Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis

By

Ameema Khan

00000364804

MS Healthcare Biotechnology

Supervisor

Prof. Dr. Peter John

Department of Healthcare Biotechnology

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

H-12, Islamabad, Pakistan

2023

Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis

A thesis submitted in partial fulfillment of the requirement for the degree of Master of Science

in

Healthcare Biotechnology

By

Ameema Khan

00000364804

Supervisor

Prof. Dr. Peter John

Department of Healthcare Biotechnology

Atta-ur-Rahman School of Applied Biosciences (ASAB)

National University of Sciences and Technology (NUST)

H-12, Islamabad, Pakistan

2023

FORM TH-4

National University of Sciences & Technology

MS THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by:

Student Name: **Ameema Khan** Reg No: **00000364804**

Titled: **Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis** be accepted in partial fulfillment of the requirements for the award of MS degree in Healthcare Biotechnology with (A grade).

Examination Committee Members

1. Name: Prof. Dr. Sobia Manzoor Signature:

2. Name: Dr. Maria Shabbir Signature:

3. Name: Prof. Dr. Muhammad Ansar Signature:

Supervisor's Name: Prof. Dr. Peter. John Signature:

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Date: 25/8/2023

Co-supervisor's name: Prof. Dr. Attya Bhatti Signature:

Prof. Dr. Attya Bhatti
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Date: 25/8/2023

Date: 07/09/2023

DR. HUSSAIN M. WAHEDI
Associate Professor
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Head of Department

COUNTERSINGED

Date: 07/09/2023

Principal
Atta ur Rahman School of
Applied Biosciences (ASAB)
NUST Islamabad

Dean/Principal

Copyright Statement

Thesis Acceptance Certificate

Certified that the final copy of MS thesis entitled **"Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis"** submitted by **Ms. Ameema Khan**, (Registration No: 00000364804), of Atta-ur-Rahman School of Applied Biosciences (ASAB) has been verified by undersigned, found complete in all respects as per NUST status/regulations and is accepted as partial fulfillment for the award of MS degree.

Supervisor: _____

Prof. Dr. Peter John

Atta-ur-Rahman School of Applied Biosciences

National University of Sciences and Technology

Dated: 25/8/2023

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Head of Department: _____

For

Prof. Dr. Sobia Manzoor

Atta-ur-Rahman School of Applied Biosciences

National University of Sciences and Technology

Dated: 07/09/2023

DR. HUSSAIN M. WAHEDI
Associate Professor
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

Principal: _____

Prof. Dr. Muhammad Asghar

Atta-ur-Rahman School of Applied Biosciences

National University of Sciences and Technology

Dated: 07/09/2023

Principal
Atta ur Rahman School of
Applied Biosciences (ASAB)
NUST Islamabad

Certificate For Plagiarism

It is certified that MS thesis titled **"Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis"** by Ameema Khan has been examined.

We undertake the follows:

i. Thesis has significant new work/knowledge as compared to already published or are under consideration to be published work. No sentence, equation, diagram, table, paragraph, or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.

ii. The work presented is original and the own work of the author (i.e., there is no plagiarism). No ideas, process, results, or words of others have been presented as own work.

iii. There is no fabrication of data or results, which have been compiled/ analyzed.

iv. There is no falsification by manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

v. The thesis has been checked using TURNITIN (copy of original report attached) and found within limits as per HEC plagiarism policy and instructions.

Signature of Student: _____

Ameema Khan

Signature of Supervisor: _____ 25/8/2023

Prof. Dr. Peter John

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied
Biosciences (ASAB), NUST Islamabad

Declaration

I certify that this research work titled **"Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis"** is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged/referred to.

Signature of Student

Ameema Khan

00000364804

Dedicated to my parents who always have my back.

Acknowledgements

All praises and glories be to my creator Allah Almighty, the most Merciful and the most Beneficent, who bestowed me with the knowledge to achieve this milestone. My most sincere thanks is to our beloved Prophet Hazrat Muhammad (P.B.U.H), my ultimate source of guidance not only in this path but throughout my life.

I would like to present my sincerest gratitude to Prof. Dr. Peter John, my MS supervisor, who has made this research possible by his valuable guidance and help. My cordial thanks to Prof. Dr. Attya Bhatti, my co-supervisor, who has always blessed me with her precious guidance and motivation.

My heartiest appreciation is for all the respected GEC members, Prof. Dr. Sobia Manzoor, HoD Healthcare Biotechnology, ASAB, NUST, Dr. Maria Shabbir and my external GEC member Prof. Dr. Muhammad Ansar for their guidance, comments, and suggestions. I am also thankful to all the faculty and staff of ASAB, NUST for always facilitating me. My acknowledgements cannot be completed without thanking Prof. Dr. Jahangir Sarwar Khan, Dr. Rubaba, Dr. Zeeshan, Dr. Khansa and Dr. Saba for their assistance in collecting samples required for this research project.

My family, Baba, Mama, Sarosh, Shazia and my friends deserve all the words of gratitude for being my constant support. I feel blessed to have such parents who have encouraged me throughout to achieve my goals and follow my dreams. Wherever I stand today is because of these two awesome people. I hope I have made you feel proud Mama Baba.

Cheers to my biggest supporter, my soulmate Emaan Imtiaz who was with me through out. We did this together and faced every challenge happily because of having each other's back. It was much easier when you were by my side dear. Bisma I can never forget to thank you for always being there to listen to my rants. Applauds to my NUST family, Sidra, Faria, Sana and Imaan for supporting me through thick and thin. I will always cherish the memories created with you all. I would like to thank all my lab fellows, Maria, Kashaf, Hafsa, Ilham, Tayyaba, Huda, Saliha, Rabia and Iraj for providing me with the best lab environment. Tackling the situations together we all rocked. We often failed at experiments but won at creating memories.

Lastly, I am grateful to everyone who has helped me at any phase of my MS. I would have done nothing without the support of all you precious people. I owe my degree to all of you.

Ameema Khan

Table of Contents

List of Figures

List of Tables

List of Acronyms

BMI	Body Mass Index
rASRM	Revised American Society for Reproductive Medicine
EFI	Endometriosis Fertility Index
IVF	*In-vitro* Fertilization
VEGF	Vascular Endothelial Growth Factor
TV-USG	Transvaginal Ultrasound
MRI	Magnetic Resonance Imaging
SNP	Single Nucleotide Polymorphisms
GnRH	Gonadotropin Releasing Hormone
GWAS	Genome Wide Association Studies
PR-A	Progesterone Receptor A
PR-B	Progesterone Receptor B
PR-C	Progesterone Receptor C
ER-β	Estrogen Receptor β
SF-1	Steroidogenic Factor-1
PGR/PR	Progesterone Receptor
NR3C3	Nuclear Receptor Subfamily 3 Group C Member 3
AF	Activation Function
kDa	Kilo Dalton
ID	Inhibition Domain
ESR1	Estrogen Receptor 1
ESR2	Estrogen Receptor 2
IL10	Interleukin 10
CSIF	Cytokine Synthesis Inhibitory Factor
IL1	Interleukin 1
TNF-α	Tumor Necrosis Factor Alpha
IFN-γ	Interferon Gamma
CD8+ T	Cytotoxic T lymphocytes
NK	Natural Killer
IL10R	Interleukin 10 Receptor
STAT3	Signal Transducer and Activator of Transcription 3
JAK 1	Janus kinase 1
TYK2	Tyrosine Kinase 2
SOCS3	Suppressor of Cytokine Signaling 3
MHC-II	Major Histocompatibility Complex II
HIV	Human Immunodeficiency Viruses
AIDS	Acquired Immune Deficiency Syndrome
SLE	Systemic Lupus Erythematosus

VEZT	Vezatin
PCR	Polymerase Chain Reaction
RNA	Ribonucleic Acid
NF-kB	Nuclear Factor kappa B
VEGAS	Versatile Gene-based Association Study
GWAS	Genome Wide Association Studies
WNT4	Wingless type MMTV integration site family member 4
GREB1	Growth Regulating Estrogen Receptor Binding 1
MSP	Methylation Specific PCR
BSP	Bisulfite Sequence PCR
UTR	Untranslated Region
NCBI	National Centre for Biotechnology Information
UniProt	Universal Protein Resource
SIFT	Sorting Intolerant from Tolerant
PolyPhen-2	Polymorphism Phenotyping version 2
PhD-SNP	Predictor of human Deleterious Single Nucleotide Polymorphisms
PANTHER	Protein Analysis Through Evolutionary Relationships
P-Value	Probability Value
pH	Potential Hydrogen
$\Delta\Delta G$	Change in Gibbs Free Energy
I-TASSER	Iterative Threading ASSEmbly Refinement
C-Score	Confidence Score
TM Score	Template Modeling Score
RMSD	Root Mean Square Deviation
PDB	Protein Data Bank
PTMs	Post Translation Modifications
IRB	Institutional Review Board
EDTA	Ethylenediaminetetraacetic acid
SDS	Sodium Dodecyle Sulphate
DNA	Deoxyribonucleic Acid
rpm	Revolutions per minute
TAE	Tris Acetate Ethylene-diamine-tetra-acetic acid
UV-Vis	Ultraviolet-Visual
ARMS-PCR	Amplification Refractory Mutation System Polymerase Chain Reaction
Tm	Melting Temperature
$MgCl_2$	Magnesium Chloride

Abstract

Endometriosis is a chronic, inflammatory, hormone dependent gynecological disease characterized by the growth and proliferation of endometrium outside of the uterus in different parts of the body. Endometriosis mainly affects the women of reproductive age, and its prevalence is approximately 10% worldwide and 6% in Pakistan. Complex etiology, painful clinical manifestations, invasive and limited diagnostic, and treatment options makes endometriosis a disease of great concern. Endometriosis has a strong genetic basis and studies have elucidated the association of SNPs of different genes with its progression, which can then be used as diagnostic and prognostic markers. The current study aims to investigate the association of rs1042838 (a missense SNP of PGR), rs1800871 (a UTR-5 SNP of IL10) and rs10859871 (located 17kb upstream of VEZT) with the risk of endometriosis using ARMS PCR. Moreover, *in-silico* analysis was performed to identify pathogenic missense SNPs of PGR and 3' and 5' UTR SNPs of IL10 gene. The results of our study revealed no significant difference between the genotype and allele frequency of patients and healthy controls. So, it can be concluded that the risk alleles of all three SNPs were not associated with susceptibility to endometriosis in our sample pool, but replication of this study with larger sample size is required for further validation of results. *In-silico* studies predicted that six missense SNPs of PGR and three regulatory SNPs of IL10 have disease causing potential which can be tested *in-vitro* for their role in endometriosis and other diseases.

Chapter 1

1. Introduction

1.1. Endometriosis

Endometriosis is a chronic, inflammatory, hormone dependent gynecological disease that affects the women of reproductive age and is characterized by the growth and proliferation of endometrial tissue present in the inner lining of uterus outside of the uterus and its cavity. In most cases the growth of the endometrial tissues spreads to pelvic cavity, fallopian tubes, ovaries, appendix, colon (França et al., 2022), cervix, rectum, small intestine, vagina, abdominal walls, bladder, and effects their normal functioning (Mohammed Rasheed & Hamid, 2020) resulting in infertility. In rare cases it can even spread to eyes, lungs and brain as well resembling the metastatic characteristics of tumors but details are still needed to be clarified (Lu & Gao, 2021).

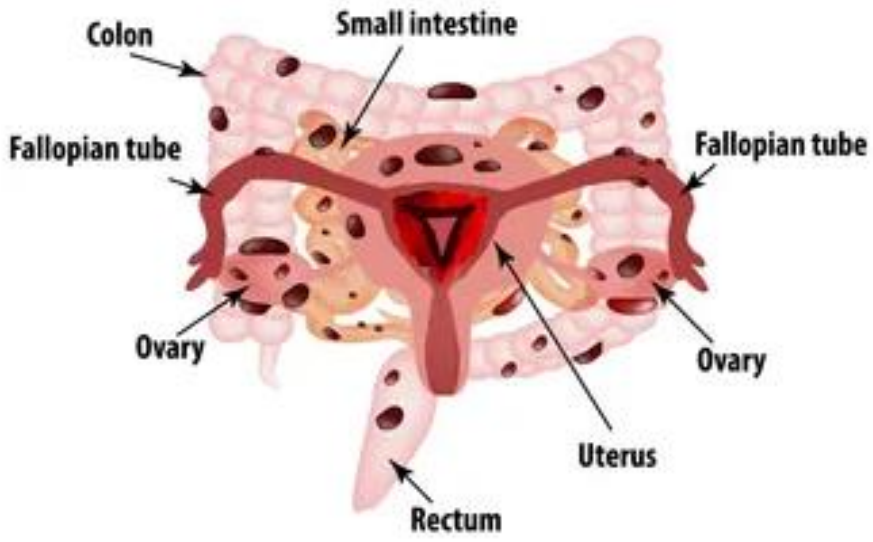

Figure 1.1. Growth of endometrial lesions in various parts of human body (Sonacare, 2018).

Women with endometriosis are further prone to diabetes, pelvic and liver diseases, cardiovascular diseases and many autoimmune disorders (Joseph & Mahale, 2019). Severe depression and anxiety has also been observed in women with endometriosis due to painful symptoms, invasive surgeries and more risk of infertility (L.-C. Chen et al., 2016).

1.2. Epidemiology

Endometriosis is a serious social, financial, and medical problem that affects a vast majority of women in between age of menarche to menopause and this disease is at its peak between ages of 25 years to 45 years. In rare cases, some women might have endometriosis even after menopause (Smolarz et al., 2021).

1.2.1. Worldwide Prevalence

Approximately 190 million women are affected with endometriosis worldwide which makes a percentage of about 10% (Mohammed Rasheed & Hamid, 2020). Prevalence of endometriosis varies between 2 to 11% in women without any symptoms and 5 to 21% among women with pain in pelvic cavity (Shafrir et al., 2018). Among women between age of 15 years to 49 years, an incidence of 0.1% is reported annually for endometriosis (Vercellini et al., 2014). Endometriosis leads to infertility in about 50% of the women diagnosed with it (Králíčková et al., 2020). Women with endometriosis are reported to be two-fold more at risk of infertility than women without it (Prescott et al., 2016). It is reported that among women, who undergo laparoscopy and laparotomy for diagnosis or treatment of various gynecological disorders, 0.1% to 53% are diagnosed with endometriosis. Furthermore, Asian women are at higher risk of endometriosis than others (Smolarz et al., 2021).

1.2.2. Prevalence in Pakistan

In Pakistan, an incidence of 6% of endometriosis is reported (Khan & Illahi, 2000). Another study reports that out of 160 patients that underwent laparoscopy for various gynecological disorders, endometriosis was diagnosed in 44 that makes a percentage of 27.5% (Tahira et al., 2019).

Despite of being a chronic disease and leading cause of infertility among women, exact prevalence of endometriosis is not clearly known because of late diagnosis and lack of proper management and treatment of disease. More studies are required for better understanding of epidemiology of endometriosis and its various aspects.

1.3. Risk Factors

There are several risk factors, genetic as well as epigenetic, exposure to which can increase the risk for endometriosis. Exposure to diethylstilbestrol, early onset of menarche, lower body weight, irregular menstrual cycle, decreased ratio of waist to hip, lower number of births by a woman are all thought to influence the development of endometriosis (Zondervan et al., 2020).

Smoking, alcohol, sedentary lifestyle, diet like red meat, fish etc., rich in fats, hormones, and dioxins can make women prone to endometriosis. Age also plays a crucial role and women between 25 to 29 years of age are at higher risk (Smolarz et al., 2021). Recently, a study has revealed that obese women have higher chance of getting endometriosis as compared to those with low body mass index (BMI) (Tang et al., 2020).

1.4. Etiopathogenesis

Several theories have been proposed regarding etiopathogenesis of endometriosis, but the exact mechanisms remain unknown. Some of the major mechanisms and factors that lead to the development of endometriosis are shown in figure 4.2.

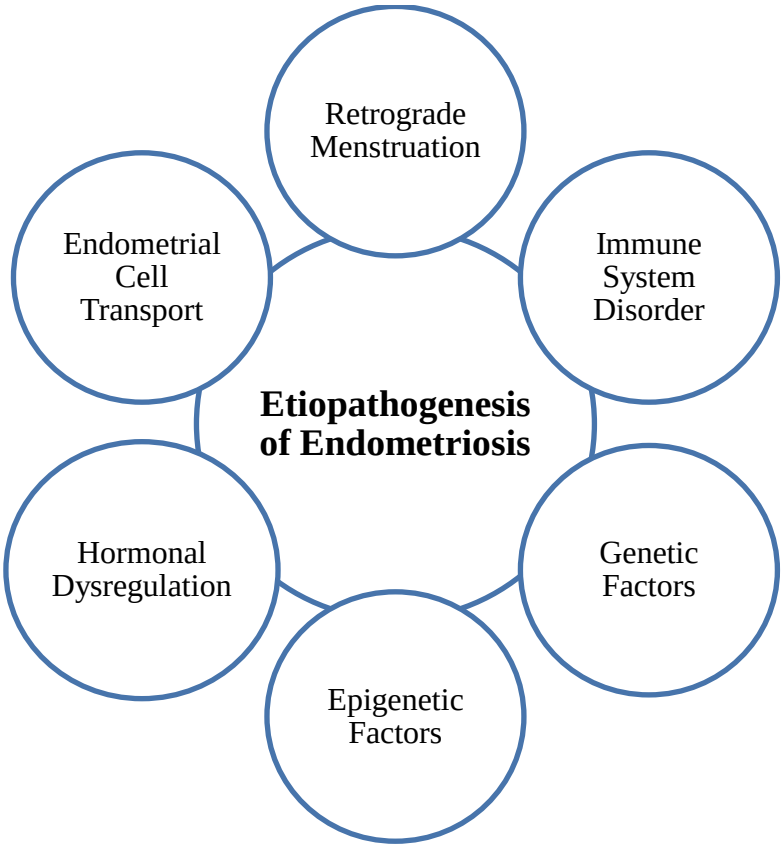

Figure 1.2. Processes involved in Etiopathogenesis of Endometriosis.

1.4.1. Retrograde Menstruation

The oldest known theory about the progression of endometrial lesions is retrograde menstruation first suggested by Sampson in 1927 (Ja, 1927). According to it, the menstrual blood and debris containing the endometrial tissue and cells flow back passing through fallopian tubes and deposit into peritoneal cavity resulting in the growth of endometrial lesions in peritoneum,

ovaries and pelvic organs (Tanbo & Fedorcsak, 2017). A study reported that first retrograde menstruation is observed at birth due to lack of necessary hormones in baby girls. This blood might not release due to tightness of cervix and thickness of cervix mucus, therefore the endometrial cells in the blood enter abdominal cavity and stay there until puberty due to lack of reproductive hormones. Once a girl enter adolescence, the endometrial implants grow and spread rapidly due to presence of hormones (Brosens & Benagiano, 2013). It should be noted that this phenomena of retrograde menstruation occurs in about 76 to 90% of women (França et al., 2022) but only 10% of them suffer from endometriosis (Bulun et al., 2019).

1.4.2. Endometrial Cell Transport

Another theory suggests that the endometrial cells might enter blood stream via capillaries in uterus and into lymphatic system through lymph nodes residing in uterus. Once they enter circulatory and lymphatic system, they travel to distant sites like brain, lungs, abdominal cavity and result in formation of endometrial lesions under influence of various reproductive hormones (Jerman & Hey-Cunningham, 2015) (França et al., 2022).

1.4.3. Immune System Disorder

Failure of immune cells to function normally can also lead to continuation and progression of endometriosis. In women with endometriosis, macrophages with reduced phagocytic activity are observed thus diminishing the natural killing activity. Macrophages instead release an increased number of inflammatory mediators i.e., cytokines, prostaglandin and chemokines which will induce inflammation. Lack of anti-inflammatory mediators is also observed which leads to further development of endometrial lesions. Disbalance in function of type 1 and type 2 helper T cells also promotes growth of endometrial tissue in areas other than inside uterus (Smolarz et al., 2021). Impairment in neutrophils' function result in the over production of VEGF which in turns promote growth of new blood vessels, thus promoting survival of endometrial lesions (Mohammed Rasheed & Hamid, 2020).

1.4.4. Hormonal Dysregulation

Steroid hormones play an important role in the pathogenesis of endometriosis. Endometrial lesions have an over production of aromatase, an estrogen synthesizing enzyme, that produces estradiol from cholesterol, along with its production from ovary, which promotes growth of these ectopic lesions. Higher levels of estrogen receptor β are observed in case of endometriosis which is expected to be due to the methylation of the promoter of the gene

encoding for it. This elevated level reduces the level of estrogen receptor α and causes progesterone resistance in endometrial lesions thus promoting their growth and causing inflammation (Vercellini et al., 2014).

Progesterone resistance is mainly caused by defects in progesterone receptor activation or due to problems in transcription of progesterone receptor gene (Bulun et al., 2019). Progesterone has anti-inflammatory potential and disruption of progesterone signaling can lead to failure of the hormone to balance the effect of estrogen hormone causing more cell proliferation. This leads to increased growth of endometrial lesions and inflammation which causes extreme pelvic pain (Marquardt et al., 2019).

1.4.5. Heritable Genetic Factors

Since many years, a lot of research is carried out to find out the heritable genetic factors associated with endometriosis. Although the genetics of endometriosis is complex, it is believed to be caused by many different factors. Studies report that many genetic and epigenetic factors together are involved in occurrence and progression of endometriosis. It has been estimated through twin studies that endometriosis has a heritability factor of about 47 to 51%, among which 26% comes from genetic variations (Saha et al., 2015). Women whose mothers had endometriosis have more than two times more chance of developing the disease as compared to women without affected mothers. Furthermore, women having affected first degree relatives are also at higher risk to form endometrial lesions (Dalsgaard et al., 2013).

Genes playing a vital role in synthesis and metabolism of steroid hormones, inflammation, normal functioning of receptors of various hormones, growth factors, regulation of cell cycle and apoptosis are manifested to be associated with endometriosis (Hansen & Eyster, 2010). Many genome wide association studies (GWAS) have been carried out, reporting about 19 single nucleotide polymorphisms (SNPs) to be associated with endometriosis. Most of the identified SNPs are associated with severe form of endometriosis rather than the mild cases and associated risk is of 1.75% (Sapkota et al., 2017). This risk percentage is very small as compared to other diseases and thus it fails to highlight high risk patients.

Most of the genetic factors identified are in intergenic regions, manifested to serve regulatory functions, details of which are not known yet. The focus is now to identify genetic factors that are related to etiopathogenesis of endometriosis (Zondervan et al., 2016).

1.4.6. Epigenetic Factors

Along with genetic factors many epigenetic factors are also responsible for the development of endometriosis. Hypermethylation of promoter of HOXA10 and progesterone receptor B (PR-B) genes effects transcription of the respective genes leading to reduced expression and causing progesterone resistance and cell invasion which promotes growth of endometrial lesions. Contrary to this, hypomethylation of promoters of estrogen receptor β (ER-β) and aromatase gene and transcription factor Steroidogenic Factor-1 (SR-1) leads to increased production of estrogen causing endometriosis (S.-W. Guo, 2009).

1.5. Clinical Manifestations

Pelvic pain and infertility are regarded as the major symptoms of endometriosis (Lu & Gao, 2021). Others include on and off uterine bleeding, dysmenorrhea (severe period cramps), dyspareunia (painful sexual intercourse), hematochezia (blood from anus with stool), dyschesia (difficulty in passing stool) (França et al., 2022). Other than these major symptoms some rare symptoms like nausea, bloating, various types of allergies, headaches, fatigue, anxiety, depression are also observed in women with endometriosis. Menorrhagia i.e. menstrual bleeding lasting for a period of time more than normal may also be observed in case of endometriosis in some women (Smolarz et al., 2021).

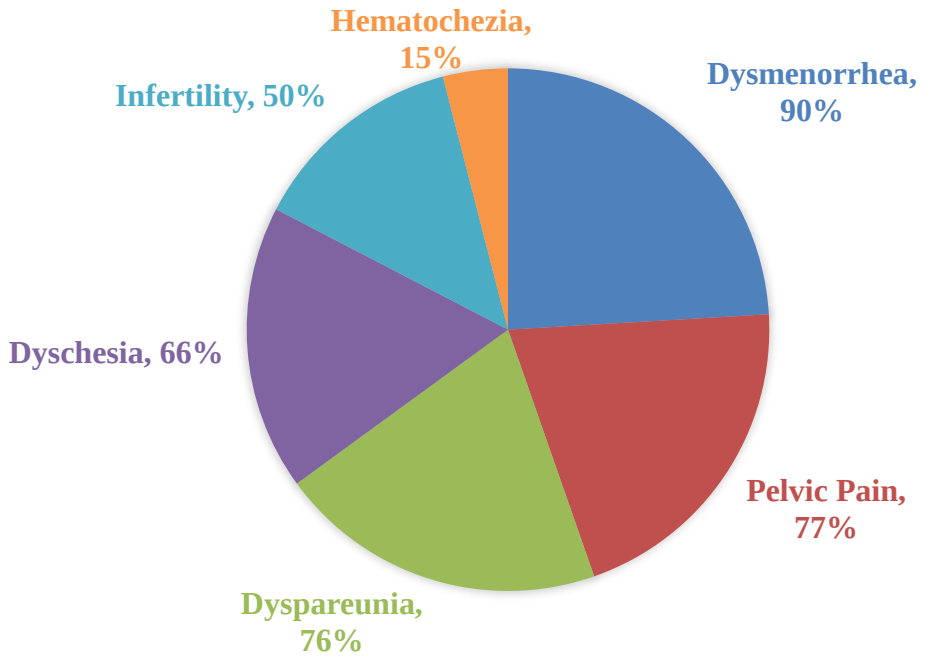

Figure 1.3. Common symptoms along with their percentage observed in patients with endometriosis (França et al., 2022).

Despite these obvious symptoms a lot of heterogeneity is observed in patients with endometriosis. Some women might be at severe stage of endometriosis but asymptomatic and some might show all the symptoms but disease progression is not that much severe (Zondervan et al., 2020).

