UNIVERSITY OF TARTU
FACULTY OF SCIENCE AND TECHNOLOGY
INSTITUTE OF MOLECULAR AND CELL BIOLOGY

Gene expression profiling of endometrial polyp

Master's thesis
(30 EAP)

Darja Tarassova

Supervisor's: Merli Saare, Ph.D
Tambet Tõnissoo, Ph.D
Amruta Pathare, Ph.D

TARTU 2024

INFOLEHT

Endomeetriumi polüübi geenisekspressiooni profileerimine

Endomeetriumi polüübid (EP) on naistel sageli esinevad healoomulised endomeetriumi koe vohandid, mille tekkimise molekulaarsed ja geneetilised põhjused on siiani ebaselged. Selles uuringus võrreldi endomeetriumi koe ja EP biopsiate transkriptoome, et tuvastada geenisekspressiooni muutusi ning analüüsiti rakulist mitmekesisust ühe-raku RNA sekveneerimise meetodil. Töö tulemused näitasid, et EP ja endomeetriumi transkriptoomi profiilid koebiopsiates on väga sarnased ja tuvastati ainult üksikud erinevused geenisekspressiooni tasemel. Polüübi üksikraku RNA sekveneerimisel kinnitati, et EP ja endomeetriumi rakuline mitmekesisus on väga sarnane, kuid tuvastati kolm geeni (*BNC2*, *CSMD1* ja *LINC01060*), mis olid erinevalt ekspresseeritud perivaskulaarses rakuklastris. Nende geenide olulisust EP patogeneesis peavad kinnitama edasised uuringud.

Märksõnad: enometriaalne polüüp, endometrium, geenisekspressioon, transkriptoom, üksikraku RNA sekveneerimine

CERCS (B220): Geneetika, tsütogeneetika

Gene expression profiling of endometrial polyp

Endometrial polyps (EP) are benign growths of endometrial tissue that occur frequently in women, the molecular and genetic causes of which are still unclear. This study compared the transcriptomes of endometrial tissue and EP biopsies to detect changes in gene expression and analyzed cellular diversity using single-cell RNA sequencing. The results of the work showed that the transcriptome profiles of EP and endometrium in tissue biopsies are very similar and only individual differences were detected at the level of gene expression. Single-cell sequencing of the polyp confirmed that the cellular diversity of the EP and endometrium is very similar, but identified three genes (*BNC2*, *CSMD1* and *LINC01060*) that were differently expressed in the perivascular cell cluster. The importance of these genes in the pathogenesis of EP needs to be confirmed by further studies.

Keywords: endometrial polyp, endometrium, gene expression, transcriptome, single-cell RNA sequencing

CERCS (B220): Genetics, cytogenetics

1. TABLE OF CONTENT

2. ABBREVIATIONS

BMI – body mass index;

CE – chronic endometritis;

DEG – differentially expressed gene;

DMEM – Dulbecco's Modified Eagle Medium;

ECM – extracellular matrix;

EMT - epithelial–mesenchymal transition;

EP – endometrial polyp;

FACS - fluorescence-activated cell sorting;

FBS – fetal bovine serum;

GEM – Gel Beads-in-emulsion;

GWAS – genome-wide association study;

HRT – hormone replacement therapy;

MC – menstrual cycle;

PCA – principal component analysis;

PCOS – poly-cystic ovarian syndrome;

qRT-PCR – quantitative real-time PCR;

RIN – RNA integrity number;

scRNA-seq – single-cell RNA sequencing;

3. INTRODUCTION

Endometrial polyps (EP) are a commonly occurring benign overgrowth of endometrial tissue in women, and their clinical significance remains undeniable. It is because these polyps can be associated with a variety of problems in the female reproductive system, including infertility, increased risk of miscarriage, menstrual cycle (MC) irregularities and dysfunctional uterine bleeding. There is also the possibility that EP may undergo malignant changes, which emphasizes the importance of their timely detection and treatment. EP can present with a variety of symptoms, such as abnormal uterine bleeding, infertility and even pain in the uterus, which significantly affects a woman's quality of life, but sometimes they can be completely asymptomatic. Statistics show that of all cases of polyps of the female genital tract, the proportion of endometrial polyps is 7.8-50% (de Azevedo et al., 2016; Dreisler et al., 2009; Tanos et al., 2017). Despite the high prevalence of EP, there is still limited research on the molecular aspects that could explain the cause and pathogenesis of EP.

There are different methods that can be applied to understand the molecular pathogenesis of the EP. One of the suitable methods for studying the molecular nature of the occurrence of EP is the study of the entire transcriptome, which can help identify endometrial overgrowth related genes. Through comparison between the endometrial polyp (EP) and healthy endometrial tissue, it is possible to identify the genes implicated in the molecular pathogenesis of EP.

Another way that can help to better understand the molecular changes inside the EP is to study cellular heterogeneity of the EP by using single-cell studies. Studying molecular changes at the single-cell level is essential for profoundly understanding the biological processes associated with EP. This method allows to look closer at the gene regulation in individual cells that can be missed during the whole tissue analysis. In this context, single-cell sequencing is a powerful tool for identifying unique subsets of cells within a polyp and identifying molecular markers associated with their development and function.

This work concentrated on two main tasks. First, to describe the molecular profile of EP through analysis of whole tissue transcriptome. By comparing EP with healthy endometrium, it is possible to identify differential expression of genes that may be associated with abnormal endometrial thickening and polyp formation. Secondly, to analyze the cellular population and diversity of EP by using single-cell RNA sequencing (scRNA-seq). By analyzing gene expression changes in different cell populations, it becomes possible to find molecular pathways driving EP formation.

4. LITERATURE REVIEW

4.1 Endometrial polyp overview

4.1.1 Overview of normal menstrual cycle

The endometrium is the inner layer of the uterus, which is necessary for the implantation of a blastocyst in the event of pregnancy and is subject to cyclical molecular, cellular and functional changes. If pregnancy does not occur, the mucous membrane is shed in the form of monthly bleeding, known as menstruation, and then a new inner layer of the uterus grows. The growth of endometrial tissue is a strictly controlled process by hormones (estrogen and progesterone), cytokines (for example, IL-8 and TNF-α)(Arici et al., 1998; Iwabe et al., 1998, 2000) and chemokines (for example, CXCL12) (Tsutsumi et al., 2011), which usually lasts from 24-35 days and, depending on changes in hormone levels, grows or degenerates.

The menstrual cycle (MC) can be divided into three main phases: menstrual, proliferative and secretory phase (Figure 1). The first day of MC is considered to be the day when bleeding begins, during which the mucous layer separates and a new one begins to re-epithelize (Ludwig & Spornitz, 1991) (Figure 1). The proliferative phase is characterized by the rapid growth of endometrial glands, stroma and blood vessels classified into early-, mid- and late proliferative phases based on histology, although the morphological changes are very subtle (Mazur & Kurman, 2005).

Ovulation is a key event in the ovarian cycle when a secondary oocyte is released from the mature Graafian follicle of the ovary into the fallopian tube, ready for potential fertilization by capacitated sperm. Ovulation occurs on the 14th day of a 28-day cycle, and after ovulation, the growth of endometrial glands and stroma continues, reaching its maximum thickness by days 21-22 (midsecretory phase, Figure 1). At this time, the endometrial glands become tortuous and secretory. From days 23 to 28, stromal cells increase in size and volume, acquiring an epithelioid shape and turning into predecidual cells. After the onset of endometrial degeneration, apoptosis occurs in the glandular structures on the 27th day, and on the next day, 28th, fibrin thrombosis forms in small vessels, which causes hemorrhage and release of red blood cells into the surrounding tissue. As a result, predecidual cells are destroyed, losing their extensive cytoplasm (Mazur & Kurman, 2005).

Usually the menstruation lasts 3-6 days, but in the case of pathology, there could be abnormal uterine bleeding occurring in the middle of the cycle, or in some cases there can be heavy bleeding during menstruation. The abnormal uterine bleeding and heavy bleeding during menstruation could be related to the presence of endometrial polyps in the uterine cavity.

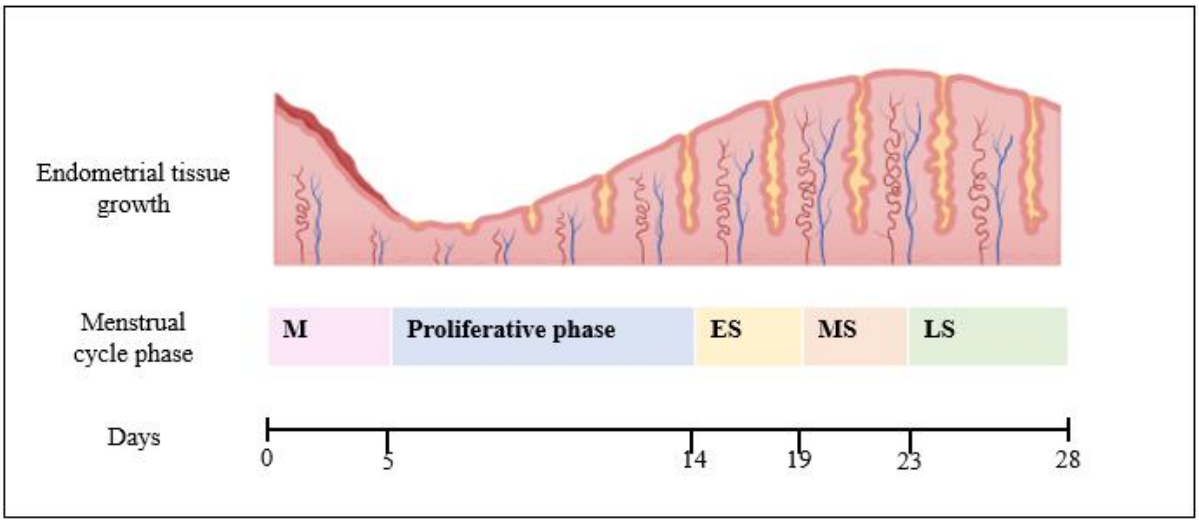

Figure 1. Endometrial growth changes during menstrual cycle phase. M – menstruation; ES-early secretory phase; MS – midsecretory phase; LS – late secretory phase.

4.1.2 Endometrial polyp etiology, risk factors and patology

Endometrial polyps (EP) are overgrowths that occur in the endometrial lining of the uterus. It develops due to the excessive growth of endometrial glands and stroma around a central blood vessel. Although EP is considered benign, a small percentage of all EP cases may retain atypical hyperplasia or carcinoma of the endometrium. However, such cases usually do not exceed 4%. (Nijkang et al., 2019; Reslová et al., 1999).

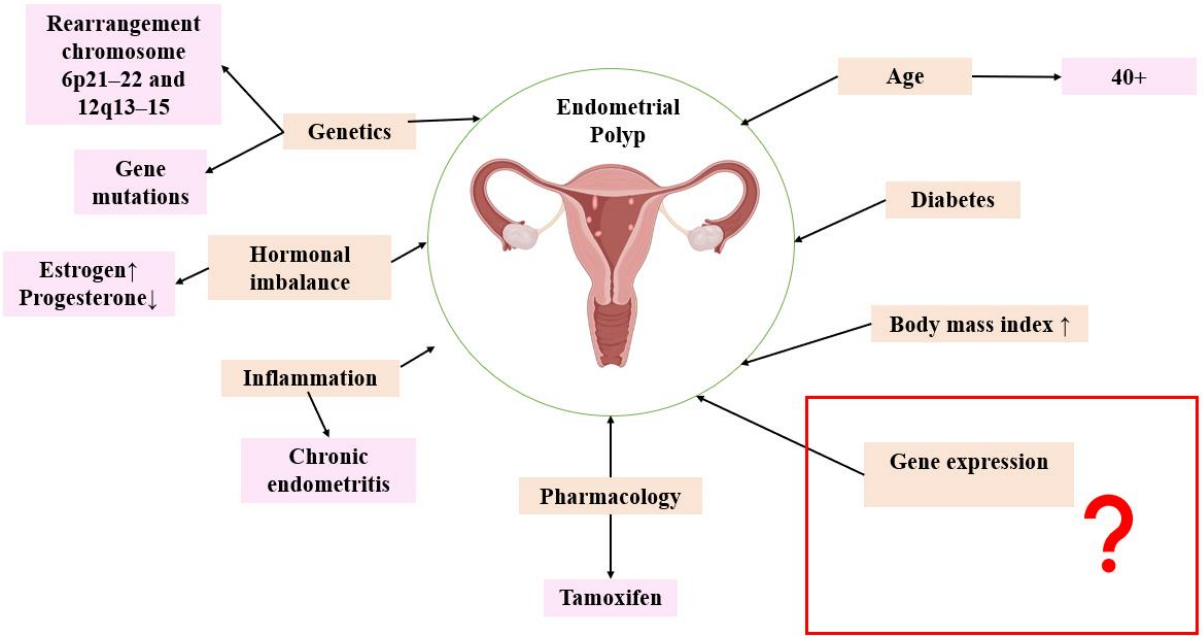

Figure 2. Graphical representation of possible factors contributing to the occurrence of endometrial polyps.

If we talk about the nature of the polyp, then mostly they are benign and are often accidentally discovered during the gynecologic examination, which suggests that most of EP does not cause any symptoms (Dreisler et al., 2009; Nijkang et al., 2019). However, not all endometrial polyps are asymptomatic, and they may cause a variety of symptoms such as uterine pain and abnormal uterine bleeding (Check et al., 2011; Nijkang et al., 2019; Tjarks & Van Voorhis, 2000).

EPs have been associated with primary and secondary infertility (Shokeir et al., 2004). The exact mechanism why EP causes infertility in women has not yet been established, but it has been suggested that EP causes a block of the cervical canal or fallopian tube, which in turn makes it difficult for sperm to reach the egg cell (Nijkang et al., 2019; Salim et al., 2011). Additionally, EPs have been shown to be accountable for recurrent implantation failure cases (Alansari & Wardle, 2012; Valle, 1983).

The etiology of EP represents a wide range of factors that are not fully understood. At the moment, several key aspects can be highlighted, including the woman's age, changes in her hormonal balance, and chronic inflammatory processes in the reproductive tract. In addition, chronic diseases such as diabetes are among the potential causes. According to recent publications, there is a connection between the use of certain medications and the development of this disease, and there is also active research into genetic factors influencing its development (Nijkang et al., 2019, Figure 2).

Pathogenesis of endometrial polyps

One of the causes for the occurrence of EP can be hormonal dysregulation. It is characterized by high exposure to estrogen and the unopposed action of progesterone. Estrogen is the primary hormone responsible for the proliferation of endometrial cells, and typically prompts the growth of the endometrial lining in preparation for embryo implantation during the follicular phase of the MC. Following this phase, progesterone acts to counterbalance estrogen's effects. However, when this hormonal balance is disrupted and progesterone fails to suppress estrogen adequately, the endometrial layer may thicken, potentially leading to the formation of polyps. (Z. Liu et al., 2010; Maia et al., 2009).

Genetic factors have also been associated with EP pathogenesis. It is known that women with EP had translocations in chromosomes 6 and 12, which led to changes in the structure of the endometrium (Vanni et al., 1993). Also, in women with EP, an increased expression of P53 and P63 proteins was observed (Nogueira et al., 2006; Vang et al., 2004). The P63 protein plays an important role in maintaining cell proliferation, apoptosis, differentiation and prevention of cell anging, in addition, it is assumed that P63 plays a significant role in the prevention of metastasis

(Bergholz & Xiao, 2012). The P53 protein regulates the cell cycle by blocking cell division under stressful conditions and preventing the development of tumors (Finlay & Levine, n.d.; Vousden & Prives, 2009). Studies have shown that in the case of hyperplasia in EP, there is an increase in p53 expression, which may indicate a protective cell response (Neto et al., 2013).

Risk factors associated with EPs

The risk of developing EPs increases with age. Women over the age of 40 are most often at risk, that is, they are women who are either in premenopause (the period that occurs from the first signs of menopause and lasts until the last menstruation) or in postmenopause (the period of menstruation absence for more than 12 months) (Check et al., 2011; Kolhe, 2018; Mansour & Chowdhury, 2022; Nijkang et al., 2019; Tjarks & Van Voorhis, 2000; Vieira et al., 2022). Also, it has been showed that postmenopausal women are even more likely to develop EPs than premenopausal or infertile women (Bettocchi et al., 2011; Di Spiezio Sardo et al., 2015; Pace et al., 1992; Salim et al., 2011). There was a noticeable trend that the likelihood of EPs increases with increasing reproductive age. In other words, the peak occurrence of EP occurs between 40 and 49 years of age ('AAGL Practice Report', 2012; Mansour & Chowdhury, 2022). But at the same time, the cause of the occurrence of EP in postmenopausal women is not yet fully known ('AAGL Practice Report', 2012).

Another risk factor for EP is obesity. Obesity is associated with increased endogenous estrogen production through elevated levels of aromatase enzymes that convert androgen in fats into estrogen. Since the obesity factor is directly correlated with poly-cystic ovarian syndrome (PCOS) and type 2 diabetes, these conditions could have the same effect of high levels of endogenous estrogen and increased incidence of EPs (L. Lu et al., 2023; Rshoud et al., 2022). However, the role of diabetes in the malignant transformation of EP is not clear (Nappi et al., 2009).

Moreover, the risk of EP development is also related to exogenous estrogen administration such as tamoxifen, which is a uterine estrogen agonist prescribed in breast cancer treatment in pre- and post-menopausal women (Powles et al., 2007). Tamoxifen has an estrogenic effect on the endometrium, increasing the incidence of EP development, hyperplasia, and endometrial cancer (M. E. Jones et al., 2012). Tamoxifen decreases the expression of estrogen receptors, increases progesterone receptors, and decreases the level of apoptotic cells in polyps, thus stimulating the unopposed proliferation (McGurgan et al., 2006). Also, there is a hypothesis that the relationship between the use of tamoxifen and the appearance of EP is due to the lack of antiproliferative activity of the hormone progesterone in the endometrial mucosa (Kossaï &

Penault-Llorca, 2020). A similar effect of continuous exogenous estrogen stimulation was observed in postmenopausal women on hormone replacement therapy (HRT), having a higher incidence of EPs. HRT is a hormonal therapy prescribed to women during menopause to relieve symptoms such as hot flashes, vaginal dryness or atrophy, osteoporosis, etc., which in turn are associated with changes in hormonal levels (G. Giordano et al., 2007; Stuenkel et al., 2015).
Chronic endometritis (inflammation of the endometrium) has also been found to contribute to EPs (Cicinelli et al., 2021). Chronic endometritis (CE) is an infectious disease that manifests itself in the form of chronic inflammation. Nappi et al. performed a meta-analysis to find a link between EP and CE, an attempt was made to pool the available data on a possible correlation between EP and CE (Nappi et al., 2009). As a result, it was found that more than half of the women participating in the study suffering from CE had EP. Of all polyps, the total proportion of CD-138-positive (CD-138 is membrane protein a marker of CE) is 70.73%. It proves the connection between CE and EP (Vitagliano et al., 2021).

