Lack of replication of four candidate SNPs implicated in human male fertility traits: a large-scale population-based study

Youichi Sato[1], Atsushi Tajima[2], Kouki Tsunematsu[1], Shiari Nozawa[3], Miki Yoshiike[3], Eitetsue Koh[4], Jiro Kanaya[4], Mikio Namiki[4], Kiyomi Matsumiya[5], Akira Tsujimura[6], Kiyoshi Komatsu[7], Naoki Itoh[8], Jiro Eguchi[9], Issei Imoto[2], Aiko Yamauchi[1], Teruaki Iwamoto[3,10]

[1]Department of Pharmaceutical Information Science, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima, 770-8505, Japan

[2]Department of Human Genetics, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima, 770-8503, Japan

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

[4]Department of Urology, Kanazawa University Graduate School of Medical Sciences, Kanazawa, 920-8641, Japan

[5]Department of Urology, Osaka Police Hospital, Osaka, 543-0035, Japan

[6]Department of Urology, Graduate School of Medicine, Faculty of Medicine, Osaka University, Osaka, 565-0871, Japan

[7]Department of Urology, Harasanshinkai Hospital, Fukuoka, 812-0033, Japan

[8]Department of Urology, Sapporo Medical University, Sapporo, 060-8543, Japan

[9]Department of Urology, Nagasaki University, Nagasaki, 852-8523, Japan

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

Address correspondence to: Youichi Sato, Ph.D.
Department of Pharmaceutical Information Science, Institute of Health Biosciences
The University of Tokushima Graduate School
1-78-1 Sho-machi, Tokushima City 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 Human Genetics, Institute of Health Biosciences
The University of Tokushima Graduate School
3-18-15 Kuramoto, Tokushima 770-8503, Japan
Phone: +81-88-633-7075; FAX: +81-88-633-7453
e-mail: tajima.atsushi@tokushima-u.ac.jp

Running title: Non-replicated loci for male fertility traits

Abstract

STUDY QUESTION: Are the four candidate loci (rs7867029, rs12870438, rs7174015, and rs724078) for human male fertility traits, identified in a genome-wide association study (GWAS) of a Hutterite population in the USA, associated with semen quality traits in a Japanese population?

SUMMARY ANSWER: These four single nucleotide polymorphisms (SNPs) rs7867029, rs12870438, rs7174015, and rs724078 have no association with semen parameters in meta-analysis of two Japanese male cohorts.

WHAT IS KNOWN ALREADY: Four (rs7867029, rs12870438, rs7174015, and rs724078) of the SNPs associated with family size or birth rate in the GWAS of a Hutterite population in the USA were associated with semen parameters in ethnically diverse men from Chicago, USA.

STUDY DESIGN, SIZE, DURATION: This is a replication study in a total of 2015 Japanese subjects, including 791 fertile men and 1224 young men from the general population.

PARTICIPANTS/MATERIALS, SETTING, METHODS:We performed a replication study in two cohorts to assess whether the SNPs rs7867029, rs12870438, rs7174015, and rs724078 are associated with sperm concentration, semen volume, total sperm numbers, total motile sperm numbers, or sperm motility. The rs12870438 SNP was detected by restriction fragment length polymorphism PCR while rs7174015, rs724078, and rs7867029 SNPs were genotyped using TaqMan probes.

MAIN RESULTS AND THE ROLE OF CHANCE: This study indicated that none of the four SNPs rs7867029, rs12870438, rs7174015, and rs724078 displayed a significant association with semen parameters in the meta-analysis of two Japanese male cohorts.

LIMITATIONS, REASONS FOR CAUTION: Only four SNPs identified in the Hutterite GWAS were examined for associations with semen quality traits in a Japanese population. In addition, the linkage disequilibrium structures around the testing markers were different between ethnic groups.

WIDER IMPLICATIONS OF THE FINDINGS: Locus mapping studies using a set of tagging SNPs across the loci will be necessary in populations with larger sample sizes in order to understand the contribution of specific genes to semen quality.

STUDY FUNDING/COMPETING INTEREST (S): 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) (23510242) (to A.T.) from the Japan Society for the Promotion of Science, the European Union (BMH4-CT96-0314) (to T. I.), and the Takeda Science Foundation (to A.T.). None of the authors has any competing interests to declare.

Keywords: replication study/ semen quality/ male fertility/Japanese population

Introduction

Many cases of male infertility are caused by spermatogenic failure such as azoospermia, oligozoospermia, or asthenozoospermia; in addition, decreased semen quality can also result in an elevated risk of male infertility. However, the genetic determinants for human semen quality are poorly understood.

To date, four genome-wide association studies (GWASs) regarding male fertility and infertility have been reported. A pilot GWAS in Caucasians (92 cases and 80 controls) showed that 20 single nucleotide polymorphisms (SNPs) were significantly ($P < 10^{-5}$) associated with azoospermia or oligozoospermia (Aston *et al*., 2009). Furthermore, two GWASs in Chinese men have revealed common variants located near *PRMT6* (which encodes protein arginine N-methyltransferase 6), *PEX10* (which encodes peroxisome biogenesis factor 10), and *SOX5* (which encodes SRY related HMG-box gene 5) and within the *HLA* region that are associated with risk for nonobstructive azoospermia (Hu *et al*., 2012; Zhao *et al*., 2012). The findings from these two Chinese GWASs have been evaluated in independent Japanese cohorts (Jinam *et al*., 2013; Sato *et al*., 2013). Lastly, a GWAS of 269 married Hutterite men in the USA, a culture that traditionally proscribes contraception and uniformly desires large families, revealed 41 SNPs that are significantly correlated with family size or birth rate ($P < 1 \times 10^{-4}$). In the subsequent validation study using 123 ethnically diverse men composed mainly of Hispanics and African Americans, nine of the 41 SNPs were also reported to be associated with reduced sperm quantity and/or function (Kosova *et al*., 2012). The associations of these nine SNPs with reduced fertility and sperm parameters remain to be confirmed in additional, larger cohorts.

Four of the nine SNPs detected in the GWAS for male fertility traits were found

to be associated with sperm concentration, semen volume, total sperm count, total motile sperm count, and/or sperm motility (Kosova *et al*., 2012; also see Table I), and they are thought to be common in the Japanese population because of their minor allele frequencies (MAFs) > 0.05 in the HapMap-JPT population. In this study, to further clarify the contribution of these four SNPs to semen quality in diverse populations, we conducted a replication study to assess whether the four SNPs were associated with sperm parameters in two large Japanese cohorts.

