33 (2022) 201063

Contents lists available at ScienceDirect

Human Gene

journal homepage: www.journals.elsevier.com/human-gene

Association of seven fundamental genetic polymorphisms in long noncoding RNA MALAT1, SOX2OT and H19 with recurrent miscarriage in Turkish-Azeri Iranian population

Parvin Hakimi [a,b], Naser Lotfalizad [c], Leyla Pabarja [d], Mohammad Reza Asadi [c], Jalal Gharesouran [c], Hossein Dehghani [e], Hani Sabaie [c], Shahla Danaie [f], Bashdar Mahmud Hussen [g], Mohammad Taheri [h,*], Maryam Rezazadeh [c,*]

[a] Molecular Medicine Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[b] Women's Reproductive Health Research Center, Tabriz University of medical sciences, Tabriz, Iran
[c] Department of Medical Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
[d] Department of Molecular Genetics, Rabe Rashidi Institute, Tabriz, Iran
[e] Department of Molecular Medicine, School of Medicine, Birjand University of Medical Sciences, Birjand, Iran
[f] Obstetrics and Gynecology Group in ACECR of East Azerbaijan, Tabriz, Iran
[g] Department of Pharmacognosy, College of Pharmacy, Hawler Medical University, Kurdistan Region, Iraq
[h] Urology and Nephrology Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
Recurrent miscarriage
lncRNA
MALAT1
SOX2OT
H19
ceRNA
rs619586
SNP

ABSTRACT

Background: Recurrent miscarriage (RM), at a rate of 1% of all couples trying to conceive, is one of the main problems in pregnancy. The root cause of the disease has not been determined, but genetic factors and lifestyle influence the process. The minor change in the expression of lncRNAs shows its significant effects on various complications, one of which can be RM. SNPs are single nucleotide polymorphisms that have the potential to alter the function and expression of lncRNAs.

Material and methods: This study investigated seven SNPs in RM with 400 cases and 400 controls in the Iranian Turkish Azeri population using ARMS-PCR, including the MALAT1 lncRNA rs591291 rs3200401, rs619586, rs664589, and rs656605, SOX2OT lncRNA rs9839776, and H19 lncRNA rs3741216 variants.

Results: MALAT1 lncRNA rs591291 rs3200401, rs619586, rs664589, and rs656605, SOX2OT lncRNA rs9839776, and H19 lncRNA rs3741216 variants. We found that the MALAT1 lncRNA rs619586G variant in the co-dominant (AA vs. AG; OR = 0.65, 95% CI = 0.47–0.9 and AA vs. GG; OR = 0.5, 95% CI = 0.26–0.99, $P = 0.0074$), the dominant (AG + GG vs. AA; OR = 0.63, 95% CI = 0.46–0.85, $P = 0.0023$) and over-dominant (AA+GG vs. AG; OR = 0.68, 95% CI = 0.5–0.93, $P = 0.016$) inherited models has an association with decreased risk of RM and the MALAT1 lncRNA rs591291- rs3200401- rs619586 - rs664589 - rs656605 (C-C-G-C-G) haplotype is most likely a haplotype prone to RM.

Conclusion: In conclusion, our findings propose that the rs619586 G variant in the Iranian Turkish Azeri may have potential protective effects, lowering the risk of RM.

1. Introduction

Recurrent Miscarriage (RM) is the name given to abortions before 20 weeks. The disorder affects 1 to 5% of pregnant women and about 5% of fertile couples (Hefler et al., 2001; Rai and Regan, 2006). Some experts consider two consecutive abortions, and others consider three or more abortions as RM (Stephenson and Kutteh, 2007). The leading cause of RM has not been identified yet; some couples have not found a cure despite many studies. However, some cases can be considered influential factors, including chromosomal disorders (Branch et al., 2010), anatomical factors (Rull et al., 2012), thrombophilia (McNamee et al., 2012), endocrine disorders (Kaur and Gupta, 2016), and immune factors (Williams, 2012). It has rarely been observed that RM occurs for a single specific reason, and in most cases, a combination of several factors

* Corresponding authors.
E-mail addresses: Mohammad.taheri@uni-jena.de (M. Taheri), Rezazadehm@tbzmed.ac.ir (M. Rezazadeh).

https://doi.org/10.1016/j.humgen.2022.201063
Received 17 February 2022; Received in revised form 29 March 2022; Accepted 30 May 2022
Available online 1 June 2022

causes this abnormality, which is why it is becoming increasingly difficult to identify the mechanism and genes involved in this complication. Among these, lncRNAs are important with regulatory and critical roles in basic mechanisms.

Long noncoding RNAs (lncRNA) with a specified length of between 200 and several kb do not encode proteins. They play an essential regulatory role in many processes such as transcription, regulation of gene expression, tumorigenesis, migration, and cell invasion (Shi et al., 2013; Ma et al., 2016; Wu et al., 2019). Blood vessel formation, the primary process in the crosstalk between mother and fetus, is one of the most critical roles regulated by lncRNAs (Wang et al., 2018a). Regulation of trophoblast cell migration and invasion by lncRNAs is also a vital pathway studied in spontaneous abortion (Qin et al., 2019). Thus, lncRNA expression regulation is crucial because of its role in cellular pathways and disorder, which can be one of the main pathogenic mechanisms in this disease. Polymorphisms are among the factors that disrupt the expression of lncRNAs (Xue et al., 2018; Zheng et al., 2018). Therefore, further studies on the association between polymorphisms in lncRNA genes in RM in other populations may help better understand the disease's underlying mechanism.