4.1.3 EP and risk of malignancy

Although most EPs are benign, but it is noted that in less than one percent of all cases, an EP can develop into a malignant neoplasm and cause endometrial cancer (G. Giordano et al., 2007). According to Wethington et al. the risk of developing endometrial cancer is approximately 1.3% in women with endometrial polyps, and approximately 0.3% of endometrial cancers are associated with the presence of EPs (Wethington et al., 2011). However, the risks of occurrence are assessed based on age, body weight, predisposition, symptoms, polyp size, etc. (Nijkang et al., 2019). Considering statistical data, postmenopausal women with EP are more susceptible to serous adenocarcinoma, since this is associated with endometrial inactivity (Nijkang et al., 2019). Vaginal bleeding, pre- or post-menopausal endometrial polyps, and the age of the women contribute to an increased risk of malignancy of the EP (Hileeto et al., 2005; S. C. Lee et al., 2010).

4.1.4 Endometrial polyp diagnosis and treatment

EPs are frequently asymptomatic and can be discovered by chance during a routine gynecological examination, or when a couple is diagnosed with infertility. Additionally, in certain instances, abnormal uterine bleeding persists as a notable symptom, characterized by bleeding occurrences unrelated to the menstrual cycle. In the case of EP, it is observed in 10-40% of women (Vieira et al., 2022).

Diagnosis of an endometrial polyp requires a medical history (or anamnesis), followed by a physical examination and diagnostic tests. Currently, there are no non-invasive or blood based tests to diagnose the EP.

Transvaginal ultrasound is a traditional method for EP visual examination and diagnosis. On the ultrasound image, polyps give a hyperechoic signal, their foci with regular contours in the uterine cavity and along the endometrial line. However, there is a chance of confusing polyps with other pathological uterine neoplasms (Vieira et al., 2022).

Another method for diagnosing EP is hysterosonography, also known as saline-infusion sonohysterography, which is an ultrasound technique. Its advantage is that it allows to more accurately characterize and determine the location of the polyp in the uterus, it is also minimally invasive, safe, and well-tolerated (Salim et al., 2011).

Diagnostic hysteroscopy is considered the gold standard for diagnosing EPs. Hysteroscopy offers direct visualization of the endometrium, thus allowing better characterization of the polyp and suggesting further treatment (Carugno et al., 2020). There is only one way to treat EP using surgery – polypectomy. Additionally, an in-office polypectomy, a minimally invasive procedure during which the polyp is removed, is a valuable adjunct to diagnosis. The main advantages of an in-office polypectomy are that there is no need to make large incisions, it minimize discomfort or pain, it is cheaper and does not require a hospital stay (Garuti et al., 2008).

4.1.5 Morphological and histological characteristics of endometrial polyps

Transvaginal ultrasonography is used to diagnose and detect EP in the uterine cavity. The sizes, as well as the shapes of polyps, are different, and also the diameter of the EP varies from a few millimeters to a couple of centimeters. There are three common types of polyps – flat, sessile polyps and pedunculated polyps (Weschler, 2002; Nijkang et al., 2019; Vieira et al., 2022), Figure 3.

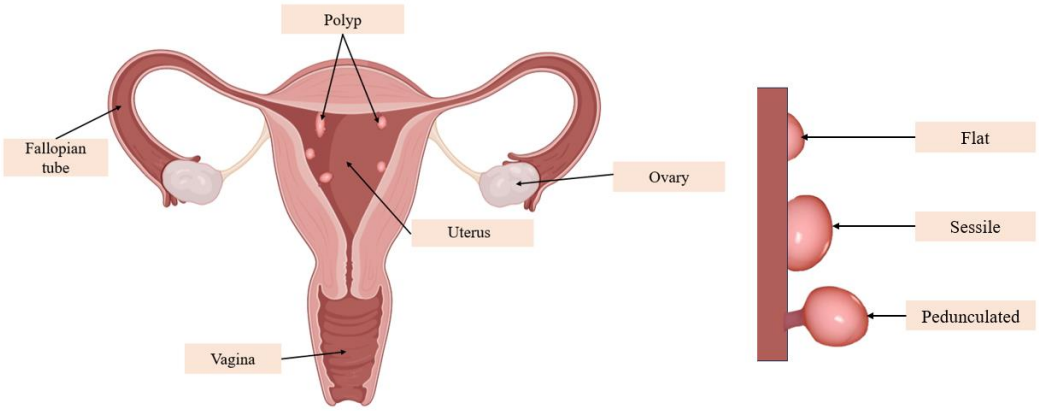

Figure 3. Endometrial polyp. Endometrial polyps have different sizes and shapes (flat, sessile and penduculated).

Morphologically EP can be divided into five categories: hyperplastic, atrophic, functional, mixed and myomatous. The most common type of polyps that is found in the endometrium is hyperplastic polyp. It is characterized by glandular proliferation, there are different shapes and sizes, and along the periphery there is a proliferative epithelium with mitotic activity. In this type of polyp, a decrease in the interglandular stroma may also be observed. Besides hyperplastic polyps, atrophic polyps can also be found. These polyps are commonly encountered, featuring dilated cystic glands lined with cuboidal epithelium devoid of mitotic activity, and separated by fibrous stroma.

Also, some EP can be functional, responsive to hormonal fluctuations, display proliferative changes or altered secretory activity compared to adjacent endometrium, often presenting with dense, edematous, or predecidual stroma. Polyps can also be mixed and demonstrate a combination of glandular structures within fibrous stroma, frequently associated with metaplasia, characteristic of the postmenopausal period.

Sometimes polyps can be myomatous and be distinguished by abundant smooth muscle tissue surrounding glands within the endometrial stroma. Metaplastic changes within myomatous polyps may present various characteristics, such as squamous, ciliated, mucinous, or eosinophilic features.

Histologically, EPs are relatively similar to normal endometrial tissue, and discriminating normal endometrial tissue from endometrial polyp is sometimes challenging and needs an experienced pathologist. In polyp histology, the fragments of normal endometrial tissue (stromal cells and endometrial glands) are usually mixed with thick-walled blood vessels ((Ciscato et al., 2020), Figure 4). In EP diffuse stromal hypercellularity can be seen that is

similar to the endometrial stroma at the proliferative stage. In addition, the polyp contains a collection of endometrial glands that form closely attached, discrete foci (Ciscato et al., 2020).

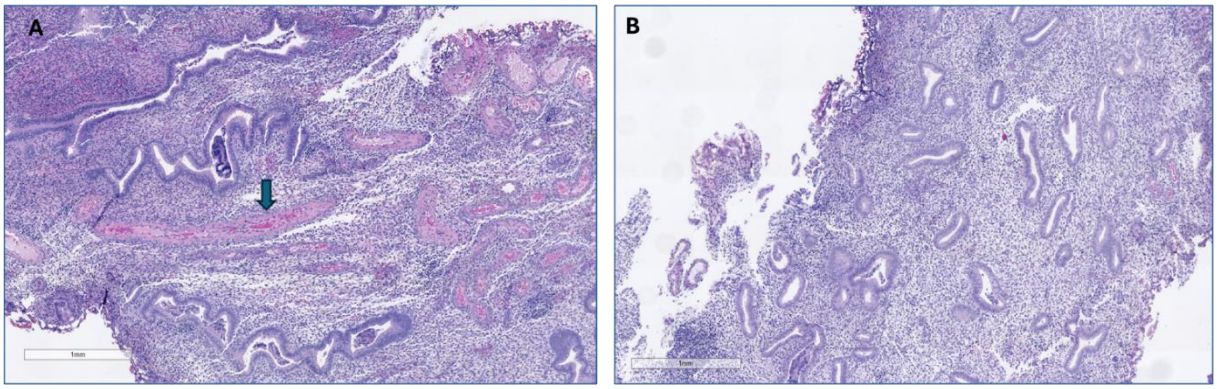

Figure 4. Illustrative figure of histology of endometrial polyp (EP) and endometrium in patient H01. **A.** EP with characteristic thick wall blood vessels (denoted with arrow). **B.** Normal endometrium tissue from the same women. Scale bar 1 mm.

4.2. A review of polyp genomic studies

4.2.1 Genomics of polyps

Genomic changes in DNA have been associated with the initiation and development of polyps in the endometrium, thereby opening the possibility of studying the basic molecular mechanisms of this pathological process. Genomic changes can cause both a one-time, and recurrent occurrence, since there is evidence that the probability of recurrence after polypectomy ranges from 2.5-44% (Cea García et al., 2020; Ceci et al., 2020; Ciscato et al., 2020; Gu et al., 2018). Lu et al. proposed that recurrent EP might be associated with gene polymorphisms, particularly focusing on *LIN28B* (M.-Y. Lu et al., 2022).. LIN28B protein is a highly conserved RNA-binding protein that is responsible for cell proliferation (Farzaneh et al., 2017). Humans produce two proteins, LIN28A and LIN28B, which are paralogs and their participation is noted in many hyperplastic diseases (Guo et al., 2021; Yang et al., 2021). Overexpression of *LIN28B* gene is observed in primary tumors (Huang et al., 2019; Lin et al., 2020; Lovnicki et al., 2020), and it was found that the *LIN28B* gene affects the recurrence of many tumors (Gadd et al., 2017). Lu et al. were the first who suggested that the *LIN28B* gene may be associated with relapse of EP after polypectomy in women of reproductive age. The study discovered a connection between polymorphism *LIN28Brs369065 C > T* and an increased risk of relapse of EP (M.-Y. Lu et al., 2022). Earlier studies have shown that *LIN28B* the gene regulates cell proliferation of the endometrial stroma through the LIN28b/let-7 axis (Canfield et al., 2019; Huang et al., 2019; H. Li et al., 2019; Wu et al., 2019; X. Zhang et al., 2020), which is an integral part of EP relapse (M.-Y. Lu et al., 2022).

Since EP is considered a common disease among women, Pathare et al. conducted a GWAS meta-analysis involving 25100 women with polyps of the female reproductive tract and 207193 women were included as controls (data from the FinGen study and the Estonian Biobank) (Pathare et al., 2024). This meta-analysis of genomic associations identified ten genomic loci with a high level of statistical significance, two of which (*rs2277339* and *rs1265005*) were associated with the *PRIM1* and *COL17A1* genes, respectively. These genes are thought to play a possible role in cell proliferation (V. A. Jones et al., 2020; W.-H. Lee et al., 2019). Also, several of the identified loci found in this study, have previously been associated with endometrial cancer and uterine fibroids, highlighting potential common mechanisms associated with tissue overgrowth and polyp development (Pathare et al., 2024).

Another meta-analysis was conducted by Kliewa et al. to identify the multivariate genetic architecture of reproductive disorders and their loading on a common genetic factor, and to determine whether this latent factor shares a common genetic architecture with female depression, including perinatal depression (PND) (Kiewa et al., 2023). The study analyzed the general etiology of diseases such as PCOS, ovarian cysts, endometriosis, dysmenorrhea, pelvic inflammatory diseases, menopause symptoms, menorrhagia, uterine fibroids, including polyps of the female reproductive tract, as well as the correlation of these diseases with depression. Kiewa et al. found the comorbidity of genital polyps and menopause (odds ratio [OR] = 9.6, CI = 9.2–10.0,p <0.001) and a strong correlation between menopausal symptoms and genital polyps (p <0.001). The analysis revealed one EP-associated gene, *EEFSEC* that in addition to the polyp was also associated with endometriosis and menorrhagia (Kiewa et al., 2023).

4.2.2 Gene mutation

It is also worth paying attention to mutations in the genes themselves, which can lead to the activation of signaling pathways associated with the overgrowth of the tissue and the development of tumors. It has been reported that people taking tamoxifen and have mutations in the gene *KRAS* in endometrial cells, can have the risk of both recurrence of EP and endometrial cancer (Tsujioka et al., 2009, 2010). *KRAS* is a GTPase that activates proteins necessary to distribute growth factors and other pathways (Wennerberg et al., 2005). Takeda et al. conducted a study on a group of premenopausal women not taking medications to exclude the influence of age groups and artificial factors to determine whether *KRAS* mutations affect the formation of EP. As a result, it was confirmed that mutations in the gene *KRAS* can lead to the formation of EP (Takeda et al., 2019). They found that slightly less than half (45.7%) of the examined EP exhibited multiple mutations in the RAS family genes, which was determined

using whole exome sequencing and targeted mutation analysis (Takeda et al., 2019). A study by Takeda et al. for the first time, highlighted the high frequency of pathogenic RAS mutations in untreated EPs, indicating a significant role of RAS mutations in the formation of multiple polyps (Takeda et al., 2019).

4.2.3 Gene expression and transcriptome studies in polyps

Genomic studies have shed first insight into EP pathogenesis, but the molecular mechanisms underlying the formation and development of EP still need to be explored. To date, only a few studies have focused on the analysis of the gene expression levels of specific genes or full transcriptome to reveal genes associated with EP development.

Transcriptome analysis reveals not only which genes are expressed but also their expression levels, providing insights that extend beyond genomic or proteomic data. Significant changes in gene expression can occur independently of genomic variations, underscoring the importance of transcriptomic investigations. While transcriptome changes provide valuable insights, their correlation with proteomic changes is often moderate to weak, emphasizing the need for integrated multi-omic approaches in biological research.

By far, some of the gene expression studies have explored the expression of specific genes, highlighting their importance in the development of EPs. For example, studying the expression of estrogen and progesterone receptors in the EP is important because these hormonal receptors are crucial in regulating the menstrual cycle. Therefore, determining the expression level of these receptors may help explain the molecular mechanisms associated with the formation of EP. Taylor et al. studied the expression of estrogen and progesterone receptors in relation to the *BCL2* and *Ki-67* gene expression (Taylor et al., 2003). Gene *BCL2* plays an important role in preventing cell death by performing an anti-apoptotic function (Amezcua et al., 1999; Pinheiro et al., 2014). Expression of *BCL2* is upregulated during the follicular phase in normal cyclic endometrium. In contrast, according to the study by Gompel et al., its expression virtually ceases by the secretory phase's onset (Gompel et al., 1994). However, high levels of BCL2 expression are observed in both complex (involving both glandular crowding and architectural complexity) and simple (without significant architectural changes in endometrial hyperplasia) hyperplasia, as noted in the study by Amezcua et al. (Amezcua et al., 1999). Ki-67 is a marker of cell mitotic activity, and an increase in its expression indicates increased cell mitotic activity and proliferation (Gerdes et al., 1984, 1991). The increased expression of *BCL2* and weak, absent expression of progesterone receptors was demonstrated in polyps in the proliferative phase of the cycle (Taylor et al., 2003). Increased expression of estrogen receptors was observed

in secretory phase polyps. This indicates that the increase in expression of *BCL2* followed by a decrease or cessation of apoptosis can be the cause of the pathogenesis of EP because the tissue does not undergo normal apoptosis during the late secretory phase of the normal menstrual cycle, and the EP is not shed along with the rest of the endometrium during menstruation (Taylor et al., 2003).

It is known that obese women have an increased risk of developing EP. In this regard, researchers studied increased body mass index (BMI) as one of the risk factors which influences the molecular mechanisms associated with the formation of EP. Based on the study by Giordano et al. the results indicate significant associations between gene expression *MKI67* and *BCL2* in EPs in postmenopausal women and their BMI. In particular, obese women have increased glandular expression *MKI67*, which highlights the influence of BMI on markers of cell proliferation. It was also found that polyps larger than 2 cm exhibit overexpression of *BCL2* in the stroma, which may indicate the role of this gene in the formation of large polyps (M. V. Giordano et al., 2020).

Polyps also tend to form at the site of inflammation and it has been suggested that diseases such as chronic endometritis or inflammation of the endocervical canal can cause a polyp. Since EP and cervical polyps belong to the group of female genital tract polyps, studies in the field of cervical canal polyps may provide valuable data to understand the molecular mechanisms of endometrial polyps. Liu et al. proposed that genes involved in inflammatory processes would have differential expressions in polyps and normal cervical mucosa (Y. Liu et al., 2012). This study analysed 509 genes, 16 of which were associated with inflammatory processes and had different expressions in polyps and cervical mucosa. Six genes, *IL-12P40, IL-17, IL-11, IFN-γ, TNF-α* and *CCR2*, had significantly increased expression in cervical polyps and the expression of *IL-1b, IL-2, IL-10, IL-15, IL-22, CCL-15, IFN-β1, TGF-β1* and *CD86* was lower in polyps than in normal tissue (Y. Liu et al., 2012). This may indicate that a decrease in anti-inflammatory mediators leads to an imbalance between anti-inflammatory and pro-inflammatory factors, which may make cervical tissue more susceptible to polyp formation (Y. Liu et al., 2012). Moreover, some genes were also found dysregulated in other types of polyps, for example, in nasal polyps, which suggests that polyps in different body sites may have similar inflammatory mechanisms (Ba et al., 2011; Molet et al., 2003; Saitoh et al., 2010; Shen et al., 2011). All of the above results indicate that the inflammatory process influences the occurrence of polyps by involving different ways of regulating signalling pathways, making the inflamed tissue more susceptible to hyperplasia (Y. Liu et al., 2012).

miRNAs are small non-coding RNAs that regulate gene expression at the post-transcriptional level, are also studied in the context of EP. A study conducted by Su et al. focused on the effect

of progesterone on the expression level of miR-320b and its effect on the target gene *MCL1* in the context of developing EPs (Y. Su et al., 2023). Altered expression of *MCL1* has been found in almost all types of breast cancer, including ER+ (estrogen receptor positive) cancer, at which there is a decrease in the expression level of miR-320 (Bolomsky et al., 2020; Louault et al., 2019). This miRNA specifically affects the gene *MCL1*, causing negative changes in its expression (Y. Su et al., 2023; Yu et al., 2016). The findings of the research revealed notable reductions in the expression of miR-320b and the progesterone receptor among patients with EP. Conversely, there was a significant increase in the expression level of *MCL1*. These results support the initial hypothesis suggesting that decreased expression of the progesterone receptor and miR-320, coupled with increased expression of *MCL1*, play a substantial role in the pathogenesis of EP (Y. Su et al., 2023). Another gene that performs an antiapoptotic function is *MCL1* across BCL-2 families (Qian et al., 2022).

To date, there is only one transcriptome study where the EP were compared with adjacent endometrial tissue (Chiu et al., 2024). Sequencing of 12 paired EP and endometrium samples showed that although the expression levels of most genes were quite similar, there were still 322 differentially expressed genes (DEGs) between the EP group and the adjacent endometrial tissue group. They found that among DEGs, 88 genes were significantly more expressed in the EP group and 234 genes had significantly higher expression levels in the adjacent endometrium (Chiu et al., 2024).

