**ORIGINAL PAPER**

# Genome-Wide Linkage and Association Study of Childhood Gender Nonconformity in Males

Alan R. Sanders[1,2] · Gary W. Beecham[3,4] · Shengru Guo[3] · Khytam Dawood[5] · Gerulf Rieger[6] · Ritesha S. Krishnappa[7] · Alana B. Kolundzija[8] · J. Michael Bailey[9] · Eden R. Martin[3,4]

## Abstract

Male sexual orientation is influenced by environmental and complex genetic factors. Childhood gender nonconformity (CGN) is one of the strongest correlates of homosexuality with substantial familiality. We studied brothers in families with two or more homosexual brothers (409 concordant sibling pairs in 384 families, as well as their heterosexual brothers), who self-recalled their CGN. To map loci for CGN, we conducted a genome-wide linkage scan (GWLS) using SNP genotypes. The strongest linkage peaks, each with significant or suggestive two-point LOD scores and multipoint LOD score support, were on chromosomes 5q31 (maximum two-point LOD = 4.45), 6q12 (maximum two-point LOD = 3.64), 7q33 (maximum two-point LOD = 3.09), and 8q24 (maximum two-point LOD = 3.67), with the latter not overlapping with previously reported strongest linkage region for male sexual orientation on pericentromeric chromosome 8. Family-based association analyses were used to identify associated variants in the linkage regions, with a cluster of SNPs (minimum association $p = 1.3 \times 10^{-8}$) found at the 5q31 linkage peak. Genome-wide, clusters of multiple SNPs in the $10^{-6}$ to $10^{-8}$ $p$-value range were found at chromosomes 5p13, 5q31, 7q32, 8p22, and 10q23, highlighting glutamate-related genes. This is the first reported GWLS and genome-wide association study on CGN. Further increasing genetic knowledge about CGN and its relationships to male sexual orientation should help advance our understanding of the biology of these associated traits.

**Keywords** Complex trait · Genome-wide linkage scan · Male sexual orientation · Childhood gender nonconformity · GWAS

## Introduction

Male sexual orientation is moderately heritable (30~40% heritability) and appears multifactorial, with evidence of multiple genetic and environmental contributions via family, twin, and segregation analyses (Alanko et al., 2010; Bailey & Bell,

✉ Alan R. Sanders
alan.sanders.md@gmail.com

1 Department of Psychiatry and Behavioral Sciences, NorthShore University HealthSystem Research Institute, 1001 University Place, Evanston, IL 60201, USA

2 Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA

3 John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA

4 Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA

5 Department of Psychology, Pennsylvania State University, University Park, PA, USA

6 Department of Psychology, University of Essex, Colchester, UK

7 Department of Psychiatry, Icahn School of Medicine at Mount Sinai, Elmhurst, NY, USA

8 Collective Impact, Washington, DC, USA

9 Department of Psychology, Northwestern University, Evanston, IL, USA

1993; Bailey & Benishay, 1993; Bailey & Pillard, 1991; Bailey et al., 1993, 1999, 2000; Buhrich et al., 1991; Hamer et al., 1993; Heston & Shields, 1968; Kallmann, 1952; Kendler et al., 2000; King & McDonald, 1992; Kirk et al., 2000; Langström et al., 2010; Pattatucci & Hamer, 1995; Pillard & Weinrich, 1986; Santtila et al., 2008; Schwartz et al., 2010; Whitam et al., 1993). Genome-wide linkage studies (GWLS) of homosexual brother pairs have been applied to the trait (Mustanski et al., 2005; Ramagopalan et al., 2010; Sanders et al., 2015), with the largest GWLS sample finding genome-wide significant linkage to the pericentromeric region of chromosome 8 (LOD = 4.08) and strong support for the previously reported linkage to Xq28 (LOD = 2.99) (Sanders et al., 2015). In addition, genome-wide association studies (GWAS) of the trait are now emerging (Drabant et al., 2012; Sanders et al., 2017); most recently, a GWAS with a greatly enlarged sample size found five loci (two in males, one in females, and two in the combined analyses) significantly associated with same-sex sexual behavior (Ganna et al., 2019).

Sexual orientation is empirically closely linked to some aspects of gender roles, including childhood play behavior and gender identity (Bailey & Zucker, 1995), and adult sex-typed behavior, particularly occupational and recreational interests (Lippa, 2008). While many studies (see below) have been retrospective, i.e., querying recalled childhood gender nonconformity (CGN), recent large-scale, population-based prospective studies have replicated the robust relationship between sexual orientation and CGN (Li et al., 2017; Xu et al., 2019, 2020, 2021). One complementary approach to prospective studies relies on analyzing behaviors depicted in home videos made during childhood, later blindly rated by rater panels; consistent with typical consolidation of gender identity, sexual orientation differences in observer rated CGN diverged around age 3 or 4 years (Rieger et al., 2008). A second approach has examined the correspondence between homosexual men's recalled childhood gender nonconformity and their mothers' memories of their sons, finding a strong association, $r = 0.69$ (Bailey et al., 1995). The strong association between sexual orientation and CGN is cross-culturally robust as well (Bailey et al., 2016). One meta-analysis of 41 retrospective studies showed that homosexuals recalled substantially more CGN than did heterosexuals (Bailey & Zucker, 1995). The effect sizes reported for both sexes are among the largest ever reported in the realm of sex-dimorphic behaviors—the Cohen's $d$ (Cohen, 1988) was 1.3 for men and 1.0 for women (Bailey & Zucker, 1995). In our family linkage sample (Sanders et al., 2015), we found moderate familiality and replicated CGN differences by sexual orientation in males (Bailey et al., 2020). Of further note, in twin studies, CGN shows moderate to high heritability in males: 37% (Knafo et al., 2005), 70% (van Beijsterveldt et al., 2006), 29% (Alanko et al., 2010), and 50% (Bailey et al., 2000). Despite these findings, there have not yet been gene mapping efforts for CGN. Here, we report the first GWLS and GWAS on CGN in males.

