

79899_Auto_Edited-check.docx

WORD COUNT

2104

TIME SUBMITTED

23-FEB-2023 05:03PM

PAPER ID

96927715

Name of Journal: *World Journal of Gastroenterology*

Manuscript NO: 79899

Manuscript Type: ORIGINAL ARTICLE

Case Control Study

1 Gut microbiota predicts the diagnosis of celiac disease in Saudi children

El Mouzan M *et al.* Gut microbiota in CeD

Abstract

BACKGROUND

Celiac disease (CeD) is a multisystem immune-mediated multifactorial condition strongly associated with the intestinal microbiota.

AIM

To evaluate the predictive power of the gut microbiota in the diagnosis of CeD and to search for important taxa that may help to distinguish CeD patients from controls.

METHODS

Microbial DNA from bacteria, viruses, and fungi, was isolated from mucosal and fecal samples of 40 children with CeD and 39 controls. All samples were sequenced using the HiSeq platform, the data were analyzed, and abundance and diversities were assessed. For this analysis, the predictive power of the microbiota was evaluated by calculating the area under the curve (AUC) using data for the entire microbiome. The Kruskal-Wallis test was used to evaluate the significance of the difference between AUCs. The Boruta logarithm, a wrapper built around the random forest classification algorithm, was used to identify important bacterial biomarkers for CeD.

RESULTS

In fecal samples, AUCs for bacterial, viral, and fungal microbiota were 52%, 58%, and 67.7% respectively, suggesting weak performance in predicting CeD. However, the combination of fecal bacteria and viruses showed a higher AUC of 81.8 %, indicating stronger predictive power in the diagnosis of CeD. In mucosal samples, AUCs for bacterial, viral, and fungal microbiota were 81.2%, 58.6%, and 35%, respectively, indicating that mucosal bacteria alone had the highest predictive power. Two bacteria, *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47*, in fecal samples and one virus, *Human_endogenous _retrovirus_K*, in mucosal samples are predicted to be “important” biomarkers, differentiating celiac from nonceliac disease groups.

Bacteroides intestinalis is known to degrade complex arabinoxylans and xylan which have a protective role in the intestinal mucosa. Similarly, several *Burkholderiales* species have been reported to produce peptidases that hydrolyze gluten peptides, with the potential to reduce the gluten content of food. Finally, a role for *Human_endogenous _retrovirus_K* in immune-mediated disease such as CeD has been reported.

CONCLUSION

The excellent predictive power of the combination of the fecal bacterial and viral microbiota with mucosal bacteria alone indicates a potential role in the diagnosis of difficult cases of CeD. *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47*, which were found to be deficient in CeD, have a potential protective role in the development of prophylactic modalities. Further studies on the role of the microbiota in general and *Human_endogenous _retrovirus_K* in particular are needed.

Key Words: Celiac disease; Microbial signature; Children; Saudi Arabia

El Mouzan M, Assiri A, Al Sarkhy A. Gut microbiota predicts the diagnosis of celiac disease in Saudi children . *World J Gastroenterol* 2023; In press

Core Tip: Celiac disease (CeD) is known to be associated with the microbiota. In this study, the combination of bacterial and viral taxa in stools and mucosal bacterial taxa were the strongest predictors of celiac disease. In addition, we report important bacterial markers, namely, *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47*, which were reduced in children with CeD, suggesting a protective role in children with CeD.

INTRODUCTION

Celiac disease (CeD) is an immune-mediated condition with multisystem clinical expression^[1,2]. The disease is distributed worldwide, and the incidence is increasing.

The global seroprevalence and biopsy-proven prevalence are estimated to be 1.4% and 0.7%, respectively^[3]. In the Kingdom of Saudi Arabia (KSA), a seroprevalence between 1.5% and 3% and a biopsy-proven prevalence of 1% are some of the highest observed rates in the world^[4,5]. The pathogenesis of CeD is multifactorial, requiring genetic susceptibility in the form of human leukocyte antigen DQ2 and DQ8 genotypes and exposure to gluten-containing food^[6]. In the KSA, high prevalence of genetic susceptibility of 47% has been reported^[7]. Although genetic susceptibility and exposure to gluten-containing food are necessary, not all genetically susceptible individuals develop CeD. Moreover, in some cases, CeD develops later in life after many years of gluten ingestion, indicating that other factors may be important in loss of tolerance to gluten and development of clinical disease^[8,9]. Microbial dysbiosis associated with CeD is thought to be one of the important environmental factors contributing to loss of tolerance to gluten and thereby playing a role in pathogenesis of CeD^[10,11]. However, identification of microbial markers and the predictive power of microbial dysbiosis in CeD have rarely been reported^[12]. Therefore, the objectives of this study were to determine the predictive power of the gut microbial community in the diagnosis of CeD and to search for taxa that may be important in differentiating children with CeD from those without CeD.