Materials and Methods

This study was approved by the ethics committees of the University of Tokushima and St. Marianna Medical University. All participants provided written informed consent.

Two Japanese cohort samples

Two Japanese cohorts, namely, 791 men of proven fertility and 1224 young men from the general population, were included in the replication study. Some of the subjects in this study have been described in previous reports (Iwamoto *et al*., 2013a, b). Briefly, fertile men were recruited from the partners of pregnant women who attended obstetric clinics in four cities in Japan (Sapporo, Kanazawa, Osaka, and Fukuoka) (Iwamoto *et al*., 2013a). The eligibility criteria for the male participants were as follows: the participants had to have been aged 20–45 years at the time of invitation by the hospital at which they were recruited, and both the man and his mother had to have been born in and living in Japan. In addition, the pregnancy of the female partner had to have been the result of conception by sexual intercourse and not by fertility treatment. Young men from the general Japanese population were recruited from university students in three study centers based in the urology departments at university hospitals in Japan (Kawasaki, Kanazawa, and Nagasaki), as previously reported (Iwamoto *et al*., 2013b). In addition, we recruited university students at a study center in Sapporo. Inclusion criteria were that the man was 18–24 years old and that both he and his mother had been born in Japan.

Physical examination and semen analysis in the two cohorts

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 and analyzed as previously described (Iwamoto *et al*., 2013a, b). Briefly, 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 extremes of temperature and were then liquefied at 37°C prior to their examination. The sperm concentration of each sample was assessed using a Bürker-Türk hemocytometer. Semen volume was measured with a graduated 5-mL syringe (Terumo; Tokyo, Japan). Sperm motility was assessed from 10 μL of well-mixed semen placed on a clean glass slide, covered, and then examined at a total magnification of 400× at 37°C. The motility assessment was repeated twice, and the average value from two samples was calculated. The sperm were assessed using the World Health Organization (WHO) motility classes A, B, C, and D (World Health Organization, 1999). In this study, sperm in classes A and B were considered as motile.

SNP selection and genotyping

Genomic DNA was extracted from the peripheral blood samples of subjects using a QIAamp DNA blood kit (Qiagen; Tokyo, Japan). From SNPs previously reported to show association with sperm concentration, semen volume, total sperm count, total motile sperm count, and/or sperm motility (Kosova *et al*., 2012), 4 SNPs (rs7867029, rs12870438, rs7174015, and rs724078) with MAFs > 0.05 in the HapMap-JPT population were selected for genotyping. These four SNPs were reportedly associated with two or more of the five sperm parameters of interest in this study at permutation-based *P*-values < 0.05 (Table I). The rs12870438 SNP was detected by

restriction fragment length polymorphism PCR using the following primer sets: 5′-GCAAACAGGAGAAGGGTGTT -3′ (forward) and 5′-GCTTTGGAGCATGTTTTCCC -3′ (reverse). DNA from each subject was amplified using Taq DNA polymerase (Promega; Tokyo, Japan) under the appropriate amplification conditions. The resulting PCR products were then digested using the *Hha*I restriction enzyme (New England Biolabs Japan Inc.; Tokyo, Japan). The digested products were separated by electrophoresis on a 2.5% agarose gel. The following fragment sizes were used for allele identification on gels: 488 bp (A-allele) and 278 + 210 bp (G-allele). The rs7174015, rs724078, and rs7867029 SNPs were genotyped using TaqMan probes rs7174015 (C_32072246_10), rs724078 (C_2500858_10), and rs7867029 (C_31364474_20; Applied Biosystems; Tokyo, Japan) with the ABI 7900HT real-time PCR system (Applied Biosystems).

Statistical analysis

Hardy–Weinberg equilibrium (HWE) was assessed in two cohort samples by using an internet-based HWE calculator (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl).

The analyses for sperm concentration, semen volume, total sperm number, and total motile sperm number were processed using square-root transformed values to minimize deviation from a normal distribution. The associations between SNPs and semen parameters were assessed using multiple linear regression with adjustments for age, BMI, and ejaculation abstinence in each of the two cohorts. Sperm motility and total motile sperm number were additionally adjusted for time from masturbation to test. The results from the two cohorts were combined in a meta-analysis using the meta package for the R version 3.0.2 statistical environment (http://www.R-project.org/). The extent of

heterogeneity among studies was quantified by the I^2 statistic (Higgins *et al.*, 2003) and statistically assessed by the Cochran's *Q* test. If there was no heterogeneity, as determined by the I^2 statistic less than 50% or a *P* value more than 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 employed.

All statistical analyses were performed using R version 3.0.2 (The R Project for Statistical Computing [http://www.r-project.org]). For replication purposes, only SNP-trait associations observed in the previous GWAS (Kosova *et al.*, 2012) were tested assuming the specific genetic models reported (n = 11 tests; Table I). Statistical significance was considered at *P*-values < 0.0045 (0.05/11) to account for multiple testing.

Results

Semen characteristics of the two cohorts

The characteristics of semen from fertile Japanese men and from young men from the general Japanese population are presented in Supplementary Table SI. As previously reported (Iwamoto *et al*, 2013b), except in the case of sperm motility, semen parameters for men from the general population were significantly lower than those for fertile men.

Association analysis of four SNPs and semen parameters in fertile men, and young men from the general population in Japan

To investigate the associations between the four SNPs (rs7867029, rs12870438, rs7174015, and rs724078) and semen parameters we genotyped these SNPs in a total of 2015 men. The allele and genotype frequencies of the four SNPs analyzed in each cohort are shown in Table II. The genotyping of the SNPs is complete except for rs12870438 (the missing genotyping rate is 0.1%), and the genotypes of all four SNPs were in HWE in the respective two cohorts ($P > 0.05$). Then, we assessed the associations between the four SNPs and semen parameters using a multiple linear regression analysis of the two cohorts. In this study, we performed an association analysis with semen parameters that were related to the minor allele in Hutterites, as reported in a previous GWAS under the association model (Kosova *et al*., 2012). Multiple linear regression analysis revealed that rs7867029 showed a trend toward a negative association with sperm motility ($\beta = -1.98$, $P = 0.026$) in young men from the general Japanese population, and rs12870438 showed a trend toward a positive association with total sperm numbers (TSN) ($\beta = 7.80$, $P = 0.028$) in fertile men (Table

III). However, none of the four SNPs reached the adjusted *P*-value for multiple testing ($P < 0.0045$). Next, to assess the strength of the association, we conducted a combined analysis using a meta-analysis of the two Japanese male cohorts. However, unlike the results of the previous study (Kosova *et al*., 2012), none of the four SNPs displayed a significant association with semen parameters. Furthermore, there were no associations observed between the four SNPs and other semen parameters in three genetic models (additive, Supplementary Table SII; recessive, Supplementary Table SIII; and dominant, Supplementary Table SIV) in the combined analysis.

