Journal of Assisted Reproduction and Genetics (2020) 37:2839–2852
https://doi.org/10.1007/s10815-020-01920-5

GENETICS

# Association study between single-nucleotide variants rs12097821, rs2477686, and rs10842262 and idiopathic male infertility risk in Serbian population with meta-analysis

Vucic Nemanja[1] · Dobrijevic Zorana[1] · Kotarac Nevena[1] · Matijasevic Suzana[1] · Vukovic Ivan[2] · Budimirovic Branko[3] · Djordjevic Mirka[3] · Savic-Pavicevic Dusanka[1] · Brajuskovic Goran[1]

Received: 13 May 2020 / Accepted: 6 August 2020 / Published online: 19 August 2020


## Abstract

**Purpose** A genome-wide association study conducted in the Han Chinese population identified three single nucleotide variants rs12097821, rs2477686, and rs10842262 as being significantly associated with non-obstructive azoospermia. Our aim was to evaluate the possible association between these susceptibility loci and idiopathic male infertility risk in the Serbian population.

**Methods** A case-control study was conducted on 431 male individuals from the Serbian population divided into two groups. The case group consisted of 208 males diagnosed with oligoasthenozoospermia or non-obstructive azoospermia, while the control group involved 223 fertile men who have fathered at least one child.

**Results** According to codominant ($P_{codom} = 0.048$, $OR_{codom} = 0.57$, 95%CI 0.35–0.92) and overdominant ($P_{overdom} = 0.017$, $OR_{overdom} = 0.62$, 95%CI 0.42–0.92) genetic models, rs10842262 was found to be associated with male infertility. Stratifying infertile men according to diagnosis yielded statistically significant results for non-obstructive azoospermia cases under multiple genetic models ($P_{codom} = 0.038$, $OR_{codom} = 0.47$, 95%CI 0.26–0.85; $P_{dom} = 0.031$, $OR_{dom} = 0.53$, 95%CI 0.30–0.94; $P_{overdom} = 0.016$, $OR_{overdom} = 0.55$, 95%CI 0.33–0.90). Minor allele C of rs2477686 genetic variant was shown to be associated with the reduced risk of oligoasthenozoospermia under the log-additive genetic model ($P = 0.03$, OR = 0.69, 95%CI 0.50–0.97). The results of the meta-analysis indicate both rs2477686 and rs10842262 to be associated with male infertility.

**Conclusion** Our results show variants rs2477686 and rs10842262 to be significantly associated with male infertility in the Serbian population. Nevertheless, case-control studies in other populations are needed to validate their association with infertility in males diagnosed with oligoasthenozoospermia and non-obstructive azoospermia.

**Keywords** Male infertility · Genetic association · rs10842262 · rs2477686 · rs12097821 · Meta-analysis

## Introduction

Infertility, the major issue in reproductive medicine worldwide, is defined as an inability to achieve pregnancy after 12 to 24 months of regular unprotected sexual intercourse [1]. The alarming estimation is that one out of seven couples suffers from disorders related to reproduction that are causing infertility. Solely or as a significant contributor, a male factor is responsible for approximately half of the cases of couples' childlessness [2]. Despite the application of modern diagnostic and prognostic workup, the etiology of male infertility remains unidentified in more than half of cases and therefore referred to as idiopathic infertility [3]. Identification of numerous genes involved in human spermatogenesis and testicular

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s10815-020-01920-5) contains supplementary material, which is available to authorized users.

✉ Brajuskovic Goran
brajuskovic@bio.bg.ac.rs

[1] Centre for Human Molecular Genetics, Faculty of Biology, University of Belgrade, Studentski trg 16, PO Box 43, Belgrade 11 158, Serbia

[2] Clinic of Urology, Clinical Center of Serbia, Belgrade, Serbia

[3] "Academian Vojin Sulovic" Centre for In Vitro Fertilisation, General Hospital Valjevo, Valjevo, Serbia

development has led to focusing research efforts on investigating the potential genetic origin of male infertility. Nevertheless, only a modest proportion of these genes have been identified as a causative factor in infertility pathogenesis. Among the most prominent genetic defects associated with male infertility are microdeletions within the azoospermia factor (AZF) locus of the Y chromosome and Klinefelter syndrome. Still, these known genetic causes account for a small portion of infertile cases, whereas most of the cases are defined as idiopathic and most likely influenced by the interaction between genetic and environmental factors. Therefore, investigators' attention in the field of the genetic basis of male infertility was focused on identifying risk-associated genetic variants through the candidate gene and genome-wide association studies (GWAS) [4].

A comprehensive GWAS conducted in the Han Chinese population has identified three genetic susceptibility loci for non-obstructive azoospermia (NOA). Within these loci, three single-nucleotide variants (SNV) rs12097821 at 1p13.3, rs2477686 at 1p36.32, and rs10842262 at 12p12.1 exhibited a significant association with NOA [5].

Genetic variant rs12097821 is located more than 200 kb upstream of the *PRMT6* gene, encoding a type I arginine methyltransferase that predominantly resides in the nucleus and is highly expressed in human testes [6]. It is shown that *PRMT6* could be downregulated by androgen receptor (AR) and could potentially play an important role in male reproduction by influencing cell migration and apoptosis of germ cells in mice [7]. Ortholog of human *PRMT5* gene, which is related to *PRMT6*, is involved in germ cell development, both in *Drosophila* and the mouse [8, 9].

The second genetic variant, rs2477686, identified as NOA susceptibility locus in Chinese GWAS, resides 47 kb downstream of peroxisome biogenesis factor 10 (*PEX10*) gene. The potential involvement of *PEX10* in the molecular basis of male infertility is supported by experimental results suggesting that functional disruption of *PEX10* ortholog gene in *Drosophila* causes spermatocyte cytokinesis failure, resulting in spermatogenesis defects [10].

Genetic variant rs10842262 is located within an intron of SRY-related HMG-box gene 5 (*SOX5*). *SOX5* encodes a transcription regulator containing a DNA-binding motif known as the high-mobility group (HMG) box, homologous to HMG box of the sex-determining region (SRY) protein [11]. This transcription regulator is highly expressed in the nucleus of round spermatids where it regulates gene expression during spermatogenesis [12]. Furthermore, other SOX family proteins have been reported to have a major role in germ cell development and regulation of adult Sertoli cells in mice and rats [13, 14].

After the initial GWAS, another case-control study was conducted in the Han Chinese population by Zhou et al. and they did not find the association of rs12097821 and rs2477686 with NOA. However, the same group showed a statistically significant association for rs10842262 in the *SOX5* gene [15]. Another study conducted in the Han Chinese population reported no statistically significant association with NOA risk when analyzing *PRMT6* rs12097821, *SOX5* rs10842262, and *PEX10* rs2477686 [16]. Also, Liu SY et al. showed that rs2477686 was significantly associated with oligoasthenozoospermia and progressive motility reduction of the sperm in Han Chinese infertile men [17]. In contrast to these results, a follow-up study of these SNVs in the Japanese population indicated that none of them was significantly associated with NOA [18].

Collectively, these association studies have improved our understanding of the molecular basis of male infertility but their inconsistent results require further validation, especially in non-Asian populations in which the association between these genetic variants and male infertility was not assessed. Therefore, we carried out a case-control study in the Serbian population followed with meta-analysis to assess whether the three genetic variants are associated with the male infertility risk. The present study is the first analysis of the association between rs12097821, rs2477686, and rs10842262 with male infertility conducted in a European population. Furthermore, the meta-analysis was included in order to assess the combined results of available case-control studies on this issue considering various potential genetic models of association, as well as the possible differences based on ethnicity and diagnosis.

