

# SCIENTIFIC REPORTS



OPEN

## Y Chromosome, Mitochondrial DNA and Childhood Behavioural Traits

Laurence J. Howe<sup>1</sup>, A. Mesut Erzurumluoglu<sup>1,2</sup>, George Davey Smith<sup>1</sup>, Santiago Rodriguez<sup>1</sup> & Evie Stergiakouli<sup>1</sup>

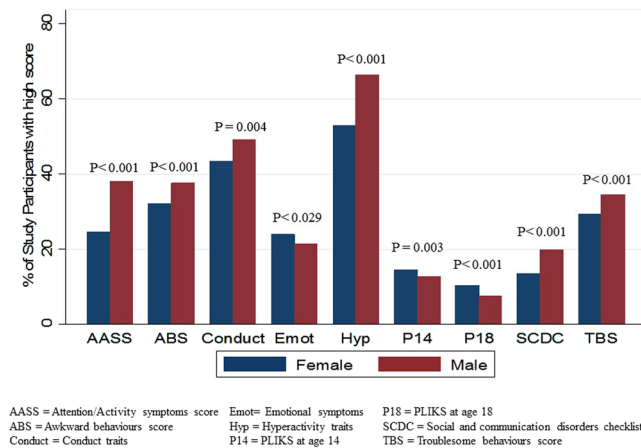
Many psychiatric traits are sexually dimorphic in terms of prevalence, age of onset, progression and prognosis; sex chromosomes could play a role in these differences. In this study we evaluated the association between Y chromosome and mitochondrial DNA haplogroups with sexually-dimorphic behavioural and psychiatric traits. The study sample included 4,211 males and 4,009 females with mitochondrial DNA haplogroups and 4,788 males with Y chromosome haplogroups who are part of the Avon Longitudinal Study of Parents and Children (ALSPAC) based in the United Kingdom. Different subsets of these populations were assessed using measures of behavioural and psychiatric traits with logistic regression being used to measure the association between haplogroups and the traits. The majority of behavioural traits in our cohort differed between males and females; however Y chromosome and mitochondrial DNA haplogroups were not associated with any of the variables. These findings suggest that if there is common variation on the Y chromosome and mitochondrial DNA associated with behavioural and psychiatric trait variation, it has a small effect.

Childhood psychiatric disorders are amongst the best examples of sexual dimorphism in disease. Males and females differ in terms of prevalence, age of onset, development and prognosis of psychiatric disorders<sup>1</sup>; both ADHD and autism are thought to be more prevalent in males<sup>2,3</sup> while psychosis is thought to have an earlier age of onset in men compared to women<sup>4</sup>. Several theories have been proposed to explain sex differences in psychiatric traits such as the “male brain hypothesis”<sup>5</sup> and “imprinting of the brain”<sup>6</sup> however research on possible contributing factors is largely limited<sup>7</sup>. Many childhood psychiatric disorders, including ADHD and autism, can be conceptualized as extremes of continuous traits found in the general population with the same risk factors contributing to both clinical disease and subclinical traits<sup>8–10</sup>. One of the genetic factors that clearly differs between males and females is the sex chromosome complement. Some Y chromosome genes are expressed in the human brain<sup>11</sup> and the Y chromosome has been associated with increased aggression and impaired parental behaviour in animal models where the sex chromosome complement has been dissociated from the hormonal milieu<sup>12</sup>. There is mixed evidence over whether patients with sex chromosome abnormalities are at increased risk of psychiatric disorders, varying by abnormality and disorder<sup>13–15</sup>.

Despite its potential involvement in sexually dimorphic disorders and traits, the Y chromosome, which is the sex-determining chromosome in humans, is largely excluded from genetic studies even of sexually dimorphic disease. The Y chromosome is passed largely unchanged from father to son and consists of the non-recombining region (NRY) and small pseudoautosomal regions at the telomeres. The lack of recombination provides the Y chromosome with the best haplotypic resolution in the human genome. However, the implication of lack of recombination is that the Y chromosome cannot be studied easily as part of a genome-wide association study (GWAS) pipeline. In contrast, it is analysed using haplogroups which are stable lineages of the Y chromosomes that share a common ancestor. The Y Chromosome Consortium (YCC) has produced a nomenclature system incorporating all verified Y chromosome single nucleotide polymorphisms (SNPs) and defining Y chromosome haplogroups<sup>16</sup>.

