



Association between long-chain non-coding RNA *SRA1* gene single-nucleotide polymorphism and polycystic ovary syndrome susceptibility.

Jifan Tan¹ · XiuLan Hao¹ · TingTing Zhao¹ · JianLan Ying¹ · Tian Li¹ · Li Cheng¹

Received: 10 June 2020 / Accepted: 6 August 2020 / Published online: 11 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Objective To study the association between single-nucleotide polymorphism (SNP) of long-chain non-coding RNA steroid receptor RNA activator (lncRNA *SRA1*) gene and polycystic ovary syndrome (PCOS) susceptibility.

Methods Sanger sequencing was used to analyze the genotypes of the lncRNA *SRA1* gene rs801460, rs10463297, and rs250426 in 315 PCOS patients and 315 control groups.

Results There was no correlation between lncRNA *SRA1* gene rs801460, rs250426 SNP, and PCOS susceptibility ($p > 0.05$). The T allele at the rs10463297 locus of the *SRA1* gene has a lower risk of PCOS than the C allele (OR = 0.63, 95%CI: 0.50–0.79, $p < 0.01$). Among people with a BMI ≥ 26.5 kg/m², when carrying the TC genotype and CC genotype at rs801460, the risk of PCOS susceptibility was lower than the TT genotype (OR = 0.54, 95%CI: 0.33–0.89, $p = 0.02$). At different ages and BMI stratifications, there was a significant association between rs10463297 SNP and PCOS susceptibility ($p < 0.05$). Multi-factor dimensionality reduction (MDR) analysis results showed that age, BMI, rs801460, rs10463297, and rs250426 interactions constitute a “high-risk combination.” PCOS susceptibility risk was 5.96 times that of a “low-risk combination” (95%CI: 4.14–8.56, $p < 0.01$). *SRA1* gene rs801460, rs10463297, rs250426 constructed TCT haplotype was associated with increased risk of PCOS susceptibility (OR = 1.66, 95%CI: 1.20–2.30, $p < 0.01$); the CTT haplotype was associated with a decreased risk of PCOS susceptibility (OR = 0.56, 95%CI: 0.36–0.87, $p = 0.01$). lncRNA *SRA1* gene rs10463297 SNP was correlated with the level of lncRNA *SRA1* in the peripheral blood leukocytes ($p < 0.01$).

Conclusion From this study, we found that the lncRNA *SRA1* gene rs10463297 SNP is associated with PCOS susceptibility.

Keywords Polycystic ovary syndrome · Single-nucleotide polymorphism · Steroid receptor RNA activator · Long noncoding RNA

Jifan Tan and XiuLan Hao contributed equally to this work.

✉ Tian Li
sandyliitian@126.com

✉ Li Cheng
cl0926@126.com

Jifan Tan
tanjifan89@163.com

XiuLan Hao
41127491@qq.com

TingTing Zhao
283418302@qq.com

JianLan Ying
229762655@qq.com

¹ Department of Obstetrics and Gynecology, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518107, China

Introduction

Polycystic ovary syndrome (PCOS) is one of the most common reproductive endocrine diseases of women of childbearing age. The prevalence of women of childbearing age in the world is as high as 10%, and it is the most common cause of anovulatory infertility [1]. The main clinical manifestations of PCOS are rare ovulation or anovulation, polycystic ovarian changes, and clinical or laboratory hyperandrogenism [2]. Studies have shown that PCOS is associated with multiple dysfunctions, including infertility, obstetric complications, diabetes, obesity, and cardiovascular disease [3]. It is a complex disease with a high degree of heterogeneity, high incidence during adolescence and childbirth, and potential high-risk complications after childbirth, affecting the health and quality of life of patients throughout their lives.

In recent years, some researchers have suggested that there is a vicious circle in PCOS patients, that is, androgen excess can cause the accumulation of visceral fat and lead to obesity; the latter cause excessive androgen production by the ovaries and/or adrenal glands directly or indirectly through autocrine, paracrine, and internal feedback or by causing insulin resistance and hyperinsulinemia [4]. So far, the exact etiology and pathophysiology of PCOS are still controversial. Some studies have found that the candidate genes related to the cause of PCOS are involved in many aspects such as the process of sex hormone action, insulin action, and chronic inflammatory factors [5–7]. However, due to the heterogeneity of genotypes and phenotypes, no gene has been recognized as a determinant gene for the onset of PCOS. It can be seen that the abnormal interaction of multiple genes/multiple regulatory pathways leads to feedback imbalances and chain reactions that may cause a high degree of heterogeneity in the clinical manifestations of PCOS. Therefore, further understanding of the molecular mechanism of the pathogenesis of PCOS will help to discover new diagnostic methods and treatment strategies.

With the continuous development of detection technology, whole genome sequencing studies have found that most genome codes produce lncRNA [8]. lncRNA can regulate gene expression at epigenetic, transcriptional, and post-transcriptional levels and plays an important role in the regulation of gene expression during cell and embryo development and maintenance of body homeostasis [9]. Studies have shown that many lncRNAs are closely related to the function of steroid hormone receptors, and the mechanism is likely to be the interactive regulation between the expression or activity of lncRNA and steroid receptor [10]. The research on lncRNA has broadened our understanding of the pathogenesis of various diseases, but there are few studies on the association between lncRNA gene polymorphism and PCOS susceptibility.

SRA1 was first discovered by Lanz et al. [11] in 1999, and it is related to steroid hormone receptor-related pathways. SRA1, as a “molecular scaffold,” cooperates with various transcription factors and co-regulators to enhance or inhibit the expression of genes regulated by steroid hormones [12]. Early research found that lncRNA SRA1 can enhance the activation of steroid hormone receptors through riboprotein complexes, but more studies have found that it is widely involved in the activation and inhibition of many nuclear receptors (such as transcription factors and non-steroidal receptors) and regulates many physiological processes, such as steroid receptor signaling pathways, steroid synthesis, adipocyte synthesis, and insulin sensitivity [12]. The above processes are involved in the pathogenesis of PCOS, which suggests that lncRNA SRA1 may play a role in the occurrence and development of PCOS.

