Biological Psychiatry: Celebrating 50 Years

Archival Report

Evidence of Assortative Mating in Autism Spectrum Disorder

Siobhan Connolly, Richard Anney, Louise Gallagher, and Elizabeth A. Heron

ABSTRACT

BACKGROUND: Assortative mating is a nonrandom mating system in which individuals with similar genotypes and/or phenotypes mate with one another more frequently than would be expected in a random mating system. Assortative mating has been hypothesized to play a role in autism spectrum disorder (ASD) in an attempt to explain some of the increase in the prevalence of ASD that has recently been observed. ASD is considered to be a heritable neurodevelopmental disorder, but there is limited understanding of its causes. Assortative mating can be explored through both phenotypic and genotypic data, but up until now it has never been investigated through genotypic measures in ASD.

METHODS: We investigated genotypically similar mating pairs using genome-wide single nucleotide polymorphism data on trio families (Autism Genome Project data [1590 parents] and Simons Simplex Collection data [1962 parents]). To determine whether or not an excess in genetic similarity was present, we employed kinship coefficients and examined spousal correlation between the principal components in both the Autism Genome Project and Simons Simplex Collection datasets. We also examined assortative mating using phenotype data on the parents to detect any correlation between ASD traits.

RESULTS: We found significant evidence of genetic similarity between the parents of ASD offspring using both methods in the Autism Genome Project dataset. In the Simons Simplex Collection, there was also significant evidence of genetic similarity between the parents when explored through spousal correlation.

CONCLUSIONS: This study gives further support to the hypothesis that positive assortative mating plays a role in ASD.

Keywords: Ancestral assortative mating, Assortative mating, Autism spectrum disorder (ASD), Genetic assortative mating, Kinship, Random mating

https://doi.org/10.1016/j.biopsych.2019.04.014

Assortative mating occurs when similar male and female individuals mate with each other more (or less) often than expected by chance (1). This similarity can be trait specific and/or genotype specific (2). When the trait has a genetic component (phenotype) and assortative mating occurs, both genetic and phenotypic assortative mating take place. Ancestral assortative mating occurs when spouse pairs are more (or less) likely to share genes of common ancestry (3). Studies have largely focused on phenotypic assortative mating, with fewer studies investigating genetic assortative mating (3–8). Here, we were interested in investigating genetic and phenotypic assortative mating in the complex disorder autism spectrum disorder (ASD).

ASD is considered to be a heritable (9) neurodevelopmental disorder characterized by patterns of repetitive behaviors and deficits in language and social behavior, but there is limited understanding of its causes. Assortative mating has been hypothesized in ASD because parents can often display characteristics of ASD. Baron-Cohen (10) noted that estimates of prevalence in 2006 [0.2% (11) and 0.44% (12)] were much higher than the traditional estimate of 0.04% and suggested that this change in prevalence could be due to individuals with ASD-like traits (high systemizers) mating with each other. A more recent estimate of the prevalence of ASD was much higher, with 1 in 68 children aged 8 years having an ASD diagnosis (1.47%) (13). Peyrot *et al.* (14) found that, in general, assortative mating could lead to a considerable increase in prevalence for disorders that are less common and are highly heritable.

Assortative mating has previously been investigated in ASD using a variety of approaches. One such approach has been to look at the educational and occupational phenotypes of parents and family members to see whether logical and systematic professions lead to an increase of ASD in the families or to an increase in severity of ASD traits among the family members (15–19). A recent study showed that systemizing traits were found to be genetically correlated with ASD (20). Nordsletten *et al.* (21) found evidence of nonrandom mating within and across 11 psychiatric disorders, including ASD, in a Swedish population that examined correlation in diagnostic status.

SEE COMMENTARY ON PAGE 250

 ISSN: 0006-3223

In addition, studies have also investigated the presence of ASD-like traits in parents of ASD offspring (22–25). It has been noted by many studies that parents of ASD offspring often present with ASD-like characteristics more so than expected (especially in multiplex [more than one affected individual per family] ASD families) (24,26–29).

There is evidence that assortative mating influences genotype frequencies that are associated with complex traits (24). If assortative mating for a trait is present and is not accounted for in a genetic study, it can confound heritability estimates (14,24,30–32). Evidence from Klei *et al.* (33) also supports this because these authors found an elevated level of heritability in pseudocontrols (nontransmitted alleles that would be expected to behave as unaffected offspring) from multiplex ASD families. If parents are more genetically similar at casual variants for ASD (i.e., positive assortative mating is taking place), then the nontransmitted alleles from the parents at these loci are more likely to be risk alleles and increase heritability estimates in pseudocontrols, although it has been argued that assortative mating may lead to only a modest increase of heritability estimates for a disorder (14).

Another concern is the assumption of random mating that is used when investigating genetic associations in ASD, which may be violated if assortative mating is taking place. Understanding the genetic etiology of ASD may be completely entwined with the understanding of assortative mating (34).

To examine genetic assortative mating in ASD, we investigated the genetic similarity of the parents (spouses) compared with the genetic similarity of nonspouse pairings (we created) restricted to male/female pairings from the same ancestral population using kinship coefficients. The investigation of the genetic similarity between parents with ASD was carried out using single nucleotide polymorphism (SNP) data from genome-wide association studies (GWASs). This approach has previously been used to examine assortative mating for traits such as height, body mass index, and education (5,7), but not to our knowledge in the context of ASD. An approach that has some similarity was used to investigate genetic assortative mating in ASD using pairwise genetic distance, in contrast to kinship coefficients, to investigate whether spouses were more related than randomly mated parents (35). Population stratification was not accounted for when estimating pairwise genetic distance, but comparisons were within a small subset of the data obtained from one population. Another study also examined ASD genetic assortative mating using summary statistics from an ASD GWAS but found no significant evidence owing to sample size and limitations of ASD known variants (8).