Mitochondrial DNA is separate from nuclear DNA and is contained within mitochondria; intracellular organelles that provide energy to the cell. Its inheritance is analogous to the Y chromosome with DNA inherited

<sup>1</sup>MRC Integrative Epidemiology Unit (IEU), University of Bristol, Barley House, Oakfield Grove, BS8 2BN, Bristol, UK. <sup>2</sup>Genetic Epidemiology Group, Department of Health Sciences, University of Leicester, Centre for Medicine University Road, LE1 7RH, Leicester, UK. Correspondence and requests for materials should be addressed to E.S. (email: [e.stergiakouli@bristol.ac.uk](mailto:e.stergiakouli@bristol.ac.uk))



**Figure 1.** Sexual dimorphism across psychiatric and behavioural traits (N>4000).

unchanged, apart from mutations, from the mother. As for the Y chromosome, haplogroups are used to study its association with disease meaning that mitochondrial DNA is often omitted from GWAS<sup>17</sup>. Mutations of mitochondrial DNA can cause neuromuscular disorders which affect around 1 in 5000 individuals<sup>18</sup>. Increased co-morbidity has been reported between these disorders and psychiatric illness<sup>19</sup> while there have been reports that mitochondrial DNA variation can increase or decrease risk to complex diseases, such as schizophrenia and Parkinson's disease<sup>20</sup>. However, there are no replicated associations between common variation on the mitochondrial DNA and specific psychiatric disorders, to date<sup>19</sup>.

Y chromosome and mitochondrial DNA haplogroups could influence behavioural traits in children from the general population and so play a role in the observed sexual dimorphism but this has not been previously investigated. In our study, we derived Y chromosome and mitochondrial DNA haplogroups in children from the Avon Longitudinal Study of Parents and Children (ALSPAC) and tested whether haplogroups are associated with parent-reported behavioural traits. Findings could have implications for the role of haplogroups in sexual dimorphism.

## Results

Y chromosome and mitochondrial DNA haplogroups were derived from children in the general population. The most common major Y chromosome haplogroups in ALSPAC males were R (72%) and I (19%) Y-chromosome haplogroups. The distribution of major Y chromosome haplogroups in ALSPAC was consistent with available data from England (Supplementary Figure 1). The most common major mitochondrial DNA haplogroups in ALSPAC were H (49%), U (13%), J (11%) and T (10%). The distribution of the major mitochondrial-DNA haplogroups in ALSPAC was consistent with available data from England (Supplementary Figure 1).

We investigated whether behavioural and psychiatric traits differed between males and females. The majority of traits showed sex differences with the exception of psychosis measured at 14 (Fig. 1); there was a higher proportion of males than of females with high number of behavioural and psychiatric traits across all categories. There was one exception, psychotic experiences, which were more commonly reported in females. The largest difference was for attention/activity trait, with 38.1% of males scoring high compared to 24.5% of females (P < 0.001).

**Primary analysis of major haplogroups.** The primary analysis involved Y chromosome and mitochondrial DNA major haplogroups. Weak association was found between major Y chromosome haplogroups and a high or low classification of number of behavioural and psychiatric traits (Table 1) or the raw number of traits (Table S2).

Weak evidence was found for association between the major mitochondrial-DNA haplogroups and a high or low classification of number of behavioural and psychiatric traits (Table 2) or the raw number of traits (Table S3).

**Secondary analysis of major haplogroup subgroups.** To further investigate for potential associations, we considered subgroups of the major Y chromosome haplogroups R and I (Table 3) in the secondary analysis. Y chromosome haplogroups and subgroups were not associated with categorical (Table S4) or continuous variables (Table S5) of behavioural and psychiatric traits.

We considered subgroups of the larger mitochondrial-DNA haplogroups H, J, K, T and U (Table 3) in the secondary analysis. Weak evidence was found of association between these haplogroups and subgroups and categorical (Table S6) or continuous variables (Table S7) of behavioural and psychiatric scores.

