cmgh
CELLULAR AND MOLECULAR GASTROENTEROLOGY AND HEPATOLOGY

RESEARCH LETTER

A Novel Noninvasive Method Based on Salivary Inflammatory Biomarkers for the Screening of Celiac Disease

Celiac disease (CD) is a chronic, immune-mediated disorder caused by an intolerance to ingested gluten present in some cereals such as wheat, rye, and barley. CD develops in genetically susceptible individuals and it has been estimated that it affects approximately 1% of Caucasians.[1,2]

The diagnosis of CD is based on the presence of intestinal symptoms together with the evaluation of genetic (HLA-DQ2+ and/or DQ8+), serologic (anti-endomysium and anti-transglutaminase autoantibodies), and histologic markers.[3] It has been argued that approximately 70% of CD cases are not diagnosed properly as a result of negative serologic results and the complex interpretation of the histologic findings in intestinal biopsy specimens, especially in older children and adult patients.[4] Moreover, diagnostic endoscopy is an expensive procedure that involves sedation or anesthesia, and no more than 1:7 people with the highest HLA risk genotype are finally diagnosed with CD.[5] However, an early and proper diagnosis of the disease is of great importance to avoid extraintestinal complications, including cardiovascular and neurologic problems or the development of certain types of cancer. Thus, additional diagnostic tests, preferentially tests that are cost effective and noninvasive, are strictly necessary.[6]

