Original Article

Failure to Identify Somatic Mutations in Monozygotic Twins Discordant for Schizophrenia by Whole Exome Sequencing

Nan Lyu[1], Li-Li Guan[1], Hong Ma[1], Xi-Jin Wang[1,2], Bao-Ming Wu[1], Fan-Hong Shang[1], Dan Wang[3], Hong Wen[3], Xin Yu[1]

[1]Peking University Sixth Hospital, Peking University Institute of Mental Health, Key Laboratory of Mental Health, Ministry of Health (Peking University), Beijing 100191, China

[2]The First Psychiatric Hospital of Harbin, Harbin, Heilongjiang 150056, China

[3]The Mental Health Centre of Sichuan Province, Mianyang, Sichuan 621000, China

Abstract

Background: Schizophrenia (SCZ) is a severe, debilitating, and complex psychiatric disorder with multiple causative factors. An increasing number of studies have determined that rare variations play an important role in its etiology. A somatic mutation is a rare form of genetic variation that occurs at an early stage of embryonic development and is thought to contribute substantially to the development of SCZ. The aim of the study was to explore the novel pathogenic somatic single nucleotide variations (SNVs) and somatic insertions and deletions (indels) of SCZ.

Methods: One Chinese family with a monozygotic (MZ) twin pair discordant for SCZ was included. Whole exome sequencing was performed in the co-twin and their parents. Rigorous filtering processes were conducted to prioritize pathogenic somatic variations, and all identified SNVs and indels were further confirmed by Sanger sequencing.

Results: One somatic SNV and two somatic indels were identified after rigorous selection processes. However, none was validated by Sanger sequencing.

Conclusions: This study is not alone in the failure to identify pathogenic somatic variations in MZ twins, suggesting that exonic somatic variations are extremely rare. Further efforts are warranted to explore the potential genetic mechanism of SCZ.

Key words: Monozygotic Twins; Schizophrenia; Somatic Mutation; Whole Exome Sequencing

Introduction

Schizophrenia (SCZ) is a severe, debilitating, and complex psychiatric disorder that contributes substantially to the global disease burden. As a multifactorial disease, a variety of genetic and environmental factors play roles in the etiology. Genome-wide association studies have identified a number of common risk loci, but all common polygenic variations combined account for only 32–36% of the genetic risk for SCZ.[1] It is thought that the so-called "missing heritability" can in part be explained by rare variations of large effect size.[2] Next generation sequencing technology enables the identification of novel rare susceptibility loci at the whole genome and exome levels. In recent years, researchers have characterized a polygenic burden arising primarily from rare, disruptive mutations distributed across many genes enriched in discrete biological processes[3-5] such as synaptic network function.

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.4103/0366-6999.178009

Discordant monozygotic (MZ) twins provide a unique window into the genetic factors underlying the disease phenotype. Several studies have shown that phenotypic discordance between MZ twins can be attributed to rare sequences differences caused by *de novo* events such as chromosomal mosaicism (e.g., the discordant karyotypes mos 47, XX, +21/46, XX and 46, XX),[6] differing copy number variations (CNVs)[7,8] and distinct single nucleotides.[9] The genetic differences between MZ twins can

Address for correspondence: Dr. Li-Li Guan, Peking University Sixth Hospital, Peking University Institute of Mental Health, Key Laboratory of Mental Health, Ministry of Health (Peking University), Beijing 100191, China
E-Mail: guanlili@bjmu.edu.cn

Received: 26-11-2015 **Edited by:** Qiang Shi

How to cite this article: Lyu N, Guan LL, Ma H, Wang XJ, Wu BM, Shang FH, Wang D, Wen H, Yu X. Failure to Identify Somatic Mutations in Monozygotic Twins Discordant for Schizophrenia by Whole Exome Sequencing. Chin Med J 2016;129:690-5.

be regarded as an extreme form of somatic mosaicism[10,11] resulting from postzygotic mutations that occurred at an early stage of embryonic development and grew into two populations of cells with genotypes different from that of the single fertilized egg.[10] The genetic comparison of discordant MZ twins represents a promising paradigm for identifying rare novel candidates that may help to account for some or all missing heritability.

