



**UNIVERSIDADE FEDERAL DO CEARÁ**

**FACULDADE DE MEDICINA**

**NÚCLEO DE PESQUISA E DESENVOLVIMENTO DE MEDICAMENTOS**

**PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA TRANSLACIONAL**

**VITÓRIA BRAGA MELO**

**ASSINATURA DE EXPRESSÃO GÊNICA EM MULHERES COM  
ENDOMETRIOSE: UMA ABORDAGEM *IN SILICO***

**FORTALEZA/CE 2025**

VITÓRIA BRAGA MELO

ASSINATURA DE EXPRESSÃO GÊNICA EM MULHERES COM ENDOMETRIOSE:  
UMA ABORDAGEM *IN SILICO*

Dissertação apresentada ao Programa de Pós-Graduação em Medicina Translacional da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Medicina Translacional.

Orientador: Profa. Dra. Cristiana Libardi Miranda Furtado  
Coorientador: Prof. Dr. Antonio Edson Rocha Oliveira

FORTALEZA/CE 2025

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- M486a    Melo, Vitoria Braga.  
          Assinatura de expressão gênica em mulheres com endometriose : Uma abordagem in silico / Vitoria  
          Braga Melo. – 2025.  
          90 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Medicina, Programa de Pós-  
          Graduação em Medicina Translacional, Fortaleza, 2025.  
          Orientação: Prof. Dr. Cristiana Libardi Miranda Furtado.  
          Coorientação: Prof. Dr. Antonio Edson Rocha Oliveira.
1. Metanálise. 2. Transcriptoma. 3. Endometriose. I. Título.

CDD 610

---

VITÓRIA BRAGA MELO

ASSINATURA DE EXPRESSÃO GÊNICA EM MULHERES COM ENDOMETRIOSE:  
UMA ABORDAGEM *IN SILICO*

Dissertação apresentada ao Programa de Pós-Graduação em Medicina Translacional da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Medicina Translacional.

Aprovada em: \_\_\_\_ / \_\_\_\_ / \_\_\_\_.

BANCA EXAMINADORA

---

Profa. Dra. Cristiana Libardi Miranda Furtado(Orientadora)  
Universidade Federal de Minas Gerais (UFMG)

---

Dra. Thalita Basso Scandolara  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Rosana Maria dos Reis  
Faculdade de Medicina de Ribeirão Preto (FMRP/USP)

---

Prof. Dr. Antonio Edson Rocha Oliveira (Coorientador)  
Universidade de Fortaleza (UNIFOR)

## DEDICATÓRIA

Dedico este trabalho a Deus, pois sem Sua presença, força e graça, nada disso teria sido possível.

Dedico também à minha família e amigos chegados, pelo amor incondicional, apoio constante e por nunca deixarem de acreditar em mim.

## **AGRADECIMENTOS**

Agradeço, primeiramente, a Deus, por me conceder força, saúde e pelas oportunidades que permitiram a conclusão desta importante etapa profissional.

À minha família, especialmente e infinitamente aos meus maravilhosos pais, José e Helena, que são minha base e meus maiores incentivadores. Obrigada por estarem ao meu lado em cada desafio, acreditando em mim mesmo nos momentos em que nem eu conseguia fazê-lo. Ao meu irmão, Felipe, pela companhia constante e pelo apoio, mesmo entre nossas inevitáveis implicâncias. À minha madrinha do coração, Penha, por sua presença tão constante e por sempre me ouvir e aconselhar com tanto carinho.

Às amigas de toda a vida, Natália, Rayanne, Aline, Fátima e Luana Braga, por toda a nossa história construída com amizade, respeito, confiança e admiração. Compartilhar nossas conquistas sempre foi motivo de alegria mútua, e assim sempre será.

Sou profundamente grata aos meus parceiros de fé e de vida, Davi, Yohana, Sam, Gabi, Brendo e John que, com humor e amor, me tiram do eixo e, ao mesmo tempo, me colocam de volta no caminho sempre que me sinto perdida. São pessoas preciosas e fundamentais em minha caminhada.

Sou profundamente grata a Cássia e Francilene, duas mulheres fenomenais, que juntas são capazes de melhorar qualquer ambiente. Cassinha e Fran não apenas me incentivaram a retornar a vida acadêmica, como também me mostraram o caminho a seguir.

Gratidão à minha brilhante orientadora, Prof<sup>ª</sup> Dra. Cristiana Libardi Miranda Furtado, agradeço imensamente pela dedicação, humanidade e disponibilidade com que me guiou. Sua postura ética e acolhedora tornou-se, para mim, um exemplo inspirador de profissional e ser humano.

Ao meu querido coorientador, Prof. Dr. Antonio Edson Rocha Oliveira, sou imensamente grata por ter dedicado tempo e paciência para me ensinar bioinformática desde o início, tornando possível não apenas a realização deste trabalho, mas também meu crescimento

acadêmico e profissional.

Um agradecimento mais que especial às minhas queridas “episisters”: Sarah Leyenne, que a vida me trouxe de volta depois de tantos anos e que agora permanece de forma definitiva; Maria Rodrigues, um presente da Fiocruz cuja amizade se tornou um alicerce indispensável; Mariana Manso, que me ensinou a importância de expressar sentimentos, seja nas lágrimas ou na coragem de arriscar; e Jéssica Mota, cuja doçura e eficiência sempre me lembram de manter minha essência. Nossa parceria começou na rotina do laboratório e se transformou em uma amizade que levo para a vida.

Gratidão imensa à Thalita Scandolaro por sempre ser fonte dos melhores conselhos e direcionamentos profissionais com sinceridade e empatia.

Agradeço a todos da família LOE, em nome da Prof<sup>a</sup> Dra. Claudia do O` Pessoa e Silvana França, que me acolheram com tanto carinho e me ajudaram sempre que precisei. Aos meus amigos queridos Aniele, Dave, Lina Castelo, Celina (Cebeautiful), Pedro, Neto, Sarah S, João Ávila, Levi, Cecília, Cainã, Sophia, Arthur, Lina Gress, João Victor e Robert, obrigada por tornarem meus dias mais leves e felizes, mesmo nos momentos mais desafiadores, inclusive nas reuniões mais tensas.

À Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), agradeço pelo apoio concedido por meio da bolsa de estudo de Mestrado, essencial para a realização desta pesquisa.

Por fim, agradeço a todos que, direta ou indiretamente, contribuíram para esta fase da minha vida. Esta conquista é, sem dúvida, coletiva e carrego cada um de vocês em meu coração.

“A análise de expressão gênica também abre caminho para o desenvolvimento de terapias direcionadas, como inibidores de vias inflamatórias ou moduladores epigenéticos, que podem melhorar os resultados clínicos (ZONDERVAN et al., 2020).”



## RESUMO

**Assinatura De Expressão Gênica Em Mulheres Com Endometriose: Uma Abordagem *In Silico*. Vitoria Braga Melo. Orientadora: Profa. Dra. Cristiana Libardi Miranda Furtado. Dissertação de Mestrado. Programa de Pós-Graduação em Medicina Translacional. Núcleo de Pesquisa e Desenvolvimento de Medicamentos, Faculdade de Medicina, UFC. Fortaleza, 2025.**

A endometriose é uma doença crônica, um problema crescente na saúde pública, que afeta aproximadamente de 10% a 15% da população feminina em idade fértil. Caracterizada pela presença de células semelhantes as endometriais fora da cavidade uterina cujos fenótipos são classificados em endometriose peritoneal superficial, endometrioma ovariano e endometriose profunda infiltrativa. Os sintomas variam de acordo com o fenótipo da doença, podendo ser desde quadros assintomáticos até quadros dor crônica e infertilidade que estão entre os principais desfechos clínicos da doença. Apresenta-se como uma doença complexa e multifatorial, com diversos genes contribuindo para a fisiopatologia da doença e a interação com fatores ambientais que podem agravar os sintomas e tratamento. Apesar de extensas pesquisas, um perfil genético abrangente para a doença ainda precisa ser estabelecido, destacando a necessidade de investigações adicionais para elucidar os mecanismos moleculares subjacentes da endometriose. Assim, o objetivo deste trabalho foi identificar assinatura de expressão gênica consensual em mulheres com endometriose. Por meio de abordagem *in silico*, realizou-se a metanálise sete conjuntos de dados (*datasets*) distintos, sendo cinco de *microarranjo* e dois de *RNA-seq* que incluíam resultados de mulheres com diferentes fenótipos, subtipos analisados, destacam-se o endométrio eutópico de mulheres com endometriose na fase secretora inicial (ESPEM), comparado ao endométrio de mulheres sem a doença (ESPN), e a forma ovariana (ODEM), caracterizada pela presença de endometriomas. Além disso, as lesões ectópicas (EEM) foram comparadas ao endométrio eutópico normal (EN), assim como casos de endometriose não especificada quanto à localização (EM). Os fenótipos de endometriose infiltrativa profunda incluem as formas que acometem a bexiga (DiEB) e o intestino (DiElN), ambas associadas a maior gravidade clínica. As lesões peritoneais, por sua vez, foram categorizadas morfológicamente em vermelhas (PeLR), pretas (PeLB) e brancas (PeLW),

representando diferentes estágios evolutivos da doença — desde formas ativas e vascularizadas até lesões cicatriciais e fibróticas. Na metanálise dos dados de microarranjo foram realizadas nove comparações, resultando em um total de 205 genes diferencialmente expressos (GDEs). Posteriormente foi feita a integração dos dados de *RNA-seq* que validaram 57 GDEs. As análises de enriquecimento funcional foram realizadas separadamente para os genes com expressão aumentada e diminuída no grupo endometriose. Os genes *CFH*, *GAS6*, *TSPAN4*, *WNT2B* e *PLSCR4* foram os cinco principais genes regulados positivamente, enquanto os genes *PIGN*, *ITGB8*, *PFAS*, *KLHL13* e *PARP1* foram os regulados negativamente. As vias enriquecidas relacionadas aos genes com expressão aumentada estavam envolvidas principalmente na ativação do sistema complemento, sinalização celular, proliferação celular, inflamação, angiogênese e biossíntese de estrogênio. Paralelamente, os genes com expressão diminuída estão ligados a vias de reparo de DNA e regulação do complexo principal de histocompatibilidade e escape imunológico. Esses resultados demonstram que a integração de dados transcriptômicos públicos permitiu a identificação de uma assinatura gênica consistente na endometriose, fornecendo novas perspectivas para a compreensão de sua biologia e possíveis alvos terapêuticos.

**Palavras-chave:** Metanálise. Transcriptoma. Endometriose.

## ABSTRACT

**Gene Expression Signature in Women with Endometriosis: An *In Silico* Approach. Vitoria Braga Melo. Advisor: Prof. Dr. Cristiana Libardi Miranda Furtado. Master's Dissertation. Graduate Program in Translational Medicine. Center for Drug Research and Development, Faculty of Medicine, Federal University of Ceará (UFC). Fortaleza, 2025.**

Endometriosis is a chronic and multifactorial disease that poses a growing public health challenge, affecting approximately 10–15% of women of reproductive age. It is defined by the presence of endometrial-like tissue outside the uterine cavity and can manifest in three main phenotypes: superficial peritoneal endometriosis, ovarian endometrioma, and deep infiltrating endometriosis. Clinical presentation varies depending on the phenotype, ranging from asymptomatic cases to chronic pelvic pain and infertility, which are among the most prevalent and debilitating outcomes. Despite considerable research efforts, the molecular mechanisms underlying endometriosis remain incompletely understood, and a comprehensive genetic profile has yet to be fully established. The pathogenesis is thought to involve complex gene–environment interactions, with numerous genes contributing to disease susceptibility, symptom severity, and treatment response. This study aimed to identify a consensual gene expression signature associated with endometriosis. Using an *in silico* approach, we conducted a meta-analysis of seven publicly available transcriptomic datasets — five based on microarranjo platforms and two on bulk RNA sequencing — encompassing various phenotypic presentations of the disease. The analyzed subtypes included the eutopic endometrium of women with endometriosis during the early secretory phase (ESPEM), compared to healthy controls (ESPN), and ovarian endometriosis (ODEM), characterized by the presence of endometriomas. Ectopic lesions (EEM) were compared to normal eutopic endometrium (EN), and non-location-specific endometriosis cases (EM) were also included. Deep infiltrating endometriosis phenotypes comprised lesions affecting the bladder (DiEB) and intestine (DiEIn), both of which are associated with increased clinical severity. Additionally, peritoneal lesions were morphologically classified as red (PeLR), black (PeLB), or white (PeLW), reflecting different stages of disease progression — from active vascularized lesions to fibrotic, scarred tissue. The meta-analysis of microarranjo data identified 205 differentially expressed genes (DEGs) across nine distinct group comparisons.

Subsequent integration with RNA-seq datasets validated 57 of these DEGs. Functional enrichment analyses were performed separately for upregulated and downregulated genes in the endometriosis group. The top five upregulated genes were *CFH*, *GAS6*, *TSPAN4*, *WNT2B*, and *PLSCR4*, while *PIGN*, *ITGB8*, *PFAS*, *KLHL13*, and *PARP1* were among the most significantly downregulated. Upregulated genes were enriched in pathways related to complement activation, cell signaling, proliferation, inflammation, angiogenesis, and estrogen biosynthesis. Conversely, downregulated genes were associated with DNA repair mechanisms, regulation of the major histocompatibility complex, and immune evasion processes. These findings demonstrate that the integration of publicly available transcriptomic data enables the identification of a robust and consistent gene expression signature in endometriosis. This signature provides valuable insights into the molecular landscape of the disease and may contribute to the development of novel diagnostic biomarkers and therapeutic targets.

**Keywords:** Meta-analysis. Transcriptome. Endometriosis.

## LISTA DE FIGURAS

Figura 1. Caminhos potenciais para a patogênese e características fisiopatológicas da endometriose pélvica. ....	23
Figura 2. Fluxograma seguido nos scripts de análises dos datasets da técnica de microarray. ....	37
Figura 3. Fluxograma seguido nos scripts de análises dos datasets da técnica de RNA-seq. Fonte própria. ....	41
Figura 4. Fluxograma seguido nos scripts de análises dos datasets da técnica de bulk RNA-seq. ....	44
Figura 5. Fluxograma de filtragem dos datasets para análise de expressão gênica. ....	48
Figura 6. Gráficos gerados pelo pacote Array Quality Metrics de todos os microarrays Affymetrix analisados. ....	50
Figura 7. Gráficos gerados pelo pacote Array Quality Metrics referente ao dataset GSE141549, Illumina. A: Comparação entre corridas. B: Gráfico de dispersão de grupos. C: Distribuições de intensidade de matriz. D: Densidade entre grupos. ....	52
Figura 8. Volcano plots obtidos mediante os resultados de GDEs de cada comparação analisada. ....	53
Figura 9. Gráfico de metanálise função votecount. A: Contagem de GDEs entre os datasets. B: GDEs em consenso entre as comparações. ....	55
Figura 10. referente a meta-análise. Grupos experimentais analisados. ....	55
Figura 11. Gráfico de coluna referente às vias enriquecidas do conjunto de GDEs obtidos na metanálise. ....	59
Figura 12. Resumo das vias biológicas enriquecidas e dos genes relacionados. ....	61
Figura 13. Gráficos gerados no relatório de qualidade do FASTQC referente aos datasets GSE135485 e GSE168902. ....	63
Figura 14. Volcano plots das análises referentes aos datasets selecionados da técnica de RNA-seq entre mulheres com endometriose e controle. ....	65
Figura 15. Diagrama de Venn, demonstrando os GDEs em consenso entre integração das técnicas. ....	66
Figura 16. Heatmap demonstrando a assinatura gênica da endometriose considerando múltiplos fatores mostra mudança no perfil de expressão da endometriose. Os valores de log2FC dos estão representados no gráfico em vermelho (aumento) e azul (diminuição). ..	67

## LISTA DE TABELAS

Tabela 1. Critérios de inclusão e exclusão dos estudos selecionados. ....	36
Tabela 2. Lista de estudos transcriptômicos selecionados.....	46
Tabela 3. Datasets de microarranjo descritos em grupos. ....	48
Tabela 4. Quantitativo total de GDEs por comparação. ....	54
Tabela 5. Conjunto de 205 GDEs obtidos em consenso na metanálise entre os datasets de microarranjo.....	56
Tabela 6. Quantitativo geral das vias biológicas encontradas em cada banco de dados. ....	58
Tabela 7. Datasets RNA-seq. Os estudos marcados em verde foram analisados e incluídos à pesquisa. ....	62
Tabela 8. Quantificação de GDEs identificados em cada um dos datasets de RNA-seq.....	65

## **LISTA DE ABREVIATURAS E SIGLAS**

ABNT	Associação Brasileira de Normas Técnicas
GEO	do inglês – Gene Expression Omnibus
NBR	Norma Brasileira Regulamentar
UFC	Universidade Federal do Ceará
NCBI	National Center for Biotechnology Information
RNA	do inglês – Ribonucleic Acid
DNA	do inglês – Deoxyribonucleic Acid
FUNCAP	Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico

## SUMÁRIO

1.	Introdução.....	17
2.	Objetivos.....	19
2.1.	Objetivos Gerais .....	19
2.2.	Objetivos Específicos.....	19
3.	Revisão de Literatura.....	20
3.1.	Etiologia.....	21
3.2.	Sintomatologia e diagnóstico .....	25
3.3.	Dor e infertilidade na endometriose.....	26
3.4.	Expressão gênica na endometriose .....	28
3.5.	Ferramentas de bioinformática para análise de expressão gênica.....	30
3.5.1.	<i>Microarranjo e RNA-seq.....</i>	<i>31</i>
3.5.3.	Enriquecimento Funcional de Vias biológicas .....	34
4.	Materiais e métodos.....	36
4.1.	Análise de transcriptoma em bancos de dados.....	36
4.1.1.	<i>Obtenção de estudos transcriptômicos relacionados à Endometriose.....</i>	<i>36</i>
4.1.2.	Identificação de assinatura gênica da endometriose através de análises de transcriptômas. ....	36
4.1.2.1.	Assinatura gênica da endometriose através de análises transcriptômicas de <i>Microarranjo</i> .....	37
4.1.2.2.	<i>Metanálise entre os resultados obtidos nas análises de Microarranjo. ....</i>	<i>39</i>
4.1.2.3.	<i>Análise de enriquecimento funcional de vias biológicas.....</i>	<i>40</i>
4.1.2.4.	<i>Validação da assinatura gênica da endometriose através de análises transcriptômicas RNA-seq .....</i>	<i>40</i>
5.	Resultados.....	46
5.1.	Estudos encontrados por meio da curadoria.....	46
5.2.	Análises de expressão gênica dos <i>datasets</i> referentes à técnica de.....	47
	<i>microarranjo. ....</i>	<i>47</i>
5.2.1.	<i>Análise de qualidade dos dados feitos com pacote Array Quality Metrics .....</i>	<i>50</i>
5.2.2.	Resultado das Análise dos genes diferencialmente expressos em cada estudo selecionado.....	53
5.3.	Resultados da Metanálise entre os dados de <i>microarranjo</i> .....	54
5.4.	Resultados do Enriquecimento Funcional .....	57
5.5.	Resultados da análise dos genes diferencialmente expressos em dados de .....	
	<i>RNA-seq .....</i>	<i>61</i>
6.	Discussão .....	69
7.	Conclusão .....	78
8.	Referências.....	79



## 1. Introdução

Endometriose constitui uma questão de saúde pública de relevância global, impactando significativamente tanto os aspectos físicos quanto psicológicos das mulheres afetadas. Os desfechos clínicos primários que incluem dor pélvica crônica e infertilidade (ZONDERVAN et al., 2020; GIUDICE; KAO et al., 2004). Por tratar-se de uma condição de etiologia complexa e multifatorial, fatores genéticos e ambientais contribuem de maneira crucial para a fisiopatologia da doença (BULUN et al., 2019; BURNEY; GIUDICE, 2012). Além disso, alterações nos mecanismos epigenéticos, que modulam a expressão gênica, estão diretamente associadas à susceptibilidade ou predisposição à endometriose (GUO, 2009; NASU et al., 2011). A expressão de genes envolvidos na regulação de vias de crescimento e diferenciação celular tem sido correlacionada à gravidade da endometriose, particularmente nos casos em que a doença ultrapassa a cavidade uterina e infiltra múltiplos tecidos da região pélvica, resultando em prognósticos desfavoráveis e no comprometimento da qualidade de vida das mulheres acometidas pela doença (BULUN et al., 2019).

Os avanços nas tecnologias ômicas, que empregam estratégias de análise em larga escala para a identificação de genes, transcritos, proteínas, metabólitos e outros biomarcadores, representam um progresso significativo na compreensão dos sistemas biológicos, especialmente nos que se refere às condições de saúde e doença humana (FASSBENDER et al., 2013). Esses marcadores estão disponíveis em bancos de dados online, que consolidam resultados de estudos variados, abrangendo diferentes grupos populacionais e biomarcadores. A utilização de abordagens *in silico* tem se consolidado como uma estratégia essencial na pesquisa biomédica, ao integrar e analisar grandes volumes de dados por meio de ferramentas computacionais. A análise de bancos de dados permitem a identificação precisa, acurada e economicamente viável de marcadores moleculares associados a diversas doenças (AGHAJANOVA et al., 2010), os quais podem ser utilizados no diagnóstico precoce, prognóstico e desenvolvimento de novas terapias (PAPPALARDO et al., 2019).

Dentre as abordagens ômicas, a transcriptômica é o campo da biologia molecular dedicado ao estudo do conjunto completo de moléculas de RNAs transcritos por uma célula, tecido ou organismo em um dado momento (MUTZ et al., 2013). Essa

abordagem permite compreender a atividade gênica em diferentes condições fisiológicas ou patológicas, identificando genes diferencialmente expressos, variantes de transcritos e redes regulatórias associadas. Com o avanço das tecnologias de alto desempenho, como o sequenciamento (*RNA-Seq*) ou microarranjos (*microarranjos*) de RNAs, tornou-se possível explorar o transcriptoma de forma mais sensível e abrangente (STARK et al., 2019).

A realização de metanálises com dados transcriptômicos é particularmente relevante no estudo da endometriose, a o integrar dados oriundos de diferentes estudos públicos, possibilitando ampliar o número de amostras analisadas e incluir diferentes populações, aumentando o poder estatístico e a sensibilidade na identificação de genes diferencialmente expressos (GDEs). Essa abordagem permite contornar limitações de estudos individuais, como pequenas coortes ou variabilidade técnica, além de fornecer uma visão mais abrangente e reproduzível dos mecanismos moleculares associados à doença. No caso da endometriose, que envolve múltiplos fenótipos clínicos e padrões moleculares diversos, a metanálise se mostra essencial para revelar assinaturas gênicas robustas, vias biológicas comuns entre estudos e potenciais biomarcadores com relevância clínica. Além disso, ao valorizar o reuso de dados públicos, essa estratégia contribui para a ciência aberta e reduz a duplicação de esforços experimentais (MA et al., 2022; RAMASAMY et al., 2008).

A integração da análise de dados de microarranjos com outras técnicas, como o RNA-seq, como uma forma de validação, adicionados a outras abordagens para refinar a pesquisa, tais como a metanálise e o enriquecimento funcional, podem ajudar a fornecer um quadro mais completo da expressão gênica na endometriose, favorecendo na identificação de biomarcadores transcriptômicos específicos e robustos com potencial de aprimorar o diagnóstico precoce e o tratamento personalizado da endometriose.

## 2. Objetivos

### 2.1. Objetivos Gerais

Identificar a assinatura de expressão gênica *in silico* e *vias biológicas associadas* a endometriose.

### 2.2. Objetivos Específicos

- Realizar a curadoria dos estudos que incluem análise de transcriptoma em mulheres com endometriose;
- Realizar análise de expressão gênica diferencial em datasets referentes a técnica de microarranjo;
- Realizar análise de expressão gênica diferencial em datasets referentes a técnica de sequenciamento de RNA (*RNA-seq*);
- Metanálise e integração dos dados de microarranjo e *RNA-seq* para identificar assinatura de expressão gênica em mulheres com endometriose;
- Avaliar as vias biológicas relacionadas aos genes diferencialmente expressos (GDEs) identificados *in silico* em mulheres com endometriose.

### 3. Revisão de Literatura

A endometriose é uma condição ginecológica crônica, associada à presença de estrogênio, e caracterizada pela presença de células semelhantes às do endométrio fora da cavidade uterina (SIGNORILE; BALDI, 2010). Essa condição afeta aproximadamente 15% das mulheres em idade fértil e cerca de 3% das mulheres após a menopausa (FUKUNAGA, 2001; MEHEDINTU et al., 2014). Dependendo da população investigada, a prevalência da endometriose pode variar de 6% a 10% das mulheres em idade reprodutiva, mais de 60% das adolescentes com dor pélvica crônica e até 50% das mulheres com infertilidade (GIUDICE, 2010). A doença é classificada em estágios, incluindo a endometriose peritoneal superficial (SUP), endometrioma ovariano (OMA) e endometriose profunda infiltrativa (DIE), sendo esta última considerada a forma mais grave. A endometriose infiltrativa acomete aproximadamente 1% das mulheres na menacme e é caracterizada por dores intensas na maioria das mulheres, podendo afetar o trato urinário e o intestino (NÁCUL; SPRITZER, 2010; BERLANDA et al., 2017).