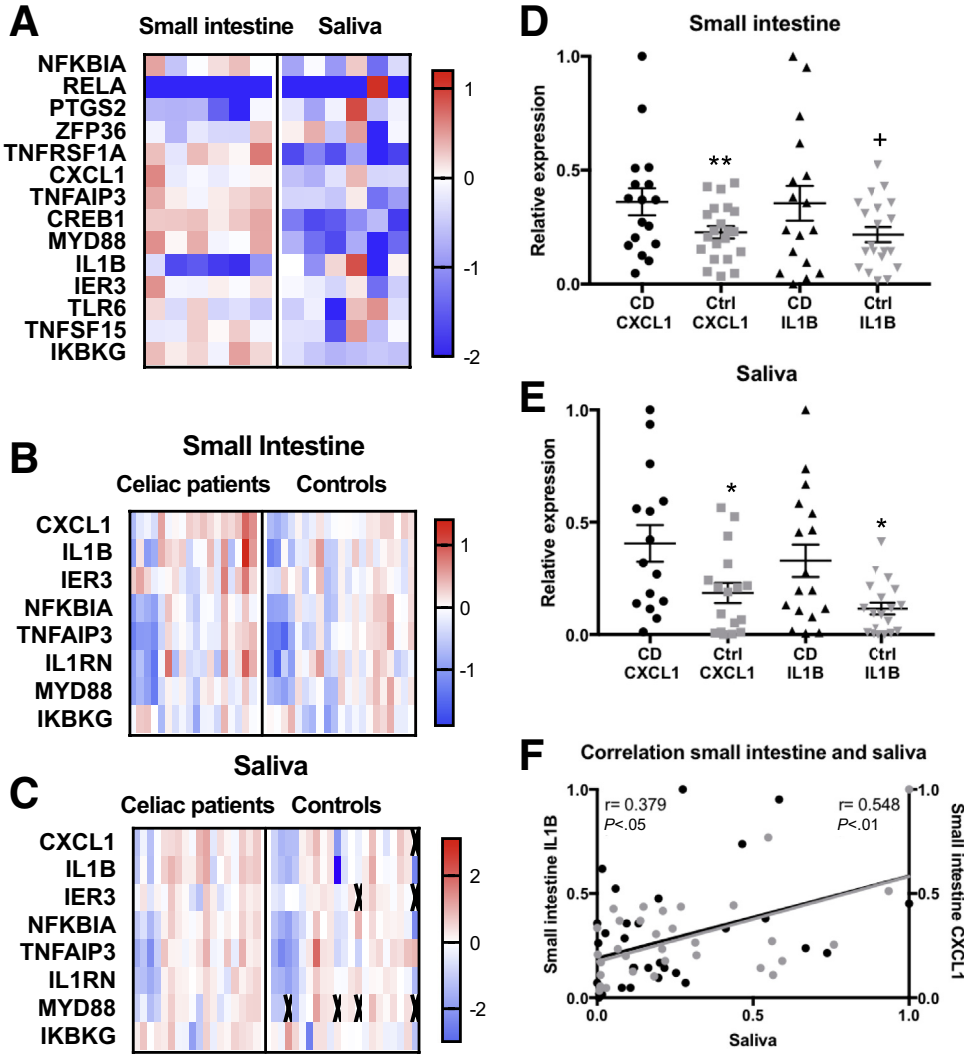


Figure 1. Heatmaps of (*A*) the inflammatory biomarkers expressed in small intestine and saliva (n = 6) and of the 8 selected genes in the (*B*) small intestine or (*C*) saliva from active celiac patients (n = 18) and controls (n = 21). Expression values were normalized against the average expression of each gene. Differential expression of the genes *CXCL1* and *IL1B* in the **(*D*)** small intestine and **(*E*)** saliva samples from active celiac patients celiac disease (CD) (N = 18) and controls (Ctrl) (n = 21). $^{**}P < .01$; $^{*}P < .05$; $^{+}P < .1$ by Student *t* test. (*F*) Pearson correlation (r) of *CXCL1* (*grey dots*) and *IL1B* (*black dots*) expression between small intestine and saliva samples (n = 39).

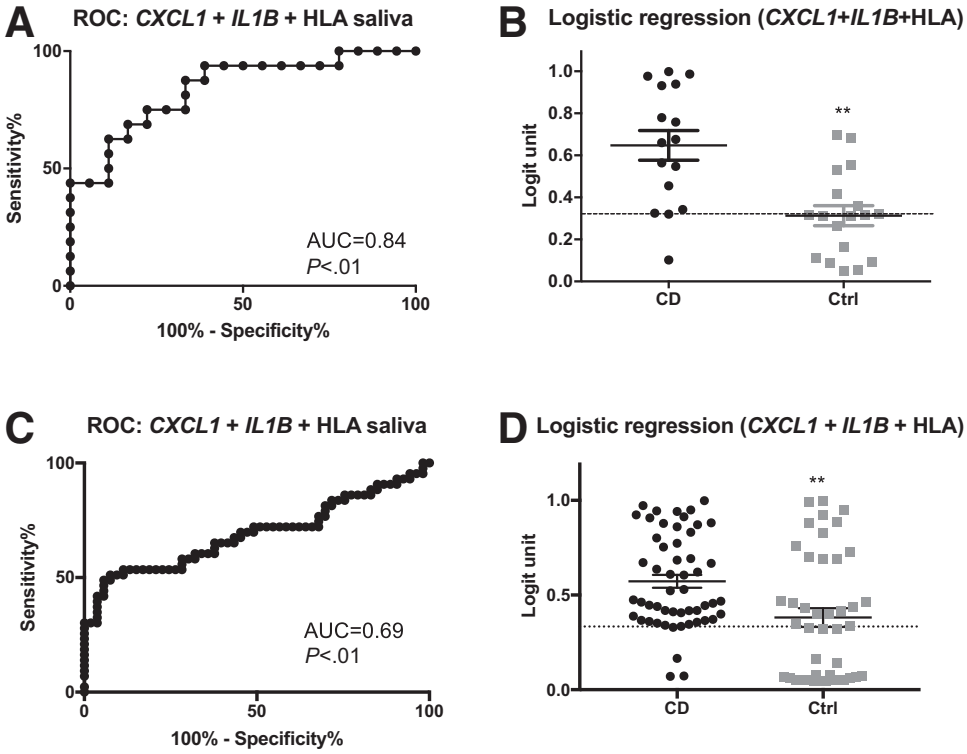


Figure 2. Receiver operating characteristic (ROC) plot of the combinatory logistic regression model including gene expression and HLA genotype in the (*A*) evaluation cohort (n = 39) and (*C*) confirmation cohort (N = 100). Comparison of the combinatory logistic regression values between celiac patients (celiac disease [CD]) and controls (Ctrl) in the **(*B*)** evaluation cohort (n = 39) and **(*D*)** confirmation cohort (n = 100). $^{**}P < .01$ by Student t test.