We also investigated ancestral assortative mating using principal components analysis (PCA) and investigated phenotypic assortative mating using ASD trait data of the parents. We carried out our analyses on two ASD GWAS datasets: the Autism Genome Project (AGP) dataset (36,37) and the Simons Simplex Collection (SSC) dataset (36).

METHODS AND MATERIALS

Data

The AGP GWAS trio family dataset was collected at sites across Europe and North America and is described elsewhere (36,37). Here we considered the stage 2 dataset consisting of 2931 multiplex and simplex (only one affected individual per family) families.

The SSC GWAS consists of data on 2591 simplex North American families [see (38–40) for further details on the SSC data]. The majority of AGP and SSC samples are of European ethnicity (see Supplemental Figures S17 and S18 for PCA plots of the offspring).

The AGP GWAS data include families grouped into two nested diagnostic categories, Strict ASD [autism diagnoses on both Autism Diagnostic Interview and Autism Diagnostic Observation Schedule instruments (41,42)] and Spectrum ASD (autism spectrum diagnosis on either the Autism Diagnostic Interview or Autism Diagnostic Observation Schedule instrument), as defined in (37). We applied the same ASD phenotype criteria as were used in the AGP (37) to define a Strict ASD phenotype within the SSC data. The Strict phenotype would be expected to have less clinical heterogeneity and, therefore, to have the potential to increase the power of identifying robust findings when compared with a broader autism diagnosis (43).

We also examined genetic assortative mating in a smaller sample comprising the Autism Genetic Resource Exchange (44). Owing to the size of the dataset that was available for analysis and to the fact that we were unable to filter families on ASD diagnostic criteria, details of the analysis of this dataset are available in the Supplement.

Quality Control Procedures

The quality control procedures follow a standard approach to trio GWAS quality control, with individuals and SNPs removed when missingness > 0.05, Hardy-Weinberg equilibrium p value $< .00001$, and minor allele frequency < 0.05 (see Figure 1 and Supplemental Table ST1). We removed families that were related and individuals with extreme levels of heterozygosity (> 2 SD). We also excluded certain linkage disequilibrium ranges that can result in confounding in certain analyses such as PCA when examining population stratification (45). We limited our analyses to complete trios because we needed both parents of the affected offspring for our analyses. After quality control, the AGP with a Strict phenotype contained 1590 trios and 712,319 SNPs, and the SSC with a Strict phenotype contained 1962 trios and 417,809 SNPs.

Statistical Methods

We used kinship coefficients to examine assortative mating in ASD. A kinship coefficient is the probability that two homologous alleles sampled at random from two individuals are identical by descent (which occurs when two DNA sequences are inherited from a common ancestor). We compared the distribution of the kinship coefficients for spouses (mother/father pairings) with the distribution of all other possible nonspouse pairings restricted to male/female pairings from the same ancestral background, which we refer to as all nonspouse pairs. This is similar to the approach taken in Domingue *et al.* (5), although we take into account population stratification using a different method. The nonspouse pairs were restricted to male/female pairings because this was an investigation of genetic assortative mating between parents of offspring with ASD, and genetic assortative mating may be