Cerca de 15% a 50% das mulheres com endometriose apresentam o fenótipo peritoneal superficial, caracterizada pela presença de três tipos de lesões que representam a evolução da doença (IMPERIALE et al., 2023). O principal sítio de localização da endometriose é intrapélvico, quando os focos se concentram nos ovários, tubas uterinas e na superfície externa do útero e intestinos, bem como no tecido de revestimento da cavidade pélvica. Também podem ser extra pélvicas, configurando-se como cutâneas quando em cicatrizes de cirurgias como cesárea, laparoscopia e histerectomia (GARCIA et al., 2009). O tecido endometriótico pode ser encontrado na cavidade pélvica, no peritônio pélvico, ovários, septo reto-vaginal, pericárdio, pleura, rim, bexiga, e, em alguns casos mais isolados, cérebro, além de poder ser raramente encontrado no trato urinário de homens (GIUDICE; KAO, 2004). Nesses casos, a ocorrência geralmente está associada à exposição prolongada a estrogênios exógenos devido a tratamentos hormonais como nos casos de câncer de próstata, ou endógeno, em casos como obesidade ou cirrose hepática (PATEYRON et al., 2020; SOBRINHO et al., 2018) podendo ocorrer também devido a distúrbios da diferenciação sexual durante o desenvolvimento embrionário (DE SOUZA et al., 2019).

As lesões de glândulas e estroma endometriais localizadas fora da cavidade endometrial são descritas como ectópicas, por oposição ao endométrio normalmente localizado ou eutópico (FREITAS et al., 2011; NASU et al., 2011). Diversos estudos têm mostrado que, ainda que histologicamente semelhantes, esses tecidos podem ser distintos tanto bioquimicamente quanto funcionalmente. Essas diferenças incluem alterações na estrutura tecidual, taxa de proliferação, componentes do sistema imune, mecanismos, taxa de apoptose, expressão de proteínas e genes específicos, bem como a produção de esteroides e citocinas (BULUN et al., 2019; BURNEY; GIUDICE, 2012; GUO, 2009). Tais distinções são fundamentais para diferenciar o tecido eutópico (localizado dentro da cavidade uterina) do ectópico (FASSBENDER et al., 2013). É aconselhável que conduta clínica não deve se limitar ao momento do diagnóstico ou focar exclusivamente nas lesões, mas sim considerar a totalidade da experiência da paciente ao longo da sua “vida com endometriose”, isso implica em um plano terapêutico individualizado, interdisciplinar e multimodal, que leve em conta não apenas os sintomas e a extensão das lesões, mas também o impacto na qualidade de vida, os desejos reprodutivos e as prioridades da paciente. Tal perspectiva é especialmente relevante frente à heterogeneidade clínica e biológica da doença, evidenciada por diferentes fenótipos e respostas variáveis aos tratamentos médicos e cirúrgicos (CHAPRON et al., 2019).

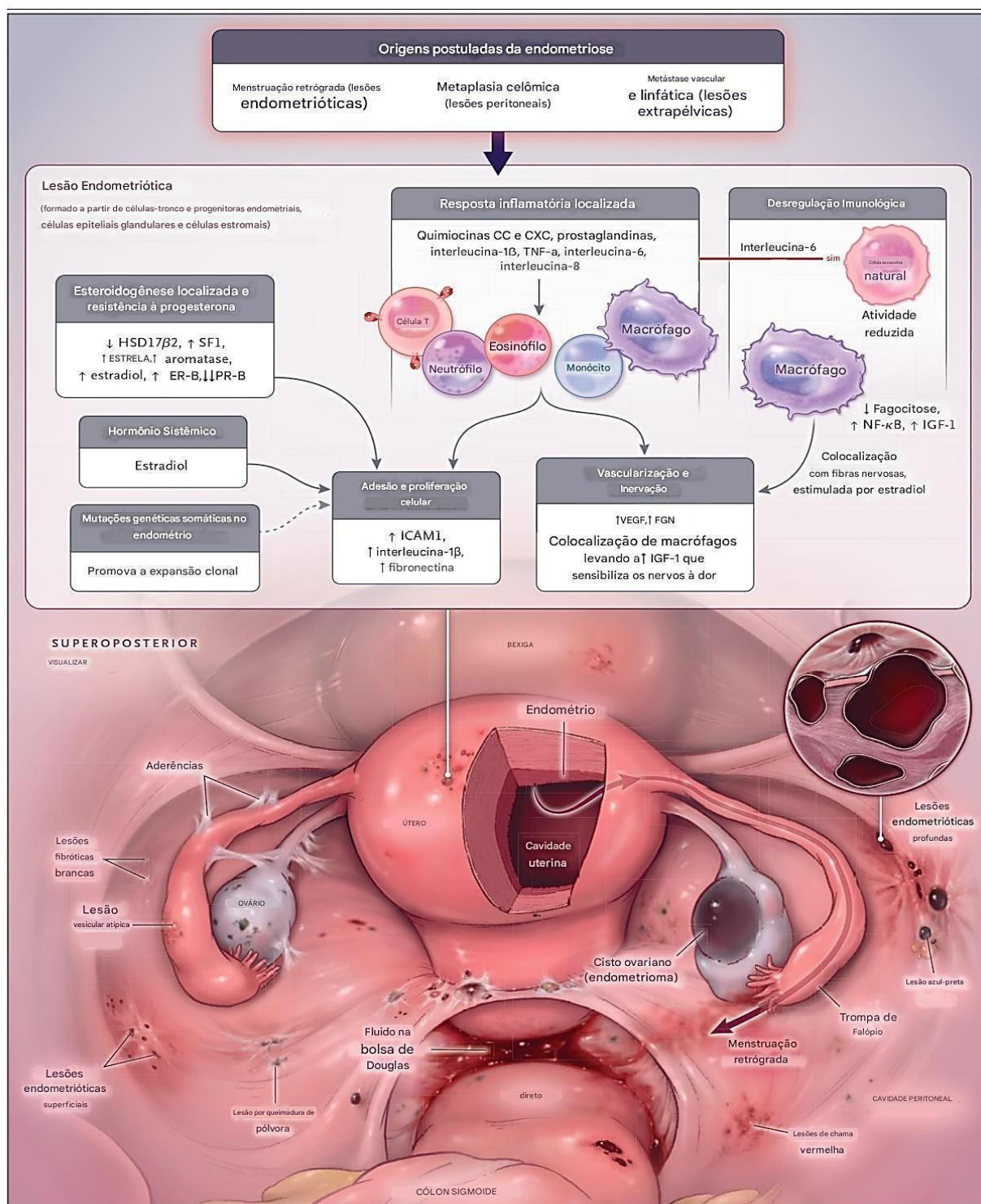
### **3.1. Etiologia**

A etiopatogenia da endometriose permanece controversa e ainda não foi completamente elucidada, embora diversas hipóteses tenham sido propostas (DE SOUSA BARBOSA et al., 2015). Dentre elas, a interação de fatores genéticos e ambientais, dentre eles imunológicos como no estudo de Amro et al. (2022), que propôs uma nova perspectiva sobre a patogênese da endometriose ao destacar o papel do microbioma peritoneal, desafiando a concepção anterior do peritônio como um ambiente estéril. Segundo os autores, microrganismos oriundos da vagina, do útero, das trompas e do trato gastrointestinal podem atingir a cavidade peritoneal por diferentes vias, incluindo a migração transmural intestinal. Uma vez presentes, essas comunidades microbianas parecem influenciar diretamente a instalação e a progressão da endometriose, promovendo um ambiente inflamatório crônico e alterando a resposta imune local. A presença de bactérias como *Shigella* e vírus como

o HPV em lesões endometrióticas sugere um possível papel etiológico ou cofator dessas infecções. Além disso, fatores externos como dieta, atividade física e o uso de suplementos (por exemplo, berberina, NAC, zinco) são apontados como moduladores do microbioma, com potencial impacto na doença. Os autores ressaltam que, embora ainda haja lacunas importantes a serem esclarecidas, a modulação da microbiota peritoneal pode representar um novo eixo terapêutico e preventivo, tornando-se um componente promissor na abordagem clínica da endometriose (AMRO et al., 2022), corroborando com na patogênese da doença, reforçando seu caráter multifatorial e poligênico (KENNEDY et al., 2005).

As teorias que buscam explicar a fisiopatologia da endometriose podem ser categorizadas em duas vertentes principais: aquelas que atribuem a origem do tecido endometrial ectópico ao próprio endométrio uterino e aquelas que propõem uma origem não uterina para essas implantações. A implantação metastática decorrente do fluxo menstrual retrógrado, a persistência de resquícios de tecido embrionário mülleriano e a metaplasia de células celômicas peritoneais são as hipóteses mais estudadas sobre a etiologia da endometriose (BURNEY et al., 2012). A interação entre essas vias inclui processos como a produção local de estrogênio, resistência à progesterona, inflamação crônica e desregulação imunológica, que complicam ainda mais o desenvolvimento e a manutenção das lesões endometrióticas. Esses fatores atuam juntos para promover a adesão celular, a proliferação, a vascularização e a inervação, levando à manifestação de vários tipos de endometriose. Essa intrincada relação enfatiza a necessidade de uma abordagem multifacetada para compreender e tratar a endometriose (ZONDERVAN et al., 2020), como o demonstrado na figura abaixo:

Figura 1. Caminhos potenciais para a patogênese e características fisiopatológicas da endometriose pélvica.



Lesões endometrióticas superficiais e profundas são estabelecidas e mantidas por meio de mecanismos moleculares interativos que promovem a adesão e proliferação celular, esteroidogênese sistêmica e localizada, resposta inflamatória localizada e desregulação imunológica, além de vascularização e inervação. A seta tracejada indica um efeito postulado. ER, receptor de estrogênio; HSD17β2, 17β-hidroxiesteróide desidrogenase 2; ICAM, molécula de adesão intercelular IGF fator de crescimento semelhante à insulina, NF-κB, fator nuclear κB, NGF, fator de crescimento nervoso, PR, receptor de

progesterona, SF1, fator esteróideogênico, STAR, proteína reguladora aguda esteróideogênica, TNF, fator de necrose tumoral e VEGF, fator de crescimento endotelial vascular (adaptada de Zondervan et al., 2020).

Para a ocorrência e manutenção da endometriose, é necessário que uma série de eventos biológicos ocorram, como o fenômeno de escape do sistema imunológico. Conforme descrito por Lebovic, Mueller e Taylor (2001), acredita-se que o estabelecimento da doença depende da capacidade das células endometriais ectópicas de neutralizar a resposta imune. Outros mecanismos são frequentemente descritos, como o processo de adesão celular mediado por moléculas de adesão, incluindo integrinas e caderinas, que desempenham um papel crucial na fixação inicial das células endometriais (BÉLIARD et al., 1997). Adicionalmente, o processo de invasão celular regulado por metaloproteinases da matriz (MMPs) e seus inibidores teciduais específicos (TIMPs), enzimas envolvidas na remodelação tecidual têm sido envolvidos na fisiopatologia da endometriose (ROGERS et al., 2009). outra característica importante é a redução na taxa de apoptose celular em mulheres com endometriose, particularmente durante as fases secretórias/menstruais tardias e proliferativas iniciais do ciclo menstrual, sugerindo que essas células podem manter atividades fisiológicas fora de seus padrões normais (DMOWSKI et al., 2001). A neovascularização do ambiente peritoneal é outro fator importante para manutenção da endometriose, que é descrito como altamente angiogênico e é observado um aumento na atividade e nas quantidades Fator de Crescimento Endotelial Vascular no fluido peritoneal e no endométrio de mulheres com endometriose (DONNEZ et al., 1998).

Embora as lesões de endometriose sejam formadas por glândulas e estroma histologicamente semelhantes ao endométrio eutópico, a presença de tecido ectópico não implica, obrigatoriamente, em um estado patológico universal, sendo necessário considerar aspectos contextuais e funcionais na sua avaliação. No entanto, diversas evidências indicam que ele apresenta alterações moleculares e funcionais distintas do endométrio intrauterino, contribuindo para o desenvolvimento e os sintomas da doença. Dentre essas alterações, destacam-se a resistência à progesterona, a hiperativação de vias inflamatórias, o aumento na produção de citocinas e prostaglandinas, além da capacidade de evadir a apoptose, estimular a angiogênese e invadir tecidos adjacentes (BURNEY; GIUDICE, 2012; NAP et al., 2004).



Adicionalmente, o microambiente imune associado ao tecido ectópico é disfuncional, favorecendo um estado inflamatório crônico (LEBOVIC et al., 2001). Vale ressaltar que existe heterogeneidade entre os focos de endometriose, pois lesões antigas e fibrosas podem exibir reduzida atividade metabólica ou mesmo inatividade, enquanto outras mantêm intensa atividade biológica (VERCELLINI et al., 2014). Portanto, embora nem sempre apresente características patológicas evidentes, o tecido ectópico está geralmente inserido em um contexto de desequilíbrio biológico que sustenta a endometriose como uma condição complexa e multifatorial.

Um dos mecanismos fundamentais na patogênese da endometriose é a proliferação celular nos focos ectópicos. Evidências demonstram que a superepressão da aromatase, enzima responsável pela biossíntese de estrogênio, promove crescimento tecidual independente da estimulação ovariana (DE ABREU et al., 2006). Esse microambiente hormonal alterado caracteriza-se não apenas pela produção ectópica de estrogênio mediada pela aromatase, mas também pela redução da 17 $\beta$ -hidroxiesteroide desidrogenase tipo 2 (17 $\beta$ -HSD2), resultando em acúmulo de estradiol e manutenção do estado inflamatório crônico (BURNEY; GIUDICE, 2012).

### **3.2. Sintomatologia e diagnóstico**

As manifestações clínicas das mulheres com endometriose tendem a ser bastante variáveis, entretanto a doença costuma estar associada a dois sintomas primordiais: dor pélvica e infertilidade (MARQUI, 2014; FREITAS et al., 2011). A dor pélvica pode ser não menstrual, de característica acíclica que pode evoluir para presença contínua, mas também pode ser relacionada com a menstruação (dismenorreia) ou com relações sexuais (dispareunia). De acordo com um estudo de Bellelis et al. (2010), que avaliou aspectos clínicos de 892 mulheres, a dor pélvica crônica é o sintoma mais prevalente presente em 56,8% das mulheres, seguido pela dispareunia com 54,7%, e a infertilidade com 40% da. Uma menor parcela das sortadoras foram assintomáticas, representando cerca de 3 a 22% do total.

O diagnóstico atual da endometriose utiliza uma conduta múltipla que combina avaliação clínica criteriosa com exames de imagem especializados, particularmente a

ultrassonografia transvaginal com preparo intestinal e/ou a ressonância magnética pélvica (GUERRIERO et al., 2021; BAZOT et al., 2017). A laparoscopia com biópsia, antes considerada obrigatória para confirmação diagnóstica, é atualmente recomendada apenas em casos específicos de dúvida diagnóstica ou falha de tratamento inicial, pois a associação de sintomas característicos com achados de imagem permite estabelecer o diagnóstico com segurança (JOHNSON et al., 2017). As principais diretrizes clínicas, incluindo NICE (2017) e ACOG (2010), recomendam iniciar tratamento sem confirmação cirúrgica quando há suspeita clínica, visando evitar atrasos no manejo. Exames laboratoriais tais como a dosagem de CA-125 (cancer antigen 125), CCR-1 (chemokine receptors 1) e interleucina-6 podem ser solicitados, visto que altas dosagens dessas substâncias possam sugerir o diagnóstico da endometriose, entretanto, esses exames não são suficientes para a obtenção de uma confirmação definitiva, nenhum apresentou até o momento sensibilidade e especificidade adequadas para uso clínico isolado (VODOLAZKAIA et al., 2012; ARAÚJO et al., 2020). Dessa forma, o diagnóstico atual da endometriose baseia-se na integração de dados clínicos, exames de imagem e resposta terapêutica, representando uma abordagem mais acessível, abrangente e menos invasiva (ZONDERVAN; BECKER; MISSMER, 2020).

### **3.3. Dor e infertilidade na endometriose**

Sabe-se que a dor é comumente relacionada à endometriose e possui característica crônica, que afetando significativamente a qualidade de vida das mulheres com endometriose. Um fator importante do diagnóstico de endometriose é a realização de anamnese e exame físico que detecte a dor pélvica. Deve-se questionar, durante a anamnese, os aspectos da dor, relacionados à localização, fatores de melhora e piora, duração, intensidade e a relação com o ciclo menstrual. Durante o exame físico, uma atenção especial deve ser dada à marcha, posturas antálgicas e faces com características de dor, observando alterações durante a palpação (CALDEIRA et al., 2008). Teoriza-se que o sintoma resulta da interação entre processos inflamatórios, hormonais, neurogênicos e psicológicos. O tecido ectópico promove uma resposta inflamatória local persistente, caracterizada pela liberação de citocinas pró-inflamatórias, como IL-1 $\beta$ , IL-6, TNF- $\alpha$ , e de prostaglandinas (PGE2), que sensibilizam terminações nervosas periféricas e aumentam a percepção dolorosa

(BARCENA DE ARELLANO et al., 2013).

As lesões endometrióticas apresentam elevada densidade de fibras nervosas sensoriais e simpáticas, o que contribui para a hipersensibilidade local (TOKUSHIGE et al., 2006). A resistência à progesterona e o hiperestrogenismo local perpetuam a inflamação e estimulam a produção de mediadores pró-dor, como o fator de crescimento nervoso (NGF), reforçando a neurogênese e a dor (BULUN et al., 2012). Paralelamente, ocorre sensibilização central, um processo no qual estímulos dolorosos contínuos promovem alterações neuroplásticas nos circuitos da dor na medula espinhal e no cérebro, levando à cronificação da dor mesmo na ausência de estímulo periférico direto (MOROTTI et al., 2017). Adicionalmente, os aspectos psicossociais, como ansiedade e depressão, e a presença de comorbidades dolorosas, como síndrome do intestino irritável e fibromialgia, também desempenham papel modulador importante na dor crônica vivenciada por muitas mulheres (VERCELLINI et al., 2014; TAYLOR et al., 2021), mostrando necessidade de uma abordagem multidisciplinar e personalizada no manejo da dor endometriótica. Estudos sugerem que, devido ao estado de inflamação crônica, ocorrem remanejamentos no cérebro, como reduções no volume de áreas envolvidas no processamento emocional e sensorial, evidenciando uma remodelação cerebral significativa. Além disso, há indicações de que a inflamação sistêmica pode levar à sensibilização central, contribuindo para a dor neuropática crônica e outros sintomas neurológicos. Essas alterações neurológicas representam uma dimensão importante da doença, destacando que a endometriose não é apenas uma condição ginecológica, mas também possui um componente neurológico que deve ser considerado no manejo clínico abrangente (TAYLOR et al., 2021).

No que se refere a infertilidade, evidências apontam a ocorrência da diminuição da taxa de fertilização, implantação e gestação em portadoras de endometriose (BARNHART et al., 2002; AL-FADHLI et al., 2006). A infertilidade está relacionada à sintomatologia da endometriose e ocorre em 30% a 50% das mulheres com a doença (HOLOCH; LESSEY, 2010). A condição de infertilidade associada à endometriose é decorrente de múltiplos fatores, pois resulta de um conjunto de mecanismos fisiopatológicos que comprometem tanto a fertilização quanto a implantação embrionária e a manutenção da gestação. Um dos principais fatores implicados é o

ambiente inflamatório crônico do peritônio, com níveis elevados de citocinas e prostaglandinas que afetam negativamente a função tubária, a qualidade dos gametas e a receptividade endometrial (BERKKANOGLU; ARICI, 2003).

As alterações anatômicas provocadas por lesões infiltrativas e aderências pélvicas contribuem para a infertilidade associada a endometriose devido a alterações na anatomia reprodutiva, dificultando a captação do oócito, o que dificulta o sucesso dos tratamentos de reprodução assistida (DUNSELMAN et al., 2014; CARPINELLO, 2017). Os endometriomas ovarianos, por sua vez, estão associados à diminuição da qualidade oocitária, provavelmente devido ao estresse oxidativo e à presença de ferro no conteúdo cístico (BENAGLIA et al., 2014). No útero, o endométrio de mulheres acometidas pela doença apresenta expressão reduzida de genes fundamentais para a implantação, como *HOXA10* e *LIF* podendo levar a falhas na receptividade endometrial (AGHAJANOVA et al., 2009). Finalmente, disfunções imunológicas, como alterações em células NK e presença de autoanticorpos, também têm sido associadas à falha de implantação e maior risco de infertilidade (ZHANG et al., 2018).

### **3.4. Expressão gênica na endometriose**

A expressão gênica aberrante do endométrio eutópico está relacionada com a falha da implantação embrionária em mulheres com endometriose, dentre eles os receptores de progesterona, alguns fatores de transcrição, enzimas que estão envolvidas no metabolismo de esteróides, entre outros (WEI et al., 2009). Alterações da expressão gênica durante a janela de implantação, processos altamente controlados para garantir a interação entre o endométrio e o embrião, são evidenciados em mulheres com endometriose (GIUDICE et al., 2002). Matsuzaki e colaboradores (2005), observaram a regulação positiva de diversos genes, em especial genes envolvidos nas vias de sinalização RAS/RAS/MAPK e PI3K, favorecendo o processo de crescimento, proliferação, diferenciação, apoptose e migração celular, beneficiando assim, o desenvolvimento dos focos endometrióticos fora da cavidade uterina (MATSUZAKI et al., 2005). Adicionalmente, também foi observado a regulação positiva de genes reguladores do ciclo celular no endométrio de mulheres com endometriose, como *CCNA2*, *CCNE1* e *CCNE2* (EYSTER et al.,

2007).

Evidências indicam que alterações nos mecanismos epigenéticos que regulam a expressão gênica, como metilação do DNA, modificações de histonas e regulação por RNAs não codificantes (ncRNAs), modulam a função de genes envolvidos na inflamação, evasão da apoptose e resistência hormonal (DYSON et al., 2014; GRIMSTAD et al., 2017). Um exemplo marcante é a hipermetilação do gene *PR-B*, responsável por codificar o receptor de progesterona tipo B, cuja redução contribui para a resistência à progesterona observada nas lesões (BULUN et al., 2012). Essa modificação epigenética pode levar à diminuição da expressão do receptor PGR-B, prejudicando a resposta do endométrio à progesterona e contribuindo para falhas na implantação embrionária e infertilidade, ainda que o tecido pareça histologicamente normal. (ROCHA-JUNIOR et al. 2019). Essas alterações no epigenoma das células endometriais mediadas por alterações hormonais sustentam a natureza crônica e recidivante da doença, além de justificar a limitada eficácia de terapias hormonais convencionais em alguns casos.

A análise de expressão gênica permite identificar quais genes estão ativos ou inativos em tecidos endometrióticos, comparando-os com tecidos saudáveis. Através de métodos como sequenciamento de RNA (*RNA-seq*) e microarranjos, os pesquisadores podem detectar padrões de expressão gênica associados à inflamação, angiogênese, resistência hormonal e proliferação celular, que são processos-chave na progressão da endometriose (BULUN et al., 2019; BORGHESE et al., 2008), mostrando que a importância dessa análise reside na sua capacidade de identificar potenciais biomarcadores diagnósticos e alvos terapêuticos. A desregulação de genes como *APLNR*, *CCL21* e *FCGR2A* tem sido associada infiltração de células imunológicas, angiogênese excessiva e invasão tecidual, características marcantes da endometriose (PENG et al., 2021). Os perfis de expressão gênica podem ajudar na estratificação com base na gravidade da doença ou na resposta ao tratamento, tornando possível uma abordagem mais personalizada. A análise de expressão gênica também abre caminho para o desenvolvimento de terapias direcionadas, como inibidores de vias inflamatórias ou moduladores epigenéticos, que podem melhorar os resultados clínicos (ZONDERVAN et al., 2020).

Identificar o perfil de expressão transcricional na endometriose é de fundamental importância para a elucidação da etiologia da doença, bem como para compreensão do sentido biológico dos genes diferencialmente expressos e sua relação com a endometriose superficial, ovariana e infiltrativa. Os RNAs codificantes e não codificantes, podem representar uma estratégia de diagnóstico precoce e minimamente invasiva, de forma a melhorar o entendimento sobre a doença e minimizar os sintomas mais agressivos oriundos do longo período até o diagnóstico.

### **3.5. Ferramentas de bioinformática para análise de expressão gênica**

A análise de dados públicos por meio de ferramentas de bioinformática têm um impacto profundo na compreensão da fisiopatologia de doenças com a inclusão de participantes de diferentes etnias, de baixo custo, permitindo a identificação de biomarcadores, alvos terapêuticos e mecanismos moleculares subjacentes a doenças complexas, através da integração e interpretação de grandes volumes de dados genômicos, transcriptômicos e epigenômicos (HASIN et al., 2017; REUTER et al., 2015). Estudos transcriptômicos têm elucidado vias críticas envolvidas em câncer (ZHANG et al., 2021), doenças autoimunes (MÁRQUEZ et al., 2016) e desordens reprodutivas, como a endometriose (ZONDERVAN et al., 2020; AGHAJANOVA et al., 2017). A associação de dados transcriptômicos com outras camadas de informação, como genômica e proteômica (multi-omics), promete avanços significativos na medicina personalizada e no desenvolvimento de terapias direcionadas (SUBRAMANIAN et al., 2020; MISRA et al., 2019). A bioinformática desempenha um papel central na análise de dados transcriptômicos, fornecendo as ferramentas e metodologias necessárias para transformar dados brutos em insights biológicos significativos (CONESA et al., 2016; STARK et al., 2019). Tanto a tecnologia de *microarranjo* quanto o sequenciamento de RNA em larga escala (*RNA-seq*) geram grandes volumes de dados, cuja interpretação exige abordagens computacionais robustas (WANG et al., 2009; STARK et al., 2019).

### 3.5.1. *Microarranjo e RNA-seq*

As análises de dados de *microarranjo* e *RNA-seq* vem sendo amplamente utilizadas na pesquisa biomédica para compreender o perfil de expressão gênica envolvidos na fisiopatologia das doenças. No caso da endometriose para explorar perfis de expressão gênica e identificar biomarcadores e vias moleculares envolvidas em desfechos específicos da doença, como a dor pélvica crônica e a infertilidade (BORGHESE et al., 2008). Essas técnicas permitem a comparação entre amostras de tecido endometrial de mulheres com e sem endometriose, ajudando a identificar GDEs que podem estar associados ao desenvolvimento e à progressão da doença (AGHAJANOVA et al., 2017; FASSBENDER et al., 2015).

