2016-03-21

Further investigations of the relation between polymorphisms in sex steroid related genes and autistic-like traits

Zettergren, Anna; Karlsson, Sara; Hovey, Daniel; Jonsson, Lina; Melke, Jonas; Anckarsäter, Henrik; Lichtenstein, Paul; Lundström, Sebastian; Westberg, Lars

Psychoneuroendocrinology. 2016 68:1-5.
http://doi.org/10.1016/j.psyneuen.2016.02.020
http://hdl.handle.net/10616/45084

Further investigations of the relation between polymorphisms in sex steroid related genes and autistic-like traits. Psychoneuroendocrinology 2016 68: 1-5.

Zettergren, A; Karlsson, S; Hovey, D; Jonsson, L; Melke, J; Anckarsäter, H; Lichtenstein, P; Lundström, S; Westberg, L.

DOI: **10.1016/j.psyneuen.2016.02.020**

Further investigations of the relation between polymorphisms in sex steroid related genes and autistic-like traits.

Anna Zettergren[a,b*], Sara Karlsson[a], Daniel Hovey[a], Lina Jonsson[a], Jonas Melke[a], Henrik Anckarsäter[c], Paul Lichtenstein[d], Sebastian Lundström[c,e] and Lars Westberg[a]

[a]Institute of Neuroscience and Physiology, Department of Pharmacology, University of Gothenburg, Sweden.

[b]Institute of Neuroscience and Physiology, Department of Psychiatry and Neurochemistry, University of Gothenburg, Sweden.

[c]Institute of Neuroscience and Physiology, Centre of Ethics, Law and Mental Health (CELAM), University of Gothenburg, Sweden.

[d]Karolinska Institutet, Department of medical epidemiology and biostatistics, Stockholm, Sweden.

[e]Institute of Neuroscience and Physiology, Gillberg Neuropsychiatry Centre, University of Gothenburg, Sweden.

Running title: Sex steroid related genes and autistic-like traits

*Corresponding author: Anna Zettergren, PhD. Institute of Neuroscience and Physiology at the Sahlgrenska Academy, University of Gothenburg, Department of Psychiatry and Neurochemistry, Wallinsgatan 6, SE 431 41 Mölndal, Sweden.
E-mail: anna.zettergren@neuro.gu.se, Tel: +46 31 3438714

Abstract

Autism spectrum disorders (ASDs) are more prevalent in boys than in girls, indicating that high levels of testosterone during early development may be a risk factor. Evidence for this hypothesis comes from studies showing associations between fetal testosterone levels, as well as indirect measures of prenatal androgenization, and ASDs and autistic-like traits (ALTs). In a recent study we reported associations between ALTs and single nucleotide polymorphisms (SNPs) in the genes encoding estrogen receptor 1 (*ESR1*), steroid-5-alpha-reductase, type 2 (*SRD5A2*) and sex hormone-binding globulin (*SHBG*) in a subset (n=1771) from the Child and Adolescent Twin Study in Sweden (CATSS). The aim of the present study was to try to replicate these findings in an additional, larger, sample of individuals from the CATSS (n=10 654), as well as to analyze additional SNPs of functional importance in *SHBG* and *SRD5A2*. No associations between the previously associated SNPs in the genes *ESR1* and *SRD5A2* and ALTs could be seen in the large replication sample. Still, our results show that two non-linked SNPs (rs6259 and rs9901675) at the *SHBG* gene locus might be of importance for language impairment problems in boys. The results of the present study do not point towards a major role for the investigated SNPs in the genes *ESR1* and *SRD5A2* in ALTs, but a possible influence of genetic variation in *SHBG*, especially for language impairment problems in boys, cannot be ruled out.

Key words: Autism spectrum disorders; Autistic-like traits; Sex steroids; Gene; Polymorphism; Association

1. Introduction

Autism spectrum disorders (ASDs) are heterogeneous neurodevelopmental disorders, characterized by social interaction impairments and communication problems, as well as restricted and repetitive behavior. ASDs have been proposed to represent the extreme end of dimensionally distributed autistic-like traits (ALTs), and it has been shown that ASDs and ALTs share genetic and environmental effects with each other (Lundstrom et al. 2012).

ASDs are about four times more prevalent in boys than in girls and it has been hypothesized that high levels of testosterone during early development may be a risk factor for these types of disorders (Baron-Cohen et al. 2011). Several studies showing fetal testosterone levels (Baron-Cohen et al. 2015; for further refs see Baron-Cohen et al. 2011), as well as indirect measures of prenatal androgenization (for refs see Baron-Cohen et al. 2011) to be associated with ASDs and ALTs strengthen this theory. Further support for the importance of sex steroids in these conditions comes from previous genetic studies, showing associations between sex steroid related genes and ASDs as well as ALTs (Chakrabarti et al. 2009).

