Gene Expression Profiling of Duodenal Biopsies Discriminates Celiac Disease Mucosa From Normal Mucosa

HANNA BRAGDE, ULF JANSSON, INGVAR JARLSFELT, AND JAN SÖDERMAN

Division of Medical Diagnostics [H.B., I.J., J.S.], Department of Pediatrics [U.J.], Ryhov County Hospital, Jönköping SE-551 85, Sweden

ABSTRACT: Celiac disease (CD) is identified by histopathologic changes in the small intestine which normalize during a gluten-free diet. The histopathologic assessment of duodenal biopsies is usually routine but can be difficult. This study investigated gene expression profiling as a diagnostic tool. A total of 109 genes were selected to reflect alterations in crypt-villi architecture, inflammatory response, and intestinal permeability and were examined for differential expression in normal mucosa compared with CD mucosa in pediatric patients. Biopsies were classified using discriminant analysis of gene expression. Fifty genes were differentially expressed, of which eight (*APOC3, CYP3A4, OCLN, MAD2L1, MKI67, CXCL11, IL17A,* and *CTLA4*) discriminated normal mucosa from CD mucosa without classification errors using leave-one-out cross-validation ($n = 39$) and identified the degree of mucosal damage. Validation using an independent set of biopsies ($n = 27$) resulted in four discrepant cases. Biopsies from two of these cases showed a patchy distribution of lesions, indicating that discriminant analysis based on single biopsies failed to identify CD mucosa. In the other two cases, serology support class according to discriminant analysis and histologic specimens were judged suboptimal but assessable. Gene expression profiling shows promise as a diagnostic tool and for follow-up of CD, but further evaluation is needed. (*Pediatr Res* **69:** 530–537, 2011)

Celiac disease (CD) is an inflammatory condition of the small intestine with a diverse range of symptoms (1), and a prevalence of approximately 1% in Western European populations (2). CD is triggered by gliadin peptides derived from gluten-containing cereals, and the inflamed small intestinal mucosa is characterized by villous atrophy, crypt hyperplasia, and infiltration of lymphocytes in the epithelium (1). In addition, patients with CD have increased permeability over the epithelial layer of the small intestine, and an altered structure of the tight junction has been demonstrated in children with CD (3).

A positive serological test for antibodies against tissue transglutaminase and/or gliadin may indicate CD (4). However, a histopathologic assessment (HA) of small intestinal biopsies regarding abnormalities characteristic of CD is required to establish a diagnosis (5,6). Diagnostic criteria also require that the patient should have a clinical remission when on a strict gluten-free diet (GFD). The mucosal damage is often graded histologically, using such established measures as the modified Marsh scale (7). The interobserver reproducibility between pathologists assessing the same small intestinal biopsy specimens shows fair to substantial agreement (8,9), with most of the disagreements found in Marsh 1-3B lesions (8) and with some of the interpathologist assessments transgressing the boundary between normal mucosa and CD mucosa (8).

Gene expression profiling shows a potential as a robust test for classification purposes, with a high interlaboratory reproducibility (10,11). A commercial test is already available for inflammatory bowel disease (12). Genes that display differential expression in CD-damaged mucosa, compared with normal small intestinal mucosa (13,14), may prove useful in developing a gene expression-based diagnostic test. Moreover, it might be possible to reveal histological and functional alterations associated with CD by analyzing genes that are differentially expressed along the crypt-villous axis (15,16) or that may reflect alterations in epithelial permeability associated with CD (17), or by monitoring genes potentially involved in the intestinal immune response in CD (18). On the basis of these concepts, we investigate the ability of differential gene expression to classify biopsy specimens using multivariate statistical methods.

MATERIALS AND METHODS

Study subjects. Diagnostic biopsy specimens were collected, after informed and written consent, from pediatric patients investigated for suspected CD. After CD diagnosis, control biopsies were collected on GFD and in some cases after reversion to a gluten-containing diet (diet duration >9 mo and for the majority >1 y; Table 1). Specimens extracted by using a pediatric Watson capsule were split into two pieces, one for RNA purification and one for HA. Regarding specimens extracted by endoscope, three to four were obtained for HA and one for RNA purification. Specimens for RNA purification were promptly immersed in prechilled RNA*later* RNA Stabilization Reagent (Qiagen, Hilden, Germany) and stored at 4°C for approximately 18 h and thereafter at −20°C awaiting RNA purification. Biopsies were assessed by a single experienced pathologist (I.J.), blinded to the serology and gene expression results, in accordance with instructions for quality assurance and standardization assembled by the Swedish society of pathology. Based on this, a ratio of two was used as a lower boundary for a normal villous/crypt ratio, and for establishing intraepithelial lymphocytosis corresponding to Marsh 1, a count of more than 25 intraepithelial lymphocytes (hematoxylin-eosin stained) per 100 epithelial cells was required. The histological changes were

Abbreviations: CD, celiac disease; C_T, threshold cycle; **DA,** discriminant analysis; **GFD,** gluten-free diet; **HA,** histopathologic assessment; **PC,** principal component; **PCA,** principal component analysis

Table 1. *Information on the study subjects, including age, gender, diet, gliadin antibodies (AGA), tissue transglutaminase antibodies (tTGA), histopathologic status (Marsh), and group assignment*