1.6. Classification

There are various classification systems for endometriosis that are either based on anatomical features of endometrial lesions or on their location. Three main classification systems for endometriosis are commonly used until the development of a proper classification system that touches all the important aspects and helps the clinicians in better treatment of the disease. These are rASRM (revised American Society for Reproductive Medicine) classification, EFI (Endometriosis Fertility Index) and Enzian classification.

Table 1.1. Classification system for endometriosis along with their underlying drawbacks (Lee et al., 2020).

Classification System	**Stages/Score**	**Description**	**Drawbacks**
revised American Society for Reproductive Medicine (rASRM)	Stage I (1-5 points)	Minimal (small number of implants in peritoneum and ovaries)	Fails to highlight the presence of implants in vagina, bladder, and intestine. Does not consider intensity of pain and infertility.
	Stage II (6-15 points)	Mild (deep implants and chocolaty cyst in ovaries)	
	Stage III (16-40 points)	Moderate (more deeper implants, cysts in ovaries and filmy adhesions in fallopian tubes)	
	Stage IV (>40 points)	Severe (Many implants, huge cysts in ovaries and dense adhesions)	
Enzian	A	Implants in vagina and rectovaginal area	Publicly less accepted due to complexity
	B	Implants in sacrouterine ligament and pelvic wall	

	C	Implants in colon and rectum	
Endometriosis Fertility Index (EFI)	0-10 points	Predicts the probability of women with endometriosis to get pregnant without IVF with 0 having the least chance to 10 having the most chance of pregnancy	More complex than other two and does not consider pain and other symptoms of endometriosis.

Other than above mentioned three classification systems, endometriosis can also be categorized based on anatomical location as follows:

i. **Peritoneal endometriosis** in which tissue implants are formed in peritoneum (lining of abdomen).
ii. **Ovarian endometriomas** in which a chocolate like cyst is formed in one or both ovaries.
iii. **Deep Infiltrating Endometriosis** is more severe and tissue implants outgrow into pelvic organs, appendix, bowels, intestines and bladder (França et al., 2022).
iv. **Extra abdominal endometriosis** is characterized by formation of lesions in regions other than abdominal cavity like thoracic region, lungs, diaphragm and brain (Horne & Missmer, 2022).

1.7. Diagnosis

Diagnosis of endometriosis is a complex task and remains unclear. Most of the diagnostic methods are invasive and detect endometriosis quite late after a course of painful symptoms. The first step in diagnosis is oral examination of pelvic and abdominal cavity by an expert but confirmation of the disease is only possible by subsequent diagnostic tests which are as follow:

1.7.1. Imaging

Transvaginal Ultrasound (TV-USG) and Magnetic Resonance Imaging (MRI) are used for the detection of endometrial lesions in ovaries, urinary tract, rectosigmoid region with a sensitivity of about 93% and specificity of 95%. However, these methods are not suitable for diagnosis of peritoneal lesions because they are deep enough to exceed the detection limit of these methods (Taylor et al., 2021).

1.7.2. Laparoscopy

Laparoscopy is the small-scale surgical examination of the pelvic and abdominal cavity and is one of the most effective diagnostic methods for endometrial lesions. It highlights the progression, appearance and size of lesions and also helps in treatment of disease (Rolla, 2019).

1.7.2. Diagnostic Biomarkers

A lot of research is carried out on various biomarkers in blood, urine, and endometrial tissues for early diagnosis of endometriosis but none of them has proved to be a successful one. CA-125 is a popular biomarker but its level elevates in many other diseases as well along with endometriosis which reduces its specificity for use in diagnosis of single disease ((Broi et al., 2019).

1.8. Treatment

Before choosing a treatment plan for endometriosis, symptoms, age, level of progression of disease, possible side effects all should be taken into consideration. Moreover, it should be noted that there is no possible cure for endometriosis and the chances of recurrence of disease are very high. Therefore, proper diet and lifestyle also play an important role along with medications and hormonal therapies to suppress symptoms and restore normal hormone balance, and surgeries to remove lesions.

1.8.1. Hormonal Therapies and Medications

Progestins and combined oral contraceptives are the first-choice medicine used in treatment of endometriosis as they are less expensive and effective against reducing pain, but they have certain side effects as well and are not effective in some patients. Gonadotropin Releasing Hormone (GnRH) agonists and antagonists are the second-choice medication, but their oral administration is not possible. A recent GnRH antagonist called Elagolix is introduced in market that can be taken through oral routes and reduces estrogen levels but has adverse effects on bone health. Other than these many other medicines like danazol, gestrinone, and progestogen have also been proven to be effective against endometriosis.

Aromatase inhibitors are the new advancement in medications for endometriosis that directly inhibit the production of estrogen hormone. For pain reduction, nonsteroidal anti-inflammatory drugs are used in combination with the above-mentioned medicines (Ferrero et al., 2018). Recently, some natural products like Curcumin and Puerarin have been proven to be effective against endometriosis (França et al., 2022).

1.8.2. Surgical Procedures

When pain medications are not effective in women with endometriosis, surgery is the last resort. Surgery is done for the complete removal of endometrial lesions located in different regions inside the body. Laparoscopy, laparotomy, and hysterectomy are all used for the surgical

removal of endometrial implants. However, they are not fully curative and there is a high probability of comeback of pain symptoms (Tanbo & Fedorcsak, 2017). It is reported that there is a chance of 15 to 20% that a second surgery might be needed within 2 years of first one and this chance is increased to 50% after 5 to 7 years (Vercellini et al., 2009). An invention is laparoscopy assisted by robots which has proven to be more accurate as compared to the one done by humans, especially in case of deep lesions, obesity. This method is safe, effective but expensive (Restaino et al., 2020).

1.8.3. Non-medical treatments for management of symptoms

Many different treatments, not having medical routes, are also tried for management of symptoms in case of endometriosis. These include homeopathic medicines, traditional Chinese herbal medicines, antioxidants for removal of oxidative stress, reflexology, acupuncture, and therapies for stress removal. Physiotherapy post-surgery also proves beneficial for pain relief and normal functioning of pelvic organs (Wójcik et al., 2022). All these non-medical treatments help in management of symptoms but cannot cure the endometrial lesions.

1.9. Objectives of Study

The main objectives of the study are as follow:

i. To check the pathogenicity of the SNPs of selected genes through *in-silico* analysis.
ii. To investigate the association of selected SNPs with endometriosis through *in-vitro* analysis.

1.10. Impact of Study

Endometriosis is a chronic disease with late diagnosis and invasive treatment options. Moreover, the rate of recurrence of disease is very high. The current study will open new doors for thorough understanding of genetics associated with endometriosis, by identifying pathogenic SNPs that may cause the disease. This can lead to the development of potential genetic markers. Genetic markers identified can act as therapeutic targets which can play an important role in diagnostics and treatment of disease.

Chapter 2

2. Literature Review

Endometriosis is a chronic disease common in women of reproductive age and leads to severe pelvic pain and in worse conditions to infertility in about 50% of the women suffering from the disease (Králíčková et al., 2020). It is a complex disease having many causation factors including hormonal, genetic, epigenetic, and immunological but still the exact etiology remains unclear. Moreover, it is a disease with severe painful symptoms and invasive as well as limited diagnostic and treatment options (Soliman et al., 2017).

2.1. Genetic basis of Endometriosis

Extensive research on endometriosis has revealed that it has a strong genetic basis and different family aggregation and twin studies have highlighted the role of many genes in progression of this disease. It is reported that 6.9% of the mothers and sisters of the women with endometriosis are also affected with the disease (Pagliardini et al., 2015). This proves a strong genetic association of disease but due to late diagnosis and management issues as well as complex mechanisms, heritability of endometriosis is not clear yet. Moreover, studies reveal that endometriosis is inherited in polygenic pattern highlighting involvement of many genes along with various environmental factors as well (Hansen & Eyster, 2010).

Single nucleotide polymorphisms in genes play a major role in the development of endometriosis like many other diseases. These are the most common type of variations in humans and can change the gene product if present in coding region or regulatory region of gene. Thus, SNPs easily highlight the candidate genes responsible for progression of disease through association analysis (Angioni et al., 2020).

Candidate gene analysis and Genome Wide Association Studies (GWAS) carried out in different regions of the world report association of GREB1,VEZT, WNT4, (Nyholt et al., 2012), ESR2, PGR, IL10 (Liaqat et al., 2013), CDKN2B-AS1, ID4(Rahmioglu et al., 2014) 1L1 (Sapkota, Low, et al., 2015), FSHB (Matalliotakis et al., 2017), TGF-β1 (Yu et al., 2017), FN1, SYN1, CCDC170 (Sapkota et al., 2017), ESR1 (Osiński, Wirstlein, et al., 2018), FSHR (André et al., 2018) (Liaqat et al., 2021), AGTR1, CCL2, MMP 1, 7 and 9, IGF 1 and 2, CYP17A1 and CD40 (Vargas et al., 2020) in progression of endometriosis. Other than this many other genes have also been reported but results are vague. Out of all these, a missense SNP rs1042838 of

PGR, rs1800871 in 5' UTR of IL10 and rs10859871 located near VEZT in an enhancer region have been tested in this study to check their association with endometriosis.

2.2. Progesterone Receptor (PGR) Gene

Progesterone Receptor (PGR) gene is an important gene expressed in *Homo sapiens* and is located on chromosome 11q22.1. PGR has 8 exons and 7 introns, and it encodes for a 933 amino acid residues long protein called progesterone receptor (PGR). Progesterone receptor is also known as PR or nuclear receptor subfamily 3 group C member 3 (NR3C3). Progesterone receptors have a vital role in human body as they are involved in the activation and nuclear translocation of progesterone which has foremost role in development and growth of tissues of reproductive system (Ghali et al., 2020).

Like all other proteins, PR consists of a carboxyl-terminal (C-terminal) that is specific for ligand binding, a DNA binding domain, a hinge region, and an amino-terminal (N-terminal) that serves regulatory function. Moreover, two activation domains or functions are also present i.e., AF-1 in N-terminal of the protein and AF-2 in C-terminal. They function to provide interaction sites to coregulatory proteins. (Z. Li et al., 2022). PR is expressed in different parts of body including uterus, ovaries, urinary tract, testis, pancreas, bone, brain and mammary gland (Kowalik et al., 2013).

2.2.1. PGR Isoforms and Function

The progesterone receptor protein mainly exists in different isoforms and various splice variants among which two isoforms named PR-A and PR-B are reported to have major role. The two isoforms are similar in structure (Sherbet, 2018) but differ by an additional 164 amino acid residue in N-terminal of PR-B that is not present in PR-A. The two isoforms have different functions in human body and are transcribed by two different promotor regions (Z. Li et al., 2022). Both isoforms serve as transcriptional factors that are activated by binding of ligand (Mote et al., 2000). PR-A having a molecular mass of 94 kDa serves in repressing transcription whereas PR-B with molecular mass of 114 kDa serves in activating transcription of progesterone depending genes (Mani & Oyola, 2012).

In a cell, when both isoforms PR-A and PR-B are activated, PR-A inhibits the activity of PR-B due to presence of inhibition domain (ID) and thus results in reduction in effect of progesterone. ID is also present in PR-B but is vanquished by an additional activation function AF-3 in PR-B. A third isoform PR-C with molecular mass of 60 kDa is also reported after its

identification in placenta of humans. The function of this isoform is still vague, but studies suggest that it helps in transcriptional activity of PR-A and PR-B by forming heterodimers with them.

In addition to these three isoforms there exist a few more isoforms and splice variants, resulting due to alternative splicing, but they don't have a functional role due to absence of various important domains in their structure. Their role is yet to be explained but studies suggest that they might result in various complications by competing with PR-A and PR-B in their binding with ligands and regulators (Kowalik et al., 2013).

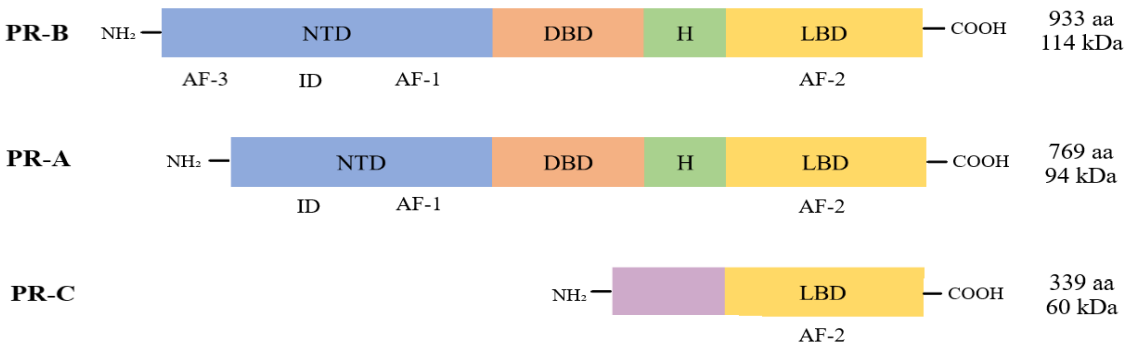

Figure 2.1. Schematic representation of isoform A, B and C of Progesterone Receptor (PR). NTD: N-Terminal Domain, DBD: DNA Binding Domain, H: Hinge, AF: Activation Function, ID: Inhibitory Domain

2.2.2. Role of PGR in Endometriosis

Exploration of the genetic profile of women with endometriosis has revealed that genes responsible for progesterone regulation are down regulated in case of the disease. Low levels of PR-A and extremely low levels of PR-B are observed in patients. In some cases, PR-B is not even detectable, the reason may be due to the hypermethylation of the promoter that initiates transcription of PR-B (Reis et al., 2020). The main reason for hypermethylation is reported to be epigenetic changes including increased release of cytokines and presence of abnormal micro-RNAs. Other than this, binding of ESR2 gene to promoter of PGR gene also results in reduced transcription of PGR gene, affecting the levels of PR-A and PR-B (Yilmaz & Bulun, 2019). A study reported the hypermethylation (about 50%) at PR-B receptor, by carrying out melting assay of high resolution using biopsy samples of patients with endometriosis, as compared to control where methylation was only 20% (Rocha-Junior et al., 2019).

PGR gene polymorphisms are also reported to play an important role in the progression of endometriosis. PGR has a high polymorphism rate and about 800 different SNPs of PGR gene have been reported (Ghali et al., 2020). In some cases, these SNPs result in no obvious effect but contrary to this few can lead to severe changes and are involved in the development of disease. Two SNPs rs1042838 and rs10895068 are divulged to be involved in growth of endometrial lesions. Thus, these SNPs can be used as diagnostic and therapeutic markers for endometriosis (Liaqat et al., 2013).

A study involving *in-silico* analysis reported connection between PGR and ESR1 genes and endometriosis, highlighting them as major susceptibility loci for risk of endometriosis. According to it down regulation of both the genes was observed in endometrial tissues in comparison of healthy tissues (Vargas et al., 2020). Another study in Iran performed genotyping using whole blood samples and tissue samples from patients with endometriosis and reported that presence of SNP rs10895068 can lead to elevated levels of PR-B isoform, influencing progesterone function and thus reducing symptoms of endometriosis. The study is quite influential but further studies are required for authentication (Mousazadeh et al., 2019).

2.2.3. Role of PGR SNP rs1042838 in Endometriosis

PGR SNP rs1042838 is a missense SNP (C>A) present in exon 4 of the gene that results in a single amino acid change from valine to leucine (V660L). This SNP is more common in European and South Asian nations in comparison of other nations (J. Li et al., 2018). It alters the flexibility of hinge region of the PR protein thus affecting its ability to form dimers and to interact with co regulators (Vang et al., 2023).

Association of this SNP has been reported with endometriosis, but results are contentious. A meta-analysis suggests the association of this SNP with endometriosis in European populations but the results in other populations are somehow vague. The reason may be due to the involvement of other gene polymorphisms as well as environmental factors (Pabalan et al., 2014). A case control study of Pakistani population reported that the mutant allele of SNP rs1042838 was expressed more in women with endometriosis as compared to the control thus highlighting its role in development of endometriosis. Furthermore, association of this SNP to endometriosis was even stronger when it existed in the form of haplotype along with another SNP rs10895068 present in 5’ UTR region of PGR gene (Liaqat et al., 2013).

2.2.4. Role of PGR in other Diseases

PGR gene polymorphisms may also be one of the causes of progression of many cancers like ovarian, gastric, endometrial and breast cancer. Higher level of PGR gene expression is observed in gastric cancer patients leading to poor prognosis of the disease (M. Li & Zhou, 2021). Elevated levels of transcription of progesterone receptor are observed in ovarian and cervical cancer cells due to SNP rs1042838. A systematic review reported association of three intronic variants of PGR gene that are rs10895054, rs1824128, rs590688, one 5' UTR variant rs10895068 and two missense variants rs1042838 and rs3740753 with breast cancer (Vang et al., 2023). PGR SNP variants rs10895068, rs590688 and rs1042838 are associated with risk of breast cancer in Tunisian women (Ghali et al., 2020).

A meta-analysis confirmed the association of rs1042838 along with Alu insertion, collectively known as PROGINS, with cancers of female reproductive tract (C. Zhou et al., 2022). PGR variants are also involved in premature births and SNPs responsible for this are rs1042838 and rs1042839 in mothers and rs1942836 in newborn (Kadivnik et al., 2022).

2.3. Interleukin 10 (IL10) Gene

In *Homo sapiens* Interleukin 10 (IL10), comprising of 7 exons, is located on chromosome 1q32.1. IL10 is also known as Cytokine Synthesis Inhibitory Factor (CSIF). It has a 178 amino acids long sequence and encodes for interleukin 10 protein, a major cytokine in human body, that regulates anti-inflammatory responses (Sharifinejad et al., 2022).

IL10 protein has a molecular weight of 20.5 kDa (UniProt, n.d.-a) and is majorly expressed in leukocytes, immune cells like macrophages, lymphocytes, monocytes and dendritic cells. In some cases, other immune cells also produce IL10 and in case of infections or tumors it is expressed in epithelial cells as well as keratinocytes (Piazzon et al., 2016).

2.3.1. IL10 Function

IL10 is a major anti-inflammatory regulator that reduces the damage due to inflammation by targeting antigen presenting cells and reducing production of IL1, TNF-α and IFN-γ which are all proinflammatory cytokines (Engelhardt & Grimbacher, 2014). IL10 acts as an inhibitor of NFkB pathway, regulator of production of antibodies and enhancer for growth and survival of B cells (Sharifinejad et al., 2022), CD8+ T cells and NK cells (Piazzon et al., 2016).

IL10 also plays an important role in activation of JAK/STAT signaling pathway. Activation of tetramer IL10R (Interleukin 10 Receptors) A and B by binding of IL10, as

homodimer with them, subsequently results in phosphorylation of signal transducer and activator of transcription 3 (STAT3). This step is mediated by Janus kinase 1 (JAK 1) and tyrosine kinase 2 (TYK2) and the phosphorylated STAT3 is then translocated into the nucleus. The activation and translocation of STAT3 is involved in regulating expression of various anti-inflammatory genes like SOCS3 and downstream cytokines (Guillot-Sestier et al., 2015).

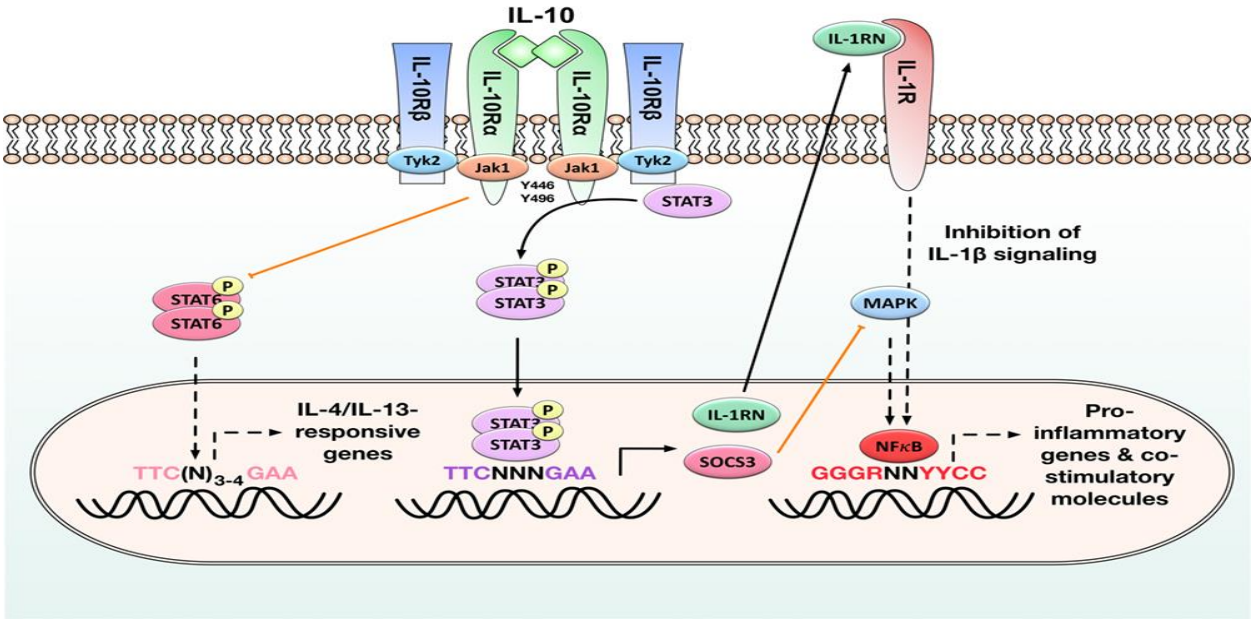

Figure 2.2. Regulation of JAK/STAT signaling pathway by IL10 (Schülke, 2018).

STAT3 signaling cascade hinders T cell activation by suppressing the expression MHC-II and other molecules involved in activation of T cells. Activation of macrophages is also regulated by STAT 3 signaling cascade activated by IL10 (Iyer & Cheng, 2012).

2.3.2. Role of IL10 in Endometriosis

Higher amount of IL10 protein was reported to be found in peritoneal fluid in women with endometriosis thus proving that growth of endometrial lesions is associated with disturbance in the release of IL10 and other proteins in peritoneal fluid (Tabibzadeh et al., 2003). Increased mRNA expression of IL10 have been reported by many studies in endometriotic lesions in different parts of human body as compared to the normal endometrium of patient as well as healthy women (Gueuvoghlanian-Silva et al., 2018) (W.-J. Zhou et al., 2019). Moreover, elevated levels of IL10 are also observed in serum of women with endometriosis (Suen et al., 2014) (Măluţan et al., 2015).

A study suggested that IL10 promotes growth of the endometrial lesions by promoting angiogenesis and suppressing immune response against them (Suen et al., 2019). According to

another study increased level of IL10 along with pro inflammatory cytokine IFN-γ collectively promote growth of endometrial lesions suggesting that balanced anti and pro inflammatory conditions are necessary for the progression of endometriosis (Qiu et al., 2020).

Due to the important role of IL10 in progression of endometriosis studies suggest association of various IL10 gene polymorphisms with risk of endometriosis. A meta-analysis explored role of three SNPs i.e., rs1800896, rs1800871 and rs1800872 in endometriosis and confirmed association of rs1800896 with risk of endometriosis (Zhong et al., 2023). Different studies report that in case of endometriosis, the elevated IL10 levels maybe 50 to 70% due to gene polymorphisms in IL10 gene and report association of −592C/A and −819T/C polymorphisms in promoter region of IL10 with endometriosis (J. Xie et al., 2009) (Malutan et al., 2017).

2.3.3. Role of IL10 SNP rs1800871 in Endometriosis

SNP rs1800871 (G>A) is present in 5’ UTR promoter region of IL10 gene in chromosome 1 of Homo sapiens. The minor allele of this SNP in homozygous form AA is reported to be associated with increased risk of endometriosis (Juo et al., 2009). A study carried out in Lahore, Pakistan also reported rs1800871 as a risk factor for endometriosis and reported association of mutant alleles of rs1800871, rs1800872 and rs1800896 present in haplotype form in promoter region of IL10 gene with the development of endometrial lesions (Liaqat et al., 2013).

2.3.4. Role of IL10 in other Diseases

Imbalance in the normal expression of Il10 is responsible for progression of various diseases. SNPs in IL10 gene and its promoter region are reported to be associated with risk of various viral infections like pneumonia induced sepsis (Mao et al., 2017), hepatitis B (Ramezani et al., 2012), hepatitis C (Ramos et al., 2012), influenza virus (Alagarasu et al., 2021), corona virus (Yessenbayeva et al., 2023), HIV infection and AIDS (Oleksyk et al., 2009). In addition to this polymorphisms in promoter region of IL10 gene are responsible for various autoimmune diseases including inflammatory bowel disease (Engelhardt & Grimbacher, 2014), rheumatoid arthritis (Padyukov et al., 2004), Behçet’s disease (Shahriyari et al., 2019), Parkinson’s disease (Ulhaq & Garcia, 2020) and Sjögren’s syndrome (Colafrancesco et al., 2019).

IL10 is considered as a crucial cytokine involved in pathogenesis of Alzheimer disease and polymorphisms in IL10 gene are associated with risk of Alzheimer (Magalhães et al., 2017).

Moreover, higher levels of IL10 are involved in causation of Systemic Lupus Erythematosus (SLE) and deficiency of IL10 is observed in patients with asthma (Iyer & Cheng, 2012). Different studies also highlight the association of IL10 gene polymorphisms with risk of Kawasaki disease (Ferdosian et al., 2021), hepatocellular carcinoma (J. Wang et al., 2019) and cervical cancer (Datta et al., 2020).

2.4. Vezatin (VEZT) Gene

Vezatin (VEZT) gene is located on chromosome 12q22 in *Homo sapiens* and has 10 exons. VEZT is recognized as an important gene having a tumor suppressing role and it encodes for a 779 amino acid residues long protein called vezatin also known as adherens junctions transmembrane protein (Y.-S. Li et al., 2015). Vezatin protein has a molecular mass of about 88 kDa and because of alternative splicing comprises of four different isoforms (UniProt).

Moreover, there are two main domains of VEZT protein, the extracellular one being short and intracellular one being long which is included in complex of adheren junctions because of its anchoring to myosin VIIa (Holdsworth-Carson et al., 2016). It is highly expressed in epithelial cells (Y.-S. Li et al., 2015), endometrium and myometrium of uterus (Angioni et al., 2020), brain (Miao et al., 2013), dendrites, and muscles of human body (Spinner et al., 2020).