Discordant MZ twin pairs have been used previously to investigate the potential mechanism of SCZ. Many studies have focused on differences in DNA methylation between discordant MZ twins,[9,12-16] and series of genes have been identified to play a role in gene networks relevant in the context of SCZ. However, individual differences in DNA methylation are correlated with DNA sequence variations.[17] The genes that affected by differential methylation between discordant twins also harbor various types of discordant sequence variations that should not be overlooked.[9] CNVs, a predominant form of DNA structural variation, have been the subject of recent focus. Maiti *et al*.[18] observed somatic deletions (14q32.11, 8q11.21) and duplication (19q13.41) in SCZ-affected twins, and Castellani *et al*.[19] observed an additional novel somatic deletion (7q11.21). Nevertheless, some studies have failed to confirm somatic CNVs in discordant twin pairs,[20,21] suggesting that there may be other underlying genetic causes of the SCZ discordant phenotype including single nucleotide variations (SNVs) and short insertions and deletions (indels), which are undetectable by the arrays used to detect CNVs.

In a recent study, Castellani *et al*.[9] performed whole genome sequencing together with genome-wide methylation analysis without the validation step of Sanger sequencing in two families with MZ twins discordant for SCZ. A total of 71 genes harboring somatic SNVs and indels were identified in affected twins. Nine missense mutations were observed in exonic regions. Exonic regions containing about 85% of known disease-related variants[22] are valuable to be explored. To the best of our knowledge, the aforementioned study is the only one to examine differences in somatic SNV and indel profiles in MZ twins discordant for SCZ. Evidence from further studies is needed. Therefore, we conducted high depth exome sequencing of a family with an MZ twin pair discordant for SCZ to explore possible relevant somatic mutations.

Methods

Ethical statement

This study was approved by the Ethics Committee of the Institute of Mental Health, the Sixth Hospital, Peking University. All subjects understood the purpose and procedure of the study and provided informed written consents prior to enrollment.

Subjects

We sequenced the exomes of one family with a pair of twins discordant for SCZ [Figure 1] from the Sichuan Province of China. We included the elder sister of the twins in the study to get more genetic information. All family members were of Chinese Han descent. The demographic and clinical data of the family are shown in Table 1. The MZ twins were 25-year-old males exhibiting highly similar physical characteristics, including height, weight, pupil color, hair color, and hair texture, and could not be easily distinguished by individuals other than their parents. Twin 1 was healthy and without mental health problems. Twin 2 had been diagnosed with SCZ and major depressive disorder. He had experienced delusions of reference and persecution, auditory hallucinations, and thought broadcasting since the age of 19 years. Then, these symptoms gradually lessened with regular medication. He began to feel depressed at the age of 24 years, characterized by diminished interest in his usual activities, a lack of energy, poor concentration, and insomnia. The depressive symptoms lasted for 6 months and then remitted. The parents and the elder sister of the twins were all unaffected. There was no family history or personal history of severe psychiatric or neurological illness.

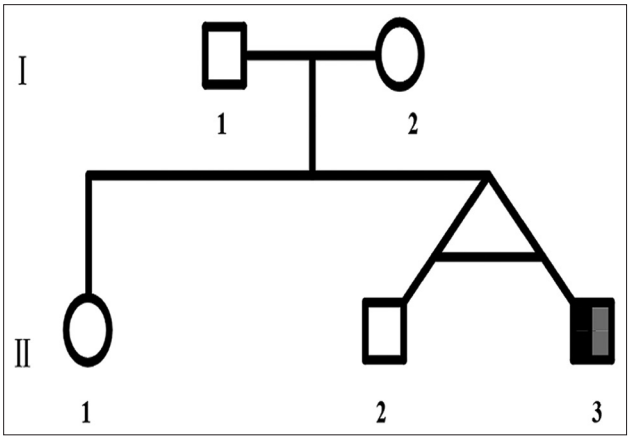

Figure 1: Pedigree chart of a family with discordant monozygotic twins. Twin 2 (II-3) was diagnosed with schizophrenia and major depressive disorder. Shaded and unshaded symbols indicate affected and unaffected individuals, respectively. Squares and circles represent males and females, respectively.