## Discussion

Research on possible factors contributing to sex differences in behavioural and psychiatric traits has been limited<sup>7</sup> and this is the first investigation of the association of Y chromosome and/or mitochondrial DNA haplogroups with behavioural and psychiatric traits in children from the general population. We report here that there were more males than females with higher scores on behavioural and psychiatric trait scales. However, psychotic experiences were more common in females than males at two different time points. The haplogroup structure of Y

Odds ratios of major Y chr haplogroup on high psychiatric trait category: OR (95% C.I.) R haplogroup is the reference			
Behavioural/ Psychiatric trait score	I	Other	Adjusted P values
Attention/ Activity symptoms score (DAWBA) N = 2400	1.08 (0.88, 1.33)	1.46 (1.02, 2.09)	0.10
Awkward behaviours score (DAWBA) N = 2390	1.06 (0.86, 1.30)	1.31 (0.92, 1.86)	0.31
Troublesome behaviours score (DAWBA) N = 2396	0.98 (0.79, 1.20)	1.03 (0.72, 1.49)	0.95
SCDC N = 2397	0.98 (0.76, 1.27)	1.28 (0.84, 1.96)	0.50
Hyperactivity traits (SDQ) N = 2461	1.06 (0.86, 1.31)	1.12 (0.78, 1.62)	0.74
Conduct traits (SDQ) N = 2462	1.34 (1.10, 1.63)	1.11 (0.78, 1.56)	0.014 > 0.006 <sup>1</sup>
Emotional symptoms (SDQ) N = 2462	1.05 (0.83, 1.33)	1.20 (0.81, 1.79)	0.64
Total behavioural traits (SDQ) N = 2459	1.01 (0.80, 1.27)	1.30 (0.89, 1.89)	0.41
PLIKSi age 14 N = 2062	0.89 (0.64, 1.22)	0.94 (0.53, 1.65)	0.76
PLIKSi age 18 N = 1270	0.99 (0.59, 1.67)	1.20 (0.50, 2.90)	0.92

**Table 1.** Odds ratios of major Y chromosome haplogroups on binary behavioural trait measures from logistic regression adjusted for paternal social class and GCSE results. 95% CI: 95% Confidence Interval; DAWBA: Development and Well-Being Assessment; SCDC: Social and communication disorders checklist; SDQ: Strengths and Difficulties Questionnaire; PLIKSi: Psychosis-Like Symptom Interview. <sup>1</sup>Smallest adjusted p value compared to multiple testing threshold.

Odds ratios of major mitochondrial DNA haplogroup on high psychiatric trait category: OR (95% C.I.) HV haplogroup is the reference						
Behavioural/ Psychiatric trait score	J	K	TR	U	Other	Adjusted P values
Attention/ Activity symptoms score (DAWBA) N = 3857	1.02 (0.80, 1.29)	0.84 (0.64, 1.10)	1.04 (0.82, 1.31)	1.00 (0.80, 1.25)	0.86 (0.65, 1.15)	0.72
Awkward behaviours score (DAWBA) N = 3841	0.93 (0.74, 1.17)	0.92 (0.72, 1.18)	0.90 (0.72, 1.13)	0.97 (0.78, 1.20)	0.84 (0.64, 1.10)	0.79
Troublesome behaviours score (DAWBA) N = 3843	0.88 (0.69, 1.11)	0.99 (0.76, 1.28)	0.99 (0.79, 1.25)	1.03 (0.83, 1.27)	1.15 (0.88, 1.51)	0.74
SCDC N = 3850	0.86 (0.63, 1.16)	0.84 (0.60, 1.18)	0.86 (0.64, 1.16)	0.96 (0.73, 1.26)	0.76 (0.52, 1.09)	0.59
Hyperactivity traits (SDQ) N = 3926	0.89 (0.71, 1.11)	1.02 (0.80, 1.30)	1.08 (0.87, 1.35)	0.93 (0.75, 1.14)	1.07 (0.82, 1.38)	0.73
Conduct traits (SDQ) N = 3925	1.21 (0.98, 1.50)	0.89 (0.70, 1.12)	1.02 (0.82, 1.26)	0.93 (0.76, 1.13)	0.92 (0.71, 1.18)	0.30
Emotional symptoms (SDQ) N = 3927	1.10 (0.87, 1.38)	0.84 (0.64, 1.11)	0.92 (0.72, 1.17)	0.95 (0.76, 1.19)	0.79 (0.59, 1.06)	0.52
Total behavioural traits (SDQ) N = 3922	1.26 (0.98, 1.63)	1.11 (0.83, 1.47)	1.19 (0.92, 1.54)	1.11 (0.87, 1.41)	0.77 (0.55, 1.08)	0.14
PLIKSi age 14 N = 3426	1.09 (0.78, 1.51)	0.88 (0.61, 1.27)	1.14 (0.84, 1.56)	0.70 (0.50, 0.98)	0.80 (0.53, 1.22)	0.19
PLIKSi age 18 N = 2369	0.95 (0.57, 1.59)	1.07 (0.64, 1.78)	1.65 (1.09, 2.50)	1.15 (0.75, 1.79)	0.47 (0.21, 1.03)	0.05 > 0.006 <sup>1</sup>