## Method

### Participants and Measures

We studied a set of families each with two or more homosexual brothers (409 concordant sibling pairs in 384 families) collected largely from community festivals for a linkage study on male sexual orientation and detailed previously (Sanders et al., 2015). The sample was predominantly of European ancestry (98%) and non-Hispanic (95%), and the average age of the brothers was 44 years old. Briefly, after obtaining written informed consent as approved by the institutional review board of NorthShore University HealthSystem, blood was collected for DNA and questionnaires were completed, including both on sexual orientation (Kinsey scales; (Kinsey et al., 1948) and CGN (Rieger et al., 2008; Zucker et al., 2006). For sexual orientation in the studied sample, homosexual men identified as homosexual and endorsed a Kinsey score of 5 or 6 for fantasy, while heterosexual men identified as heterosexual and endorsed a Kinsey score of 0 or 1 for fantasy. For CGN, we used a scale with 23 items: thirteen items from The Recalled Childhood Gender Identity/Gender Role Questionnaire (items 1, 2, 3, 5, 7, 8, 10, 11, 15, 18, 19, 20, and 21) (Zucker et al., 2006); all seven items from the Childhood Gender Nonconformity Scale (Men) (Rieger et al., 2008), and three items created for this study. The three new items were: (1) "Looking back, I think that others must have found me" (answers ranging from 1 = very masculine to 5 = very feminine); (2) "When I was a child, peers" (answers ranging from 1 = frequently commented on my masculinity to 5 = frequently commented on my femininity); and (3) "When I was a child, adults" (answers ranging from 1 = frequently commented on my masculinity to 5 = frequently commented on my femininity). All items were rescaled so that higher scores indicated greater CGN. Finally, items were standardized and then averaged, with higher scores indicating higher CGN and achieving a reliability (coefficient alpha) of 0.91 as detailed in our recent family study (Bailey et al., 2020). We had CGN scores for 824 of the 826 genotyped brothers. Since the CGN scores were not normally distributed (Shapiro–Wilk test; (Shapiro & Wilk, 1965), $p < 0.0001$), prior to genetic analyses below, we transformed CGN scores via sqrt(2 + CGN) to normalize them (i.e., normCGN, which yielded a Shapiro–Wilk test result of $p = 0.61$).

### Analyses

Genotyping (Affymetrix 5.0 SNP array) and rigorous quality control (QC) steps were previously detailed (Sanders et al., 2015, 2017). Briefly, these included (1) removal of SNPs (minor allele frequency, MAF < 0.05; missingness ≥ 1%; Hardy–Weinberg equilibrium [HWE] deviation $p < 10^{-6}$), and (2) removal of samples (missingness > 5%; failing

checks for duplications and relatedness; ancestry outliers via principal component analysis, PCA). Following the QC filter application in the larger dataset (Sanders et al., 2015, 2017), the families with individuals with CGN scores were extracted for analysis. We conducted our two-point and multipoint GWLS with MERLIN (Abecasis et al., 2002) using the variance components linkage analysis approach with CGN scores as a quantitative trait. For multipoint analysis, we used a set of 45,451 (44,778 autosomal, 673 chromosome X) SNPs that had been pruned for linkage disequilibrium (LD, $r^2 > 0.16$) using PLINK v1.9 (Purcell et al., 2007) and MAF < 0.1. Prior to performing GWAS analyses, we imputed using the IMPUTE2 software (Howie et al., 2009) with the 1,000 Genomes Project Phase 3 reference data(Abecasis et al., 2010) (removing SNPs with an information score < 0.6, MAF < 0.05). We employed the R package, Genome-Wide Association analyses with Family (GWAF) (Chen & Yang, 2010) using the first three principal components (PCs) as covariates to conduct our GWAS of CGN on a final QC'd SNP dataset containing a total of 6,240,683 retained SNPs (210,033 typed and 6,030,650 imputed).