Diagnosis and assessment

All family members completed the Structured Clinical Interview from the Diagnostic and Statistical Manual-IV-TR Axis I Disorders-Patient.[23] Diagnoses were made by an experienced psychiatrist according to Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision criteria; and were confirmed by one senior psychiatrist.

We used three scales, including the Positive and Negative Syndrome Scale,[24] the Scale for the Assessment of Positive Symptoms,[25] and the Scale for the Assessment of Negative Symptoms[26] to rate symptom severity. Overall individual and social functioning were assessed by the Personal and Social Performance scale[27] and the Global Assessment of Functioning.[28,29]

The Wechsler Intelligence Scale for Adult-Chinese Revised[30] was used to estimate intelligence quotient. A standardized

Table 1: Summary of demographic and clinical family data

Individual	Age (years)	Onset age (years)	Diagnosis (SCID-I/P)	Education (years)	PANSS	SAPS	SANS	PSP	GAF	IQW
I-1	53	–	Normal	9	–	–	–	–	85	90
I-2	50	–	Normal	9	–	–	–	–	81	81
II-1	31	–	Normal	9	–	–	–	–	85	98
II-2	25	–	Normal	9	–	–	–	–	85	99
II-3	25	19	SCZ, MD	9	49*	4†	19‡	68	61	107

*The score of PANSS is composed of three components: positive score (P=8), negative score (N=11), and general psychopathology score (G=30); †the items related to apathy, slow response, poverty of thought, and lack of social contact were scored in SANS; ‡the items related to delusion were scored in SAPS. SCID-I/P: Structured Clinical Interview from the Diagnostic and Statistical Manual-IV-TR Axis I Disorders-Patient; PANSS: Positive and Negative Syndrome Scale; SAPS: Scale for the Assessment of Positive Symptoms; SANS: Scale for the Assessment of Negative Symptoms; PSP: Personal and Social Performance scale; GAF: Global Assessment of Functioning; IQW: Intelligence quotient estimated by Wechsler Intelligence Scale for Adult-Chinese Revised; SCZ: Schizophrenia; MD: Major depressive disorder; –: Not available.

questionnaire was designed to collect demographic data and detailed information regarding the childhood environment, developmental history, and family history of the individuals.

Genomic DNA extraction and exome capture

Genomic DNA was extracted from leukocytes obtained from peripheral blood using standard methods. Whole exome sequencing was performed for four members of the family (I-1, I-2, II-2, and II-3). Targeted enrichment was conducted using the NimbleGen SeqCap EZ Exome + Untranslated Regions (UTR) Library (Roche NimbleGen, Madison, WI, USA), and exon-enriched DNA was sequenced on the Illumina HiSeq2500 sequencing platform (Illumina, San Diego, USA) according to the manufacturer's instructions to obtain 125 bp paired-end reads.

Read mapping and variant calling

FASTQ files (i.e., the raw data) were filtered to remove low-quality reads, and the remaining high-quality reads were mapped onto a human reference genome (hg19) using the Burrows–Wheeler Aligner version 0.6.2 (http://sourceforge.net/projets/bio-bwa/). Variants such as SNVs and indels were called using a pipeline according to the Genome Analysis Toolkit version 3.1 software (https://www.broadinstitute.org/gatk/). Parameters of VariantFiltration for SNVs were set as follows: QD <2.0, MQ <40.0, FS >60.0, HaplotypeScore >13.0, MQRankSum < −12.5, and ReadPosRankSum < −8.0. Those for indels were set as follows: QD <2.0, FS >200.0, and ReadPosRankSum < −20.0.

Variant annotation and prioritization

ANNOVAR software (http://annovar.openbioinformatics.org/en/latest/)[31] was used to annotate the variants. Sorting intolerant from tolerant (SIFT) (http://sift.jcvi.org/) and polymorphism phenotyping 2 (PolyPhen2) (http://genetics.bwh.harvard.edu/pph 2/) were used to assess the pathogenicity of protein-altering variants. Evolutionary conservation was evaluated using genomic evolutionary rate profiling (GERP++) (http://mendel.stanford.edu/SidowLab/downloads/gerp/).