**Table 2.** Odds ratios of major Mitochondrial DNA haplogroups on binary behavioural trait measures from logistic regression adjusted for maternal and paternal social classes and GCSE results. 95% CI: 95% Confidence Interval; DAWBA: Development and Well-Being Assessment; SCDC: Social and communication disorders checklist; SDQ: Strengths and Difficulties Questionnaire; PLIKSi: Psychosis-Like Symptom Interview. <sup>1</sup>Smallest adjusted p value compared to multiple testing threshold.

chromosome and mitochondrial DNA in ALSPAC is consistent with a population of English origin. Y chromosome or mitochondrial DNA haplogroups were not associated with behavioural and psychiatric traits in ALSPAC.

There is convincing evidence that the Y chromosome is implicated in regulating differences in behaviour between the sexes in animal models. However, the biological mechanisms behind this are poorly understood. The difficulty in including the Y chromosome in GWAS is not contributing to improving understanding of the role of Y chromosome. The most convincing report of an association of the Y chromosome with a sexually dimorphic phenotype has been found in relation to coronary heart disease. In a cohort of British men haplogroup I was associated with a 50% increase in risk of cardiovascular disease with differential expression of macrophages between haplogroup I and other haplogroups<sup>21</sup>.

Behavioural problems and psychiatric disorders in children and adults are highly sexually dimorphic. However, there are few reports of Y chromosome haplogroup investigations in psychiatric phenotypes. A previous study<sup>22</sup> found weak evidence of an association between Y chromosome haplogroups and autism and our findings are consistent with this. In a study of Y chromosome haplogroups and aggression in a cohort of men from Pakistan, there were no differences in mean scores of aggression across the five different haplogroups captured in the study<sup>23</sup> while there are reports of patients with Turner syndrome (XO) and Klinefelter syndrome (XXY) exhibiting higher rates of autism and there is weak evidence that individuals with 47, XYY syndrome have

Y chromosome		Mitochondrial DNA	
Haplogroup	Count (%)	Haplogroup	Count (%)
<b>R</b>	<b>3452 (72.1)</b>	<b>H</b>	<b>4109 (49.9)</b>
R1a1	210 (4.4)	H1	1592 (19.4)
R1b1	328 (6.9)	H2	826 (10.0)
R1b1b2	1528 (31.9)	H3	387 (4.7)
R1b1b2g	1061 (22.2)	H4	168 (2.0)
R1b1b2h	296 (6.2)	H5	252 (3.1)
Q and other R	29 (0.6)	H6	235 (2.9)
<b>IJ</b>	<b>1039 (21.7)</b>	Other HV	649 (7.8)
I1	460 (9.6)	<b>J</b>	<b>885 (10.8)</b>
I2	150 (3.1)	J1	717 (8.7)
I2b	228 (4.8)	J2	168 (2.1)
Other I	71 (1.5)	<b>K</b>	<b>706 (8.6)</b>
J	130 (2.7)	K1	569 (6.9)
<b>Other</b>	<b>297 (6.2)</b>	Other K	137 (1.7)
E	143 (3.0)	<b>T</b>	<b>845 (10.3)</b>
CFGH	110 (2.3)	T2	668 (8.1)
LTNO	44 (0.9)	Other TR	177 (2.2)
		<b>U</b>	<b>1084 (13.2)</b>
		U4	165 (2.0)
		U5	722 (8.8)
		Other U	197 (2.4)
		<b>Other</b>	<b>591 (7.2)</b>
		M (M, C, D, L)	67 (0.8)
		N (N, A, I, W, X)	524 (6.4)
<b>Total:</b>	<b>4788</b>	<b>Total:</b>	<b>8220</b>

**Table 3.** Major haplogroups and subgroups.

a higher risk of developing bipolar disorder and schizophrenia<sup>13,15</sup>. Mitochondrial DNA has been investigated in relation to schizophrenia<sup>24</sup> and bipolar disorder<sup>25</sup> although conclusions have been limited with small sample sizes and inconsistent reporting<sup>26</sup>.

We investigated Y chromosome and mitochondrial DNA haplogroups in children from a sample representative of the general population. The main strength of our study was the use of a large population of children from a homogeneous population. The ALSPAC study participants come from a small geographic region in the South West of England. Only participants of European origin were included in the study as part of the quality control protocol of the GWAS. All of our analyses were performed within the same population and did not require a control group, which could have resulted in issues with population stratification due to imperfect matching of cases and controls. This can be an important issue in Y chromosome analysis because even populations from the same country can have different Y chromosome haplogroup structure. We were also able to access data on different behavioural and psychiatric trait scales, such as the SDQ and DAWBA, which have been used extensively in population cohorts for research purposes.