## Results

The normCGN score distributions are plotted by sexual orientation (Fig. 1), with a recapitulation of the patterns previously seen (Bailey & Zucker, 1995), namely that scores were more variable for homosexual men (M = 1.41, variance = 0.046) and clustered more tightly in the gender conforming region of the distribution for heterosexual men (M = 1.02, variance = 0.017)

with the intergroup differences being significant ($p < 0.0001$, $t$-test). In the GWLS, we detected suggestive two-point linkage (LOD ≥ 2.2) (Lander & Kruglyak, 1995) for 382 SNPs (Supplementary Table 1) with ten SNPs exceeding the threshold for genome-wide significance (LOD ≥ 3.6) (Lander & Kruglyak, 1995). Of these ten SNPs, a cluster of five SNPs (rs2349010, rs11242393, rs2074349, rs13172798, rs17171566) at chromosome 5q31 were intronic in *KLHL3* (*kelch like family member 3*), one (rs7841264) at 8q24 was intronic in *CASC8* (*cancer susceptibility 8*), and the other four (rs555920, rs877688, rs4416661, rs2967925) were intergenic. However, we note that linkage signals are imprecise (especially for traits manifesting complex genetics) and thus, larger regions containing additional genes are implicated. Our multipoint nonparametric GWLS results (Fig. 2) show the four strongest peaks (multipoint LOD > 1.8) located on chromosomes 5 (136 ~ 152 cM, based on multipoint drop-1 LOD support interval), 6 (79 ~ 86 cM), 7 (139 ~ 158 cM), and 8 (127 ~ 138 cM); three of these peaks (chromosomes 5, 6, and 8) contain SNPs with genome-wide significant two-point LOD scores, with respective two-point LOD maxima being 4.45 (rs17171566), 3.64 (rs4416661), and 3.67 (rs7841264). We note that these locations do not overlap with previously reported linkage peaks for male sexual orientation, such as at pericentromeric chromosome 8 (Sanders et al., 2015).

Our GWAS (Fig. 3) showed several regions of multiple SNPs in the $10^{-6}$ to $10^{-8}$ $p$-value range, including two regions with SNPs reaching genome-wide significance ($5 \times 10^{-8}$). For our GWAS QC, we achieved a $\lambda_{1000} = 1.046$ (i.e., a low genomic inflation factor; Supplementary Fig. 1 for quantile–quantile plot). The top regions (Supplementary Table 1)

**Fig. 1** NormCGN distributions plotted by sexual orientation. normCGN score bins (x-axis) are plotted against percentage of men (y-axis) in two panels. The top panel of 33 heterosexual men and the bottom panel of 791 homosexual men
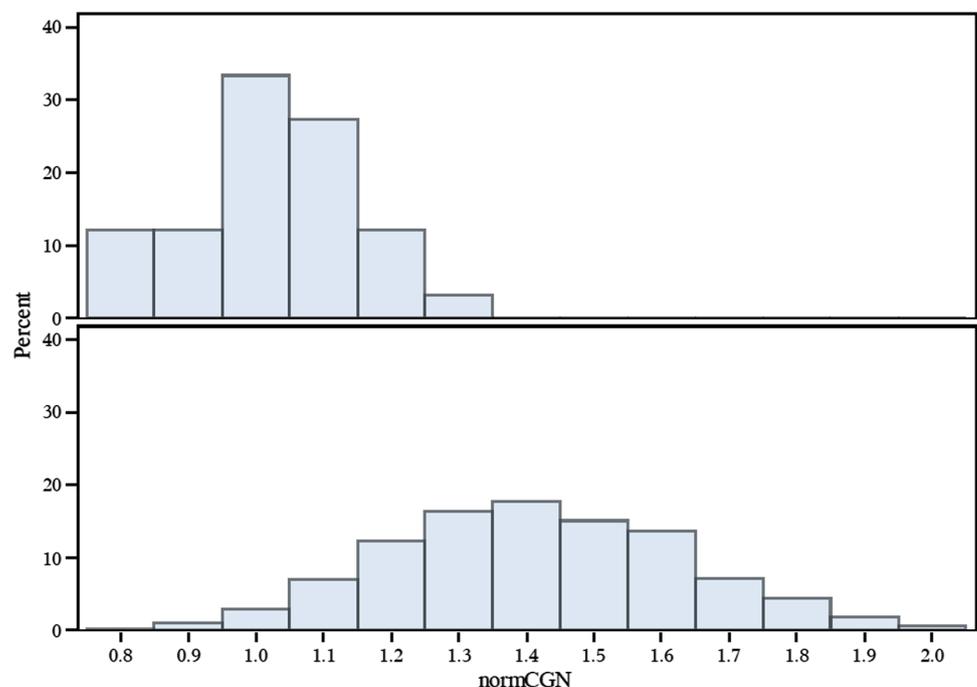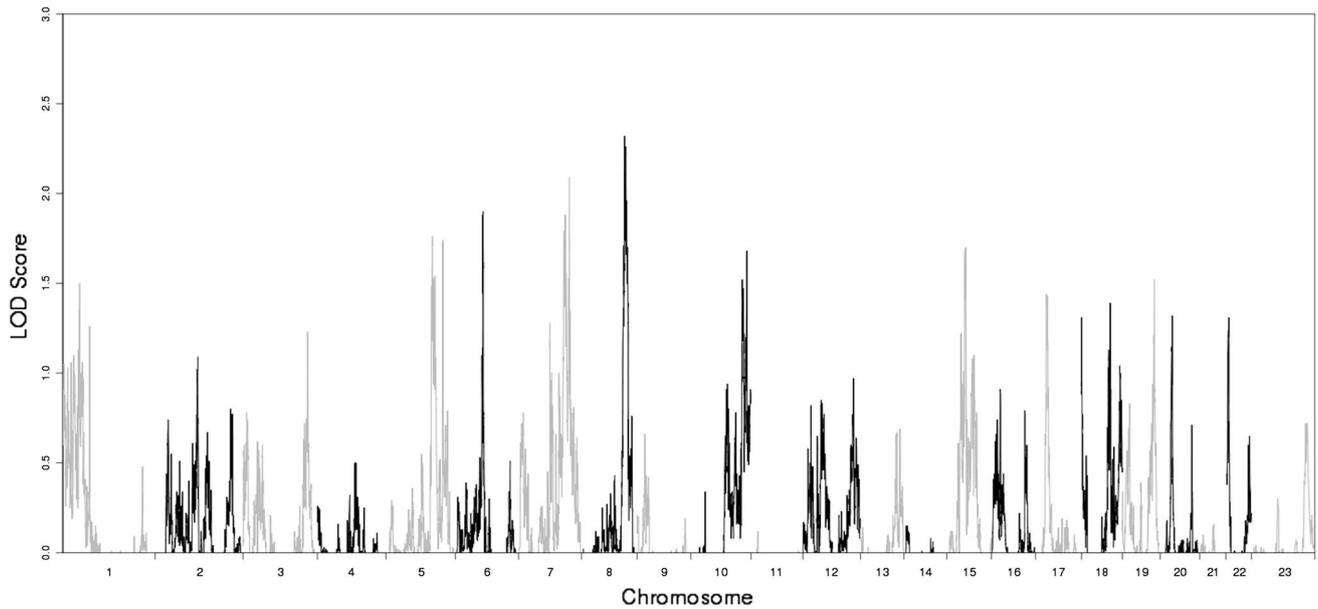
**Fig. 2** Multipoint LOD scores plotted by chromosomal positions for the GWLS of normCGN. Adjacent chromosomes are separated by alternating black and gray lines. Maximum multipoint LOD scores are at ~129 cM on chromosome 8q24