Discussion

Recently, four (rs7867029, rs12870438, rs7174015, and rs724078) of the 41 SNPs correlated with family size or birth rate ($P < 1 \times 10^{-4}$) in the GWAS of 269 Hutterite men in the USA were found to be associated with sperm concentration, semen volume, total sperm count, total motile sperm count, or sperm motility in 123 ethnically diverse men from Chicago (Kosova *et al*., 2012). Additionally, we recently showed that of the four SNPs, rs7867029, rs7174015, and rs12870438 were significantly associated with the risk of developing oligozoospermia, and rs12870438 was also associated with azoospermia (Sato *et al*., submitted). In the present study, there was limited evidence of a significant association ($P < 0.05$) between these four SNPs and one or more of the five semen parameters in each of two Japanese replication cohorts, whereas none of the four SNPs displayed a significant association with any of the semen parameters. The current replication meta-analysis is well-powered to detect associations of semen quality trait loci with modest effect sizes because this provides > 80% power for SNPs that explain 1% or higher of total phenotypic variance. The observed heterogeneity of the SNP-trait associations between previous (Kosova *et al*., 2012) and this study, as well as between the two Japanese cohorts, may be attributed to potential biases in selection of the study subjects; in the previous study, most of the 123 subjects were Hispanic (58.5%) and had been referred for infertility evaluation at the University of Illinois Andrology Laboratory, Chicago, IL, USA, while 791 men of proven fertility and 1224 general controls were separately recruited for population-based assessment of semen quality in Japanese men.

Several limitations of this study should be noted. In this study, only the SNPs identified originally in the Hutterite GWAS were examined for associations with semen quality traits under three genetic models in a Japanese population. Owing to between-

population differences in linkage disequilibrium (LD) structures around the SNPs examined, the tested SNPs may not be in high LD with unidentified true causal variants in Japanese subjects (Supplementary Figures 1-4). The low LD between the genotyped SNPs and causal variants could increase the likelihood of false-negative findings, through the lowering of the statistical power of the analysis (Clarke *et al*., 2007). The differences in the extent of LD and the underlying haplotype structures between populations could also affect the fit of the specified genetic model to the data obtained, and the direction of the effect for the associated allele. This may account for apparent inconsistencies in the model fitting between previous and this studies. To overcome these limitations, fine-scale LD mapping of the fertility trait loci using a set of tagging SNPs across the loci will be necessary in populations with larger sample sizes. The locus mapping studies will allow for a better understanding of susceptibility genes contributing to semen quality in humans.

Acknowledgements

We thank all the volunteers who participated in this study. We are grateful to the late Prof. Yutaka Nakahori for collecting blood samples from the participants. We also thank Prof. Toyomasa Katagiri for his assistance with the AB GeneAmp PCR system 9700.

Authors' roles

Y.S. and A.T.: study design and data analysis; Y.S. and K.T.: genotyping; S.N., M.Y., E.K., J.K., M.N., K.M., A.T., K.K., N.I., J.E., and T.I.: cohort collection and characterization; Y.S., A.T., K.T., S.N., M.Y., E.K., J.K., M.N., K.M., A.T., K.K., N.I., J.E., I.I., A.Y., and T.I.: preparation and approval of the final version of the manuscript.

Funding

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) (23510242) (to A.T.) from the Japan Society for the Promotion of Science, the European Union (BMH4-CT96-0314) (to T.I.), and the Takeda Science Foundation (to A.T.).

Conflicts of interest

None declared.

References

Aston KI, Carrell DT. Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. *J Androl* 2009;**30**:711–725.

Clarke GM, Carter KW, Palmer LJ, Morris AP, Cardon LR. Fine mapping versus replication in whole-genome association studies. *Am J Hum Genet* 2007;**81**:995–1005.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M et al. The structure of haplotype blocks in the human genome. *Science* 2002;**296**:2225–2229.

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557–560.

Hu Z, Xia Y, Guo X, Dai J, Li H, Hu H, Jiang Y, Lu F, Wu Y, Yang X *et al*. A genome-wide association study in Chinese men identifies three risk loci for non-obstructive azoospermia. *Nat Genet* 2012;**44**:183–186.

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

Iwamoto T, Nozawa S, Mieno MN, Yamakawa K, Baba K, Yoshiike M, Namiki M, Koh E, Kanaya J, Okuyama A *et al*. Semen quality of 1559 young men from four cities in

Japan: a cross-sectional population-based study. *BMJ Open* 2013b;**3**:e002222.

Jinam TA, Nakaoka H, Hosomichi K, Mitsunaga S, Okada H, Tanaka A, Tanaka K, Inoue I. HLA-DPB1*04:01 allele is associated with non-obstructive azoospermia in Japanese patients. *Hum Genet* 2013;**132**:1405–1411.

Kosova G, Scott NM, Niederberger C, Prins GS, Ober C. Genome-wide association study identifies candidate genes for male fertility traits in humans. *Am J Hum Genet* 2012;**90**:950–961.

Sato Y, Jinam T, Iwamoto T, Yamauchi A, Imoto I, Inoue I, Tajima A. Replication study and meta-analysis of human nonobstructive azoospermia in Japanese populations. *Biol Reprod* 2013;**88**:87.

Sato Y, Tajima A, Tsunematsu K, Nozawa S, Yoshiike M, Koh E, Kanaya J, Namiki M, Matsumiya K, Tsujimura A *et al*. An association study of four candidate loci for human male fertility traits with male infertility. *Hum Reprod* submitted.

World Health Organization. WHO Laboratory Manual for the Examination of Human Semen and Sperm-cervical Mucus Interaction, 4th edn. Cambridge: Cambridge University Press; 1999.

Zhao H, Xu J, Zhang H, Sun J, Sun Y, Wang Z, Liu J, Ding Q, Lu S, Shi R *et al*. A genome-wide association study reveals that variants within the HLA region are

associated with risk for nonobstructive azoospermia. *Am J Hum Genet* 2012;**90**:900–906.

Legends of Supplementary Figures 1–4

Supplementary Figure S1. Comparisons of linkage disequilibrium (LD) patterns of rs7867029 for male fertility traits among human populations.

