Special Edition on "Frontiers of Biosciences in Sustainable Development (FBSD)"
CIBTech Journal of Zoology ISSN: 2319–3883; http://www.cibtech.org.htm
2023 Vol.12/S1, pp.63-80/Akthar and Powell
***Research Article** (Open Access)*

INVESTIGATING SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES RELATED TO ENDOMETRIAL CANCER AND TYPE 2 DIABETES

***N. Syed Akthar[1] and L. Powell[2]**
[1]Department of Medical, Veterinary and Life Sciences
University of Birmingham, England
[2]Medical Genetics and Genomics, Academic Unit of Medical Genetics and Pathology,
Laboratory Medicine, Queen Elizabeth University Hospital,
The University of Glasgow,
Glasgow
**Author for correspondence: syednafeesa0101@gmail.com*

ABSTRACT
Single Nucleotide Polymorphisms (SNPs) are genetic variants in the DNA and can function as important biological markers by locating disease-associated genes. Endometrial cancer (EC) is the most prevalent gynaecological malignancy in the world causing approximately one lakh deaths per year. The incidence rate of EC has been increasing yearly by 2.5% and doubled in the last three decades. Type 2 diabetes (T2D) is a prevalent chronic disease that results from genetic and environmental interactions and is a major health problem, with populations affected worldwide. T2D has similar risk factors to EC, such as obesity and a sedentary lifestyle. T2D and obesity are major risk factors for EC. While associations between EC and T2D have been made in several studies, the associations remain uncertain. SNPs in P*olymerase Epsilon* (*POLE*) and *Catenin Beta-1* (*CTNNB1*) genes associated with EC and T2D were identified using online databases and investigated using PCR, RFLP-PCR, TaqMan (qPCR) assay and DNA sequencing. The genotypes of the SNPs from the TaqMan assay were heterozygous for *POLE* rs139075637 with AT alleles identified, which are contradictory to the previously published data and homozygous for *CTNNB1* rs121913228 with TT alleles identified. In the *POLE* and *CTNNB1* variants, no DNA changes were observed in the sequencing results. However, additional DNA changes were observed in rs1057519945 and rs121913399 in normal Uterus and adenocarcinoma endometroid samples. The study results indicate the presence of DNA changes in the genes and there might be similarities between EC and T2D. The results from this pilot study are important to allow for more in-depth research on a large scale to investigate further the potential links between these SNPs EC and T2D.

***Keywords:** Catenin Beta-1, Polymerase epsilon, Single nucleotide polymorphism, The cancer genome atlas, Endometrial cancer and Type II diabetes*

INTRODUCTION
1.1) Cancer
Cancer is one of the most dreaded diseases across the globe that is responsible for approximately one in six deaths according to the World Health Organisation (WHO) (Cancer, WHO, 2022). Cancer is not a single disease, thus a group of diseases that has the potential to originate in any tissue or organ due to the uncontrollable growth of abnormal cells. The major cause of death due to cancer is metastasis, where the malignant cells start spreading to other organs disrupting their natural function. And the common type of cancer prevalent in women is endometrial cancer in which malignant cells form in the endometrial tissues of the uterus. Endometrial cancer (EC) is the 4th and 6th most common cancer in the UK and worldwide in women (Uterine Cancer Statistics | Cancer Research UK, 2022). The cases of endometrial cancer and the mortality rates have been increasing over the years.

1.2) Endometrial cancer (EC)

The endometrium is the lining of the uterus's inner epithelial layer, composed of glandular and luminal epithelia, vascular smooth muscle cells and stromal fibroblasts. The development of cancerous cells in this lining leads to endometrial carcinoma, carcinosarcoma or adenocarcinoma. These cancers can be cured or treated if detected in the early stages by removing the uterus. However, things become complicated when cancer starts to spread to other parts of the body through metastasis. The EC is graded as lower grade (grade 1 and 2) and higher-grade cancers (grade 3). Grade 1 cancer has tumours with 95% or more of the cancer tissue forming glands while grade 2 has 50% to 94%. Grades 1 and 2 are type 1 EC as they are not very aggressive and don't spread quickly to other tissues. Grade 3 has fewer cancer tissue forming glands and is a type 2 EC, as grade 3 is very aggressive, grows and spreads to other tissues quickly. More information about the types and grades can be looked at in Table 1. The symptoms of EC can include post-menopausal bleeding, vaginal bleeding and pelvic pain. The various causes of EC can be infertility, post and late menopause, taking oestrogen-related therapies, and other health ailments like diabetes and an unhealthy lifestyle. The factors that can affect the risk of developing EC include age, BMI levels, irregular hormone levels and family history (lynch syndrome).

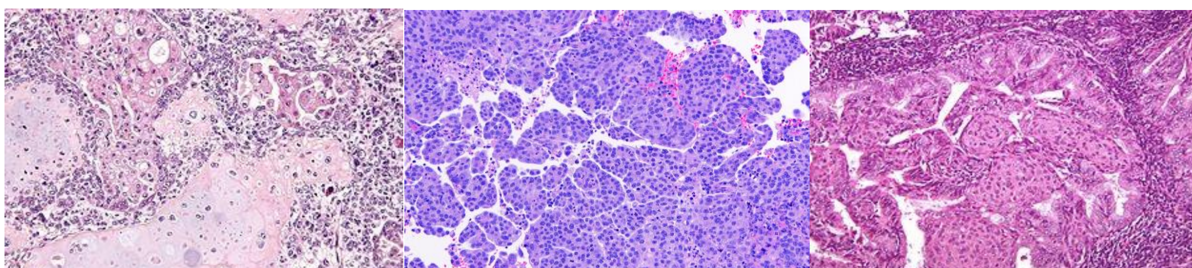
Types of Endometrial Cancer	**Type 1 Endometrial Cancer**		**Type 2 Endometrial Cancer**
Grades of Endometrial Cancer	Grade 1	Grade 2	Grade 3
Tissue-forming glands in tumours	95%	50% to 94%	>50%
Cell appearance	Normal	Abnormal	Very abnormal
Cell growth	Grow slowly	Moderately	Very quickly
Cell type	Non-aggressive	Non-aggressive	Very Aggressive
Spread type	Doesn't spread quickly	Spreads slowly	Spreads quickly
Grade type	low grade	moderate grade	high grade
Differentiation	well differentiated	moderately differentiated	poorly differentiated

Table 1: Types and grades of EC. The table displays the types and grades of Endometrial cancer and the characteristics of cells in grades 1, 2 and 3.

The three types of endometrium cancer samples whose genomic DNA (gDNA) are used for research in this project are discussed in the following paragraphs.

Endometrial carcinosarcoma (ECS) is a rare and aggressive type of endometrial cancer, comprising less than 5% of uterine malignancies and containing features of carcinoma and sarcoma, arising from a single malignant

epithelial clone. While diagnosing, ECS is generally treated as grade 2 endometrial cancer. The ECS forms at the junction of the endometrial epithelial tissues and the connective tissues of the uterus and hence the name carcinosarcoma. Endometrial serous adenocarcinoma (ESC) is a rare type 2 cancer developing from the lining of the womb accounting for 10% of endometrial cancers. ESCs mostly arise in postmenopausal women and the tumours spread to lymph nodes and other parts of the body. Endometrial adenocarcinoma endometrioid is the most prevalent type of malignancy in women that occurs mainly during postmenopausal age, comprising 75% of uterine malignancies. Adenocarcinoma forms in the glandular tissue, the tissues that make up the uterus's endometrium lining. The glandular tissues are made up of glandular epithelial cells. The three types of EC investigated are endometrial carcinosarcoma, endometrial serous adenocarcinoma and endometrial adenocarcinoma endometroid.



A. Carcinosarcoma B. Serous adenocarcinoma C. Adenocarcinoma

Figure 1(1A, 1B and 1C): The images of pathological slides of endometrial carcinosarcoma, endometrial serous adenocarcinoma and endometrial adenocarcinoma endometroid respectively.