However, the correlation between the lncRNA SRA1 gene polymorphism and the occurrence and development of PCOS are unclear. In this study, we selected three common SNP sites

(minor allele frequency > 0.1) in the lncRNA *SRA1* gene to study and analyze its association with PCOS susceptibility. Among them, rs801460 and rs10463297 belong to the intron variant of lncRNA *SRA1* gene. Yan et al. [13] found that rs801460 and rs10463297 SNP were significantly associated with estrogen receptor (ER) positive status. It was suggested that the lncRNA *SRA1* gene polymorphism may be significantly related to the occurrence and development of breast cancer. The rs250426 locus belongs to stop gained. We assume that the functional SNP in the lncRNA *SRA1* gene may be related to PCOS susceptibility.

Based on Haploview 4.1 software, we selected rs801460, rs10463297, and rs250426 of lncRNA *SRA1* gene for research. The purpose of this study was to analyze the relationship between the rs801460, rs10463297, and rs250426 SNPs of the lncRNA *SRA1* gene and the susceptibility of PCOS to provide support for the diagnosis and prevention of PCOS in clinical work.

Materials and methods

Subject

We recruited 315 patients with PCOS from the Seventh Affiliated Hospital of Sun Yat-sen University between January 2017 and January 2020. The diagnosis of PCOS patients refers to the diagnostic standards recommended by the European Society of Human Reproduction and Embryology and the American Society of Reproductive Medicine (ESHRE / ASRM) in 2003 [14]. Two of the following three criteria were met: (1) menstrual disorders (rare menstruation or amenorrhea, less than 9 cycles per year); (2) clinical or biochemical hyperandrogenism; and (3) ultrasound examination shows multiple ovaries cystic changes (at least 12 follicles of 2–9 mm in one or both ovaries or a volume greater than 10 ml in one ovary). Patients with PCOS have 12 or more follicles in each ovary, a diameter of 2 ± 9 mm, and/or increased ovarian volume (> 10 ml). Exclusion criteria are as follows: (1) patients with congenital adrenal hyperplasia; (2) patients with Cushing syndrome; (3) PCOS subjects with androgen-secreting tumors; (4) patients with diabetes and glucose intolerance are excluded; (5) patients with PCOS caused by metabolic disorders; and (6) cancer patients. We recruited healthy women with normal endocrine function and no metabolic diseases as a control group. The subjects in the control group had a normal menstrual cycle and normal ovulation without PCOS-related endocrine dysfunction. PCOS patients and controls did not receive hormone therapy within 3 months before the study.

We obtained the age at menarche (AAM) of PCOS patients and the control group through face-to-face communication. The clinical data of the subjects were complete. On the 3rd–5th day of the menstrual cycle, 5 ml of peripheral blood was

obtained. After anti-coagulation with ethylenediaminetetraacetic acid (EDTA), and centrifugation at 3000 r/min for 15 min, the serum was separated and stored in a refrigerator at -80°C . The contents of follicle stimulating hormone (FSH), luteinizing hormone (LH), total testosterone, prolactin, thyroid stimulating hormone (TSH), and estradiol (E₂) were measured by radioimmunoassay (RIA), and the coefficient of variation within and between measurements is controlled within 10%. All subjects signed an informed consent form, and the study protocol was approved by the ethics committee of the Seventh Affiliated Hospital of Sun Yat-sen University.

Genotype analysis

We collected peripheral blood leukocytes and extracted the genomic DNA using QIAamp DNA Blood Mini Kit 250 (QIAGEN, Germany, Hilden). Using genomic DNA as a template, PCR was used to amplify DNA fragments containing the rs801460 and rs10463297 and rs250426 locus of the lncRNA *SRA1* gene. The sequence of rs801460 amplification primer is 5'-AGGGCTACTCTCTGAGACAACT-3' (forward primer) and 5'-GCCTAGCATATTGTCTGGCG-3' (reverse primer). rs10463297 amplification primer sequences is 5'-CAGACACCTCGTTACCCAGTC-3' (forward primer) and 5'-GGGCGA ACTAGACAGCAGTG-3' (reverse primer). The sequence of the amplification primer at rs250426 is 5'-TCTCCGTCTGTCTCCGAGC-3' (forward primer) and 5'-CTGCTTACCAAGAGGGTCGC-3' (reverse primer). The PCR reaction mixture contained 50 ng genomic DNA, 1 μL forward primer, 1 μL reverse primer, 1 U/ μL Taq DNA polymerase, 100 μM dNTP, and 2 mM Mg²⁺. The PCR reaction was first performed at 94 $^{\circ}\text{C}$ for 5 min; and then amplified at 98 $^{\circ}\text{C}$ for 10 s, 58 $^{\circ}\text{C}$ for 15 s, and 72 $^{\circ}\text{C}$ for 2 min and 30 cycles of amplification; and finally extended at 72 $^{\circ}\text{C}$ for 5 min. The PCR products were sequenced by BioSun Biotechnology Co., Ltd. (Hangzhou, China). The sequencing results were compared with the sequences in the dbSNP database (<https://www.ncbi.nlm.nih.gov/snp/>), and the SNP locus genotypes were determined according to the comparison results.

Quantitative real-time PCR (qRT-PCR)

We used TRIzol (Invitrogen, CA, USA) to extract total RNA from serum leukocytes. Using the extracted RNA as a template, reverse transcription PCR was performed to synthesize cDNA using PrimeScript RT reagent Kit (Takara, Shiga, Japan) according to the supplier's instructions. SYBR® Premix Ex Taq™ II (Takara Bio, Inc., Shiga, Japan) was used to detect lncRNA *SRA1* levels on the Rotor Gene 3000 system (Corbet Research, Sydney, Australia), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a control. The sequence of lncRNA *SRA1* primers is 5'-CCTGGACGTGTCTCAACTGG-3' (forward primer) and

5'-CCCGGAACTCCACTGTTAGC-3' (reverse primer). GAPDH primers is 5'-GGGAACTGTGGCGTGAT-3' (forward primer) and 5'-GAGTGGGTGTCGCTGTTGA-3' (reverse primer). The expression level of lncRNA *SRA1* relative to GAPDH is represented by $2^{-\Delta\Delta\text{Ct}}$, and 3 accessory wells are repeated for each sample.

