

Genetically Predicted Inflammatory Proteins Mediate the Association Between Gut Microbiota and Preterm Delivery: A Mendelian Randomization Study

Min Zhang^{1*}, Xiaodan Chen¹, Yan Zhang², Jia Huang¹, Ling Chen¹

¹Department of Preventive Healthcare, Datuan Community Health Service Center, Pudong New District, Shanghai, China

²Department of Nursing, Datuan Community Health Service Center, Pudong New District, Shanghai, China

Min Zhang: zhm395010878@gmail.com

Xiaodan Chen: 1716245607@qq.com

Yan Zhang: 1076354318@qq.com

Jia Huang: 609285765@qq.com

Ling Chen: 369778246@qq.com

***Correspondence:** Min Zhang. Department of Preventive Healthcare, Datuan Community Health Service Center. 169 Yongding South Street, Pudong New District, Shanghai 201311, China. Email: zhm395010878@gmail.com

ORCID: <https://orcid.org/0009-0002-6879-9106>.

Running head: Elusimicrobiaceae and Thioalkalivibrionaceae Associations with Preterm Delivery.

Abstract

Aims

Inflammatory proteins and unique gut microbiota profiles characterize preterm delivery (PTD). Nevertheless, the comprehensive understanding of gut microbiota and inflammatory proteins of PTD remains unclear. This study aimed to investigate the causal relationship between gut microbiota and PTD and identify the inflammatory proteins as potential mediators.

Methods and results

The exposure genome-wide association studies (GWAS) data were sourced from the GWAS Catalog, while the outcome GWAS data were obtained from the Early Growth Genetics (EGG) Consortium. The study used 473 types of gut microbiota, 91 types of inflammatory proteins, and PTD from GWAS. We then performed two-sample Mendelian randomization (TSMR) and bidirectional Mendelian randomization (BDMR) analyses to explore the causal relationships between gut microbiota, inflammatory proteins, and PTD. Additionally, we conducted two-step Mendelian randomization (2SMR) to identify potential mediating inflammatory proteins in this process. MR analysis identified 26 gut microbiota and 6 types of inflammatory proteins causally associated with PTD. Furthermore, there was no strong evidence that genetically predicted PTD affected these gut microbiota and inflammatory proteins. Further, 2SMR analysis revealed that the association between Elusimicrobiaceae and PTD was mediated by the C-C motif chemokine 23 (CCL23), accounting for 5.09% (95%CI; 4.1%-8.7%) of the association. Similarly, the relationship between Thioalkalivibrionaceae and PTD was mediated by the Interleukin-20 receptor subunit alpha (IL-20RA), which accounted for 16.88% (95%CI; 12.77%-20.99%) of the association.

Conclusions

Our results reveal that Elusimicrobiaceae and Thioalkalivibrionaceae were significantly associated with PTD, with mediation occurring via CCL23 and IL-20RA, respectively.

Impact Statement

This study establishes a causal link between specific gut microbiota, inflammatory proteins, and PTD through MR analyses. The findings indicate that targeting the pathways involving Elusimicrobiaceae - CCL23 - PTD and Thioalkalivibrionaceae - IL20RA - PTD may provide promising interventions for preventing and treating PTD.

Keywords: gut microbiota, inflammatory proteins, preterm delivery, mendelian randomization, association

1. Introduction

Preterm delivery (PTD) refers to the delivery of a baby before 37 weeks of gestation, which can have significant implications for the health of newborns, increasing their risk of various complications, including respiratory distress, infections, feeding difficulties, and developmental disorders. According to the World Health Organization, approximately 15 million infants are born preterm each year globally.¹ The rates of PTD vary across different countries and regions, with some low-income countries experiencing particularly high rates.² The economic burden of PTD on families and society is substantial, as preterm infants often require extended medical care, which can place significant financial strain on families.³

The gut microbiota influences maternal nutrient absorption and metabolism, which are crucial for fetal development.⁴ An imbalance in gut microbiota may impair the absorption of essential nutrients, adversely affecting fetal growth and increasing the risk of PTD. Mounting evidence suggests that the

delicate balance of the gut microbial community plays a pivotal role in the complex interplay between inflammation and the risk of PTD. Dysbiosis, characterized by a reduction in microbial diversity and an imbalance in microbial composition, has been associated with the onset of systemic low-grade inflammation.⁵ Inflammatory mediators produced by the dysbiotic gut microbiota can disrupt the maternal-fetal interface, potentially triggering preterm labor. Lipopolysaccharides (LPS), a potent pro-inflammatory molecule derived from the cell walls of certain gut bacteria, can enter the circulatory system and elicit a robust immune response, contributing to the development of PTD.⁶⁻⁷

In conclusion, the gut microbiota, inflammatory mediators, and PTD are intricately linked.⁸ Understanding the complex mechanisms by which gut dysbiosis, inflammatory pathways, and metabolic alterations interact to influence pregnancy outcomes can provide valuable insights into potential therapeutic strategies aimed at modulating the microbiome and mitigating the risk of PTD. Nevertheless, the comprehensive understanding of gut microbiota and inflammatory proteins of PTD remains unclear.

Mendelian randomization (MR) analysis is a robust epidemiological research strategy based on Mendelian inheritance.⁹⁻¹⁰ This technique allows for estimating causal relationships using genetic variants as instrumental variables (IVs), thereby addressing confounding variables, measurement errors, and reverse causation.¹¹ We applied three large-scale genome-wide association studies (GWAS) to study the relationship between 473 gut microbiota and 91 inflammatory proteins and PTD using MR. The aim is to elucidate the causal relationship between gut microbiota, inflammatory proteins, and PTD and to gain a clearer understanding of how gut microbiota affects PTD through its effects on inflammatory proteins.

2. Materials and Methods

2.1 Study design

In this study, we employed a range of Mendelian Randomization (MR) methodologies, including two-sample MR (TSMR), bidirectional MR (BDMR), and two-step MR (2SMR), to investigate the causal relationships between gut microbiota, inflammatory proteins, and PTD. By employing BDMR, we avoid the potential for reverse causation, ensuring that our findings accurately reflect the directionality of the relationships.¹²

Our study design is based on three key assumptions essential for the causal interpretation of MR estimates.¹³ We ensured the reliability of our findings by using genetic variations or single nucleotide polymorphisms (SNPs) as IVs that meet three crucial criteria: (i) the genetic IVs are strongly associated with exposure, (ii) the genetic IVs are not associated with confounders linked to the selected exposure and outcome, and (iii) the genetic IVs influence the outcome only through the exposure.¹⁴ Additionally, we completed the STROBE-MR checklist to ensure the integrity of this observational MR study;¹⁵⁻¹⁶ The relevant details are presented in [Supplementary Table S1](#), and a clear overview of our study design is presented in [Figure 1](#).