In a recent study (Zettergren et al. 2013), we investigated possible associations between ALTs and 29 SNPs in eight genes related to sex steroids, in a subset of individuals (n=1771) from the Child and Adolescent Twin Study in Sweden (CATSS). The results indicated that three genes might be of importance; estrogen receptor 1 (*ESR1*), steroid-5-alpha-reductase, type 2 (*SRD5A2*) and sex hormone- binding globulin (*SHBG*). In the first part of the present study we investigated if our previous findings could be replicated in an additional, larger, sample of individuals from the CATSS (n=10 655). In the second part of the study, we analysed single nucleotide polymorphisms (SNPs) of functional importance (related to circulating levels of androgens and/or SHBG) at the *SHBG* locus (Ohlsson et al. 2011; Coviello et al. 2012;

Prescott et al. 2012), as well as a SNP in *SRD5A2* related to the activity of the enzyme (Makridakis et al. 2000).

2. Material and Methods

2.1 Subjects and Measurements

Individuals, born between 1992 and 2002, included in the current study belong to two different subsets of the CATSS; discovery sample (n=1771) and replication sample (n=10 654). In our recent paper, only the discovery sample was analyzed. In the present work the replication sample was used for investigations of previous findings. For analyses of SNPs not included in our previous study the two samples (discovery and replication) were collapsed into a total sample (n=12425 individuals; n=6222 boys and n=6203 girls). The total sample consisted of 2301 monozygotic (MZ) and 3869 dizygotic twin pairs, as well as 85 subjects included without their co-twin. One hundred and seven subjects (24 from the discovery sample and 83 from the replication sample) were excluded from the statistical analyses due to documented brain damage or a known genetic syndrome. When the individuals in CATSS were 9 or 12 years old, their parents responded to a telephone interview containing, among other things, the questionnaire Autism–Tics, AD/HD, and other Co-morbidities inventory (A-TAC) (Larson et al. 2010). The A-TAC is a validated instrument developed to assess neurodevelopmental problems and coexisting disorders in epidemiological settings. It can be used both in a categorical and continuous fashion. The 17 items on ALTs (six corresponding to language impairment, six to social interaction impairment and five to restricted and repetitive behavior) had three response categories; ”no” (coded 0), “yes, to some extent” (coded 0.5), and “yes” (coded 1.0). The measure of total ASD score is the sum of these 17 items. In the total sample (CATSS-12425) the range of the total ASD score is 0-17, and the mean (sd) is 0.90 (1.70) in boys and 0.55 (1.17) in girls. The CATSS study has been approved

by the Ethical review Board at Karolinska Institutet and informed consent was provided by all participants.

2.2 DNA extraction and genotyping of polymorphisms

DNA was extracted from saliva samples using OraGene® DNA self-collection kit (DNA Genotek, Inc.,Ottawa, Ontario, Canada). Three SNPs (rs2747648 in *ESR1*, rs523349 in *SRD5A2* and rs6259 in *SHBG*) were chosen, based on associations with ALTs found in our previous study in the discovery sample, for genotyping in the replication sample. Additional functional SNPs located at the *SHBG* locus (rs9901675, rs1625895, rs6258, rs727428, rs1641537, rs12150660) and in *SRD5A2* (rs9282858) were chosen for genotyping in the total sample. The polymorphisms were genotyped with KASPar® PCR SNP genotyping system (KBiosciences, Herts, UK) and the genotyping success rate was >95% for all SNPs except rs12150660 with a success rate of approximately 90%. All SNPs were in Hardy-Weinberg equilibrium, except rs1641537 that was excluded from further analyses. Further information about the genotyped SNPs can be found in Supplementary table 1.

2.3 Statistical analysis

Statistical associations between SNPs and continuous measures of ALTs, including the modules described above, were investigated using a linear mixed effect model in the “Proc mixed” procedure of SAS 9.3 (SAS Institute, Inc., Cary, NC). This model made it possible to adjust for the dependent nature of the twin observations. Mean scores and standard deviations, as well as the standard errors presented in Figure 1, were calculated using the “Proc means” procedure in SAS 9.3. Firstly, an additive genetic model was assumed and if an association was clearly driven by the uncommon homozygote (relevant for rs6259 and rs9901675) a recessive genetic model was also tested. For two SNPs, rs6258 and rs9282858, the uncommon

homozygote was present in very few individuals and the analyses were performed after these individuals were collapsed with the heterozygotes. To control for multiple testing Bonferroni correction was performed. The replication study included 9 independent tests (3 SNPs and 3 ALT-modules), resulting in a corrected p-value limit of 0.006. Since the associations we attempted to replicate were all sex-specific, the replication analyses were not corrected for tests in two sexes. The analyses of additional SNPs in the second part of the study included 36 independent tests (6 SNPs, 2 sexes and 3 ALT-modules), resulting in a corrected p-value limit of 0.001.