Statistical analysis

In this study, SPSS 22.0 (SPSS Inc., Chicago, IL, USA) was used for statistical analysis. The statistical analysis of the categorical variable [n(%)] uses the χ^2 test; continuous variables (mean \pm SD) use *t* test or ANOVA. We used χ^2 test to evaluate whether the genotype frequencies of rs801460, rs10463297, and rs250426 of lncRNA *SRA1* gene were in Hardy-Weinberg equilibrium (HWE). Logistic regression was used to analyze the correlation between rs801460, rs10463297, and rs250426 SNPs of the lncRNA *SRA1* gene and PCOS susceptibility and calculate the odds ratio (OR) and 95% confidence interval (CI). Bonferroni's correction was applied to adjust age, BMI, and age at menarche (AAM). Multi-factor dimensionality reduction (MDR) was used to further evaluate the interaction between the *SRA1* gene rs801460, rs10463297, and rs250426 SNPs with the age and body mass index (BMI). Haploview 4.1 was used to analyze the linkage disequilibrium of *SRA1* gene rs801460, rs10463297, and rs250426 locus. The receiver operating curve (ROC) was used to evaluate the diagnostic value of PCOS in lncRNA *SRA1* levels in peripheral blood leukocytes. All statistical tests are two-sided tests; $p < 0.05$ indicates that the difference is statistically significant.

Results

Characteristics of the study population

The clinical characteristics of 315 PCOS patients and 315 control groups are shown in Table 1. The age of PCOS patients was not significantly different from that of the control group, but the BMI of PCOS patients was much higher than that of the control group ($p < 0.01$). Compared with the control group, PCOS patients had higher levels of LH, LH/FSH, total testosterone, TSH, and E₂ ($p < 0.01$), but there was no statistically significant difference between FSH, prolactin, AAM, and the control group ($p > 0.05$).

SRA1 gene polymorphism and PCOS susceptibility

The distribution of genotypes and allele frequencies of *SRA1* gene rs801460, rs10463297, and rs250426 in PCOS patients and controls are shown in Table 2. In the control group, rs801460 locus, rs10463297 locus, and rs250426 locus genotype were in Hardy-Weinberg equilibrium (rs801460 locus

Table 1 Characteristics of PCOS cases and non-PCOS controls

Parameters	PCOS (n = 315)	Control (n = 315)	p
Age (years, mean ± SD)	30.60 ± 6.71	31.32 ± 6.97	0.19
BMI (kg/m ² , mean ± SD)	28.05 ± 3.33	24.95 ± 2.48	<0.01
FSH (IU/L, mean ± SD)	7.53 ± 3.62	7.50 ± 2.66	0.91
LH (IU/L, mean ± SD)	15.24 ± 5.87	4.94 ± 2.14	<0.01
LH/FSH	0.95 ± 0.42	0.74 ± 0.41	<0.01
Total testosterone (ng/mL, mean ± SD)	0.95 ± 0.42	0.53 ± 0.20	<0.01
Prolactin (ng/mL, mean ± SD)	15.39 ± 5.51	14.69 ± 3.17	0.06
TSH (μIU/mL, mean ± SD)	2.63 ± 0.82	1.52 ± 0.43	<0.01
AAM (years, mean ± SD)	14.54 ± 2.50	14.24 ± 2.32	0.12
E ₂ (pg/mL, mean ± SD)	221.46 ± 76.65	164.37 ± 72.53	<0.01

PCOS polycystic ovary syndrome, BMI body mass index, FSH follicle stimulating hormone, LH luteinizing hormone, TSH thyroid stimulating hormone, E₂ estradiol, AAM age at menarche

$p = 0.17$, rs10463297 locus $p = 0.18$, rs250426 locus $p = 0.16$). We did not find an association between the rs801460 and rs250426 SNP and PCOS ($p > 0.05$). *SRAI* gene rs10463297 locus CT genotype, TT genotype, and dominant model (CT + TT vs. CC) were associated with decreased risk

of PCOS susceptibility (OR = 0.55, 95%CI: 0.39–0.77, $p < 0.01$; OR = 0.49, 95% CI: 0.31–0.78, $p < 0.01$; OR = 0.53, 95%CI: 0.39–0.73, $p < 0.01$); the T allele had a lower risk of PCOS than the C allele (OR = 0.63, 95%CI: 0.50–0.79, $p < 0.01$).

Table 2 Comparison of genotype and allele frequency of *SRAI* gene rs801460, rs10463297, and rs250426 in PCOS patients and control group

	PCOS (n = 315)	Control (n = 315)	HWE p	OR (95%CI)*	p
rs801460					
TT	131 (41.59%)	121 (38.41%)	0.17	1.00 (reference)	
TC	125 (39.68%)	139 (44.13%)		0.83 (0.59–1.17)	0.34
CC	59 (18.73%)	55 (17.46%)		0.99 (0.64–1.54)	0.97
Dominance model				0.88 (0.64–1.21)	0.46
Recessive model				1.09 (0.73–1.64)	0.76
T	387 (61.43%)	381 (60.48%)		1.00 (reference)	
C	243 (38.57%)	249 (39.52%)		0.96 (0.77–1.21)	0.77
rs10463297					
CC	164 (52.06%)	115 (36.51%)	0.18	1.00 (reference)	
CT	110 (34.92%)	141 (44.76%)		0.55 (0.39–0.77)	<0.01
TT	41 (13.02%)	59 (18.73%)		0.49 (0.31–0.78)	<0.01
Dominance model				0.53 (0.39–0.73)	<0.01
Recessive model				0.65 (0.42–1.01)	0.06
C	438 (69.52%)	371 (58.89%)		1.00 (reference)	
T	192 (30.48%)	259 (41.11%)		0.63 (0.50–0.79)	<0.01
rs250426					
TT	205 (65.08%)	223 (70.79%)	0.16	1.00 (reference)	
TG	89 (28.25%)	80 (25.40%)		1.21 (0.85–1.73)	0.34
GG	21 (6.67%)	12 (3.81%)		1.90 (0.91–3.97)	0.12
Dominance model				1.30 (0.93–1.82)	0.15
Recessive model				1.80 (0.87–3.73)	0.15
T	499 (79.21%)	526 (83.49%)		1.00 (reference)	
G	131 (20.79%)	104 (16.51%)		1.33 (0.99–1.77)	0.06