1.3) Type 2 diabetes (T2D)

Type 2 diabetes (T2D) is a serious condition in which the blood glucose levels become too high in an individual and are caused due to insensitivity of the body to insulin. As a result of which, the glucose cannot be carried to the individual cells causing them to starve and, in extreme cases, cause cell deaths. The risk factors for T2D are obesity or overweight, and a family history of diabetes. BMI (body mass index) is the ratio or measure of body weight compared to an individual's height. This measurement can help to determine the obesity of an individual and the measures that can be taken to fix the condition. Increased levels of BMI have the chance of developing T2D. Insulin resistance (IR) is a condition in which the cells do not respond to the hormone insulin in the body. IR is developed with T2D. When the pancreas does not produce Insulin hormone in the required quantities, or the body doesn't take up enough insulin, diabetes is caused due to hormonal imbalance in the body. And the menopausal hormonal changes influence the blood sugar level in women.

1.4) Endometrial cancer and Type 2 diabetes

Hormonal imbalances can be caused by obesity due to which the hormones such as leptin and insulin influence our appetite, metabolism and body fat distribution. Obesity can be a cause to trigger T2D and potentially the onset of endometrial cancer. Obesity and low physical activity are the same risk factors shared by EC and T2D. Insulin resistance and abnormal functioning of hormones can be possible associations between T2D, and EC. IR associated with T2D is one of the major risk factors of EC although the association remains unclear, several studies have linked them to play a role in promoting diabetes-related cancers.

1.5) SNPs (Single Nucleotide polymorphisms)

Single nucleotide polymorphisms (SNPs) are genetic variants in DNA and approximately 4.5 million SNPs are present in an individual's genome. SNPs are represented by the difference in nucleotides in an individual's DNA and occur throughout the DNA. Most SNPs appear to have no effect on the development and health of an individual. However, some SNPs can have a major impact on one's health or development. SNPs can be linked to drug response, increased risk of developing diseases and susceptibility to various factors on an environmental basis, such as toxins. SNPs can also be referred to as point variations, some of which might

affect the function of a particular gene. However, critical changes might lead to a different gene function, which can disrupt the whole cell's function and eventually the tissues' function and even lead to cancer.

1.6) Selection of SNPs

For this project SNPs within *polymerase epsilon* (*POLE*) and *Catenin Beta 1* (*CTNNB1*) associated with endometrial cancer were identified using TCGA (The cancer genome atlas) (Zhou et al., 2021). TCGA is a vast resource tool of bioinformatics that has the largest genomic data collection for cancer (Tomczak et al., 2015). The TCGA database contains information about more than 30 types of cancer and contains molecular data from sequencing analysis and clinical information about participants and sample**s.**

1.7) POLE (polymerase epsilon)

Polymerase epsilon (*POLE*) is a replicative DNA and a gene coding protein. This gene helps encode a specific part, the catalytic subunit of DNA polymerase epsilon. DNA polymerase epsilon is one of the four types of DNA polymerases in eukaryotic cells, that is responsible for possible replication of DNA (chromosomal) and DNA repair. Variations in this gene have been linked to immunodeficiency, colorectal cancer and endometrial cancer.

1.8) CTNNB1(Catenin Beta-1)

Catenin Beta 1 (CTNNB1) is a gene that encodes for a specific part of a complex protein that composes Adherens junctions, which is necessary for creating and maintaining epithelial layers of cells. Adherens junctions are essential as they control adhesion between the cells and regulate the overall growth of cells (Travaglino et al., 2022). The *CTNNB1* gene-encoded protein anchors the actin cytoskeleton to the cell membrane and might be responsible for the inhibition of various cell signalling pathways by binding to growth factor receptors and by activating signalling mediators (Ledinek et al., 2022). Variations of this gene have been shown to develop ovarian, and endometrial cancer.

1.9) Aims

To perform a pilot study to investigate possible SNPs related to EC using TCGA and to determine any possible link with similar risk profile diseases such as T2D. The primary aim of this project is to identify SNPs within *POLE* and *CTNNB*1 related to EC using TCGA. To perform the SNP genotyping of EC and T2D samples to observe and identify any link between EC and T2D. To explore if there is any potential impact on the future clinical diagnosis. The project will be carried out using bioinformatics resources TCGA, PubMed, Ensembl, and Primer3, to identify SNPs and design primers. Investigate these SNPs using PCR (Polymerase chain reaction) based techniques, RFLP (Restriction fragment length polymorphism)- PCR, ARMS (Amplification-Refractory Mutation System)- PCR, TaqMan (genotyping qPCR) assays and DNA sanger sequencing.

1.10) Rational

The mortality rate of EC is increasing every year and has a dire need to be addressed. The simple SNP markers can be used in diagnosing patients quickly and for providing better treatment at an early stage. SNPs are said to have a key role in enabling the personalised medicine concept and if the SNPs are identified, possibly drug or gene therapy can be designed to target them.

1.11) Hypothesis

The hypothesis of this study is the identification of SNPs associated with EC-related genes using TCGA and investigation of the findings for risk indicators and understanding of the potential impact on clinical diagnosis. If the SNPs are identified indicating the risks, then a drug or therapy can be designed to target them through large-scale studies for patient treatment.

1.12) Outcomes

The intended outcomes of this project are the identification of SNPs related to EC using TCGA and investigating these SNPs in relevant EC samples using PCR-based technologies-RFLP PCR, ARMS PCR, TaqMan (qPCR) and Sequencing. And investigating these SNPs in relevant samples from T2D, a disease with a similar disease profile as EC. Critically considering the experimental evidence together with peer-reviewed publications to evaluate the links between these SNPs, EC and T2D**.**

MATERIALS AND METHODS

The following methods include all the online software, database tools and websites used to select the SNPs and design the primers.

2.1) The cancer genome atlas (TCGA)

The suitable SNPs for *DNA Polymerase Epsilon (POLE)* and *Catenin Beta-1 (CTNNB1)* variants were selected using the TCGA database for this project (Radenbaugh *et al.,* 2014). In the National Cancer Institute Genomic Data Commons (GDC) data portal, the primary site - uterus, related to endometrial cancer was selected, which contains 2664 cases (portal.gdc.cancer.gov, 2022). In the uterus data portal, parameters such as disease types (adenomas, adenocarcinomas, mucinous and serous neoplasms), genes *(POLE* and *CTNNB1*) and type (single-base substitution) were selected to filter out the required data.

2.2) National Center for Biotechnology Information (NCBI)

NCBI, a part of the National Institutes of Health (NIH) provides biomedical and genomic information. The SNP database with dbSNP information contains data on variations (Chang et al., 2017). The selected SNP was investigated for their position, alleles, frequency, HGVS data, genomic view data, gene consequences and clinical significance, including the publications on the variant. The clinVar (the database that collects human variant information in human health) results show in detail information about the location of variation, conditions, and clinical significance including interpretations and evidence regarding the variant.

2.3) Ensembl

In the Ensembl genome database, the selected SNPs were investigated in the human species for the flanking sequences. The result shows the details of variant SNP, genomic context, gene and regulations data and flanking sequences, population genetics, phenotype data and phylogenetic context. The flanking sequence of the SNP was obtained, the display of the sequence was configured to achieve the desired result of the flanking sequence that was used for the Primer3 tool.

2.4) Primer3

The flanking sequence with the missense variant was loaded on the primer3 tool website to obtain primers (Park et al., 2020). The primer3 web output produces forward and reverse primers with the product size, length of the primers, melting temperatures, GC content, secondary structure, and self and pair complementarity data. The reverse primer obtained should be complemented using the **reverse complement tool**. Primers with secondary structure percentage should be avoided as they bind to themselves instead of DNA resulting in no DNA fragments.

2.5) NEB cutter

A unique and non-repetitive recognition site was identified in the selected amplicon size to perform restriction digest. The missense variant in the recognition sequence was changed with the ancestral allele of that variant from the population genetics section in Ensembl. The enzyme was selected using the NEB cutter tool. For better restriction digest, the recognition sequence should be shorter in length for frequent digestions and six base cutters should be selected for better cleavage. Additional information about the enzyme was viewed on the **REBASE** neb website.