Protein-protein Interaction pathway analysis of DEGs revealed changes in polyp-associated protein pathways: the WNT signaling pathway and pathways linked to muscle function. While studying changes in the WNT signalling pathways, it became clear that the *DKK1* and *DKKL1* genes were responsible for pathway activation in the polyp group, while the *GPC3*, *GREM1*, *RSPO3*, *SFRP5* and *WNT10B* were downregulated (Chiu et al., 2024). Authors concluded that regulation of the expression of *DKK1*, *GPC3*, *BMP7*, *FGF17* and *WNT10B* genes was altered in EP, which means that the WNT signaling pathway was changed and this gene expression dysregulation might be a reason for polyps tissue proliferation. They also reported that genes responsible for the contraction of smooth muscles of the papillae were almost all downregulated in the EP, including *ACTA2*, *ACTG2*, *KCNMB1*, *KCNMB2*, *MYL9*, *PPP1R12B*, and *TAGS* (Chiu et al., 2024). This is important to consider because endometrial vascular smooth muscle cells continually proliferate during the MC, which in turn is an auxiliary mechanism for the angiogenesis of endometrial arterioles, which in turn controls rapid blood flow by regulating arteriolar diameter (Mulvany & Aalkjaer, 1990). If this control is lost, this leads to heavy and prolonged bleeding, which is also seen sometimes in EP. Taken together, Chiu et al. concluded that changes in WNT signaling and vascular smooth muscle pathways are probable contributors

to excessive proliferation and impaired vascular and stromal development. Dysfunction in these crucial homeostatic pathways may result in abnormal bleeding and infertility, commonly observed clinically in association with EPs (Chiu et al., 2024).

4.3 Methods to investigate transcriptome

In multicellular organisms, most cells have the same genetic material but can differ significantly in the transcriptome. The difference may lie in the pattern and degree of gene expression activity at a particular point of time. The transcriptome includes not only information about coding RNA molecules but also about regulatory molecules, which exhibit a wide range of variability. Transcriptome can be analyzed using different RNA sequencing (RNA-seq) methodologies. and each method is suitable for its specific research purpose. For example, total RNA sequencing is a method in which the entire transcriptome both coding and non-coding RNAs are sequenced, and the heterogeneity of gene expression in tissues, cells, etc., can be assessed. Also, transcriptome can be evaluated using tRNA sequencing and rRNA sequencing but both these methods are rarely used. Still, tRNA sequencing helps identify genetic variations and modifications and rRNA sequencing aims to study the assessment and composition of ribosomes. Targeted RNA sequencing can be used to analyze specific regions, genes, exons or other parts of the transcriptome and small RNA sequencing can provide data about small RNAs such as miRNA, siRNA and piwi-interacting RNA, which can also be sequenced to study gene regulation and other cellular processes.

mRNA sequencing is the most used transcriptome technique that analyzes all the mRNA molecules produced by the tissues and cells. The mRNA molecule is the result of the transcription of genes and serves as a main indicator of their level of expression, making it an important molecular marker for the analysis of genetic activity. The method enables to cover a larger spectrum of genes and transcripts in a sample, which allows to isolate genes for further study. Another advantage is that the method has high sensitivity and resolution, making it possible to pay attention to low-expressed genes. The Illumina platform and the TruSeq Stranded mRNA kit offer powerful tools for transcriptome analysis (Figure 5). The fundamental approach involves enriching mRNA libraries by selecting molecules with poly-A tails prior to cDNA synthesis.

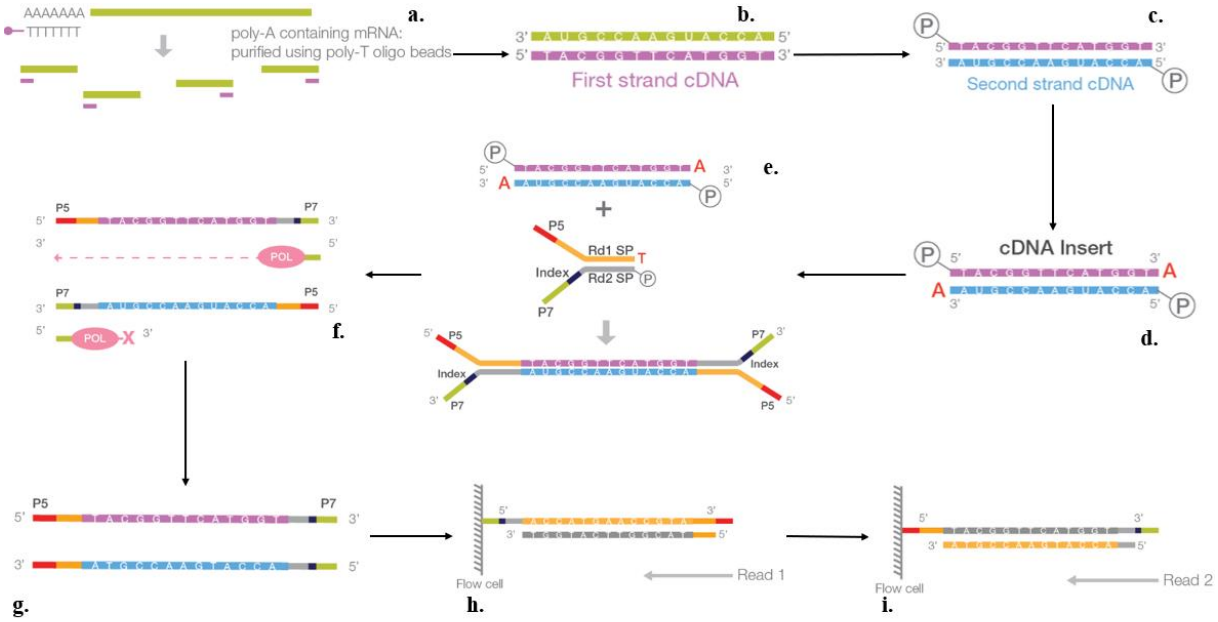

Figure 5. Graphical description of TruSeq Stranded mRNA. **a.** Purifying and fragmenting mRNA. **b.** Synthesizing first strand cDNA. **c.** Synthesizing second strand cDNA. **d.** Adenylating 3'ends. **e.** Ligating adapters. **f.** Enrich DNA fragments. **g.** LS final library. **h.** Cluster generation and Read 1 sequencing. **i.** Paired-end Turnaround and Read 2 sequencing (source: TruSeq stranded mRNA kit by Illumina).

Transcriptome can also be analyzed at the single-cell level. scRNA-seq is a relatively new molecular biology technique that provides gene expression information from individual cells in a heterogeneous population (Tang et al., 2009). While the mass sequencing of RNA isolated from the whole tissue allows us to determine only the average expression of genetic material from a large group of cells, but it does not provide a transcriptome profile of individual cells, and cell to cell interactions (Saliba et al., 2014). Because of its broad capabilities, scRNA-seq is a valuable tool for analyzing the cellular composition and gene expression of cells isolated from EP.

scRNA-seq protocols contain different steps – dissociation of the tissue into single cells, isolation of single cells followed by a reverse transcription reaction where the RNA is converted into cDNA and labelled with barcodes (Klein et al., 2015). Each barcode for each cDNA is unique, and upon completion of sequencing, it is possible to identify from which cell a particular transcript originates (Klein et al., 2015; Macosko et al., 2015). After reverse transcription, the same processes occur as during bulk sequencing: amplification, library creation and sequencing.

10X Genomics Chromium is a scRNA-seq execution system which is used for the isolation of single cells. Its main difference from other systems is the use of Gel beads-in-emulsion (GEM) technology. GEMs are tiny droplets that contain a gel bead, which serves as a reaction vessel

for individual genomic reactions. During sample preparation, genomic material from cells is combined with reagents and portioned into individual GEMs using microfluid technology. Further, each GEMs with genomic material undergo barcoding. As a result, every GEM gets a unique molecular barcode, which is necessary for future analysis, as they enable the assignment of sequencing reads back to their original cells. Library preparation starts after barcoding. GEMs with genomic material go through typical library preparation steps: fragmentation, adapter ligation and amplification. After library preparation, the libraries from multiple GEMs are pooled together. The final step is data analysing, during which, the unique molecular barcodes are utilized to demultiplex the reads, assigning them to their original cells. The graphical description for better understanding is presented in Figure 6.

The advantage of 10X Genomics Chromium technology is that the system has high throughput and does not require pre-cell sorting by fluorescence-activated cell sorting (FACS) (Chen et al., 2019). Another advantage of the technology is its capacity to process up to eight different samples simultaneously. This feature makes it suitable for analyzing multiple elements concurrently. For example, it allows for the simultaneous processing of material obtained from both the endometrium and the EP. Another important advantage of the 10X Genomics technology is that subsequent cell processing (reverse transcription, cDNA amplification, and library construction) can be carried out simultaneously in one tube, which makes this technology simpler and more convenient compared to others (Chen et al., 2019).

10X Genomics Chromium System also has some limitation. For example, limited control over cell input may result in rare cell populations being underrepresented or neglected altogether (Chen et al., 2019). Also, unlike the FACS-based method, it is impossible to determine which cells are collected for further processing and quality control (Chen et al., 2019).

Taken together, scRNA-seq is suitable for studying gene expression in the EP because the sequencing can help identify the cell types that make up the polyp. It can also help identify rare types of cells, or types that are unique to a polyp, and characterize them. scRNA-seq can also help identify cellular differences between the EP and the endometrium itself, thereby identifying the gene or genes behind the formation of the EP.

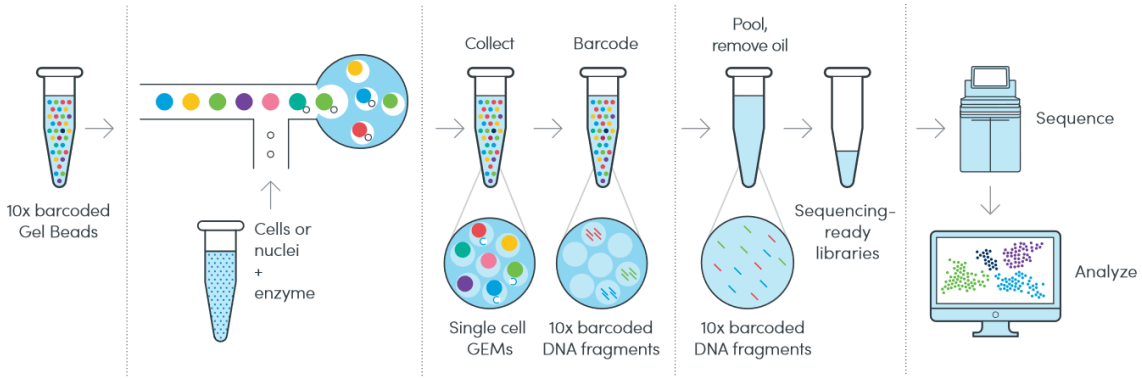

Figure 6. A graphical description of 10X Genomics procedure (Source: 10X Genomics).

5.EXPERIMENTAL PART

5.1 AIMS OF THE THESIS

The general aim of the study was to understand the gene expression alterations that could be lead to endometrial tissue overgrowth and formation of endometrial polyps (EP). The specific objectives of the thesis are as follows:

1. To compare endometrial tissue and EP whole biopsies transcriptome to reveal gene expression changes related to EP development.
2. To describe and analyzed the cellular heterogeneity and gene expression levels of the specific cell population of the EP and endometrium at the single-cell level.

5.2 MATERIALS AND METHODS

The Research Ethics Committee of University of Tartu approved the study protocols and the informed consent forms (No 221/M-31, 340T-12). All participants signed a written informed consent.

5.2.1 Study participants

Endometrial and EP biopsies were collected from women at the Women's Clinic of the University of Tartu. All women were diagnosed with an endometrial polyp during the hysteroscopy procedure. The average age of the women was 39.8 ± 6.1 years, average weight 65.8 ± 6.8 kg, average height 166.8 ± 5.7 cm and, average BMI 23.9 ± 3.7 kg/m^2. For the whole tissue transcriptome study, biopsies from both endometrium and polyps were obtained from twelve women, leading to a total of 24 samples. None of the participants smoked and had not received hormonal treatments for at least three months prior to the time of hysteroscopy. Some women have concomitant diagnoses such as endometriosis (n=2), schizophrenia (n=1), migraine (n=1) and iron deficiency anemia (n=1). Practically, all women had a regular menstrual cycle and only one had irregular MC. Additionally, three endometrium biopsies and five endometrial polyp (EP) biopsies were collected for single-cell sequencing. The average age of the women was 36.8 ±5.8 years and average BMI 24.9 ± 5.1 kg/m^2.

5.2.2 Collection and processing of tissue samples

During hysteroscopic surgery, endometrial polyps and endometrial tissue were excised, and they were collected and transported in a hypothermic preservation medium: HypoThermosol FRS Preservation Solution (Sigma-Aldrich) at +4°C. Endometrial and EP tissue samples were divided into pieces. A piece of tissue was placed into RNA later solution at +4°C for 24 hours and stored at -80°C for whole transcriptome assay. Another parts of the tissue were preserved in the cryopreservation medium containing 1x Dulbecco's Modified Eagle's Medium (DMEM, Gibico, Thermo Fisher Scientific), 30% fetal bovine serum (FBS, Biowest, Riverside, MO, USA) and 7.5% Dimethyl Sulfoxide Hybri-Max (Sigma- Aldrich). Cryotubes were placed in Nalgene Cryo 1°C degree (Thermo Scientific), at -80°C freezer overnight and then transferred to liquid nitrogen for further use. Cryopreserved samples were used for scRNA-seq study.

5.2.3 RNA isolation

miRNAeasy Mini kit (Qiagen, Germany) was used to extract RNA from biopsies according to the manufacturer's protocol. Briefly, 260 µl of RTL reagent (Qiagen, Germany) was added to

endometrial and EP biopsies and biopsies were disrupted by TissueRuptor (Qiagen, Germany). Further RNA extraction was performed using a column-based technique as per the manufacturer's instruction.

The extracted RNA was quantified using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific) at 260 nm.

RNA quality (RNA integrity number RIN) was assessed with Qubit RNA IQ Assay (Thermo Fisher Scientific). Samples with RIN ≥ 7 were considered for further transcriptome and real-time PCR analysis. RNA isolation and quality assessment were done by Darja Tarassova.

5.2.4 Whole transcriptome study

The methodology and study design of whole transcriptome mRNA sequencing of EP and paired endometrial tissue is presented in Figure 7. For a comparative assessment of transcriptomic changes in EP tissue and paired endometrial tissue, mRNA sequencing was performed on 24 samples collected from 12 women diagnosed with EP. After the isolation of RNA from the tissues, the RNA sequencing library was prepared using the TruSeq stranded mRNA kit (Illumina, USA) and TruSeq RNA CD Index plate, according to the manufacturer's instructions (Figure 8 – library preparation). Briefly, the isolated RNA concentration was diluted with nuclease-free ultrapure water to a final concentration of 500 ng in 50 µL. The library preparation involved steps, such as purification and fragmentation of mRNA, synthesis of first strand cDNA, synthesis of second strand cDNA, purification of cDNA, adenylation of 3'ends, ligation of index adaptors, cleanup of ligated fragments, amplification and cleanup of DNA fragments, check the quality and quantity of libraries and normalization and pooling of libraries. The prepared cDNA library was pooled and quantified using the Agilent High Sensitive DNA 1000 kit. The resulting libraries were single-end sequenced using an Illumina NextSeq 1000/2000 sequencer according to the manufacturer's instructions. Differential gene expression was analyzed by Illumina BaseSpace and Dragon software, which employs the Deseq method for differential expression. The library was prepared by Amruta Pathare and Darja Tarassova.

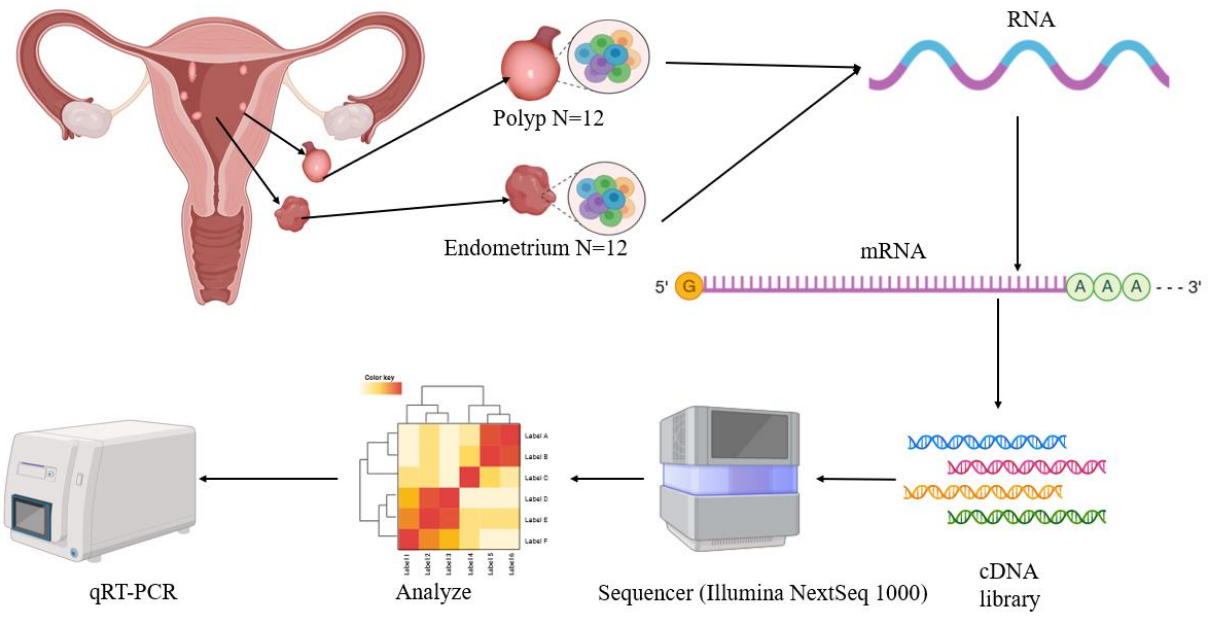

Figure 7. The graphical description ofstudy methodology and design. mRNA sequencing was performed for EP andpaired endometrial tissue. Isolated total RNA sample was converted into a cDNA library using the TruSeq Stranded mRNA kit by Illumina, sequenced by using Illumina NextSeq 1000 and analyzed.