were on chromosomes 5p13 (minimum $p = 4.4 \times 10^{-7}$, rs3832338), 5q31 (minimum $p = 1.3 \times 10^{-8}$, rs113946051), 7q32 (minimum $p = 3.1 \times 10^{-7}$, rs3757757), 8p22 (minimum $p = 1.2 \times 10^{-6}$, rs17488730), and 10q23 (minimum $p = 1.6 \times 10^{-8}$, rs1008912). There are a number of genes of potential relevance to CGN in and around these regions, as described below. Regional association plots for the top linkage regions are displayed in Supplementary Figs. 2

(chromosome 5), 3 (chromosome 6), 4 (chromosome 7), and 5 (chromosome 8). Of note, the linkage peak on chromosome 5q31 also contains a cluster of associated ($10^{-6} < p < 10^{-8}$ p-value) SNPs (Supplementary Table 2, Supplementary Fig. 2). In addition, we display a regional association plot for chromosome 10q23 (Supplementary Fig. 6), which though not in a top linkage region did show genome-wide significant association for 9 SNPs.
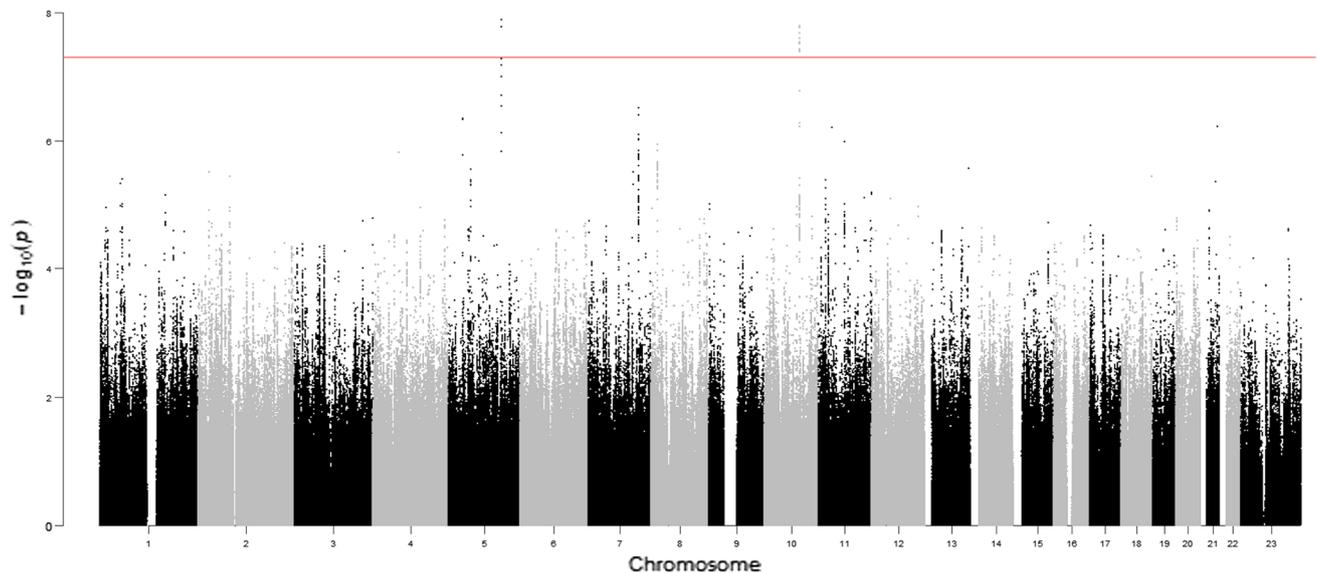