2.2 Data sources

Gut microbiota association studies (GWAS) data utilized in this research were obtained from a study of the genetic characteristics of gut microbiota. The original GWAS was conducted on 2,801 microbial taxa and 7,967,866 human genetic variants from 5,959 individuals enrolled in the FR02 cohort. GWAS summary data for gut microbiota was downloaded from the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>), and the GWAS Catalog accession numbers range from **GCST90032172 to GCST90032644**. More detailed information about the GWAS data can be obtained from their study.¹⁷

The GWAS data for inflammatory proteins were sourced from the GWAS Catalog, originally generated by Zhao JH et al.,¹⁸ with GWAS Catalog accession numbers ranging from **GCST90274758 to GCST90274848**. This comprehensive study identified genetic associations for 91 circulating inflammatory proteins across 11 cohorts, encompassing a total of 14,824 participants, predominantly of European ancestry.

Data on PTD has been contributed by the EGG Consortium and has been downloaded from www.egg-consortium.org.¹⁹ PTD was defined as a spontaneous delivery <259 days (37 completed gestational weeks) or by using the ICD-10 O60 code, and controls as a delivery occurring between 273 and 294 days (39 and 42 gestational weeks). Maternal GWAS meta-analysis of PTD in 233,290 women (n cases=15,419), all of whom are of recent European ancestry. Data from 23andMe is not included in the meta-analysis.

2.3 Selection of IVs and data harmonization

To adhere to the stringent criteria based on the three principal MR assumptions and to mitigate horizontal pleiotropy, only independent genome-wide significant SNPs were employed as IVs for the exposure. The IVs must be closely related to the exposure (gut microbiota, inflammatory proteins, PTD), and SNPs significantly associated with the occurrence were selected at the whole-genome level ($P < 5 \times 10^{-8}$, $r^2 < 0.001$, window size=10kb). If there are too few SNPs Included, the inclusion criteria can be changed to ($P < 5 \times 10^{-5}$ or $P < 1 \times 10^{-6}$, $r^2 < 0.001$, window size=10kb). Additionally, we calculated the F statistics of the IVs to assess the extent of weak instrument bias. To reduce the bias caused by weak working variables, the working variables with $F > 10$ are retained. The formula for calculating the F value and R^2 is as follows.

$$R^2 = 2 \times \beta_{\text{exp}}^2 \times \text{eaf}_{\text{exp}} (1 - \text{eaf}_{\text{exp}}) / [2 \times \beta_{\text{exp}}^2 \times \text{eaf}_{\text{exp}} \times (1 - \text{eaf}_{\text{exp}}) + 2 \times n_{\text{exp}} \times \text{eaf}_{\text{exp}} \times (1 - \text{eaf}_{\text{exp}}) \times \text{se}_{\text{exp}}^2],$$
$$F = R^2 \times [n_{\text{exp}} - 2 / (1 - R^2)].$$

Finally, after data filtering, horizontal pleiotropy analysis was conducted. If there is horizontal pleiotropy, the MR-PRESSO test is performed.²⁰ The outliers obtained from the test were removed before proceeding with further analysis.

2.4 Mediation analyses link “gut microbiota–inflammatory proteins–PTD”

Firstly, We performed BDMR analyses to investigate the causal relationship between gut microbiota and PTD, estimating the total effect (β) of this relationship. We used the inverse variance weighted (IVW) method to estimate effects, reporting $\beta \pm \text{SE}$ for continuous outcomes and OR (95%CI) for binary outcomes. In brief, the IVW method meta-analyzed SNP-specific Wald estimates by dividing the SNP-outcome association by the SNP-exposure association, using random effects to derive the final causal effect estimate. Additionally, we used MR-Egger and weighted median methods as

complementary approaches to IVW, providing a more comprehensive understanding of the causal relationship.

Then, we proceeded to investigate the causal relationship between inflammatory proteins and PTD using the same BDMR approach. We employed the inverse variance weighted (IVW) method, along with MR-Egger and weighted median methods, to ensure a robust understanding of this relationship.

Finally, we conducted a mediation analysis using a 2SMR study to explore whether inflammatory proteins mediate the causal effect from the gut microbiota to PTD.²¹ The overall effect was decomposed into a direct effect (without mediators) and an indirect effect (through mediators). The total effect of gut microbiota on PTD was decomposed into direct effects of gut microbiota on PTD (overall effect) and indirect effects mediated by inflammatory proteins through the mediator (mediation effect). Mediation effects were calculated through $\beta_1 \times \beta_2$: (i) the causal effect of the mediator (91 inflammatory proteins) on PTD (β_2) and (ii) the causal effect of the exposure (2 significant gut microbiota on PTD in primary MR analysis) on the mediator (β_1). We calculated the percentage mediated by the mediating effect by dividing the indirect effect by the total effect. Meanwhile, 95% CIs were calculated with the delta method as previously reported.²²⁻²³

2.5 Sensitivity analysis

The directional association between each identified SNP and both the exposure and outcome variables was assessed using MR Steiger filtering.²⁴ This method measures the degree to which the variation in exposure and outcomes can be attributed to instrumental SNPs and determines if the variability in outcomes is less than that in exposure. Horizontal pleiotropy was further investigated via the MR-Egger approach, which utilizes weighted linear regression with an unconstrained intercept.²⁵ This intercept acts as an indicator of the average pleiotropic effect across genetic variations, reflecting the typical direct influence of a variant on the outcome variable. If the intercept significantly deviated from zero (MR-Egger intercept $p < 0.05$), it indicated the presence of horizontal pleiotropy. Moreover, Cochrane's Q-test was used to assess heterogeneity, with lower p-values suggesting increased heterogeneity and a higher probability of directional pleiotropy. Leave-one-out analyses were also performed to identify potential SNP outliers.²⁶

2.6 Statistical analysis

We used the results calculated by the IVW method as the final main result, and a significance threshold of $p < 0.05$ was applied to the MR analysis, where P-values below this threshold were considered statistically significant. All statistical analyses and data visualizations were performed using R software (R Foundation, Vienna, Austria), with the TwoSampleMR (<https://github.com/MRCIEU/TwoSampleMR>) package for 2SMR analysis and the PNG (Boutell, Netherlands) package for data visualization.²⁷

3. RESULTS

3.1 Selection of IVs

After screening, there were 26 different types of gut microbiota and 6 different types of inflammatory proteins with potential causal relationships with PTD. The F-statistics for all IVs were above 10,

indicating no evidence of weak instrument bias.²⁸ After the Bonferroni adjustment, the p-values were all below the Bonferroni threshold.