O *microarranjo* é uma técnica de hibridização que permite medir a expressão gênica em larga escala, possibilitando a detecção de regiões específicas do genoma previamente definidas por sondas e sua expressão será mensurada por fluorescência emitida da ligação dos mRNAs presentes na amostras aos genes/sondas conhecidos presentes no chip. Já a técnica de *RNA-seq*, todo o material genético presente na amostra é amplificado, podendo ser proveniente de um tecido (*RNA-seq*) ou de um grupo de célula selecionadas (*single-cell*), obtendo um grande volume de informações (GONÇALVES et al., 2015). A análise dos dados gerados requer etapas complexas de processamento e normalização para remoção de vieses (IRIZARRY et al., 2003; GAUTIER et al., 2004).

Ferramentas de bioinformática como o ambiente de software livre R são amplamente utilizadas nas análises de dados ômicos, especialmente quando associadas a outras ferramentas, como o Bioconductor, um repositório de mais de 2.000 pacotes de programação R para análises genômicas (GENTLEMAN et al., 2004; HUBER et al., 2015). Dentre os pacotes presentes no Bioconductor, o limma é muito usado para realizar análise diferencial de expressão gênica em *microarranjos* e *RNA-seq* (RITCHIE et al., 2015; LAW et al., 2016). O pacote affy é usado no pré-processamento de dados de *microarranjos* da plataforma Affymetrix, facilitando a normalização, controle de qualidade e identificação de genes diferencialmente expressos (GAUTIER et al., 2004; CARVALHO et al., 2007). A normalização é particularmente crítica para corrigir vieses técnicos, como variações na hibridização e

intensidade de fluorescência. Métodos como RMA (Robust Multi-array Average) (IRIZARRY et al., 2003) e GCRMA (WU et al., 2004) são amplamente utilizados para garantir a comparabilidade entre amostras (TT, 2016; ATHAR et al., 2019).

O RNA-seq é uma tecnologia de sequenciamento de nova geração (NGS) que revolucionou a análise da expressão gênica por sua alta resolução e sensibilidade (WANG et al., 2009; STARK et al., 2019). Esta técnica permite não apenas a quantificação precisa dos níveis de RNA, mas também a detecção de novas isoformas de genes e eventos de splicing alternativo, que são frequentemente desregulados em doenças complexas como a endometriose (DYSON et al., 2014). Diferente das análises de *microarranjo*, o *RNA-seq* permite a identificação de RNAs sem conhecimento prévio das regiões de interesse, apesar de painéis específicos de microRNAs, longos RNAs não codificantes, painéis serem empregados nos mais diversos campos de investigação. Estudos utilizando abordagens de *RNA-seq* têm elucidado perfis de expressão em diferentes subtipos de tecido endometrial e estágios da doença, identificando genes e vias moleculares cruciais para a patogênese (BORGHESE et al., 2008).

A análise de dados de *RNA-seq* envolve um pipeline bioinformático complexo e multi-etapas (CONESA et al., 2016; STARK et al., 2019). Ferramentas como o FastQC são essenciais para avaliação da qualidade das *reads* brutas, identificando problemas como baixa qualidade de bases ou contaminação por adaptadores (BROWN et al., 2017). Alinhadores como STAR (DOBIN et al., 2013) e HISAT2 (KIM et al., 2019) são amplamente utilizados para mapeamento eficiente contra genomas de referência, enquanto programas como featureCounts (LIAO et al., 2014) e HTSeq (ANDERS et al., 2015) realizam a quantificação da expressão gênica. Para análise diferencial de expressão gênica, pacotes estatísticos especializados como edgeR (ROBINSON et al., 2010; MCCARTHY et al., 2012) e DESeq2 (LOVE et al., 2014) empregam modelos estatísticos avançados que consideram a variabilidade biológica e técnica inerente aos dados de *RNA-seq*. Estes métodos têm sido particularmente úteis na identificação de biomarcadores potenciais para endometriose (AGHAJANOVA et al., 2017; FASSBENDER et al., 2015).



### 3.5.2. Metanálise

A metanálise é uma abordagem estatística poderosa que combina resultados de múltiplos estudos para identificar padrões consistentes de expressão gênica associados a condições biológicas específicas, como doenças ou tratamentos (HONG; BREITLING, 2008). Essa técnica permite aumentar o poder estatístico e a confiabilidade das descobertas, especialmente quando estudos individuais têm tamanhos amostrais pequenos ou apresentam heterogeneidade técnica e biológica (CAHAN et al., 2007; HAYNES et al., 2017).

A metanálise de GDEs geralmente envolve a integração de dados de *microarranjo* ou/e *RNA-seq* de diferentes plataformas, seguida da normalização cruzada e combinação de estatísticas de expressão gênica, como valores de  $p$  ou *fold-changes* (TAMINAU et al., 2012). Estudos metodológicos recentes destacam a importância de abordagens robustas para corrigir vieses técnicos e biológicos, incluindo diferenças em protocolos experimentais, efeitos de lote e heterogeneidade entre populações (RAMASAMY et al., 2008; TSENG et al., 2012; LUO et al., 2021). Métodos avançados como o REM (random-effects model) e a ponderação por qualidade de estudo têm sido amplamente adotados para lidar com essa variabilidade (CHOI et al., 2017).

No contexto de ferramentas computacionais especializadas, o metavolcanoR se destaca para visualização de resultados de Metanálises em estudos de expressão gênica (CADORE et al., 2024). Este pacote permite a criação de gráficos volcano plot para Metanálises, integrando simultaneamente a significância estatística (valores de  $p$  ajustados) com a magnitude das mudanças (*fold-changes*) em estudos combinados. Tais visualizações facilitam a identificação de assinaturas transcricionais robustas e a validação cruzada de resultados entre coortes independentes, aumentando substancialmente a confiança nas descobertas (ZHOU et al., 2019).

Apesar dos avanços recentes na pesquisa genômica, a expressão gênica associada à endometriose ainda é pouco explorada na literatura. Uma busca sistemática revelou apenas três estudos que atenderam aos critérios de inclusão estabelecidos. Kang et al. (2025) conduziram uma metanálise para correlacionar a fisiopatologia da endometriose com perfis de expressão gênica, utilizando

exclusivamente dados de microarranjo (três *datasets* públicos). Embora o estudo tenha identificado vias moleculares potencialmente envolvidas, a limitação ao *microarranjo* e o pequeno número de conjuntos de dados podem ter restringido a generalização dos resultados.

Em outra abordagem, Poli-Neto et al. (2020) investigaram o perfil imunológico do endométrio eutópico em diferentes estádios da doença, mas também se limitaram a quatro *datasets* de *microarranjo*. Apesar de revelarem assinaturas transcricionais relacionadas à resposta imune, a ausência de dados de *RNA-seq* ou validação funcional reduz a robustez das conclusões. O único estudo que integrou dados de *RNA-seq* e *microarranjo* (VARGAS et al., 2022) focou apenas nos genes relacionados às diferenças transcricionais no endométrio durante a janela de implantação em mulheres com e sem endometriose. Essa abordagem amplia o espectro de análise, a heterogeneidade metodológica e o tamanho amostral, permitindo a identificação de um grande número de transcritos. A ausência de estudos na literatura destaca uma escassez de evidências consolidadas, reforçando a necessidade da realização de um estudo mais robusto acerca do perfil transcricional da endometriose, objetivo do presente estudo.

### 3.5.3. Enriquecimento Funcional de Vias biológicas

O enriquecimento funcional de vias biológicas é uma abordagem computacional essencial para interpretar listas de genes diferencialmente expressos, permitindo a identificação de vias biológicas, processos celulares e funções moleculares estatisticamente significativas (SUBRAMANIAN et al., 2005). Entre as ferramentas disponíveis, o EnrichR destaca-se por sua ampla adoção na comunidade científica, oferecendo uma interface integrada com mais de 180 bases de dados de anotação, incluindo *Gene Ontology* (GO), KEGG, Reactome e BioPlanet (CHEN et al., 2013; KULESHOV et al., 2016). O EnrichR emprega métodos estatísticos robustos, como o teste exato de Fisher, para determinar o enriquecimento de termos biológicos, além de fornecer visualizações intuitivas, como gráficos de barras e redes de interação (KULESHOV et al., 2016). Estudos recentes demonstram sua aplicabilidade em pesquisas genômicas, incluindo a identificação de vias desreguladas em doenças complexas (e.g., câncer e doenças neurodegenerativas) e a descoberta de potenciais

alvos terapêuticos (XIE et al., 2021). Além disso, sua integração com linguagens de programação, como R e Python, por meio de bibliotecas dedicadas (e.g., enrichR em R), facilita a automação de análises de enriquecimento em larga escala, consolidando-o como uma importante ferramenta para estudos de biologia de sistemas (KULESHOV et al., 2016; XIE et al., 2021).

## 4. Materiais e métodos

### 4.1. Análise de transcriptoma em bancos de dados

#### 4.1.1. Obtenção de estudos transcriptômicos relacionados à Endometriose.

Para a identificação dos estudos de transcriptômica foi realizada uma curadoria para selecionar estudos que incluíram amostras de endométrio eutópico e/ou ectópico. Os conjuntos de dados de *microarranjo* e *RNA-seq* foram coletados do repositório GEO, e selecionados mediante aos critérios de inclusão e exclusão definidos na tabela 1.

Tabela 1. Critérios de inclusão e exclusão dos estudos selecionados.

Critérios de inclusão	Critérios de exclusão
Estudos relacionados à endometriose;	Estudos baseados exclusivamente em modelos animais ou linhas celulares ( <i>in vitro</i> )
Organismo <i>Homo sapiens</i> ;	Estudos sem grupo controle (endométrio eutópico ou tecido saudável equivalente)
Tecnologia de <i>Microarranjo</i> ou <i>RNA-seq</i> ;	Estudos com dados incompletos ou sem acesso público
Estudos <i>in vivo</i> de amostras do endométrio de pacientes	Estudos com menos de 3 amostras por grupo experimental
Apresentar no mínimo três replicatas para cada grupo;	
Fornecer as tabelas de contagens brutas.	

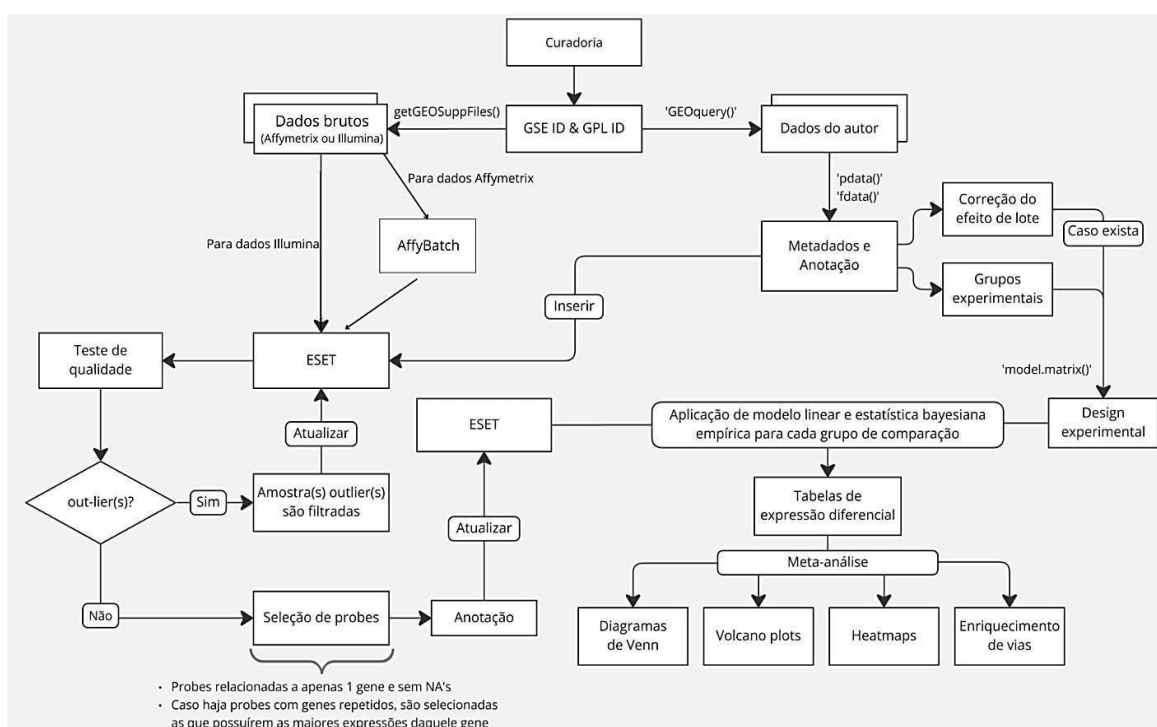
#### 4.1.2. Identificação de assinatura gênica da endometriose através de análises de transcriptômas.

Diferentes abordagens foram realizadas para a obtenção dos valores de expressão referentes a cada tecnologia.

#### 4.1.2.1. Assinatura gênica da endometriose através de análises transcriptômicas de *Microarranjo*.

Todos os *datasets* foram analisados na linguagem de programação R por meio dos dados brutos, onde foi feita uma nova normalização e a comparação com a normalização feita pelo autor do estudo. As etapas foram descritas conforme o indicado na figura 2.

Figura 2. Fluxograma seguido nos scripts de análises dos *datasets* da técnica de *microarray*.



Para os dados de *microarranjo*, os valores de expressão bruta fornecidos pelos autores foram obtidos através do pacote *GeoQuery* (DAVIS AND MELTZER 2007), os metadados de cada estudo foram obtidos de acordo com as informações cadastradas pelos autores de cada amostra no GEO. Para acessar e baixar esses dados, utiliza-se a função *getGEOSuppFiles()*, que permite a obtenção de dados de plataformas. Para os dados dos estudos que utilizaram a plataforma *Affymetrix*, os dados brutos foram obtidos diretamente do GEO, com o uso do pacote *GEOQuery*, baixando arquivos *.CEL* extraídos e organizados em diretórios específicos, e foi normalizado através do pacote *affy* e então convertido em um *ExpressionSet* (ESET), função do pacote *limma*, que contém as informações de expressão gênica normalizadas e

anotadas. No caso de estudos provenientes da Illumina, os dados são obtidos também por meio do pacote *GEOQuery*, porém, a normalização dos dados é feita em com uso da função *necq*, e posteriormente é convertido em *ESET*, ambos do pacote *limma* sem a necessidade de um pacote intermediário para formatação.

O controle de qualidade foi realizado avaliando os dados com o pacote *ArrayQualityMetrics* (KAUFFMANN et al., 2009), gerando um relatório HTML interativo, permitindo a detecção de amostras problemáticas através da análise de métricas como intensidade de sinal, correlação entre réplicas, proporção de sondas presentes/ausentes e distribuição de intensidades por meio de gráficos, e a necessidade de normalização ou correção de artefatos. Este último envolveu a normalização via RMA (Robust Multi-array Average), que incluiu correção de background, normalização e conversão dos dados para escala de log base 2. Gráficos como boxplots, densidades e MDS (Multidimensional Scaling) foram utilizados para inspeção visual da qualidade dos dados, identificando possíveis outliers que poderiam distorcer os resultados. Para reduzir redundâncias, a sonda mais expressiva de cada gene foi selecionada, resultando em uma matriz de expressão filtrada e mais robusta.

Os metadados das amostras, contendo informações como grupos experimentais e características fenotípicas, foram organizados em uma tabela de acordo com o interesse, enquanto as anotações das sondas (*probes*) são carregadas a partir de um arquivo complementar ou diretamente do objeto “*gse*”. Esta etapa envolve uma rigorosa filtragem para a remoção de sondas que estão associadas a múltiplos genes, sem anotação (células vazias) ou aquelas com valores descritos como “NA”, que simboliza um valor omissa na estatística. Esta etapa é feita como forma de garantir a qualidade das análises subsequentes. Em seguida, o próximo passo é definir os grupos experimentais que serão comparados. Isso é feito selecionando uma coluna específica dos metadados, e ajustando os rótulos para refletir claramente as condições biológicas de interesse.

Após o tratamento dos dados, ambos os modelos de estudos seguem a mesma análise de expressão diferencial, que é construída a partir de uma matriz de design experimental que especifica os contrastes de interesse, como comparações entre grupos de tratamento e controle. Um modelo linear é ajustado para cada gene

utilizando `lmFit`, seguido por um processo de suavização empírica de Bayes (`eBayes`) que melhora as estimativas de variância, especialmente importante para experimentos com poucas réplicas. Os GDEs são identificados com base em limiares estatísticos rigorosos (ajustados para múltiplos testes) e magnitude de efeito ( $\text{adj.P.Val} < 0.05$  e  $|\log\text{FC}| \geq 0.6$ ). Os resultados são sistematicamente organizados em tabelas detalhadas contendo valores de  $\log\text{FC}$ , p-valor ajustados e intervalos de confiança, que são exportados para arquivos tabulares.

A etapa de visualização inclui a geração de um “volcano plot”, utilizando pacote `ggplot2` (WICKHAM et al., 2016), que permite a rápida identificação visual dos genes mais significativos e com maiores magnitudes de alteração. Os top 10 genes regulados positivamente e negativamente são destacados no gráfico e rotulados com seus símbolos, facilitando a interpretação biológica. Um *heatmap* hierárquico é criado para mostrar os padrões de expressão dos genes mais relevantes, agrupando automaticamente as amostras por similaridade e condições experimentais. Adicionalmente, diagramas de Venn são utilizados para comparar os conjuntos de genes diferencialmente expressos entre diferentes contrastes, revelando sobreposições e diferenças importantes entre as condições testadas.

#### **4.1.2.2. Metanálise entre os resultados obtidos nas análises de Microarranjo.**

Após as análises de expressão gênica diferencial conduzidas com os dados de *microarranjo*, foi realizada a metanálise a partir das listas de tabelas contendo os genes, valores de  $\log$  Fold Change ( $\log\text{FC}$ ) e p-value gerados de cada *dataset* selecionado. O pacote `MetaVolcanoR` do R foi empregado com o objetivo de identificar genes expressos em consenso entre os conjuntos de dados analisados (ZHOU et al., 2019). Em seguida, a função `voteCount_mv` foi aplicada para identificar genes consistentemente regulados, adotando como critérios de significância um p-value  $\leq 0,05$  e um  $|\log\text{FC}| \geq 0,6$  equivalente a fold-change de 1,5). Essa função contabilizou quantos estudos classificaram cada gene com regulação positiva ( $\log\text{FC}$  positivo) ou regulação negativa ( $\log\text{FC}$  negativo), gerando um *dataframe* com as contagens de concordância. A partir da lista dos GDEs em comum, foram selecionados os dez principais genes com regulação positiva e regulação negativa respectivamente para

que estejam evidenciados na visualização dos gráficos produzidos, tais como o *volcano plot* e o *heatmap*.

#### **4.1.2.3. Análise de enriquecimento funcional de vias biológicas**

Para compreender o significado biológico dos genes diferencialmente expressos, foram realizadas análises de enriquecimento funcional utilizando a ferramenta EnrichR (CHEN et al., 2013) na versão online através do site <https://maayanlab.cloud/Enrichr/>, onde foi inserida a lista dos genes diferencialmente expressos identificados em consenso através da metanálise separadamente entre os genes regulados positivamente e negativamente. As bases de dados selecionadas para a verificação de processos biológicas, vias de sinalização, metabólicos e funções moleculares foram GO Biological Process, KEGG Pathways, Reactome Pathways, Wikipathway e MsigDB\_Hallmark, pois a combinação dos dados obtidos entre essas bases de dados fornecem uma visão mais completa, detalhada e robusta, contextualizando o conhecimento biológico acerca dos resultados encontrados.

As vias biológicas consideradas relevantes foram selecionadas com base em três critérios principais: (1) significância estatística, considerando os termos com resultados de valores-p ajustados para falsas descobertas (do inglês, False Discovery Rate, FDR) inferiores a 0,05; (2) sobreposição de genes (overlap), exigindo a presença de pelo menos três genes em comum entre a lista de genes de interesse e os conjuntos de genes das vias biológicas presentes nas bases de dados analisadas; e (3) relevância biológica, priorizando resultados que estivessem alinhados com o conhecimento prévio do sistema biológico em estudo, particularmente no contexto da endometriose.

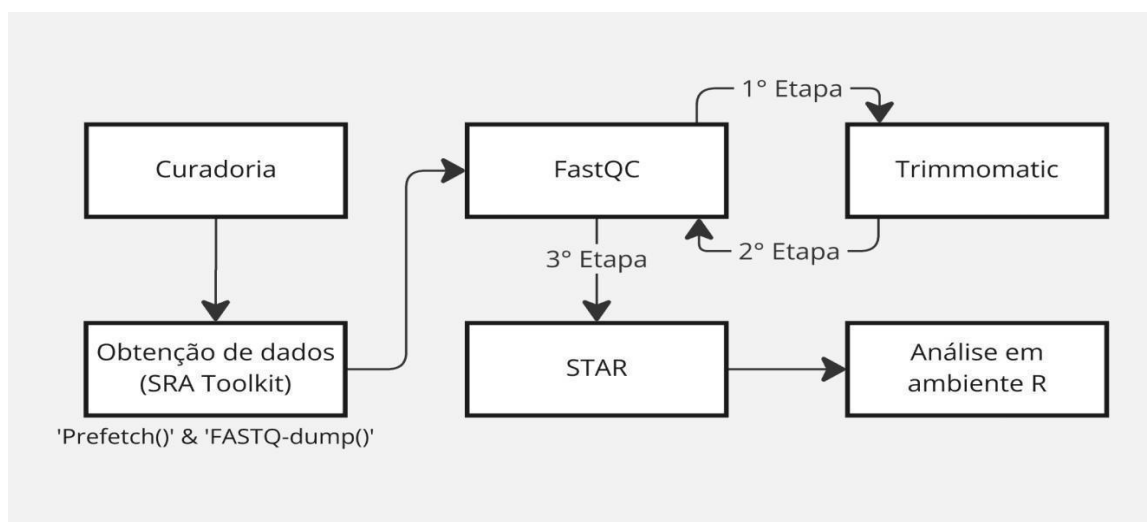
#### **4.1.2.4. Validação da assinatura gênica da endometriose através de análises transcriptômicas RNA-seq**

Assim como nos dados de *microarranjo*, os *datasets* de *RNA-seq* foram obtidos no GEO, porém, por se tratar de uma técnica cujo resultado final (*output*) tem um grande volume de dados, o tratamento dessas informações seguiu diferentes abordagens. A primeira etapa da análise foi feita em *prompt* de comando em ambiente



*LINUX*. Já no ambiente R, para a análise de expressão diferencial, o pipeline foi feito a partir do uso do pacote edgeR (ROBINSON et al., 2010), e foram considerados como GDEs os genes que apresentaram valores de P ajustado (adj. p-val)  $< 0,05$  e uma variação na expressão (log2FoldChange) igual ou superior a 1,0 para genes regulados positivamente, ou igual ou inferior a -1,0 para genes regulados negativamente.

Figura 3. Fluxograma seguido nos scripts de análises dos datasets da técnica de RNA-seq. Fonte própria.



O *workflow* descrito na figura 3 apresenta as etapas iniciais das análises para o tratamento dos dados de *RNA-seq*, realizadas na execução do estudo. O processo começa com a curadoria e obtenção dos dados brutos, utilizando a ferramenta *SRA Toolkit* (LEINONEN et al., 2011), a partir de duas etapas principais: (1) o comando *prefetch* foi empregado para download dos arquivos no formato nativo .sra a partir dos identificadores de acesso, utilizando conexão FTP com o banco de dados do NCBI e parâmetros de otimização (--max-size 50G para arquivos grandes e --transport ascp para transferência acelerada quando disponível); (2) a conversão para formato *FASTQ* foi realizada através do comando *fastq-dump* com os parâmetros "--split-files" para separação de reads paired-end, "--gzip" para compressão direta da saída, e "--skip-technical" para remoção de reads técnicas, garantindo arquivos finais no formato "\_R1.fastq.gz" e "\_R2.fastq.gz".