**Fig. 3** Manhattan plot of GWAS for normCGN. Plot of negative $\log_{10}$ of the *p*-values for the single SNP association analysis, ordered along the x-axis by chromosomal position

## Discussion

In this first GWLS on CGN in males, we found genome-wide significant linkage with multipoint support for several linkage regions, most notably at chromosomes 5q31 and 8q24 (Fig. 2, Supplementary Table 1). The strongest multipoint linkage peaks for CGN (Fig. 2) did not align with the strongest such signals from earlier GWLS on male sexual orientation (Sanders et al., 2015). This was not unexpected since while CGN and sexual orientation are associated phenotypes, they are far from being the same and both are traits with complex genetics, and thus, one would not necessarily expect largely overlapping linkage or association patterns. We note that one of the top multipoint peaks from the GWLS (chromosome 5q31, Supplementary Fig. 2) also contains a cluster of 10 associated ($10^{-6} < p < 10^{-8}$ $p$ value) SNPs from the GWAS, 2 of which are genome-wide significant associations, thus with both linkage and association positional evidence. However, none of the genes in the immediate region of this cluster have obvious putative connections to CGN.

This initial GWAS report on CGN had some interesting findings as well. Compared to the previous GWAS on male sexual orientation on the same dataset (Sanders et al., 2017), the current CGN GWAS had substantially more regions with SNPs associated at a level of $10^{-6} < p < 10^{-8}$, including two loci (5q31, 10q23) breaching genome-wide significance (Fig. 3, Supplementary Table 2). Possible explanations include a potentially stronger genetic contribution for CGN (versus sexual orientation) and enhanced statistical power for a quantitative measure with CGN (versus a categorical approach for sexual orientation). A recent large association meta-analysis of same-sex sexual behavior found five genome-wide significant loci (Ganna et al., 2019); however, none of those loci overlap with the top GWLS or GWAS findings for CGN in the current study.

We found two loci (5q31, 10q23) with SNPs reaching genome-wide significance ($p < 5 \times 10^{-8}$) for association with CGN and detected several additional regions (Fig. 3, Supplementary Table 2) with promising findings ($10^{-6} < p < 10^{-8}$ association p-values). These regions contain a number of genes of putative relevance to the trait, some of which we highlight next. At the 5p13 SNP cluster, the nearest gene is *SLC1A3*, a brain expressed glutamate transporter which has been implicated in some behavioral phenotypes, e.g., attention deficit hyperactivity disorder, mood disorders, cortico-limbic connectivity during affective regulation (Huang et al., 2019; Medina et al., 2016; Poletti et al., 2018; van Amen-Hellebrekers et al., 2016). The 10q23 SNP cluster overlaps with *GRID1*, which encodes a glutamate receptor channel subunit, and has also been implicated in various behavioral phenotypes (e.g., mood disorders; Fallin et al., 2005; Zhang et al., 2018) and when deleted in the mouse leads to changes in emotional and social behaviors (Yadav et al.,

2012). The SNPs in the 7q32 cluster fall within (3'UTR, synonymous coding) and near *LRRC4*, which has been implicated in autism spectrum disorders (Du et al., 2020; Um et al., 2018). When deleted ($Lrrc4^{-/-}$) in the mouse, N-Methyl-D-aspartate receptor (NMDAR, an ionotropic glutamate receptor)-dependent synaptic plasticity in the hippocampus was decreased, and these mice displayed mild social interaction deficits, increased self-grooming, and modest anxiety-like behaviors, which were reversed by pharmacological NMDAR activation (Um et al., 2018). Thus, three of the top associated SNP clusters involve glutamate-related genes which have separate evidence of relevance to other behavioral traits, some of which vary in prevalence by gender (e.g., mood disorder; (Sanders et al., 2010) and references therein) in the general population.

Gene mapping challenges include those inherent to GWLS and GWAS of traits manifesting complex genetics such as CGN, as well as limitations in statistical power given the sample size. We discuss power limitations further in the supplementary text but note here that for traits manifesting complex genetics (such as CGN), contributory genetic variants generally have individually small effects, leading to challenges in generating replicable findings. Other limitations include the current study being on a predominantly European ancestry sample and only on males, using retrospective recall of CGN rather than prospective ratings, and not including a replication sample. Replication and extension efforts are somewhat hampered in that relevant survey questions are often not included in large biobank samples such as for CGN; however, there are more sexuality data-points becoming available in some instances (e.g., sexual orientation and gender identity questions in allofus.nih.gov). Additional and larger studies in the future should provide further insight into genetic contributions to CGN and also to its relationship with sexual orientation.