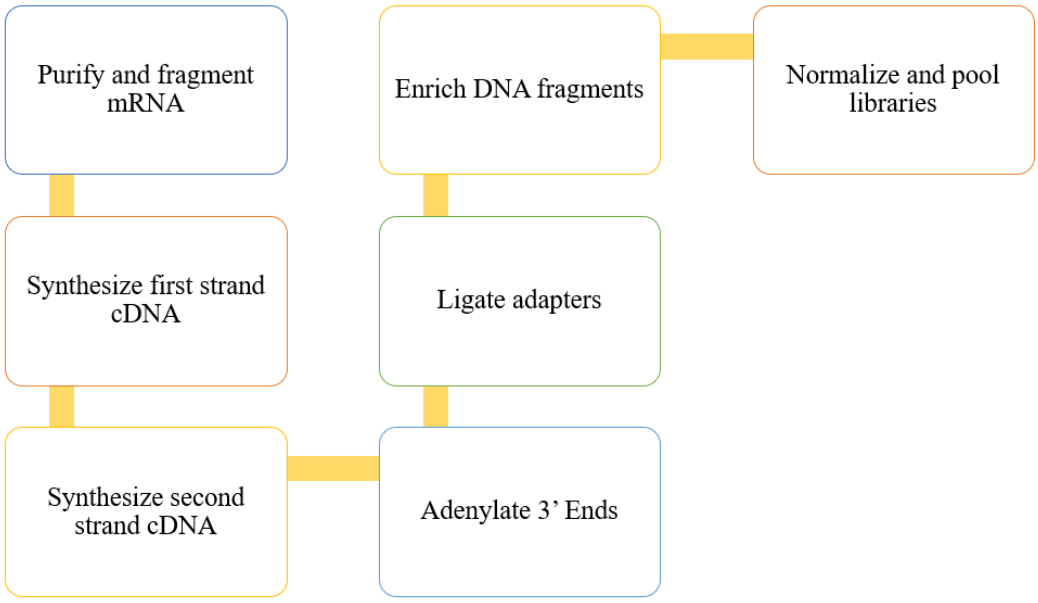

Figure 8. TruSeq Stranded mRNA library preparation.

5.2.5 Whole transcriptome data analysis

DEseq differential gene expression analysis was performed using Illumina BaseSpace. Briefly, fastq files from a sequencing run were generated using the BCL Convert application (v2.1.0), which were then processed by the DRAGEN FASTQ Toolkit application (v1.0.0) to trim adapter sequences, and DRAGEN FastQC + MultiQC (v3.9.5) was used. Further steps such as alignment with hg38 Alt-Masked v2, normalization and differential expression analysis by DEseq were performed using DRAGEN RNA (v3.10.4). Differentially expressed genes

between endometrial and polyp tissues having a logFold change >1.0 and FDR P value <0.05 were considered significant.

5.2.6 cDNA synthesis

First strand cDNA was synthesized from extracted RNA (250ng) by reverse transcription using RevertAid First Strand cDNA Synthesis Kit (Thermo-Fisher Scientific Inc. MA, USA) per the manufacturer's recommendation. cDNA was stored at -80°C until used for qRT-PCR. The reaction volume was 20 µl, which contained 250 ng of RNA and 15 µl of water.

5.2.7 Quantitative real-time PCR

Quantitative real-time polymerase chain reaction (qRT-PCR) was performed on the significantly differentially expressed genes (DEGs) from the transcriptome study results. The primers used for qRT-PCR (Table 1) were designed using the Primer 3 program (http://www.primer3.ut.ee) and sequence binding was checked using the GenomeTester 1.3 program (http://www.bioinfo.ut.ee/genometester). *SDHA* (succinate dehydrogenase flavoprotein complex subunit A) was used as an endogenous control in the experiment.
qRT-PCR was performed using 1 × HOT FIREPol EvaGreen qPCR Supermix, (Solis BioDyne, Estonia) according to the manufacturer's instructions. qRT-PCR was performed in a reaction volume of 20 µl, where the reaction mixture contained 14.5 µl ddH2O, 4 µl Master mix SYBER Green (5X), 0.5 µl primer (10 pmol) and 1 µl cDNA. All reactions were performed in duplicates. The qRT-PCR procedure was done by Darja Tarassova
The reaction was performed on a CFX Opus 96 Real-Time PCR System (Bio-Rad) using the following program:

1. 95°C for 10 minutes

Followed by 40 cycles:

1. 95°C for 15 seconds
2. 60°C for 1 minutes

Melting Curve:

1. 65°C to 95°C increment 0.5°C for 5 seconds

Table 1. Primers used for transcriptome validation study

Gene	Sequence (5'-3') Fw	Sequence (3'-5') Rev
*SDHA**	TGGGAACAAGAGGGCATCTG	CCACCACTGCATCAAATTCAT
KMT2B	CCTGATTCTGGGGTCTCTGG	GAGTACCCGGCTGCCAAA
DLEC1	CGGTGTAGTAATCTTCTGGGGA	TATTGCGGAAAATGAGCGGG
RAB3C	CTCTCTGACATTTTGTCGCAGAT	GGAAGACGAGCGGGTCATC

* *SDHA* - housekeeping gene used in previous studies (Vandesompele et al., 2002)

5.2.8 Single-cell RNA sequencing study

A graphical description of EP and endometrial tissue scRNA-seq study design is presented in Figure 9. Briefly, tissues were dissociated into single cells and prepared to perform scRNA-seq using 10X Genomics technology. All subsequent procedures were performed according to the manufacturer's instructions.

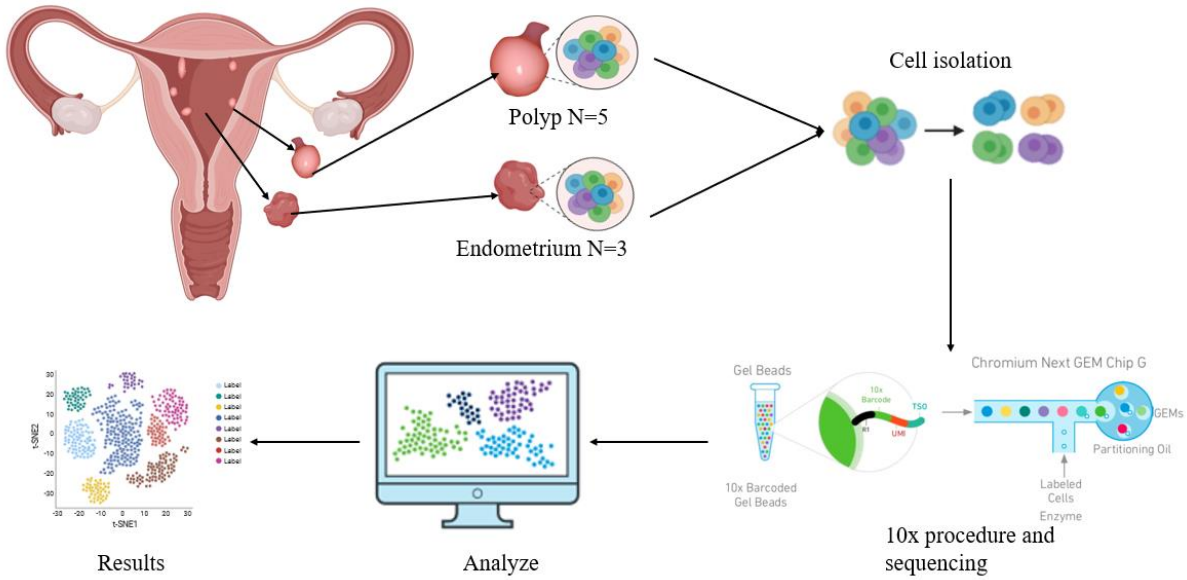

Figure 9. Graphical description of scRNA-seq study design using 10X Genomics.

5.2.9 Dissociation of tissue into single-cell suspension

Endometrial and polyp tissue dissociation into single cells was done according to already published protocols (Sarsenova et al., 2024) with modification. As polyp tissue is considerably thicker, the tissue dissociation was optimized (described below). Briefly, cryopreserved tissues were thawed in a 37°C water bath and were washed twice with a pre-warmed DMEM medium (Phenol red free, ccFBS 10% + antibiotics/antifungal, Corning 21521016) to eliminate cryoprotectant and excess blood. Enzymatic dissociation using collagenase, DNase, and Dispase II (2.5mg/ml collagenase from Clostridium histolyticum, Sigma C2674; 0.25mg/ml

DNase, AppliChem GmbH A3778 and 10mg/ml Dispase II, Gibco 17105-041 in 10ml of the DMEM medium) was performed for approximately 30 min at 37°C on a rotating shaker. Undigested tissue fragments were settled by gravity and re-digested enzymatically for an additional 15 min. The resulting single-cell suspension was filtered (30-µm cell strainer), pelleted by centrifugation at 300×g for 5 min, and treated with ACK lysis buffer (Gibco, A10492-01, following the manufacturer's instructions) to remove red blood cells. The cell counts and viability of the cells was evaluated by Corning ® Cell Counter by using Trypan blue dye. Live cells having viability >90% were enriched, washed with 0.04% BSA (Sigma A4503-50G) in Dulbecco's (D)PBS (1×, Gibco, 14190-144) on ice to eliminate ambient RNAs, and resuspended in 0.04% BSA in DPBS to attain a concentration of 700-1200 cells/µl as per the 10x Genomics protocol.

For tissue dissociation optimization, different tissue sizes, different dissociation enzymes and incubation time were used to optimize the final protocol and get cells with good viability (over 90%). Firstly, for standardization, different cryopreserved tissue samples with different weights (20 mg to 120 mg) were processed to get the desired concentration of cells as well as good viability. Tissues weighing 50 mg to 100 mg were shown to have a desired viability of over 90%, and a desired concentration of 1000 cells/µl was achieved. Tissues over 100 mg were observed to be clumped together and needed to be diluted more, because of which there is a probability of losing diverse cell populations. Further, based on the literature, we used different cocktails of enzymes such as only Collagenase, Collagenase and Dispase II collagenase and TypLE (Gibco, Cat No:12604013) at 37°C and protease (Bacillus licheniformic protease, Sigma P5380) at 4°C. The viability with cold protease was not sufficient, and the tissue was not completely digested even after 1 hour of incubation. Also, the tissue incubation time with enzymes was optimized to allow the complete digestion of tissue while retaining cell viability. The incubation time for 30 min, 45 min, 1 hour and 1.5 hours were evaluated. The combination of Collagenase and Dispace II with incubation time of a total 45 min digested the entire tissue without having any undigested glands remaining behind. After dissociation, due to the complete digestion of the whole tissue and some of the fibrotic parts, the minute cell debris or fibrotic threads were observed under the microscope. To get rid of it, different cell strainer sizes such as 30 µm, 40 µm, and 50 µm were used. Tissue dissociation optimization and final single-cell suspension for 10X Genomics were made by Darja Tarassova.

5.2.10 10X Genomics experiment

Single-cell RNA libraries were generated utilizing the 10x Chromium Next GEM Single Cell 3' reagent v3.1 kit (Dual index, 10x Genomics, CG000315 Rev C), inclusive of single cell 3' gel bead kit, library construction kit, and chip G kit, following the manufacturer's protocol (10x Genomics). Aiming for a targeted cell recovery of 3000 cells per sample, single-cell suspensions, single-cell 3' gel beads, and master mix for reverse transcription reagents were loaded onto a 10x Chromium microfluidic chip. The samples underwent processing in a 10x Chromium controller to generate GEM. Subsequent procedures, encompassing cell lysis, first-strand cDNA synthesis, amplification, and purification, were performed in accordance with the manufacturer's protocol to yield barcoded full-length cDNA. Library preparation steps were performed for all samples using the library construction kit to mitigate batch effects. Quality assessment of cDNA and single-cell libraries was performed using Agilent 4150 tape station (Agilent Technologies). Dual indexed single-cell libraries were combined, and pair-end sequencing was conducted targeting 35000 reads per cell employing NovaSeq PE150 (Illumina). The 10X experiment was done by Merli Saare, Amruta Pathare and assisted by Darja Tarassova.

5.2.11 scRNA-seq 10X Genomics data analysis

The 10x Genomics scRNA-seq data was aligned with refdata-gex-GRCh38-2020-A reference genome downloaded from the 10x website, and count matrices for each sample were created using the Cell Ranger (v7.0.0) software.

The 10x Genomics scRNA-seq data were analyzed with the CRAN package Seurat. In the data processing procedure, cells were retained based on criteria including a UMI count of more than 500 and a number of genes more than 200. High-complexity cell types (> 0.80) were kept, and cells with more than 15% of mitochondrial reads were filtered out. Genes with zero counts were removed by keeping only genes that were expressed within 3 or more cells. The Sctransform method of normalizing, estimating the variance of the raw filtered data, and identifying the most variable genes was applied to regress out the source of variation by mitochondrial expression and cell cycle phase. Next, the Sctransformed samples underwent Principal Component Analysis (PCA), and a minimum number (n = 24) of PCs were identified. Based on the resulting PCs, Uniform Manifold Approximation and Projection (UMAP) clustering, FindNeighbors, and FindClusters with 0.1 resolution steps from Seurat were run. The pre-processed data was submitted to the DoubletFinder (McGinnis et al., 2019) R package with default parameters. Identified cell doublets were then removed from each sample before integrating all nine

samples. 5000 highly variable shared features were used to integrate the samples to identify shared subpopulations across EP and endometrium. Canonical correlation analysis (CCA) was performed to identify shared sources of variation between the groups. In the second step of integration, mutual nearest neighbors (MNN) were identified, and incorrect anchors (cells) were filtered out to integrate the samples across conditions. After integration, dimensionality reduction techniques, PCA, and UMAP were used to visualize the integrated data. FindClusters with resolutions ranging from 0.2 to 1.4 from Seurat were used to find clusters in the integrated data. The UMAP technique was used to visualize the cell clusters. This integrated dataset with identified cell clusters was used for further downstream analysis. The data was analyzed by Ankita Sunil Lawarde.

5.2.12 Cell cluster annotation

Two approaches were employed to identify and label the major cell types in the integrated dataset. Firstly, the FindAllMarkers function from the Seurat package was utilized to detect differentially expressed genes within each identified cell class, using default parameters. Marker genes were shortlisted based on criteria including log2fold change (FC) > 1, adjusted p-value < 0.05, PCT.1 ≥ 0.7 (indicating higher expression of a marker in a specific cluster), and PCT.2 ≤ 0.3 (indicating lower expression of the same marker in other clusters). Statistical analyses, encompassing examination of cell proportions, comparison of total cell numbers, and cell numbers by cell cycle phases between EP and endometrium, were conducted using Fisher's exact test. Cell cluster annotation was done by Amruta Pathare.

5.2.13 Differential gene expression analysis

Differential gene expression analysis between the cell clusters of EP and endometrium and for each cell cycle phase was carried out using DESeq2 (v.1.40.1) R package (Love et al., 2014). The data was analyzed by Ankita Sunil Lawarde.

5.2.14 Statistical data analysis

The t-test was used to analyze the difference in *KMT2B* and *DLEC1* expression between endometrial tissue and EP in the whole tissue transcriptome part. A p-value <0.05 was considered statistically significant. Statistical analysis was performed using Microsoft Office Excel software. Gene expression levels were measured in duplicates and adjusted by the levels of the housekeeping gene *SDHA*, with ΔCt serving as a measuring unit for gene expression,

demonstrating the difference in expression between the housekeeping gene (*SDHA*) and either *KMT2B* or *DLEC1*.

5.3 RESULTS

5.3.1 Results of whole transcriptome analysis

Twelve EP and paired endometrial tissues were sequenced using the TruSeq stranded mRNA kit (Illumina, USA). Data analysis revealed a total number of 16729 genes after filtering. The most highly expressed genes in the EP and endometrium were *EEF1A1* and *COL3A1*. Other highly expressed top-list genes are presented in Table 2. The top list of highly expressed genes was identical between the EP and endometrium.

Principal component analysis (PCA) of all transcriptome data was used to see if samples clustered according to tissue types (endometrium and EP) and menstrual cycle phase at the time of collection, (proliferative and secretory phase) (Figures 10 A). Samples with similar trends in gene expression profiles should group together and form distinct clusters in the graph. Based on PCA clustering data, the overall gene expression profiles of EP and endometrium did not show obvious clustering according to their disease status (Figure 10 A), indicating that the gene expression profiles of these two tissue types are very similar.

To determine the effect of the different phases of MC on the gene expression in EPs, all EP and endometrial samples, irrespective of the MC were compared. Data analysis resulted in two DEGs, an upregulated gene *KMT2B* (p-value<0.01) and a down-regulated gene *COL9A1* (p-value<0.01) found in EP. Additionally, it was observed that two samples (H3, P48) exhibited different gene expression profiles than the entire study group and these were removed as outliers for further analysis (Figure 10B, PCA after outlier removal). After removing outliers besides *KMT2B*, we identified four additional DEGs with upregulation of *RAB3C* genes (p =0.02) and downregulation of *COL9A1* (p =0.01), *GPR22* (p =0.02) and *IGKV1-33* (p =0.02).

To eliminate potential biases related to the MC, EP (n=9) and endometrium (n=9) from the proliferative phase were compared. Data analysis revealed four genes with significantly different expression levels. Gene *KMT2B* (p=0.02) and *DLEC1* (p-value<0.01) were upregulated, and *RAB3C* (p-value<0.01) and *COL9A1*(p-value<0.01) were downregulated. The DEG in EP and their sequencing read counts are presented in Table 3.

We also analyzed only the samples from secretory phase (three pairs of EP and endometrium) of the MC. A secretory phase comparison of EPs versus endometrium revealed 5 DEGs, *AL137782.1, XPNPEP3, AL731577.2, C9orf129* and *SLC25A21*. Amongst them, *C9orf129* is a pseudogene, whereas two non-coding RNAs, *AL137782.1* and *AL731577.2*. Protein coding genes *SLC25A21* and *XPNPEP3* are involved in mitochondrial functions.

In addition, to understand the effect of the menstrual phases on gene expression, we compared EPs in secretory versus proliferative phases. The same comparison was performed in

endometrial tissues to ensure that they undergo the cyclic transcriptome changes from proliferative to secretory phase, like healthy endometrial tissues The comparison in EPs showed 88 DEGs, while the comparison between secretory and proliferative endometrium revealed 232 DEGs.