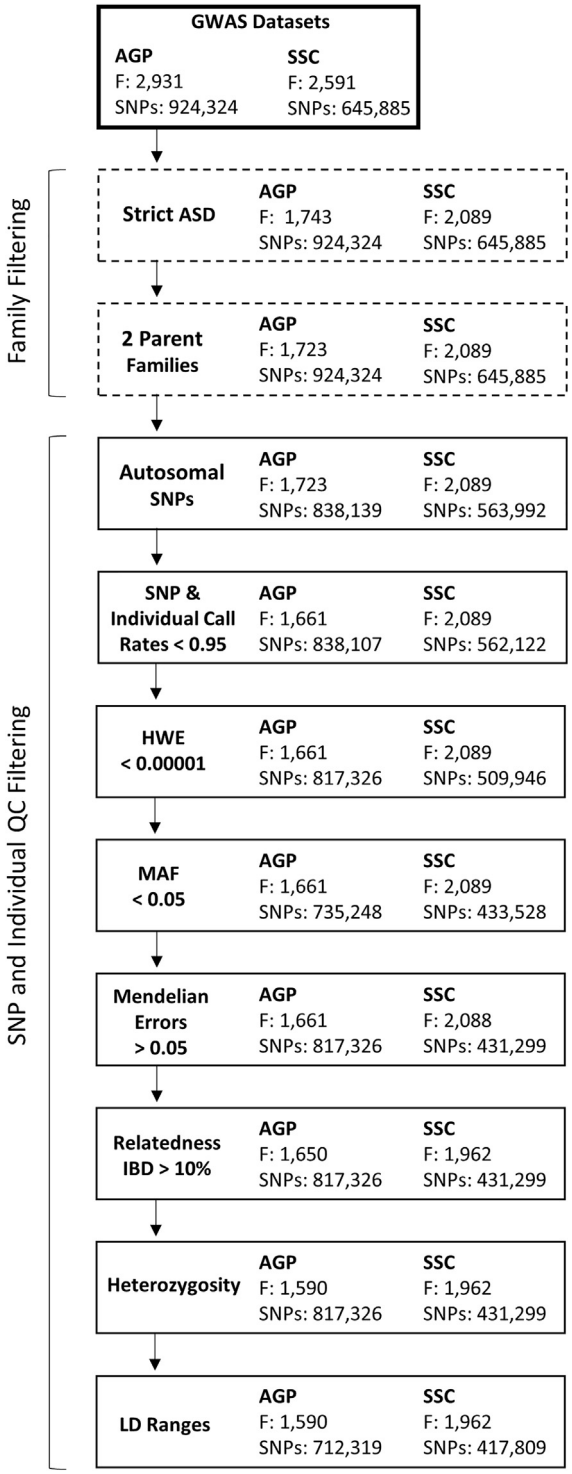

Figure 1. Quality control (QC) procedure for the Autism Genome Project (AGP) and Simons Simplex Collection (SSC) datasets. ASD, autism spectrum disorder; F, number of families; GWAS, genome-wide association study; HWE, Hardy-Weinberg equilibrium; IBD, identical by descent; LD, linkage disequilibrium; MAF, minor allele frequency; SNP, single nucleotide polymorphism.

different for same-sex couples (5). We filtered the nonspouse pairs on ancestral population because creating nonspouse pairings across different ethnicities will result in individuals who look less genetically similar when comparing them with spouses from the same ethnicity. This will lead to false positive findings for genetic assortative mating [see (46) and the Supplement].

The first step in our approach was to use the ADMIXTURE software (47) to estimate the amount of admixture (when individuals from genetically different populations mate and produce offspring) in the samples with the 11 populations in the HapMap3 dataset (48) as a reference (–supervised). Note the HapMap3 dataset used has had individuals removed owing to cryptic relatedness [see (49) for further details]. This allowed us to identify the different populations that are contained within the ASD datasets. We then removed spouses who did not mate within the same ancestral population, and only spouse pairs where the individuals had similar proportions of ancestry to each other and to others in their population remained (within 2 SD of the mean of the proportion for each ancestral population).

In the second step, within these ancestral populations, we compared the distribution of the spouses' kinship coefficients with the distribution of the nonspouse pairs' kinship coefficients, where all kinship coefficients were estimated using the KING software (50) (see Supplement). In this software, the kinship calculation accounts for heterogeneity between samples (for both the spouses and nonspouse pairs, see the Supplement). The quantiles (from 0.001 to 0.999 in increments of 0.001) for the spouse pairs' kinship coefficients were calculated and then mapped to the kinship coefficients for the nonspouse pairs; results were plotted in the same manner as Domingue *et al.* (5). The 45° line indicates the null hypothesis that the genetic similarity among spouse pairs matches the genetic similarity among nonspouse pairs. If the genetic similarity among spouses differs from the genetic similarity of nonspouse pairs, this is captured by departure from the 45° line and we calculate the area (and 95% confidence intervals [CIs] using 10,000 bootstrap replications) between this curve and the 45° line.

Before analyzing the ASD datasets, we tested this approach in a subset of the HapMap dataset (51) that contains 101 spouse pairs from three different populations, which we refer to as the HapMap Spouse dataset (see the Supplement). For this dataset, when fully accounting for population stratification, we found no significant evidence of genetic similarity among the spouses, as would be expected (see the Supplement).

The next approach investigated ancestral assortative mating using the method of Sebro *et al.* (3): examining the correlation between the principal components (PCs) from the PCA of the genetic datasets. The PCs reflect the population structure within a dataset due to ancestry (52), and the estimation of correlation between the spouses' PCs reflects the degree to which the spouses are mating within their ancestral populations (3). PCAs were carried out using the EIGENSTRAT program (53) on pruned sets of SNPs (using PLINK software with a window size of 50 SNPs, with the window shift set to 5

and an R^2 threshold of .25) with high call rates greater than 0.999 (AGP dataset 124,547 SNPs and SSC dataset 115,734 SNPs after pruning).

We also investigated assortative mating among the parents using ASD trait information obtained from the Broad Autism Phenotype Questionnaire (BAPQ) (specifically the subscales Aloof, Rigid, Pragmatic Language, and Total BAPQ) (54) and the Social Responsive Scale–Adult version (SRS-A) (specifically the subscales Awareness, Cognition, Mannerisms, Motivation, and Communication and the combined Total SRS score) (55). The BAPQ includes both self-report and informant-report versions that can be combined to give a more accurate result (Best Score), whereas the SRS-A is generally an informant-report for the spouse or partner. For the SSC dataset, BAPQ data are available for 1945 (99.1%) of the spouses and SRS-A data are available for 1953 (99.5%) of the spouses. However, the sample sizes were limited in the AGP, where 270 spouses (17.0%) have complete BAPQ data and 428 spouses (26.9%) have SRS-A data.

RESULTS

The Autism Genome Project

We investigated the genetic similarity of the parents in the AGP data by comparing the kinship coefficients of the spouses with the kinship coefficients of all other possible nonspouse pairs. We required that both individuals in each spouse pair and nonspouse pair were from the same ancestral population to avoid spurious results of assortative mating. We removed spouses who did not mate within the same ancestral population, resulting in 1092 spouse pairs that belonged to one of six populations (see Supplemental Figure S10).

The genetic similarity between spouse pairs compared with nonspouse pairs is shown in Figure 2. Here the intersection of the vertical and horizontal lines represents where the median value (vertical line) of genetic similarity among spouses corresponds to the 0.52996 quantile (horizontal line) of all other possible nonspouse pairs (within ancestral populations). The interpretation of this is that spouses are more genetically similar than all possible nonspouse pairs. To calculate the degree of this increase in similarity among spouses, we calculated the area of the shaded region above the 45° line (0.0251; 95% CI = 0.0110, 0.0389). Because this 95% CI does not contain 0 (the null hypothesis is that the area is 0, i.e., no assortative mating), this shows significant evidence that positive genetic assortative mating could be taking place in the AGP dataset.