**Declarations**

## References

Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., ... McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. *Nature, 467*, 1061–1073. https://doi.org/10.1038/nature09534

Abecasis, G. R., Cherny, S. S., Cookson, W. O., & Cardon, L. R. (2002). Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genetics, 30*, 97–101. https://doi.org/10.1038/ng786

Alanko, K., Santtila, P., Harlaar, N., Witting, K., Varjonen, M., Jern, P., ... Sandnabba, N. K. (2010). Common genetic effects of gender atypical behavior in childhood and sexual orientation in adulthood: A study of Finnish twins. *Archives of Sexual Behavior, 39*, 81–92. https://doi.org/10.1007/s10508-008-9457-3

Bailey, J. M., & Bell, A. P. (1993). Familiality of female and male homosexuality. *Behavior Genetics, 23*, 313–322. https://doi.org/10.1007/BF01067431

Bailey, J. M., & Benishay, D. S. (1993). Familial aggregation of female sexual orientation. *American Journal of Psychiatry, 150*, 272–277. https://doi.org/10.1176/ajp.150.2.272

Bailey, J. M., Dunne, M. P., & Martin, N. G. (2000). Genetic and environmental influences on sexual orientation and its correlates in an Australian twin sample. *Journal of Personality and Social Psychology, 78*, 524–536. https://doi.org/10.1037//0022-3514.78.3.524

Bailey, J. M., Nothnagel, J., & Wolfe, M. (1995). Retrospective measured individual differences in childhood sex-typed behavior among gay men: Correspondence between self- and maternal reports. *Archives of Sexual Behavior, 24*, 613–622. https://doi.org/10.1007/BF01542183

Bailey, J. M., & Pillard, R. C. (1991). A genetic study of male sexual orientation. *Archives of General Psychiatry, 48*, 1089–1096. https://doi.org/10.1001/archpsyc.1991.01810360053008

Bailey, J. M., Pillard, R. C., Dawood, K., Miller, M. B., Trivedi, S., Farrer, L. A., & Murphy, R. L. (1999). A family history study of male sexual orientation using three independent samples. *Behavior Genetics, 29*, 79–86. https://doi.org/10.1023/a:1021652204405

Bailey, J. M., Pillard, R. C., Neale, M. C., & Agyei, Y. (1993). Heritable factors influence sexual orientation in women. *Archives of General Psychiatry, 50*, 217–223. https://doi.org/10.1001/archpsyc.1993.01820150067007

Bailey, J. M., Rieger, G., Krishnappa, R. S., Kolundzija, A. B., Dawood, K., & Sanders, A. R. (2020). Familiality of gender nonconformity among homosexual men. *Archives of Sexual Behavior, 49*, 2461–2468. https://doi.org/10.1007/s10508-020-01626-w

Bailey, J. M., Vasey, P. L., Diamond, L. M., Breedlove, S. M., Vilain, E., & Epprecht, M. (2016). Sexual orientation, controversy, and science. *Psychological Science in the Public Interest, 17*, 45–101. https://doi.org/10.1177/1529100616637616

Bailey, J. M., & Zucker, K. J. (1995). Childhood sex-typed behavior and sexual orientation: A conceptual analysis and quantitative review. *Developmental Psychology, 31*, 43–55. https://doi.org/10.1037/0012-1649.31.1.43

Buhrich, N., Bailey, J. M., & Martin, N. G. (1991). Sexual orientation, sexual identity, and sex-dimorphic behaviors in male twins. *Behavior Genetics, 21*, 75–96. https://doi.org/10.1007/BF01067668

Chen, M. H., & Yang, Q. (2010). GWAF: An R package for genome-wide association analyses with family data. *Bioinformatics, 26*, 580–581. https://doi.org/10.1093/bioinformatics/btp710

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Drabant, E. M., Kiefer, A. K., Eriksson, N., Mountain, J. L., Francke, U., Tung, J. Y., . . . Do, C. B. (2012). *Genome wide association study of sexual orientation in a large, web-based cohort* (Poster 2100W). Presented at the meeting of the American Society of Human Genetics, San Francisco, CA. Retrieved from https://blog.23andme.com/wp-content/uploads/2012/11/Drabant-Poster-v7.pdf

Du, Y., Li, Z., Liu, Z., Zhang, N., Wang, R., Li, F., ... Wu, J. (2020). Nonrandom occurrence of multiple de novo coding variants in a proband indicates the existence of an oligogenic model in autism. *Genetics in Medicine, 22*, 170–180. https://doi.org/10.1038/s41436-019-0610-2

Fallin, M. D., Lasseter, V. K., Avramopoulos, D., Nicodemus, K. K., Wolyniec, P. S., McGrath, J. A., ... Pulver, A. E. (2005). Bipolar I disorder and schizophrenia: A 440-single-nucleotide polymorphism screen of 64 candidate genes among Ashkenazi Jewish case-parent trios. *American Journal of Human Genetics, 77*, 918–936. https://doi.org/10.1086/497703

Ganna, A., Verweij, K. J. H., Nivard, M. G., Maier, R., Wedow, R., Busch, A. S., & Zietsch, B. P. (2019). Large-scale GWAS reveals insights into the genetic architecture of same-sex sexual behavior. *Science, 365*. https://doi.org/10.1126/science.aat7693