One of the limitations of our study was the small number of individuals in some of the haplogroups. However, this is expected due to the rarity of certain Y chromosome haplogroups in European populations. Another limitation is that the effect of genetic variation on psychiatric traits may vary at different ages and may appear during adolescence which we were not able to test due to the data not being available. Another challenge in the study was presented by the distribution of behavioural and psychiatric traits in the general population. A large number of individuals did not have any symptoms. Therefore, the distribution of the traits was exceedingly left skewed and transforming to something approaching normality was not possible. To overcome this, we dichotomized the variables and also performed sensitivity analyses with raw continuous scores. Also, we only investigated common variants and not copy-number variation or rare variants.

Although we do not report an association between the Y chromosome haplogroups and mitochondrial DNA haplogroups and risk of behavioural and psychiatric problems in the general population, we consider that our study has important implications; it highlights that if there are common variants on the Y chromosome and mitochondrial DNA associated with behavioural problems, they are of small effect size and would require large collaborative efforts to be identified. Other possible explanations for sexual dimorphism in psychiatric traits are gene-environment interactions<sup>27</sup> and gene-sex interactions which were not investigated in this study.

Given that the DNA markers on the Y chromosome and mitochondrial DNA are available on most GWAS chips but have not been largely utilised, such a collaborative effort would involve extracting relevant data from multiple studies rather than new genotyping efforts and could be an important step in understanding non-autosomal variation. Collaboration would be even more important for the Y chromosome because only half of any genotyped sample is usually male and can be used for analysis.

## Methods

**Study Participants.** The Avon Longitudinal Study of Parents and Children (ALSPAC) is a prospective birth cohort which recruited pregnant women with expected delivery dates between April 1991 and December 1992 from Bristol UK. 14,541 pregnant women were initially enrolled with 14,062 children born. Detailed information on health and development of children and their parents were collected from regular clinic visits and completion of questionnaires. A detailed description of the cohort has been published previously<sup>28, 29</sup>. The study website contains details of all the data that is available through a fully searchable data dictionary: <http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/>. The study was carried out in accordance with the ethical standards of the Helsinki Declaration and ethical approval was obtained from the ALSPAC Law and Ethics Committee and the Local Ethics Committees. Informed consent was obtained from all subjects.

**Genetic measures.** *ALSPAC genotyping.* A total of 9,912 participants were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping platform by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. PLINK software (v1.07) was used to carry out quality control (QC) measures<sup>30</sup>. After QC, 8,365 unrelated individuals were available for analysis. Information on QC performed is contained in the Supplementary Methods.

*Y chromosome.* For Y-DNA haplogroup determination, the Y-chromosomal SNPs of all 5,085 male participants in the dataset were used. The pseudo-autosomal SNPs (coded as chromosome 25) were removed from the analysis using the PLINK software package<sup>30</sup>. The resulting Y chromosomal genotypes (816 SNPs) of each individual were then piped in to the Y-Fitter (v0.2) software (maps genotype data to the Y-DNA phylogenetic tree built by Karafet *et al.*<sup>31</sup>, available online at [sourceforge.net/projects/yfitter](http://sourceforge.net/projects/yfitter)) and their respective Y-DNA haplogroup was determined. After removal of individuals with ‘False’ haplogroup determinations (i.e. ones which did not have enough SNPs to reliably determine haplogroup), we were left with 5,080 individuals. Remaining individuals with a haplogroup result which began with the letter R (e.g. R1b1) were clustered in to a single group named ‘Clade R’ and likewise the same was done with the haplogroups beginning with the other letters. Major haplogroups were defined as containing multiple haplogroups that are closely related to utilize information on less common haplogroups. After QC our data-set contained 4,788 males with derived Y chromosome haplogroups.

*Mitochondrial DNA.* 7,554 custom mitochondrial probes, targeting 2,824 unique mtDNA positions, were included on the Illumina HumanHap550 quad genome-wide SNP genotyping platform. A total of 1062 probes passed QC for the batch that was genotyped by Laboratory Corporation of America (n = 7590), whilst 629 probes passed QC for the batch genotyped by the Sanger Institute (n = 775). Haplogroup assignment was performed as described by Kloss-Brandstatter *et al.*<sup>32</sup>, and samples with a quality score of more than 90% were used for our analysis. Major haplogroups were defined as containing multiple haplogroups that are closely related to utilize information on less common haplogroups. After QC and removing individuals with withdrawn consent; our data-set contained 4,211 males and 4,009 females with derived mitochondrial DNA haplogroups. More information on how the mitochondrial DNA haplogroups were derived is contained in the Supplementary Methods.