3.2 Causal association between gut microbiota and PTD

When evaluating the causal association of gut microbiota with PTD, 26 types of gut microbiota show potential causal relationships. Relevant details can be found in [Supplementary Table S2](#). Among them, the abundances in stool of Thioalkalivibrionaceae (OR=0.816; 95%CI=0.690-0.966; p=0.018), Syntrophorhabdaceae (OR=0.832; 95%CI=0.714-0.970; p=0.018), Dorea (OR=0.883; 95%CI=0.789-0.988; p=0.030), Johnsonella ignava (OR=0.884; 95%CI=0.793-0.986; p=0.027), Butyricimonas sp900258545 (OR=0.924; 95%CI=0.874-0.977; p=0.005), Turicibacter (OR=0.936; 95%CI=0.887-0.987; p=0.014), Turicibacteraceae (OR=0.950; 95%CI=0.903-0.999; p=0.045), and CAG-882 sp003486385 (OR=0.952; 95%CI=0.918-0.987; p=0.008) show a negative association with PTD. This indicates that an increase in the abundance of these microbes leads to a decreased risk of PTD.

Conversely, the following types of gut microbiota show a positive association with PTD: Bifidobacterium bifidum (OR=1.049; 95%CI=1.013-1.087; p=0.008), CAG-448 (OR=1.051; 95%CI=1.005-1.099; p=0.031), Succinivibrio (OR=1.059; 95%CI=1.011-1.110; p=0.016), Coprobacillus (OR=1.065; 95%CI=1.009-1.125; p=0.022), Succinivibrionaceae (OR=1.066; 95%CI=1.004-1.132; p=0.036), Lawsonibacter sp900066645 (OR=1.082; 95%CI=1.004-1.165; p=0.039), CAG-698 (OR=1.092; 95%CI=1.010-1.181; p=0.027), Brachyspira (OR=1.094; 95%CI=1.008-1.188; p=0.032), Treponema D (OR=1.095; 95%CI=1.011-1.187; p=0.026), UBA11963 (OR=1.105; 95%CI=1.019-1.198; p=0.015), Elusimicrobiaceae (OR=1.106; 95%CI=1.012-1.209; p=0.027), Prevotella sp900318625 (OR=1.111; 95%CI=1.030-1.197; p=0.006), UBA7748 sp900314535 (OR=1.120; 95%CI=1.016-1.233; p=0.022), Fournierella massiliensis (OR=1.121; 95%CI=1.010-1.245; p=0.032), Blautia A sp002159835 (OR=1.123; 95%CI=1.035-1.219; p=0.005), CAG-465 sp000433135 (OR=1.125; 95%CI=1.042-1.215; p=0.003), An7 (OR=1.139; 95%CI=1.028-1.262; p=0.013), and Thermoplasmatota (OR=1.190; 95%CI=1.038-1.364; p=0.012).

When evaluating the causal effects of PTD on the gut microbiota, it was observed that all p-values were greater than 0.05, suggesting that PTD does not have an effect on the gut microbiota under consideration. The relevant details are presented in [Supplementary Table S3](#). The final analysis reveals potential causal relationships between 26 types of gut microbiota and PTD, as illustrated in [Figure 2](#).

3.3 Causal association between inflammatory proteins and PTD

When evaluating the causal association of inflammatory proteins with PTD, 6 types of inflammatory proteins show potential causal relationships. Relevant details are provided in [Supplementary Table S4](#). Among them, three inflammatory proteins exhibit a negative association with PTD: Interleukin-20 receptor subunit alpha (OR=0.842; 95%CI=0.738-0.960; p=0.010), C-C motif chemokine 19 (OR=0.907; 95%CI=0.847-0.971; p=0.005), and C-C motif chemokine 23 (OR=0.945; 95%CI=0.893-0.999; p=0.047). This indicates that an increase in these inflammatory proteins is associated with a decreased risk of PTD, suggesting that these inflammatory proteins may be protective factors against PTD.

The other three types of inflammatory proteins show a positive correlation with PTD. The IVW analysis results for these proteins are as follows: Interleukin-15 receptor subunit alpha (OR=1.058; 95%CI=1.000-1.119; p=0.050), C-X-C motif chemokine 10 (OR=1.100; 95%CI=1.008-1.199; p=0.032), and Interleukin-17A (OR=1.118; 95%CI=1.001-1.249; p=0.047) indicating that these inflammatory factors are risk factors for PTD.

When evaluating the causal effects of PTD on the inflammatory proteins, it was observed that all p-values were greater than 0.05, suggesting that PTD does not have an effect on the inflammatory proteins under consideration. The relevant details are provided in [Supplementary Table S5](#). The final analysis reveals potential causal relationships between 6 types of inflammatory proteins and PTD, as illustrated in [Figure 3](#).

3.4 Mediation MR analyses of inflammatory proteins

In the 2SMR analysis, 26 types of gut microbiota and 6 types of inflammatory proteins were found to be causally associated with PTD, of which the abundance of Elusimicrobiaceae in stool was significantly associated with the C-C motif chemokine 23 levels (CCL23). Similarly, the abundance of Thioalkalivibrionaceae in stool was significantly associated with the Interleukin-20 receptor subunit alpha levels (IL-20RA). The results are presented in [Supplementary Table S6](#). Genetically predicted Elusimicrobiaceae abundance in stool was negatively associated with the CCL23 (IVW; $\beta=-0.090$, SE=0.042, P=0.034). Genetically predicted Thioalkalivibrionaceae abundance in stool was positively associated with the IL-20RA levels (MR Egger; $\beta=0.411$, SE=0.188, P=0.032, IVW; $\beta=0.199$, SE=0.088, P=0.023).

When evaluating the causal effects of inflammatory proteins on gut microbiota, it was observed that all p-values were greater than 0.05, suggesting that inflammatory proteins do not affect the gut microbiota under consideration. The relevant details are provided in [Supplementary Table S7](#).

We found that the abundance of Elusimicrobiaceae in stool was associated with decreased levels of CCL23, and this decrease was associated with an increased risk of PTD. Our study showed that CCL23 accounted for 5.09% (95%CI; 4.1%-8.7%) of the increased risk associated with the abundance of Elusimicrobiaceae in stool concerning the CCL23 and PTD. These results preliminarily illustrate that Elusimicrobiaceae abundance in stool may increase the risk of PTD partially by diminishing the effects of the CCL23 on PTD.