## Materials and methods

### Subjects

Our study group consisted of 208 infertile male subjects and 223 male controls from the Serbian population. Infertile men were treated at Clinical Centre Serbia, Belgrade, Serbia, and General Hospital Valjevo, Serbia, in the period from 2016 to 2018. Men within the infertile group were diagnosed with oligoasthenozoospermia ($n = 114$) and non-obstructive azoospermia (NOA) ($n = 94$). Diagnoses were established based on a minimum of two semen analyses according to guidelines given by the WHO [19]. The patient's serum hormone levels and semen parameters are shown in Table 1. Infertile men who had a history of any of the following conditions, orchitis, obstruction of *vas deferens*, chromosomal abnormalities, or microdeletions of the AZF region on the Y chromosome, were excluded as potential subjects for this study. The control group consisted of fertile men chosen from the collections of DNA samples from the Center for Human Molecular Genetics. The main inclusion criterion for controls was that they fathered at least one child. Written informed consent was obtained from all participants. The study was approved by the

**Table 1** Patient's serum hormone levels and semen parameters

| Clinical characteristics (mean) | Oligoasthenozoospermia ($n = 114$) | Azoospermia ($n = 94$) |
|---|---|---|
| Serum hormones | | |
| FSH (mlU/ml) | 15.14 | 17.36 |
| LH (mlU/ml) | 7.62 | 7.53 |
| PRL (pmol/L) | 342.11 | 247.24 |
| T (nmol/L) | 9 | 8.05 |
| Inhibin B (pg/ml) | 82.11 | 78.86 |
| Sperm parameters | | |
| Volume (ml) | 4.09 | 2.94 |
| Total sperm number (× 106) | 6.66 | / |
| Sperm density(× 106/ml) | 2.33 | / |
| Motility (%) | 11 | / |

*FSH* follicle-stimulating hormone, *LH* luteinizing hormone, *PRL* prolactin, *T* testosterone

Ethics Committees of Clinical Center of Serbia, Belgrade, Serbia (Number 68/5), and General Hospital Valjevo, Serbia (Number OBV-02-1170), while all the experiments were conducted in accordance with the Helsinki Declaration of 1975.

## Genotype determination

The commercial kits for DNA isolation (Qiagen, Germany and Zymoresearch, USA) were used for the extraction of genomic DNA from peripheral blood samples following the manufacturer's protocol. Quality of genomic DNA was analyzed by electrophoresis on 0.8% agarose gel, and DNA was quantified using Qubit dsDNA BR Assay Kit on Qubit® 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA).

Genotyping of three single-nucleotide variants was performed by PCR restriction fragment length polymorphism (PCR-RFLP) method. Primer sequences for all the genetic variants tested were designed using Primer3 software [20]. Sequences of oligonucleotide primers used for PCR amplification are listed in Online Resource 1.

The 15 μl PCR reaction mixture contained 1.5 μl of genomic DNA per sample (concentration ranging between 1 and 20 ng/μl), 0.2 μM of forward and reverse PCR primer each, 200 μM of deoxyribonucleoside triphosphates (dNTPs, Fermentas, Hanover, MD, USA), 1.5 μl 10× PCR buffer A (with 15 mM $MgCl_2$) (Kapa Biosystems, Woburn, MA, USA), 0.04 U/μl of Taq DNA polymerase (Kapa Biosystems), and nuclease-free water (Serva, Westbury, NY, USA). The PCR protocol consisted of denaturation at 97 °C for 3 min followed by 35 cycles of 95 °C for 60 s, 60 °C for 60 s, and 72 °C for 60 s and the final extension 10 min at 72 °C. Electrophoresis on a 1.5% agarose gel stained with ethidium bromide was performed using 5 μl of the amplified reaction mixture for each sample.

Products of PCR reaction were subjected to overnight digestion at the optimum temperature corresponding to specific restriction enzymes that were used (Table 2). Digested products were separated by electrophoresis on 3% agarose gel.

Regenotyping of rs10842262 was performed with bidirectional Sanger sequencing using the BigDye® Terminator v3.1 Cycle Sequencing kit (Life Technologies, USA) followed by capillary gel electrophoresis on Applied Biosystems™ 3100 Genetic Analyzer (Waltham, MA, USA). Sequencing data were analyzed with BioEdit Sequence Alignment Editor (Fig. 1) [21].

## Association analysis

The exact test was applied for the assessment of deviation from Hardy-Weinberg equilibrium (HWE). Allele and genotype frequencies were calculated with the exact test implemented in the SNPstats software package [22].

Given the binary response variable, logistic regression was used to calculate 95% confidence intervals (CI) and the odds ratios (OR) estimation and a $p$ value $< 0.05$ was set as a significance level. Association between alleles and genotypes of the selected genetic variants and infertility risk was examined under the following genetic models: dominant, recessive, codominant, overdominant, and log-additive. Association analyses were performed using R package SNPassoc (version 1.9-2) and Akaike information criterion (AIC) was used to determine the best-fitting genetic model [23, 24].

## Meta-analysis

### Publication search

The literature included in this meta-analysis was selected from PubMed database using the search strategy based on combinations of keywords: "infertility" or "azoospermia" or "asthenozoospermia" or "oligoasthenozoospermia" or "oligospermia," gene names, protein names and aliases, or

**Table 2** Details for the PCR-RFLP method used for genotyping of the selected SNVs

| Genetic variant | Restriction enzyme | Digestion temperature | Minor/major alleles | Digestion fragments length |
|---|---|---|---|---|
| rs12097821 | *AluI* | $60^0$ C | G/T | 100 bp, 78 bp, 39 bp |
| rs2477686 | *FspBI* | $37^0$ C | G/C | 172 bp, 74 bp |
| rs10842262 | *FspBI* | $37^0$ C | C/G | 137 bp, 88 bp |

genetic variant ID number ("*PRMT6*" or "rs12097821," "*PEX10*" or "rs2477686," "*SOX5*" or "rs10842262") with "association" or "polymorphism" or "SNP" or "variant" or "variants" without language restriction. References cited in retrieved original studies, as well as in review articles and previous meta-analyses, were examined for additional studies suitable for inclusion in this meta-analysis. Selected articles were published before October 2018.

### Inclusion and exclusion criteria and data extraction

Eligible studies met the following criteria: (a) analysis of association between a genetic variant and idiopathic male infertility; (b) case-control study design; (c) provided sufficient data about genotype and allele counts to calculate risk estimates (OR with 95% CIs and *P* values); (d) provided detailed information about the recruitment of participants, diagnostic protocols, genotyping, statistical analysis, and other relevant methodological data. Studies that included several case-control groups were segregated into multiple study panels which were independently included in the meta-analysis.

The data extracted from selected studies included the following: first author's last name, year of publication, the country from which participants were recruited, ethnicity, source of controls, methods, sample size, genotype, and allele counts. The main exclusion criterion was the presence of major errors in study design or interpretation and presentation of results. Data extraction was performed independently by two authors. In case of any disagreement, a consensus was reached by consultation between all authors. Genetic variants analyzed in three or more eligible studies were considered relevant for the meta-analysis. Also, we decided to perform the meta-analysis with and without the exclusion of the results from the initial GWAS, since it was predicted to have the large overall influence on the results of meta-analysis, due to the large sample size and the associated study power.

### Statistical analysis

Statistical software OpenMeta-analyst (The Center for Evidence-based Medicine, Brown University, Providence, RI, USA) was used for performing heterogeneity tests and meta-analyses [25]. Estimates of ORs and their 95% CIs were calculated by using the fixed-effect or the random-effects model, depending on the results of heterogeneity tests. For assessing heterogeneity of results across studies, Cochran's *Q* statistic test was used, combined with inconsistency index ($I^2$). Heterogeneity was considered significant at $P < 0.1$, while $I^2 = 0$–$25\%$ suggested no heterogeneity, $I^2 = 25$–$50\%$ moderate heterogeneity, $I^2 = 50$–$75\%$ large heterogeneity, and $I^2 = 75$–$100\%$ extreme heterogeneity. Random-effects model was selected for meta-analysis when heterogeneity tests yielded significant results ($P < 0.1$ or $I^2 \geq 50\%$). For the fixed-effects model, the Mantel-Haenszel method of weighting was used, while for pooling results under the random-effects model, the method proposed by DerSimonian and Laird was applied [26, 27]. Estimates of ORs and its 95% CIs were calculated for each SNP using a fixed-effects or random-effects model based on the results of heterogeneity tests. Separate meta-analyses were performed for the allelic, dominant, recessive, and overdominant genetic models. Selected studies were classified according to diagnosis and ethnicity of participants and meta-analysis was performed if two or more studies corresponded to a classification subgroup. Publication bias was assessed by visual inspection of Funnel plots and by using Egger's test. A *P* value of Egger's test less than 0.05 was considered representative of statistically significant publication bias.