A etapa seguinte envolve a avaliação da qualidade das reads brutas por meio

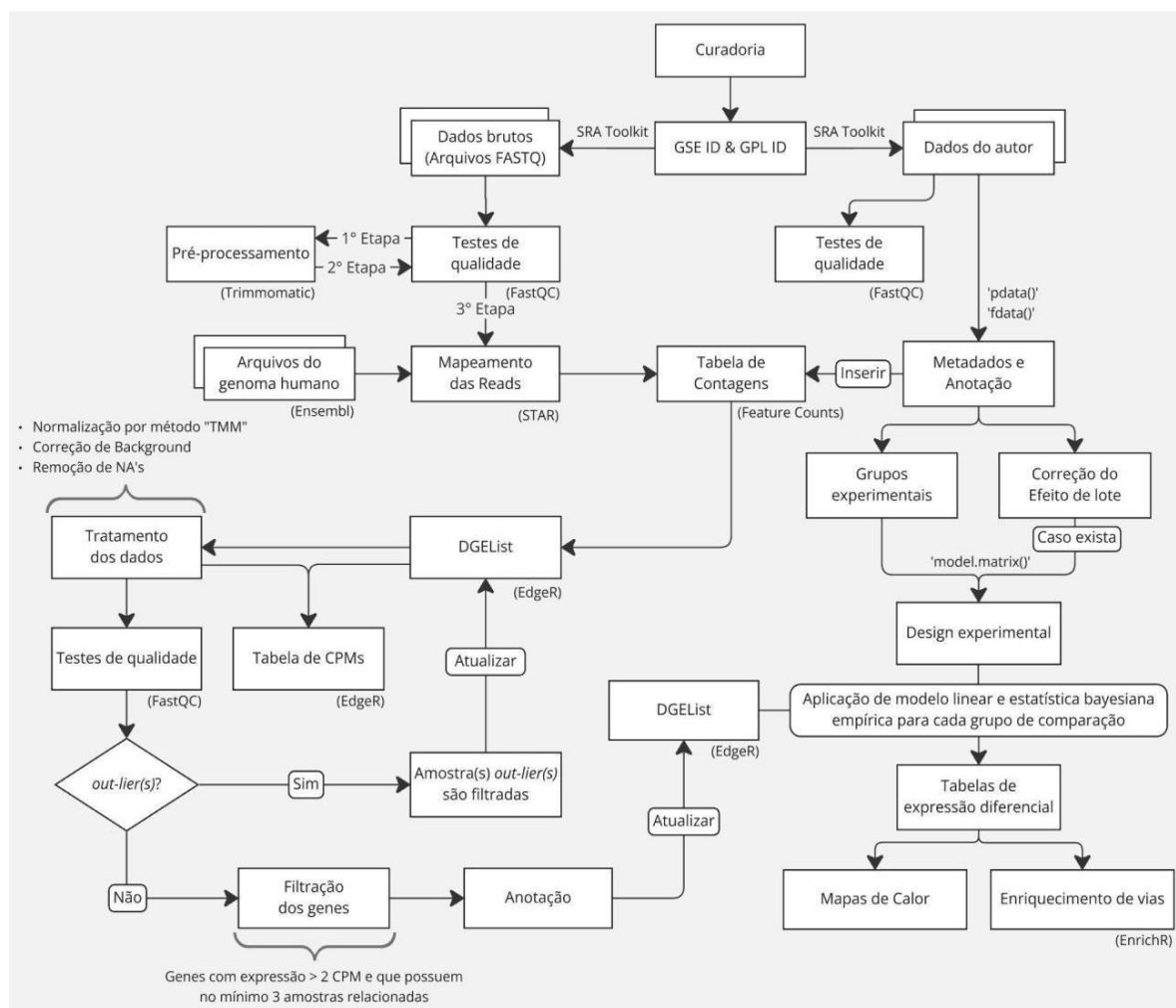
do programa FastQC (ANDREWS, 2010), ferramenta que analisa métricas críticas como: (1) escores de qualidade por base (Phred  $Q \geq 30$  como padrão de corte), (2) distribuição de nucleotídeos (para detecção de viés técnico), (3) conteúdo de GC (comparado com a referência do organismo), (4) presença de adaptadores (limite de tolerância de 5%) e (5) taxa de duplicação. Os relatórios HTML gerados foram interpretados visualmente por meio de um gráfico de *BoxWhisker*, com ênfase no perfil de qualidade ao longo das reads, onde quedas nas extremidades (>20% de bases com  $Q < 30$ ) ou contaminantes (adaptadores >5%) indicaram necessidade de *trimming* (cortar/remover) com a ferramenta Trimmomatic.

O Trimmomatic (BOLGER et al., 2014) foi utilizado como uma etapa de pré-processamento, pois realiza o controle de qualidade dos dados brutos de sequenciamento. O pipeline consistiu em: (1) remoção de adaptadores utilizando o parâmetro ILLUMINACLIP com arquivo de adaptadores específicos; (2) corte de bases de baixa qualidade ( $Q < 20$ ) nas extremidades das reads (LEADING:20, TRAILING:20); (3) filtragem por janela deslizante de 4 bases (SLIDINGWINDOW:4:20), removendo regiões onde a qualidade média fosse inferior a  $Q_{20}$ ; e (4) seleção final de reads com comprimento mínimo sendo considerado sempre com metade do original (MINLEN). No caso de dados Single-End, o output possui um arquivo e Paired-End, o output possui quatro arquivos de saída. O FastQC foi reexecutado para validação, garantindo que os dados atendessem aos critérios mínimos (70% das reads com  $Q \geq 30$  em  $\geq 75\%$  de suas posições) antes do alinhamento. Essa abordagem, baseada em Andrews (2010), é fundamental para evitar artefatos técnicos em análises *downstream* de *RNA-seq*, garantindo a integridade e qualidade dos dados para etapas subsequentes de alinhamento e análise de expressão gênica diferencial.

Após o pré-processamento, a 3ª etapa consiste no alinhamento das reads ao genoma de referência utilizando o STAR (DOBIN et al., 2013), um alinhador otimizado para *RNA-seq* que emprega um algoritmo de suffix arrays comprimidos para mapeamento eficiente contra o genoma de referência. O processo iniciou-se com a construção de um índice genômico personalizado contendo as sequências do genoma (GRCh38.p14). O alinhamento de reads de *RNA-seq* foi realizado utilizando genoma de referência e anotações obtidos do Ensembl, seguindo um pipeline padronizado.

Inicialmente, os arquivos FASTA do genoma primário e GTF com anotações foram baixados diretamente do repositório do Ensembl. O genoma foi então indexado utilizando o STAR com parâmetros otimizados: `--sjdbOverhang 100` (para reads de 101 bp), `--genomeSAindexNbases 14` (para genomas grandes) e inclusão das anotações de splicing via `--sjdbGTFfile`. O alinhamento propriamente dito foi executado em modo duas-passagens (`--twopassMode Basic`) com os seguintes critérios: `--outFilterMultimapNmax 20` para reads multi-mapeáveis, `--outSAMtype BAM SortedByCoordinate` para saída ordenada, e `--quantMode GeneCounts` para geração simultânea de contagens gênicas. A qualidade do alinhamento foi avaliada através de métricas como porcentagem de reads mapeadas (>80%) e concordância com junções anotadas (>85%). Os arquivos BAM (ou SAM quando especificado) resultantes foram processados posteriormente para quantificação de expressão gênica, garantindo rastreabilidade completa desde a origem dos dados de referência até os resultados finais de alinhamento.

Figura 4. Fluxograma seguido nos scripts de análises dos datasets da técnica de bulk RNA-seq.



No ambiente R, os dados são convertidos em um objeto `DGELIST` do pacote `EdgeR` (ROBINSON et al., 2010). Esse objeto é uma estrutura de dados específica para análise de expressão gênica, que facilita o tratamento e a manipulação dos dados, permitindo que os dados sejam organizados de maneira eficiente, incluindo informações como contagens de reads, metadados das amostras e anotações dos genes, feitas com o uso do pacote `Rsubread`, usando a função `FeatureCounts` é utilizado para calcular as contagens de reads associadas a cada gene, enquanto as funções `pdata()` e `fdata()` do `EdgeR` são empregadas para atualizar as informações dos grupos experimentais e das anotações dos genes. A próxima etapa é a filtragem de genes, onde são mantidos apenas aqueles que apresentam expressão superior a 2 CPM (Contagens Por Milhão) e que estão presentes em pelo menos três amostras. Essa filtragem é essencial para eliminar genes de baixa expressão, que podem não

ser biologicamente relevantes ou que podem introduzir ruído na análise. A filtragem ajuda a garantir que apenas genes com expressão significativa sejam considerados nas etapas subsequentes, melhorando a confiabilidade dos resultados.

A correção do efeito de lote é utilizada para diminuir a variabilidade introduzida por experimentos realizados em diferentes lotes ou condições. A correção do efeito de lote é feita utilizando uma abordagem de modelo linear, que ajusta os dados para remover variações técnicas não relacionadas às condições biológicas de interesse. Essa etapa é fundamental para garantir que as diferenças observadas na expressão gênica sejam realmente biológicas e não artefatos técnicos. A análise estatística foi feita por meio de modelos lineares e estatística bayesiana empírica para comparar a expressão gênica entre os grupos experimentais. Esses métodos permitem identificar genes diferencialmente expressos com controle de erros associados a múltiplos testes. Os resultados são organizados em tabelas de expressão diferencial, que listam os genes significativamente regulados, juntamente com suas estatísticas associadas (como *valores-p* e *fold changes*).

Foi realizada a associação entre os resultados obtidos na metanálise dos diferentes datasets de *microarranjo* através do pacote '*MetavolcanoR*', obtendo uma nova tabela de genes em consenso entre as análises feitas em todas as comparações dos *datasets* das técnicas de *microarranjo* e *RNA-seq*. Os resultados foram interpretados e visualizados utilizando ferramentas como heatmaps), que permitem a visualização dos padrões de expressão gênica entre as amostras. Além do diagrama de Venn, com objetivo de demonstrar didaticamente o os genes comuns entre os conjuntos de dados usados nas comparações.

## 5. Resultados

### 5.1. Estudos encontrados por meio da curadoria

A curadoria dos dados foi realizada no site do NCBI, por meio da aba de busca '*GEO Datasets*', utilizando os termos '*endometriosis*' e '*Homo sapiens*'. Foram aplicados filtros específicos para selecionar estudos com as técnicas de interesse: '*expression profiling by array*' (*microarranjo*) e '*expression profiling by high throughput sequencing*' (*RNA-seq*). Os estudos foram selecionados com base nos critérios de inclusão e exclusão previamente estabelecidos (Tabela 1), resultando em uma lista de 35 estudos contendo dados de expressão gênica, conforme descrito na Tabela 2. Dentre esses, 19 estudos utilizaram a tecnologia de *microarranjo* e 16 utilizaram *RNA-seq*, totalizando 930 amostras de *microarranjo* e 301 amostras de *RNA-seq*, conforme detalhado na Tabela 2.

Tabela 2. Lista de estudos transcriptômicos selecionados.

GSE ID	Tecnologia	Tipo de RNA	Nº de amostras
GSE141549	<i>Microarranjo</i>	mRNA	408
GSE165004	<i>Microarranjo</i>	mRNA	72
GSE86534	<i>Microarranjo</i>	ncRNA/mRNA	12
GSE120103	<i>Microarranjo</i>	ncRNA/mRNA	36
GSE111974	<i>Microarranjo</i>	mRNA	48
GSE94414	<i>Microarranjo</i>	mRNA	15
GSE68870	<i>Microarranjo</i>	ncRNA/mRNA	8
GSE78851	<i>Microarranjo</i>	mRNA	8
GSE77182	<i>Microarranjo</i>	ncRNA/mRNA	5
GSE74373	<i>Microarranjo</i>	ncRNA/mRNA	6
GSE57733	<i>Microarranjo</i>	mRNA	6
GSE51981	<i>Microarranjo</i>	mRNA	148
GSE25628	<i>Microarranjo</i>	mRNA	22
GSE37837	<i>Microarranjo</i>	mRNA	36
GSE23339	<i>Microarranjo</i>	mRNA	19
GSE11691	<i>Microarranjo</i>	mRNA	18
GSE6364	<i>Microarranjo</i>	mRNA	37
GSE7305	<i>Microarranjo</i>	mRNA	20

GSE124010	<i>Microarranjo</i>	miRNAs	6
GSE185392	<i>RNA-seq</i>	mRNA	20
GSE168902	<i>RNA-seq</i>	mRNA	6
GSE157153	<i>RNA-seq</i>	mRNA	66
GSE153740	<i>RNA-seq</i>	mRNA	8
GSE153739	<i>RNA-seq</i>	mRNA	7
GSE168214	<i>RNA-seq</i>	mRNA/ncRNA	6
GSE130435	<i>RNA-seq</i>	mRNA	18
GSE135485	<i>RNA-seq</i>	mRNA	58
GSE134056	<i>RNA-seq</i>	mRNA	38
GSE99949	<i>RNA-seq</i>	mRNA	8
GSE185273	<i>RNA-seq</i>	ncRNA	6
GSE153813	<i>RNA-seq</i>	miRNA	9
GSE121406	<i>RNA-seq</i>	miRNA	8
GSE105765	<i>RNA-seq</i>	miRNA	16
GSE105764	<i>RNA-seq</i>	ncRNA	16
GSE56414	<i>RNA-seq</i>	miRNA	11

---

*GSE, GEO Series*

## 5.2. Análises de expressão gênica dos *datasets* referentes à técnica de *microarranjo*.

Os dados dos estudos feitos com a técnica de *microarranjo* foram analisados em um primeiro momento seguindo todo processo descrito no fluxograma mostrado na figura 2. Durante a análise de cada *dataset*, foram observados detalhes técnicos além dos critérios de inclusão e exclusão pré-estabelecidos, sendo verificadas as informações de cada estudo selecionado e a avaliação da qualidade dos dados obtidos. Tais detalhes envolvem a escolha das plataformas onde os dados foram obtidos, os formatos dos arquivos, se há quantidade de amostras significativas por grupo a ser analisado, se as amostras são provenientes de locais relevantes e de interesse para a pesquisa e, por fim, a qualidade das amostras propriamente ditas (etapa feita durante a execução do *pipeline*). Dos 19 *datasets* de *microarranjo* selecionados durante a curadoria, apenas cinco cumpriram os critérios de qualidade necessários para o prosseguimento para a conclusão da investigação dos genes diferencialmente expressos.

Figura 5. Fluxograma de filtragem dos datasets para análise de expressão gênica.



Tabela 3. Datasets de microarranjo descritos em grupos.

GSE_ID	Tecnologia	Plataforma	Organismo	Tipo de RNA	Amostra	N/Grupo
GSE25628	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio ectópico	7
GSE25628	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio eutópico	9
GSE25628	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio normal	6
GSE11691	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio ectópico	9
GSE11691	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio ectópico	9
GSE6364	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio proliferativo de paciente com endometriose	6
GSE6364	microarranjo	Affymetrix	Homo sapiens	mRNA	endométrio proliferativo de paciente normal	5
GSE6364	microarranjo	Affymetrix	Homo sapiens	mRNA	Endométrio secretor inicial de paciente com endometriose	6
GSE6364	microarranjo	Affymetrix	Homo sapiens	mRNA	Endométrio secretor inicial de paciente normal	3

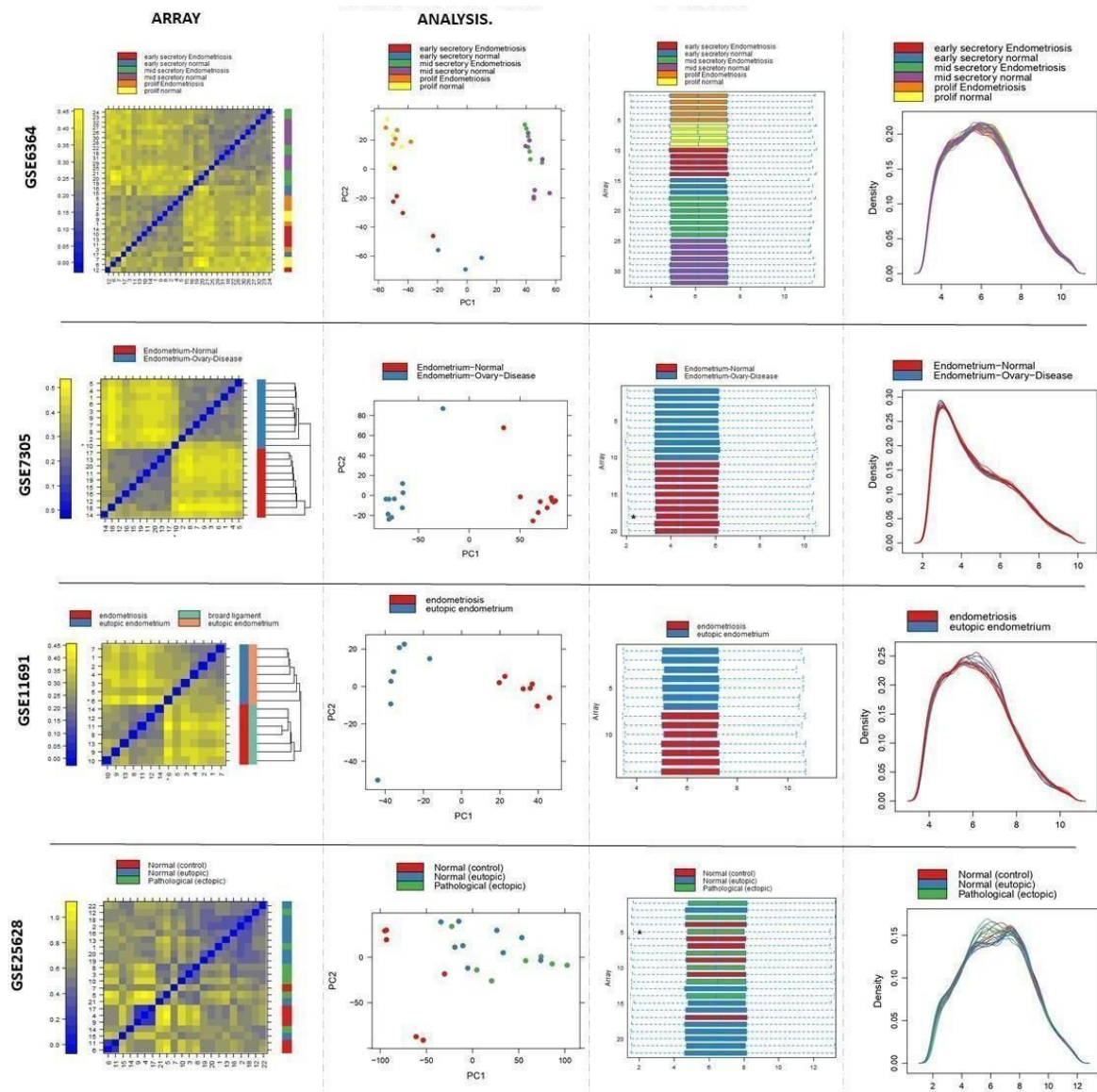


GSE6364	microarranjo	Affymetrix	Homo sapiens	mRNA	Endométrio secretor médio de paciente com endometriose	9
GSE6364	microarranjo	Affymetrix	Homo sapiens	mRNA	Endométrio secretor médio de paciente normal	8
GSE7305	microarranjo	Affymetrix	Homo sapiens	mRNA	Endometrio/doença ovariana	10
GSE7305	microarranjo	Affymetrix	Homo sapiens	mRNA	Endometrio-Normal	10
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	patient endometrium (PE) - - controle	104
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	peritoneum (PP)	39
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	Deep infiltrating endometriosis lesion Intestine (DiElN) -- Lesão de endometriose profunda infiltrante Intestino	39
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	Deep infiltrating endometriosis lesion Bexiga (DiEB) - Lesão de endometriose profunda infiltrante	3
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	Ovarian endometrioma (OMA) – endometrioma ovariano	28
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	Rectovaginal lesion (REV) – lesão rectovginal	22
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	control endometrium (CE) – controle de endometrio	43
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	control peritoneum (CP) – controle de peritonio	24
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	Peritoneal endometriosis lesion (PeL) – lesão de endometriose peritonia	79
GSE141549	microarranjo	Illumina	Homo sapiens	mRNA	Sacruterine ligament lesion (Sul) – lesão no ligamento socrouterino	27

---

### 5.2.1. Análise de qualidade dos dados feitos com pacote Array Quality Metrics

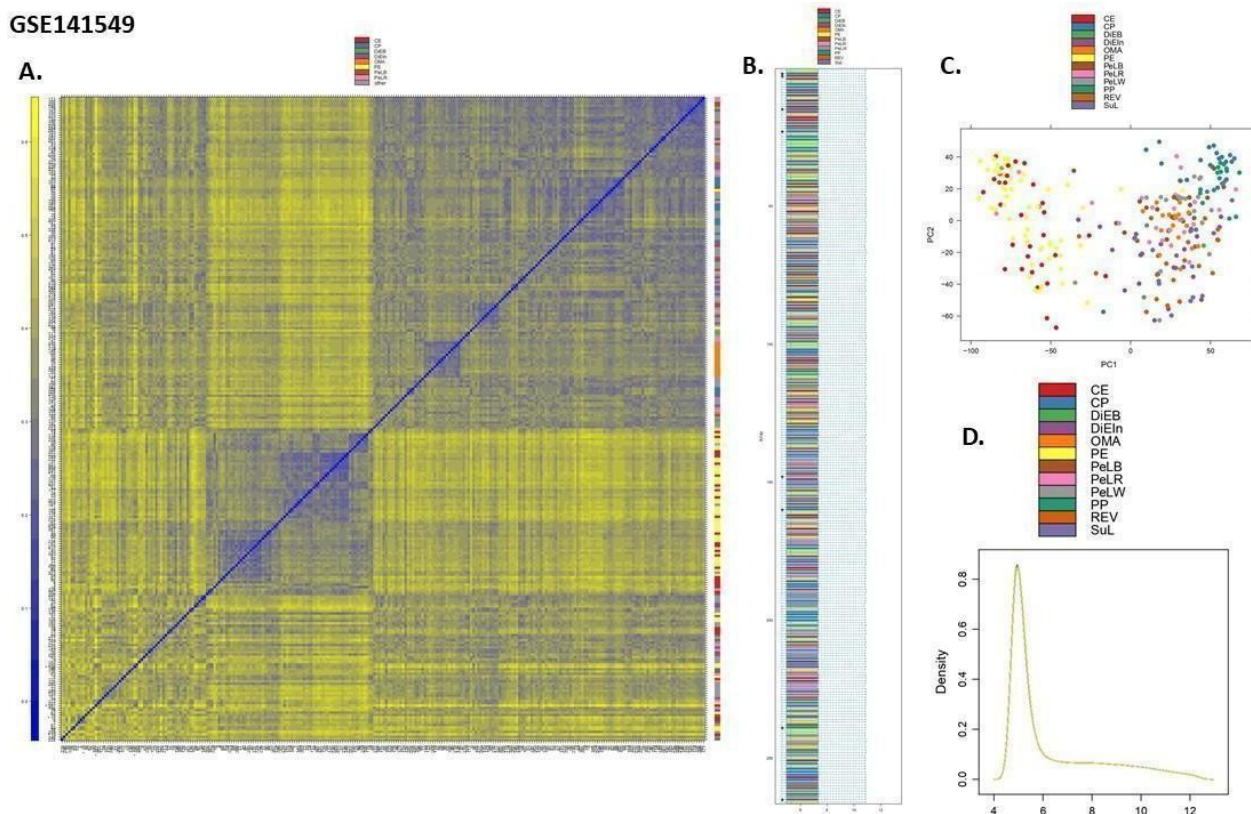
Figura 6. Gráficos gerados pelo pacote Array Quality Metrics de todos os microarrays Affymetrix analisados.



Antes da realização das análises de expressão diferencial, foram conduzidas abordagens exploratórias para avaliação da qualidade e estrutura dos dados em cada um dos estudos selecionados (Figura 6). Para isso, foram geradas quatro representações gráficas principais: (A) matriz de distância entre as amostras, (B) análise de componentes principais (PCA), (C) boxplots de distribuição de expressão e (D) gráficos de densidade.

Os *datasets* GSE6364, GSE7305, GSE11691 e GSE25628 foram obtidos da plataforma *Affymetrix* e apresentaram qualidades semelhantes. A matriz de distância (Figura 6A) permitiu visualizar a similaridade entre os perfis de expressão gênica das amostras, destacando o agrupamento por condição experimental. Em geral, observou-se uma forte coesão entre as amostras do mesmo grupo, evidenciada pelos agrupamentos hierárquicos e pelos blocos azulados ao longo da diagonal principal. Na análise de componentes principais (Figura 6B), cada ponto representa uma amostra individual, com coloração correspondente à sua classificação experimental. A separação entre os grupos ao longo dos eixos PC1 e PC2 indicou padrões de expressão gênica diferenciados entre as condições analisadas, como nos estudos GSE7305 e GSE11691, onde houve clara distinção entre tecido endometrial normal e lesões de endometriose. Os boxplots (Figura 6C) demonstraram uma distribuição uniforme das intensidades de expressão gênica após normalização com medianas alinhadas e amplitude de expressão semelhante entre as amostras dentro de cada estudo, indicando dados adequados para análise comparativa. Os gráficos de densidade (Figura 6D) complementaram essa avaliação, evidenciando curvas de distribuição com formato semelhante entre os grupos, o que sugere ausência de viés sistemático entre os conjuntos de dados. Essas análises confirmaram a qualidade técnica das amostras e a adequação dos dados para a condução das análises subsequentes de expressão diferencial e metanálise.

Figura 7. Gráficos gerados pelo pacote Array Quality Metrics referente ao dataset GSE141549, Illumina. A: Comparação entre corridas. B: Gráfico de dispersão de grupos. C: Distribuições de intensidade de matriz. D: Densidade entre grupos.