Case identification	Age/ gender	Gluten-free diet	AGA	tTGA	Marsh*	Group
1	2/F	No	<7	<7	0	Reference
2†	1/F	No	<7	<7	0	Reference
3†	2/M	No	<7	<7	0	Reference
4†	2/F	No	<7	<7	0	Reference
5†	1/F	No	<7	<7	0	Reference
6†	3/M	No	<7	15	0	Reference
7	3/M	No	<7	<7	0	Reference
8	12/F	No	<7	16	0	Reference
9	7/F	Yes	<7	<7	0	CD, control biopsy
10	4/F	Yes	<7	<7	0	CD, control biopsy
11	7/M	Yes	<7	<7	0	CD, control biopsy
12	12/F	Yes	<7	<7	0	CD, control biopsy
13	15/M	Yes	8	154	0	CD, control biopsy
14	16/F	Yes	<7	7	0	CD, control biopsy
15	6/F	Yes	<7	9	0	CD, control biopsy
16	15/F	Yes	<7	<7	2	CD, control biopsy
17	17/F	No	<7	63	2	CD
18†	3/F	No	<7	52	3A	CD
19†	4/F	No	<7	77	3A	CD
20†	3/F	No	<7	75	3A	CD
21†	3/F	No	456	2619	3A	CD
22†	6/M	No	<7	13	3A	CD
23	14/M	No	<7	96	3B	CD, control biopsy
24	3/M	No	44	2230	3B	CD
25†	3/F	No	62	1990	3B	CD
26	4/M	No	<7	53	3B	CD
27	4/F	No	<7	31	3B	CD
28	5/M	No	15	641	3B	CD
29	9/F	No	29	1024	3B	CD
30	14/M	No	<7	15	3B	CD
31†	1/M	No	142	1317	3C	CD
32†	1/M	No	1023	960	3C	CD
33†	1/F	No	3047	9819	3C	CD
34†	1/M	No	361	101	3C	CD
35†	3/F	No	1913	5692	3C	CD
36†	2/F	No	368	4283	3C	CD
37	2/F	No	53	847	3C	CD
38	3/F	No	117	6832	3C	CD
39	2/M	No	13	260	3C	CD
40	8/F	No	47	574	3C	CD
41	8/F	No	<7	<7	0	Reference
42	8/M	No	<7	<7	0	Reference
43	16/F	No	<7	11	0	Reference
44	9/F	No	<7	23	0	Reference
45	2/M	No	<7	<7	0	Reference
46	10/F	No	<7	<7	0	Reference
47	11/M	No	<7	<7	0	Reference
48	12/F	No	<7	33	0	Reference
49	12/M	No	9	37	0	Reference
50	8/F	Yes	<7	<7	0	CD, control biopsy
51	17/F	Yes	<7	<7	0	CD, control biopsy
52	12/F	Yes	<7	28	0	CD, control biopsy
53	8/F	Yes	<7	<7	0	CD, control biopsy
54	5/F	No	8	<7	0	CD, control biopsy
55	16/F	Yes	<7	<7	0	CD, control biopsy
56	15/F	No	<7	15	0	CD, control biopsy
57	4/F	Yes	<7	<7	1	CD, control biopsy
58	8/F	Yes	<7	<7	2	CD, control biopsy
59	14/F	No	<7	21	2	CD
60	14/F	No	<7	27	2	CD
61	14/F	No	<7	72	2	CD, control biopsy
62	2/M	No	15	524	3A	CD
63	10/F	No	15	699	3A	CD

(Continued)

Table 1. (*Continued*)

Case identification	Age/gender	Gluten-free diet	AGA	tTGA	Marsh*	Group
64	17/F	No	NA‡	NA‡	3B	CD
65	14/M	No	119	2430	3B	CD
66	3/F	No	82	982	3B	CD
67	1/M	No	417	1763	3C	CD

* Highest detected Marsh grade in the histopathologic examination.
† Biopsy specimen extracted using a pediatric Watson capsule, otherwise by endoscopy.
‡ Not available.

reported according to the modified Marsh scale (0, 1, 2, 3A, 3B, or 3C) (7). Each case was classified according to the highest Marsh grade detected among the biopsies. The quality by which specimens from each case could be assessed was determined by optimal orientation of at least four contiguous villous-crypt units. Based on quality, biopsies were divided into adequate, suboptimal but assessable, or nonassessable.

Discriminant functions were established using data from cases 1–40. Data from cases 41–67 were used for an independent validation (Table 1). The study was conducted under the approval of the Regional Ethical Review Board in Linköping.

RNA purification and reverse transcription. RNA was purified using the RNeasy Protect Mini Kit (Qiagen) according to the manufacturer's instructions and eluted in a buffer (1 mM Tris-HCl and 0.1 mM EDTA). The concentration and purity of the RNA were determined using Nanodrop ND-1000 (Thermo Fisher Scientific Inc., Waltham, MA), and integrity was assessed using RNA Integrity Number (19). RNasin Plus RNase Inhibitor (Promega Corporation, Madison, WI) was added to the RNA. One microgram or 2 μg of RNA was reverse transcribed in a total volume of 20 or 40 μL, respectively, using the High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. The resulting cDNA and the remaining RNA were stored at −80°C.

Gene expression analysis. Gene expression was analyzed on the 7500 Fast real-time PCR system using Taqman Gene Expression Assays (Table S1, http://links.lww.com/PDR/A69), Taqman Fast Universal PCR Master Mix, and the recommended Fast Protocol (Applied Biosystems) in a total volume of 15 μL. The gene expression analysis was based on reverse transcription replicates from each case. A total of 5 or 10 ng equivalents of cDNA were used per reaction. The work flow was controlled for contamination, and gene expression assays were controlled for amplification of genomic DNA.

Data analysis of gene expression data. Eleven potential reference genes (Human Endogenous Control Plate; Applied Biosystems) were evaluated for low sample-to-sample variation using the GeNorm and NormFinder algorithms in the Genex software (MultiD Analyses, Göteborg, Sweden) and cDNA obtained from reference population biopsies ($n = 2$) and cases with varying degrees of mucosal damage ($n = 4$).