2.4.1. VEZT Function

Vezatin protein plays a pivotal role in the synthesis and maintenance of adherent junctions by service of its long and short domain as bridge between myosin VIIa and complex of catenin, cadherin, and actin. Thus, it is involved in maintenance of polarity and elevation of adherence among epithelial cells (Y.-S. Li et al., 2015). Moreover, vezatin is also crucial for nourishment and sustenance of neuromuscular synapse junction (Koppel et al., 2019), post synaptic components and dendrites (Danglot et al., 2012).

Normal implantation of embryo requires adhesion of epithelial cells which is regulated by vezatin protein. A study in mouse reported that lack of vezatin gene expression was deleterious for the embryo implantation and growth. Absence of vezatin protein led to reduced amount of E-cadherin protein as it is essential for regulation of this protein. Due to poor adherence, presence of abnormal embryo or no presence at all was seen in mice after blastocyte stage. This proved that absence of vezatin in mice, due to knock out of VEZT gene, clearly led to the deterioration of cells because of poor adhesion between cells (Hyenne et al., 2007).

Vezatin is also involved in maintaining the cell adhesion of hair cells in the inner ear. A study revealed that presence of malfunctioned vezatin in hair cells of inner ear can lead to hearing loss if exposed to loud sound (Bahloul et al., 2009).

2.4.2. Role of VEZT in Endometriosis

As VEZT plays a major role in cell adhesion and embryo implantation and is efficiently expressed in myometrium and endometrium of uterus, it can have a major role in progression of endometriosis as well. Studies suggest increased expression of VEZT in women with endometriosis (Meola et al., 2010). Another study suggests an increased expression of VEZT in epithelial cells during luteal phase of menstrual cycle. There is also an increase in expression of RNA in cadherin-catenin complex during this phase as compared to other phases of menstrual cycles. As progesterone is the major hormone released during luteal phase, there might be a chance that progesterone is responsible for up regulation of VEZT gene in this phase leading to increased cell adhesions. But studies reveal no such connection of VEZT gene and its regulatory region with progesterone and its receptor. However, VEZT gene has a binding site for transcription factor NF-kB which is responsible for inflammation and thus varied levels of NF-kB are involved in development of endometrial lesions in regions other than uterus. Studies predict that fluctuations in NF-kB might be involved in up regulating VEZT expression along with abnormal levels of progesterone and thus VEZT can somehow also lead to progression of endometriosis (Angioni et al., 2020).

Many studies all around the world have reported the association of VEZT gene with endometriosis. VEGAS (Versatile Gene-based Association Study), a program for testing GWAS studies based on linkage equilibrium and size of gene, predicted WNT4, VEZT and GREB1 located on chromosome 1, 12 and 2 respectively as the top three genes to be associated with endometriosis. SNP rs10859871 near VEZT gene is majorly responsible for increased risk of endometriosis (Nyholt et al., 2012). More precisely a region on chromosome 12 near VEZT gene is reported to be a cause of development of endometriosis according to studies conducted in different populations. SNP rs4762347 in 3’ UTR of VEZT gene is reported to result in abnormal expression of VEZT in patients with endometriosis as compared to controls (Luong et al., 2015).

2.4.3. Role of SNP rs10859871 near VEZT gene in Endometriosis

SNP rs10859871 (A>C) is located approximately 17kb upstream of VEZT gene in an enhancer region on chromosome 12 of *Homo sapiens*. Due to its presence in enhancer region it

affects transcription of VEZT. Association of this SNP with endometriosis has been reported by many studies. Results from eight GWAS studies prove strong association of this SNP with all stages of endometriosis (Rahmioglu et al., 2014) (Cardoso et al., 2020). Different studies confirmed the association of mutant CC genotype of rs10859871 with endometriosis in women from different regions of world including Italy (Pagliardini et al., 2015), Australia and New Zealand (Holdsworth-Carson et al., 2016), Russia (Pshenichnyuk et al., 2017)

A meta-analysis of Belgian population foretold the association of this SNP with only stage III and IV of endometriosis (Sapkota, Fassbender, et al., 2015). A study conducted in Greece divulged that AC genotype of SNP rs10859871 is associated with risk of endometriosis rather than CC genotype (Matalliotakis et al., 2017). Contrary to this, studies conducted on Mediterranean (Angioni et al., 2020) and Polish Caucasian (Osiński, Mostowska, et al., 2018) populations showed no association of this SNP with endometriosis.

2.4.4. Role of VEZT in other Diseases

A study in China reported role of VEZT gene in gastric cancer. Reduced levels of expression of VEZT gene were observed in gastric tissues of patients and in some there was no expression at all. Hypermethylation of promoter of VEZT gene was identified as the culprit for this reduced expression by performing Methylation Specific PCR (MSP) and Bisulfite Sequence PCR (BSP) in cell lines of gastric cancer. It was also observed that by using the inhibitor for DNA methylation on cell lines, expression level of VEZT gene was significantly increased thus highlighting a treatment option for gastric cancer (X. Guo et al., 2011). Another study by same lab reported that methylation of promoter of VEZT gene is induced by *Helicobacter pylori* infection (Miao et al., 2013).

Many studies also reported that restoration of normal expression of VEZT gene can have tumor suppressing role in gastric cancer (Miao et al., 2013) (D. Xie et al., 2017). Another study suggested that VEZT can be used as potential biomarker for diagnosis of gastric cancer (Beheshtizadeh & Moslemi, 2017). Elevated expression of vezatin encoded by VEZT gene was found in tissues from brain of epilepsy patients thus it was reported that up regulation of vezatin might be a reason for seizures and thus causes epilepsy. Knock down of vezatin vanquished seizures in mice with epilepsy thus proving the association of overexpression of vezatin with seizures (Y. Wang et al., 2021).

Chapter 3

3. Methodology

3.1. Selection of SNPs

SNPs of three different genes reported to be associated with endometriosis i.e., PGR, IL10 (Liaqat et al., 2013) and VEZT (Holdsworth-Carson et al., 2016) were selected. The details about the three SNPs selected for this study are as follow:

Table 3.1. Location and type of selected SNPs of PGR, IL10 and VEZT.

Gene	SNP ID	Location	Variant Type
PGR	rs1042838	chr11:101062681 (GRCh38.p14)	Missense
IL10	rs1800871	chr1:206773289 (GRCh38.p14)	5 Prime UTR
VEZT	rs10859871	chr12:95318100 (GRCh38.p14)	Intron

3.2. In-silico Analysis of Missense SNPs of PGR

Since the selected SNP of PGR gene is a missense SNP, *in-silico* analysis for filtering all the deleterious missense SNPs of PGR gene was performed.

3.2.1. Retrieval of SNPs

The data of SNPs of PGR gene was obtained from the National Centre for Biotechnology Information (NCBI) dbSNP https://www.ncbi.nlm.nih.gov/snp/ (accessed on 03-01-2023). The data set included information like SNP IDs, chromosome location and allele change of the SNPs.

3.2.2. Retrieval of Protein Sequence

The canonical protein sequence of Progesterone Receptor was retrieved in FASTA format from UniProt https://www.uniprot.org/. The protein sequence comprised of 933 amino acids.

3.2.3. Identification of Deleterious Missense SNPs

The missense SNPs retrieved from the NCBI were submitted to SIFT (Sorting Intolerant from Tolerant) https://sift.bii.a-star.edu.sg/. SIFT is a sequence homology based tool that predicts that whether a change in amino acid due to a single nucleotide change effects protein function or not (Kumar et al., 2009). In Sift, a score of less than 0.05 is considered to be deleterious. The missense SNPs that were predicted to be deleterious by SIFT were further submitted to four more online software including PolyPhen-2, PhD-SNP, SNPs&GO and PANTHER. PolyPhen-2 (Polymorphism Phenotyping v2) http://genetics.bwh.harvard.edu/pph2/ predicts that whether an amino acid substitution functionally or structurally effects human protein or not by comparing

wild type and mutant alleles (Adzhubei et al., 2010). PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms) https://snps.biofold.org/phd-snp/phd-snp.html anticipates the effect of SNP on protein stability and tells us that whether an SNP can be disease related or not (Capriotti et al., 2006). SNP&GO https://snps-and-go.biocomp.unibo.it/snps-and-go/ characterizes the submitted amino acid change into neutral or disease causing by exploring functional annotation of protein (Calabrese et al., 2009). PANTHER (Protein Analysis Through Evolutionary Relationships) http://www.pantherdb.org/tools/csnpScoreForm.jsp predicts an amino acid change to be deleterious based upon the conservation time of wild type amino acid i.e. the longer an amino acid has been conserved over generations, the more deleterious the change at its position would be (Mi et al., 2021). The missense SNPs that were predicted to be damaging by at most three of the above-mentioned tools were selected for further analysis.

3.2.4. Assessment of the Structural and Functional effect of Deleterious Missense SNPs

To assess the structural and functional effect of deleterious missense SNPs on protein MutPred2 web server http://mutpred.mutdb.org/ was used. MutPred2 is an online tool that appraises the deleterious effect of amino acid change in protein sequence by predicting the functional and structural changes that an amino acid substitution causes to protein (Pejaver et al., 2020). Protein sequence in FASTA format along with list of amino acid substitutions were submitted to the software. A P-value less than 0.05 is contemplated as confident and less than 0.01 as very confident.

3.2.5. Evolutionary Conservation Analysis

To predict the effect of missense SNPs on conserved amino acid in protein sequence ConSurf https://consurf.tau.ac.il/consurf_index.php was used. ConSurf determines the rate of evolution of amino acids by generation of phylogenetic relationship among homologous sequences (Yariv et al., 2023). Protein FASTA sequence was submitted to ConSurf and conservation score of 1 to 9 for all the amino acids in protein sequence was obtained, with score of 9 being the most conserved and 1 being the least conserved. It also gives information that whether an amino acid residue is buried or exposed, functional or structural (Ashkenazy et al., 2016).

3.2.6. Predicting the Effect of Deleterious Missense SNPs on Protein Stability

Protein stability is an important feature in normal protein functioning. To predict the effect of change in amino acid sequence, due to missense SNP, on protein stability I-Mutant2.0

https://folding.biofold.org/i-mutant/i-mutant2.0.html was used. I-Mutant 2.0 is an online tool that efficiently predicts that whether an amino acid change increases or decreases the stability of protein (Capriotti et al., 2005). Protein sequence in FASTA format was submitted to I-Mutant2.0 along with amino acid substitution information at default setting i.e., temperature 25°C and pH 7 and ΔΔG value as well as reliability index (RI) was obtained. Positive ΔΔG value (ΔΔG value>0) means protein stability is increased and negative ΔΔG value (ΔΔG value<0) means protein stability is decreased (Capriotti et al., 2005).

3.2.7. 3D Protein Modelling

The missense SNPs filtered using the above-mentioned tools were further subjected to protein modelling. 3D structures of wild PGR protein and its mutant were designed using I-TASSER (Iterative Threading ASSEmbly Refinement). I-TASSER https://zhanggroup.org/I-TASSER/ is one of the best online tool available for 3D protein modelling that first uses multiple threading method that recognizes the templates from the Protein Data Bank (PDB) and then models the 3D protein structure homologous to the amino acid sequence provided in FASTA format (Yang & Zhang, 2015). I-TASSER predicts five protein models for each amino acid sequence along with their Confidence Scores (C-Scores). C-Scores give us information about the overall quality of 3D protein model (Roy et al., 2010). Protein Structure with highest C-Score was selected for further analysis.

3.2.8. Validation of 3D Protein Structures

3D protein structures with highest C-Score as per I-TASSER were further validated for good quality using ERRAT. ERRAT https://saves.mbi.ucla.edu/ is a structure validation tool that validates a protein structure quality based on atomic interactions (Colovos & Yeates, 1993). ERRAT calculates an overall quality factor which > 50 is considered acceptable (Omar et al., 2018).

3.2.9. Comparison of the Wild and Mutant 3D Protein Structures

To compare the wild and mutant protein structures they were superimposed using TM-align. TM-align https://zhanggroup.org/TM-align/ identifies correct alignment between the wild protein structure and its mutants and produces the Template Modeling score (TM score) and the Root Mean Square Deviation (RMSD) values that highlights the differences between the two protein structures. TM score lies between 0 and 1 where 1 means perfectly aligned proteins and vice versa. Greater RMSD value means more difference between wild and mutant protein

structures (Zhang & Skolnick, 2005). Protein models designed using I-TASSER were submitted to TM-align in PDB format and their TM score and RMSD values were obtained.

3.2.10. Visualization of the Wild and Mutant Protein Models

For visualization and further analysis of the structure of wild and mutant amino acid residues PyMOL version 2.5.5 was used. PyMOL is an outstanding software for visualization of protein models as well as analyzing various chemical properties of the structures. Moreover, BIOVIA Discovery Studio Visualizer, another tool for macromolecule visualization, was also used for analysis of wild and mutant protein structures.

3.2.11. Post Translation Modification (PTM) Analysis

Post Translation Modifications (PTMs) are very crucial for normal functioning and interactions of protein. The gain and loss of PTM sites, due to amino acid change, because of presence of deleterious missense SNP were predicted. Five main PTMs i.e., Methylation, Phosphorylation, Ubiquitination, Glycosylation and SUMOylation were predicted using different tools. Arginine and lysine are the two main targets of methylation and for prediction of these sites GPS-MSP 1.0 http://msp.biocuckoo.org/online.php (Deng et al., 2017) was used. In case of phosphorylation, the target of kinases are serine, threonine and tyrosine which were predicted using GPS 6.0 http://gps.biocuckoo.cn/online.php (M. Chen et al., 2023). Similarly PTM sites for ubiquitination and SUMOylation were foretold by GPS-Uber http://gpsuber.biocuckoo.cn/online.php (C. Wang et al., 2022) and GPS-SUMO 2.0 https://sumo.biocuckoo.cn/online.php (Zhao et al., 2014) respectively. For N-linked glycosylation, NetNGlyc 1.0 https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/ (Gupta & Brunak, 2002) and for O-linked glycosylation NetOGlyc 4.0 https://services.healthtech.dtu.dk/services/NetOGlyc-4.0/ (Steentoft et al., 2013) were used.

3.2.12. Protein-Protein Interaction Analysis

A protein interacts with many other proteins and this interaction is necessary for regulation of many important biological processes crucial for life. This interaction, if disrupted, may lead to the development of various diseases. For prediction of protein-protein interactions of PGR STRING https://string-db.org/ was used. STRING gives information about the physical as well as functional connection among various proteins in the form of interaction pathway (Szklarczyk et al., 2022). Name of the protein and organism name was added to get the interaction pathway.

3.3. In-silico Analysis of Regulatory SNPs of IL10 and rs10859871 near VEZT

Since the selected SNP of IL10 gene is a UTR-5 variant, *in-silico* analysis for filtering all the pathogenic UTR-5 and UTR-3 SNPs of Il10 gene was performed. Moreover, the third intronic SNP rs10859871 located 17kb upstream of VEZT was also tested for its pathogenicity.

3.3.1. Retrieval of SNPs

The data of regulatory SNPs located in 5' and 3' UTR regions of IL10 gene and rs10859871 was obtained from Ensembl https://asia.ensembl.org/index.html (accessed on 19-05-2023). The data set included information like SNP IDs, chromosome location and allele change of the SNPs.

3.3.2. Identification of Deleterious SNPs

The SNPs retrieved from Ensemble were submitted to RegulomeDB and SNPinfo (FuncPred) to check whether the SNPs are disease causing or not. RegulomeDB https://regulomedb.org/regulome-search/ is an online tool that predicts the functional effect of regulatory SNPs and gives information that whether the SNPs effect the binding of transcription factor and expression of gene or not. It divides SNPs into six categories ranging from 1 to 6 with 1 having the most functional effect on binding and expression to 6 being the least (Boyle et al., 2012). SNPinfo (FuncPred) https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html is another online web server that helps in filtering the pathogenic SNPs and helps in selection of SNPs for association analysis with diseases (Xu & Taylor, 2009). rsIDs were submitted in RegulomeDB and SNPinfo for assessment of deleterious SNPs.

3.4. *In-vitro* Analysis

3.4.1 Study Subjects and Sample Collection

The underlying case control study was conducted to analyze the role of PGR (rs1042838), IL10 (rs1800871) and VEZT (rs10859871) gene polymorphisms in patients with endometriosis in comparison with normal control. Informed written consent was taken from all the patients and normal controls who participated in this study. A total of 38 human subjects were included in this study among which 19 were clinically diagnosed endometriosis' patients and 19 were normal controls. Samples were collected from Holy Family Hospital Rawalpindi, Benazir Bhutto Hospital Rawalpindi, District Headquarter Hospital Rawalpindi and Research and Diagnostic laboratory ASAB over a period of 8 months.

3.4.2. Criteria of Inclusion and Exclusion

The following inclusion and exclusion criteria was observed.

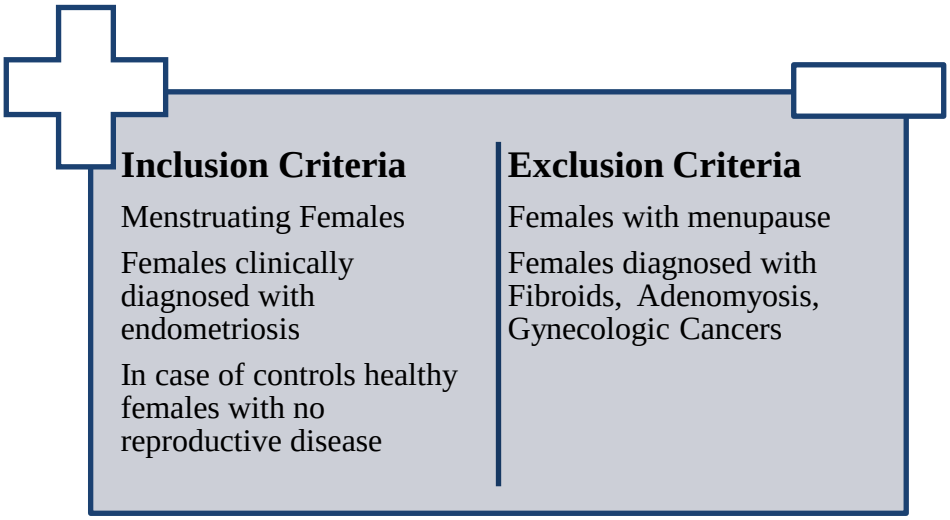

Figure 3.1. Inclusion and exclusion criteria.

3.4.3. Collection and Storage of Blood Samples

Blood samples from the subjects were collected in 3ml Ethylenediaminetetraacetic acid (EDTA) blood collecting tubes by hospital staff. EDTA tubes were properly labelled with name and age of patients and healthy controls as well as date of collection of blood sample. EDTA tubes containing blood samples were transported to the laboratory by placing them in icebox and were placed at 4°C in refrigerator.

3.4.4. Extraction of Genomic DNA by Phenol-Chloroform Method

DNA was extracted from the whole blood samples by using Phenol-Chloroform method of DNA extraction. It is a two-day protocol which is cost effective as well as a reliable method of DNA extraction giving satisfactory results. All the glassware and plastic wares used in DNA extraction were properly washed, cleaned, and autoclaved before use.

3.4.4.1. Preparation of Solutions for DNA Extraction

Composition and function of all the solutions used in Phenol-Chloroform method of DNA extraction are as follow:

Table 3.2. Solutions required for DNA extraction and their functions.

Component	Molarity	Quantity	Function
Solution A			
Sucrose	0.32 M	109.55 g	Facilitates the release of

Tris (pH 7.5)	10 mM	12.114 g	cell components including DNA by causing cell lysis and breakage of cell membrane
Magnesium Chloride	5 mM	0.476 g	
Triton X-100 (1% V/V)	-	10 ml	
Autoclaved Distilled Water	-	Upto 1000 ml	
Solution B			
Tris (pH 7.5)	10 mM	12.114 g	Results in the DNA precipitation and separation of the proteins
Sodium Chloride	400 mM	23.37 g	
Ethylene Diamine Tetra Acetic Acid (EDTA)	2 mM	0.58 g	
Autoclaved Distilled Water	-	Up to 1000 ml	
Solution C			
Phenol	-	250 µl	Helps in the separation of DNA from protein and other cell debris by forming aqueous layer
Solution D			
Chloroform	-	48 mL	Involved in DNA purification by stabilizing the coagulated proteins and reducing foaming
Iso-amyl Alcohol	-	2 mL	
20% SDS solution			
Sodium Dodecyle Sulphate (SDS)		20 g	Involved in denaturation of proteins and lipids and separates them from DNA
Autoclaved Distilled Water		Up to 100 ml	
3M Sodium Acetate Solution			
Sodium Acetate	3 M	12.3 g	Involved in precipitation of DNA by neutralizing the negative charges on phosphate backbone
Autoclaved Distilled Water	-	Up to 50 ml	

3.4.4.2. Reagents for DNA Extraction

Other chemicals needed in Phenol-Chloroform method of DNA Extraction and their functions are as follow:

Table 3.3. Reagents required for DNA extraction and their function.

Component	**Function**
Proteinase K	Involved in digestion of proteins
Isopropanol	Involved in precipitation of DNA by removing the solvation shell that surrounds the DNA
98% Ethanol	Involved in precipitation of DNA
PCR Water	Used for storage of DNA as it prevents degradation of DNA

3.4.4.3. Protocol of DNA Extraction

Day 1:

i. Blood in EDTA tube was kept at room temperature for some time and mixed well by inverting the tube several times. 750 µl of blood was then added in a 1.5 ml microcentrifuge tube and same amount of solution A i.e., 750 µl was added in it.
ii. The microcentrifuge tube was inverted 4-6 times and then kept at room temperature for 10 minutes.
iii. The mixture in microcentrifuge tube was then centrifuged at 13,000 revolutions per minute (rpm) in a microcentrifuge for 10 minutes.
iv. After centrifugation, the supernatant was carefully discarded, and the nuclear pallet was resuspended in 400 µl of solution B.
v. The nuclear pallet was dissolved completely in solution B by continuous tapping or slight vertexing.
vi. The mixture was again centrifuged at 13,000 revolutions per minute (rpm) for 10 minutes.
vii. After centrifugation the supernatant was again discarded and 400 µl of solution B, 15 µl of 20% SDS solution and 8 µl of Proteinase K was added in microcentrifuge tube.
viii. The mixture was then incubated at 37°C overnight.

Day 2

i. A fresh mixture of 250 µl each of solution C and solution D was prepared in a microcentrifuge tube and then was added to the microcentrifuge tube incubated overnight.
ii. The resulting mixture was centrifuged at 13,000 revolutions per minute (rpm) for 10 minutes.
iii. After centrifugation the aqueous layer formed, that contains the DNA, was transferred into a new microcentrifuge tube.
iv. 500 µl of solution D was added in the separated aqueous layer and it was then centrifuged at 13000 revolutions per minute (rpm) for 10 minutes.
v. The resulting aqueous layer was again transferred to a new microcentrifuge tube and then 55 µl of 3M sodium acetate solution and 500 µl of isopropanol was added in it.
vi. DNA was allowed to precipitate by inverting the tube several times and then mixture was again centrifuged for 10 minutes at 13000 revolutions per minute (rpm).
vii. After centrifugation the supernatant was discarded and 200 µl of chilled 98% ethanol was added in the microcentrifuge tube containing DNA pellet.
viii. Centrifugation was done for 8 minutes at 13000 revolutions per minute (rpm).
ix. After centrifugation the ethanol was evaporated by air drying and DNA pellet was submerged in 100 µl of PCR water and stored at -20°C.

3.4.5. Gel Electrophoresis of DNA

To analyze the quality of extracted DNA 2% (w/v) agarose gel electrophoresis was performed. The details of preparation of buffer solutions and reagents required for gel electrophoresis are as follow:

3.4.5.1. 0.5 M EDTA Solution

0.5 M EDTA solution was prepared by dissolving 186.12 grams of EDTA in distilled water to a final volume of 1000 ml which was then used in preparation of TAE buffer for gel electrophoresis. pH of 0.5M EDTA solution should be 8.

3.4.5.2. 50X Tris Acetate Ethylene-diamine-tetra-acetic acid (TAE) Buffer

50X stock solution of Tris Acetate Ethylene-diamine-tetra-acetic acid (TAE) Buffer was prepared as follow:

Table 3.4. Components required for preparation of TAE buffer.

Component	**Quantity**
Tris Base	242 g
0.5 M EDTA Solution	100 ml
Glacial Acetic Acid	57.1 ml
Deionized water	Up to 1000 ml

pH of 50 X TAE buffer was adjusted between 8.2 to 8.4. After preparation 50X TAE buffer was autoclaved and stored at room temperature.

3.4.5.3. 1X Tris Acetate Ethylene-diamine-tetra-acetic acid (TAE) Buffer

1X TAE buffer was prepared from 50X TAE buffer by taking 20 ml of 50X TAE buffer and adding 980 mL of distilled water in it.

3.4.5.4. Protocol

i. 2g of agarose was weighed by using electronic weighing balance and was added into 100 ml of 1X TAE buffer.
ii. The mixture was dissolved by heating it in microwave oven for approximately 2 minutes.
iii. After slight cooling, 4 µl of ethidium bromide was added for staining of DNA.
iv. The gel solution was then poured into gel casting tray and was allowed to solidify at room temperature. Wells for sample loading were made with the help of comb.
v. When solidified, gel was placed in electrophoresis tank which was filled with 1X TAE buffer.
vi. 2 µl of loading dye was used to stain 6 µl of extracted DNA sample each, which was then loaded into gel wells.
vii. 1 kbp DNA ladder was also loaded in gel well for the purpose of determining the length of DNA sample.
viii. Gel electrophoresis was performed at 80 volts for 30 minutes.
ix. After completion of gel electrophoresis, gel was placed in UV Transilluminator as well as ChemiDoc and DNA bands were analyzed and compared with DNA ladder.