Oral mucosa is the first part of the gastrointestinal system and it has been suggested that it could resemble its immunopathologic characteristics, being a much more accessible tissue for CD diagnosis.[7] Given that saliva is the most accessible body fluid, it has been broadly scrutinized for biomarkers of noninvasive diagnosis of a wide range of disorders, including inflammatory bowel disease.[8] Although some attempts for CD screening in saliva have been performed (ie, antibody detection),[9] gene expression analyses of CD-related inflammatory cytokines have not been assessed in this fluid.

Considering that CD is highly underdiagnosed and that actual tests are expensive and invasive, our aim was to analyze the expression of inflammatory genes in the intestine and saliva of celiac patients and controls to find salivary biomarkers that resemble the status of the intestinal epithelium and that could be used for diagnosis. In addition, we wanted to leverage the saliva collection to set up HLA genotyping in this fluid, strengthening the predictive power of the gene expression signature.

To select potential biomarkers, we quantified the expression of 92 inflammatory genes in intestinal and saliva samples from 6 individuals (3 celiac patients and 3 controls). Fourteen of the genes tested were expressed in all the samples, in both saliva and small intestine (Figure 1*A*). The 8 inflammatory genes with the highest and most reproducible expression levels were selected for subsequent analyses in intestinal and saliva samples from another 18 celiac patients and 21 non-celiac individuals. All genes were expressed in the intestine, although in 3 of the saliva samples 1 to 3 genes could not be detected (Figure 1*B* and *C*). To evaluate if the inflammatory gene expression in saliva resembles the status of the celiac intestinal epithelia, we compared gene expression between celiac and non-celiac patients in both tissues. We found that genes *CXCL1* and *IL1B* were up-regulated in CD biopsy specimens (Figure 1*D*). Likewise, *CXCL1* and *IL1B* presented increased expression in the saliva of patients (Figure 1*E*). Correlation analyses between intestine and saliva showed a statistically significant correlation between the levels of these genes in both tissues (Figure 1*F*), suggesting that the celiac-related gene expression changes of the intestine can be assessed in saliva. Investigating whether other inflammatory conditions of the gut or mouth also present these or other alterations in saliva gene expression would be of great interest.

Subsequently, we evaluated whether these salivary biomarkers could have clinical utility to differentiate between CD patients and non-celiac controls in the same sample set. Receiver operating characteristic curve analyses yielded area under the curve (AUC) values greater than 0.7 for both genes (Supplementary Figure 1*A* and *B*). In addition, a logistic regression model was built using the combination of both biomarkers. The combinatory logistic regression model yielded a receiver operating characteristic plot AUC value of 0.69 (Supplementary Figure 1*C* and

D). Finally, the addition of the HLA genotyping (CD risk, HLA-DQ2/DQ8+; or nonrisk)[10] in the same saliva samples improved all 3 models, giving a 0.84 AUC value ($P < .01$; 95% CI, 0.7064–0.9742) (Figure 2*A* and *B*), with 81% sensitivity and 67% specificity in distinguishing celiac patients from nonceliac individuals.

To validate if this model could be useful for a noninvasive diagnosis of CD we performed a blind analysis of 100 saliva samples of individuals who were evaluated by the gastroenterologists as potential celiac patients. We assigned a disease/non-disease classification to each patient based on the values of the combined expression of the 2 biomarkers plus the presence of the HLA risk genotype. We correctly classified 73% of the individuals (Figure 2*D*) with 91% sensitivity and 51% specificity, thus confirming the validity of the technique for CD screening. If only those patients being prediagnosed as celiac using this model were subjected to endoscopy and biopsy acquisition, we would improve the positivity rate from the actual 1:7 to 2:3.