Several rigorous filtering steps were performed to prioritize SNVs and indels: (1) variants that were present in the affected twin and that were unshared by the healthy twin and the parents were included; (2) the variants exhibited more than 20 × coverage; (3) variants for which the percentage of reads supporting the newly identified allele (defined as mutation frequency) was >20%, with at least three reads supporting the mutation, were chosen;[5] (4) variants in splicing regions and protein-coding regions, such as missense, frameshift, stop-loss and stop-gain mutations were included; (5) reserve variants with a minor allele frequency ≤1% in a public database such as dbSNP 138, 1000 Genomes Project Phase 3 (released May 2013), the National Heart, Lung, and Blood Institute Exome Sequencing Project (ESP 6500), or the Complete Genomics CG46 database were chosen; (6) select variants were predicted to be damaging if they had a SIFT[32] score <0.05 and a PolyPhen2[33] score >0.85; and (7) conservative variants (GERP++[34] score >2.0) were selected.

Variant validation

Sanger sequencing was performed on an ABI 3730XL Genetic Analyzer (PE Applied Biosystems, Forest City, CA, USA) for four members of the family (I-1, I-2, II-2, and II-3) to validate the variants prioritized by the filtering steps. Forward and reverse primer sequences for the candidate loci are listed in Table 2.

Results

A quality summary for whole exome sequencing is provided in Table 3. An average of 96.9% of the target regions (96.5 Mb) were captured in each exome, and the mean depth of the target region was 136×. In addition, 94.9% of the captured target exons were covered by ten or more reads. A total of 580,849 variants (511,951 SNVs and 68,898 indels) were called in the family, of which 84,545 variants (83,131 SNVs and 1414 indels) were in exonic regions and 18 variants (14 SNVs and 4 indels) were in splicing regions.

There were 127,762 and 127,186 SNVs present in each twin, respectively, and 124,185 (97.4%) of the SNVs were shared between the twins, supporting their monozygosity. Twin 1 and twin 2 shared 79.7% of the total SNPs with their father and 80.0% with their mother, providing proof of the biological relationship of the family.

Three candidate somatic variations passed our rigorous filtering steps [Figure 2], including 1 SNV in *CEP57*, 1

Table 2: List of candidate variants in the affected twin after filtering

Gene	Chromosome	Type	Start position (hg19)	Location region	Reference allele	Mutant allele	Primer sequences
CEP57	11	SNV	95528678	Splicing	A	T	F:5'-TCTTCAGTAGAAATTCTGGAACCT-3' R: 5'-TGGAGGAACATGATGTACAGCTC-3'
CTAGE5	14	Deletion	39784005	Splicing	AT	–	F: 5'-CGTTCCATTTAGAAAGCGAGCC-3' R: 5'-GCCACATGCAACACCTCAAA-3'
ORC2	2	Insertion	201778717	Splicing	–	A	F: 5'-CCTTGCTCAGCAGTGGATCT-3' R: 5'-AATGCGGATAGGCCAACTAC-3'

SNV: Single nucleotide variation; F: Forward primer; R: Reverse primer.

Table 3: Whole exome sequencing quality report data

Items	I-1	I-2	II-2	II-3
Target regions (Mb)	96.5	96.5	96.5	96.5
Map rate (%)	98.6	99.0	98.8	98.6
Coverage of target regions (%)	97.0	96.9	96.9	96.9
Average read depth (×)	146.4	144.6	121.2	132.0
Percentage of target regions with coverage >10× (%)	95.1	94.9	94.8	94.8
Percentage of target regions with coverage >20× (%)	93.6	93.2	93.0	93.0
Percentage of target regions with coverage >30× (%)	92.0	91.5	91.0	91.0

Mb: Million bases.

deletion in *CTAGE5*, and 1 insertion in *ORC2* [Table 2]. Unfortunately, none of the variants was confirmed by Sanger DNA sequence analysis, as shown in Figure 3.

Discussion

This study pioneers the application of a family with MZ twins discordant for SCZ to explore pathogenic SNVs and indels using whole exome sequencing. Sanger sequencing was performed in our study to confirm all candidates. One somatic SNV and 2 somatic indels with the potential to be pathogenic in the affected twin were identified. However, none of these variations was validated via Sanger sequencing.