3.4.6. Quantification of DNA

ThermoScientific Nanodrop 2000 UV-Vis Spectrophotometer and NanoDrop 2000™ software was used to quantify the extracted DNA samples. 1 µl of PCR water was first placed on pedestal which acts as blank and then 1 µl of DNA sample was loaded to analyze the absorbance

ratio at 260 nm wavelength which was kept as standard. The purity of DNA was determined based upon the absorbance ratio of 260/280nm and a ratio of ~1.8 is considered as in case of DNA.

3.4.7 Primer Designing

Primers were designed manually for each gene polymorphism to carry out allele specific Amplification Refractory Mutation System PCR (ARMS-PCR). FASTA sequence of each SNP was retrieved from Ensemble. For each SNP three primers were designed as shown in table 3.5.

Table 3.5. Primers designed for ARMS PCR.

Gene	SNP	Variation	Primer Sequence (5' to 3')	Tm	GC Content	Product Size
PGR	rs1042838	C/A	Forward primer wild: CTTTCATTTGGAACGCCCA**C** Forward primer mutant: CTTTCATTTGGAACGCCCA**A** Reverse primer: GTATAATTACCCACACCTCA	52.4 50.4 48.3	50 45 40	362
IL10	rs1800871	G/A	Forward primer: AGACTTCTCCTTGCTAACTTA Reverse primer wild: ACCCTTGTACAGGTGATGTAA**C** Reverse primer mutant: ACCCTTGTACAGGTGATGTAA**T**	48.5 53.6 51.7	38 45 41	392
VEZT	rs10859871	A/C	Forward primer wild: CTATTCAGCTCATTTCTTTGGA**A** Forward primer mutant: CTATTCAGCTCATTTCTTTGGA**C** Reverse primer: ACAATGGTGTAACTGGTAGGGT	50.2 51.9 53.6	35 39 45	221

In the case of PGR and VEZT two forward primers and one common reverse primer whereas in the case of IL10 two reverse primers and a forward common primer were designed. Two primers differ from each other only by last nucleotide at 3' end which contains the polymorphic nucleotide. The wildtype nucleotide was placed at end of one primer and the mutant

nucleotide was placed at end of second primer. OligoCalc software was used to determine the GC content, melting temperatures (Tm), hair-pin formation and self-complementarity of primers. Primer specificity was determined by using Primer-Blast as well as UCSC in-silico PCR software which gives information about amplicon size as well as shows the target size where primers bind.

3.4.8. Preparation of Working Dilutions of Primers

Primers were received in a lyophilized state, so first 100 µM stock was prepared for each primer. The amount of nuclease free water added into the lyophilized primers was determined by multiplying the nmol of primer by 10. After preparation of stock working dilutions of primers were made to avoid contamination and reduce freeze/thaw cycles of stock. Working dilutions were made by adding 10 µl of primer from stock and 90 µl of nuclease free water in a sterile microcentrifuge tube.

3.4.9. Allele Specific Amplification Refractory Mutation System PCR (ARMS-PCR)

Allele Specific Amplification Refractory Mutation System PCR (ARMS-PCR) was carried out on extracted DNA samples of subjects to validate the presence of SNP. For each SNP, two separate PCR reactions were carried out. Common primer was added in both the reactions whereas one reaction was carried forward using wildtype primer and other by using mutant primer. If wild type SNP is present in sample only that PCR reaction will amplify DNA that contains wildtype primer and same will be in case of mutant SNP. In this way the presence of SNP will be indicated in samples. A reaction mixture of 20 µl was made in each PCR tube. Reagents of each PCR tube shown in table 3.6.

Table 3.6. Components of PCR reaction mixture.

Component	**Quantity**
PCR Water	12 µl
Forward Primer	1 µl
Reverse Primer	1 µl
DNA Template	2 µl
PCR Master Mix (contains DNA Polymerase, $MgCl_2$, PCR Buffer, dNTPs)	4 µl
Total Volume	20 µl

The reaction mixture was prepared in Biosafety cabinet to avoid contamination and on ice to prevent activation of DNA polymerase enzyme. After preparation, reaction mixture was

short spun for few seconds for removal of air bubbles and homogenous mixing. The PCR tubes were then placed in a Thermocycler for amplification. First the primers were optimized using gradient PCR and then optimized primers were used to amplify sample DNA and detect SNP. The optimized conditions for PCR reaction for PGR (rs1042838), IL10 (rs1800871) and VEZT (rs10859871) are as follow:

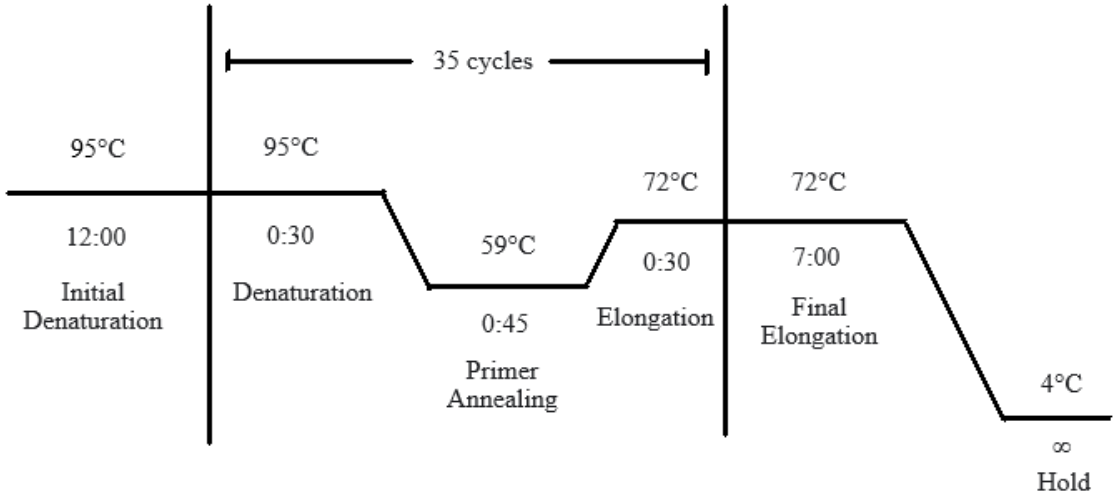

Figure 3.2. Schematic representation of PCR Profile of PGR (rs1042838).

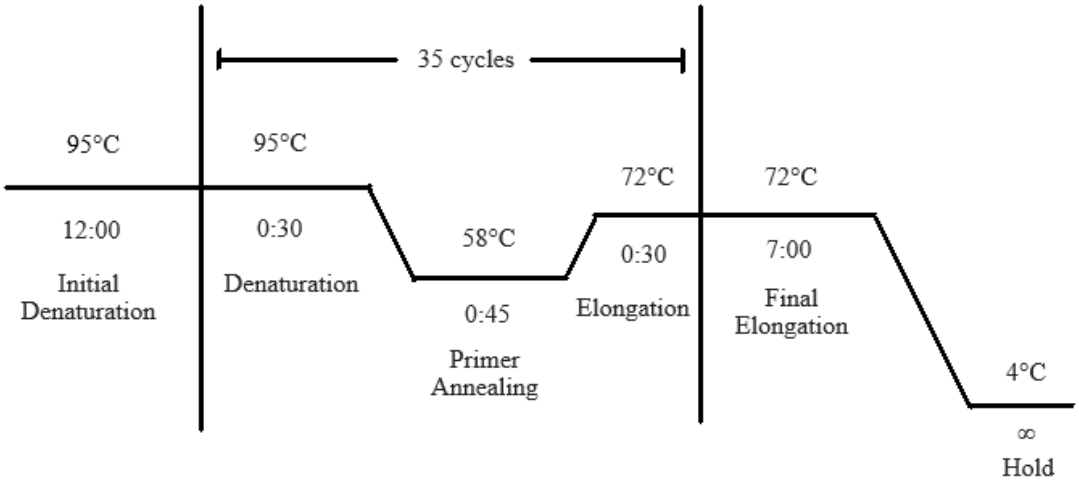

Figure 3.3. Schematic representation of PCR Profile of IL10 (rs1800871).

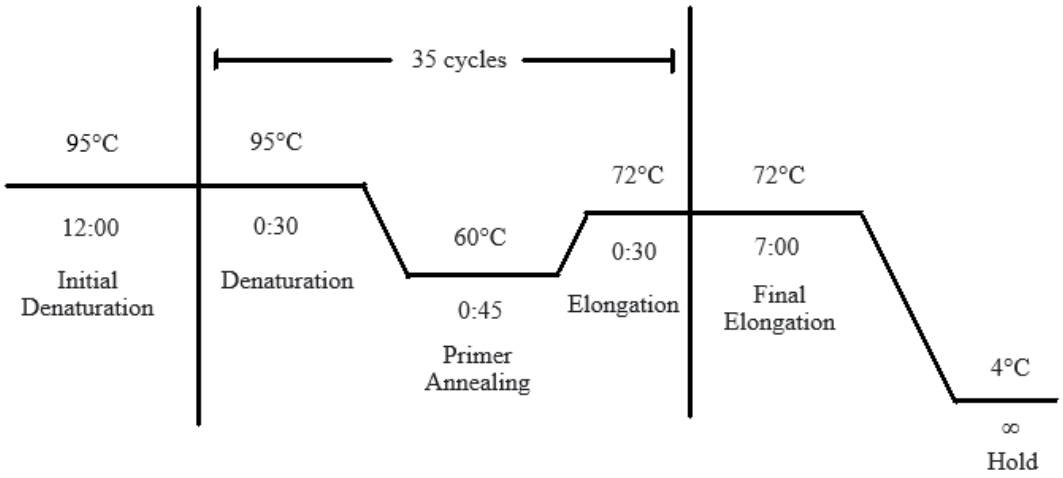

Figure 3.4. Schematic representation of PCR Profile of VEZT (rs10859871).

3.4.10. Gel Electrophoresis for Analysis of PCR Products

For analysis of ARMS PCR products 2 % gel electrophoresis was performed again with the same protocol mentioned before but with a change that 50 bp ladder was used and gel was run for 45 minutes at 80 volts. After that gel was visualized using UV Transilluminator and ChemiDoc and PCR products were analyzed and compared with ladder.

3.4.11. Statistical Analysis

Statistical analysis was performed using GraphPad Prism version 10. Genotype and allele frequencies were calculated and Chi square test x^2 was applied to test the association of SNPs with risk of endometriosis. P values; *P: <0.05, **P: <0.01, ***P: <0.001 were considered as significant.

Chapter 4

4. Results

4.1. In-silico Analysis of Missense SNPs of PGR

4.1.1. Retrieved SNPs of PGR

A total of 38956 SNPs of PGR gene were retrieved from dbSNP. Out of these 34,994 were intronic SNPs, 790 were missense SNPs, 448 were synonymous SNPs while remaining were either regulatory SNPs or non-specified. The percentages of different types of SNPs of PGR are shown in figure 4.1. Missense SNPs were further analyzed to detect deleterious SNPs in PGR gene that can cause disease.

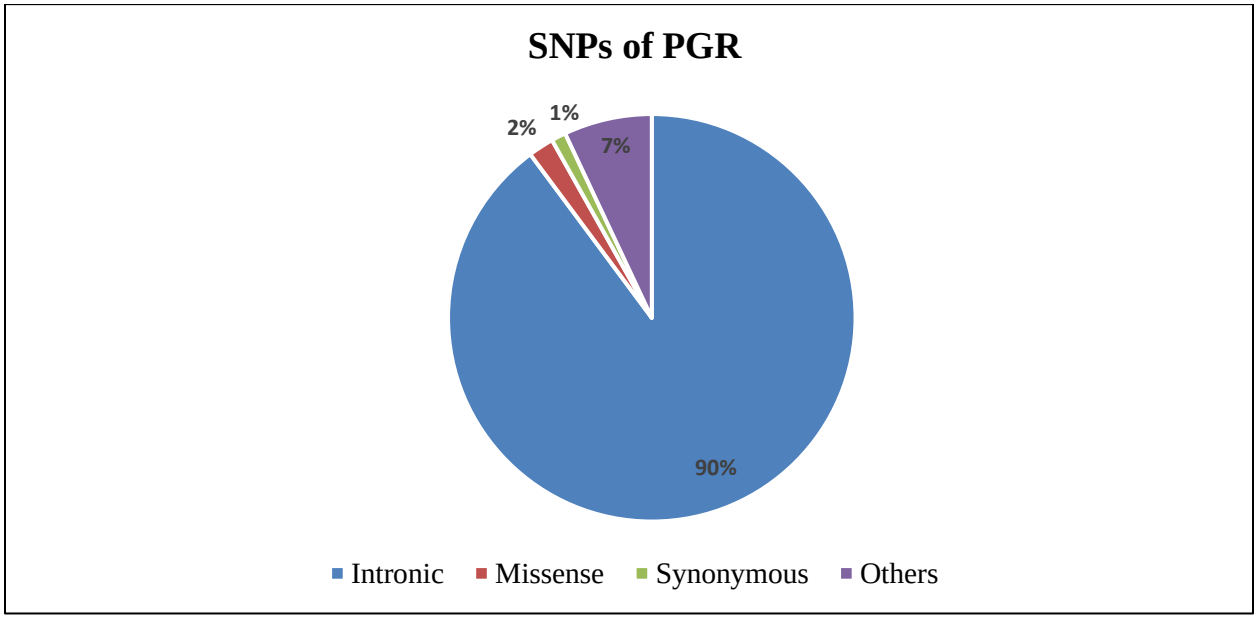

Figure 4.1. Percentage of different types of SNPs of PGR.

4.1.2. Deleterious Missense SNPs of PGR

All the 790 missense SNPs retrieved from dbSNP were first submitted to SIFT. 27 SNPs having SIFT score of less than 0.05 were predicted as deleterious by SIFT as shown in table 4.1. The selected SNP for current study i.e., rs1042838 was predicted as tolerated by SIFT.

Table 4.1. Results of analysis of missense SNPs of PGR by SIFT.

Sr. No.	SNP ID	Allele Change	Amino Acid Change	Sift Score	Prediction
1.	rs2020874	C>A	R625I	0.001	Deleterious
2.	rs2020880	G>A	S865L	0.026	Deleterious

3.	rs11571143	C>T	A50T	0.014	Deleterious
4.	rs11571145	G>A	P186L	0.004	Deleterious
5.	rs111810948	C>G	Q682H	0.02	Deleterious
6.	rs112862471	T>C	I681M	0.003	Deleterious
7.	rs112862471	T>C	I783M	0.003	Deleterious
8.	rs141417204	T>C	K769E	0.026	Deleterious
9.	rs143150313	C>T	M692I	0.042	Deleterious
10.	rs144880156	G>A	L582F	0.004	Deleterious
11.	rs146822380	G>C	T706R	0.008	Deleterious
12.	rs147932536	G>C	D107E	0.047	Deleterious
13.	rs149186732	C>T	R740Q	0.004	Deleterious
14.	rs191163476	C>T	G125R	0.016	Deleterious
15.	rs199796213	G>T	D746E	0	Deleterious
16.	rs199841157	A>G	S20P	0	Deleterious
17.	rs200145151	G>T	P61H	0.003	Deleterious
18.	rs200987911	G>A	R242W	0.008	Deleterious
19.	rs201384017	T>C	H599R	0.03	Deleterious
20.	rs202123793	G>A	A6V	0.008	Deleterious
21.	rs36986385	G>T	S54Y	0.002	Deleterious
22.	rs370034247	C>A	D357Y	0.008	Deleterious
23.	rs371165895	C>T	R788Q	0.041	Deleterious
24.	rs372324844	G>A	R167W	0.002	Deleterious
25.	rs373694750	T>C	K731R	0.045	Deleterious
26.	rs374220674	C>A	S554I	0.049	Deleterious
27.	rs376371918	A>G	I744T	0.002	Deleterious
28.	rs1042838	C>A	V660L	0.336	Tolerated

27 SNPs predicted to be deleterious by SIFT and selected SNP for current study (rs1042838) were then further analyzed by PolyPhen-2, PhD-SNP, SNPs&GO and PANTHER. Out of 27 SNPs Polyphen-2 predicted 15 SNPs to be probably damaging having score between 0.9 to 1 and 4 to be possibly damaging with score ranging from 0.4 to 0.8 shown in table 4.2.

PhD-SNP added 7 SNPs in disease causing category and SNPs&GO predicted only 3 SNPs to be disease causing shown in table 4.3 and 4.4 respectively. PhD-SNP and SNPs&GO give us a reliability index for each SNP and the greater the reliability index, the more the chances of SNP to be pathogenic. PANTHER predicted 5 SNPs to be probably damaging (Pdel>0.5) and 22 to be possibly damaging (Pdel=0.5) shown in table 4.5. Pdel in PANTHER predictions mean probability of deleterious effect, greater the Pdel value more the chance of SNP to be deleterious.

rs1042838 was predicted to be neutral by PolyPhen-2, PhD-SNP and SNPs&GO and possibly damaging by PANTHER. As the majority of *in-silico* tools did not predict it as pathogenic, it was considered as benign having less disease causing potential and was not carried forward for further *in-silico* analysis.

Table 4.2. Results of analysis of missense SNPs of PGR by PolyPhen-2.

Sr. No.	**SNP ID**	**Allele Change**	**Amino Acid Change**	**PolyPhen-2 Score**	**Effect**
1.	rs2020874	C>A	R625I	1	Probably Damaging
2.	rs2020880	G>A	S865L	0.598	Possibly Damaging
3.	rs11571143	C>T	A50T	0.032	Benign
4.	rs11571145	G>A	P186L	0.322	Benign
5.	rs111810948	C>G	Q682H	0.454	Possibly Damaging
6.	rs112862471	T>C	I681M	0.019	Benign
7.	rs112862471	T>C	I783M	1	Probably Damaging
8.	rs141417204	T>C	K769E	0.988	Probably Damaging
9.	rs143150313	C>T	M692I	0.002	Benign
10.	rs144880156	G>A	L582F	1	Probably Damaging
11.	rs146822380	G>C	T706R	0.972	Probably Damaging
12.	rs147932536	G>C	D107E	0.999	Probably Damaging
13.	rs149186732	C>T	R740Q	1	Probably Damaging
14.	rs191163476	C>T	G125R	0.068	Benign
15.	rs199796213	G>T	D746E	0.997	Probably Damaging
16.	rs199841157	A>G	S20P	0.057	Benign
17.	rs200145151	G>T	P61H	0.993	Probably Damaging

18.	rs200987911	G>A	R242W	0.027	Benign
19.	rs201384017	T>C	H599R	0.987	Probably Damaging
20.	rs202123793	G>A	A6V	0.924	Possibly Damaging
21.	rs36986385	G>T	S54Y	0.986	Probably Damaging
22.	rs370034247	C>A	D357Y	0.995	Probably Damaging
23.	rs371165895	C>T	R788Q	0.639	Possibly Damaging
24.	rs372324844	G>A	R167W	1	Probably Damaging
25.	rs373694750	T>C	K731R	0.317	Benign
26.	rs374220674	C>A	S554I	0.999	Probably Damaging
27.	rs376371918	A>G	I744T	0.999	Probably Damaging
28.	rs1042838	C>A	V660L	0.006	Benign

Table 4.3. Results of analysis of missense SNPs of PGR by PhD-SNP.

Sr. No.	**SNP ID**	**Allele Change**	**Amino Acid Change**	**PhD-SNP Effect**	**Reliability Index**
1.	rs2020874	C>A	R625I	Neutral	5
2.	rs2020880	G>A	S865L	Neutral	2
3.	rs11571143	C>T	A50T	Neutral	7
4.	rs11571145	G>A	P186L	Neutral	5
5.	rs111810948	C>G	Q682H	Neutral	2
6.	rs112862471	T>C	I681M	Neutral	6
7.	rs112862471	T>C	I783M	Neutral	3
8.	rs141417204	T>C	K769E	Disease	3
9.	rs143150313	C>T	M692I	Disease	1
10.	rs144880156	G>A	L582F	Neutral	3
11.	rs146822380	G>C	T706R	Disease	0
12.	rs147932536	G>C	D107E	Neutral	8
13.	rs149186732	C>T	R740Q	Disease	0
14.	rs191163476	C>T	G125R	Neutral	4
15.	rs199796213	G>T	D746E	Disease	5
16.	rs199841157	A>G	S20P	Neutral	4

17.	rs200145151	G>T	P61H	Neutral	3
18.	rs200987911	G>A	R242W	Neutral	4
19.	rs201384017	T>C	H599R	Neutral	7
20.	rs202123793	G>A	A6V	Neutral	7
21.	rs36986385	G>T	S54Y	Neutral	4
22.	rs370034247	C>A	D357Y	Disease	0
23.	rs371165895	C>T	R788Q	Neutral	6
24.	rs372324844	G>A	R167W	Neutral	4
25.	rs373694750	T>C	K731R	Neutral	6
26.	rs374220674	C>A	S554I	Neutral	1
27.	rs376371918	A>G	I744T	Disease	4
28.	rs1042838	C>A	V660L	Neutral	6

Table 4.4. Results of analysis of missense SNPs of PGR by SNPs&Go.

Sr. No.	**SNP ID**	**Allele Change**	**Amino Acid Change**	**SNPs&Go Prediction**	**Reliability Index**
1.	rs2020874	C>A	R625I	Neutral	1
2.	rs2020880	G>A	S865L	Neutral	3
3.	rs11571143	C>T	A50T	Neutral	8
4.	rs11571145	G>A	P186L	Neutral	5
5.	rs111810948	C>G	Q682H	Neutral	6
6.	rs112862471	T>C	I681M	Neutral	8
7.	rs112862471	T>C	I783M	Neutral	4
8.	rs141417204	T>C	K769E	Neutral	1
9.	rs143150313	C>T	M692I	Neutral	8
10.	rs144880156	G>A	L582F	Neutral	2
11.	rs146822380	G>C	T706R	Neutral	4
12.	rs147932536	G>C	D107E	Neutral	8
13.	rs149186732	C>T	R740Q	Disease	3
14.	rs191163476	C>T	G125R	Neutral	6
15.	rs199796213	G>T	D746E	Disease	6

16.	rs199841157	A>G	S20P	Neutral	7
17.	rs200145151	G>T	P61H	Neutral	6
18.	rs200987911	G>A	R242W	Neutral	4
19.	rs201384017	T>C	H599R	Neutral	4
20.	rs202123793	G>A	A6V	Neutral	6
21.	rs36986385	G>T	S54Y	Neutral	5
22.	rs370034247	C>A	D357Y	Neutral	3
23.	rs371165895	C>T	R788Q	Neutral	3
24.	rs372324844	G>A	R167W	Neutral	2
25.	rs373694750	T>C	K731R	Neutral	4
26.	rs374220674	C>A	S554I	Neutral	8
27.	rs376371918	A>G	I744T	Disease	0
28.	rs1042838	C>A	V660L	Neutral	9

Table 4.5. Results of analysis of missense SNPs of PGR by PANTHER.

Sr. No.	**SNP ID**	**Allele Change**	**Amino Acid Change**	**PANTHER Prediction**	**Pdel**
1.	rs2020874	C>A	R625I	Possibly Damaging	0.5
2.	rs2020880	G>A	S865L	Possibly Damaging	0.5
3.	rs11571143	C>T	A50T	Possibly Damaging	0.5
4.	rs11571145	G>A	P186L	Possibly Damaging	0.5
5.	rs111810948	C>G	Q682H	Possibly Damaging	0.5
6.	rs112862471	T>C	I681M	Possibly Damaging	0.5
7.	rs112862471	T>C	I783M	Possibly Damaging	0.5
8.	rs141417204	T>C	K769E	Probably Damaging	0.57
9.	rs143150313	C>T	M692I	Possibly Damaging	0.5
10.	rs144880156	G>A	L582F	Possibly Damaging	0.5
11.	rs146822380	G>C	T706R	Possibly Damaging	0.5
12.	rs147932536	G>C	D107E	Possibly Damaging	0.5
13.	rs149186732	C>T	R740Q	Possibly Damaging	0.5
14.	rs191163476	C>T	G125R	Probably Damaging	0.57

15.	rs199796213	G>T	D746E	Probably Damaging	0.85
16.	rs199841157	A>G	S20P	Possibly Damaging	0.5
17.	rs200145151	G>T	P61H	Possibly Damaging	0.5
18.	rs200987911	G>A	R242W	Possibly Damaging	0.5
19.	rs201384017	T>C	H599R	Probably Damaging	0.57
20.	rs202123793	G>A	A6V	Possibly Damaging	0.5
21.	rs36986385	G>T	S54Y	Possibly Damaging	0.5
22.	rs370034247	C>A	D357Y	Possibly Damaging	0.5
23.	rs371165895	C>T	R788Q	Possibly Damaging	0.5
24.	rs372324844	G>A	R167W	Possibly Damaging	0.5
25.	rs373694750	T>C	K731R	Probably Damaging	0.57
26.	rs374220674	C>A	S554I	Possibly Damaging	0.5
27.	rs376371918	A>G	I744T	Possibly Damaging	0.5
28.	rs1042838	C>A	V660L	Possibly Damaging	0.5

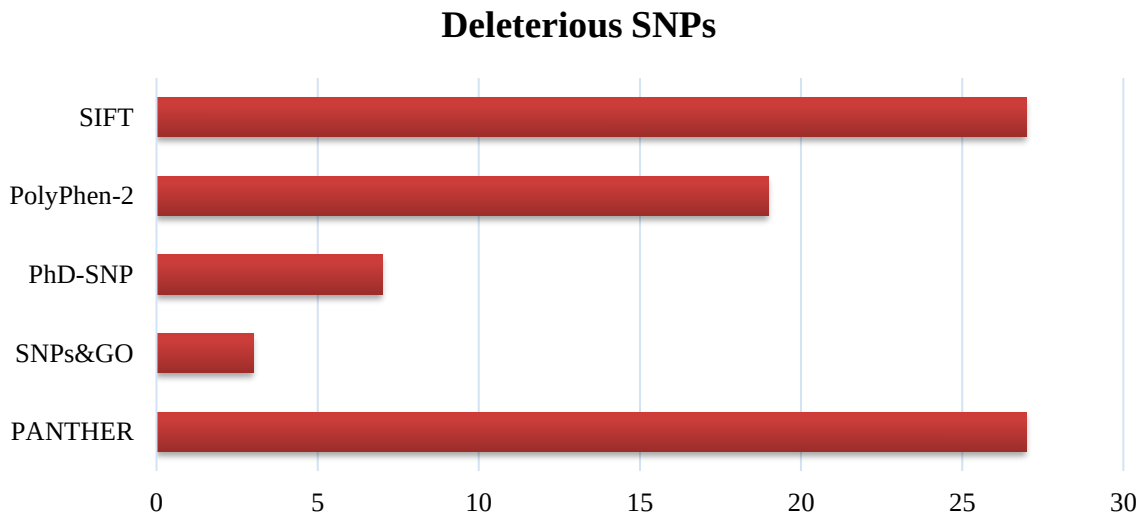

Figure 4.2. Deleterious missense SNPs of PGR predicted by SIFT, PolyPhen-2, PhD-SNP, SNPs&GO and PANTHER.