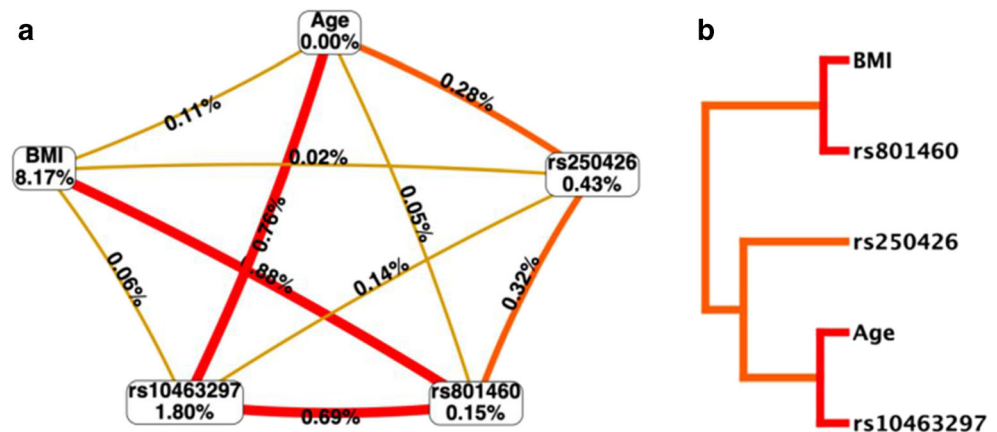
*Data were calculated by logistic regression analysis with adjusted for age, BMI, and age at menarche. PCOS polycystic ovary syndrome, HWE Hardy-Weinberg equilibrium, OR odds ratio, CI confidence interval

Table 3 Stratified analysis of *SRAI* gene polymorphism and PCOS susceptibility

	PCOS (n = 315)	Control (n = 315)	OR (95%CI)*	p
rs801460				
Age (years)				
< 30.9				
TT	72 (43.37%)	63 (37.95%)	1.00 (reference)	
TC + CC	94 (56.63%)	103 (62.05%)	0.80 (0.52–1.24)	0.37
≥ 30.9				
TT	59 (39.60%)	58 (38.93%)	1.00 (reference)	
TC + CC	90 (60.40%)	91 (61.07%)	0.97 (0.61–1.55)	0.91
BMI (kg/m ²)				
< 26.5				
TT	35 (33.65%)	88 (42.11%)	1.00 (reference)	
TC + CC	69 (66.35%)	121 (57.89%)	1.43 (0.88–2.34)	0.19
≥ 26.5				
TT	96 (45.50%)	33 (31.13%)	1.00 (reference)	
TC + CC	115 (54.50%)	73 (68.87%)	0.54 (0.33–0.89)	0.02
rs10463297				
Age (years)				
< 30.9				
CC	86 (51.81%)	66 (39.76%)	1.00 (reference)	
CT + TT	80 (48.19%)	100 (60.24%)	0.61 (0.40–0.95)	0.04
≥ 30.9				
CC	78 (52.35%)	49 (32.89%)	1.00 (reference)	
CT + TT	71 (47.65%)	100 (67.11%)	0.45 (0.28–0.71)	< 0.01
BMI(kg/m ²)				
< 26.5				
CC	55 (52.88%)	80 (38.28%)	1.00 (reference)	
CT + TT	49 (47.12%)	129 (61.72%)	0.55 (0.34–0.89)	0.02
≥ 26.5				
CC	109 (51.66%)	35 (33.02%)	1.00 (reference)	
CT + TT	102 (48.34%)	71 (66.98%)	0.46 (0.28–0.75)	< 0.01
rs250426				
Age(years)				
< 30.9				
TT	107 (64.46%)	123 (74.10%)	1.00 (reference)	
TG + GG	59 (35.54%)	43 (25.90%)	1.58 (0.99–2.53)	0.07
≥ 30.9				
TT	98 (65.77%)	100 (67.11%)	1.00 (reference)	
TG + GG	51 (34.23%)	49 (32.89%)	1.06 (0.66–1.72)	0.90
BMI (kg/m ²)				
< 26.5				
TT	78 (75.00%)	151 (72.25%)	1.00 (reference)	
TG + GG	26 (25.00%)	58 (27.75%)	0.87 (0.51–1.49)	0.70
≥ 26.5				
TT	127 (60.19%)	72 (67.92%)	1.00 (reference)	
TG + GG	84 (39.81%)	34 (32.08%)	1.40 (0.86–2.29)	0.22

*Data were calculated by logistic regression analysis with adjusted for age, BMI, and age at menarche. *PCOS* polycystic ovary syndrome, *OR* odds ratio, *CI* confidence interval, *BMI* body mass index

Fig. 1 MDR analysis of *SRA1* gene SNP interaction with age and BMI. **a** Circular graph. The data at the vertices represents the contribution to PCOS. The online data represents the strength of the interaction between the factors at the two vertices. **b** Dendrogram, strong interactions get together, and weak interactions stay away from each other



Stratified analysis of *SRA1* gene polymorphism and PCOS susceptibility

We further conducted stratified analysis to assess the association between *SRA1* gene polymorphism and PCOS susceptibility (Table 3). We found that among people with BMI ≥ 26.5 kg/m², when carrying the TC genotype and CC genotype at rs801460, the risk of PCOS susceptibility was lower than that of the TT genotype (OR = 0.54, 95% CI: 0.33–0.89, $p = 0.02$), there was no association between rs801460 gene polymorphism and PCOS susceptibility ($p > 0.05$) in other stratifications. At different ages and BMI stratifications, there was a significant association between rs10463297 SNP and PCOS susceptibility ($p < 0.05$). At different ages and different BMI stratifications, no correlation was found between the rs250426 SNP and PCOS susceptibility ($p > 0.05$).

Interaction between SNP of *SRA1* gene and age and BMI

We used multi-factor dimensionality reduction (MDR) to further evaluate the interaction of the *SRA1* gene rs801460, rs10463297, and rs250426 SNP with the age and BMI of the subject (Fig. 1). We found that BMI, rs10463297 SNP, and PCOS susceptibility are significantly related; *SRA1* rs801460, rs10463297, and rs250426 SNP interact with age

and BMI (Fig. 1a). There was a strong interaction between BMI and rs801460 and age and rs10463297 (Fig. 1b). The best model of interaction was composed of age, BMI, rs801460, rs10463297, and rs250426. Training accuracy was 0.7092; the risk of PCOS susceptibility of “high-risk combination” was 5.96 times of “low-risk combination” (95%CI: 4.14–8.56, $p < 0.01$) (Table 4).