## Results

Genotype distributions of rs12097821, rs10842262, and rs2477686 were compared between groups of 208 infertile men and 223 fertile control subjects. Genotyping was successfully performed in 95.1% (410/431) of study participants for rs12097821 and rs10842262, and in 94.4% (407/431) of subjects for rs2477686.

The distributions of rs12097821 and rs2477686 genotypes in the control group were compatible with HWE, whereas genotype distributions for rs10842262 were not consistent with HWE. Table 3 shows genotype frequencies of three genetic variants in infertile subjects and fertile controls, as well as the results of tests for association with the idiopathic infertility risk.

Although genotype distributions for rs10842262 were not consistent with HWE, comparison of genotype frequencies between infertile men and the control group showed statistical

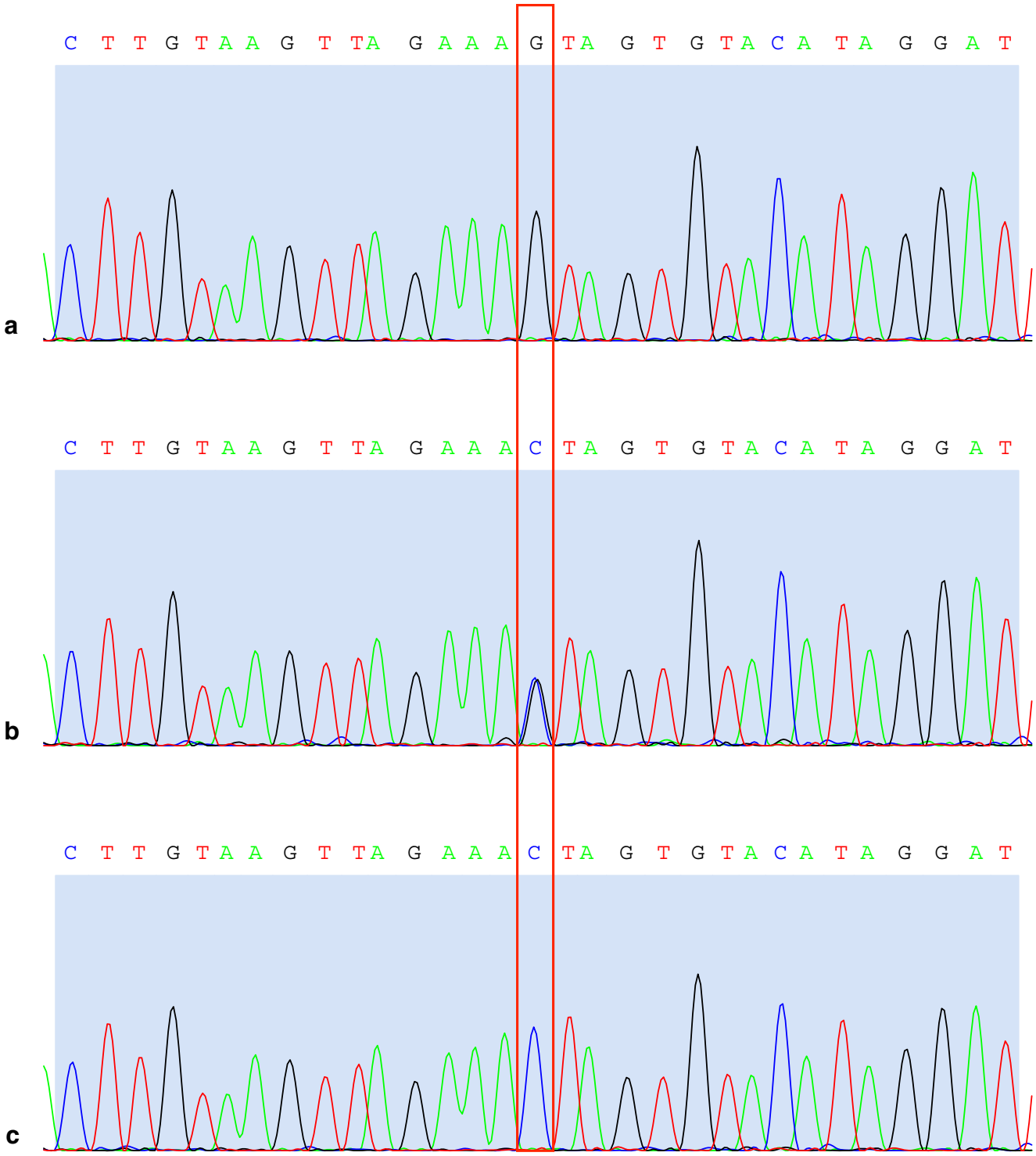


**Fig. 1** Chromatograms of Sanger sequencing showing: rs10842262 minor allele G homozygous genotype (a), rs10842262 G/C heterozygous genotype (b), and rs10842262 major allele C homozygous genotype (c)

significance, with the decreased risk of male infertility associated with heterozygous genotype according to overdominant ($P_{overdom} = 0.017$, $OR_{overdom} = 0.62$, 95%CI 0.42–0.92) and codominant ($P_{codom} = 0.048$, $OR_{codom} = 0.57$, 95%CI 0.35–0.92) genetic models. Stratifying infertile cases according to diagnosis yielded statistically significant results for association between this genetic variant and NOA under multiple genetic models ($P_{codom} = 0.038$, $OR_{codom} = 0.47$, 95%CI 0.26–0.85; $P_{dom} = 0.031$, $OR_{dom} = 0.53$, 95%CI 0.30–0.94; $P_{overdom} = 0.016$, $OR_{overdom} = 0.55$, 95%CI 0.33–0.90)