Threshold cycle (C_T) values were established using version 1.4 of the 7500 Fast System sequence detection software (Applied Biosystems). The baseline was calculated using data from cycles 3 to 15, and thresholds were adjusted to the log-linear range and set to the same level for all samples within one assay. Missing C_T values, due to low copy numbers, were replaced by the highest C_T value available for the gene in question, increased by one cycle. The resulting C_T values were normalized against the selected reference genes, and the relative expression was determined for each gene in relation to the sample with the lowest expression ($2^{(C_T\text{max} - C_T\text{sample})}$) (Genex).

Statistical methods. Statistical analysis were performed using Statistica version 9.1 (Statsoft, Tulsa, OK). P values <0.05 were considered significant (no adjustment for multiple testing). Differential expression was investigated using normalized C_T values and Mann-Whitney, and the relation between relative gene expression and Marsh grade was investigated using Spearman rank correlation. Genes with significant differential expression were used for unsupervised grouping by means of hierarchical clustering (unweighted pair-group average, Euclidian distances) and principal component analysis (PCA). Before discriminant analysis (DA), the predictor variables most strongly related to the dependent variable (disease status) were identified by segmenting the range of values for each predictor (normalized gene expression) into 10 categories and performing Pearson's χ^2 test of independence. This approach was used to select the 16 best predictors for two groups (Marsh 0 *versus* Marsh 2-3C) and for three groups (Marsh 0 *versus* Marsh 2-3B *versus* Marsh 3C). Using the selected genes as input, the best subsets were selected using a general DA model and either the two-group classification scenario or the three-group classification scenario. Subsets were evaluated using Wilk's lambda, the number of predictors, and representation (villi/crypt, epithelial

permeability, and immune response). The classification accuracies of the algorithms were assessed by leave-one-out cross-validation ($n = 39$) and on an independent set of 27 biopsies.

RESULTS

HA. Biopsies from a total of 70 cases were available for assessment. Specimens from three cases were judged to be nonassessable. The remaining 67 cases were included in the study (Table 1). Specimens from 18 of these cases were judged to be suboptimal but assessable, whereas the rest were adequate for assessment. A patchy distribution of mucosal changes was found in cases 17, 22, 61, and 64.

*Screening for differentially expressed genes.* *HPRT1* and *PGK1* were identified as suitable reference genes based on low variation (Table S2, http://links.lww.com/PDR/A69).

From a total of 109 genes (Table S1, http://links.lww.com/PDR/A69), 52 were selected for further analysis based on a significant difference in expression in Marsh 3A-3C mucosa ($n = 10$) compared with Marsh 0 mucosa (reference population, $n = 4$). Reference biopsies and Marsh 0 control biopsies show evidence of differential expression for eight of these 52 genes (data not shown). Although this observation is of interest for further investigation, we used the HA as a reference and treated all Marsh 0 mucosas as one group in this study.

According to the dataset consisting of cases 1–40 (Table 1), *FGF7* and *CDK2* were not differentially expressed [Marsh 0 ($n = 15$) *versus* Marsh 2-3C ($n = 25$)] and were therefore excluded, resulting in a total of 50 genes in the subsequent analyses.

Unsupervised multivariate grouping of study subjects. Relationships among study subjects (Table 1, cases 1–40) were investigated using PCA. A scatter plot of the first two principal components (PCs) resulted in a separation of cases with histologically normal mucosa (Marsh 0), whereas cases with histopathologic changes showed a gradual transition from Marsh 2 to Marsh 3C (Fig. 1). However, case 22 (Marsh 3A, Table 1) clustered together with Marsh 0 cases. These groupings, based on gene expression data, were confirmed by hierarchical cluster analysis (Fig. 2). One main cluster harbored all Marsh 0 cases. In addition to the oppositely grouped case 22, this main cluster also contained a case with a Marsh 2 mucosa (Table 1, case 17). The other main cluster harbored only cases with Marsh 2-3C mucosas and was divided into two subclusters, with the majority of Marsh 3C cases in one of the subclusters (Fig. 2). By selecting the 12 variables making the largest contribution to PC1 and PC2, respectively, a total of 16 genes were selected (Table 2; top three contributors for each PC illustrated in Fig. 3), representing 83% of the variation in PC1 and 86% of the variation in PC2. Using these 16 genes, the general pattern of grouping
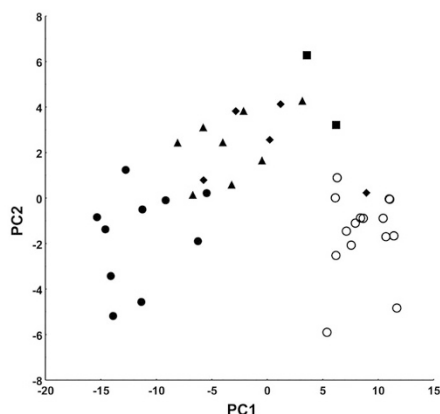
Figure 1. PCA based on 50 genes displaying a significantly different expression in Marsh 2-3C mucosa compared with Marsh 0 mucosa. ○, Marsh 0; ■, Marsh 2; ◆, Marsh 3A; ▲, Marsh 3B; ●, Marsh 3C.
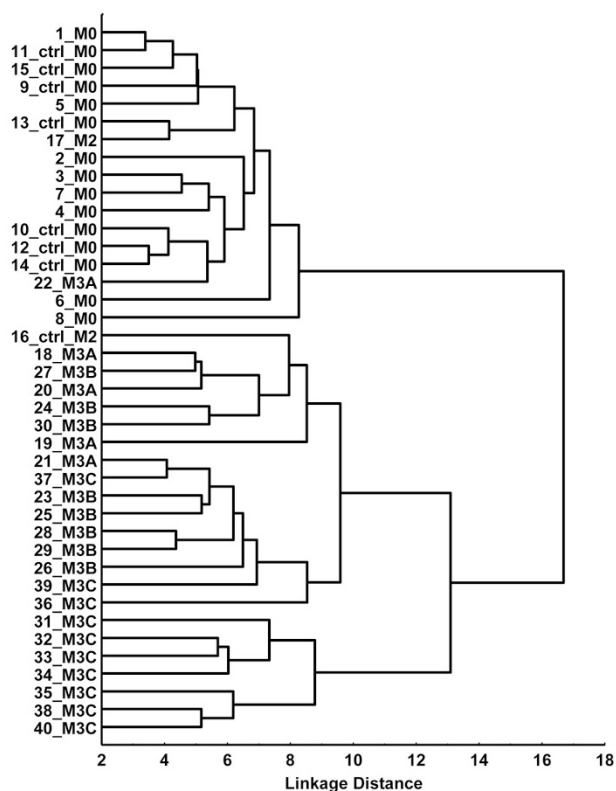
Figure 2. Hierarchical clustering based on 50 genes differentially expressed in Marsh 2-3C mucosa, compared with Marsh 0 mucosa. Sample identification is based on case number combined with mucosal status (M, Marsh; ctrl, control biopsy).