Metastasis Associated Lung Adenocarcinoma Transcript1 (MALAT1) is an eight thousand nucleotide-length lncRNA most expressed in the lungs, pancreas, and nervous system and highly expressed in cancers such as lung cancer, hepatocellular carcinoma, breast, and pancreatic cancer (Wang et al., 2018a). Inflammation is one of the positively associated processes with RM, and many studies have shown the involvement of MALAT1 in inflammatory processes (Vitagliano et al., 2017). Instead, MALAT1 expression in villus samples of RM patients showed a significant decrease compared to controls, indicating that regulation of MALAT1 expression is one of the primary mechanisms in RM (Wang et al., 2018a). A study of MALAT1 gene-specific polymorphisms of rs591291, rs619586, and rs3200401 in 2019 on RM in the Chinese population revealed the protective role of the rs619586 G variant (Che et al., 2019). SOX2 overlapping transcript (SOX2OT) Another lncRNA located on the sox2 gene. The role of SOX2OT in many cancers, including breast cancer and colorectal cancer, has been studied (Askarian-Amiri et al., 2014; Liu et al., 2016). A study on the southern China population in 2019 considered the effect of the rs9839776CT variant on increasing carrier risk for RM (Fang et al., 2019). LncRNA H19 is an imprinted gene expressed by the maternal allele and is expressed in many tissues during the fetal period, but it loses its expression in most tissues except the endometrium and ovaries in adulthood (Zeng et al., 2020). The role of H19 in infertility is proved through its reduced expression and the effect on implantation in infertile women (Korucuoglu et al., 2010). This study analyzes lncRNA MALAT1 rs591291, rs619586, rs3200401, rs664589, and rs656605 lncRNA SOX2OT rs9839776 and lncRNA H19 rs3741216 in a case-control study with 400 cases and 400 controls. The relationship between these polymorphisms and RM in the Iranian Turkish Azeri population has been studied.

2. Materials and methods

2.1. Ethics statement

This study was approved by the Research Ethics Committee (REC) of Tabriz University of Medical Sciences (TUOMS) (66328). All-female subjects who participated in this study signed and approved a written informed consent form.

2.2. Study subjects

In the present study, a total of 400 patients with RM with two or more spontaneous abortions of an unknown cause with the same male partner and 400 healthy controls (with two months of normal pregnancy and no history of abortion) were recruited in Tabriz Children's Medical Center and Hospital, between May 2017 and April 2021. The control group matched the case group in terms of age, and all study subjects were selected from a homogeneous population. All couples with RM had chromosomal abnormalities removed from the study. There were no autoimmune diseases, metabolic disorders, arterial or venous thrombosis, hypertension, endocrine disorders, uterine abnormalities, hepatic or renal dysfunction, or fetal chromosomal abnormalities in the case or control groups.

2.3. SNP genotyping and DNA extraction

Each subject's peripheral blood was taken from a vein and placed in a tube containing EDTA. Using a DNA extraction kit (Geneall, Korea) and following the manufacturer's instructions, genomic DNA samples were extracted from 200 μL of peripheral blood leukocyte samples obtained from all participants. The DNA was kept in a 20 °C freezer until the genotype analysis. To select SNPs, we first filtered the SNPs for lncRNA H19, lncRNA MALAT1, and lncRNA SOX2OT based on the database (dbSNP) and with the Global Minimum Allele Frequency (MAF) criterion>5%. Then, tetra-primer amplification refractory mutation system (ARMS)-PCR was used for genotyping specific single nucleotide polymorphisms. Table-S1 summarizes the ARMS PCR primer set and sequences, each primer's temperature, and product size in the supplementary information. In a total volume of 25 μL, PCR reaction was carried out with 7.5 μL of ddH2O, 12.5 μL of Taq DNA Polymerase 2× Master Mix RED (2 mM MgCl), 100 ng of DNA template, one μL of each outer primer, and ten pmol of each primer. The PCR products were electrophoresed on a 2% agarose gel. Bse1I (Thermo Scientific, United States) restriction enzyme was used to detect the rs619586 polymorphism. The reactions were incubated at 55 °C for 16 h. In rs6999622 polymorphism, normal individuals showed 322 bp and 160 bp bands, and mutant individuals had 322 bp and 207 bp bands, while heterozygotes showed all three fragments displayed in Figure-1S. The S1-S6 figures in the supplementary information show the Tetra-primer amplification refractory mutation system (ARMS)-PCR analysis of rs371216, rs591291, rs3200401, rs664589, rs656605, and rs9839776. In all PCR reactions, negative controls had no DNA pattern. Enzyme digestions were completed with both positive and negative controls. The experiments were blindly performed by a technician, meaning that the samples were coded and the disease status was unknown.

2.4. Bioinformatics analysis

All variants were evaluated for the effect of creating or losing miRNA action sites in the lncSNP2 database. Gaining the action site of miRNAs can involve lncRNAs and miRNAs in a ceRNA network. Conversely, the

Table 1
bioinformatics analysis using lncRNASNP2 database (http://bioinfo.life.hust.edu.cn/lncRNASNP).

lncRNA	SNP	Effect on miRNA target	miRNA
MALAT1	rs619586	Gain	miR-214-3p* miR-3619-3p miR-761-3p
MALAT1	rs664589	Gain	miR-122-3p
MALAT1	rs3200401	Loss	miR-1324 miR-3661 miR-4499
MALAT1	rs656605	Gain	miR-6855-3p miR-371b-5p miR-4513 miR-5002-3p miR-373-5p miR-3191-3p miR-548ah-5p
H19	rs3741216	Gain	miR-130b-3p* miR-301a-3p

loss of the action site of miRNA can disrupt a network of ceRNAs. The results of the analysis are summarized in Table 1.

2.5. Statistical analyses

Statistical analyses were performed on the collected data using SPSS 18 software (statistical package for the social sciences Inc., Chicago, IL, USA). The Chi-square ($\chi 2$) test was used to compare the frequencies of alleles and genotypes. The odds ratio (OR) was used in four inheritance models to create an association between each genotype and RM (co-dominant, dominant, recessive, and over-dominant). Using the Bonferroni correction test, *P* values were adjusted for multiple comparisons. *Statistical significance* was defined as a P value of less than 0.05. Haplotype association of MALAT1 polymorphisms with disease calculated using the SNPStats (Solé et al., 2006) (https://www.snpstats.net/start.htm). Hardy-Weinberg equilibrium (HWE) for the control group was calculated using the goodness-of exact test (Edwards, 2008).

