World Journal of

Gastroenterology

World Journal of Gastroenterology

Contents

Contents

ABOUT COVER

Editorial Board Member of *World Journal of Gastroenterology*, Osamu Toyoshima, MD, PhD, Director, Professor, Department of Gastroenterology, Toyoshima Endoscopy Clinic, 6-17-5 Seijo, Setagaya-ku, Tokyo 157-0066, Japan. t@ichou.com

AIMS AND SCOPE

The primary aim of *World Journal of Gastroenterology* (*WJG*, *World J Gastroenterol*) is to provide scholars and readers from various fields of gastroenterology and hepatology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online. *WJG* mainly publishes articles reporting research results and findings obtained in the field of gastroenterology and hepatology and covering a wide range of topics including gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, gastrointestinal oncology, and pediatric gastroenterology.

INDEXING/ABSTRACTING

The *WJG* is now indexed in Current Contents®/Clinical Medicine, Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports®, Index Medicus, MEDLINE, PubMed, PubMed Central, and Scopus. The 2021 edition of Journal Citation Report® cites the 2020 impact factor (IF) for *WJG* as 5.742; Journal Citation Indicator: 0.79; IF without journal self cites: 5.590; 5-year IF: 5.044; Ranking: 28 among 92 journals in gastroenterology and hepatology; and Quartile category: Q2. The *WJG*'s CiteScore for 2020 is 6.9 and Scopus CiteScore rank 2020: Gastroenterology is 19/136.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Jia-Hui Li*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Ze-Mao Gong.*

Celiac disease: From genetics to epigenetics

Elisa Gnodi, Raffaella Meneveri, Donatella Barisani

ORCID number: Elisa Gnodi 0000-0002-4390-6633; Raffaella Meneveri 0000-0003-0791-5951; Donatella Barisani 0000-0003-4592-6221.

Elisa Gnodi, Raffaella Meneveri, Donatella Barisani, School of Medicine and Surgery, University of Milano-Bicocca, Monza 20900, Italy

Corresponding author: Donatella Barisani, MD, MSc, Associate Professor, School of Medicine and Surgery, University of Milano-Bicocca, Via Cadore 48, Monza 20900, Italy. donatella.barisani@unimib.it

Abstract

Celiac disease (CeD) is a multifactorial autoimmune disorder spread worldwide. The exposure to gluten, a protein found in cereals like wheat, barley and rye, is the main environmental factor involved in its pathogenesis. Even if the genetic predisposition represented by HLA-DQ2 or HLA-DQ8 haplotypes is widely recognised as mandatory for CeD development, it is not enough to explain the total predisposition for the disease. Furthermore, the onset of CeD comprehend a wide spectrum of symptoms, that often leads to a delay in CeD diagnosis. To overcome this deficiency and help detecting people with increased risk for CeD, also clarifying CeD traits linked to disease familiarity, different studies have tried to make light on other predisposing elements. These were in many cases genetic variants shared with other autoimmune diseases. Since inherited traits can be regulated by epigenetic modifications, also induced by environmental factors, the most recent studies focused on the potential involvement of epigenetics in CeD. Epigenetic factors can in fact modulate gene expression with many mechanisms, generating more or less stable changes in gene expression without affecting the DNA sequence. Here we analyze the different epigenetic modifications in CeD, in particular DNA methylation, histone modifications, non-coding RNAs and RNA methylation. Special attention is dedicated to the additional predispositions to CeD, the involvement of epigenetics in developing CeD complications, the pathogenic pathways modulated by epigenetic factors such as microRNAs and the potential use of epigenetic profiling as biomarker to discriminate different classes of patients.

Key Words: Celiac disease; Epigenetics; DNA methylation; Histone modifications; Long non-coding RNAs; MicroRNAs

Core Tip: Currently identified genes account only for half of celiac disease (CeD)

predisposition. An important role could be played by epigenetics, inheritable traits without DNA sequence alterations, which could be influenced by gluten exposure. DNA methylation, histone modifications and non-coding RNAs act on different gene expression steps, from gene transcription to post-translational ones. Epigenetic changes can be additional predisposition factors or specific of CeD stages (active disease, gluten-free diet) as recently reported. Analysis of epigenetic data and their integration with transcriptome (by machine learning) can help to stratify patients, or discover new players in CeD pathogenesis, possible focus of novel therapeutic approaches.

INTRODUCTION

Celiac disease (CeD) is a multifactorial autoimmune enteropathy that develops in genetically predisposed subjects carrying the HLA-DQ2 or HLA-DQ8 haplotype. The prevalence of these HLA haplotypes in the general population is around 30%-40%, suggesting that they are necessary, but not sufficient, to induce CeD[1]. Regional variances have been observed in the frequency of these haplotypes in CeD, with the HLA-DQ8 that ranges from about 2%-4% in western countries to 25%-30% in the Iranian population and in the middle East[2]. Indeed, CeD prevalence worldwide is about 1%-2%, with environmental factors that contribute to the regional differences, like the exposure to gluten, the known CeD exogenous antigen[1]. The clinical presentation of CeD is quite heterogeneous, ranging from classical intestinal-related symptoms (diarrhoea, failure to thrive) to non-intestinal manifestations (anaemia, dermatitis, osteoporosis), and such diverse clinical picture may delay, in some cases, the prompt diagnosis[3]. Even if the main mechanisms by which gluten peptides cause CeD intestinal lesion are now quite established, the reason why only a few of the genetically predisposed subjects develop the disease still needs to be clarified. In fact, geneticists calculated that the presence of a specific HLA accounted only for about 40% of the genetic predisposition, leaving most of the genes involved in the development of the disorder still unknown[4]. Thus the initial focus of researchers was to identify further genes that could constitute a "genetic background" predisposing to the disease. Few studies used the classical linkage analysis approach, but the obtained results were specific of a single population and could not be replicated[5], or the identified genomic region did not harbour genes that appeared to be involved in CeD pathogenesis[6]. The identification of additional predisposing genes was obtained using a different approach, *i.e.*, Genome Wide Association Studies (GWAS) performed on populations of different geographical origin. These studies identified 39 Loci associated with CeD development, but also confirmed the role of the Human Leukocyte Antigen (HLA) region. Additional analyses revealed that some loci included more than one gene associated with CeD, thus raising the number of the involved polymorphisms (single nucleotide polymorphisms, SNPs) up to 57. Although these studies were quite extensive, the identified loci were not able to completely explain CeD genetic predisposition, since HLA + all these additional genes accounted for about 54% of the hereditability of CeD[7]. Interestingly, a great number of these SNPs involved genes with an immune function either in the intestine or in the thymus, further supporting the idea that an alteration of the immune response represents an essential step in CeD predisposition and pathogenesis. Moreover, a more recent paper was able, integrating genomic and transcriptomic data, to prioritize genes involved in CeD, and to identify TRAF-type zinc finger domain containing 1 (TRAFD1) as a master regulator of genes involved in interferon (IFN)γ signaling and MHC I antigen processing/presentation[8]. The different combination of these SNPs could also be associated with a different phenotype, as recently reported by Cerquieira *et al*[9], who identified the TLR7/ TLR8 Locus associated with disease onset before 7 years of age, whereas SH2B3/ATXN2, ITGA4/UBE2E3 and IL2/IL21 Loci were associated with later development of CeD and a more severe small bowel mucosal damage. In addition, SH2B3/ATXN2 was associated with type 1 diabetes; in fact some of the

identified SNPs are described as predisposing factors also for other autoimmune disorders like type 1 diabetes and Crohn's disease[7,10], almost suggesting the presence of a common genetic background predisposing to autoimmunity. It must also be noted that, among the identified SNPs, only 5% were present in coding regions, and about 5% and 9% in 5' and 3' untranslated regions, respectively[7]. This means that 81% of the identified SNPs were either in intergenic or intronic regions, suggesting that their role could be to regulate gene expression, possibly through the interaction with transcription factors or proteins able to regulate chromatin status, *i.e.*, epigenetic modifications. Epigenetics has gained much attention in recent years, but its involvement in CeD still needs to be well characterized. First of all, this term has been used to define a wide range of concepts, and the 2008 Cold Spring Harbor meeting defined epigenetics as *"the study of a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence"*[11]. DNA methylations, histone modifications resulting in chromatin remodelling and non-coding RNAs are the three main categories that fall into this statement. The latter one also includes microRNAs, that have been found deregulated in CeD patients, suggesting that the regulation of the expression of target mRNAs by post-transcriptional modifications is relevant in CeD pathogenesis[12-14]. In this direction, RNA sequencing studies are currently trying to identify different gene expression signatures that could help stratify patients based on CeD stage/presentation, or highlight new pathways implicated in CeD development[15-17].

In this review we will provide an update on the current knowledge about epigenetics in CeD and investigate the possible role of epigenetic profiling in patient classification and/or in determining the risk in subjects with familiarity to CeD.

