Genome-wide association study of semen volume, sperm concentration, testis size, and plasma inhibin B levels

Youichi Sato[1,*], Atsushi Tajima[2,*], Misaki Kiguchi[1], Suzu Kogusuri[1], Aki Fujii[1], Takehiro Sato[2], Shiari Nozawa[3], Miki Yoshiike[3], Makiko Mieno[4], Kosuke Kojo[5,6], Masahiro Uchida[5,7], Haruki Tsuchiya[5], Kazumitu Yamasaki[5,7], Issei Imoto[8], Teruaki Iwamoto[3,5,9]

[1] Department of Pharmaceutical Information Science, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, 770-8505, Japan

[2] Department of Bioinformatics and Genomics, Graduate School of Advanced Preventive Medical Sciences, Kanazawa University, Ishikawa, 920-8640, Japan

[3] Department of Urology, St. Marianna University School of Medicine, Kanagawa, 216-8511, Japan

[4] Department of Medical Informatics, Center for Information, Jichi Medical University, Tochigi, 329-0498, Japan

[5] Center for Infertility and IVF, International University of Health and Welfare Hospital, Tochigi, 329-2763, Japan

[6] Department of Urology, University of Tsukuba Hospital, Ibaraki, 305-8576, Japan

[7] Urology department, Tsukuba Gakuen Hospital, Ibaraki, 305-0854, Japan

[8] Department of Molecular Genetics, Aichi Cancer Center Research, Aichi, 464-8681, Japan

[9] Department of Male Infertility, Reproduction Center, Sanno Hospital, Tokyo, 107-0052, Japan

* Address correspondence to: Youichi Sato, Ph.D.

Department of Pharmaceutical Information Science, Institute of Biomedical Sciences, Tokushima University Graduate School

1-78-1 Sho-machi, Tokushima, Tokushima 770-8505, Japan

Phone: +81-88-633-7253; FAX: +81-88-633-7253

E-mail: youichi.sato@tokushima-u.ac.jp

or

Atsushi Tajima, Ph.D.

Department of Bioinformatics and Genomics, Graduate School of Advanced Preventive Medical Sciences, Kanazawa University

13-1 Takara-machi, Kanazawa, Ishikawa 920-8640, Japan

Phone: +81-76-265-2715; FAX: +81-76-234-4249

E-mail: atajima@med.kanazawa-u.ac.jp

Abstract

Semen quality is affected by environmental factors, endocrine function abnormalities, and genetic factors. A GWAS recently identified *ERBB4* at 2q34 as a genetic locus associated with sperm motility. However, GWASs for human semen volume and sperm concentration have not been conducted. In addition, testis size also reportedly correlates with semen quality, and it is important to identify genes that affect testis size. Reproductive hormones also play an important role in spermatogenesis. To date, genetic loci associated with plasma testosterone, sex hormone-binding globulin (SHBG), follicle stimulating hormone (FSH), and luteinizing hormone (LH) levels have been identified using GWASs. However, GWASs have not identified any relevant loci for plasma inhibin B levels. We conducted a two-stage GWAS using 811 Japanese men in a discovery stage followed by a replication stage using an additional 721 Japanese men. The results of the discovery and replication stages were combined into a meta-analysis. After setting a suggestive significance threshold for *P* values $< 5 \times 10^{-6}$ in the discovery stage, we identified ten regions with SNPs (semen volume: one, sperm concentration: three, testes size: two, and inhibin B: four). We selected only the most significant SNP in each region for replication genotyping. Combined discovery and replication results in the meta-analysis showed that the locus 12q21.31 associated with plasma inhibin B levels (rs11116724) had the most significant association ($P = 5.7 \times 10^{-8}$). The *LRRIQ1* and *TSPAN19* genes are located in the 12q21.31 region. This study provides new susceptibility variants that contribute to plasma inhibin B levels.

Introduction

Examination of male infertility requires a semen test. The World Health Organization [1] has established the test criteria, which include semen volume, sperm concentration, and sperm motility. Thus, semen quality affects fertility. Semen quality is affected by environmental factors like smoking and stress, endocrine function abnormalities including testosterone, and genetic factors. Regarding genetic factors, polymorphisms of circadian regulator (*CLOCK*) [2] and steroid 5α-reductase 2 (*SRD5A2*) [3] genes are reportedly associated with human semen volume. It has been reported that protamine 1 (*PRM1*), *PRM2* [4], deleted in azoospermia-like (*DAZL*) [5], *CLOCK* [2], cytochrome P450 family 19 subfamily A peptide 1 (*CYP19A1*) [6], histidine-rich glycoprotein (*HRG*) [7], and voltage-dependent anion channel (*VDAC*) [8] gene polymorphisms show significant association with human semen concentration. Genome-wide association studies (GWASs) are widely used to identify variants associated with diseases and quantitative traits. We used a GWAS to identify *ERBB4* at 2q34 as a locus associated with sperm motility [9]. However, GWASs for human semen volume and sperm concentration have not been conducted yet, and the related genetic factors are poorly understood. In addition, testis size has also been reported to correlate with sperm density, total sperm count, and total motile sperm count [10]. Thus it is important to identify the genes that affect testis size.