3. Results

No relationships between the previously associated SNPs in the genes *ESR1* (rs2747648) and *SRD5A* (rs523349) and ALTs (total score or scores for the three sub-modules language impairment, social interaction impairment and restrictive and repetitive behavior) could be seen in the large replication sample of over 10 600 individuals or in the total sample of over 12 400 subjects (see Supplementary table 2). However, the previously investigated SNP rs6259 (Asp356Asn), in *SHBG* was found to be related to the language impairment module in boys (p=0.006 for recessive model), which is a replication of a nominal association in our previous study (see Table 1). The association in the replication sample reached the Bonferroni corrected p-value, and decreased in the total sample, CATSS 12425 (Table 1). The other two sub-modules, social interaction impairments and restricted and repetitive behavior, were found not to be associated with this *SHBG* SNP in the replication sample.

In addition to the above mentioned gene variants we investigated additional SNPs of functional importance in *SHBG* and *SRD5A2*. One of the SNPs at the *SHBG* locus (rs9901675) was found to be nominally associated (not reaching the Bonferroni-corrected p-

value limit) with language impairments in boys (p=0.03 for additive model and p=0.04 for AA vs AG+GG in a recessive model). The mean ALT-score for language impairments was 0.56 for AA-carriers, 0.23 for AG-carriers and 0.28 for GG-carriers. None of the other investigated SNPs in *SHBG* and *SRD5A2* were found to be related to any ALT-score (see Supplementary table 2).

Furthermore, as both rs6259 and rs9901675 of the *SHBG* locus were associated with language impairments specifically in boys, the combined contribution of the two "high-risk" genotypes (AA of each SNP) were investigated. The comparison of language impairment scores between carriers and non-carriers of either AA-genotype yielded a highly significant p-value of 0.00006 (see Figure 1, panel A). Noteworthy, no individual carries high-risk genotypes (AA) of both rs6259 and rs9901675. A histogram showing the distribution of language impairment scores in boys carrying or not carrying a high-risk genotype of rs6259 or rs9901675 at the *SHBG* locus is presented in Figure 1, panel B.

Figure 1. A: Associations between the *SHBG*-SNPs rs6259 and rs9901675 in combination and language impairment scores. B: Distribution of language impairment scores in boys carrying or not carrying a high-risk genotype of rs6259 or rs9901675 at the *SHBG* locus.

4. Discussion

The results of the present study do not, despite our previous findings, point towards a role for the investigated SNPs in the genes *ESR1* and *SRD5A2* in ALTs. More specifically, previous analyses of our discovery sample comprising 1771 individuals from the normal population revealed highly significant associations (surviving Bonferroni correction) between rs2747248 in *ESR1* and restrictive and repetitive behavior in boys, and rs523349 in *SRD5A2* and social interaction impairments in girls (Zettergren et al. 2013). However, none of these associations

could be seen in a large replication-sample, and no associations were found between these SNPs and the total ASD-score or any of the other sub-module scores.

Still, our present result shows that two different SNPs (rs6259 and rs9901675) at the *SHBG* gene locus might be of importance for language impairment problems in boys, and for one of these SNPs (rs6259) this result is a replication of a finding in our previous study. Both SNPs have been strongly associated with circulating levels of SHBG and testosterone in large GWA-studies (Ohlsson et al. 2011; Coviello et al. 2012; Jin et al. 2012; Prescott et al. 2012), and rs6259 are known to affect protein function and have been found to be associated with several disorders and traits related to SHBG-levels, such as polycystic ovary syndrome and type 2 diabetes (Ding et al. 2009; Martinez-Garcia et al. 2012). Noteworthy in the context of genes regulating sex hormone levels in association with language problems is the previous report of an important role for estrogen signalling in human cognitive functions implicated in reading, speech and language (Tammimies et al. 2012).

Analyses of proteins in serum of individuals with ASDs compared to controls have shown altered levels of SHBG in both men and women (increased in male and decreased in female patients) (Steeb et al. 2014). Apart from our studies, only one genetic study specifically on sex steroid related genes included SNPs in the *SHBG*-gene in ASDs and/or ALTs (Chakrabarti et al. 2009). In that study no associations between the SNPs rs6259 or rs6257 and Asperger syndrome and/or ALTs in the general population were found. However, compared to our study Chakrabarti and colleagues used another measure (Autism Spectrum Quotient) of ALTs, did not specifically investigate language impairments, and analyzed a smaller population (349 individuals from the general population), probably lacking the statistical power to detect the associations identified in the present study.

It is of interest to note that a SNP (rs6259) of functional importance in the *SHBG* gene associates with language impairments in boys in two independent populations. Strikingly, another SNP (rs9901675) of functional importance for the same gene also associates, although nominally, with language problems specifically in boys. Intriguingly, the associations with these two SNPs are to be considered independent of each other, since no individual carries the "high-risk" genotype (related to higher language impairment score) of both SNPs. The unrelatedness between the two SNPs is further supported by others, who have reported very low linkage (based on r^2-values) between these two genetic variants (Coviello et al. 2012).

The replication sample used in the present work originates from the same study-population (CATSS) as the smaller discovery sample used in our previous study of sex-steroid-related genes in ALTs. Still, there are differences between the two samples, which might explain part of the divergent results seen in our studies. In the discovery sample the mean-values for the ALT-scores are higher than in the replication sample, since a larger percentage of individuals reaching the cut-off score for prediction of a screening diagnosis are included in the discovery sample. This "over-representation" of clinical cases may have had an influence on the results found when analyzing ALTs in the discovery sample.