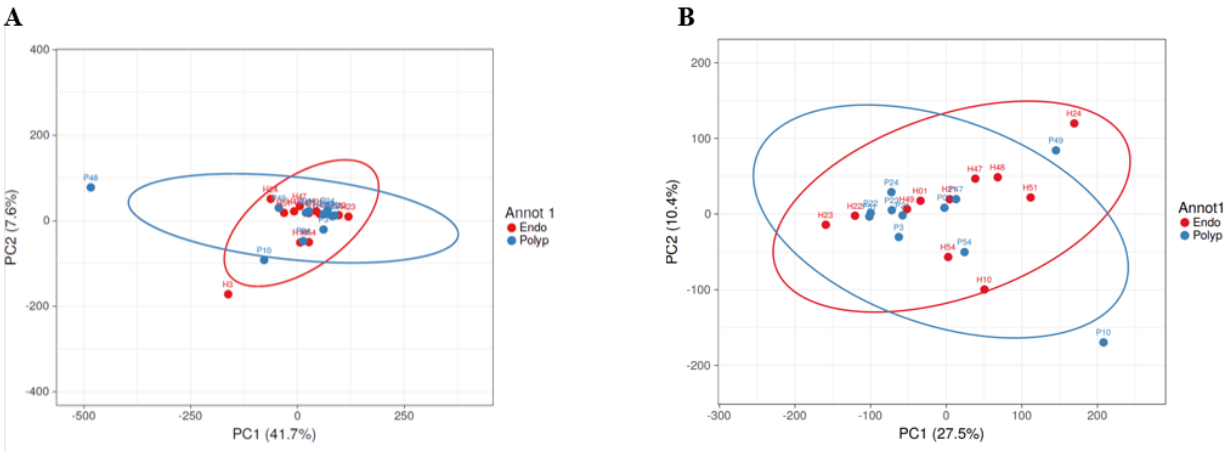

Figure 10. Principal component analysis (PCA) of endometrial polyp (EP) and endometrium. **A.** PCA of all studied EP and endometrium. **B.** PCA after removing outliers (H3 and P48).

Tabel 2. Top 10 of the most abundantly expressed genes in endometrium and EP

Endometrium	Average read counts	Endometrial Polyp	Average read counts
EEF1A1	178226.7	*EEF1A1*	199048.8
COL3A1	137336.2	*COL3A1*	122384.1
COL1A1	89919.4	*COL1A1*	86269.3
SFRP4	65975.7	*SFRP4*	64153.9
COL1A2	50415.7	*COL1A2*	48112.1
ACTB	48983.3	*ACTB*	47270.0
TPT1	45946.8	*TPT1*	50266.5
SPARC	44098.0	*SPARC*	39734.3
EEF1A1P5	41215.8	*EEF1A1P5*	49479.5
VIM	35027.4	*VIM*	37981.9

Table 3. Sequencing read counts in diferentially expressed genes between the EP and endometrium samples.

Gene	H01	H23	H21	H22	H47	H48	H49	H54	H24	H3	H10	H51
DLEC1	17	2	11	3	14	3	8	8	10	33	21	44
RAB3C	10	6	2	0	12	148	305	20	1	7	1	2
COL9A1	1	0	21	5	45	0	6	2	15	4	3	10
KMT2B	311	120	312	174	599	526	316	472	624	1069	379	464

Gene	P01	P23	P21	P22	P47	P48	P49	P54	P24	P3	P10	P51
DLEC1	48	4	9	9	17	149	45	57	5	28	58	9
RAB3C	0	2	1	0	4	5	1	0	1	2	0	2
COL9A1	1	7	10	7	6	1	0	0	1	1	3	8
KMT2B	294	281	265	224	508	1718	287	587	315	302	753	255

H - endometrium sample (n=12); P – Endometrial polyp tissue (n=12).

Three potential candidates were selected for further validation study by qRT-PCR: *KMT2B*, *RAB3C*, and *DLEC1*. These genes exhibited differential expression between the EP and endometrium and had statistical significance, which makes them promising genes for subsequent comparative analysis (Figure 11).

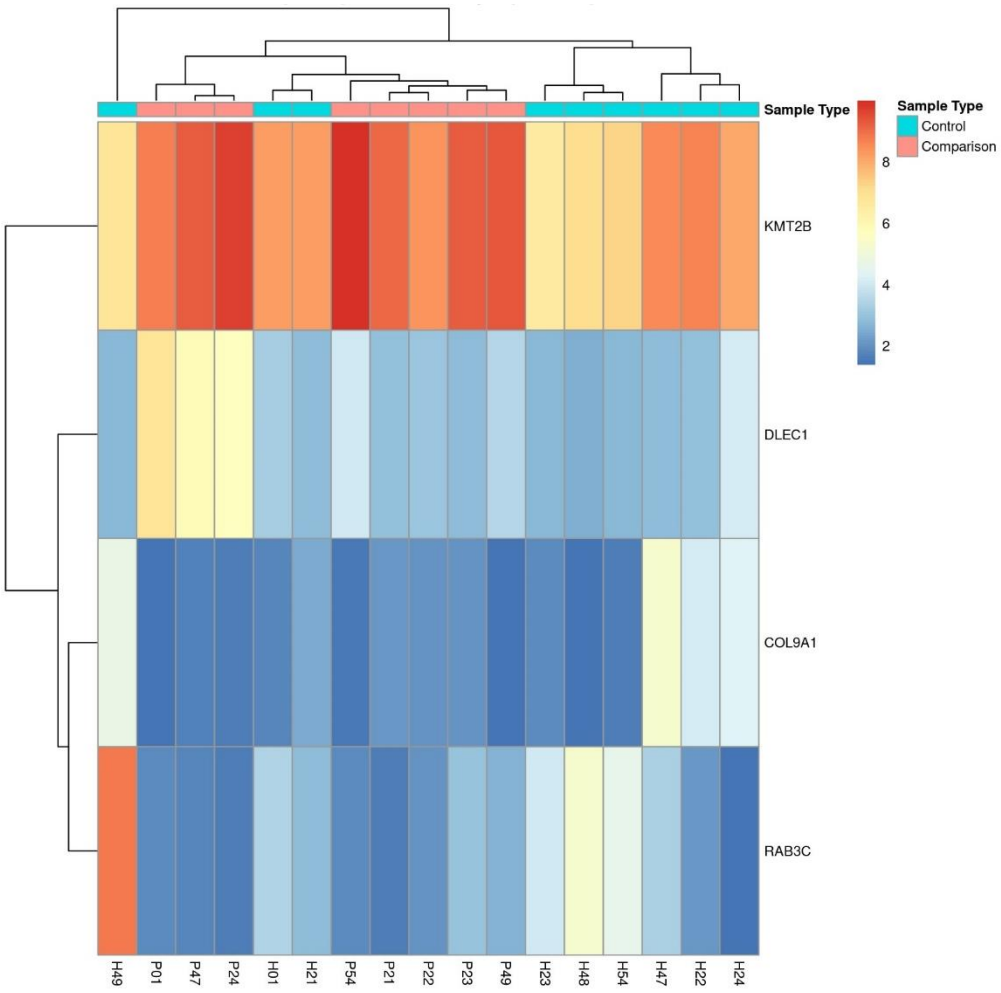

Figure 11. Heatmap of top 4 the most expressed genes in EP of proliferative phase.

5.3.2 qRT-PCR results

The expression levels of *KMT2B, DLEC1, RAB3C* genes, and housekeeping gene *SDHA* were determined by qRT-PCR in 12 endometrial and 12 EP samples. Normalized gene expression level (ΔCT) differences between the cells from EP and endometrium were evaluated by Student's two-sided t-test, and $P < 0.05$ was considered a statistically significant difference.

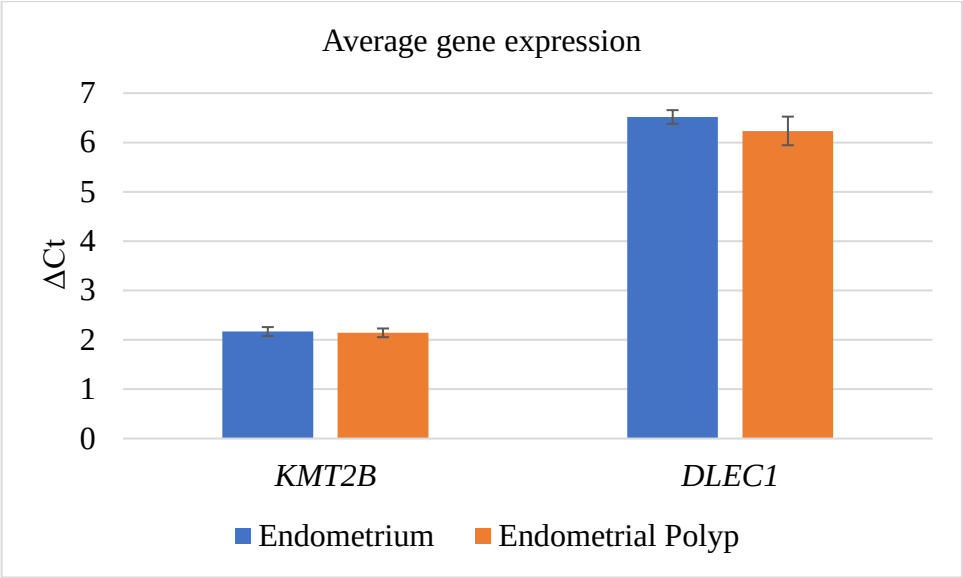

Figure 12. Relative *KMT2B* and *DLEC1* mRNA expression of level in the endometrium and EP. EP (n=12), endometrium (n=12), experiment was done three times.

The results of qRT-PCR analysis indicate that there are no statistically significant differences in the average gene expression levels for the *KMT2B* and *DLEC1* genes between endometrial and EP tissues ($p = 0.8$ for *KMT2B* and $p = 0.4$ for *DLEC1*) (Figure 12). This suggests that the expression levels of both genes remain almost the same in both endometrial tissue and EP.

5.3.3 Results of single-cell RNA sequencing

scRNA-seq was done in 4 endometrium and 5 EP samples. A total of 16688 reads were obtained from scRNA-seq EP. After the quality control analysis and filtration of doublets, 5721 cells from endometrial tissues and 10967 cells from EP tissues were processed for further analysis. The detailed information of samples and results is presented in Table 4.

Table 4. Single-cell RNA sequencing (scRNA-seq) general data analysis

Sample	H01	H21	H23	P01	P16	P21	P22	P33	
Estimate Number of Cells	4092	1629	1760	1961	2566	2909	694	1077	Total: 16688
Mean Reads per Cell	42263	119686	84751	79726	67486	29307	25648	27359	Average: 105 023.25
Total Genes Detected	29901	28663	29109	29578	29888	29307	25648	27359	Average: 32 179.38

The cells in both the EP and endometrium were categorized and grouped into 20 clusters based on the expression of specific marker genes for each cell type (Figure 13). These clusters were then further merged into nine main clusters, including stromal, endothelial, perivascular, immune, epithelial, ciliated cells, macrophages, B cells, and an unknown cluster (Figure 14). Each cluster was defined by genes that were expressed differently compared to the rest of the clusters. The unidentified cluster contained cells lacking expression of marker genes assigning them to any specific cluster.

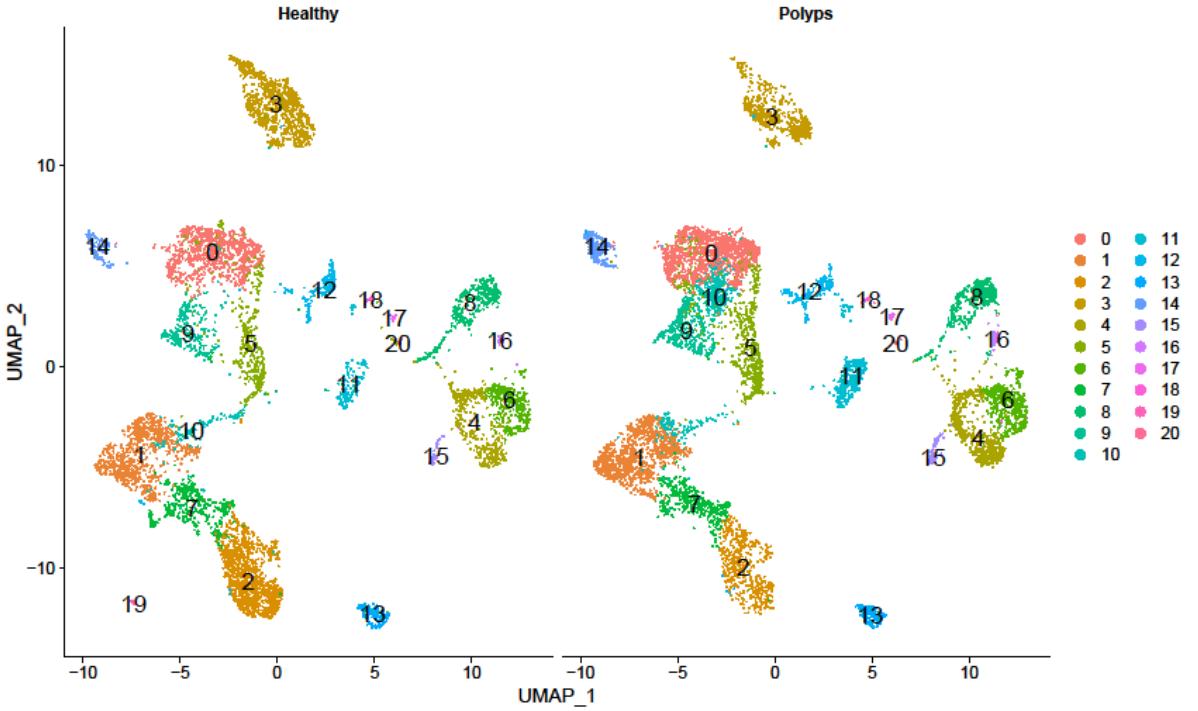

Figure 13. The cells in both the EP and endometrium were categorized and grouped into 20 clusters based on the expression of specific marker genes for each cell type. Healthy indicates endometrium cells and polyps indicates EP. Numbers indicate cluster index number.

To further understand whether cellular distribution between the clusters of EP and endometrium is similar or different, the total number of cells obtained from the biopsy was taken into account and the percentage of cells in each cluster was calculated. Analysis of cluster sizes revealed that the cell clusters of the endometrium and EP do not differ significantly (T-test, $p>0.05$). Thus, it can be argued that the cell clusters of the endometrium and EP have a similar cellular population.

Differential gene expression analysis was done between different clusters and between endometrium and EP. There were no statistically significant DEGs while different clusters were compared to each other in the majority of clusters excluding the perivascular cell cluster. In the perivascular cluster three DEGs: *BNC2*, *CSMD1* and *LINC01060*, were found ($p<0.05$).

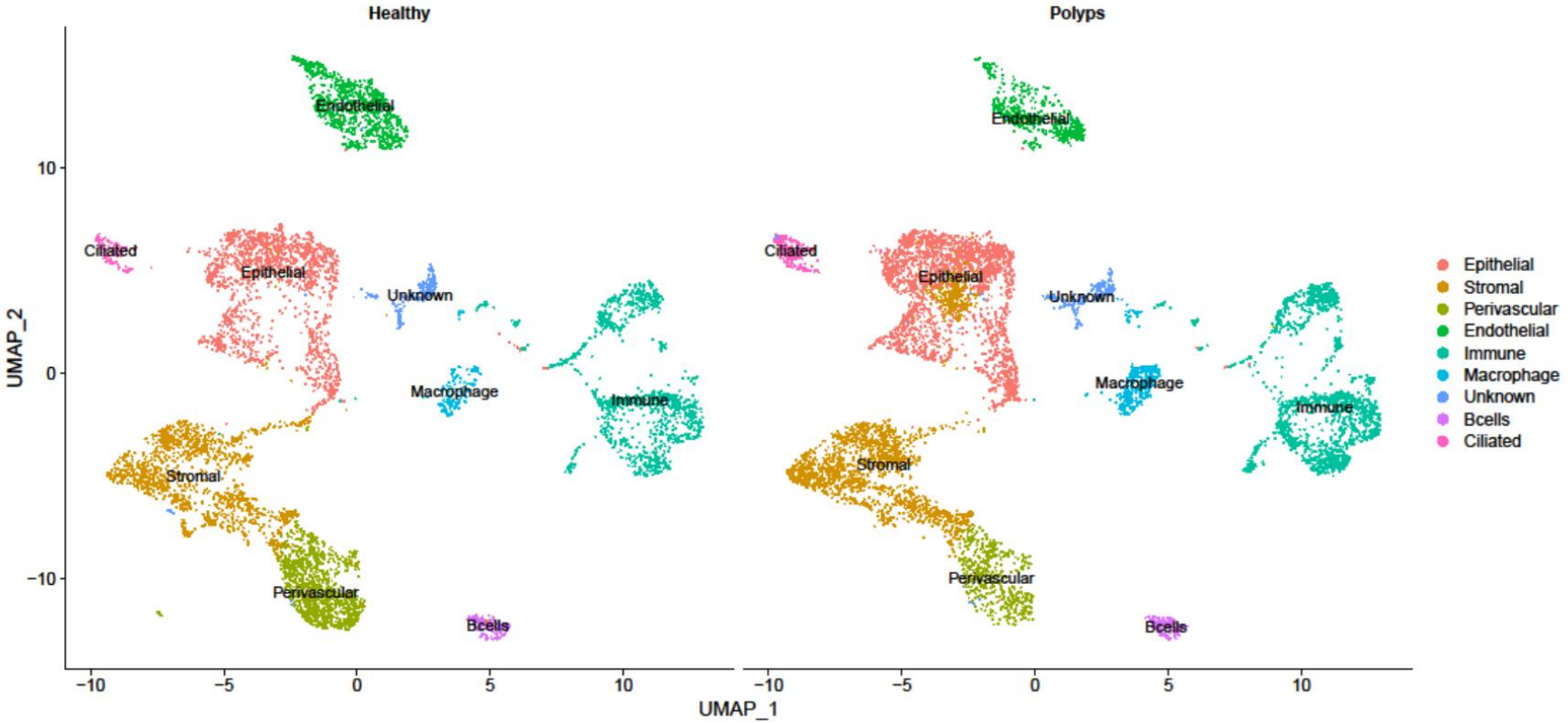

Figure 14. Single-cell distribution between different clusters in endometrium and EP. Comparative analysis of endometrium and EP cells after reclustering. Healthy – endometrium; Polyps – endometrial polyp (EP).

DISCUSSION

An endometrial polyp is a overgrowth of endometrial tissue that occurs in women of reproductive age. The disease itself can be either asymptomatic or cause pain in the uterus and abnormal uterine bleeding, which significantly worsens a woman's quality of life. In addition, EP can cause infertility by preventing either the attachment of the embryo causing inflammation, or the penetration of sperm into the fallopian tube for subsequent fertilization. Despite the fact that EP is a fairly common disease (7.8% to 34.9% in the general population (Lenci et al., 2014; Malik & Zeb, 2022)), the reasons why EP develops and what are the causes of endometrial tissue overgrowth are still poorly understood. There are already a few hints from previous studies that EP pathogenesis could be associated with genomic and molecular alteration of endometrial tissue. The aim of the current study was to find molecular changes in EP that could be associated with endometrial overgrowth and EP development. Therefore, in the first part of this study, the transcriptome signatures of whole EP biopsies and endometrial tissue were compared to identify changes in gene expression that could cause the development of EP. In the second part of this study, scRNA-seq of EP and endometrium was performed to understand the cellular heterogeneity and transcriptome changes inside specific cell clusters.

