

# The Genetics of Sexuality and Aggression (GSA) Twin Samples in Finland

Ada Johansson,<sup>1</sup> Patrick Jern,<sup>1</sup> Pekka Santtila,<sup>1</sup> Bettina von der Pahlen,<sup>1</sup> Elias Eriksson,<sup>2</sup>  
Lars Westberg,<sup>2</sup> Henrik Nyman,<sup>3</sup> Johan Pensar,<sup>3</sup> Jukka Corander,<sup>3</sup> and N. Kenneth Sandnabba<sup>1</sup>

<sup>1</sup>*Department of Psychology and Logopedics, Abo Akademi University, Turku, Finland*

<sup>2</sup>*Department of Pharmacology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Sweden*

<sup>3</sup>*Department of Mathematics, Abo Akademi University, Turku, Finland*

The Genetics of Sexuality and Aggression (GSA) project was launched at the Abo Akademi University in Turku, Finland in 2005 and has so far undertaken two major population-based data collections involving twins and siblings of twins. To date, it consists of about 14,000 individuals (including 1,147 informative monozygotic twin pairs, 1,042 informative same-sex dizygotic twin pairs, 741 informative opposite-sex dizygotic twin pairs). Participants have been recruited through the Central Population Registry of Finland and were 18–49 years of age at the time of the data collections. Saliva samples for DNA genotyping ( $n = 4,278$ ) and testosterone analyses ( $n = 1,168$ ) were collected in 2006. The primary focus of the data collections has been on sexuality (both sexual functioning and sexual behavior) and aggressive behavior. This paper provides an overview of the data collections as well as an outline of the phenotypes and biological data assembled within the project. A detailed overview of publications can be found at the project's Web site: <http://www.cebfg.fi/>.

■ **Keywords:** aggression, environmental, genetic, population based, sexuality, twin study

The Genetics of Sexuality and Aggression (GSA) project in Finland has completed two major data collections since its inception in 2005. The major research goals of the GSA project are to conduct investigations with genetically sensitive designs, in large, population-based samples, on phenotypes related to sexuality and aggression. What makes this data bank unique is its focus on sexuality and aggression-related phenotypes that have rarely been studied in population-based samples. For several of these phenotypes, the GSA samples are the largest genetically informative samples in the world. Most of our research has focused on sexuality and gender identity-related phenotypes, such as sexual function and dysfunction, sexual behavior and its variations, and body image. A number of relevant psychiatric phenotypes have also been measured (e.g., behaviors and attitudes related to eating, and psychopathology such as anxiety and depression). Furthermore, the research site has a long history of conducting aggression research with genetically informative designs (in both mice and humans), which is now continued within the framework of the GSA project. For example, a major research goal of the GSA project is to examine genetic effects on aggressive behavior,

including identifying genes that moderate the association between alcohol use and aggressive behavior.

The project is conducted at the Department of Psychology and Logopedics at the Abo Akademi University in Turku, Finland. From the second data collection (carried out in 2006) onwards, collaborations with the Department of Pharmacology at the Sahlgrenska Academy, University of Gothenburg, Sweden, have been undertaken in order to collect biological markers. For a subset of the second data collection, genotype data is available, as well as data on free testosterone (T) levels, both collected by means of non-invasive saliva sampling. For the genotype analyses, we implemented a candidate gene approach, that is, polymorphisms were chosen for genotyping based on theoretical hypotheses or previous associations between them

RECEIVED 31 August 2012; ACCEPTED 10 October 2012. First published online 16 November 2012.