LITERATURE SEARCH

The literature search was carried out on PubMed, searching the key words "epigenetics", "histones -acetylation, -methylation, -phosphorylation, -ubiquitination", "chromatine", "non-coding RNA", "lncRNA", "miRNA", "microRNA", "RNA methylation" AND "celiac disease". The reviewed papers were original research articles published from 2010 to October 2021. All articles that were considered relevant to the purpose of the present review were included, whereas the ones that did not add novelty or did not give clear results were excluded during the critical revision of the literature.

DNA METHYLATION

DNA methylation is the most known epigenetic modification, that sees the family of DNA methyltransferase (DNMTs) enzymes transfer methyl-groups to the C-5 residue of cytosine pyrimidine ring. This phenomenon occurs in particular in CpG-rich regions (CpG islands), mainly localized in promoters and regulatory regions. Being the methylated DNA physically less accessible for transcription, this causes the inactivation of the nearby genes[18]. The majority of gene promoters are normally methylated to guarantee a fine regulation of transcription in different tissues or cell types. Like most of the epigenetic modifications, it does not directly alter the DNA sequence, but a typical methylation profile can usually be inherited. Hypo-methylation or hyper-methylation of specific genomic loci has been reported to predispose to disease and cancer[19]; for this reason, the first studies on DNA methylation in CeD analyzed the predisposition to develop small bowel adenocarcinomas. In particular, Bergmann *et al*[20] found a high level of CpGs methylation and microsatellite instability correlated to the loss of MLH1 expression in three different small bowel carcinomas in CeD patients, whereas this feature was not present in non-CeD patients. A similar finding was observed by Diosdado *et al*[21], who also detected an hyper-methylation of the *APC* gene promoter that caused defects in the mismatch repair mechanisms in these patients. Further studies were conducted by Rizzo *et al*[22], who were able to identify four different CpG Island Methylator Phenotypes, with two of them being specific for CeD patients.

Although these data highlight a typical methylation pattern in CeD complications, the identification of a methylation profile associated with the predisposition for developing CeD is essential. Since GWAS pointed out the presence of many risk variants common to other autoimmune diseases, it has been hypothesised that they could also share common methylation patterns. Hammaker *et al*[23] focused on the

specific risk variant rs906868, shared between CeD and rheumatoid arthritis and mapping in the promoter of LBH, gene with a regulatory function on the Wnt pathway, essential for differentiation and development. Interestingly, the authors detected a differential methylation pattern in the two diseases, thus suggesting the possibility of a disease-specific profile[23]. Methylation can also occur only on one allele, thus not causing a complete silencing of the gene but its modulation. Hutchinson *et al*[24] investigated allele-specific methylation (ASM) on CpG islands localized throughout the genome, detecting four of them as implicated in complex diseases. In particular, rs2762051, a C/T risk variant that maps in the non-coding RNA DLEU1, undergoes ASM in CeD[24]. It is interesting to note that this variant belongs to the group of SNPs previously identified as significantly associated to CeD predis-position[7], thus further supporting the need for an integration between the genetic and epigenetic profile. A specific gene methylation could also be associated to a different phenotype, in particular if the methylation site is able to influence the expression of disease-related proteins. In fact, a lower level of methylation was detected in different IFNγ CpGs in patients with ulcerative colitis (UC) and a more severe phenotype[25], suggesting that methylation can regulate this UC pivotal pathogenetic pathway. Fernandez-Jimenez *et al*[26] demonstrated that this can also happen in CeD, since they observed not only a different methylation profile in the genes of the core NFkB pathway in active CeD subjects as compared to controls, but also an inter-mediate pattern in CeD patients on gluten-free diet (GFD). Although it was not clear if gliadin could directly affect the methylation of these specific genes, long exposure to gliadin could be able to affect the methylation pattern, since proline-rich peptides seem to act as opioid-like molecules causing a modulation in glutathione activity and DNA methylation[27].

It is also known that methylation profiles can be cell type-specific, thus a single cell methylation analysis could provide essential information. However, the methods employed to obtain single cell preparations can sometimes alter the DNA methylation patterns as described by Jenke *et al*[28], thus pointing out the importance of considering this aspect in designing this kind of studies. Another potential bias of methylation studies regards the preferential amplification of an allele or a strand, based on its methylation degree. For this reason Ochoa *et al*[29] realized a new Bayesan calibration method and validated it also on CeD patients samples, and the most recent studies on methylation in CeD considered these techniques. Fernandez-Jimenez *et al* [30] compared the whole methylome of the epithelium and immune cells from CeD and controls biopsies, finding a cell-specific methylation pattern, with 43 (11 hypo-, 32 hyper-methylated) and 310 (40 hypo-, 270 hyper-methylated) differentially methylated positions (DMP) in the epithelium and in immune cells respectively. It is important to note that only a portion of the DMPs was shared between the two cellular compo-nents, highlighting a cell-specific, disease-driven modification. Cielo *et al*[31] used a different approach, separating the epithelium from the lamina propria and analysing the expression and methylation of genes known to be altered in CeD. Among the candidate genes, they found a decreased methylation in SB2H3 and IL21, that coupled with their increased expression in the epithelium and lamina propria respectively, but only in active CeD compared to controls. Since these genes take part in the pathways mainly involved in inflammation and barrier integrity in CeD, identifying a different signature in patients and controls may partially explain the abnormal response to food antigens and help patient stratification[31]. Finally, the most recent study in this field by Hearn *et al*[32] compared methylation profiles in saliva samples obtained from CeD patients on a GFD and controls with the use of a beadchip array, finding a different methylation pattern in HLA-DQB1 and in specific loci near *SLC17A3* gene in a pilot cohort, data not confirmed in a larger study group. Being this collection method non-invasive, it could be a great tool for patient discrimination, especially in the screening of predisposed subjects, once a different and discriminatory methylation profile has been clearly established. Moreover, the patient histological classification needs to be taken into account, even if a partial concordance in methylation patterns in saliva and intestine has been observed in a small group of patients[33]. In this direction, encouraging results come from studies on inflammatory bowel diseases (IBDs) from Howell *et al*[34], that were able to statistically discriminate different classes of patients by analysing the transcriptome and the methylation status in the intestinal epithelium, correlating them with the disease outcome. Indeed, methylation profiling can be an asset to study different aspects of CeD, both to understand its pathogenesis and as a biomarker. A graphical representation of DNA methylation is present in Figure 1, whereas its highlights in CeD are reported in Table 1.

Table 1 DNA methylation features in celiac disease

Predisposition to CeD	What	Result	Highlight in CeD
Allele-specific methylation (ASM)[24]		ASM in rs2762051 in *DLEU1* gene	Linked to CeD phenotype
Rs906868 in *LBH* gene promoter [23]	Risk variant shared with RA	Disease-specific methylation	Different methylation in rs906868 can predispose to CeD or RA → influence on Wnt signalling
Methylation in HLA region in CeD[30]	Specific patterns in epithelial and immune cells	Genotype-independent methylation (except for HLA-DPB2)	Methylation patterns in HLA region → CeD predisposition
Methylation profiling in HLA-DQB1 and SLC17A3[32]	Bead-chip on saliva samples	Different methylation profiles, not confirmed in bigger cohort	Potential methylation-based screening Major validation needed
Opioid like-effect of gliadin[27]		Modulation of glutathione and DNA methylation	Predisposition to inflammation and oxidation
CeD pathogenesis			
Methylation in NFkB-related genes[26]	NFkB pathway↑↑	Disruption of regulatory equilibrium	Co-methylation patterns typical of active CeD in NFkB pathway genes
Cell-specific methylation[30]		Epithelium → 43 DMP Immune cells → 310 DMP	Cell-specific methylation signature & gene expression in CeD *vs* controls
Different methylation of SB2H3, IL-21, cREL and TNFAIP3[31]	Epithelium and lamina propria - specificity	Correlation with pro-inflammatory(↑) and cell adhesion(↓) pathways	Methylation of SB2H3 → epithelium Methylation of IL21 → lamina propria Typical of CeD samples.
Tumor development			
↑CpGs methylation[20]	Microsatellite instability	↓MLH1 expression	Typical in CeD-related small bowel adenocarcinoma
MLH1 deregulation[21]	*APC* gene hyper-methylation	Chromosomal aberrations/microsatellite instability	Defects in mismatch repair Typical in CeD-related small bowel adenocarcinomas
Microsatellite instability[22]	Typical CpGs methylator phenotype	Methylation profiling → phenotypical classification	Mesenchymal and immune phenotypes are common in CeD-related small bowel carcinoma

CeD: Celiac disease; ASM: Allele specific methylation; RA: Rheumatoid arthritis; DMP: Differentially methylated positions; CpGs: CpG islands.

HISTONE MODIFICATIONS

A further step in gene expression regulation in the nucleus is represented by histone modifications. Histones are key proteins of chromatin, able to modify its structure making the DNA more or less accessible to transcription. The most common histone modifications are acetylation, phosphorylation, ubiquitination and methylation. Gene expression analysis studies performed on CeD biopsies compared to controls, pointed out a differential regulation of histone-modifying enzymes, thus suggesting the presence of a disease-related epigenetic signature involving histone modifications[16, 17].