2.6) SNP Check gene tool

The primer pairs were investigated for other SNPs present apart from the SNPs investigated. This was achieved by adding the primer pair with the location of the SNP in the SNP Check gene tool and the input results delivered the details of the SNPs present and other additional SNPs present in the primer pair. The SNPs should be investigated to avoid additional SNPs, self-complementarity, and primer dimer. These SNPs may cause instability during PCR and yield inaccurate results in sequencing.

2.7) Primer1

The Primer1 tool was used to design Amplification-Refractory Mutation System (ARMS) primers. The flanking sequence was obtained, and the missense variant was replaced with the ancestral allele and loaded in the primer1 tool. The position of the variant and the two major alleles were added to the primer1. The GC content to 70% and melting temperatures were changed and others should be left at default settings. The results yielded a long list of output sets (see figure 2). One set of primers was selected for the experiment.

rs762000608 G>A

Forward inner primer (G allele):	Melting temperature
475 GTCATGAAGCAGAGAGGATGGCGAACG 501	73
Reverse inner primer (A allele):	
527 AGGTGATGCACTTCTACCGCTGGCGTT 501	73
Forward outer primer (5' - 3'):	
342 CTGGCCATGTCTCTGGTTCTGGGGAGTA 369	73
Reverse outer primer (5' - 3'):	
687 AAACCTGTGCCCATTTCAGTTGGAGCGT 660	73

Product size for G allele: 214

Product size for A allele: 186

Product size of two outer primers: 346

Figure 2: Primer 1 tool output. The output result image of the primer 1 tool for the rs762000608 SNP. The output result displays forward inner, reverse inner, forward outer and reverse outer primers with melting temperatures and product sizes for the alleles and primers.

The following include the methods followed in the laboratory for the project.

2.8) Samples

The 4 types of EC gDNA samples (See table 2 for samples) were selected from AMBSIO. The manufacturer described the samples were obtained from females aged 51 to 66 belonging to diverse races**.**

The T2D gDNA samples (See table 2 for samples) were selected from Warren type 2 diabetes DNA panel 2 manufactured by the European Collection of Cell Cultures (ECACC). The manufacturer described the samples were obtained from unrelated females aged 61 to 75 belonging to the European Caucasian population.

2.9) Optimisation of primers

The primers were reconstituted, and TE buffer was added to prepare working stock primers. To set up the PCR, samples were prepared using, 2µl DNA and 48µl MasterMix (2µl forward primer, 2µl reverse primer, 19µl water and 25µl MegaMix blue (Microzone)) for each reaction. For negative control, 2µl water and 48µl MasterMix were added. The tubes were given a brief spin before loading them into the PCR machine (Initial denaturation at 93°C for 5mins, denaturing cycle at 93°C for 45 secs for 30 cycles, annealing for 45 secs (temperature dependant on primer annealing temperature), and extension at 72°C for 1 min, final extension at 72°C for 5 minutes with 4°C infinite cool hold). A gradient PCR was used with temperatures ranging from 55°C - 70°C to establish the optimal conditions to produce the desired PCR fragment.

2.10) Amplification-Refractory Mutation System (ARMS) Primers

The primers were reconstituted and working stock primers were prepared from initial primer stock. To perform PCR, MasterMix 1 and MasterMix 2 were prepared for two sets of the same samples (See table 3). To one set of samples, MasterMix 1 should be added and to the other set of the same samples, MasterMix 2 should be added to amplify the allele. The primers were optimised till the fragments were observed on the gel. The same

set of samples were loaded side by side to analyse the fragments on the gel. Alternatively, MasterMix was made with four primers together with water and MegaMix blue (Microzone) and electrophoresed using 1.5% gel to analyse the fragments.

Sample type	Simplified name	Sample number	Age	Sex	Panel
Normal uterus	NU	CD563141	66	F	N/A
Carcinosarcoma	ECS	CD565240	64	F	N/A
Adenocarcinoma endometroid	AC	CD563144	51	F	N/A
Adenocarcinoma serous	ESC	CD564116	61	F	N/A
T2D	T2D1	WR2520 A6	62	F	Warren T2D DNA panel 2
T2D	T2D2	WR2094 B1	68	F	Warren T2D DNA panel 2
T2D	T2D3	WR1563 B3	66	F	Warren T2D DNA panel 2
T2D	T2D4	WR2075 C4	63	F	Warren T2D DNA panel 2
T2D	T2D5	WR1640 C6	65	F	Warren T2D DNA panel 2
T2D	T2D6	WR1739 C7	75	F	Warren T2D DNA panel 2
T2D	T2D7	WR1836 C8	70	F	Warren T2D DNA panel 2
T2D	T2D8	WR1970 C11	66	F	Warren T2D DNA panel 2
T2D	T2D9	WR2137 D2	62	F	Warren T2D DNA panel 2
T2D	T2D10	WR1825 D10	61	F	Warren T2D DNA panel 2
T2D	T2D11	WR2062 D12	60	F	Warren T2D DNA panel 2
T2D	T2D12	WR1018 E3	70	F	Warren T2D DNA panel 2
T2D	T2D13	WR1687 E11	64	F	Warren T2D DNA panel 2
T2D	T2D14	WR1358 F5	74	F	Warren T2D DNA panel 2
T2D	T2D15	WR1558 G7	73	F	Warren T2D DNA panel 2
T2D	T2D16	WR1880 H6	66	F	Warren T2D DNA panel 2
T2D	T2D17	WR1857 H8	69	F	Warren T2D DNA panel 2

Table 2: EC and T2D samples. The table displays the list of EC and T2D samples with simplified names, sample numbers, age, sex and panel details. The samples belonged to females aged 51 to 75. (T2D= Type 2 Diabetes; F= Female)

MasterMix 1	MasterMix 2
Forward inner primer	Forward outer primer
Reverse outer primer	Reverse inner primer
MegaMix blue	MegaMix blue
Sterile water	Sterile water

Table 3: ARMS PCR MasterMix. The table displays the constituents of MasterMix 1 and MasterMix 2 for ARMS PCR. MasterMix 1 was made of forward inner primer, reverse outer primer, water and megamix blue. MasterMix 2 was made of forward outer primer, reverse inner primer, sterile water and megamix blue

2.11) Restriction Fragment Length Polymorphism Assay (RFLP)

PCR-restriction fragment length polymorphism assay was performed for the samples that have successful fragments seen in the gel. To perform restriction digest, restriction digest Mastermix was prepared using PCR product, restriction enzyme (depending upon the unit concentration of the enzyme by NEB), 10 X NE. Buffer and sterile water. The tubes were briefly centrifuged and incubated at the given temperature and time (Manufacturer guidelines). Electrophorese was performed using a 3% gel to separate the digested fragments.

2.12) Sequencing

Sequencing PCR was performed, with MasterMix comprising of 12.5µl of MegaMix-gold (Microzone), 2µl each of forward primer and reverse primer and 6.5µl of sterile water. 23µl of MasterMix and 2µl of DNA were used for the PCR reaction. Electrophorese was performed using a 3 % gel with tracking dye (Thermo Scientific 6X TriTrack DNA loading dye) to analyse the fragments. The process of Exosap and sequencing was carried out. The ExoSAP-IT reagent (Thermo Scientific) was added to clean the PCR product and the samples were set for the exoSAP programme (37°C for 4 mins, 80°C for 1 min and held at 4°C) (Bell, n.d.). To the primers, M13 tags were added as the tags initiate the synthesis of DNA on single-stranded templates and the tubes were sent away for sequencing.

2.13) Agarose Gel 1.5 and 3%

The gel was prepared using 1.5gm or 3gm agarose powder with 100ml TAE buffer, and 3µl SYBR safe stain (Invitrogen by Thermo Fisher Scientific). The PCR run samples and enzyme digested samples were loaded in the wells along with a ladder (Thermo scientific GeneRuler 100bp DNA ladder) and run at 100V, 50V or 35V for 30 minutes to visualise the fragments and digested fragments. For PCR primers and ARMS primers, 1.5% agarose gel was used to resolve DNA fragments and for sequencing PCR and restriction digest samples 3% agarose gel was used to separate the fragments (Gaudet et al., n.d.).

