Bioinformatics Analysis of Noval SNPs Polymorphism with Haplotypes of Mucin-1 GeneAssociated with Infertility

Asmaa M. Salih Almohaidi[1], Abeer Abdulhameed Faisal[2],Mohanad Waheeb Mahdi[3] and Shumaila Ijaz[4]

[1]*Biomedical Application Branch, Institute of Laser Application for Postgraduate Study, University of Baghdad, 10071 Baghdad, Iraq*

[2]*Department of Biology, College of Science for Women, University of Baghdad, 10071 Baghdad, Iraq*

[3]*Department of Biology, College of Education for Pure Sciences, University of Diyala, 32001 Baqubah, Iraq*

[4]*School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, 518060 Shenzhen, China*

asmaams_bio@csw.uobaghdad.edu.iq, abeer.abd2202m@csw.uobaghdad.edu.iq, mohanad@uodiyala.edu.iq, shumailaijaz1996@yahoo.com

Keywords: Infertility, Polymorphism and Haplotype, Newly Snps, MUC1 Gene.

Abstract: Mucin one (*MUC1*) is a protein coding gene which located on long arm of chromosome 1 in region2 band 2 (1q22). The current novel study examines the relation of many new variation *MUC1* polymorphisms, haplotype and infertility in females. The study included one hundred and twenty women, sixty women with infertility and sixty healthy fertile women. This study analysis sequencing of specific segments (876 bp) at exon (3-6) of *MUC1* gene to examine polymorphism for new specific suggested SNPs and haplotype. The current study investigated fife new suggested SNPs in *MUC1* gene. The new suggested SNP 155187497 GG homozygous genotype showed a significant protective association with infertility with a 0.31 odds ratio while the GA heterozygous genotype showed a significant risky association with a 3.05 odds ratio. In contrast, the AA homozygous genotype exhibited nonsignificant risky effects with 3.05 odds ratio associated with infertility. These finding suggest that the G allele acted as a protective allele, while the A allele may be associated with the infertility. The other four new suggested SNPs 155187477, 155187668, 155187698 and 155187793 of *MUC1* gene showed no significant differences between infertile and fertile women. In addition, three suggested SNPs (155187668, 155187698 and 155187793) and two suggested SNPs (155187477 and 155187668) in the *MUC1* gene demonstrated significant linkage disequilibrium LD with one another.

1 INTRODUCTION

Mucins are expansive glycoproteins with a substantial molecular weight that offer lubrication to the epithelial surfaces of the respiratory, gastrointestinal, and reproductive systems [1]. The Mucin genes are a group of 24 genes, known as *MUC1* to *MUC24*, that belong to the same family [2]. One of them is *MUC1*. *MUC1* is the predominant mucin present in the endometrial epithelium of several species [3].

MUC1 gene is located on the long arm of chromosome 1 in region band 2(1q22), and comprises 11 exons/introns, it encodes for a 122102 Da protein that is made up of 1255 amino acids and a liner DNA [4]. This gene is responsible for producing a glycoprotein that is connected to the upper surface of the gastrointestinal epithelium. The glycoprotein is membrane-bound and contains a transmembrane domain. It has crucial functions in protecting epithelial surfaces from external irritants [5], [6].

MUC1 plays a significant role in the pathogenesis of immunity of diseases [7], specifically in the advancement of cancer and the spread of cancer cells to other parts of the body [8], [9]. The abnormal expression of MUC1 is linked to the progression, infiltration, and spread of several types of malignancies, such as colorectal cancer, liver cancer, lung cancer, and gastric cancer [7]-[10].

World Health Organization (WHO) defines infertility as a medical condition that occurs when a person is unable to produce a clinical pregnancy while engaging in frequent unprotected sexual intercourse for 12 months or more. There are multiple factors are significant contributors to infertility such as hormonal and genetics [11], [12]. Research has

identified multiple genetic variables that have a role in causing infertility among individuals in the population [13], [14].

There are previous local Iraqi studies that have confirmed that the *MUC1 gene* is related to infertility [15], [16], therefore present study tries to explore the relationship between *MUC1* gene polymorphism and haplotype with infertility. Despite extensive research on physiology and genetics, the underlying mechanisms of infertility remain incompletely comprehended. The present study aims to examine the relationship between infertility.

2 MATERIALS AND METHODS

2.1 Study Design

The current study involved 120 women, 60 women with infertility with an age range from (18 – to 45 years) as well as sixty healthy fertile women whose ages matched the patient group with at least one childbirth. The samples were collected from AL-Elwiya Educational Hospital, Rusafa Health Department, Iraqi Ministry of Health (permission number 117181, issued on June 8, 2023), and from the College of Science for Women/ University of Baghdad. The sample collection period extended from August 2023 until December 2023. Blood was collected during the early follicular phase (days 2-5 of the menstrual cycle). All females included in this study were within the reproductive age range and free from endocrine or other systemic diseases, while those with endocrine diseases and other diseases were excluded.

2.2 Collection of Blood Samples

Blood samples were collected from each female for both groups, as two ml of blood was added directly into an EDTA-containing tube for the genotyping study.