There were no whole tissue transcriptome studies in EPs at the time this study was designed and performed. Therefore, the study design covered 12 paired endometrial and EP biopsies taken at different phases (proliferative and secretory) of the MC. After transcriptome data analysis we found only a few differentially expressed genes between the EP and endometrium. A closer look to these genes indicated that although they were statistically significantly differentially expressed, the expression levels of all these genes were relatively low. Still, two of these genes (*DLEC1* and *KMT2B*) were decided to select for validation by qRT-PCR.

DLEC1 - deleted in lung and esophageal cancer 1 gene, is a tumor suppressor gene whose reduced expression is observed in various types of human cancers. There is evidence that hypermethylation of the *DLEC1* region may cause decreased expression (Kwong et al., 2006; Navarro et al., 2012; Wang et al., 2012; Ying et al., 2009; Q. Zhang et al., 2010). It was also found that *DLEC1* is expressed in spermatids and spermatozoa. Okitsu et al. disrupted Dlec1 in mice and found that the deletion leads to impaired spermatogenesis and, as a result, infertility (Okitsu et al., 2020).

Another gene, *KMT2B*, is the cause of dystonia (*KMT2B*-related dystonia), which is an autosomal dominant disease that manifests as lower limb dystonia with gait disturbance and abnormal foot position from childhood (Gorman et al., 2018; Lange et al., 2017; Zech et al., 2016, 2017). The *KMT2B* gene itself is responsible for encoding a histone

lysinematyltransferase, which is involved in the methylation of histone H3 and lysine 4 (H3K4), as well as in epigenetic modifications associated with active gene transcription (Shao et al., 2014). In addition to dystonia, *KMT2B* has been shown to promote breast cancer cell proliferation and colon cancer stemness and chemosensitivity (C.-H. Su et al., 2016; Y. Zhang et al., 2023). Zhao et al. also identified the role of *KMT2B* in the development of cervical cancer. Overexpression of *KMT2B* promoted tumor growth and increased vascular density and also promoted cell migration and invasion (Zhao et al., 2023). Based on these data, it could be assumed that the expression of *KMT2B* can lead to tissue proliferation in the cervical canal and contribute to the appearance of a polyp (cervical polyp is one of the types of polyps of the female reproductive tract.

Unfortunately, qRT-PCR analyses did not confirm sequencing results, and no significant difference in the relative expression levels of *DLEC1* and *KMT2B* mRNA was found in endometrial and EP tissues (Figure 12).

The menstrual cycle phase could be one of the factors that significantly changes the transcriptome profile in endometrial tissue. As there is no clear data on whether the EP exhibits a similar molecular profile during the MC, this study included samples from both MC phases. As expected, numerous DEGs (n=232) were identified when comparing proliferative and secretory endometrium samples. Additionally, 88 DEGs (adjusted $p>0.05$) were found when comparing EP samples from different phases of the menstrual cycle, indicating that the transcriptomes of both the endometrium and EP are influenced by the MC.

At the time of this study, a transcriptome study by Chiu et al. was published (Chiu et al., 2024). Although the study design was very similar, the results of their study provide different results. They found 322 DEG genes in EP and these identified genes were involved in the regulation of WNT pathways and muscle contraction. Based on the results, authors found alterations in WNT signaling and vascular smooth muscle pathways were found and authors asumed that these pathways contribute to excessive proliferation and impaired vascular/stromal development. None of the genes that were detected in this study was found in Chiu et al. study and *vice versa*. There might be various reasons why significant molecular changes were not detected in EP tissues in this study. For example, it could be that despite the healthy appearance of the patient's nearby endometrial tissue, there could already be molecular changes and therefore our sequencing failed to detect significant differences in the gene expression level of paired endometrial tissue and EP biopsies. As this study did not include healthy women without EP, future studies should also compare gene expression in the endometrial tissue of a woman without a diagnosis of EP. Also, discrepancies between results of this study and Chiu et al., study could be related to the sequencing platforms used. Chiu et al. used Universal Plus mRNA-

seq with NuQuant (Tecan, Morgan Hill, CA, USA), while in this study TruSeq stranded mRNA platform from Illumina was used. It could be that both studies used different sub-types (flat, sessile, pedunculate) of EP. There is no data describing molecular signatures in different types of polys. As the sub-type was recorded in neither of the studies, this could be one of the reasons for the disconcordance between the outcomes of the study. Also, it is worth mentioning that Chiu et al. did not validate their DEGs, but they selected for validation only highly expressed genes with no statistical significance.

Moreover, some of EP could be malignant (based on the literature 1.3% in women with EP (Wethington et al., 2011)) and it could be that this study included only polyps with no malignant potential.

Chiu et al. studied EPs exclusively in infertile patients, while this study included a mixed population. Molecularly dysregulated pathways associated with infertility might contribute to EP development. However, since our study did not specifically focus on infertility, these genes and pathways may not have been prominently identified.

This study and study by Chiu et al. analysed the EP tissue and adjacent "healthy" endometrium of the same women. To make a final conclusion about the EP molecular signature future studies should also include healthy endometrium in the study design as visually healthy endometrium of women with EP could already have molecular changes that are not present in healthy women.

This study is the first to use scRNA-seq on the 10X Genomics platform to gain deeper insight into EP cellular distribution and heterogeneity. After data analysis, this study found that cellular distribution between the EP and endometrium was largely similar ($p>0.05$) and nine main cell clusters (stromal, endothelial, perivascular, immune, epithelial, ciliated cells, macrophages, B cells, and an unknown cluster) were found. The high similarity between EP and endometrium is also confirmed by the cluster-vise analysis between these two tissue types and only differences were detected in the perivascular cell cluster, in which 3 DEGs (*BNC2, CSMD1, LINC01060*) were identified. The exact role of perivascular cells in EP pathogenesis is currently unclear. In the normal endometrium, perivascular cells, including pericytes, are located near blood vessels, possess the mesenchymal stem cells (MSCs) phenotypes and play a role in regenerating capacities and development of vasculature (S. Li & Ding, 2021). Pericytes are also important in the formation of fibroblast-like cells and in maintaining the extracellular matrix (ECM) (S. Li & Ding, 2021). Any disturbances in the process of angiogenesis or degradation of the ECM would result in aberrant vasculature and abnormal bleeding (S. Li & Ding, 2021). Thus, as EPs are characterized by the thickening of blood vessel walls, the differential genes in perivascular cell cluster could provide some hints about endometrial tissue overgrowth, dysregulated vascularisation and EP formation. Amongst the DEGs found in perivascular cells,

none of them has been associated with EP. *BNC2* is also reported to be associated with liver fibrogenesis (Bobowski-Gerard et al., 2022). Although the expression of *BNC2* has also been observed in the uterus, it can have a role in fibrinogenesis related to EP. The *CSMD1* is a tumour suppressor in cancer and has many functions, such as development, complement control, neurodevelopmental disorders and cancer progression, but the specific mechanisms in which this gene is involved have not yet been identified (Blom, 2017; Burgess et al., 2021; Donohoe et al., 2013; Escudero-Esparza et al., 2013; Kamal et al., 2017; W.-H. Lee et al., 2019; Ma et al., 2009). However, its role in perivascular cluster in relation to EP is still unknown. *LINC01060* gene is associated with EMT-related lncRNA and is highly expressed in osteosarcoma (Y. Zhang et al., 2023). *In vitro* and *in vivo* knockdown of *LINC01060* resulted in the suppression of hyperproliferation and migration, thus inhibiting tumor growth and metastasis (Y. Zhang et al., 2023). Furthermore, the *LINC01060* lncRNA was found to be involved in the Erk5 MAPK signalling pathway, which has been previously reported in the regulation of vascular endothelial growth factor expression for embryonic angiogenesis (L. Li et al., 2020; Sohn et al., 2002). Since this lncRNA shares a common pathway for the regulation of angiogenesis as well as cancer processes, its differential expression in the perivascular cluster highlights the mechanism for polyp development.

Taken together, this study found that transcriptomes of EP and endometrium were highly similar and only few DEG were found. Further studies in EPs, focusing on the lncRNA and even miRNA for the regulation of gene expression, would offer valuable insights. Thus, based on the overall results of scRNA-seq, we can infer that although the transcriptome profile and cellular populations of EPs and endometrial tissues are similar, three DEGs in the perivascular cluster further elucidated the potential mechanisms underlying EP development. Therefore, investigating the putative role of perivascular cells in endometrial overgrowth and the pathogenesis of EPs warrants future large-scale studies, including analysis of specific subtypes of EP.

SUMMARY

Endometrial polyps (EP) are commonly occurring benign overgrowth of endometrial tissue in women. It is a common condition among premenopausal and menopausal women, but it can also occur in women of reproductive age. A number of reasons contribute to its appearance, mainly age, hormonal imbalance, chronic diseases, inflammation, medications, increased BMI, and genetics. Despite all the information available about EP, the specific molecular and genetic reasons behind the occurrence of EP are still not known.

The primary objective of this study was to identify potential changes in gene expression by comparing the transcriptome of endometrial tissue and EP biopsies. scRNA-seq of EP and endometrium was also performed to identify specific cell populations and assess EP cellular heterogeneity and gene expression levels at the single-cell level.

In this whole-tissue transcriptome work of the study, endometrial and EP biopsies from 12 women taken during different phases of the MC (proliferative and secretory) were analyzed. After analyzing the transcriptome data, we found that the gene expression profiles of the endometrium and EP were very similar and identified only a few genes that were statistically significantly differentially expressed. A closer look at these genes revealed that although they were expressed statistically significantly differently, the expression levels of all of these genes were relatively low. However, it was decided to select two of these genes (*DLEC1* and *KMT2B*) for testing by qRT-PCR. Unfortunately, qRT-PCR analyzes did not confirm the sequencing results and did not show a significant difference.

Although the study of the EP and endometrial transcriptome did not show the expected results, scRNA-seq using the 10X Genomics platform identified three DEGs in the perivascular cluster: *BNC2*, *CSMD1* and *LINC01060* ($p<0.05$). This is the first study of EP using scRNA-seq to date that has shown results. The obtained genes are subject to further study to determine their possible involvement in developing EP.

RESÜMEE

Endomeetriumi polüübi geenieksпressiooni profileerimine

Darja Tarassova resümee

Endomeetriumi polüübid (EP) on naistel sageli esinevad healoomulised endomeetriumi koe vohandid, mis esinevad nii reproduktiivses eas naistel kui ka enne ja peale menopausi. EP tekkimisele aitavad kaasa mitmed põhjused, peamiselt vanus, hormonaalsed häired, kroonilised haigused, põletikud, ravimid, suurenenud kehamassiindeks ja geneetika. Vaatamata sellele, et EP-d on väga sagedased, ei teata siiani, et miks nad tekivad ja millised on molekulaarsed ja geneetilised põhjused, mis tingivad endomeetriumi koe vohamise.

Selle uuringu peamine eesmärk oli tuvastada võimalikke muutusi geenieksпressioonis, võrreldes endomeetriumi koe ja EP biopsiate transkriptoomi. Viidi läbi ka EP ja endomeetriumi ühe-raku RNA sekveneerimine, et tuvastada spetsiifilisi rakupopulatsioone ja hinnata EP rakulist heterogeensust ja geenieksпressiooni tasemeid üksikute rakkude tasandil.

Selles uuringu täiskoe transkriptoomi töös analüüsiti 12 naise endomeetriumit ja EP biopsiat, mis võeti menstruaaltsükli erinevates faasides (proliferatiivne ja sekretoorne). Pärast transkriptsiooniandmete analüüsimist leiti, et endomeetriumi ja EP geenieksпressiooni profiilid on väga sarnased ja tuvastati ainult üksikud geenid, mis olid statistiliselt oluliselt erinevalt ekspresseeritud. Neid geene lähemalt uurides selgus, et kuigi need ekspresseerusid statistiliselt oluliselt erinevalt, olid kõigi nende geenide ekspressioonitasemed suhteliselt madalad. Siiski otsustati valida kaks neist geenidest (*DLEC1* ja *KMT2B*) testimiseks qRT-PCR abil. Kahjuks ei kinnitanud qRT-PCR analüüsid sekveneerimistulemusi ega näidanud olulist erinevust.

Kuna EP ja endomeetriumikoe transkriptoomi uurimine ei andnud oodatud tulemusi viidi läbi 10X Genomics platvormi kasutav üksikraku RNA sekveneerimine, kasutades kolme endomeetriumit ja viite EP. Sekveneerimistulemused näitasid, et EP ja endomeetriumi rakuline mitmekesisus on väga sarnane, kuid EP-des tuvastati kolm geeni (*BNC2*, *CSMD1* ja *LINC01060*), mis olid erinevalt ekspresseeritud perivaskulaarses rakuklastris.

Kokkuvõtvalt, käesoleva töö tulemusena leiti, et EP ja endomeetriumi molekulaarsed profiilid on väga sarnased ja võib arvata, et muutused koe transkriptoomis ei mõjuta EP kujunemist. Perivaskulaarsete rakkude roll EP kujunemises ja leitud geenid vajavad täiendavaid uuringuid.

REFERENCES

AAGL Practice Report: Practice Guidelines for the Diagnosis and Management of Endometrial Polyps. (2012). *Journal of Minimally Invasive Gynecology*, *19*(1), 3–10. https://doi.org/10.1016/j.jmig.2011.09.003

Alansari, L. M., & Wardle, P. (2012). Endometrial polyps and subfertility. *Human Fertility (Cambridge, England)*, *15*(3), 129–133. https://doi.org/10.3109/14647273.2012.711499

Amezcua, C. A., Zheng, W., Muderspach, L. I., & Felix, J. C. (1999). Down-Regulation of Bcl-2 Is a Potential Marker of the Efficacy of Progestin Therapy in the Treatment of Endometrial Hyperplasia. *Gynecologic Oncology*, *73*(1), 126–136. https://doi.org/10.1006/gyno.1998.5336

Arici, A., Seli, E., Zeyneloglu, H. B., Senturk, L. M., Oral, E., & Olive, D. L. (1998). Interleukin-8 Induces Proliferation of Endometrial Stromal Cells: A Potential Autocrine Growth Factor1. *The Journal of Clinical Endocrinology & Metabolism*, *83*(4), 1201–1205. https://doi.org/10.1210/jcem.83.4.4743

Ba, L., Zhang, N., Meng, J., Zhang, J., Lin, P., Zhou, P., Liu, S., & Bachert, C. (2011). The association between bacterial colonization and inflammatory pattern in Chinese chronic rhinosinusitis patients with nasal polyps. *Allergy*, *66*(10), 1296–1303. https://doi.org/10.1111/j.1398-9995.2011.02637.x

Bergholz, J., & Xiao, Z.-X. (2012). Role of p63 in Development, Tumorigenesis and Cancer Progression. *Cancer Microenvironment*, *5*(3), 311–322. https://doi.org/10.1007/s12307-012-0116-9

Bettocchi, S., Achilarre, M. T., Ceci, O., & Luigi, S. (2011). Fertility-enhancing hysteroscopic surgery. *Seminars in Reproductive Medicine*, *29*(2), 75–82. https://doi.org/10.1055/s-0031-1272469

Blom, A. M. (2017). The role of complement inhibitors beyond controlling inflammation. *Journal of Internal Medicine*, *282*(2), 116–128. https://doi.org/10.1111/joim.12606

Bobowski-Gerard, M., Boulet, C., Zummo, F. P., Dubois-Chevalier, J., Gheeraert, C., Bou Saleh, M., Strub, J.-M., Farce, A., Ploton, M., Guille, L., Vandel, J., Bongiovanni, A., Very, N., Woitrain, E., Deprince, A., Lalloyer, F., Bauge, E., Ferri, L., Ntandja-Wandji, L.-C., … Eeckhoute, J. (2022). Functional genomics uncovers the transcription factor BNC2 as required for myofibroblastic activation in fibrosis. *Nature Communications*, *13*(1), 5324. https://doi.org/10.1038/s41467-022-33063-9

Bolomsky, A., Vogler, M., Köse, M. C., Heckman, C. A., Ehx, G., Ludwig, H., & Caers, J. (2020). MCL-1 inhibitors, fast-lane development of a new class of anti-cancer agents. *Journal of Hematology & Oncology*, *13*(1), 173. https://doi.org/10.1186/s13045-020-01007-9

Burgess, S. J., Gibbs, H., Toomes, C., Coletta, P. L., & Bell, S. M. (2021). The Role of Csmd1 during Mammary Gland Development. *Genes*, *12*(2), 162. https://doi.org/10.3390/genes12020162

Canfield, J., Arlier, S., Mong, E. F., Lockhart, J., VanWye, J., Guzeloglu-Kayisli, O., Schatz, F., Magness, R. R., Lockwood, C. J., Tsibris, J. C. M., Kayisli, U. A., & Totary-Jain, H. (2019). Decreased LIN28B in preeclampsia impairs human trophoblast differentiation and migration. *The FASEB Journal*, *33*(2), 2759–2769. https://doi.org/10.1096/fj.201801163R

Carugno, J., Di Spiezio Sardo, A., Alonso Pacheco, L., Haimovich, S., Campo, R. L., Angelis, C., Bradley, L., Bettocchi, S., Arias, A., Isaacson, K., Okohue, J., Farrugia, M., Kumar, A., Xue, X., Cavalcanti, L., Laganà, A. S., & Grimbizis, G. (2020). COVID-19 Pandemic. Impact on Hysteroscopic Procedures: A Consensus Statement from the Global Congress of Hysteroscopy Scientific Committee. *Journal of Minimally Invasive Gynecology*, *27*. https://doi.org/10.1016/j.jmig.2020.04.023

Cea García, J., Jiménez Caraballo, A., Ríos Vallejo, M. del M., & Zapardiel, I. (2020). Retrospective Cohort Study on the Symptomatic Recurrence Pattern after Hysteroscopic Polypectomy. *Gynecology and Minimally Invasive Therapy*, *9*(4), 209–214. https://doi.org/10.4103/GMIT.GMIT_102_19

Ceci, O., Franchini, M., Cardinale, S., Cannone, R., Giarrè, G., Fascilla, F. D., Bettocchi, S., & Cicinelli, E. (2020). Comparison of endometrial polyp recurrence in fertile women after office hysteroscopic endometrial polypectomy using two widely spread techniques. *The Journal of Obstetrics and Gynaecology Research*, *46*(10), 2084–2091. https://doi.org/10.1111/jog.14400

Check, J. H., Bostick-Smith, C. A., Choe, J. K., Amui, J., & Brasile, D. (2011). Matched controlled study to evaluate the effect of endometrial polyps on pregnancy and implantation rates following in vitro fertilization-embryo transfer (IVF-ET). *Clinical and Experimental Obstetrics & Gynecology*, *38*(3), 206–208.