remained the same (data not shown). According to both PCA and hierarchical clustering, case 22 was positioned in the opposite group compared with its Marsh classification and was therefore excluded from the dataset on which variable selection and discriminant functions were established.

Expression characteristics of differentially expressed genes. A reduced gene expression for the villi category was observed with increasing mucosal damage (Fig. 3). The ratios between the mean relative expression in all Marsh 2-3C ($n = 24$, excluding case 22) and all Marsh 0 ($n = 15$) cases for the villi gene category ranged from 0.10 to 0.62, where *CYP3A4*

(0.10) and *APOA1* (0.11) exhibited the largest differences in mean expression. An increased gene expression for the crypt category was observed with increasing Marsh grade (Fig. 3), except for *SEMA3F*. For the crypt genes, the Marsh 2-3C/ Marsh 0 ratios ranged from 0.65 to 2.11, where *RRM2* (2.11) and *MKI67* (1.59) displayed the largest ratios. Among genes related to epithelial permeability, the gene expression decreased with increasing mucosal damage (Fig. 3). The ratios ranged from 0.59 to 0.80, with the largest difference in mean expression found in *CLDN15* (0.59). In the immune-response-related category, the mean expression ratios ranged from 0.46 to 15.10, and the largest ratios were found for *IL17A* (15.10), *CXCL11* (11.67), and *IFNG* (8.79) (Fig. 3). The highest mean expressions were associated with the Marsh 2-3C group for all genes in this category except for *IL2*, *IL18*, and *TLR3*. In the category of other potential marker genes, the ratios ranged from 0.14 to 4.87, with the largest ratios observed for *TM4SF4* (0.14) and *SELE* (4.87).

Variable selection and discriminant analysis. The dependent variable (disease status) was classified into two groups (Marsh 0 *versus* Marsh 2-3C) or three groups (Marsh 0 *versus* Marsh 2-3B *versus* Marsh 3C), and for each of these two classification scenarios, the 16 predictors that related most strongly to disease status were selected (Table 2) and combined into a set of 24 genes. Among these 24 genes, the best subsets for DA were identified. Subsets of eight genes were included in the selected classification algorithms for two and for three groups, with the only difference found in the villi genes (Table 2).

These two subsets of genes were investigated using leave-one-out cross-validation on the classification scenario from which the subset originated. In addition, because the two subsets of genes only differed with respect to villi genes with similar expression profiles (Fig. 3), the gene subset originating from the three-group classification scenario was also investigated in the two-group classification scenario. Using either of the two subsets of genes and leave-one-out cross-validation, the classification of the biopsy left out as either Marsh 0 or Marsh 2-3C was 100% in agreement with the HA for all 39 cases. For the three-group classification scenario, all but one classification agreed with the HA. The discrepant result concerned case 21 (Marsh 3A), which was classified as Marsh 3C by DA. Case 22, which had previously been excluded from the dataset, was classified as Marsh 0.

The performance of the classification algorithm derived from the three-group classification scenario was further investigated using an independent set of biopsies from 27 cases (Table 1, cases 41–67). Class according to DA agreed with class according to the HA in 23 of the 27 cases in the two-group classification scenario (Table 3). The four discrepancies concerned case 49 (Marsh 0 according to histopathology), which was classified as Marsh 2-3C by DA, and case 58 (Marsh 2 according to histopathology), case 61 (Marsh 2 according to histopathology), and case 64 (Marsh 3B according to histopathology), which were identified as Marsh 0 by DA. Two additional discrepancies were observed using the three-group classification scenario and concerned case 62 (Marsh 3A according to histopathology) and case 65 (Marsh

Table 2. *Genes selected using PCA, Pearson's χ^2 test of independence and general discriminant models*

PCA	Pearson's χ^2 M0/M2-3C	Pearson's χ^2 M0/M2-3B/M3C	Discriminant analysis M0/M2-3B/M3C	Discriminant analysis M0/M2-M3C
*APOA1**	*APOA1*	*ALPI*	*APOC3*	*APOA1*
APOA4	*CTLA4* (soluble)	*APOA1*	*CYP3A4*	*APOA4*
APOB	*CTLA4* (membrane)	*APOA4*	*MAD2L1*	*MAD2L1*
APOC3†*	*CXCL11*	*APOB*	*MKI67*	*MKI67*
CTLA4 (soluble)	*CYP3A4*	*APOC3*	*OCLN*	*OCLN*
CXCL11†	*GZMB*	*CYP3A4*	*CTLA4* (membrane)	*CTLA4* (membrane)
*CYP3A4**	*IFNG*	*ENPP3*	*CXCL11*	*CXCL11*
CYP4F2	*IL17A*	*GZMB*	*IL17A*	*IL17A*
ENPP3	*IL8*	*IL2*		
GZMB	*LCT*	*IL8*		
IFNG†	*MAD2L1*	*MKI67*		
IL17A	*MKI67*	*OCLN*		
LCT	*OCLN*	*SEMA3F*		
NOS2	*RRM2*	*TJP1*		
TM4SF4	*TM4SF4*	*TM4SF4*		
TNFRSF9	*TNFRSF9*	*TNFRSF9*		