2.3 Genomic DNA Extraction and Genotyping

The DNA was extracted from the whole blood of the study group subjects by using the protocol in EasyPure® Blood Genomic DNA Kit (TransGen, biotech. EE121-01). The genotyping of the *MUC1* was carried out using polymerase chain reaction (PCR) and sequencing. The PCR cycles for the *MUC1* (876 bp segment), reaction began at 94°C for 1 minute, then 35 cycles of denaturation at 94°C for 30 seconds, annealing at 56°C for 30 seconds, extension at 72°C for 30 seconds and a final extension at 72°C for 5 minutes. PCR amplifications were detected in a total volume of 26 µL consisting of 6 µL genomic DNA, 4.5 µL D.W., 12.5 µL master mix, and 1.5 µL of each primer. The primers were designed by first author using Primer 3plus, V4, and double-checked by the University Code of Student Conduct (UCSC) programs, and with their reference sequences in the National Center for Biotechnology Information (NCBI) database. They were synthesized and lyophilized by Alpha DNA Ltd. (Canada). The primers were used in DNA sequencing of specific segments ((876 bp) at exon (3-6) of MUC1 gene which located on chromosome 1 q22 as shown in Table 1.

Table 1: Specific primer sequences for DNA sequencing.

Primer	Sequence (5′→3′ direction)	Templet length	Reference
MUC1 Human Sequencing Primer			
Forward	TCCCAGCACC GACTACTACC	876 bp	Design by first author
Reverse	CAGCTGCCCGT AGTTCTTTC		

2.4 DNA Sequencing

The purified PCR products of the analyzed *MUC1* gene were sent to Macrogen Company in Korea for DNA sequencing. Furthermore, the nucleotide sequences were compared to the information in the gene bank of the National Center for Biotechnology Information (NCBI) website databases using the BLAST search tool.

2.5 Statistical Analysis

The statistical significance of the P values ($P \leq 0.05$) was determined using Fisher's exact test, while the Odds Ratio was calculated through a particular $\chi 2$ analysis using the WINPEPI computer application (version 11.63), the odds ratio was used to estimate the likelihood of specific genotypes or alleles being associated with infertility. The Hardy-Weinberg equilibrium was assessed using a chi-squared test conducted using OEGE-Online Encyclopedia for Genetic Epidemiology studies [17].

3 RESULTS AND DISCUSSION

Within study samples, there is a variant in *MUC1* identified. This *MUC1* SNP was not reported in the National Center for Biotechnology Information (NCBI) dbSNP database[1].

3.1 Genotype Result

The region exon 3-6 (from position 155187333 to 155188213) of the *MUC1* gene were amplified under optimum conditions (Fig. 1) by using a specific primer, then amplified segment with product size 876bp measurement of the query (sequence results) and subject (gene sequence on NCBI). Table 2 shows the registered new position in nucleotide suggested SNPs (155187477 C < G, 155187497 A < G, 155187668 A < G, 155187698 T <C, and 155187793 C < A).

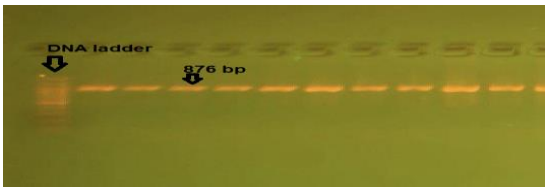

Figure 1: Gel electrophoresis for PCR product of *MUC1* gene (876bp) with DNA ladder on agarose gel concentration (2%) in (70 volt/cm2, 1 hour).

The present study is regarded as a groundbreaking investigation of the new suggested SNPs in the *MUC1* gene within the Iraqi population. This study examines the overall number of patients and control subjects, as well as the expected Frequencies of *MUC1* Genotypes Using Hardy-Weinberg Equilibrium for the expected frequencies of genotypes and the genotype frequencies based on the number of patients.

3.2 Polymorphism Suggested SNP 155187477 C > G

Table 3 shows the Expected Frequencies of *MUC1* Genotypes Using Hardy-Weinberg Equilibrium for the expected frequencies of genotypes. The distribution of suggested SNP 155187477 polymorphisms deviates from the Hardy-Weinberg Equilibrium, the observed genotype frequencies had significant differences than those Expected. This result may be because the control samples were gathered from individuals who are related, or it may indicate that this specific locus is subject to evolutionary selection in the Iraqi population. at the same time the patient's group may be under the effect of infertility which it deviates from HWE.

The GG genotype may be considered a common genotype in the Iraqi female population because the total observed record patients with control groups was 102 while the other genotype GC genotype recorded in 4 females while the CC genotype in 14 females.

Table 4 shows the comparison of the Genotype and Allele Frequencies of suggested SNP 155187477 polymorphism, there were no significant disparities in the distribution of genotypes and the frequency of alleles between the patients and control groups. The odds ratio for the GC genotype was 3.11, but for the GG genotype it was just 0.77, while the odds ratio for the CC genotype was 1.00, The findings indicate that females with the GC and CC genotypes may have a higher vulnerability to infertility, whereas the GG genotype acts as a protective factor, reducing this susceptibility. Additionally, the odds ratio for the C allele was 1.16, suggesting a possible association with the disease. Conversely, the odds ratio for the G allele was 0.86, showing a decrease in susceptibility and tend to be a protective effect.

3.3 Polymorphism Suggested SNP 155187497 A > G

The frequency of genotypes and alleles of suggested SNP 155187497 polymorphisms consistent with the Hardy–Weinberg Equilibrium among the patient and control group Table 5. The total observed of the GG genotype may be considered a common genotype in the Iraqi female population because the total observed record (patients with control groups) was 81 females.

Table 2: Sequencing ID at gene bank, score, expectations, and compatibility of DNA sequences obtained.