Analysis of linkage disequilibrium of *SRA1* gene SNP

In this study, Haploview 4.1 was used to analyze the linkage disequilibrium of *SRA1* gene rs801460, rs10463297, and rs250426 (Fig. 2). Four haplotypes can be constructed based on the analysis results. We found that the TCT haplotype was associated with an increased risk of PCOS susceptibility (OR = 1.66, 95%CI: 1.20–2.30, $p < 0.01$); the CTT haplotype was associated with a reduced risk of PCOS susceptibility (OR = 0.56, 95%CI: 0.36–0.87, $p = 0.01$) (Table 5).

False-positive report probability (FPRP)

The false-positive report rate (FPRP) values at different prior probability levels are shown in Table 6. Considering that this study was an exploratory study and the sample size was small, we set the critical value of FPRP value to 0.5. When the prior probability was set to 0.1, the prior probability calculated in the

Table 4 MDR analysis results of *SRA1* gene rs801460, rs10463297, and rs250426 locus SNP interaction with age and BMI

Model	Training accuracy	OR (95%CI)*	p	CVC
BMI	0.6667	4.00 (1.40–11.41)	< 0.01	10/10
Age, BMI	0.6667	4.00 (2.82–5.67)	< 0.01	10/10
BMI, rs10463297, rs250426	0.6753	4.40 (3.09–6.28)	< 0.01	5/10
BMI,rs801460, rs10463297, rs250426	0.6887	4.96 (3.47–7.09)	< 0.01	6/10
Age, BMI, rs801460, rs10463297, rs250426	0.7092	5.96 (4.14–8.56)	< 0.01	10/10

*Data were calculated by logistic regression analysis with adjusted for age, BMI, and age at menarche. OR odds ratio, CI confidence interval, BMI body mass index, CVC cross-validation consistency, MDR multi-factor dimensionality reduction

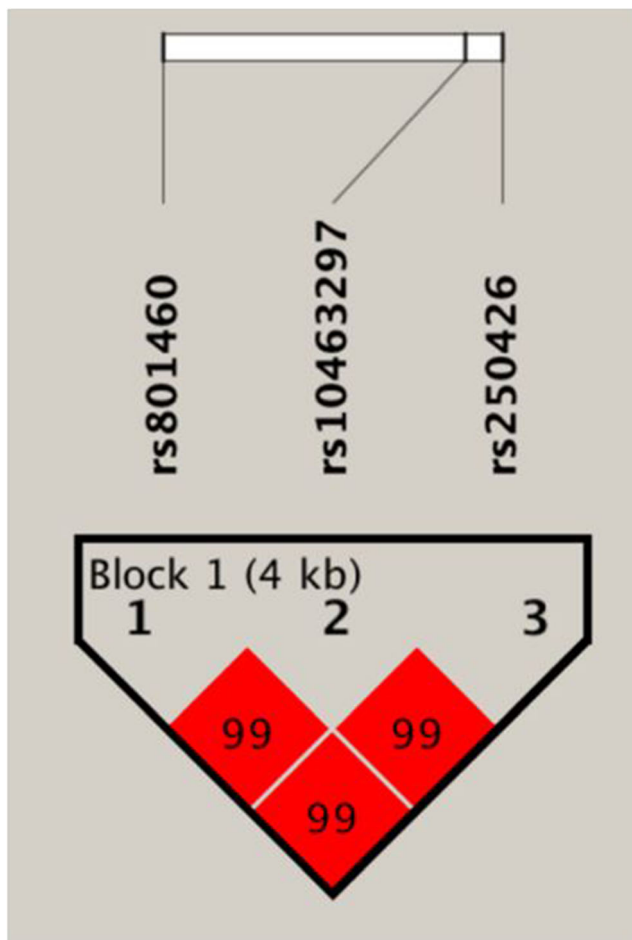


Fig. 2 Haploview 4.1 software analyzes the linkage disequilibrium map of *SRA1* gene rs801460, rs10463297, and rs250426 locus. The data in the red squares represent the value of $100D'$; $D' = 0$ means that the SNP sites are in complete linkage equilibrium. $D' = 1$ means that there is a complete linkage disequilibrium between SNP locus

rs10463297 locus dominant model (CT + TT vs. CC) with age \geq 30.9 years was 0.500, the prior probability values of other calculations are all < 0.500 , which showed that the correlation between SNP and PCOS in this study has certain credibility.

Level of lncRNA SRA1 in peripheral blood leukocytes

QRT-PCR was used to detect lncRNA SRA1 levels in peripheral blood leukocytes of 315 PCOS patients and 315 control

groups. The results showed that the level of lncRNA SRA1 in peripheral blood leukocytes of PCOS patients was significantly higher than that of the control group ($p < 0.01$, Fig. 3a). The receiver operating curve (ROC) was used to analyze and evaluate the diagnostic value of lncRNA SRA1 levels in peripheral blood leukocytes for PCOS. The results showed that the cut-off was 0.35, sensitivity was 49.52% (95%CI: 44.04%–55.02%), specificity was 66.35% (95%CI: 60.96%–71.34%), and the area under the curve (AUC) was 0.59 (95 %CI: 0.55–0.64, $p < 0.01$) (Fig. 3b).

Correlation between lncRNA SRA1 gene SNP and lncRNA SRA1 level in peripheral blood leukocytes

Further analysis of the correlation between the lncRNA *SRA1* gene SNP and the level of lncRNA SRA1 in peripheral blood leukocytes. The results showed that there was no statistically significant difference in the levels of lncRNA SRA1 in peripheral blood leukocytes of PCOS patients with different genotypes at rs801460 locus ($p = 0.31$); however, the difference in lncRNA SRA1 levels in peripheral blood leukocytes with different genotypes at rs801460 was significant ($p = 0.02$) in the control group (Fig. 4a). Both PCOS patients and control subjects had statistically significant differences in the levels of lncRNA SRA1 in peripheral blood leukocytes with different genotypes at rs10463297 locus ($p < 0.01$) (Fig. 4B). The rs250426 gene polymorphism was not significantly correlated with lncRNA SRA1 levels in peripheral blood leukocytes of PCOS patients and controls ($p = 0.65$, $p = 0.81$) (Fig. 4c).