ADDRESS FOR CORRESPONDENCE: Ada Johansson, Department of Psychology and Logopedics, Abo Akademi University, FI-20500 Turku, Finland. E-mail: [ada.johansson@abo.fi](mailto:ada.johansson@abo.fi)

**TABLE 1**  
**Number of Participants in the Two Genetics of Sexuality and Aggression (GSA) Samples**

	Sample I (2005) <sup>a</sup>	Sample II (2006) <sup>b</sup>
Target sample <sup>c</sup>	All Finnish MZ and DZ twins in age range 33–43	All Finnish MZ and DZ twins in age range 18–33 years and their over 18-year-old siblings
Final sample size (response rate)	3,558 (36%)	10,524 (45%)
MZ		
Men (complete twin pairs)	<i>n</i> = 389 (91)	<i>n</i> = 727 (210)
Women (complete twin pairs)	<i>n</i> = 693 (247)	<i>n</i> = 1,474 (599)
DZS		
Men (complete twin pairs)	<i>n</i> = 549 (110)	<i>n</i> = 752 (192)
Women (complete twin pairs)	<i>n</i> = 934 (269)	<i>n</i> = 1,329 (471)
DZO		
Men	<i>n</i> = 298	<i>n</i> = 790
Women	<i>n</i> = 396	<i>n</i> = 1,293
Complete twin pairs	203	538
Undefined twins		
Men (complete twin pairs)	<i>n</i> = 77 (29)	<i>n</i> = 57 (18)
Women (complete twin pairs)	<i>n</i> = 222 (103)	<i>n</i> = 109 (39)
Non-twin siblings	None	<i>n</i> = 3,993
Biological samples	None	<i>n</i> = 4,278 DNA saliva samples <i>n</i> = 1,168 Hormone saliva samples

Note: MZ = monozygotic twin, DZ = Dizygotic twin; DZS = same-sex dizygotic twins, DZO = opposite-sex dizygotic twins.

<sup>a</sup>Zygosity of the twins was determined using two standard questionnaire items inquiring about physical resemblance (Sarna et al., 1978).

<sup>b</sup>Zygosity was determined using two standard questionnaire items inquiring about physical resemblance (Sarna et al., 1978) as well as for a subset of the individuals using single nucleotide polymorphic markers extracted from DNA saliva samples. For more detailed information regarding determination of zygosity in the two samples, the reader is advised to consult the method section.

<sup>c</sup>No overlap between the samples.

and relevant phenotypes. Polymorphisms with known functionality were prioritized, and for polymorphisms with unknown functionality, those with known associations to traits or behavior variables were chosen in order to increase the likelihood of the studied polymorphisms being functional. In addition, for genes in which numerous polymorphisms were genotyped, polymorphisms were chosen also with the coverage of the gene in question in mind. At present, approximately 130 single nucleotide polymorphisms (SNPs) have been genotyped, including polymorphisms related to the oxytocin and vasopressin systems (Jern et al., 2012c; Johansson et al., 2012b), sex steroids, and the dopamine and serotonin systems (Jern et al., 2012b). In addition, a number of repeat polymorphisms have been genotyped (e.g., DAT1; Santtila et al., 2010, and the 5-HTTLPR; Jern et al., 2012a).

The research plans for both major data collections have been approved by the Ethics Committee of the Abo Akademi University in accordance with the 1964 Declaration of Helsinki. A more detailed description of our research group, including its publications, can be found at the project Web site: <http://www.cebfg.fi/>.

## Samples and Recruitment Procedure

The GSA project consists of two population-based samples, described in detail below. The participants of both samples were identified using the Central Population Registry of Finland, which is a government-based registry of all Finnish citizens, including personal information regarding, for example, full name, address, native language, family

relations, and date of birth and death (Population Register Centre, [www.vrk.fi/en](http://www.vrk.fi/en)). The register is based on statutory notifications made by private individuals and public authorities and is maintained by the Population Register Centre and local offices that operate under the Ministry of Finance. For more detailed information regarding the register, such as regulations and principles of data disclosure, see [www.vrk.fi/en](http://www.vrk.fi/en) and the Description of Data File document available at <http://www.vrk.fi/default.aspx?id=39>. Conducting data collection in Finland is possible also for foreign research institutions, but must be made according to relevant legislation, specifically regarding data protection. Typically, however, non-Finnish researchers will collect names and addresses from the Central Population Registry through Finnish collaborators, such as universities or the National Institute for Health and Welfare Terveyden ja hyvinvoinnin laitos (THL).