The missense SNPs predicted to be disease causing by three out of four of the tools i.e., PolyPhen-2, PhD-SNP, SNPs&GO and PANTHER were shortlisted for further analysis and were considered as deleterious missense SNPs of PGR gene as shown in table 4.6.

Table 4.6. Shortlisted deleterious missense SNPs of PGR gene.

Sr. No.	SNP ID	Allele Change	Amino Acid Change	Prediction
1.	rs141417204	T>C	K769E	Deleterious
2.	rs146822380	G>C	T706R	Deleterious
3.	rs149186732	C>T	R740Q	Deleterious
4.	rs199796213	G>T	D746E	Deleterious
5.	rs370034247	C>A	D357Y	Deleterious
6.	rs376371918	A>G	I744T	Deleterious

4.1.3. Structural and Functional Effect of Deleterious Missense SNPs

Four out of six shortlisted deleterious missense SNPs i.e., K769E, D746E, D357Y AND I744T were predicted to have structural and functional effect on protein by MutPred2 as they had MutPred2 score > 0.5 which is considered to be pathogenic. K769E was predicted to alter ordered interface and result in gain in loop, and loss of stand and allosteric site at W765. D357Y was predicted to alter disordered interface and result in gain of sulfation at D357. I744T and D746E were predicted to alter transmembrane protein.

4.1.4. Evolutionary Conservation Analysis

All the amino acids present in the protein sequence of PGR underwent evolutionary conservation analysis using ConSurf. The results for the entire protein sequence are shown in figure 4.3. It provides a conservation score ranging from 1 to 9 for all the amino acid positions, 1 being least conserved and 9 being highly conserved, along with exposed, buried, functional and structural status of residues. The evolutionary status for the six deleterious missense SNPs is shown in table 4.7. Four out of six SNPs are present in highly conserved regions. SNPs present in highly conserved regions can have more deleterious effect as compared to less conserved ones.

Table 4.7. ConSurf prediction for deleterious missense SNPs of PGR.

Sr. No.	SNP ID	Allele Change	Amino Acid Change	Conservation Score	Prediction
1.	rs141417204	T>C	K769E	9	Functional residue (highly conserved and exposed)
2.	rs146822380	G>C	T706R	6	Exposed residue
3.	rs149186732	C>T	R740Q	9	Functional residue (highly

					conserved and exposed)
4.	rs199796213	G>T	D746E	9	Functional residue (highly conserved and exposed)
5.	rs370034247	C>A	D357Y	5	Exposed residue
6.	rs376371918	A>G	I744T	9	Structural residue (highly conserved and buried)

Figure 4.3. Evolutionary Conservation Analysis of PGR.

4.1.5. Effect of Deleterious Missense SNPs on Protein Stability

I-Mutant 2.0 was used to check the effect of deleterious SNPs on protein stability of PGR. The results for all the six shortlisted SNPs along with their ΔΔG value are mentioned in table 4.8.

Table 4.8. Effect of deleterious missense SNPs on protein stability.

Sr. No.	**SNP ID**	**Allele Change**	**Amino Acid Change**	**Stability**	**Reliability Index**	**ΔΔG value**
1.	rs141417204	T>C	K769E	Decrease	1	-0.4
2.	rs146822380	G>C	T706R	Decrease	2	-0.41
3.	rs149186732	C>T	R740Q	Decrease	9	-1.65
4.	rs199796213	G>T	D746E	Increase	0	-0.6
5.	rs370034247	C>A	D357Y	Increase	0	-0.16
6.	rs376371918	A>G	I744T	Decrease	9	-3.55

Positive ΔΔG value (ΔΔG value>0) means protein stability is increased and negative ΔΔG value (ΔΔG>0) means protein stability is decreased. D746E and D357Y were predicted to increase the stability of protein whereas all others decreased stability. It should be noted that as D746E and D357Y increased protein stability they should have positive ΔΔG value but they have negative ΔΔG value. It is because sometimes prediction regarding stability can be contrary to the value of ΔΔG and this happens when reliability index is low, and this is true in case of D746E and D357Y i.e., reliability index=0.

4.1.6. 3D Protein Modelling

3D protein models of wild PGR protein and all its mutant structures of six SNPs shortlisted to be deleterious were designed using I-TASSER. I-TASSER designed five different models for wild protein and each of the six mutants. The model having the highest C-score was selected for further analysis. C-score predicts the overall quality of the protein models, and higher C-score means better quality. C-score of the selected models of wild and mutant proteins of PGR are enlisted in table 4.9.

Figure 4.4. Wild PGR protein structure having 933 amino acid residues.

4.1.7. Validation of 3D Protein Structures

Wild and mutant protein models designed by I-TASSER were validated for good quality using ERRAT. ERRAT calculates an overall quality factor which > 50 is considered acceptable. The ERRAT score for all the models is shown in table 4.9. ERRAT score for all the wild and mutant structures designed was above 70 predicting the good quality of protein models.

Table 4.9. C-score, ERRAT score, TM-score and RMSD values of wild and mutant structures of PGR.

Sr. No.	Model	C-Score	ERRAT Confidence Score	TM-Score	RMSD Value
1.	Wild	-0.21	77.055	-	-
2.	K769E	-0.66	75.346	0.94559	1.30
3.	T706R	-0.73	75.573	0.94568	1.30
4.	R740Q	-0.57	77.739	0.99467	0.82
5.	D746E	-0.25	74.776	0.99675	0.62
6.	D357Y	-0.16	71.644	0.99733	0.54
7.	I744T	-0.18	70.857	0.99609	0.68

4.1.8. Visualization and Superimposition of the Wild and Mutant 3D Protein Structures

The six mutant structures were one by one superimposed on wild protein structure using TM-Align to highlight the differences and deviation of mutant structure from wild. TM-Score and RMSD values predicted by TM-Align are mentioned in table 4.9. TM-score highlights the similarities between wild and mutant structure and its value lies between 0 and 1 where 1 means perfectly aligned proteins and vice versa. All the six mutant structures of PGR had a TM-score

greater than 0.9 predicting that mutant structures are almost like the wild one. RMSD value predicts the structural deviation of mutant structure form its wild and greater the RMSD value, more is the difference between wild and mutant protein structure and greater are the chances of the SNP to be pathogenic affecting structure and function of the protein. Two mutant structures K769E and T706R had the highest RMSD value i.e., 1.30 which increases their chances of being more pathogenic.

Finally, the wild and mutant structures were further superimposed using PyMOL to highlight the change in amino acids showing the difference between two structures. The superimposed wild and mutant structures for K769E, T706R, R740Q, D746E, D357Y and I744T are shown in figure 4.5, 4.6, 4.7. 4.8, 4.9 and 4.10 respectively.

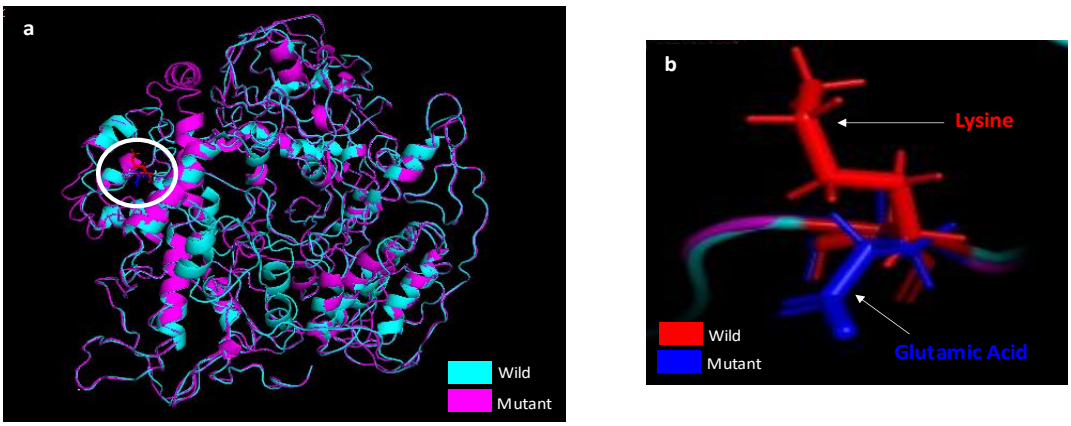

Figure 4.5. a) Superimposition of PGR wild and mutant K769E. b) Highlighted amino acid change from Lysine (wild) to Glutamic acid (mutant).

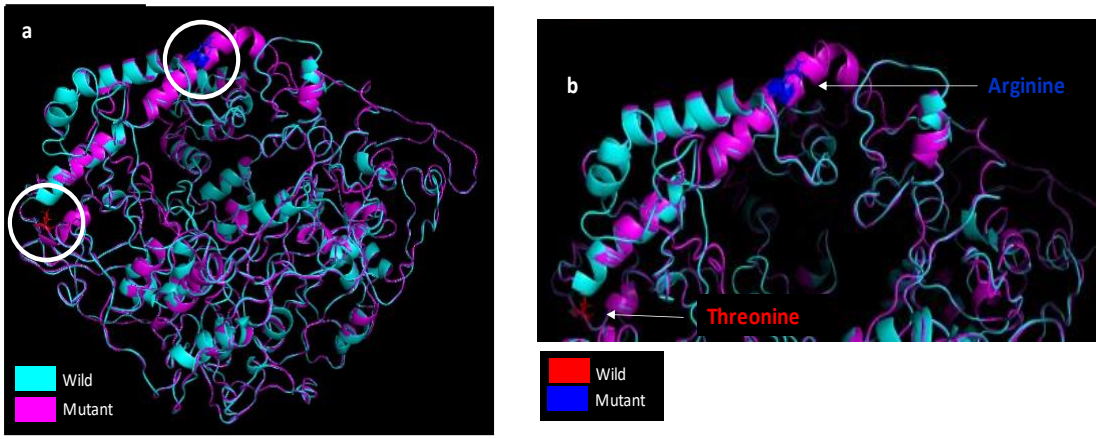

Figure 4.6. a) Superimposition of PGR wild and mutant T706R. b) Highlighted amino acid change from threonine (wild) to Arginine (mutant).

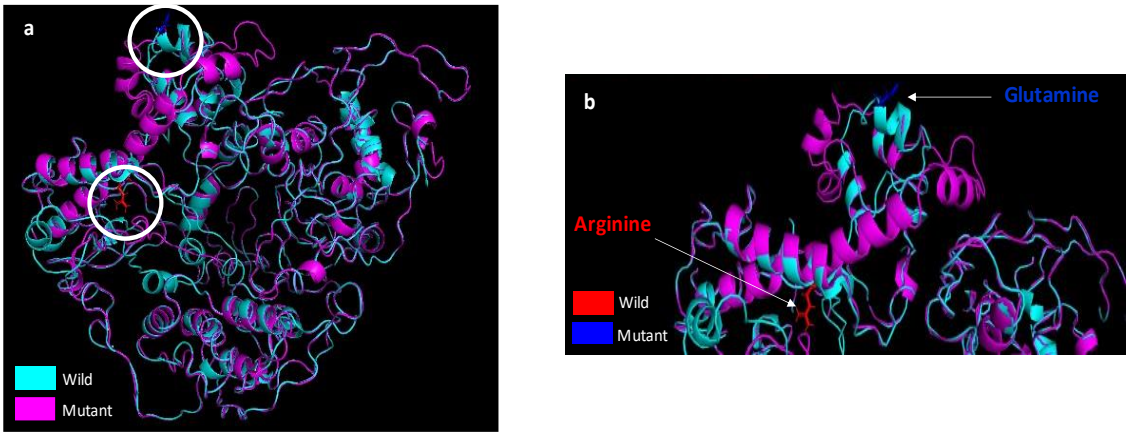

Figure 4.7. a) Superimposition of PGR wild and mutant R740Q. b) Highlighted amino acid change from Arginine (wild) to Glutamine (mutant).

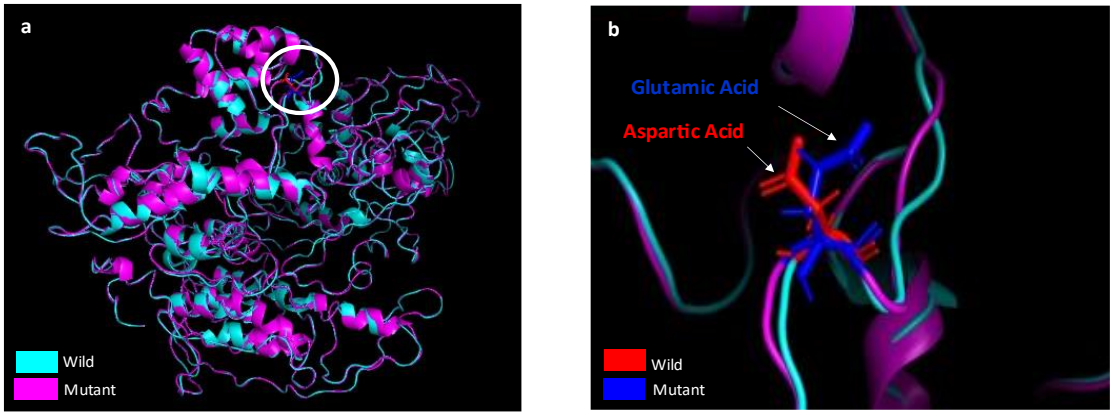

Figure 4.8. a) Superimposition of PGR wild and mutant D746E. b) Highlighted amino acid change from Aspartic acid (wild) to Glutamic acid (mutant).

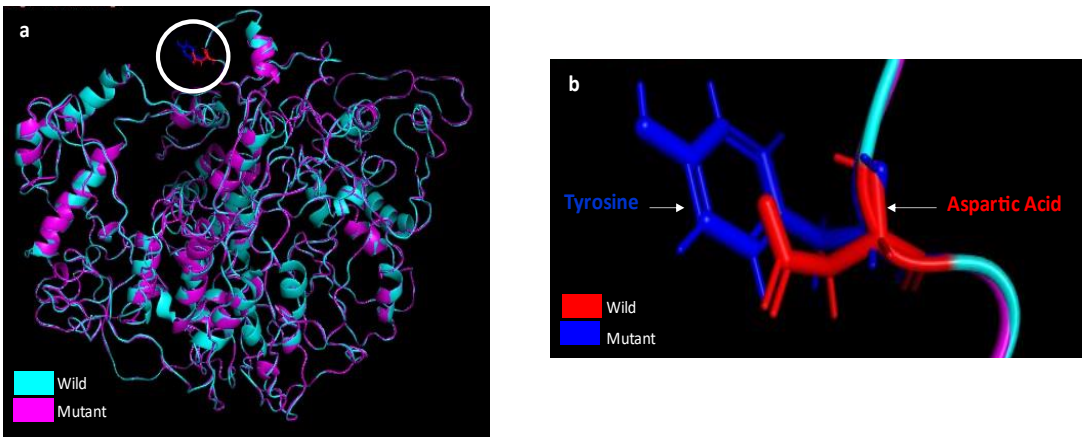

Figure 4.9. a) Superimposition of PGR wild and mutant D357Y. b) Highlighted amino acid change from Aspartic acid (wild) to Tyrosine (mutant).

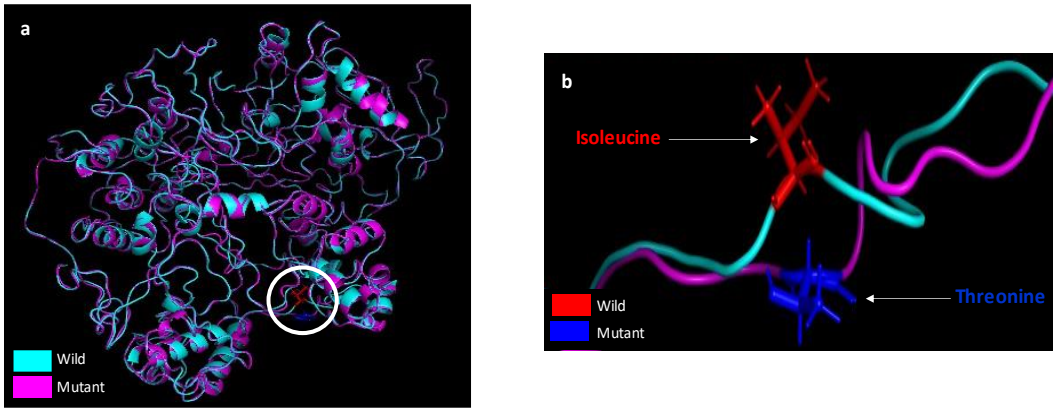

Figure 4.10. a) Superimposition of PGR wild and mutant I744T. b) Highlighted amino acid change from Isoleucine (wild) to Threonine (mutant).

4.1.9. Post Translation Modification (PTM) Analysis

Post Translational Modifications (PTMs) are important for protein signaling, folding, stability, attachment, and regulation. Methylation, phosphorylation, ubiquitination, glycosylation and SUMOylation sites in wild PGR protein and all six mutants were predicted. Moreover, the gain and loss in these PTM sites because of amino acid change at exact location or nearby were also predicted. No methylation, ubiquitination and N-glycosylation sites were present in PGR wild protein and its mutants as predicted by GPS-MSP 1.0, GPS-Uber and NetNGlyc 1.0 respectively. Six SUMOlyation sites were predicted by GPS-SUMO 2.0 in wild protein and same were predicted in all the six mutant proteins at same position without any change. For O-linked glycosylation sites prediction NetOGlyc 4.0 was used, and 82 sites were predicted to be O-glycosylated in wild PGR protein. Mutants K769E and D746E showed gain in O-glycosylation site at position 713, T706R showed gain at position 549 and 714 and loss at position 706 and D357Y exhibited loss of O-glycosylation site at position 456. Gain in O-glycosylation site at position 549 was shown by mutants I744T AND R740Q, with R740Q showing gain at position 713 as well.

In the case of phosphorylation 55 potential sites were predicted by GPS 6.0 in wild protein. Mutant K769E resulted in gain of phosphorylation site at position 768 and 772 and loss of phosphorylation site at position 796. Moreover, there was loss of phosphorylation at position 328 and gain at position 349 and 350 in mutant D357Y as compared to wild. Mutant R740Q showed gain of phosphorylation site at position 733 whereas mutant T706R showed loss of

phosphorylation sites at position 700, 706 and 733. Mutants D746E and I744T showed no change in phosphorylation sites as compared to wild.

4.1.10. Protein-Protein Interaction Analysis

The interaction of PGR protein with different other proteins was checked using STRING. Protein-protein interaction is very crucial for various important biological processes as it regulates different pathways and this interaction if disrupted can cause various diseases. STRING predicted interaction of PGR with 10 other proteins as shown in figure 4.11. PGR interacted the most with SRC having an interaction score of 0.999. It also showed interaction with STAT3, NCOA1, NCOA3 NCOR2, HSPA4, HSP90AA1, FKBP5, MAPK3 and ERBB4. The predicted deleterious SNPs can alter the interaction of PGR with all the mentioned proteins and thus can disturb the normal pathways leading to disease progression.

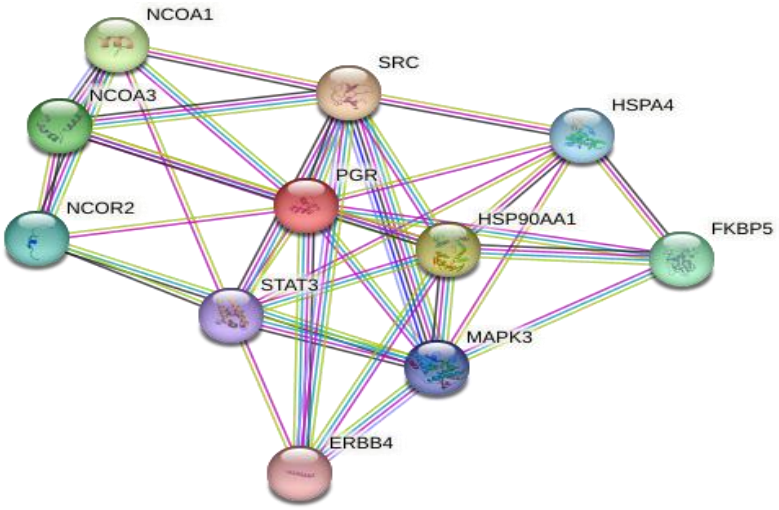

Figure 4.11. Protein-protein interaction of PGR predicted by STRING.

4.2. *In-silico* Analysis of Regulatory SNPs of IL10 and rs10859871 near VEZT

4.2.1. Deleterious 5' and 3' UTR SNPs of IL10

All the 3' and 5' UTR SNPs retrieved from Ensembl were submitted to RegulomeDB for assessment of their effect on transcription factor binding and expression of gene. Only 2 SNPs from UTR-5 region and 1 from UTR-3 region were predicted to have functional effect by RegulomeDB. Afterwards all the SNPs were submitted to SNPinfo (FuncPred) to check their disease-causing potential. A total of 26 SNPs, 12 in UTR-5 region affecting transcription factor binding site and 14 in UTR-3 region affecting miRNA binding site, were predicted by SNPinfo as disease causing SNPs. The SNPs predicted to have functional and disease-causing effect by

both software as shown in table 4.10 had the greatest potential to be pathogenic SNPs leading to progression of various diseases including endometriosis.

Table 4.10. Deleterious regulatory SNPs of IL10 predicted by RegulomeDB and SNPinfo (FuncPred).

Sr. No.	SNP ID	Variant Type	Allele Change	RegulomeDB Score	RegulomeDB Prediction	SNPinfo Prediction
1.	rs1800871	5' UTR	G>A	1b	Likely to affect transcription factor binding and gene expression	Effect on transcription factor biding site
2.	rs1800872	5' UTR	G>T	1f		Effect on transcription factor biding site
3.	rs3024498	3'UTR	T>C	1f		Effect on miRNA binding site

4.2.2. Deleterious effect of rs10859871 near VEZT

Deleterious effect of rs10859871 located 17kb upstream of VEZT was predicted using RegulomeDB and SNPinfo (FuncPred). RegulomeDB gave a score of 1f to rs10859871 predicting it to affect the transcription factor binding and gene expression whereas SNPinfo predicted no functional effect.

4.3. *In-vitro* Analysis

A case control study was conducted to check the association of rs1042838 a missense SNP of PGR, rs1800871 a UTR-5 SNP of Il10 and rs10859871 located approximately 17kb upstream of VEZT with the risk of endometriosis. A total of 38 whole blood samples (19 diseases samples and 19 healthy controls) were collected from Holy Family Hospital Rawalpindi, Benazir Bhutto Hospital Rawalpindi, District Headquarter Hospital Rawalpindi and Research and Diagnostic laboratory ASAB. DNA was extracted from the whole blood samples using Phenol-Chloroform method of DNA extraction. The quality and quantity of extracted DNA was determined through 2% agarose gel electrophoresis and ThermoScientific Nanodrop 2000 UV-Vis Spectrophotometer and NanoDrop 2000™ Software respectively. The pure DNA samples having absorbance ratio ~1.8 were further analyzed by performing allele specific ARMS PCR to screen the samples for the presence of polymorphism. 2% agarose gel electrophoresis was performed to analyze the PCR products. Statistical analysis was performed on the data

obtained from ARMS PCR to check any significant association of the three polymorphisms with endometriosis.

4.3.1. Association Analysis of PGR rs1042838 Polymorphism

PGR rs1042838 is a missense SNP (C>A), where C is the ancestral allele and A is the risk allele. It is present in exon 4 of the gene on chromosome 11q22.1 and results in a single amino acid change from valine to leucine (V660L). For this SNP, 2% agarose gel electrophoresis results showing PCR products (amplicon size=362 bp) of diseased samples are shown in figure 4.12 and of healthy controls in figure 4.13. The Chi square x^2 test values as well as observed genotype and allele frequencies are mentioned in table 4.10. Moreover, graphs for genotype and allele frequency of experimental vs control group are shown in figure 4.14 and 4.15 respectively. There was no significant difference observed between the frequency of ancestral C allele and risk A allele in experimental vs control group. Ancestral Allele C was more prevalent in both control and case groups as compared to risk allele A.

The Chi square x^2 test was performed to check the association of rs1042838 with endometriosis using GraphPad Prism version 10. The Chi square value obtained i.e., $x^2 = 2.143$ with p-value of 0.3425 and degree of freedom 2 indicates that the observed genotypes are in Hardy-Weinberg equilibrium with expected genotypes and there is no significant association of the SNP rs1042838 of PGR gene with risk of endometriosis.

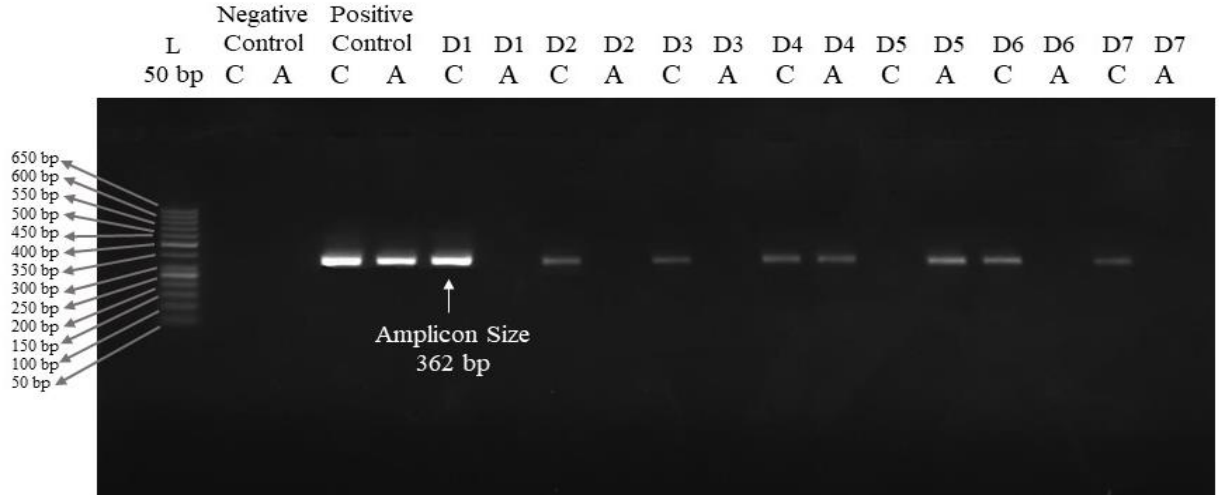

Figure 4.12. 2% agarose gel showing PCR products of diseased samples for PGR rs1042838 polymorphism.