Selections for the Pearson's χ^2 test and the general discriminant models are based on either two groups (Marsh 0 (M0) *vs* Marsh 2-3C (M2-3C)) or three groups (M0 *vs* Marsh 2-3B (M2-3B) *vs* Marsh 3C (M3C)).

* Top three contributors to principal component 1, illustrated in Figure 3.

† Top three contributors to principal component 2, illustrated in Figure 3.

3B according to histopathology), which were both classified as Marsh 3C by DA (Table 3).

DISCUSSION

In this study, we have explored putative gene expression markers suitable for discriminating biopsy specimens extracted from histologically normal small intestinal mucosa from those extracted from a mucosa displaying features characteristic of CD. Our approach identified eight genes that seem to be suitable for discrimination based on gene expression profiling (*APOC3*, *CYP3A4*, *OCLN*, *MAD2L1*, *MKI67*, *CXCL11*, *IL17A*, and membrane-bound *CTLA4*).

High discriminatory power for this gene set was indicated by leave-one-out cross-validation. Validation on an independent set of biopsies resulted in four discrepancies using a classification based on either two or three groups. Two of these four discrepancies concerned cases with a patchy distribution of lesions (cases 61 and 64) in biopsies of adequate quality for HA, indicating that DA based on single biopsies failed to identify CD mucosa. For a fraction of pediatric patients diagnosed with CD, a patchy distribution of mucosal alterations has been described (20). Serology for case 61 supports the results from histopathology (Table 1), whereas no serology results are available for case 64. The other two discrepancies involved suboptimal specimens for HA (cases 49 and 58). The serology supports the results from DA in these cases (Table 1). An additional two discrepancies (cases 62 and 65) were observed using the three-group classification scenario, one of which involved suboptimal histologic specimens (case 65). Based on the entire dataset (cases 1–67, excluding cases with a patchy mucosa and the Marsh 1 case) and leave-one-out cross-validation, class according to DA agreed with class according to the HA for case 62 (data not shown). However, for case 65, DA still disagreed with class according to the HA.

Given these observations, multiple biopsies would probably yield more accurate results with respect to true CD status of the small intestinal mucosa, not only for the HA but also for gene expression profiling.

The posterior probabilities of the single Marsh 1 specimen (Table 3) were strikingly distributed between classes, thus suggesting the possibility of distinct classification.

In the intestine, *APOA1* (21), *APOA4* (21), *APOC3* (22), and *CYP3A4* (22) represent genes that are predominantly expressed in villous enterocytes and thus potential markers for villous atrophy. Higher expression levels from these genes were found in Marsh 0 and Marsh 2 mucosa, compared with Marsh 3A-3C mucosa (Fig. 3), being consistent with the destruction of villi present in Marsh 3 mucosa. Significant decreases in *CYP3A4* protein expression have previously been observed in pediatric patients with active CD and in patients subjected to gluten rechallenge (23). An altered plasma apolipoprotein profile has been observed in children with CD, with normalization on a GFD (24).

MAD2L1 and *MKI67* are preferentially expressed in crypt cells (22). Marsh 2-3C mucosa (characterized by crypt hyperplasia) exhibited an increased expression of these genes (Fig. 3). Both the *MAD2L1* and the *MKI67* proteins have been associated with cell proliferation (25,26), and increased histologic staining of *MKI67* protein in crypts is a potential marker of CD with minimal morphologic change (27).

OCLN encodes a transmembrane tight junction protein (occludin) and contributes to the epithelial barrier (17). Marsh 2-3C mucosa exhibited a decreased expression of this gene (Fig. 3). On exposure to gliadin, a larger progressive decrease in *OCLN* expression has been observed in intestinal tissue from patients with CD, compared with controls (28).

