic characterization and genealogical tracing in an Afrikaner Schizophrenia Database. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 124B, 20–28.
- Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., Wong, W.S., Sigurdsson, G., Walters, G.B., Steinberg, S., Helgason, H., Thorleifsson, G., Gudbjartsson, D.F., Helgason, A., Magnusson, O.T., Thorsteinsdottir, U., Stefansson, K., 2012. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 488, 471–475.
- McClellan, J., King, M.C., 2010. Genetic heterogeneity in human disease 141, 210–217. *Cell* 141, 210–217.
- Nurnberger, J.I., Blehar, M.C., Kaufmann, C., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., Severe, J.B., Malaepina, D., Reich, T., 1994. Diagnostic interview for genetic studies. Rationale, unique features and training. NIMH Genetics Initiative. *Archives of General Psychiatry* 51, 849–859.
- Rodriguez-Murillo, L., Gogos, J.A., Karayiorgou, M., 2012. The genetic architecture of schizophrenia: new mutations and emerging paradigms. *Annual Review of Medicine* 63, 63–80.
- Roos, J.L., Pretorius, H.W., Karayiorgou, M., 2009. Clinical characteristics of an Afrikaner founder population recruited for a schizophrenia genetic study. *Annals of the New York Academy of Sciences* 1151, 85–101.
- Schneider, L.C., Struening, E.L., 1983. SLOF: a behavioural rating scale for assessing the mentally ill. *Social Work Research and Abstracts* 19, 9–21.
- Sobin, C., Blundell, M.L., Conry, A., Weiller, F., Gavigan, C., Haiman, C., Karayiorgou, M., 2001. Early non-psychotic deviant behaviour in schizophrenia: a possible endophenotypic marker for genetic studies. *Psychiatry Research* 101, 101–113.
- Sobin, C., Roos, J.L., Pretorius, H.W., Lundy, L.S., Karayiorgou, M., 2003. A comparison study of early non-psychotic deviant behaviour in Afrikaner and US patients with schizophrenia and schizoaffective disorder. *Psychiatry Research* 117, 113–125.
- Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E., Hansen, T., Jakobsen, K.D., Muglia, P., Francks, C., Matthews, P.M., Gylfason, A., Halldorsson, B.V., Gudbjartsson, D., Thorgeirsson, T.E., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., Bjornsson, A., Mattiasdottir, S., Blondal, T., Haraldsson, M., Magnusdottir, B.B., Giegling, I., Möller, H.J., Hartmann, A., Shianna, K.V., Ge, D., Need, A.C., Crombie, C., Fraser, G., Walker, N., Lonnqvist, J., Suvisaari, J., Tuulio-Henriksson, A., Paunio, T., Touloupoulou, T., Bramon, E., Di Forti, M., Murray, R., Ruggeri, M., Vassos, E., Tosato, S., Walshe, M., Li, T., Vasilescu, C., Mühleisen, T.W., Wang, A.G., Ullum, H., Djurovic, S., Melle, I., Olesen, J., Kiemeny, L.A., Franke, B., GROUPSabbati, C., Freimer, N.B., Gulcher, J.R., Thorsteinsdottir, U., Kong, A., Andreassen, O.E., Ophoff, R.A., Georgi, A., Rietschel, M., Werge, T., Petursson, H., Goldstein, D.B., N.B.st, M.M., Peltonen, L., Collier, D.A., St Clair, D., Stefansson, K., 2008. Large recurrent microdeletions associated with schizophrenia. *Nature* 445, 232–236.
- Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A., Stray, S.M., Rippey, C.F., Roccanova, P., Makarov, V., Lakshmi, B., Findling, R.L., Sikich, L., Stromberg, T., Merriman, B., Gogtay, N., Butler, P., Eckstrand, K., Noory, L., Gochman, P., Long, R., Chen, Z., Davis, S., Baker, C., Eichler, E.E., Meltzer, P.S., Nelson, S.F., Singleton, A.B., Lee, M.K., Rapoport, J.L., King, M.C., Sebat, J., 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320, 539–543.
- Wiersma, D., Wanderling, J., Dragomirecka, E., Ganey, K., Harrison, G., An Der Heiden, W., Nienhuis, F.J., Walsh, D., 2000. Social disability in schizophrenia: its development and prediction over 15 years in incidence cohorts in six European centres. *Psychological Medicine* 30, 1155–1167.
- Xu, B., Roos, J.L., Levy, S., Van Rensburg, E.J., Gogos, J.A., Karayiorgou, M., 2008. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature Genetics* 40, 880–885.
- Xu, B., Woodroffe, A., Rodriguez-Murillo, L., Roos, J.L., Van Rensburg, E.J., Abecasis, G.R., Gogos, J.A., Karayiorgou, M., 2009. Elucidating the genetic architecture of familial schizophrenia using rare copy number variant and linkage scans. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16746–16751.
- Xu, B., Roos, J.L., Dexeimer, P., Boane, B., Plummer, B., Levy, S., Gogos, J.A., Karayiorgou, M., 2011. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nature Genetics* 43, 864–868.
- Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., Karayiorgou, M., 2012. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nature Genetics* 44, 1365–1369.