CAMBRIDGE
UNIVERSITY PRESS

Article

Mitochondrial DNA Copy Number and Heteroplasmy in Monozygotic Twins Discordant for Schizophrenia

Phyo W. Win[1], Shiva M. Singh[2] and Christina A. Castellani[1,3,4]

[1]Department of Pathology and Laboratory Medicine, Schulich School of Medicine and Dentistry, Western University, London, Canada, [2]Department of Biology, Western University, London, Canada, [3]Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, Western University, London, Canada and [4]McKusick-Nathans Institute, Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Abstract

Schizophrenia (SZ) is a severe, complex, and common mental disorder with high heritability (80%), an adult age of onset, and high discordance (~50%) in monozygotic twins (MZ). Extensive studies on familial and non-familial cases have implicated a number of segregating mutations and de novo changes in SZ that may include changes to the mitochondrial genome. Yet, no single universally causal variant has been identified, highlighting its extensive genetic heterogeneity. This report specifically focuses on the assessment of changes in the mitochondrial genome in a unique set of monozygotic twins discordant (MZD) for SZ using blood. Genomic DNA from six pairs of MZD twins and two sets of parents ($N = 16$) was hybridized to the Affymetrix Human SNP Array 6.0 to assess mitochondrial DNA copy number (mtDNA-CN). Whole genome sequencing (WGS) and quantitative polymerase chain reaction (qPCR) was performed for a subset of MZD pairs and their parents and was also used to derive mtDNA-CN estimates. The WGS data were further analyzed to generate heteroplasmy (HP) estimates. Our results show that mtDNA-CN estimates for within-pair and mother-child differences were smaller than comparisons involving unrelated individuals, as expected. MZD twins showed discordance in mtDNA-CN estimates and displayed concordance in directionality of differences for mtDNA-CN across all technologies. Further, qPCR performed better than Affymetrix in estimating mtDNA-CN based on relatedness. No reliable differences in HP were detected between MZD twins. The within-MZD differences in mtDNA-CN observed represent postzygotic somatic changes that may contribute to discordance of MZ twins for diseases, including SZ.

Keywords: Monozygotic twins; mitochondrial DNA copy number; heteroplasmy; schizophrenia

Transmission of genetic information is carried out by two sets of genomes, nuclear and mitochondrial. The majority of genetic information that is contained in the nucleus follows a Mendelian pattern of inheritance. In contrast, the relatively small (16.5 kbp), circular, double-stranded DNA (mtDNA) that functions as the mitochondrial genome is primarily maternally transmitted and randomly distributed into daughter cells during cell division. mtDNA encodes for 13 peptide subunits, 22 tRNAs, and 2 rRNAs that are essential for mitochondrial protein synthesis and partake in several essential functions. These include energy production, cell signaling, calcium homeostasis, reactive oxygen species production, and apoptosis (Kujoth et al., 2006). Ultimately, they may affect energy reserves, mortality, development and aging, and contribute to disease risk.

The mitochondrial genome is packaged as a double-membraned organelle in the cytoplasm, an individual mitochondrion may carry multiple copies of mtDNA depending on the cell type, which can range on average from 10s to 1000s (Filograna et al.,

2021), and a cell normally contains multiple mitochondria. What determines this organization is not fully understood. What is known, however, is that variable copies of the mitochondrial genome (mtDNA-CN) in a cell, together with any mtDNA mutation(s) (heteroplasmy), affect a cell's overall function (Wallace & Chalkia, 2013). To this end, mtDNA-CN establishes the *quantity* of mitochondrial DNA, while heteroplasmy (HP), which captures the *quality,* refers to the ratio of mutated to normal mtDNA, and represents a crucial threshold, which may lead to pathological outcomes (Kopinski et al., 2019; Stefano & Kream, 2016). Specifically, mtDNA-CN has been shown to have a negative correlation with age and is higher on average in women than men (Ashar et al., 2015; Mengel-From et al., 2014). On the other hand, heteroplasmy increases with age, which reflects mutational burden (López-Otín et al., 2013; Mengel-From et al., 2014; Zhang et al., 2017). Interestingly, mtDNA-CN and heteroplasmy have been shown to be associated with a large number of complex multifactorial diseases (Ashar et al., 2015; Ashar et al., 2017; Hong et al., 2020). There is emerging interest in assessing the involvement of mitochondria in an ever-increasing number of complex diseases, including schizophrenia.

Schizophrenia (SZ) is a severe mental disorder with psychotic symptoms that include delusions and social withdrawal. Psychosis

does not usually start until late adolescence and is associated with a reduction of 15−30 years in life expectancy (Olfson et al., 2015; Thornicroft, 2011). SZ, a complex disease, has high heritability (∼80%), and reduced concordance (∼50%) in monozygotic twins (Gejman et al., 2010), a pattern that is compatible with a complex polygenic etiology influenced by environmental factors (Cardno & Gottesman, 2000; Sullivan et al., 2003). Attempts to identify genes and genetic and environmental determinants of the disease have been challenging. The results show that genetic, epigenetic, and environmental insults contribute to this devastating disorder. A recent large-scale genomewide association study (GWAS) has found common variants associated with SZ at 270 unique loci in 69,369 SZ patients and 236,642 controls. The study also highlights 19 genes based on protein-coding or UTR variation, and 130 genes in total, using fine-mapping and functional genomic data (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2020; Trubetskoy et al., 2022). Yet, additional novel experimental approaches and technologies are needed to resolve the causes of schizophrenia. Here we will consider the potential involvement of the mitochondrial genome in SZ in a novel experimental design that uses monozygotic twins discordant (MZD) for the disease (Castellani et al., 2017; Tang et al., 2017).

Monozygotic twins (MZ), also referred to as identical twins, have historically been used for studies assessing genetic and environmental contributions to diseases. Due to the unique characteristic of a single zygote splitting into two embryos with a shared genome, differences in a disease phenotype in MZ pairs can be reasoned to be a source of difference from environmental or de novo events independently operating on each individual. As a result, MZ twins have been used to ascertain the genetic and environmental contributions to complex diseases, including SZ (Castellani et al., 2015; Li et al., 2017; Melka et al., 2015; Singh et al., 2020; Tang et al., 2017; Wrede et al., 2015). Over the years, these studies have focused on the nuclear genome. More recently, there has been some attempt to assess the involvement of mtDNA between MZD twins for a number of disorders. These have included sleep, multiple sclerosis, SZ, obesity, and childhood intelligence, to name a few (Avital et al., 2012; Bijnens et al., 2019; Heinonen et al., 2015; Li et al., 2017; Maeda et al., 2016; Souren et al., 2016; Wang et al., 2015; Wrede et al., 2015). For example, decreased mtDNA-CN was observed with sleep efficiency, childhood intelligence, and obesity, in blood mtDNA-CN, placental mtDNA content, and adipose tissue mtDNA respectively (Bijnens et al., 2019; Heinonen et al., 2015; Wrede et al., 2015). However, association studies on multiple sclerosis and SZ involving mtDNA heteroplasmy have shown little to no discordance for mtDNA heteroplasmy calls between discordant MZ twins. Current findings remain to be validated by larger sample sizes and greater sequencing depth, and a role for mitochondria in neurodevelopment and manifestation of SZ is an avenue for exploration. In fact, a link to developing brain function in a rat model of SZ has shown the importance of mitochondrial dynamics to this disease through mitochondrial transplantation (Ene et al., 2023). In addition, mitochondrial dysfunction in synaptic and metabolic pathways have been implicated but not well established in the pathophysiology of SZ (Hjelm et al., 2015; Park & Park, 2012; Stork & Renshaw, 2005). The objective of this report is to evaluate the distribution of mitochondrial DNA copy number (mtDNA-CN) and heteroplasmy in schizophrenia by assessing MZ twins discordant for SZ, and their parents where possible, using