**Table 3** Association of three genetic variants and male infertility

| Genetic variant | Genetic model | Genotype | Cases | Controls | Logistic regression | | |
|---|---|---|---|---|---|---|---|
| | | | | | OR (CI 95%) | *P* value | AIC |
| rs10842262 | Codominant | C/C | 54 (26.5%) | 38 (18.4%) | 1.00 | *0.048* | 568.3 |
| | | G/C | 102 (50%) | 127 (61.6%) | *0.57 (0.35–0.92)* | | |
| | | G/G | 48 (23.5%) | 41 (19.9%) | *0.82 (0.46–1.48)* | | |
| | Dominant | C/C | 54 (26.5%) | 38 (18.4%) | 1.00 | 0.051 | 568.6 |
| | | G/C-G/G | 150 (73.5%) | 168 (81.5%) | 0.63 (0.39–1.01) | | |
| | Recessive | C/C-G/C | 156 (76.5%) | 165 (80.1%) | 1.00 | 0.37 | 571.6 |
| | | G/G | 48 (23.5%) | 41 (19.9%) | 1.24 (0.77–1.98) | | |
| | Overdominant | C/C-G/G | 102 (50%) | 79 (38.4%) | 1.00 | *0.017* | 566.7 |
| | | G/C | 102 (50%) | 127 (61.6%) | *0.62 (0.42–0.92)* | | |
| | Log-additive | – | – | – | 0.91 (0.68–1.21) | 0.5 | 571.9 |
| *rs2477686* | Codominant | G/G | 66 (32.5%) | 51 (25%) | 1.00 | 0.23 | 567.3 |
| | | C/G | 94 (46.3%) | 102 (50%) | 0.71 (0.45–1.13) | | |
| | | C/C | 43 (21.2%) | 51 (25%) | 0.65 (0.38–1.12) | | |
| | Dominant | G/G | 66 (32.5%) | 51 (25%) | 1.00 | 0.094 | 565.4 |
| | | C/G-C/C | 137 (67.5%) | 153 (75%) | 0.69 (0.45–1.07) | | |
| | Recessive | G/G-C/G | 160 (78.8%) | 153 (75%) | 1.00 | 0.36 | 567.4 |
| | | C/C | 43 (21.2%) | 51 (25%) | 0.81 (0.51–1.28) | | |
| | Overdominant | G/G-C/C | 109 (53.7%) | 102 (50%) | 1.00 | 0.46 | 567.7 |
| | | C/G | 94 (46.3%) | 102 (50%) | 0.86 (0.58–1.27) | | |
| | Log-additive | – | – | – | 0.80 (0.61–1.05) | 0.11 | 565.7 |
| *rs12097821* | Codominant | G/G | 152 (74.2%) | 161 (78.5%) | 1.00 | 0.57 | 573.2 |
| | | T/G | 51 (24.9%) | 42 (20.5%) | 1.29 (0.81–2.05) | | |
| | | T/T | 2 (1%) | 2 (1%) | 1.06 (0.15–7.61) | | |
| | Dominant | G/G | 152 (74.2%) | 161 (78.5%) | 1.00 | 0.3 | 571.3 |
| | | T/G-T/T | 53 (25.9%) | 44 (21.5%) | 1.28 (0.81–2.02) | | |
| | Recessive | G/G-T/G | 203 (99%) | 203 (99%) | 1.00 | 1 | 572.4 |
| | | T/T | 2 (1%) | 2 (1%) | 1.00 (0.14–7.17) | | |
| | Overdominant | G/G-T/T | 154 (75.1%) | 163 (79.5%) | 1.00 | 0.29 | 571.3 |
| | | T/G | 51 (24.9%) | 42 (20.5%) | 1.29 (0.81–2.04) | | |
| | Log-additive | – | – | – | 1.24 (0.81–1.90) | 0.33 | 571.4 |

*OR* odds ratio, *CI* confidence interval, *AIC* Akaike information criteria

(Table 4), while for association with oligoasthenozoospermia results remained insignificant (Table 5).

When considering the possible association between rs2477686 and male infertility risk, the comparison of genotype frequencies between men diagnosed with idiopathic infertility and the control group also showed no statistically significant difference for all genetic models tested. Nevertheless, a statistical trend of significance was reached for association under a dominant genetic model ($P_{dom} = 0.094$) (Table 3). Also, minor allele C of this genetic variant was shown to be associated with the reduced risk of oligoasthenozoospermia under the log-additive genetic model ($P = 0.03$, OR = 0.69, 95%CI 0.50–0.97) (Table 5).

Genotype frequencies of rs12097821 were not found to differ significantly between infertile subjects and the control group. Tests for the association of this SNV with the risk of male infertility yielded *p* values of 0.29 and 0.3 for the overdominant and dominant genetic models, respectively, which were the best-fitting according to AIC score. Also, when infertile subjects were stratified according to diagnosis, this genetic variant was not found to be associated with NOA (Table 4), nor with oligoasthenozoospermia (Table 5).

While conducting meta-analysis, a total of 5 records were retrieved through initial database searching by using multiple combinations of relevant keywords and by removing the duplicate records. One of these studies was the initial GWAS which we decided to exclude in order to eliminate the potential bias. All of the studies analyzed multiple genetic variants, while one of them included several stages of the study, based on which multiple study panels were formed [18]. Also, two panels were formed from the present case-control study results based on patient diagnosis. Even though a study by Liu et al. [17] included cases with different diagnoses, multiple study panels were not formed since the number of men in different groups after stratification was small.

**Table 4** Association of three genetic variants and male infertility in the group of patients diagnosed with non-obstructive azoospermia

| Genetic variant | Genetic model | Genotype | Cases | Controls | Logistic regression | | |
|---|---|---|---|---|---|---|---|
| | | | | | OR (CI 95%) | *P* value | AIC |
| *rs10842262* | Codominant | C/C | 28 (29.8%) | 38 (18.4%) | 1.00 | *0.038* | 372.5 |
| | | G/C | 44 (46.8%) | 127 (61.6%) | *0.47 (0.26–0.85)* | | |
| | | G/G | 22 (23.4%) | 41 (19.9%) | *0.73 (0.36–1.48)* | | |
| | Dominant | C/C | 28 (29.8%) | 38 (18.4%) | 1.00 | *0.031* | 372.4 |
| | | G/C-G/G | 66 (70.2%) | 168 (81.5%) | **0**.*53 (0.30–0.94)* | | |
| | Recessive | C/C-G/C | 72 (76.6%) | 165 (80.1%) | 1.00 | 0.49 | 376.6 |
| | | G/G | 22 (23.4%) | 41 (19.9%) | 1.23 (0.68–2.21) | | |
| | Overdominant | C/C-G/G | 50 (53.2%) | 79 (38.4%) | 1.00 | *0.016* | 371.3 |
| | | G/C | 44 (46.8%) | 127 (61.6%) | *0.55 (0.33–0.90)* | | |
| | Log-additive | – | – | – | 0.83 (0.57–1.21) | 0.34 | 376.1 |
| *rs2477686* | | | | | | | |
| *rs2477686* | Codominant | G/G | 28 (30.4%) | 51 (25%) | 1.00 | 0.45 | 371.3 |
| | | C/G | 39 (42.4%) | 102 (50%) | 0.70 (0.39–1.26) | | |
| | | C/C | 25 (27.2%) | 51 (25%) | 0.89 (0.46–1.74) | | |
| | Dominant | G/G | 28 (30.4%) | 51 (25%) | 1.00 | 0.33 | 369.9 |
| | | C/G-C/C | 64 (69.6%) | 153 (75%) | 0.76 (0.44–1.31) | | |
| | Recessive | G/G-C/G | 67 (72.8%) | 153 (75%) | 1.00 | 0.69 | 370.7 |
| | | C/C | 25 (27.2%) | 51 (25%) | 1.12 (0.64–1.96) | | |
| | Overdominant | G/G-C/C | 53 (57.6%) | 102 (50%) | 1.00 | 0.22 | 369.4 |
| | | C/G | 39 (42.4%) | 102 (50%) | 0.74 (0.45–1.21) | | |
| | Log-additive | – | – | – | 0.94 (0.67–1.32) | 0.72 | 370.8 |
| *rs12097821* | Codominant | G/G | 70 (74.5%) | 161 (78.5%) | 1.00 | 0.74 | 377.7 |
| | | T/G | 23 (24.5%) | 42 (20.5%) | 1.26 (0.70–2.25) | | |
| | | T/T | 1 (1.1%) | 2 (1%) | 1.15 (0.10–12.89) | | |
| | Dominant | G/G | 70 (74.5%) | 161 (78.5%) | 1.00 | 0.44 | 375.7 |
| | | T/G-T/T | 24 (25.5%) | 44 (21.5%) | 1.25 (0.71–2.22) | | |
| | Recessive | G/G-T/G | 93 (98.9%) | 203 (99%) | 1.00 | 0.94 | 376.3 |
| | | T/T | 1 (1.1%) | 2 (1%) | 1.09 (0.10–12.19) | | |
| | Overdominant | G/G-T/T | 71 (75.5%) | 163 (79.5%) | 1.00 | 0.44 | 375.7 |
| | | T/G | 23 (24.5%) | 42 (20.5%) | 1.26 (0.70–2.24) | | |
| | Log-additive | – | – | – | 1.22 (0.72–2.08) | 0.46 | 375.7 |

*OR* odds ratio, *CI* confidence interval, *AIC* Akaike information criteria

Finally, 5 studies (besides Hu et al. [5]), based on which 7 eligible independent study panels, were formed for rs12097821 and rs2477686, and included in the meta-analysis. For rs10842262, the number of study panels included in the meta-analysis was 6, since the number of relevant studies was 4.