Histone acetylation, governed by acetyltransferase (HATs) and deacetylase (HDACs) enzymes, usually result in active gene transcription due to chromatin relaxation[35]. Zorro *et al*[36] analyzed the transcriptomic and epigenetic responses to IL-15, IFNβ and IL-21 stimulation in primary cytotoxic T lymphocytes derived from CeD biopsies. Specific transcriptomic patterns were identified for the different cytokine stimulation, but also different levels of histone acetylation (H3K27ac) were detected. Interestingly, the increase in transcription after IFNβ stimulation was associated, in about 60% of the cases, with an increase in H3K27ac in promoter and enhancer regions, whereas this was not the case after IL-15 stimulation, pointing towards the presence of a different regulatory mechanism. In fact, upon IL-15 stimulation a relevant number of differentially expressed genes were ncRNAs, suggesting their potential regulatory role independently from H3K27ac modifications
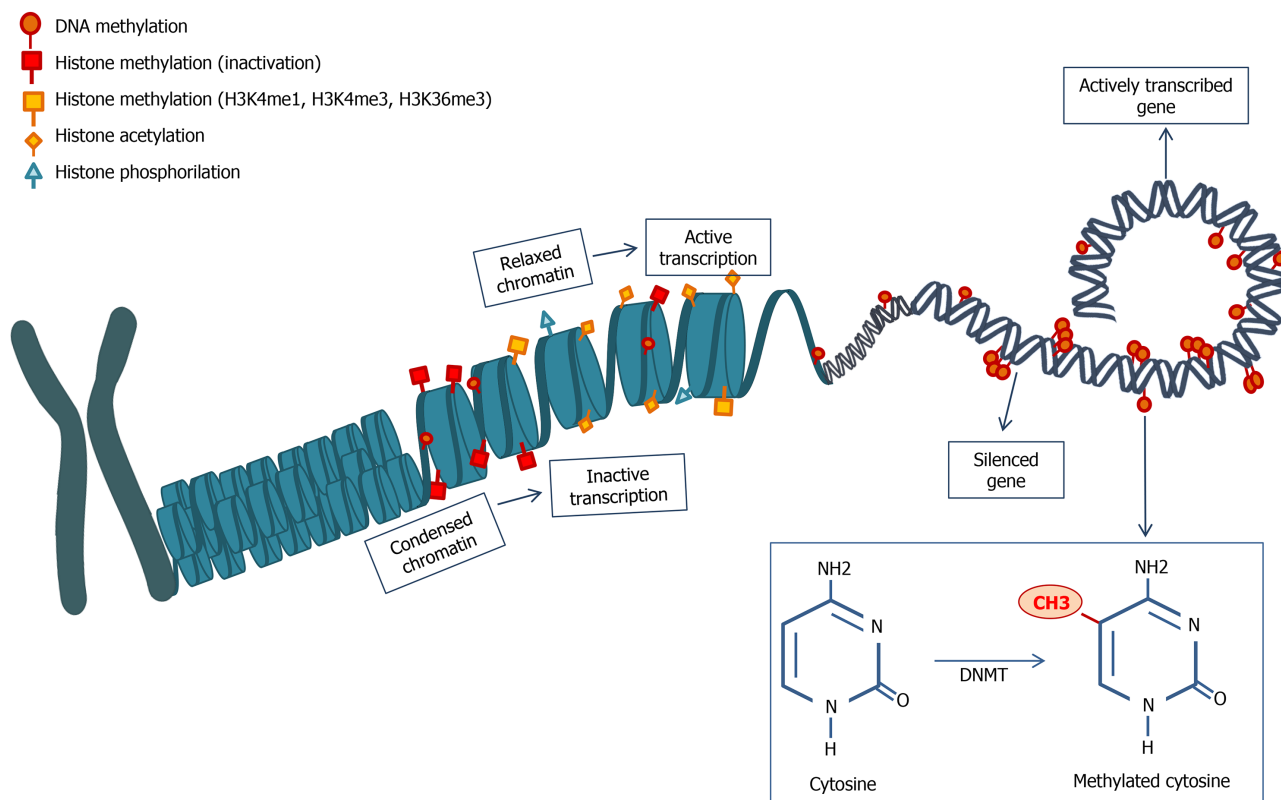
Figure 1 Schematic representation of DNA methylation and histone modifications. Histone modifications are many and determine a different chromatin status, modifying DNA accessibility and interfering with gene transcription. Methylation can also happen directly on the DNA sequence, mainly resulting in gene silencing. The image is original and was created with the use of Servier Medical Art modified templates, licensed under a Creative Common Attribution 3.0 Unported License (https://smart.servier.com). DNMT: DNA methyltransferase.

[36].

Histone methylation is more complex and mainly linked to repression, but the specific methylations H3K4me1, H3K4me3 and H3K36me3 are commonly found in actively transcribed regions[35]. In fact, H3K4me3 was used as a marker of active transcription by Gutierrez-Achury *et al*[37]; the authors firstly identified SNPs associated with both CeD and rheumatoid arthritis, but also SNPs specific for each disease. The combination of these data with H3K4me3 profiles available in public databases allowed the authors to detect similar histone enrichment corresponding to the common SNPs in less specialized immune cell types, whereas disease-specific SNPs overlapped with H3K4me3 profiles in more specialized cells[37]. H3K36me3 signature was investigated by Moffitt *et al*[38] in the enteropathy associated T cell lymphoma (EATL), the most common neoplastic consequence of CeD. The histone-lysine N-methyltransferase SETD2 is the responsible for this kind of methylation and it was silenced in 32% of EATL in CeD subjects. In fact, *in vivo* studies with a knock out model for SETD2 in T cells carried a decreased H3K36me3 pattern along with γδ-T cells expansion[38]. H3K27me3 was instead used as gene silencing signature by Oittinen *et al*[39] in determining the involvement of the polycomb repressive complex 2 (PRC2) in controlling Wnt signalling in the intestine. They found a H3K27me3 pattern which differed according to the cell position/differentiation along the crypt/villus axis, and involved also genes related to proliferation and differentiation in the epithelium. Thus PRC2-driven tri-methylation is important to maintain the homeostasis driven by Wnt; however, since Wnt pathway deregulation has been connected to CeD pathogenesis, they also hypothesise that a Wnt/PRC2 disrupted axis could cause the development of crypts hyperplasia in CeD[39] (Table 2). However, among the external factors that can influence the epigenetic regulation, the interaction with the microbiota needs also to be considered, since changes in the gut bacterial populations can affect the activity of enzymes involved in epigenetic regulation[40,41].

Table 2 Celiac disease-relevant histone modifications

Modification	Focus	Result	Relevance in CeD
H3K27ac [36]	Activation and enhancing of transcription	IL-15, IFNβ and IL-21 induce specific acetylation profiles in CTLs	Strong association between H3K27ac and gene expression IFNβ-induced
			IL-15-derived changes related to lncRNAs
H3K4me3 [37]	Active transcription marker	Shared variants between CeD and RA	Similar histone enrichment in shared variants in simple cells	Different histone enrichment in disease-related specialized cells
H3K36me3 [38]	Active transcription marker	SETD2 silenced in 32% of EATL	↓H3K36me3 → γδT cells expansion	Predisposition to lymphomagenesis
H3K27me3 [39]	Gene silencing	PRC2–driven trimethylation	Villi → H3K27me3 on proliferation and differentiation genes	PRC2 methylation help maintenance of Wnt homeostasis → deregulation linked to CeD crypts hyperplasia
		Crypts → H3K27me3 on nutrient transport and cell killing genes	

CeD: Celiac disease; CTLs: Cytotoxic T lymphocytes; LncRNAs: Long non-coding RNAs; ECM: Extracellular matrix; RA: Rheumatoid arthritis; EATL: Enteropathy associated T cell lymphoma; PRC2: Polycomb repressive complex 2.

LONG NON-CODING RNAS

Long non-coding RNAs (lncRNAs) have emerged in recent years as a class of transcripts with a wide spectrum of mechanisms of action that can affect gene expression regulation both at transcriptional and post-transcriptional level, in the nucleus and in the cytosol. Their genomic location is various, either being intergenic or within the introns of coding genes. Among their many mechanisms of action, the most frequent are direct DNA sequence interaction, transcription factor sequestration, chromatin rearrangement, regulation of histone modifications and microRNA sponging (Figure 2)[42,43]. Being implicated in so many regulatory patterns, they have been studied in cancer and in many other diseases, also with an autoimmune background[44].

The pathogenesis of CeD involves several different mechanisms, *i.e.*, the passage of the gluten peptides across the intestinal barrier, but also the activation of innate and adaptive immune response. LncRNAs can influence all these processes, but the expression of these RNAs can, in turn, be affected by the presence of SNPs.

A typical aspect in CeD pathogenesis is the loosening of the tight junctions (TJ), that increases the intestinal permeability[45]. Jauregi-Miguel *et al*[46] identified, in a genomic region associated with CeD, a lncRNA named RP4-587D13.2. This is located in an intron of *MAGI2* gene, a scaffold protein present in the tight junction plaque and able to regulate epithelial integrity. Interestingly, RP4-587D13.2 can regulate MAGI2 expression, altering the downstream TJ-related proteins like CLDN1 and ZAK. RP4-587D13.2 and MAGI2 resulted also downregulated in CeD samples (both with active CeD and on a GFD) and it seems that gliadin stimulation can reinforce this effects[46].