Hormones are physiologically active substances that are synthesized and secreted by endocrine glands, such as the thyroid and pituitary glands. They act on target cells via blood and other pathways. Very small hormone amounts can have regulatory effects on digestive absorption, circulation, respiration, immunity, and metabolism, and they also contribute to homeostasis. The major reproductive hormones, follicle

stimulating hormone (FSH) and luteinizing hormone (LH), are produced by the pituitary gland after stimulation by gonadotropin-releasing hormone. In the testes, secreted LH induces testosterone secretion by Leydig cells, and FSH promotes spermatogenesis in Sertoli cells in the presence of testosterone [11,12]. Approximately 50 to 60% of testosterone in circulation binds to sex hormone-binding globulin (SHBG), which decreases its biological activity [13]. Inhibin B produced in Sertoli cells acts on the hypothalamus pituitary system in a negative feedback loop along with estradiol to inhibit FSH production. These hormones play an important role in sexual development, spermatogenesis, and testicular function maintenance in men [14,15]. Since these hormone values are heritable, genetic factors are thought to exist [16]. GWASs have shown that genetic loci associated with reproductive hormone levels exist in some populations. *SHBG* (rs12150660, rs6258, rs727428, and rs2075230), family with sequence similarity 9 member B (*FAM9B*) (rs5934505), and Jumonji domain-containing 1C (*JMJD1C*) (rs10822184) are loci significantly associated with testosterone levels in European and Chinese populations [17-19]. *SHBG* (rs12150660, rs6258, rs2075230, and rs12150660), protein arginine methyltransferase 6 (*PRMT6*) (rs17496332), glucokinase regulator (*GCKR*) (rs780093), zinc finger and BTB domain containing 10 (*ZBTB10*) (rs440837), *JMJD1C* (rs7910927), solute carrier organic anion transporter family member 1B1 (*SLCO1B1*) (rs4149056), nuclear receptor subfamily 2 group F member 2 (*NR2F2*) (rs8023580), zinc finger protein 652 (*ZNF652*) (rs2411984), teratocarcinoma-derived growth factor 3 (*TDGF3*) (rs1573036), and BAI1 associated protein 2 like 1 (*BAIAP2L1*) (rs3779195) are loci significantly associated with SHBG levels in European and Chinese populations [17,19,20]. *CYP19A1* (rs2414095) and follicle stimulating hormone subunit beta (*FSHB*) (rs11031005) are loci significantly associated with FSH

levels in the Chinese population and United Kingdom twin data [17,21]. *FSHB* (rs11031002) is a locus significantly associated with LH levels in United Kingdom twin data [21]. However, a GWAS focusing on reproductive hormones in the Japanese population has not been conducted, and inhibin B level-related loci have not yet been identified using GWASs.

Here, we conducted a GWAS in 811 Japanese men to clarify the genetic determinants of human semen volume, sperm concentration, testis size, and plasma inhibin B levels for which genetic loci have not yet been identified. The associations were validated in an additional Japanese man. For SNPs that were indicated to be related, we also investigated the association with other parameters and male infertility. In addition, we conducted the validation analyses of previously known SNPs that have shown genome-wide significant associations with testosterone, SHBG, FSH or LH associated in non-Japanese populations, only using Japanese men in the discovery stage of the current GWAS.

Materials and Methods

Subjects

This study was approved by the ethics committees each university and hospital. All participants provided written informed consent.

We performed a two-stage genetic association study. The discovery stage included 811 men (20.7 ± 1.7 years old, mean ± SD) from the general Japanese population who were recruited from university students in three study centers based in departments of urology at university hospitals in Japan (Kawasaki, Kanazawa, and Nagasaki) as previously reported [22]. The replication stage included 721 men (31.1 ± 4.8 years old, mean ± SD) of proven fertility recruited from the partners of pregnant women who attended obstetric clinics in four cities in Japan (Sapporo, Kanazawa, Osaka, and Fukuoka) [23]. The characteristics for the two stage subjects are summarized in Table 1.

Clinical trait measurements

Clinical trait measurement for these subjects has been described in previous reports [22,23]. Briefly, age, body weight, height, and ejaculation abstinence period were self-reported. BMI (kg/m^2) was calculated from body weight and height. Semen samples were obtained once by masturbation after sexual abstinence for at least 48 h and were ejaculated into clean, wide-necked, sterile, nontoxic collection containers. The samples were protected from extreme temperatures and liquefied at 37°C before examination. Semen volume was measured with a graduated 5 ml syringe (Terumo; Tokyo, Japan). The sperm concentration of each sample was assessed using a Bürker-Türk hemocytometer. Testis sizes were measured using a Prader Orchidometer (Pharmacia & Upjohn, Copenhagen, Denmark). An average of the left and right testis was reported as

the testicular size. Blood samples were drawn from the cubital vein of each subject usually in the mornings to reduce the effect of diurnal variations in hormone levels. Inhibin B was measured using a specific two-sided enzyme immunometric assay (Serotec, Kidlington, UK). Testosterone, SHBG, FSH, and LH levels were determined using a time-resolved immunofluorometric assay (Delfia, Wallac, Turku, Finland) at the Department of Growth and Reproduction, Rigshospitalet, in Copenhagen, Denmark.

Genotyping, quality control and imputation

Genomic DNA was extracted from the peripheral blood samples of subjects using a QIAamp DNA blood kit (Qiagen; Tokyo, Japan). Previously genotyped and imputed data were used in the discovery stage [9]. Briefly, 816 men were genotyped using the Illumina HumanCore V1.0 DNA Analysis Kit (Illumina; Tokyo, Japan) following the manufacturer's instructions. 298,930 SNPs were genotyped. The quality control of genotyped SNPs and samples was conducted using the PLINK version 1.07 software package (http://pngu.mgh.harvard.edu/~purcell/plink/) [24]. Of the 816 samples, four were excluded because there were duplicates or familial relationships and one was excluded because it was genetic outlier based on a principle component analysis-based method using the genotyped data of the HapMap CHB and JPT as internal controls. Finally, 811 samples were included for genome-wide association analysis. For genotype imputation, we filtered out 56,462 SNPs from our data set according to the following criteria: without reference SNP (rs) ID, monomorphic in our sample set, with call rates lower than 0.98, or genotyped by duplicated probes. There were no SNPs excluded based on the Hardy-Weinberg equilibrium test ($P \geq 1.0\times10^{-6}$). Genotype data were flipped to the forward strand with conform-gt, which is the utility program of BEAGLE 4.1 [25,26],

using genotype data for Asian samples (JPT and CHB) of the 1000 Genomes Project [27,28] as a reference panel. Imputation was performed using BEAGLE 4.1. 1000 Genomes Project Phase 3 v5 was used as a reference panel. SNPs with R^2 lower than 0.8 and all indels were excluded from imputed genotype data. Finally, we obtained genotypes for 3,901,256 SNPs, which were used for subsequent association analyses.