Similarly, we found that the abundance of Thioalkalivibrionaceae in stool was associated with increased levels of IL-20RA, and this increase was associated with a decreased risk of PTD. Our study showed that the IL-20RA accounted for 16.88% (95%CI; 12.77%-20.99%) of the increased risk associated with the abundance of Thioalkalivibrionaceae in stool concerning the IL-20RA and PTD. These results preliminarily illustrate that Thioalkalivibrionaceae abundance in stool may decrease the risk of PTD partially by enhancing the effects of the IL-20RA on PTD. The results are presented in [Supplementary Table S8](#), and the schematic representation is shown in [Figure 4](#).

3.5 Sensitivity analysis

To address potential pleiotropy in our causal effect estimates, we performed multiple sensitivity analyses. Cochran's Q-test and funnel plot analysis revealed no evidence of heterogeneity or

asymmetry among the SNPs involved in the causal relationship. Furthermore, leave-one-out analysis validated the impact of each SNP on the overall causal estimates. Moreover, re-analysis of the MR study after excluding individual SNPs consistently yielded similar results, indicating that all SNPs contributed significantly to the establishment of the causal relationship, with relevant details provided in [Supplementary Table S9](#).

4. Discussion

In the present study, we investigated the causal association among 473 gut microbiota phenotypes, 91 inflammatory proteins, and PTD using large-scale genetic data and MR analysis. To our knowledge, this is the first MR analysis to explore the causal relationships among various gut microbiota phenotypes, inflammatory proteins, and PTD. Through stringent inclusion criteria and sensitivity analyses, we have identified potential causal links between 26 distinct gut microbiota types and PTD, with 2 specific inflammatory proteins potentially serving as mediators in this pathway.

Our study highlights significant associations between various types of gut microbiota and PTD, contributing to the body of evidence regarding the role of the gut microbiome in PTD. We identified 26 gut microbiota species with potential causal relationships to PTD, which opens avenues for further exploration. Several taxa exhibited negative associations with PTD. Specifically, higher abundances in the stool of Thioalkalivibrionaceae, Syntrophorhabdaceae, Dorea, Johnsonella ignava, Butyricimonas sp900258545, Turicibacter, Turicibacteraceae, and CAG-882 sp003486385 were linked to a reduced risk of PTD. This suggests that higher abundances of these microbes may confer a protective effect against PTD. The observed odds ratios (OR) for these communities (e.g., Thioalkalivibrionaceae OR=0.816) indicate a statistically significant reduced risk of PTD, highlighting the potential role of gut microbiota in modulating pregnancy outcomes. This observation aligns with a growing literature that suggests healthy gut microbiota may positively influence pregnancy by enhancing immune function and reducing inflammation.²⁹⁻³⁰ Conversely, certain taxa demonstrated positive associations with PTD, including Bifidobacterium bifidum, CAG-448, Succinivibrio, Coprobacillus, Succinivibrionaceae, Lawsonibacter sp900066645, CAG-698, Brachyspira, Treponema D, UBA11963, Elusimicrobiaceae, Prevotella sp900318625, UBA7748 sp900314535, Fournierella massiliensis, Blautia A sp002159835, CAG-465 sp000433135, An7, and Thermoplasmata. This underscores the potential of these taxa to contribute to increased PTD risk. It is possible that dysbiosis, characterized by an abundance of pro-inflammatory or pathogenic microbes, may disrupt normal physiological processes during pregnancy, thereby raising the likelihood of PTD.³¹⁻³² The mechanisms underlying these associations warrant further investigation, including exploring how these microbial communities could influence metabolic pathways, maternal immune responses, and the overall uterine environment. Moreover, our results indicate that PTD itself does not appear to have a significant impact on the above gut microbiota, as evidenced by non-significant p-values across the evaluated microbiota types. This suggests a unidirectional relationship where gut microbiota potentially influences the course of pregnancy rather than the reverse. This finding emphasizes the necessity to consider gut health in prenatal care strategies and interventions aimed at preventing PTD.

The findings of this study shed light on the complex interplay between inflammatory proteins and PTD, highlighting the potential of specific inflammatory markers as predictive biomarkers and therapeutic targets for PTD. Our analysis identified 6 distinct types of inflammatory proteins associated with PTD, with several demonstrating noteworthy causal relationships. Notably, a subset of these proteins, including IL-20RA, C-C motif chemokine 19, and CCL23, exhibited a negative association with PTD. The observed OR for these proteins (e.g., IL-20RA OR=0.842) indicate a statistically significant reduced risk of PTD, suggesting that higher levels of these inflammatory factors may confer a protective effect. These findings align with emerging evidence that suggests certain inflammatory pathways may play a beneficial role in pregnancy by modulating immune responses and maintaining a healthy intrauterine environment. For instance, IL-20RA has been linked to anti-inflammatory and tissue-repairing functions, which could contribute to the observed protective effect against PTD.³³ Similarly, the chemokines C-C motif chemokine 19 and CCL23 have been implicated in the recruitment and regulation of immune cells, potentially promoting a balanced inflammatory state during pregnancy.³⁴ In contrast, the remaining three inflammatory proteins, including Interleukin-15 receptor subunit alpha, C-X-C motif chemokine 10, and Interleukin-17A, exhibited a positive association with PTD. The elevated odds ratios (e.g., Interleukin-17A OR=1.118) for these factors suggest that they may contribute to an increased risk of PTD. These pro-inflammatory mediators have been linked to various pathological processes, such as uterine inflammation, placental dysfunction, and premature cervical ripening, which could potentially disrupt the normal course of pregnancy and lead to PTD.³⁵ Our study similarly found that PTD itself does not have a significant impact on the inflammatory proteins under investigation, as evidenced by non-significant p-values. This finding suggests a unidirectional relationship, where specific inflammatory profiles may precede and potentially influence the risk of PTD rather than the reverse. This observation underscores the importance of considering inflammatory biomarkers in the context of prenatal care and the development of targeted interventions.