Although one previously reported study has been successful in the identification of somatic variations in MZ twins discordant for SCZ using high-throughput sequencing,[9] we are not alone in the failure to identify convincing somatic mutations that could be responsible for the discordant phenotypes between MZ twins. Many studies using a similar design have likewise failed. For instance, whole genome or exome sequencing studies in MZ twins discordant for Crohn's disease,[35] multiple sclerosis,[36] amyotrophic lateral sclerosis,[37] congenital heart defect,[38] and congenital hypothyroidism[39] have failed to identify significant variants. These negative results support the notion that somatic variants between MZ twins are extremely rare.

The rate of early postzygotic mutation is exceptionally low.[40] In a recent study, only 1 and 8 somatic SNVs were identified in two pairs of healthy MZ twins, respectively, at the whole genomic level, and none was located in coding regions. The rate of early postzygotic SNVs is estimated to be 0.04×10^{-9} and 0.34×10^{-9} per generation per position[40] among two twin pairs, which equates to one-thirtieth and one-third of the gremlin mutation rate. Thus, the occurrence of somatic *de novo* mutation is so rare that it is not easily observed, especially in exonic regions, which account for only 1–2% of the genome.

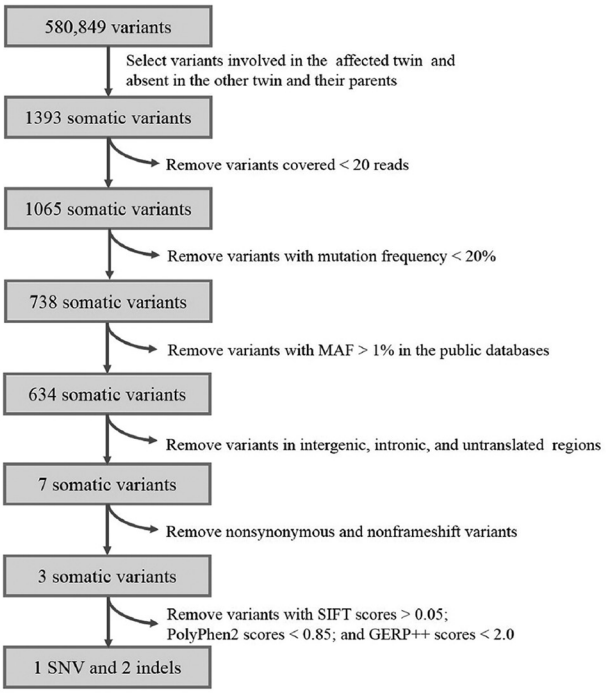

Figure 2: Illustration of the variant reduction procedure. Three *de novo* candidates were identified. MAF: Minor allele frequency; SIFT: Sorting intolerant from tolerant; PolyPhen: Polymorphism phenotyping; GERP: Genomic evolutionary rate profiling; SNV: Single nucleotide variation.

Given that definitive differences of SNVs and indels to explain the SCZ discordance were absent in our twin pair, other possibilities should be considered. First, rare and large *de novo* CNVs have emerged as an important genomic factor in psychiatric disease and occur eight times more frequently in patients with sporadic SCZ;[41,42] we could not exclude the possibility that the co-twins carry different CNV profiles. Epigenetic differences could also explain the discordance in our twin pair. Methylation differences between identical twins have been demonstrated as early as the newborn stage[43] and accumulate with age,[44] possibly resulting in phenotypic differences. Integration DNA sequencing and other