Ricaño-Ponce *et al*[47] identified genes located in close proximity to autoimmune-related SNPs, revealing that 42 of these SNPs could specifically affect the expression of 53 non-coding RNAs. In particular, two lncRNAs were associated to CeD-specific SNPs, namely AP002954.4 and AC104820.2. In both cases the SNPs were able to affect the expression of these lncRNAs, which had a role in the immune response[47]. Interestingly, AC104820.2 had already been detected upregulated in active CeD patients intestinal mucosa by Plaza-Izurieta *et al*[48], who also identified the SNP rs1018326 as associated with this lncRNA. These data thus support the involvement of lncRNAs in the immune response in CeD, however there are aspects that must be kept in mind, as also reported by Hrdlickova *et al*[49]. In fact, although the authors suggested the lncRNAs RP3-395 M20.9 and IL21-AS to be relevant in CeD, and linked them to TNF and IL-21 pathways, they also reported that the same lncRNAs are involved in different autoimmune disorders. In addition, some lncRNAs were detected only in some specific cell subtypes, underlying the need for a cell-specific analysis[49].

Castellanos-Rubio *et al*[50] identified lnc13 as associated to the susceptibility to celiac disease, demonstrating also the functional role of this lncRNA. Lnc13 binds the nuclear ribonucleoprotein hnRNPD and histone deacetylase 1, forming a complex that acts as a repressor of inflammatory gene expression. The presence of inflammatory stimuli activate the NFkB pathway that reduces lnc13, thus removing the repression
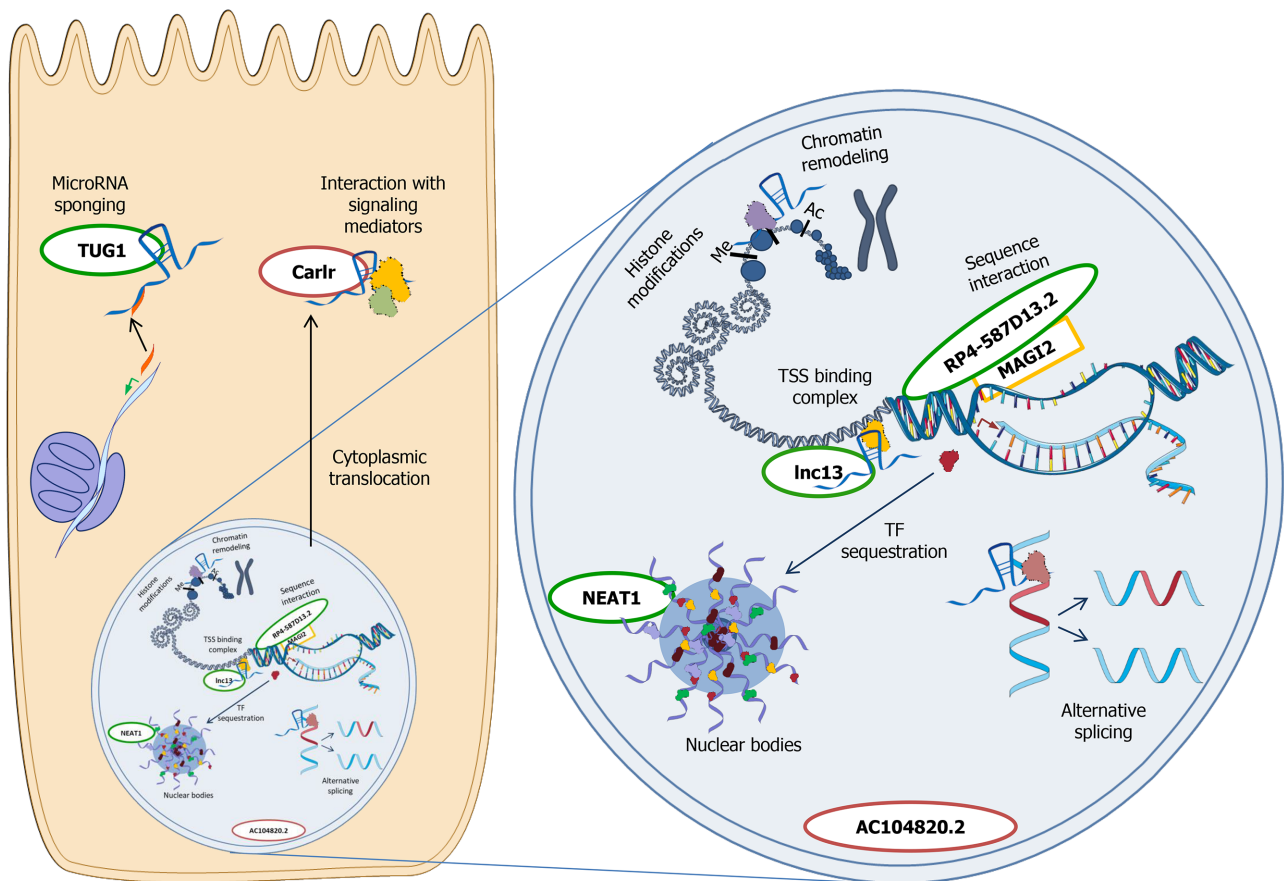
Figure 2 Main mechanisms of action of long non-coding RNAs in the cell. Cytoplasm - microRNA sponging: microRNAs are sequestered from their mRNA targets, resulting in mRNA translation; Interaction with signalling mediators: binding to pathway mediators can modulate downstream targets activation. Nucleus - chromatin remodeling: Long non-coding RNAs (lncRNAs) interact with chromatin and modify its conformation; histone modifications: lncRNAs influence the activity of the enzymes responsible for histone modifications; sequence interaction: lncRNAs act on nearby genes transcription; transcription start site (TSS) binding complex: lncRNAs bind ribonucleoproteins, interfering with gene transcription; transcription factor (TF) sequestration in nuclear bodies: TF are moved to nuclear bodies from the promoter region of target genes, influencing gene transcription. Nuclear bodies like paraspeckles consist in a lncRNA scaffold and target proteins; alternative splicing: lncRNAs determine a preferential splicing in favour of an isoform respect to another one. The so-far investigated lncRNAs in celiac disease (CeD) are reported near their known mechanism of action, respecting their main cellular localization. lncRNAs in green were reported to be downregulated in CeD, whereas the ones in red were found upregulated in a specific cell compartment. The image is original and was created with the use of Servier Medical Art modified templates, licensed under a Creative Common Attribution 3.0 Unported License (https://smart.servier.com). TSS: Transcription start site; TF: Transcription factor.

on gene transcription. Interestingly, in CeD patients, not only lnc13 expression is reduced, but there is also a SNP that generates a lnc13 variant that binds hnRNPD less effectively, possibly contributing to CeD development[50]. NFkB activation is central in CeD-driven inflammation, and it is able to induce the expression of another lncRNA, *i.e.*, Carlr. This molecule is localized in the nucleus, but it translocates in the cytosol after the activation of the NFkB pathway, and it seems essential for the induction of downstream genes such as IL-1β in macrophages. However, differently from what expected, Carlr was downregulated in total biopsies from active CeD patients, although its cytosolic fraction was increased; this observation thus needs further evaluation to determine the real role of this lncRNA in the intestinal tissue[51].

Finally, it is known that IL-15 production has a central role in CeD pathogenesis [52]. Zorro *et al*[36] in their work highlighted a subset of ncRNAs that were downregulated by IL-15 stimulation, suggesting the involvement of this class of transcripts in IL-15 response. A recent study by our group demonstrated that gliadin induces TUG1 and NEAT1 expression in biopsies obtained from CeD patients on gluten-free diet, mechanism that depends on innate immunity activation in the case of NEAT1. In fact, IL-15 can trigger NEAT1 expression both *ex vivo* and in vitro, through the newly identified axis IL-15/STAT3/NEAT1[53]. The effects of NEAT1 upregulation deserve more investigation, since this lncRNA localizes in the nucleus and is able to bind transcription factors, thus sequestering them and, in turn, regulating the expression of several genes, such as IL-8[54].

RNA METHYLATION

A further regulation can be played by RNA modifications. In fact, recent findings are pointing out that direct sequence methylation can also happen at RNA level. Differently from DNA, RNA can undergo many more kinds of methylation (up to 72), involving both mRNA and non-coding RNAs. Previous studies found RNA methylation linked to biological and immune processes, as well as complex pathologies like obesity[55]. Up to now, the main limitation in RNA methylation studies was the difficulty to perform a site-specific detection and quantification. Using a recently developed method that allows to detect m6A, the most frequent RNA modification, Olazagoitia-Garmendia *et al*[56] characterised the CeD-associated SNP rs3087898, located in the 5' untranslated region of exportin1 gene, a modulator of NFkB pathway. The mRNA derived from the CeD-associated variant was preferentially methylated with respect to the normal one, and this modification was able to increase the protein production, resulting in an inflammatory microenvironment. Besides, gliadin was able to increase this modification and exportin1 protein levels, both *in vitro* and in CeD patients samples[56]. Altogether, these studies suggest that attention should be given also to this kind of modifications when focusing on gene expression regulation in CeD.