Marsh 2-3C mucosa exhibited an increased level of *CXCL11* expression, a result in conformity with previous studies (18). The *CXCL11* encoded chemokine has been re-
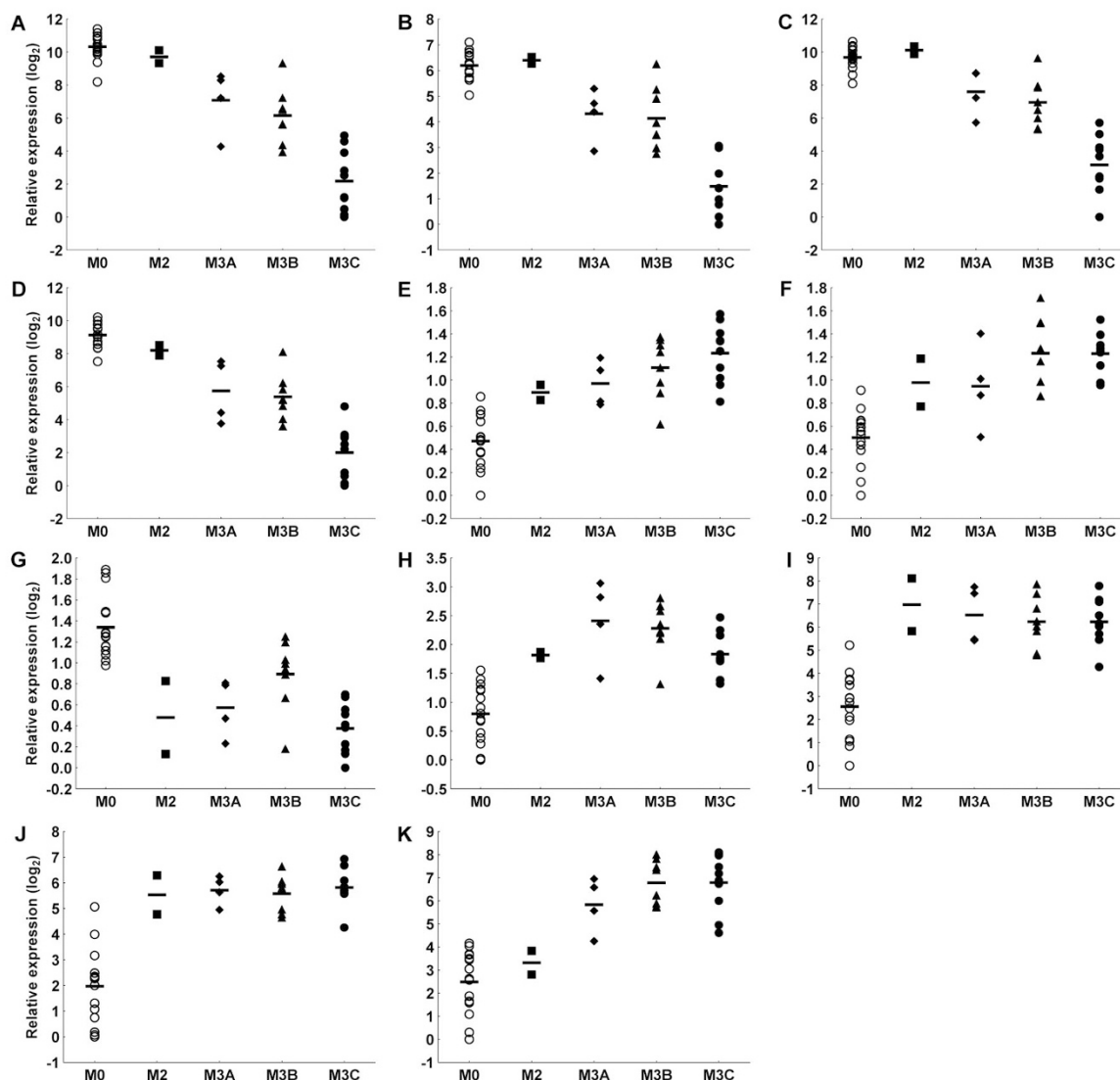
Figure 3. Relative gene expression. Mean expression is marked with a bar for each Marsh grade. Villi genes: (*A*) *APOA1*, Spearman rank correlation coefficient (r_s) = −0.91; (*B*) *APOA4*, r_s = −0.85; (*C*) *APOC3*, r_s = −0.87; (*D*) *CYP3A4*, r_s = −0.91. Crypt genes: (*E*) *MAD2L1*, r_s = 0.82; (*F*) *MKI67*, r_s = 0.78. Epithelial permeability gene: (*G*) *OCLN*, r_s = −0.75. Immune response-related genes: (*H*) *CTLA4* (membrane bound isoform), r_s = 0.63; (*I*) *CXCL11*, r_s = 0.70; (*J*) *IFNG*, r_s = 0.73; (*K*) *IL17A*, r_s = 0.83. All *p* values for the correlations were <0.001.

vealed as a potent antagonist of the CCR3 receptor (29), recently identified as a possible risk factor for CD (30). IL-17 (encoded by *IL17A*) has been associated with a number of chronic inflammatory and autoimmune diseases (31). In line with this study, enhanced *IL17A* expression has been observed in the intestinal mucosa of patients with CD (32). The protein encoded by *CTLA4* exists in a membrane-bound variant and a soluble variant, participates in the regulation of T lymphocyte-mediated inflammatory responses, and has been genetically linked to CD (33). Increased levels of soluble CTLA-4 have been detected in serum from patients with active CD, with a decrease on GFD (34). Transcripts corresponding to both membrane-bound and soluble CTLA-4 show increased expression levels in Marsh 2-3C biopsies (Fig. 3; Table S1, http://links.lww.com/PDR/A69).

The expressions of several of the genes show no distinction between Marsh 3A and Marsh 3B (Fig. 3). This is particularly

interesting because a simplified classification has been proposed, sorting Marsh 3A and 3B into one class and Marsh 1 and Marsh 2 into another class, resulting in an increase in the overall kappa values from 0.35 for the modified Marsh classification to 0.55 for the simplified classification (8). In this study, we have encountered only a few Marsh 2 mucosas and therefore have chosen to group Marsh 2 together with Marsh 3A-3B.

A simplified classification consisting of fewer grades of mucosal lesion in conjunction with a classification based on a gene expression panel might provide more accurate and reproducible results compared with current practice in diagnosing and monitoring CD. Furthermore, issues related to biopsy orientation could be avoided. Quantitative real-time PCR is a readily accessible technique that can be used to rapidly assess the expression of multiple genes. A classification based on gene expression analysis of three biopsies per patient would entail a laboratory service cost comparable with that of a HA.