In conclusion, in a large sample of young individuals from Sweden we find no evidence of importance for the previously associated SNPs in the genes *ESR1* and *SRD5A2* in ALTs. Still, the genetic variation in these genes is not fully covered by the SNPs investigated in our studies. Furthermore, we cannot rule out a possible influence of genetic variation in the gene *SHBG* on these traits and further investigations of this gene, as well as more comprehensive genotyping of other genes related to sex-steroids, in ALTs are warranted.

References

Baron-Cohen, S., B. Auyeung, B. Norgaard-Pedersen, D. M. Hougaard, M. W. Abdallah, L. Melgaard, A. S. Cohen, B. Chakrabarti, L. Ruta and M. V. Lombardo, 2015. Elevated fetal steroidogenic activity in autism. Mol Psychiatry 20(3): 369-376.

Baron-Cohen, S., M. V. Lombardo, B. Auyeung, E. Ashwin, B. Chakrabarti and R. Knickmeyer, 2011. Why are autism spectrum conditions more prevalent in males? PLoS Biol 9(6): e1001081.

Chakrabarti, B., F. Dudbridge, L. Kent, S. Wheelwright, G. Hill-Cawthorne, C. Allison, S. Banerjee-Basu and S. Baron-Cohen, 2009. Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. Autism Res 2(3): 157-177.

Coviello, A. D., R. Haring, M. Wellons, D. Vaidya, T. Lehtimaki, S. Keildson, K. L. Lunetta, C. He, M. Fornage, V. Lagou, M. Mangino, N. C. Onland-Moret, B. Chen, J. Eriksson, M. Garcia, Y. M. Liu, A. Koster, K. Lohman, L. P. Lyytikainen, A. K. Petersen, et al., 2012. A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple Loci implicated in sex steroid hormone regulation. PLoS Genet 8(7): e1002805.

Ding, E. L., Y. Song, J. E. Manson, D. J. Hunter, C. C. Lee, N. Rifai, J. E. Buring, J. M. Gaziano and S. Liu, 2009. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. N Engl J Med 361(12): 1152-1163.

Jin, G., J. Sun, S. T. Kim, J. Feng, Z. Wang, S. Tao, Z. Chen, L. Purcell, S. Smith, W. B. Isaacs, R. S. Rittmaster, S. L. Zheng, L. D. Condreay and J. Xu, 2012. Genome-wide association study identifies a new locus JMJD1C at 10q21 that may influence serum androgen levels in men. Hum Mol Genet 21(23): 5222-5228.

Larson, T., H. Anckarsater, C. Gillberg, O. Stahlberg, E. Carlstrom, B. Kadesjo, M. Rastam and P. Lichtenstein, 2010. The autism--tics, AD/HD and other comorbidities inventory (A-TAC): further validation of a telephone interview for epidemiological research. BMC Psychiatry 10: 1.

Lundstrom, S., Z. Chang, M. Rastam, C. Gillberg, H. Larsson, H. Anckarsater and P. Lichtenstein, 2012. Autism spectrum disorders and autistic like traits: similar etiology in the extreme end and the normal variation. Arch Gen Psychiatry 69(1): 46-52.

Makridakis, N. M., E. di Salle and J. K. Reichardt, 2000. Biochemical and pharmacogenetic dissection of human steroid 5 alpha-reductase type II. Pharmacogenetics 10(5): 407-413.

Martinez-Garcia, M. A., A. Gambineri, M. Alpanes, R. Sanchon, R. Pasquali and H. F. Escobar-Morreale, 2012. Common variants in the sex hormone-binding globulin gene (SHBG) and polycystic ovary syndrome (PCOS) in Mediterranean women. Hum Reprod 27(12): 3569-3576.

Ohlsson, C., H. Wallaschofski, K. L. Lunetta, L. Stolk, J. R. Perry, A. Koster, A. K. Petersen, J. Eriksson, T. Lehtimaki, I. T. Huhtaniemi, G. L. Hammond, M. Maggio, A. D. Coviello, L. Ferrucci, M. Heier, A. Hofman, K. L. Holliday, J. O. Jansson, M. Kahonen, D. Karasik, et al., 2011. Genetic determinants of serum testosterone concentrations in men. PLoS Genet 7(10): e1002313.

Prescott, J., D. J. Thompson, P. Kraft, S. J. Chanock, T. Audley, J. Brown, J. Leyland, E. Folkerd, D. Doody, S. E. Hankinson, D. J. Hunter, K. B. Jacobs, M. Dowsett, D. G. Cox, D. F. Easton and I. De Vivo, 2012. Genome-wide association study of circulating estradiol, testosterone, and sex hormone-binding globulin in postmenopausal women. PLoS One 7(6): e37815.