L=ladder, D= diseased samples (endometriosis). Two bands indicate that the study subject is heterozygous having both alleles whereas one band indicates study subject is homozygous having either C or A allele.

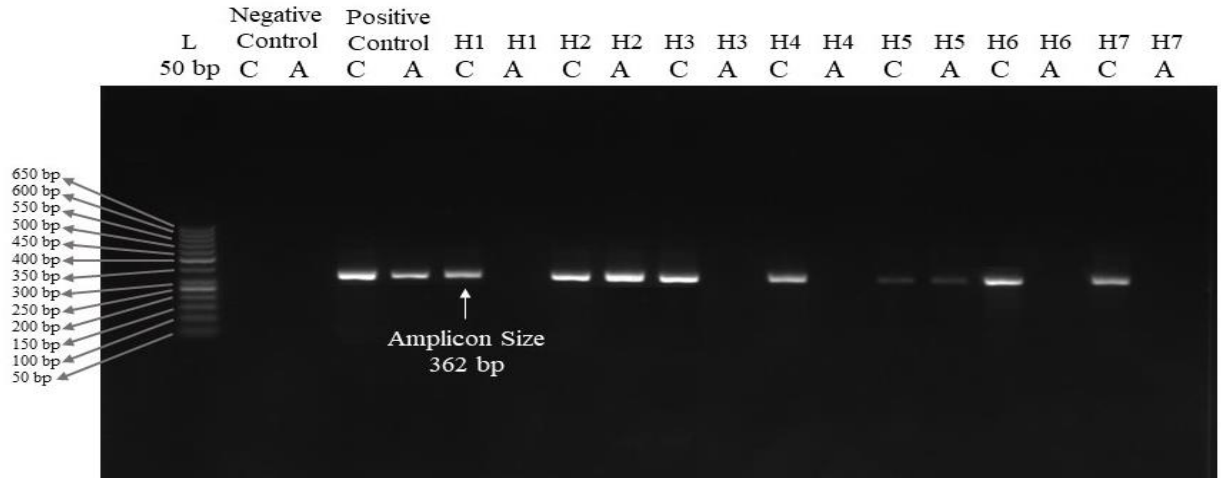

Figure 4.13. 2% agarose gel showing PCR products of healthy controls for PGR rs1042838 polymorphism.

L=ladder, H= healthy controls. Two bands indicate that the study subject is heterozygous having both alleles whereas one band indicates study subject is homozygous having either C or A allele.

Table 4.11. Chi square x2 test, genotype, and allele frequencies of PGR rs1042838 polymorphism.

Cases/Controls PGR	**Genotype Frequency**						**Allele Frequency**			
	CC		**CA**		**AA**		**C**		**A**	
	n	**%**	**N**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Cases n=19 (Endometriosis Patients)	15	78.9	3	15.8	1	5.3	33	86.8	5	13.2
Controls n=19 (Healthy Females)	13	68.4	6	31.6	0	0	32	84.2	6	15.8
x^2	2.143						0.1063			
p-value	0.3425						0.7444			
Df	2						1			

Note: x^2= Chi Square test value, p-value= probability value, Df= Degree of freedom

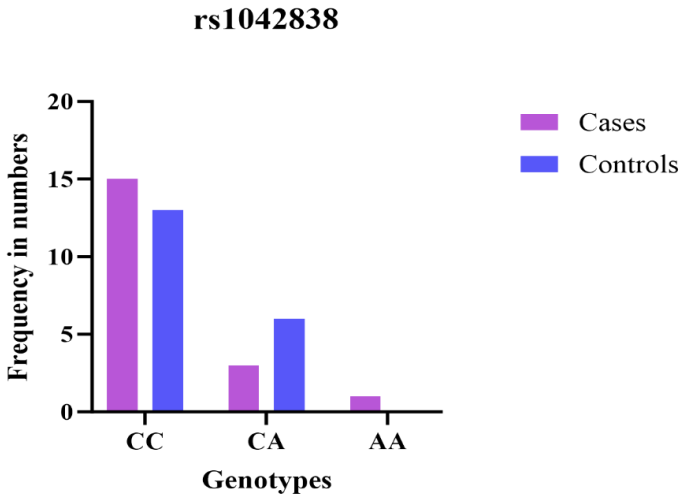

Figure 4.14. Genotype distribution of PGR rs1042838 polymorphism.

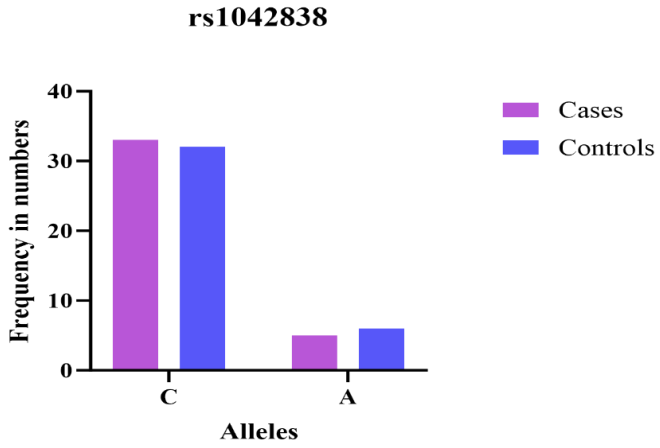

Figure 4.15. Allele distribution of PGR rs1042838 polymorphism.

4.3.2. Association Analysis of IL10 UTR-5 rs1800871 Polymorphism

rs1800871 (G>A), where G is ancestral allele and A is the risk allele, is in 5' UTR region of IL10 gene on chromosome 1q32.1. For this SNP, 2% agarose gel electrophoresis results showing PCR products (amplicon size=392 bp) of diseased samples are shown in figure 4.16 and of healthy controls in figure 4.17. The Chi square x^2 test values, observed genotype and allele frequencies are mentioned in table 4.11. In addition, graphs for genotype and allele frequency of experimental vs control group are shown in figures 4.18 and 4.19 respectively. There was no significant difference observed between the frequency of ancestral G allele and risk A allele in experimental vs control group. Both alleles in heterozygous form GA were more prevalent in both disease and control samples as compared to homozygous GG or AA alleles.

The Chi square x^2 test was performed to check the association of rs1800871 with endometriosis using GraphPad Prism version 10. The Chi square value obtained i.e., $x^2 = 0.2345$ with p-value of 0.8894 and degree of freedom 2 indicates that the observed genotypes are in Hardy-Weinberg equilibrium with expected genotypes and there is no significant association of the SNP rs1800871 in UTR-5 if IL10 gene with risk of endometriosis.

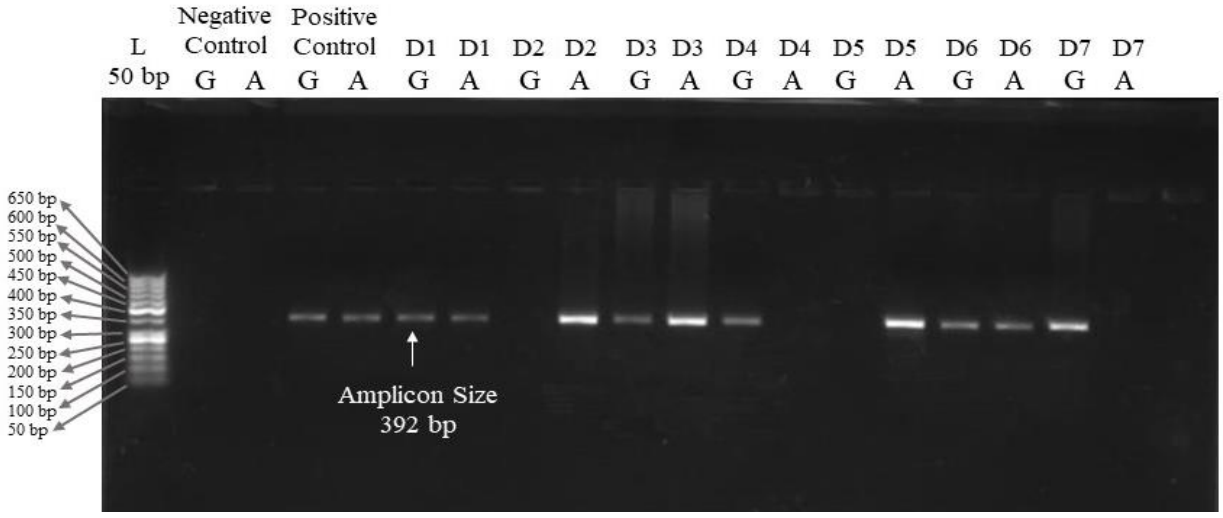

Figure 4.16. 2% agarose gel showing PCR products of diseased samples for IL10 UTR-5 rs1800871 polymorphism.

L=ladder, D= diseased samples (endometriosis). Two bands indicate that the study subject is heterozygous having both alleles whereas one band indicates study subject is homozygous having either G or A allele.

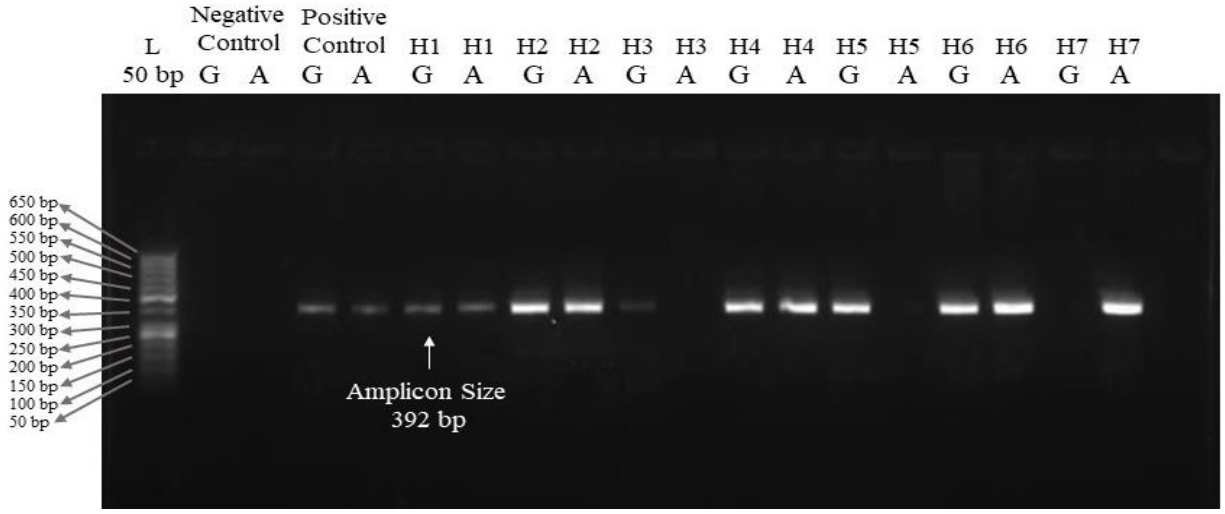

Figure 4.17. 2% agarose gel showing PCR products of healthy controls for IL10 UTR-5 rs1800871 polymorphism.

L=ladder, H= healthy controls. Two bands indicate that the study subject is heterozygous having both alleles whereas one band indicates study subject is homozygous having either G or A allele.

Table 4.12. Chi square x2 test, genotype, and allele frequencies of IL10 UTR-5 rs1800871 polymorphism.

Cases/Controls IL10	**Genotype Frequency**						**Allele Frequency**			
	GG		**GA**		**AA**		**G**		**A**	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Cases n=19 (Endometriosis Patients)	2	10.5	15	79	2	10.5	19	50	19	50
Controls n=19 (Healthy Females)	3	15.8	14	73.7	2	10.5	20	52.6	18	47.4
x^2	0.2345						0.05267			
p-value	0.8894						0.8185			
Df	2						1			

Note: x^2= Chi Square test value, p-value= probability value, Df= Degree of freedom

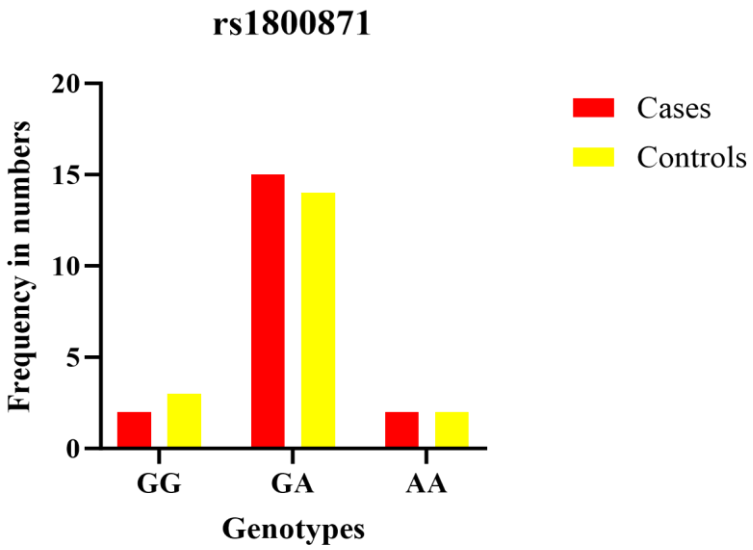

Figure 4.18. Genotype distribution of IL10 UTR-5 rs1800871 polymorphism.

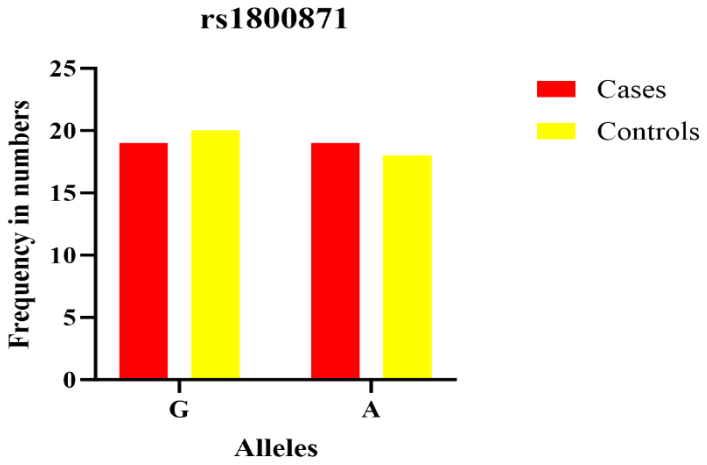

Figure 4.19. Allele distribution of IL10 UTR-5 rs1800871 polymorphism.

4.3.3. Association Analysis of rs10859871 Polymorphism near VEZT

rs10859871 (A>C), where A is ancestral allele and C is the risk allele, is located approximately 17kb upstream of VEZT gene on chromosome 12. For this SNP, 2% agarose gel results showing PCR products (amplicon size=221 bp) of diseased samples are shown in figure 4.20 and of healthy controls in figure 4.21. The Chi square x^2 test values as well as observed genotype and allele frequencies are mentioned in table 4.12. Graphs for genotype and allele frequency of experimental vs control group are shown in figure 4.22 and 4.23 respectively. There was no significant difference observed between the frequency of ancestral A allele and risk C allele in experimental vs control group. CA genotype in heterozygous form was prevalent in both cases and controls as compared to homozygous CC and AA genotypes. Moreover, CC genotype was observed slightly more in cases as compared to controls.

The Chi square x^2 test was performed to check the association of rs10859871 with endometriosis using GraphPad Prism version 10. The Chi square value obtained i.e., $x^2 = 2.095$ with p-value of 0.3508 and d 2 indicates that the Hardy Weinberg equilibrium conditions have been met and there is no significant association of the SNP rs10859871 near VEZT gene with susceptibility to endometriosis.

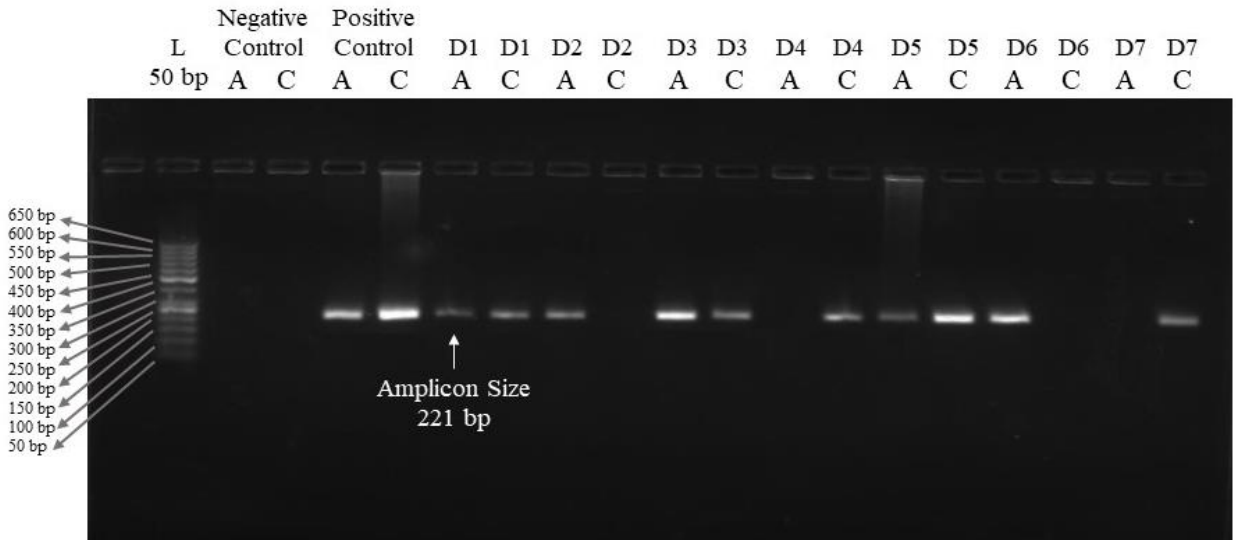

Figure 4.20. 2% agarose gel showing PCR products of diseased samples for rs10859871 polymorphism near VEZT.

L=ladder, D= diseased samples (endometriosis). Two bands indicate that the study subject is heterozygous having both alleles whereas one band indicates study subject is homozygous having either A or C allele.

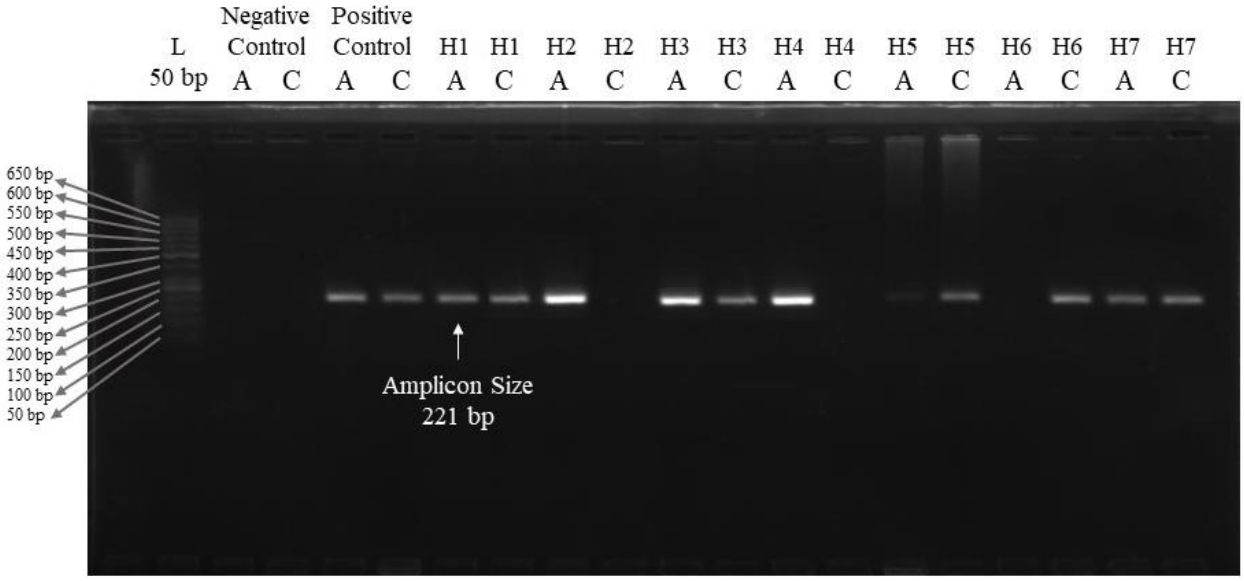

Figure 4.21. 2% agarose gel showing PCR products of healthy controls for rs10859871 polymorphism near VEZT.

L=ladder, H= healthy controls. Two bands indicate that the study subject is heterozygous having both alleles whereas one band indicates study subject is homozygous having either A or C allele.

Table 4.13. Chi square x2 test, genotype, and allele frequencies of rs10859871 polymorphism near VEZT.

Cases/Controls VEZT	**Genotype Frequency**						**Allele Frequency**			
	AA		AC		CC		A		C	
	N	%	n	%	n	%	N	%	n	%
Cases n=19 (Endometriosis Patients)	4	21.1	10	52.6	5	26.3	18	47.4	20	52.6
Controls n=19 (Healthy Females)	3	15.8	14	73.7	2	10.5	20	52.6	18	47.4
x^2	2.095						0.2105			
p-value	0.3508						0.6464			
Df	2						1			

Note: x^2= Chi Square test value, p-value= probability value, Df= Degree of freedom

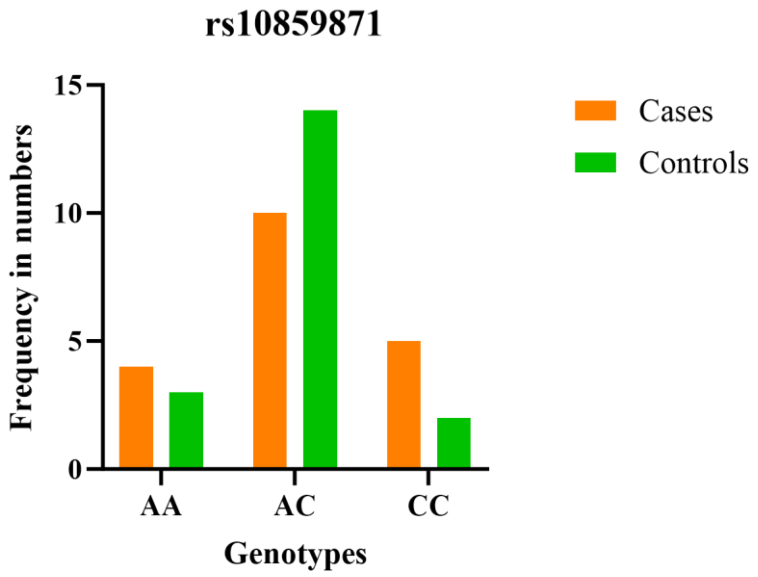

Figure 4.22. Genotype distribution of rs10859871 polymorphism near VEZT.

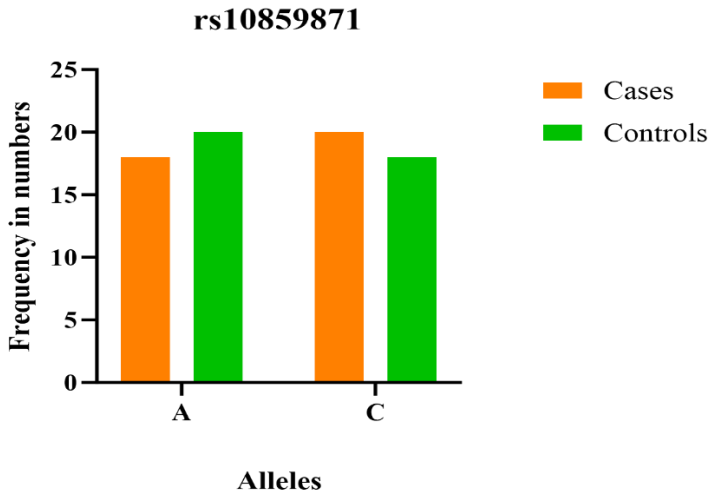

Figure 4.23. Allele distribution of rs10859871 polymorphism near VEZT.

Chapter 5

5. Discussion

Endometriosis is a chronic reproductive disorder characterized by the growth and proliferation of endometrial tissue in different parts of the body other than uterus. Approximately 10% of the women worldwide (Mohammed Rasheed & Hamid, 2020) and 6% of the women in Pakistan (Khan & Illahi, 2000) suffer from this disease and it is most prevalent between ages of 25 to 45 years (Smolarz et al., 2021). However, the exact prevalence cannot be estimated due to the late diagnosis and poor management of the disease. Endometriosis is recognized as a disease with complex and unclear etiology involving retrograde menstruation, endometrial cell transport as well as various genetic, epigenetic, hormonal, and immunological factors. The progression of endometriosis lead to severe symptoms, pelvic pain and infertility being the major ones (Soliman et al., 2017).

Endometriosis is believed to have a strong genetic basis. Studies reveal polygenic inheritance pattern of endometriosis (Hansen & Eyster, 2010) and involvement of single nucleotide polymorphisms in development of disease. Candidate gene analysis and GWAS studies have reported involvement of various genes in progression of endometriosis and about 19 SNPs are identified to be associated with the disease (Sapkota et al., 2017).