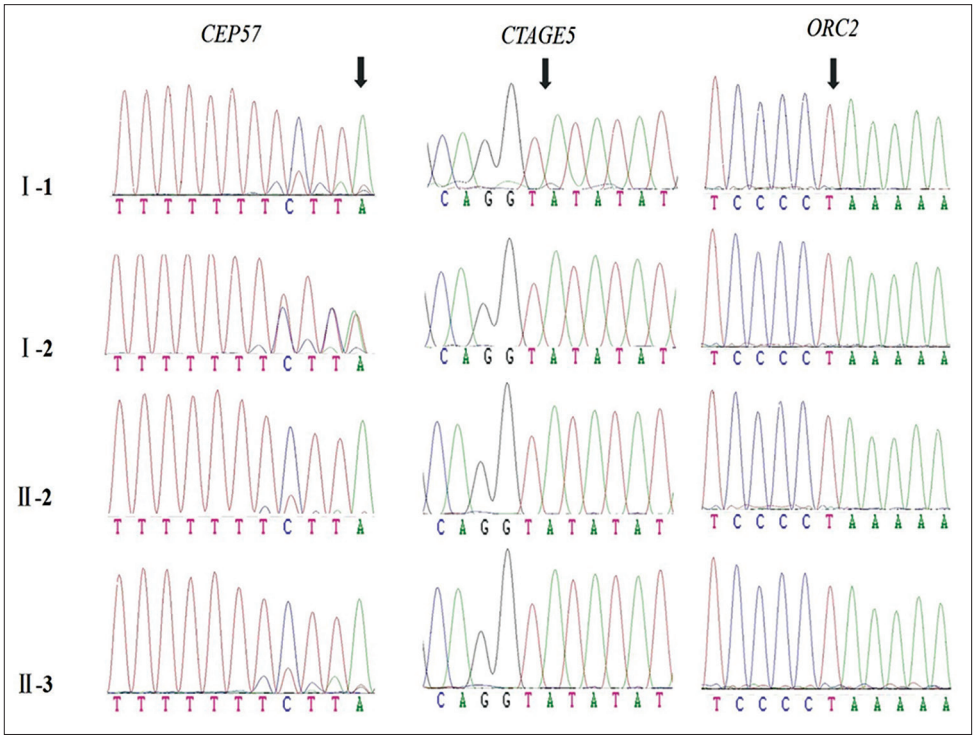

Figure 3: Validation of the mutations by Sanger sequencing of the family. We failed to confirm 1 SNV and 2 indels. SNV: Single nucleotide variation; Indels: Insertions and deletions.

technologies, such as CNV and epigenetic analysis, to explore the potential mechanisms of SCZ should be considered. Third, the tissue-specific somatic variation could explain discordance. Somatic variations affect various cell types and tissues depending on the time at which the mutation occurs.[45] In our twin pair, we did not detect somatic variations in leukocytes, but this does not mean that other tissues, such as the central nervous system (CNS) and other tissues originating from the ectodermal layer, do not harbor pathogenic somatic mutations. Thus, the detection of brain tissue or hair follicle (which is of ectodermal origin) should be considered in the future.

Experimental limitations should also be considered. Whole exome sequencing primarily detects protein coding regions and a small fraction of sequences adjacent to coding exons. Thus, most noncoding regions, such as intergenic, intronic, and 3'- or 5'-UTR, are not covered. It is possible that genetic differences might exist in regions outside the assessable coding sequences.

In conclusion, we did not identify convincing pathogenic SNVs and indels in coding regions by whole exome sequencing in the family with MZ twins discordant for SCZ. This finding suggests that the pathogenic genetic difference between MZ twins is too rare to be detected. Further efforts are needed to investigate the potential mechanisms of SCZ, such as noncoding DNA, CNS-specific somatic variations, and different profiles of CNVs and epigenetics.

Financial support and sponsorship

This work was supported by a grant from the National Natural Science Foundation of China (No. 81201039).

Conflicts of interest

There are no conflicts of interest.

References

1. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, *et al*. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009;460:748-52. doi: 10.1038/nature08185.
2. Rodriguez-Murillo L, Gogos JA, Karayiorgou M. The genetic architecture of schizophrenia: New mutations and emerging paradigms. Annu Rev Med 2012;63:63-80. doi: 10.1146/annurev-med-072010-091100.
3. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, *et al*. A polygenic burden of rare disruptive mutations in schizophrenia. Nature 2014;506:185-90. doi: 10.1038/nature12975.
4. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, *et al*. *De novo* mutations in schizophrenia implicate synaptic networks. Nature 2014;506:179-84. doi: 10.1038/nature12929.
5. Girard SL, Gauthier J, Noreau A, Xiong L, Zhou S, Jouan L, *et al*. Increased exonic *de novo* mutation rate in individuals with schizophrenia. Nat Genet 2011;43:860-3. doi: 10.1038/ng.886.
6. Choi SA, Ko JM, Shin CH, Yang SW, Choi JS, Oh SK. Monozygotic twin discordant for Down syndrome: Mos 47, XX, 21/46, XX and 46, XX. Eur J Pediatr 2013;172:1117-20. doi: 10.1007/s00431-013-1998-3.
7. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, Diaz de Ståhl T, *et al*. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. Am J Hum Genet 2008;82:763-71. doi: 10.1016/j.ajhg.2007.12.011.
8. Breckpot J, Thienpont B, Gewillig M, Allegaert K, Vermeesch JR, Devriendt K. Differences in copy number variation between discordant monozygotic twins as a model for exploring chromosomal mosaicism in congenital heart defects. Mol Syndromol 2012;2:81-7. doi: 10.1159/000335284.