Steeb, H., J. M. Ramsey, P. C. Guest, P. Stocki, J. D. Cooper, H. Rahmoune, E. Ingudomnukul, B. Auyeung, L. Ruta, S. Baron-Cohen and S. Bahn, 2014. Serum proteomic analysis identifies sex-specific differences in lipid metabolism and inflammation profiles in adults diagnosed with Asperger syndrome. Mol Autism 5(1): 4.

Tammimies, K., I. Tapia-Paez, J. Ruegg, G. Rosin, J. Kere, J. A. Gustafsson and I. Nalvarte, 2012. The rs3743205 SNP is important for the regulation of the dyslexia candidate gene DYX1C1 by estrogen receptor beta and DNA methylation. Mol Endocrinol 26(4): 619-629.

Zettergren, A., L. Jonsson, D. Johansson, J. Melke, S. Lundstrom, H. Anckarsater, P. Lichtenstein and L. Westberg, 2013. Associations between polymorphisms in sex steroid related genes and autistic-like traits. Psychoneuroendocrinology 38(11): 2575-2584.

Table 1. Associations between SNPs of the *SHBG* locus and autistic-like traits in the discovery sample (CATSS-1771), replication sample (CATSS-10654) and total sample (CATSS-12425).

				ALT modules		
			Total ALT Score	Language impairment	Social interaction impairment	Restricted and repetitive behavior
CATSS-1771[1]						
SHBG:rs6259 (Asp356Asn)		N	mean score (sd)	mean score (sd)	mean score (sd)	mean score (sd)
All	A/A	26	1.35 (3.31)	0.42 (1.17)	0.44 (0.98)	0.48 (1.23)
	A/G	381	0.71 (1.36)	0.26 (0.58)	0.26 (0.55)	0.20 (0.51)
	G/G	1318	1.01 (1.93)	0.34 (0.72)	0.35 (0.78)	0.31 (0.73)
	pa		**0.029**	0.057	0.172	**0.019**
	pb		0.288	0.350	0.454	0.224
Boys	A/A	14	2.29 (4.33)	0.75 (1.54)	0.75 (1.27)	0.79 (1.61)
	A/G	188	0.96 (1.57)	0.37 (0.69)	0.31 (0.62)	0.28 (0.58)
	G/G	664	1.29 (2.24)	0.44 (0.83)	0.44 (0.91)	0.41 (0.84)
	pa		**0.0028**	**0.016**	**0.020**	**0.011**
	pb		**0.0058**	**0.016**	**0.032**	**0.020**
Girls	A/A	12	0.25 (0.50)	0.04 (0.14)	0.08 (0.19)	0.13 (0.31)
	A/G	193	0.47 (1.07)	0.15 (0.41)	0.21 (0.47)	0.12 (0.41)
	G/G	654	0.72 (1.48)	0.25 (0.56)	0.26 (0.60)	0.21 (0.58)
	pa		0.186	0.072	0.498	0.296
	pb		0.382	0.267	0.340	0.862
[1]Results from our previous study in a discovery sample consisting of 1771 individuals (Zettergren et al, 2013).						
CATSS-10654[2]						
SHBG:rs6259 (Asp356Asn)		N	mean score (sd)	mean score (sd)	mean score (sd)	mean score (sd)
All	A/A	150	0.65 (1.45)	0.31 (0.71)	0.19 (0.55)	0.16 (0.43)
	A/G	2009	0.67 (1.45)	0.21 (0.55)	0.23 (0.56)	0.22 (0.57)
	G/G	8096	0.65 (1.26)	0.21 (0.51)	0.23 (0.53)	0.21 (0.53)
	pa		0.789	0.132	0.785	0.413
	pb		0.914	**0.044**	0.525	0.234
Boys	A/A	82	0.91 (1.79)	0.47 (0.90)	0.25 (0.64)	0.20 (0.51)
	A/G	992	0.81 (1.73)	0.25 (0.63)	0.27 (0.63)	0.28 (0.68)
	G/G	4060	0.81 (1.45)	0.26 (0.57)	0.27 (0.59)	0.28 (0.62)
	pa		0.877	**0.022**	0.970	0.497
	pb		0.614	**0.006***	0.806	0.246
Girls	A/A	68	0.33 (0.78)	0.11 (0.27)	0.11 (0.39)	0.11 (0.28)
	A/G	1017	0.54 (1.09)	0.18 (0.45)	0.20 (0.48)	0.16 (0.44)
	G/G	4036	0.50 (1.01)	0.17 (0.44)	0.19 (0.46)	0.14 (0.40)
	pa		0.291	0.448	0.473	0.552
	pb		0.249	0.303	0.272	0.421