In the replication stage, SNPs were genotyped using TaqMan probes (rs13400448, C_31177579_10; rs1501607, C_1490652_10; rs7724206, C_27428161_10; rs1761601, C_1803412_10; rs8179990, C_29718786_10; rs6955111, Custom; rs2346597, C_15778404_10; rs4870827, C_28925_10; rs2440587, C_16240853_10; rs12817187, C_32172890_10; rs149500, C_ 2025086_10; rs11116719, C_ 2025119_10; and rs11116724, C_12125451_10; Applied Biosystems; Tokyo, Japan) with the ABI 7900HT real-time PCR system (Applied Biosystems).

rs12817187 and rs11116724 were directly genotyped in 200 randomly selected samples of discovery subjects to confirm the concordance of imputed results. The concordance of typing results between genotyped and imputed was 100%.

Statistical analysis

The analyses for semen volume and sperm concentration were processed using square-root-transformed. Testosterone, SHBG, FSH, and LH were processed using natural log-transformed to minimize deviation from a normal distribution.

In the discovery and replication stage, the associations between SNPs and testis size, and hormone values were assessed using a standardized multiple linear regression under an additive genetic model with adjustments for age and BMI using PLINK or the R version 3.5.0 software package (http://www.R-project.org/). Semen volume and sperm

concentration were additionally adjusted for ejaculation abstinence, and sperm motility was additionally adjusted for ejaculation abstinence and time from masturbation to semen evaluation. We set a suggestive threshold of *P* values $< 5 \times 10^{-6}$ in the discovery stage. The results from the discovery and replications were combined in a meta-analysis using the meta package for the R software. The extent of heterogeneity among studies was quantified using the I^2 statistic [29] and statistically assessed using Cochran's *Q* test. If there was no heterogeneity based on an I^2 statistic < 50% or a *P*-value > 0.1, a fixed-effects model using the inverse variance method was used. Otherwise, the random-effects model using the DerSimonian and Laird method was used.

The Manhattan and quantile-quantile plots were generated using qqman package for the R software, and the regional plot was created by LocusZoom using the 1000 Genomes project Asian (ASN) data (Nov 2014) [30]. HaploReg V 4.1 (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php) was used for functional annotation analysis of variants [31]. Significantly expressed quantitative trait loci (eQTLs) were searched on GTEx Portal database (http://www.gtexportal.org/home/) [32].

Results

GWAS for semen volume, sperm concentration, testis size, and plasma inhibin B levels

We conducted a two-stage GWAS to identify loci associated with human semen volume, sperm concentration, testis size, and plasma inhibin B levels. Manhattan plots and quantile-quantile plots of the GWAS were presented in Figure 1 and Supplementary Figure S1, respectively. The genomic-control inflation factors (λ) of semen volume, sperm concentration, testis size, and inhibin B level were 1.00650, 1.00875, 1.00000, and 1.01398, respectively, indicating the unlikelihood of inflation of the false-positive association. We could not find any SNPs that reached the genome-wide significance level. Therefore, we set a suggestive threshold of *P* values < 5.0×10^{-6}. The 2p16.1 locus was suggestively associated with semen volume, the 3q26.31, 5q23.1, and 10p15.1 loci were suggestively associated with sperm concentration, the 3q13.11 and 7q22.1 loci were suggestively associated with testis size, and the 4q13.1, 8q24.13, 9q22.31, and 12q21.31 loci were suggestively associated with plasma inhibin B levels (Supplementary Table S1).

We conducted a replication study for the most significant SNPs (rs13400448 at 2p16.1, rs1501607 at 3q26.31, rs7724206 at 5q23.1, rs1761601 at 10p15.1, rs8179990 at 3q13.11, rs6955111 at 7q22.1, rs2346597 at 4q13.1, rs4870827 at 8q24.13, rs2440587 at 9q22.31, and rs12817187 at 12q21.31) at each locus. In the replication stage with 721 men, rs12817187 showed a significant association with plasma inhibin B levels (β = 0.17, $P = 3.9\times 10^{-3}$) (Table 2). However, none of the SNPs displayed any association with semen volume, sperm concentration, and testis size. When we combined the discovery and replication stages using meta-analysis, we found that rs12817187 had a suggestive association (β = 0.21, $P = 9.5 \times 10^{-8}$) with plasma inhibin B levels (Table 2). The variance

in inhibin B level based on rs12817187 was 1.8%.

To confirm the most significant SNP in the region where rs12817187 is located, we selected rs1494500, rs11116719, and rs11116724, which are SNPs with *P* values < 1.0×10^{-5} and not completely in linkage disequilibrium with the rs12817187 SNP that indicated association with plasma inhibin B levels (Supplementary Table S2), and conducted a replication study. When we combined the discovery and replication stages using meta-analysis, we found that rs11116724 was most significantly associated ($\beta = 0.22$, $P = 5.7 \times 10^{-8}$) with plasma inhibin B levels (Table 3).

Association of rs11116724 with other parameters and male infertility

We further analyzed the association of the rs11116724 SNP with other parameters (including semen volume, sperm concentration, testis size, plasma testosterone, SHBG, FSH, and LH levels) and risk of male infertility (including nonobstructive azoospermia, oligozoospermia, and asthenozoospermia) (Supplementary Materials and Methods). rs11116724 was marginally associated with sperm concentration ($\beta = 0.081$, $P = 0.042$) and testis size ($\beta = 0.12$, $P = 0.036$) (Supplementary Table S3). However, these associations were not significant after adjusting for plasma inhibin B levels (data not shown). This is mainly because plasma inhibin B levels are positively correlated with sperm concentration and testis size (Supplementary Table S4). rs11116724 did not indicate significant association with male infertility risk (Supplementary Table S5).