Score	Expect	Identities	Gaps	Strand
1507 bits (816)	0-0	828/833 (99%)	4/833 (0%)	Plus/Plus

Table 3: Expected frequencies of suggested SNP 155187477 Genotypes in *MUC1* gene using hardy-weinberg equilibrium for the expected frequencies of genotypes.

Group	Genotypes				χ^2
155187477		GG	GC	CC	
Control	Observed	52	1	7	51.21 NC
	Expected	45.93	13.12	0.93	
Patients	Observed	50	3	7	37.86 NC
	Expected	44.2	14.59	1.2	
Total observed		102	4	14	

NC: distribution is not consistent with Hardy Weinberg's law at the level of significance: $X^2 > 3.84$.

Table 4: Comparison of the genotype and allele frequencies of suggested SNP 155187477 polymorphism in *MUC1* gene between groups study.

Genotypes 155187477	Patients [n 60, (%)]	Control [n 60, (%)]	OR	(95% C.I.)	P value
GG	50 (83.33 %)	52 (86.67%)	0.77	0.28 to 2.1	0.61
GC	3 (5.00%)	1 (1.67%)	3.11	0.31 to 30.73	0.33
CC	7 (11.67%)	7 (11.67%)	1 (reference)	0.32 to 3.04	……
Allele G	103 (85.83%)	105 (88.33%)	0.86	0.41 to 1.82	0.70
Allele C	17 (14.17%)	15 (11.67%)	1.16	0.54 to 2.43	0.70

Non-significant P> 0.05, OR: Odd Ratio, CI: confidence interval.

Table 5: Expected Frequencies of suggested SNP 155187497 genotypes in *MUC1* gene using hardy-weinberg equilibrium for the expected frequencies of genotypes.

Group	Genotypes				χ^2
155187497		GG	GA	AA	
Control	Observed	48	12	0	0.74 C
	Expected	48.6	10.8	0.6	
Patients	Observed	33	26	1	2.67 C
	Expected	35.26	21.46	3.26	
Total observed		81	38	1	

C: Distribution consistent with Hardy Weinberg's law at the level of significance: $X^2 < 3.84$.

Table 6: Comparison of the genotype and Allele Frequencies of suggested SNP 155187477 polymorphism in *MUC1* gene between groups study.

Genotypes 155187497	Patients [n 60, (%)]	Control [n 60, (%)]	OR	(95% C.I.)	P value
GG	33 (55.00%)	48 (80.00%)	0.31	0.13 to 0.64	0.004**
GA	26 (43.33%)	12 (20.00%)	3.05	1.35 to 6.89	0.007*
AA	1 (1.67%)	0	3.05	0.12 to 76.39	0.49
Allele G	92 (75.83%)	108 (90.00%)	0.35	0.17 to 0.72	0.005**
Allele A	28 (24.17%)	12 (10.00%)	2.73	1.31 to 5.69	0.007*

Non-significant P> 0.05, significant differences $P \leq 0.05$*, high significant differences $P \leq 0.005$**, OR: Odd Ratio, CI: confidence interval.

From Table 6, the suggested SNP 155187497 polymorphism. There were high significant differences in the GG and significant differences in GA while no significant differences in the AA genotypes distribution and allele frequency between patients and control groups and the two genotypes GA and AA have odds ratio > 1 (3.05 for each of them), thus they are acting as risk factor. But GG genotype the odds ratio was 0.31 therefore GG genotype acts as a protective factor. The distribution of allele G shows highly significant differences between the groups study (p= 0.005), while the distribution of allele A shows significant differences between patients and control groups (p= 0.007). The

odds ratio for the G allele was 0.35 while the odds ratio for the A allele was 2.73 indicating that the A allele could have the susceptibility to association with the disease but the G allele decreases this susceptibility by acting as a protective agent against infertility.

3.4 Polymorphism Suggested SNP 155187668 A > G

Table 7 shows the Expected Frequencies of Genotypes in *MUC1* Genotypes Using Hardy-Weinberg Equilibrium for the expected frequencies of genotypes. The frequency of genotypes and alleles of suggested SNP 155187668 polymorphisms goes with the Hardy–Weinberg Equilibrium among the control and patients' groups. The total observed of the GG genotype may be considered a common genotype in the Iraqi female population because the total observed record (patients with control groups) was 90 females while 30 females have GA.

The suggested SNP 155187668 polymorphism has been clarified in Table 8, there were no significant disparities in the distribution of genotypes and the frequency of alleles between the patients and control groups. The odds ratio for the GA genotype was 1.16, but for the GG genotype it was just 0.84, this result indicates that females with the GA genotype may have a higher vulnerability to infertility, whereas the GG genotype acts as a protective factor, reducing this susceptibility. Additionally, the odds ratio for the A allele was 1.16, suggesting a possible association with the disease. Conversely, the odds ratio for the G allele was 0.85, showing a decrease in susceptibility and tend to be a protective effect to infertility.

Table 7: Expected frequencies of suggested SNP 155187668 Genotypes in *MUC1* gene using hardy-weinberg equilibrium for the expected frequencies of genotypes.

Group	Genotypes				χ^2
155187668		GG	GA	AA	
Control	Observed	46	14	0	1.04 C
	Expected	46.81	12.36	0.81	
Patients	Observed	44	16	0	1.42 C
	Expected	45.06	13.86	1.06	
Total observed		90	30	0	

C: Distribution consistent with Hardy Weinberg's law at the level of significance: $X^2 < 3.84$, 3.84.