MICRORNAS

MicroRNAs (miRNAs) are short RNA sequences (20-25 nucleotides) accounting for about 30% of gene expression regulation, affecting various processes, including chromatin conformation and transcription, as well as mRNA stability and translation. Usually they present an inverse relationship with their targets, inducing their downregulation.

miRNAs can also be detected in the plasma, and for this reason they have been evaluated in a large number of different disorders, in the attempt to identify specific biomarkers for early diagnosis or follow up. Initial papers on CeD screened the duodenal tissue to identify a specific signature representative of the events in the mucosa, and to determine whether the "signature" miRNAs were returning to normal levels in patients on GFD. Capuano *et al*[57] evaluated biopsies of children with CeD at diagnosis or on GFD, identifying about 20% of the analyzed miRNAs as differentially regulated. The authors then focused on the upregulation of miR-449a[57], which targets Notch1 and KLF4, genes involved in intestinal cell proliferation and differentiation. Indeed Notch1 expression was significantly downregulated in CeD biopsies, confirming the regulatory role of miRNAs also on proliferation in CeD. Other miRNAs were also downregulated, *i.e.*, miR-124a, miR-189, miR-299-5p and miR-379, similarly to what reported in Crohn's disease[58].

Studies on adult CeD patients identified different subset of miRNAs according to the clinical picture or the severity of intestinal damage[12,14]. A significant downregulation of miR-31-5p was observed in all the CeD groups, whereas miR-192-3p and miR-192-5p were downregulated in CeD patients with anemia and a severe histological lesion, respectively. miR-192-5p targeted two different molecules involved in the innate immunity, *i.e.*, NOD2 and CXCL2, that were upregulated in CeD patients, in particular in severe cases[12]. On the other hand, miR-31-5p had as a target Foxp3, essential for Treg development; again, a significant inverse correlation was observed between the miRNA and the target mRNA. These changes were induced by gliadin, as demonstrated by *in vitro* stimulation of biopsies of patients on GFD[12]. It is interesting to note that the data on miR-192 are similar to those observed in IBDs, in which miR-192 regulated NOD2 expression[59].

However, a more comprehensive approach was needed, and a recent paper correlated, in the duodenal mucosa of celiac subjects, the miRNA and mRNA expression patterns obtained by sequencing[17]. The data analysis revealed the presence of a complex network involving various pathways known to be deregulated in CeD, such as immunity (interferon), suggesting that CeD-associated miRNAs play a central role in causing the intestinal damage.

The identification of a specific miRNA plasma signature remains elusive at the present time. An analysis performed in pediatric CeDs revealed the presence of a trend similar to that observed in duodenal biopsies for miR-192-5p, miR-31-5p and miR-21-5p. However none of these miRNAs could be defined as a marker able to evaluate the recovery of the mucosa on GFD, since there was no return to normal levels (miR192-5p) or a clear cut-off could not be established (miR-31-5p and miR-21-5p)[13]. A

further study performed by Amr *et al*[60] analyzed the same miRNAs, *i.e.*, miR-21 and miR-31, identifying a cutoff value and determining a sensitivity and specificity of 82.4% and 80.8% for miR-21 and 93.8% and 72% for miR-31. Another recent paper confirmed the value of plasmatic miR-21 as a marker, with a sensitivity and specificity of 0.65 and 0.83, but better results were obtained considering miR-155 (0.94 and 0.87, respectively)[61]. Although encouraging, these results need further validation in larger cohorts; moreover even for miR-21 the identified cutoff values were quite dissimilar, thus suggesting the need to determine the cutoff according to the method and equipment employed for the analysis. Last but not least, the data obtained in CeD patients will have to be compared with those obtained from subjects with other inflammatory disorders such as IBDs, to confirm the specificity of the findings.

MACHINE LEARNING

As described, the evaluation of the epigenome generates large datasets, and the data increases when the integration of different datasets (such as those derived from methylome and long/short non-coding RNA transcriptome) is required. The generation of a large quantity of data allows to move towards a data-driven analysis, *i.e.*, machine learning or artificial intelligence (these terms are now regarded as interchangeable). Machine learning can, through statistical methods and algorithms, make classifications or predictions of input data, such as recognize a specific endoscopy pattern and classify the patient as celiac or non-celiac. In simpler machine learning approaches supervised learning is commonly used, *i.e.*, the use of labeled datasets designed to train or "supervise" algorithms into classifying data or predicting outcomes. In the medical field this obviously requires training datasets selected by specialized clinicians (for example the identification of endoscopic images characteristic of celiac disease). Deep learning is a subset of machine learning, in which the algorithms are more complex and include more than one layer in the neuronal network. For deep learning the presence of a labeled training dataset is not required, since these algorithms can analyze data in the raw form and determine the features that distinguish the various groups.

With regards to CeD, the machine learning approach has been originally used to analyze endoscopic data, either those generated by upper endoscopy[62] or by videocapsule and, in this latter case, the deep learning approach was able to reach a high sensitivity and specificity in the diagnosis of CeD using the video capsule images [63,64]. Although these data seem quite encouraging, it must be kept in mind that the gastrointestinal tract is a complex environment, and several factors can affect the obtained images, as discussed in depth by Hegenbart *et al*[65]. The deep learning approach could be useful also in the evaluation of duodenal biopsies, as demonstrated by different authors that were able to develop artificial intelligence-based methods for the correct classification of duodenal samples[66-68]. Machine learning has also been employed to improve the diagnosis based on B/T cell repertoire[69,70]; although the data obtained in the latter case are quite interesting, this approach is not easily applicable for routine diagnosis.

A machine learning approach could have a very important clinical application should it be able to predict who, within a risk population, will develop the disease. An encouraging study that analyzes this aspect has recently been published by Piccialli *et al*[71], who tested different models in order to predict the development of an overt CeD in childen with potential CeD. The use of a machine learning approach towards basic science data obtained from duodenal biopsies of CeD patients has recently been published by Wolf *et al*[16], who analyzed not only the trascriptomic profiles but also the expression of some lncRNAs and miRNAs. Moreover, the authors focused on gene expression signatures which associate with transcription factor (TF)-activity and chromatin state, as well as DNA and histone methylation pattern, thus providing an indirect measure of epigenetic modifications. This comprehensive analysis revealed, for example, a different role of miRNAs and lncRNAs in the regulation of gene expression, correlating the latter category with more subtle adjustments of the transcription under control of epigenetic mechanisms.

CONCLUSION

Understanding the epigenetics of celiac disease remains mandatory in order to completely clarify the pathogenetic processes behind its development and the
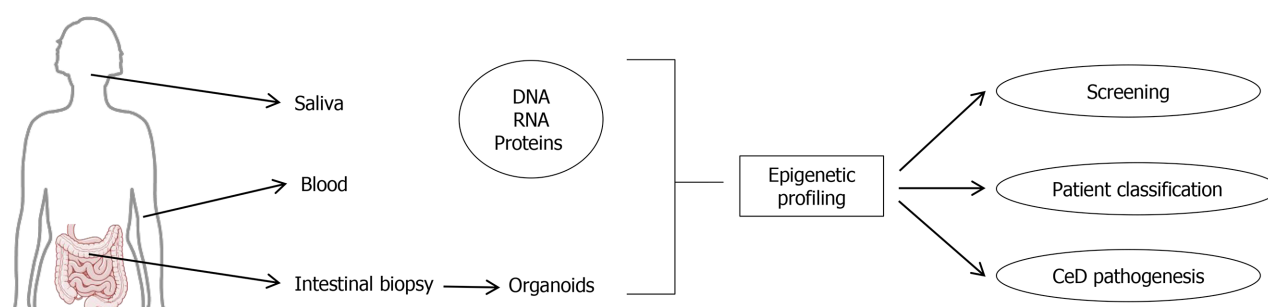
Figure 3 Potential use of epigenetics studies in celiac disease. The image is original and was created with the use of Servier Medical Art modified templates, licensed under a Creative Common Attribution 3.0 Unported License (https://smart.servier.com). CeD: Celiac disease.

spectrum of its manifestations. The data here reviewed demonstrate that epigenetic changes occur in celiac disease, however further data are still needed before the identification of these changes could be used to screen or better categorize the patients, as depicted in Figure 3. In particular it will be essential to verify that the characteristics detected at a peripheral level (*i.e.*, blood or saliva) correspond to those present in the intestinal mucosa. In this regard, data seem encouraging for miRNAs, but the identification of a specific signature will probably require the combination of more than one of them to provide an excellent sensitivity and specificity.