Only Finnish citizens with Finnish as a native language were targeted in both waves of data collection. A summary of the two samples can be seen in Table 1. Both same- and opposite-sex twins, and in the latter data collection, siblings of twins, were contacted.

The first data collection was conducted during 2005 and targeted all Finnish-speaking twin pairs born by the end of 1971, currently residing in Finland. Twin pairs were sampled according to their date of birth from the above mentioned date backwards until 2,000 male–male, 2,000 female–female, and 1,000 opposite-sex pairs had been identified. This resulted in a target sample of 10,000 individuals representative of the Finnish population within this age range. A questionnaire with a paid return envelope was sent

in the beginning of 2005, with reminder letters being sent after a few weeks, followed again after a few weeks by a second posting of the questionnaire. Questionnaires were finally returned by 3,604 respondents, resulting in an overall response rate of 36%. The response rate was lower for male (27%) when compared to female (45%) respondents. The responses of 24 men and 22 women were discarded as their questionnaires were incompletely filled in, resulting in a final sample size of 2,245 women and 1,313 men. The mean age of the sample was 37.57 years ( $SD = 2.94$ , range 33–43 years). No biological samples were collected from participants in Sample I.

The second data collection was conducted in 2006 and this time targeted all Finnish-speaking twin pairs residing in Finland and born between July 22, 1973 and March 1, 1988 (i.e., between 18–33 years old at the time of data collection), as well as their siblings of at least 18 years of age. A total of 23,577 adults met the criteria described above. First, a letter of inquiry was sent to all potential participants in March 2006. The purpose of the study was described in a cover letter and the participants were requested to answer the survey on the Internet or wait for a paper version of the survey to be sent to them. The participants could also indicate whether in addition to answering the questionnaire they wanted to participate by giving samples of saliva for DNA and hormone analyses. A total of 958 individuals stated explicitly that they did not want to participate in the study. Next, the questionnaire was sent to those participants who at that time had not yet responded online or explicitly stated that they did not want to participate. A reminder letter was sent out in the end of July 2006 to individuals who had not responded to the questionnaire. Four travel vouchers valued between €500 and €1,500 were assigned by lot to participants who responded to the questionnaire.

A total of 10,524 participants, of which 6,531 were twin individuals, responded to the survey, yielding an overall participation rate of 45%. Again, the response rate was higher for women (57%) than for men (33%). The response rates were similar between twin individuals (46%) and siblings (43%) to the twins. The mean age for the twins was 24.97 years ( $SD = 4.01$ , age range 18–33) and for the siblings 28.58 ( $SD = 5.97$ , age range 18–49). Detailed information about the zygosity of the twins can be seen in Table 1.

At the same time as questionnaires were sent to participants, kits for obtaining saliva samples for DNA and hormone analyses were sent to those participants who had indicated willingness to give such samples. Using samples of saliva is not as invasive for the participants as collecting blood samples, and the method used has been shown to be a reliable way to collect human genomic DNA containing less bacterial contamination than other oral collection methods (Birnbom et al., 2008). All male participants who had indicated willingness to take part in the DNA analyses were sent the DNA collection kit (Oragene DNA self-collection kits; DNA Genotek, Inc., Kanata, Ontario, Canada). The

participants were advised to follow the manufacturer's instructions and to deposit approximately 2 mL of saliva into the collection cup. As a limited number of DNA kits were available and the number of female respondents who had given consent to give samples of saliva was higher than the corresponding number of male respondents, not all eligible women were selected for this part of the study. Female participants were selected according to whether they had at least one same-sex sibling who, by the beginning of June, had also agreed to participate by giving a sample of DNA. A total of 6,482 kits for obtaining samples of DNA were sent out. A reminder letter was sent in August to those who had not returned their DNA kits. A total of 4,278 kits were finally returned, yielding a 66% return rate for the DNA samples. For the analyses of samples of hormone same-sex twin pairs were prioritized. Test tubes were sent to 1,918 participants who had indicated willingness to take part in the hormone analyses. Of these, 1,168 individuals returned the hormone samples (61% return rate). Two collection tubes were sent to each participant and they were advised to provide saliva samples in the morning after waking up, preferably before 9 a.m. They were also advised not to eat, drink, brush their teeth, or take any medication prior to giving the samples. Approximately 2 mL of saliva was to be deposited in each of the two tubes. In addition, the participants were asked to respond to a number of questions related to the time point of giving the samples (e.g., what the time was when the samples were given, use of alcohol during the last 24 hours, use of cigarettes and other tobacco products, and questions regarding the use of hormonal contraception for women).