Table 1. Summary of participant twin pairs discordant for schizophrenia and their respective parents

Twin pair	Individual	Age	Sex	Affymetrix 6.0	WGS	qPCR
1	1A	20	M	YES	NA	YES
	1B	20	M			
Family 2	2A	43	F	YES	YES	YES
	2B	43	F			
	Father	80	M	YES	NA	YES
	Mother	76	F			
Family 3	3A	53	F	YES	YES	NA
	3B	53	F			
	Father	82	M	YES	YES	NA
	Mother	74	F			
4	4A	22	F	YES	NA	YES
	4B	22	F			
5	5A	36	M	YES	NA	NA
	5B	36	M			
6	6A	42	M	YES	NA	YES
	6B	42	M			

Note: A, affected twin; B, unaffected twin; NA, not available; WGS, whole genome sequencing; qPCR, quantitative polymerase chain reaction.

microarrays (Affymetrix Human SNP Array 6.0), complete genome sequences (WGS), and quantitative polymerase chain reaction (qPCR). The findings argue that even with a powerful experimental model involving MZD twins, relatively large numbers of well-characterized patients will be needed to establish any direct role for SZ-related mitochondrial aberrations in humans.

Materials and Methods

Monozygotic Twin Subjects and Parents

All human subjects were diagnosed and recruited by Dr Richard O'Reilly under a grant from the Canadian Institutes of Health Research. They provided written consent to participate in the study, which has been described previously (Castellani et al., 2015; O'Reilly et al., 2013). In brief, genomic DNA from whole blood was extracted from six pairs of MZ twins discordant for SZ and two sets of parents using the 5 Prime Perfect Pure DNA Blood kit, following manufacturer's protocol ($N = 16$). A summary table of MZ twins and their respective parents with age, sex, and technology applied are listed in Table 1. Participants ranged from 20 to 82 years of age with an equal ratio of males and females.

Microarray mtDNA-CN Estimation

DNA from whole blood of all MZ twins discordant for SZ and their respective parents ($N = 16$) was hybridized to the Affymetrix Human SNP Array 6.0. The array results were used to assess mtDNA-CN, defined as the ratio of total autosomal reads to total mtDNA reads. The java-based program *Genvisis* was used to generate mtDNA-CN estimates adjusted for principal components, age, and sex (Yang et al., 2021). *Genvisis* involves using log R ratios (LRR) to obtain population-normalized intensity values for each SNP on the Affymetrix microarray. Mitochondrial and

autosomal markers are then selected for inclusion in the analysis. The method involves several quality filters and corrections, including correcting LRRs for GC content and removing mitochondrial markers that do not have a perfect match to the annotated location or have off-target matches. The resulting estimates capture batch effects, quality metrics, and latent confounding variables. The difference in mtDNA-CN estimates between MZD twins and their respective mothers (mother-child), and unrelated individual comparisons were generated. In the case of unrelated individuals, the difference in mtDNA-CN estimates for the unrelated group was compared between each individual to individuals outside of their respective families; therefore, comparisons did not have any familial relations—for example, twin 1A to twin 2B.

Whole Genome Sequencing, Alignment, mtDNA-CN Estimation, and Heteroplasmy Calling

Whole genome sequencing for a subset of six subjects comprising two MZ twin pairs discordant for SZ and one set of respective parents was generated from Complete Genomics Inc. (Mountain View, CA). CGA tools version 1.8.0 build 3 was used to align each individual genome sequence with hg19 as reference. The revised Cambridge Reference Sequence (NC_012920.1) was used as reference for mitochondrial genome alignment (Andrews et al., 1999). The haplogroup of each individual was determined with Haplogrep 3 version 3.2.1 (Schönherr et al., 2023). Two distinct programs, *Mutect2* and *Mutserve,* were used for heteroplasmy calling (Diroma et al., 2021; McKenna et al., 2010). *Mutect2* and *Mutserve* were run and mitochondrial heteroplasmic variants were called at 1%, 3% and 5% threshold level and an output variant call format (VCF) file for each heteroplasmy level were generated as text files (Weissensteiner et al., 2016). Mutect2 filtering criteria included setting: -initial-tumor-lod to 0, –tumor-lod-to-emit to 0, –af-of-alleles-not-in-resource to 4e-3, and –pruning-lod-threshold to -4, where LOD represents the logarithm of the odds (mitochondria mode). The alignment, base, and mapping quality scores cutoff was set to a minimum of 10 for both programs. *fastMitoCalc* was used to estimate mtDNA-CN from whole genome-sequencing data and is based on the proportion of sequencing coverage to underlying DNA copy number for autosomal and mitochondrial DNA (Qian et al., 2017). mtDNA-CN was defined as the average coverage of mtDNA divided by the average coverage of autosomal DNA, multiplied by 2.

Quantitative Polymerase Chain Reaction mtDNA-CN Estimation

qPCR-based estimates of mtDNA-CN for a subset of twins comprising four discordant twin pairs and one set of respective parents were generated. A modified multiplexed real-time qPCR assay from a previously described method was used (Hsieh et al., 2018; Yang et al., 2021). More specifically, the mtDNA target was the mtDNA control region (D-loop), a noncoding region in the mtDNA genome. The nuclear DNA (nDNA) target was the albumin gene, which is a single-copy gene present in all human cells. This method uses 20 ng of DNA per sample as input for the qPCR reaction. The Viia7 qPCR machine (ThermoFisher Scientific) alongside the provided Viia7 Software version 1.2.2 was used to determine Ct values. All samples were run in triplicates and deltaCt was calculated using the R software.
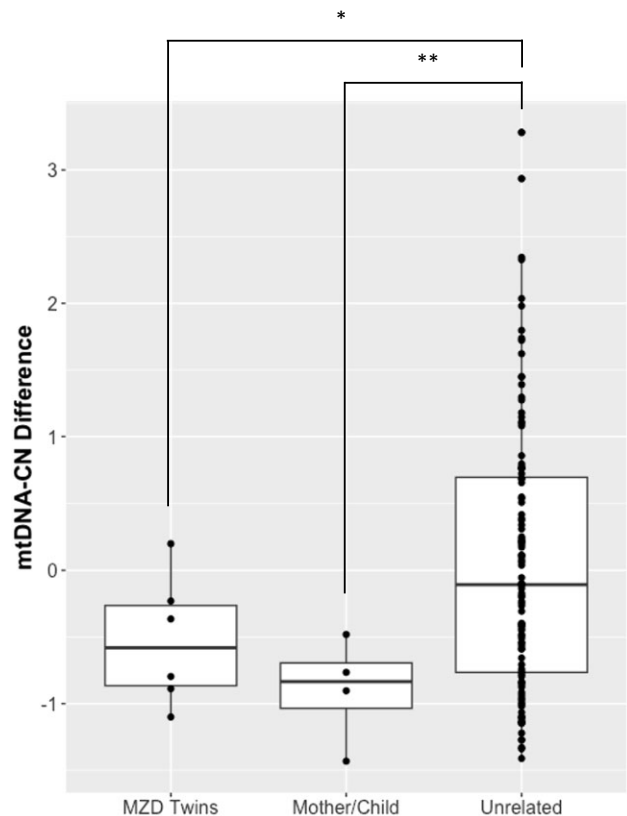

Figure 1. Affymetrix array mtDNA-CN estimates difference between individuals. Mean comparison of the groups showed significant difference for MZD Twins (*$p < .05$) and Mother/Child (**$p < .01$) when compared to unrelated samples.