REFERENCES

1   **Lebwohl B**, Sanders DS, Green PHR. Coeliac disease. *Lancet* 2018; **391**: 70-81 [PMID: 28760445 DOI: 10.1016/S0140-6736(17)31796-8]

2   **Rostami-Nejad M**, Romanos J, Rostami K, Ganji A, Ehsani-Ardakani MJ, Bakhshipour AR, Zojaji H, Mohebbi SR, Zali MR, Wijmenga C. Allele and haplotype frequencies for HLA-DQ in Iranian celiac disease patients. *World J Gastroenterol* 2014; **20**: 6302-6308 [PMID: 24876751 DOI: 10.3748/wjg.v20.i20.6302]

3   **Elli L**, Ferretti F, Orlando S, Vecchi M, Monguzzi E, Roncoroni L, Schuppan D. Management of celiac disease in daily clinical practice. *Eur J Intern Med* 2019; **61**: 15-24 [PMID: 30528262 DOI: 10.1016/j.ejim.2018.11.012]

4   **Petronzelli F**, Bonamico M, Ferrante P, Grillo R, Mora B, Mariani P, Apollonio I, Gemme G, Mazzilli MC. Genetic contribution of the HLA region to the familial clustering of coeliac disease. *Ann Hum Genet* 1997; **61**: 307-317 [PMID: 9365784 DOI: 10.1046/j.1469-1809.1997.6140307.x]

5   **Zhong F**, McCombs CC, Olson JM, Elston RC, Stevens FM, McCarthy CF, Michalski JP. An autosomal screen for genes that predispose to celiac disease in the western counties of Ireland. *Nat Genet* 1996; **14**: 329-333 [PMID: 8896565 DOI: 10.1038/ng1196-329]

6   **Greco L**, Corazza G, Babron MC, Clot F, Fulchignoni-Lataud MC, Percopo S, Zavattari P, Bouguerra F, Dib C, Tosi R, Troncone R, Ventura A, Mantavoni W, Magazzù G, Gatti R, Lazzari R, Giunta A, Perri F, Iacono G, Cardi E, de Virgiliis S, Cataldo F, De Angelis G, Musumeci S, Clerget-Darpoux F. Genome search in celiac disease. *Am J Hum Genet* 1998; **62**: 669-675 [PMID: 9497251 DOI: 10.1086/301754]

7   **Kumar V**, Wijmenga C, Withoff S. From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases. *Semin Immunopathol* 2012; **34**: 567-580 [PMID: 22580835 DOI: 10.1007/s00281-012-0312-1]

8   **van der Graaf A**, Zorro MM, Claringbould A, Võsa U, Aguirre-Gamboa R, Li C, Mooiweer J, Ricaño-Ponce I, Borek Z, Koning F, Kooy-Winkelaar Y, Sollid LM, Qiao SW, Kumar V, Li Y, Franke L, Withoff S, Wijmenga C, Sanna S, Jonkers I; BIOS Consortium. Systematic Prioritization of Candidate Genes in Disease Loci Identifies *TRAFD1* as a Master Regulator of IFNγ Signaling in Celiac Disease. *Front Genet* 2020; **11**: 562434 [PMID: 33569077 DOI: 10.3389/fgene.2020.562434]

9   **Cerqueira JXM**, Saavalainen P, Kurppa K, Laurikka P, Huhtala H, Nykter M, L E Koskinen L, Yohannes DA, Kilpeläinen E, Shcherban A, Palotie A, Kaukinen K, Lindfors K. Independent and cumulative coeliac disease-susceptibility loci are associated with distinct disease phenotypes. *J Hum Genet* 2021; **66**: 613-623 [PMID: 33446885 DOI: 10.1038/s10038-020-00888-5]

10   **Trynka G**, Wijmenga C, van Heel DA. A genetic perspective on coeliac disease. *Trends Mol Med* 2010; **16**: 537-550 [PMID: 20947431 DOI: 10.1016/j.molmed.2010.09.003]

11   **Berger SL**, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. *Genes Dev* 2009; **23**: 781-783 [PMID: 19339683 DOI: 10.1101/gad.1787609]

12   **Magni S**, Buoli Comani G, Elli L, Vanessi S, Ballarini E, Nicolini G, Rusconi M, Castoldi M, Meneveri R, Muckenthaler MU, Bardella MT, Barisani D. miRNAs affect the expression of innate and adaptive immunity proteins in celiac disease. *Am J Gastroenterol* 2014; **109**: 1662-1674 [PMID: 25070052 DOI: 10.1038/ajg.2014.203]

13   **Buoli Comani G**, Panceri R, Dinelli M, Biondi A, Mancuso C, Meneveri R, Barisani D. miRNA-regulated gene expression differs in celiac disease patients according to the age of presentation. *Genes Nutr* 2015; **10**: 482 [PMID: 26233308 DOI: 10.1007/s12263-015-0482-2]

14   **Vaira V**, Roncoroni L, Barisani D, Gaudioso G, Bosari S, Bulfamante G, Doneda L, Conte D, Tomba C, Bardella MT, Ferrero S, Locatelli M, Elli L. microRNA profiles in coeliac patients distinguish different clinical phenotypes and are modulated by gliadin peptides in primary duodenal fibroblasts. *Clin Sci (Lond)* 2014; **126**: 417-423 [PMID: 24063611 DOI: 10.1042/CS20130248]

15   **Bragde H**, Jansson U, Fredrikson M, Grodzinsky E, Söderman J. Celiac disease biomarkers identified by transcriptome analysis of small intestinal biopsies. *Cell Mol Life Sci* 2018; **75**: 4385-4401 [PMID: 30097691 DOI: 10.1007/s00018-018-2898-5]

16   **Wolf J**, Willscher E, Loeffler-Wirth H, Schmidt M, Flemming G, Zurek M, Uhlig HH, Händel N, Binder H. Deciphering the Transcriptomic Heterogeneity of Duodenal Coeliac Disease Biopsies. *Int J Mol Sci* 2021; **22** [PMID: 33806322 DOI: 10.3390/ijms22052551]

17   **Tan IL**, Barisani D, Panceri R, Modderman R, Visschedijk M, Weersma RK, Wijmenga C, Jonkers I, Coutinho de Almeida R, Withoff S.   A Combined mRNA- and miRNA-Sequencing Approach Reveals miRNAs as Potential Regulators of the Small Intestinal Transcriptome in Celiac Disease. Int J Mol Sci. 2021; 22:11382. [PMID: 34768815 DOI: 10.3390/ijms222111382].

18   **Jin B**, Li Y, Robertson KD. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer* 2011; **2**: 607-617 [PMID: 21941617 DOI: 10.1177/1947601910393957]

19   **Tost J**. DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker. *Mol Biotechnol* 2010; **44**: 71-81 [PMID: 19842073 DOI: 10.1007/s12033-009-9216-2]

20   **Bergmann F**, Singh S, Michel S, Kahlert C, Schirmacher P, Helmke B, Von Knebel Doeberitz M, Kloor M, Bläker H. Small bowel adenocarcinomas in celiac disease follow the CIM-MSI pathway. *Oncol Rep* 2010; **24**: 1535-1539 [PMID: 21042749 DOI: 10.3892/or_00001015]

21   **Diosdado B**, Buffart TE, Watkins R, Carvalho B, Ylstra B, Tijssen M, Bolijn AS, Lewis F, Maude K, Verbeke C, Nagtegaal ID, Grabsch H, Mulder CJ, Quirke P, Howdle P, Meijer GA. High-resolution array comparative genomic hybridization in sporadic and celiac disease-related small bowel adenocarcinomas. *Clin Cancer Res* 2010; **16**: 1391-1401 [PMID: 20179237 DOI: 10.1158/1078-0432.CCR-09-1773]

22   **Rizzo F**, Vanoli A, Sahnane N, Cerutti R, Trapani D, Rinaldi A, Sellitto A, Ciacci C, Volta U, Villanacci V, Calabrò A, Arpa G, Luinetti O, Paulli M, Solcia E, Di Sabatino A, Sessa F, Weisz A, Furlan D. Small-bowel carcinomas associated with celiac disease: transcriptomic profiling shows predominance of microsatellite instability-immune and mesenchymal subtypes. *Virchows Arch* 2020; **476**: 711-723 [PMID: 31696360 DOI: 10.1007/s00428-019-02675-w]

23   **Hammaker D**, Whitaker JW, Maeshima K, Boyle DL, Ekwall AH, Wang W, Firestein GS. LBH Gene Transcription Regulation by the Interplay an Enhancer Risk Allele and DNA Methylation in Rheumatoid Arthritis. *Arthritis Rheumatol* 2016; **68**: 2637-2645 [PMID: 27159840 DOI: 10.1002/art.39746]

24   **Hutchinson JN**, Raj T, Fagerness J, Stahl E, Viloria FT, Gimelbrant A, Seddon J, Daly M, Chess A, Plenge R. Allele-specific methylation occurs at genetic variants associated with complex disease. *PLoS One* 2014; **9**: e98464 [PMID: 24911414 DOI: 10.1371/journal.pone.0098464]

25   **Gonsky R**, Deem RL, Landers CJ, Derkowski CA, Berel D, McGovern DP, Targan SR. Distinct IFNG methylation in a subset of ulcerative colitis patients based on reactivity to microbial antigens. *Inflamm Bowel Dis* 2011; **17**: 171-178 [PMID: 20848535 DOI: 10.1002/ibd.21352]