## Zygosity Determination

Zygosity of the twins in the GSA samples was mainly determined using two questionnaire items inquiring about physical resemblance (Sarna et al., 1978). The questions read as follows: (1) 'During childhood, were you and your twin partner as like as "two peas in a pod" or were you not more alike than siblings in general?' and (2) 'Were you and your twin partner so similar in appearance at school age that people had difficulty in telling you apart?' The response alternatives for the first question were: *Like two peas in a pod*, *Of ordinary family likeness*, and *Don't know*. For the second question, the response alternatives were: *No*, *Yes*, and *Don't remember*. Twins were classified as monozygotic (MZ) if both twins in a pair answered *Like two peas in a pod* to the first question and *Yes* to the second question and as dizygotic (DZ) if they responded *Of ordinary family likeness* to the first question and *No* to the second (for more detailed information, see Sarna et al., 1978). Questionnaire-based procedures to determine zygosity of twins show acceptable reliability (Christiansen et al., 2003; Eisen et al., 1989; Sarna et al., 1978). In Sample I, zygosity was determined using the above mentioned method. In Sample II, zygosity was determined based on the above mentioned questionnaire items

for the majority of the individuals. As we recently received DNA genotype information for a subset of Sample II, zygosity could for a subset of twin pairs ( $n = 775$  twin pairs) be determined based on genotype data. Note that, since genotype-based zygosity was determined recently, studies that have been published prior to this date have relied on questionnaire-based zygosity determination for the entire second sample. The availability of genotype data has also allowed us to calculate the probability of a twin pair being MZ and DZ given the similarities in their genotypes at SNPs, as well as estimate the reliability of the questionnaire-based zygosity determination in our sample.

In order to calculate the probability of a twin pair being MZ or DZ, genotype data of 129 SNPs was used. Using a Bayesian approach, the probability for a twin pair to be MZ/DZ given the genotypes of the twins as well as given a specific error rate in the determination of genotypes (which was in this case set to 0.002 as estimated by the laboratory) was calculated. Upon request, we are happy to send a more detailed account of the method to interested readers. We chose a probability of 0.95 for the twin pair being MZ as a cut-off for assigning a twin MZ status, the rest of the twin pairs were determined to be DZ. For a majority of the 775 twin pairs for whom zygosity could be determined using genotype information, the probability for being MZ was estimated to be either below 0.000001 (i.e., DZ) or above 0.99999 (i.e., MZ). For 13 pairs, the probability was estimated to be between these figures. Of these pairs, eight were assigned MZ status ( $p(\text{MZ}|\text{genotypes}) > .95$ ), and five DZ status ( $p(\text{MZ}|\text{genotypes}) < .47$ ). The percentage of twin pairs with correctly determined zygosity based on the questionnaire items was 91% in this subpopulation. This is in the lower range of reported accuracies of questionnaire-based zygosity determination (91–98% according to Christiansen et al., 2003). The majority of the misclassifications were twins who, based on the questionnaire items, had been classified as monozygotic when they were in fact dizygotic, according to their genotypes. When examining the accuracies of the questionnaire-based zygosity determinations in men and women separately, it was seen that whereas the difference in accuracies between MZ (94%) and DZ (98%) male twins was quite small, there was a large difference between MZ (84%) and DZ (97%) accuracies for women. Since women were overrepresented (74%) in the sample for which DNA-based zygosity determination was carried out, this may have had an effect on the overall accuracy for this particular sample. The influence such a misclassification could have on heritability estimates in quantitative genetic twin studies is that it could lead to an underestimation of heritability.