All three genetic variants tested (rs12097821, rs2477686, and rs10842262) were found to be associated with idiopathic male infertility in the meta-analysis which included the initial GWAS by Hu et al. [5]. For rs12097821, *P* values for both allelic and recessive genetic models were less than 0.001 ($OR_{allelic} = 1.173$, 95%CI 1.080–1.275; $OR_{rec} = 1.406$, 95%CI 1.224–1.615, results not shown). Nevertheless, when the mentioned GWA study was excluded, the association of rs12097821 with the risk of infertility occurrence did not reach statistical significance for any genetic model tested (Fig. 2). Also, results remained statistically insignificant in the group of infertile men diagnosed with NOA, as well as in the group of men with Asian ancestry (results not shown).

On the other hand, even after the exclusion of the results obtained by Hu et al. [5], minor allele G of rs2477686 was found to increase the risk of infertility occurrence under allelic and recessive genetic models ($P_{allelic} = 0.009$, $OR_{allelic} = 1.169$, 95%CI 1.040–1.315; $P_{rec} = 0.008$, $OR_{rec} = 1.472$, 95%CI 1.109–1.955), while for the association under the dominant genetic model, statistical trend of significance was reached ($P_{dom} = 0.069$) (Fig. 3). Also, a significant result for association under allelic genetic model was obtained for the subgroup of studies that included infertile men diagnosed with oligozoospermia, oligoasthenozoospermia, or asthenozoospermia. Still, assuming the same genetic model, only the statistical trend of significance was reached for the association of

**Table 5** Association of three genetic variants and male infertility in the group of patients diagnosed with non-obstructive oligoasthenozoospermia

| Genetic variant | Genetic model | Genotype | Cases | Controls | Logistic regression | | |
|---|---|---|---|---|---|---|---|
| | | | | | OR (CI 95%) | *P* value | AIC |
| *rs10842262* | Codominant | G/G | 26 (23.6%) | 41 (19.9%) | 1.00 | 0.3 | 412.1 |
| | | G/C | 58 (52.7%) | 127 (61.6%) | 0.72 (0.40–1.29) | | |
| | | C/C | 26 (23.6%) | 38 (18.4%) | 1.08 (0.54–2.17) | | |
| | Dominant | G/G | 26 (23.6%) | 41 (19.9%) | 1.00 | 0.44 | 411.8 |
| | | G/C-C/C | 84 (76.4%) | 165 (80.1%) | 0.80 (0.46–1.40) | | |
| | Recessive | G/G-G/C | 84 (76.4%) | 168 (81.5%) | 1.00 | 0.28 | 411.3 |
| | | C/C | 26 (23.6%) | 38 (18.4%) | 1.37 (0.78–2.40) | | |
| | Overdominant | G/G-C/C | 52 (47.3%) | 79 (38.4%) | 1.00 | 0.13 | 410.1 |
| | | G/C | 58 (52.7%) | 127 (61.6%) | 0.69 (0.43–1.11) | | |
| | Log-additive | – | – | – | 1.04 (0.72–1.48) | 0.85 | 412.4 |
| *rs2477686* | Codominant | G/G | 38 (34.2%) | 51 (25%) | 1.00 | 0.092 | 410 |
| | | C/G | 55 (49.5%) | 102 (50%) | 0.72 (0.42–1.23) | | |
| | | C/C | 18 (16.2%) | 51 (25%) | 0.47 (0.24–0.94) | | |
| | Dominant | G/G | 38 (34.2%) | 51 (25%) | 1.00 | 0.084 | 409.8 |
| | | C/G-C/C | 73 (65.8%) | 153 (75%) | 0.64 (0.39–1.06) | | |
| | Recessive | G/G-C/G | 93 (83.8%) | 153 (75%) | 1.00 | 0.067 | 409.5 |
| | | C/C | 18 (16.2%) | 51 (25%) | 0.58 (0.32–1.05) | | |
| | Overdominant | G/G-C/C | 56 (50.5%) | 102 (50%) | 1.00 | 0.94 | 412.8 |
| | | C/G | 55 (49.5%) | 102 (50%) | 0.98 (0.62–1.56) | | |
| | Log-additive | – | – | – | *0.69 (0.50–0.97)* | *0.03* | 408.1 |
| *rs12097821* | Codominant | G/G | 82 (73.9%) | 161 (78.5%) | 1.00 | 0.63 | 414.8 |
| | | T/G | 28 (25.2%) | 42 (20.5%) | 1.31 (0.76–2.26) | | |
| | | T/T | 1 (0.9%) | 2 (1%) | 0.98 (0.09–10.99) | | |
| | Dominant | G/G | 82 (73.9%) | 161 (78.5%) | 1.00 | 0.35 | 412.8 |
| | | T/G-T/T | 29 (26.1%) | 44 (21.5%) | 1.29 (0.75–2.22) | | |
| | Recessive | G/G-T/G | 110 (99.1%) | 203 (99%) | 1.00 | 0.95 | 413.7 |
| | | T/T | 1 (0.9%) | 2 (1%) | 0.92 (0.08–10.29) | | |
| | Overdominant | G/G-T/T | 83 (74.8%) | 163 (79.5%) | 1.00 | 0.34 | 412.8 |
| | | T/G | 28 (25.2%) | 42 (20.5%) | 1.31 (0.76–2.26) | | |
| | Log-additive | – | – | – | 1.25 (0.75–2.07) | 0.39 | 412.9 |

*OR* odds ratio, *CI* confidence interval, *AIC* Akaike information criteria

rs2477686 with NOA. The same genetic variant was found to be associated with idiopathic male infertility in Asians under the allelic genetic model ($P = 0.049$) (results not shown).

Similarly as for rs2477686, for rs10842262 minor allele G, association with the increased risk of infertility occurrence was found for allelic and recessive genetic models even after the exclusion of the GWAS results ($P_{allelic} = 0.004$, $OR_{allelic} = 1.149$, 95%CI 1.046–1.262; $P_{rec} = 0.004$, $OR_{rec} = 1.314$, 95%CI 1.089–1.587) (Fig. 4). The results remained statistically significant in the subgroup of cases and controls with Asian ancestry, as well as in the NOA subgroup. Also, statistical significance in the Asian subgroup was reached for dominant genetic model (results not shown).

Visual inspection of Funnel plots showed no obvious asymmetry which would suggest the presence of publication bias. Furthermore, the results of Egger's tests showed no evidence of publication bias for any genetic model tested (results not shown).

## Discussion

A recent GWAS conducted in the Han Chinese population has identified three novel SNVs associated with NOA susceptibility [5]. Afterwards, several validation studies have been conducted in Asian populations providing discordant results on the potential association of these genetic variants with male infertility. In order to further elucidate the effect of these variants on idiopathic male infertility, we have performed a case-control study involving the group of infertile men from the Serbian population, diagnosed with NOA and oligoasthenozoospermia. Since previous studies have been conducted in Han Chinese and