Collection site information was available for the AGP spouses, and we undertook an additional analysis where we compared spouses with nonspouse pairs within the same site and ancestral population (see Figure 3). Here the genetic relatedness estimate at the 0.5 quantile (vertical line) of spousal pairs corresponds to the 0.5576 quantile (horizontal line) of all nonspouse pairs within the site and ancestral population. The shaded area gives an estimate of assortative mating and is equal to 0.0451 here (95% CI = 0.0304, 0.0600). This also indicates that the spouses are more genetically similar than would be expected.

We also examined ancestral assortative mating in the AGP dataset by estimating the correlation between the spouses' PCs (3). Figure 4 shows strong correlations between mothers and fathers for PC1 (this PC separates the Europeans from the non-Europeans; see Supplemental Figure S17 for a PCA plot of the offspring) at 0.843 (p value $<$ .0001). This shows significant evidence for spouse pairs mating within their ancestral populations and hence shows evidence of ancestral assortative mating. For results from PC2 and PC3, see the Supplement.

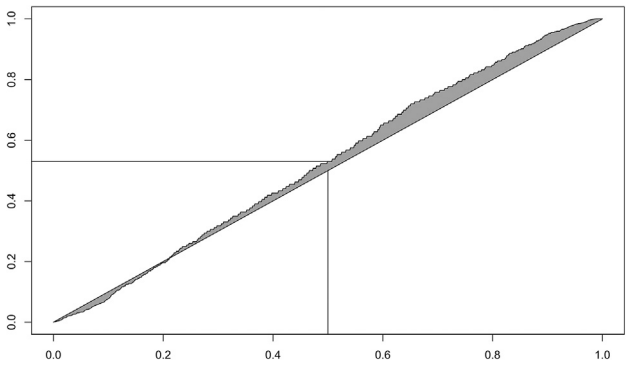

Figure 2. Assortative mating in Autism Genome Project dataset. The x-axis represents quantiles of the distribution of kinship coefficients between spouse pairs. The y-axis represents quantiles of the distribution of kinship coefficients between all other nonspouse pairs. The shaded area gives an estimate of assortative mating and is equal to 0.0251 here (95% confidence interval = 0.0110, 0.0389). The genetic relatedness estimate at the 0.5 quantile (vertical line) of spousal pairs corresponds to the 0.52996 quantile (horizontal line) of all other nonspouse pairs.

When investigating the correlation between mothers and fathers for ASD traits using the BAPQ and SRS-A, we found significant correlations between the parents on all ASD subscales for the AGP data apart from the Aloof subscale from the BAPQ and the Motivation subscale from the SRS (see Supplemental Tables ST2 and ST3).

The Simons Simplex Collection

The same procedures carried out on the AGP dataset were carried out on the SSC dataset. Because we wanted to

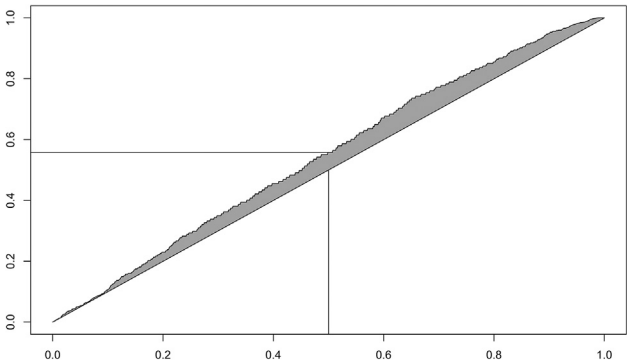

Figure 3. Assortative mating in Autism Genome Project dataset within site. The x-axis represents quantiles of the distribution of kinship coefficients between spouse pairs. The y-axis represents quantiles of the distribution of kinship coefficients between all other nonspouse pairs within the same site and ancestral population. The shaded area gives an estimate of assortative mating and is equal to 0.0451 here (95% confidence interval = 0.0304, 0.0600). The genetic relatedness estimate at the 0.5 quantile (vertical line) of spousal pairs corresponds to the 0.5576 quantile (horizontal line) of all other nonspouse pairs.

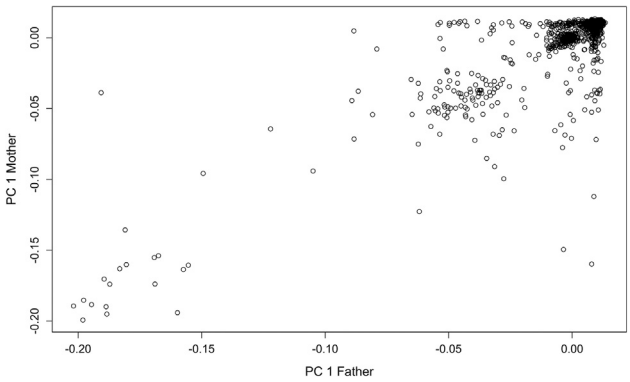

Figure 4. Principal component 1 (PC 1) of mothers vs. fathers in the Autism Genome Project dataset. The spousal correlation for PC 1 is .843.

compare spouses with all possible nonspouse pairs in the SSC dataset, we again removed spouses who did not mate within the same population, resulting in 1221 spouse pairs that belonged to one of six populations (see Supplemental Figure S11). We used the kinship coefficients to estimate the genetic similarity between spouse pairs compared with all nonspouse pairs (see Figure 5). The genetic relatedness estimate at the 0.5 quantile (vertical line) of spousal pairs corresponds to the 0.5032 quantile (horizontal line) of all nonspouse pairs. The shaded area gives an estimate of assortative mating and is equal to −0.0062 (95% CI = −0.0187, 0.0065). Unlike for the AGP data, we see a negative value for the area, indicating that the spouses are less similar than the nonspouse pairs. But this is not significant; therefore, there is no evidence of assortative mating in the SSC dataset using this method. There was no site data available for the SSC to conduct any site-related analyses.