## Phenotypes

As mentioned earlier, the GSA project aims to cover a broad range of, specifically, phenotypes related to sexuality, but also traits related to aggressive behavior. In addition, a num-

ber of background items were included, as well as phenotypes related to other aspects of human behavior and variables, as well as variables usually considered environmental (however, needless to say, gene-environment correlations can always be at play). The phenotypes covered can be seen in Table 2. Besides investigating complex phenotypic relationships between different variables, we are interested in estimating overall genetic and environmental effects on traits, as well as identifying specific polymorphisms affecting aspects of human sexual and aggressive behavior. Furthermore, one aim is to examine interplay between genes and environmental factors on sexuality and aggression-related traits in the form of gene-environment correlations (Westerglund et al., 2012) as well as gene-environment interactions (Johansson et al., 2012a; Johansson et al., 2012b).

## Statistical Methods

Quantitative genetic twin models are conducted using the Mx statistical package (Neale et al., 2003), or using the OpenMx software for use with R (Boker et al., 2011, 2012). Model fitting is conducted using full information likelihood estimation, which allows for inclusion of singletons, that is, when only information from one twin of a twin pair is available, as well as the inclusion of singleton siblings. When analyzing data from the second sample, a maximum of 3 male and 3 female non-twin siblings were included in the twin-model fitting analyzes in order to reduce complexity of the scripts. Opposite-sex dizygotic twins are included, except if analyses are conducted only using participants of one gender.

In genetic association analyses between polymorphisms and phenotypes, we use statistical techniques to take into account interdependence between family members, such as the generalized estimating equations method of SPSS, or a linear mixed-effects model using the lme4 package (Bates & Maechler, 2010) in R 2.10.1 (R Development Core Team, 2006).

## Representativeness of the Samples

Both samples were population based, meaning that high generalizability of the samples to the general Finnish population can be expected. However, whenever the response rate is below 100%, a question regarding potential differences between responders and non-responders arises. The response rates (36% and 45%) can in this case be considered somewhat below the true response rates because of false address information, as approximately 15% of Finns move each year (Statistics Finland, 2005). An estimation of the proportion of respondents who never received their survey is, however, difficult to make since the postal service was not obliged to return deliveries that did not reach the recipients.

When interpreting the response rates of the two data collections, it should be kept in mind that the

**TABLE 2**  
**Phenotypes Covered in the Two Genetics of Sexuality and Aggression (GSA) Samples**

Phenotype	Sample I (2005)	Sample II (2006)
Background sections:		
Children	X	X
Family conditions (e.g., biological or adoptive parents)	X	X
Brothers/sisters	X	
University degree of parents		X
Length/weight		X
Pregnancy/abortion/ miscarriage		X
Use of hormonal contraceptives		X
Diabetes/Parkinson/HIV		X
Medication for high blood pressure/depression		X
Sexuality-related sections:		
Sexual orientation	X	X
Partnered and non-partnered sexual activities	X	X
Socio-sexual orientation	X	X
Relationship status	X	X
Transvestism/transsexualism	X	X
Sexual excitation	X	X
Age preferences of sexual partners	X	X
Male and female sexual dysfunctions	X	X
Age at first intercourse	X	X
Relationship quality	X	
Mate poaching	X	
Childhood gender identity	X	
Sexual behavior in childhood	X	
Attitudes toward sexuality in childhood family		X
Sexual socialization		X
Sexual functioning in relationship		X
Sexual distress		X
Feelings and sexual functioning during the first attempted intercourse		X
Intersexuality		X
Paraphilic behavior and interest		X
Aggression-related sections:		
Psychopathy	X	X
Sexual coercion	X	X
Victimization of sexual coercion	X	X
Male interest in sexual coercion		X
Anger		X
Aggressive behavior		X
Sections related to substance use:		
Alcohol use	X	X
Use of anabolic steroids		X
Smoking		X
Miscellaneous:		
Anxiety and depression	X	X
Body image (including sexual body image)	X	X
Childhood trauma and neglect	X	X
Parenting style	X	
Physical closeness with family members	X	
Differential experiences between siblings	X	
Personality	X	
Disordered eating		X