Here we present a systematic study profiling inflammatory markers in saliva samples of CD patients, enhancing the prospect of an important role for salivary diagnostics in the detection of gastrointestinal pathologies. Moreover, we have used the same starting material for both gene expression analyses and HLA genotyping, limiting the collection to a single noninvasive sample. We found that the use of this combination as a prediagnostic approach would reduce the number of patients subjected to endoscopy and biopsy acquisition and improve the positivity rate, with the subsequent savings in medical care and patient well-being.

M. SEBASTIAN-DELACRUZ[1,2]
A. OLAZAGOITIA-GARMENDIA[1,2]
A. HUERTA MADRIGAL[3]
K. GARCIA-ETXEBARRIA[4]
L. M. MENDOZA[1]
N. FERNANDEZ-JIMENEZ[1,2]
Z. GARCIA CASALES[5]
E. DE LA CALLE NAVARRO[5]
A. E. CALVO[5]
M. LEGARDA[2,6]
C. TUTAU[2,6]
I. IRASTORZA[2,6]
L. BUJANDA[7]
J. R. BILBAO[1,2,8]
A. CASTELLANOS-RUBIO[1,2,7,9]

[1]Department of Genetics, Physical Anthropology and Animal Physiology, [6]Department of Pediatrics, University of the Basque Country (UPV/EHU), Leioa, Spain
[2]Biocruces Bizkaia Health Research Institute, Barakaldo, Spain; [3]Enfermedades Digestivas, Hospital de Galdakao-Usansolo, Galdakao, Spain
[4]Biodonostia, Gastrointestinal Genetics Group, [7]Department of Gastroenterology, Biodonostia Health Research Institute, Universidad del País Vasco, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas, San Sebastian, Spain
[5]Hospital de Txagorritxu, Vitoria-Gasteiz, Spain
[8]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas, Instituto de Salud Carlos III, Madrid, Spain
[9]Ikerbasque, Basque Foundation for Science, Bilbao, Spain
Corresponding author: Ainara Castellanos-Rubio, PhD, Department of Genetics, Physical Anthropology and Animal Physiology Maria Goyri building, Lab 1.11 University of the Basque Country (UPV/EHU) Leioa, 48940, Spain, e-mail: ainara.castellanos@ehu.eus.

References

1. Green PHR, et al. N Engl J Med 2007;357:1731–1743.
2. Jabri B, et al. Nat Rev Immunol 2009;9:858–870.
3. Husby S, et al. J Pediatr Gastroenterol Nutr 2012;54:136–160.
4. Lionetti E, et al. Best Pract Res Clin Gastroenterol 2015;29:365–379.
5. Romanos J, et al. J Gastroenterol Hepatol 2010;1:11–18.
6. Kaukinen K. Nat Rev Gastroenterol Hepatol 2021;18:87–88.
7. Ellis HJ, et al. Clin Exp Immunol 2000;120:10–11.
8. Nijakowski K, et al. Int J Mol Sci 2020;21:7477.
9. Adornetto G, et al. Anal Bioanal Chem 2015;407:7189–7196.
10. Kaukinen K, et al. Am J Gastroenterol 2002;97:695–699.

Abbreviations used in this letter: **AUC, area under the curve; CD, celiac disease; PCR, polymerase chain reaction.**

Most current article

2352-345X
https://doi.org/10.1016/j.jcmgh.2021.05.016

Received April 6, 2021. Accepted May 24, 2021.

Author contributions
A. Castellanos-Rubio designed the research; M. Sebastian-delaCruz, Olazagoitia-Garmendia, L. M. Mendoza, and N. Fernandez-Jimenez performed the experiments; A. Huerta Madrigal, Z. Garcia Casales, E. de la Calle Navarro, A. E. Calvo, M. Legarda, C. Tutau, I. Irastorza, and L. Bujanda recruited patients and collected human samples; M. Sebastian-delaCruz, K. Garcia-Etxebarria, J. R. Bilbao, and A. Castellanos-Rubio analyzed the data; and M. Sebastian-delaCruz, J. R. Bilbao, and A. Castellanos-Rubio wrote the paper. All authors read and approved the final version of the manuscript.

Conflicts of interest
The authors disclose no conflicts.