2.14) TaqMan assay

The TaqMan assay (SNP genotyping assay) can be used to achieve accurate results for a larger number of samples in less time (Shen et al., n.d.). This assay was performed using the StepOnePlus instrument (Applied Biosystems by ThermoFisher Scientific). The TaqMan master mix (MM) was prepared according to the manufactory's specifications and added to wells in the 96-well plate. For each sample, 2µl was added of patient DNA (one well per sample), sterile dH2O as the negative control in triplicate and G11(AT alleles) as the positive control in one well. The plate was sealed and centrifuged for 3 minutes at 1500 rpm to eliminate bubbles. The program was used with qPCR run default settings.

2.15) Analysis

Once the results arrive, the sequenced data should be analysed using the **mutation surveyor** system. The TaqMan assay results should be analysed upon the completion of the run.

RESULTS AND DISCUSSION

3.1) Optimisation of primers through temperature gradient PCR

The primer pairs were optimised through temperature gradient PCR. The primer annealing temperatures were 58°C for rs1057519943, rs1057519945, rs121913413 and 59°C for rs121913399 and 60°C for rs762000608 and the expected PCR product sizes can be seen in Table 4. PCR products were confirmed to be of the expected size by electrophoresis. It was not possible to optimise the primer set rs1057519943 SNP using temperatures from 54°C to 64°C, no fragments were visible, however, primer dimers were formed (See image) due to the presence of secondary structure in the reverse primer. Therefore, the primers were redesigned.

The annealing temperatures for the ARMS primers were 60°C for rs762000608 and the expected PCR product sizes can be seen in table 5. PCR products were confirmed to be of the expected size by electrophoresis. It was not possible to optimise the ARMS primer set for rs121913412 SNP using temperatures from 54°C to 64°C, no fragments were visible, however, primer dimers were formed.

SNP and Gene variant	Simplified Primer name	Forward sequence Reverse sequence	Tm	Expected size	Expected digest size	Allele testing
rs105751943 *POLE* Set 1	*POLE* 943	CTCGAAAGGGGAGGGTACAG TGATCAGCTGCTCTCTTAGGT	--	494bp	273;218	T
rs105751943 *POLE* Set 2	*POLE* 943	CCCACACACTCTGCCTAGA CTGCTCTCTTAGGTATGTCAGTT	58°C	588bp	398;188	T
rs105751945 *POLE*	*POLE* 945	ACACACAGTAAGGAGACCGG ATGTGGGGCCTTGGTGTGA	58°C	472bp	107;117	T
rs121913399 *CTNNB1*	*CTNNB1* 399	ATCACTGAGCTAACCCTGGC GTTCTCAAAACTGCATTCTGACT	59°C	445bp	286;177	G
rs121913413 *CTNNB1*	*CTNNB1* 413	ATCACTGAGCTAACCCTGGC GTTCTCAAAACTGCATTCTGACT	58°C	405bp	153;114	C

Table 4: RFLP PCR primers. The table displays the SNPs and gene variants (*POLE* and *CTNNB1* SNPs) with simplified primer names, forward and reverse sequences, melting temperatures of the primers, expected size of the PCR product, and expected digest size and alleles tested.

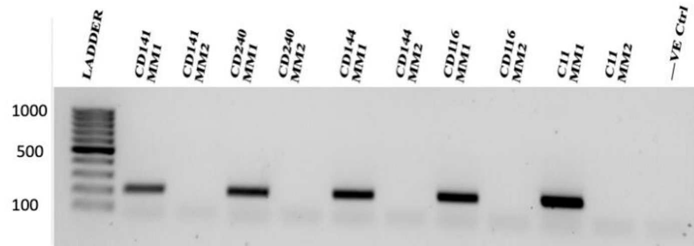


Figure 2: ARMS PCR Gel. The gel image of samples was electrophoresed after running PCR with ARMS primers. The fragments were seen in lane 1(samples with MasterMix1) at 214bp. And there were no fragments present in lane 2 (samples with MasterMix2).

Primer sequence	SNP and Gene variant	
	POLE **rs762000608**	***CTNNB1*** **rs121913412**
Forward inner	GTCATGAAGCAGAGAGGATGGCGAACG	ACTCTGGAATCCATTCTGGTGCCATTA
Reverse inner	AGGTGATGCACTTCTACCGCTGGCGTT	CCTTTACCACTCAGAGAAGGAGCTGTAGG
Forward outer	CTGGCCATGTCTCTGGTTCTGGGGAGTA	GCTTTTCTTGGCTGTCTTTCAGATTTGA
Reverse outer	AAACCTGTGCCCATTTCAGTTGGAGCGT	TGCATTCTGACTTTCAGTAAGGCAATGA
FI+RO size	214	192
All primers expected sizes	346	405

Table 5: ARMS PCR primers. The table displays the primer sequences of SNP and gene variants (*POLE* and *CTNNB1* SNPs) along with the expected size of the primers.
The sequencing primers were optimised using the same annealing temperatures as in table 6. PCR products were confirmed to be of the expected size by electrophoresis.

	Sequencing primers	
SNP and Gene variant	**Forward sequencing primers**	**Reverse sequencing primers**
rs105751943 *POLE*	CTCGAAAGGGGAGGGTACA	TGATCAGCTGCTCTCTTAGGT
rs105751943 *POLE*	CCCACACACTCTGCCTAGA	CTGCTCTCTTAGGTATGTCAGTT
rs105751945 *POLE*	ACACACAGTAAGGAGACCGG	ATGTGGGGCCTTGGTGTGA
rs121913399 *CTNNB1*	ATCACTGAGCTAACCCTGGC	GTTCTCAAAACTGCATTCTGACT
rs121913413 *CTNNB1*	ATCACTGAGCTAACCCTGGC	GTTCTCAAAACTGCATTCTGACT
M13	GTAGCGCGACGGCCAGT	CAGGGCGCAGCGATGAC

Table 6: Sequencing primers. The table displays the forward and reverse sequencing primers of *POLE* and *CTNNB1* SNPs.

3.2) PCR amplification of gDNA using the primers

The PCR was performed on gDNA samples of Endometrial Cancer and type 2 diabetes using the primers in table 4. The gDNA samples (EC and T2D, see table 2) rs1057519945 had fragments at the expected band size of 472bp see figures 3 and 4. For the primer set rs121913399, the fragments were seen at 445bp (expected size) see figure 5.

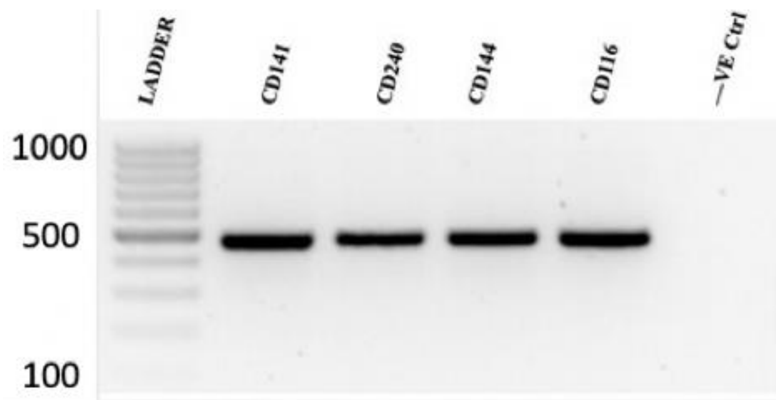


Figure 3: rs1057519945 PCR amplified samples gel. The gel image of EC samples that were amplified using the rs1057519945 set primers. The fragments were visible at 472 bp.