The comprehensive analysis of our study has uncovered intricate relationships between gut microbiota, inflammatory proteins, and PTD, providing valuable insights into the complex interplay of these factors. The identification of specific microbial taxa and inflammatory proteins with causal associations to PTD risk is a significant step forward in understanding the underlying mechanisms driving this adverse pregnancy outcome. Notably, the abundance of Elusimicrobiaceae and Thioalkalivibrionaceae in stool exhibited significant associations with specific inflammatory proteins, namely CCL23 and IL-20RA, respectively. The negative association between genetically predicted Elusimicrobiaceae abundance and CCL23 levels suggests that higher levels of this microbiota may correlate with decreased levels of this pro-inflammatory chemokine. Our analysis indicated that this decrease was associated with an increased risk of PTD, with CCL23 accounting for 5.09% of the increased risk linked to the abundance of Elusimicrobiaceae. This finding reveals a potential mechanism by which diminished levels of CCL23 may contribute to PTD risk, suggesting that Elusimicrobiaceae abundance in stool may impair the protective effects of this inflammatory marker. Conversely, the analysis revealed that the abundance of Thioalkalivibrionaceae in stool was positively correlated with increased levels of IL-20RA, which in turn was associated with a decreased risk of PTD. This relationship was elucidated by our finding that IL-20RA accounted for 16.88% of the increased risk associated with Thioalkalivibrionaceae abundance concerning PTD. This illustrates that higher levels of Thioalkalivibrionaceae may mitigate the risk of PTD by enhancing the protective

effects of IL-20RA. These results suggest foundational patterns in the interactions between gut microbiota and inflammatory markers in the context of pregnancy and PTD. This reinforces the hypothesis that modulation of gut microbiota can influence systemic inflammatory responses and thereby impact pregnancy outcomes.³⁶ In addition, when we evaluated the causal effects of these two inflammatory proteins on gut microbiota, all p-values were greater than 0.05, indicating that the inflammatory proteins we studied do not significantly affect the composition of gut microbiota. This unidirectional relationship suggests that gut microbiota may play a more pronounced role in affecting inflammatory profiles rather than being influenced by them, which has implications for understanding the etiology of PTD.

The present study has several strengths that can guide future directions in PTD research. First, our work comprises the first systematic investigation to examine the causal relationship between gut microbiota, inflammatory proteins, and PTD. This unique contribution is underscored by the comprehensive analysis of 473 distinct types of gut microbiota and 91 inflammatory proteins. By using an MR design, the study effectively mitigated issues related to reverse causality and residual confounding variables. Furthermore, thorough sensitivity analyses were conducted to eliminate the potential influence of genetic polymorphisms, thereby enhancing the validity of the causal inferences drawn from the study. Second, we have identified a set of gut microbiota and inflammatory proteins that exhibit a strong correlation with the risk of PTD. These findings present potential biomarkers for non-invasive stool testing in PTD that warrant validation through subsequent experiments. Finally, we have proposed a potential axis of Elusimicrobiaceae–CCL23–PTD, as well as another axis Thioalkalivibrionaceae–IL-20RA–PTD, which may serve as a foundation for drug development strategies in PTD research. Further investigation utilizing *in vitro* and *in vivo* models is necessary to validate these proposed axes and to develop targeted therapies accordingly, which may serve as a foundation for drug development strategies in PTD research.

This study conducted TSMR, BDMR, and 2SMR analysis utilizing large GWAS datasets, demonstrating high statistical efficiency. The conclusions drawn in our study are grounded in genetic IVs, with causal inference conducted through multiple MR analysis methods. The findings are robust and unaffected by horizontal pleiotropy and other confounding factors. However, it is important to acknowledge several limitations in our study. Firstly, our analysis was performed using the European population, limiting its application to everyone worldwide.³⁷ Second, the dataset on PTD is general and lacks subtype information, which may not accurately reflect the characteristics of specific subtypes within PTD. Third, using summary-level statistics rather than individual-level data in our analysis limits our ability to investigate causal relationships within specific subgroups, such as females and males. Moreover, despite efforts to detect and remove outlier variants, the potential influence of horizontal pleiotropy on our findings cannot be completely ruled out.³⁸ Finally, our research indicates a modest genetic prediction of PTD, mediated by the CCL23 at 5.09% and the IL-20RA at 16.88%, suggesting the need for further investigation into additional mediators.

5. Conclusions

This study demonstrates a causal relationship between specific gut microbiota and inflammatory proteins and PTD through MR analyses. Notably, Elusimicrobiaceae and Thioalkalivibrionaceae were significantly associated with PTD, with mediation occurring via CCL23 and IL-20RA, respectively. The findings suggest that targeting these microbiota and inflammatory pathways could offer potential avenues for intervention in preventing PTD. However, further research is needed to explore the underlying mechanisms and validate these associations in clinical settings.

6. ETHICAL APPROVAL

This research utilized publicly accessible data from GWAS studies. Each individual study included in the GWAS was approved by the respective Institutional Review Board, and participants or their authorized representatives provided informed consent.

7. AUTHORS CONTRIBUTIONS

Min Zhang conducted the data analysis and interpretation and wrote the manuscript. Xiaodan Chen and Yan Zhang were responsible for proposing the research hypothesis. Jia Huang and Ling Chen were responsible for conceptualizing the study.

8. FUNDING

This work was supported by Shanghai Pudong New District Health Commission Health Science and Technology Project (Grant numbers PW2021A-76).

9. AVAILABILITY OF DATA AND MATERIALS

The analysis utilized publicly available datasets. Detailed information on all original contributions can be found in the "Data Sources" section, including specific download links and accession numbers. Readers can refer to this section for access. For further inquiries, please contact the corresponding author.

10. CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

11. References

1 Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller AB, Kinney M, Lawn J. Born too soon: the global epidemiology of 15 million preterm births. *Reprod Health*. 2013;10 Suppl 1::S2.