In the current study, out of SNPs of all the genes reported to be associated with endometriosis, SNPs located in or near three genes i.e., a missense SNP rs1042838 of PGR, a UTR-5 SNP rs1800871 of IL10 and rs10859871 located 17kb upstream of VEZT were selected to check their role in endometriosis. All the three SNPs are reported to be associated with endometriosis according to studies conducted in different parts of the world, but some studies also nullified their role in endometriosis.

In the case of PGR since the selected SNP was a non-synonymous exonic SNP, an *in-silico* analysis was performed to filter all the missense SNPs of PGR for the deleterious and disease-causing effect. A total of 790 missense SNPs were retrieved from NCBI, and these SNPs were tested for deleterious effect first by SIFT and then by using four other tools i.e., Polyhen-2, PhD-SNP, SNPs&GO and PANTHER. The SNP rs1042838 (V660L) selected for association analysis with endometriosis in current study was predicted to be neutral by all the SNP filtering software except PANTHER. So overall rs1042838 was considered as benign and was not carried

forward for further *in-silico* analysis. However, six different missense SNPs of PGR gene i.e., rs141417204 (K769E), rs146822380 (T706R), rs149186732 (R740Q), rs199796213 (D746E), rs370034247 (D357Y) and rs376371918 (I744T) were predicted to be deleterious and disease causing by at most three of the above-mentioned software along with SIFT.

The six SNPs predicted to be damaging were further analyzed to see their effect on structure and function and most of them can alter protein structure or function as predicted by Mutpred2. Four out of the six SNPs filtered with exception of T706R and D357Y were in highly conserved region as predicted by Consurf which increases their chance of being deleterious and disease causing as amino acids located in highly conserved regions are most crucial for normal functioning of protein and changes to these amino acids can cause damaging effects. K769E, R740Q and D746E were predicted to affect the function of protein whereas I744T can have structural effects. Moreover, K769E, T706R, R740Q, I744T were predicted to decrease the stability of protein by I-Mutant 2.0. The protein structure of wild and all six mutants were obtained using I-TASSER, were validated as good quality models by ERRAT and all the mutants were superimposed using TM-Align on wild protein to predict the difference in structure. TM-score of superimpositions of all the mutants was greater than 0.9 predicting a very slight difference in structure of wild and mutant protein structures. This highlights that the six deleterious SNPs do not result in drastic change in structure of protein and are more inclined towards effecting the function of protein. However, RMSD value of K769E and T706R was predicted as 1.30 by TM-Align making them the topmost candidates capable of having damaging effect on protein structure and function.

PTM analysis predicted no potential sites for methylation, ubiquitination, and N-glycosylation in wild and mutant protein structures. Six SUMOlyation sites were predicted but no change was observed due to presence of SNPs. 82 O-glycosylation and 55 phosphorylation sites were predicted in wild protein. Amino acid change due to all the six deleterious SNPs resulted in gain of O-glycosylation and phosphorylation sites at certain position and loss as well. This proves that all the six deleterious SNPs are not only damaging to structure and function of protein but can also alter PTM sites affecting protein signaling, folding, stability, attachment, and regulation. Thus, the six filtered SNPs may have potential role in progression of various diseases including endometriosis, but in-vitro testing is required to validate the results.

In-silico analysis of regulatory SNPs lying in UTR-5 and UTR-3 region of IL10 and rs10859871 near VEZT was also performed using two online tools RegulomeDB and SNPinfo (Funcpred). Out of all the regulatory SNPs of IL10, three were predicted to be deleterious by both web servers including the one selected for current study i.e., rs1800871 along with rs1800872 in UTR-5 region and rs3024498 in UTR-3 region. All the three SNPs were predicted to be capable of altering the gene expression by affecting the transcription factor binding site in case of rs1800871 and rs1800872 and by affecting the miRNA binding site in case of rs3024498. Moreover, rs10859871, an intronic SNP located near VEZT, was predicted to affect gene expression and thus categorized as deleterious by RegulomeDB only.

In *in-vitro* analysis, a case control study was carried out to check the role of three selected SNPs i.e., rs1042838, rs1800871 and rs10859871 in endometriosis. Allele specific ARMS PCR was performed to screen the samples, collected from various hospitals of Islamabad and Rawalpindi, in comparison with healthy controls for the presence of polymorphism. The results were then statistically analyzed using GraphPad Prism version 10 for any significant association with endometriosis. No significant difference was observed in frequency of ancestral and risk allele in both experimental and control groups for all the three SNPs. Moreover, the Chi square test confirmed that the observed genotypes were in Hardy Weinberg equilibrium with expected genotypes and there was no significant association of the three selected SNPs with risk of endometriosis.

These results were contrary to the findings of studies which reported a significant association of risk allele A of rs1042838 of PGR and rs1800871 of IL10 (Liaqat et al., 2013) as well as risk allele C of rs10859871 (Holdsworth-Carson et al., 2016) (Pshenichnyuk et al., 2017) with endometriosis. *In-silico* tools also predicted rs1800871 and rs10859871 as pathogenic but in the case of rs1042838, our results were consistent with the *in-silico* predictions. Moreover, a study conducted in Australia also reported no association of rs1042838 (Treloar et al., 2005) and studies conducted on Mediterranean (Angioni et al., 2020) and Polish Caucasian (Osiński, Mostowska, et al., 2018) populations reported no association of rs10859871 near VEZT with higher risk of endometriosis. One of the reasons for no significant association might be small sample size because of late diagnosis, less awareness and poor management of disease. A study designed to replicate our research over a larger sample size can further validate the results.

Conclusion

The present study demonstrates no significant association of rs1042838 (a missense SNP of PGR), rs1800871 (a UTR-5 SNP of IL10) and rs10859871 (located 17kb upstream of VEZT) with endometriosis in our studied samples. This result was consistent with our *in-silico* findings in case of rs1042838 which was predicted to be benign but was contrary in case of rs1800871 and rs10859871 which were predicted to be deleterious. The present results need to be further validated by replication of this research in larger sample population. Moreover, the *in-silico* analysis predicted additional polymorphisms located in exonic region of PGR and 5' and 3' UTR region of IL10 having disease causing potential as they can affect the normal structure and function of the respective proteins. These deleterious SNPs can be involved in the risk of endometriosis and pathogenesis of other diseases as well, but *in-vitro* validation is required for confirmation of these results.

Future Perspectives

For better understanding and more significant results the current study could be replicated with larger sample size and can be further taken forward for expression analysis and gene sequencing to confirm the role of PGR, IL10 and VEZT gene polymorphisms in endometriosis. Furthermore, extensive studies highlighting the clear role of PGR, IL10 and VEZT in progression of endometriosis can be helpful. The potential role of pathogenic SNPs of IL10 and PGR, filtered via *in-silico* analysis, in endometriosis and other related diseases can be tested via *in-vitro* analysis and if validated can be used for personalized medicines and prognostic and diagnostic markers for respective diseases. Moreover, they can be used to develop screening tests for diagnosis and can be drug targets for various therapeutic strategies.

References

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Alagarasu, K., Kaushal, H., Shinde, P., Kakade, M., Chaudhary, U., Padbidri, V., Sangle, S. A., Salvi, S., Bavdekar, A. R., D'costa, P., & Choudhary, M. L. (2021). TNFA and IL10 Polymorphisms and IL-6 and IL-10 Levels Influence Disease Severity in Influenza A(H1N1)pdm09 Virus Infected Patients. *Genes*, *12*(12), Article 12. https://doi.org/10.3390/genes12121914

André, G. M., Martins Trevisan, C., Pedruzzi, I. N., Fernandes, R. F. M., Oliveira, R., Christofolini, D. M., Bianco, B., & Barbosa, C. P. (2018). The Impact of FSHR Gene Polymorphisms Ala307Thr and Asn680Ser in the Endometriosis Development. *DNA and Cell Biology*, *37*(6), 584–591. https://doi.org/10.1089/dna.2017.4093

Angioni, S., D'Alterio, M. N., Coiana, A., Anni, F., Gessa, S., & Deiana, D. (2020). Genetic Characterization of Endometriosis Patients: Review of the Literature and a Prospective Cohort Study on a Mediterranean Population. *International Journal of Molecular Sciences*, *21*(5), Article 5. https://doi.org/10.3390/ijms21051765

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., & Ben-Tal, N. (2016). ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research*, *44*(W1), W344–W350. https://doi.org/10.1093/nar/gkw408

Bahloul, A., Simmler, M.-C., Michel, V., Leibovici, M., Perfettini, I., Roux, I., Weil, D., Nouaille, S., Zuo, J., Zadro, C., Licastro, D., Gasparini, P., Avan, P., Hardelin, J.-P., & Petit, C. (2009). Vezatin, an integral membrane protein of adherens junctions, is required for the sound resilience of cochlear hair cells. *EMBO Molecular Medicine*, *1*(2), 125–138. https://doi.org/10.1002/emmm.200900015

Beheshtizadeh, M., & Moslemi, E. (2017). Analysis of G3BP1 and VEZT Expression in Gastric Cancer and Their Possible Correlation with Tumor Clinicopathological Factors. *Journal of Gastric Cancer*, *17*(1), 43–51. https://doi.org/10.5230/jgc.2017.17.e5

Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., Karczewski, K. J., Park, J., Hitz, B. C., Weng, S., Cherry, J. M., & Snyder, M. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research*, *22*(9), 1790–1797. https://doi.org/10.1101/gr.137323.112

Broi, M. G. D., Ferriani, R. A., & Navarro, P. A. (2019). Ethiopathogenic mechanisms of endometriosis-related infertility. *JBRA Assisted Reproduction*, *23*(3), 273–280. https://doi.org/10.5935/1518-0557.20190029

Brosens, I., & Benagiano, G. (2013). Is neonatal uterine bleeding involved in the pathogenesis of endometriosis as a source of stem cells? *Fertility and Sterility*, *100*(3), 622–623. https://doi.org/10.1016/j.fertnstert.2013.04.046

Bulun, S. E., Yilmaz, B. D., Sison, C., Miyazaki, K., Bernardi, L., Liu, S., Kohlmeier, A., Yin, P., Milad, M., & Wei, J. (2019). Endometriosis. *Endocrine Reviews*, *40*(4), 1048–1079. https://doi.org/10.1210/er.2018-00242

Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P. L., & Casadio, R. (2009). Functional annotations improve the predictive score of human disease-related mutations in proteins. *Human Mutation*, *30*(8), 1237–1244. https://doi.org/10.1002/humu.21047

Capriotti, E., Calabrese, R., & Casadio, R. (2006). Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics*, *22*(22), 2729–2734. https://doi.org/10.1093/bioinformatics/btl423

Capriotti, E., Fariselli, P., & Casadio, R. (2005). I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research*, *33*(Web Server issue), W306–W310. https://doi.org/10.1093/nar/gki375

Cardoso, J. V., Perini, J. A., Machado, D. E., Pinto, R., & Medeiros, R. (2020). Systematic review of genome-wide association studies on susceptibility to endometriosis. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *255*, 74–82. https://doi.org/10.1016/j.ejogrb.2020.10.017

Chen, L.-C., Hsu, J.-W., Huang, K.-L., Bai, Y.-M., Su, T.-P., Li, C.-T., Yang, A. C., Chang, W.-H., Chen, T.-J., Tsai, S.-J., & Chen, M.-H. (2016). Risk of developing major depression and anxiety disorders among women with endometriosis: A longitudinal follow-up study. *Journal of Affective Disorders*, *190*, 282–285. https://doi.org/10.1016/j.jad.2015.10.030

Chen, M., Zhang, W., Gou, Y., Xu, D., Wei, Y., Liu, D., Han, C., Huang, X., Li, C., Ning, W., Peng, D., & Xue, Y. (2023). GPS 6.0: An updated server for prediction of kinase-specific phosphorylation sites in proteins. *Nucleic Acids Research*, *51*(W1), W243–W250. https://doi.org/10.1093/nar/gkad383

Colafrancesco, S., Ciccacci, C., Priori, R., Latini, A., Picarelli, G., Arienzo, F., Novelli, G., Valesini, G., Perricone, C., & Borgiani, P. (2019). *STAT4*, *TRAF3IP2*, *IL10*, and *HCP5* Polymorphisms in Sjögren's Syndrome: Association with Disease Susceptibility and Clinical Aspects. *Journal of Immunology Research*, *2019*, e7682827. https://doi.org/10.1155/2019/7682827

Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science : A Publication of the Protein Society*, *2*(9), 1511–1519.

Dalsgaard, T., Hjordt Hansen, M. V., Hartwell, D., & Lidegaard, O. (2013). Reproductive prognosis in daughters of women with and without endometriosis. *Human Reproduction (Oxford, England)*, *28*(8), 2284–2288. https://doi.org/10.1093/humrep/det231

Danglot, L., Freret, T., Le Roux, N., Nême, N. N., Burgo, A., Hyenne, V., Roumier, A., Contremoulins, V., Dauphin, F., Bizot, J.-C., Vodjdani, G., Gaspar, P., Boulouard, M., Poncer, J.-C., Galli, T., & Simmler, M.-C. (2012). Vezatin Is Essential for Dendritic Spine Morphogenesis and Functional Synaptic Maturation. *The Journal of Neuroscience*, *32*(26), 9007–9022. https://doi.org/10.1523/JNEUROSCI.3084-11.2012

Datta, A., Tuz Zahora, F., Abdul Aziz, Md., Sarowar Uddin, M., Ferdous, M., Shalahuddin Millat, Md., Shahid Sarwar, Md., Abdul Barek, Md., Jafrin, S., Nahar, S., & Safiqul Islam, M. (2020). Association study of IL10 gene polymorphisms (rs1800872 and rs1800896) with cervical cancer in the Bangladeshi women. *International Immunopharmacology*, *89*, 107091. https://doi.org/10.1016/j.intimp.2020.107091

Deng, W., Wang, Y., Ma, L., Zhang, Y., Ullah, S., & Xue, Y. (2017). Computational prediction of methylation types of covalently modified lysine and arginine residues in proteins. *Briefings in Bioinformatics*, *18*(4), 647–658. https://doi.org/10.1093/bib/bbw041

Engelhardt, K. R., & Grimbacher, B. (2014). IL-10 in Humans: Lessons from the Gut, IL-10/IL-10 Receptor Deficiencies, and IL-10 Polymorphisms. In S. Fillatreau & A. O'Garra (Eds.), *Interleukin-10 in Health and Disease* (Vol. 380, pp. 1–18). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43492-5_1

Ferdosian, F., Dastgheib, S. A., Morovati-Sharifabad, M., Lookzadeh, M. H., Noorishadkam, M., Mirjalili, S. R., Akbarian-Bafghi, M. J., & Neamatzadeh, H. (2021). Cumulative Evidence for Association between IL-10 Polymorphisms and Kawasaki Disease Susceptibility: A Systematic Review and Meta-Analysis. *Fetal and Pediatric Pathology*, *40*(2), 153–165. https://doi.org/10.1080/15513815.2019.1686789

Ferrero, S., Barra, F., & Leone Roberti Maggiore, U. (2018). Current and Emerging Therapeutics for the Management of Endometriosis. *Drugs*, *78*(10), 995–1012. https://doi.org/10.1007/s40265-018-0928-0

França, P. R. de C., Lontra, A. C. P., & Fernandes, P. D. (2022). Endometriosis: A Disease with Few Direct Treatment Options. *Molecules*, *27*(13), 4034. https://doi.org/10.3390/molecules27134034

Ghali, R. M., Al-Mutawa, M. A., Ebrahim, B. H., Jrah, H. H., Zaied, S., Bhiri, H., Hmila, F., Mahjoub, T., & Almawi, W. Y. (2020). Progesterone Receptor (PGR) Gene Variants Associated with Breast Cancer and Associated Features: A Case-Control Study. *Pathology & Oncology Research*, *26*(1), 141–147. https://doi.org/10.1007/s12253-017-0379-z

Gueuvoghlanian-Silva, B. Y., Bellelis, P., Barbeiro, D. F., Hernandes, C., & Podgaec, S. (2018). Treg and NK cells related cytokines are associated with deep rectosigmoid endometriosis and clinical symptoms related to the disease. *Journal of Reproductive Immunology*, *126*, 32–38. https://doi.org/10.1016/j.jri.2018.02.003

Guillot-Sestier, M.-V., Doty, K. R., Gate, D., Rodriguez, J., Leung, B. P., Rezai-Zadeh, K., & Town, T. (2015). Il10 Deficiency Rebalances Innate Immunity to Mitigate Alzheimer-Like Pathology. *Neuron*, *85*(3), 534–548. https://doi.org/10.1016/j.neuron.2014.12.068

Guo, S.-W. (2009). Epigenetics of endometriosis. *Molecular Human Reproduction*, *15*(10), 587–607. https://doi.org/10.1093/molehr/gap064

Guo, X., Jing, C., Li, L., Zhang, L., Shi, Y., Wang, J., Liu, J., & Li, C. (2011). Down-regulation of VEZT gene expression in human gastric cancer involves promoter methylation and miR-43c. *Biochemical and Biophysical Research Communications*, *404*(2), 622–627. https://doi.org/10.1016/j.bbrc.2010.12.026

Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 310–322.

Hansen, K. A., & Eyster, K. M. (2010). Genetics and Genomics of Endometriosis. *Clinical Obstetrics and Gynecology*, *53*(2), 403–412. https://doi.org/10.1097/GRF.0b013e3181db7ca1

Holdsworth-Carson, S. J., Fung, J. N., Luong, H. T. T., Sapkota, Y., Bowdler, L. M., Wallace, L., Teh, W. T., Powell, J. E., Girling, J. E., Healey, M., Montgomery, G. W., & Rogers, P. A. W. (2016). Endometrial vezatin and its association with endometriosis risk. *Human Reproduction*, *31*(5), 999–1013. https://doi.org/10.1093/humrep/dew047

Horne, A. W., & Missmer, S. A. (2022). Pathophysiology, diagnosis, and management of endometriosis. *BMJ*, *379*, e070750. https://doi.org/10.1136/bmj-2022-070750

Hyenne, V., Souilhol, C., Cohen-Tannoudji, M., Cereghini, S., Petit, C., Langa, F., Maro, B., & Simmler, M.-C. (2007). Conditional knock-out reveals that zygotic vezatin-null mouse embryos die at implantation. *Mechanisms of Development*, *124*(6), 449–462. https://doi.org/10.1016/j.mod.2007.03.004

Iyer, S. S., & Cheng, G. (2012). Role of Interleukin 10 Transcriptional Regulation in Inflammation and Autoimmune Disease. *Critical Reviews™ in Immunology*, *32*(1), 23–63. https://doi.org/10.1615/CritRevImmunol.v32.i1.30

Ja, S. (1927). Peritoneal endometriosis due to menstrual dissemination of endometrial tissue into the peritoneal cavity. *Am J Obstet Gynecol*, *14*, 422–469.

Jerman, L. F., & Hey-Cunningham, A. J. (2015). The Role of the Lymphatic System in Endometriosis: A Comprehensive Review of the Literature1. *Biology of Reproduction*, *92*(3), 64, 1–10. https://doi.org/10.1095/biolreprod.114.124313

Joseph, S., & Mahale, S. D. (2019). Endometriosis Knowledgebase: A gene-based resource on endometriosis. *Database: The Journal of Biological Databases and Curation*, *2019*, baz062. https://doi.org/10.1093/database/baz062

Juo, S.-H. H., Wu, R., Lin, C.-S., Wu, M.-T., Lee, J.-N., & Tsai, E.-M. (2009). A functional promoter polymorphism in interleukin-10 gene influences susceptibility to endometriosis. *Fertility and Sterility*, *92*(4), 1228–1233. https://doi.org/10.1016/j.fertnstert.2008.08.015

Kadivnik, M., Kralik, K., Muller-Vranješ, A., Vučemilović-Jurić, V., Šijanović, S., & Wagner, J. (2022). Progesterone receptor genetic variants in pregnant women and fetuses as possible predictors of spontaneous premature birth: A preliminary case–control study. *Journal of Obstetrics and Gynaecology Research*, *48*(5), 1099–1109. https://doi.org/10.1111/jog.15194

Khan, A., & Illahi, N. (2000). *PREVALENCE AND PRESENTATION OF ENDOMETRIOSIS IN PATIENTS ADMITTED IN NISHTAR HOSPITAL, MULTAN*. https://www.semanticscholar.org/paper/PREVALENCE-AND-PRESENTATION-OF-ENDOMETRIOSIS-IN-IN-Khan-Illahi/6f21f8fa2e424cafa5030208aef56a75571bcad4

Koppel, N., Friese, M. B., Cardasis, H. L., Neubert, T. A., & Burden, S. J. (2019). Vezatin is required for the maturation of the neuromuscular synapse. *Molecular Biology of the Cell*, *30*(20), 2571–2583. https://doi.org/10.1091/mbc.E19-06-0313

Kowalik, M. K., Rekawiecki, R., & Kotwica, J. (2013). The putative roles of nuclear and membrane-bound progesterone receptors in the female reproductive tract. *Reproductive Biology*, *13*(4), 279–289. https://doi.org/10.1016/j.repbio.2013.09.001

Králíčková, M., Laganà, A. S., Ghezzi, F., & Vetvicka, V. (2020). Endometriosis and risk of ovarian cancer: What do we know? *Archives of Gynecology and Obstetrics*, *301*(1), 1–10. https://doi.org/10.1007/s00404-019-05358-8

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, *4*(7), 1073–1081. https://doi.org/10.1038/nprot.2009.86

Lee, S.-Y., Koo, Y.-J., & Lee, D.-H. (2020). Classification of endometriosis. *Yeungnam University Journal of Medicine*, *38*(1), 10–18. https://doi.org/10.12701/yujm.2020.00444

Li, J., Hong, X., Mesiano, S., Muglia, L. J., Wang, X., Snyder, M., Stevenson, D. K., & Shaw, G. M. (2018). Natural Selection Has Differentiated the Progesterone Receptor among Human Populations. *American Journal of Human Genetics*, *103*(1), 45–57. https://doi.org/10.1016/j.ajhg.2018.05.009

Li, M., & Zhou, C. (2021). Progesterone receptor gene serves as a prognostic biomarker associated with immune infiltration in gastric cancer: A bioinformatics analysis. *Translational Cancer Research*, *10*(6), 2663–2677. https://doi.org/10.21037/tcr-21-218

Li, Y.-S., Chen, Y.-Z., Guo, X.-B., Liu, X., & Li, L.-P. (2015). VEZT as a novel independent prognostic factor in gastric cancer. *Cancer Biomarkers*, *15*(4), 375–380. https://doi.org/10.3233/CBM-150476

Li, Z., Wei, H., Li, S., Wu, P., & Mao, X. (2022). The Role of Progesterone Receptors in Breast Cancer. *Drug Design, Development and Therapy*, *16*, 305–314. https://doi.org/10.2147/DDDT.S336643

Liaqat, I., Arifa, N. A., Asif, S., & Lone, K. P. (2021). Association of FSHR gene polymorphisms with endometriosis in women visiting tertiary-care hospitals of Lahore, Pakistan. *JPMA. The Journal of the Pakistan Medical Association*, *71*(4), 1118–1122. https://doi.org/10.47391/JPMA.836

Liaqat, I., Jahan, N., Lone, K. P., Pakstis, A., & Taylor, H. S. (2013). Genetic Polymorphisms Associated with Endometriosis in Pakistani Women. *Journal of Endometriosis and Pelvic Pain Disorders*, *5*(4), 134–143. https://doi.org/10.5301/je.5000165

Lu, Z., & Gao, Y. (2021). Screening differentially expressed genes between endometriosis and ovarian cancer to find new biomarkers for endometriosis. *Annals of Medicine*, *53*(1), 1377–1389. https://doi.org/10.1080/07853890.2021.1966087

Luong, H. TT., Painter, J. N., Sapkota, Y., Nyholt, D. R., Rogers, P. A., & Montgomery, G. W. (2015). AB028. Identifying the functional role of VEZT gene for endometriosis risk. *Annals of Translational Medicine*, *3*(Suppl 2), AB028. https://doi.org/10.3978/j.issn.2305-5839.2015.AB028

Magalhães, C. A., Carvalho, M. das G., Sousa, L. P. de, Caramelli, P., & Gomes, K. B. (2017). Alzheimer's disease and cytokine IL-10 gene polymorphisms: Is there an association? *Arquivos de Neuro-Psiquiatria*, *75*, 649–656. https://doi.org/10.1590/0004-282X20170110

Malutan, A. M., Drugan, C., Walch, K., Drugan, T., Ciortea, R., & Mihu, D. (2017). The association between interleukin-10 (IL-10) −592C/A, −819T/C, −1082G/A promoter polymorphisms and endometriosis. *Archives of Gynecology and Obstetrics*, *295*(2), 503–510. https://doi.org/10.1007/s00404-016-4269-5

Măluţan, A. M., Drugan, T., Ciortea, R., Mocan-Hognogi, R. F., Bucuri, C., Rada, M. P., & Mihu, D. (2015). Serum anti-inflammatory cytokines for the evaluation of inflammatory status in endometriosis. *Journal of Research in Medical Sciences : The Official Journal of Isfahan University of Medical Sciences*, *20*(7), 668–674. https://doi.org/10.4103/1735-1995.166215

Mani, S., & Oyola, M. (2012). Progesterone Signaling Mechanisms in Brain and Behavior. *Frontiers in Endocrinology*, *3*. https://www.frontiersin.org/articles/10.3389/fendo.2012.00007

Mao, Z.-R., Zhang, S.-L., & Feng, B. (2017). Association of IL-10 (-819T/C, -592A/C and -1082A/G) and IL-6 -174G/C gene polymorphism and the risk of pneumonia-induced sepsis. *Biomarkers*, *22*(2), 106–112. https://doi.org/10.1080/1354750X.2016.1210677

Marquardt, R. M., Kim, T. H., Shin, J.-H., & Jeong, J.-W. (2019). Progesterone and Estrogen Signaling in the Endometrium: What Goes Wrong in Endometriosis? *International Journal of Molecular Sciences*, *20*(15), 3822. https://doi.org/10.3390/ijms20153822

Matalliotakis, M., Zervou, M. I., Matalliotaki, C., Rahmioglu, N., Koumantakis, G., Kalogiannidis, I., Prapas, I., Zondervan, K., Spandidos, D. A., Matalliotakis, I., & Goulielmos, G. N. (2017). The role of gene polymorphisms in endometriosis. *Molecular Medicine Reports*, *16*(5), 5881–5886. https://doi.org/10.3892/mmr.2017.7398

Meola, J., Rosa e Silva, J. C., Dentillo, D. B., da Silva, W. A., Veiga-Castelli, L. C., de Souza Bernardes, L. A., Ferriani, R. A., de Paz, C. C. P., Giuliatti, S., & Martelli, L. (2010). Differentially expressed genes in eutopic and ectopic endometrium of women with endometriosis. *Fertility and Sterility*, *93*(6), 1750–1773. https://doi.org/10.1016/j.fertnstert.2008.12.058

Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L.-P., Mushayamaha, T., & Thomas, P. D. (2021). PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Research*, *49*(D1), D394–D403. https://doi.org/10.1093/nar/gkaa1106

Miao, R., Guo, X., Zhi, Q., Shi, Y., Li, L., Mao, X., Zhang, L., & Li, C. (2013). VEZT, a Novel Putative Tumor Suppressor, Suppresses the Growth and Tumorigenicity of Gastric Cancer. *PLOS ONE*, *8*(9), e74409. https://doi.org/10.1371/journal.pone.0074409

Mohammed Rasheed, H. A., & Hamid, P. (2020). Inflammation to Infertility: Panoramic View on Endometriosis. *Cureus*, *12*(11), e11516. https://doi.org/10.7759/cureus.11516

Mote, P. A., Balleine, R. L., McGowan, E. M., & Clarke, C. L. (2000). Heterogeneity of progesterone receptors A and B expression in human endometrial glands and stroma. *Human Reproduction*, *15*(suppl 3), 48–56. https://doi.org/10.1093/humrep/15.suppl_3.48

Mousazadeh, S., Ghaheri, A., Shahhoseini, M., Aflatoonian, R., & Afsharian, P. (2019). Differential expression of progesterone receptor isoforms related to PGR +331g/a polymorphism in endometriosis: A case-control study. *International Journal of Reproductive Biomedicine*, *17*(3), 185–194. https://doi.org/10.18502/ijrm.v17i3.4517

Nyholt, D. R., Low, S.-K., Anderson, C. A., Painter, J. N., Uno, S., Morris, A. P., MacGregor, S., Gordon, S. D., Henders, A. K., Martin, N. G., Attia, J., Holliday, E. G., McEvoy, M., Scott, R. J., Kennedy, S. H., Treloar, S. A., Missmer, S. A., Adachi, S., Tanaka, K., … Montgomery, G. W. (2012). Genome-wide association meta-analysis identifies new endometriosis risk loci. *Nature Genetics*, *44*(12), Article 12. https://doi.org/10.1038/ng.2445

Oleksyk, T. K., Shrestha, S., Truelove, A. L., Goedert, J. J., Donfield, S. M., Phair, J., Mehta, S., O'Brien, S. J., & Smith, M. W. (2009). Extended IL10 haplotypes and their association with HIV progression to AIDS. *Genes & Immunity*, *10*(4), Article 4. https://doi.org/10.1038/gene.2009.9

Omar, S., Mohd Tap, F., Shameli, K., Rasit Ali, R., Che Jusoh, N., & Ahmad Khairudin, N. B. (2018). Sequence analysis and comparative modelling of nucleocapsid protein from pseudomonas stutzeri. *IOP Conference Series Materials Science and Engineering (Online)*, *458*(1), 7.