3. Results

In the current study, we genotyped 7 SNPs associated with lncRNAs. The rs591291, rs3200401, rs619586, rs664589, and rs656605 variants are located on the MALAT1 lncRNA. The SOX2OT lncRNA rs9839776 and the H19 lncRNA rs3741216 variants are also genotyped from 7 SNPs. Statistical analysis revealed no significant deviation from Hardy-Weinberg equilibrium for rs3200401, rs664589, rs656605, rs9839776, and rs3741216 polymorphisms, except rs591291, which showed a significant departure from Hardy–Weinberg equilibrium in both patients ($p = 0.0038$) and controls ($p = 0.002$) and rs619586 in all subjects ($p = 0.0083$) (Table 2). Next, four inheritance models were used to calculate the genotype-RM risk associations (co-dominant, dominant, over-dominant, and recessive models). There was no significant association between RM and the MALAT1 rs591291, rs3200401, rs664589, and rs656605 genotypes in any of the inheritance models, as well as the SOX2OT rs9839776 and the H19 rs3741216 genotypes (Table 3). However, there was a significant association between the rs619586 SNP and RM risk. For the rs619586 A > G polymorphism, G allele carriers have a significantly decreased RM risk (AA vs. AG; OR = 0.65, 95% CI = 0.47–0.9 and AA vs. GG; OR = 0.5, 95% CI = 0.26–0.99, $P = 0.0074$) in the co-dominant model. Similarly, significant associations were observed in the dominant (AG + GG vs. AA; OR = 0.63, 95% CI = 0.46–0.85, $P = 0.0023$) and over-dominant (AA+GG vs. AG; OR = 0.68, 95% CI = 0.5–0.93, $P = 0.016$) models.

Haplotype association of MALAT1 polymorphisms with disease analysis showed that rs591291- rs3200401-rs619586-rs664589-rs656605 (C-C-G-C-A) haplotype with *p*-value = 0.0009, OR = 0.35, 95% CI = 0.19–0.65 has a higher frequency in controls (0.0916) than in patients (0.027). On the other hand, rs591291-rs3200401-rs619586-rs664589-rs656605 (C-C-G-C-G) haplotype with a p-value = 0.038, OR = 0.35, and 95%CI = 1.06–6.88 has a higher frequency in patients than controls. (Table 4).

4. Discussion

Although no specific cause for RM has been identified, studies should continue because about 1% of couples who attempt to conceive are affected by RM (Kling et al., 2018). Indeed, the factors influencing the process of causing this complication can be rooted in genetic factors and lifestyle (Ng et al., 2021). This study tries to address the genetic dimension that predisposes to RM. This study investigated seven specific polymorphisms in MALAT1, SOX2OT, and H19 lncRNAs to determine the RM susceptibility in the Turkish Azeri population as a case-control study.

A case-control study frequently compares the prevalence of a particular disease between people with normal alleles and people with variant alleles, yielding an odds ratio (OR). Single-nucleotide polymorphism, the most common type of allele variation, consists of a major allele (M) and a minor allele (m). As a result, the genotype can be major allele homozygote (MM), heterozygote (Mm), or minor allele homozygote (mm). Each genotype is assigned odds, and a pair of odds produces an OR. The primary method of estimating an OR is to summarize data using a two-by-two contingency. As a result, the three types of genotypes are frequently transformed into two variables. A dominant model, for example, compares MM versus Mm + mm, whereas a recessive model compares MM + Mm versus mm. Assuming the heterozygote has the most significant impact, an over-dominant model compares MM + mm versus Mm. Co-dominant models, which include additive and multiplicative models, hypothesize that MM, Mm, and mm are associated with the lowest, intermediate, and highest risk or associated with the highest, intermediate, and lowest risk, respectively (Thakkinstian et al., 2005; Attia et al., 2003).

MALAT1 is one of the lncRNAs that has been related to diseases, and a growing body of data shows that MALAT1 is involved in several pathological processes (Zhao et al., 2018a). Several studies have connected MALAT1 gene polymorphisms to illness vulnerability. For example, the MALAT1 rs619586 G polymorphism has been associated with a reduced incidence of hepatocellular carcinoma (Liu et al., 2012) and colorectal cancer (Zhao et al., 2018b). A case-control study in the Chinese population also observed that the MALAT1 rs619586 AG and GG genotypes are associated with a reduced incidence of coronary atherosclerotic heart disease and may have a protective role in preventing coronary artery disease atherosclerotic heart disease (Wang et al., 2018b). Simultaneously, a study in the Chinese Han population revealed that the lncRNA MALAT1 rs619586 AG genotype and the rs3200401 CT genotype are associated with a decreased risk of breast cancer and that carriers of the rs619586 G variant have lower MALAT1 expression rather than carriers of the rs619586AA genotype (Peng et al., 2018).

Most notably, the p53-MALAT1 axis was identified in RM (Wang et al., 2019). RM patients have elevated levels of the cell cycle regulator P53. P53 repressed MALAT1 in placenta cells, influencing the normal function of placental cells. One of the possible reasons for RM could be a dysregulated p53-MALAT1 axis (Wang et al., 2019). Similarly, another study found that MALAT1 expression levels were lower in RM patients' villous specimens (Wu et al., 2021). MALAT1 aided trophoblast cell migration and invasion by suppressing CRY2 protein expression. MALAT1 recruited FBXW7 to reduce the stability of the CRY2 protein. MALAT1 may recruit the E3 ubiquitin ligase FBXW7 in order to induce CRY2 ubiquitin-mediated degradation and to involve in trophoblast migration and invasion. MALAT1 deficiency in trophoblasts could be linked to RM (Wu et al., 2021).

In this regard, MALAT1 lncRNA polymorphisms rs591291, rs3200401, and rs619586 have been studied in research on the Chinese population. The MALAT1 gene rs619586 G variant reduced the incidence of RM in a southern Chinese community and was also a protective factor against vulnerability to RM (Che et al., 2019).

Table 2
exact test for Hardy-Weinberg equilibrium (*P* Value).

	rs591291	rs3200401	rs619586	rs664589	rs656605	rs9839776	rs3741216
All subjects	<0.0001	0.68	0.0083	1	0.12	1	0.21
Controls	0.002	0.38	0.075	0.61	0.3	0.88	0.82
Patient's	0.0038	0.89	0.074	0.82	0.25	0.89	0.18

Table 3
Association study of rs591291, rs3200401, rs619586, rs664589, rs656605, rs9839776, and rs3741216 polymorphisms and RM.