Table 3. *Posterior probabilities for the classifications of an independent set of 27 biopsies by discriminant functions based on either two groups (Marsh 0 (M0) vs Marsh 2-3C (M2-3C)) or three groups (M0 vs Marsh 2-3B (M2-3B) vs Marsh 3C (M3C))*

Case identification	Marsh*	Posterior probabilities		Posterior probabilities		
		M0	M2-3C	M0	M2-3B	M3C
41	0	1.00	0.00	1.00	0.00	0.00
42	0	1.00	0.00	1.00	0.00	0.00
43	0	1.00	0.00	1.00	0.00	0.00
44	0	1.00	0.00	1.00	0.00	0.00
45	0	1.00	0.00	1.00	0.00	0.00
46	0	1.00	0.00	1.00	0.00	0.00
47	0	1.00	0.00	1.00	0.00	0.00
48	0	1.00	0.00	1.00	0.00	0.00
49†‡	0	0.01	0.99	0.00	0.00	1.00
50	0	1.00	0.00	1.00	0.00	0.00
51	0	1.00	0.00	0.98	0.02	0.00
52	0	1.00	0.00	1.00	0.00	0.00
53	0	1.00	0.00	1.00	0.00	0.00
54	0	1.00	0.00	1.00	0.00	0.00
55	0	1.00	0.00	1.00	0.00	0.00
56	0	1.00	0.00	1.00	0.00	0.00
57	1	0.48	0.52	0.41	0.59	0.00
58†‡	2	1.00	0.00	1.00	0.00	0.00
59	2	0.35	0.65	0.13	0.87	0.00
60	2	0.00	1.00	0.00	1.00	0.00
61†‡	2	1.00	0.00	1.00	0.00	0.00
62‡	3A	0.00	1.00	0.00	0.17	0.83
63	3A	0.00	1.00	0.00	0.62	0.38
64†‡	3B	0.97	0.03	0.99	0.01	0.00
65‡	3B	0.05	0.95	0.00	0.00	1.00
66	3B	0.00	1.00	0.00	0.96	0.04
67	3C	0.00	1.00	0.00	0.00	1.00

* Results from histopathologic assessments.

† Discriminant analysis disagree with the histopathologic assessment, two-group classification scenario.

‡ Discriminant analysis disagree with the histopathologic assessment, three-group classification scenario.

Although the results of this study emphasize the need for multiple specimens from the small intestine for the evaluation of intestinal damage, we are confident that it is possible to establish DA based on gene expression as a diagnostic tool for CD as well as for monitoring the effects of GFD. However, DA based on the selected genes needs to be further validated.

REFERENCES

1. Green PH, Jabri B 2003 Coeliac disease. Lancet 362:383–391
2. Dubé C, Rostom A, Sy R, Cranney A, Saloojee N, Garritty C, Sampson M, Zhang L, Yazdi F, Mamaladze V, Pan I, Macneil J, Mack D, Patel D, Moher D 2005 The prevalence of celiac disease in average-risk and at-risk Western European populations: a systematic review. Gastroenterology 128:S57–S67
3. Schulzke JD, Bentzel CJ, Schulzke I, Riecken EO, Fromm M 1998 Epithelial tight junction structure in the jejunum of children with acute and treated celiac sprue. Pediatr Res 43:435–441
4. Green PH, Rostami K, Marsh MN 2005 Diagnosis of coeliac disease. Best Pract Res Clin Gastroenterol 19:389–400
5. Hill ID, Dirks MH, Liptak GS, Colletti RB, Fasano A, Guandalini S, Hoffenberg EJ, Horvath K, Murray JA, Pivor M, Seidman EG 2005 Guideline for the diagnosis and treatment of celiac disease in children: recommendations of the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. J Pediatr Gastroenterol Nutr 40:1–19
6. Walker-Smith JA on behalf of the Working Group of European Society of Paediatric Gastroenterology and Nutrition 1990 Revised criteria for diagnosis of coeliac disease. Report of Working Group of European Society of Paediatric Gastroenterology and Nutrition. Arch Dis Child 65:909–911
7. Rostami K, Kerckhaert J, Tiemessen R, von Blomberg BM, Meijer JW, Mulder CJ 1999 Sensitivity of antiendomysium and antigliadin antibodies in untreated celiac disease: disappointing in clinical practice. Am J Gastroenterol 94:888–894
8. Corazza GR, Villanacci V, Zambelli C, Milione M, Luinetti O, Vindigni C, Chioda C, Albarello L, Bartolini D, Donato F 2007 Comparison of the interobserver reproducibility with different histologic criteria used in celiac disease. Clin Gastroenterol Hepatol 5:838–843
9. Weile B, Hansen BF, Hagerstrand I, Hansen JP, Krasilnikoff PA 2000 Interobserver variation in diagnosing coeliac disease. A joint study by Danish and Swedish pathologists. APMIS 108:380–384
10. Kohlmann A, Haschke-Becher E, Wimmer B, Huber-Wechselberger A, Meyer-Monard S, Huxol H, Siegler U, Rossier M, Matthes T, Rebsamen M, Chiappe A, Diemand A, Rauhut S, Johnson A, Liu WM, Williams PM, Wieczorek L, Haferlach T 2008 Intraplatform reproducibility and technical precision of gene expression profiling in 4 laboratories investigating 160 leukemia samples: the DACH study. Clin Chem 54:1705–1715
11. Kohlmann A, Kipps TJ, Rassenti LZ, Downing JR, Shurtleff SA, Mills KI, Gilkes AF, Hofmann WK, Basso G, Dell'orto MC, Foa R, Chiaretti S, De Vos J, Rauhut S, Papenhausen PR, Hernandez JM, Lumbreras E, Yeoh AE, Koay ES, Li R, Liu WM, Williams PM, Wieczorek L, Haferlach T 2008 An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase. Br J Haematol 142:802–807
12. von Stein P, Lofberg R, Kuznetsov NV, Gielen AW, Persson JO, Sundberg R, Hellstrom K, Eriksson A, Befrits R, Ost A, von Stein OD 2008 Multigene analysis can discriminate between ulcerative colitis, Crohn's disease, and irritable bowel syndrome. Gastroenterology 134:1869–1881; quiz 2153–2154
13. Juuti-Uusitalo K, Maki M, Kaukinen K, Collin P, Visakorpi T, Vihinen M, Kainulainen H 2004 cDNA microarray analysis of gene expression in coeliac disease jejunal biopsy samples. J Autoimmun 22:249–265
14. Diosdado B, Wapenaar MC, Franke L, Duran KJ, Goerres MJ, Hadithi M, Crusius JB, Meijer JW, Duggan DJ, Mulder CJ, Holstege FC, Wijmenga C 2004 A microarray screen for novel candidate genes in coeliac disease pathogenesis. Gut 53:944–951
15. Olsen L, Hansen M, Ekstrom CT, Troelsen JT, Olsen J 2004 CVD: the intestinal crypt/villus in situ hybridization database. Bioinformatics 20:1327–1328
16. Mariadason JM, Nicholas C, L'Italien KE, Zhuang M, Smartt HJ, Heerdt BG, Yang W, Corner GA, Wilson AJ, Klampfer L, Arango D, Augenlicht LH 2005 Gene