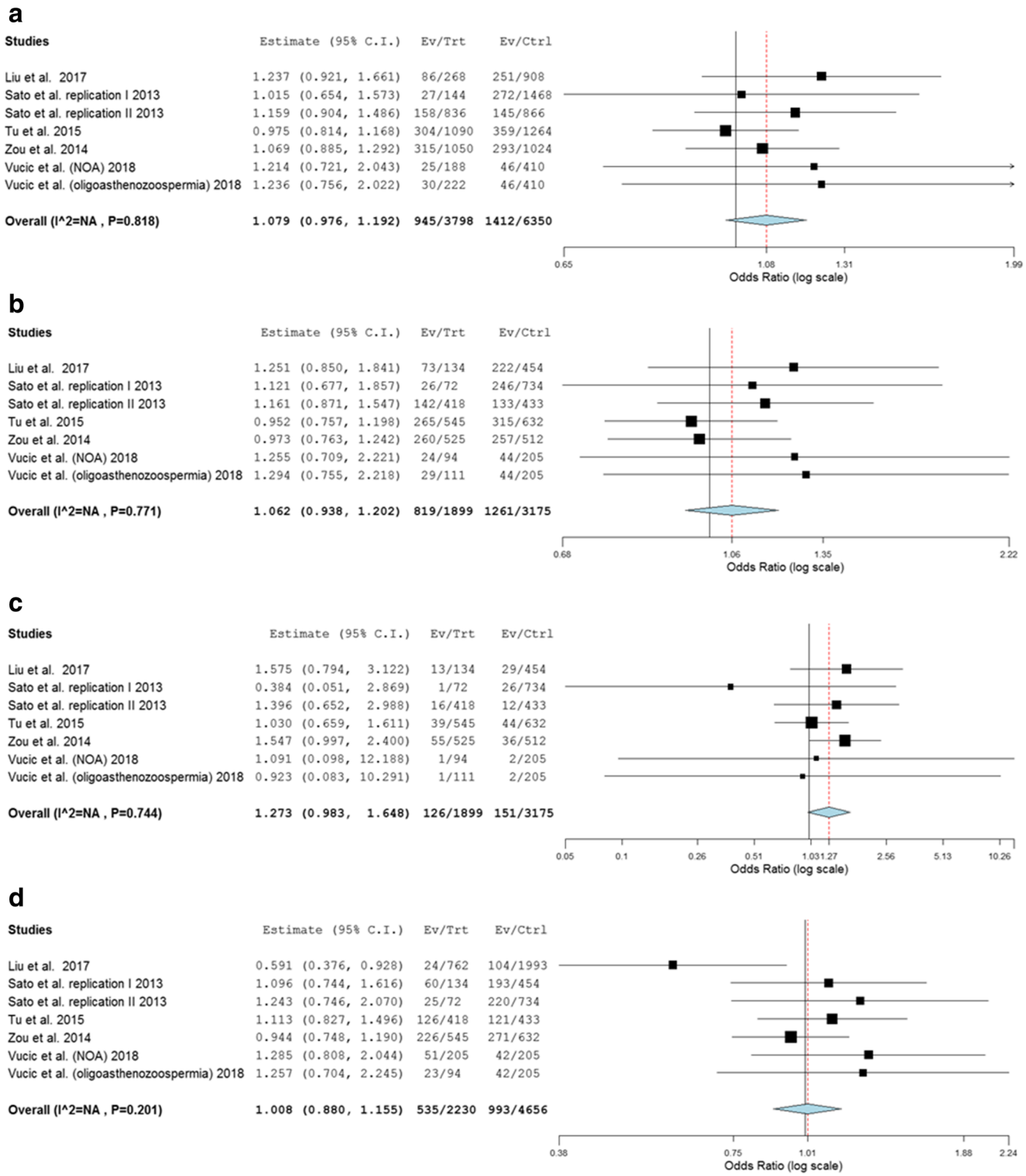


**Fig. 2** Meta-analysis of the association between rs12097821 and male infertility. **a** Allelic model. **b** Dominant model. **c** Recessive model. **d** Overdominant model. The results of different studies presented as ORs, with 95% CI, and the overall effect with 95% CI are shown in the forest plot. The size of the square symbol representing the study's result is proportional to the weight assigned to the study. *P* values presented are derived from heterogeneity tests

Japanese populations [15–18], this is the first case-control study on this issue carried out in a population of European descent.

The lack of evidence suggesting the association of rs12097821 with male infertility in the Serbian population contrasts the results obtained in GWAS conducted

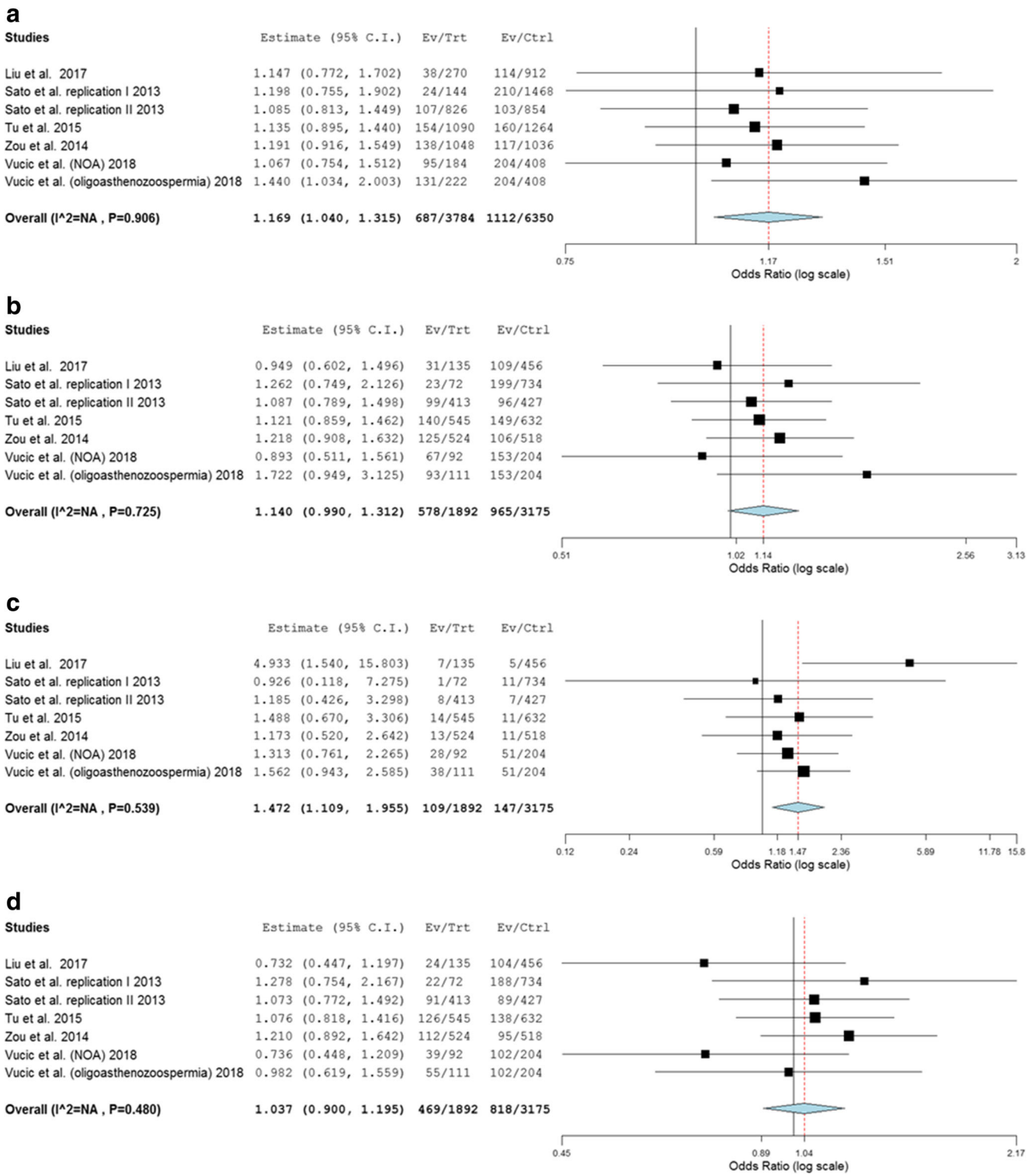


**Fig. 3** Meta-analysis of the association between rs2477686 and male infertility. **a** Allelic model. **b** Dominant model. **c** Recessive model. **d** Overdominant model. The results of different studies presented as ORs, with 95% CI, and the overall effect with 95% CI are shown in the forest plot. The size of the square symbol representing the study's result is proportional to the weight assigned to the study. *P* values presented are derived from heterogeneity tests

by Hu et al. [5]. Nevertheless, our results regarding the effect of this genetic variant on male infertility risk are consistent with the results of all other previous studies on this issue [15–18]. On the other hand, most of these