Statistical Analysis

Unpaired *t* tests were used to determine differences in mtDNA-CN means between groups. The Shapiro-Wilk test was used to test for normality of all variables before performing *t* tests to ensure that the assumptions of normality were met. All statistical analyses were performed using R (version 4.1.2). Additionally, using the known correlate of mitochondrial relatedness to mtDNA-CN, the effectiveness of different mtDNA-CN estimation techniques was compared with a linear regression of mtDNA-CN estimates for each technology to relatedness. The WGS mtDNA-CN estimates were not used in this regression analysis due to the small sample size which could undermine the validity of the comparison.

Results

Monozygotic Twins Discordant for Schizophrenia Differ for mtDNA-CN

The demographics of MZD twins and parental samples are summarized in Table 1. In this cohort, we identified differences in mtDNA-CN using three different technologies: Affymetrix, WGS and qPCR. The results generated by the Affymetrix microarray, for which we have results for all twin pairs, highlights the difference across individuals of differing relationships of MZ twins, mother-child and unrelated individuals (Figure 1). However, it is important to note that these conclusions are drawn from a study with a limited sample size of 16 individuals, comprising 6 twin pairs and 2
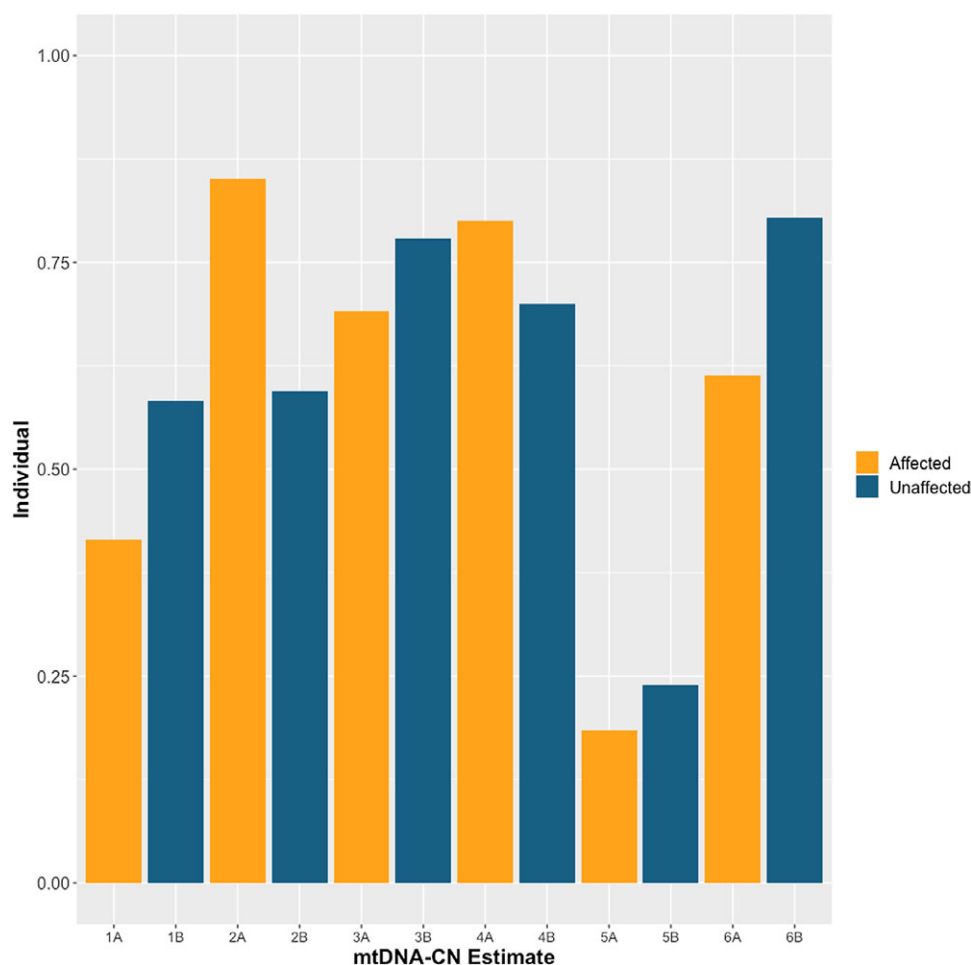
Figure 2. Affymetrix array mtDNA-CN estimates for each individual twin across 6 monozygotic discordant (MZD) twin pairs. SZ affected co-twin (A), Unaffected (B).

sets of parents. It shows that the familial comparisons display smaller differences while the unrelated individuals show extensive variability, as expected. Here, the MZ within-pair differences were smaller than unrelated individual comparisons ($N = 16$, $p = .03$, effect size $= .60$). Additionally, the mother-child to unrelated group comparison also showed a significant difference ($n = 4$, $p = .01$, effect size $= .95$). Further, a trend was seen in mother-child to MZD twin comparison where a closer similarity was seen in mtDNA-CN estimates between a mother and child as compared to within MZD twins ($n = 4$, $p = .23$, effect size $= .81$), though this difference was not statistically significant. More interesting are the comparisons that specifically relate to within MZD pair comparisons. Interestingly, the comparison varies across the six twin pairs studied (Figure 2). In all twin pairs, Affymetrix mtDNA-CN estimation resulted in a difference in mtDNA-CN between each pair. In the case of qPCR ($n = 10$) and WGS ($n = 6$), where only a subset of the MZD twins was run, the results also showed a similar trend, whereby identical twins displayed discordant mtDNA-CN measures.

Direction of mtDNA-CN Differences is Consistent Across Array, WGS and qPCR Technologies

MZD twins exhibit shared directionality in mtDNA-CN estimation across three different mtDNA-CN estimation techniques (Figure 3). In some cases, the magnitude of the mtDNA-CN estimate showed numerical differences across different technologies; however, the direction of mtDNA-CN estimation between MZD twins was consistent. As an example, twin pair 2, for which mtDNA-CN estimates from all three technologies were generated, a lower mtDNA-CN estimate in the affected twin was observed compared to the unaffected co-twin. Also, for twin pairs 3, 5 and 6, where more than one technology was available, there is a similar trend, and unlike twin pair 2, the affected twin displays higher mtDNA-CN than the unaffected co-twin. Such results may argue for the potential involvement of mtDNA-CN in twin discordance for SZ.