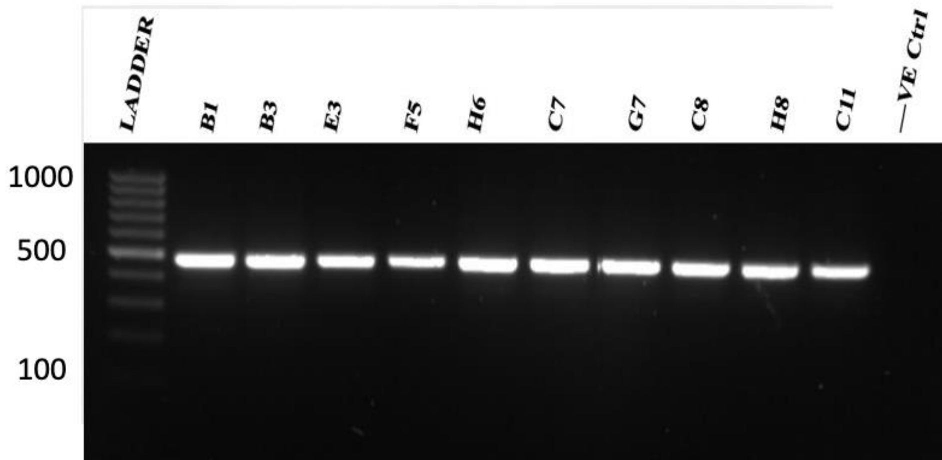


Figure 4: rs1057519945 PCR amplified samples gel. The gel image of T2D samples that were amplified using the rs1057519945 set primers. The fragments were visible at 472 bp.

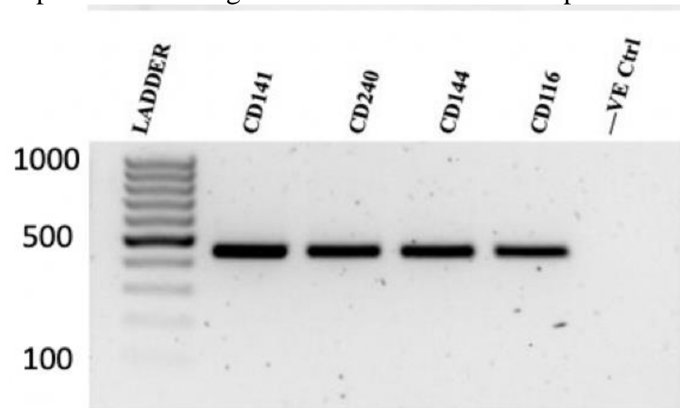


Figure 5: rs121913399 PCR amplified samples gel. The gel image of EC samples that were amplified using the rs121913399 set primers. The fragments were visible at 445bp.

3.3) PCR RFLP analysis

The RFLP PCR was performed to digest the fragments at the recognition site using the restriction enzyme

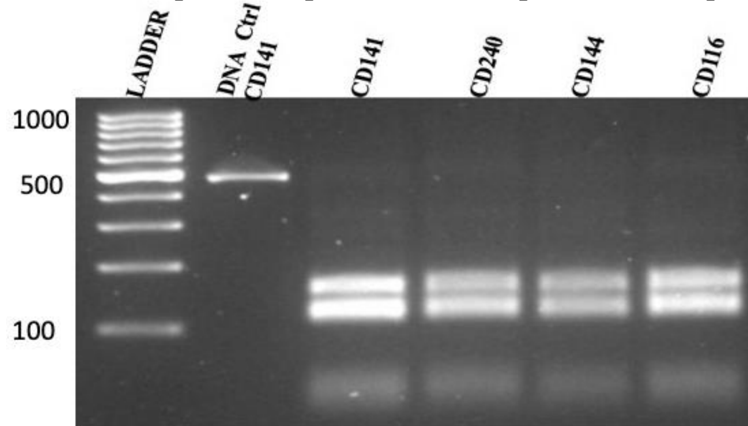


Figure 6: PCR RFLP gel. The gel image of EC samples amplified with rs1057519945 underwent an RFLP assay. The digested fragments are visible at 107bp and 117bp.

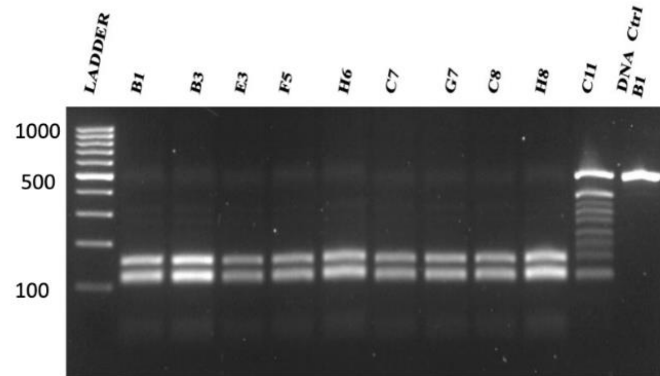


Figure 7: PCR RFLP gel. The gel image of T2D samples amplified with rs1057519945 underwent an RFLP assay. The digested fragments are visible at 107bp and 117bp.

For rs1057519943, no digestion was observed as the BspHI enzyme could not digest the allele at the recognition site. For rs1057519945, digestion was observed fragments were separated as the BtsCI enzyme digested the T allele at the recognition site and the fragments of 107bp and 117bp in size were visible (see figures 6 and 7). For rs121913399, no digestion was observed as the BfaI enzyme could not digest the allele. For rs121913413, DNA the fragments were separated, however, the allele that's being digested by the AluI enzyme was unclear as the enzyme digests A, C and T alleles. The fragments were of size 153bp and 114bp after the digestion.

3.4) DNA sequencing analysis

The sequencing was performed to investigate the presence of any changes in the gene. The samples were made with the mega mix gold master mix and were run on PCR to visualise the fragments on the gel. The exoSap and sequencing process was completed and sent for sequencing.

The sequencing was performed for SNPs of *POLE* rs1057519945, *CTNNB1* rs121913399 and rs121913413 with endometrial cancer samples, type 2 diabetes samples and normal panel DNA. The change for *POLE* rs1057519945 was expected at the location, c.1231G>T (p. Val411Leu), however, there was no change at this position in any of the samples. There was a polymorphism observed in CD563141 normal cancer DNA sample, which is an rs4883555 *POLE* intron variant (12:133250118 (GRCh37)) with sequence name GRCh37.p13 Chr 12 and change NC_000012.11: g.133250118C>T (See figure 8).

Rs1057519945

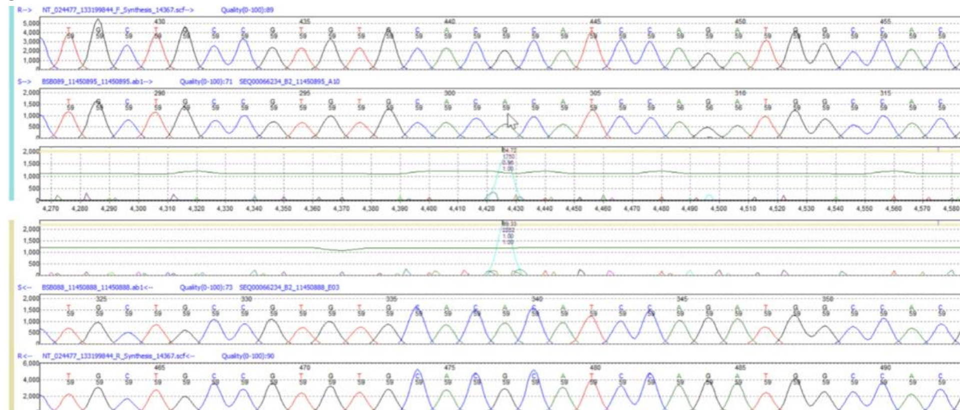


Figure 8: Sequencing of CD563141 normal cancer DNA sample with rs1057519945. The sequencing image of CD563141 normal uterus cancer DNA sample with rs1057519945. A change was observed at GRCh37.p13 Chr 12 and change NC_000012.11: g.133250118C>T.