The white rectangle with black vertical bars represents a region in and around the fertility trait locus examined. The identifier of the single nucleotide polymorphism (SNP) genotyped and the genomic position are shown. The triangle represents the LD map calculated from SNPs with minor allele frequencies ≥ 0.05 in the respective populations: JPT, Japanese in Tokyo, Japan; CEU, Utah residents with ancestry from northern and western Europe; MEX, Mexican ancestry in Los Angeles, CA, USA; ASW, African ancestry in Southwest USA. The color of each square in the triangle expresses the extent of LD: black, $r^2 = 1$; shades of grey, $0 < r^2 < 1$; white $r^2 = 0$. LD blocks according to the definition of Gabriel et al. (2002) are indicated by bold black lines. The recombination rates obtained the HapMap database and the RefSeq genes within the region are shown in the panel above.

Supplementary Figure S2. Comparisons of linkage disequilibrium (LD) patterns of

rs12870438 for male fertility traits among human populations.

See legend to Supplementary Figure S1 for description of symbols.

Supplementary Figure S3. Comparisons of linkage disequilibrium (LD) patterns of rs7174015 for male fertility traits among human populations.

See legend to Supplementary Figure S1 for description of symbols.

Supplementary Figure S4. Comparisons of linkage disequilibrium (LD) patterns of rs724078 for male fertility traits among human populations.

See legend to Supplementary Figure S1 for description of symbols.

Table I. Summary of a previous genome-wide association study*

SNP	Chr	Position (NCBI Build 36.3)	Closest Genes[a]	Location	Previous GWAS*						
					Allele[b]	Model	Semen parameters in the Chicago men (Permutation *P*-value)					Other associated traits
							Conc.	Vol.	TSN	TMSN	Motility (%)	
rs7867029	9	80,210,238	*PSAT1*	dwnst.	G	Dominant	**0.042**	0.86	0.11	0.061	**0.0040**	FS; Avg. Veloc.; Mean ALH
rs12870438	13	42,378,205	*EPSTI1*	intron	A	Recessive	**0.0050**	0.50	**0.024**	**0.023**	0.11	FS; Avg. Veloc.; Mean ALH
rs7174015	15	48,504,360	*USP8*	intron	T	Recessive	0.080	**0.016**	**0.0011**	**0.0056**	0.35	FS; Avg. Veloc.; Mean ALH
rs724078	6	29,597,027	*MAS1L, UBD*	upst., dwnst.	T	Recessive	0.14	0.13	**0.023**	**0.018**	**0.041**	BR; Mean ALH

*(Kosova *et al*., 2012)

SNP: single nucleotide polymorphism, Chr, chromosome; dwnst., downstream; upst., upstream; Conc., sperm concentration; Vol., semen volume; TSN, total sperm numbers; TMSN, total motile sperm numbers; FS, family size; BR, birth rate; Avg. Veloc, average velocity; ALH, amplitude of lateral head displacement.

[a]Gene names: *PSAT1*, phosphoserine aminotransferase 1; *EPSTI1*, epithelial stromal interaction 1; *USP8*, ubiquitin specific peptidase 8; *MAS1L*, MAS1 oncogene-like; *UBD*, ubiquitin D.

[b]"Allele" indicates the Hutterite minor allele reported in previous GWAS (Kosova *et al*., 2012).

Bold numbers indicate statistical significance ($P < 0.05$) in previous GWAS (Kosova *et al*., 2012).

Table II. Allele frequencies in this study, previous GWAS* and HapMap populations

SNP	Allele[a]	Freq. (Genotypes[b]) in this study		Freq. in Previous GWAS		Freq. in HapMap populations (phase 3)			
		Fertile	Young	Hutterites	Chicago men	ASW	CEU	JPT	MEX
rs7867029	G	0.20 (27/256/508)	0.19 (43/389/792)	0.09	0.25	0.306	0.106	0.227	0.122
rs12870438	A	0.098 (4/148/638)	0.10 (14/221/988)	0.17	0.17	0.071	0.403	0.065	0.230
rs7174015	T	0.54 (226/396/169)	0.55 (365/608/251)	0.36	0.57	0.592	0.438	0.535	0.520
rs724078	T	0.29 (61/334/396)	0.29 (99/517/608)	0.27	0.47	0.582	0.274	0.285	0.460

Freq., allele frequencies; ASW, African ancestry in Southwest USA; CEU, Utah residents with Northern and Western European ancestry from the CEPH collection; JPT, Japanese in Tokyo, Japan; MEX, Mexican ancestry in Los Angeles, CA, USA.
[a]"Allele" indicates the Hutterite minor allele reported in previous GWAS *(Kosova *et al*., 2012).
[b]"Genotypes" indicate genotype counts (2/1/0).

Table III. An association analysis under the previously reported model* between four SNPs and semen parameters in fertile men, and young men from the general population in Japan

SNP	Model[a]	Semen Parameter	Fertile		Young		Combined		Heterogeneity	
			β (SE)	*P*	β (SE)	*P*	β (SE) [model][b]	P_{meta}	P_{het}	I^2(%)
rs7867029	Dominant	Conc.	0.24 (0.27)	0.37	-0.10 (0.18)	0.58	0.0075 (0.15) [F]	0.96	0.29	10.8
		Motility (%)	4.4 (2.60)	0.091	**-1.98 (0.89)**	**0.026**	0.74 (3.16) [R]	0.81	0.020	81.5
rs12870438	Recessive	Conc.	3.55 (1.81)	0.051	-0.049 (0.82)	0.95	1.39 (1.76) [R]	0.43	0.070	69.4
		TSN	**7.80 (3.54)**	**0.028**	0.82 (1.42)	0.56	3.55 (3.41) [R]	0.30	0.067	70.2
		TMSN	2.59 (2.78)	0.35	0.73 (1.16)	0.53	1.01 (1.07) [F]	0.35	0.54	0.0
rs7174015	Recessive	Vol.	0.013 (0.034)	0.71	0.0090 (0.025)	0.72	0.010 (0.020) [F]	0.61	0.93	0.0
		TSN	-0.082 (0.56)	0.88	-0.051 (0.33)	0.88	-0.059 (0.28) [F]	0.84	0.96	0.0
		TMSN	0.063 (0.40)	0.87	0.026 (0.27)	0.92	0.037 (0.22) [F]	0.87	0.94	0.0
rs724078	Recessive	TSN	0.24 (0.94)	0.80	-0.62 (0.55)	0.26	-0.40 (0.48) [F]	0.41	0.43	0.0
		TMSN	0.45 (0.75)	0.54	-0.42 (0.45)	0.35	-0.19 (0.39) [F]	0.63	0.32	0.9
		Motility (%)	4.55 (2.79)	0.10	0.38 (1.56)	0.81	1.37 (1.36) [F]	0.31	0.19	41.2

*(Kosova *et al*., 2012)

Data are shown as the estimated liner regression statistic β, SE, and *P*-value with adjustments for age, BMI, and ejaculation abstinence. Motility and total motile sperm number were additionally adjusted for time from masturbation to test. The sperm concentration, semen volume, total sperm number, and total motile sperm number were processed using square-root-transformed values. Bold numbers indicate *P*-values of < 0.05.