9. Castellani CA, Melka MG, Gui JL, O'Reilly RL, Singh SM. Integration of DNA sequence and DNA methylation changes in monozygotic twin pairs discordant for schizophrenia. Schizophr Res 2015;169:433-40. doi: 10.1016/j.schres.2015.09.021.
10. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. Nat Rev Genet 2002;3:748-58. doi: 10.1038/nrg906.
11. Silva S, Martins Y, Matias A, Blickstein I. Why are monozygotic twins different? J Perinat Med 2011;39:195-202. doi: 10.1515/jpm.2010.140.
12. Melka MG, Castellani CA, O'Reilly R, Singh SM. Insights into the origin of DNA methylation differences between monozygotic twins discordant for schizophrenia. J Mol Psychiatry 2015;3:7. doi: 10.1186/s40303-015-0013-5.
13. Castellani CA, Laufer BI, Melka MG, Diehl EJ, O'Reilly RL, Singh SM, *et al.* DNA methylation differences in monozygotic twin pairs discordant for schizophrenia identifies psychosis related genes and networks. BMC Med Genomics 2015;8:17. doi: 10.1186/s12920-015-0093-1.
14. Petronis A, Gottesman II, Kan P, Kennedy JL, Basile VS, Paterson AD, *et al.* Monozygotic twins exhibit numerous epigenetic differences: Clues to twin discordance? Schizophr Bull 2003;29:169-78. doi: 10.1093/oxfordjournals.schbul.a006988.
15. Dempster EL, Pidsley R, Schalkwyk LC, Owens S, Georgiades A, Kane F, *et al.* Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. Hum Mol Genet 2011;20:4786-96. doi: 10.1093/hmg/ddr416.
16. Kato T, Iwamoto K, Kakiuchi C, Kuratomi G, Okazaki Y. Genetic or epigenetic difference causing discordance between monozygotic twins as a clue to molecular basis of mental disorders. Mol Psychiatry 2005;10:622-30. doi: 10.1038/sj.mp. 4001662.
17. Teh AL, Pan H, Chen L, Ong ML, Dogra S, Wong J, *et al.* The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. Genome Res 2014;24:1064-74. doi: 10.1101/gr.171439.113.
18. Maiti S, Kumar KH, Castellani CA, O'Reilly R, Singh SM. Ontogenetic *de novo* copy number variations (CNVs) as a source of genetic individuality: Studies on two families with MZD twins for schizophrenia. PLoS One 2011;6:e17125. doi: 10.1371/journal.pone.0017125.
19. Castellani CA, Awamleh Z, Melka MG, O'Reilly RL, Singh SM. Copy number variation distribution in six monozygotic twin pairs discordant for schizophrenia. Twin Res Hum Genet 2014;17:108-20. doi: 10.1017/thg.2014.6.
20. Ono S, Imamura A, Tasaki S, Kurotaki N, Ozawa H, Yoshiura K, *et al.* Failure to confirm CNVs as of aetiological significance in twin pairs discordant for schizophrenia. Twin Res Hum Genet 2010;13:455-60. doi: 10.1375/twin.13.5.455.
21. Bloom RJ, Kähler AK, Collins AL, Chen G, Cannon TD, Hultman C, *et al.* Comprehensive analysis of copy number variation in monozygotic twins discordant for bipolar disorder or schizophrenia. Schizophr Res 2013;146:289-90. doi: 10.1016/j.schres.2013.02.021.
22. van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. Trends Genet 2014;30:418-26. doi: 10.1016/j.tig.2014.07.001.
23. First MB, Spitzer RL, Williams JB. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P). New York: Biometrics Research, New York State: Psychiatric Institute; 2002.
24. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13:261-76. doi: 10.1093/schbul/13.2.261
25. Andreasen NC. Scale for the Assessment of Positive Symptoms (SAPS). Iowa City:University of Iowa; 1984.
26. Andreasen NC. Scale for the Assessment of Negative Symptoms (SANS). Iowa City: University of Iowa; 1984.