Table 8: Comparison of the genotype and allele frequencies of suggested SNP 155187668 polymorphism in *MUC1* gene between groups study.

Genotypes 155187668	Patients [n 60, (%)]	Control [n 60, (%)]	OR	(95% C.I.)	P value
GG	44 (73.33%)	46 (76.67%)	0.84	0.36 to 1.92	0.67
GA	16 (26.67%)	14 (23.33%)	1.16	0.52 to 2.73	0.67
AA	0	0	…….	……	……...
Allele G	104 (86.67%)	106 (88.33%)	0.85	0.39 to 1.84	0.69
Allele A	16 (13.33%)	14 (11.67%)	1.16	0.54 to 2.5	0.69

Non-significant P> 0.05, OR: Odd Ratio, CI: confidence interval.

Table 9: Expected frequencies of suggested SNP 155187698 Genotypes in *MUC1* gene using hardy-weinberg equilibrium for the expected frequencies of genotypes.

Group	Genotypes				χ^2
155187698		CC	CT	TT	
Control	Observed	49	11	0	0.61 C
	Expected	49.50	9.99	0.5	
Patients	Observed	51	9	0	0.39 C
	Expected	51.33	8.32	0.33	
Total observed		100	20	0	

C: Distribution consistent with Hardy Weinberg's law at the level of significance: $X^2 < 3.84$, 3.84.

Table 10: Comparison of the genotype and allele frequencies of suggested SNP 155187698 polymorphism in *MUC1* gene between groups study.

Genotypes 155187698	Patients [n 60, (%)]	Control [n 60, (%)]	OR	(95% C.I.)	P value
CC	51 (85.00%)	49 (81.67%)	1.27	0.48 to 3.37	0.62
CT	9 (15.005)	11 (18.33%)	0.78	0.29 to 2.06	0.62
TT	0	0	……	……	……..
Allele C	111 (92.50%)	109 (90.83%)	1.24	0.49 to 3.12	0.64
Allele T	9 (7.50%)	11 (9.17%)	0.8	0.32 to 2.01	0.64

Non-significant P> 0.05, OR: Odd Ratio, CI: confidence interval.

Table 11: Expected Frequencies of suggested SNP 155187793 Genotypes in *MUC1* gene using hardy-weinberg equilibrium for the expected frequencies of genotypes.

Group	Genotypes				χ^2
155187793		AA	AC	CC	
Control	Observed	33	26	1	2.67 C
	Expected	35.26	21.46	3.26	
Patients	Observed	35	22	3	0.04 C
	Expected	35.26	21.46	3.26	
Total observed		68	48	4	

C: Distribution consistent with Hardy Weinberg's law at the level of significance: $X^2 < 3.84$

Table 12: Comparison of the Genotype and Allele Frequencies of suggested SNP 155187793 polymorphism in *MUC1* gene between groups study.

Genotypes 155187793	Patients [n 60, (%)]	Control [n 60, (%)]	OR	(95% C.I.)	P value
AA	35 (58.33%)	33 (55.00%)	1.14	0.55 to 2.35	0.71
AC	22 (36.67%)	26 (43.33%)	0.76	0.36 to 1.57	0.46
CC	3 (5.00%)	1 (1.67%)	3.11	0.31 to 30.73	0.33
Allele A	92 (76.67%)	92 (76.67%)	1.00 (Reference)	0.54 to 1.81	……
Allele C	28 (23.33%)	28 (23.33%)	1.00 (Reference)	0.54 to 1.81	……

Non-significant P> 0.05, OR: Odd Ratio, CI: confidence interval.

3.5 Polymorphism Suggested SNP 155187698 T >C

The frequency of genotypes and alleles of suggested SNP 155187698 polymorphisms goes with Hardy–Weinberg Equilibrium among the control and patients' groups Table 9. The total observed of the CC genotype may be considered a common genotype in the Iraqi female population because the total observed record in the two groups was 100 females.

Table 10 shows the suggested SNP 155187698 polymorphism. There were no significant disparities in the distribution of genotypes and the frequency of alleles between the patients and control groups. The odds ratio for the CT genotype was 0.78, but for the CC genotype it was 1.27, the findings indicate that females with the CC genotype may have a higher vulnerability to infertility, whereas the CT genotype acts as a protective factor, reducing this susceptibility. Additionally, the odds ratio for the C allele was 1.24, suggesting a possible association with the disease. Conversely, the odds ratio for the T allele was 0.8, showing a decrease in susceptibility and a protective effect.

3.6 Polymorphism Suggested SNP 155187793 C > A

Finally, in Table 11 the frequency of genotypes and alleles of suggested SNP 155187793 polymorphisms

goes with Hardy–Weinberg Equilibrium between both two groups. The total observed AA genotype could be considered a common genotype in the Iraqi female population because the total observed number of patients and control groups was 68 while the other genotypes were less than.

The suggested SNP 155187793 polymorphism is shown in Table 12. There were no significant differences in the distribution of genotypes and the frequency of alleles between the patients and control groups. The odds ratio for the AC genotype was 0.76, but for the AA and CC genotypes, it was (1.14 and 3.11) respectively. The findings indicate that females with the AA and CC genotypes could have susceptibility to association with the disease, whereas the AC genotype acts as a protective factor, reducing this susceptibility. Additionally, the odds ratio for both the A allele and C allele was 1.00, suggesting a possible association with the disease.