[a]"Model" indicates the genetic model for the minor allele in Hutterite reported in previous GWAS (Kosova *et al*., 2012).

[b]The β-coefficient and its SE were summarized using an inverse variance-weighted meta-analysis under fixed-effects model [F] or the DerSimonian and Laird method under random-effects model [R].

P_{het}, *P* value for heterogeneity.

Supplementary Table SI. Semen characteristics of fertile men, and young men from the general population in Japan.

	Fertile (n=791)	Young (n=1224)	*P*
Age (years)	31.2 ± 4.8	20.8 ± 1.7	< 0.0001
BMI (kg/m^2)	23.3 ± 3.0	21.6 ± 2.5	< 0.0001
Ejaculation abstinence (hours)	193.8 ± 324.7	78.5 ± 39.1	< 0.0001
Conc. (×10^6/ml)	105.1 ± 83.2	76.3 ± 57.6	< 0.0001
Vol. (ml)	3.1 ± 1.5	2.9 ± 1.4	0.0005
TSN (×10^6)	315.8 ± 293.7	207.5 ± 178.4	< 0.0001
TMSN (×10^6)	181.7 ± 165.0	122.7 ± 100.3	< 0.0001
Motility (%)	59.8 ± 20.8	59.2 ± 15.1	0.525

Data are represented as mean ± standard deviation. *P* values were obtained with Student's unpaired *t*-test.
The sperm concentration, semen volume, total sperm number, and total motile sperm number were processed using square-root-transformed values.
Conc., sperm concentration; Vol., semen volume; TSN, total sperm numbers; TMSN, total motile sperm numbers.

Supplementary Table SII. An association analysis under additive model between four SNPs and semen parameters in fertile men, and young men from the general population in Japan

	Semen	Fertile		Young		Combined		Heterogeneity	
SNP	Parameter	β (SE)	*P*	β (SE)	*P*	β (SE) [model][a]	P_{meta}	P_{het}	I^2(%)
rs7867029	Conc.	0.14 (0.23)	0.55	-0.064 (0.16)	0.68	-0.00064 (0.13) [F]	1.00	0.47	0.0
	Vol.	**0.065 (0.027)**	**0.017**	0.016 (0.021)	0.43	0.037 (0.024) [R]	0.12	0.15	51.5
	TSN	0.65 (0.45)	0.15	-0.014 (0.27)	0.96	0.16 (0.23) [F]	0.49	0.21	36.9
	TMSN	0.68 (0.36)	0.057	-0.12 (0.22)	0.059	0.23 (0.40) [R]	0.56	0.057	72.3
	Motility (%)	2.27 (1.34)	0.089	**-1.51 (0.77)**	**0.049**	0.22 (1.88) [R]	0.91	0.014	83.4
rs12870438	Conc.	0.59 (0.31)	0.060	0.040 (0.20)	0.84	0.26 (0.27) [R]	0.33	0.14	54.2
	Vol.	0.050 (0.037)	0.17	-0.0019 (0.026)	0.94	-0.018 (0.021) [F]	0.39	0.29	12.0
	TSN	0.18 (0.62)	0.76	0.045 (0.35)	0.90	0.079 (0.30) [F]	0.79	0.84	0.0
	TMSN	0.036 (0.48)	0.94	-0.051 (0.29)	0.86	-0.028 (0.25) [F]	0.91	0.88	0.0
	Motility (%)	0.59 (1.81)	0.75	-0.28 (0.99)	0.77	-0.085 (0.87) [F]	0.92	0.67	0.0
rs7174015	Conc.	-0.11 (0.18)	0.56	-0.0030 (0.12)	0.98	-0.035 (0.10) [F]	0.73	0.64	0.0
	Vol.	0.032 (0.022)	0.13	0.0037 (0.016)	0.82	0.014 (0.013) [F]	0.28	0.29	11.6
	TSN	0.12 (0.36)	0.75	0.020 (0.21)	0.93	0.045 (0.18) [F]	0.79	0.84	0.0
	TMSN	0.087 (0.28)	0.76	0.037 (0.18)	0.83	0.051 (0.15) [F]	0.73	0.88	0.0
	Motility (%)	0.79 (1.06)	0.46	0.040 (0.61)	0.95	0.22 (0.53) [F]	0.92	0.54	0.0
rs724078	Conc.	0.11 (0.20)	0.59	-0.089 (0.14)	0.52	-0.027 (0.11) [F]	0.81	0.42	0.0
	Vol.	-0.013 (0.024)	0.60	0.023 (0.018)	0.20	0.010 (0.014) [F]	0.47	0.23	29.7
	TSN	0.15 (0.40)	0.71	-0.020 (0.24)	0.93	0.025 (0.20) [F]	0.90	0.72	0.0
	TMSN	-0.068 (0.32)	0.83	0.078 (0.19)	0.69	0.037 (0.17) [F]	0.82	0.69	0.0
	Motility (%)	-0.031 (1.18)	0.98	0.59 (0.67)	0.38	0.44 (0.58) [F]	0.45	0.65	0.0

Data are shown as the estimated liner regression statistic β, standard error (SE), and *P*-value with adjustments for age, BMI, and ejaculation abstinence. Motility and total motile sperm number were additionally adjusted for time from masturbation to test. The sperm concentration, semen volume, total sperm number, and total motile sperm number were processed using square-root-transformed values. Bold numbers indicate *P*-values of < 0.05.

Conc., sperm concentration; Vol., semen volume; TSN, total sperm numbers; TMSN, total motile sperm numbers; P_{het}, *P* value for heterogeneity.

[a]The β-coefficient and its SE were summarized using an inverse variance-weighted meta-analysis under fixed-effects model [F] or the DerSimonian and Laird method under random-effects model [R].