- 2 Liu L, Johnson HL, Cousens S, et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet*. 2012;379(9832):2151-2161.
- 3 McCormick MC. The contribution of low birth weight to infant mortality and childhood morbidity. *N Engl J Med*. 1985;312(2):82-90.
- 4 Sajdel-Sulkowska EM. The impact of maternal gut microbiota during pregnancy on fetal gut-brain axis development and life-long health outcomes. *Microorganisms*. 2023;11(9):2199.
- 5 Di Vincenzo F, Del Gaudio A, Petito V, Lopetuso LR, Scaldaferrri F. Gut microbiota, intestinal permeability, and systemic inflammation: a narrative review. *Intern Emerg Med*. 2024;19(2):275-293.
- 6 Miller FA, Sacco A, David AL, Boyle AK. Interventions for infection and inflammation-induced preterm birth: a preclinical systematic review. *Reprod Sci*. 2023;30(2):361-379.
- 7 Jia B, Tang L, Liu H, et al. Potential roles of the interactions between gut microbiota and metabolites in LPS-induced intrauterine inflammation and associated preterm birth. *J Transl Med*. 2024;22(1):7.
- 8 Li L, Yang J, Liu T, Shi Y. Role of the gut-microbiota-metabolite-brain axis in the pathogenesis of preterm brain injury. *Biomed Pharmacother*. 2023;165:115243.
- 9 Lovegrove CE, Howles SA, Furniss D, Holmes MV. Causal inference in health and disease: a review of the principles and applications of Mendelian randomization. *J Bone Miner Res*. 2024;39:1086-1098.
- 10 Power GM, Sanderson E, Pagoni P, et al. Methodological approaches, challenges, and opportunities in the application of Mendelian randomization to lifecourse epidemiology: a systematic literature review. *Eur J Epidemiol*. 2024;39(5):501-520.
- 11 Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408.
- 12 Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation? *Eur J Epidemiol*. 2021;36(3):253-257.
- 13 Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133-63.
- 14 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89-98.
- 15 Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using Mendelian randomization (STROBE-MR): explanation and elaboration. *BMJ*. 2021;375:n2233.
- 16 Skvanková VW, Richmond RC, Woolf BAR, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA*. 2021;326(16):1614-1621.
- 17 Qin Y, Havulinna AS, Liu Y, et al. Combined effects of host genetics and diet on human gut microbiota and incident disease in a single population cohort. *Nat Genet*. 2022;54(2):134-142.
- 18 Zhao JH, Stacey D, et al. Genetics of circulating inflammatory proteins identifies drivers of immune-mediated disease risk and therapeutic targets. *Nat Immunol*. 2023;24(9):1540-1551. doi: 10.1038/s41590-023-01588-w.
- 19 Solé-Navais P, Flatley C, Steinhorsdottir V, et al. Genetic effects on the timing of parturition and links to fetal birth weight. *Nat Genet*. 2023;55(4):559-567.
- 20 Ong JS, MacGregor S. Implementing MR-PRESSO and GCTA-GSMR for pleiotropy assessment in Mendelian randomization studies from a practitioner's perspective. *Genet Epidemiol*.

- 2019;43(6):609-616.
- 21 Burgess S, Daniel RM, Butterworth AS, et al. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol*. 2015;44(2):484-95.
 - 22 Carter AR, Sanderson E, Hammerton G, et al. Mendelian randomization for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol*. 2021;36(5):465-478.
 - 23 Yuan J, Xiong X, Zhang B, et al. Genetically predicted C-reactive protein mediates the association between rheumatoid arthritis and atlantoaxial subluxation. *Front Endocrinol (Lausanne)*. 2022;13:1054206.
 - 24 Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017;13(11):e1007081.
 - 25 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512-25.
 - 26 Sekula P, Del Greco MF, Pattaro C, Köttgen A. Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol*. 2016;27(11):3253-3265.
 - 27 Rasooly D, Patel CJ. Conducting a reproducible Mendelian randomization analysis using the R Analytic Statistical Environment. *Curr Protoc Hum Genet*. 2019;101(1):e82.
 - 28 Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res*. 2017;26(5):2333-2355.
 - 29 Yao Y, Cai X, Ye Y, et al. The role of microbiota in infant health: from early life to adulthood. *Front Immunol*. 2021;12:708472.
 - 30 Koren O, Konnikova L, Brodin P, Mysorekar IU, Collado MC. The maternal gut microbiome in pregnancy: implications for the developing immune system. *Nat Rev Gastroenterol Hepatol*. 2024;21(1):35-45.
 - 31 Zhao M, Chu J, Feng S, et al. Immunological mechanisms of inflammatory diseases caused by gut microbiota dysbiosis: a review. *Biomed Pharmacother*. 2023;164:114985.
 - 32 de Oliveira GLV, Cardoso CRB, Taneja V, Fasano A. Editorial: Intestinal dysbiosis in inflammatory diseases. *Front Immunol*. 2021;12:727485.
 - 33 Ouyang W, O'Garra A. IL-10 family cytokines IL-10 and IL-22: from basic science to clinical translation. *Immunity*. 2019;50(4):871-891.
 - 34 Bjerrum JT, Hansen M, Olsen J, Nielsen OH. Genome-wide gene expression analysis of mucosal colonic biopsies and isolated colonocytes suggests a continuous inflammatory state in the lamina propria of patients with quiescent ulcerative colitis. *Inflamm Bowel Dis*. 2010;16(6):999-1007.
 - 35 Wadhwa PD, Culhane JF, Rauh V, et al. Stress, infection and preterm birth: a biobehavioural perspective. *Paediatr Perinat Epidemiol*. 2001;15 Suppl 2:17-29.
 - 36 Di Simone N, Santamaria Ortiz A, Specchia M, Tersigni C, Villa P, Gasbarrini A, Scambia G, D'Ippolito S. Recent insights on the maternal microbiota: impact on pregnancy outcomes. *Front Immunol*. 2020;11:528202.
 - 37 Davey Smith G, Holmes MV, Davies NM, et al. Mendel's laws, Mendelian randomization and causal inference in observational data: substantive and nomenclatural issues. *Eur J Epidemiol*. 2020;35(2):99-111.
 - 38 Cinelli C, LaPierre N, Hill BL, Sankararaman S, Eskin E. Robust Mendelian randomization in the presence of residual population stratification, batch effects and horizontal pleiotropy. *Nat Commun*. 2022;13(1):1093.

GWAS Summary Data

Gut microbiota abundance in stool (Exposure)
5959 Finnish individuals, FINRISK 2002, Qin Y et al., 2022

Preterm delivery (Outcomes)
Cases = 18797, controls = 260246, European individuals from Moba et al., EGG consortium, Sole-Navais P et al., 2023

Inflammatory Proteins (Mediator)
14824 European individuals from INTERVA et al., Zhao JH et al, 2023