qPCR Outperforms Affymetrix in Estimating mtDNA-CN

Using the expected correlate for mtDNA-CN of relatedness, we found expected trends of mtDNA-CN with mitochondrial relatedness (Twin pairs $= 1$, Unrelated $= 0$; Figure 4). The effectiveness of qPCR and Affymetrix technologies in estimating mtDNA-CN was determined by assessing the strength of these known relationships. Specifically, we found that as relatedness increases, mtDNA-CN concordance decreases. As outlined in the methods, WGS was not utilized in this part of the analysis due to the small sample size for estimating expected trends. qPCR performed best at mtDNA-CN estimation based on this relatedness metric (qPCR $R^2 = 0.74$, Affymetrix $R^2 = 0.07$).
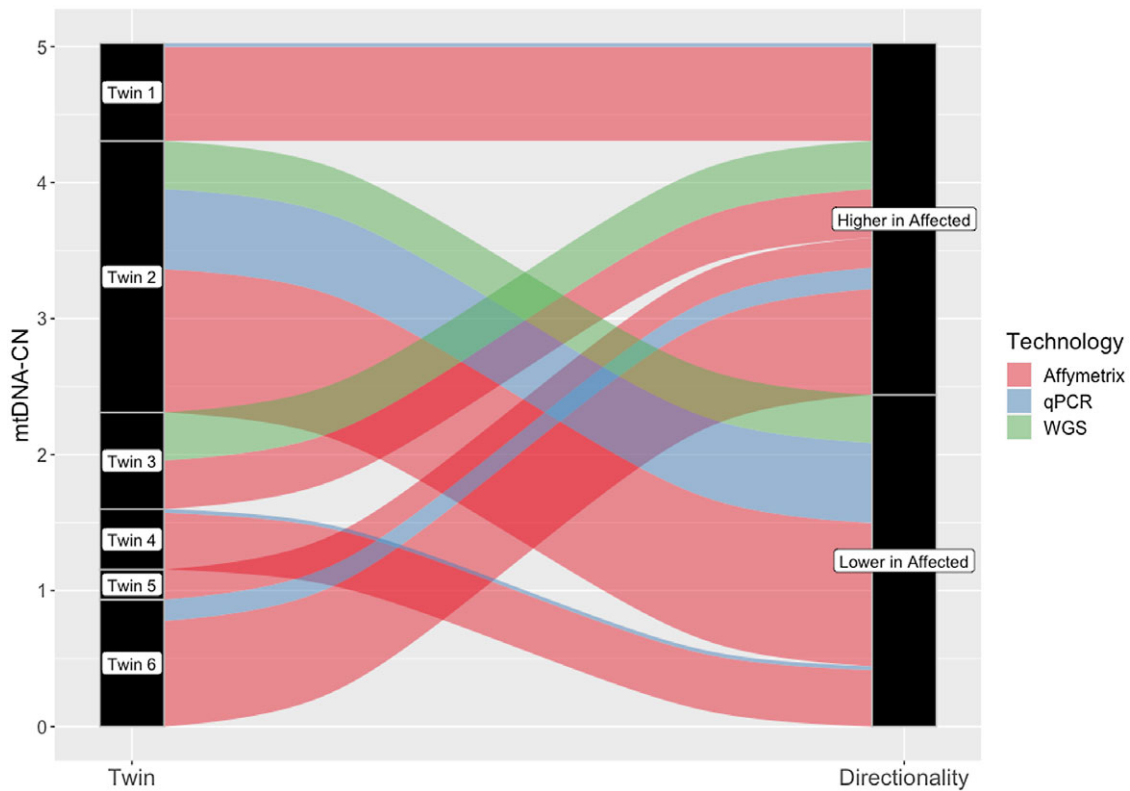
Figure 3. Alluvial plot indicating the directionality of the mtDNA-CN differences between affected and unaffected twins as estimated by Affymetrix, qPCR and WGS. Directionality of mtDNA-CN differences was concordant across all available technologies. Line thickness reflects the magnitude of the mtDNA-CN difference between co-twins.
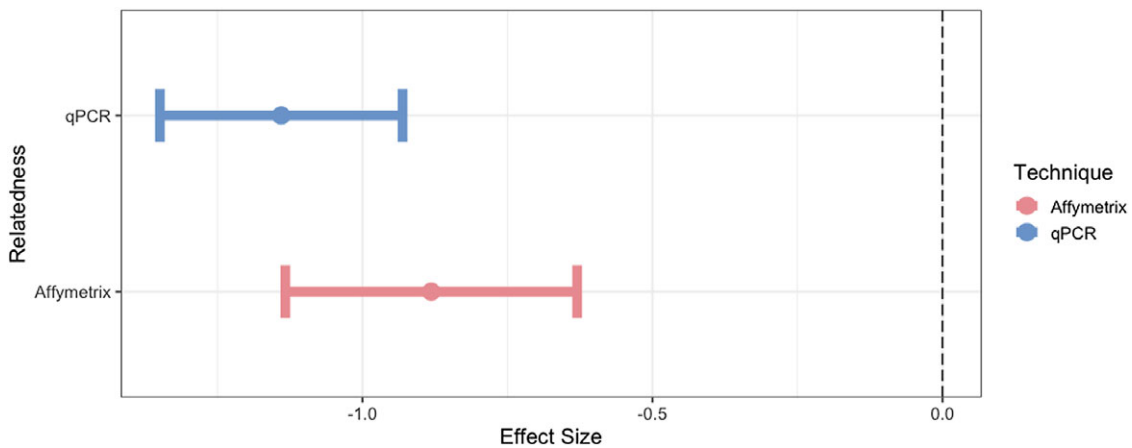

Figure 4. The effect size and 95% confidence interval for the association of mtDNA-CN with its known correlate, relatedness. Only unrelated individuals and identical twins are included in this analysis.

Absence of mtDNA Heteroplasmy Differences Between MZD Twins

Heteroplasmy calling was performed using WGS data for two twin pairs (Twin 2 and 3), using two independent programs, *Mutserve* and *Mutect2*. The cut-off criteria for calling a mutation was a call that was discordant from reference and had a base quality above 10. *Mutect2* did not identify any heteroplasmic calls at the specified thresholds of 1%, 3% and 5%. In *Mutserve* analysis, eight heteroplasmic sites were identified in twin pair 2 and no sites were identified in twin pair 3. In twin pair 2, four heteroplasmy calls were identified by *Mutserve* to be

unshared between twins, three of the heteroplasmies were unique to the affected twin (75%) and only one (25%) was unique to the unaffected twin (Table 2). Heteroplasmic calls at 1%, 3% and 5% thresholds were generated, and the results are available in Table 3. However, with further in-depth review of the raw sequencing reads at these locations through IGV visualization (Robinson et al., 2011), all unique heteroplasmy calls between co-twins called by either program could be attributed to sites lost during quality control steps in one twin but not in the co-twin and calls in the co-twin that were just below the heteroplasmic threshold (Figure 5). We conclude that these unique calls reflect differences in filtering/calling criteria rather than true

Table 2. Mutserve mitochondrial heteroplasmy calls for unique and shared heteroplasmy in discordant twin pair 2 aligned to reference rCRS with calls found in the affected twin in bold type and unaffected twin in roman type. Unique heteroplasmies are shown first followed by shared heteroplasmies. Heteroplasmy calls not found in the major Haplogrep assigned haplogroup (U4b1a3a) are indicated with 'X'.