Functional annotations

Regional plots of the region 400 kb upstream or downstream of rs11116724 are

shown in Figure 2. rs11116724 is located in the intron of leucine rich repeats and the IQ motif containing 1 (*LRRIQ1*) gene, and the *LRRIQ1* and tetraspanin 19 (*TSPAN19*) genes are located in the 12q21.31 region.

We used the HaploReg database to obtain putative functional annotations of these loci [32]. The summary of SNPs in high LD compared to rs11116724 based on the HaploReg database is shown in Supplementary Table S6. Of the 103 SNPs in high LD, rs7312075, rs3765044 and rs17012533 were missense variants in the *LRRIQ1* gene. We predicted the effect of three SNPs on LRRIQ1 function using two *in silico* SNP prediction algorithms: Sorting Intolerant From Tolerant (SIFT) [33] and PolyPhen-2 [34]. rs7312075 was one of three SNPs classified as damaging using SIFT analysis (Supplementary Table S7). Although the most significant SNP, rs11116724, was not associated with promoter histone marks, enhancer histone marks, DNase I hypersensitive regions, and protein binding, 33 other SNPs were associated with promoter histone marks, enhancer histone marks, DNase I hypersensitive regions in various cell types and tissues, and/or protein binding. 93 SNPs, including rs11116724, could alter nucleotide sequences of several regulatory motifs. In addition, we found that rs11116724 was a significant eQTL for *TSPAN19* in the testes. The effect allele (A) of rs11116724 was associated with decreased *TSPAN19* ($P = 1.1 \times 10^{-5}$) expression in the testes according to the GTEx Portal V8 (Supplementary Figure S2) [32].

GWAS for testosterone, SHBG, FSH, and LH in Japanese

Next, we conducted a GWAS for testosterone, SHBG, FSH, and LH, whose loci were identified by previous GWASs, on 811 Japanese men. Manhattan plots and quantile-quantile plots of the GWAS are presented in Supplementary Figures S3 and S4,

respectively. However, none of the SNPs reached the genome-wide significance level ($P < 5.0 \times 10^{-8}$). In addition, the results of the association analysis between hormone levels in Japanese men and reported SNPs that reached the genome-wide significance levels were summarized in Supplementary Table S8. rs2075230, which was reported to be strongly associated with SHBG levels ($P = 4.75 \times 10^{-19}$) in Chinese men, showed the highest significance ($P = 3.2 \times 10^{-5}$) in Japanese men (Supplementary Table S8).

Discussion

In this GWAS for human semen volume, sperm concentration, testis size, and plasma inhibin B levels, we showed that locus 12q21.31 is suggestively associated with plasma inhibin B levels. This association was also validated at different ages between the same ethnic groups.

The *LRRIQ1* and *TSPAN19* genes are located in the 12q21.31 region. *LRRIQ1* is most strongly expressed in the testes and strongly localized to the seminiferous ducts in the testes. LRRIQ1's function is not well understood, but it interacts with 28 proteins like activin A receptor type 1 (ACVR1), activin A receptor type 2A (ACVR2A), and activin A receptor type 2B (ACVR2B) [35]. Activin has the opposite effect to inhibin; it promotes FSH secretion. Activin receptor was identified in 1991 [36], and it was reported that inhibin also functions via a common receptor with activin [37]. Although it is not known how the interaction of the activin receptor and LRRIQ1 affects inhibin B levels, it may be that differences in inhibin B levels occur in response to signal changes from the receptor. *TSPAN19*, located in the 12q21.31 region, is a member of the tetraspanin family with four transmembrane domains, but its function is not well understood. *TSPAN19* is strongly expressed in the testes and lungs and specifically localized to Leydig cells in the testes. Since inhibin B is secreted by Sertoli cells in the testes, LRRIQ1 and/or TSPAN19 may somehow contribute to inhibin B secretion. We could not find significant eQTL for *LRRIQ1* in the testes, but three SNPs (rs7312075, rs3765044, and rs17012533) in high LD with the most significant SNP, rs11116724, were missense variants in the *LRRIQ1* gene. rs7312075 was one of three SNPs to be classified as damaging based on SIFT and PolyPhen-2 analyses. Therefore, rs7312075 in high LD with rs11116724 damages LRRIQ1 function and may affect inhibin B secretion. For TSPAN19, the effect allele of

rs11116724 was associated with decreased *TSPAN19* expression in the testes according to the GTEx Portal V8. On the other hand, the effect allele of rs11116724 was associated with increased inhibin B levels. Thus, it is suggested that the effect allele of rs11116724 increases inhibin B expression by decreasing TSPAN19 expression.

On the other hand, we could not identify the loci related to semen volume, sperm concentration, or testis size. No single SNP had a large effect size for these phenotypes. Thus, multi-factor inheritance or environmental factors may have an influence.

We could not identify novel SNPs that reached the genome-wide significance level for testosterone, SHBG, FSH, and LH in Japanese men. Some of the reported SNPs that reached the genome-wide significance levels showed association, but they did not reach the genome-wide significance level in Japanese men. This may be due to differences in sample size and/or ethnic group.

In conclusion, this GWAS identified the novel locus 12q21.31 as suggestively associated with plasma inhibin B levels. *LRRIQ1* and *TSPAN19* genes are located in the 12q21.31 region. *LRRIQ1* and *TSPAN19* are expressed specifically in the testes. The functions of these genes are unknown, but it has been suggested that these genes contribute to inhibin B secretion since inhibin B is secreted from Sertoli cells. Future studies are needed to identify functional variant(s) directly related to plasma inhibin B levels at this locus by targeted resequencing of the two candidate genes. In addition, future replication studies in diverse populations with larger sample sizes are essential since our association did not reach the genome-wide significance level.