Note: There might be discrepancies in the items used to cover the phenotypes between the two samples. There was no overlap in participants between the two samples.

questionnaires covered sensitive topics related to sexuality and aggressive behavior, among others. The response rates can be seen as comparable to other sexuality (e.g., Bailey et al., 2000; Långström & Zucker, 2005) and aggression-related survey studies (e.g., Hall Smith et al., 2002). However, some differences between responders and non-responders could exist. Dunne et al. (1997) found that responders to a sexuality survey conducted through telephone interviews were somewhat better educated and had less conservative sexual attitudes compared to non-responders. Also, responders were more novelty-seeking and had somewhat elevated lev-

els of major depression and alcohol dependence, as well as an earlier age at first sexual intercourse and higher rates of sexual abuse. With regards to aggressive behavior, Vink et al. (2004) estimated non-response rates in a family sample by using data from respondents as proxy for the data from their non-responding family members. They noted that those from less cooperative families tended to show, for example, higher levels of aggressive behavior and alcohol problems. These effects were not, however, significant after correction for multiple testing, and the authors concluded that aggression and alcohol-related data in studies

with modest response rates is relatively unbiased (response rate in the study was 32.3% for twins and 40.2% for siblings; Vink et al., 2004).

We used a similar approach to investigate differences between responders and potential non-responders using the second GSA sample. We used individuals who had exited the online survey prematurely, leaving the last third of the questionnaire unanswered ( $n = 365$ ) as a proxy for potential non-responders, and compared their levels on sexuality and aggression-related measures to those participants who had completed the questionnaire ( $n = 4,418$ ). Participants who completed the questionnaire reported sexual interest in subjects of a somewhat broader age range ( $M = 2.26$ ,  $SE = 0.02$ , of 11 possible age categories from 0–6 years to 61 years or above), in comparison to individuals who did not finish the questionnaire,  $M = 1.78$ ,  $SE = 0.15$ ,  $F(1, 4434) = 10.304$ ,  $p = .001$ . The same was seen for sexual fantasies during masturbation, (responders  $M = 2.12$ ,  $SE = 0.02$ , non-responders  $M = 1.71$ ,  $SE = 0.17$ ,  $F(1, 3809) = 5.548$ ,  $p = .019$ ). Responders showed a higher mean age at first intercourse,  $M = 17.53$ ,  $SD = 0.045$ , than presumptive non-responders,  $M = 16.99$ ,  $SD = 0.25$ ,  $F(1, 3923) = 4.584$ ,  $p = .032$ . In addition, for women, a significant difference was seen between responders,  $M = 0.92$ ,  $SD = 0.27$ , and non-responders,  $M = 0.84$ ,  $SD = 0.37$ , in the number of participants who had experienced intercourse in comparison to those who had not,  $\chi^2 = 5.761$ ,  $df = 1$ ,  $p = .028$ . We wanted to be liberal in estimating differences between these two groups, and therefore correlation between members of the same family was not taken into account. In addition, we did not correct for multiple testing for the same reason. All in all, differences between the groups were examined for 26 phenotypes for men and 25 phenotypes for women. No other differences were found between the groups, suggesting that there would not be large differences between responders and non-responders. It should be kept in mind, however, that there might be differences between individuals who begin responding to a questionnaire but exit it before completion, and those not even considering responding. Furthermore, the GSA samples are comparable to other Finnish population-based samples on different characteristics such as mean age of first intercourse (Mustanski et al., 2007), and rates of sexual abuse (Sariola & Uutela, 1994). The level of education of the participants' parents was comparable or slightly higher than in the general population in Finland. Altogether, 19% of the mothers and 16% of the fathers of 20- to 25-year-old participants had a university degree, compared to an approximation of 15% for both genders in a similar age range in the general population (Statistics Finland, 2011).

## Collaborations

In order to obtain critical guidance and advice on our research activities, and strategic planning for future research

directions, our research group has an International Scientific Advisory Board consisting of three international experts from relevant fields. These are Prof. Nick Martin (Genetic Epidemiology, Molecular Epidemiology and Neurogenetics Laboratories, Queensland Institute of Medical Research, Brisbane, Australia), Prof. Julia R. Heiman (Director of the Kinsey Institute for Research in Sex, Gender and Reproduction at Indiana University, Bloomington, IN, USA), and Prof. John Archer (University of Central Lancashire, UK).