We examined ancestral assortative mating in the SSC by analyzing the correlation between PC1 for the mothers and PC1 for the fathers (again, as for the AGP dataset, this principal component separates the Europeans from the non-Europeans; see Supplemental Figure S18 for PCA of the offspring). Figure 6 displays a strong correlation of 0.802 (p value $<$ .0001) between the mothers and fathers. For results from PC2 and PC3, see the Supplement. This shows significant evidence of ancestral assortative mating, similar to the results for the AGP dataset.

When investigating the correlation between mothers and fathers for ASD traits using the BAPQ in the SSC dataset, there was no evidence of correlation (see Supplemental Figure S24). With the SRS-A, we found significant correlations between the parents on all subscales except the Awareness and Motivation subscales for the SSC data, although the correlation values are lower than the AGP's SRS-A results (see Supplemental Table ST4).

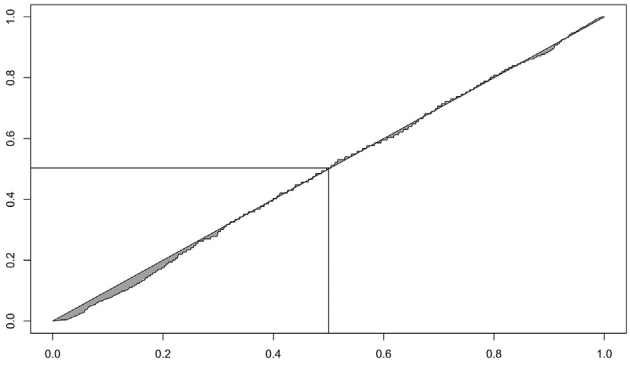

Figure 5. Assortative mating in Simons Simplex Collection dataset. The x-axis represents quantiles of the distribution of kinship coefficients between spouse pairs. The y-axis represents quantiles of the distribution of kinship coefficients between all other nonspouse pairs. The shaded area gives an estimate of assortative mating and is equal to −0.00622 here (95% confidence interval = −0.0187, 0.0065). The genetic relatedness estimate at the 0.5 quantile (vertical line) of spousal pairs corresponds to the 0.5032 quantile (horizontal line) of all nonspouse pairs.

DISCUSSION

Assortative mating has been hypothesized to play a role in ASD, suggesting that people with ASD-like traits mate with one another more frequently than would be expected. There has been evidence to suggest that assortative mating in ASD can affect the prevalence and heritability estimates (14,33).

We also tested this kinship coefficient approach on the HapMap Spouse dataset to assess whether we could account for population structure and avoid finding false evidence of assortative mating. We found no significant results for assortative mating in the HapMap Spouse dataset when restricting the nonspouse pairings to male/female pairings within their ancestral population (see the Supplement). If we did not take into account the population substructure correctly in this dataset, this led to very strong findings of assortative mating in the HapMap Spouse dataset (see the Supplement) (46). We do note that this dataset is relatively smaller than the two ASD datasets analyzed here and hence would be expected to have lower power to detect such effects. In addition, the ancestral populations are known in the HapMap Spouse dataset, unlike in the AGP and SSC datasets.

We examined kinship coefficients to investigate genetic assortative mating in the AGP and SSC datasets. We found significant evidence that the kinship coefficients for spouses were more similar when compared with those for nonspouse

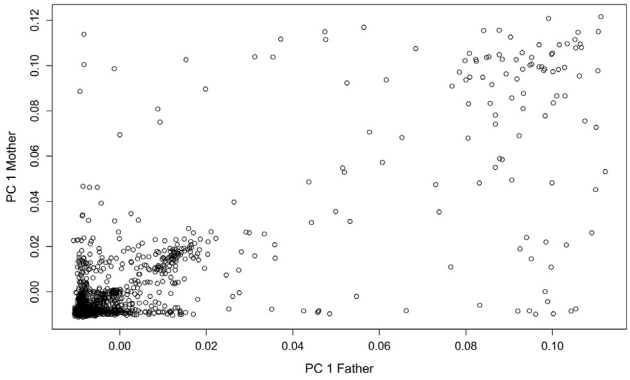

Figure 6. Principal component 1 (PC 1) of mothers vs. fathers in the Simons Simplex Collection dataset. The spousal correlation for PC 1 is 0.802.

pairs within the same ancestral population in the AGP data, although we did not find any evidence of this in the SSC data.

Because the spouse pairs and nonspouse pairs needed to be from the same ancestral population for our analyses, we retained only spouse pairs where the individuals had similar proportions of ancestry to each other and to others in their population, reducing our sample size. In particular, in the SSC dataset, 741 spouse pairs were removed (compared with 498 in the AGP). This, we suspect, would have reduced the power to detect the genetic similarity among the spouse pairs in the SSC dataset. Having stated this, when assortative mating is detected as in the AGP dataset, it is less likely to be confounded with population stratification, although we cannot fully rule this out.

Other populations have shown evidence of ancestral assortative mating when spousal correlation of the PCs on the genetic data has been investigated (3,28,56). We also present evidence of this in the two ASD datasets by identifying a significant correlation between the spouses' PCs.