Correlation between lncRNA SRA1 level in peripheral blood leukocytes and obesity

We analyzed the correlation between BMI and lncRNA SRA1 levels in peripheral blood leukocytes of PCOS patients and control subjects. The results showed that the levels of lncRNA SRA1 in peripheral blood leukocytes of PCOS patients and control subjects were positively correlated with BMI ($r = 0.88$, $p < 0.01$; $r = 0.84$, $p < 0.01$; Fig. 5a, b). The average BMI of all subjects in this study was 26.5 kg/m²; we define BMI \geq 26.5 kg/m² as obese, BMI < 26.5 kg/m² as non-obesity; the results showed that in PCOS patients and control groups, the

Table 5 Correlation between the haplotypes of rs801460, rs10463297, and rs250426 of *SRA1* gene and PCOS susceptibility risk

Haplotype [#]	PCOS(n = 315)	Control(n = 315)	OR(95%CI)	<i>p</i>
CTG	107 (33.90%)	116 (36.80%)	0.88 (0.64–1.22)	0.51
TCT	143 (45.50%)	105 (33.20%)	1.66 (1.20–2.30)	< 0.01
CTT	38 (12.20%)	62 (19.80%)	0.56 (0.36–0.87)	0.01
TTT	27 (8.50%)	32 (10.30%)	0.83 (0.48–1.42)	0.58

[#] rs801460, rs10463297, rs250426 locus haplotype. PCOS polycystic ovary syndrome, OR odds ratio, CI confidence interval

Table 6 FPRP value of the results of the study on the correlation between lncRNA SRA1 gene SNP and PCOS susceptibility

Genotype	OR(95%CI)	Statistical power	Prior probability		
			0.1	0.01	0.001
rs10463297 CT vs. CC	0.55 (0.39–0.77)	0.857	0.450	0.900	0.989
rs10463297 TT vs. CC	0.49 (0.31–0.78)	0.910	0.479	0.910	0.990
rs10463297 CT + TT vs. CC	0.53 (0.39–0.73)	0.922	0.459	0.903	0.990
rs10463297 C allele vs. T allele	0.63 (0.50–0.79)	0.985	0.417	0.887	0.988
rs801460 TC + CC vs. TT					
BMI \geq 26.5 kg/m ²	0.54 (0.33–0.89)	0.942	0.455	0.902	0.989
rs10463297 CT + TT vs. CC					
Age < 30.9 years	0.61 (0.40–0.95)	0.886	0.425	0.890	0.988
Age \geq 30.9 years	0.45 (0.28–0.71)	0.876	0.500	0.917	0.991
BMI < 26.5 kg/m ²	0.55 (0.34–0.89)	0.931	0.450	0.900	0.989
BMI \geq 26.5 kg/m ²	0.46 (0.28–0.75)	0.895	0.495	0.915	0.991
TCT Haplotype#	1.66 (1.20–2.30)	0.902	0.213	0.749	0.968
CTT Haplotype#	0.56 (0.36–0.87)	0.877	0.446	0.898	0.989

OR odds ratio, CI confidence interval. # rs801460, rs10463297, rs250426 site formed haplotype

level of lncRNA SRA1 in peripheral blood leukocytes of obese people was relatively high ($p < 0.01$, Fig. 5c, d).

Discussion

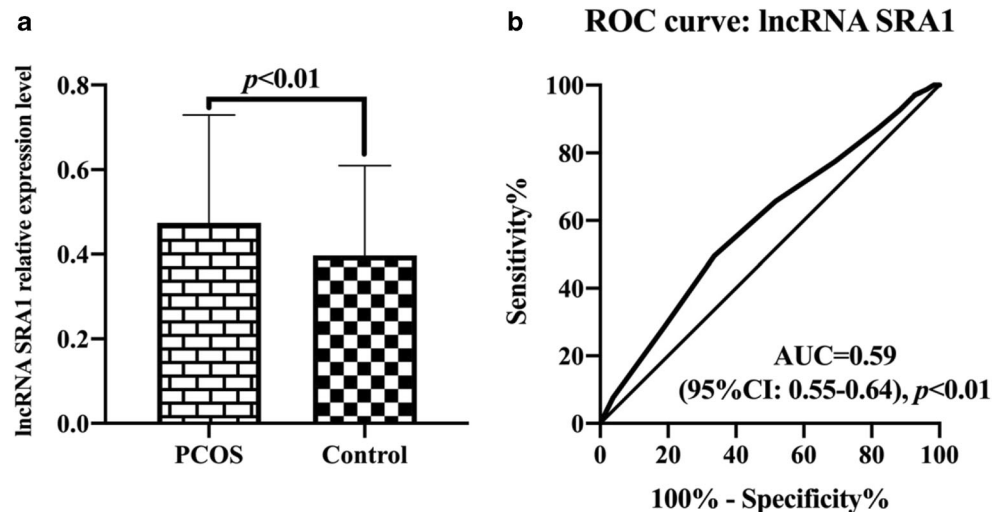
In this study, we used a case-control study to find that the SNP of the lncRNA SRA1 gene rs10463297 SNP was associated with PCOS susceptibility. We found that the TCT haplotype constructed at rs801460, rs10463297, and rs250426 at the SRA1 gene was associated with an increased risk of PCOS susceptibility (OR = 1.66, 95%CI: 1.20–2.30, $p < 0.01$). The CTT haplotype was associated with a reduced risk of PCOS susceptibility (OR = 0.56, 95%CI: 0.36–0.87, $p = 0.01$). Further research found that there was a significant correlation

between the lncRNA SRA1 gene rs10463297 SNP and the level of lncRNA SRA1 in peripheral blood leukocytes.

With the change of natural environment and psychosocial factors, the number of infertility and PCOS patients in young fertility groups has increased significantly [15]. In recent years, with the change of China's fertility policy, the demand for second children of older women whose fertility has declined significantly has increased, and this group of people has significantly increased their attention to reproductive health.

lncRNA was originally regarded as the “noise” of genome transcription, as a by-product of RNA polymerase II transcription, and had no biological function [16, 17]. Steroid hormone receptors are members of the DNA transcription factor superfamily. They play a key role in biological processes such as reproductive endocrinology by binding to ligands in formal functions [18, 19]. A variety of lncRNAs are closely related to

Fig. 3 Detection of lncRNA SRA1 levels in peripheral blood leukocytes. **a** Comparison of lncRNA SRA1 levels in peripheral blood leukocytes of PCOS patients and control group. **b** Receiver operating curve (ROC) for diagnosis of PCOS in lncRNA SRA1 levels of peripheral blood leukocytes