Mutserve call: Shared or unique heteroplasmy	Position	Sequencing depth	Reference allele	Alternate allele	Variant level	Not found in haplogroup U4b1a3a	Conclusion following inspection of underlying calls
Unique	**117**	**300**	**T**	**A**	**0.017**	**X**	**Heteroplasmic in both twins (3%)**
	7261	**656**	**T**	**C**	**0.011**	**X**	**Heteroplasmic in both twins (2%)**
	12344	**308**	**T**	**A**	**0.023**	**X**	**Heteroplasmic in both twins (5%)**
	14775	201	T	C	0.025	X	Heteroplasmic in both twins (5%)
Shared	**6047**	**532**	**A**	**G**	**0.998**		**Heteroplasmic in both twins (2%)**
	6047	526	A	G	0.989		Heteroplasmic in both twins (2%)
	7705	**148**	**T**	**C**	**0.98**		**Heteroplasmic in both twins (2%)**
	7705	159	T	C	0.962		Heteroplasmic in both twins (2%)
	14620	**511**	**C**	**T**	**0.978**		**Heteroplasmic in both twins (3%)**
	14620	515	C	T	0.983		Heteroplasmic in both twins (3%)
	11332	**546**	**C**	**T**	**0.993**		**Heteroplasmic in both twins (4%)**
	11332	514	C	T	0.984		Heteroplasmic in both twins (4%)

Table 3. Total heteroplasmy call count for Mutserve at a heteroplasmy calling threshold of 1%, 3% and 5% for each individual. Unshared call counts between twins are indicated in brackets.

	Mutserve heteroplasmy level and count		
Individual	1%	3%	5%
2A	7 (3)	7 (3)	10 (3)
2B	5 (1)	5 (1)	7 (1)
3A	0 (0)	1 (0)	8 (0)
3B	0 (0)	5 (0)	5 (0)
Father 3A/3B	5	5	11
Mother 3A/3B	3	3	4

unique heteroplasmies between identical twins. The haplogroup for twin pair 2 was determined to be U4b1a3a and the haplogroup for twin pair 3 was determined to be L3. Both individuals from twin pair 2 have at least one heteroplasmic call that is not found in their major haplogroup. Given that these heteroplasmies are shared between both the affected and unaffected twin and display the same level of heteroplasmic load in both twins, it is most likely that these variants represent biological variation that arose prior to the twinning event.

Discussion

The etiology of SZ is complex, with additional heterogeneity in clinical presentation. Although the nuclear genome of MZ twins discordant for SZ would theoretically be near identical, with rare postzygotic changes involving somatic mutations (Singh et al., 2020), it is possible that non-Mendelian genetic factors could contribute to the phenotypic outcome of SZ, including the contribution of the mitochondrial genome (Bi et al., 2016). Furthermore, since the mitochondrial genome is inherited maternally, studies of mothers and their respective MZ offspring provides a prospect to trace the emergence of variation in the mitochondrial genome while controlling for variation in the nuclear genome. This study used next generation sequencing, genotyping arrays, and qPCR to assess mtDNA-CN and heteroplasmy variation in MZ twins discordant for SZ with their respective parents in a subset of twins.

mtDNA-CN has been shown to be associated with age, cardiovascular disease risk, diabetes and all-cause mortality, among other diseases (Ashar et al., 2015; Ashar et al., 2017; Su et al., 2022). More specifically, in discordant MZ twin studies, mtDNA-CN variation has been shown to be associated with sleep quality, obesity, and childhood intelligence (Avital et al., 2012; Bijnens et al., 2019; Detjen et al., 2007; Wrede et al., 2015). Therefore, mtDNA-CN could also be an important avenue for establishing non-Mendelian factors to SZ pathophysiology. Using MZ twins discordant for SZ, the results included in this report indicate that MZ twins discordant for SZ may display differences in mtDNA-CN and the direction of these differences involving affected and unaffected co-twins is independent of the technology used. Interestingly, a trend with a stronger concordance in mother-twin mtDNA-CN estimates for each twin compared to within-twin differences is noteworthy. However, this result must be taken with caution due to the small sample size and marginal statistical difference; therefore, it is critical to interpret these findings with caution and acknowledge the potential impact of the small sample size on the generalizability of the results. A possible reason for this trend could be explained by the fact that throughout the lifespan of the affected SZ twin, environmental and random genetic factors could have accumulated with inherited familial risk leading to the discordant SZ phenotype (Singh et al., 2020). With that, larger variation in mtDNA-CN between discordant MZ twins compared to mother-child could suggest a role for mtDNA-CN as one of the factors that causes a threshold 'push' towards the phenotypic SZ outcome (Singh et al., 2020). Given the complex nature of SZ, which involves multiple familial genetic, epigenetic and environmental factors, caution should be taken when interpreting these results (Hjelm et al., 2015; Pardiñas et al., 2018; Singh et al., 2020). It is also probable that the degree of effect of mtDNA-CN on SZ is relatively small compared to other established genetic risk factors, which strongly requires that the sample size for future studies be expanded upon to control for these aforementioned factors.
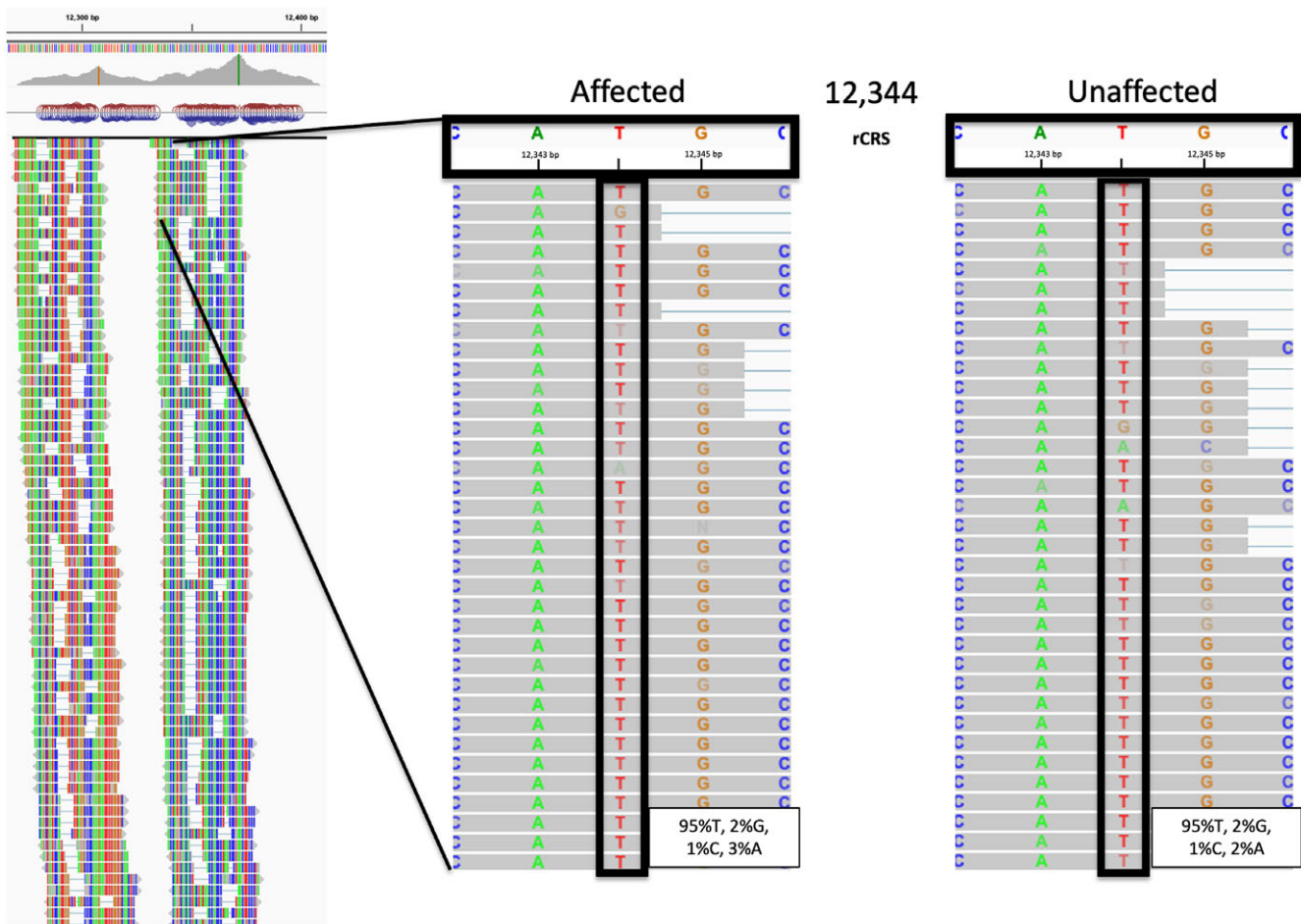
Figure 5. IGV software screenshot of WGS alignments at mitochondrial location 12,344 for affected (left) and unaffected (right) twin pair 2 aligned to the rCRS reference genome. The affected twin had 95% T, 2% G, 1% C, 3% A base pair calls (330 T, 6G, 2C, 10A), a read depth of 348, and mean base and mapping quality of 17.9 and 37.5, respectively. The unaffected twin had 95% T, 2% G, 1% C, 2% A base pair calls (277T, 6G, 3C, 6A), a read depth of 567, and mean base and mapping quality of 19.1 and 37.9, respectively.