When investigating the ancestral assortative mating, it is impossible to tease apart how much of the spousal correlation is attributed to proximity of the spouses with each other. For instance, in the AGP dataset that is collected at many sites across Europe and North America, some countries will have less admixture than others and it is more likely that individuals from these populations will mate with other individuals from the same ancestral background based on proximity. The SSC data, on the other hand, were collected only at sites in North America, which has more admixture between different ethnicities, yet we still see strong spousal correlations on the PCs.

For the AGP dataset, where collection site data were available (not available for the SSC data), we randomly paired the spouses within ancestral population and collection site as a proxy for proximity, obtaining significant findings here also (see Figure 3). This gives us more confidence that we are accounting for as much of the population stratification as possible, but we acknowledge that subtle population stratification may still be present.

We found no significant evidence of assortative mating in the SSC dataset using the kinship coefficients approach. In addition, for the phenotypic analyses, in examining the ASD traits in the parents through the BAPQ and SRS-A, the findings in the SSC dataset were not as strong as those in the AGP dataset. These findings could be due to the differences in the ascertainment for the AGP and SSC datasets. It is worth noting that the SSC dataset had strict recruitment criteria, including only simplex families in the study. Furthermore, the parents of these families were also screened in the recruitment process for ASD traits, a design that inherently enriches for rare and de novo mutations (38). The AGP dataset had no such criteria and contains both multiplex and simplex families [approximately 38% of families are simplex (33)]. Previous studies have investigated the correlation between ASD parents for ASD traits using these instruments, and our results for the AGP data show similar correlations of .40 (25). This is in contrast to results obtained for the SSC, where the SRS-A correlation values were lower and there was no sign of a linear relationship between the parents for the BAPQ data (see Supplemental Figure S24).

Evidence has been shown that differences in the heritability estimates for ASD between multiplex and simplex families exist (33). The results from Klei *et al.*, and the stronger evidence of assortative mating in the AGP datasets that we have shown here, indicate that the genetic mechanisms differ between multiplex families and simplex families (33,57,58). Understanding these differences, and the effects of the broader autism phenotypes present in parents of ASD offspring, will be imperative for understanding the etiology of ASD (28).

We acknowledge that the methods used here are not the only possible approaches to investigate assortative mating in a population. Other methods such as polygene risk scores could offer a means of exploring assortative mating among the parents for ASD risk variants. However, owing to currently still relatively small sample sizes being available for ASD genotype studies, evidence for genetic variants associated with ASD is limited. For instance, the largest GWAS to date in ASD identified few additional findings (59). This, coupled with the nature of this complex disorder having many variants with small effect sizes (37), suggests that an approach using polygene risk scores would currently lack the power needed to find significant evidence. A previous study encountered these issues when trying to use such an approach in ASD (8). However, such an approach may, in the future, offer another avenue for exploring assortative mating in ASD.

In conclusion, we found evidence to suggest that genetic assortative mating is taking place in the AGP dataset and that there is no evidence of this when investigating a simplex family cohort, the SSC dataset. We also identified significant evidence of correlations between parents with certain ASD traits in both the AGP and SSC datasets, although this evidence of phenotype assortative mating is weaker in the SSC dataset. Further investigations into assortative mating in ASD are warranted because it can confound heritability estimates and increase prevalence estimates of the disorder. This study also further emphasizes the different etiology that may be taking place between simplex and multiplex ASD families. It would certainly be of interest to investigate assortative mating at SNPs associated with ASD traits, although the literature is not yet available to support this work owing to the complexity of ASD, but we anticipate that this will be possible in the near future.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the Autism Genetic Resource Exchange (AGRE) and the AGP funders: Autism Speaks (USA), the Medical Research Council (UK), the Health Research Board (Ireland), the National Institutes of Health (USA), Genome Canada, the Canadian Institutes of Health Research, and the Hilibrand Foundation. SC received funding from a Trinity College Dublin Studentship Award.

We thank all the families for participating in the AGP, all the families at the participating SSC sites, the AGP principal investigators, and the SSC principal investigators (A. Beaudet, R. Bernier, J. Constantino, E. Cook, E. Fombonne, D. Geschwind, R. Goin-Kochel, E. Hanson, D. Grice, A. Klin, D. Ledbetter, C. Lord, C. Martin, D. Martin, R. Maxim, J. Miles, O. Ousley, K. Pelphrey, B. Peterson, J. Piggot, C. Saulnier, M. State, W. Stone, J. Sutcliffe, C. Walsh, Z. Warren, and E. Wijsman), and we appreciate obtaining access to phenotypic data on Simons Foundation Autism Research Initiative (SFARI) Base. We are grateful to D.N. Jones for pointing out many helpful

publications. We are grateful to reviewers for very helpful comments and suggestions. In particular, we acknowledge S. Sanders for his help with the SSC data. We gratefully acknowledge the resources provided by Autism Genetic Resource Exchange (AGRE) and the participating AGRE families. The Autism Genetic Resource Exchange is a program of Autism Speaks. We acknowledge the support of the Trinity Centre for High Performance Computing.

The authors report no other biomedical financial interests or potential conflicts of interest.

Approved researchers can obtain the SSC population dataset described in this study by applying at https://base.sfari.org. The AGP datasets can be obtained from the database of Genotypes and Phenotypes (dbGaP) at http://www.ncbi.nlm.nih.gov/gap.

ARTICLE INFORMATION

From the Neuropsychiatric Genetics Research Group (SC, RA, LG, EAH), Department of Psychiatry, Trinity College Dublin, Trinity Centre for Health Sciences, Dublin, and Computer Science and Mathematics Department (SC), Dundalk Institute of Technology, Dundalk, Ireland; Institute of Psychological Medicine and Clinical Neurosciences (RA), Cardiff University School of Medicine, Cathays, Cardiff, United Kingdom.