Gene	Locus	Model	Genotype	Controls	Patients	Odds Ratio	P value
MALAT1	rs591291	Allele	C	600 (75%)	580 (72.5%)	1	0.26
			T	200 (25%)	220 (27.5%)	1.14 (0.91–1.42)	
		Codominant	CC	237 (59.2%)	222 (55.5%)	1.00	0.55
			CT	126 (31.5%)	136 (34%)	1.15 (0.85–1.56)	
			TT	37 (9.2%)	42 (10.5%)	1.21 (0.75–1.96)	
		Dominant	CC	237 (59.2%)	222 (55.5%)	1.00	0.28
			CT + TT	163 (40.8%)	178 (44.5%)	1.17 (0.88–1.54)	
		Recessive	CC + CT	363 (90.8%)	358 (89.5%)	1.00	0.55
			TT	37 (9.2%)	42 (10.5%)	1.15 (0.72–1.83)	
		Overdominant	CC + TT	274 (68.5%)	264 (66%)	1.00	0.45
			CT	126 (31.5%)	136 (34%)	1.12 (0.83–1.51)	
		Log-Additive				1.12 (0.91–1.38)	0.29
	rs3200401	Allele	C	624 (78%)	620 (77.5%)	1	0.81
			T	176 (22%)	180 (22.5%)	1.03 (0.81–1.3)	
		Codominant	CC	240 (60%)	241 (60.2%)	1.00	0.67
			CT	144 (36%)	138 (34.5%)	0.95 (0.71–1.28)	
			TT	16 (4%)	21 (5.3%)	1.31 (0.67–2.57)	
		Dominant	CC	240 (60%)	241 (60.2%)	1.00	0.94
			CT + TT	160 (40%)	159 (39.8%)	0.99 (0.75–1.31)	
		Recessive	CC + CT	384 (96%)	379 (94.8%)	1.00	0.4
			TT	16 (4%)	21 (5.3%)	1.33 (0.68–2.59)	
		Overdominant	CC + TT	256 (64%)	262 (65.5%)	1.00	0.66
			CT	144 (36%)	138 (34.5%)	0.94 (0.7–1.25)	
		Log-Additive				1.03 (0.81–1.31)	0.81
	rs619586	Allele	A	630 (78.8%)	680 (85%)	1	0.001
			G	170 (22.2%)	120 (15%)	0.65 (0.5–0.85)	
		Codominant	AA	254 (63.5%)	294 (73.5%)	1.00	0.0074
			AG	122 (30.5%)	92 (23%)	0.65 (0.47–0.9)	
			GG	24 (6%)	14 (3.5%)	0.5 (0.26–0.99)	
		Dominant	AA	254 (63.5%)	294 (73.5%)	1.00	0.0023
			AG + GG	146 (36.5%)	106 (26.5%)	0.63 (0.46–0.85)	
		Recessive	AA+AG	376 (94%)	386 (96.5%)	1.00	0.095
			GG	24 (6%)	14 (3.5%)	0.57 (0.29–1.12)	
		Overdominant	AA+GG	278 (69.5%)	308 (77%)	1.00	0.016
			AG	122 (30.5%)	92 (23%)	0.68 (0.5–0.93)	
		Log-Additive				0.68 (0.53–0.87)	0.0019
	rs664589	Allele	C	712 (89%)	700 (87.5%)	1	0.35
			G	88 (11%)	100 (12.5%)	1.16 (0.85–1.57)	
		Codominant	CC	318 (79.5%)	305 (76.3%)	1	0.46
			CG	76 (19%)	90 (22.5%)	1.23 (0.88–1.74)	
			GG	6 (1.5%)	5 (1.2%)	0.87 (0.26–2.88)	
		Dominant	CC	318 (79.5%)	305 (76.3%)	1.00	0.27
			CG + GG	82 (20.5%)	95 (23.8%)	1.21 (0.86–1.69)	
		Recessive	CC + CG	394 (98.5%)	395 (98.8%)	1.00	0.76
			GG	6 (1.5%)	5 (1.2%)	0.83 (0.25–2.75)	
		Overdominant	CC + GG	324 (81%)	310 (77.5%)	1.00	0.22
			CG	76 (19%)	90 (22.5%)	1.24 (0.88–1.74)	
		Log-Additive				1.16 (0.85–1.57)	0.35

Gene	Locus	Model	Genotype	Controls	Patients	Odds Ratio	P value
SOXOT	rs656605	Allele	A	736 (92%)	722 (90.2%)	1	0.22
			G	64 (8%)	78 (9.8%)	1.24 (0.88–1.76)	
		Codominant	AA	340 (85%)	328 (82%)	1.00	0.49
			AG	56 (14%)	66 (16.5%)	1.22 (0.83–1.8)	
			GG	4 (1%)	6 (1.5%)	1.55 (0.43–5.56)	
		Dominant	AA	340 (85%)	328 (82%)	1.00	0.25
			AG + GG	60 (15%)	72 (18%)	1.24 (0.86–1.81)	
		Recessive	AA+AG	396 (99%)	394 (98.5%)	1.00	0.52
			GG	4 (1%)	6 (1.5%)	1.51 (0.42–5.38)	
		Overdominant	AA+GG	344 (86%)	334 (83.5%)	1.00	0.33
			AG	56 (14%)	66 (16.5%)	1.21 (0.82–1.79)	
		Log-Additive				1.23 (0.88–1.72)	0.23
	rs9839776	Allele	C	641 (80.1%)	620 (77.5%)	1	0.2
			T	159 (19.9)	180 (22.5%)	1.17 (0.92–1.49)	
		Codominant	CC	256 (64%)	241 (60.2%)	1.00	0.41
			CT	129 (32.2%)	138 (34.5)	1.14 (0.84–1.53)	
			TT	15 (3.8%)	21 (5.2%)	1.49 (0.75–2.95)	
		Dominant	CC	256 (64%)	241 (60.2%)	1.00	0.27
			CT + TT	144 (36%)	159 (39.8%)	1.17 (0.88–1.56)	
		Recessive	CC + CT	385 (96.2%)	379 (94.8%)	1.00	0.31
			TT	15 (3.8%)	21 (5.2%)	1.42 (0.72–2.8)	
		Overdominant	CC + TT	271 (67.8%)	262 (65.5%)	1.00	0.5

(*continued on next page*)