MATERIALS AND METHODS

Shotgun metagenomic analysis of fecal and duodenal mucosal samples from children with new onset CeD was performed for bacteria, viruses, and fungi. Briefly, there were 40 children with CeD. Patients were eligible if they had confirmed CeD by standard criteria and not received antibiotics for at least 6 mo. They were enrolled as they presented to clinics. There were 39 non-CeD controls, including 20 school children who were clinically healthy who provided stool samples and 19 from whom tissue samples from the second part of the duodenum were collected during diagnostic endoscopy performed for clinical indications. The children with CeD and controls were enrolled in the study after consent/assent. All children were recruited from King Khaled

University Hospital and King Fahad Medical City, both institutions are in Riyadh (KSA). Microbial DNA was isolated and sequenced using the HiSeq platform. The results included abundance and diversity analyses for bacteria, viruses, and fungi that were recently reported^[13-15].

For the purpose of this report, receiver operating characteristic (ROC) analysis with calculation of the area under the curve (AUC) was used to assess the predictive power of the gut microbiota in the diagnosis of CeD. ROC analysis and calculation of the AUC for discrimination using data regarding microbial communities, including bacteria, viruses, and fungi, in stool and mucosa were performed. Boruta analysis was used to identify important taxa that may differentiate children with CeD from non-CeD controls. In brief, the Boruta logarithm is a wrapper built around the random forest classification algorithm implemented in the R package random forest^[16]. The Boruta process consists of assigning an 'importance' score to each variable and identifying a threshold above which the variables are deemed important and below which they are not. This process is repeated to establish reproducibility and robustness and therefore generates many 'importance' scores for each taxon. Species-level relative abundance data were used to generate shadow variables to predict taxa that may be important in distinguishing celiac from nonceliac groups^[17]. Sensitivity and specificity were calculated based on the output from a random forest classifier^[18]. Bioinformatics and statistical analyses were performed by specialists at Cosmos ID, United States (<https://www.cosmosid.com/>).

RESULTS

The areas under the curve for the bacterial, viral, and fungal microbiota in fecal samples were 52%, 58%, and 67.7%, respectively, suggesting poor performance for each, in predicting CeD. However, the combination of fecal bacteria and viruses revealed a higher AUC of 81.8 %, indicating a stronger predictive power for the diagnosis of CeD (Figure 1). Nevertheless, the difference between the AUC of the bacterial microbiome

alone and the combined bacterial and viral microbiomes showed borderline significance ($P = 0.05211$).

The areas under the curve for the bacterial, viral, and fungal microbiota in mucosal samples were 81.2%, 58.6%, and 35%, respectively, indicating the the highest predictive power for mucosal bacteria alone (Figure 2). The difference in AUC between mucosal and fecal bacteria was significant ($P = 0.01885$).

The scores for the confirmed important variables are summarized in Table 1, including the mean, median, minimum, and maximum importance values. The 'Decision' column indicates 'Confirmed' for the microbiota above the threshold set by Boruta analysis. The microbiota including two bacteria (*Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47*) and one virus (*Human_endogenous _retrovirus_K*), was confirmed to be important for distinguishing between CeD and non-CeD groups.

The results of the Boruta random forest algorithm analysis are illustrated in Figure 3. Two bacteria (*Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47*) in fecal samples and one virus in mucosal samples (*Human_endogenous _retrovirus_K*) were predicted to be "important" for differentiating celiac from nonceliac disease groups. No fungal species were found to be important.

DISCUSSION

The role of the microbiota in predicting diseases in general has been largely reported. However, to our knowledge, the sensitivity and specificity of the gut microbiota in distinguishing CeD from non-CeD has not been reported thus far. In this study, the finding of a high AUC of the combination of bacteria and viruses in fecal samples (81.8%) indicates excellent predictive power with potential use in the diagnoses of difficult cases of CeD. Similarly, the significantly higher AUC for bacteria (81.2%) in the mucosal than in the fecal samples indicates stronger predictive power for mucosal bacteria ($P = 0.01885$). However, this was not surprising, as CeD is mainly a mucosal small bowel disease. These findings may have potential applications in the diagnosis of difficult cases of CeD.