**Behavioural measures.** *Strengths and Difficulties Questionnaire.* The Strengths and Difficulties Questionnaire (SDQ) is a behavioural screening questionnaire with high reliability and validity which includes questions on five domains: emotional symptoms, conduct problems, hyperactivity symptoms, peer relationship problems and prosocial behaviours ranging from 0 to 10 each<sup>33</sup>. A total difficulties score can be obtained by summing the following subscales: emotional symptoms, conduct problems, hyperactivity score and peer relationship problems (not including prosocial behaviour for which a higher score indicates less behavioural problems). SDQ total difficulties score ranges from 0 to 40 ([www.sdqinfo.com](http://www.sdqinfo.com)). Mother’s reports on their children’s behavioural problems were obtained at 7 years of age to maximise the number of individuals available for analysis. Cut-offs were used to dichotomise each of the five SDQ subscales as well as total difficulties score according to previous studies.

*Development and Well-Being Assessment.* The Development and Well-Being Assessment (DAWBA) is a semi-structured interview used to diagnose psychiatric disorders in children<sup>34</sup> and calculate the number of psychiatric disorder traits. Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional/Defiant Disorder (ODD) and Conduct Disorder (CD) traits were assessed in ALSPAC when the participants were aged 7 years and 7 months old using the parent-completed DAWBA attention/activity symptoms score, awkward behaviours score and troublesome behaviours score respectively. For each item, parents marked boxes to say whether their child showed the behaviour; these were coded 0 for “no,” 1 for “a little more than others,” and 2 for “a lot more than others.” A total trait score was calculated by summing these responses. Scores on measures with <30% missing items were mean imputed.

*Skuse social cognition score.* The Skuse social cognition scale (SCDC)<sup>35</sup> is designed to summarise the main features of the social cognition behaviour of individuals and was employed to assess autistic traits in children from ALSPAC. This was in the form of a parent-completed screening questionnaire on the child at 7 years and 7 months.

*Psychosis-Like Symptom Interview.* The semi-structured Psychosis-Like Symptom Interview (PLIKSi)<sup>36, 37</sup>, which draws on principles of standardized clinical examination developed for the Schedule for Clinical Assessment in



Neuropsychiatry (SCAN), was used to assess psychotic experiences at ages 12 and 18 years. The PLIKSi allows rating of 12 psychotic experiences including hallucinations (visual and auditory), delusions (spied on, persecution, thoughts read, reference, control, grandiosity, other) and experiences of thought interference (broadcasting, insertion and withdrawal). The interviewers were psychology graduates trained in assessment using the SCAN psychosis section and using the PLIKSi. Psychotic experiences were rated as not present, suspected or definitely psychotic. Unclear responses were always 'rated down' and symptoms were rated as definite only when a clear example was provided. Individuals were deemed as having a psychotic experience if rated as having 1 or more definite or suspected psychotic experience compared to no psychotic experiences.

**Statistical Analysis.** Most of the continuous variables were heavily zero-skewed; the majority of participants scored zero. We used categorical variables as appropriate normalisation was not possible in the majority of cases. Cut-offs were used to dichotomise each of the variables according to previous studies and the distribution of the variables<sup>33</sup>. The attention/activity symptoms score and SCDC were stratified into ( $\leq 5$  &  $> 5$ ), the awkward behaviours score, troublesome behaviours score and two PLIKS measures were stratified into ( $\leq 0$  &  $> 0$ ), the hyperactivity traits score and emotional symptoms score were stratified into ( $\leq 2$  &  $> 2$ ), the conduct traits score was stratified into ( $\leq 1$  &  $> 1$ ) and the total behavioural score was stratified into ( $\leq 10$  &  $> 10$ ). Analysis on both categorical and continuous variables are presented for comparison.