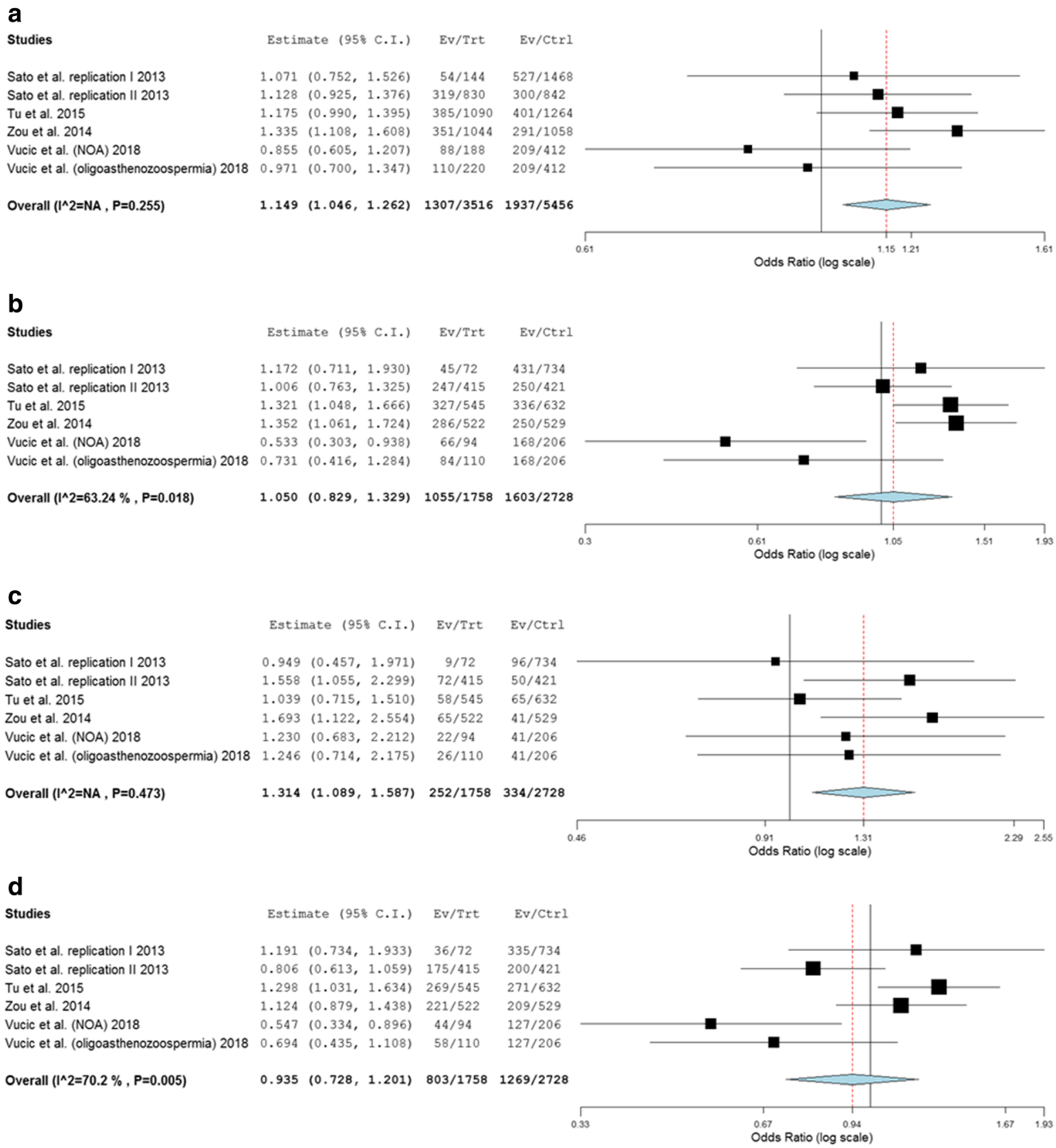


**Fig. 4** Meta-analysis of the association between rs10842262 and male infertility. **a** Allelic model. **b** Dominant model. **c** Recessive model. **d** Overdominant model. The results of different studies presented as ORs, with 95% CI, and the overall effect with 95% CI are shown in the forest plot. The size of the square symbol representing the study's result is proportional to the weight assigned to the study. *P* values presented are derived from heterogeneity tests

studies did not test multiple genetic models of association, which could have led to the identification of other associations.

The results of Liu SY et al. showed a statistically significant association of rs2477686 with male infertility in cases diagnosed with oligozoospermia, oligoasthenozoospermia,

and asthenozoospermia, as well as progressive motility reduction of the sperm associated with GG genotype of the same variant [17]. This is also the single previous study that included the infertile men with oligozoospermia, oligoasthenozoospermia, and asthenozoospermia, while the other studies assessed only the association of analyzed genetic variants with NOA. Even though the participants were not stratified based on their diagnosis into multiple study panels, their results are supported by our findings which also show the protective effect of the C allele. Also, our results on the potential association of rs2477686 with NOA are consistent with the results from other validation studies conducted in Chinese and Japanese populations [15, 16, 18].

As for the potential association of rs10842262 with male infertility, both initial GWAS and another study conducted in the Chinese population provided evidence to support this hypothesis when analyzing allele and genotype frequencies in controls and men diagnosed with NOA [5, 15]. Furthermore, in the study by Tu et al., statistical trend of significance was obtained for the association of rs10842262 with NOA [16], while Sato et al. found no evidence to support the tested association [18]. Our study also showed statistically significant differences in rs10842262 genotype distributions among infertile men and controls. Also, the significant results were found when only the infertile men diagnosed with NOA were included in the association test, consistent with the previous findings in both GWAS and Hu et al. Still, genotype frequencies for rs10842262 in our study were not consistent with HWE for which reason the results were examined for miscalling genotypes. Therefore, for rs10842262, 144 control samples and 96 cases chosen by random sampling were retyped by bidirectional Sanger sequencing followed by capillary gel electrophoresis to assess the validity of proper SNP analysis by the PCR-RFLP method. Our results showed no obvious errors in the genotyping procedure. Another factor with possible influence on deviations from HWE could be natural selection. Therefore, one of the explanations for deviations from HWE could be the effect of the tested genetic variant on male fertility, which is the violation of one of the conditions of Hardy-Weinberg law [26, 28, 29].

In the meta-analysis, the association of rs12097821 with male infertility in general, as well as with NOA, was found for all genetic models tested, suggesting the potential large overall effect of the GWAS on quantitative synthesis results due to high weights assigned to this study and the panels formed based on study phases. Therefore, we have decided to repeat the data synthesis with the exclusion of the results obtained by Hu et al. [5]. Also, as mentioned earlier, in other previous studies, the association under dominant, recessive, and overdominant models was not assessed, but genotype counts were still included in the meta-analysis. The results obtained in the present meta-analysis without the exclusion of GWAS results are similar to the previous one conducted by Tu et al. [16], with only minor change in the overall OR for allelic genetic model. This would suggest the minor effect of the results from case-control study by Liu et al. [17] and the present study on the meta-analysis results, since these two studies were conducted recently, after the previous meta-analysis. The results obtained for other genetic models could not be compared with the previous ones since Tu et al. [16] conducted quantitative data synthesis only for allelic model. Still, after the exclusion of GWAS results [5], statistical significance was lost for all the comparisons, yielding inconsistencies with Tu et al. [16] and suggesting the predominant effect of GWAS on the results of data synthesis. Therefore, additional studies on the association of rs12097821 with male infertility are needed in order to make better assessment of the effect of this genetic variant.