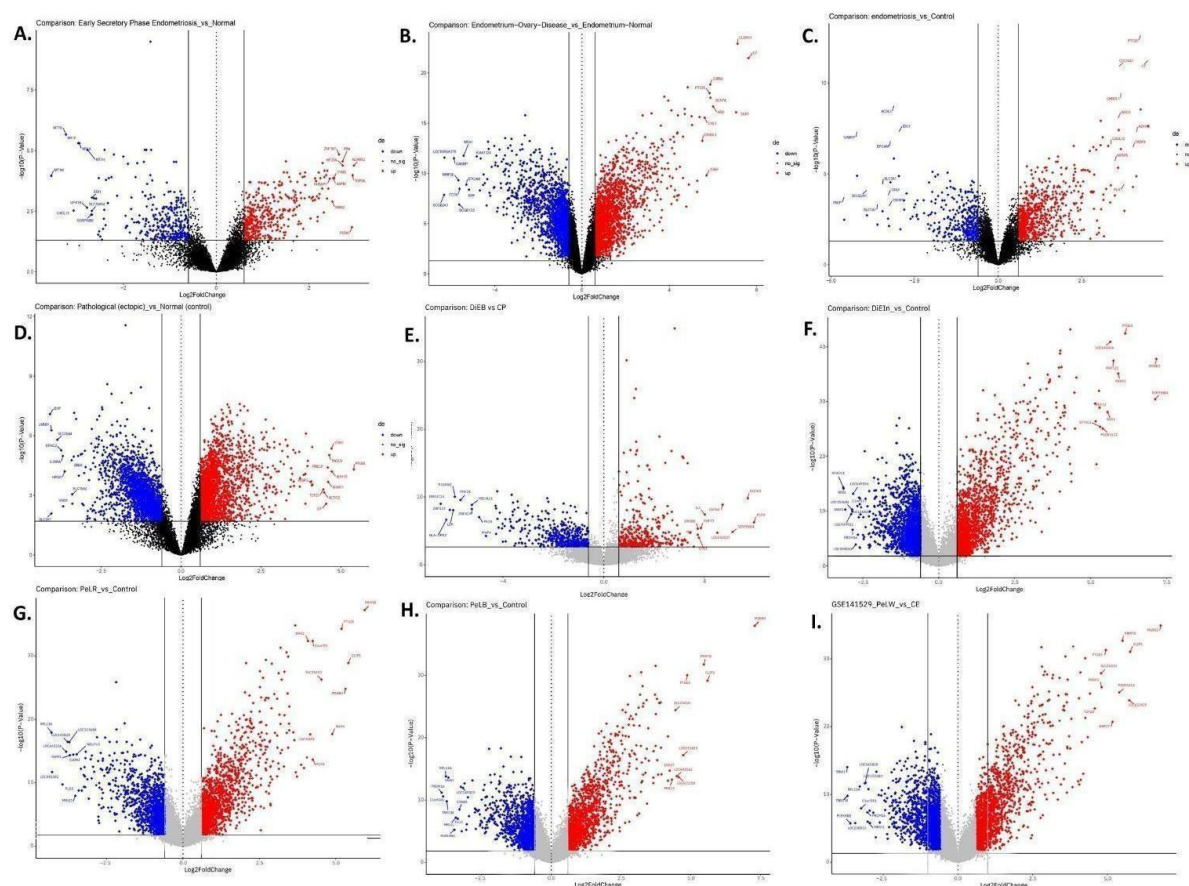
O GSE141549 foi o único dado de *microarranjo* da plataforma *Illumina* e apresentou particularidades importantes que merecem destaque. A matriz de distância entre amostras (Figura 7A) revelou um padrão altamente heterogêneo, com ausência de agrupamentos claros e predominância de variação entre as amostras. Tal comportamento sugere grande diversidade biológica ou possível presença de fatores técnicos não controlados, como efeitos de lote (batch effect) ou heterogeneidade amostral. Na análise de componentes principais (Figura 7B), observou-se ampla dispersão das amostras, sem formação de clusters nítidos entre os grupos experimentais, indicando alta variabilidade intra e intergrupos, podendo surgir a existência de subgrupos dentro da mesma condição experimental ou diferenças sutis na expressão gênica entre os grupos analisados. Os boxplots (Figura 7C) evidenciaram boa normalização geral, porém com variações consideráveis na amplitude de expressão entre algumas amostras. Tais discrepâncias podem sinalizar a presença de *outliers* ou subgrupos com perfis de expressão distintos. Além disso, os gráficos de densidade (Figura 7D) apresentaram uma distribuição assimétrica, com inclinação acentuada à esquerda, indicando uma distribuição não normal da

expressão gênica em parte das amostras. Os resultados de qualidade reforçam a necessidade do uso de estratégias como análise por subgrupos, correção de efeitos de *batch*, remoção de *outliers* e aplicação de métodos estatísticos robustos para garantir a confiabilidade dos resultados obtidos a partir deste conjunto de dados.

### 5.2.2. Resultado das Análise dos genes diferencialmente expressos em cada estudo selecionado.

A partir da análise de expressão dos dados entre os grupos de interesse, foram obtidas tabelas contendo a listagem dos GDEs regulados positivamente e negativamente para cada comparação. Com a tabela de cada comparação, foi produzido *volcanoplot* de cada análise, como demonstrado na figura 8. O quantitativo total dos GDEs encontrados em cada análise estão descritos na tabela 4

Figura 8. Volcano plots obtidos mediante os resultados de GDEs de cada comparação analisada.



Legenda das comparações: A. Early Secretory Phase Endometriosis (ESPEM) vs Early Secretory Phase Normal (ESPN), B. Ovary Disease/Endometriosis (ODEM) vs Normal (N), C. Ectopic Endometriosis (EEM) vs 45 Eutopic Normal (EN), D. Endometriosis (EM) vs Eutopic Normal (EN), E. Deep Infiltrating



Endometriosis - Blader (DiEB) vs Control Endometrium (CE), F. Deep Infiltrating Endometriosis - Intestine (DiEIn) vs Control Endometrium (CE), G. Peritoneal Lesion - Red (PeLR) vs Control Endometrium (CE), H. Peritoneal Lesion - Black (PeLB) vs Control Endometrium (CE), I. Peritoneal Lesion - White (PeLW) vs Control Endometrium (CE).

Tabela 4. Quantitativo total de GDEs por comparação.

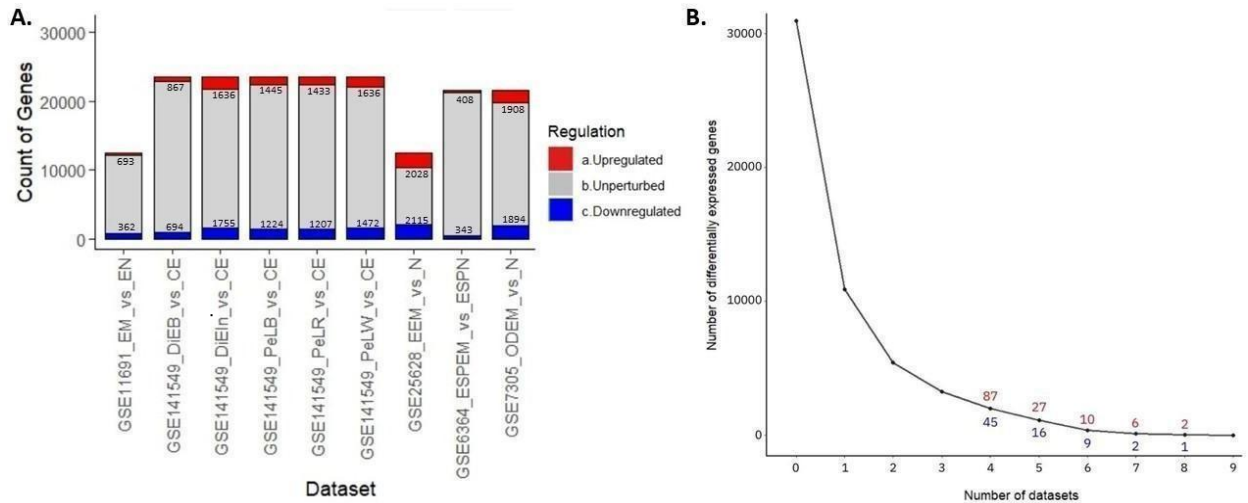
Nome Da Comparação		Up	Down	Dataset
<b>A.</b>	ESPEM_vs_ESPN	408	343	GSE6364
<b>B.</b>	ODEM_vs_N	1908	1894	GSE7305
<b>C.</b>	EEM_vs_EN	693	362	GSE11691
<b>D.</b>	EM_vs_EN	2028	2115	GSE25628
<b>E.</b>	DiEB_vs_CE	867	694	GSE141549
<b>F.</b>	DiEIn_vs_CE	1636	1755	
<b>G.</b>	PeLR_vs_CE	1433	1207	
<b>H.</b>	PeLB_vs_CE	1445	1224	
<b>I.</b>	PeLW_vs_CE	1636	1472	

*Legenda das comparações:* **A.** Early Secretory Phase Endometriosis (ESPEM) vs Early Secretory Phase Normal (ESPN), **B.** Ovary Disease/Endometriosis (ODEM) vs Normal (N), **C.** Ectopic Endometriosis (EEM) vs Eutopic Normal (EN), **D.** Endometriosis (EM) vs Eutopic Normal (EN), **E.** Deep Infiltrating Endometriosis - Blader (DiEB) vs Control Endometrium (CE), **F.** Deep Infiltrating Endometriosis - Intestine (DiEIn) vs Control Endometrium (CE), **G.** Peritoneal Lesion - Red (PeLR) vs Control Endometrium (CE), **H.** Peritoneal Lesion - Black (PeLB) vs Control Endometrium (CE), **I.** Peritoneal Lesion - White (PeLW) vs Control Endometrium (CE).

### 5.3. Resultados da Metanálise entre os dados de *microarranjo*

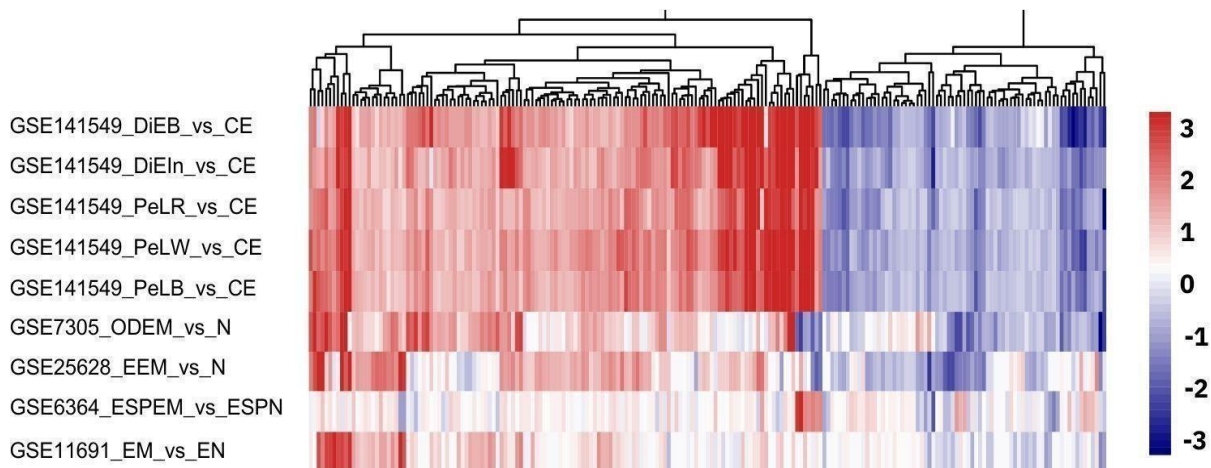
Foi realizada uma metanálise dos dados de microarranjo envolvendo as nove comparações de interesse, com o objetivo de identificar o consenso dos GDEs entre os diferentes estudos, utilizando o pacote *MetaVolcanoR*. Foram considerados GDEs os genes que apresentaram padrão consistentes pelo método de Fisher (função *combining*). Adicionalmente, a função *votecount* foi usada com o objetivo de confirmar o método de Fisher, a metodologia considerou como GDEs os genes que apresentaram padrão consistente de regulação positiva (up) ou negativa (down) em, no mínimo, cinco das comparações analisadas, conforme o demonstrado na figura 9.

Figura 9. Gráfico de metanálise função *votecount*. A: Contagem de GDEs entre os datasets. B: GDEs em consenso entre as comparações.



O número de genes presentes nos datasets analisados variou entre 12.548 e 23.588. Após a análise de expressão gênica diferencial, o total de genes up-regulados variaram de 408 (GSE6364) a 2.028 (GSE25628). De forma semelhante, os genes down-regulados variaram de 343 (GSE6364) a 2.115 (GSE25628), conforme ilustrado na Figura 9A. Além disso, a análise de interseção entre os *datasets* revelou um consenso mínimo de 87 genes positivamente regulados e 45 genes negativamente regulados compartilhados por, no mínimo, quatro comparações, de acordo com o apresentado na Figura 9B.

Figura 10. referente a meta-análise. Grupos experimentais analisados



Grupos experimentais analisados: Early Secretory Phase Endometriosis (ESPEM) vs Early Secretory Phase Normal (ESPN), Ovary Disease/Endometriosis (ODEM) vs Normal (N), Ectopic Endometriosis (EEM) vs Eutopic Norma(EN), Endometriosis (EM) vs Eutopic Normal (EN), Deep Infiltrating Endometriosis – Blader (DiEB) vs Control Endometrium (CE), Deep Infiltrating Endometriosis – Intestine

(DiElN) vs Control Endometrium (CE), Peritoneal Lesion - White (PeLW) vs Control Endometrium (CE), Peritoneal Lesion - Red (PeLR) vs Control Endometrium (CE) e Peritoneal Lesion - Black (PeLB) vs Control Endometrium (CE).

A metanálise resultou na identificação de 205 genes em consenso entre os datasets analisados, dos quais 132 apresentaram expressão aumentada (assinatura gênica positiva) e 73 apresentaram expressão reduzida (assinatura gênica negativa). Esses conjuntos de genes estão evidentes na Figura 10, na qual observam-se a formação de clusters de expressão semelhantes entre os estudos. Vale destacar que grande parte dos genes consensuais tem origem no dataset GSE141549, por conta das comparações entre os cinco grupos experimentais avaliados; contudo, observou-se concordância expressiva também em relação aos demais *datasets* incluídos na análise. A única exceção foi o dataset GSE6364, que apresentou um perfil de expressão discrepante, com baixa alteração gênica em comparação ao grupo controle, comportamento observado desde a análise de qualidade, a lista completa dos GDEs identificados está descrita na tabela 5.

Tabela 5. Conjunto de 205 GDEs obtidos em consenso na metanálise entre os *datasets* de *microarranjo*.

UP					DOWN				
1	CFH	51	CD53	101	SRF	1	CLDN3	51	RAD21
2	GAS6	52	CLEC7A	102	STEAP1	2	HMGB3	52	RAD51C
3	HSPA12A	53	COPZ2	103	STOX2	3	AHCY	53	RBL1
4	KLF2	54	CPEB4	104	SULT1A2	4	EEF1E1	54	RNF183
5	PPP1R12B	55	CUTA	105	SYTL4	5	ERBB3	55	SCNN1G
6	TSPAN4	56	CYR61	106	TGFBR1	6	ITGB8	56	SEC23B
7	AKAP12	57	DCTN3	107	TIMP4	7	KRT23	57	TRAK1
8	ANGPT1	58	DDX41	108	TLE2	8	LRBA	58	TTC27
9	BGN	59	EDC3	109	TMEM120A	9	NUP62CL	59	WDR5B
10	CD34	60	EGR2	110	TMEM133	10	OCLN	60	YES1
11	CSF1R	61	EPHB2	111	TMEM97	11	RAD54B	61	YME1L1
12	DCN	62	EPM2A	112	TNFAIP8L3	12	SPA17	62	ZNF217
13	GAS7	63	FPGS	113	TRAF5	13	STX19	63	ZNF334
14	GLI3	64	GIMAP8	114	TRPC1	14	TEKT1	64	ACSM3
15	HCLS1	65	GNA14	115	TSSC4	15	TFB2M	65	CTSC
16	LARP6	66	HSD17B11	116	VPS16	16	TLR3	66	DNM1
17	LHFP	67	IDH3G	117	VPS4A	17	TP53I3	67	FBXO5
18	LMNA	68	IL7R	118	WASF3	18	UBE2E3	68	KIF4A
19	LYZ	69	KLF9	119	WBSCR16	19	HPN	69	NEK1
20	MAGI2	70	LRCH2	120	WFS1	20	ADSSL1	70	PFAS
21	MMP2	71	LRPAP1	121	CSGALNACT2	21	ANAPC10	71	TLR5



22	MYH10	72	LTBP3	122	DAPK1	22	ATP6V1A	72	ZNF226
23	NDUFA4L2	73	MAP4	123	DIAPH3	23	C14orf159	73	PAFAH1B3
24	NET1	74	MAPRE2	124	MARCKSL1	24	C17orf58		
25	PCP4	75	MFAP5	125	MEST	25	C2orf49		
26	PDE1A	76	MPDZ	126	MKI67	26	CCDC6		
27	PDGFRL	77	MSRB3	127	PENK	27	CDKN2A		
28	PGM5	78	NBR1	128	PPARG	28	CMAS		
29	PLSCR4	79	NQO2	129	RNF13	29	CMTM7		
30	SPARCL1	80	PDE7B	130	ZNF367	30	DPY19L4		
31	STAT5A	81	PDZRN4	131	CCNA2	31	ELP3		
32	TENC1	82	PHF1	132	DHFR	32	EPB41L4B		
33	WNT2B	83	PLA2R1			33	GALNT12		
34	CLIP1	84	PLEKHO1			34	GSTP1		
35	ABI3BP	85	PPP2R4			35	H2AFY2		
36	ACOT8	86	PTGES			36	HIC1		
37	ADAMTSL2	87	PTPRB			37	KIAA0196		
38	ANKRD50	88	PUF60			38	KLHL13		
39	ANO6	89	RAB3IL1			39	KLHL9		
40	ANPEP	90	RALGDS			40	LRPPRC		
41	APOL3	91	RPUSD2			41	NDUFS3		
42	ARHGAP18	92	RRAS			42	NEBL		
43	ASPA	93	RUSC2			43	NIPSNAP3A		
44	ATP1A2	94	SCG2			44	PARP1		
45	BAD	95	SEMA3C			45	PBX2		
46	BEGAIN	96	SGCD			46	PDIA6		
47	C3AR1	97	SH3TC1			47	PDZK1IP1		
48	CALCOCO1	98	SLC25A37			48	PHYHD1		
49	CALHM2	99	SLC41A3			49	PIGN		
50	CAMK1	100	SLC43A3			50	PLCH1		

#### 5.4. Resultados do Enriquecimento Funcional

Com o conjunto de genes obtido na metanálise, foi realizada a análise de enriquecimento funcional utilizando o *software EnrichR*. Os genes diferencialmente expressos (GDEs) foram organizados em listas separadas de positivamente regulados e negativamente regulados, e avaliou-se sua associação com processos biológicos relacionados à doença em cinco bases de dados: *GO Biological Process*, *KEGG Pathways*, *Reactome Pathways*, *WikiPathways* e *MSigDB\_Hallmark*. As vias biológicas mais relevantes foram selecionadas com base em sua significância estatística, número de genes sobrepostos (*overlaps*) e consistência entre vias semelhantes identificadas nos diferentes bancos, considerando-se a fisiopatologia da

doença em estudo e a concordância com as informações encontradas na literatura.

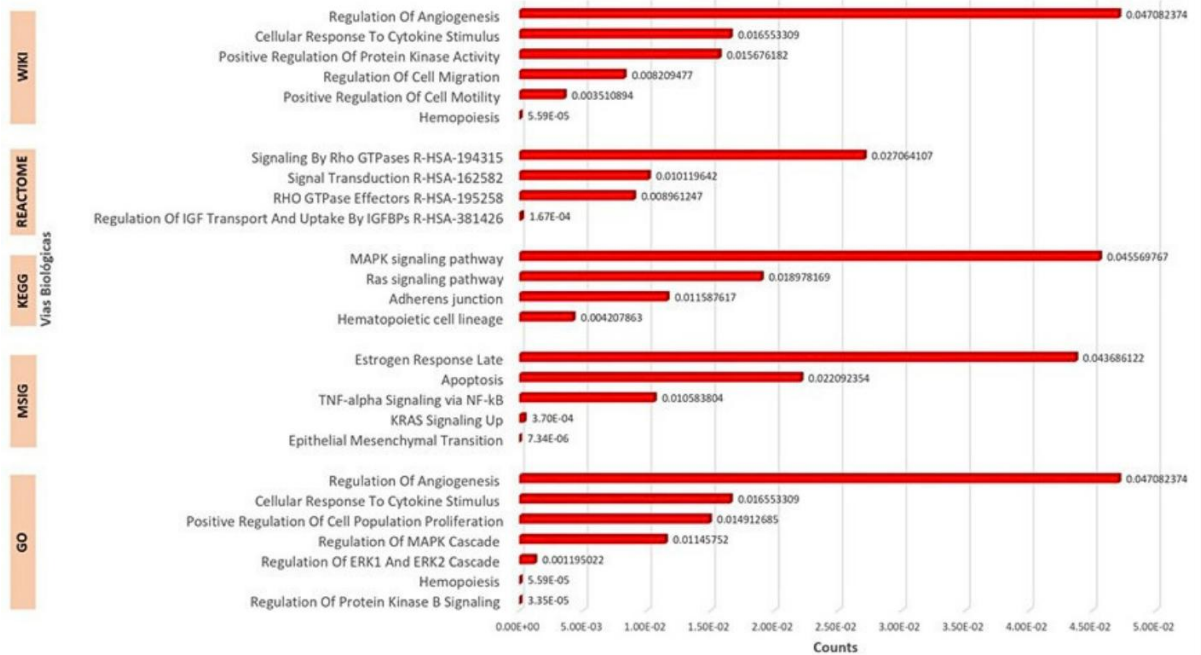
Tabela 6. Quantitativo geral das vias biológicas encontradas em cada banco de dados.

<b>Banco de Dados</b>	<b>UP</b>	<b>DOWN</b>
<i>WikiPathways</i>	22	4
<i>Reactome Pathways</i>	13	24
<i>KEGG Pathways</i>	17	9
<i>MSigDB_Hallmark</i>	9	4
<i>GO Biological Process</i>	74	17
<b>Total</b>	<b>135</b>	<b>58</b>

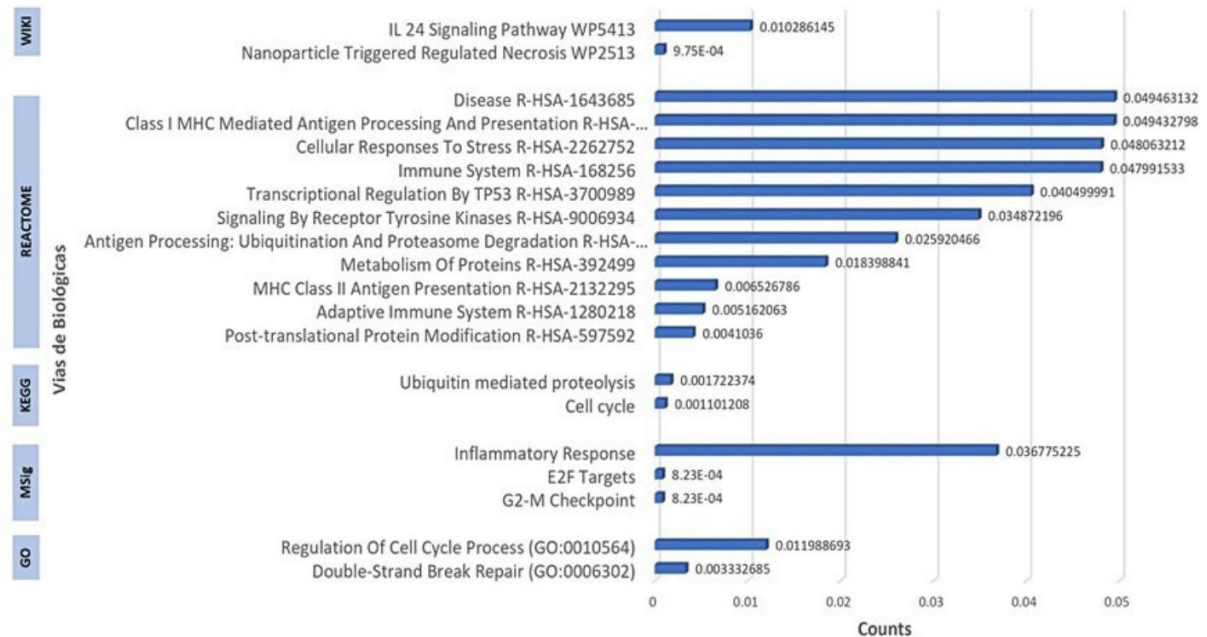
Das vias biológicas encontradas em cada um dos cinco banco de dados (tabela 6), foram consideradas em consenso entre eles e a literatura as vias descritas na figura 11.

Figura 11. Gráfico de coluna referente às vias enriquecidas do conjunto de GDEs obtidos na metanálise.

## A. POSITIVAMENTE REGULADOS



## B. NEGATIVAMENTE REGULADOS



A análise de enriquecimento funcional dos genes regulados positivamente (figura 11A) revelou envolvimento significativo em processos associados à sinalização celular, proliferação e diferenciação. Dentre os principais termos identificados no banco *GO Biological Process*, destacou-se a via de regulação da sinalização da proteína quinase B (AKT), importante para a sobrevivência celular, metabolismo e crescimento tecidual. A hematopoiese também foi enriquecida, indicando um possível recrutamento de células imunes ou progenitoras em tecidos ectópicos. Vias associadas à regulação da cascata ERK1/ERK2 e da via MAPK, ambas ligadas à proliferação celular e inflamação, foram identificadas como significativamente ativadas. Além disso, a regulação positiva da proliferação celular apareceu como um termo relevante, o que corrobora o caráter hiperproliferativo observado em lesões endometrióticas e à angiogênese, compatível com o microambiente vascularizado da doença. Além disso, houve indução de vias relacionadas à cicatrização (migração celular, reparo de membrana e ativação da cascata de coagulação) e ao processo inflamatório (produção de citocinas como IL-2 e IL-9, ativação de células dendríticas, linfócitos T e INF- $\alpha$ ). Destaca-se ainda ativação de vias ligadas à biossíntese de estrogênio, hormônio-chave na progressão da endometriose, e à sinalização da dor (formação de neurônios nociceptivos), o que pode explicar sintomas clínicos como dor pélvica.