Chen, G., Ning, B., & Shi, T. (2019). Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Frontiers in Genetics*, *10*, 317. https://doi.org/10.3389/fgene.2019.00317

Chiu, C. S.-C., Yeh, L.-Y., Pan, S.-H., & Li, S.-H. (2024). Transcriptomic Analysis Reveals Intrinsic Abnormalities in Endometrial Polyps. *International Journal of Molecular Sciences*, *25*(5), Article 5. https://doi.org/10.3390/ijms25052557

Cicinelli, E., Vitagliano, A., Loizzi, V., De Ziegler, D., Fanelli, M., Bettocchi, S., Nardelli, C., Trojano, G., Cicinelli, R., Minervini, C. F., Leronni, D., & Viggiano, L. (2021). Altered Gene Expression Encoding Cytochines, Grow Factors and Cell Cycle Regulators in the Endometrium of Women with Chronic Endometritis. *Diagnostics (Basel, Switzerland)*, *11*(3), 471. https://doi.org/10.3390/diagnostics11030471

Ciscato, A., Zare, S. Y., & Fadare, O. (2020). The significance of recurrence in endometrial polyps: A clinicopathologic analysis. *Human Pathology*, *100*, 38–44. https://doi.org/10.1016/j.humpath.2020.03.005

Di Spiezio Sardo, A., Calagna, G., Guida, M., Perino, A., & Nappi, C. (2015). Hysteroscopy and treatment of uterine polyps. *Best Practice & Research. Clinical Obstetrics & Gynaecology*, *29*(7), 908–919. https://doi.org/10.1016/j.bpobgyn.2015.06.005

Donohoe, G., Walters, J., Hargreaves, A., Rose, E. J., Morris, D. W., Fahey, C., Bellini, S., Cummins, E., Giegling, I., Hartmann, A. M., Möller, H.-J., Muglia, P., Owen, M. J., Gill, M., O'Donovan, M. C., Tropea, D., Rujescu, D., & Corvin, A. (2013). Neuropsychological effects of the CSMD1 genome-wide associated schizophrenia risk variant rs10503253. *Genes, Brain, and Behavior*, *12*(2), 203–209. https://doi.org/10.1111/gbb.12016

Dreisler, E., Stampe Sorensen, S., Ibsen, P. H., & Lose, G. (2009). Prevalence of endometrial polyps and abnormal uterine bleeding in a Danish population aged 20-74 years. *Ultrasound in Obstetrics & Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, *33*(1), 102–108. https://doi.org/10.1002/uog.6259

Escudero-Esparza, A., Kalchishkova, N., Kurbasic, E., Jiang, W. G., & Blom, A. M. (2013). The novel complement inhibitor human CUB and Sushi multiple domains 1 (CSMD1) protein promotes factor I-mediated degradation of C4b and C3b and inhibits the membrane attack complex assembly. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *27*(12), 5083–5093. https://doi.org/10.1096/fj.13-230706

Farzaneh, M., Attari, F., & Khoshnam, S. E. (2017). Concise Review: LIN28/let-7 Signaling, a Critical Double-Negative Feedback Loop During Pluripotency, Reprogramming, and Tumorigenicity. *Cellular Reprogramming*, *19*(5), 289–293. https://doi.org/10.1089/cell.2017.0015

Finlay, C. A., & Levine, J. (n.d.). *The ~53 Proto-Oncogene Can Act as a Suppressor of Transformation*.

Gadd, S., Huff, V., Walz, A. L., Ooms, A. H. A. G., Armstrong, A. E., Gerhard, D. S., Smith, M. A., Guidry Auvil, J. M., Meerzaman, D., Chen, Q.-R., Hsu, C. H., Yan, C., Nguyen, C., Hu, Y., Hermida, L. C., Davidsen, T., Gesuwan, P., Ma, Y., Zong, Z., … Perlman, E. J. (2017). A Children's Oncology Group and TARGET Initiative Exploring the Genetic Landscape of Wilms Tumor. *Nature Genetics*, *49*(10), 1487–1494. https://doi.org/10.1038/ng.3940

Garuti, G., Centinaio, G., & Luerti, M. (2008). Outpatient hysteroscopic polypectomy in postmenopausal women: A comparison between mechanical and electrosurgical resection. *Journal of Minimally Invasive Gynecology*, *15*(5), 595–600. https://doi.org/10.1016/j.jmig.2008.07.001

Gerdes, J., Lemke, H., Baisch, H., Wacker, H. H., Schwab, U., & Stein, H. (1984). Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *Journal of Immunology (Baltimore, Md.: 1950)*, *133*(4), 1710–1715.

Gerdes, J., Li, L., Schlueter, C., Duchrow, M., Wohlenberg, C., Gerlach, C., Stahmer, I., Kloth, S., Brandt, E., & Flad, H. D. (1991). Immunobiochemical and molecular biologic characterization of the cell proliferation-associated nuclear antigen that is defined by monoclonal antibody Ki-67. *The American Journal of Pathology*, *138*(4), 867–873.

Giordano, G., Gnetti, L., Merisio, C., & Melpignano, M. (2007). Postmenopausal status, hypertension and obesity as risk factors for malignant transformation in endometrial polyps. *Maturitas*, *56*(2), 190–197. https://doi.org/10.1016/j.maturitas.2006.08.002

Giordano, M. V., Lucas, H. D. S., Fiorelli, R. K. A., Giordano, L. A., Giordano, M. G., Baracat, E. C., & Júnior, J. M. S. (2020). Expression levels of BCL2 and MKI67 in endometrial polyps in postmenopausal women and their correlation with obesity. *Molecular and Clinical Oncology*, *13*(6), 69. https://doi.org/10.3892/mco.2020.2139

Gompel, A., Sabourin, J. C., Martin, A., Yaneva, H., Audouin, J., Decroix, Y., & Poitout, P. (1994). Bcl-2 expression in normal endometrium during the menstrual cycle. *The American Journal of Pathology*, *144*(6), 1195–1202.

Gorman, K. M., Meyer, E., & Kurian, M. A. (2018). Review of the phenotype of early-onset generalised progressive dystonia due to mutations in KMT2B. *European Journal of Paediatric Neurology: EJPN: Official Journal of the European Paediatric Neurology Society*, *22*(2), 245–256. https://doi.org/10.1016/j.ejpn.2017.11.009

Gu, F., Zhang, H., Ruan, S., Li, J., Liu, X., Xu, Y., & Zhou, C. (2018). High number of endometrial polyps is a strong predictor of recurrence: Findings of a prospective cohort study in reproductive-age women. *Fertility and Sterility*, *109*(3), 493–500. https://doi.org/10.1016/j.fertnstert.2017.11.029

Guo, H., Liao, Y., Lin, A., Lin, H., Huang, X., Ruan, J., Yuan, L., & Zhuo, Z. (2021). Association Between LIN28A Gene Polymorphisms and Glioma Susceptibility in Chinese Children. *Cancer Control : Journal of the Moffitt Cancer Center*, *28*, 10732748211040009. https://doi.org/10.1177/10732748211040009

Hileeto, D., Fadare, O., Martel, M., & Zheng, W. (2005). Age dependent association of endometrial polyps with increased risk of cancer involvement. *World Journal of Surgical Oncology*, *3*(1), 8. https://doi.org/10.1186/1477-7819-3-8

Huang, J., Cao, D., Sha, J., Zhu, X., & Han, S. (2019). DLL3 is regulated by LIN28B and miR-518d-5p and regulates cell proliferation, migration and chemotherapy response in advanced small cell lung cancer. *Biochemical and Biophysical Research Communications*, *514*(3), 853–860. https://doi.org/10.1016/j.bbrc.2019.04.130

Iwabe, T., Harada, T., Tsudo, T., Nagano, Y., Yoshida, S., Tanikawa, M., & Terakawa, N. (2000). Tumor Necrosis Factor-α Promotes Proliferation of Endometriotic Stromal Cells by Inducing Interleukin-8 Gene and Protein Expression. *The Journal of Clinical Endocrinology & Metabolism*, *85*(2), 824–829. https://doi.org/10.1210/jcem.85.2.6335

Iwabe, T., Harada, T., Tsudo, T., Tanikawa, M., Onohara, Y., & Terakawa, N. (1998). Pathogenetic Significance of Increased Levels of Interleukin-8 in the Peritoneal Fluid of Patients with Endometriosis. *Fertility and Sterility*, *69*(5), 924–930. https://doi.org/10.1016/S0015-0282(98)00049-1

Jones, M. E., van Leeuwen, F. E., Hoogendoorn, W. E., Mourits, M. J., Hollema, H., van Boven, H., Press, M. F., Bernstein, L., & Swerdlow, A. J. (2012). Endometrial cancer survival after breast cancer in relation to tamoxifen treatment: Pooled results from three countries. *Breast Cancer Research : BCR*, *14*(3), R91. https://doi.org/10.1186/bcr3206

Jones, V. A., Patel, P. M., Gibson, F. T., Cordova, A., & Amber, K. T. (2020). The Role of Collagen XVII in Cancer: Squamous Cell Carcinoma and Beyond. *Frontiers in Oncology*, *10*, 352. https://doi.org/10.3389/fonc.2020.00352

Kamal, M., Holliday, D. L., Morrison, E. E., Speirs, V., Toomes, C., & Bell, S. M. (2017). Loss of CSMD1 expression disrupts mammary duct formation while enhancing proliferation,

migration and invasion. *Oncology Reports*, *38*(1), 283–292. https://doi.org/10.3892/or.2017.5656

Kiewa, J., Mortlock, S., Meltzer-Brody, S., Middeldorp, C., Wray, N. R., & Byrne, E. M. (2023). A Common Genetic Factor Underlies Genetic Risk for Gynaecological and Reproductive Disorders and Is Correlated with Risk to Depression. *Neuroendocrinology*, *113*(10), 1059–1075. https://doi.org/10.1159/000533413

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., & Kirschner, M. W. (2015). Droplet barcoding for single cell transcriptomics applied to embryonic stem cells. *Cell*, *161*(5), 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044

Kolhe, S. (2018). Management of abnormal uterine bleeding—Focus on ambulatory hysteroscopy. *International Journal of Women's Health*, *10*, 127–136. https://doi.org/10.2147/IJWH.S98579

Kossaï, M., & Penault-Llorca, F. (2020). Role of Hormones in Common Benign Uterine Lesions: Endometrial Polyps, Leiomyomas, and Adenomyosis. In L. Deligdisch-Schor & A. Mareş Miceli (Eds.), *Hormonal Pathology of the Uterus* (Vol. 1242, pp. 37–58). Springer International Publishing. https://doi.org/10.1007/978-3-030-38474-6_3

Kwong, J., Lee, J.-Y., Wong, K.-K., Zhou, X., Wong, D. T. W., Lo, K.-W., Welch, W. R., Berkowitz, R. S., & Mok, S. C. (2006). Candidate Tumor-Suppressor Gene *DLEC1* Is Frequently Downregulated by Promoter Hypermethylation and Histone Hypoacetylation in Human Epithelial Ovarian Cancer. *Neoplasia*, *8*(4), 268–278. https://doi.org/10.1593/neo.05502

Lange, L. M., Tunc, S., Tennstedt, S., Münchau, A., Klein, C., Assmann, B., & Lohmann, K. (2017). A novel, in-frame KMT2B deletion in a patient with apparently isolated, generalized dystonia. *Movement Disorders: Official Journal of the Movement Disorder Society*, *32*(10), 1495–1497. https://doi.org/10.1002/mds.27137

Lee, S. C., Kaunitz, A. M., Sanchez-Ramos, L., & Rhatigan, R. M. (2010). The oncogenic potential of endometrial polyps: A systematic review and meta-analysis. *Obstetrics and Gynecology*, *116*(5), 1197–1205. https://doi.org/10.1097/AOG.0b013e3181f74864

Lee, W.-H., Chen, L.-C., Lee, C.-J., Huang, C.-C., Ho, Y.-S., Yang, P.-S., Ho, C.-T., Chang, H.-L., Lin, I.-H., Chang, H.-W., Liu, Y.-R., Wu, C.-H., & Tu, S.-H. (2019). DNA primase polypeptide 1 (PRIM1) involves in estrogen-induced breast cancer formation through activation

of the G2/M cell cycle checkpoint. *International Journal of Cancer*, *144*(3), 615–630. https://doi.org/10.1002/ijc.31788

Lenci, M. A., do Nascimento, V. A. L., Grandini, A. B., Fahmy, W. M., Depes, D. de B., Baracat, F. F., & Lopes, R. G. C. (2014). Premalignant and malignant lesions in endometrial polyps in patients undergoing hysteroscopic polypectomy. *Einstein*, *12*(1), 16–21. https://doi.org/10.1590/S1679-45082014AO2764

Li, H., Liu, D., Liu, L., Huang, S., Ma, A., & Zhang, X. (2019). The role of HOTAIR/miR-152-3p/LIN28B in regulating the progression of endometrial squamous carcinoma. *Archives of Medical Science : AMS*, *17*(2), 434–448. https://doi.org/10.5114/aoms.2019.89632

Li, S., & Ding, L. (2021). Endometrial Perivascular Progenitor Cells and Uterus Regeneration. *Journal of Personalized Medicine*, *11*(6), Article 6. https://doi.org/10.3390/jpm11060477

Lin, S.-L., Duan, H., Wang, S., & Li, J.-J. (2020). Overexpression of Lin28B Promoted the Proliferation of Adenomyotic Smooth Muscle Cells of the Junctional Zone via Regulating Let-7a. *Reproductive Sciences*, *27*(5), 1156–1163. https://doi.org/10.1007/s43032-019-00107-3

Liu, Y., Zhang, Y., Fu, J., & Tan, W. (2012). Inflammation-Related Gene Expression Profiles of Endocervical Polyps. *Journal of Interferon & Cytokine Research*, *32*(5), 191–197. https://doi.org/10.1089/jir.2011.0066

Liu, Z., Kuokkanen, S., & Pal, L. (2010). Steroid Hormone Receptor Profile of Premenopausal Endometrial Polyps. *Reproductive Sciences*, *17*(4), 377–383. https://doi.org/10.1177/1933719109356803

Louault, K., Bonneaud, T. L., Séveno, C., Gomez-Bougie, P., Nguyen, F., Gautier, F., Bourgeois, N., Loussouarn, D., Kerdraon, O., Barillé-Nion, S., Jézéquel, P., Campone, M., Amiot, M., Juin, P. P., & Souazé, F. (2019). Interactions between cancer-associated fibroblasts and tumor cells promote MCL-1 dependency in estrogen receptor-positive breast cancers. *Oncogene*, *38*(17), 3261–3273. https://doi.org/10.1038/s41388-018-0635-z

Lovnicki, J., Gan, Y., Feng, T., Li, Y., Xie, N., Ho, C.-H., Lee, A. R., Chen, X., Nappi, L., Han, B., Fazli, L., Huang, J., Gleave, M. E., & Dong, X. (2020). LIN28B promotes the development of neuroendocrine prostate cancer. *The Journal of Clinical Investigation*, *130*(10), 5338–5348. https://doi.org/10.1172/JCI135373

Lu, L., Luo, J., Deng, J., Huang, C., & Li, C. (2023). Polycystic ovary syndrome is associated with a higher risk of premalignant and malignant endometrial polyps in premenopausal women:

A retrospective study in a tertiary teaching hospital. *BMC Women's Health*, *23*(1), 127. https://doi.org/10.1186/s12905-023-02269-4

Lu, M.-Y., Li, X.-H., Niu, J.-L., & Liu, B. (2022). LIN28B Polymorphisms Confer a Higher Postoperative Recurrence Risk in Reproductive-Age Women with Endometrial Polyps. *Disease Markers*, *2022*, 4824357. https://doi.org/10.1155/2022/4824357

Ludwig, H., & Spornitz, U. M. (1991). Microarchitecture of the human endometrium by scanning electron microscopy: Menstrual desquamation and remodeling. *Annals of the New York Academy of Sciences*, *622*, 28–46. https://doi.org/10.1111/j.1749-6632.1991.tb37848.x

Ma, C., Quesnelle, K. M., Sparano, A., Rao, S., Park, M. S., Cohen, M. A., Wang, Y., Samanta, M., Kumar, M. S., Aziz, M. U., Naylor, T. L., Weber, B. L., Fakharzadeh, S. S., Weinstein, G. S., Vachani, A., Feldman, M. D., & Brose, M. S. (2009). Characterization CSMD1 in a large set of primary lung, head and neck, breast and skin cancer tissues. *Cancer Biology & Therapy*, *8*(10), 907–916. https://doi.org/10.4161/cbt.8.10.8132

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., & McCarroll, S. A. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, *161*(5), 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002

Maia, H., Casoy, J., & Valente Filho, J. (2009). Is aromatase expression in the endometrium the cause of endometriosis and related infertility? *Gynecological Endocrinology*, *25*(4), 253–257. https://doi.org/10.1080/09513590802627647

Malik, A., & Zeb, L. (2022). To Determine the Frequency of Endometrial Polyps in Patients Presenting with Irregular per Vaginal Bleeding. *Fortune Journal of Health Sciences*, *5*(1), 198–204.

Mansour, T., & Chowdhury, Y. S. (2022). Endometrial Polyp. In *StatPearls [Internet]*. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK557824/

Mazur, M. T., & Kurman, R. J. (2005). *Diagnosis of endometrial biopsies and curettings: A practical approach* (2. ed). Springer.