Table 3 (*continued*)

Gene	Locus	Model	Genotype	Controls	Patients	Odds Ratio	P value
			CT	129 (32.2%)	138 (34.5%)	1.11 (0.82–1.48)	
		Log-Additive				1.17 (0.92–1.49)	0.2
		Allele	T	698 (87.2%)	697 (87.1%)	1	0.94
			A	102 (12.8%)	103 (12.9%)	1.01 (0.75–1.35)	
			TT	305 (76.2%)	307 (76.8%)	1.00	0.71
		Codominant	TA	88 (22%)	83 (20.8%)	0.94 (0.67–1.32)	
			AA	7 (1.8%)	10 (2.5%)	1.42 (0.53–3.78)	
			TT	305 (76.2%)	307 (76.8%)	1.00	0.87
H19	rs3741216	Dominant	TA + AA	95 (23.8%)	93 (23.2%)	0.97 (0.7–1.35)	
			TT + TA	393 (98.2%)	390 (97.5%)	1.00	0.46
		Recessive	AA	7 (1.8%)	10 (2.5%)	1.44 (0.54–3.82)	
			TT + AA	312 (78%)	317 (79.2%)	1.00	0.67
		Overdominant	TA	88 (22%)	83 (20.8%)	0.93 (0.66–1.3)	
		Log-Additive				1.01 (0.76–1.35)	0.94

Table 4
Haplotype association of *MALAT1* polymorphisms with disease (calculated using the SNPStats (https://www.snpstats.net/start.htm).

rs591291	rs3200401	rs619586	rs664589	rs656605	Frequency in control group	Frequency in patients' group	OR (95%CI)	P value
C	C	A	C	A	0.3984	0.4151	1.00	–
T	C	A	C	A	0.1073	0.1472	1.2 (0.82–1.76)	0.34
C	T	A	C	A	0.1019	0.1008	0.91 (0.59–1.41)	0.68
C	C	G	C	A	0.0916	0.027	0.35 (0.19–0.65)	0.0009
C	C	A	G	A	0.0494	0.0585	1.06 (0.59–1.91)	0.84
T	T	A	C	A	0.0425	0.0479	1.12 (0.61–2.05)	0.71
C	C	A	C	G	0.0286	0.0345	1.01 (0.51–1.98)	0.98
T	C	G	C	A	0.0372	0.0217	0.62 (0.27–1.47)	0.28
C	C	G	C	G	0.0145	0.0352	2.7 (1.06–6.88)	0.038
C	T	G	C	A	0.0254	0.0093	0.53 (0.15–1.9)	0.33
T	T	G	C	A	0.014	0.0163	1.07 (0.42–2.73)	0.89
T	C	A	G	A	0.0163	0.015	1.32 (0.31–5.67)	0.71
C	T	A	G	A	0.0047	0.0047	6.15 (0.68–55.95)	0.11
C	C	G	G	A	0.0203	0.0056	0.36 (0. 06–1.96)	0.24

MALAT1 plays a vital role in splicing and epigenetic regulation (Yang et al., 2011). Given the role of MALAT1 in inflammation and its association with RM (Vitagliano et al., 2017) to reduced expression in VILLUS specimens in patients with RM (Wang et al., 2018a), its role as an essential lncRNA in the process of RM should be considered.

In our study, five MALAT1 SNPs were examined, which include rs591291, rs3200401, rs619586, rs664589, and rs656605. The rs619586G variant in the co-dominant (AA vs. AG; OR = 0.65, 95% CI = 0.47–0.9 and AA vs. GG; OR = 0.5, 95% CI = 0.26–0.99, P = 0.0074), the dominant (AG + GG vs. AA; OR = 0.63, 95% CI = 0.46–0.85, P = 0.0023) and over-dominant (AA+GG vs. AG; OR = 0.68, 95% CI = 0.5–0.93, P = 0.016) inherited models has a protective effect against RM and probably reduces the risk of disease and plays an important role in protecting against disease. No significant results were obtained from the other four MALAT1 variants. Haplotype association of MALAT1 polymorphisms with disease analysis showed that rs591291-rs3200401-rs619586-rs664589-rs656605 (C-C-G-C-G) haplotype with *p*-value = 0.038, OR = 0.35, 95%CI = 1.06–6.88 with a higher frequency in patients than controls, It is most likely a haplotype prone to RM, while the rs591291-rs3200401-rs619586-rs664589-rs656605 (CCGCA) haplotype with *p*-value = 0.0009, OR = 0.35, 95% CI = 0.19–0.65 with more controls of patients are associated with a reduced risk of developing RM.

There are currently just a few study papers on SOX2OT gene polymorphisms (Boraska et al., 2014). Rs9839776 is an intron variant in the SOX2OT gene associated with disease susceptibility in numerous studies (Tang et al., 2017). A genome-wide association study revealed that rs9839776 in SOX2OT is likely associated with anorexia nervosa susceptibility (Boraska et al., 2014). Likewise, SOX2OT expression levels correlated with the prognosis of patients with ovarian cancer, and SOX2OT enhanced cell proliferation and cell migration in ovarian cancer cells (Han et al., 2018). Interestingly, a similar study on breast cancer revealed that the SOX2OT SNP rs9839776 is substantially related to an increased risk of breast cancer and higher SOX2OT expression levels, indicating that rs9839776 promotes the beginning of breast cancer via changing SOX2OT expression (Tang et al., 2017). On the other hand, another study has identified a link between breast cancer and reproductive risk factors (Nafissi et al., 2018) involved in abortion (Erlandsson et al., 2003). In a study of SOX2OT rs9839776 lncRNA in the Chinese population at risk of developing RM, this variant was identified as a prognostic marker in the southern Chinese population (Fang et al., 2019). Despite the study conducted in China on the SOX2OT variant, we could not obtain a similar result for the rs9839776 variant in the Turkish Azeri population.