CATSS-12425[3]						
SHBG: rs6259 (Asp356Asn)		N	mean score (sd)	mean score (sd)	mean score (sd)	mean score (sd)
All	A/A	176	0.75 (1.84)	0.32 (0.79)	0.22 (0.63)	0.20 (0.61)
	A/G	2385	0.67 (1.43)	0.22 (0.55)	0.23 (0.55)	0.21 (0.56)
	G/G	9398	0.70 (1.37)	0.23 (0.054)	0.24 (0.57)	0.22 (0.56)
	pa		0.67	**0.050**	0.82	0.69
	pb		0.52	**0.024**	0.86	0.58
Boys	A/A	96	1.11 (2.35)	0.51 (1.01)	0.32 (0.77)	0.28 (0.79)
	A/G	1176	0.83 (1.70)	0.27 (0.64)	0.27 (0.62)	0.28 (0.66)
	G/G	4715	0.87 (1.59)	0.28 (0.61)	0.29 (0.64)	0.29 (0.65)
	pa		0.23	**0.001***	0.61	0.80
	pb		0.13	**0.0005***	0.54	0.83
Girls	A/A	80	0.31 (0.74)	0.10 (0.25)	0.10 (0.37)	0.11 (0.28)
	A/G	1209	0.52 (1.08)	0.17 (0.44)	0.20 (0.47)	0.14 (0.43)
	G/G	4683	0.52 (1.09)	0.17 (0.45)	0.20 (0.48)	0.15 (0.43)
	pa		0.36	0.36	0.35	0.70
	pb		0.15	0.15	0.15	0.41

[2]Results from analyses in a replication sample consisting of 10654 individuals.

[3]Results from analyses in the total sample consisting of 12425 (1771+10654) individuals.

pa: p-values after comparisons between all genotypes. pb: p-values after comparisons based on a recessive model: AA vs AG+GG

Significant p-values ($p \leq 0.05$) are shown in bold. p-values surviving correction for Bonferroni are marked with an asterisk (*).

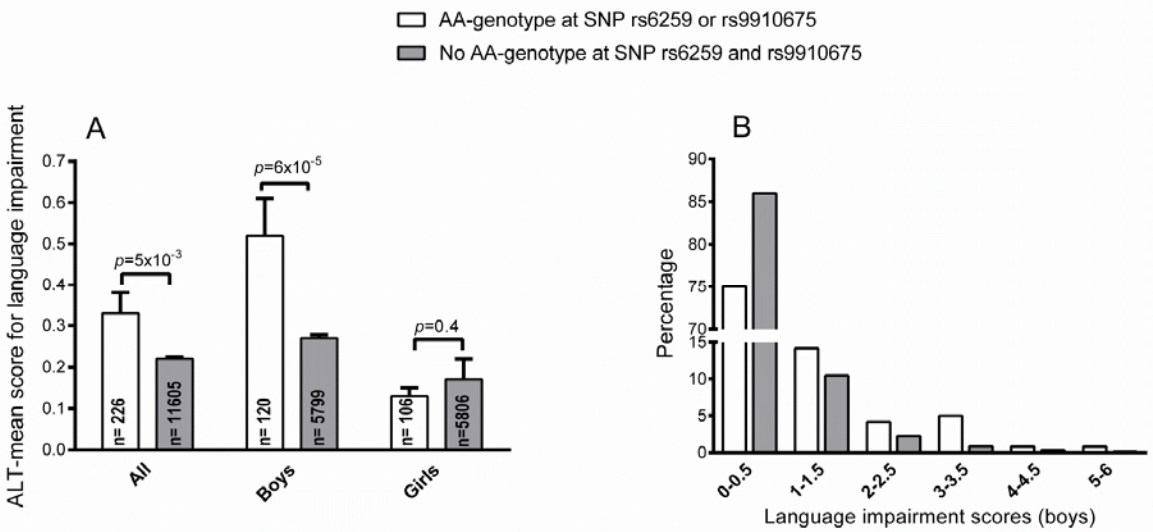
AA-genotype at SNP rs6259 or rs9910675
No AA-genotype at SNP rs6259 and rs9910675
A
ALT-mean score for language impairment
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
p=5x10-3
p=6x10-5
p=0.4
n= 226
n= 11605
n= 120
n= 5799
n= 106
n=5806
All
Boys
Girls
B
Percentage
90
85
80
75
70
15
10
5
0
0-0.5
1-1.5
2-2.5
3-3.5
4-4.5
5-6
Language impairment scores (boys)

Fig 1.

Table S1. Information about analyzed polymorphisms.

Gene locus	Chrom	SNP	Location	MAF CATSS	MAF 1000Genomes*
SHBG	17p13	rs9901675	5’	0.06	0.08
SHBG	17p13	rs12150660	Intron 1	0.24	0.20
SHBG	17p13	rs6258	Exon 4	0.01	0.01
SHBG	17p13	rs6259	Exon 8	0.11	0.13
SHBG	17p13	rs727428	3’	0.45	0.46
SHBG	17p13	rs1625895	3’	0.12	0.12
SRD5A2	2p23	rs523349	Exon 1	0.32	0.28
SRD5A2	2p23	rs9282858	Exon 1	0.02	0.07
ESR1	6q25	rs2747648	3’-UTR	0.03	0.04

*The 1000Genomes data base can be found at: http://www.1000genomes.org/. The presented MAF refers to the CEU (Utah residents with Northern and Western European ancestry) population.