Funding
This study was supported by a grant from the Spanish Ministry of Science, Universities and Innovation (SAF2017-91873-EXP), a grant from the Department of Health from the Basque Government (EJ-2017111082), and a research fellowship from the Asociación de Celiacos y Sensibles al Gluten de Madrid (A.C.R.). Also supported by a predoctoral fellowship from the University of the Basque Country (M.S.dlC.) and the Basque Government (A.O.G.).

Supplementary Materials and Methods

Patients and Study Design

This study was approved by the Basque Country Clinical Research Ethics Board (CEIC-E ref. PI2018007) and analyses were performed after informed consent was obtained from all subjects or their parents. Celiac disease in pediatric patients was diagnosed according to the European Society of Pediatric Gastroenterology Hematology and Nutrition criteria in force at the time of recruitment, including antigliadin, anti-endomysium, and anti–transglutaminase antibody determinations, as well as a confirmatory small-bowel biopsy, and analyses were performed after informed consent was obtained from all subjects or their parents. All newly diagnosed adult CD patients had increased transglutaminase antibody titers and showed characteristic small intestinal histopathologic abnormalities, including villous atrophy, crypt hyperplasia, and intraepithelial lymphocytosis. Healthy control individuals were matched for age and sex, none of them suffered from any immunologic disorder. The information on individual characteristics, such as age and sex, and clinical characteristics, is presented in Supplementary Table 1.

RNA Extraction

For intestinal biopsy samples RNA extraction was performed using the NucleoSpin RNA Kit (740984.50; Macherey Nagel, Düren, Germany). Saliva samples were collected with DNA/RNA Shield collection tube with swab (R1107-E; Zymo Research, Irvine, CA) after confirming that patients did not eat at least 2 hours before the sample collection. RNA was extracted from 500 uL saliva supernatant using the Direct-zol RNA Miniprep kit with TRI reagent (R2053; Zymo Research).

TaqMan Low-Density Arrays

A total of 125 ng RNA was used for the retrotranscription reaction using the iScript cDNA Synthesis Kit (1708890; Bio-Rad, Hercules, CA). Samples were prepared and mixed with Gene Expression Master Mix (4369016; ThermoFisher Applied Biosystems, Foster City, CA) following the manufacturer's instruction to load them in the TaqMan Low Density Arrays (Applied Biosystems). The *RPLP0* gene was used as endogenous control both in biopsy specimens and saliva samples. Reactions were run in the 7900HT Fast Real-Time Polymerase Chain Reaction System using the TaqMan Low-Density Arrays Block Module (Applied Biosystems). All quantitative polymerase chain reaction (PCR) measurements were performed in triplicate and expression levels were analyzed using the 2–delta delta Ct method. All TaqMan Gene expression Assays loaded in the TaqMan low-density arrays are listed in Supplementary Table 2.