H19 is another lncRNA expressed only from the maternally inherited chromosome (Zemel et al., 1992) and is strongly involved in normal embryo development and placental cell differentiation (Nordin et al., 2014). H19 LncRNA is involved in endometriosis by acting on the H19/Let-7/IGF1R regulatory axis as ceRNA (Ghazal et al., 2015). Endometriosis patients have reduced H19 expression at the eutopic endometrium level, resulting in increased let-7 bioavailability. This allows for significant inhibition of Igf1r and, as a result, reduced IGF1 signaling, resulting in decreased stromal cell proliferation. Changes in the H19/Let-7/IGF1R-mediated regulation of endometrial cell proliferation may pose a promising mechanism for infertility in endometriosis patients (Ghazal et al., 2015). In a study of 230 cases and 240 controls performed on a population in southeast Iran on breast cancer, the H19 rs3741216 variant was significantly associated with a decreased risk of breast cancer (Hassanzarei et al., 2017). The lncRNA H19 rs3741216 variant also showed no significant association with RM.

According to bioinformatics analyses, the SNPs mentioned in Table 1 might be implicated in the ceRNA network. Competing endogenous RNA (ceRNA) can control other RNA transcripts by competing with shared microRNAs (Fig. 1). The ceRNA network connects coding and non-coding RNAs, particularly lncRNAs, via microRNA (miRNA)

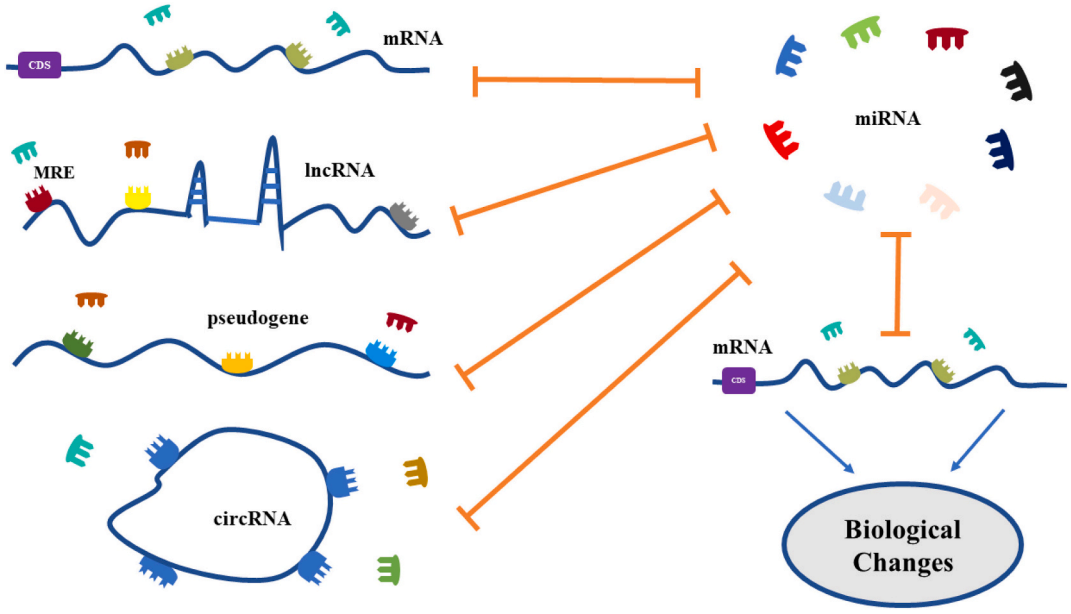

Fig. 1. Schematic network of the ceRNA model. By sponging shared miRNAs, all transcriptome components, such as lncRNAs, circRNAs, pseudogenes, and mRNAs, can co-regulate and interact. As a result, ceRNA dysregulation and biological changes can result from differentially expressed transcripts.

complementary sequences known as miRNA response elements (MREs) and creates a sizeable regulatory network in the transcriptome (Salmena et al., 2011). A study in 2021, using transcriptome sequencing, revealed several RM-related ceRNA regulatory axes (Huang et al., 2021). This study points to the potential of these regulatory axes in the pathogenesis of RM and forms a network of these regulatory axes that includes 31 lncRNAs and three mRNAs (NTNG2, GRIA1 and AQP1) and miR-210-5p (Huang et al., 2021). In this regard, in the present study, the polymorphisms studied on lncRNAs can potentially create action sites for miRNAs that can cause the formation of ceRNA axes. Rs619586 can be effective by establishing a site for miR-214-3p via the MALAT1/miR-214-3p/XBP1 ceRNA network (Duan et al., 2015) and rs3741216 can be effective by establishing a site for miR-130-3p via the H19/miR-130-3p/DSG1 ceRNA network (Li et al., 2017). According to Table 1, rs619586, rs664589, rs656605, and rs3741216 created the miRNAs' action sites on lncRNA, while rs3200401 caused the loss of the miRNAs' action sites, potentially disrupting a ceRNA network including the MALAT1 lncRNA, miR-1324, and miR-3661. Furthermore, a study revealed that miR-214-3p up-regulation is linked to RM (Al-Rubaye et al., 2021), whereas miR-122-3p up-regulation is linked to ectopic pregnancy (Sullivan-Pyke et al., 2018).

MALAT1 lncRNA rs591291, rs3200401, rs619586, rs664589, rs656605, SOX2OT lncRNA rs9839776, and H19 lncRNA rs3741216 variations were studied in patients with RM in northern Iran for the first time. Given that five of the seven lncRNA SNPs studied induce gain or loss of the miRNA action site, they may disrupt multiple ceRNA networks and modify many regulatory processes.

First, we propose that research similar to ours be done with greater statistical power in other demographics and geographical components. Second, more research may be conducted on SNPs that have the ability to induce or lose miRNA action sites and therefore become engaged in the ceRNA regulation network.

Finally, the MALAT1 rs619586 polymorphism was identified as an SNP with a substantial protective impact on the Iranian Azeri Turkish population. However, further research with big enough sample sizes and additional bioinformatics studies may be critical in identifying beneficial polymorphisms.

Author contribution statement

BMH, MR, MA, JG wrote the draft and revised it. MT designed and supervised the study. HS, PK, NL and NP performed the experiment. SD and HD collected the samples and data. All the authors read and approved the submitted version.

Declaration of Competing Interest

The authors declare they have no conflict of Interest.

Acknowledgment

This study was finacially supported by Shahid Beheshti University of Medical Sciences.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humgen.2022.201063.