MUC1 gene showed highly significant differences in variation suggested SNP 155187497 between the fertile and infertile women so this gene may relate to the infertile Iraqi population. This result is in line with a previous Iraqi study which showed that the SNP rs1611770 in the *MUC1* gene was significantly association with infertility[18]. Also, this study is in agreement with a prior investigation conducted in Baghdad city that found the polymorphism of HNF1A SNP rs2464196 polymorphism G>A on the risk of Polycystic Ovary Syndrome in a Sample of Iraqi Women. [19]. Comparison of the differences in present study results regarding newly suggested SNPs in the *MUC1* gene (155187477 C < G, 155187497 A < G, 155187668 A < G, 155187698 T<C and 155187793 C < A) genotypes frequencies distribution in patients and control with previous study results in infertility may be related to racial and ethnic differences in *MUC1* gene polymorphisms and in exposure to mutagens that increase or decrease infertility risk as well as the interaction of *MUC1* gene with additional genomic mutations to causes development infertility [20].

The current study provides genetic information confirming the role of *MUC1* in infertility. *MUC1* codes for a large epithelial apical surface glycoprotein that serves as a barrier to embryo implantation by functioning as an anti-adhesive molecule on the uterine epithelium. MUC1 expression must be meticulously managed for effective implantation when the expression is inadequate amounts can impede the implantation [21], same result if the expression of *MUC1* is high level prevents trophoblast adhesion to the endometrial surface, which is critical for implantation [22]. The author's previous study has proposed the expression of *MUC1* in the uterus of women diagnosed with failure embryo implantation [22], [23] This has led to the hypothesis that altered expression of *MUC1* may predispose women to infertility.

Furthermore, the current study emphasizes there was no significant difference between the suggested SNPs (155187477, 155187668, 155187698, and 155187793) of *MUC1* gene polymorphism between the fertile and infertile women. This result agrees with a previous Iraqi study which showed that the distribution of the SNPs in *MUC1* rs139620330 and rs145224844 genotype and allele frequency are no significant differences between the fertile and the infertile groups [16]. In addition, this result deviates from previous studies about *VDR* gene polymorphism and its roles in Arabian population [24]. However, some studies have succeeded in showing the relationship between the MUC1 gene polymorphisms and increased risks of cancer [25], [26] in different populations [27].

3.7 Haplotype and Linkage Disequilibrium Analyses

Haplotype and linkage disequilibrium (LD) analysis of newly suggested SNP (155187477 C < G, 155187497 A < G, 155187668 A < G, 155187698 T <C and 155187793 C < A) in the *MUC1* gene was performed by using SHESIS plus software to investigate their association with increased risk of infertility. The term LD was initially introduced to denote alterations in genetic variation within a population over time. Recombination events will disrupt regions of contiguous chromosomes, leading to linkage equilibrium of alleles, or the complete independence of each allele, over generations in a population engaged in random mating [28].

The current study shows the association between some suggested SNPs in *MUC1* gene polymorphisms has been analyzed by using linkage disequilibrium (LD). Figure 2 shows that two *MUC1* suggested SNPs (155187668 and 155187698) have a high linkage disequilibrium (LD) with a D’ of 0.99. Additionally, other tested genetics suggested SNPs (155187668 and 155187793) have moderate linkage disequilibrium (LD) with a D’ of 0.70. While the other suggested SNPs (155187477 and 155187668) have low linkage disequilibrium (LD) with a D’ of 0.63. Thus, these findings suggest that three suggested SNPs (155187668, 155187698, and 155187793) and two suggested SNPs (155187477 and 155187668) constitute a single haplotype block. Current analysis

can lead to the conclusion that present variation in MUC1 could be related to infertility and linked with each other which can increase susceptibility to infertility.

Moreover, polymorphism of three suggested SNPs (155187668, 155187698, and 155187793) and two suggested SNPs (155187477 and 155187668) in the current study consistent with the previous study about TET2 gene which emphasized that there was strong linkage disequilibrium between TET2 SNPs rs34402524 and rs2454206 which lead to being more strongly associated with an increased risk of chronic myeloid leukemia CML [29]. While disagreeing with the author's result of the previous study about the *MUC* 17 gene there were no significant differences between the SNPs rs11979706, rs10246021,rs6966570, and rs4729655 on MUC17gene and progress endometriosis [30].

Table 13 summarizes *MUC1* gene haplotype frequencies and risk association to infertility, the current study showed there were no significant differences (P> 0.05) in haplotype between the patients and control groups. The presence of haplotypes. C G G C A (OR = 0.68, 95% Cl: 0.27~2.33, P = 0.68), G G A C A (OR = 0.59, 95% Cl: 0.22~1.59, P = 0.29), G G G C A (OR = 0.97, 95% Cl: 0.56~1.67, P = 0.92) and G G G C C (OR = 0.52, 95% Cl: 0.22~1.21, P = 0.21) in MUC1 gene were significantly associated with a decrease in risk of infertility. While, the haplotype C G G C C (OR = 5.27, 95% Cl: 0.67~41.37, P= 0.08), G A A C A (OR = 2.59, 95% Cl: 0.44~15.08, P = 0.27), G A G C A (OR = 1.45, 95% Cl: 0.44~4.74, P = 0.53) and G A G C C (OR = 3.15, 95% Cl: 0.72~13.83, P = 0.11), the most prevalent haplotypes among patients, so tend to act as a risk factor because it increases susceptibility to infertility in females.