Address correspondence to Siobhan Connolly, Ph.D., Computer Science and Mathematics Dept., Dundalk Institute of Technology, Marshes Upper, Dundalk, Co. Louth, A91 K584, Ireland; E-mail: connols4@tcd.ie.

Received Jun 6, 2017; revised Apr 2, 2019; accepted Apr 3, 2019.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2019.04.014.

REFERENCES

1. Schneider KA, Peischl S (2011): Evolution of assortative mating in a population expressing dominance. PLoS One 6:e16821.
2. De S, Pal SK, Ghosh A (1998): Genotypic and phenotypic assortative mating in genetic algorithm. Inf Sci 105:209–226.
3. Sebro R, Hoffman TJ, Lange C, Rogus JJ, Risch NJ (2010): Testing for non-random mating: Evidence for ancestry-related assortative mating in the Framingham Heart Study. Genet Epidemiol 34:674–679.
4. Génin E, Ober C, Weitkamp L, Thomson G (2000): A robust test for assortative mating. Eur J Hum Genet 8:119–124.
5. Domingue BW, Fletcher J, Conley D, Boardman JD (2014): Genetic and educational assortative mating among US adults. Proc Natl Acad Sci U S A 111:7996–8000.
6. Conley D, Laidley T, Belsky DW, Fletcher JM, Boardman JD, Domingue BW (2016): Assortative mating and differential fertility by phenotype and genotype across the 20th century. Proc Natl Acad Sci U S A 113:6647–6652.
7. Robinson MR, Kleinman A, Graff M, Vinkhuyzen AAE, Couper D, Miller MB, *et al.* (2017): Genetic evidence of assortative mating in humans. Nat Hum Behav 1:0016.
8. Yengo L, Robinson MR, Keller MC, Kemper KE, Yang Y, Trzaskowski M, *et al.* (2018): Imprint of assortative mating on the human genome. Nat Hum Behav 2:948–954.
9. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, *et al.* (2014): Most genetic risk for autism resides with common variation. Nat Genet 46:881–885.
10. Baron-Cohen S (2006): Two new theories of autism: Hyper-systemising and assortative mating. Archives Dis Child 91:2–5.
11. Williams JG, Higgins JPT, Brayne CEG (2006): Systematic review of prevalence studies of autism spectrum disorders. Arch Dis Child 91:8–15.
12. Harrison MJ, O'Hare AE, Campbell H, Adamson A, McNeillage J (2006): Prevalence of autistic spectrum disorders in Lothian, Scotland: An estimate using the "capture–recapture" technique. Arch Dis Child 91:16–19.
13. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators, Centers for Disease Control and Prevention (2012): Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveill Summ 61:1–19.
14. Peyrot WJ, Robinson MR, Penninx BWJH, Wray NR (2016): Exploring boundaries for the genetic consequences of assortative mating for psychiatric traits. JAMA Psychiatry 73:1189–1195.
15. Baron-Cohen S, Hammer J (1997): Parents of children with Asperger syndrome: What is the cognitive phenotype? J Cogn Neurosci 9:548–554.
16. Silberman S (2001): The geek syndrome. Wired 9.
17. Baron-Cohen S, Wheelwright S, Burtenshaw A, Hobson E (2007): Mathematical talent is linked to autism. Hum Nat 18:125–131.
18. Roelfsema MT, Hoekstra RA, Allison C, Wheelwright S, Brayne C, Matthews FE, Baron-Cohen S (2012): Are autism spectrum conditions more prevalent in an information technology region? A school-based study of three regions in The Netherlands. J Autism Dev Disord 42:734–739.
19. Dickerson AS, Pearson DA, Loveland KA, Rahbar MH, Filipek PA (2014): Role of parental occupation in autism spectrum disorder diagnosis and severity. Res Autism Spectr Disord 8:997–1007.
20. Warrier V, Toro R, Chakrabarti B, the iPSYCH-Broad Autism Group, Børglum AD, Grove J, *et al.* (2017): Systemizing is genetically correlated with autism and is genetically distinct from social autistic traits [published online ahead of print Dec 3]. bioRxiv.
21. Nordsletten AE, Larsson H, Crowley JJ, Almqvist C, Lichtenstein P, Mataix-Cols D (2016): Patterns of nonrandom mating within and across 11 major psychiatric disorders. JAMA Psychiatry 73:354–361.
22. Baron-Cohen S (2006): The hyper-systemizing, assortative mating theory of autism. Prog Neuropsychopharmacol Biol Psychiatry 30:865–872.
23. Happé F, Briskman J, Frith U (2001): Exploring the cognitive phenotype of autism: weak "central coherence" in parents and siblings of children with autism: II. Real-life skills and preferences. J Child Psychol Psychiatry 42:309–316.
24. Constantino JN, Todd RD (2005): Intergenerational transmission of subthreshold autistic traits in the general population. Biol Psychiatry 57:655–660.
25. Lyall K, Constantino JN, Weisskopf MG, Roberts AL, Ascherio A, Santangelo SL (2014): Parental social responsiveness and risk of autism spectrum disorder in offspring. JAMA Psychiatry 71:936–942.
26. Gerdts JA, Bernier R, Dawson G, Estes A (2013): The broader autism phenotype in simplex and multiplex families. J Autism Dev Disord 43:1597–1605.
27. Virkud YV, Todd RD, Abbacchi AM, Zhang Y, Constantino JN (2009): Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. Am J Med Genet B Neuropsychiatr Genet 150B:328–334.
28. Page J, Constantino JN, Zambrana K, Martin E, Tunc I, Zhang Y, *et al.* (2016): Quantitative autistic trait measurements index background genetic risk for ASD in Hispanic families. Mol Autism 7:39.
29. Rubenstein E, Chawla D (2018): Broader autism phenotype in parents of children with autism: A systematic review of percentage estimates. J Child Fam Stud 27:1705–1720.
30. Browning SR, Browning BL (2011): Population structure can inflate SNP-based heritability estimates. Am J Hum Genet 89:191–193.
31. Sebro R, Risch NJ (2012): A brief note on the resemblance between relatives in the presence of population stratification. Heredity 108:563–568.
32. Sebro R, Peloso GM, Dupuis J, Risch NJ, Freedman M, Groop L, *et al.* (2017): Structured mating: Patterns and implications. PLoS Genet 13:e1006655.
33. Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, *et al.* (2012): Common genetic variants, acting additively, are a major source of risk for autism. Mol Autism 3:9.
34. Brosnan M, Walker I (2009): A preliminary investigation into the potential role of waist hip ratio (WHR) preference within the assortative mating hypothesis of autistic spectrum disorders. J Autism Dev Disord 39:164–171.
35. Zhu Z, Lu X, Yuan D, Huang S (2017): Close genetic relationships between a spousal pair with autism-affected children and high minor allele content in cases in autism-associated SNPs. Genomics 109:9–15.