Chi squared tests were used to test for sex differences in categorical variables with linear regression used for continuous variables. Chi squared tests were used to test for associations between haplogroups and categorical variables with logistic regression employed when confounders were included. As a sensitivity analysis, multiple regression was used to test for association between haplogroups and the original continuous variables whilst adjusting for confounders. The tetrachoric correlation between the binary variables was computed to determine if the variables were independent with a 0.8 cut-off. The Bonferroni-Holm method was used to adjust the number of independent variables for multiple testing<sup>38</sup>. No pairwise correlation exceeded the threshold of 0.8 therefore the variables were treated as independent (Table S1) and multiple testing correlation was performed assuming 10 independent variables.

The pattern of inheritance means that individuals with different haplogroups may differ phenotypically, specifically in terms of social class and outcomes associated with it, such as education. The most appropriate variables to gauge social class of a child would be the social class and the educational level of the parents. General certificate of secondary education (GCSE) results were used as a measure of a child's cognitive abilities. The Y chromosome is inherited from the father; therefore paternal social class was included in the Y chromosome analysis. Mitochondrial DNA is inherited from the mother; thus maternal social class was included in the analysis along with paternal social class for mitochondrial DNA. Y chromosome analysis was only performed on male participants while male and female participants were included in mitochondrial DNA analysis.

Initially just the major haplogroups e.g. "R" and "I" for the Y chromosome and "H" and "U" for the mitochondrial DNA were analysed but subsequent analysis investigated subgroups of these major haplogroups. Subgroups were investigated if their frequency was  $> 2\%$ . Y chromosome and mitochondrial DNA haplogroups were grouped together using online phylogenetic trees<sup>39,40</sup>.

Pie-charts were created using ALSPAC haplogroup data as well as haplogroup data for England from Eupedia ([www.eupedia.com](http://www.eupedia.com)) and STATA 13 was used for all data analysis.

## References

- Ober, C., Loiselle, D. A. & Gilad, Y. Sex-specific genetic architecture of human disease. *Nature Reviews Genetics* **9**, 911–922 (2008).
- Rucklidge, J. J. Gender differences in attention-deficit/hyperactivity disorder. *Psychiatric Clinics of North America* **33**, 357–373 (2010).
- Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B. & Baron-Cohen, S. Sex/gender differences and autism: setting the scene for future research. *Journal of the American Academy of Child & Adolescent Psychiatry* **54**, 11–24 (2015).
- Ochoa, S., Usall, J., Cobo, J., Labad, X. & Kulkarni, J. Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review. *Schizophrenia research and treatment* **2012** (2012).
- Baron-Cohen, S. The extreme male brain theory of autism. *Trends in cognitive sciences* **6**, 248–254 (2002).
- Badcock, C. & Crespi, B. Battle of the sexes may set the brain. *Nature* **454**, 1054–1055 (2008).
- Taylor, M. J. *et al.* Is There a Female Protective Effect Against Attention-Deficit/Hyperactivity Disorder? Evidence From Two Representative Twin Samples. *Journal of the American Academy of Child & Adolescent Psychiatry* **55**, 504–512. e502 (2016).
- Levy, F., Hay, D. A., McSTEPHEN, M., Wood, C. & Waldman, I. Attention-deficit hyperactivity disorder: a category or a continuum? Genetic analysis of a large-scale twin study. *Journal of the American Academy of Child & Adolescent Psychiatry* **36**, 737–744 (1997).
- Stergiakouli, E. *et al.* Shared genetic influences between attention-deficit/hyperactivity disorder (ADHD) traits in children and clinical ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry* **54**, 322–327 (2015).
- Robinson, E. B. *et al.* Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nature genetics* **48**, 552–555 (2016).
- Xu, J., Burgoyne, P. S. & Arnold, A. P. Sex differences in sex chromosome gene expression in mouse brain. *Human molecular genetics* **11**, 1409–1419 (2002).
- Gatewood, J. D. *et al.* Sex chromosome complement and gonadal sex influence aggressive and parental behaviors in mice. *The Journal of neuroscience* **26**, 2335–2342 (2006).
- Mors, O., Mortensen, P. & Ewald, H. No evidence of increased risk for schizophrenia or bipolar affective disorder in persons with aneuploidies of the sex chromosomes. *Psychological Medicine* **31**, 425–430 (2001).
- DeLisi, L. E. *et al.* Schizophrenia and sex chromosome anomalies. *Schizophrenia Bulletin* **20**, 495–505 (1994).
- Werling, D. M. & Geschwind, D. H. Sex differences in autism spectrum disorders. *Current opinion in neurology* **26**, 146 (2013).
- Consortium, Y. C. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome research* **12**, 339–348 (2002).
- Chinnery, P. F. & Schon, E. A. Mitochondria. *Journal of Neurology, Neurosurgery & Psychiatry* **74**, 1188–1199 (2003).
- Bannwarth, S. *et al.* Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders. *Journal of medical genetics*, jmedgenet-2013-101604 (2013).
- Anglin, R. E., Mazurek, M. F., Tarnopolsky, M. A. & Rosebush, P. I. The mitochondrial genome and psychiatric illness. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **159**, 749–759 (2012).
- Hudson, G., Gomez-Duran, A., Wilson, I. J. & Chinnery, P. F. Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *PLoS Genet* **10**, e1004369 (2014).