Acknowledgements

We are grateful to the late Prof. Yutaka Nakahori and Profs. Eitetsue Koh, Jiro Kanaya, Mikio Namiki, Kiyomi Matsumiya, Akira Tsujimura, Kiyoshi Komatsu, Naoki Itoh, and Jiro Eguchi for collecting blood samples from participants. We also thank Prof. Toyomasa Katagiri for his assistance with the AB GeneAmp PCR system 9700. This study was supported in part by the Ministry of Health and Welfare of Japan (1013201) (to T.I.), Grant-in-Aids for Scientific Research (C) (26462461) (to Y.S.), (23510242) (to A.T) and Grant-in-Aids for Scientific Research (B) (17H04331) (to Y.S), (15H04320) (to A.T) from the Japan Society for the Promotion of Science, and the European Union (BMH4-CT96-0314) (to T.I.). We would like to thank Editage (www.editage.com) for English language editing.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

1. World Health Organization Department of Reproductive Health and Research. WHO laboratory manual for the examination and processing of human semen, fifth edition. WHO. 2010.

2. Zhang J, Ding X, Li Y, Xia Y, Nie J, Yi C, et al. Association of CLOCK gene variants with semen quality in idiopathic infertile Han-Chinese males. Reprod Biomed Online. 2012;25:536–42.

3. Zhao D, Wu W, Xu B, Niu X, Cui H, Zhang Y, et al. Variants in the SRD5A2 gene are associated with quality of semen. Mol Med Rep. 2012;6:639–44.

4. Tüttelmann F, Krenková P, Römer S, Nestorovic AR, Ljujic M, Stambergová A, et al. A common haplotype of protamine 1 and 2 genes is associated with higher sperm counts. Int J Androl. 2010;33:e240–8.

5. Teng YN, Chang YP, Tseng JT, Kuo PH, Lee IW, Lee MS, et al. A single-nucleotide polymorphism of the DAZL gene promoter confers susceptibility to spermatogenic failure in the Taiwanese Han. Hum Reprod 2012;27:2857–65.

6. Sato Y, Tajima A, Katsurayama M, Nozawa S, Yoshiike M, Koh E, et al. A replication study of a candidate locus for follicle-stimulating hormone levels and association analysis for semen quality traits in Japanese men. J Hum Genet. 2016;61:911–5.

7. Lindgren KE, Nordqvist S, Kårehed K, Sundström-Poromaa I, Åkerud H. The effect of a specific histidine-rich glycoprotein polymorphism on male infertility and semen parameters. Reprod Biomed Online. 2016;33:180–8.

8. Pan L, Liu Q, Li J, Wu W, Wang X, Zhao D, Ma J. Association of the VDAC3 gene polymorphism with sperm count in Han-Chinese population with idiopathic male infertility. Oncotarget 2017;8:45242–8.

9. Sato Y, Tajima A, Sato T, Nozawa S, Yoshiike M, Imoto I, et al. Genome-wide association study identifies ERBB4 on 2q34 as a novel locus associated with sperm motility in Japanese men. J Med Genet. 2018;55:415–21.

10. Sakamoto H, Yajima T, Nagata M, Okumura T, Suzuki K, Ogawa Y. Relationship between testicular size by ultrasonography and testicular function: measurement of testicular length, width, and depth in patients with infertility. Int J Urol. 2008;15:529–33.

11. Belchetz PE, Plant TM, Nakai EJ, Knobil E. Hypophysial responses to continuous and intermittent delivery of hypopthalamic gonadotropin-releasing hormone. Science. 1978;202:631–3.

12. Smith MA, Vale WW. Desensitization to gonadotropin-releasing hormone observed in superfused pituitary cells on Cytodex beads. Endocrinology. 1981;108:752–9.

13. Kaufman JM, Vermeulen A. The decline of androgen levels in elderly men and its

clinical and therapeutic implications. Endocr Rev. 2005;26:833–76.

14. McLachlan RI, O'Donnell L, Meachem. SJ, Stanton PG, De Kretser DM, Pratis K, et al. Identification of specific sites of hormonal regulation in spermatogenesis in rats, monkeys and man. Recent Prog Horm Res. 2002;57:149–79.

15. Schlatt S, Ehmcke J. Regulation of spermatogenesis: An evolutionary biologist's perspective. Semin Cell Dev Biol. 2014;29C:2–16.

16. Kuijper EA, Lambalk CB, Boomsma DI, van der Sluis S, Blankenstein MA, de Geus EJ, et al. Heritability of reproductive hormones in adult male twins. Hum Reprod. 2007;22:2153–9.

17. Ohlsson C, Wallaschofski H, Lunetta KL, Stolk L, Perry JR, Koster A, et al. Genetic determinants of serum testosterone concentrations in men. PLoS Genet. 2011;7:e1002313.

18. Jin G, Sun J, Kim ST, Feng J, Wang Z, Tao S, et al. Genome-wide association study identifies a new locus JMJD1C at 10q21 that may influence serum androgen levels in men. Hum Mol Genet. 2012;21:5222–8.

19. Chen Z, Tao S, Gao Y, Zhang J, Hu Y, Mo L, et al. Genome-wide association study of sex hormones, gonadotropins and sex hormone-binding protein in Chinese men. J Med Genet. 2013;50:794–801.

20. Coviello AD, Haring R, Wellons M, Vaidya D, Lehtimäki T, Keildson S, et al. A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple Loci implicated in sex steroid hormone regulation. PLoS Genet. 2012;8:e1002805.

21. Ruth KS, Campbell PJ, Chew S, Lim EM, Hadlow N, Stuckey BG, et al. Genome-wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. Eur J Hum Genet. 2016;24:284–90.

22. Iwamoto T, Nozawa S, Mieno MN, Yamakawa K, Baba K, Yoshiike M, et al. Semen quality of 1559 young men from four cities in Japan: a cross-sectional population-based study. BMJ Open. 2013;3:e002222.