References

Al-Rubaye, S., Ghaderian, S.M.H., Salehpour, S., Salmani, T., Vojdani, S., Yaseen, R., et al., 2021. Aberrant expression of BAX, MEG3, and miR-214-3P genes in recurrent pregnancy loss. In: Gynecological endocrinology : the official journal of the International Society of Gynecological Endocrinology, pp. 1–5.

Askarian-Amiri, M.E., Seyfoddin, V., Smart, C.E., Wang, J., Kim, J.E., Hansji, H., et al., 2014. Emerging role of long non-coding RNA SOX2OT in SOX2 regulation in breast cancer. PLoS One 9 (7), e102140.

Attia, J., Thakkinstian, A., D'Este, C., 2003. Meta-analyses of molecular association studies: methodologic lessons for genetic epidemiology. J. Clin. Epidemiol. 56 (4), 297–303.

Boraska, V., Franklin, C.S., Floyd, J.A., Thornton, L.M., Huckins, L.M., Southam, L., et al., 2014. A genome-wide association study of anorexia nervosa. Mol. Psychiatry 19 (10), 1085–1094.

Branch, D.W., Gibson, M., Silver, R.M., 2010. Clinical practice. Recurrent miscarriage. N. Engl. J. Med. 363 (18), 1740–1747.

Che, D., Yang, Y., Xu, Y., Fang, Z., Pi, L., Fu, L., et al., 2019. The lncRNA MALAT1 rs619586 G variant confers decreased susceptibility to recurrent miscarriage. Front. Physiol. 10, 385.

Duan, Q., Yang, L., Gong, W., Chaugai, S., Wang, F., Chen, C., et al., 2015. MicroRNA-214 is up-regulated in heart failure patients and suppresses XBP1-mediated endothelial cells angiogenesis. J. Cell. Physiol. 230.