21. Charchar, F. J. *et al.* Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *The Lancet* **379**, 915–922 (2012).
22. Jamain, S. *et al.* Y chromosome haplogroups in autistic subjects. *Molecular psychiatry* **7**, 217 (2002).
23. Shah, S. S., Ayub, Q., Firasat, S., Kaiser, F. & Mehdi, S. Y haplogroups and aggressive behavior in a Pakistani ethnic group. *Aggressive behavior* **35**, 68–74 (2009).
24. Wang, G.-x. *et al.* Mitochondrial haplogroups and hypervariable region polymorphisms in schizophrenia: a case-control study. *Psychiatry research* **209**, 279–283 (2013).
25. Kato, T., Kunugi, H., Nanko, S. & Kato, N. Mitochondrial DNA polymorphisms in bipolar disorder. *Journal of affective disorders* **62**, 151–164 (2001).
26. Anglin, R. E., Garside, S. L., Tarnopolsky, M. A., Mazurek, M. F. & Rosebush, P. I. The psychiatric manifestations of mitochondrial disorders: a case and review of the literature. *The Journal of clinical psychiatry* **73**, 506–512 (2012).
27. Caspi, A. & Moffitt, T. E. Gene–environment interactions in psychiatry: joining forces with neuroscience. *Nature Reviews Neuroscience* **7**, 583–590 (2006).
28. Boyd, A. *et al.* Cohort profile: the ‘children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. *International journal of epidemiology*, dys064 (2012).
29. Fraser, A. *et al.* Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International journal of epidemiology* **42**, 97–110 (2013).
30. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* **81**, 559–575 (2007).
31. Karafet, T. M. *et al.* New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome research* **18**, 830–838 (2008).
32. Kloss-Brandstätter, A. *et al.* HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Human mutation* **32**, 25–32 (2011).
33. Goodman, R., Ford, T., Simmons, H., Gatward, R. & Meltzer, H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry* **177**, 534–539 (2000).
34. Goodman, R., Ford, T., Richards, H., Gatward, R. & Meltzer, H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of child psychology and psychiatry* **41**, 645–655 (2000).
35. Skuse, D. *et al.* Evidence from Turner’s syndrome of an imprinted X-linked locus affecting cognitive function. *Nature* (1997).
36. Horwood, J. *et al.* IQ and non-clinical psychotic symptoms in 12-year-olds: results from the ALSPAC birth cohort. *British Journal of Psychiatry* **193**, 185–191, doi:10.1192/bjp.bp.108.051904 (2008).
37. Zammit, S. *et al.* Psychotic experiences and psychotic disorders at age 18 in relation to psychotic experiences at age 12 in a longitudinal population-based cohort study. *American Journal of Psychiatry* **170**, 742–750, doi:10.1176/appi.ajp.2013.12060768 (2013).
38. Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65–70 (1979).
39. <http://www.isogg.org/tree/>. Y-DNA Haplogroup Tree 2016, Version: 11.100 (2016).
40. Van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human mutation* **30**, E386–E394 (2009).

## Acknowledgements

This work was supported by the Medical Research Council and the University of Bristol [MRC Integrative Epidemiology Unit (IEU) MC\_UU\_12013/1–9]. The authors are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at the Wellcome Trust Sanger Institute and LabCorp (Laboratory Corporation of America) supported by 23andMe. We thank Drs. David Evans, Beate St Pourcain, Susan Ring, John Kemp and George McMahon for their contribution to QC of ALSPAC GWAS data. We especially thank Dr. John Kemp for his contribution to the QC of Y chr genetic data.

## Author Contributions

L.H. performed the statistical analysis on the haplogroups and psychiatric traits and M.E. derived the Y chromosome and mitochondrial D.N.A. haplogroups under the guidance of G.D.S., S.R. and E.S. L.H. and M.E. wrote the manuscript under the guidance of G.D.S., S.R. and E.S. All authors read and gave valuable suggestions on the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:10.1038/s41598-017-10871-4

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017