23. Iwamoto T, Nozawa S, Yoshiike M, Namiki M, Koh E, Kanaya J, et al. Semen quality of fertile Japanese men: a cross-sectional population-based study of 792 men. BMJ Open. 2013;3:e002223.

254 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.

25. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81:1084–97.

26. Browning BL, Browning SR. Genotype imputation with millions of reference samples. Am J Hum Genet. 2016;98:116–26.

287 The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526:68–74.

28. The 1000 Genomes Project Consortium. An integrated map of structural variation in 2,504 human genomes. Nature 2015;526:75–81.

29. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.

30. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336–7.

31. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40 (Database issue):D930–4.

32. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45:580–5.

33. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31:3812–4.

34. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013;Chapter7:Unit7.20.

35. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al. Architecture of the human interactome defines protein communities and disease networks. Nature 2017;545:505–9.

36. Mathews LS, Vale WW. Expression cloning of an activin receptor, a predicted transmembrane serine kinase. Cell 1991;65:973–2.

37. Lewis KA, Gray PC, Blount AL, MacConell LA, Wiater E, Bilezikjian LM, et al. Betaglycan binds inhibin and can mediate functional antagonism of activin signaling. Nature 2000;404:411–4.

Figure legends

Figure 1. Manhattan plot from GWAS for semen volume (A), sperm concentration (B), testes size (C), and Inhibin B (D). The negative $\log_{10}$-transformed *P* values (y axis) of SNPs are shown according to their position on the chromosome. The horizontal line represents suggestive (*P* value = 5×10^{-6}) (dotted line) and genome-wide (*P* value = 5×10^{-8}) (solid line) threshold.

Figure 2. Regional plot of the associations between SNPs at the region 400 kb upstream or downstream of lead SNP of inhibin B. The negative $\log_{10}$-transformed *P* values (y axis) of the genotyped and imputed SNPs are shown according to their position. Purple diamond indicates the lead SNP, and color indicates LD strength with the lead SNP. The right Y-axis shows the recombination rate estimated from the 1000 Genomes project Asian (ASN) data (Nov 2014).

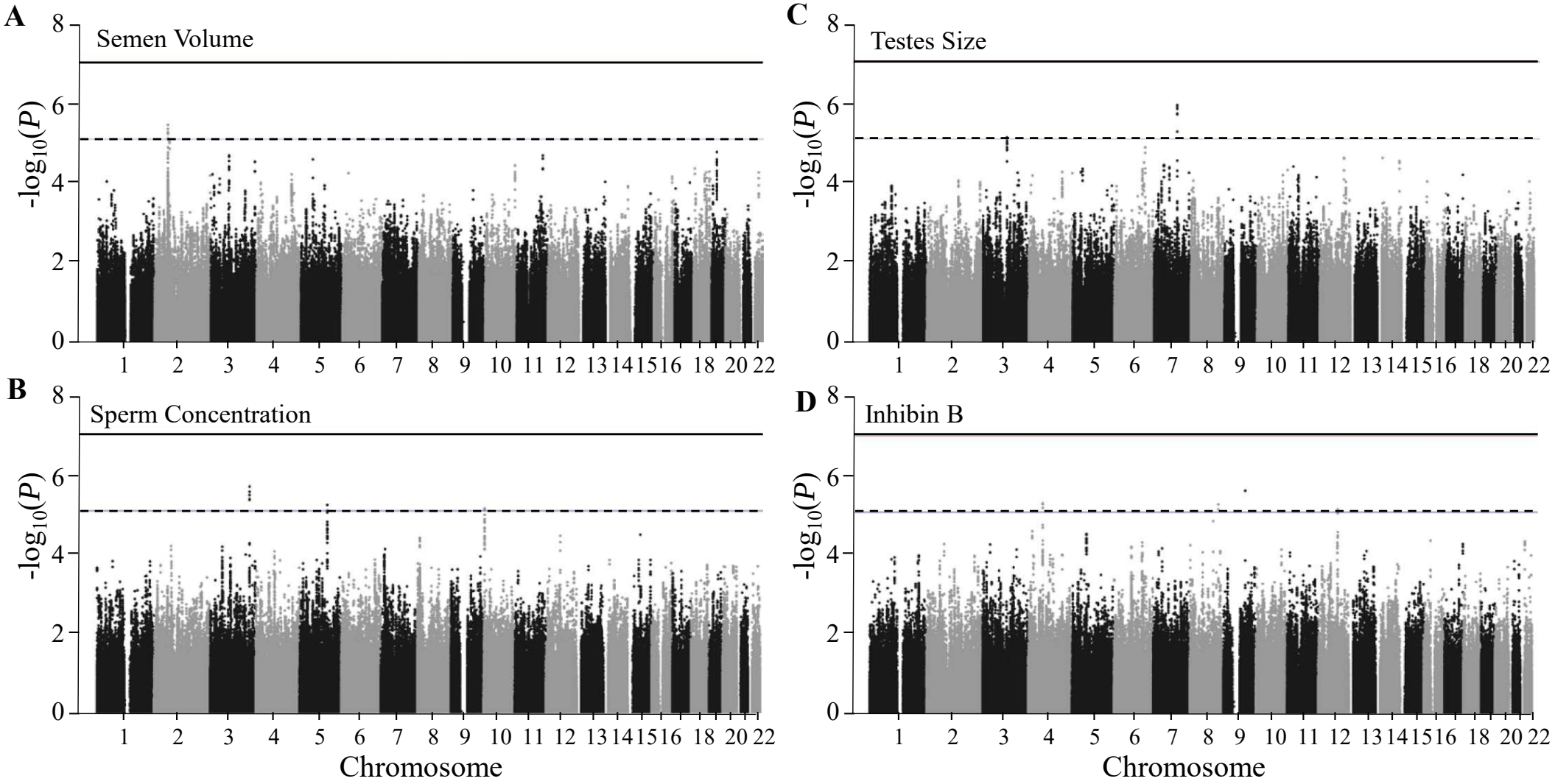
A
Semen Volume
B
Sperm Concentration
C
Testes Size
D
Inhibin B
-log10(P)
0
2
4
6
8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 20 22
Chromosome