Additionally, although WGS, microarray and qPCR provided a reliable measure of mtDNA-CN, our results favor qPCR over the Affymetrix technology for mtDNA-CN estimation. Further, to accurately quantify mtDNA-CN, it is crucial to consider the impact of DNA extraction methods on estimation. It has been shown that different extraction methods can lead to variations in mtDNA-CN measurements, likely arising from nonuniform changes in nDNA and mtDNA ratios. Studies utilizing different extraction methods have demonstrated that mtDNA-CN estimated from phenol-chloroform-isoamyl alcohol-based methods are significantly less variable than those isolated from traditional silica-based column selection methods (Guo et al., 2009; Longchamps et al., 2020). Additionally, repeated freeze-thaw cycles of samples has been found to decrease overall mtDNA-CN estimates, while whole blood samples yielded higher mtDNA-CN estimates as compared to buffy coat samples (Andreu et al., 2009). Taken together, optimizing these factors are crucial for obtaining reliable and consistent measurements of mtDNA-CN. Due to low sample sizes, we did not assess WGS in this comparison. As a result of differing sample size availability between measured mtDNA-CN techniques of microarray and qPCR and contrasting performance results to previous studies with larger sample sizes, these current findings should be interpreted with caution. In this study we identified consistent mtDNA-CN differences with shared directionally between MZD twins across different technological measures.

Further understanding of factors that contribute to the differences in mtDNA-CN between discordant twins will be expected to provide insights into the etiology of SZ while identifying novel genetic and environmental factors that influence MZD twin dynamics. One such feature may be mitochondrial heteroplasmy, which we observed in this subset of MZ twins discordant for SZ to show complete concordance; however, we were also limited by both sample size and read depth.

The results included in this report show no clear unique heteroplasmy calls in the affected SZ twin compared to the unaffected co-twin. Our findings are consistent with earlier research on the relationship between mtDNA heteroplasmy in discordant SZ MZ twins (Li et al., 2017). The high concordance in mtDNA heteroplasmy between the discordant MZ twins may point to a relatively equal distribution of mtDNA during embryonic segregation and similar modulating effect from the nuclear genetic background. Additionally, this may also suggest other environmental and/or epigenetic factors are contributing to the discordance in SZ between MZ twins (Brown, 2011; Dempster et al., 2011; Fraga et al., 2005). However, due to the complex nature of mtDNA heteroplasmy in SZ, interaction with the nuclear genome, and postzygotic genetic contributions to SZ, a multifaceted approach to the role of mtDNA heteroplasmy in SZ is required (Pardiñas et al., 2018; Singh et al., 2020). Contamination can also be a significant factor leading to the presence of heteroplasmic

mutations in individuals. In studies involving small sample sizes or employing high-throughput sequencing technologies, the risk of contamination increases (Yao et al., 2007). As a result, when analyzing the heteroplasmic mutations specific to a predetermined haplogroup variation may appear as a result of natural biological variation. The four non-U4b1a3a heteroplasmies shared by both twins do not contribute to a shared alternative haplogroup, further supporting the unlikelihood that they represent contamination. Implementing rigorous quality control measures and carefully assessing lineage and phylogenetic relationships aids in interpretation. In the case of this report, the existence of variants in both co-twins increases the likelihood that they represent true biological variation, though the potential for contamination cannot be exclusively eliminated.

In conclusion, we did not find a discernible difference in mtDNA heteroplasmy between MZ twins discordant for SZ based on our cut-off criteria using whole genome sequencing technology. However, we identified differences in mtDNA-CN that may or may not be related to twin discordance for SZ. Future research looking at mtDNA-CN and heteroplasmy in a larger cohort of MZ twin discordant for SZ and their families would further our understanding on how unique, shared, or changing heteroplasmy, and mtDNA-CN variation affects SZ risk.

References

Andreu, A. L., Martinez, R., Marti, R., & García-Arumí, E. (2009). Quantification of mitochondrial DNA copy number: Pre-analytical factors. *Mitochondrion*, 9, 242–246. https://doi.org/10.1016/j.mito.2009.02.006

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., & Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics*, 23, Article 2. https://doi.org/10.1038/13779

Ashar, F. N., Moes, A., Moore, A. Z., Grove, M. L., Chaves, P. H. M., Coresh, J., Newman, A. B., Matteini, A. M., Bandeen-Roche, K., Boerwinkle, E., Walston, J. D., & Arking, D. E. (2015). Association of mitochondrial DNA levels with frailty and all-cause mortality. *Journal of Molecular Medicine*, 93, 177–186. https://doi.org/10.1007/s00109-014-1233-3

Ashar, F. N., Zhang, Y., Longchamps, R. J., Lane, J., Moes, A., Grove, M. L., Mychaleckyj, J. C., Taylor, K. D., Coresh, J., Rotter, J. I., Boerwinkle, E., Pankratz, N., Guallar, E., & Arking, D. E. (2017). Association of mitochondrial DNA copy number with cardiovascular disease. *JAMA Cardiology*, 2, 1247–1255. https://doi.org/10.1001/jamacardio.2017.3683

Avital, G., Buchshtav, M., Zhidkov, I., Tuval Feder, J., Dadon, S., Rubin, E., Glass, D., Spector, T. D., & Mishmar, D. (2012). Mitochondrial DNA heteroplasmy in diabetes and normal adults: Role of acquired and inherited mutational patterns in twins. *Human Molecular Genetics*, 21, 4214–4224. https://doi.org/10.1093/hmg/dds245

Bi, R., Tang, J., Zhang, W., Li, X., Chen, S.-Y., Yu, D., Chen, X., & Yao, Y.-G. (2016). Mitochondrial genome variations and functional characterization in Han Chinese families with schizophrenia. *Schizophrenia Research*, 171, 200–206. https://doi.org/10.1016/j.schres.2016.01.011

Bijnens, E. M., Derom, C., Weyers, S., Janssen, B. G., Thiery, E., & Nawrot, T. S. (2019). Placental mitochondrial DNA content is associated with childhood intelligence. *Journal of Translational Medicine*, 17, 361. https://doi.org/10.1186/s12967-019-2105-y

Brown, A. S. (2011). The environment and susceptibility to schizophrenia. *Progress in Neurobiology*, 93, 23–58. https://doi.org/10.1016/j.pneurobio.2010.09.003

Cardno, A. G., & Gottesman, I. I. (2000). Twin studies of schizophrenia: From bow-and-arrow concordances to star wars Mx and functional genomics. *American Journal of Medical Genetics*, 97, 12–17.