Entre os genes down-regulados (figura 11B), observou-se o enriquecimento de vias associadas à manutenção da estabilidade genômica e progressão do ciclo celular. Notavelmente, a via de reparo de quebras duplas no *DNA* foi um dos termos mais significativos, incluindo genes como *RAD51C* e *PARP1*, sugerindo uma possível deficiência em mecanismos de reparo do DNA. Além disso, a análise identificou o termo checkpoint *G2-M* no banco *MSigDB\_Hallmark*, refletindo uma possível desregulação da transição mitótica, bem como a via de alvos de *E2F*, indicando uma supressão de genes envolvidos no controle da progressão do ciclo celular e proliferação. Esses achados sugerem uma possível perda da capacidade proliferativa controlada em certos contextos celulares ou tecidos acometidos. Também foi observada diminuição na defesa imune inata (receptores *Toll-like*) e na redução nas vias de ativação do complexo de histocompatibilidade (*MHC I e II*), sugerindo escape imunológico.

Figura 12. Resumo das vias biológicas enriquecidas e dos genes relacionados.



### 5.5. Resultados da análise dos genes diferencialmente expressos em dados de *RNA-seq*

Com o objetivo de validar os GDEs identificados na metanálise inicial com dados de *microarranjo*, foram incorporados ao estudo datasets de *RNA-seq*, permitindo a identificação de genes em consenso entre as diferentes plataformas tecnológicas. Essa abordagem integrativa proporcionou uma filtragem mais rigorosa dos genes efetivamente associados à endometriose, aumentando a confiabilidade dos resultados.

Assim como nas análises dos *datasets* de *microarranjo*, os dados de *RNA-seq* de cada estudo também foram avaliados considerando não apenas os critérios de inclusão e exclusão previamente definidos, mas também aspectos técnicos relevantes. Para cada estudo selecionado, foram verificadas informações detalhadas e foi realizada uma avaliação da qualidade dos dados. Esses aspectos incluem a plataforma utilizada para a geração dos dados, os formatos dos arquivos, a quantidade de amostras por grupo (avaliando se é estatisticamente adequada, em triplicata), a origem das amostras (considerando sua relevância para os objetivos da pesquisa) e, por fim, a qualidade das amostras em si. Esta última etapa foi realizada durante a execução do pipeline.

Tabela 7. Datasets RNA-seq. Os estudos marcados em verde foram analisados e incluídos à pesquisa.

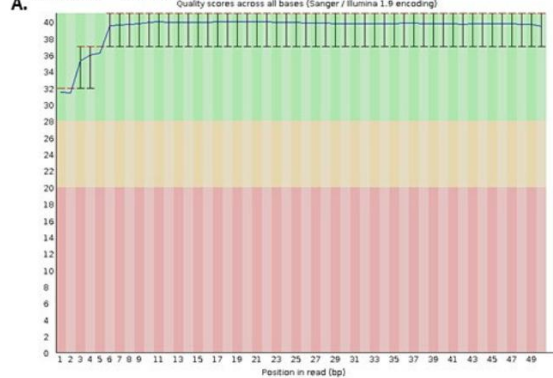
<b>Dataset</b>	<b>Plataforma</b>	<b>tipo de RNA</b>	<b>Nº de amostra</b>	<b>Analisados</b>
GSE185392	Illumina NovaSeq 6000	mRNA	20	Não
GSE168902	Illumina NovaSeq 6000	mRNA	6	Sim
GSE157153	Ion Torrent Proton	mRNA	66	Não
GSE153740	Illumina NextSeq 500	mRNA	8	Sim
GSE153739	Illumina NextSeq 500	mRNA	7	Não
GSE168214	Illumina HiSeq 4000	lncRNA	6	Não
GSE130435	Illumina HiSeq 4000	mRNA	18	Não
GSE135485	Illumina HiSeq 3000	mRNA	58	Sim
GSE134056	Illumina NextSeq 500	mRNA	36	Não
GSE99949	Ion Torrent PGM	mRNA	8	Não
GSE185273	Illumina NextSeq 500	ncRNA	6	Não
GSE153813	Illumina NextSeq 500	miRNA	9	Não
GSE121406	Illumina NextSeq 500	miRNA	8	Não
GSE105765	Illumina HiSeq 2000	miRNA	16	Não
GSE105764	Illumina HiSeq 4000	ncRNA	16	Não
GSE56414	Illumina Genome Analyzer	miRNA	11	Não

Ao executar as análises conforme o estipulado no fluxograma descrito nas figuras 3 e 4, foi realizada a análise de qualidade de todas as amostras dos seus respectivos datasets selecionados como satisfatórios para a análise, conforme os critérios definidos no tópico 4.1.2.4. Como forma de exemplificar os gráficos de qualidade obtidos por meio da ferramenta *FASTQC* (figura 13).

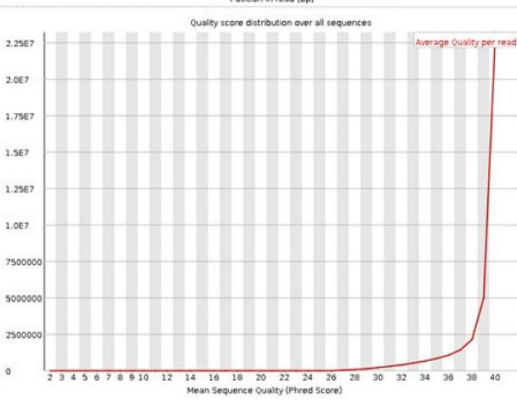
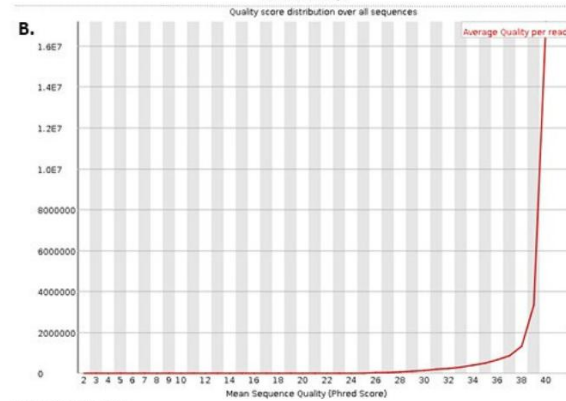
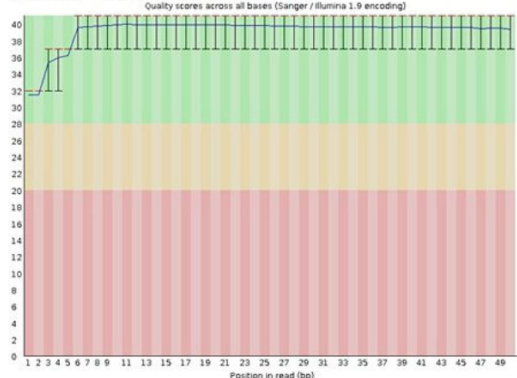
Figura 13. Gráficos gerados no relatório de qualidade do FASTQC referente aos datasets GSE135485 e GSE168902

#### GSE135485

##### Amostra – SRR9915625 (ENDOMETRIOSE)

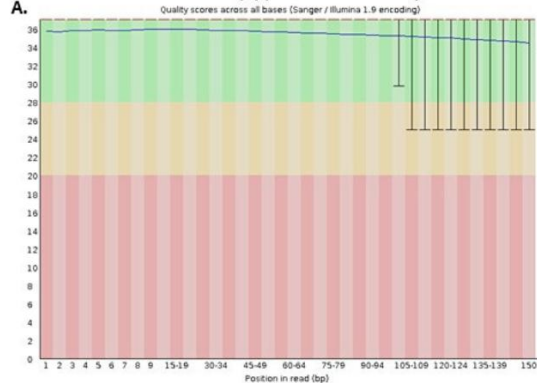


##### Amostra – SRR9915682 (CONTROLE)

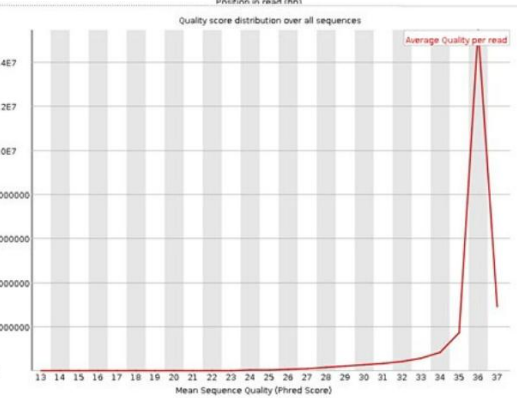
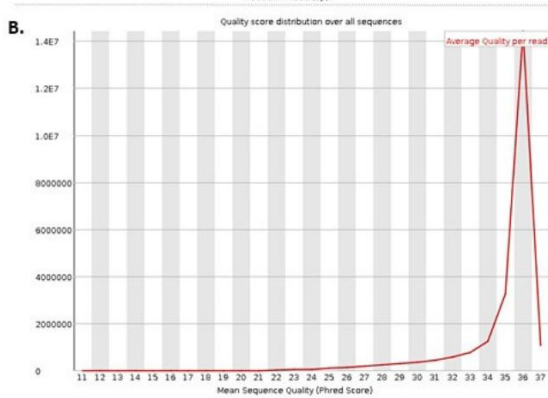
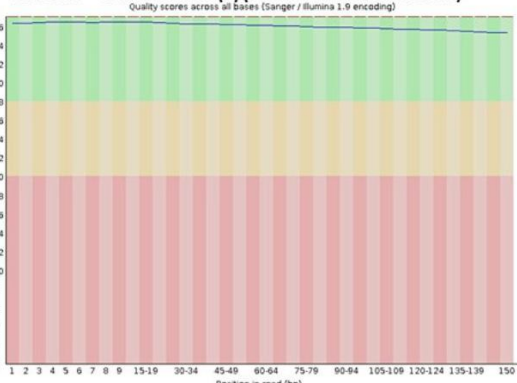


#### GSE168902

##### Amostra – SRR13963617(2) (ENDOMETRIOSE)



##### Amostra – SRR13963612(1) (CONTROLE - LEIOMIOMA)



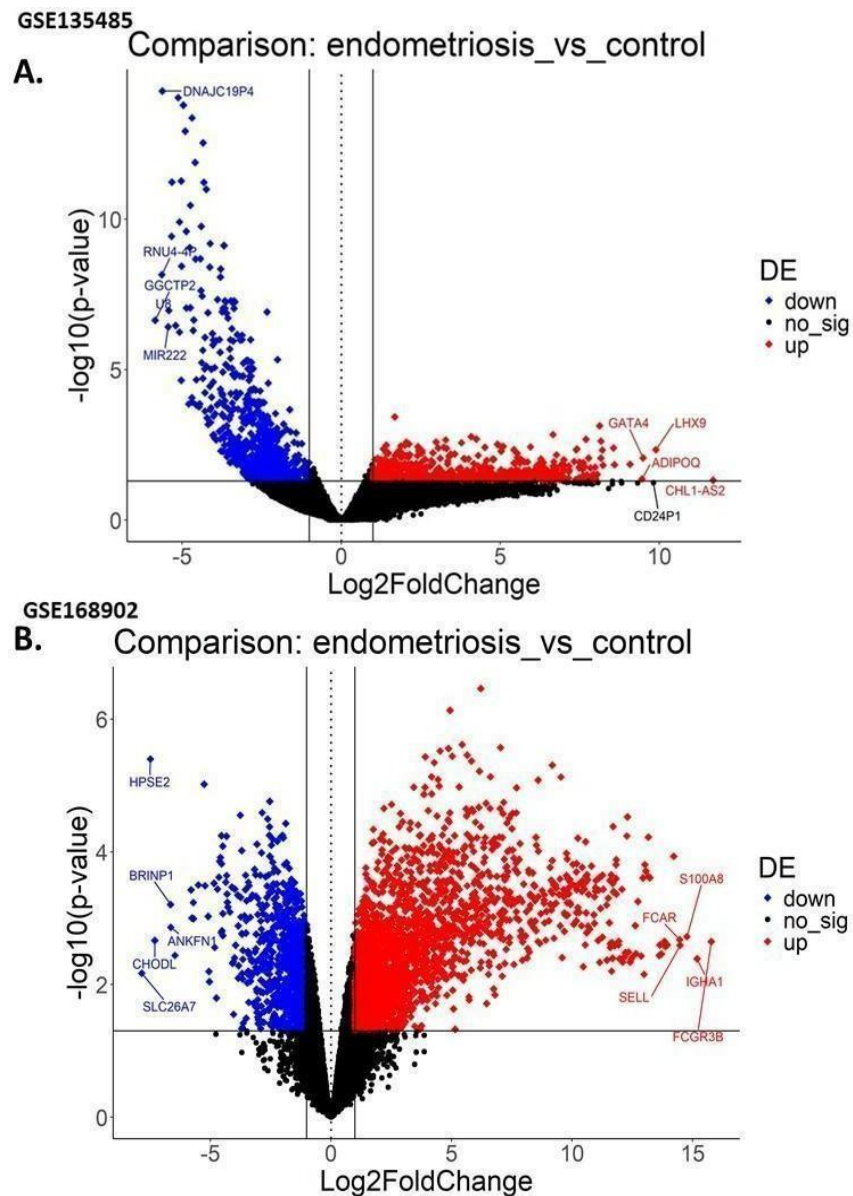
A avaliação da qualidade das sequências por meio dos gráficos gerados pelo

FastQC (figura 13) revelou que as amostras dos estudos GSE135485 e GSE168902 apresentam, em sua maioria, perfis de alta qualidade. No estudo GSE135485, tanto a amostra de endometriose (SRR9915625) quanto a de controle (SRR9915682) demonstraram qualidade consistente ao longo de toda a extensão das leituras, com valores de Phred majoritariamente acima de 30 (faixa verde), indicando alta confiabilidade dos dados. Já no estudo GSE168902, observou-se que a amostra controle (SRR13963612 – leiomioma) também apresentou excelente qualidade em todas as posições da leitura. No entanto, a amostra de endometriose (SRR13963617) mostrou uma leve queda na qualidade nas posições finais da leitura, com maior variabilidade nas últimas bases, embora a média de qualidade ainda permaneça acima dos níveis considerados aceitáveis, e possuem qualidade adequada para análises transcriptômicas. Tais resultados foram semelhantes em todas as demais amostras dos *datasets* analisados.

As análises de genes diferencialmente expressos foram feitas (figura 14) por meios da lista de GDEs identificados em cada dataset, foi feito um volcano plot demonstrando os genes regulados positivamente, demonstrados em vermelho, e os genes regulados negativamente, demonstrado em azul. A tabela com o quantitativo dos GDEs encontrados está descrito na tabela 8.



Figura 14. Volcano plots das análises referentes aos datasets selecionados da técnica de RNA-seq entre mulheres com endometriose e controle.



Legendas:(A) GSE135485 (B) GSE168902. DE, Diferencialmente Expresso; down, regulados negativamente; no sig, Sem significância estatística; up, regulados positivamente.

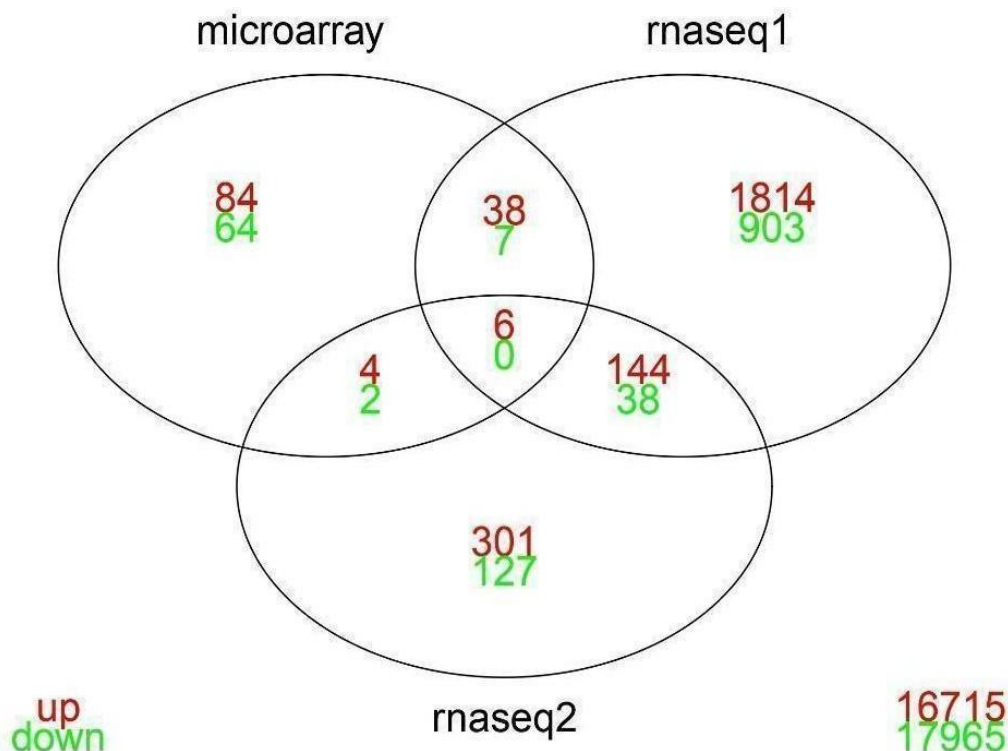
Tabela 8. Quantificação de GDEs identificados em cada um dos datasets de RNA-seq.

Nome da comparação	Up	Down	Dataset	Tecnologia
EaL_vs_CE	2002	348	GSE168902	RNA-seq
OEM_vs_CE	455	167	GSE135485	

Os grupos experimentais: Endometriosis and Leyomioma (EaL) e Ovary Edometriosis (OEM) foram comparados contra os grupos Control Endometrium (CE).

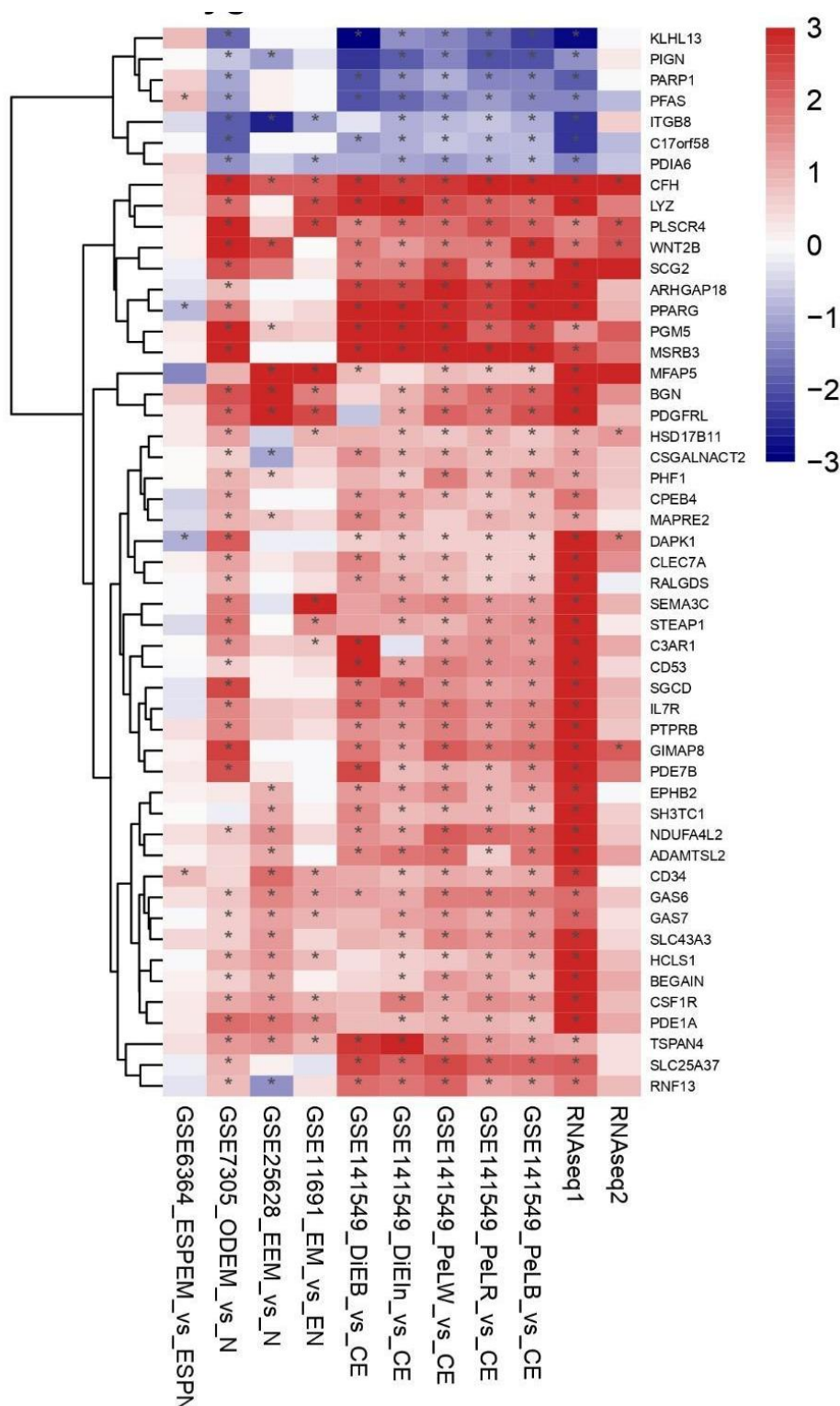
### 5.6. Análise de Consenso entre os GDEs Microarranjo e RNA-seq

Figura 15. Diagrama de Venn, demonstrando os GDEs em consenso entre integração das técnicas.



O conjunto de GDEs obtidos na metanálise foi analisado para identificar os genes em consenso entre as diferentes plataformas (Figura 15). Entre os dados de microarranjo e *RNA-seq* 1 verificou-se 38 genes com regulação positiva e 7 com regulação negativa em comum. Já entre microarranjo e *RNA-seq* 2, apenas 4 genes regulação positiva e 2 regulação negativa foram compartilhados. De forma notável, *RNA-seq* 1 e *RNA-seq* 2 apresentaram maior convergência entre si, com 144 genes com expressão positiva e 38 negativa compartilhados, o que possivelmente reflete a similaridade entre as plataformas tecnológicas utilizadas. A interseção entre os três conjuntos revelou seis genes regulados positivamente em comum, representando alvos potencialmente robustos e recorrentes na endometriose, por terem sido identificados de forma consistente em diferentes abordagens analíticas, ressaltando tanto a variabilidade metodológica inerente às diferentes plataformas transcriptômicas quanto a relevância de abordagens integrativas na identificação de biomarcadores confiáveis.

Figura 16. Heatmap demonstrando a assinatura gênica da endometriose considerando múltiplos fatores mostra mudança no perfil de expressão da endometriose. Os valores de  $\log_2FC$  dos estão representados no gráfico em vermelho (aumento) e azul (diminuição).



A estratégia de cruzamento entre dados de *microarranjo* e *RNA-seq* foi feita e se demonstrou robusta, também demonstrada por meio de um *heatmap*, destacando alterações transcricionais consistentes entre as plataformas. A concordância

observada entre as diferentes técnicas reforça a relevância biológica dos GDEs identificados, posicionando-os como potenciais biomarcadores ou alvos terapêuticos para a endometriose. Esses resultados não apenas validam os achados iniciais da metanálise, mas também fornecem uma base mais sólida para investigações futuras sobre os mecanismos moleculares envolvidos na doença.

Mesmo com diferentes condições de endometriose, diferentes plataformas e possíveis vieses entre as técnicas, a análise de consenso entre os dados de *microarranjo* e *RNA-seq* dos genes (figura 16) mais consistentemente como regulados positivamente destacam-se *CFH*, *GAS6*, *TSPAN4*, *WNT2B* e *PLSCR4*. No conjunto de genes com regulação negativa em múltiplas análises, destacam-se *PIGN*, *ITGB8*, *PFAS*, *KLHL13*, *PARP1*. A recorrência desses genes ao longo das diferentes plataformas transcriptômicas, condições experimentais distintas, reforça sua relevância biológica, configurando-os como potenciais biomarcadores transcricionais robustos para a doença.

Esse conjunto de genes pode servir como base para análises funcionais posteriores e para o desenvolvimento de biomarcadores diagnóstico-terapêuticos mais específicos para os diferentes tipos de endometriose. Esses resultados evidenciam a importância de abordagens de comparação de diferentes técnicas na identificação de alvos moleculares confiáveis, especialmente diante da heterogeneidade clínica e molecular da endometriose.

## 6. Discussão

A endometriose é uma doença crônica que envolve interações gênicas complexas e um componente ambiental que modificam o potencial proliferativo das células contribuindo para os diferentes fenótipos associados e desfechos clínicos. Neste estudo, aplicamos uma metanálise abrangente de dados transcriptômicos públicos, realizando a integração de dados entre as técnicas de análise da expressão gênica de microarranjo e *RNA-seq*, em plataformas distintas, permitindo a identificação da assinatura diferencial de expressão gênica. Ao reunir e comparar 35 datasets (com mais de 1.200 amostras no total), e posterior filtragem e análise de expressão diferencial de sete desses datasets, foi possível estabelecer um perfil transcricional em consenso da doença, destacando tanto GDEs quanto vias biológicas com expressão alterada em diferentes contextos experimentais.

Essa abordagem superou a heterogeneidade inerente aos estudos isolados e proporcionou uma análise mais robusta dos mecanismos moleculares associados à endometriose. A metanálise entre as comparações dos grupos experimentais dos *datasets* provenientes da técnica de *microarranjo* identificou 205 genes em consenso, dos quais 132 estavam consistentemente regulados positivamente e 73 regulados negativamente entre os grupos com endometriose e controles. Posteriormente, foi realizada a validação da metanálise por meio da integração entre os GDEs identificados nas análises de expressão diferencial de *RNA-seq*, na qual foi validado um painel de 57 genes. Os principais genes detectados foram, como *CFH*, *GAS6*, *TSPAN4*, *WNT2B* e *PLSCR4* entre o conjunto dos genes positivamente regulados, além de *PIGN*, *ITGB8*, *PFAS*, *KLHL13* e *PARP1* genes regulados negativamente. Dentre as vias biológicas mais significativas tanto estatisticamente, quanto clinicamente, encontram-se enriquecidas relacionadas aos genes com expressão aumentada as vias envolvidas principalmente na ativação do sistema complemento, sinalização celular, proliferação celular, inflamação, angiogênese e biossíntese de estrogênio. Paralelamente, os genes com expressão diminuída estão ligados a vias de reparo de DNA e regulação do complexo principal de histocompatibilidade e escape imunológico. Esses resultados demonstram que a integração de dados transcriptômicos públicos permitiu a identificação de uma

assinatura gênica consistente na endometriose, fornecendo novas perspectivas para a compreensão de sua biologia e possíveis alvos terapêuticos.