Supplementary Table SIII. An association analysis under recessive model between four SNPs and semen parameters in fertile men, and young men from the general population in Japan

	Semen	Fertile		Young		Combined		Heterogeneity	
SNP	Parameter	β (SE)	*P*	β (SE)	*P*	β (SE) [model][a]	P_{meta}	P_{het}	I^2(%)
rs7867029	Conc.	-0.40 (0.71)	0.57	0.098 (0.47)	0.84	-0.057 (0.39) [F]	0.89	0.56	0.0
	Vol.	0.069 (0.083)	0.41	0.083 (0.062)	0.18	0.078 (0.049) [F]	0.12	0.89	0.0
	TSN	-0.55 (1.39)	0.69	0.37 (0.82)	0.65	0.13 (0.71) [F]	0.85	0.57	0.0
	TMSN	-0.85 (1.09)	0.43	0.33 (0.67)	0.63	0.0019 (0.57) [F]	1.00	0.36	0.0
	Motility (%)	-3.21 (4.07)	0.43	-0.31 (2.31)	0.89	-1.01 (2.01) [F]	0.61	0.54	0.0
rs12870438	Conc.	3.55 (1.81)	0.051	-0.049 (0.82)	0.95	1.39 (1.76) [R]	0.43	0.070	69.4
	Vol.	0.20 (0.21)	0.35	0.073 (0.11)	0.50	0.098 (0.095) [F]	0.30	0.59	0.0
	TSN	**7.80 (3.54)**	**0.028**	0.82 (1.42)	0.56	3.55 (3.41) [R]	0.30	0.067	70.2
	TMSN	2.59 (2.78)	0.35	0.73 (1.16)	0.53	1.01 (1.07) [F]	0.35	0.54	0.0
	Motility (%)	-15.96 (10.40)	0.13	2.42 (3.98)	0.54	-4.26 (8.84) [R]	0.63	0.099	63.3
rs7174015	Conc.	-0.029 (0.29)	0.92	-0.092 (0.19)	0.63	-0.073 (0.16) [F]	0.65	0.85	0.0
	Vol.	0.013 (0.034)	0.71	0.0090 (0.025)	0.72	0.010 (0.020) [F]	0.61	0.93	0.0
	TSN	-0.082 (0.56)	0.88	-0.051 (0.33)	0.88	-0.059 (0.28) [F]	0.84	0.96	0.0
	TMSN	0.063 (0.40)	0.87	0.026 (0.27)	0.92	0.037 (0.22) [F]	0.87	0.94	0.0
	Motility (%)	1.07 (1.65)	0.52	0.12 (0.93)	0.90	0.35 (0.81) [F]	0.67	0.61	0.0
rs724078	Conc.	0.29 (0.48)	0.55	**-0.71 (0.32)**	**0.025**	-0.28 (0.50) [R]	0.57	0.084	66.6
	Vol.	-0.053 (0.056)	0.35	0.060 (0.042)	0.15	0.010 (0.056) [R]	0.85	0.11	61.6
	TSN	0.24 (0.94)	0.80	-0.62 (0.55)	0.26	-0.40 (0.48) [F]	0.41	0.43	0.0
	TMSN	0.45 (0.75)	0.54	-0.42 (0.45)	0.35	-0.19 (0.39) [F]	0.63	0.32	0.9
	Motility (%)	4.55 (2.79)	0.10	0.38 (1.56)	0.81	1.37 (1.36) [F]	0.31	0.19	41.2

Data are shown as the estimated liner regression statistic β, standard error (SE), and *P*-value with adjustments for age, BMI, and ejaculation abstinence. Motility and total motile sperm number were additionally adjusted for time from masturbation to test. The sperm concentration, semen volume, total sperm number, and total motile sperm number were processed using square-root-transformed values. Bold numbers indicate *P*-values of < 0.05.

Conc., sperm concentration; Vol., semen volume; TSN, total sperm numbers; TMSN, total motile sperm numbers; P_{het}, *P* value for heterogeneity.

[a]The β-coefficient and its SE were summarized using an inverse variance-weighted meta-analysis under fixed-effects model [F] or the DerSimonian and Laird method under random-effects model [R].

Supplementary Table SIV. An association analysis under dominant model between four SNPs and semen parameters in fertile men, and young men from the general population in Japan