McGinnis, C. S., Murrow, L. M., & Gartner, Z. J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Systems*, *8*(4), 329-337.e4. https://doi.org/10.1016/j.cels.2019.03.003

McGurgan, P., Taylor, L. J., Duffy, S. R., & O'Donovan, P. J. (2006). Does tamoxifen therapy affect the hormone receptor expression and cell proliferation indices of endometrial polyps? An immunohistochemical comparison of endometrial polyps from postmenopausal women exposed and not exposed to tamoxifen. *Maturitas*, *54*(3), 252–259. https://doi.org/10.1016/j.maturitas.2005.11.007

Molet, S. M., Hamid, Q. A., & Hamilos, D. L. (2003). IL-11 and IL-17 expression in nasal polyps: Relationship to collagen deposition and suppression by intranasal fluticasone propionate. *The Laryngoscope*, *113*(10), 1803–1812. https://doi.org/10.1097/00005537-200310000-00027

Mulvany, M. J., & Aalkjaer, C. (1990). Structure and function of small arteries. *Physiological Reviews*, *70*(4), 921–961. https://doi.org/10.1152/physrev.1990.70.4.921

Nappi, L., Indraccolo, U., Sardo, A. D. S., Gentile, G., Palombino, K., Castaldi, M. A., Spinelli, M., & Greco, P. (2009). Are Diabetes, Hypertension, and Obesity Independent Risk Factors for Endometrial Polyps? *Journal of Minimally Invasive Gynecology*, *16*(2), 157–162. https://doi.org/10.1016/j.jmig.2008.11.004

Navarro, A., Yin, P., Monsivais, D., Lin, S. M., Du, P., Wei, J.-J., & Bulun, S. E. (2012). Genome-Wide DNA Methylation Indicates Silencing of Tumor Suppressor Genes in Uterine Leiomyoma. *PLOS ONE*, *7*(3), e33284. https://doi.org/10.1371/journal.pone.0033284

Neto, L. C. de A., Soares, J. M., Giusa-Chiferi, M. G., Gonçalves, W. J., & Baracat, E. C. (2013). Expression of p53 protein in the endometrial polyp in postmenopausal women. *European Journal of Gynaecological Oncology*, *34*(6), 509–512.

Nijkang, N. P., Anderson, L., Markham, R., & Manconi, F. (2019). Endometrial polyps: Pathogenesis, sequelae and treatment. *SAGE Open Medicine*, *7*, 205031211984824. https://doi.org/10.1177/2050312119848247

Nogueira, A. A., Sant'Ana de Almeida, E. C., Poli Neto, O. B., Zambelli Ramalho, L. N., Rosa e Silva, J. C., & Candido dos Reis, F. J. (2006). Immunohistochemical expression of p63 in endometrial polyps: Evidence that a basal cell immunophenotype is maintained. *Menopause (New York, N.Y.)*, *13*(5), 826–830. https://doi.org/10.1097/01.gme.0000242274.32278.a2

Okitsu, Y., Nagano, M., Yamagata, T., Ito, C., Toshimori, K., Dohra, H., Fujii, W., & Yogo, K. (2020). Dlec1 is required for spermatogenesis and male fertility in mice. *Scientific Reports*, *10*(1), 18883. https://doi.org/10.1038/s41598-020-75957-y

Pace, S., Grassi, A., Franceschini, P., Villani, C., Frega, A., & Lotti, G. (1992). [Diagnostic and therapeutic aspects of endometrial polyps]. *Minerva Ginecologica*, *44*(5), 227–231.

Pathare, A., Pujol-Gualdo, N., Rukins, V., Džigurski, J., Peters, M., Team, E., Mägi, R., Salumets, A., Saare, M., & Laisk, T. (2024). *Large-scale genome-wide association study to determine the genetic underpinnings of female genital tract polyps*. https://doi.org/10.1101/2024.01.29.24301773

Pinheiro, A., Antunes, A., Andrade, L., De Brot, L., Pinto-Neto, A. M., & Costa-Paiva, L. (2014). Expression of hormone receptors, Bcl-2, Cox-2 and Ki67 in benign endometrial polyps and their association with obesity. *Molecular Medicine Reports*, *9*(6), 2335–2341. https://doi.org/10.3892/mmr.2014.2125

Powles, T. J., Ashley, S., Tidy, A., Smith, I. E., & Dowsett, M. (2007). Twenty-year follow-up of the Royal Marsden randomized, double-blinded tamoxifen breast cancer prevention trial. *Journal of the National Cancer Institute*, *99*(4), 283–290. https://doi.org/10.1093/jnci/djk050

Qian, S., Wei, Z., Yang, W., Huang, J., Yang, Y., & Wang, J. (2022). The role of BCL-2 family proteins in regulating apoptosis and cancer therapy. *Frontiers in Oncology*, *12*, 985363. https://doi.org/10.3389/fonc.2022.985363

Reslová, T., Tosner, J., Resl, M., Kugler, R., & Vávrová, I. (1999). Endometrial polyps. A clinical study of 245 cases. *Archives of Gynecology and Obstetrics*, *262*(3–4), 133–139. https://doi.org/10.1007/s004040050241

Rshoud, F. Al., Omari, B. A., Qudsi, A., & Salhiyeh, A. A. (2022). Polycystic ovarian syndrome – association and risk factors between endometrial polyp and infertility. A retrospective study. *Przegląd Menopauzalny = Menopause Review*, *21*(2), 106–110. https://doi.org/10.5114/pm.2022.116478

Saitoh, T., Kusunoki, T., Yao, T., Kawano, K., Kojima, Y., Miyahara, K., Onoda, J., Yokoi, H., & Ikeda, K. (2010). Role of interleukin-17A in the eosinophil accumulation and mucosal remodeling in chronic rhinosinusitis with nasal polyps associated with asthma. *International Archives of Allergy and Immunology*, *151*(1), 8–16. https://doi.org/10.1159/000232566

Saliba, A.-E., Westermann, A. J., Gorski, S. A., & Vogel, J. (2014). Single-cell RNA-seq: Advances and future challenges. *Nucleic Acids Research*, *42*(14), 8845–8860. https://doi.org/10.1093/nar/gku555

Salim, S., Won, H., Nesbitt-Hawes, E., Campbell, N., & Abbott, J. (2011). Diagnosis and management of endometrial polyps: A critical review of the literature. *Journal of Minimally Invasive Gynecology*, *18*(5), 569–581. https://doi.org/10.1016/j.jmig.2011.05.018

Sarsenova, M., Lawarde, A., Pathare, A. D. S., Saare, M., Modhukur, V., Soplepmann, P., Terasmaa, A., Käämbre, T., Gemzell-Danielsson, K., Lalitkumar, P. G. L., Salumets, A., & Peters, M. (2024). *Endometriotic lesions exhibit distinct metabolic signature compared to paired eutopic endometrium at the single-cell level*. https://doi.org/10.1101/2024.02.27.580606

Shao, G.-B., Chen, J.-C., Zhang, L.-P., Huang, P., Lu, H.-Y., Jin, J., Gong, A.-H., & Sang, J.-R. (2014). Dynamic patterns of histone H3 lysine 4 methyltransferases and demethylases during mouse preimplantation development. *In Vitro Cellular & Developmental Biology. Animal*, *50*(7), 603–613. https://doi.org/10.1007/s11626-014-9741-6

Shen, Y., Tang, X.-Y., Yang, Y.-C., Ke, X., Kou, W., Pan, C.-K., & Hong, S.-L. (2011). Impaired balance of Th17/Treg in patients with nasal polyposis. *Scandinavian Journal of Immunology*, *74*(2), 176–185. https://doi.org/10.1111/j.1365-3083.2011.02546.x

Shokeir, T. A., Shalan, H. M., & El-Shafei, M. M. (2004). Significance of endometrial polyps detected hysteroscopically in eumenorrheic infertile women. *The Journal of Obstetrics and Gynaecology Research*, *30*(2), 84–89. https://doi.org/10.1111/j.1447-0756.2003.00163.x

Stuenkel, C. A., Davis, S. R., Gompel, A., Lumsden, M. A., Murad, M. H., Pinkerton, J. V., & Santen, R. J. (2015). Treatment of Symptoms of the Menopause: An Endocrine Society Clinical Practice Guideline. *The Journal of Clinical Endocrinology & Metabolism*, *100*(11), 3975–4011. https://doi.org/10.1210/jc.2015-2236

Su, C.-H., Lin, I.-H., Tzeng, T.-Y., Hsieh, W.-T., & Hsu, M.-T. (2016). Regulation of IL-20 Expression by Estradiol through KMT2B-Mediated Epigenetic Modification. *PLoS ONE*, *11*(11), e0166090. https://doi.org/10.1371/journal.pone.0166090

Su, Y., Feng, W., & Shi, H. (2023). Treatment with progesterone attenuates proliferation of endometrial polyps (EP) via regulation of expression of miR-320b and its target gene, MCL1. *Archives of Medical Science*, *19*(6), 1934–1939. https://doi.org/10.5114/aoms/171308

Takeda, T., Banno, K., Kobayashi, Y., Adachi, M., Yanokura, M., Tominaga, E., Kosaki, K., & Aoki, D. (2019). Mutations of RAS genes in endometrial polyps. *Oncology Reports*, *42*(6), 2303–2308. https://doi.org/10.3892/or.2019.7353

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K., & Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, *6*(5), 377–382. https://doi.org/10.1038/nmeth.1315

Taylor, L. J., Jackson, T. L., Reid, J. G., & Duffy, S. R. G. (2003). The differential expression of oestrogen receptors, progesterone receptors, Bcl-2 and Ki67 in endometrial polyps. *BJOG: An International Journal of Obstetrics & Gynaecology*, *110*(9), 794–798. https://doi.org/10.1111/j.1471-0528.2003.02098.x

Tjarks, M., & Van Voorhis, B. J. (2000). Treatment of endometrial polyps. *Obstetrics and Gynecology*, *96*(6), 886–889. https://doi.org/10.1016/s0029-7844(00)01059-0

Tsujioka, H., Hachisuga, T., Fukuoka, M., Ueda, T., Miyahara, D., Horiuchi, S., Shirota, K., Yoshizato, T., Emoto, M., Miyamoto, S., & Kawarabayashi, T. (2009). Monitoring of Endometrial K- *ras* Mutation in Tamoxifen-Treated Patients With Breast Cancer. *International Journal of Gynecologic Cancer*, *19*(6), 1052–1056. https://doi.org/10.1111/IGC.0b013e3181a8b0aa

Tsujioka, H., Hachisuga, T., Hikita, S., Ueda, T., Yotsumoto, F., Shirota, K., Yoshizato, T., Kawarabayashi, T., Kuroki, M., & Miyamoto, S. (2010). Apoptosis as a Possible Candidate Mechanism for Removal of Tamoxifen-related Endometrial Cells with KRAS Mutations. *Anticancer Research*, *30*(8), 3119–3123.

Tsutsumi, A., Okada, H., Nakamoto, T., Okamoto, R., Yasuda, K., & Kanzaki, H. (2011). Estrogen induces stromal cell-derived factor 1 (SDF-1/CXCL12) production in human endometrial stromal cells: A possible role of endometrial epithelial cell growth. *Fertility and Sterility*, *95*(1), 444–447. https://doi.org/10.1016/j.fertnstert.2010.08.037

Valle, R. F. (1983). Hysteroscopy for gynecologic diagnosis. *Clinical Obstetrics and Gynecology*, *26*(2), 253–276. https://doi.org/10.1097/00003081-198306000-00007

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*(7), research0034.1. https://doi.org/10.1186/gb-2002-3-7-research0034

Vang, R., Barner, R., Wheeler, D. T., & Strauss, B. L. (2004). Immunohistochemical staining for Ki-67 and p53 helps distinguish endometrial Arias-Stella reaction from high-grade carcinoma, including clear cell carcinoma. *International Journal of Gynecological Pathology:*

Official Journal of the International Society of Gynecological Pathologists, *23*(3), 223–233. https://doi.org/10.1097/01.pgp.0000130106.13533.38

Vanni, R., Dal Cin, P., Marras, S., Moerman, P., Andria, M., Valdes, E., Deprest, J., & Van Den Berghe, H. (1993). Endometrial polyp: Another benign tumor characterized by 12q13–q15 changes. *Cancer Genetics and Cytogenetics*, *68*(1), 32–33. https://doi.org/10.1016/0165-4608(93)90070-3

Vieira, M. D. C., Vitagliano, A., Rossette, M. C., Neto, L. C. D. A., Gallo, A., & Sardo, A. D. S. (2022). Endometrial Polyps: Update Overview on Etiology, Diagnosis, Natural History and Treatment. *Clinical and Experimental Obstetrics & Gynecology*, *49*(10), 232. https://doi.org/10.31083/j.ceog4910232

Vitagliano, A., Cialdella, M., Cicinelli, R., Santarsiero, C. M., Greco, P., Buzzaccarini, G., Noventa, M., & Cicinelli, E. (2021). Association between Endometrial Polyps and Chronic Endometritis: Is It Time for a Paradigm Shift in the Pathophysiology of Endometrial Polyps in Pre-Menopausal Women? Results of a Systematic Review and Meta-Analysis. *Diagnostics*, *11*(12), 2182. https://doi.org/10.3390/diagnostics11122182

Vousden, K. H., & Prives, C. (2009). Blinded by the Light: The Growing Complexity of p53. *Cell*, *137*(3), 413–431. https://doi.org/10.1016/j.cell.2009.04.037

Wang, Z., Li, L., Su, X., Gao, Z., Srivastava, G., Murray, P. G., Ambinder, R., & Tao, Q. (2012). Epigenetic silencing of the 3p22 tumor suppressor DLEC1 by promoter CpG methylation in non-Hodgkin and Hodgkin lymphomas. *Journal of Translational Medicine*, *10*(1), 209. https://doi.org/10.1186/1479-5876-10-209

Wennerberg, K., Rossman, K. L., & Der, C. J. (2005). The Ras superfamily at a glance. *Journal of Cell Science*, *118*(5), 843–846. https://doi.org/10.1242/jcs.01660

Wethington, S. L., Herzog, T. J., Burke, W. M., Sun, X., Lerner, J. P., Lewin, S. N., & Wright, J. D. (2011). Risk and predictors of malignancy in women with endometrial polyps. *Annals of Surgical Oncology*, *18*(13), 3819–3823. https://doi.org/10.1245/s10434-011-1815-z

Wu, J., Feng, X., Du, Y., Luan, B., Yu, H., Yu, Y., Wu, L., & Zhao, H. (2019). β-catenin/LIN28B promotes the proliferation of human choriocarcinoma cells via Let-7a repression. *Acta Biochimica et Biophysica Sinica*, *51*(5), 455–462. https://doi.org/10.1093/abbs/gmz027

Yang, Z., Deng, Y., Zhang, K., Bai, Y., Zhu, J., Zhang, J., Cheng, J., Li, L., He, J., & Wang, W. (2021). LIN28A polymorphisms and hepatoblastoma susceptibility in Chinese children. *Journal of Cancer*, *12*(5), 1373–1378. https://doi.org/10.7150/jca.52621

Ying, J., Poon, F. F., Yu, J., Geng, H., Wong, A. H. Y., Qiu, G.-H., Goh, H. K., Rha, S. Y., Tian, L., Chan, A. T. C., Sung, J. J. Y., & Tao, Q. (2009). DLEC1 is a functional 3p22.3 tumour suppressor silenced by promoter CpG methylation in colon and gastric cancers. *British Journal of Cancer*, *100*(4), 663–669. https://doi.org/10.1038/sj.bjc.6604888

Yu, J., Wang, J.-G., Zhang, L., Yang, H.-P., Wang, L., Ding, D., Chen, Q., Yang, W.-L., Ren, K.-H., Zhou, D.-M., Zou, Q., Jin, Y.-T., & Liu, X.-P. (2016). MicroRNA-320a inhibits breast cancer metastasis by targeting metadherin. *Oncotarget*, *7*(25), 38612–38625. https://doi.org/10.18632/oncotarget.9572

Zech, M., Boesch, S., Maier, E. M., Borggraefe, I., Vill, K., Laccone, F., Pilshofer, V., Ceballos-Baumann, A., Alhaddad, B., Berutti, R., Poewe, W., Haack, T. B., Haslinger, B., Strom, T. M., & Winkelmann, J. (2016). Haploinsufficiency of KMT2B, Encoding the Lysine-Specific Histone Methyltransferase 2B, Results in Early-Onset Generalized Dystonia. *American Journal of Human Genetics*, *99*(6), 1377–1387. https://doi.org/10.1016/j.ajhg.2016.10.010

Zech, M., Jech, R., Havránková, P., Fečíková, A., Berutti, R., Urgošík, D., Kemlink, D., Strom, T. M., Roth, J., Růžička, E., & Winkelmann, J. (2017). KMT2B rare missense variants in generalized dystonia. *Movement Disorders: Official Journal of the Movement Disorder Society*, *32*(7), 1087–1091. https://doi.org/10.1002/mds.27026

Zhang, Q., Ying, J., Li, J., Fan, Y., Poon, F. F., Ng, K. M., Tao, Q., & Jin, J. (2010). Aberrant Promoter Methylation of DLEC1, a Critical 3p22 Tumor Suppressor for Renal Cell Carcinoma, is Associated With More Advanced Tumor Stage. *Journal of Urology*, *184*(2), 731–737. https://doi.org/10.1016/j.juro.2010.03.108

Zhang, X., Li, M., Sun, G., Bai, Y., Lv, D., & Liu, C. (2020). MiR-563 restrains cell proliferation via targeting LIN28B in human lung cancer. *Thoracic Cancer*, *11*(1), 55–61. https://doi.org/10.1111/1759-7714.13257

Zhang, Y., Kang, M., Zhang, B., Meng, F., Song, J., Kaneko, H., Shimamoto, F., & Tang, B. (2023). Retraction Note: m6A modification-mediated CBX8 induction regulates stemness and chemosensitivity of colon cancer via upregulation of LGR5. *Molecular Cancer*, *22*, 166. https://doi.org/10.1186/s12943-023-01874-z

Zhao, D., Yuan, H., Fang, Y., Gao, J., Li, H., Li, M., Cong, H., Zhang, C., Liang, Y., Li, J., Yang, H., Yao, M., Du, M., Tu, H., & Gan, Y. (2023). Histone Methyltransferase KMT2B Promotes Metastasis and Angiogenesis of Cervical Cancer by Upregulating EGF Expression. *International Journal of Biological Sciences*, *19*(1), 34–49. https://doi.org/10.7150/ijbs.72381

ACKNOWLEDGMENTS

First of all, I would like to express my deepest gratitude to my supervisors, Merli Saare, Amruta Pathare and Tambet Tõnissoo. I appreciate and value the fact that they allowed me to become part of such a complex, very important and very interesting project to study an endometrial polyp and write my graduate work based on this. Your valuable advice, patience and professionalism were invaluable and important to me. Only thanks to You were I able to overcome all the difficulties that stood in the way of creating this work.

I really appreciate the strength, knowledge, and laboratory skills invested in me. In this work, and I promise that I will carry this through the years and will apply it beyond the university doors.

I would also like to express my gratitude to the people who were an irreplaceable part of this project Maire Peters, Ankita Sunil Lawarde, Andres Salumets.

LIHTLITSENTS

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Darja Tarassova,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Endomeetriumi polüübi geenieksпressiooni profileerimine“ (ingl.k. Gene expression profiling of endometrial polyp), mille juhendajad on Merli Saare (Ph.D), Tambet Tõnissoo (Ph.D), Amruta Pathare (Ph.D) reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Darja Tarassova
28.05.2024