27. Tianmei S, Liang S, Yun'ai S, Chenghua T, Jun Y, Jia C, *et al.* The Chinese version of the personal and social performance scale (PSP): Validity and reliability. Psychiatry Res 2011;185:275-9. doi: 10.1016/j.psychres.2010.05.001.
28. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976;33:766-71. doi:10.1001/archpsyc.1976.01770060086012.
29. Startup M, Jackson MC, Bendix S. The concurrent validity of the Global Assessment of Functioning (GAF). Br J Clin Psychol 2002;41(Pt 4):417-22. doi: 10.1348/014466502760387533.
30. Yaoxian G. Manual for the Wechsler Adult Intelligence Scale: Chinese Revision, Changsha Hunan. China: Hunan Medical College; 1982.
31. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164. doi: 10.1093/nar/gkq603.
32. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812-4. doi: 10.1093/nar/gkg509.
33. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, *et al.* A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9. doi: 10.1038/nmeth0410-248.
34. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol 2010;6:e1001025. doi: 10.1371/journal.pcbi.1001025.
35. Petersen BS, Spehlmann ME, Raedler A, Stade B, Thomsen I, Rabionet R, *et al.* Whole genome and exome sequencing of monozygotic twins discordant for Crohn's disease. BMC Genomics 2014;15:564. doi: 10.1186/1471-2164-15-564.
36. Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, Miller NA, *et al.* Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 2010;464:1351-6. doi: 10.1038/nature08990.
37. Meltz Steinberg K, Nicholas TJ, Koboldt DC, Yu B, Mardis E, Pamphlett R. Whole genome analyses reveal no pathogenetic single nucleotide or structural differences between monozygotic twins discordant for amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 2015;16:385-92. doi: 10.3109/21678421.2015.1040029.
38. Chaiyasap P, Kulawonganunchai S, Srichomthong C, Tongsima S, Suphapeetiporn K, Shotelersuk V. Whole genome and exome sequencing of monozygotic twins with trisomy 21, discordant for a congenital heart defect and epilepsy. PLoS One 2014;9:e100191. doi: 10.1371/journal.pone.0100191.
39. Magne F, Serpa R, Van Vliet G, Samuels ME, Deladoëy J. Somatic mutations are not observed by exome sequencing of lymphocyte DNA from monozygotic twins discordant for congenital hypothyroidism due to thyroid dysgenesis. Horm Res Paediatr 2015;83:79-85. doi: 10.1159/000365393.
40. Dal GM, Ergüner B, Sagiroglu MS, Yüksel B, Onat OE, Alkan C, *et al.* Early postzygotic mutations contribute to *de novo* variation in a healthy monozygotic twin pair. J Med Genet 2014;51:455-9. doi: 10.1136/jmedgenet-2013-102197.
41. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of *de novo* copy number mutations with sporadic schizophrenia. Nat Genet 2008;40:880-5. doi: 10.1038/ng.162.
42. Sebat J, Levy DL, McCarthy SE. Rare structural variants in schizophrenia: One disorder, multiple mutations; one mutation, multiple disorders. Trends Genet 2009;25:528-35. doi: 10.1016/j.tig.2009.10.004.
43. Ollikainen M, Smith KR, Joo EJ, Ng HK, Andronikos R, Novakovic B, *et al.* DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. Hum Mol Genet 2010;19:4176-88. doi: 10.1093/hmg/ddq336.
44. Lévesque ML, Casey KF, Szyf M, Ismaylova E, Ly V, Verner MP, *et al.* Genome-wide DNA methylation variability in adolescent monozygotic twins followed since birth. Epigenetics 2014;9:1410-21. doi: 10.4161/15592294.2014.970060.
45. Freed D, Stevens EL, Pevsner J. Somatic mosaicism in the human genome. Genes (Basel) 2014;5:1064-94. doi: 10.3390/genes5041064.