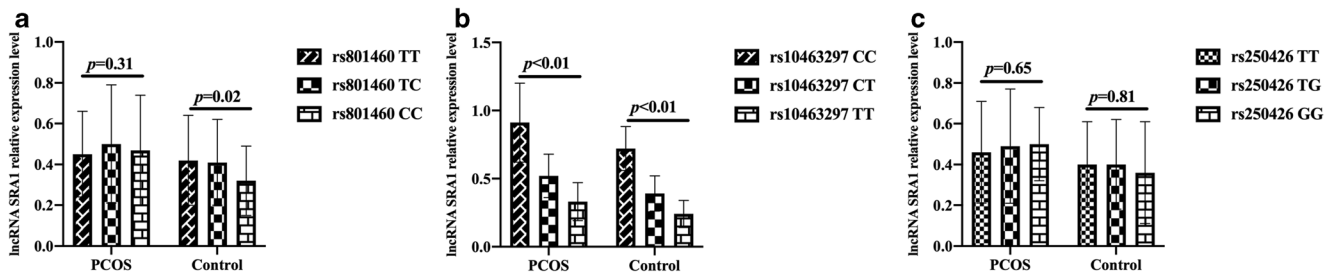


Fig. 4 Correlation between IncRNA *SRA1* gene SNP and IncRNA *SRA1* level in peripheral blood leukocytes. **a** Comparison of IncRNA *SRA1* levels in peripheral blood leukocytes with different genotypes at rs801460 locus in PCOS patients and the control group. **b** Comparison of IncRNA *SRA1* levels in peripheral blood leukocytes with different

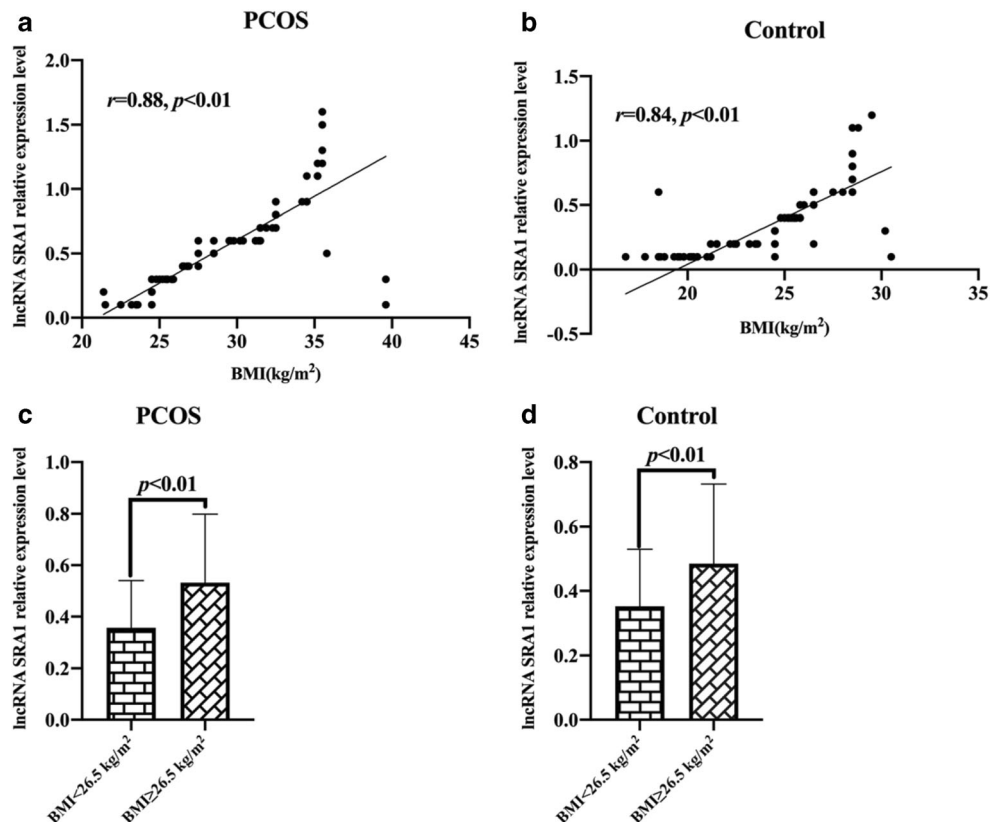
genotypes at rs10463297 locus in PCOS patients and control group. **c** Comparison of IncRNA *SRA1* levels in peripheral blood leukocytes with different genotypes at rs250426 locus in PCOS patients and the control group

the function of steroid hormone receptors, and the mechanism is probably the interactive regulation between lncRNAs and the expression or activity of steroid receptors. For example, CTBP1-AS is an androgen-regulated lncRNA and a natural antisense transcript of the androgen receptor co-suppressor factor CTBP1 [20].

In this study, we found that the expression level of IncRNA *SRA1* in peripheral blood leukocytes of PCOS patients increased, and the level of IncRNA *SRA1* in peripheral blood leukocytes of PCOS patients and controls were positively correlated with BMI. The latest research shows that IncRNA *SRA1* acts as a transcriptional coactivator of PPAR γ and promotes the differentiation of

adipocytes in vitro [21]. The expression level of IncRNA *SRA1* in mouse adipose tissue was significantly higher than that of other organs. After knocking out IncRNA *SRA1*, mice were resistant to high-fat diet-induced obesity, resulting in a decrease in fat mass and an increase in meat content. The results of this study are similar to this study, indicating that IncRNA *SRA1* may be an important regulator of fat cell synthesis and function. In addition, the results of this study also show that the detection of IncRNA *SRA1* levels in peripheral blood leukocytes has potential value in the diagnosis of PCOS. It is speculated that IncRNA *SRA1* may become a potential target for PCOS diagnosis and treatment.

Fig. 5 Correlation between IncRNA *SRA1* levels in peripheral blood leukocytes and BMI. **a** Correlation between IncRNA *SRA1* level and BMI in peripheral blood leukocytes of PCOS patients. **b** Correlation between IncRNA *SRA1* level and BMI in peripheral blood leukocytes of controls. **c** Comparison of IncRNA *SRA1* levels in peripheral blood leukocytes of PCOS patients with different BMI. **d** Comparison of IncRNA *SRA1* levels in peripheral blood leukocytes of control groups with different BMI



In this study, we found that there was no association between the lncRNA *SRA1* gene rs801460 and rs250426 SNPs and PCOS, while the T allele carriers at the rs10463297 locus had a lower risk of PCOS than the C allele carriers. We calculated the minimum sample size using the rs10463297 allele frequency as a reference. We found that the minimum sample size required for both PCOS patients and the control group in this study were 315 and 315, which is consistent with the sample size selected in this study. In addition, the calculation results of the FPRP showed that there was a real correlation between the SNP and PCOS studied in this study. This shows that the results of this study have a certain degree of objective credibility.