In matters relating to molecular genetics and hormone analyses, a collaboration with the Department of Pharmacology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Sweden was established. Other collaborators include Prof. Jukka Corander and his research group at the Department of Mathematics at the Abo Akademi University. Prof. Corander's group provides expertise in biostatistics and quantitative analyses of complex data sets.

## Ongoing and Planned Data Collections

At present, a separate study investigating male sexual dysfunction is being undertaken. This study involves collecting data from a clinical group of patients who have sought medical help for ejaculatory disorders, as well as a sub-sample of the men who have participated in the GSA second data collection. This data collection is expected to be completed during the fall of 2012.

## Acknowledgments

The GSA project has been funded by the following major grants: Grants No. 210298, 212703, 136263, and 138291 from the Academy of Finland; and a Center of Excellence Grant No. 21/22/05 from the Stiftelsen för Åbo Akademi Foundation. The authors would like to acknowledge Mathias Hellsten, M. Psych., for having conducted preliminary analyses.

## References

- 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.
- Bates, D., & Maechler, M. (2010). *lme4: Linear mixed-effects models using S4 classes*. R-package version 0.999375–33. Accessed on 1<sup>st</sup> November, 2012. Retrieved from <http://lme4.r-forge.r-project.org/>
- Birnboim, H. C., Iwasiow, R. M., & James, C. M. P. (2008). *Human genomic DNA content of saliva samples collected with the Oragene self-collection kit* (DNA genotek PD-WP-011). Accessed on 1<sup>st</sup> November, 2012. Retrieved from <http://www.dnagenotek.com/US/pdf/PD-WP-011.pdf>
- Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R., ... Fox, J. (2011). OpenMx: An open