When it comes to rs2477686, similarly as for the rs12097821, it was presumed that the previous results of the meta-analysis by Tu et al. [16] could be influenced by the large effect of the study conducted by Hu et al. [5]. Nevertheless, the statistical significance was preserved for the associations under allelic and recessive genetic models even after the exclusion of GWAS results. Also, the association of the minor allele G with the male infertility occurrence was still statistically significant in the subgroup of Asian descent. Furthermore, statistically significant association under allelic genetic model was also found for a subgroup of studies that included infertile men with oligozoospermia, oligoasthenozoospermia, or asthenozoospermia. Although this subgroup included only two data sets, from *Liu et al.* and the present study, both case-control comparisons yielded significant associations. Nevertheless, other case-control studies, both in Asian and European populations, are needed to validate the association of rs2477686 with oligoasthenozoospermia, as well as to test the association of this genetic variant with NOA.

As for rs10842262, statistically significant results were obtained for the recessive and allelic genetic model, which is expected considering the fact that most of the studies have shown the association of this genetic variant with NOA, or with male infertility in general. Also, this is consistent with the results of the previous meta-analysis [18], which lacked the results from our present case-control study in the data synthesis and which included GWAS results.

We are aware that the main limitation of this study is its relatively small sample size, especially when considering the number of infertile men selected into subgroups according to diagnosis. Therefore, in order to confirm the results obtained in the present study, a further increase in the sample size is required. The observed discordances in results of previous studies and the present one can be explained by differences in ethnic backgrounds, as well as in study designs, selection, and recruitment of participants, exclusion criteria, and

genotyping procedures. Since this is the first European population in which the supposed associations were tested, we could have compared our data with previous ones found only in Asians. Also, the association of these genetic variants with oligoasthenozoospermia was assessed only in one previous study. Therefore, additional studies would provide the results necessary for making conclusions about the potential associations of the studied genetic variants with idiopathic male infertility in population with different ethnic backgrounds, as well as about the associations with oligoasthenozoospermia.

**Acknowledgments** We thank our colleagues from "Academian Vojin Sulovic" Centre for In Vitro Fertilisation, General Hospital Valjevo and Clinic of Urology, Clinical Center of Serbia who greatly assisted the research.

**Code availability** Custom code for association testing is available from the corresponding author on reasonable request.

**Authors' contributions** All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Vucic Nemanja, Matijasevic Suzana, Kotarac Nevena, and Dobrijevic Zorana. The first draft of the manuscript was written by Vucic Nemanja and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding information** The study was financially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Contract no. 451-03-68/2020-14/ 200178).

**Data availability** The dataset generated and analyzed during the current study are available from the corresponding author on reasonable request.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethics approval** The study was approved by the Ethics Committees of Clinical Center of Serbia, Belgrade, Serbia (Number 68/5), and General Hospital Valjevo, Serbia (Number OBV-02-1170), while all the experiments were conducted in accordance with the Helsinki Declaration of 1975.

**Consent to participate** Informed consent was obtained from all individual participants included in the study.

**Consent for publication** Not applicable.

## References

1. Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, Sullivan E, Vanderpoel S, International Committee for Monitoring Assisted Reproductive Technology, World Health Organization International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. Fertil Steril 2009;92(5):1520–1524.
2. Barratt CLR, Bjorndahl L, De Jonge CJ, Lamb DJ, Osorio Martini F, McLachlan R, et al. The diagnosis of male infertility: an analysis of the evidence to support the development of global WHO guidance-challenges and future research opportunities. Hum Reprod Update. 2017;23(6):660–80.
3. Krausz C. Male infertility: pathogenesis and clinical diagnosis. Best Pract Res Clin Endocrinol Metab. 2011;25(2):271–85.
4. Aston KI. Genetic susceptibility to male infertility: news from genome-wide association studies. Andrology. 2014;2(3):315–21.
5. Hu Z, Xia Y, Guo X, Dai J, Li H, Hu H, et al. A genome-wide association study in Chinese men identifies three risk loci for non-obstructive azoospermia. Nat Genet. 2011;44(2):183–6.
6. Frankel A, Yadav N, Lee J, Branscombe TL, Clarke S, Bedford MT. The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity. J Biol Chem. 2002;277(5):3537–43.
7. Luo M, Li Y, Guo H, Lin S, Chen J, Ma Q, et al. Protein arginine methyltransferase 6 involved in germ cell viability during spermatogenesis and down-regulated by the androgen receptor. Int J Mol Sci. 2015;16(12):29467–81.
8. Gonsalvez GB, Rajendra TK, Tian L, Matera AG. The Sm-protein methyltransferase, dart5, is essential for germ-cell specification and maintenance. Curr Biol. 2006;16(11):1077–89.
9. Ancelin K, Lange UC, Hajkova P, Schneider R, Bannister AJ, Kouzarides T, et al. Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol. 2006;8(6):623–30.
10. Chen H, Liu Z, Huang X. Drosophila models of peroxisomal biogenesis disorder: peroxins are required for spermatogenesis and very-long-chain fatty acid metabolism. Hum Mol Genet. 2010;19(3):494–505.
11. Wunderle VM, Critcher R, Ashworth A, Goodfellow PN. Cloning and characterization of SOX5, a new member of the human SOX gene family. Genomics. 1996;36(2):354–8.
12. Budde LM, Wu C, Tilman C, Douglas I, Ghosh S. Regulation of IkappaBbeta expression in testis. Mol Biol Cell. 2002;13(12):4179–94.
13. Frojdman K, Harley VR, Pelliniemi LJ. Sox9 protein in rat sertoli cells is age and stage dependent. Histochem Cell Biol. 2000;113(1):31–6.
14. O'Bryan MK, Takada S, Kennedy CL, Scott G, Harada S, Ray MK, et al. Sox8 is a critical regulator of adult Sertoli cell function and male fertility. Dev Biol. 2008;316(2):359–70.
15. Zou S, Li Z, Wang Y, Chen T, Song P, Chen J, et al. Association study between polymorphisms of PRMT6, PEX10, SOX5, and nonobstructive azoospermia in the Han Chinese population. Biol Reprod. 2014;90(5):96.
16. Tu W, Liu Y, Shen Y, Yan Y, Wang X, Yang D, et al. Genome-wide loci linked to non-obstructive azoospermia susceptibility may be independent of reduced sperm production in males with normozoospermia. Biol Reprod. 2015;92(2):41.
17. Liu SY, Zhang CJ, Peng HY, Sun H, Lin KQ, Huang XQ, et al. Strong association of SLC1A1 and DPF3 gene variants with idiopathic male infertility in Han Chinese. Asian J Androl. 2017;19(4):486–92.
18. Sato Y, Jinam T, Iwamoto T, Yamauchi A, Imoto I, Inoue I, et al. Replication study and meta-analysis of human nonobstructive azoospermia in Japanese populations. Biol Reprod. 2013;88(4):87.
19. Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, Behre HM, et al. World Health Organization reference values for human semen characteristics. Hum Reprod Update. 2010;16(3):231–45.
20. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3–new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115.

21. Hall TA, editor. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium. 1999;41:95–8.
22. Sole X, Guino E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. Bioinformatics. 2006;22(15):1928–9.
23. Bozdogan H. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika. 1987;52(3):345–70.
24. Gonzalez JR, Armengol L, Sole X, Guino E, Mercader JM, Estivill X, et al. SNPassoc: an R package to perform whole genome association studies. Bioinformatics. 2007;23(5):644–5.
25. Wallace BC, Dahabreh IJ, Trikalinos TA, Lau J, Trow P, Schmid CH. Closing the gap between methodologists and end-users: R as a computational back-end. J Stat Softw. 2012;49(5):1–15.
26. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719–48.
27. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88.
28. Schaid DJ, Jacobsen SJ. Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions. Am J Epidemiol. 1999;149(8):706–11.
29. Neale B, Ferreira M, Medland S, Posthuma D. Statistical genetics: gene mapping through linkage and association. Abingdon-on-Thames: Taylor & Francis; 2007.