Edwards, A.W., 2008. G. H. Hardy (1908) and hardy-Weinberg equilibrium. Genetics 179 (3), 1143–1150.
Erlandsson, G., Montgomery, S.M., Cnattingius, S., Ekbom, A., 2003. Abortions and breast cancer: record-based case-control study. Int. J. Cancer 103 (5), 676–679.
Fang, Z., Che, D., Qing, S., Li, Q., Men, H., Yuan, L., et al., 2019. The lncRNA SOX2OT rs9839776 C>T polymorphism indicates recurrent miscarriage susceptibility in a southern Chinese population. Mediat. Inflamm. 2019, 9684703.
Ghazal, S., McKinnon, B., Zhou, J., Mueller, M., Men, Y., Yang, L., et al., 2015. H19 lncRNA alters stromal cell growth via IGF signaling in the endometrium of women with endometriosis. EMBO Mol Med. 7 (8), 996–1003.
Han, L., Zhang, W., Zhang, B., Zhan, L., 2018. Long non-coding RNA SOX2OT promotes cell proliferation and motility in human ovarian cancer. Exp Ther Med. 15 (2), 2182–2188.
Hassanzarei, S., Hashemi, M., Sattarifard, H., Hashemi, S.M., Bahari, G., 2017. Genetic polymorphisms in long noncoding RNA H19 are associated with breast cancer susceptibility in Iranian population. Meta Gene. 14, 1–5.
Hefler, L.A., Tempfer, C.B., Unfried, G., Schneeberger, C., Lessl, K., Nagele, F., et al., 2001. A polymorphism of the interleukin-1beta gene and idiopathic recurrent miscarriage. Fertil. Steril. 76 (2), 377–379.
Huang, Y., Hao, J., Liao, Y., Zhou, L., Wang, K., Zou, H., et al., 2021. Transcriptome sequencing identified the ceRNA network associated with recurrent spontaneous abortion. BMC Med. Genet. 14 (1), 278.
Kaur, R., Gupta, K., 2016. Endocrine dysfunction and recurrent spontaneous abortion: an overview. International Journal of Applied and Basic Medical Research. 6 (2), 79–83.
Kling, C., Hedderich, J., Kabelitz, D., 2018. Fertility after recurrent miscarriages: results of an observational cohort study. Arch. Gynecol. Obstet. 297 (1), 205–219.
Korucuoglu, U., Biri, A.A., Konac, E., Alp, E., Onen, I.H., Ilhan, M.N., et al., 2010. Expression of the imprinted IGF2 and H19 genes in the endometrium of cases with unexplained infertility. Eur. J. Obstet. Gynecol. Reprod. Biol. 149 (1), 77–81.
Li, C.-x., Li, H.-g., Huang, L.-t., Kong, Y.-w., Chen, F.-y., Liang, J.-y., et al., 2017. H19 lncRNA regulates keratinocyte differentiation by targeting miR-130b-3p. Cell Death Dis. 8, e3174.
Liu, Y., Pan, S., Liu, L., Zhai, X., Liu, J., Wen, J., et al., 2012. A genetic variant in long non-coding RNA HULC contributes to risk of HBV-related hepatocellular carcinoma in a Chinese population. PLoS One 7 (4), e35145.
Liu, S., Xu, B., Yan, D., 2016. Enhanced expression of long non-coding RNA Sox2ot promoted cell proliferation and motility in colorectal cancer. Minerva Med. 107 (5), 279–286.
Ma, C., Shi, X., Zhu, Q., Li, Q., Liu, Y., Yao, Y., et al., 2016. The growth arrest-specific transcript 5 (GAS5): a pivotal tumor suppressor long noncoding RNA in human cancers. Tumour Biology 37 (2), 1437–1444.
McNamee, K., Dawood, F., Farquharson, R., 2012. Recurrent miscarriage and thrombophilia: an update. Current opinion in obstetrics & gynecology. 24 (4), 229–234.
Nafissi, N., Faraji, M., Hosseini, M., Shojaee, L., Ziaei, F., Akbari, M.E., et al., 2018. Relationships between reproductive risk factors for breast Cancer and tumor molecular subtypes. Asian Pacific journal of cancer prevention : APJCP. 19 (7), 1767–1770.
Ng, K.Y.B., Cherian, G., Kermack, A.J., Bailey, S., Macklon, N., Sunkara, S.K., et al., 2021. Systematic review and meta-analysis of female lifestyle factors and risk of recurrent pregnancy loss. Sci. Rep. 11 (1), 7081.
Nordin, M., Bergman, D., Halje, M., Engström, W., Ward, A., 2014. Epigenetic regulation of the Igf2/H19 gene cluster. Cell Prolif. 47 (3), 189–199.
Peng, R., Luo, C., Guo, Q., Cao, J., Yang, Q., Dong, K., et al., 2018. Association analyses of genetic variants in long non-coding RNA MALAT1 with breast cancer susceptibility and mRNA expression of MALAT1 in Chinese Han population. Gene. 642, 241–248.
Qin, X., Chen, Y., Chen, S., Liu, X., Zeng, W., Tian, F., et al., 2019. Plasmacytoma variant translocation 1 (PVT1) regulates trophoblast viability, proliferation, and migration and is downregulated in spontaneous abortion. American Journal of Reproductive Immunology (New York NY: 1989) 81 (1), e13048.
Rai, R., Regan, L., 2006. Recurrent miscarriage. Lancet (London, England). 368 (9535), 601–611.
Rull, K., Nagirnaja, L., Laan, M., 2012. Genetics of recurrent miscarriage: challenges, current knowledge, future directions. Front. Genet. 3, 34.
Salmena, L., Poliseno, L., Tay, Y., Kats, L., Pandolfi, P.P., 2011. A ceRNA hypothesis: the Rosetta stone of a hidden RNA language? Cell. 146 (3), 353–358.
Shi, X., Sun, M., Liu, H., Yao, Y., Song, Y., 2013. Long non-coding RNAs: a new frontier in the study of human diseases. Cancer Lett. 339 (2), 159–166.
Solé, X., Guinó, E., Valls, J., Iniesta, R., Moreno, V., 2006. SNPStats: a web tool for the analysis of association studies. Bioinformatics. 22 (15), 1928–1929.
Stephenson, M., Kutteh, W., 2007. Evaluation and management of recurrent early pregnancy loss. Clin. Obstet. Gynecol. 50 (1), 132–145.
Sullivan-Pyke, C., Sansone, S., Jou, J., Koelper, N.C., Stentz, N.C., Takacs, P., et al., 2018. Small non-coding RNA as a serum biomarker for non-viable and ectopic pregnancy. Fertil. Steril. 110, e258.
Tang, X., Gao, Y., Yu, L., Lu, Y., Zhou, G., Cheng, L., et al., 2017. Correlations between lncRNA-SOX2OT polymorphism and susceptibility to breast cancer in a Chinese population. Biomark. Med 11 (3), 277–284.
Thakkinstian, A., McElduff, P., D'Este, C., Duffy, D., Attia, J., 2005. A method for meta-analysis of molecular association studies. Stat. Med. 24 (9), 1291–1306.
Vitagliano, A., Noventa, M., Gizzo, S., 2017. Autoimmunity, systemic inflammation, and their correlation with repeated implantation failure and recurrent miscarriage: is chronic endometritis the missing piece of the jigsaw? American Journal of Reproductive Immunology (New York NY: 1989) 77 (1).
Wang, Y., Liu, H.Z., Liu, Y., Wang, H.J., Pang, W.W., Zhang, J.J., 2018a. Downregulated MALAT1 relates to recurrent pregnancy loss via sponging miRNAs. Kaohsiung J. Med. Sci. 34 (9), 503–510.
Wang, G., Li, Y., Peng, Y., Tang, J., Li, H., 2018b. Association of polymorphisms in MALAT1 with risk of coronary atherosclerotic heart disease in a Chinese population. Lipids Health Dis. 17 (1), 75.
Wang, Y., Liu, H.-Z., Liu, Y., Wang, H.-J., Pang, W.-W., Zhang, J.-J., 2019. Disordered p53-MALAT1 pathway is associated with recurrent miscarriage. Kaohsiung J. Med. Sci. 35 (2), 87–94.
Williams, Z., 2012. Inducing tolerance to pregnancy. N. Engl. J. Med. 367 (12), 1159–1161.
Wu, Y., Zhou, Y., Huan, L., Xu, L., Shen, M., Huang, S., et al., 2019. LncRNA MIR22HG inhibits growth, migration and invasion through regulating the miR-10a-5p/NCOR2 axis in hepatocellular carcinoma cells. Cancer Sci. 110 (3), 973–984.
Wu, L., Liu, Q., Fan, C., Yi, X., Cheng, B., 2021. MALAT1 recruited the E3 ubiquitin ligase FBXW7 to induce CRY2 ubiquitin-mediated degradation and participated in trophoblast migration and invasion. J. Cell. Physiol. 236 (3), 2169–2177.
Xue, Z., Cui, C., Liao, Z., Xia, S., Zhang, P., Qin, J., et al., 2018. Identification of LncRNA Linc00513 containing lupus-associated genetic variants as a novel regulator of interferon signaling pathway. Front. Immunol. 9, 2967.
Yang, L., Lin, C., Liu, W., Zhang, J., Ohgi, K.A., Grinstein, J.D., et al., 2011. ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. Cell. 147 (4), 773–788.
Zemel, S., Bartolomei, M.S., Tilghman, S.M., 1992. Physical linkage of two mammalian imprinted genes, H19 and insulin-like growth factor 2. Nat. Genet. 2 (1), 61–65.
Zeng, H., He, D., Xie, H., Zhao, Y., Peng, Z., Deng, H., et al., 2020. H19 regulates angiogenic capacity of extravillous trophoblasts by H19/miR-106a-5p/VEGFA axis. Arch. Gynecol. Obstet. 301 (3), 671–679.
Zhao, M., Wang, S., Li, Q., Ji, Q., Guo, P., Liu, X., 2018a. MALAT1: a long non-coding RNA highly associated with human cancers. Oncol. Lett. 16 (1), 19–26.
Zhao, K., Jin, S., Wei, B., Cao, S., Xiong, Z., 2018b. Association study of genetic variation of lncRNA MALAT1 with carcinogenesis of colorectal cancer. Cancer Manag. Res. 10, 6257–6261.
Zheng, Z., Liu, S., Wang, C., Han, X., 2018. A functional polymorphism rs145204276 in the promoter of long noncoding RNA GAS5 is associated with an increased risk of ischemic stroke. Journal of stroke and cerebrovascular diseases 27 (12), 3535–3541.