36. Anney R, Klei L, Pinto D, Regan R, Conroy J, Magalhaes TR, *et al.* (2010): A genome-wide scan for common alleles affecting risk for autism. Hum Mol Genet 19:4072–4082.
37. Anney R, Klei L, Pinto D, Almeida J, Bacchelli E, Baird G, *et al.* (2012): Individual common variants exert weak effects on the risk for autism spectrum disorders. Hum Mol Genet 21:4781–4792.
38. Fischbach GD, Lord C (2010): The Simons Simplex Collection: A resource for identification of autism genetic risk factors. Neuron 68:192–195.
39. Sanders SJ, Ercan-Sencicek GA, Hus V, Luo R, Murtha MT, Moreno-De Luca D, *et al.* (2011): Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron 70:863–885.
40. Chaste P, Klei L, Sanders SJ, Hus V, Murtha MT, Lowe JK, *et al.* (2015): A genomewide association study of autism using the Simons Simplex Collection: Does reducing phenotypic heterogeneity in autism increase genetic homogeneity? Biol Psychiatry 77:775–784.
41. Lord C, Rutter M, Le Couteur A (1994): Autism Diagnostic Interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24:659–685.
42. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, *et al.* (2000): The Autism Diagnostic Observation Schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord Biol Psychiatry 30:205–223.
43. Geschwind DH, Sowinski J, Lord C, Iversen P, Shestack J, Jones P, *et al.* (2001): The Autism Genetic Resource Exchange: A resource for the study of autism and related neuropsychiatric conditions. Am J Hum Genet 69:463–466.
44. Waltes R, Gfesser J, Haslinger D, Schneider-Momm K, Biscaldi M, Voran A, *et al.* (2014): Common EIF4E variants modulate risk for autism spectrum disorders in the high-functioning range. J Neural Transm (Vienna) 121:1107–1116.
45. Price AL, Weale ME, Patterson N, Myers SR, Need AC, Shianna KV, *et al.* (2008): Long-range LD can confound genome scans in admixed populations. Am J Hum Genet 83:132–135; author reply 135–139.
46. Abdellaoui A, Verweij KJH, Zietsch BP (2014): No evidence for genetic assortative mating beyond that due to population stratification. Proc Natl Acad Sci U S A 111:E4137.
47. Alexander DH, Novembre J, Lange K (2009): Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664.
48. Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, *et al.* (2010): Integrating common and rare genetic variation in diverse human populations. Nature 467:52–58.
49. Stevens EL, Baugher JD, Shirley MD, Frelin LP, Pevsner J (2012): Unexpected relationships and inbreeding in HapMap phase III populations. PLoS One 7:e49575.
50. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM (2010): Robust relationship inference in genome-wide association studies. Bioinformatics 26:2867–2873.
51. International HapMap Consortium (2003): The International HapMap Project. Nature 426:789–796.
52. Lee S, Epstein MP, Duncan R, Lin X (2012): Sparse principal component analysis for identifying ancestry-informative markers in genome-wide association studies. Genet Epidemiol 36:293–302.
53. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006): Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909.
54. Hurley RSE, Losh M, Parlier M, Reznick JS, Piven J (2007): The Broad Autism Phenotype Questionnaire. J Autism Dev Disord 37:1679–1690.
55. Constantino JN, Gruber CP (2005): The Social Responsiveness Scale Manual. Los Angeles: Western Psychological Services.
56. Abdellaoui A, Hottenga JJ, de Knijff P, Nivard MG, Xiao X, Scheet P, *et al.* (2013): Population structure, migration, and diversifying selection in The Netherlands. Eur J Hum Genet 21:1277–1285.
57. Bernier R, Gerdts J, Munson J, Dawson G, Estes A (2012): Evidence for broader autism phenotype characteristics in parents from multiple-incidence autism families. Autism Res 5:13–20.
58. Chaste P, Leboyer M (2012): Autism risk factors: Genes, environment, and gene–environment interactions. Dialogues Clin Neurosci 14:281–292.
59. Autism Spectrum Disorders Working Group of the Psychiatric Genomics Consortium, Anney RJL, Ripke S, Anttila V, Grove J, Holmans P, *et al.* (2017): Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. Mol Autism 8:21.