Fig 1

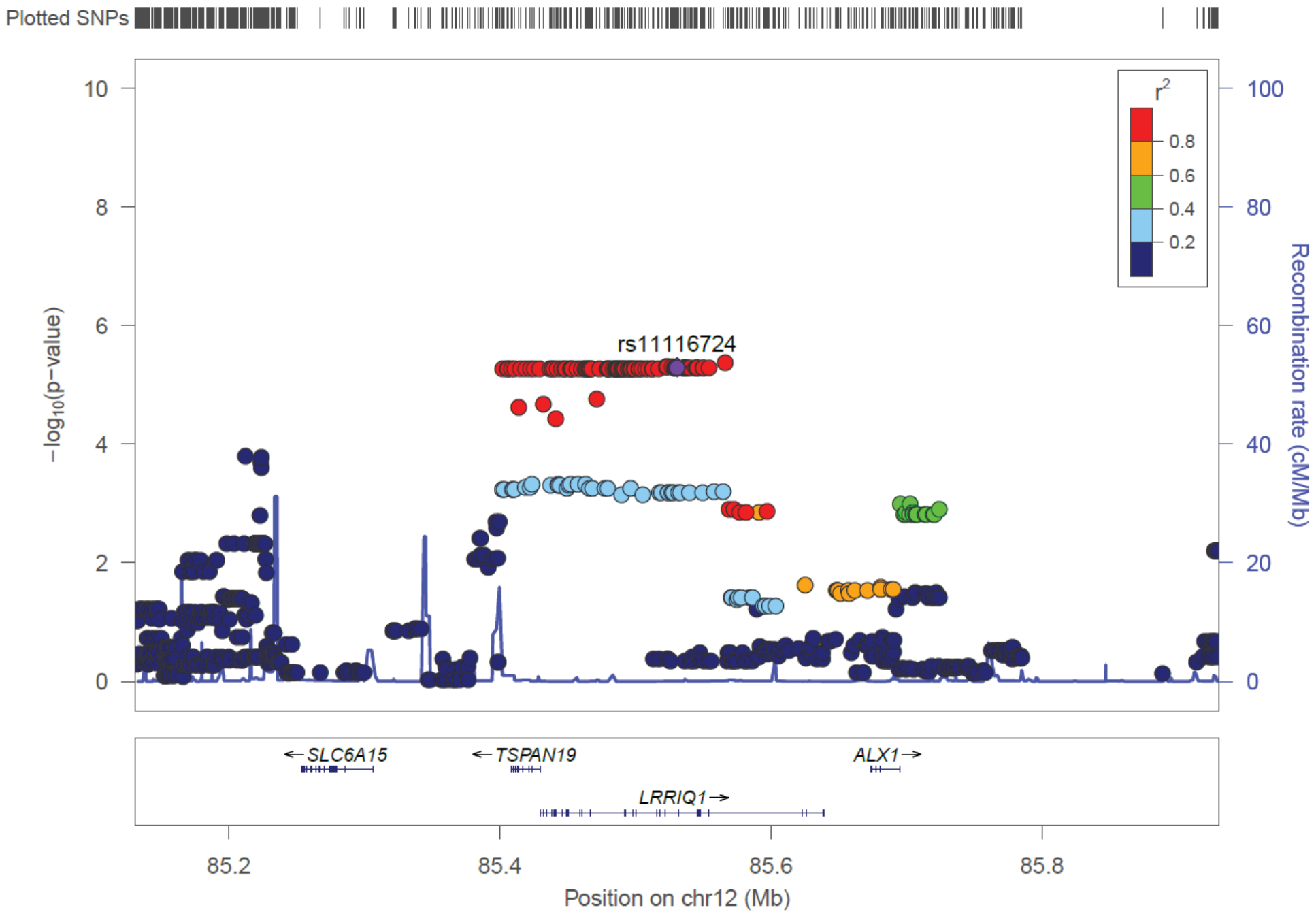
Plotted SNPs
rs11116724
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
SLC6A15
TSPAN19
ALX1
LRRIQ1
Position on chr12 (Mb)

Fig 2

Table 1. Characteristics of subjects

	Discovery (*N*=811)	Replication (*N*=721)
Age (years)	20.7 ± 1.7	31.1 ± 4.8
BMI (kg/m^2)	21.5 ± 2.5	23.3 ± 3.1
Ejaculation abstinence (hours)	79.8 ± 39.5	199.8 ± 338.2
Semen volume (ml)	2.9 ± 1.4	3.1 ± 1.5
Sperm concentration (×10^6/ml)	72.1 ± 56.9	106.7 ± 83.3
Testis size (ml)	20.8 ± 4.3	20.0 ± 3.6
Inhibin-B (pg/ml)	197.8 ± 63.9	177.0 ± 55.3
Testosterone (nmol/l)	23.1 ± 7.5	19.6 ± 6.7
SHBG (nmol/l)	26.5 ± 10.3	33.3 ± 14.6
FSH (U/l)	2.4 ± 1.2	4.1 ± 2.1
LH (U/l)	3.2 ± 1.3	3.7 ± 1.5

Data are presented as mean ± standard deviation. Testis size was calculated from an average of the left and right testis.

Abbreviations: BMI, body mass index; SHBG, sex hormone-binding globulin; FSH, follicle-stimulating hormone; LH, luteinizing hormone.