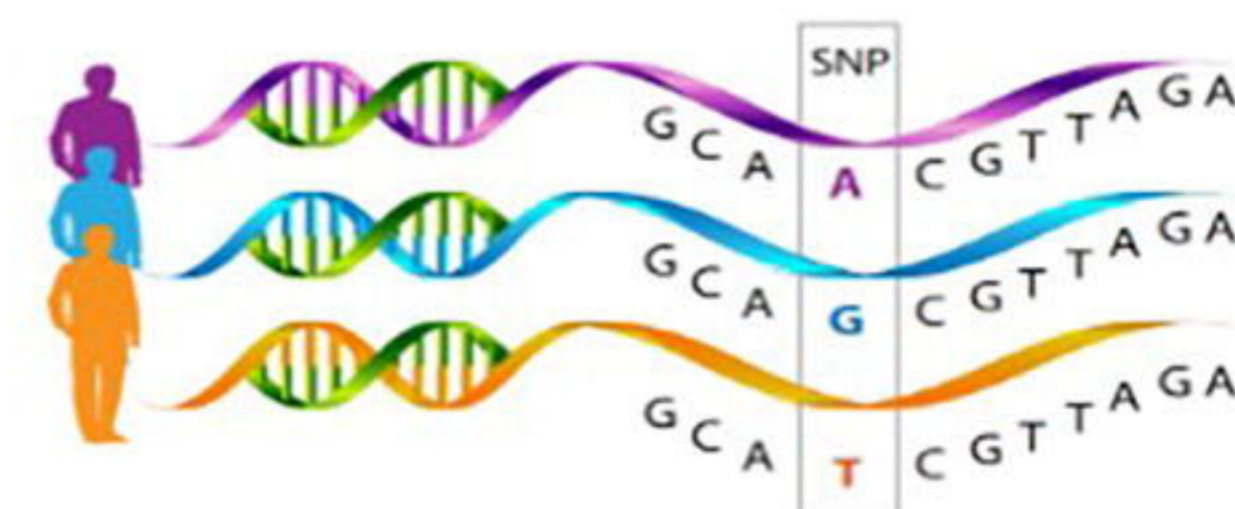
Data Quality Control

Verify the accuracy and consistency of the data



Instrumental variables

Single nucleotide polymorphism



Assumption 1

Exposure

Gut microbiota abundance in stool



Mediator

Inflammatory Proteins



Outcome

Preterm delivery



Confoundings

outside factor skews the relationship between exposure and outcome



Assumption 3

Assumption 2

Assumption 1: the genetic IVs are strongly associated with exposure

Assumption 2: the genetic IVs are not associated with confounders linked to the selected exposure and outcome

Assumption 3: the genetic IVs influence the outcome only through the exposure

Selection of genetic IVs

- Associated with gut microbiota abundance in stool ($P < 5 \times 10^{-5}$), blood metabolites ($P < 5 \times 10^{-8}$), childhood obesity ($P < 5 \times 10^{-5}$)
- LD $R^2 < 0.001$, window size=10kb
- Exclude the ambiguous SNPs with non-concordant alleles, palindromic SNPs and SNPs with a minor allele frequency < 0.01
- Remove potential pleiotropic SNPs by MR-PRESSO
- Remove SNPs associated with confounder

Mendelian randomization

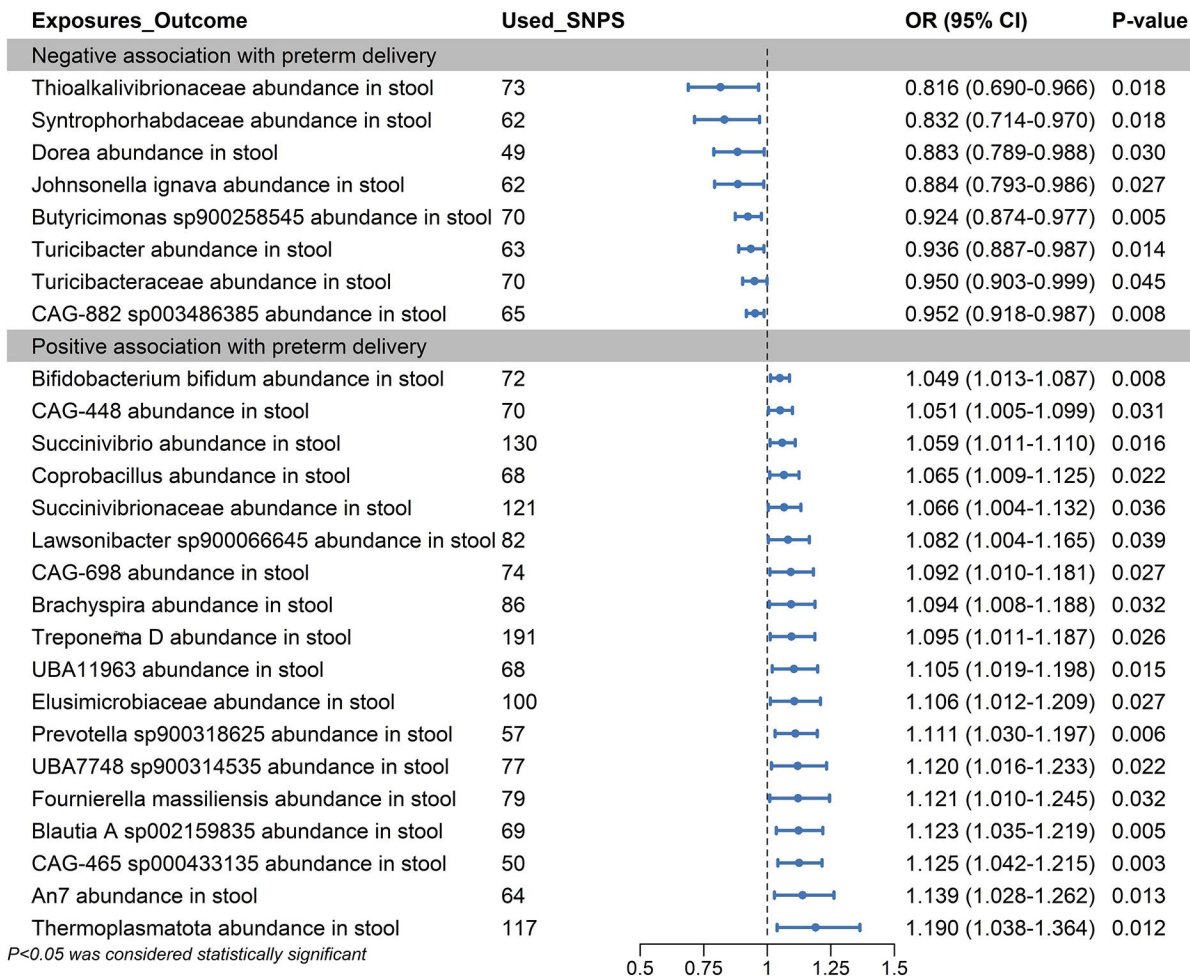
- Inverse variance weighted
- MR Egger
- Weighted median
- Simple mode
- Weighted mode

Sensitivity analysis

- Cochran's Q test (Heterogeneity)
- F statistic (Strength of IVs)
- R^2 (The proportion of variance)
- MR-Egger intercept (Pleiotropic test)
- MR-PRESSO (Outlier test)
- Leave-one-out

Two-step MR analysis
Mediation effect = $\beta_1 * \beta_2$

Figure 1. Overall study design plan. GWAS, genome-wide association studies; SNPs: single-nucleotide polymorphisms; β : the total causal effects of exposure on outcome was calculated using 2SMR method; β_1 : the causal effects of exposure on mediators were calculated using 2SMR method. β_2 : the causal effects of mediators on outcome were calculated using MVMR method. β' : the direct causal effects of exposure on outcome were calculated using the formula $\beta' = \beta - (\beta_1 * \beta_2)$.



p<0.05 was considered statistically significant

Figure 2. Mendelian randomization analyses show causal effects between gut microbiota and preterm delivery. The forest plot shows significant causal associations with $p < 0.05$ and the estimated odds ratio (OR) with 95% confidence intervals (CI).