The *CTNNB1* rs121913399 had no change present at the location NM_001904.4(*CTNNB1*): c.100G>C (p. Gly34Arg), however, there was a change seen in the CD563144 adenocarcinoma endometroid sample, which

is an rs121913403 *CTNNB1* missense variant NM_001904.4(*CTNNB1*): c.110C>T (p. Ser37Phe) (See figure 9). The sequencing data of SNP rs121913413 had no changes present in the expected location NM_001904.4(*CTNNB1*): c.122C>T (p. Thr41Ile), and there were no other changes seen in any of the samples rs121913399

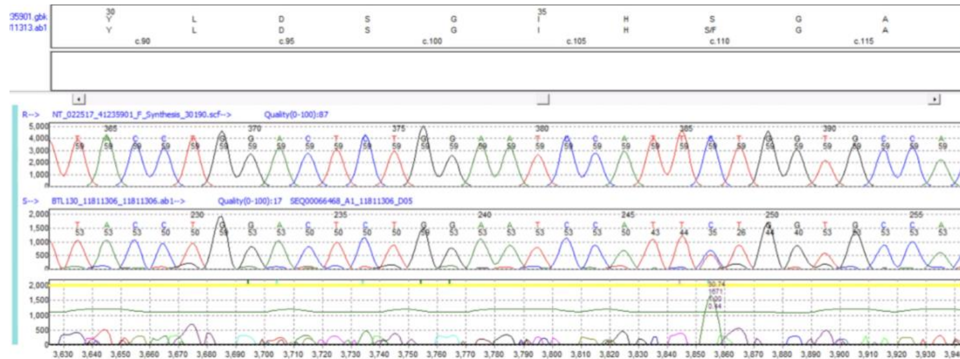


Figure 9: Sequencing of CD563144 adenocarcinoma endometroid sample with rs121913399. The sequencing image of CD563144 adenocarcinoma endometroid sample with rs121913399. A change was observed at NM_001904.4(*CTNNB1*): c.122C>T (p. Thr41Ile) position.

3.5) TaqMan SNP Genotyping assay for allelic discrimination of unknown DNA samples

The TaqMan PCR-based assay was performed to genotype the SNP and determined the allelic discrimination of unknown samples. To perform the TaqMan assay, the step One plus machine was set up with the layout design, and parameters and qPCR run settings were changed to 65 cycles for the good formation of groups and with default temperature. The allelic master mix was prepared and added to the 96well plate. Then negative control(water), positive control (G11) and DNA samples were added according to the sample layout design to the wells. The plate was loaded into the machine for the qPCR run. Following the run, the data was analysed. The allelic discrimination plot for the *POLE* SNP rs139075637 for the unknown DNA samples was heterozygous with A and T alleles forming green-coloured circular clusters and black-coloured undetermined negative controls (Figure 10A). The allelic discrimination plot for the *CTNNB*1 SNP rs121913228 for the unknown DNA samples was homozygous with T alleles forming red-coloured circular clusters and square-shaped, black-coloured negative controls (Figure 10B)

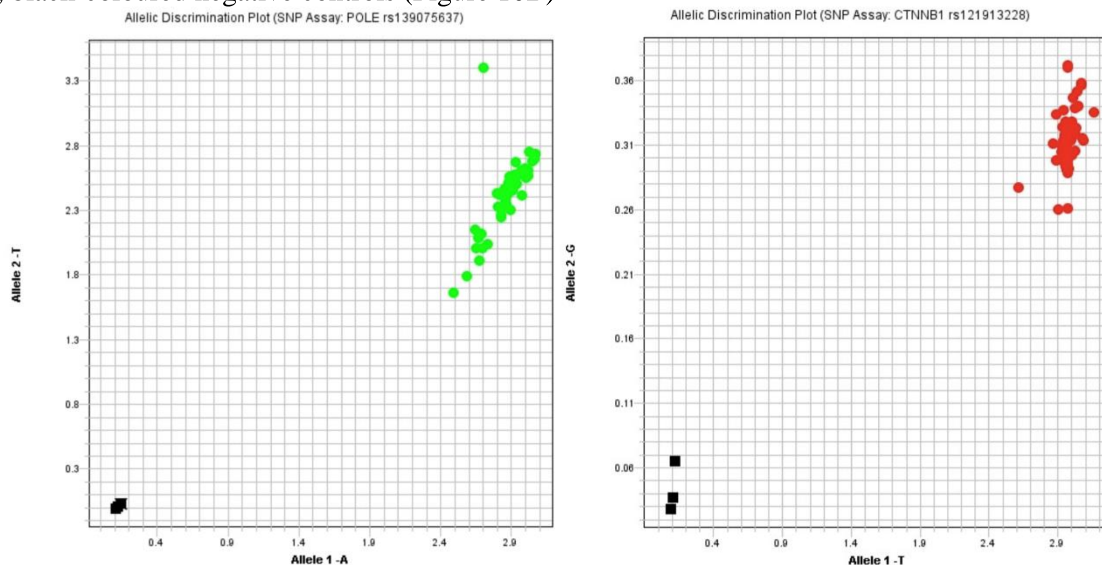


Fig. 10A) Heterozygous Allele 1 -A/Allele 2 –T Fig. 10B) Homozygous Allele 1 –TT [Experimental results performed by Xu Ning]

Figure 10: Allelic discrimination plot – A) The figure shows the **allelic discrimination plot** obtained for *POLE* SNP rs139075637 using the TaqMan genotyping assay. The green-coloured clustered circles present on the graph plot represent heterozygous genotypes (AT) and the black-coloured square cluster represents undetermined no template control. **B)** The figure shows the **allelic discrimination plot** obtained for *CTNNB1* SNP rs121913228 using the TaqMan genotyping assay. The red-coloured clustered circles present on the top left of the graph plot represent heterozygous genotypes (AT) and the black-coloured square cluster towards the bottom right end represents undetermined no template control.

Discussion

4.1) SNP identification

In this project, the SNPs from EC-related genes were identified using the search terms stated in section 2 and the primers were designed using Primer 1 and Primer 3 tools. Primer pairs for the relevant SNPs (see table 4 for SNPs), were optimised for gDNA of EC and T2D samples (see table 2 for samples).

4.2) RFLP assay analysis

4.2.1) rs1057519945: For SNP rs1057519945, the BtsCI enzyme digests for the allele T at the recognition site (TCTTCATCCACC) producing digested gDNA fragments of 107 bp and 117 bp (see fig 6. for restriction digest gel). Although allele T is not the ancestral allele (C), the allele T is present in the EC and T2D gDNA samples. The T allele with ClinVar Accession Id: RCV000560987.1 is associated with the Hereditary cancer-predisposing syndrome with Uncertain clinical Significance (Ensembl, 2021).

4.2.2) rs1057519943: No digestion was observed for SNP rs1057519943, and the ancestral allele of this SNP is G. As the BspHI enzyme digests for the T allele, this suggests that the T allele was not present in the gDNA samples. Although the T allele was not present, the T allele is associated with uterine Carcinosarcoma with potential pathogenic clinical significance with ClinVar Accession Id: RCV000438366.1 (Ensembl, 2021). The reason for no digestion in rs1057519943 can be due to the allele belonging to the minor frequency set in the population, and the sample for the same could have been unavailable. The sample size can be increased to have the frequency of the allele present in the samples.

4.2.3) rs121913399: For SNP rs121913399 no digestion of fragments using the BfaI enzyme (targeted for the A allele) was observed, suggesting that the allele was not present in the EC gDNA samples. Although allele A was not an ancestral allele, this was not as expected. Allele G is the ancestral allele with the highest population and 0.9999 frequency count followed by allele C with 0.01 frequency (Ensembl,2021). Allele A had no frequency count in the population and genetics information of rs121913399 in Ensembl. Moreover, there is no clinical significance of allele A related to endometrial cancer, however, alleles C and G have potential pathogenic clinical significance (Ensembl, 2021). The reason for no digestion in rs121913399 is due to the allele having no frequency in the population. The alleles G and C can be investigated further as they have been linked to EC.

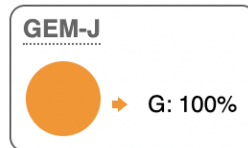


Figure 11: The allele frequency from population genetics. The allele frequency of the G allele from population genetics, Ensembl. The whole population had only allele G frequency in the rs121913399 variant.