Entre os genes com expressão aumentada, destacou-se o gene *GAS6*, envolvido na ativação do receptor tirosina quinase AXL, foi previamente associado a processos inflamatórios, fibrogênicos e de invasão celular em lesões ectópicas (HONDA et al., 2008). Estudos recentes demonstram que o eixo *GAS6*–AXL promove a diferenciação de células estromais em miofibroblastos por meio da via NF- $\kappa$ B, contribuindo para a fibrose e remodelamento tecidual (SHASHA et al., 2025). Esse eixo também ativa a via PI3K-Akt, favorecendo a sobrevivência celular ectópica. O *CFH*, componente da cascata do complemento, sugere uma modulação da imunidade inata, potencialmente conferindo evasão imune às células ectópicas (BAE et al., 2022). *TSPAN4*, *WNT2B* e *PLSCR4* estão implicados na adesão, remodelamento da matriz extracelular e sinalização celular, processos essenciais à implantação e persistência das lesões.

Apesar de ainda escassas, algumas evidências apontam para o possível envolvimento desses genes na fisiopatologia da endometriose. *WNT2B* pertence à família Wnt, amplamente reconhecida por sua atuação em processos de diferenciação celular, remodelamento tecidual e fibrose, com dados sugerindo sua contribuição para a proliferação de células estromais endometriais e progressão de lesões endometrióticas (BULUN et al., 2019; WU et al., 2022). *PLSCR4*, por sua vez, codifica uma escramblase de fosfolipídios com possível função na sinalização de membrana e regulação de receptores, sendo sua regulação positiva relacionada a processos de migração celular e possível invasividade das lesões (ZHANG et al., 2021). Para *TSPAN4*, não há evidência direta de associação com endometriose, mas sua participação em processos de adesão celular e remodelamento da matriz extracelular em outros contextos patológicos justifica sua investigação como potencial modulador da adesividade celular em tecido ectópico (LIU et al., 2020).

Os genes com expressão reduzida, destacam-se *PARP1* (Poly(ADP-ribose) polymerase 1), amplamente relatado na literatura, especialmente no que se refere sua participação em diversos tipos câncer e mais especificamente, no de ovário, onde

exerce função central no câncer de ovário, especialmente no carcinoma seroso de alto grau, o subtipo mais frequente e agressivo da doença. Sua principal função está relacionada ao reparo de quebras simples de fita do DNA, via excisão de base (base excision repair), sendo fundamental para a manutenção da integridade genômica (POMMIER et al., 2016). Em tumores com deficiência na via de recombinação homóloga, como aqueles associados a mutações nos genes *BRCA1/2*, a inibição de *PARP1* leva ao acúmulo de lesões de DNA irreparáveis, culminando em morte celular por letalidade sintética (LORD; ASHWORTH, 2017). Além de sua atuação no reparo do DNA, *PARP1* também regula vias de transcrição de genes relacionados à sobrevivência celular, angiogênese, inflamação e interação com o microambiente tumoral, ampliando seu impacto na biologia tumoral (POMMIER et al., 2016; KONSTANTINOPOULOS et al., 2015).

*PARP1* e *KLHL13* destacam-se pela participação no reparo de DNA e controle do ciclo celular e estresse oxidativo, sugerindo um prejuízo na manutenção da integridade genômica nas lesões endometrióticas (KACPERCZYK-BARTNIK, J. et al., 2022; BIRT, J. A. et al., 2013). O gene *ITGB8*, que codifica uma integrina  $\beta 8$  envolvida na ativação do fator de crescimento TGF- $\beta$ , é essencial para processos de adesão celular, modulação imune e implantação embrionária. Estudos indicam que sua expressão é crítica para permitir a ativação funcional das proteínas SMAD2/3, por meio de sua fosforilação e translocação nuclear durante a implantação, sendo que sua redução em pacientes com endometriose pode comprometer a receptividade endometrial e, conseqüentemente, contribuir para a infertilidade associada à doença (KUMAR et al., 2015; KATIYAR et al., 2018).

O gene *PFAS* apresenta função essencial no metabolismo de purinas, sua down-regulação pode impactar a homeostase celular e imunológica, favorecendo um ambiente propício à sobrevivência de células ectópicas, ainda que mais estudos sejam necessários para elucidar seu papel específico na endometriose (AO et al., 2024). Por sua vez, *PIGN*, embora não tenha relatos posteriores sobre seu envolvimento na fisiopatologia da doença, sabe-se que tem como função a biossíntese de âncoras GPI (glicosilfosfatidilinositol), é para a fixação de proteínas na membrana celular. A expressão reduzida de *PIGN* pode prejudicar a apresentação de proteínas de superfície, como receptores e moléculas imunomoduladoras, contribuindo para o

escape imunológico das células endometriais ectópicas (THE HUMAN PROTEIN ATLAS, 2024).

A análise de enriquecimento funcional dos genes down-regulados revelou um padrão marcante de supressão de vias relacionadas à resposta imune adaptativa, tais como “Apresentação do antígeno MHC de classe I e II” e “sistema imunológico adaptativo”. Esses resultados sugerem que, em lesões de endometriose, há uma quebra na capacidade de apresentação antigênica e, conseqüentemente, um escape imunológico que favorece a permanência de células ectópicas. Essa hipótese é compatível com achados prévios na literatura, que apontam alterações funcionais em células NK, T e APCs no microambiente peritoneal de mulheres com endometriose (AHN et al., 2015; ABRAMIUK, et al., 2022).

Além disso, observou-se a inibição de vias de reparo do DNA, como reparo de fita-dupla de DNA, bem como checkpoint mitótico (G2-M Checkpoint) e alvos de E2F (E2F Targets). A regulação negativa de genes como *PARP1* e *RAD51C* pode comprometer a resposta ao dano genético, promovendo instabilidade genômica e acúmulo de mutações. Esse cenário reforça a noção de que, embora a endometriose seja uma doença benigna, pode apresentar características pró-neoplásicas em alguns subtipos, como os endometriomas ovarianos, especialmente em casos com longo tempo de evolução (ANGLESIO et al., 2015; KLEMMT; STARZINSKI-POWITZ, 2018).

Além disso, as vias de metabolismo proteico e modificação pós-traducional, como “Modificação pós-traducional de proteínas” e “proteólise mediada por ubiquitina”, também se mostraram inibidas. A disfunção nesse eixo afeta a renovação proteica, interfere na regulação de receptores de membrana e prejudica respostas ao estresse celular, favorecendo a sobrevivência de células disfuncionais. Do ponto de vista terapêutico, essas vias podem representar alvos alternativos ainda pouco explorados no contexto da endometriose. Outro ponto relevante foi a regulação negativa de vias de sinalização celular e supressão tumoral, como IL-24 Signaling e Transcriptional Regulation by TP53. A proteína p53 atua como regulador da apoptose, controle do ciclo celular e estabilidade genômica. Sua inibição pode refletir um ambiente permissivo à proliferação anormal e resistência à morte celular, fatores que contribuem para a persistência das lesões endometrióticas e sua progressão (ZHANG



et al., 2020).

Adicionalmente, os genes superexpressos foram associados principalmente à ativação de vias relacionadas à sinalização celular, proliferação, angiogênese, resposta inflamatória e biossíntese de estrogênio. Foram observadas ativações de sinalização de AKT, cascata de MAPK, cascata de ERK1/2, sinalização de IL-2 e IL-9 e biossíntese de estrogênio. Esses dados são compatíveis com a fisiopatologia da doença, que envolve um ambiente local inflamatório, hiperproliferativo e estrogênio-dependente (BULUN et al., 2019; VIGANO et al., 2018). O enriquecimento de termos como angiogenesis, migração celular e cicatrização, também sustenta o papel do microambiente remodelador na manutenção das lesões, corroborando com o descrito na literatura.

No que se refere a dor na endometriose, observou-se a ativação significativa de vias associadas à inflamação e ao remodelamento tecidual, como TNF- $\alpha$ /NF- $\kappa$ B, MAPK, Ras e KRAS, todas implicadas na amplificação de sinais pró-inflamatórios e na modulação da resposta nociceptiva. Essas vias promovem a liberação de mediadores inflamatórios como IL-1 $\beta$ , IL-6, TNF- $\alpha$  e prostaglandinas, além de fatores neurotróficos como o NGF (Nerve Growth Factor), os quais favorecem a inervação sensorial aberrante e a sensibilização de nociceptores periféricos (ASANTE et al., 2011; CHEN et al., 2025). A ativação de processos como transição epitelial-mesenquimal (EMT) e angiogênese também sugere a criação de um microambiente permissivo à neuroangiogênese, fenômeno que contribui para o aumento da densidade nervosa nas lesões, correlacionando-se clinicamente com maior intensidade de dor (AGOSTINIS et al., 2021; STRATTON & BERKLEY et al., 2011). Esses achados reforçam a hipótese de que a dor na endometriose decorre não apenas da presença anatômica das lesões, mas sobretudo de sua atividade molecular inflamatória e neurogênica, o que justifica a variabilidade clínica entre mulheres com lesões semelhantes. Nesse contexto, o perfil de expressão gênica pode oferecer subsídios para estratificação fenotípica da dor e direcionamento de terapias alvo.

Os resultados encontrados no contexto da infertilidade na endometriose corroborou com o que se tem na literatura, onde desregulação de vias de reparo do DNA (*PARP1*, *RAD51C*) compromete a estabilidade genômica, prejudicando a

qualidade oocitária e embrionária. Simultaneamente, o microambiente inflamatório, mediado por NF- $\kappa$ B, MAPK/ERK e citocinas (IL-6, TNF- $\alpha$ ), cria condições adversas para implantação e desenvolvimento embrionário. O estresse oxidativo, exacerbado pelo acúmulo de ferro e ativação de ERK1/2/Nrf2, danifica tanto o tecido ovariano quanto endometrial. Adicionalmente, alterações na sinalização estrogênio-dependente e na apoptose (via PI3K/AKT/TP53) perturbam a decidualização e renovação endometrial. Esses mecanismos combinados explicam a infertilidade associada à doença e destacam a necessidade de terapias multifocais que restauram a homeostase reprodutiva (ZHANG et al., 2023; QI et al., 2025). A regulação negativa das vias imunes pode contribuir para a infertilidade associada à endometriose, pela falha na renovação endometrial e na interação materno-fetal (GIUDICE; KAISER, 2006), já no que se refere à supressão de vias de reparo e apoptose pode justificar a resistência terapêutica observada em portadoras da endometriose, especialmente à terapia hormonal.

A interseção entre plataformas identificou genes com expressão consistente, como *CFH*, *GAS6*, *TSPAN4*, *WNT2B* e *PLSCR4* (up), além de *PIGN*, *ITGB8*, *PFAS*, *KLHL13* e *PARP1* (down). A recorrência consistente desses 10 genes em múltiplos estudos e plataformas de diferentes tecnologias, reforça sua relevância como marcadores moleculares estáveis da endometriose, com potencial aplicação em diagnósticos moleculares, prognóstico ou mesmo como alvos terapêuticos. A identificação de alvos moleculares consistentes oferece caminhos para o desenvolvimento de biomarcadores diagnósticos e terapias personalizadas, em especial aquelas baseadas na modulação imune, no controle do ciclo celular ou na reparação do DNA.

Na literatura atual, uma busca criteriosa por estudos com metodologia semelhante à adotada nesta dissertação resultou na identificação de apenas três metanálises relevantes. Kang (2025) foi o único estudo a buscar um perfil de expressão gênica integrando diferentes datasets de endometriose; no entanto, sua análise restringiu-se a apenas três conjuntos de dados de microarranjo (GSE6364, GSE73622-MS, GSE73622-SF e GSE141549-GPL13376), dos quais dois foram também incorporados no presente estudo. A segunda metanálise identificada, realizada por Poli-Neto (2020), limitou sua abordagem a dados obtidos exclusivamente

por microarranjo, com foco direcionado a transcritos associados a processos imunológicos, não explorando outras vias potencialmente envolvidas na fisiopatologia da doença. Por fim, Vargas et al. (2022) realizaram uma integração metodológica entre microarranjo e RNA-Seq, validando parte dos achados; contudo, direcionaram sua investigação unicamente para aspectos relacionados à infertilidade em mulheres com endometriose.

Dessa forma, o presente trabalho apresenta um diferencial significativo ao propor uma abordagem integrativa mais abrangente, tanto na seleção criteriosa e ampliada dos datasets utilizados, quanto na robustez da análise integrada entre plataformas distintas (microarranjo e RNA-Seq). A estratégia aqui aplicada permitiu a identificação de um painel consistente de genes diferencialmente expressos, obtido por consenso entre as diferentes metodologias, ampliando o potencial de identificação de marcadores moleculares robustos associados à endometriose em seus múltiplos contextos fisiopatológicos.

Apesar dos avanços obtidos, o presente estudo apresenta algumas limitações inerentes à heterogeneidade dos dados disponíveis nos bancos públicos, incluindo a ausência ou incompletude de informações clínicas detalhadas, o que dificultou a estratificação das amostras segundo os distintos fenótipos e apresentações clínicas da endometriose. Tal limitação inviabilizou, neste momento, análises diferenciais entre subgrupos clínicos nos diversos estudos integrados. Como perspectiva futura, pretende-se validar experimentalmente os principais genes diferencialmente expressos identificados, bem como integrar dados de lncRNAs aos resultados obtidos, com o objetivo de ampliar a compreensão dos mecanismos epigenéticos envolvidos na regulação gênica associada à doença.

Os achados aqui obtidos têm importantes implicações clínicas. A integração dos dados de RNA-Seq corroborou de maneira consistente a maioria dos achados provenientes da análise de *microarranjos*, mesmo diante das diferenças metodológicas e da variabilidade dos dados. Este achado reforça a robustez dos genes identificados como potenciais biomarcadores moleculares da endometriose, independentemente da plataforma tecnológica utilizada, e destaca o potencial dessas

assinaturas no aprimoramento do diagnóstico e compreensão dos mecanismos fisiopatológicos da doença.

## 7. Conclusão

- Foi realizada uma metanálise abrangente de dados transcriptômicos públicos de endometriose, integrando plataformas de microarranjo e RNA-Seq, totalizando mais de 1.200 amostras inicialmente curadas e posteriormente filtradas conforme rigorosos critérios de qualidade.

- A integração dos dados permitiu a identificação de uma assinatura gênica robusta composta por 205 genes diferencialmente expressos na metanálise de microarranjo, dos quais 132 foram regulados positivamente e 73 negativamente.

- A validação integrada com os dados de RNA-Seq confirmou um painel de 57 genes em consenso entre as plataformas, destacando um conjunto de 10 genes recorrentes (CFH, GAS6, TSPAN4, WNT2B, PLSCR4 — upregulados; e PIGN, ITGB8, PFAS, KLHL13, PARP1 — downregulados), considerados potenciais biomarcadores moleculares estáveis da endometriose.

- A análise de enriquecimento funcional dos genes upregulados revelou participação em processos relacionados à proliferação celular, inflamação, angiogênese, biossíntese de estrogênio, neurogênese e cicatrização, aspectos consistentes com a fisiopatologia proliferativa e inflamatória da endometriose.

- Já os genes downregulados mostraram supressão de vias ligadas ao reparo de DNA, escape imunológico, instabilidade genômica, checkpoint mitótico e resposta imune adaptativa, indicando possível comprometimento da homeostase genômica e imunológica nas lesões endometrióticas.

- Os achados reforçam o papel de múltiplos mecanismos fisiopatológicos interconectados na doença, incluindo processos inflamatórios crônicos, desregulação hormonal, remodelamento tecidual, alterações imunes e prejuízo no reparo do DNA, que podem explicar não apenas a progressão da doença, mas também sintomas como infertilidade e dor pélvica crônica.

- O estudo destacou a importância da integração multiplataforma na superação da heterogeneidade técnica e biológica dos dados transcriptômicos públicos, permitindo a identificação de alvos moleculares consistentes, com relevância potencial tanto diagnóstica quanto terapêutica.

- Como limitação, destaca-se a escassez de informações clínicas detalhadas nos *datasets* públicos, dificultando a estratificação por fenótipos clínicos de endometriose e impedindo, neste momento, análises específicas por subgrupos.

- Como perspectivas futuras, propõe-se a validação experimental dos genes identificados, a incorporação de dados de lncRNAs e a exploração funcional dos biomarcadores encontrados, visando aprofundar a compreensão dos mecanismos epigenéticos e abrir caminhos para o desenvolvimento de terapias personalizadas.

## 8. Referências

- ABRAMIUK, M.; GRYWALSKA, E.; MAŁKOWSKA, P.; SIERAWSKA, O.; HRYNKIEWICZ, R.; NIEDŹWIEDZKA-RYSTWEJ, P. The role of the immune system in the development of endometriosis. *Cells*, Basel, v. 11, n. 13, p. 2028, 2022. Disponível em: <https://doi.org/10.3390/cells11132028>.
- ACOG - AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGISTS. *ACOG Committee Opinion No. 444: Management of endometriosis*. *Obstetrics & Gynecology*, v. 116, n. 1, p. 223-236, 2010.
- AGHAJANOVA, L. et al. Altered Gene Expression Profiling in Endometrium: Evidence for Progesterone Resistance. *Semin Reprod Med* 2010; 28:51-58. DOI:<http://dx.doi.org/10.1055/s-0029-1242994>.
- AGHAJANOVA, L., BURNEY, R.O., TRAN, N.D., GIUDICE, L.C. (2017). mRNA and miRNA Biomarkers for Endometriosis. In: D'Hooghe, T. (eds) *Biomarkers for Endometriosis*. Springer, Cham. [https://doi.org/10.1007/978-3-319-59856-7\\_9](https://doi.org/10.1007/978-3-319-59856-7_9)
- AGOSTINIS, C. et al. Immunological basis of the endometriosis: the complement system as a potential therapeutic target. *Frontiers in Immunology*, v. 11, art. 599117, 2021. Disponível em: <https://doi.org/10.3389/fimmu.2020.599117>.
- AHN, S. H. et al. Pathophysiology and immune dysfunction in endometriosis. *Biomed Res Int*, 2015, 2015: 795976. <https://doi.org/10.1155/2015/795976>
- AL-FADHLI, R. et al. Effects of different stages of endometriosis on the outcome of in vitro fertilization. *Journal of Obstetrics and Gynaecology Canada*, v. 28, n. 10, p. 888-891, 2006. DOI:10.1016/S1701-2163(16)32279-7
- ANDERS, S. et al. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, v. 31, n. 2, p. 166-169, 2015. <https://doi.org/10.1093/bioinformatics/btu638>
- ANDREWS, S. et al. FastQC: a quality control tool for high throughput sequence data. 2010. Disponível em: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- ANGLESIO, M. S. et al. Cancer-associated mutations in endometriosis without cancer. *New England Journal of Medicine*, v. 376, n. 19, p. 1835-1848, maio 2017. Disponível em: <https://doi.org/10.1056/NEJMoa1614814>.
- AMRO, B. et al. New understanding of diagnosis, treatment and prevention of endometriosis. *International Journal of Environmental Research and Public Health*, v. 19, n. 11, p. 6725, 2022. DOI: <https://doi.org/10.3390/ijerph19116725>.
- ARAÚJO, F. C.; SCHMIDT, D. B. Endometriose um problema de saúde pública: revisão de literatura. *Revista Saúde e Desenvolvimento*, v. 14, n. 18, 2020. Disponível em: <https://www.revistasuninter.com/revistasauade/index.php/saudeDesenvolvimento/>

article/view/1021.

ASANTE, Albert; TAYLOR, Robert N. Endometriosis: the role of neuroangiogenesis. *Annual review of physiology*, v. 73, n. 1, p. 163-182, 2011.

BARNHART, K. et al. Effect of endometriosis on in vitro fertilization. *Fertility and Sterility*, v. 77, n. 6, p. 1148-1155, 2002. DOI:10.1016/S0015-0282(02)03112-6

BARRETT, T. et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research*, v. 41, n. D1, p. D991-D995, 2013. <https://doi.org/10.1093/nar/gks1193>

BAZOT, M. et al. Diagnostic accuracy of physical examination, transvaginal sonography, rectal endoscopic sonography, and magnetic resonance imaging to diagnose deep infiltrating endometriosis. *Fertility and Sterility*, v. 92, n. 6, p. 1825-1833, 2017. <https://doi.org/10.1016/j.fertnstert.2008.09.005>

BÉLIARD, A. et al. Reduction of apoptosis and proliferation in endometriosis. *Fertility and Sterility*, v. 82, n. 1, p. 80-85, 2004. DOI:10.1016/j.fertnstert.2003.11.037

BELLELIS, P. et al. Epidemiological and clinical aspects of pelvic endometriosis - a case series. *Revista da Associação Médica Brasileira*, v. 56, n. 4, p. 467-471, 2010.

DOI:10.1590/S0104-42302010000400022

BERKKANOGLU, M.; ARICI, A. Immunology and endometriosis. *American Journal of Reproductive Immunology*, v. 50, n. 1, p. 48–59, 2003. DOI: <https://doi.org/10.1034/j.1600-0897.2003.00031.x>.

BERLANDA, N. et al. Surgery versus hormonal therapy for deep endometriosis: is it a choice of the physician. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, v. 209, p. 67-71, 2017. DOI:10.1016/j.ejogrb.2016.12.007

BIRMINGHAM, A. Practice Committee of the American Society for Reproductive Medicine. Endometriosis and infertility. *Fertility and Sterility*, v. 86, n. 5, suppl. 1, p. S156-S160, 2006. DOI:10.1016/j.fertnstert.2006.08.013

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, v. 30, n. 15, p. 2114–2120, ago. 2014.

Disponível em: <https://doi.org/10.1093/bioinformatics/btu170>.

BORGHESE, Bruno et al. Research Resource: Gene Expression Profile for Ectopic Versus Eutopic Endometrium Provides New Insights into Endometriosis Oncogenic Potential. *Molecular Endocrinology*, v. 22, n. 11, p. 2557–2562, nov. 2008. Disponível em: <https://doi.org/10.1210/me.2008-0322>.

BROWN, Joseph; PIRRUNG, Meg; McCUE, Lee Ann. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool.

*Bioinformatics*, v. 33, n. 19, p. 3137-3139, out. 2017. Disponível em:



<https://doi.org/10.1093/bioinformatics/btx373>.

BULUN, S. E. et al. Endometriosis. *Endocrine Reviews*, v. 40, n. 4, p. 1048-1079, 2019. DOI:10.1210/er.2018-00242

BURNEY, R. O.; GIUDICE, L. C. Pathogenesis and pathophysiology of endometriosis. *Fertility and Sterility*, v. 98, n. 3, p. 511-519, 2012. DOI:10.1016/j.fertnstert.2012.06.029

CADORE, N. A., LORD, V. O., RECAMONDE-MENDOZA, M., KOWALSKI, T. W., & VIANNA, F. S. L. (2024). Meta-analysis of Transcriptomic Data from Lung Autopsy and Cellular Models of SARS-CoV-2 Infection. *Biochemical Genetics*, 62(2), 892–914. <https://doi.org/10.1007/s10528-023-10453-2>

CALDEIRA, R. P. et al. Tratamento terapêutico multi-profissional para endometriose com dor pélvica. *Universitas: Ciências da Saúde*, v. 6, n. 1, p. 69-83, 2008. Disponível em: <http://www.publicacoesacademicas.uniceub.br/index.php/cienciasaude/article/view/644>

CAHAN, Patrick et al. Meta-analysis of microarranjo results: challenges, opportunities, and recommendations for standardization. *Gene*, v. 401, n. 1-2, p. 12-18, 2007. Disponível em: <https://doi.org/10.1016/j.gene.2007.06.016>.

CARVALHO, B. S. et al. Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. *Biostatistics*, v. 8, n. 2, p. 485-499, 2007. <https://doi.org/10.1093/biostatistics/kxl042>

CHAPRON, Charles; MARCELLIN, Louis; BORGESE, Bruno; SANTULLI, Pietro. Rethinking mechanisms, diagnosis and management of endometriosis. *Nature Reviews Endocrinology*, [S.l.], v. 15, n. 11, p. 666–682, nov. 2019. DOI: <https://doi.org/10.1038/s41574-019-0245-z>.

CHEN, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, v. 14, n. 1, p. 128, 2013. <https://doi.org/10.1186/1471-2105-14-128>

CHEN, Y.; LI, T. Unveiling the mechanisms of pain in endometriosis: comprehensive analysis of inflammatory sensitization and therapeutic potential. *International Journal of Molecular Sciences*, v. 26, n. 4, p. 1770, 2025. Disponível em: <https://doi.org/10.3390/ijms26041770>.

CHOI, J. K. et al. Statistical methods for meta-analysis of microarranjo data: a comparative review. *Briefings in Bioinformatics*, v. 18, n. 2, p. 204-213, 2017. <https://doi.org/10.1093/bib/bbw060>

CLOUGH, E.; BARRETT, T. The Gene Expression Omnibus database. *Methods in Molecular Biology*, v. 1418, p. 93-110, 2016. <https://doi.org/10.1007/978-1->

4939-3578-9\_5

CONESA, A. et al. A survey of best practices for *RNA-seq* data analysis. *Genome Biology*, v. 17, n. 1, p. 13, 2016. <https://doi.org/10.1186/s13059-016-0881-8>

DAVIS, S; MELTZER, P. S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*, v. 23, n. 14, p. 1846–1847, jul. 2007. Disponível em: <https://doi.org/10.1093/bioinformatics/btm254>.