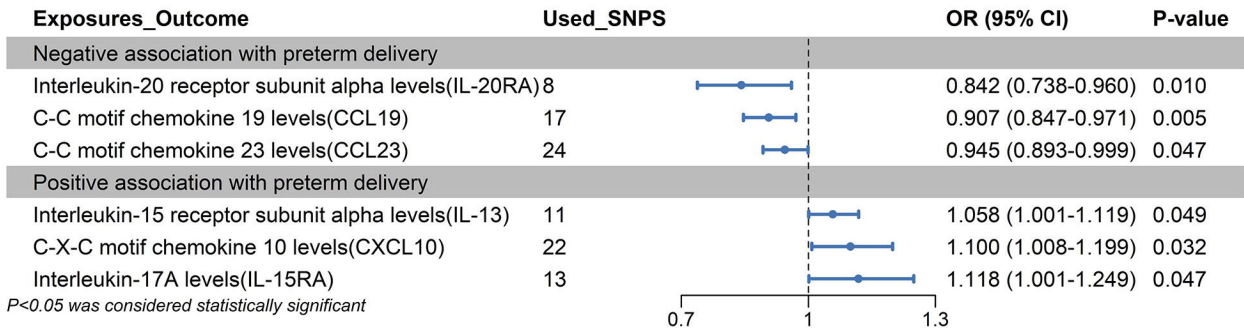


Figure 3. Mendelian randomization analyses show causal effects between inflammatory proteins and preterm delivery. The forest plot shows significant causal associations with $p < 0.05$ and the estimated odds ratio (OR) with 95% confidence intervals (CI).

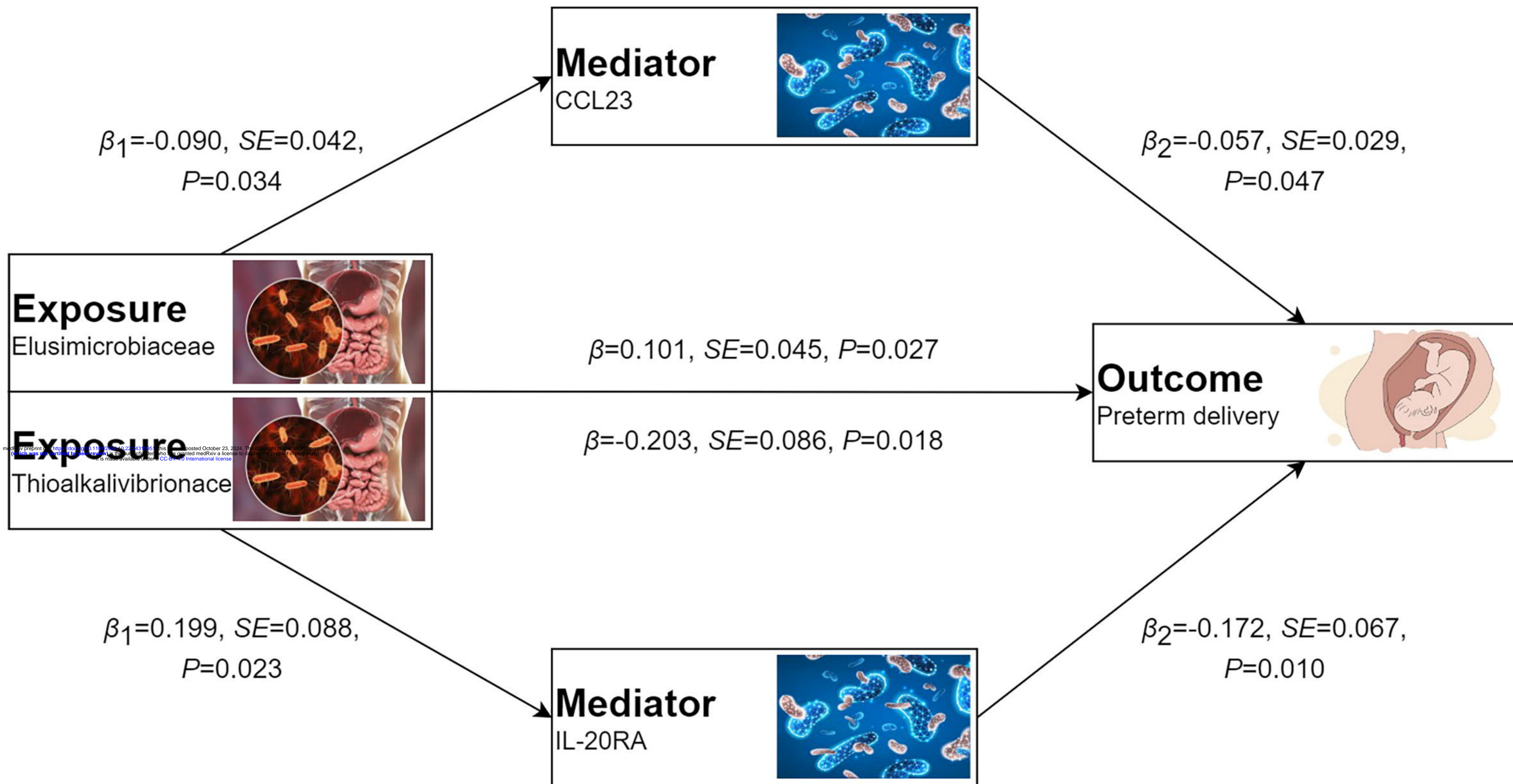


Figure 4. Schematic diagram of the mediating effect of Inflammatory Proteins between Gut microbiota and Preterm delivery.