26   **Fernandez-Jimenez N**, Castellanos-Rubio A, Plaza-Izurieta L, Irastorza I, Elcoroaristizabal X, Jauregi-Miguel A, Lopez-Euba T, Tutau C, de Pancorbo MM, Vitoria JC, Bilbao JR. Coregulation and modulation of NFκB-related genes in celiac disease: uncovered aspects of gut mucosal inflammation. *Hum Mol Genet* 2014; **23**: 1298-1310 [PMID: 24163129 DOI: 10.1093/hmg/ddt520]

27   **Trivedi MS**, Shah JS, Al-Mughairy S, Hodgson NW, Simms B, Trooskens GA, Van Criekinge W, Deth RC. Food-derived opioid peptides inhibit cysteine uptake with redox and epigenetic consequences. *J Nutr Biochem* 2014; **25**: 1011-1018 [PMID: 25018147 DOI: 10.1016/j.jnutbio.2014.05.004]

28   **Jenke AC**, Postberg J, Raine T, Nayak KM, Molitor M, Wirth S, Kaser A, Parkes M, Heuschkel RB, Orth V, Zilbauer M. DNA methylation analysis in the intestinal epithelium-effect of cell separation on gene expression and methylation profile. *PLoS One* 2013; **8**: e55636 [PMID: 23409010 DOI: 10.1371/journal.pone.0055636]

29   **Ochoa E**, Zuber V, Fernandez-Jimenez N, Bilbao JR, Clark GR, Maher ER, Bottolo L. MethylCal: Bayesian calibration of methylation levels. *Nucleic Acids Res* 2019; **47**: e81 [PMID: 31049595 DOI: 10.1093/nar/gkz325]

30   **Fernandez-Jimenez N**, Garcia-Etxebarria K, Plaza-Izurieta L, Romero-Garmendia I, Jauregi-Miguel

A, Legarda M, Ecsedi S, Castellanos-Rubio A, Cahais V, Cuenin C, Degli Esposti D, Irastorza I, Hernandez-Vargas H, Herceg Z, Bilbao JR. The methylome of the celiac intestinal epithelium harbours genotype-independent alterations in the HLA region. *Sci Rep* 2019; **9**: 1298. [PMID: 30718669 DOI: 10.1038/s41598-018-37746-6]

31   **Cielo D**, Galatola M, Fernandez-Jimenez N, De Leo L, Garcia-Etxebarria K, Loganes C, Tommasini A, Not T, Auricchio R, Greco L, Bilbao JR. Combined Analysis of Methylation and Gene Expression Profiles in Separate Compartments of Small Bowel Mucosa Identified Celiac Disease Patients' Signatures. *Sci Rep* 2019; **9**: 10020 [PMID: 31292504 DOI: 10.1038/s41598-019-46468-2]

32   **Hearn NL**, Chiu CL, Lind JM. Comparison of DNA methylation profiles from saliva in Coeliac disease and non-coeliac disease individuals. *BMC Med Genomics* 2020; **13**: 16 [PMID: 32014011 DOI: 10.1186/s12920-020-0670-9]

33   **Hearn NL**, Coleman AS, Ho V, Chiu CL, Lind JM. Comparing DNA methylation profiles in saliva and intestinal mucosa. *BMC Genomics* 2019; **20**: 163 [PMID: 30819108 DOI: 10.1186/s12864-019-5553-0]

34   **Howell KJ**, Kraiczy J, Nayak KM, Gasparetto M, Ross A, Lee C, Mak TN, Koo BK, Kumar N, Lawley T, Sinha A, Rosenstiel P, Heuschkel R, Stegle O, Zilbauer M. DNA Methylation and Transcription Patterns in Intestinal Epithelial Cells From Pediatric Patients With Inflammatory Bowel Diseases Differentiate Disease Subtypes and Associate With Outcome. *Gastroenterology* 2018; **154**: 585-598 [PMID: 29031501 DOI: 10.1053/j.gastro.2017.10.007]

35   **Bannister AJ**, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res* 2011; **21**: 381-395 [PMID: 21321607 DOI: 10.1038/cr.2011.22]

36   **Zorro MM**, Aguirre-Gamboa R, Mayassi T, Ciszewski C, Barisani D, Hu S, Weersma RK, Withoff S, Li Y, Wijmenga C, Jabri B, Jonkers IH. Tissue alarmins and adaptive cytokine induce dynamic and distinct transcriptional responses in tissue-resident intraepithelial cytotoxic T lymphocytes. *J Autoimmun* 2020; **108**: 102422 [PMID: 32033836 DOI: 10.1016/j.jaut.2020.102422]

37   **Gutierrez-Achury J**, Zorro MM, Ricaño-Ponce I, Zhernakova DV; Coeliac Disease Immunochip Consortium, RACI Consortium, Diogo D, Raychaudhuri S, Franke L, Trynka G, Wijmenga C, Zhernakova A. Functional implications of disease-specific variants in loci jointly associated with coeliac disease and rheumatoid arthritis. *Hum Mol Genet* 2016; **25**: 180-190 [PMID: 26546613 DOI: 10.1093/hmg/ddv455]

38   **Moffitt AB**, Ondrejka SL, McKinney M, Rempel RE, Goodlad JR, Teh CH, Leppa S, Mannisto S, Kovanen PE, Tse E, Au-Yeung RKH, Kwong YL, Srivastava G, Iqbal J, Yu J, Naresh K, Villa D, Gascoyne RD, Said J, Czader MB, Chadburn A, Richards KL, Rajagopalan D, Davis NS, Smith EC, Palus BC, Tzeng TJ, Healy JA, Lugar PL, Datta J, Love C, Levy S, Dunson DB, Zhuang Y, Hsi ED, Dave SS. Enteropathy-associated T cell lymphoma subtypes are characterized by loss of function of SETD2. *J Exp Med* 2017; **214**: 1371-1386 [PMID: 28424246 DOI: 10.1084/jem.20160894]

39   **Oittinen M**, Popp A, Kurppa K, Lindfors K, Mäki M, Kaikkonen MU, Viiri K. Polycomb Repressive Complex 2 Enacts Wnt Signaling in Intestinal Homeostasis and Contributes to the Instigation of Stemness in Diseases Entailing Epithelial Hyperplasia or Neoplasia. *Stem Cells* 2017; **35**: 445-457 [PMID: 27570105 DOI: 10.1002/stem.2479]

40   **Fellows R**, Varga-Weisz P. Chromatin dynamics and histone modifications in intestinal microbiota-host crosstalk. *Mol Metab* 2020; **38**: 100925 [PMID: 31992511 DOI: 10.1016/j.molmet.2019.12.005]

41   **Shock T**, Badang L, Ferguson B, Martinez-Guryn K. The interplay between diet, gut microbes, and host epigenetics in health and disease. *J Nutr Biochem* 2021; **95**: 108631 [PMID: 33789148 DOI: 10.1016/j.jnutbio.2021.108631]

42   **Yao RW**, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. *Nat Cell Biol* 2019; **21**: 542-551 [PMID: 31048766 DOI: 10.1038/s41556-019-0311-8]

43   **Ma L**, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. *RNA Biol* 2013; **10**: 925-933 [PMID: 23696037 DOI: 10.4161/rna.24604]

44   **Wu GC**, Pan HF, Leng RX, Wang DG, Li XP, Li XM, Ye DQ. Emerging role of long noncoding RNAs in autoimmune diseases. *Autoimmun Rev* 2015; **14**: 798-805 [PMID: 25989481 DOI: 10.1016/j.autrev.2015.05.004]

45   **Fasano A**. Zonulin, regulation of tight junctions, and autoimmune diseases. *Ann N Y Acad Sci* 2012; **1258**: 25-33 [PMID: 22731712 DOI: 10.1111/j.1749-6632.2012.06538.x]

46   **Jauregi-Miguel A**, Santin I, Garcia-Etxebarria K, Olazagoitia-Garmendia A, Romero-Garmendia I, Sebastian-delaCruz M, Irastorza I; Spanish Consortium for the Genetics of Celiac Disease, Castellanos-Rubio A, Bilbao JR. MAGI2 Gene Region and Celiac Disease. *Front Nutr* 2019; **6**: 187 [PMID: 31921880 DOI: 10.3389/fnut.2019.00187]

47   **Ricaño-Ponce I**, Zhernakova DV, Deelen P, Luo O, Li X, Isaacs A, Karjalainen J, Di Tommaso J, Borek ZA, Zorro MM, Gutierrez-Achury J, Uitterlinden AG, Hofman A, van Meurs J; BIOS Consortium; Lifelines Cohort Study, Netea MG, Jonkers IH, Withoff S, van Duijn CM, Li Y, Ruan Y, Franke L, Wijmenga C, Kumar V. Refined mapping of autoimmune disease associated genetic variants with gene expression suggests an important role for non-coding RNAs. *J Autoimmun* 2016; **68**: 62-74 [PMID: 26898941 DOI: 10.1016/j.jaut.2016.01.002]