Castellani, C. A., Laufer, B. I., Melka, M. G., Diehl, E. J., O'Reilly, R. L., & Singh, S. M. (2015). DNA methylation differences in monozygotic twin pairs discordant for schizophrenia identifies psychosis related genes and networks. *BMC Medical Genomics*, 8, 17. https://doi.org/10.1186/s12920-015-0093-1

Castellani, C. A., Melka, M. G., Gui, J. L., Gallo, A. J., O'Reilly, R. L., & Singh, S. M. (2017). Post-zygotic genomic changes in glutamate and dopamine pathway genes may explain discordance of monozygotic twins for schizophrenia. *Clinical and Translational Medicine*, 6, 43. https://doi.org/10.1186/s40169-017-0174-1

Dempster, E. L., Pidsley, R., Schalkwyk, L. C., Owens, S., Georgiades, A., Kane, F., Kalidindi, S., Picchioni, M., Kravariti, E., Toulopoulou, T., Murray, R. M., & Mill, J. (2011). Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. *Human Molecular Genetics*, 20, 4786–4796. https://doi.org/10.1093/hmg/ddr416

Detjen, A. K., Tinschert, S., Kaufmann, D., Algermissen, B., Nürnberg, P., & Schuelke, M. (2007). Analysis of mitochondrial DNA in discordant monozygotic twins with neurofibromatosis type 1. *Twin Research and Human Genetics*, 10, 486–495. https://doi.org/10.1375/twin.10.3.486

Diroma, M. A., Modi, A., Lari, M., Sineo, L., Caramelli, D., & Vai, S. (2021). New insights into mitochondrial DNA reconstruction and variant detection in ancient samples. *Frontiers in Genetics*, 12, 619950. https://doi.org/10.3389/fgene.2021.619950

Ene, H. M., Karry, R., Farfara, D., & Ben-Shachar, D. (2023). Mitochondria play an essential role in the trajectory of adolescent neurodevelopment and behavior in adulthood: Evidence from a schizophrenia rat model. *Molecular Psychiatry*, 28, Article 3. https://doi.org/10.1038/s41380-022-01865-4

Filograna, R., Mennuni, M., Alsina, D., & Larsson, N.-G. (2021). Mitochondrial DNA copy number in human disease: The more the better? *FEBS Letters*, 595, 976–1002. https://doi.org/10.1002/1873-3468.14021

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y.-Z., … Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences*, 102, 10604–10609. https://doi.org/10.1073/pnas.0500398102

Gejman, P., Sanders, A., & Duan, J. (2010). The role of genetics in the etiology of schizophrenia. *The Psychiatric Clinics of North America*, 33, 35–66. https://doi.org/10.1016/j.psc.2009.12.003

Guo, W., Jiang, L., Bhasin, S., Khan, S. M., & Swerdlow, R. H. (2009). DNA extraction procedures meaningfully influence qPCR-based mtDNA copy number determination. *Mitochondrion*, 9, 261–265. https://doi.org/10.1016/j.mito.2009.03.003

Heinonen, S., Buzkova, J., Muniandy, M., Kaksonen, R., Ollikainen, M., Ismail, K., Hakkarainen, A., Lundbom, J., Lundbom, N., Vuolteenaho, K., Moilanen, E., Kaprio, J., Rissanen, A., Suomalainen, A., & Pietiläinen, K. H. (2015). Impaired mitochondrial biogenesis in adipose tissue in acquired obesity. *Diabetes*, 64, 3135–3145. https://doi.org/10.2337/db14-1937

Hjelm, B. E., Rollins, B., Mamdani, F., Lauterborn, J. C., Kirov, G., Lynch, G., Gall, C. M., Sequeira, A., & Vawter, M. P. (2015). Evidence of

mitochondrial dysfunction within the complex genetic etiology of schizophrenia. *Molecular Neuropsychiatry*, 1, 201–219. https://doi.org/10.1159/000441252

Hong, Y. S., Longchamps, R. J., Zhao, D., Castellani, C. A., Loehr, L. R., Chang, P. P., Matsushita, K., Grove, M. L., Boerwinkle, E., Arking, D. E., & Guallar, E. (2020). Mitochondrial DNA copy number and incident heart failure: The Atherosclerosis Risk in Communities (ARIC) Study. *Circulation*, 141, 1823–1825. https://doi.org/10.1161/CIRCULATIONAHA.120.046001

Hsieh, A. Y. Y., Budd, M., Deng, D., Gadawska, I., & Côté, H. C. F. (2018). A monochrome multiplex real-time quantitative PCR assay for the measurement of mitochondrial DNA content. *The Journal of Molecular Diagnostics*, 20, 612–620. https://doi.org/10.1016/j.jmoldx.2018.05.001

Kopinski, P. K., Janssen, K. A., Schaefer, P. M., Trefely, S., Perry, C. E., Potluri, P., Tintos-Hernandez, J. A., Singh, L. N., Karch, K. R., Campbell, S. L., Doan, M. T., Jiang, H., Nissim, I., Nakamaru-Ogiso, E., Wellen, K. E., Snyder, N. W., Garcia, B. A., & Wallace, D. C. (2019). Regulation of nuclear epigenome by mitochondrial DNA heteroplasmy. *Proceedings of the National Academy of Sciences*, 116, 16028–16035. https://doi.org/10.1073/pnas.1906896116

Kujoth, G. C., Leeuwenburgh, C., & Prolla, T. A. (2006). Mitochondrial DNA mutations and apoptosis in mammalian aging. *Cancer Research*, 66, 7386–7389. https://doi.org/10.1158/0008-5472.CAN-05-4670

Li, H., Bi, R., Fan, Y., Wu, Y., Tang, Y., Li, Z., He, Y., Zhou, J., Tang, J., Chen, X., & Yao, Y.-G. (2017). MtDNA Heteroplasmy in Monozygotic Twins Discordant for Schizophrenia. *Molecular Neurobiology*, 54, 4343–4352. https://doi.org/10.1007/s12035-016-9996-x

Longchamps, R. J., Castellani, C. A., Yang, S. Y., Newcomb, C. E., Sumpter, J. A., Lane, J., Grove, M. L., Guallar, E., Pankratz, N., Taylor, K. D., Rotter, J. I., Boerwinkle, E., & Arking, D. E. (2020). Evaluation of mitochondrial DNA copy number estimation techniques. *Plos One*, 15, e0228166. https://doi.org/10.1371/journal.pone.0228166

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The hallmarks of aging. *Cell*, 153, 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039

Maeda, K., Kawai, H., Sanada, M., Terashima, T., Ogawa, N., Idehara, R., Makiishi, T., Yasuda, H., Sato, S., Hoshi, K., Yahikozawa, H., Nishi, K., Itoh, Y., Ogasawara, K., Tomita, K., Indo, H. P., & Majima, H. J. (2016). Clinical phenotype and segregation of mitochondrial 3243A>G mutation in 2 pairs of monozygotic twins. *JAMA Neurology*, 73, 990–993. https://doi.org/10.1001/jamaneurol.2016.0886