Osiński, M., Mostowska, A., Wirstlein, P., Wender-Ożegowska, E., Jagodziński, P. P., & Szczepańska, M. (2018). The assessment of GWAS — identified polymorphisms associated with infertility risk in Polish women with endometriosis. *Ginekologia Polska*, *89*(6), Article 6. https://doi.org/10.5603/GP.a2018.0052

Osiński, M., Wirstlein, P., Wender-Ożegowska, E., Mikołajczyk, M., Jagodziński, P. P., & Szczepańska, M. (2018). HSD3B2, HSD17B1, HSD17B2, ESR1, ESR2 and AR

expression in infertile women with endometriosis. *Ginekologia Polska*, *89*(3), Article 3. https://doi.org/10.5603/GP.a2018.0022

Pabalan, N., Salvador, A., Jarjanazi, H., Christofolini, D. M., Barbosa, C. P., & Bianco, B. (2014). Association of the progesterone receptor gene polymorphism (PROGINS) with endometriosis: A meta-analysis. *Archives of Gynecology and Obstetrics*, *290*(5), 1015–1022. https://doi.org/10.1007/s00404-014-3308-3

Padyukov, L., Hytönen, A.-M., Smolnikova, M., Hahn-Zoric, M., Nilsson, N., Hanson, L.-A., Tarkowski, A., & Klareskog, L. (2004). Polymorphism in promoter region of IL10 gene is associated with rheumatoid arthritis in women. *The Journal of Rheumatology*, *31*(3), 422–425.

Pagliardini, L., Gentilini, D., Sanchez, A. M., Candiani, M., Viganò, P., & Di Blasio, A. M. (2015). Replication and meta-analysis of previous genome-wide association studies confirm vezatin as the locus with the strongest evidence for association with endometriosis. *Human Reproduction*, *30*(4), 987–993. https://doi.org/10.1093/humrep/dev022

Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K. A., Lin, G. N., Nam, H.-J., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., & Radivojac, P. (2020). Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19669-x

Piazzon, M. C., Lutfalla, G., & Forlenza, M. (2016). IL10, A Tale of an Evolutionarily Conserved Cytokine across Vertebrates. *Critical Reviews in Immunology*, *36*(2), 99–129. https://doi.org/10.1615/CritRevImmunol.2016017480

Prescott, J., Farland, L. V., Tobias, D. K., Gaskins, A. J., Spiegelman, D., Chavarro, J. E., Rich-Edwards, J. W., Barbieri, R. L., & Missmer, S. A. (2016). A prospective cohort study of endometriosis and subsequent risk of infertility. *Human Reproduction*, *31*(7), 1475–1482. https://doi.org/10.1093/humrep/dew085

Pshenichnyuk, E. Y., Юрьевна, П. Е., Kuznetsova, M. V., Владимировна, К. М., Bourmenskaya, O. V., Владимировна, Б. О., Kochetkova, T. O., Олеговна, К. Т., Nepsha, O. S., Сергеевна, Н. О., Trofimov, D. Y., Юрьевич, Т. Д., Adamyan, L. V., & Владимировна, А. Л. (2017). AN ASSOCIATION BETWEEN THE FREQUENCIES OF SINGLE NUCLEOTIDE POLYMORPHISMS IN THE ZNF366 AND VEZT GENES AND THE RISK OF EXTERNAL GENITAL ENDOMETRIOSIS: DATA ON A RUSSIAN POPULATION. *Obstetrics and Gynecology*, *6*, Article 6. https://doi.org/10.18565/aig.2017.6.64-72

Qiu, X.-M., Lai, Z.-Z., Ha, S.-Y., Yang, H.-L., Liu, L.-B., Wang, Y., Shi, J.-W., Ruan, L.-Y., Ye, J.-F., Wu, J.-N., Fu, Q., Yi, X.-F., Chang, K.-K., & Li, M.-Q. (2020). IL-2 and IL-27 synergistically promote growth and invasion of endometriotic stromal cells by maintaining the balance of IFN-γ and IL-10 in endometriosis. *Reproduction*, *159*(3), 251–260. https://doi.org/10.1530/REP-19-0411

Rahmioglu, N., Nyholt, D. R., Morris, A. P., Missmer, S. A., Montgomery, G. W., & Zondervan, K. T. (2014). Genetic variants underlying risk of endometriosis: Insights from meta-analysis of eight genome-wide association and replication datasets. *Human Reproduction Update*, *20*(5), 702–716. https://doi.org/10.1093/humupd/dmu015

Ramezani, A., Banifazl, M., Mamishi, S., Sofian, M., Eslamifar, A., & Aghakhani, A. (2012). The Influence of Human Leukocyte Antigen and IL-10 Gene Polymorphisms on Hepatitis

B Virus Outcome. *Hepatitis Monthly*, *12*(5), 320–325. https://doi.org/10.5812/hepatmon.6094

Ramos, J. A., Silva, R., Hoffmann, L., Ramos, A. L. A., Cabello, P. H., Ürményi, T. P., Villella-Nogueira, C. A., Lewis-Ximenez, L., & Rondinelli, E. (2012). Association of IL-10, IL-4, and IL-28B gene polymorphisms with spontaneous clearance of hepatitis C virus in a population from Rio de Janeiro. *BMC Research Notes*, *5*(1), 508. https://doi.org/10.1186/1756-0500-5-508

Reis, F. M., Coutinho, L. M., Vannuccini, S., Batteux, F., Chapron, C., & Petraglia, F. (2020). Progesterone receptor ligands for the treatment of endometriosis: The mechanisms behind therapeutic success and failure. *Human Reproduction Update*, *26*(4), 565–585. https://doi.org/10.1093/humupd/dmaa009

Restaino, S., Mereu, L., Finelli, A., Spina, M. R., Marini, G., Catena, U., Turco, L. C., Moroni, R., Milani, M., Cela, V., Scambia, G., & Fanfani, F. (2020). Robotic surgery vs laparoscopic surgery in patients with diagnosis of endometriosis: A systematic review and meta-analysis. *Journal of Robotic Surgery*, *14*(5), 687–694. https://doi.org/10.1007/s11701-020-01061-y

Rocha-Junior, C. V., Da Broi, M. G., Miranda-Furtado, C. L., Navarro, P. A., Ferriani, R. A., & Meola, J. (2019). Progesterone Receptor B (PGR-B) Is Partially Methylated in Eutopic Endometrium From Infertile Women With Endometriosis. *Reproductive Sciences*, *26*(12), 1568–1574. https://doi.org/10.1177/1933719119828078

Rolla, E. (2019). Endometriosis: Advances and controversies in classification, pathogenesis, diagnosis, and treatment. *F1000Research*, *8*, F1000 Faculty Rev-529. https://doi.org/10.12688/f1000research.14817.1

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols*, *5*(4), Article 4. https://doi.org/10.1038/nprot.2010.5

Saha, R., Pettersson, H. J., Svedberg, P., Olovsson, M., Bergqvist, A., Marions, L., Tornvall, P., & Kuja-Halkola, R. (2015). Heritability of endometriosis. *Fertility and Sterility*, *104*(4), 947–952. https://doi.org/10.1016/j.fertnstert.2015.06.035

Sapkota, Y., Fassbender, A., Bowdler, L., Fung, J. N., Peterse, D., O, D., Montgomery, G. W., Nyholt, D. R., & D'Hooghe, T. M. (2015). Independent Replication and Meta-Analysis for Endometriosis Risk Loci. *Twin Research and Human Genetics*, *18*(5), 518–525. https://doi.org/10.1017/thg.2015.61

Sapkota, Y., Low, S.-K., Attia, J., Gordon, S. D., Henders, A. K., Holliday, E. G., MacGregor, S., Martin, N. G., McEvoy, M., Morris, A. P., Takahashi, A., Scott, R. J., Kubo, M., Zondervan, K. T., Montgomery, G. W., & Nyholt, D. R. (2015). Association between endometriosis and the interleukin 1A (IL1A) locus. *Human Reproduction (Oxford, England)*, *30*(1), 239–248. https://doi.org/10.1093/humrep/deu267

Sapkota, Y., Steinthorsdottir, V., Morris, A. P., Fassbender, A., Rahmioglu, N., De Vivo, I., Buring, J. E., Zhang, F., Edwards, T. L., Jones, S., O, D., Peterse, D., Rexrode, K. M., Ridker, P. M., Schork, A. J., MacGregor, S., Martin, N. G., Becker, C. M., Adachi, S., … Nyholt, D. R. (2017). Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nature Communications*, *8*, 15539. https://doi.org/10.1038/ncomms15539

Schülke, S. (2018). Induction of Interleukin-10 Producing Dendritic Cells As a Tool to Suppress Allergen-Specific T Helper 2 Responses. *Frontiers in Immunology*, *9*. https://www.frontiersin.org/articles/10.3389/fimmu.2018.00455

Shafrir, A. L., Farland, L. V., Shah, D. K., Harris, H. R., Kvaskoff, M., Zondervan, K., & Missmer, S. A. (2018). Risk for and consequences of endometriosis: A critical epidemiologic review. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *51*, 1–15. https://doi.org/10.1016/j.bpobgyn.2018.06.001

Shahriyari, E., Vahedi, L., Roshanipour, N., Jafarabadi, M. A., Khamaneh, A., & Laleh, M. G. (2019). Exploring the association of IL-10 polymorphisms in Behcet's disease: A systematic review and meta-analysis. *Journal of Inflammation*, *16*(1), 26. https://doi.org/10.1186/s12950-019-0230-2

Sharifinejad, N., Zaki-Dizaji, M., Sepahvandi, R., Fayyaz, F., dos Santos Vilela, M. M., ElGhazali, G., Abolhassani, H., Ochs, H. D., & Azizi, G. (2022). The clinical, molecular, and therapeutic features of patients with IL10/IL10R deficiency: A systematic review. *Clinical and Experimental Immunology*, *208*(3), 281–291. https://doi.org/10.1093/cei/uxac040

Sherbet, G. (2018). PGR (progesterone receptor). *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, *11*. https://doi.org/10.4267/2042/68995

Smolarz, B., Szyłło, K., & Romanowicz, H. (2021). Endometriosis: Epidemiology, Classification, Pathogenesis, Treatment and Genetics (Review of Literature). *International Journal of Molecular Sciences*, *22*(19), Article 19. https://doi.org/10.3390/ijms221910554

Soliman, A. M., Fuldeore, M., & Snabes, M. C. (2017). Factors Associated with Time to Endometriosis Diagnosis in the United States. *Journal of Women's Health*, *26*(7), 788–797. https://doi.org/10.1089/jwh.2016.6003

Sonacare. (2018). *Sonacare*. Sonacare. https://sonacare.com.au/svg-%2F-endometriosis-scan

Spinner, M. A., Pinter, K., Drerup, C. M., & Herman, T. G. (2020). A Conserved Role for Vezatin Proteins in Cargo-Specific Regulation of Retrograde Axonal Transport. *Genetics*, *216*(2), 431–445. https://doi.org/10.1534/genetics.120.303499

Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T.-B. G., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., & Clausen, H. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO Journal*, *32*(10), 1478–1488. https://doi.org/10.1038/emboj.2013.79

Suen, J.-L., Chang, Y., Chiu, P.-R., Hsieh, T.-H., Hsi, E., Chen, Y.-C., Chen, Y.-F., & Tsai, E.-M. (2014). Serum Level of IL-10 Is Increased in Patients with Endometriosis, and IL-10 Promotes the Growth of Lesions in a Murine Model. *The American Journal of Pathology*, *184*(2), 464–471. https://doi.org/10.1016/j.ajpath.2013.10.023

Suen, J.-L., Chang, Y., Shiu, Y.-S., Hsu, C.-Y., Sharma, P., Chiu, C.-C., Chen, Y.-J., Hour, T.-C., & Tsai, E.-M. (2019). IL-10 from plasmacytoid dendritic cells promotes angiogenesis in the early stage of endometriosis. *The Journal of Pathology*, *249*(4), 485–497. https://doi.org/10.1002/path.5339

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2022). The STRING database in 2023: Protein–protein association networks and functional

enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, *51*(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000

Tabibzadeh, S., Becker, J. L., & Parsons, A. K. (2003). Endometriosis is associated with alterations in the relative abundance of proteins and IL-10 in the peritoneal fluid. *Frontiers in Bioscience: A Journal and Virtual Library*, *8*, a70-78. https://doi.org/10.2741/1045

Tahira, T., Niaz, A., & Khalid, A. (2019). Prevalence of Endometriosis in Patients with Subfertility. *Annals of Punjab Medical College (APMC)*, *13*(3), Article 3. https://doi.org/10.29054/apmc/2019.753

Tanbo, T., & Fedorcsak, P. (2017). Endometriosis-associated infertility: Aspects of pathophysiological mechanisms and treatment options. *Acta Obstetricia et Gynecologica Scandinavica*, *96*(6), 659–667. https://doi.org/10.1111/aogs.13082

Tang, Y., Zhao, M., Lin, L., Gao, Y., Chen, G. Q., Chen, S., & Chen, Q. (2020). Is body mass index associated with the incidence of endometriosis and the severity of dysmenorrhoea: A case–control study in China? *BMJ Open*, *10*(9), e037095. https://doi.org/10.1136/bmjopen-2020-037095

Taylor, H. S., Kotlyar, A. M., & Flores, V. A. (2021). Endometriosis is a chronic systemic disease: Clinical challenges and novel innovations. *The Lancet*, *397*(10276), 839–852. https://doi.org/10.1016/S0140-6736(21)00389-5

Treloar, S. A., Zhao, Z. Z., Armitage, T., Duffy, D. L., Wicks, J., O'Connor, D. T., Martin, N. G., & Montgomery, G. W. (2005). Association between polymorphisms in the progesterone receptor gene and endometriosis. *Molecular Human Reproduction*, *11*(9), 641–647. https://doi.org/10.1093/molehr/gah221

Ulhaq, Z. S., & Garcia, C. P. (2020). Inflammation-related gene polymorphisms associated with Parkinson's disease: An updated meta-analysis. *Egyptian Journal of Medical Human Genetics*, *21*(1), 14. https://doi.org/10.1186/s43042-020-00056-6

UniProt. (n.d.-a). *IL10—Interleukin-10—Homo sapiens (Human) | UniProtKB | UniProt*. Retrieved July 15, 2023, from https://www.uniprot.org/uniprotkb/P22301/entry

UniProt. (n.d.-b). *VEZT - Vezatin—Homo sapiens (Human) | UniProtKB | UniProt*. Retrieved July 9, 2023, from https://www.uniprot.org/uniprotkb/Q9HBM0/entry#sequences

Vang, A., Salem, K., & Fowler, A. M. (2023). Progesterone Receptor Gene Polymorphisms and Breast Cancer Risk. *Endocrinology*, *164*(4), bqad020. https://doi.org/10.1210/endocr/bqad020

Vargas, E., Aghajanova, L., Gemzell-Danielsson, K., Altmäe, S., & Esteban, F. J. (2020). Cross-disorder analysis of endometriosis and its comorbid diseases reveals shared genes and molecular pathways and proposes putative biomarkers of endometriosis. *Reproductive BioMedicine Online*, *40*(2), 305–318. https://doi.org/10.1016/j.rbmo.2019.11.003

Vercellini, P., Barbara, G., Abbiati, A., Somigliana, E., Viganò, P., & Fedele, L. (2009). Repetitive surgery for recurrent symptomatic endometriosis: What to do? *European Journal of Obstetrics & Gynecology and Reproductive Biology*, *146*(1), 15–21. https://doi.org/10.1016/j.ejogrb.2009.05.007

Vercellini, P., Viganò, P., Somigliana, E., & Fedele, L. (2014). Endometriosis: Pathogenesis and treatment. *Nature Reviews Endocrinology*, *10*(5), 261–275. https://doi.org/10.1038/nrendo.2013.255

Wang, C., Tan, X., Tang, D., Gou, Y., Han, C., Ning, W., Lin, S., Zhang, W., Chen, M., Peng, D., & Xue, Y. (2022). GPS-Uber: A hybrid-learning framework for prediction of general and

E3-specific lysine ubiquitination sites. *Briefings in Bioinformatics*, *23*(2), bbab574. https://doi.org/10.1093/bib/bbab574

Wang, J., Wang, Z., & Tan, T. (2019). Association of CTLA-4, TNF alpha and IL 10 polymorphisms with susceptibility to hepatocellular carcinoma. *Scandinavian Journal of Immunology*, *90*(6), e12819. https://doi.org/10.1111/sji.12819

Wang, Y., Yuan, J., Yu, X., Liu, X., Tan, C., Chen, Y., & Xu, T. (2021). Vezatin regulates seizures by controlling AMPAR-mediated synaptic activity. *Cell Death & Disease*, *12*(10), Article 10. https://doi.org/10.1038/s41419-021-04233-2

Wójcik, M., Szczepaniak, R., & Placek, K. (2022). Physiotherapy Management in Endometriosis. *International Journal of Environmental Research and Public Health*, *19*(23), 16148. https://doi.org/10.3390/ijerph192316148

Xie, D., Shang, L., Peng, L., & Li, L. (2017). Up-regulation of VEZT by small activating RNA inhibits the proliferation, invasion and migration of gastric cancer cells. *Biochemical and Biophysical Research Communications*, *482*(4), 542–548. https://doi.org/10.1016/j.bbrc.2016.11.071

Xie, J., Wang, S., He, B., Pan, Y., Li, Y., Zeng, Q., Jiang, H., & Chen, J. (2009). Association of estrogen receptor alpha and interleukin-10 gene polymorphisms with endometriosis in a Chinese population. *Fertility and Sterility*, *92*(1), 54–60. https://doi.org/10.1016/j.fertnstert.2008.04.069

Xu, Z., & Taylor, J. A. (2009). SNPinfo: Integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Research*, *37*(Web Server issue), W600-605. https://doi.org/10.1093/nar/gkp290

Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, *43*(W1), W174–W181. https://doi.org/10.1093/nar/gkv342

Yariv, B., Yariv, E., Kessel, A., Masrati, G., Chorin, A. B., Martz, E., Mayrose, I., Pupko, T., & Ben-Tal, N. (2023). Using evolutionary data to make sense of macromolecules with a "face-lifted" ConSurf. *Protein Science*, *32*(3), e4582. https://doi.org/10.1002/pro.4582

Yessenbayeva, A., Apsalikov, B., Massabayeva, M., Kazymov, M., Shakhanova, A., Mussazhanova, Z., Kadyrova, I., Aukenov, N., & Shaimardanov, N. (2023). Biomarkers of immunothrombosis and polymorphisms of IL2, IL6, and IL10 genes as predictors of the severity of COVID-19 in a Kazakh population. *PLOS ONE*, *18*(6), e0288139. https://doi.org/10.1371/journal.pone.0288139

Yilmaz, B. D., & Bulun, S. E. (2019). Endometriosis and nuclear receptors. *Human Reproduction Update*, *25*(4), 473–485. https://doi.org/10.1093/humupd/dmz005

Yu, Y.-X., Xiu, Y.-L., Chen, X., & Li, Y.-L. (2017). Transforming Growth Factor-beta 1 Involved in the Pathogenesis of Endometriosis through Regulating Expression of Vascular Endothelial Growth Factor under Hypoxia. *Chinese Medical Journal*, *130*(08), 950–956. https://doi.org/10.4103/0366-6999.204112

Zhang, Y., & Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, *33*(7), 2302–2309. https://doi.org/10.1093/nar/gki524

Zhao, Q., Xie, Y., Zheng, Y., Jiang, S., Liu, W., Mu, W., Liu, Z., Zhao, Y., Xue, Y., & Ren, J. (2014). GPS-SUMO: A tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Research*, *42*(W1), W325–W330. https://doi.org/10.1093/nar/gku383

Zhong, S., Liang, Y., Wu, Z., & Wei, L. (2023). Association between polymorphisms of cytokine genes and endometriosis: A comprehensive systematic review and meta-analysis. *Journal of Reproductive Immunology*, *158*, 103969. https://doi.org/10.1016/j.jri.2023.103969

Zhou, C., Zou, X., Wen, X., & Guo, Z. (2022). Association of the PROGINS PgR polymorphism with susceptibility to female reproductive cancer: A meta-analysis of 30 studies. *PLoS ONE*, *17*(7), e0271265. https://doi.org/10.1371/journal.pone.0271265

Zhou, W.-J., Yang, H.-L., Shao, J., Mei, J., Chang, K.-K., Zhu, R., & Li, M.-Q. (2019). Anti-inflammatory cytokines in endometriosis. *Cellular and Molecular Life Sciences*, *76*(11), 2111–2132. https://doi.org/10.1007/s00018-019-03056-x

Zondervan, K. T., Becker, C. M., & Missmer, S. A. (2020). Endometriosis. *New England Journal of Medicine*, *382*(13), 1244–1256. https://doi.org/10.1056/NEJMra1810764

Zondervan, K. T., Rahmioglu, N., Morris, A. P., Nyholt, D. R., Montgomery, G. W., Becker, C. M., & Missmer, S. A. (2016). Beyond endometriosis GWAS: From Genomics to Phenomics to the Patient. *Seminars in Reproductive Medicine*, *34*(4), 242–254. https://doi.org/10.1055/s-0036-1585408

turnitin

Digital Receipt

Prof. Dr. Peter John
Deptt of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB), NUST Islamabad

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Ameema Khan
Assignment title:	Quick Submit
Submission title:	Molecular Screening of PGR, IL10 and VEZT Gene Polymorphi...
File name:	Thesis.pdf
File size:	1.78M
Page count:	84
Word count:	18,099
Character count:	94,121
Submission date:	15-Aug-2023 06:03AM (UTC-0700)
Submission ID:	2146190207

Molecular Screening of *PGR*, *IL10* and *VEZT* Gene Polymorphisms in Patients with Endometriosis

By

Ameema Khan

0000036480l4

MS Healthcare Biotechnology

Supervisor

Prof. Dr. Peter John

Department of Healthcare Biotechnology
Atta-ur-Rahman School of Applied Biosciences (ASAB)
National University of Sciences and Technology (NUST)
H-12, Islamabad, Pakistan
2023

Molecular Screening of PGR, IL10 and VEZT Gene Polymorphisms in Patients with Endometriosis

ORIGINALITY REPORT

13%	8%	9%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Higher Education Commission Pakistan Student Paper	3%
2	"Endometriosis", Springer Science and Business Media LLC, 2014 Publication	1%
3	www.mdpi.com Internet Source	1%
4	link.springer.com Internet Source	<1%
5	encyclopedia.pub Internet Source	<1%
6	Philips, Anju K, Auli Sirén, Kristiina Avela, Mirja Somer, Maarit Peippo, Minna Ahvenainen, Fatma Doagu, Maria Arvio, Helena Kääriäinen, Hilde Van Esch, Guy Froyen, Stefan A Haas, Hao Hu, Vera M Kalscheuer, and Irma Järvelä. "X-exome sequencing in Finnish families with Intellectual Disability - four novel mutations	<1%