Table S2. Associations between SNPs of *SHBG*, *ESR1* and *SRD5A2* and autistic-like traits in the total sample (CATSS-12425).

					ALT modules		
CATSS-12425[1]				Total ALT score	Language impairment	Social interaction impairment	Restricted and repetitive behavior
SHBG:rs9901675							
	All	A/A	50	0.72 (1.18)	0.39 (0.74)	0.20 (0.36)	0.13 (0.31)
		A/G	1223	0.62 (1.22)	0.20 (0.50)	0.22 (0.52)	0.19 (0.50)
		G/G	10695	0.70 (1.39)	0.23 (0.55)	0.24 (0.57)	0.22 (0.56)
		p		0.36	0.11	0.55	0.15
	Boys	A/A	24	1.02 (1.49)	0.56 (0.97)	0.29 (0.46)	0.16 (0.28)
		A/G	610	0.73 (1.31)	0.23 (0.51)	0.25 (0,58)	0.24 (0.54)
		G/G	5356	0.88 (1.64)	0.28 (0.63)	0.29 (0.64)	0.29 (0.66)
		p		0.20	0.03	0.52	0.17
	Girls	A/A	26	0.44 (0.73)	0.23 (0.40)	0.11 (0.21)	0.09 (0.34)
		A/G	613	0.50 (1.11)	0.16 (0.48)	0.19 (0.45)	0.14 (0.45)
		G/G	5339	0.52 (1.07)	0.17 (0.44)	0.19 (0.48)	0.15 (0.42)
		p		0.96	0.80	0.75	0.84
SHBG:rs12150660							
	All	GG	6307	0.72 (1.45)	0.24 (0.57)	0.25 (0.59)	0.22 (0.57)
		TG	3903	0.69 (1.37)	0.22 (0.53)	0.24 (0.56)	0.22 (0.56)
		TT	625	0.79 (1.57)	0.25 (0.62)	0.29 (0.66)	0.25 (0.60)
		p		0.88	0.51	0.68	0.68
	Boys	GG	3212	0.89 (1.69)	0.30 (0.65)	0.29 (0.66)	0.29 (0.67)
		TG	1930	0.88 (1.59)	0.26 (0.60)	0.30 (0.63)	0.30 (0.64)
		TT	319	1.02 (1.94)	0.34 (0.73)	0.32 (0.76)	0.35 (0.74)
		p		0.82	0.40	0.93	0.40
	Girls	GG	3102	0.53 (1.12)	0.17 (0.46)	0.20 (0.49)	0.15 (0.45)
		TG	1973	0.52 (1.09)	0.17 (0.44)	0.19 (0.46)	0.15 (0.43)
		TT	306	0.56 (1.01)	0.16 (0.45)	0.25 (0.53)	0.14 (0.40)
		p		0.82	0.88	0.31	0.89
SHBG:rs6258							
	All	CC	11814	0.70 (1.39)	0.23 (0.55)	0.24 (0.56)	0.22 (0.56)
		CT+TT	181	0.61 (1.23)	0.18 (0.39)	0.21 (0.55)	0.21 (0.51)
		p		0.57	0.21	0.71	0.92
	Boys	CC	5912	0.87 (1.62)	0.28 (0.62)	0.29 (0.64)	0.29 (0.65)
		CT+TT	95	0.67 (1.45)	0.21 (0.42)	0.22 (0.64)	0.24 (0.56)
		p		0.30	0.20	0.35	0.61
	Girls	CC	5902	0.52 (1.08)	0.17 (0.45)	0.19 (0.47)	0.14 (0.43)
		CT+TT	86	0.54 (0.92)	0.15 (0.36)	0.20 (0.43)	0.18 (0.45)
		p		0.83	0.46	0.75	0.29
SHBG:rs727428			n	mean score (sd)	mean score (sd)	mean score (sd)	mean score (sd)
	All	CC	3643	0.69 (1.3)	0.23 (0.5)	0.24 (0.57)	0.21 (0.54)
		CT	5870	0.68 (1.4)	0.22 (0.5)	0.24 (0.57)	0.22 (0.56)