However, each SNP may significantly influence phenotypic alterations, while an individual SNP in DNA may not directly affect protein structure or function, although they possess the capacity to modify gene expression levels. The isolated impact of a single regulatory SNP may be negligible alone; however, when combined with other SNPs as a collective unit that governs gene expression or function, the cumulative effect might alter the phenotype by modifying gene structure and expression [31]. So, there are multiple previous studies showed that the association between some SNPs in different genes with development the of different diseases [32]. The Potential of Single Nucleotide Polymorphisms as Biomarkers and Their Association with the Increased Risk of different Diseases [33], [34] and modify immuno response [35] all that could related with Haplotype [36], [37], Haplotype is crucial for various genetic analyses from sequencing [38] and a tool is a powerful resource for analyzing hybrid or recombinant diploid or polyploid genomes and identifying parental ancestry for sub-genomic regions [39], [40]. Therefore, inherited current haplotype could create an infertile new offspring.

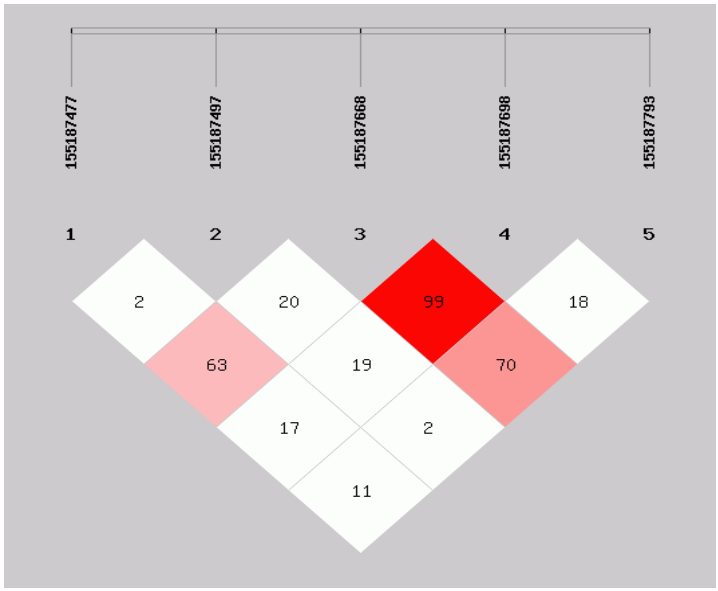

Figure 2: Linkage disequilibrium estimated between suggested SNPs in mucin one gene. The blocks indicate haplotype blocks and the text above the horizontal numbers is the suggested SNP names. The values in the red boxes are pair-wise suggested SNP correlations (D′). The red-to-white gradient reflects higher to lower LD values.

Table 13: Analysis of haplotypes in the *MUC1* gene.

Haplotypes	Patients Frequency	Control frequency	Fisher's p	Odds ratio	95%CI
C G G C A	6.42(0.05)	8.19(0.07)	0.68	0.80	[0.27~2.33]
C G G C C	5.41(0.04)	1.11(0.01)	0.08	5.27	[0.67~41.37]
G A A C A	4.35(0.04)	1.79(0.02)	0.27	2.59	[0.44~15.08]
G A G C A	6.90(0.06)	5.02(0.04)	0.53	1.45	[0.44~4.74]
G A G C C	7.09(0.06)	2.45(0.02)	0.11	3.15	[0.72~13.83]
G G A C A	7.01(0.06)	11.64(0.09)	0.29	0.59	[0.22~1.59]
G G G C A	57.61(0.48)	60.47(0.50)	0.92	0.97	[0.56~1.67]
G G G C C	9.38(0.08)	17.27(0.14)	0.12	0.52	[0.22~1.21]

Frequency<0.03 in both patients and control groups has been dropped, OR: Odd Ratio, CI: confidence interval.

4 CONCLUSIONS

This novel study identified a significant association between the new NO. 155187497 genotype variations GG and GA of the MUC1 gene and infertility within a sample of the Iraqi population. Additionally, the polymorphisms of three suggested SNPs (155187668, 155187698, and 155187793) and two suggested SNPs (155187477 and 155187668) in the MUC1 gene demonstrated nonsignificant linkage disequilibrium with one another, except suggested SNPs (155107668 and 15518698) show high LD, therefore suggesting both of them may contribute to infertility when inherited together, potentially increasing susceptibility in future female's generations. Additional even that the all SNPs showed non-significant linked among them but needs more studies to clarify their linkage as one block, because the haplotypes emphasize on the role of haplotype block as protective or risky factors. Finally, this whole work led to the hypothesis that altered expression of MUC1 may associate with current Polymorphism and haplotypes, so may predispose women to infertility. However, MUC1could be a novel marker for susceptible for infertility, which will shed a spark light on using it to treat the endometrium problem related with infertility and waring from familial infertility inherited.

REFERENCES

[1] M. T. Borchers, M. P. Carty, and G. D. Leikauf, "Regulation of human airway mucins by acrolein and inflammatory mediators," American Journal of Physiology-Lung Cellular and Molecular Physiology, vol. 276, no. 4, pp. L549-L555, 1999.

[2] S. Kaur, S. Kumar, N. Momi, A. R. Sasson, and S. K. Batra, "Mucins in pancreatic cancer and its microenvironment," Nature Reviews Gastroenterology & Hepatology, vol. 10, no. 10, pp. 607-620, 2013.