Reverse-Transcription Quantitative PCR

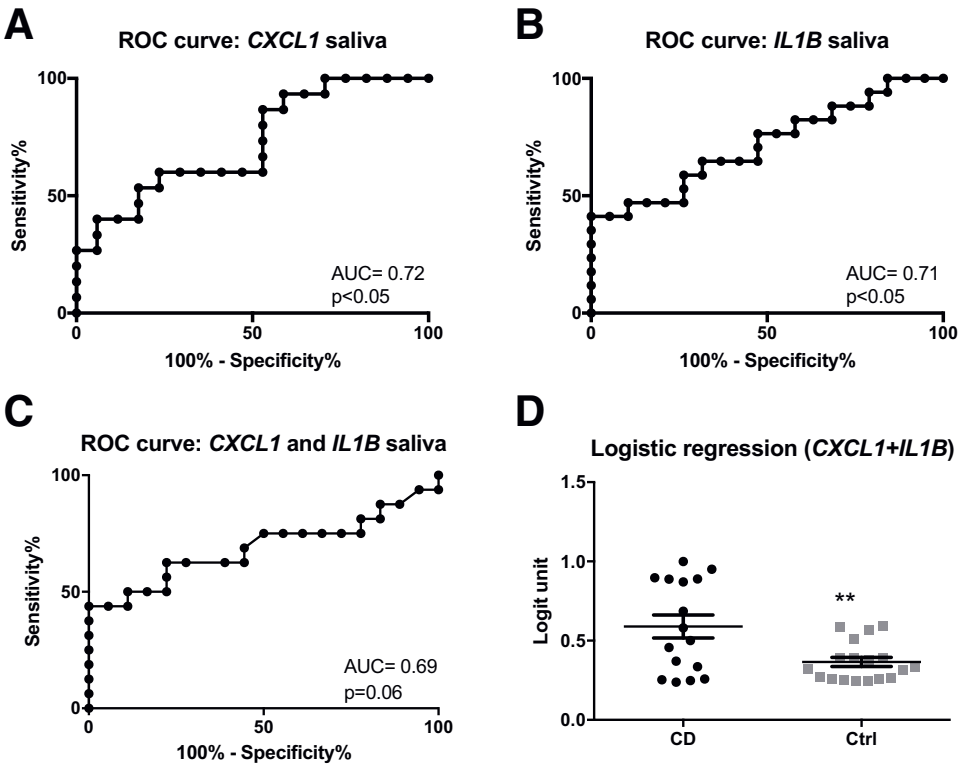
For the validation of the biomarkers, 50 ng RNA extracted from saliva was used for One Step Reverse-Transcription Quantitative PCR using the Quantitec Probe Reverse-Transcription-PCR kit (204443; Qiagen, Hilden, Germany). The *RPLP0* gene was used as endogenous control. Reactions were run in a Bio-Rad CFX384. All quantitative PCR measurements were performed in duplicate and expression levels were analyzed using the 2–ΔΔCt method. The previously used TaqMan assays were purchased and used independently.

Biomarker Evaluation and Predictive Model

For each biomarker, the receiver operating characteristic curve was constructed and the AUC value was computed by numeric integration of the receiver operating characteristic curve. The logistic regression method was used for predicting the validity of the biomarker. For this model, the predicted probability for each subject was obtained and also was used to construct a receiver operating characteristic curve. The standard error of the AUC and the 95% CI for the receiver operating characteristic curves were computed and the sensitivity and specificity for each biomarker and for the biomarker combination were estimated by identifying the cut-off point of the predicted probability that yielded the highest sum of sensitivity and specificity.

Allelic Discrimination for HLA Genotyping

DNA extraction was performed using 50 uL saliva supernatant by an isopropanol–ethanol precipitation-based method. 30uL of isopropanol (I9516-500ML; Sigma-Aldrich, St. Louis, MO) were added to the saliva sample for nucleic acid precipitation followed by cold centrifugation at 15,000×g for 15 minutes. After washing the pellet with 70% ethanol (ET00041000; Scharlab S.L., Sentmenat, Barcelona, Spain) at room temperature, the DNA fraction was precipitated by cold centrifugation at 15,000 ×g for 10 minutes. The pellet was air-dried and redissolved in DNase- and RNase-free water. HLA-DQ2 genotyping was performed using the genotyping assay for rs2187668 single-nucleotide polymorphism and TaqMan Genotyping Master Mix (4371353; Applied Biosystems, Foster City, CA). HLA-DQ8 genotyping was performed using the genotyping assay for rs7454108 single-nucleotide polymorphism and rhAmp Genotyping Master Mix (1076014; IDT, Irvine, CA), with rhAmp Reporter Mix with reference (1076020; IDT). Reactions were run in a Bio-Rad CFX384 and allelic discrimination was performed using Bio-Rad CFX Maestro Software.



Supplementary Figure 1. Receiver operating characteristic (ROC) plot of the individual gene expression values of (*A*) *CXCL1* and (*B*) *IL1B* in the evaluation cohort (n = 39). (*C*) ROC plot of the combinatory logistic regression model of the expression of both genes in the evaluation cohort (n = 39). (*D*) Comparison of the combinatory logistic regression values using gene expression between celiac disease patients (CD) and controls (Ctrl) in the evaluation cohort (n = 39). $^{**}P < .01$ based on an unpaired Student *t* test. AUC, area under the curve.