Hamer, D. H., Hu, S., Magnuson, V. L., Hu, N., & Pattatucci, A. M. (1993). A linkage between DNA markers on the X chromosome and male sexual orientation. *Science, 261*, 321–327. https://doi.org/10.1126/science.8332896

Heston, L. L., & Shields, J. (1968). Homosexuality in twins: A family study and a registry study. *Archives of General Psychiatry, 18*, 149–160. https://doi.org/10.1001/archpsyc.1968.01740020021003

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genetics, 5*, e1000529. https://doi.org/10.1371/journal.pgen.1000529

Huang, X., Zhang, Q., Chen, X., Gu, X., Wang, M., & Wu, J. (2019). A functional variant in SLC1A3 influences ADHD risk by disrupting a hsa-miR-3171 binding site: A two-stage association study. *Genes, Brain, and Behavior, 18*, e12574. https://doi.org/10.1111/gbb.12574

Kallmann, F. J. (1952). Twin and sibship study of overt male homosexuality. *American Journal of Human Genetics, 4*, 136–146.

Kendler, K. S., Thornton, L. M., Gilman, S. E., & Kessler, R. C. (2000). Sexual orientation in a U.S. national sample of twin and nontwin sibling pairs. *American Journal of Psychiatry, 157*, 1843–1846. https://doi.org/10.1176/appi.ajp.157.11.1843

King, M., & McDonald, E. (1992). Homosexuals who are twins: A study of 46 probands. *British Journal of Psychiatry, 160*, 407–409. https://doi.org/10.1192/bjp.160.3.407

Kinsey, A. C., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male*. W. B. Saunders Company.

Kirk, K. M., Bailey, J. M., Dunne, M. P., & Martin, N. G. (2000). Measurement models for sexual orientation in a community twin sample. *Behavior Genetics, 30*, 345–356. https://doi.org/10.1023/a:1026557719181

Knafo, A., Iervolino, A. C., & Plomin, R. (2005). Masculine girls and feminine boys: Genetic and environmental contributions to atypical gender development in early childhood. *Journal of Personality and Social Psychology, 88*, 400–412. https://doi.org/10.1037/0022-3514.88.2.400

Lander, E., & Kruglyak, L. (1995). Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nature Genetics, 11*, 241–247. https://doi.org/10.1038/ng1195-241

Langström, N., Rahman, Q., Carlström, E., & Lichtenstein, P. (2010). Genetic and environmental effects on same-sex sexual behavior: A population study of twins in Sweden. *Archives of Sexual Behavior, 39*, 75–80. https://doi.org/10.1007/s10508-008-9386-1

Li, G., Kung, K. T., & Hines, M. (2017). Childhood gender-typed behavior and adolescent sexual orientation: A longitudinal population-based study. *Developmental Psychology, 53*, 764–777. https://doi.org/10.1037/dev0000281

Lippa, R. A. (2008). Sex differences and sexual orientation differences in personality: Findings from the BBC Internet survey. *Archives of Sexual Behavior, 37*, 173–187. https://doi.org/10.1007/s10508-007-9267-z

Medina, A., Watson, S. J., Bunney, W., Jr., Myers, R. M., Schatzberg, A., Barchas, J., ... Thompson, R. C. (2016). Evidence for alterations of the glial syncytial function in major depressive disorder. *Journal of Psychiatric Research, 72*, 15–21. https://doi.org/10.1016/j.jpsychires.2015.10.010

Mustanski, B. S., Dupree, M. G., Nievergelt, C. M., Bocklandt, S., Schork, N. J., & Hamer, D. H. (2005). A genomewide scan of male sexual orientation. *Human Genetics, 116*, 272–278. https://doi.org/10.1007/s00439-004-1241-4

Pattatucci, A. M. L., & Hamer, D. H. (1995). Development and familiality of sexual orientation in females. *Behavior Genetics, 25*, 407–420. https://doi.org/10.1007/BF02253370

Pillard, R. C., & Weinrich, J. D. (1986). Evidence of familial nature of male homosexuality. *Archives of General Psychiatry, 43*, 808–812. https://doi.org/10.1001/archpsyc.1986.01800080094012

Poletti, S., Riberto, M., Vai, B., Ghiglino, D., Lorenzi, C., Vitali, A., ... Benedetti, F. (2018). A glutamate transporter EAAT1 gene variant influences amygdala functional connectivity in bipolar disorder. *Journal of Molecular Neuroscience, 65*, 536–545. https://doi.org/10.1007/s12031-018-1138-7

Purcell, S., Cherny, S. S., & Sham, P. C. (2003). Genetic Power Calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics, 19*, 149–150. https://doi.org/10.1093/bioinformatics/19.1.149

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*, 559–575. https://doi.org/10.1086/519795

Ramagopalan, S. V., Dyment, D. A., Handunnetthi, L., Rice, G. P., & Ebers, G. C. (2010). A genome-wide scan of male sexual orientation. *Journal of Human Genetics, 55*, 131–132. https://doi.org/10.1038/jhg.2009.135