Table 2. Summary results for the most significant SNPs in loci identified the association for semen volume, sperm concentration, testis size, and plasma inhibin B levels in the discovery stage

SNP (effect/reference)	Chr.	Position (hg19)	Stage	Genotype	EAF	β_{STD} (SE)	*P*	*P*het[a]	I^2 (%)[b]
Semen volume									
rs13400448	2p16.1	57,135,651	Discovery	144/391/276	0.42	-0.23 (0.048)	2.1×10^{-6}		
(C/T)			Replication	117/336/268	0.40	-0.081 (0.053)	0.13		
			Meta-analysis			-0.16 (0.074)	0.034	0.039	76.6
Sperm concentration									
rs1501607	3q26.31	170,942,533	Discovery	86/348/377	0.32	0.24 (0.051)	1.6×10^{-6}		
(T/C)			Replication	91/314/316	0.34	-0.0093 (0.054)	0.86		
			Meta-analysis			0.12 (0.12)	0.35	0.0006	91.5
rs7724206	5q23.1	119,614,887	Discovery	105/379/327	0.36	0.23 (0.050)	3.5×10^{-6}		
(C/A)			Replication	94/324/303	0.36	-0.00095 (0.054)	0.99		
			Meta-analysis			0.12 (0.12)	0.32	0.0015	90.1
rs1761601	10p15.1	5,200,059	Discovery	9/168/634	0.11	-0.35 (0.075)	4.3×10^{-6}		
(T/C)			Replication	2/153/566	0.11	-0.14 (0.088)	0.12		
			Meta-analysis			-0.25 (0.11)	0.019	0.068	70.0
Testes size									
rs8179990	3q13.11	103,776,410	Discovery	53/306/452	0.25	0.25 (0.055)	4.7×10^{-6}		
(T/C)			Replication	59/294/368	0.29	0.0047 (0.57)	0.94		
			Meta-analysis			0.13 (0.12)	0.30	0.0018	89.7
rs6955111	7q22.1	99,582,199	Discovery	21/207/583	0.15	-0.33 (0.066)	6.6×10^{-7}		
(C/T)			Replication	7/173/541	0.13	-0.036 (0.080)	0.65		
			Meta-analysis			-0.19 (0.15)	0.20	0.0046	87.6
Inhibin B									
rs2346597	4q13.1	63,939,334	Discovery	22/201/588	0.15	0.32 (0.067)	2.9×10^{-6}		
(C/T)			Replication	21/207/493	0.17	0.0043 (0.068)	0.95		
			Meta-analysis			0.16 (0.16)	0.30	0.0012	90.5
rs4870827	8q24.13	123,946,041	Discovery	176/423/212	0.48	-0.23 (0.50)	3.2×10^{-6}		
(G/A)			Replication	125/389/207	0.44	0.080 (0.054)	0.14		
			Meta-analysis			-0.078 (0.16)	0.62	<0.0001	94.5
rs2440587	9q22.31	94,138,821	Discovery	57/297/457	0.25	0.27 (0.055)	1.4×10^{-6}		
(T/C)			Replication	58/266/397	0.27	-0.023 (0.057)	0.69		
			Meta-analysis			0.12 (0.15)	0.40	0.0002	92.6
rs12817187	12q21.31	85,566,208	Discovery	53/322/436	0.27	0.26 (0.056)	4.2×10^{-6}		
(A/G)			Replication	57/282/382	0.27	0.17 (0.057)	3.9×10^{-3}		
			Meta-analysis			0.21 (0.040)	9.5×10^{-8}	0.24	27.6

Data are shown as the estimated standardized linear regression statistic β_{STD}, SE, and *P* value using an additive genetic model with adjustments for age and BMI. Semen volume and sperm concentration were additionally adjusted for ejaculation abstinence. The semen volume and sperm concentration were processed using square-root-transformed values.

[a]Cochran's Q test heterogeneity *P* value.

[b] I^2 heterogeneity index.

Abbreviations: SNP, single-nucleotide polymorphism; Chr, chromosome; EAF, effect allele frequency; β_{STD}, standardized regression coefficient.

Table 3. Summary results for SNPs at the 12q21.31 locus identified the association for plasma inhibin B levels in the discovery stage

SNP (effect/reference)	Position (hg19)	Stage	Genotype	EAF	β_{STD} (SE)	*P*	*P*het[b]	I^2 (%)[c]
rs1494500	85,402,328	Discovery	55/319/437	0.26	0.25 (0.056)	5.3×10^{-6}		
(C/T)		Replication	61/273/387	0.27	0.16 (0.056)	5.2×10^{-3}		
		Meta-analysis			0.21 (0.040)	1.7×10^{-7}	0.22	33.6
rs11116719	85,522,932	Discovery	55/320/436	0.27	0.26 (0.056)	4.8×10^{-6}		
(T/C)		Replication	55/283/383	0.27	0.16 (0.057)	5.4×10^{-3}		
		Meta-analysis			0.21 (0.040)	1.5×10^{-7}	0.23	30.1
rs11116724	85,530,856	Discovery	55/321/435	0.27	0.26 (0.056)	5.0×10^{-6}		
(A/G)		Replication	59/275/387	0.27	0.17 (0.057)	2.2×10^{-3}		
		Meta-analysis			0.22 (0.040)	5.7×10^{-8}	0.30	6.7
rs12817187[a]	85,566,208	Discovery	53/322/436	0.27	0.26 (0.056)	4.2×10^{-6}		
(A/G)		Replication	57/282/382	0.27	0.17 (0.057)	3.9×10^{-3}		
		Meta-analysis			0.21 (0.040)	9.5×10^{-8}	0.24	27.6

Data are shown as the estimated standardized linear regression statistic β_{STD}, SE, and *P* value using an additive genetic model with adjustments for age and BMI.

[a]rs12817187 is the most significant SNP in discovery stage.

[b]Cochran's Q test heterogeneity *P* value.

[c]I^2 heterogeneity index.

Abbreviations: SNP, single-nucleotide polymorphism; Chr, chromosome; EAF, effect allele frequency; β_{STD}, standardized regression coefficient.