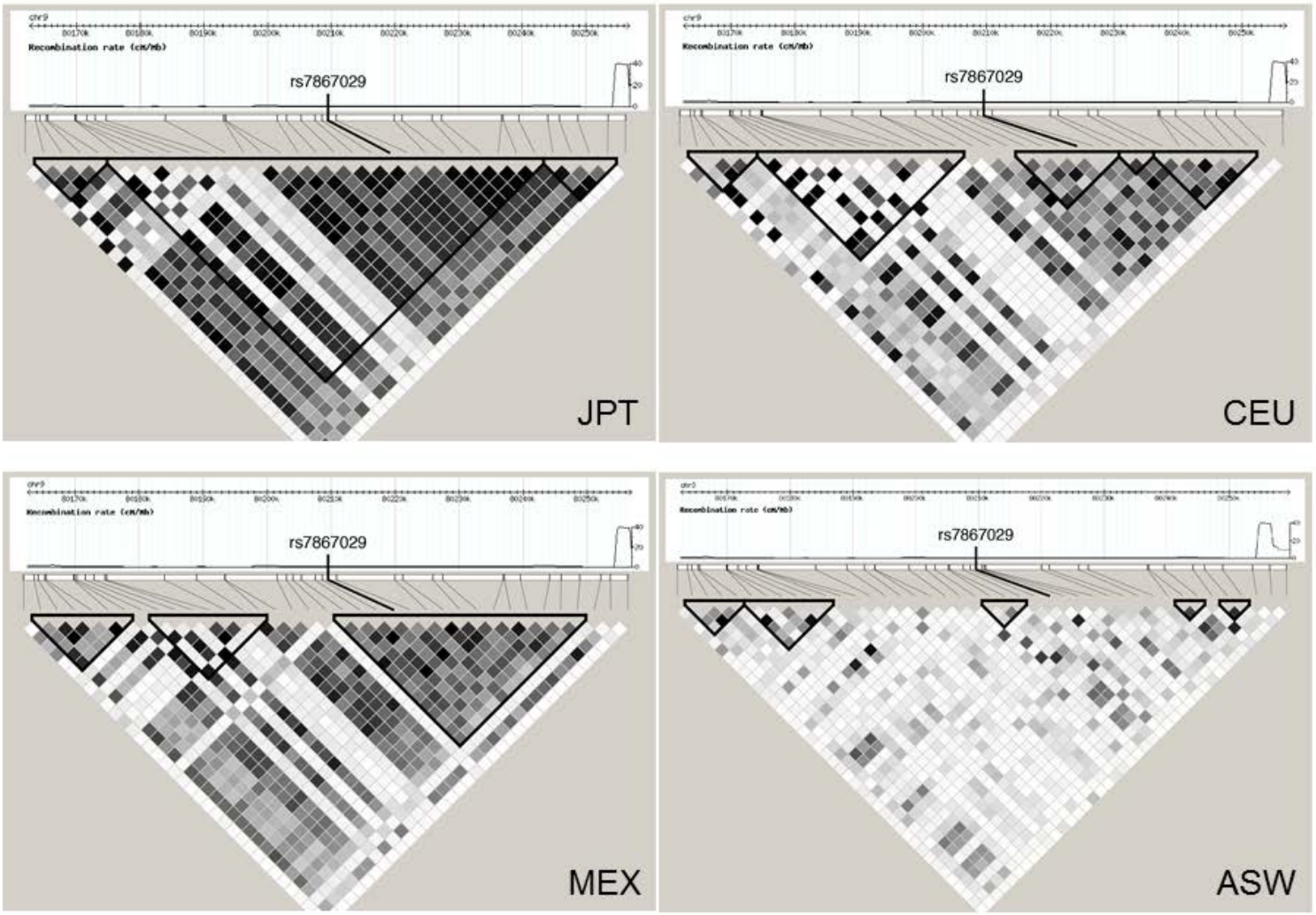
	Semen	Fertile		Young		Combined		Heterogeneity	
SNP	Parameter	β (SE)	*P*	β (SE)	*P*	β (SE) [model][a]	P_{meta}	P_{het}	I^2(%)
rs7867029	Conc.	0.24 (0.27)	0.37	-0.10 (0.18)	0.58	0.0075 (0.15) [F]	0.96	0.29	10.8
	Vol.	**0.077 (0.031)**	**0.014**	0.0095 (0.024)	0.69	0.040 (0.034) [R]	0.23	0.86	66.0
	TSN	0.95 (0.53)	0.070	-0.074 (0.32)	0.81	0.35 (0.51) [R]	0.49	0.094	64.4
	TMSN	**1.033 (0.41)**	**0.012**	-0.21 (0.26)	0.42	0.37 (0.62) [R]	0.55	0.011	84.7
	Motility (%)	4.4 (2.60)	0.091	**-1.98 (0.89)**	**0.026**	0.74 (3.16) [R]	0.81	0.020	81.5
rs12870438	Conc.	0.53 (0.33)	0.11	0.051 (0.22)	0.82	0.20 (0.18) [F]	0.28	0.23	31.3
	Vol.	-0.061 (0.038)	0.11	-0.0076 (0.029)	0.79	-0.027 (0.023) [F]	0.28	0.27	19.3
	TSN	-0.054 (0.64)	0.93	-0.0055 (0.38)	0.99	-0.018 (0.33) [F]	0.96	0.95	0.0
	TMSN	-0.045 (0.50)	0.93	-0.11 (0.31)	0.72	-0.095 (0.27) [F]	0.72	0.91	0.0
	Motility (%)	1.16 (1.88)	0.54	-0.52 (1.08)	0.63	-0.10 (0.94) [F]	0.91	0.44	0.0
rs7174015	Conc.	-0.28 (0.32)	0.38	0.11 (0.22)	0.61	-0.014 (0.18) [F]	0.94	0.31	2.1
	Vol.	**0.079 (0.037)**	**0.031**	-0.0002 (0.028)	0.99	0.036 (0.040) [R]	0.36	0.085	66.3
	TSN	0.44 (0.62)	0.48	0.13 (0.37)	0.73	0.21 (0.32) [F]	0.51	0.66	0.0
	TMSN	0.41 (0.48)	0.40	0.079 (0.31)	0.80	0.17 (0.26) [F]	0.51	0.57	0.0
	Motility (%)	1.00 (1.82)	0.58	-0.032 (1.05)	0.98	0.23 (0.91) [F]	0.80	0.62	0.0
rs724078	Conc.	0.093 (0.26)	0.72	0.069 (0.17)	0.69	0.076 (0.14) [F]	0.60	0.94	0.0
	Vol.	-0.049 (0.030)	0.87	0.019 (0.023)	0.40	0.011 (0.018) [F]	0.56	0.52	0.0
	TSN	0.17 (0.50)	0.74	0.15 (0.30)	0.61	0.16 (0.26) [F]	0.55	0.98	0.0
	TMSN	-0.24 (0.40)	0.55	0.25 (0.25)	0.31	0.12 (0.21) [F]	0.58	0.30	8.5
	Motility (%)	-1.33 (1.48)	0.37	0.84 (0.85)	0.32	0.31 (0.74) [F]	0.68	0.20	38.3

Data are shown as the estimated liner regression statistic β, standard error (SE), and *P*-value with adjustments for age, BMI, and ejaculation abstinence. Motility and total motile sperm number were additionally adjusted for time from masturbation to test. The sperm concentration, semen volume, total sperm number, and total motile sperm number were processed using square-root-transformed values. Bold numbers indicate *P*-values of < 0.05.