		TT	2409	0.72 (1.4)	0.25 (0.57)	0.50 (0.57)	0.22 (0.57)
		p		0.82	0.51	0.92	0.88
	Boys	CC	1822	0.87 (1.6)	0.29 (0.6)	0.3 (0.64)	0.29 (0.64)
		CT	2952	0.84 (1.6)	0.2 (0.60)	0.28 (0.64)	0.29 (0.66)
		TT	1193	0.93 (1.7)	0.3 (0.7)	0.31 (0.67)	0.29 (0.68)
		p		0.49	0.14	0.55	0.98
	Girls	CC	1821	0.51 (1.06)	0.17 (0.44)	0.20 (0.48)	0.14 (0.41)
		CT	2918	0.52 (1.1)	0.18 (0.47)	0.19 (0.48)	0.15 (0.43)
		TT	1216	0.52 (1.02)	0.17 (0.42)	0.18 (0.44)	0.15 (0.44)
		p		0.88	0.83	0.77	0.84
SHBG:rs1625895							
	All	CC	9310	0.69 (1.3)	0.23 (0.55)	0.24 (0.56)	0.21 (0.55)
		CT	2398	0.69 (1.3)	0.22 (0.53)	0.24 (0.58)	0.22 (0.58)
		TT	206	0.81 (1.4)	0.25 (0.57)	0.24 (0.52)	0.29 (0.61)
		p		0.55	0.54	0.97	0.23
	Boys	CC	4690	0.85 (1.61)	0.28 (0.63)	0.28 (0.63)	0.28 (0.63)
		CT	1177	0.91 (1.67)	0.28 (0.60)	0.31 (0.68)	0.31 (0.70)
		TT	97	1.0 (1.5)	0.31 (0.62)	0.32 (0.59)	0.34 (0.68)
		p		0.40	0.79	0.35	0.42
	Girls	CC	4620	0.53 (1.10)	0.17 (0.45)	0.20 (0.48)	0.14 (0.43)
		CT	1221	0.47 (1.0)	0.15 (0.43)	0.18 (0.46)	0.14 (0.40)
		TT	109	0.64 (1.18)	0.21 (0.51)	0.18 (0.44)	0.24 (0.53)
				0.39	0.33	0.39	0.11
SRD5A2:rs523349							
	All	CC	5594	0.71 (1.41)	0.23 (0.56)	0.25 (0.58)	0.22 (0.57)
		GC	5182	0.69 (1.38)	0.23 (0.54)	0.24 (0.55)	0.22 (0.55)
		GG	1199	0.67 (1.33)	0.21 (0.53)	0.24 (0.56)	0.21 (0.53)
		p		0.77	0.73	0.65	0.91
	Boys	CC	2810	0.90 (1.69)	0.29 (0.64)	0.30 (0.68)	0.30 (0.67)
		GC	2564	0.87 (1.61)	0.28 (0.62)	0.28 (0.63)	0.29 (0.65)
		GG	622	0.76 (1.40)	0.24 (0.58)	0.26 (0.58)	0.25 (0.56)
		p		0.22	0.38	0.39	0.22
	Girls	CC	2784	0.51 (1.03)	0.16 (0.44)	0.19 (0.46)	0.14 (0.42)
		GC	2618	0.51 (1.08)	0.17 (0.44)	0.19 (0.47)	0.14 (0.41)
		GG	577	0.58 (1.25)	0.18 (0.47)	0.21 (0.54)	0.18 (0.50)
		p		0.28	0.56	0.49	0.20
SRD5A2:rs9282858							
	All	GG	11605	0.70 (1.41)	0.23 (0.55)	0.24 (0.57)	0.22 (0.56)
		GA+AA	485	0.61 (0.99)	0.19 (0.48)	0.22 (0.45)	0.18 (0.41)
		p		0.18	0.14	0.51	0.16
	Boys	GG	5828	0.88 (1.64)	0.28 (0.63)	0.29 (0.65)	0.29 (0.66)
		GA+AA	228	0.70 (1.06)	0.22 (0.52)	0.24 (0.47)	0.23 (0.49)
		p		0.13	0.11	0.32	0.18
	Girls	GG	5777	0.52 (1.09)	0.17 (0.45)	0.19 (0.48)	0.15 (0.43)
		GA+AA	257	0.52 (0.90)	0.17 (0.44)	0.21 (0.44)	0.13 (0.32)
		p		0.97	0.82	0.63	0.73

ESR1:rs2747648							
	All	CC	16	1.40 (1.85)	0.43 (0.57)	0.53 (0.74)	0.43 (0.77)
		CT	687	0.78 (1.69)	0.25 (0.64)	0.28 (0.68)	0.24 (0.60)
		TT	11436	0.69 (1.38)	0.22 (0.54)	0.24 (0.56)	0.22 (0.56)
		p		0.06	0.33	0.03	0.16
	Boys	CC	7	2.14 (2.42)	0.57 (0.67)	0.71 (0.85)	0.85 (1.02)
		CT	341	1.01 (2.05)	0.31 (0.76)	0.36 (0.80)	0.33 (0.74)
		TT	5722	0.86 (1.61)	0.28 (0.62)	0.28 (0.64)	0.29 (0.65)
		p		0.07	0.45	0.06	0.06
	Girls	CC	9	0.83 (1.08)	0.33 (0.5)	0.38 (0.65)	0.11 (0.22)
		CT	344	0.54 (1.19)	0.18 (0.49)	0.20 (0.52)	0.15 (0.40)
		TT	5714	0.52 (1.08)	0.17 (0.45)	0.19 (0.47)	0.15 (0.43)
		p		0.55	0.52	0.40	0.77

[1]Results from analyses of SNPs in the total sample consisting of 12425 individuals.

p: p-values after comparisons between all genotypes.

Few individuals (9 and 2 respectively) carry the uncommon homozygotes of rs9282858 and rs6258 and therefore those were collapsed with the heterozygotes.