[3] S. Yonezawa, M. Goto, N. Yamada, M. Higashi, and M. Nomoto, "Expression profiles of MUC1, MUC2, and MUC4 mucins in human neoplasms and their relationship with biological behavior," Proteomics, vol. 8, no. 16, pp. 3329-3341, 2008.

[4] P. Liu and M. Zeng, "Role of MUC1 rs4072037 polymorphism in gastric cancer: a meta-analysis," International Journal of Clinical and Experimental Pathology, vol. 13, no. 3, pp. 465-470, 2020.

[5] S. J. Gendler, "MUC1, the renaissance molecule," Journal of Mammary Gland Biology and Neoplasia, vol. 6, pp. 339-353, 2001.

[6] H. Sun, X. Wu, F. Wu, Y. Li, Z. Yu, X. Chen, et al., "Associations of genetic variants in the PSCA, MUC1 and PLCE1 genes with stomach cancer susceptibility in a Chinese population," 2015.

[7] Y. H. Sheng, J. M. Davies, R. Wang, K. Y. Wong, R. Giri, Y. Yang, et al., "MUC1-mediated macrophage activation promotes colitis-associated colorectal cancer via activating the interleukin-6/signal transducer and activator of transcription 3 axis," Cellular and Molecular Gastroenterology and Hepatology, vol. 14, no. 4, pp. 789-811, 2022.

[8] Y. Lan, W. Ni, and G. Tai, "Expression of MUC1 in different tumours and its clinical significance," Molecular and Clinical Oncology, vol. 17, no. 6, pp. 1-10, 2022.

[9] Y. Yin, C. Yang, J. Xu, Y. Luo, Q. Xia, and K. He, "MUC1 promotes lung metastases of liver cancer by impairing anti-tumor immunity," Discover Oncology, vol. 14, no. 1, p. 18, 2023.

[10] Y. Kim, R. Pecha, T. Keihanian, M. Mercado, S. Pena-Munoz, K. Lang, et al., "MUC1 expressions and its prognostic values in US gastric cancer patients," Cancers, vol. 15, p. 998, 2023.

[11] M. W. M. Alzubadiy, A. M. S. Almohaidi, A. A. Sultan, and L. Q. Abdulhameed, "Evaluation of E-selectin rs5367 C/T polymorphism in Iraqi diabetic foot patients," Journal of Physics: Conference Series, vol. 1294, p. 062021, 2019, doi: 10.1088/1742-6596/1294/6/062021.

[12] R. A. Majeed, A. F. Shihab, and A. H. Al-Assei, "The association of T45G polymorphism in the adiponectin gene with some hormonal parameters in Iraqi women with polycystic ovary syndrome," Medico-Legal Update, vol. 20, no. 2, pp. 775-781, 2020.

[13] M. Sharief, "Correlation of estrogen receptor alpha serum level with gene polymorphism and its effect on women with unexplained infertility, Basra, Iraq," Archives of Razi Institute, vol. 78, no. 2, p. 775, 2023.

[14] J. Liu, Z. Tan, J. He, T. Jin, Y. Han, L. Hu, et al., "Two novel mutations in PADI6 and TLE6 genes cause female infertility due to arrest in embryonic development," Journal of Assisted Reproduction and Genetics, vol. 38, no. 6, pp. 1551-1559, 2021.

[15] R. H. Saeed, A. M. S. Almohaidi, and I. J. Hammadi Al-Janabi, "MUC1 gene expression in infertile Iraqi Arab women," Biochemical & Cellular Archives, vol. 21, no. 2, 2021.

[16] S. F. Mirza, R. H. Saeed, A. M. S. Almohaidi, and I. J. H. Al-Janabi, "Evaluation mucin 1 polymorphism and expression with infertility in Iraqi females," RES Militaris, vol. 12, no. 2, pp. 6916-6927, 2022.

[17] S. Rodriguez, T. R. Gaunt, and I. N. Day, "Hardy-Weinberg equilibrium testing of biological ascertainment for Mendelian randomization studies," American Journal of Epidemiology, vol. 169, no. 4, pp. 505-514, 2009.

[18] I. H. D. Al-Janabi, A. M. S. Almohaidi, and I. J. H. Al-Janabi, "The rs1611770 SNP polymorphism and expression of mucin one in infertile females in Baghdad," Baghdad Science Journal, 2024.

[19] E. A. Toama and H. A. Jasim, "The effect of HNF1A gene polymorphism on risk of polycystic ovary syndrome in a sample of Iraqi women," Iraqi Journal of Biotechnology, vol. 21, no. 2, pp. 96-102, 2022.

[20] J. H. Kim, T. H. Kim, Y. S. Kim, W. C. Jang, A. Ryu, J. Y. Hwang, et al., "Mucin gene polymorphisms are associated with endometriosis in Korean women," Archives of Gynecology and Obstetrics, vol. 301, pp. 801-807, 2020.

[21] N. Dharmaraj, P. Chapela, M. Morgado, S. Hawkins, B. Lessey, S. Young, et al., "Expression of the transmembrane mucins, MUC1, MUC4 and MUC16, in normal endometrium and in endometriosis," Human Reproduction, vol. 29, no. 8, pp. 1730-1738, 2014.

[22] L. W. Francis, S. N. Yao, L. C. Powell, S. Griffiths, A. Berquand, T. Piasecki, et al., "Highly glycosylated MUC1 mediates high affinity L-selectin binding at the human endometrial surface," Journal of Nanobiotechnology, vol. 19, p. 1, 2021.