Identification of the “important” specific microbiota in CeD in the form of *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47* has not been previously reported. *Bacteroides intestinalis* belongs to the *Bacteroides* genus, members of which are known to degrade complex arabinoxylans and xylan from dietary fibers, including wheat, rye, oat, and barley^[19]. These degradation products, including butyrate and ferulic acid, have been shown to have a protective role in the intestinal mucosa^[20,21]. *Burkholderiales bacterium 1-1-47* is an unclassified bacterium belonging to the order *Burkholderiales*, class *Betaproteobacteria* and phylum *Proteobacteria*^[23]. Several *Burkholderiales* species and *Burkholderia gladioli* in particular have been reported to produce peptidases that hydrolyze gluten peptides, with the potential to reduce the gluten content of food^[24,25]. Accordingly, reports of decreased abundance of both *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47* in samples from children with CeD^[13], indicate a potential protective role against the effects of gluten-containing grains.

Confirmation of *Human_endogenous_retrovirus_K* virus as important in differentiating CeD from non-CeD groups is interesting. This group of viruses has been suggested to have a role in immunity and autoimmune disorders. They can contribute to host protection or to damage, suggesting a subtle balance between the persistence of human endogenous retroviruses expression and maintenance of a basal immune alert^[26]. Although a recent study found increased expression in children with CeD^[27], identification of *Human_endogenous_retrovirus_K* as important and the significantly reduced abundance in children with CeD^[13], should prompt further investigation of the role of these viruses in children with CeD.

Study limitation

This study had a relatively small sample size. In addition, the non-CeD controls were not completely healthy although they do not have CeD as TTG-A, endoscopy, and duodenal tissue histopathology were normal. However, the relatively small size might be partially compensated for by the use of the shotgun metagenomic analysis. Since this

is the first report on microbiota accuracy and identification of important bacteria and viruses, but not fungi, further studies with larger sample sizes are needed.

CONCLUSION

The high AUCs of mucosal bacteria and the combination of fecal bacteria and viruses indicate a ¹ potential role in the diagnosis of difficult cases of CeD. In addition, identification of important bacteria as decreased abundances of *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47* in children with CeD, suggests a protective role with the potential for the development of preventive and adjuvant microbial therapy for CeD. The importance of *Human_endogenous _retrovirus_K* is interesting. However, further studies with larger sample sizes, are needed to improve our understanding of the role of the microbiota in CeD.

ARTICLE HIGHLIGHTS

Research background

Dysbiosis associated with celiac disease (CeD) is well known and beneficial and harmful associations have been reported.

Research motivation

The role of the microbiota in predicting CeD has rarely been described.

Research objectives

To search for a microbial signature that may help in the diagnosis and prevention of CeD.

Research methods

Metagenomic analysis of microbial DNA in mucosa and stool of children with newly diagnosed CeD. Calculation of the area under the curve to evaluate the predictive power

of the whole microbiota and use of random forest analysis to identify important microbes in distinguishing CeD groups from controls.

Research results

Very high discriminatory power of combined bacteria and viruses (81.8%) in fecal samples and bacteria only in mucosal samples (81.2%). ¹ *Bacteroides intestinalis* and *Burkholderiales bacterium 1-1-47* in fecal samples were deemed important.

Research conclusions

The excellent predictive power of microbiota may help in the diagnosis of difficult cases of CeD. The identification of important specific bacterial species that are reduced in CeD may have a potential protective role.

Research perspectives

Future research in this area with larger sample sizes is needed to clarify the role of microbiota in the diagnosis and prevention of CeD.

5%

SIMILARITY INDEX

PRIMARY SOURCES

- 1** "NASPGHAN Annual Meeting Abstracts", Journal of Pediatric Gastroenterology & Nutrition, 2022 73 words — 3%
Crossref
 - 2** Nicole Grandi, Enzo Tramontano. "Human Endogenous Retroviruses Are Ancient Acquired Elements Still Shaping Innate Immune Responses", Frontiers in Immunology, 2018 20 words — 1%
Crossref
 - 3** pubmed.ncbi.nlm.nih.gov 12 words — 1%
Internet
-

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES < 12 WORDS

EXCLUDE MATCHES < 12 WORDS