expression profiling of intestinal epithelial cell maturation along the crypt-villus axis. Gastroenterology 128:1081–1088

17. Turner JR 2009 Intestinal mucosal barrier function in health and disease. Nat Rev Immunol 9:799–809

18. Diosdado B, van Bakel H, Strengman E, Franke L, van Oort E, Mulder CJ, Wijmenga C, Wapenaar MC 2007 Neutrophil recruitment and barrier impairment in celiac disease: a genomic study. Clin Gastroenterol Hepatol 5:574–581

19. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T 2006 The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 7:3

20. Bonamico M, Mariani P, Thanasi E, Ferri M, Nenna R, Tiberti C, Mora B, Mazzilli MC, Magliocca FM 2004 Patchy villous atrophy of the duodenum in childhood celiac disease. J Pediatr Gastroenterol Nutr 38:204–207

21. Green PH, Lefkowitch JH, Glickman RM, Riley JW, Quinet E, Blum CB 1982 Apolipoprotein localization and quantitation in the human intestine. Gastroenterology 83:1223–1230

22. Gassler N, Newrzella D, Bohm C, Lyer S, Li L, Sorgenfrei O, van Laer L, Sido B, Mollenhauer J, Poustka A, Schirmacher P, Gretz N 2006 Molecular characterisation of non-absorptive and absorptive enterocytes in human small intestine. Gut 55:1084–1089

23. Johnson TN, Tanner MS, Taylor CJ, Tucker GT 2001 Enterocytic CYP3A4 in a paediatric population: developmental changes and the effect of coeliac disease and cystic fibrosis. Br J Clin Pharmacol 51:451–460

24. Rosenthal E, Hoffman R, Aviram M, Benderly A, Erde P, Brook JG 1990 Serum lipoprotein profile in children with celiac disease. J Pediatr Gastroenterol Nutr 11:58–62

25. Skinner JJ, Wood S, Shorter J, Englander SW, Black BE 2008 The Mad2 partial unfolding model: regulating mitosis through Mad2 conformational switching. J Cell Biol 183:761–768

26. Gerdes J, Lemke H, Baisch H, Wacker HH, Schwab U, Stein H 1984 Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. J Immunol 133:1710–1715

27. Settakorn J, Leong AS 2004 Immunohistologic parameters in minimal morphologic change duodenal biopsies from patients with clinically suspected gluten-sensitive enteropathy. Appl Immunohistochem Mol Morphol 12:198–204

28. Drago S, El Asmar R, Di Pierro M, Grazia Clemente M, Tripathi A, Sapone A, Thakar M, Iacono G, Carroccio A, D'Agate C, Not T, Zampini L, Catassi C, Fasano A 2006 Gliadin, zonulin and gut permeability: Effects on celiac and non-celiac intestinal mucosa and intestinal cell lines. Scand J Gastroenterol 41:408–419

29. Loetscher P, Pellegrino A, Gong JH, Mattioli I, Loetscher M, Bardi G, Baggiolini M, Clark-Lewis I 2001 The ligands of CXC chemokine receptor 3, I-TAC, Mig, and IP10, are natural antagonists for CCR3. J Biol Chem 276:2986–2991

30. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M, Romanos J, Dinesen LC, Ryan AW, Panesar D, Gwilliam R, Takeuchi F, McLaren WM, Holmes GK, Howdle PD, Walters JR, Sanders DS, Playford RJ, Trynka G, Mulder CJ, Mearin ML, Verbeek WH, Trimble V, Stevens FM, O'Morain C, Kennedy NP, Kelleher D, Pennington DJ, Strachan DP, McArdle WL, Mein CA, Wapenaar MC, Deloukas P, McGinnis R, McManus R, Wijmenga C, van Heel DA 2008 Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40:395–402

31. Miossec P, Korn T, Kuchroo VK 2009 Interleukin-17 and type 17 helper T cells. N Engl J Med 361:888–898

32. Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, Ramon Bilbao J 2009 TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. Autoimmunity 42:69–73

33. Hunt KA, McGovern DP, Kumar PJ, Ghosh S, Travis SP, Walters JR, Jewell DP, Playford RJ, van Heel DA 2005 A common CTLA4 haplotype associated with coeliac disease. Eur J Hum Genet 13:440–444

34. Simone R, Brizzolara R, Chiappori A, Milintenda-Floriani F, Natale C, Greco L, Schiavo M, Bagnasco M, Pesce G, Saverino D 2009 A functional soluble form of CTLA-4 is present in the serum of celiac patients and correlates with mucosal injury. Int Immunol 21:1037–1045