DE ABREU, L. G. et al. Reduced aromatase activity in granulosa cells of women with endometriosis undergoing assisted reproduction techniques. *Gynecological Endocrinology*, v. 22, n. 8, p. 432-436, 2006. DOI:10.1080/09513590600890266

DE SOUSA BARBOSA, D. A. et al. Endometriose e seu impacto na fertilidade feminina. *Saúde & Ciência em Ação*, v. 1, n. 1, p. 43-56, 2015. Disponível em: <http://periodicos.uea.edu.br/index.php/cienciaemsaude/article/view/333>

DMOWSKI, W. P. et al. Apoptosis in endometrial glandular and stromal cells in women with and without endometriosis. *Human Reproduction*, v. 16, n. 7, p. 1802-1808, 2001. DOI:10.1093/humrep/16.7.1802

DOBIN, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, v. 29, n. 1, p. 15-21, 2013. <https://doi.org/10.1093/bioinformatics/bts635>

DONNEZ, J. et al. Vascular endothelial growth factor (VEGF) in endometriosis. *Human Reproduction*, v. 13, n. 6, p. 1686-1690, 1998. DOI:10.1093/humrep/13.6.1686

DUNSELMAN, G. A. J. et al. ESHRE guideline: management of women with endometriosis. *Human Reproduction*, Oxford, v. 29, n. 3, p. 400–412, mar. 2014. Disponível em: <https://doi.org/10.1093/humrep/det457>.

DYSON, M. T. et al. Genome-wide DNA methylation analysis predicts an epigenetic switch for GATA factor expression in endometriosis. *PLoS Genetics*, v. 10, n. 3, e1004158, 2014. <https://doi.org/10.1371/journal.pgen.1004158>

EDGAR, R. et al. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, v. 30, n. 1, p. 207-210, 2002. <https://doi.org/10.1093/nar/30.1.207>

EYSTER, K. M. et al. Whole genome deoxyribonucleic acid microarranjo analysis of gene expression in ectopic versus eutopic endometrium. *Fertility and Sterility*, v. 88, n. 6, p. 1505-1533, 2007. DOI:10.1016/j.fertnstert.2007.01.056

FASSBENDER, A. et al. Biomarkers of endometriosis. *Fertility and Sterility*, v. 99, n. 4, p. 1135-1145, 2013. DOI:10.1016/j.fertnstert.2013.01.097

FASSBENDER, A. et al. Update on biomarkers for the detection of

endometriosis. *BioMed Research International*, v. 2015, 130854, 2015.  
<https://doi.org/10.1155/2015/130854>

FREITAS, F. et al. *Rotinas em ginecologia*. 6. ed. Porto Alegre: Artmed, 2011. p. 144.

FUKUNAGA, M. Uterus-like mass in the uterine cervix: superficial cervical endometriosis with florid smooth muscle metaplasia? *Virchows Archiv*, v. 438, n. 3, p. 302-305, 2001. DOI:10.1007/s004280000318

GARCIA, A. M. C. et al. Endometriose cutânea umbilical: relato de caso e revisão de literatura. *Arquivos Catarinenses de Medicina*, v. 38, n. 1, p. 254-256, 2009. Disponível em: <http://www.acm.org.br/revista/pdf/artigos/656.pdf>

GAUTIER, L. et al. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, v. 20, n. 3, p. 307-315, 2004.  
<https://doi.org/10.1093/bioinformatics/btg405>

GENTLEMAN, R. C. et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, v. 5, n. 10, R80, 2004. <https://doi.org/10.1186/gb-2004-5-10-r80>

GIUDICE, L. C. et al. The molecular basis for implantation failure in endometriosis: on the road to discovery. *Annals of the New York Academy of Sciences*, v. 955, p. 252-264, 2002. DOI:10.1111/j.1749-6632.2002.tb02784.x

GIUDICE, L. C.; KAO, L. C. Endometriosis. *The Lancet*, v. 364, n. 9447, p. 1789-1799, 2004. DOI:10.1016/S0140-6736(04)17403-5

GIUDICE, L. C.; KAISER, L. R. Endometriosis. In: Jameson, J. L. et al. *Principles and Practice of Endocrinology and Metabolism*. 3. ed. Philadelphia: Lippincott Williams & Wilkins, 2006.

GIUDICE, L. C. Endometriosis. *New England Journal of Medicine*, v. 362, n. 25, p. 2389-2398, 2010. DOI:10.1056/NEJMcp1000274

GONÇALVES, KATIA CRISTIANE BRUMATTI. Estudo de bioinformática aplicado à análise de expressão gênica utilizando dados oriundos de sequenciamento por tecnologia de "Next-Generation" em animais controle e em modelos de epilepsia do lobo temporal mesial. 2015. 142 f. Dissertação (Mestrado em Ciências Médicas) - Faculdade de Ciências Médicas, Universidade Estadual de Campinas, Campinas, 2015.

GUINDALINI, C.; TUFIK, S. Uso de microarranjos na busca de perfis de expressão gênica: aplicação no estudo de fenótipos complexos. *Brazilian Journal of Psychiatry*, v. 29, n. 4, dez. 2007. Disponível em:  
<https://doi.org/10.1590/S1516-44462007000400014>.

GUERRIERO, S. et al. Transvaginal ultrasound vs magnetic resonance imaging for diagnosing deep infiltrating endometriosis: systematic review and meta-

analysis. *Ultrasound in Obstetrics & Gynecology*, v. 51, n. 5, p. 586-595, 2021  
 GRIMSTAD, FRANCES W. MD\*; DECHERNEY, ALAN MD†. A Review of the Epigenetic Contributions to Endometriosis. *Clinical Obstetrics and Gynecology* 60(3):p 467-476, September 2017. | DOI: 10.1097/GRF.0000000000000298  
 GUO, S. W. Epigenetics of endometriosis. *Molecular Human Reproduction*, v. 15, n. 10, p. 587-607, 2009. DOI:10.1093/molehr/gap063

HASIN, Y. et al. Multi-omics approaches to disease. *Genome Biology*, v. 18, n. 1, p. 83, 2017. <https://doi.org/10.1186/s13059-017-1215-1>

HAYNES, W. A. et al. Empowering multi-cohort gene expression analysis to increase reproducibility. *Pacific Symposium on Biocomputing*, v. 22, p. 144-153, 2017. [https://doi.org/10.1142/9789813207813\\_0015](https://doi.org/10.1142/9789813207813_0015)

HOLOCH, K. J.; LESSEY, B. A. Endometriosis and infertility. *Clinical Obstetrics and Gynecology*, v. 53, n. 2, p. 429-438, 2010. DOI:10.1097/GRF.0b013e3181db7d71

HONG, F.; BREITLING, R. A comparison of meta-analysis methods for detecting differentially expressed genes in microarranjo experiments. *Bioinformatics*, v. 24, n. 3, p. 374-382, 2008. <https://doi.org/10.1093/bioinformatics/btm620>

HONDA, H.; BARRUETO, F. F.; GOGUSEV, J. et al. Análise serial da expressão gênica revela expressão diferencial entre endometriose e endométrio normal: possíveis papéis para AXL e SHC1 na patogênese da endometriose. *Reproductive Biology and Endocrinology*, v. 6, p. 59, 2008. Disponível em: <https://doi.org/10.1186/1477-7827-6-59>.

HUBER, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, v. 12, n. 2, p. 115-121, 2015. <https://doi.org/10.1038/nmeth.3252>

IMPERIALE L, NISOLLE M, NOËL JC, FASTREZ M. Three Types of Endometriosis: Pathogenesis, Diagnosis and Treatment. *State of the Art. J Clin Med*. 2023 Jan 28;12(3):994. doi: 10.3390/jcm12030994.

IRIZARRY, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, v. 4, n. 2, p. 249-264, 2003. <https://doi.org/10.1093/biostatistics/4.2.249>

JOHNSON, N. P.; HUMMELSHOJ, L.; ABRAO, M. S.; ADAMSON, G. D.; ALLAIRE, C.; AMELUNG, V. et al. Consensus on current management of endometriosis. *Human Reproduction, Oxford*, v. 28, n. 6, p. 1552–1568, jun. 2013. Disponível em: <https://doi.org/10.1093/humrep/det050>. Acesso em: 16 jun. 2025.

KANG, J. et al. Identification of Endometriosis Pathophysiologic-Related Genes Based on Meta-Analysis and Bayesian Approach. *International Journal of Molecular Sciences*, [S. l.], v. 26, n. 1, p. 424, 2025. <https://doi.org/10.3390/ijms26010424>.

KAUFFMANN, A GENTLEMAN, R; HUBER, W. arrayQualityMetrics—a bioconductor package for quality assessment of microarranjo data. *Bioinformatics*, [S. l.], v. 25, n. 3, p. 415-416, 2009. Disponível em: <https://doi.org/10.1093/bioinformatics/btn647>.

KENNEDY, S. et al. ESHRE guideline for the diagnosis and treatment of endometriosis. *Human Reproduction*, v. 20, n. 10, p. 2698-2704, 2005. DOI:10.1093/humrep/dei135

KIM, D. et al. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, v. 37, n. 8, p. 907-915, 2019. <https://doi.org/10.1038/s41587-019-0201-4>

KLEMMT, P. A. B.; STARZINSKI-POWITZ, A. Molecular and cellular pathogenesis of endometriosis. *Current Women's Health Reviews*, v. 14, n. 2, p. 106–116, 2018.

KONINCKX, Philippe R.; USSIA, Anastasia; ADAMYAN, Leila; WATTIEZ, Arnaud; DONNEZ, Jacques. Deep endometriosis: definition, diagnosis, and treatment. *Fertility and Sterility*, v. 98, n. 3, p. 564–571, 2012. Disponível em: <https://doi.org/10.1016/j.fertnstert.2012.07.1061>.

KULESHOV, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, v. 44, n. W1, p. W90-W97, 2016. <https://doi.org/10.1093/nar/gkw377>

LAW, C. W. et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research*, v. 5, 1408, 2016. <https://doi.org/10.12688/f1000research.9005.1>

LEBOVIC, D. I. et al. Immunobiology of endometriosis. *Fertility and Sterility*, v. 75, n. 1, p. 1-10, 2001. DOI:10.1016/S0015-0282(00)01630-7

LEINONEN, Rasko; SUGAWARA, Hideaki; SHUMWAY, Martin; em nome da International Nucleotide Sequence Database Collaboration. The Sequence Read Archive. *Nucleic Acids Research*, v. 39, supl. 1, p. D19–D21, 1 jan. 2011. Disponível em: <https://doi.org/10.1093/nar/gkq1019>.

LI, Q.; TSENG, G. C. An integrative meta-analysis approach for identifying differentially expressed genes across microarranjo studies. *Bioinformatics*, v. 27, n. 6, p. 843-850, 2011. <https://doi.org/10.1093/bioinformatics/btr002>

LIAO, Y. et al. featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. *Bioinformatics*, v. 30, n. 7, p. 923-930, 2014. <https://doi.org/10.1093/bioinformatics/btt656>

LOVE, M. I. et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, v. 15, n. 12, p. 550, 2014. <https://doi.org/10.1186/s13059-014-0550-8>

LUO, S. et al. MiR-301a-3p Advances IRAK1-Mediated Differentiation of Th17

Cells to Promote the Progression of Systemic Lupus Erythematosus via Targeting PELI1. *Journal of Healthcare Engineering*, v. 2021, 5599523, 2021. DOI:10.1155/2021/5599523

MÁRQUEZ, A. et al. Meta-analysis of immunochip data of four autoimmune diseases reveals novel single-disease and cross-phenotype associations. *Genome Medicine*, v. 8, n. 1, p. 97, 2018. <https://doi.org/10.1186/s13073-016-0350-8>

MARQUI, A. B. T. Endometriose: do diagnóstico ao tratamento. *Revista Enfermagem Atenção à Saúde*, v. 3, n. 1, p. 97-105, 2014. Disponível em: <http://seer.uftm.edu.br/revistaeletronica/index.php/enfer/article/view/1041>

MATSUZAKI, S. et al. DNA microarranjo analysis of gene expression profiles in deep endometriosis using laser capture microdissection. *Molecular Human Reproduction*, v. 10, n. 10, p. 719-728, 2004. DOI:10.1093/molehr/gah097

MATSUZAKI, Sachiko; DARCHA, Claude. Co-operation between the AKT and ERK signaling pathways may support growth of deep endometriosis in a fibrotic microenvironment in vitro. *Human Reproduction*, v. 30, n. 7, p. 1606–1616, jul. 2015. Disponível em: <https://doi.org/10.1093/humrep/dev108>.

MCCARTHY, D. J. et al. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. *Nucleic Acids Research*, v. 40, n. 10, p. 4288-4297, 2012. <https://doi.org/10.1093/nar/gks042>

McLaren J. Vascular endothelial growth factor and endometriotic angiogenesis. *Hum Reprod Update*. 2000;6(1):45–55.

MEHEDINTU, C. et al. Endometriosis still a challenge. *Journal of Medicine and Life*, v. 7, n. 3, p. 349-357, 2014. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4233437/>

MISRA, B. B.; LANGEFELD, C.; OLIVIER, M.; COX, L. A. Integrated omics: tools, advances and future approaches. *Journal of Molecular Endocrinology*, v. 62, n. 1, p. R21-R45, 2019. Disponível em: <https://doi.org/10.1530/JME-18-0055>.

MOROTTI, M.; VINCENT, K.; BECKER, C. M. Mechanisms of pain in endometriosis. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, v. 209, p. 8-13, 2017. Disponível em: <https://doi.org/10.1016/j.ejogrb.2016.07.497>.

MUTZ, Kai-Oliver et al. Transcriptome analysis using next-generation sequencing. *Current Opinion in Biotechnology*, v. 24, n. 1, p. 22–30, 2013. DOI: 10.1016/j.copbio.2012.09.004.

NÁCUL, A. P.; SPRITZER, P. M. Aspectos atuais do diagnóstico e tratamento da

endometriose. *Revista Brasileira de Ginecologia e Obstetrícia*, v. 32, n. 6, p. 298-307, 2010. DOI:10.1590/S0100-72032010000600006

NASU, K. et al. Aberrant DNA methylation status in endometriosis: epigenetics as the pathogenesis, biomarker and therapeutic target. *Journal of Obstetrics and Gynaecology Research*, v. 37, n. 7, p. 683-695, 2011. DOI:10.1111/j.1447-0756.2011.01663.

NAP, Annemiek W. et al. Tissue integrity is essential for ectopic implantation of human endometrium in the chicken chorioallantoic membrane. *Human Reproduction*, Oxford, v. 18, n. 1, p. 30–34, jan. 2003. Disponível em: <https://doi.org/10.1093/humrep/deg033>.

NICE - NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE. *Endometriosis: diagnosis and management*. NICE Guideline NG73, 2017.

NNOAHAM, K. E. et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertility and Sterility*, v. 96, n. 2, p. 366-373, 2011. DOI:10.1016/j.fertnstert.2011.05.090

PAPPALARDO, Francesco; RUSSO, Giulia; TSHINANU, Flora Musuamba; VICECONTI, Marco. In silico clinical trials: concepts and early adoptions. *Briefings in Bioinformatics*, Oxford, v. 20, n. 5, p. 1699–1708, set. 2019. Disponível em: <https://doi.org/10.1093/bib/bby043>.

PENG, Ying; PENG, Cheng; FANG, Zheng; CHEN, Gang. Bioinformatics Analysis Identifies Molecular Markers Regulating Development and Progression of Endometriosis and Potential Therapeutic Drugs. *Frontiers in Genetics*, v. 12, 2021. Disponível em: <https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2021.622683>.

POLI-NETO, O. B. et al. Transcriptome meta-analysis reveals differences of immune profile between eutopic endometrium from stage I-II and III-IV endometriosis independently of hormonal milieu. *Scientific Reports*, [S. l.], v. 10, n. 1, p. 313, 2020. Disponível em: <https://doi.org/10.1038/s41598-019-57207-y>.

QI, Q. , LI Y, CHEN Z, LUO Z, ZHOU T, ZHOU J, ZHANG Y, CHEN S AND WANG L. Update on the pathogenesis of endometriosis-related infertility based on contemporary evidence. *Frontiers in Endocrinology*, v. 16, p. 1558271, 2025. Disponível em: <https://doi.org/10.3389/fendo.2025.1558271>.

RAMASAMY, A. et al. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Medicine*, v. 5, n. 9, e184, 2008. <https://doi.org/10.1371/journal.pmed.0050184>  
REUTER, J. A. et al. High-throughput sequencing technologies. *Molecular Cell*, v. 58, n. 4, p. 586-597, 2015. <https://doi.org/10.1016/j.molcel.2015.05.004>

RITCHIE, M. E. et al. limma powers differential expression analyses for RNA-

sequencing and microarranjo studies. *Nucleic Acids Research*, v. 43, n. 7, e47, 2015. <https://doi.org/10.1093/nar/gkv007>

ROBINSON, M. D. et al. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, v. 26, n. 1, p. 139-140, 2010. <https://doi.org/10.1093/bioinformatics/btp616>

ROCHA-JUNIOR, Carlos Valério et al. Progesterone receptor B (PGR-B) is partially methylated in eutopic endometrium from infertile women with endometriosis. *Reproductive Sciences*, [S. l.], v. 26, n. 12, p. 1568–1574, 2019. DOI: 10.1177/1933719119828078.

ROGERS, P. A. W. et al. Priorities for endometriosis research: recommendations from an international consensus workshop. *Reproductive Sciences*, v. 16, n. 4, p. 335-346, 2009. DOI:10.1177/1933719108330568

SIGNORILE, P. G.; BALDI, A. Endometriosis: new concepts in the pathogenesis. *International Journal of Biochemistry and Cell Biology*, v. 42, n. 6, p. 778-780, 2010. DOI:10.1016/j.biocel.2010.03.008

STARK, R. et al. RNA sequencing: the teenage years. *Nature Reviews Genetics*, v. 20, n. 11, p. 631-656, 2019. <https://doi.org/10.1038/s41576-019-0150-2>

STRATTON, Pamela; BERKLEY, Karen J. Chronic pelvic pain and endometriosis: translational evidence of the relationship and implications. *Human Reproduction Update*, v. 17, n. 3, p. 327–346, maio/jun. 2011. Disponível em: <https://doi.org/10.1093/humupd/dmq050>.

SUBRAMANIAN, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, v. 102, n. 43, p. 15545-15550, 2005. <https://doi.org/10.1073/pnas.0506580102>

SUBRAMANIAN, I. et al. Multi-omics data integration, interpretation, and its application. *Bioinformatics and Biology Insights*, v. 14, p. 1-24, 2020. <https://doi.org/10.1177/1177932219899051>

SUDA, KAZUAKI et al. Clonal expansion and diversification of cancer-associated mutations in endometriosis and normal endometrium. *Cell Reports*, v. 24, n. 7, p. 1777–1789, 2018. Disponível em: <https://doi.org/10.1016/j.celrep.2018.07.003>.

TAMINAU, J. et al. Unlocking the potential of publicly available microarranjo data using inSilicoDb and inSilicoMerging R/Bioconductor packages. *BMC Bioinformatics*, v. 13, n. 1, p. 1-10, 2012. <https://doi.org/10.1186/1471-2105-13-335>

TAYLOR, Hugh S.; KOTLYAR, Alexander M.; FLORES, Valerie A. Endometriosis is a chronic systemic disease: clinical challenges and novel innovations. *The Lancet*, v. 397, n. 10289, p. 839-852, 2021. Disponível em: <https://thelancet.com>.



TOKUSHIGE, N.; MARKHAM, R.; RUSSELL, P.; FRASER, I. S. High density of small nerve fibres in the functional layer of the endometrium in women with endometriosis. *Human Reproduction*, Oxford, v. 21, n. 3, p. 782–787, mar. 2006. Disponível em: <https://doi.org/10.1093/humrep/dei368>.

TSENG, G. C. et al. Comprehensive literature review and statistical considerations for microarranjo meta-analysis. *Nucleic Acids Research*, v. 40, n. 9, p. 3785-3799, 2012. <https://doi.org/10.1093/nar/gks1265>

VANNUCCINI, S. et al. Mental health, pain symptoms and systemic comorbidities in women with endometriosis: a cross-sectional study. *Journal of Psychosomatic Obstetrics & Gynecology*, v. 39, n. 4, p. 315–320, 2017. Disponível em: <https://doi.org/10.1080/0167482X.2017.1386171>.

VARGAS, E. et al. The mid-secretory endometrial transcriptomic landscape in endometriosis: a meta-analysis. *Human Reproduction Open*, [S. l.], v. 2022, n. 2, p. hoac016, 2022. Disponível em: <https://doi.org/10.1093/hropen/hoac016>.

VERCELLINI, P.; SOMIGLIANA, E.; VIGANÒ, P.; ABBIATI, A.; BARBARA, G.; FEDELE, L. Chronic pelvic pain in women: etiology, pathogenesis and diagnostic approach. *Gynecological Endocrinology*, v. 25, n. 3, p. 149–158, 2009. Disponível em: <https://doi.org/10.1080/09513590802549858>.

VERCELLINI, P. et al. Endometriosis: pathogenesis and treatment. *Nature Reviews Endocrinology*, v. 10, n. 5, p. 261-275, 2014.

Vigano, M Candiani, A Monno, E Giacomini, P Vercellini, E Somigliana, Time to redefine endometriosis including its pro-fibrotic nature, *Human Reproduction*, Volume 33, Issue 3, March 2018, Pages 347–352, <https://doi.org/10.1093/humrep/dex354>

VODOLAZKAIA, A.; EL-AALAMAT, Y.; POPOVIC, D.; MIHALYI, A.; BOSSUYT, X.; KYAMA, C. M.; FASSBENDER, A.; BOKOR, A.; SCHOLS, D.; HUSKENS, D.; MEULEMAN, C.; PEERAER, K.; TOMASSETTI, C.; GEVAERT, O.; WAELEKENS, E.; KASRAN, A.; DE MOOR, B.; D'HOOGHE, T. M. Evaluation of a panel of 28 biomarkers for the non-invasive diagnosis of endometriosis. *Human Reproduction*, Oxford, v. 27, n. 9, p. 2698–2711, 2012. Disponível em: <https://doi.org/10.1093/humrep/des234>.

WANG, Z. et al. Characterization of human epigenomes. *Current Opinion in Genetics & Development*, v. 19, n. 2, p. 127-134, 2009. DOI:10.1016/j.gde.2009.02.001

WANG, Z. et al. RNA-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, v. 10, n. 1, p. 57-63, 2009. <https://doi.org/10.1038/nrg2484>

WEI, Q. et al. Reduced expression of biomarkers associated with the implantation window in women with endometriosis. *Fertility and Sterility*, v. 91, n. 5, p. 1686-1691, 2009. DOI:10.1016/j.fertnstert.2008.02.121

WICKHAM, Hadley. Programming with ggplot2. In: WICKHAM, Hadley. ggplot2. Use R!. Cham: Springer, 2016. Disponível em: [https://doi.org/10.1007/978-3-319-24277-4\\_12](https://doi.org/10.1007/978-3-319-24277-4_12). Acesso em: 19 jun. 2025.

WU, Shasha et al. The KSR1/MEK/ERK signaling pathway promotes the progression of intrauterine adhesions. *Cellular Signalling*, v. 131, 111730, 2025. Disponível em: <https://doi.org/10.1016/j.cellsig.2025.111730>. Acesso em: 16 jun. 2025.

WU Y, YANG R, LAN J, LIN H, JIAO X AND ZHANG Q. Ovarian endometrioma negatively impacts oocyte quality and quantity but not pregnancy outcomes in women undergoing IVF/ICSI treatment: a retrospective cohort study. *Frontiers in Endocrinology*, v. 12, p. 739228, 2021. Disponível em: <https://doi.org/10.3389/fendo.2021.739228>.

WU, Z. et al. A model-based background adjustment for oligonucleotide expression arrays. *Journal of the American Statistical Association*, v. 99, n. 468, p. 909-917, 2004. <https://doi.org/10.1198/016214504000000683>

YANG, Y. M.; YANG, W. X. Epithelial-to-mesenchymal transition in the development of endometriosis. *Oncotarget*, v. 8, n. 25, p. 41679–41689, 20 jun. 2017. DOI: <https://doi.org/10.18632/oncotarget.16472>. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5522315/>.

XIE, Z. et al. Gene set knowledge discovery with Enrichr. *Current Protocols*, v. 1, n. 3, p. e90, 2021. <https://doi.org/10.1002/cpz1.90>

ZHANG, Tao; DE CAROLIS, Caterina; MAN, Gene Chi Wai; WANG, Chi Chiu. The link between immunity, autoimmunity and endometriosis: a literature update. *Autoimmunity Reviews*, v. 17, n. 10, p. 945–955, 2018. Disponível em: <https://doi.org/10.1016/j.autrev.2018.03.017>.

ZHANG, Y. et al. Transcriptional landscape and clinical utility of enhancer RNAs in cancer. *Nature Communications*, v. 12, n. 1, p. 1058, 2021. <https://doi.org/10.1038/s41467-021-21350-w>

ZHANG, M. , XU, T., TONG, D., LI, S., YU, X., LIU, B., JIANG, L., & LIU, K. Research advances in endometriosis-related signaling pathways: A review. *Biomedicine & Pharmacotherapy*, v. 164, p. 114909, 2023. Disponível em: <https://doi.org/10.1016/j.biopha.2023.114909>.

ZONDERVAN, K. T. et al. Endometriosis. *New England Journal of Medicine*, v. 382, n. 13, p. 1244-1256, 2020. DOI:10.1056/NEJMra1810764