48   **Plaza-Izurieta L**, Fernandez-Jimenez N, Irastorza I, Jauregi-Miguel A, Romero-Garmendia I, Vitoria JC, Bilbao JR. Expression analysis in intestinal mucosa reveals complex relations among genes under the association peaks in celiac disease. *Eur J Hum Genet* 2015; **23**: 1100-1105 [PMID: 25388004 DOI: 10.1038/ejhg.2014.244]

49   **Hrdlickova B**, Kumar V, Kanduri K, Zhernakova DV, Tripathi S, Karjalainen J, Lund RJ, Li Y,

Ullah U, Modderman R, Abdulahad W, Lähdesmäki H, Franke L, Lahesmaa R, Wijmenga C, Withoff S. Expression profiles of long non-coding RNAs located in autoimmune disease-associated regions reveal immune cell-type specificity. *Genome Med* 2014; **6**: 88 [PMID: 25419237 DOI: 10.1186/s13073-014-0088-0]

50   **Castellanos-Rubio A**, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, Schneider R, Kiledjian M, Bilbao JR, Ghosh S. A long noncoding RNA associated with susceptibility to celiac disease. *Science* 2016; **352**: 91-95 [PMID: 27034373 DOI: 10.1126/science.aad0467]

51   **Castellanos-Rubio A**, Kratchmarov R, Sebastian M, Garcia-Etxebarria K, Garcia L, Irastorza I, Ghosh S. Cytoplasmic Form of Carlr lncRNA Facilitates Inflammatory Gene Expression upon NF-κB Activation. *J Immunol* 2017; **199**: 581-588 [PMID: 28626066 DOI: 10.4049/jimmunol.1700023]

52   **Abadie V**, Jabri B. IL-15: a central regulator of celiac disease immunopathology. *Immunol Rev* 2014; **260**: 221-234 [PMID: 24942692 DOI: 10.1111/imr.12191]

53   **Gnodi E**, Mancuso C, Elli L, Ballarini E, Meneveri R, Beaulieu JF, Barisani D. Gliadin, through the Activation of Innate Immunity, Triggers lncRNA NEAT1 Expression in Celiac Disease Duodenal Mucosa. *Int J Mol Sci* 2021; **22** [PMID: 33525473 DOI: 10.3390/ijms22031289]

54   **Imamura K**, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, Kato A, Kawaguchi Y, Sato H, Yoneda M, Kai C, Yada T, Suzuki Y, Yamada T, Ozawa T, Kaneki K, Inoue T, Kobayashi M, Kodama T, Wada Y, Sekimizu K, Akimitsu N. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol Cell* 2014; **53**: 393-406 [PMID: 24507715 DOI: 10.1016/j.molcel.2014.01.009]

55   **Mongan NP**, Emes RD, Archer N. Detection and analysis of RNA methylation. *F1000Res* 2019; **8** [PMID: 31069058 DOI: 10.12688/f1000research.17956.1]

56   **Olazagoitia-Garmendia A**, Zhang L, Mera P, Godbout JK, Sebastian-DelaCruz M, Garcia-Santisteban I, Mendoza LM, Huerta A, Irastorza I, Bhagat G, Green PH, Herrero L, Serra D, Rodriguez JA, Verdu EF, He C, Bilbao JR, Castellanos-Rubio A. Gluten-induced RNA methylation changes regulate intestinal inflammation *via* allele-specific *XPO1* translation in epithelial cells. *Gut* 2021 [PMID: 33526437 DOI: 10.1136/gutjnl-2020-322566]

57   **Capuano M**, Iaffaldano L, Tinto N, Montanaro D, Capobianco V, Izzo V, Tucci F, Troncone G, Greco L, Sacchetti L. MicroRNA-449a overexpression, reduced NOTCH1 signals and scarce goblet cells characterize the small intestine of celiac patients. *PLoS One* 2011; **6**: e29094 [PMID: 22194996 DOI: 10.1371/journal.pone.0029094]

58   **Jensen MD**, Andersen RF, Christensen H, Nathan T, Kjeldsen J, Madsen JS. Circulating microRNAs as biomarkers of adult Crohn's disease. *Eur J Gastroenterol Hepatol* 2015; **27**: 1038-1044 [PMID: 26230660 DOI: 10.1097/MEG.0000000000000430]

59   **Kalla R**, Ventham NT, Kennedy NA, Quintana JF, Nimmo ER, Buck AH, Satsangi J. MicroRNAs: new players in IBD. *Gut* 2015; **64**: 504-517 [PMID: 25475103 DOI: 10.1136/gutjnl-2014-307891]

60   **Amr KS**, Bayoumi FS, Eissa E, Abu-Zekry M. Circulating microRNAs as potential non-invasive biomarkers in pediatric patients with celiac disease. *Eur Ann Allergy Clin Immunol* 2019; **51**: 159-164 [PMID: 30983306 DOI: 10.23822/EurAnnACI.1764-1489.90]

61   **Bascuñán KA**, Pérez-Bravo F, Gaudioso G, Vaira V, Roncoroni L, Elli L, Monguzzi E, Araya M. A miRNA-Based Blood and Mucosal Approach for Detecting and Monitoring Celiac Disease. *Dig Dis Sci* 2020; **65**: 1982-1991 [PMID: 31781909 DOI: 10.1007/s10620-019-05966-z]

62   **Gadermayr M**, Wimmer G, Kogler H, Vécsei A, Merhof D, Uhl A. Automated classification of celiac disease during upper endoscopy: Status quo and quo vadis. *Comput Biol Med* 2018; **102**: 221-226 [PMID: 29739614 DOI: 10.1016/j.compbiomed.2018.04.020]

63   **Zhou T**, Han G, Li BN, Lin Z, Ciaccio EJ, Green PH, Qin J. Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method. *Comput Biol Med* 2017; **85**: 1-6 [PMID: 28412572 DOI: 10.1016/j.compbiomed.2017.03.031]

64   **Wang X**, Qian H, Ciaccio EJ, Lewis SK, Bhagat G, Green PH, Xu S, Huang L, Gao R, Liu Y. Celiac disease diagnosis from videocapsule endoscopy images with residual learning and deep feature extraction. *Comput Methods Programs Biomed* 2020; **187**: 105236 [PMID: 31786452 DOI: 10.1016/j.cmpb.2019.105236]

65   **Hegenbart S**, Uhl A, Vécsei A. Survey on computer aided decision support for diagnosis of celiac disease. *Comput Biol Med* 2015; **65**: 348-358 [PMID: 25770906 DOI: 10.1016/j.compbiomed.2015.02.007]

66   **Syed S**, Al-Boni M, Khan MN, Sadiq K, Iqbal NT, Moskaluk CA, Kelly P, Amadi B, Ali SA, Moore SR, Brown DE. Assessment of Machine Learning Detection of Environmental Enteropathy and Celiac Disease in Children. *JAMA Netw Open* 2019; **2**: e195822 [PMID: 31199451 DOI: 10.1001/jamanetworkopen.2019.5822]

67   **Syed S**, Ehsan L, Shrivastava A, Sengupta S, Khan M, Kowsari K, Guleria S, Sali R, Kant K, Kang SJ, Sadiq K, Iqbal NT, Cheng L, Moskaluk CA, Kelly P, Amadi BC, Asad Ali S, Moore SR, Brown DE. Artificial Intelligence-based Analytics for Diagnosis of Small Bowel Enteropathies and Black Box Feature Detection. *J Pediatr Gastroenterol Nutr* 2021; **72**: 833-841 [PMID: 33534362 DOI: 10.1097/MPG.0000000000003057]

68   **Koh JEW**, De Michele S, Sudarshan VK, Jahmunah V, Ciaccio EJ, Ooi CP, Gururajan R, Oh SL, Lewis SK, Green PH, Bhagat G, Acharya UR. Automated interpretation of biopsy images for the detection of celiac disease using a machine learning approach. *Comput Methods Programs Biomed* 2021; **203**: 106010 [PMID: 33831693 DOI: 10.1016/j.cmpb.2021.106010]

69   **Foers AD**, Shoukat MS, Welsh OE, Donovan K, Petry R, Evans SC, FitzPatrick ME, Collins N,

Klenerman P, Fowler A, Soilleux EJ. Classification of intestinal T-cell receptor repertoires using machine learning methods can identify patients with coeliac disease regardless of dietary gluten status. *J Pathol* 2021; **253**: 279-291 [PMID: 33225446 DOI: 10.1002/path.5592]

70   **Shemesh O**, Polak P, Lundin KEA, Sollid LM, Yaari G. Machine Learning Analysis of Naïve B-Cell Receptor Repertoires Stratifies Celiac Disease Patients and Controls. *Front Immunol* 2021; **12**: 627813 [PMID: 33790900 DOI: 10.3389/fimmu.2021.627813]

71   **Piccialli F**, Calabrò F, Crisci D, Cuomo S, Prezioso E, Mandile R, Troncone R, Greco L, Auricchio R. Precision medicine and machine learning towards the prediction of the outcome of potential celiac disease. *Sci Rep* 2021; **11**: 5683 [PMID: 33707543 DOI: 10.1038/s41598-021-84951-x]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com