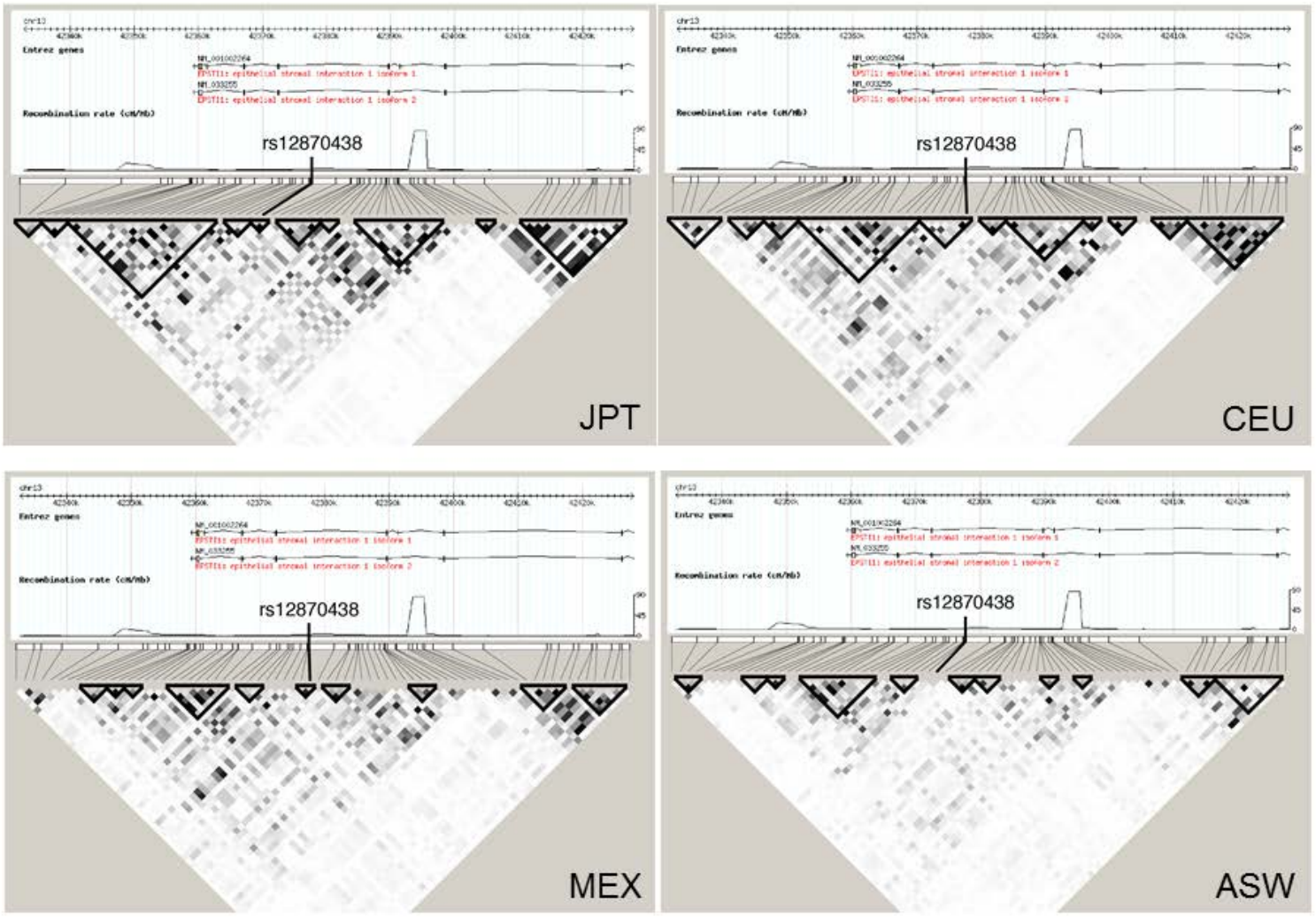
Conc., sperm concentration; Vol., semen volume; TSN, total sperm numbers; TMSN, total motile sperm numbers; P_{het}, *P* value for heterogeneity.

[a]The β-coefficient and its SE were summarized using an inverse variance-weighted meta-analysis under fixed-effects model [F] or the DerSimonian and Laird method under random-effects model [R].

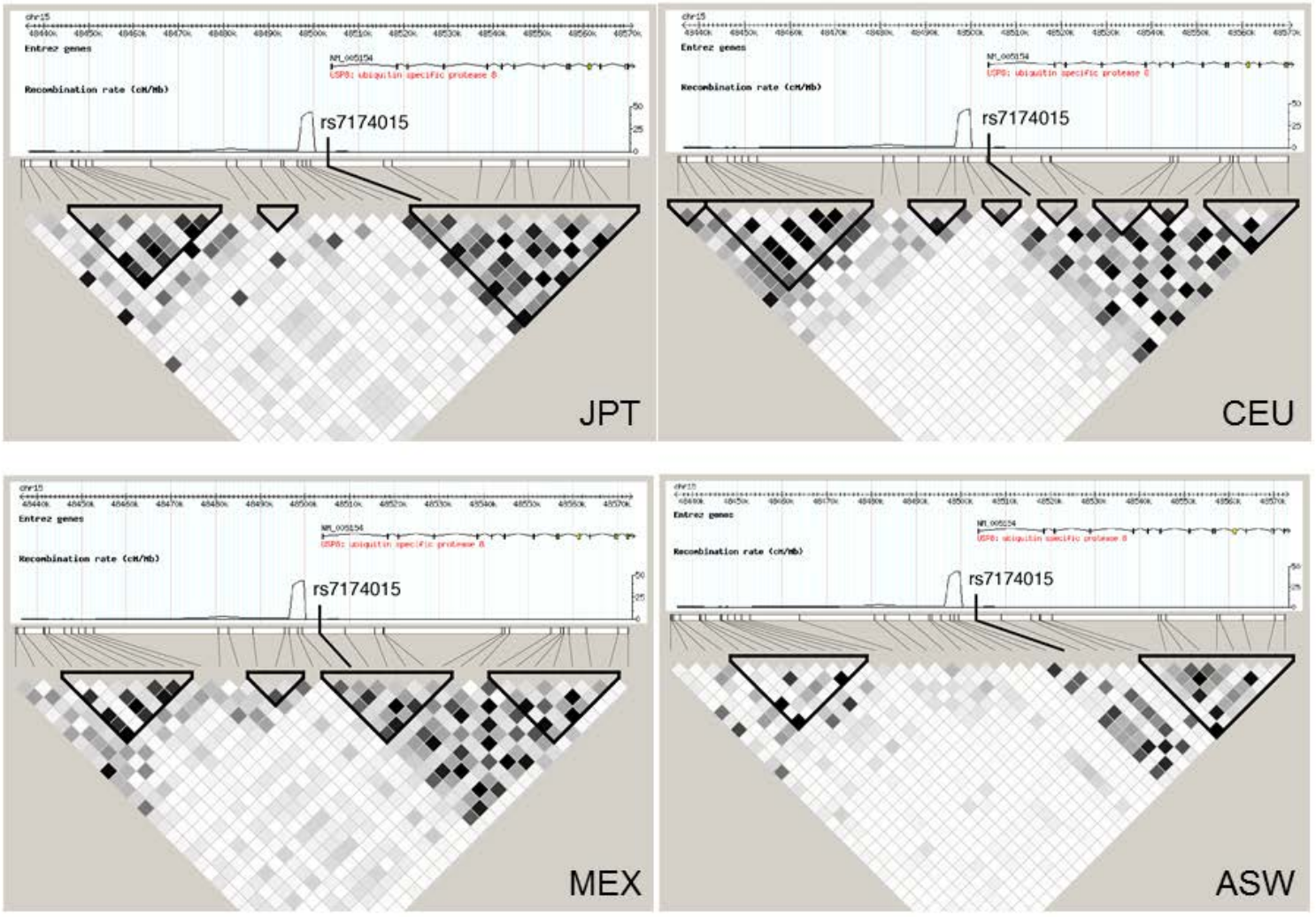
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4