- source extended structural equation modeling framework. *Psychometrika*, 76, 306–317.
- Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R., . . . Brandmaier, A. (2012). *OpenMx 1.2 User Guide*. Charlottesville, VA: Department of Psychology, University of Virginia.
- Christiansen, L., Frederiksen, H., Schousboe, K., Skytthe, A., von Wurmb-Schwark, N., Chirstensen, K., & Kyvik, K. (2003). Age and sex differences in the validity of questionnaire-based zygosity in twins. *Twin Research*, 6, 275–278.
- Dunne, M. P., Martin, N. G., Bailey, J. M., Heath, A. C., Buchholz, K. K., Madden, P. A. F., & Statham, D. J. (1997). Participation bias in a sexuality survey: Psychological and behavioural characteristics of responders and non-responders. *International Journal of Epidemiology*, 26, 844–854.
- Eisen, S., Neuman, R., Goldberg, J., Rice, J., & True, W. (1989). Determining zygosity in the Vietnam era twin registry: An approach using questionnaires. *Clinical Genetics*, 35, 423–432.
- Hall Smith, P., Thornton, G. E., DeVellis, R., Earo, J., & Coker, A. L. (2002). A population-based study of the prevalence and distinctiveness of battering, physical assault, and sexual assault in intimate relationships. *Violence Against Women*, 8, 1208–1232.
- Jern, P., Eriksson, E., & Westberg, L. (2012a). A reassessment of the possible effects of the serotonin transporter gene linked polymorphism 5-HTTLPR on premature ejaculation. *Archives of Sexual Behavior*. Advance online publication.
- Jern, P., Westberg, L., Johansson, A., Gunst, A., Eriksson, E., Sandnabba, N. K., & Santtila, P. (2012b). A study of possible associations between single nucleotide polymorphisms in serotonin receptor 1A, 1B, and 2C genes and self-reported ejaculation latency time. *Journal of Sexual Medicine*, 9, 866–872.
- Jern, P., Westberg, L., Johansson, A., Jonsson, L., Corander, J., Sandnabba, N. K., & Santtila, P. (2012c). Are single nucleotide polymorphisms in the oxytocin and vasopressin 1A/1B receptor genes likely candidates for variation in ejaculatory function? *British Journal of Urology International*. Advance online publication.
- Johansson, A., Bergman, H., Corander, J., Waldman, I. D., Karrani, N., Salo, B., . . . Westberg, L. (2012a). Alcohol and aggressive behavior in men — moderating effects of oxytocin receptor gene (OXTR) polymorphisms. *Genes, Brain and Behavior*, 11, 214–221.
- Johansson, A., Westberg, L., Sandnabba, K., Jern, P., Salo, B., & Santtila, P. (2012b). Associations between oxytocin receptor gene (OXTR) polymorphisms and self-reported aggressive behavior and anger: Interactions with alcohol consumption. *Psychoneuroendocrinology*, 37, 1546–1556.
- Långström, N., & Zucker, K. J. (2005). Transvestic fetishism in the general population: prevalence and correlates. *Journal of Sex and Marital Therapy*, 31, 87–95.
- Mustanski, B., Viken, R.J., Kaprio, J., Winter, T., & Rose, R. J. (2007). Sexual behavior in young adulthood: A population-based twin study. *Health Psychology*, 26, 610–617.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.
- Population Register Centre. (n.d). *Description of data file*. Accessed on 20<sup>th</sup> August, 2012. Retrieved from <http://www.vrk.fi/default.aspx?id=39>
- R Development Core Team. (2006). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Accessed on 20<sup>th</sup> August, 2012. Retrieved from <http://www.R-project.org/>
- Santtila, P., Jern, P., Westberg, L., Walum, H., Eriksson, E., & Sandnabba, N. K. (2010). The dopamine transporter gene (DAT1) polymorphism is associated with premature ejaculation. *Journal of Sexual Medicine*, 7, 1538–1546.
- Sariola, H., & Uutela, A. (1994). The prevalence of child sexual abuse in Finland. *Child Abuse & Neglect*, 18, 825–833.
- Sarna, S., Kaprio, J., Sistonen, P., & Koskenvuo, M. (1978). Diagnosis of twin zygosity by mailed questionnaires. *Human Heredity*, 28, 241–254.
- Statistics Finland. (2005). *Population. 2005*. Accessed on 12<sup>th</sup> April, 2007. Retrieved from <http://www.stat.fi>.
- Statistics Finland. (2011). *Perusasteen jälkeisiä tutkintoja suorittanut väestö koulutusasteen, maakunnan, sukupuolen ja ikäryhmän mukaan 2007–2010*. Accessed on 8<sup>th</sup> October, 2012. Retrieved from [http://pxweb2.stat.fi/Dialog/varval.asp?ma=030\\_vkour\\_tau\\_103\\_fi&ti=Perusasteen+j% E4lkeisi%E4+tutkintoja+suorittanut+v%E4est%F6+koulutusasteen%2C+maakunnan%2C+sukupuolen+ja+ik%E4ryhm%E4n+mukaan+2007%2D2010&path=../Database/StatFin/kou/vkour/&lang=3&multilang=fi](http://pxweb2.stat.fi/Dialog/varval.asp?ma=030_vkour_tau_103_fi&ti=Perusasteen+j% E4lkeisi%E4+tutkintoja+suorittanut+v%E4est%F6+koulutusasteen%2C+maakunnan%2C+sukupuolen+ja+ik%E4ryhm%E4n+mukaan+2007%2D2010&path=../Database/StatFin/kou/vkour/&lang=3&multilang=fi)
- Vink, J. M., Willemsen, G., Stubbe, J. H., Middeldorp, C. M., Ligthart, R., Baas, K. D., . . . Boomsma, D. I. (2004). Estimating non-response bias in family studies: Application to mental health and lifestyle. *European Journal of Epidemiology*, 19, 623–630.
- Westerlund, M., Santtila, P., Johansson, A., Jern, P., & Sandnabba, N. K. (2012). What steers them to the ‘wrong’ crowd? Genetic influence on adolescents’ peer-group sexual attitudes. *European Journal of Developmental Psychology*, 9, 645–664.