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

Melka, M. G., Castellani, C. A., O'Reilly, R., & Singh, S. M. (2015). Insights into the origin of DNA methylation differences between monozygotic twins discordant for schizophrenia. *Journal of Molecular Psychiatry*, 3, 7. https://doi.org/10.1186/s40303-015-0013-5

Mengel-From, J., Thinggaard, M., Dalgård, C., Kyvik, K. O., Christensen, K., & Christiansen, L. (2014). Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Human Genetics*, 133, 1149–1159. https://doi.org/10.1007/s00439-014-1458-9

Olfson, M., Gerhard, T., Huang, C., Crystal, S., & Stroup, T. S. (2015). Premature mortality among adults with schizophrenia in the United States. *JAMA Psychiatry*, 72, 1172–1181. https://doi.org/10.1001/jamapsychiatry.2015.1737

O'Reilly, R., Torrey, E. F., Rao, J., & Singh, S. (2013). Monozygotic twins with early-onset schizophrenia and late-onset bipolar disorder: A case report. *Journal of Medical Case Reports*, 7, 134. https://doi.org/10.1186/1752-1947-7-134

Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., . . . Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and

in regions under strong background selection. *Nature Genetics*, 50, 381–389. https://doi.org/10.1038/s41588-018-0059-2

Park, C., & Park, S. K. (2012). Molecular links between mitochondrial dysfunctions and schizophrenia. *Molecules and Cells*, 33, 105–110. https://doi.org/10.1007/s10059-012-2284-3

Qian, Y., Butler, T. J., Opsahl-Ong, K., Giroux, N. S., Sidore, C., Nagaraja, R., Cucca, F., Ferrucci, L., Abecasis, G. R., Schlessinger, D., & Ding, J. (2017). fastMitoCalc: An ultra-fast program to estimate mitochondrial DNA copy number from whole-genome sequences. *Bioinformatics*, 33, 1399–1401. https://doi.org/10.1093/bioinformatics/btw835

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29, 24–26. https://doi.org/10.1038/nbt.1754

Schönherr, S., Weissensteiner, H., Kronenberg, F., & Forer, L. (2023). Haplogrep 3 — An interactive haplogroup classification and analysis platform. *Nucleic Acids Research*, gkad284. https://doi.org/10.1093/nar/gkad284

Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, S., Walters, J. T., & O'Donovan, M. C. (2020). Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *medRxiv*. https://doi.org/10.1101/2020.09.12.20192922

Singh, S. M., Castellani, C. A., & Hill, K. A. (2020). Postzygotic somatic mutations in the human brain expand the threshold-liability model of schizophrenia. *Frontiers in Psychiatry*, 11, 587162. https://doi.org/10.3389/fpsyt.2020.587162

Souren, N. Y. P., Gerdes, L. A., Kümpfel, T., Lutsik, P., Klopstock, T., Hohlfeld, R., & Walter, J. (2016). Mitochondrial DNA variation and heteroplasmy in monozygotic twins clinically discordant for multiple sclerosis. *Human Mutation*, 37, 765–775. https://doi.org/10.1002/humu.23003

Stefano, G. B., & Kream, R. M. (2016). Mitochondrial DNA heteroplasmy in human health and disease. *Biomedical Reports*, 4, 259–262. https://doi.org/10.3892/br.2016.590

Stork, C., & Renshaw, P. F. (2005). Mitochondrial dysfunction in bipolar disorder: Evidence from magnetic resonance spectroscopy research. *Molecular Psychiatry*, 10, 900–919. https://doi.org/10.1038/sj.mp.4001711

Su, L.-Y., Li, Y., Liu, Q., Jiao, L., Shen, J., Yang, L.-X., Yuan, T.-F., & Yao, Y.-G. (2022). Decreased peripheral mtDNA in methamphetamine use disorder. *Science China Life Sciences*, 65, 648–650. https://doi.org/10.1007/s11427-021-2027-1

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait: Evidence from a meta-analysis of twin studies. *Archives of General Psychiatry*, 60, 1187–1192.

Tang, J., Fan, Y., Li, H., Xiang, Q., Zhang, D.-F., Li, Z., He, Y., Liao, Y., Wang, Y., He, F., Zhang, F., Shugart, Y. Y., Liu, C., Tang, Y., Chan, R. C. K., Wang, C.-Y., Yao, Y.-G., & Chen, X. (2017). Whole-genome sequencing of monozygotic twins discordant for schizophrenia indicates multiple genetic risk factors for schizophrenia. *Journal of Genetics and Genomics*, 44, 295–306. https://doi.org/10.1016/j.jgg.2017.05.005

Thornicroft, G. (2011). Physical health disparities and mental illness: The scandal of premature mortality. *The British Journal of Psychiatry*, 199, 441–442. https://doi.org/10.1192/bjp.bp.111.092718

Trubetskoy, V., Pardiñas, A. F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T. B., Bryois, J., Chen, C.-Y., Dennison, C. A., Hall, L. S., Lam, M., Watanabe, K., Frei, O., Ge, T., Harwood, J. C., Koopmans, F., Magnusson, S., Richards, A. L., Sidorenko, J., . . . van Os, J. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*, 604, 502–508. https://doi.org/10.1038/s41586-022-04434-5

Wallace, D. C., & Chalkia, D. (2013). Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harbor Perspectives in Biology*, 5, a021220. https://doi.org/10.1101/cshperspect.a021220

Wang, Z., Zhu, R., Zhang, S., Bian, Y., Lu, D., & Li, C. (2015). Differentiating between monozygotic twins through next-generation mitochondrial genome sequencing. *Analytical Biochemistry*, 490, 1–6. https://doi.org/10.1016/j.ab.2015.08.024

Weissensteiner, H., Forer, L., Fuchsberger, C., Schöpf, B., Kloss-Brandstätter, A., Specht, G., Kronenberg, F., & Schönherr, S. (2016). mtDNA-Server: Next-generation sequencing data analysis of human

mitochondrial DNA in the cloud. *Nucleic Acids Research*, *44*, W64–W69. https://doi.org/10.1093/nar/gkw247

Wrede, J. E., Mengel-From, J., Buchwald, D., Vitiello, M. V., Bamshad, M., Noonan, C., Christiansen, L., Christensen, K., & Watson, N. F. (2015). Mitochondrial DNA copy number in sleep duration discordant monozygotic twins. *Sleep*, *38*, 1655–1658. https://doi.org/10.5665/sleep.5068

Yang, S. Y., Newcomb, C. E., Battle, S. L., Hsieh, A. Y., Chapman, H. L., Côté, H. C. F., & Arking, D. E. (2021). Rare instances of non-random dropout with the monochrome multiplex qPCR assay for mitochondrial DNA copy number. *bioRxiv*. https://doi.org/10.1101/2021.10.11.463983

Yao, Y.-G., Bandelt, H.-J., & Young, N. S. (2007). External contamination in single cell mtDNA analysis. *PloS One*, *2*, e681. https://doi.org/10.1371/journal.pone.0000681

Zhang, R., Wang, Y., Ye, K., Picard, M., & Gu, Z. (2017). Independent impacts of aging on mitochondrial DNA quantity and quality in humans. *BMC Genomics*, *18*, 890. https://doi.org/10.1186/s12864-017-4287-0