Interestingly, we found that the risk of PCOS susceptibility was lower than that of TT genotype when carrying rs801460 TC genotype and CC genotype in people with BMI \geq 26.5 kg/m². Therefore, we speculated that the association between rs801460 SNP and PCOS susceptibility was related to BMI, indicating that rs801460 may be a disease-associated SNP in obese PCOS patients. We also found that the risk of PCOS susceptibility of the “high-risk combination” composed of the interaction of age, BMI, rs801460, rs10463297, and rs250426 locus was 5.96 times that of the “low-risk combination.” This also suggested that the interaction between SNP and environmental factors was related to the occurrence of PCOS.

In addition, our analysis of lncRNA *SRA1* levels in peripheral blood leukocytes showed that the difference in lncRNA *SRA1* levels in peripheral blood leukocytes of the control group at the rs801460 locus was significant, but we did not find this correlation in PCOS patients. Whether it was related to BMI stratification needs further study. However, both in PCOS patients and controls, the differences in the levels of lncRNA *SRA1* in peripheral blood leukocytes with different genotypes at the rs10463297 locus were statistically significant. Whether they were related to the occurrence of PCOS needs further verification in *in vivo* models.

Although some new findings have been made in this study, there are still some shortcomings in this study that need further study. First of all, the correlation between SNP and the structure and function of lncRNA *SRA1* were not studied in this study. We know that lncRNA functions through secondary structure, but there is currently no direct evidence of the effect of SNP at rs801460, rs10463297, and rs250426 on the structure of lncRNA *SRA1*. Secondly, the specific cause of the association between lncRNA *SRA1* rs10463297 SNP and PCOS susceptibility needs to be further verified *in vivo*. In addition, we know that the regulatory network in the body is intricate, the genes and environmental factors that may be involved are complex, and further research on the joint effects of other genes and environmental factors is needed. Therefore, it is necessary to further study the role of other genes and environmental factors’ interaction in the occurrence and development of PCOS.

Conclusion

From this study, we found that the lncRNA *SRA1* gene rs10463297 locus SNP is associated with PCOS susceptibility. The lncRNA *SRA1* gene rs10463297 locus SNP is significantly correlated with lncRNA *SRA1* levels in peripheral blood leukocytes. The role of more other genetic polymorphisms and environmental interactions in the occurrence and development of PCOS needs further study.

Funding information This work was supported by Guangdong Natural Science Foundation (grant no. 2018A030313735).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Goodarzi MO, Dumesic DA, Chazenbalk G, Azziz R. Polycystic ovary syndrome: etiology, pathogenesis and diagnosis. *Nat Rev Endocrinol*. 2011;7:219–31.
2. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, et al. Positions statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an androgen excess society guideline. *J Clin Endocrinol Metab*. 2006;91:4237–45.
3. Legro RS, Arslanian SA, Ehrmann DA, Hoeger KM, Murad MH, Pasquali R, et al. Diagnosis and treatment of polycystic ovary syndrome: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2013;98:4565–92.
4. Murri M, Insenser M, Fernandez-Duran E, San-Millan JL, Escobar-Morreale HF. Effects of polycystic ovary syndrome (PCOS), sex hormones, and obesity on circulating miRNA-21, miRNA-27b, miRNA-103, and miRNA-155 expression. *J Clin Endocrinol Metab*. 2013;98:E1835–44.
5. Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Shi Y, et al. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet*. 2011;43:55–9.
6. Louwers YV, Stolk L, Uitterlinden AG, Laven JS. Cross-ethnic meta-analysis of genetic variants for polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2013;98:E2006–12.
7. Shi Y, Zhao H, Shi Y, Cao Y, Yang D, Li Z, et al. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet*. 2012;44:1020–5.
8. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet*. 2014;15:7–21.
9. Roberts TC, Morris KV, Weinberg MS. Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*. 2014;9:13–20.
10. Sun M, Kraus WL. From discovery to function: the expanding roles of long noncoding RNAs in physiology and disease. *Endocr Rev*. 2015;36:25–64.
11. Lanz RB, McKenna NJ, Onate SA, Albrecht U, Wong J, Tsai SY, et al. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell*. 1999;97:17–27.

12. Colley SM, Leedman PJ. Steroid receptor RNA activator - a nuclear receptor coregulator with multiple partners: insights and challenges. *Biochimie*. 2011;93:1966–72.
13. Yan R, Wang K, Peng R, Wang S, Cao J, Wang P, et al. Genetic variants in lncRNA SRA and risk of breast cancer. *Oncotarget*. 2016;7:22486–96.
14. Rotterdam EA-SPCWG. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril*. 2004;81:19–25.
15. Lopez-Pajares V. Long non-coding RNA regulation of gene expression during differentiation. *Pflugers Arch*. 2016;468:971–81.
16. Chatterjee D, Sanchez AM, Goldgur Y, Shuman S, Schwer B. Transcription of lncRNA prt, clustered prt RNA sites for Mmi1 binding, and RNA polymerase II CTD phospho-sites govern the repression of *pho1* gene expression under phosphate-replete conditions in fission yeast. *RNA*. 2016;22:1011–25.
17. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet*. 2016;17:47–62.
18. Schwartz N, Verma A, Bivens CB, Schwartz Z, Boyan BD. Rapid steroid hormone actions via membrane receptors. *Biochim Biophys Acta*. 2016;1863:2289–98.
19. Stanisic V, Lonard DM, O'Malley BW. Modulation of steroid hormone receptor activity. *Prog Brain Res*. 2010;181:153–76.
20. Takayama K, Tsutsumi S, Katayama S, Okayama T, Horie-Inoue K, Ikeda K, et al. Integration of cap analysis of gene expression and chromatin immunoprecipitation analysis on array reveals genome-wide androgen receptor signaling in prostate cancer cells. *Oncogene*. 2011;30:619–30.
21. Liu S, Sheng L, Miao H, Saunders TL, Macdougald OA, Koenig RJ, et al. SRA gene knockout protects against diet-induced obesity and improves glucose tolerance. *J Biol Chem*. 2014;289:13000–9.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.