4.3. DNA sequencing

The sequencing was performed for rs1057519945, rs121913399 and rs121913413 SNP primers (see table 3. for SNPs) with EC and T2D gDNA samples (see table 3 for samples). No sequence changes were present in the expected location in the samples in all 3 SNP primers (See figure for the sequencing results). Nevertheless, two different unexpected changes were identified in the CD563141 sample (rs1057519945) and the CD563144 sample (rs121913399).

4.3.1 rs4883555: The *POLE* transcript (molecule type) intron variant c.1359+43G>A with rs ID, rs4883555 (chr12:132673532 (GRCh38.p13)) was observed (See figure for the change) when the primer pair for SNP rs1057519945 was used to sequence the CD563141 normal Uterus g DNA EC sample. The variant has GMAF (global minor allele frequency), EVS (exome variant server), ExAC (exome aggregation consortium), CSVS (CIBERER Spanish Variant Server) and inner allelic frequency (See table for the frequencies) and this information about rs4883555 belongs to colorectal cancer (Esteban-Jurado et al., 2017). And women with colorectal cancer (Lynch syndrome) have a risk of developing EC at around 40-60% approximately (Nyiraneza et al., 2010). As rs4883555 is present in the CD563141 EC sample, there might be a direct link between rs4883555 to EC due to lynch syndrome. The rs4883555 has one publication, and there is limited information. The rs4883555 SNP has ancestral allele C. There is little published data on rs4883555, and it was an additional finding to the SNP that was being investigated.

Variant	c.1359+43 G>A
Location	intronic
rs ID	rs4883555
GMAF	0.4235
EVS	0.4302
ExAC	0.4957
CSVS	0.313
Inner allelic frequency	0.365

Table 7: rs4883555 variant details. The table displays the variant details of rs4883555 related to colorectal cancer. The location, ID, GMAF, EVS, ExAC, CSVS and inner allelic frequency of the variant are given.

4.3.2 rs121913403: CTNNB1 gene missense variant rs121913403 was observed (See figure for the change) when the primer pair for SNP rs121913399 was used to sequence the CD563144 adenocarcinoma endometroid sample. The missense variant rs121913403 has a change, NP_001895.1: p. Ser37Phe at the location NM_001904.4: c.110C>T. Potential pathogenic clinical significance of malignancy of the uterine cervix with somatic mutations was reported for this variation with accession ID RCV000427490.1 (Chang et al., 2015). This single base substitution mutation may have a moderate functional impact according to the Ensembl Variant Effect Predictor (VEP), which had a deleterious effect on the SIFT (Sorting Intolerant from Tolerant) with a score of 0 and a probably damaging score of PolyPhen with 0.994 (TCGA, 2022).

4.4 TaqMan genotyping assay

The genotyping assay for the *POLE* gene was done using the SNP rs139075637, and the gDNA samples were heterozygous with A and T alleles (See figure 10). The *POLE* missense variant rs139075637 has a 0.999

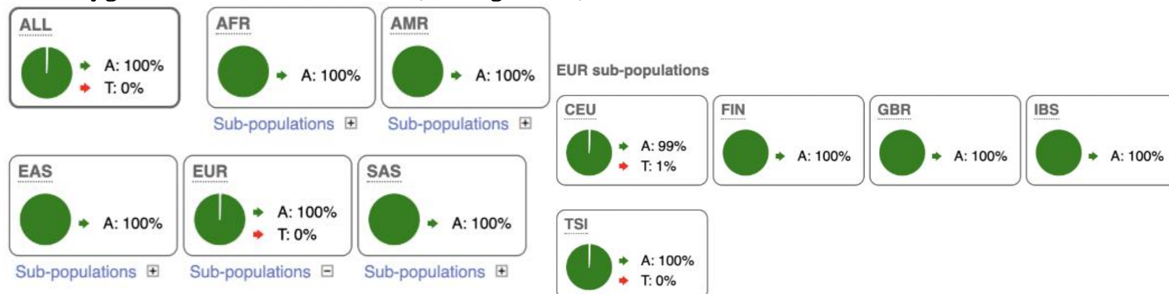


Figure 12: The overall allelic frequency of the rs139075637. The allele A is present in all the populations worldwide with 0.99 frequency and allele T negligible with 0 frequency.

frequency of allele A and 0.001 frequency of allele T in the population genetics, with A being the ancestral allele and T being the second most frequent allele (Ensembl, 2021). Data about the SNP establishes that the samples should have been homozygous (AA), however, the samples are heterozygous (AT). The results might be false positive results and can be sequenced further to achieve and confirm the true positive results (Liu et al., 2016). As the limited data suggests the heterozygous allele frequency in the population, the genotyping assay can be further redesigned according to the sample set to obtain true positive results.
For the gene *CTNNB1*, the assay was done using SNP rs121913228, and the gDNA samples were homozygous with TT alleles (See figure 10). The variant has the T allele as ancestral, and the frequency of the other alleles is not provided (Ensembl, 2021). Surprisingly, there is no data on the allele frequency in the databases except for the T allele confirming the homozygosity of the samples.

4.5 EC and T2D

Type 2 diabetes is a risk factor for endometrial cancer and has a potential role in increased mortality of EC patients (Zabuliene et al., 2021). Abnormal BMI levels and T2D can increase the risk levels of developing EC to 62% when compared to those without obesity (Njoku et al., 2022). Moreover, T2D can cause an independent risk for EC apart from sharing the same risk factors as EC (Urpilainen et al., 2021). The common risk factors of EC and T2D include age, sex, obesity and physical activity (Friberg et al., n.d.). Furthermore, the risk for EC is higher in people with T2D than in those without T2D. Although major associations of T2D with risk factors in developing EC and mortality have been outlined, there is limited evidence supporting the associations (Tsilidis et al., 2015). The effect of T2D on EC can be dependent on the tumour information, health factors and comorbidities of the patients. The experiments should be conducted with a well-designed structure on a large scale in populations with EC and T2D to gain strong evidence on the associations between EC and T2D in EC mortality. Furthermore, the studies that confirm the increased risk of EC from T2D should be continued with a large sample size to validate the findings.

4.6 Limitations of the study

This study has some limitations that are discussed below.
The main limitation of the study was EC samples, as there was one type of sample for each type of EC, the frequency of finding the SNP in the samples was reduced and produced inconclusive results. The sample size was small due to which the power of the study was reduced. As the experiment was designed to be a pilot study the experimental results can be further investigated to obtain statistically significant findings with a large. However, the experiment was designed to be a pilot study, and the experimental results can be further investigated to obtain statistically significant findings with an increase in sample size.
Using Primer 1 the primer sets obtained for SNP rs105751943, required redesigning due to the presence of secondary structure (17.82%) in the reverse primer which resulted in PCR fragments of unexpected size. The presence of secondary structures in primers results in yielding no product by binding to themselves instead of annealing to the DNA template.
It was not possible to determine the optimised annealing temperature for the ARMS primer pair for SNP rs121913412 although optimised at different temperatures ranging from 54°C to 64°C. The rs121913413 RFLP had digestion, however, the allele being digested by the AluI enzyme was not recognised as the enzyme digested C, A and T alleles.

4.7) Future research

The incidental changes found in the SNPs should be explored to authenticate the findings and investigated further for possible impact on clinical diagnosis. The sample size of the study should be increased to obtain statistically significant results. Increasing the types of samples of EC and T2D will have a higher frequency of finding the DNA changes. Redesigning the methods to effectively carry out the study in a longer period.