Ridge, P. G., Mukherjee, S., Crane, P. K., Kauwe, J. S., & Alzheimer's Disease Genetics Consortium. (2013). Alzheimer's disease: Analyzing the missing heritability. *PLoS ONE, 8*, e79771. https://doi.org/10.1371/journal.pone.0079771

Rieger, G., Linsenmeier, J. A., Gygax, L., & Bailey, J. M. (2008). Sexual orientation and childhood gender nonconformity: Evidence from home videos. *Developmental Psychology, 44*, 46–58. https://doi.org/10.1037/0012-1649.44.1.46

Sanders, A. R., Beecham, G. W., Guo, S., Dawood, K., Rieger, G., Badner, J. A., ... Martin, E. R. (2017). Genome-wide association study of male sexual orientation. *Scientific Reports, 7*, 16950. https://doi.org/10.1038/s41598-017-15736-4

Sanders, A. R., Levinson, D. F., Duan, J., Dennis, J. M., Li, R., Kendler, K. S., ... Gejman, P. V. (2010). The internet-based MGS2 control sample: Self report of mental illness. *American Journal of Psychiatry, 167*, 854–865. https://doi.org/10.1176/appi.ajp.2010.09071050

Sanders, A. R., Martin, E. R., Beecham, G. W., Guo, S., Dawood, K., Rieger, G., ... Bailey, J. M. (2015). Genome-wide scan demonstrates significant linkage for male sexual orientation. *Psychological Medicine, 45*, 1379–1388. https://doi.org/10.1017/S0033291714002451

Santtila, P., Sandnabba, N. K., Harlaar, N., Varjonen, M., Alanko, K., & von der Pahlen, B. (2008). Potential for homosexual response is prevalent and genetic. *Biological Psychology, 77*, 102–105. https://doi.org/10.1016/j.biopsycho.2007.08.006

Schwartz, G., Kim, R. M., Kolundzija, A. B., Rieger, G., & Sanders, A. R. (2010). Biodemographic and physical correlates of sexual orientation in men. *Archives of Sexual Behavior, 39*, 93–109. https://doi.org/10.1007/s10508-009-9499-1

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*, 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Um, S. M., Ha, S., Lee, H., Kim, J., Kim, K., Shin, W., ... Kim, E. (2018). NGL-2 deletion leads to autistic-like behaviors responsive to NMDAR modulation. *Cell Reports, 23*, 3839–3851. https://doi.org/10.1016/j.celrep.2018.05.087

van Amen-Hellebrekers, C. J., Jansen, S., Pfundt, R., Schuurs-Hoeijmakers, J. H., Koolen, D. A., Marcelis, C. L., ... de Vries, B. B. (2016). Duplications of SLC1A3: Associated with ADHD and autism. *European Journal of Medical Genetics, 59*, 373–376. https://doi.org/10.1016/j.ejmg.2016.06.003

van Beijsterveldt, C. E., Hudziak, J. J., & Boomsma, D. I. (2006). Genetic and environmental influences on cross-gender behavior and relation to behavior problems: A study of Dutch twins at ages 7 and 10 years. *Archives of Sexual Behavior, 35*, 647–658. https://doi.org/10.1007/s10508-006-9072-0

Whitam, F. L., Diamond, M., & Martin, J. (1993). Homosexual orientation in twins: A report on 61 pairs and three triplet sets. *Archives of Sexual Behavior, 22*, 187–206. https://doi.org/10.1007/BF01541765

Xu, Y., Norton, S., & Rahman, Q. (2019). Early life conditions and adolescent sexual orientation: A prospective birth cohort study. *Developmental Psychology, 55*, 1226–1243. https://doi.org/10.1037/dev0000704

Xu, Y., Norton, S., & Rahman, Q. (2020). Childhood maltreatment, gender nonconformity, and adolescent sexual orientation: A prospective birth cohort study. *Child Development, 91*, e984–e994. https://doi.org/10.1111/cdev.13317

Xu, Y., Norton, S., & Rahman, Q. (2021). Childhood gender nonconformity and the stability of self-reported sexual orientation from adolescence to young adulthood in a birth cohort. *Developmental Psychology, 57*, 557–569. https://doi.org/10.1037/dev0001164

Yadav, R., Gupta, S. C., Hillman, B. G., Bhatt, J. M., Stairs, D. J., & Dravid, S. M. (2012). Deletion of glutamate delta-1 receptor in mouse leads to aberrant emotional and social behaviors. *PLoS ONE, 7*, e32969. https://doi.org/10.1371/journal.pone.0032969

Zhang, T., Hou, L., Chen, D. T., McMahon, F. J., Wang, J. C., & Rice, J. P. (2018). Exome sequencing of a large family identifies potential candidate genes contributing risk to bipolar disorder. *Gene, 645*, 119–123. https://doi.org/10.1016/j.gene.2017.12.025

Zucker, K. J., Mitchell, J. N., Bradley, S. J., Tkachuk, J., Cantor, J. M., & Allin, S. M. (2006). The Recalled Childhood Gender Identity/Gender Role Questionnaire: Psychometric properties. *Sex Roles, 54*, 469–483. https://doi.org/10.1007/s11199-006-9019-x