[23] N. Pujol Gualdo, EBR Team, R. Mägi, and T. Laisk, "Genome-wide association study meta-analysis supports association between MUC1 and ectopic pregnancy," medRxiv, 2022.

[24] A. Almohaidi, R. Saeed, S. Mirza, and E. Hassan, "Polymorphism of genes in Iraqi females with type 2 diabetes mellitus," Proceedings of International Conference on Applied Innovation in IT, 2025, pp. 825-832, doi: 10.25673/120611.

[25] N.-L. T. Nguyen, N.-D. T. Dang, Q.-H. Dang, V.-C. Tran, H.-L. Vo, M. Yamaguchi, et al., "Polymorphism of MUC1 gene in Vietnamese gastric cancer patients: a multicenter case-control study," Frontiers in Oncology, vol. 11, p. 694977, 2021.

[26] N.-L. T. Nguyen, N. D. T. Dang, Q. V. Van Vu, A. K. Dang, and T.-V. Ta, "A model for gastric cancer risk prediction based on MUC1 polymorphisms and health-risk behaviors in a Vietnamese population," In Vivo, vol. 37, no. 5, pp. 2347-2356, 2023.

[27] R. Shekarriz, H. Jabbari, R. Alikhani, R. Alizadeh-Navaei, and M. B. Hashemi-Soteh, "Association between MUC1 rs4072037 polymorphism and Helicobacter pylori in patients with gastric cancer," Caspian Journal of Internal Medicine, vol. 15, no. 1, pp. 132-137, 2024.

[28] G. Hettiarachchi and A. A. Komar, "GWAS to identify SNPs associated with common diseases and individual risk: genome wide association studies (GWAS) to identify SNPs associated with common diseases and individual risk," in Single Nucleotide Polymorphisms: Human Variation and a Coming Revolution in Biology and Medicine, Springer, 2022, pp. 51-76.

[29] N. H. Ismail, I. A. Abdulhassan, and A. H. M. Al-Faisal, "The association of TET2 gene polymorphisms (rs34402524 and rs2454206) and their haplotypes with response to treatment in chronic myeloid leukemia patients," Iraqi Journal of Science, 2024.

[30] M. Qiao, H. Zhang, Y. Xue, and L. Yang, "Relationship between MUC17 gene polymorphisms and endometriosis in Central Plains Chinese women," Clinical and Experimental Obstetrics & Gynecology, vol. 49, no. 10, p. 225, 2022.

[31] J. Liu, R. Xing, J. Shao, and S. Jiao, "Relationship between MUC4 variants and metastatic recurrence in colorectal cancer," International Journal of General Medicine, 2023, pp. 5077-5087.

[32] A. A. Ahmed and A. H. Ad'hiah, "Interleukin-37 gene polymorphism and susceptibility to coronavirus disease 19 among Iraqi patients," Meta Gene, vol. 31, p. 100989, 2022.

[33] H. Wolski, G. Kurzawińska, M. Ożarowski, A. E. Mrozikiewicz, K. Drews, T. M. Karpiński, et al., "Vitamin D receptor gene polymorphisms and haplotypes in the etiology of recurrent miscarriages," Scientific Reports, vol. 11, no. 1, p. 4646, 2021.

[34] P. Salehipour, F. Rezagholizadeh, M. Mahdiannasser, R. Kazerani, and M. H. Modarressi, "Association of OLR1 gene polymorphisms with the risk of coronary artery disease: a systematic review and meta-analysis," Heart & Lung, vol. 50, no. 2, pp. 334-343, 2021.

[35] M. T. AbdelGhafar, R. A. El-Kholy, T. A. Elbedewy, A. A. Allam, R. A. E. Eissa, S. M. Samy, et al., "Impact of CD40 gene polymorphisms on the risk of immune thrombocytopenic purpura," Gene, vol. 736, p. 144419, 2020.

[36] M. J. Sullivan, N. L. Bachmann, P. Timms, and A. Polkinghorne, "HapFlow: visualizing haplotypes in sequencing data," Bioinformatics, vol. 32, no. 3, pp. 441-443, 2016, doi: 10.1093/bioinformatics/btv551.

[37] B. D. P. Sitinjak, N. Murdaya, T. A. Rachman, N. Zakiyah, and M. I. Barliana, "The potential of single nucleotide polymorphisms (SNPs) as biomarkers and their association with the increased risk of coronary heart disease: a systematic review," Vascular Health and Risk Management, vol. 19, pp. 289-301, 2023, doi: 10.2147/VHRM.S405039.

[38] R. A. Farrer, "Haplotype Tools: a toolkit for accurately identifying recombination and recombinant genotypes," BMC Bioinformatics, vol. 22, no. 560, pp. 1-15, 2021, doi: 10.1186/s12859-021-04473-1.

[39] Z. Zhang, B. Zhu, Y. Luo, J. Shi, S. Lian, J. Hao, et al., "Direct inference of haplotypes from sequencing data," Bioinformatics Advances, vol. 00, no. 195, pp. 1-8, 2025, doi: 10.1093/bioadv/vbaf195.

[40] M. R. Hoehe, "Haplotypes and the systematic analysis of genetic variation in genes and genomes," Pharmacogenomics, vol. 4, no. 5, pp. 547-570, 2005, doi: 10.2217/14622416.4.5.547.