CONCLUSION

This study aimed to investigate clinically relevant SNPs in *POLE* and *CTNNB1* genes related to EC and T2D by selecting SNPs and performing PCR, RFLP-PCR, and TaqMan (qPCR) assay and DNA sequencing techniques. The RFLP-PCR results were in line with the expected results as seen in Ensembl (100 genomes

project). The sequencing results observed no DNA changes in the specific SNPs. However, other DNA changes (intron and missense variants) in rs1057519945 and rs121913399 belonging to colorectal and cervical cancer were observed. The genotypes obtained from the qPCR results for rs139075637 did not match the database results. The hypothesis that the project results will increase understanding of the potential impact on clinical diagnosis was not supported. Rather the changes that were observed can be investigated further for the risk indications and clinical impact on diagnosis. Though the experimental results differed from expected, the results can be used further for investigation and future research. The study can be improved by redesigning the primers with new parameters and choosing the SNPs and gDNA samples using TCGA in large numbers to observe the mutations and to obtain more accurate data. The variants can be explored further for the link between EC and T2D and investigate their potential prognostic significance to reduce the EC mortality rate.

ACKNOWLEDGEMENT
I would like to express my deepest and most sincere gratitude to my supervisor Dr Leagh Powell for her continuous guidance, support, invaluable patience and feedback. The Knowledge, expertise and motivation given throughout the journey are enormous. I would like to extend my sincere thanks to Ms Frances Cousins for the help and support during the dissertation period and for making the work easier. Special thanks to Ms Saeeda Bhatti for helping with the analysis of the results. I would like to acknowledge the biotechnology staff for their hard work and assistance with the project. I am also grateful to my family and friends for their moral support and motivation during this period. Finally, I am thankful to God for providing me with this opportunity and for all the blessings.

REFERENCES
Bell, J. R. (n.d.). *A Simple Way to Treat PCR Products Prior to Sequencing Using ExoSAP-IT ®*. www.biotechniques.com
Cancer [Online]. Available https://www.who.int/news-room/fact-sheets/detail/cancer [Accessed 8 July 2022]
Chang, M. T., Asthana, S., Gao, S. P., Lee, B. H., Chapman, J. S., Kandoth, C., Gao, J., Socci, N. D., Solit, D. B., Olshen, A. BSchultz, N., & Taylor, B. S. (2015). Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nature Biotechnology*, **34(2)**. https://doi.org/10.1038/nbt.3391
Chang, Y.S., Huang, H.D., Yeh, K.T. and Chang, J.G., (2017). Identification of novel mutations in endometrial cancer patients by whole-exome sequencing. International Journal of Oncology, **50(5)**, pp.1778-1784. https://doi.org/10.3892/ijo.2017.3919
Fiona Cunningham, James E. Allen, Jamie Allen, Jorge Alvarez-Jarreta, M. Ridwan Amode, Irina M. Armean ,Olanrewaju Austine-Orimoloye, et.al., Ensembl (2022). Nucleic Acids Res. 2022, **50(1)** 988-D995, PubMed PMID: 34791404. doi:10.1093/nar/gkab1049
Friberg, E., Orsini, N., Mantzoros, C. S., & Wolk, A. (n.d.). *Diabetes mellitus and risk of endometrial cancer: a meta-analysis*. https://doi.org/10.1007/s00125-007-0681-5
Gaudet, M., Fara, A.-G., Beritognolo, I., & Sabatti, M. (n.d.). *Allele-Specific PCR in SNP Genotyping*. https://doi.org/10.1007/978-1-60327-411-1_26
Jansen, A. M., van Wezel, T., EWM van den Akker, B., Ventayol Garcia, M., Ruano, D., Tops, C. M., Wagner, A., Letteboer, T. G., Gómez-García, E. B., Devilee, P., Wijnen, J. T., Hes, F. J., & Morreau, H. (2016). Combined mismatch repair and POLE/ POLD1 defects explain unresolved suspected Lynch syndrome cancers. *European Journal of Human Genetics*, **24**, 1089–1092. https://doi.org/10.1038/ejhg.2015.252
Ledinek, Ž., Sobočan, M., & Knez, J. (2022). *The Role of CTNNB1 in Endometrial Cancer*. https://doi.org/10.1155/2022/1442441
Liu, W., Glubb, D., Bai, S., Gaedigk, A., Riffel, A. K., Dehghani, M., Hartshorne, T., Floyd, K. C., Steven Leeder, J., & Rosenblatt, K. P. (2016). CYP2D7 Sequence Variation Interferes with TaqMan CYP2D6 * 15

and * 35 Genotyping. *Frontiers in Pharmacology | Www.Frontiersin.Org*, *1*, 312. https://doi.org/10.3389/fphar.2015.00312

Njoku, K., Agnew, H. J., Crosbie, E. J., Romano, A., & Urpilainen, E. (2022). Impact of Type 2 Diabetes Mellitus on Endometrial Cancer Survival: A Prospective Database Analysis. *Article*, **12, 1**. https://doi.org/10.3389/fonc.2022.899262

Nyiraneza, C., Marbaix, E., Smets, M., Galant, C., Sempoux, C., & Dahan, K. (2010). *High risk for neoplastic transformation of endometriosis in a Carrier of Lynch Syndrome*. https://doi.org/10.1007/s10689-010-9321-1

Park, M., Won, J., Choi, B. Y., & Lee, C. J. (2020). Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Experimental and Molecular Medicine*, **52(6)**, 963–977. https://doi.org/10.1038/s12276-020-0452-7

Radenbaugh, A. J., Singer, M., Ewing, A., Stuart, J. M., Collisson, E. A., Zhu, J., & Haussler, D. (2014). *RADIA: RNA and DNA Integrated Analysis for Somatic Mutation Detection*. https://doi.org/10.1371/journal.pone.0111516

Sharma, A., & Changotra, H. (2018). Mutagenic primer-based PCR-RFLP assay for genotyping IRGM gene promoter variant rs4958843 (C/T) Background: Single-nucleotide polymorphisms play an important role in the suscepti. *J Clin Lab Anal*, **32**. https://doi.org/10.1002/jcla.22346

Shen, Q., Abdullah, K. G., & Wang, Q. K. (n.d.). *The TaqMan Method for SNP Genotyping*. https://doi.org/10.1007/978-1-60327-411-1_19

The results from the GDC data portal, here are in whole or part based upon data generated by the TCGA Research Network [Online]. Available: https://www.cancer.gov/tcga [Accessed 21 July 2022]

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*, **19(1A)**, 68–77. https://doi.org/10.5114/wo.2014.47136

Travaglino, A., Raffone, A., Raimondo, D., Reppuccia, S., Ruggiero, A., Arena, A., Casadio, P., Zullo, F., Insabato, L., Seracchioli, R., & Mollo, A. (2022). *Prognostic significance of CTNNB1 mutation in early-stage endometrial carcinoma: a systematic review and meta-analysis*. **306**, 423–431. https://doi.org/10.1007/s00404-021-06385-0

Urpilainen, E., Arima, R., Karihtala, P., Puistola, U. and Ahtikoski, A., 2021. Metformin Associates with Aggressive Features of Endometrial Cancer in Women with Type 2 Diabetes. Anticancer Research, **41(2)**, pp.821-828. https://doi.org/10.21873/anticanres.14834

Uterine cancer statistics | Cancer Research UK (2018) [Online]. *Available*: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/uterine-cancer [Accessed 19 July 2022]

Zabuliene, L., Steponaviciene, L., Linkeviciute-Ulinskiene, D., Stukas, R., Arlauskas, R., Vanseviciute-Petkeviciene, R., & Smailyte, G. (2021). Clinical Medicine Risk of Endometrial Cancer in Women with Diabetes: A Population-Based Retrospective Cohort Study. *Journal of Clinical Medicine,* **10**. https://doi.org/10.3390/jcm10163453

Zhou, H., Chen, L., Lei, Y., Li, T., Li, H., & Cheng, X. (2021). Integrated analysis of tumor mutation burden and immune infiltrates in endometrial cancer. *Current Problems in Cancer*, ***45*(2)**, 100660. https://doi.org/10.1016/J.CURRPROBLCANCER.2020.100660

Copyright: *© 2023 by the Authors, published by Centre for Info Bio Technology. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license (https://creativecommons.org/licenses/by-nc/4.0/], which permit unrestricted use, distribution, and reproduction in any medium, for non-commercial purpose, provided the original work is properly cited.*